



Facultad de Ingeniería
Carrera de Ingeniería Informática

**”Exploración basada en aprendizaje automático
para el pronóstico académico en Ingeniería
Informática”**

Trabajo de diploma para optar por el título de Ingeniería en Informática.

Diplomante: Carlos Aguilar Aguilar

Tutores: MSc. Viviana R. Toledo Rivero.

Cienfuegos

Curso 2014-2015

Dedicatoria:

Dedico éste trabajo de diploma a mis padres Gregorio Aguilar Alarcón y Marcelina Aguilar Mendoza, a mi hermano Yunion Aguilar Aguilar a los cuales quiero mucho; por todo los sacrificios que han hecho para que yo este donde estoy ; complaciéndome siempre. A ellos y a toda mi familia y amigos por apoyarme y ayudarme siempre que lo he necesitado. Y por que no, a todo aquel que se sienta identificado con esta profesión y quiera seguir mi ejemplo, también al que no.

Agradecimientos:

- ✓ *A mi familia y amigos, los que están y los que ya no se encuentran entre nosotros por creer y confiar siempre en mí, estar presentes cuando los necesito y por el amor y cariño incondicional que me brindan.*
- ✓ *A mi tutora Viviana Toledo Rivero que me ha ayudado muchísimo con la tesis, trabajado junto conmigo, la cual ha sido una madre para mí en todo este periodo.*
- ✓ *A la Secretaría Docente de la Facultad de Ingeniería en las figuras de sus trabajadoras, que tan amablemente nos trataron y nos facilitaron siempre los datos, sin los cuales no habría sido posible la realización de este trabajo*
- ✓ *A mis amigos del aula y a todos mis compañeros y conocidos.*
- ✓ *A todos los profesores que me han preparado en conocimientos, enseñanzas y me han dado su ayuda.*

Resumen:

En esta investigación se aborda el problema de obtener un pronóstico académico de estudiantes de primer y segundo año que cursan la carrera de Ingeniería Informática en la Universidad de Cienfuegos. El trabajo se aborda desde la perspectiva de las técnicas de aprendizaje automático aplicando para la depuración de los datos, selección de rasgos con diferentes evaluadores; se prueban algoritmos que implementan redes neuronales como MultiLayerPerceptron y árbol de decisión C4.5 en su variante J48 para decidir cuál modelo resulta mejor. Sobre los rasgos de mayor peso se realiza un pronóstico usando SMOreg. Para la prueba de los algoritmos se utiliza las implementaciones del software WEKA. Los resultados obtenidos confirman la utilidad de los algoritmos empleados en problemas predictivos, aportan modelos útiles y arrojan información valiosa para el trabajo docente-metodológico en la carrera.

Palabras Claves:

Minería de Datos, Aprendizaje Automático, WEKA.

Abstract:

This research faces the problem of to get an academic forecast from students of Informatics Engineering learning at first and second year of the degree at the Cienfuegos's University, Cuba using machine learning tools. In particular we explored datasets with algorithms MultiLayerPerceptron, J48 and SMOReg using WEKA like software tools and got models in order to predictive who has success or not, taking account the results (marks) in subjects. In addition during the process it got the most important subjects that have high weight to determine it. The results confirm the goodness of algorithms employed and are useful to the principals and professors of the career.

Key Words:

Data Mining, Machine Learning, WEKA.

Índice de contenido

Introducción:	1
Capítulo I: Estado del Arte	7
1.1 Introducción	7
1.2 El concepto de pronóstico. Su intención en este trabajo.	7
1.2.1 Clasificación de los pronósticos respecto al tiempo que abarcan.	7
1.2.2 Precisión del pronóstico.	8
1.2.3 Áreas donde se utiliza el pronóstico.....	8
1.2.4 Importancia de realizar un pronóstico académico.....	10
1.3 Técnicas de Minería de Datos apropiadas a la predicción.	14
1.3.1 Técnicas de minería de datos	15
1.3.2 Aprendizaje supervisado y no supervisado.....	16
1.3.3 Tipos de modelos	16
1.4 Selección de la herramienta de software a utilizar.....	18
1.5 Conclusiones parciales.....	21
Capítulo II: Aprendizaje supervisado y Diseño de Experimentos.....	22
2.1 Introducción	22
2.2 Introducción al aprendizaje supervisado y la minería de datos.....	22
2.2.1 Aprendizaje no supervisado.....	23
2.2.2 Aprendizaje supervisado	23
2.3 Redes neuronales artificiales.....	24
2.3.1 Perceptrón Multicapa (MLP).....	25
2.4 Árboles de Decisión.	28
2.5 Pronóstico.....	31
2.6 Diseño de Experimentos.....	32
2.6.1 Datos a utilizar.	32
2.6.2 Diseño del experimento.....	37

2.7 Conclusiones parciales	40
Capítulo III Análisis de los resultados.	41
3.1 Introducción.	41
3.2 Selección de rasgos	41
3.3 Entrenamiento de la red neuronal.....	46
3.3.1 Prueba de la red entrenada.	49
3.4 Entrenamiento del árbol de decisión.	52
3.4.1 Prueba de los modelos de árboles.	59
3.5 Redes neuronales versus Árboles.	60
3.6 Pronósticos.	65
3.7 Conclusiones parciales.	68
Conclusiones.....	69
Recomendaciones.....	71
Glosario de Términos.....	81
Anexos.	82
Anexo A.....	82
Anexo B.....	100
Anexo C.	109
Anexo D.	158
Anexo E.....	165
Anexo F.....	214
Anexo G.	228
Anexo H.	235
Anexo I.	238

Índice de tablas

Tabla 1 Atributos del modelo 1er año	34
Tabla 2 Atributos del modelo 2do año	37
Tabla 3 Variantes utilizadas para las redes MLP	39
Tabla 4 Métodos aplicados al modelo 1er año	42
Tabla 5 Métodos aplicados al modelo 2do año	43
Tabla 6 Métodos aplicados al modelo 2do año	45
Tabla 7 Resultados de los modelos de 1er año	47
Tabla 8 Resultados de los modelos de 2do año	48
Tabla 9 Resultados de los modelos 1er y 2do año	48
Tabla 10 Resultados del modelo 1er año con entrenamiento y conjunto de prueba	49
Tabla 11 Resultados del modelo 1er año selrasg con entrenamiento y conjunto de prueba	50
Tabla 12 Resultados del modelo 2do año con entrenamiento y conjunto de prueba....	50
Tabla 13 Resultados del modelo 2do año selrasg con entrenamiento y conjunto de prueba	51
Tabla 14 Resultados del modelo 1er y 2do año con entrenamiento y conjunto de prueba	51
Tabla 15 Resultados del modelo 1er y 2do año con entrenamiento y conjunto de prueba	52
Tabla 16 Comparación de los resultados de los árboles en los modelos de 1er año ...	52
Tabla 17 Comparación de los resultados de los árboles en los modelos de 2do año ..	55
Tabla 18 Comparación de los resultados de los árboles en los modelos de 1er y 2do año	57
Tabla 19 Resultados de los árboles con conjunto de prueba para el modelo 1er año .	59
Tabla 20 Resultados de los árboles con conjunto de prueba para el modelo 1er año selrasg	59
Tabla 21 Resultados de los árboles con conjunto de prueba para el modelo 2do año	59
Tabla 22 Resultados de los árboles con conjunto de prueba para el modelo 2do año selrasg	59

Tabla 23 Resultados de los árboles con conjunto de prueba para el modelo 1er y 2do año	60
Tabla 24 Resultados de los árboles con conjunto de prueba para el modelo 1er y 2do año selrasg	60
Tabla 25 %bien clasificados por variantes en los modelos de 1er año	62
Tabla 26 %bien clasificados por variantes en los modelos de 2do año	63
Tabla 27 %bien clasificados por variantes en los modelos de 1er y 2do año	65
Tabla 28 Rasgos decisores en los modelos	65

Índice de figuras

Figura 1 Árbol de decisión para el modelo 1er año	54
Figura 2 Árbol de decisión para el modelo 2do año	56
Figura 3 Árboles de decisión para el modelo 1er y 2do año	58
Figura 4 Parámetros especificados para 1er año	61
Figura 5 Parámetros especificados para 2do año.....	63
Figura 6 Parámetros especificados para 1er y 2do año.....	64
Figura 7 Pronóstico para Matemática I, Introducción a la Programación y Matemática II	66
Figura 8 Pronóstico para Matemática III, Base de Datos y Estructura de Datos I.....	67
Figura 9 Pronóstico combinado para 1er y 2do año	68

Introducción:

El proceso docente educativo presenta interrogantes que en muchas ocasiones son complejas, el nivel de éxito de los estudiantes es una de las más comunes.

Esto se debe a que la medida de este nivel depende de diferentes factores, ya sea los individuales de cada estudiante que viene dado por sus propios factores sociales y educacionales e incluso por otros externos como los mecanismos docentes creados en los diferentes centros de enseñanza donde matriculó.

En la Educación Superior es incluso más complicado medir este nivel debido a que en Cuba existen muchas vías para acceder a la Universidad e ingresan a las aulas una gran cantidad de estudiantes a los que se da un tratamiento homogéneo desde el punto de vista de su enseñanza a pesar de que no se obtienen iguales resultados en cuanto al aprendizaje. Esto es particularmente cierto en la carrera de Ingeniería Informática en la Universidad de Cienfuegos.

Es común escuchar en los últimos años, la opinión de que los estudiantes que egresan de la educación superior no saben pensar, sobre todo en las carreras de ingeniería donde resolver de forma óptima disímiles problemas es imprescindible, al mismo tiempo se precisa hacerlo desde la independencia y la autonomía en el desarrollo profesional. Sin embargo, algunos estudiantes alcanzan niveles de desempeño satisfactorios por lo que las observaciones realizadas inclinan a pensar que existen diferentes factores que determinan este resultado, los cuales son desconocidos y requieren ser estudiados y caracterizados.

Las medidas profilácticas a nivel de colectivo de año, el curso introductorio al ingreso y los exámenes de diagnóstico –en el caso del primer año-, así como los diferentes cortes evaluativos a lo largo de los semestres parecen no bastar para garantizar resultados exitosos para todos los estudiantes, y esta tendencia es más visible en el caso de primer y segundo años de la carrera, donde con más frecuencia se presentan bajas y/o licencias por motivos académicos. Un buen aprovechamiento en los primeros años debería garantizar la correcta inserción del estudiante en el ciclo especializado –a partir de tercer año-, y el tratamiento adecuado desde el punto de vista de su aprendizaje.

En un intento por explicar los factores que determinan el éxito o fracaso de un estudiante de ingeniería y en ocasiones de un grupo de estudiantes, se observan las series de promoción histórica de la carrera de Ingeniería Informática en la Universidad de Cienfuegos, constatándose que desde sus inicios hasta la fecha el índice comenzó con una tendencia a aumentar durante los cuatro primeros cursos, pero desde entonces a la fecha ha mostrado una tendencia sostenida a disminuir. Se constata la necesidad de identificar

posibles factores que desde lo académico permita la toma de decisiones para en un más corto que mediano plazo revertir tal situación.

No existen dos estudiantes iguales. Cada uno posee sus fortalezas y debilidades, diferentes intereses, sentido de responsabilidad, motivación, preparación, etc. Los autores Felder y Silverman [1] afirman que los estudiantes aprenden de muchas maneras: viendo y escuchando, reflexionando y actuando, razonando lógica e intuitivamente, memorizando y visualizando, construyendo analogías y modelos matemáticos. De la misma manera los métodos de enseñanza son diversos. Algunos profesores dictan conferencias, otros demuestran o discuten, algunos se centran en principios y otros en aplicaciones.

La convergencia total de ambos enfoques es sino imposible, bastante difícil, por cuanto los estudiantes se encuentran en grupos recibiendo sus clases pero mostrando un comportamiento individual en cuanto a su forma de aprender, y por su parte el profesor exhibe un comportamiento en el aula que independientemente de los métodos de enseñanza que use está también marcado por su propio estilo de aprendizaje y de enseñanza.

Según diversos autores, cuando el estilo de aprendizaje de la mayoría de los estudiantes de una clase no coincide con el estilo de enseñanza del profesor, los estudiantes no se sienten cómodos en el proceso de enseñanza. Lo que aprende un estudiante en el aula está dado, en parte, por la preparación previa que posea y por su capacidad innata, aunque también se debe tener en cuenta sus atributos como aprendiz, es decir, la manera en que él aprende, y cómo el estilo de enseñanza del profesor se ajusta a esas características.

Este tema, por la importancia que en el plano educativo tiene, ha sido estudiado por la comunidad científica y aplicado a distintos niveles de enseñanza. Como resultado se han obtenido diversos modelos que permiten determinar estilos de aprendizaje, uno de estos es el formulado por Felder y Silverman[1] diseñado para los estudiantes de ingeniería, clasificando a los estudiantes según sus preferencias en una categoría u otra, de un total de cuatro posibles dimensiones.

En ocasiones, un estudiante puede presentar características que lo colocan en un estilo en un momento dado y si se mide nuevamente un tiempo después, presentar características que lo colocan en otro. En tanto el aprendizaje como la enseñanza son procesos dinámicos donde intervienen estudiantes y profesores, esto nos conduce a pensar que si los profesores son capaces de adaptar su estilo de enseñanza a los estilos de aprendizaje de tal forma que no afecte negativamente el rendimiento del estudiante o su actitud frente a los contenidos se favorece la interacción entre ambos sujetos y el proceso sale beneficiado. El profesor debe tener en cuenta cuáles de las cuatro

dimensiones merecen ser potenciadas desde la enseñanza, pero para ello debe conocerlas cuando comienza el curso para usarlas en su provecho.

Ejemplos de estudios de este tipo, es el presentado por Durán y Costagauta [2] y Quiñones [3], este último desarrollado en la carrera de Ingeniería Informática de la Universidad de Cienfuegos. Como resultado de este último trabajo, se clasifica el grupo en subgrupos, atendiendo a los estilos predominantes y se sugiere adaptar el estilo de enseñanza al estilo de aprendizaje predominante en el grupo, de forma que sea propicio para que el estudiante entienda mejor el contenido, pero no de forma individual, sino en forma grupal. En caso de que el estudiante no este enmarcado en el estilo general debe impulsarse a este con estrategias que sean de su agrado y le permitan adaptarse mejor al estilo formulado. El fundamento de esta idea consiste en la necesidad de que el estudiante posea habilidades que caracterizan cada estilo [4] .

En un trabajo posterior de Chang [5] se determinó el estilo de enseñanza de los profesores, en la propia carrera, usando una modificación del modelo de Felder y Silverman, antes citado; como resultado principal se obtiene que no concuerdan estilos de enseñanza y aprendizaje de docentes y dicentes, lo que conduce a replantearnos la interrogante inicial en cuanto a qué factores podrían incidir en el nivel de éxito –o fracaso- alcanzado por nuestros estudiantes, fundamentalmente en los dos primeros años de la carrera..

Como requisito inicial, en este problema no se tiene una idea clara de los resultados a que nos pueden conducir los datos, y se cuenta además con un buen número de datos para explorar. Se tienen en Secretaría Docente de la Facultad de Ingeniería los resultados académicos alcanzados por nuestros estudiantes desde los inicios de la carrera sin que se hayan explorado buscando posibles nexos entre el nivel de éxito en un curso académico y los resultados que se alcanzan en cada una de las asignaturas. En situaciones de este tipo es muy utilizada la minería de datos.

La Minería de Datos (DM, Data Mining por sus siglas en inglés) consiste en la extracción no trivial de información que reside de manera implícita en los datos. Dicha información podrá resultar potencialmente útil para algún proceso. En otras palabras, la minería de datos prepara, sondea y explora los datos para sacar la información oculta en ellos.

Las bases de la minería de datos se encuentran en la inteligencia artificial y en el análisis estadístico. Mediante los modelos extraídos utilizando técnicas de minería de datos se aborda la solución a problemas de predicción, clasificación y segmentación.

Programas de computadoras modernos gestionan automáticamente grandes cantidades de datos buscando patrones o regularidades. Si patrones sólidos son encontrados, entonces se podrá generalizar y obtener predicciones

precisas sobre datos futuros. Muchos patrones podrían ser banales y nada interesantes, otros serían espurios, o coincidencias accidentales del conjunto de datos utilizado. Los datos reales son imperfectos, algunos serán falseados, otros estarán ausentes. Cualquier cosa descubierta será inexacta, habrá excepciones a cada regla y casos no cubiertos por ninguna regla, luego se necesitan algoritmos robustos para enfrentar datos imperfectos y extraer regularidades que son inexactas pero útiles. [6]

Algunos trabajos que usan la capacidad predictiva de las técnicas de minería de datos y que preceden al nuestro son, entre otros: predicción del rendimiento académico de los estudiantes argentinos ingresantes de la carrera de Bioquímica [7], o más recientemente: predicción de resultados académicos en la asignatura SO haciendo uso de las redes neuronales [8]; y otro titulado predicción de resultados académicos en las asignaturas Estructuras de Datos 1 y 2 [9], ambos desarrollados para la carrera de Ingeniería Informática, en la CUJAE, Habana, Cuba.

Como se constata el problema es actual y no solo localizado en la Universidad de Cienfuegos, sino que en otras universidades cubanas donde se estudia la carrera, los estudiantes presentan problemas en asignaturas consideradas importantes por su aporte al ingeniero en Informática. Estos trabajos abordan la problemática desde una o dos asignaturas que se consideran importantes como precedentes para otra, ubicada en otro semestre del mismo año - como es la ED2-, o en el año próximo como es el caso de SO. Nuestra intención es más general, tratando de descubrir factores que pueden afectar el rendimiento en un año y que por tanto comprometen el aprobado de ese año e impiden el paso para el año siguiente.

Como se observa, en el caso que nos ocupa, la minería de datos es ideal para intentar caracterizar el comportamiento de los estudiantes de Ingeniería Informática en cuanto a su rendimiento académico, dados los resultados históricos que por años se han ido obteniendo sin que se hayan explorado. Interesa fundamentalmente en la carrera descubrir los principales factores que determinan el éxito en un curso dado, siendo el problema de mayor importancia durante los dos primeros años de la carrera. Tales resultados aportarían suficiente información para la carrera, en aras de explicar el fenómeno de la eficiencia, y quizás otros que surjan en el marco de este propio análisis.

Teniendo en cuenta todo lo anterior se define como **problema a resolver**:

¿Cómo obtener un pronóstico académico de un estudiante de primer o segundo año de la carrera de Ingeniería informática?

Se identifica como **objeto de estudio** de este trabajo: el proceso de análisis del resultado de las evaluaciones obtenidas por los estudiantes usando minería de datos.

En este proceso se analizan los datos a partir de técnicas de aprendizaje automático, se seleccionan los algoritmos para realizar la exploración e interpretación de los datos y se escoge el que proporcione los resultados que mejor se adecuen a lo que se desea obtener por lo que se define como **campo de acción**: La exploración de las evaluaciones obtenidas por estudiantes de primero y segundo años de la carrera Ingeniería Informática de la UCf, utilizando algoritmos de minería de datos basados en aprendizaje automático.

Idea a defender

La aplicación de técnicas de minería de datos basadas en Aprendizaje Automático a la información relacionada con los resultados de las evaluaciones de los estudiantes de primer y segundo años de Ingeniería Informática permitirá afirmar con cierto grado de certeza si el estudiante tendrá éxito o no en sus primeros años.

De este modo se define como **objetivo general**: Aplicar técnicas de minería de datos basadas en Aprendizaje Automático a los resultados académicos de los estudiantes de primer y segundo años de Ingeniería Informática.

Se han trazado además, los siguientes **objetivos específicos**:

- Determinar el Estado del Arte sobre las técnicas de Minería de Datos y su aplicación en diferentes áreas de la ciencia.
- Aplicar las técnicas de aprendizaje automático a la determinación del pronóstico académico.

Las **tareas** a realizar para cumplir los objetivos propuestos son las siguientes:

- Análisis de tendencias y metodologías actuales utilizadas en minería de datos.
- Revisión bibliográfica de las técnicas de aprendizaje automático utilizadas.
- Selección de los algoritmos a utilizar.
- Obtención de los datos para el experimento.
- Aplicación de los algoritmos seleccionados a los juegos de datos.
- Análisis comparativo de los resultados obtenidos y selección del algoritmo que mejor se ajusta al problema planteado.

Para una mejor comprensión del trabajo, el mismo se ha estructurado de la forma siguiente: Introducción, Capítulos, Conclusiones y Recomendaciones.

Los capítulos del desarrollo se reseñan a continuación:

Capítulo 1: Estado del Arte: Principales conceptos asociados a la predicción y principales técnicas de MD utilizadas.

Capítulo 2: Aprendizaje Automático y Diseño de los Experimentos: Datos y variables seleccionadas para la exploración. Herramienta y principales algoritmos utilizados; justificación de su uso.

Capítulo 3: Análisis de los Resultados: Comparación de los resultados obtenidos a partir de la aplicación de algoritmos seleccionados para el aprendizaje y selección del que mejor se adecuen al objetivo de la investigación.

Capítulo I: Estado del Arte

1.1 Introducción

En el presente capítulo se presentan los conceptos relacionados con el pronóstico y lo que se ha dado en llamar en este trabajo pronóstico académico, la importancia que tiene para mejorar la calidad del proceso de enseñanza aprendizaje. Además, se referencian las principales técnicas de Minería de Datos basadas en Aprendizaje Automático que permiten realizar predicciones.

1.2 El concepto de pronóstico. Su intención en este trabajo.

Pronóstico según el diccionario de la lengua española, Larousse significa "conjetura acerca de lo que puede suceder, señal por donde se conjetura una cosa futura". En su segunda acepción es considerada **Predicción**. [10]

El pronóstico o predicción constituye una de las esencias claves de la ciencia, de una teoría científica o de un modelo científico. Así, el éxito se mide por el éxito o acierto que tengan sus predicciones. La predicción en el contexto científico es una declaración precisa de lo que ocurrirá en determinadas condiciones especificadas. Se puede expresar a través del silogismo: "Si A es cierto, entonces B también será cierto."

El método científico concluye con la prueba de afirmaciones que son consecuencias lógicas del corpus de las teorías científicas. Generalmente esto se hace a través de experimentos que deben poder repetirse o mediante estudios observacionales rigurosos. Una teoría científica cuyas aseveraciones no son corroboradas por las observaciones, por las pruebas o por experimentos probablemente será rechazada.

Las teorías que generan muchas predicciones que resultan de gran valor – tanto por su interés científico como por sus aplicaciones- se confirman o se falsean fácilmente y, en muchos campos científicos, las más deseables son aquellas que, con número bajo de principios básicos, predicen un gran número de sucesos.

Diversas técnicas modernas, ofrecen métodos y herramientas para la realización de pronósticos o predicciones en diferentes campos de la ciencia.

Por su parte, el **pronóstico estadístico** es definido como el proceso de estimación en situaciones de incertidumbre.

1.2.1 Clasificación de los pronósticos respecto al tiempo que abarcan.

Existen pocas esferas de la vida moderna donde no se utilicen los pronósticos con fines de mejorar la eficiencia de la gestión, siendo esta de la más diversa índole.

Atendiendo al tiempo que abarcan los pronósticos se pueden clasificar en: [11]

- **Pronósticos a corto plazo:** En las empresas modernas, este tipo de pronóstico se efectúa cada mes o menos, y su tiempo de planeación tiene vigencia de un año. Se utiliza para programas de abastecimiento, producción, asignación de mano de obra a las plantillas de trabajadores, y planificación de los departamentos de fabricación, entre otros.
- **Pronósticos a mediano plazo:** Abarca un lapso de seis meses a tres años. Estos se utilizan para estimar planes de ventas, producción, flujos de efectivo y elaboración de presupuestos.
- **Pronósticos a largo plazo:** Este tipo de pronóstico se utiliza en la planificación de nuevas inversiones, lanzamiento de nuevos productos y tendencias tecnológicas de materiales, procesos y productos, así como en la preparación de proyectos. El tiempo de duración es de tres años o más.

1.2.2 Precisión del pronóstico.

El trabajo de pronosticar envuelve la consideración de su precisión. Esto significa manejar niveles de error permisibles. El error del pronóstico es la diferencia entre el valor real y el pronosticado del período correspondiente.

Si consideramos t como el período de tiempo, Y el valor real para el período y F el valor que se había pronosticado, entonces E_t

$$E_t = Y_t - F_t$$

Es el error del pronóstico del período t .

Mientras mayor sea el error cometido, menos calidad tiene el pronóstico.

1.2.3 Áreas donde se utiliza el pronóstico.

La mayoría de las organizaciones confían en los pronósticos como vía para la planificación de sus inversiones [12]. Entre ellas se cuentan las empresas modernas, hospitales, meteorología, etc. Normalmente cuando se refiere a un área específica se acompaña de un adjetivo que clarifica su significado. Presentamos una breve descripción de algunos que permite identificar sus principales características:

- **Pronóstico en Medicina.**

En medicina, el **pronóstico**, generalmente de una enfermedad, es el conjunto de datos que posee la ciencia médica sobre la probabilidad de que ocurran determinadas situaciones en el transcurso del tiempo o historia natural de la enfermedad. Es la predicción de los sucesos que ocurrirán en el desarrollo de una enfermedad en términos estadísticos.

Entre sus principales características se tienen:

- Este conjunto de datos se obtienen de los múltiples estudios clínicos realizados sobre una enfermedad en concreto, generalmente series de casos. En estos estudios se obtienen unas variables llamadas

factores pronóstico que el médico debe averiguar en un individuo en particular a través de la sintomatología y de las pruebas diagnósticas realizadas. Según estos factores pronóstico se establece un **tipo de tratamiento**.

- Un pronóstico se puede expresar de una forma cualitativa o cuantitativa. Con respecto a la primera son comunes los términos "buen, mal pronóstico o intermedio", o "leve, moderado o grave". El término "pronóstico reservado" es un "pronóstico incierto" o desconocido, que puede hacer sospechar lo peor. Con respecto a la forma cuantitativa de expresar el pronóstico, se utilizan a menudo porcentajes y tasas de supervivencia o mortalidad.

- **Pronóstico del tiempo.**

El **pronóstico del tiempo** (conocido también como **pronóstico meteorológico, predicción del tiempo, boletín meteorológico, boletín climatológico, boletín del tiempo, predicción meteorológica o pronóstico climatológico**) es la aplicación de tecnología y de ciencia para predecir el estado de la atmósfera para un período futuro y una localidad o región dada.

La historia del pronóstico del tiempo es milenaria, aunque los paradigmas y las técnicas usadas han cambiado significativamente. Los pronósticos se hacen colectando tantos datos como sea posible acerca del estado de la atmósfera (particularmente temperatura, humedad, viento) y usando conocidos procesos atmosféricos (a través de la meteorología) para determinar los patrones futuros atmosféricos. Sin embargo, la naturaleza caótica de la atmósfera y el entendimiento incompleto de los procesos hacen que los pronósticos sean menos seguros al incrementarse el rango temporal del pronóstico. Es por ello que actualmente se usa en muchos países –entre los que se encuentra Cuba- un pronóstico a corto plazo con un margen de seguridad de alrededor de 72 horas, que permite pronosticar con un grado de certeza mayor que si se utilizaran pronósticos a más largo plazo.

Las dos personalidades que más acreditaron el nacimiento del pronóstico como ciencia fueron Francis Beaufort (recordado por la escala Beaufort) y su protegido Robert Fitzroy (desarrollador del barómetro Fitzroy). Ambos eran influyentes en la Naval Británica y en los círculos gubernamentales, y fueron ridiculizados por la prensa cuando dieron a conocer sus avances, aunque sus trabajos ganaron fuerza académica, y fueron aceptados por la Armada Británica, formando la base del conocimiento actual del pronóstico del tiempo.

Predicciones modernas del tiempo ayudan a evacuaciones y potencialmente salva vidas y mitiga daños. [13]

- **Pronóstico deportivo:**

Los **pronósticos deportivos**, también llamados *tips* o *picks*, son el fruto de un trabajo de análisis estadísticos y de la experiencia e intuición de los pronosticadores y generan un importante volumen de negocio en el sector de las apuestas online.

La figura de los tipsters se ha potenciado con la inclusión de las casas de apuestas en Internet pero no es más que una ampliación del clásico corredor de apuestas.

Los tipsters actuales usan programas propios estadísticos muy completos que generan un estudio de probabilidades. Además, al especializarse en los deportes, siguen muy de cerca otras variables no matemáticas como la motivación, conflictos internos de los equipos o jugadores, trayectorias o lesiones. La programación que usan estos profesionales suele ser diseñada por ellos mismos y consiste en una base de datos cruzada de decenas de variables, en función de cada deporte y suele generar informes incidiendo en determinadas variables que considera más decisivas y el resultado final es siempre matizado por otro número de variables menos matemáticas y asociadas a la actualidad deportiva.

Dado que los datos se vuelcan en apuestas normalmente ganadoras, las rentabilidades de estos profesionales suele ser bastante alta, aunque el deporte siempre guarda sorpresas y existe un porcentaje de riesgo, en general, no muy elevado. Los tipsters profesionales suelen exhibir su trayectoria de rentabilidad en sus websites y cobran por sus servicios. Existen incluso pronosticadores que sólo admiten un número cerrado de clientes, y cuyos servicios pueden llegar a los 300 o 400 euros mensuales.

1.2.4 Pronóstico académico.

El término pronóstico académico ha sido usado con diferentes intenciones en el ámbito escolar y en su mayoría vinculado al de rendimiento académico (RA). El rendimiento académico, a su vez, ha sido definido de múltiples formas, pero la posición más decantada, apoya una definición operativa del mismo, [14], [15] y un somero examen de los estudios empíricos dedicados al análisis del RA [16] muestra claramente una tendencia universal a la operacionalización del constructo, vinculada fundamentalmente en los países latinoamericanos a las calificaciones escolares, mientras que en otros países, sobre todo los anglosajones, está identificada con los resultados alcanzados en test de rendimiento escolar (testing).

Uno de los test de pronóstico académico más citado es el Test APT de Pronóstico Académico [17], el cual tiene como objetivo la evaluación de la capacidad de aprovechamiento del escolar, es una prueba de aplicación individual o colectiva para niños a partir de los 11 años, no obstante, la edad óptima de aplicación de la Batería APT es la comprendida entre los trece y los

quince años. Es un instrumento que consta de tres subconjuntos que pueden ser aplicados y valorados de forma independiente:

- 1- Razonamiento abstracto (APT-RA)
- 2- Factor Numérico (APT-N)
- 3- Factor Verbal (APT-V).

Se aplica de forma colectiva y ellos tres en su conjunto, aprecian las aptitudes que más altamente correlacionan con el éxito académico en enseñanzas medias. Se ha aplicado y tipificado también en poblaciones profesionales. Existe una forma paralela del Factor RA.

Sin embargo, las calificaciones escolares en las áreas y/o disciplinas fundamentales, tanto por su relevancia legal y académica como por su trascendencia social y personal en la orientación de los alumnos [18], parecen constituir la mejor definición del RA [19]. La unanimidad coincidente de la investigación empírica no deja lugar a dudas.

Es de destacar que en el análisis de los antecedentes de la investigación sobre RA se encuentra un escaso número de trabajos dedicados al bachillerato – nivel pre-universitario en Cuba-, y a la enseñanza superior, mientras que predominan casi absolutamente los dedicados a la enseñanza general – primaria y secundaria en nuestro país; sobre todo en los niveles secundarios. Sin duda, factores como la gratuidad y la obligatoriedad de esta etapa suscitan un mayor interés social en los países donde la educación no es gratis, y una mayor dedicación de recursos; por otro lado, la investigación educativa es más asequible en la enseñanza general por la falta de especialización del currículo, si se compara con la relativa complejidad estructural y el importante salto cualitativo que supone la especialización de las disciplinas en el bachillerato y aún más en los niveles superiores. Estas razones justifican y apoyan el estudio del RA en nuestro nivel.

Es importante señalar que numerosos estudios avalan como mejor predictor del RA el propio RA anterior (en estudios para bachillerato, Gómez [20]; MEC, [21]; Reparaz, [22]; entre otros). Si bien este resultado no aporta ninguna idea novedosa, sí resalta la homogeneidad del constructo RA consigo mismo, llegando a alcanzar elevadísimas correlaciones [23].

En este estudio se pretende abordar la predicción del RA en la carrera de Ingeniería Informática en la Universidad de Cienfuegos, empleando variables como el propio RA de estudiantes de cursos anteriores lo cual no ha sido explorado en nuestro país.

Luego, en el contexto de este trabajo, el **pronóstico**, al que en lo sucesivo se llama **pronóstico académico** se refiere al pronóstico de un estudiante, y es el conjunto de datos que se registran y exploran del estudiante, sobre la probabilidad de que ocurran determinadas situaciones en el transcurso del tiempo.- que en este caso será de un curso escolar-.

Es la predicción de los sucesos que ocurrirán en el desarrollo del estudiante durante un curso escolar en términos de rendimiento en sus asignaturas.

Puede darse en dos formas distintas:

- Puede tratarse de un pronóstico general que permita dado un estudiante -al que hipotéticamente se llama X-, y su rendimiento en algunas materias, predecir el comportamiento en función de si aprueba o no, el año que cursa. Este resultado se obtendría sobre la base de datos históricos de estudiantes que lo han precedido y cuyo comportamiento ya ha sido analizado y registrado.
- Puede tratarse de un pronóstico para la materia –o materias- que se cursan en determinado año, y también teniendo en cuenta los datos históricos que se registran, determinar el comportamiento futuro de la misma en un lapso de tiempo.

Fundamentalmente el estudiante que vence con éxitos todas las asignaturas de un curso, termina aprobado, y pasa al curso siguiente. Existen un grupo de estudiantes que al no vencer los objetivos de las asignaturas, no aprueban.

En la educación superior cubana, por la flexibilidad del sistema de ingreso a las universidades se cuenta con una matrícula diversa en las aulas. Paralelo a esta situación se tiene un reglamento que dirige el trabajo docente-metodológico y que determina un sistema de evaluación flexible, posibilitando que el estudiante aunque no apruebe el año pueda continuar en el sistema, y dentro de este venza sus dificultades.

De esta situación se derivan fundamentalmente **cuatro** casos:

- El estudiante que **aprueba** todas sus asignaturas y continúa al año superior.
- El estudiante que aprueba casi todas sus asignaturas y continúa al año superior con **arrastre** de hasta dos de las asignaturas del año precedente, siempre y cuando sean de distintos semestres.
- El estudiante que no aprueba dos asignaturas de un mismo semestre o más de dos de cualquier semestre, que continúa en el sistema **repitiendo** el año que cursaba.
- El estudiante que **no aprueba** y causa baja, abandonando la universidad.

A los efectos de simplificar el trabajo, se consideran en un primer momento de análisis de los resultados académicos, solamente las **dos posibilidades** siguientes:

- El estudiante **aprueba**. Se corresponde al primero de los cuatro casos anteriores.
- El estudiante **no aprueba**, que se corresponde con los restantes tres casos.

Aunque existe la posibilidad –según artículo 47 del ROD del MES [24]- de que un estudiante desaprobe hasta dos asignaturas del año que cursa pertenecientes a distintos semestres, y promueva al año inmediato superior matriculando las asignaturas pendientes como arrastre en el siguiente curso, a los efectos de los objetivos de nuestra investigación esos estudiantes se

consideran desaprobados, dado que no vencen los objetivos del año, y se registran en la base de datos con la categoría no aprueba.

En los últimos cursos académicos, y durante los dos primeros años de la carrera fundamentalmente, se observa con frecuencia mayor a la deseada que se presentan los últimos tres casos, lo que provoca análisis de la situación docente a nivel de colectivo de años y carrera, en aras de encontrar factores desencadenantes.

En estudios realizados en años anteriores se determinaron los estilos de aprendizaje de estudiantes y los de enseñanza de los profesores, detectándose que ambos difieren, lo que explica solo en parte los resultados que se obtienen. No se cuenta con ningún estudio al interior de los resultados de los estudiantes que expliquen cuál o cuáles materias determinan con mayor fuerza los resultados que motivan la baja del sistema, o determinan su permanencia repitiendo o con arrastres, lo que motiva este trabajo exploratorio al interior de los datos recopilados, buscando posibles factores de riesgo.

Al comenzar a explorar estos datos, se obtendrá de forma paralela resultados en cuanto a las asignaturas que podrían aportar datos de valor. No se conoce a priori qué podría resultar de esta minería –como es característico de los trabajos de minería- pero se pondrá toda la intencionalidad para obtener, siempre que sea posible, conocimiento.

De encontrarse factores pronóstico que determinen un patrón de riesgo se podría en el año académico del estudiante, determinar un tipo de tratamiento que facilite su tránsito al año siguiente sin demoras.

Por ser en primer y segundo años donde con mayor frecuencia se presentan los problemas de la calidad y de la eficiencia académica, tal como se ha mencionado antes, es sobre estos datos que se trabaja en la presente investigación.

1.2.4.1 Importancia de realizar un pronóstico académico.

El conocimiento, análisis y medición de los resultados de los estudiantes, y la influencia que puedan tener estos en su desarrollo como estudiante, proporciona gran ayuda en el proceso docente educativo donde la forma de aprender del estudiante, e incluso la del profesor influyen notablemente en la formación académica.

Este es un tema complejo, sin embargo, en las últimas dos décadas se ha debatido ampliamente y está siendo común el análisis de diferentes indicadores para adaptar el proceso de enseñanza aprendizaje de forma cada vez más asequible al estudiante.

Los documentos educativos, en múltiples ocasiones, hablan de la importancia que tiene que los docentes conozcan las dificultades de los estudiantes, y en particular las que presentan en sus materias, sin embargo muchos educadores ignoran las características particulares del aprendizaje de sus alumnos respecto a su materia.

Conocer las características del aprendizaje en una materia y si presenta o no dificultades para ser vencida ofrece al profesor las siguientes ventajas: [25]

- Orientar mejor el aprendizaje de cada alumno si se conoce cómo aprenden. Es decir, que la selección de las estrategias didácticas y estilo de enseñanza será más efectivo.
- La aplicación en el aula de los estilos de aprendizaje concretos, como una vía científica para individualizar la instrucción.
- Si la meta educativa es lograr que el alumno aprenda a aprender, entonces se debe apostar por ayudarlo a conocer y mejorar sus propias debilidades en cuanto a asimilación.

Para el alumno también resultaría ventajoso saber: [25]

- cómo controlar su propio aprendizaje.
- cómo diagnosticar sus puntos fuertes y débiles como alumno.
- conocer en qué condiciones aprende mejor.
- cómo aprender de la experiencia de cada día.
- cómo superar las dificultades que se le presentan en su proceso de aprendizaje.

El conocimiento del alumno permitiría a los docentes utilizar las estrategias, los medios y los recursos a su alcance para favorecer y fomentar la mejoría en las habilidades con mayores dificultades, y desarrollar en ellos la adaptabilidad y flexibilidad en el aprendizaje, de cara a un mundo laboral próximo al que tendrán que amoldarse continuamente en su futuro profesional.

Lo anterior revela la importancia de conocer los factores de riesgo en un año académico en particular, y ganar conciencia sobre las diferencias entre los estudiantes y sus resultados individuales para aplicar conscientemente las estrategias de enseñanza que se tienen disponibles y como deben enmarcarse dentro de nuestro contexto docente.

1.3 Técnicas de Minería de Datos apropiadas a la predicción.

El aumento del volumen y variedad de información que se encuentra informatizada en bases de datos digitales y otras fuentes ha crecido espectacularmente en las últimas décadas. Gran parte de esta información es histórica, es decir, representa transacciones o situaciones que se han producido. Aparte de su función de "memoria de la organización", la información histórica es útil para explicar el pasado, entender el presente y predecir la información futura. La mayoría de las decisiones de empresas, organizaciones e instituciones se basan también en información sobre experiencias pasadas extraídas de fuentes muy diversas. Además, ya que los datos pueden proceder de fuentes diversas y pertenecer a diferentes dominios, parece clara la inminente necesidad de analizar los mismos para la obtención de información útil para la organización.

En muchas situaciones, el método tradicional de convertir los datos en conocimiento consiste en un análisis e interpretación realizada de forma manual. El especialista en la materia, digamos por ejemplo un médico, analiza los datos y elabora un informe o hipótesis que refleja las tendencias o pautas de los mismos. Por ejemplo, un grupo de médicos puede analizar la evolución de enfermedades infecto-contagiosas entre la población para determinar el rango de edad más frecuente de las personas afectadas. Este conocimiento, validado convenientemente, puede ser usado en este caso por la autoridad sanitaria competente para establecer políticas de vacunaciones.

Esta forma de actuar es lenta, cara y altamente subjetiva. De hecho, el análisis manual es impracticable en dominios donde el volumen de los datos crece exponencialmente: la enorme abundancia de datos desborda la capacidad humana de comprenderlos sin la ayuda de herramientas potentes. Consecuentemente, muchas decisiones importantes se realizan, no sobre la base de la gran cantidad de datos disponibles, sino siguiendo la propia intuición del usuario al no disponer de las herramientas necesarias. Éste es el principal cometido de la minería de datos: resolver problemas analizando los datos presentes en las bases de datos.

1.3.1 Técnicas de minería de datos

Como ya se ha comentado, las técnicas de la minería de datos provienen de la Inteligencia artificial y de la estadística, dichas técnicas, no son más que algoritmos, más o menos sofisticados que se aplican sobre un conjunto de datos para obtener resultados.

Entre las técnicas más representativas se encuentran:

- **Redes neuronales.**- Son un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso de los animales. Se trata de un sistema de interconexión de neuronas en una red que colabora para producir un estímulo de salida. Algunos ejemplos de red neuronal son:
 - El Perceptrón.
 - El Perceptrón multicapa.
 - Los Mapas Autoorganizados, también conocidos como redes de Kohonen.
- **Regresión lineal.**- Es la más utilizada para formar relaciones entre datos. Rápida y eficaz pero insuficiente en espacios multidimensionales donde puedan relacionarse más de dos variables.
- **Árboles de decisión.**- Un árbol de decisión es un modelo de predicción utilizado en el ámbito de la inteligencia artificial, dada una base de datos se construyen estos diagramas de construcciones lógicas, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que suceden de forma sucesiva, para la resolución de un problema. Ejemplos:

- Algoritmo ID3.
- Algoritmo C4.5.
- **Modelos estadísticos.**- Es una expresión simbólica en forma de igualdad o ecuación que se emplea en todos los diseños experimentales y en la regresión para indicar los diferentes factores que modifican la variable de respuesta.
- **Agrupamiento o *Clustering*.**- Es un procedimiento de agrupación de una serie de vectores según criterios habitualmente de distancia; se tratará de disponer los vectores de entrada de forma que estén más cercanos aquellos que tengan características comunes. Ejemplos:
 - Algoritmo K-means.
 - Algoritmo K-medoids.
- **Reglas de asociación.**- Se utilizan para descubrir hechos que ocurren en común dentro de un determinado conjunto de datos.

1.3.2 Aprendizaje supervisado y no supervisado.

Según el objetivo del análisis de los datos, los algoritmos utilizados se clasifican en supervisados y no supervisados [26]:

- **Algoritmos supervisados** (o predictivos): predicen un dato (o un conjunto de ellos) desconocido a priori, a partir de otros conocidos.
- **Algoritmos no supervisados** (o del descubrimiento del conocimiento): se descubren patrones y tendencias en los datos.

1.3.3 Tipos de modelos

La minería de datos tiene como objetivo analizar los datos para extraer conocimiento. Este conocimiento puede ser en forma de relaciones, patrones o reglas inferidos de los datos y (previamente) desconocidos, o bien en forma de una descripción más concisa (es decir, un resumen de los mismos). Estas relaciones o resúmenes constituyen el modelo de los datos analizados. Existen muchas formas diferentes de representar los modelos y cada una de ellas determina el tipo de técnica que puede usarse para inferirlos.

En la práctica, los modelos pueden ser de dos tipos:

- **Predictivos:** Los modelos predictivos pretenden estimar valores futuros o desconocidos de variables de interés, que denominamos *variables objetivo* o *dependientes*, usando otras variables o campos de la base de datos, a las que nos referimos como *variables independientes* o *predictivas*.
- **Descriptivos:** Los modelos descriptivos, en cambio, identifican patrones que explican o resumen los datos, es decir, sirven para explorar las propiedades de los datos examinados, no para predecir nuevos datos.

Cada tarea puede ser realizada usando distintas técnicas. Por ejemplo, los modelos inferidos por los árboles de decisión y las redes neuronales (por citar dos técnicas de las más conocidas y utilizadas) pueden inferir modelos predictivos. Igualmente, para una misma técnica se han desarrollado diferentes algoritmos que difieren en la forma y criterios concretos con los que se construye el modelo.

Por ser de interés para este trabajo, se caracterizan dos técnicas de las más utilizadas.

1.3.3.1 Árboles de decisión.

Los árboles de decisión son una serie de decisiones o condiciones organizadas en forma jerárquica, a modo de árbol. Son muy útiles para encontrar estructuras en espacios de varias dimensiones, y en problemas que mezclen datos categóricos y numéricos. Esta técnica se usa en tareas de clasificación, agrupamiento y regresión. Los árboles de decisión usados para predecir variables categóricas reciben el nombre de árboles de clasificación, ya que distribuyen las instancias en clases. Cuando los árboles de decisión se usan para predecir variables continuas se llaman árboles de regresión.

1.3.3.2 Redes neuronales artificiales.

Las redes neuronales artificiales son todo un paradigma de computación muy potente que permite modelar problemas complejos en los que puede haber interacciones no lineales entre variables. Como los árboles de decisión, las redes neuronales pueden usarse en problemas de clasificación, de regresión y de agrupamiento.

Una red neuronal puede verse como un grafo dirigido con muchos nodos (elementos del proceso) y arcos entre ellos (sus interconexiones). Cada uno de estos elementos funciona independientemente de los demás, usando datos locales (la entrada y la salida del nodo) para dirigir su procesamiento.

La organización más popular de una red neuronal consta de una capa de entrada, en la que cada nodo corresponde a una variable independiente a examinar, unos nodos internos organizados en una o varias capas ocultas y una capa de salida con los nodos de salida (los posibles valores de las variables objetivo). Cada nodo de la capa de entrada está conectado a cada nodo de la capa oculta. Los nodos de la capa oculta pueden estar conectados a nodos de otra capa oculta o a los nodos de la capa de salida. Cada arco está etiquetado por un peso de conexión (w) y en cada nodo hay una función de activación que indica el efecto de ese nodo sobre los datos que entran en él.

Para usar una red neuronal ya entrenada se introducen los valores de los atributos de una instancia en los nodos de entrada y los nodos de salida determinan la predicción para dicha instancia.

Los pesos de conexión son parámetros desconocidos que deben estimarse por un método de entrenamiento. El método más comúnmente utilizado es el de propagación hacia atrás (*backpropagation*). La idea básica del método es reducir el valor de error de la salida de la red.

Las redes neuronales tienen una gran capacidad de generalización para problemas no lineales, aunque requieren bastantes datos para su entrenamiento. Su mayor desventaja es, no obstante, que el modelo aprendido es difícilmente comprensible.

- **Aplicaciones.**

Las características de las RNA las hacen bastante apropiadas para aplicaciones en las que no se dispone a priori de un modelo identificable que pueda ser programado, pero se dispone de un conjunto básico de ejemplos de entrada (previamente clasificados o no). Asimismo, son altamente robustas tanto al ruido como a la disfunción de elementos concretos y son fácilmente paralelizables.

Esto incluye problemas de clasificación y reconocimiento de patrones de voz, imágenes, señales, etc. Asimismo se han utilizado para encontrar patrones de fraude económico, hacer predicciones en el mercado financiero, hacer predicciones de tiempo atmosférico, etc.

También se pueden utilizar cuando no existen modelos matemáticos precisos o algoritmos con complejidad razonable, por ejemplo la red de Kohonen ha sido aplicada con un éxito más que razonable al clásico problema del viajante (un problema para el que no se conoce solución algorítmica de complejidad polinómica).

Otro tipo especial de redes neuronales artificiales se ha aplicado en conjunción con los algoritmos genéticos (AG) para crear controladores para robots. La disciplina que trata la evolución de redes neuronales mediante algoritmos genéticos se denomina Robótica Evolutiva. En este tipo de aplicación el genoma del AG lo constituyen los parámetros de la red (topología, algoritmo de aprendizaje, funciones de activación, etc.) y la adecuación de la red viene dada por la adecuación del comportamiento exhibido por el robot controlado (normalmente una simulación de dicho comportamiento).

1.4 Selección de la herramienta de software a utilizar.

En la actualidad existe gran cantidad de programas informáticos que tienen implementados de forma eficiente algoritmos que soportan el trabajo con redes neuronales y árboles de decisión. La selección de cualquiera de ellos necesita considerar un conjunto de indicadores.

Entre los indicadores a considerar hay uno que resulta de mucha actualidad, y es considerar si es software libre, lo que garantiza la legalidad de su uso. Por otro lado convendría que sea multiplataforma, y aunque no por último menos importante debe ser una herramienta que sea de dominio del investigador.

Teniendo en cuenta estos aspectos, fueron consideradas como candidatas: **KEEL, Clementine, KNIME, MatLab y WEKA.**

Entre todas ellas **MatLab y Clementine** fueron las primeras descartadas a pesar de todas sus ventajas porque están bajo una licencia de uso.

KNIME (o Konstanz Information Miner) parece ser según la literatura consultada, una excelente opción, KNIME es una plataforma de minería de datos que permite el desarrollo de modelos en un entorno visual. Está construido bajo la plataforma Eclipse y programado, esencialmente, en Java. Está concebido como una herramienta gráfica y dispone de una serie de nodos (que encapsulan distintos tipos de algoritmos) y flechas (que representan flujos de datos) que se despliegan y combinan de manera gráfica e interactiva.

Los nodos implementan distintos tipos de acciones que pueden ejecutarse sobre una tabla de datos, entre otras se mencionan las siguientes:

- Manipulación de filas, columnas, etc., como muestreos, transformaciones, agrupaciones, etc.
- Visualización (histogramas, etc.).
- Creación de modelos estadísticos y de minería de datos, como árboles de decisión, máquinas de vector soporte, regresiones, etc.
- Validación de modelos, como curvas ROC, etc.
- *Scoring* o aplicación de dichos modelos sobre conjuntos nuevos de datos.

El carácter abierto de la herramienta hace posible su extensión mediante la creación de nuevos nodos que implementen algoritmos a la medida del usuario. Además, existe la posibilidad de llamar directa y transparentemente a Weka y o de incorporar de manera sencilla código desarrollado en R o Python/Jython.

KNIME integra diversos componentes para aprendizaje automático y minería de datos a través de su concepto de fraccionamiento de datos (*data pipelining*) modular. La interfaz gráfica de usuario permite el montaje fácil y rápido de nodos para preprocesamiento de datos (ETL: extracción, transformación, carga), para el análisis de datos y modelado y visualización. KNIME es desde 2006 utilizado en la investigación farmacéutica, pero también se utiliza en otras áreas, como: análisis de datos de cliente de CRM, inteligencia de negocio y análisis de datos financieros. [27].

KNIME es una herramienta de código abierto que puede ser descargada y utilizada gratuitamente bajo los términos de la licencia GPLv3 con una excepción que permite que otros usuarios utilicen el bien definido nodo de API para añadir extensiones de propiedad [28]. Esto también permite a los proveedores comerciales de software agregar envolturas enlazando sus herramientas hacia KNIME.

La compañía que lo desarrolla ofrece, adicionalmente, la posibilidad de contratar servicios de soporte en varios niveles, además de brindar servicios de consultoría y formación.

KEEL (siglas de Knowledge Extraction based on Evolutionary Learning; en español, Extracción de Conocimiento basado en Aprendizaje Evolutivo) es un

conjunto de herramientas de software de aprendizaje automático, desarrollados bajo el proyecto español nacional TIC2002-04036-C05, TIN2005-08386-C05 y TIN2008-06681-C06.

KEEL como KNIME es una herramienta de software no comercial, diseñado para resolver una amplia gama de problemas de minería de datos (regresión, clasificación, asociación, agrupamiento, etc.) mediante la generación de experimentos estándar y educativos, y para evaluar algoritmos evolutivos [29]. Dispone de una amplia librería de algoritmos de Sistemas Difusos Evolutivos basados en diferentes esquemas: Michigan, Pittsburgh, IRL y GCCL. También dispone de módulos de tratamiento de datos y análisis estadísticos [30]. Tiene una colección de bibliotecas para técnicas de procesamiento previo y posterior para manipulación de datos, métodos de computación suave en conocimiento de extracción y aprendizaje y para proporcionar métodos de investigación científica [31].

Los programas implementados se aplican en amplia investigación y metas educativas como: aprendizaje de regla de difusión evolutiva, ajuste de regla de Mamdani, redes neurales artificiales genéticas, y aprendizaje de sistemas de clasificación (Learning Classifier Systems) [32].

Tanto KEEL como KNIME tienen en su contra que no son entornos conocidos por el investigador, y que su uso requería de un tiempo y esfuerzos que desbordan el marco de este trabajo.

Por todo lo anterior, finalmente el seleccionado para el experimento fue el software **WEKA**. Posee múltiples funcionalidades y ventajas que hace que cada vez gane más seguidores.

Es de libre distribución, esto permite actualizar su código fuente para incorporar nuevas utilidades o modificar las ya existentes, lo cual ha impulsado que sea una de las suites más utilizadas en el área durante los últimos años. [33], [34]

Sus ventajas principales son:

- Es fácilmente manejable por el usuario.
- Está disponible libremente bajo la licencia pública general de GNU.
- Está completamente implementado en Java y puede correr en casi cualquier plataforma, por lo que permite una mayor portabilidad.
- Soporta varios procesos como preprocesamiento de datos, clustering, clasificación, regresión, visualización y selección.
- Permite la entrada de datos numéricos, reales, entre otros.
- Proporciona acceso a bases de datos vía SQL gracias a la conexión JDBC (*Java Database Connectivity*) y puede procesar el resultado devuelto por una consulta hecha a la base de datos.

Para su uso se requiere que los datos sujetos a análisis se almacenen con un cierto formato, conocido como ARFF (Attribute-Relation File Format).

Estos deben estar dispuestos en el fichero de la forma siguiente: cada instancia en una fila, y con los atributos separados por comas. El formato de un fichero ARFF (Attribute-Relation File Format) sigue en general, la estructura siguiente:

```
% comentarios
% describe la relación llamada NOMBRE_RELACION
@relation NOMBRE_RELACION
% describe los atributos r1, r2, etc. de tipo real.
@attribute r1 real
@attribute r2 real
...
% describe los atributos i1, i2, etc. de tipo entero.
@attribute i1 integer
@attribute i2 integer
...
% describe los atributos s1, s2, etc. de tipo simbólico.
@attribute s1 {v1_s1, v2_s1,...vn_s1}
@attribute s2 {v1_s1, v2_s1,...vn_s1}
...
% la sección de datos, que se escriben separados por coma,
% en el orden en que fueron declarados los atributos
@data
DATOS
```

Por tanto, los atributos pueden ser principalmente de dos tipos: numéricos de tipo real o entero (indicado con las palabra *real* o *integer* tras el nombre del atributo), y simbólicos, en cuyo caso se especifican los valores posibles que puede tomar entre llaves.

Además de todas las ventajas ya citadas, WEKA es apropiado para los fines de la investigación, y adicionalmente, ha sido usado previamente por el investigador.

1.5 Conclusiones parciales.

La minería de datos y sus técnicas de aprendizaje automático es apropiada para analizar datos con fines de extraer conocimiento, que puede estar en forma de relaciones que constituyen un modelo para los datos. Tanto los árboles de decisión como las redes neuronales soportan modelos predictivos, es decir permiten estimar valores futuros o desconocidos de variables de interés, usando otras variables o campos de la base de datos.

La intención de realizar un pronóstico en la carrera, al que se ha denominado pronóstico académico, puede ser realizado usando RN y AD, entre otros, por lo que se decide su uso, para el trabajo con los datos, y la obtención de resultados, después de la revisión de literatura, se decide usar el WEKA, que soporta ambas técnicas, es de muy fácil uso, y tiene una interface gráfica muy adecuada a los propósitos de la investigación.

Capítulo II: Aprendizaje supervisado y Diseño de Experimentos.

2.1 Introducción

En el presente capítulo se profundiza en los elementos teóricos esbozados en el capítulo anterior, se presentan las herramientas de la Minería de Datos a utilizar en esta investigación, en particular los algoritmos de redes neuronales y de árboles de decisión, ventajas, desventajas y principios de funcionamiento. Se presenta el diseño de experimentos, se detalla el proceso de obtención de los datos a explorar, finalmente se exponen las variables de interés, y se presenta un esquema de cómo se lleva a cabo el análisis de los datos recopilados.

2.2 Introducción al aprendizaje supervisado y la minería de datos.

En el capítulo anterior se define la minería de datos como *el proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos*. La tarea fundamental de la minería de datos es encontrar modelos inteligibles a partir de los datos. Para que este proceso sea efectivo debería ser automático o semiautomático (asistido) y el uso de los patrones descubiertos debería ayudar a tomar decisiones más seguras que reporten, por tanto, algún beneficio. La minería de datos integra fundamentalmente técnicas de aprendizaje automático. [2]

El aprendizaje automático es un área de la inteligencia artificial que se encarga del desarrollo de algoritmos y técnicas que permitan al ordenador "aprender". A un nivel general existen dos tipos de aprendizaje, inductivo y deductivo.

El mayor esfuerzo que se realiza en el área del aprendizaje automático es el de extraer información de los datos de manera automática. El **aprendizaje inductivo** es un tipo especial de aprendizaje que parte de casos particulares (ejemplos) y obtiene casos generales (reglas o modelos) que generalizan o abstraen la evidencia. Este es precisamente el tipo de aprendizaje que se emplea en este trabajo.

El aprendizaje automático tiene una gran variedad de aplicaciones en áreas como el procesamiento del lenguaje natural, diagnósticos, bioinformática, detección de fraudes en tarjetas bancarias, clasificación de cadenas de ADN, reconocimiento del habla y la escritura, reconocimiento de objetos en visión artificial, clasificaciones psicológicas basadas en la personalidad y otros rasgos, etc.

Existen varios tipos de técnicas de aprendizaje automático inductivo que generalmente se clasifican en dos tipos: aprendizaje no supervisado y supervisado, como se mencionó en el capítulo 1.

2.2.1 Aprendizaje no supervisado

El aprendizaje no supervisado, también conocido como *clustering*, es un tipo de aprendizaje inductivo que parte de un conjunto de ejemplos, o instancias, de los que **no se conoce su clase** y se desea encontrar las clases más probables en las que se pueden agrupar dichos ejemplos. Para ello se divide un conjunto de ejemplos en subconjuntos disjuntos de manera que la similitud entre los objetos en el mismo grupo sea alta y la similitud entre objetos pertenecientes a distintos grupos sea baja.

A cada uno de estos grupos se les llama *cluster*. Mientras que algunos algoritmos calculan el número de grupos en los que partir el conjunto de ejemplos, otros necesitan este dato como parámetro de entrada. En este caso las características específicas de cada grupo no se conocen a priori, dando lugar muchas veces a resultados inesperados [35].

En este caso, se conoce la clase a la que pertenece cada ejemplo o instancia, dado que al registrar las evaluaciones obtenidas por cada estudiante se conoce si aprueba o no el año académico en cuestión, por lo que se trata de **aprendizaje supervisado**.

2.2.2 Aprendizaje supervisado

Las técnicas o algoritmos de aprendizaje supervisado tratan de aprender una categoría determinada de los objetos de la evidencia que se denomina clase. Un ejemplo típico es el del *spam*, en este caso los objetos son correos electrónicos y los algoritmos de aprendizaje usados para este problema tratan de aprender qué correos son *spam* y distinguirlos de aquéllos que no lo son. También son llamados algoritmos de clasificación o clasificadores. El nombre de aprendizaje supervisado viene porque para crear un modelo necesitan un grupo de instancias ya clasificadas, a las que se llama conjunto de entrenamiento. Intuitivamente, entrenar un clasificador, significa construir los bordes entre las clases. Los clasificadores asignan objetos a las clases en función de los bordes aprendidos. Los tipos de algoritmos de clasificación son muy variados, entre ellos destacan los probabilísticos (basados en el teorema de Bayes), los basados en reglas, los árboles de decisión, y los que representan algún tipo de función como son la regresión logística o las redes neuronales. [36]

De los dos tipos de aprendizaje se utiliza el aprendizaje supervisado, dado que se desea observar el comportamiento individualmente, más que en grupo.

Dentro de las técnicas de aprendizaje supervisado, entre las más populares, por sus posibilidades de uso en los más disímiles contextos se encuentran las redes neuronales y los árboles de decisión.

En general suelen utilizarse para el reconocimiento de patrones, la compresión de información, la reducción de la dimensionalidad, el agrupamiento, la clasificación, la visualización, etc. En la práctica, a lo largo de los años se han utilizado con mucho éxito para las predicciones, por citar solo algunos ejemplos su aplicación para la predicción de la calidad del aire [37], para el rendimiento de la caña de azúcar[38], en estudios de bibliometría[39], para predecir la quiebra bancaria[40], para pronosticar el egreso de pacientes con trauma de cráneo grave[41], y muchos otros que el lector interesado puede consultar en las referencias [42]–[53]

A continuación detallan las técnicas de redes neuronales artificiales y árboles de decisión, para dejar sentadas las bases de su selección para explorar los datos en este trabajo.

2.3 Redes neuronales artificiales.

Como se apuntó antes las redes neuronales artificiales (RNA) son sistemas conexionistas que dependiendo del tipo de arquitectura neuronal, pueden ser usadas en diferentes aplicaciones, muchas de estas tareas lo son también de la minería de datos. Las RNA se pueden utilizar, por tanto, como una herramienta para llevar a cabo minería de datos.

Dentro de las RNA, se selecciona las que emplean aprendizaje supervisado, y particularmente se profundiza en el estudio del Perceptrón Multicapa.

Desde su surgimiento las neuronas artificiales se consideran dispositivos simples de cálculo que, a partir de un vector de entrada procedente del exterior o de otras neuronas el cual contiene una información determinada (valor, etc.), proporcionan una única respuesta o salida. En general, una red neuronal artificial (RNA) es un conjunto de estas neuronas artificiales que se puede traducir como un número determinado de procesadores simples que se comunican con otros por medio de conexiones ponderadas. Puede decirse que son sistemas que hacen uso de los principios y formas que imitan simplificadaamente la organización del cerebro humano.

Las RNA al margen de "parecerse" al cerebro presentan una serie de características propias del cerebro. Por ejemplo, aprenden de la experiencia, generalizan de ejemplos previos a ejemplos nuevos y abstraen las características principales de una serie de datos.

- Aprender: adquirir el conocimiento de una cosa por medio del estudio, ejercicio o experiencia. Las RNA pueden cambiar su comportamiento en función del entorno. Se les muestra un conjunto de entradas y ellas mismas se ajustan para producir unas salidas consistentes.

- Generalizar: extender o ampliar una cosa. Las RNA generalizan automáticamente debido a su propia estructura y naturaleza. Estas redes pueden ofrecer, dentro de un margen, respuestas correctas a entradas que presentan pequeñas variaciones debido a los efectos de ruido o distorsión.
- Abstraer: aislar mentalmente o considerar por separado las cualidades de un objeto. Algunas RNA son capaces de abstraer la esencia de un conjunto de entradas que aparentemente no presentan aspectos comunes o relativos[42].

La forma en que desarrollan su trabajo es esencialmente distinta de la utilizada por las computadoras convencionales. Los procesadores microscópicos del cerebro (neuronas) operan en paralelo y presentan cualitativamente más ruido que los elementos que forman a las computadoras. No ejecutan un programa fijo con base en un conjunto previamente especificado [45], [46], [54].

La capacidad de aprendizaje adaptativo es una de las características más atractivas de las redes neuronales. Esto es, aprenden a llevar a cabo ciertas tareas mediante un entrenamiento con ejemplos ilustrativos. Como las redes neuronales pueden aprender a diferenciar patrones mediante ejemplos y entrenamientos, no es necesario elaborar modelos a priori ni necesidad de especificar funciones de distribución de probabilidades [55].

Existen alrededor de cincuenta tipos de modelos de RNA, de los cuales citamos los más usados en las aplicaciones prácticas: Perceptrón, Perceptrón Multicapa, Mapa de Kohonen, ADALINE, MADALINE, Hopfield, Máquina de Boltzman, etc.

Para esta investigación se emplea el modelo Perceptrón Multicapa (MLP).

2.3.1 Perceptrón Multicapa (MLP)

Uno de los primeros modelos de redes neuronales fue el Perceptrón (PS), el cual surge en el año 1962 y su creación se debe a Rosenblatt. El Perceptrón es un modelo unidireccional formado por dos capas, una de entrada y otra de salida. La función de activación de esta neurona de salida es de tipo escalón. Este modelo en su algoritmo de aprendizaje determina de manera automática el peso sináptico que clasifica un subconjunto de patrones a partir de un conjunto de ejemplos etiquetados, por lo que puede utilizarse como clasificador, además precisa los pesos en proporción a la diferencia existente entre la salida actual de la red y la salida que se quiere obtener, con el objetivo de minimizar el error actual de la red. [41].

El PS tiene la desventaja que solo se basa en funciones linealmente separables. Esta deficiencia motivó a Minsky y Papert a la investigación del Perceptrón y en el año 1969 proponen la idea de adicionarle a la red capas ocultas, surgiendo así un nuevo modelo de red neuronal artificial, la arquitectura del Perceptrón Multicapa (MLP por sus siglas en inglés: MultiLayer Perceptron).

Hoy en día es posible mostrar que muchos conjuntos de datos que no son linealmente separables pueden ser modelados mediante el empleo del Perceptrón Multicapa, es decir una red neuronal en forma de cascada, que tiene una o más capas ocultas.

Aunque esta potencialidad del MLP se descubrió pronto, se tardó bastante tiempo en encontrar un método o regla de aprendizaje apropiada para construidas a partir de ejemplos. Esta regla parece que fue descubierta de manera independiente varias veces, y no existe acuerdo de la fecha exacta ni de su descubridor, pero fue popularizada principalmente por el Grupo POP (*Parallel Distributed Processing*)[56], bajo el nombre de Retropropagación o Propagación hacia atrás.

Respecto al uso de la red o de la activación, la activación se propaga en la red a través de los pesos desde la capa de entrada hacia la capa intermedia donde se aplica alguna función de activación a las entradas que le llegan. Entonces la activación se propaga a través de los pesos hacia la capa de salida.

Por tanto, si se piensa en el aprendizaje, hay que actualizar dos conjuntos de pesos: aquellos entre la capa oculta o intermedia y la de salida, y aquellos entre la capa de entrada y la capa intermedia. El error debido al primer conjunto de pesos se calcula empleando el método del error cuadrático medio. Entonces se propaga hacia atrás la parte del error debido a los errores que tienen lugar en el segundo conjunto de pesos y se asigna el error proporcional a los pesos que lo causan.

Se puede utilizar cualquier número de capas ocultas que queramos ya que el método es bastante general. Sin embargo, un factor a tener en cuenta es normalmente el tiempo de entrenamiento, el cual puede ser excesivo para arquitecturas con muchas capas. Además, se ha demostrado que redes con una única capa oculta son capaces de aproximar cualquier función continua (o incluso cualquier función con sólo un número finito de discontinuidades), en el caso de utilizar funciones de activación diferenciables (no lineales) en la capa oculta.

2.3.1.1 Algoritmo de retropropagación.

El algoritmo de retropropagación es el primer algoritmo eficaz para el aprendizaje de MLP.

El algoritmo consiste en minimizar un error (comúnmente cuadrático) por medio de descenso de gradiente, por lo que la parte esencial del algoritmo es el cálculo de las derivadas parciales de dicho error con respecto a los parámetros de la red neuronal.

Funciones de activación

En el algoritmo no se especifica la función de activación utilizada, Las funciones de activación más populares son la función sigmoide

$$f(x) = \frac{1}{1 + e^{-x}}$$

y la función tangente hiperbólica, $\tanh(x)$, ambas diferenciables.

Momento y velocidad de convergencia

El método de retropropagación básico no es conocido por su rapidez de convergencia. Si simplemente incrementamos la velocidad de aprendizaje, el efecto será el de introducir inestabilidad en la regla de aprendizaje causando oscilaciones violentas en los pesos aprendidos. Es posible acelerar el método básico de varias maneras. Lo más simple es aplicar un término, denominado momento, al cambio de los pesos. La idea básica es incrementar si el nuevo cambio de pesos coincide con la dirección de los cambios anteriores y se reduce en caso contrario.

El algoritmo se puede variar ligeramente mediante el cambio de estos y otros parámetros como el del número de neuronas ocultas.

Se analiza con mayor detalle como hace esto el algoritmo implementado en Weka, herramienta que es la seleccionada para la exploración de los datos en este trabajo.

2.3.1.2 Variaciones de la retropropagación en Weka.

Uno de los aspectos más complejos es determinar el número de neuronas ocultas, ya que pocas neuronas pueden hacer que el concepto a aprender no se pueda capturar mientras que demasiadas neuronas pueden sobreajustar los datos. Existen variantes del algoritmo de retropropagación que van añadiendo neuronas ocultas a medida que se van necesitando, o ajustan el tamaño por un análisis de los datos, o bien siguen reglas más o menos sencillas para hacerlo. Muchas implementaciones en paquetes de minería de datos incluyen estas variaciones, ya que son mucho más sencillas de cara al usuario.

El algoritmo de retropropagación (**WEKA.classifiers.neural.NeuralNetwork**) del sistema WEKA permite, entre otras opciones la selección:

- del momento ("*momentum*"),
- de la velocidad de convergencia (mediante los parámetros "*learning rate*" y "*decay*"),
- el número de *épocas* de aprendizaje ("*trainingTime*"),
- la inicialización de la semilla aleatoria ("*randomSeed*"), etc.

Todos estos parámetros tienen valores por defecto y un usuario novel sólo debería modificarlos en el caso de que no obtenga buenos resultados con ellos. Quizá, lo más curioso son las reglas y opciones que utiliza para determinar el número de nodos en la capa oculta. Aunque este valor se puede especificar numéricamente de una manera directa, también se puede expresar según el número de atributos y de clases, ya que se ha comprobado que, en muchas ocasiones, la cantidad de nodos necesarios depende de estos factores. Por ejemplo, algunas de estas opciones especiales en la versión 3.2 y superiores de WEKA son:

- 'a', que genera un número de neuronas ocultas igual a $(at + el) / 2$, donde at es el número de atributos (entradas) y el es el número de valores de la clase,
- 'i', que genera un número de neuronas ocultas igual a (at) ,
- 'o', que genera un número de neuronas ocultas igual a (el) y
- 't', que genera un número de neuronas ocultas igual a $(at + el)$.

Pueden parecer unas reglas muy triviales, pero la configuración por defecto, la 'a', suele dar buenos resultados en general.

Respecto a Weka como una herramienta general que incorpora técnicas de redes neuronales, es muy recomendable para los principiantes, ya que permite utilizar parámetros y topologías por defecto, con resultados muy aceptables. Las herramientas específicas son más recomendables cuando ya se tiene un mayor conocimiento y experiencia sobre redes neuronales y se desea obtener todo su potencial, mediante una elección precisa de los algoritmos, las topologías y los parámetros.[57]. Esta es otra de las razones que avalan la selección de Weka para este trabajo.

2.4 Árboles de Decisión.

Los árboles de decisión construyen un "modelo comprensible", "hipótesis" o "representación" de la regularidad existente en los datos.

El término "comprensible" hace referencia al hecho de que estos modelos se pueden expresar de una manera simbólica, en forma de conjunto de condiciones (a diferencia de otros métodos, como las redes neuronales) y, por tanto, pueden tener como resultado modelos inteligibles para los seres humanos (y también para sistemas semi-automáticos que procesen reglas).

De todos los métodos de aprendizaje, los sistemas de aprendizaje basados en árboles de decisión son quizás el método más fácil de utilizar y de entender. Un árbol de decisión es un conjunto de condiciones organizadas en una estructura jerárquica, de tal manera que la decisión final a tomar se puede determinar siguiendo las condiciones que se cumplen desde la raíz del árbol hasta alguna de sus hojas. Los árboles de decisión se utilizan desde hace siglos, y son especialmente apropiados para expresar procedimientos médicos, legales, comerciales, estratégicos, matemáticos, lógicos, etc.

Una de las grandes ventajas de los árboles de decisión es que, en su forma más general, las opciones posibles a partir de una determinada condición son excluyentes. Esto permite analizar una situación y, siguiendo el árbol de decisión apropiadamente, llegar a una sola acción o decisión a tomar.

Lo que interesa es construir árboles de decisión a partir de los datos. Para ello existen métodos y algoritmos de *aprendizaje* de árboles de decisión, y es importante conocer cómo funcionan y, fundamentalmente, cómo utilizados convenientemente para obtener modelos precisos, fiables y, en última medida, inteligibles.

Sistemas por partición: árboles de decisión para clasificación

La tarea de aprendizaje para la cual los árboles de decisión se adecuan mejor es la clasificación. De hecho, clasificar es determinar de entre varias clases a qué clase pertenece un objeto; la estructura de condición y ramificación de un árbol de decisión es idónea para este problema. La característica más importante del problema de la clasificación es que se asume que las clases son *disjuntas*, es decir, una instancia es de la clase *a* o de la clase *b*, pero no puede ser al mismo tiempo de las clases *a* y *b*.

Un ejemplo de clasificación sería determinar si un estudiante aprueba o no, o es una cosa o la otra. En este sentido, la clasificación se diferencia de la categorización, donde se permiten más de una clase, etiqueta o categoría para cada instancia.

Debido al hecho de que la clasificación trata con clases o etiquetas disjuntas, un árbol de decisión conducirá un ejemplo hasta una y sólo una hoja, asignando, por tanto, una única clase al ejemplo. Para ello, las particiones existentes en el árbol deben ser también disjuntas. Es decir, cada instancia cumple o no cumple una condición.

Además, dicha propiedad es exhaustiva, es decir, una de las dos condiciones se debe cumplir.

Esta propiedad dio lugar al esquema básico de los primeros algoritmos de aprendizaje de árboles de decisión; el espacio de instancias se iba *partiendo* de arriba a abajo, utilizando cada vez una partición, es decir, un conjunto de condiciones excluyentes y exhaustivas.

Estos algoritmos se llaman algoritmos de *partición* o algoritmos de "divide y vencerás". Otra característica importante de los primeros algoritmos de aprendizaje de árboles de decisión es que una vez elegida la partición dicha partición no se podía cambiar, aunque más tarde se pensara que había sido una mala elección. Por tanto, *uno* de los aspectos más importantes en los sistemas de aprendizaje de árboles de decisión es el denominado *criterio de partición*, ya que una mala elección de la partición (especialmente en las partes superiores del árbol) generará un peor árbol.

Simplemente, el algoritmo va construyendo el árbol (desde el árbol que sólo contiene la raíz) añadiendo particiones y los hijos resultantes de cada partición. Lógicamente, en cada partición, los ejemplos se van dividiendo entre los hijos. Finalmente, se llega a la situación en la que todos los ejemplos que caen en los nodos inferiores son de la misma clase y esa rama ya no sigue creciendo. La única condición que hay que exigir es que las particiones al menos separen ejemplos en distintos hijos, con lo que la cardinalidad de los nodos irá disminuyendo a medida que se desciende en el árbol.

Luego lo más importante para que un algoritmo funcione bien es considerar:

- Las particiones.
- Criterio de selección de particiones.

Esto es lo que diferencia fundamentalmente los distintos algoritmos de "partición" existentes hasta la fecha, como CART[58], ID3 [59], [60], C4.5[61], ASSISTANT[62], etc.

Basándose en diferentes particiones, en un criterio de partición y otras extensiones, han aparecido numerosos algoritmos o sistemas de aprendizaje de árboles de decisión. Entre los más populares se encuentran, entre otros:

- CART[58] y derivados: son métodos "divide y vencerás" que construyen árboles binarios y se basan en el criterio de partición GINI y que sirve tanto para clasificación como para regresión. La poda se basa en una estimación de la complejidad del error ("error-complexity"). Generalmente se pueden encontrar en paquetes de minería de datos con el nombre C&RT.
- ID3[59], [60], C4.5[61] y derivados (Assistant[62]): son métodos "divide y vencerás" y están basados en criterios de partición derivados de la ganancia (GainRatio). Tienen poda basada en reglas u otros mecanismos más sofisticados. Contiene métodos de colapsado de ramas y muchas otras mejoras. Existe un libro que describe concienzudamente la implementación y características del C4.5[61], la implementación se puede descargar gratuitamente o se encuentra en numerosas librerías. Por ejemplo, una versión más avanzada, la J4.8, se distribuye con la librería WEKA. Existe, además, una versión comercial presumiblemente mejor que las anteriores denominada C5 (See5) y comercializada por el propio Quinlan directamente (RuleQuest) o a través de paquetes de minería de datos (Clementine).
- IND[63], LMDT [64] y otros sistemas híbridos: incorporan características de varios sistemas o añaden otras técnicas de aprendizaje en la construcción de árboles de decisión: regresión lineal, perceptrones, etc.

Como dijimos en la introducción de este capítulo, toda esta variabilidad es consecuencia del éxito de los árboles de decisión y los sistemas de reglas en el aprendizaje automático y muy especialmente en la minería de datos.

Las ventajas principales de las técnicas vistas son:

- Son aplicables a varias tareas de minería de datos: clasificación, regresión, agrupamiento y estimación de probabilidades.
- Tratan con atributos numéricos (continuos) y nominales (discretos).
- Muchos de ellos son eficientes y existen variantes escalables a grandes volúmenes de datos (tanto para muchos atributos como para muchos ejemplos).
- Son fáciles de usar.
- Son tolerantes al ruido, a atributos no significativos y a valores faltantes.
- Existen multitud de sistemas, muchos de ellos gratuitos.

Lógicamente, también tienen algunas desventajas. En general, no son tan precisos como otros métodos, como pueden ser las redes neuronales o las

máquinas de vectores de soporte. Además, son *débiles* (se conocen como *weak learners*), en el sentido de que, debido a su carácter voraz, son bastante dependientes de la muestra de ejemplos; dos muestras distintas sobre la misma distribución pueden dar dos árboles bastante diferentes. No obstante, es precisamente esta "debilidad" o "variabilidad" la que permite que los árboles de decisión sean frecuentemente utilizados con técnicas de combinación, como el *boosting* y el *bagging*.

2.5 Pronóstico.

La tecnología nos ha permitido capturar y almacenar grandes cantidades de datos. Encontrar patrones, tendencias y anomalías en los conjuntos de datos, y resumirlos con simples modelos cuantitativos es uno de los grandes cambios de la era de la información: convertir los datos en información y la información en conocimiento.

Hoy se encuentran libros como el Data Mining: Practical Machine Learning Tools and Techniques[65], que hacen una excelente tarea explicando cómo preparar los datos, construir los modelos, y cómo evaluar las cualidades predictivas del modelo. Presentan además una amplia variedad de modelos, como son árboles de decisión, reglas de asociación, modelos lineales, clusters, redes neuronales, etc. Y todo ello sobre una base práctica, al discutir los principales conceptos de la minería de datos apoyado en la utilización de algoritmos que han sido implementados en Weka.

Esto permite adquirir descripciones estructurales desde los ejemplos (instancias) que pueden ser usadas para la predicción, la explicación y la comprensión. Algunas aplicaciones se enfocan sobre la predicción: un pronóstico de lo que sucederá en nuevas situaciones, desde los datos que describen lo que sucedió en el pasado. Pero igualmente, el resultado del aprendizaje es una descripción real de una estructura que puede ser usada para clasificar ejemplos, y esta descripción soporta la explicación, la comprensión y la predicción. Con este sentido se utilizan en este trabajo las redes neuronales y los árboles de decisión.

No obstante, el Weka, a partir de su versión 3.2, incorpora una herramienta que permite realizar pronósticos –al estilo de cómo lo entienden los modelos de la estadística-, sobre variables de tipo numérico[65]. Resulta interesante también obtener para los atributos de mayor peso que arroje nuestro modelo, el pronóstico de su comportamiento futuro teniendo en cuenta su comportamiento pasado.

En este sentido se usa el algoritmo SMOreg y se especifican algunos parámetros necesarios para la configuración. Entre ellos:

- Número de unidades de tiempo a pronosticar.
- Periodicidad.
- Intervalo de confianza.

- Métrica de evaluación.

En el diseño del experimento se especifican los valores usados y las variables estudiadas con fines de pronóstico.

2.6 Diseño de Experimento.

2.6.1 Datos a utilizar.

Se dispone de registros de los resultados académicos de los estudiantes de Ingeniería Informática durante los cursos 2010-11, 2011-12, 2012-13 y 2013-14 en la Universidad de Cienfuegos.

Se seleccionan para explorar los datos de los estudiantes de 1er. y 2do. Años de la carrera.

La base de casos tiene un total de 114 estudiantes para el primer año, y para el segundo un total de 120. De cada caso se tienen 16 variables en el caso de primer año y en el de segundo 18, que constituyen todos aquellos factores considerados como de posible riesgo para que el estudiante presente dificultades académicas. Concretamente se registran los resultados que alcanzan en cada una de las asignaturas para cada año.

Adicionalmente se considera un modelo combinado, que integra las asignaturas de los dos cursos objetos de interés para el estudio (33 variables), y un total de 134 instancias. Para la obtención de este conjunto de datos, se utilizan los datos de los cursos mencionados y las matrículas de 1er y 2do año del curso 2009-10, y una parte de la matrícula del primero año del curso 2008-09.

Estas variables fueron elegidas según los resultados registrados en Secretaría Docente de la Facultad de Ingeniería.

A continuación se presenta la definición de las variables para cada modelo.

Para el Modelo Primer Año.

Variable	Tipo	Operacionalización.	
		Escala	Descripción
Mat I	Númérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado) 0(eximido)	Resultado alcanzado en Matemática I.
FyS	Númérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Filosofía y Sociedad.

		0(eximido)	
FInf	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Fundamentos de Informática.
IP	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Introducción a la Programación.
EFI	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Educación Física I.
IdEx I	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado) 0(eximido)	Resultado alcanzado en Idioma Extranjero I.
Alg y GA	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Álgebra y Geometría Analítica.
HCuba	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Historia de Cuba.
IdExt II	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado) 0(eximido)	Resultado alcanzado en Idioma Extranjero II.
DPOO	Numérico	2(mal) 3(regular) 4(bien) 5(excelente)	Resultado alcanzado en Diseño y Programación Orientada a Objetos.

		1(convalidado)	
EF II	N Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Educación Física II.
IGS	N Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Ingeniería y Gestión de Software.
MD	N Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Matemática Discreta.
Mat II	N Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Matemática II.
EPCap	N Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado) 0(eximido)	Resultado alcanzado en Economía Política del Capitalismo.
Aprueba	N Nominal	Si(limpio) No(causa baja, repite o arrastra)	Determina si un estudiante aprueba el año(limpio) y/o desaprueba (causando baja, repite o arrastra)

Tabla 1 Atributos del modelo 1er año

Para el Modelo Segundo Año.

Variable	Tipo	Operacionalización.	
		Escala	Descripción
Mat III	N Numérico	2(mal) 3(regular) 4(bien)	Resultado alcanzado en Matemática III

		5(excelente) 1(convalidado) Np(no presentado)	
EPCAP	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado) 0(eximido)	Resultado alcanzado en Econ. Pol. Cap.
ED1	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en E. Datos 1
IIA	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Introducción a la IA.
IDEX3	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Idioma Ext. III.
SegNac	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado) 0(eximido)	Resultado alcanzado en Seguridad Nacional.
EF3	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Educación Física III
Elec1	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Electiva I.
MatIV	Numérico	2(mal) 3(regular)	Resultado alcanzado en

		4(bien) 5(excelente) 1(convalidado) 0(eximido)	Matemática III
Acomp	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Arquitectura de Computadoras
ED2	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en E. Datos 2
BD	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Bases de Datos
TSP	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Teoría Sociopolítica.
Idext4	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado)	Resultado alcanzado en Idioma Ext. IV.
DNac	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado) 0(eximido)	Resultado alcanzado en Defensa Nacional..
EFiv	Numérico	2(mal) 3(regular) 4(bien) 5(excelente) 1(convalidado) 0(eximido)	Resultado alcanzado en Educación Física IV.
PProf	Numérico	2(mal) 3(regular)	Resultado alcanzado en

		4(bien) 5(excelente) 1(convalidado) 0(eximido)	Práctica Profesional de 2do. Año.
Aprueba	Nominal	Si(limpio) No(causa baja, repite o arrastra)	Determina si un estudiante aprueba el año(limpio) y/o desaprueba (causando baja, repite o arrastra)

Tabla 2 Atributos del modelo 2do año

Para el Modelo Primero y Segundo Año.

Es la unión de los modelos Primero y Segundo, en un modelo combinado que contendrá las variables de ambos modelos (16 +18 -1) dado que la variable objetivo Aprueba que aparece en ambos se utiliza una sola vez en este modelo, y la intersección de las instancias de ambos, más los del curso 2009-10 y un subconjunto de la matrícula del primer año del curso 2008-09.

2.6.2 Diseño del experimento.

Para el trabajo de exploración con los datos se confeccionan 3 modelos iniciales, uno para el primer año de la carrera, uno para el segundo año de la carrera y un tercer modelo donde se combinan los resultados del primer y segundo año de la carrera.

En el caso de los modelos de primer año y segundo año respectivamente el objetivo es determinar los factores de riesgo por cada año que determinan el aprobado o no, obteniendo un pronóstico académico en cada caso, de aquellos factores que resulten de mayor peso (importancia).

La idea de unir en un solo modelo ambos años, es explorar en el ciclo básico qué factores, de conjunto podrían influir en el rendimiento académico de un estudiante de nuevo ingreso, en un período de tiempo mayor.

Se procede con cada modelo de la siguiente forma:

- Se aplica selección de rasgos para depurar los rasgos más significativos en el modelo.
- Se entrena una RN utilizando varias configuraciones, y se observan los resultados.
- Se comparan los resultados de la RN escogiendo aquella que mejor resultados arroja en ambos modelos.
- Se obtiene un árbol de decisión para cada modelo.
- Se comparan los resultados de ambos métodos, RN o AD y se selecciona el mejor en cada caso.

- Se realiza un pronóstico para los factores de mayor peso, obtenidos con el método de AD.

2.6.2.1. Conjuntos de entrenamiento.

Los conjuntos de entrenamiento utilizados en el experimento se pueden consultar en el Anexo A: Bases de Casos en Weka.

- Modelo Primer Año.
- Modelo Segundo Año.
- Modelo Primer y Segundo Año.

2.6.2.2. Topología Propuesta para la Red Neuronal.

Se utiliza un Perceptrón Multicapa con 16 neuronas en la capa de entrada correspondientes a cada una de las variables del problema, para el caso del modelo de primer año, y con 18 neuronas en el caso de segundo año, en el primer momento.

Se seleccionan los rasgos más representativos, y se repite el entrenamiento.

En la capa oculta se utilizan diferentes variantes, y en la capa de salida, dos neuronas, que determinan si el estudiante aprueba o no.

En la siguiente tabla se muestran distintas configuraciones de la red utilizadas para probar todos los modelos.

Se realizaron 7 combinaciones de parámetros con el fin de obtener diferentes modelos, y seleccionar de entre todos, el que mejor resultados arroje.

Parámetros	V 1	V2	V3	V4	V5	V6	V7
GUI	F	F	F	F	F	F	F
autoBuild	V	V	V	V	V	V	V
Debug	F	F	F	F	F	F	F
hiddenLayers	2	i	o	2	2	2	2
No. neuronas en hL.	12,10			12,10	40,20	50,30	20,23
learningRate	0.3	0.7	0.7	0.3	0.7	0.7	0.7
Momentum	0.2	0.2	0.2	0.2	0.2	0.2	0.2
NuméricoToBinaryFilter	V	V	V	V	V	V	V
normalizeAttributes	V	V	V	V	V	V	V
normalizeNumericClass	V	V	V	V	V	V	V
Reset	V	V	V	V	V	V	V
randomSeed	0	0	0	0	0	0	0
trainingTime	500	500	500	500	500	500	500
validationThreshold	20	20	20	20	20	20	20
No. Folds	3	3	3	5	3	3	3

Tabla 3 Variantes utilizadas para las redes MLP

2.6.2.3. Topología Propuesta para Árboles de decisión.

Para el entrenamiento del árbol de decisión se selecciona el algoritmo J48, y se especifican los siguientes parámetros:

- **binarySplits:** false, (en la construcción de los árboles si se utiliza “splits” binario se divide los atributos nominales).
- **confidenceFactor:** 0.25, como factor de confianza utilizado para la poda (valores más pequeños incurrir en más de poda).
- **debug:** false, pues no se precisa información adicional a la consola.
- **minNumObj:** 2, como número mínimo de instancias por hoja.
- **nunFolds:** 3
- **reducedErrorPruning:** false, si se reduce el error de poda se utiliza en lugar de la poda C.4.5.
- **saveInstanceData:** false.
- **seed:** 1.
- **poda:**
- **subtreeRaising:** true.
- **unpruned:** false
- **useLaplace:** false

Se trabaja con los modos de evaluación cross validation y training set. El modo training set puede ser útil dado que nos dice cuán bueno podría llegar a ser el modelo ante datos nuevos.

2.6.2.4. Pronóstico.

Como resultado del entrenamiento de los árboles de decisión se obtienen tanto en primero como en segundo años, un grupo de variables que resultan significativas, y que se escogen para realizar un pronóstico sobre la base de la siguientes configuraciones.

2.6.2.4.1 Estableciendo el “número de unidades a pronosticar”

Se utiliza el valor 15, refiriéndose a la cantidad de estudiantes a pronosticar según resultados históricos con los que se cuenta.

2.6.2.4.2 Estableciendo la “periodicidad”.

En nuestro ejemplo sería anual, basado en el atributo curso académico. Se deja en blanco porque está relacionado con el anterior.

2.6.2.4.3 Definiendo el intervalo de confianza.

Se establece 95% de confianza, lo cual nos dice que el 95 % de nuestros datos caen en el intervalo cuando se utilizan el cálculo de intervalos de confianza sobre las predicciones que el sistema hace.

2.7 Conclusiones parciales

En este capítulo se han establecido formalmente las bases teóricas de los métodos a utilizar para explorar los datos, apoyado en las bondades de las redes neuronales y los árboles de decisión como métodos de aprendizaje supervisado y predictivos, y se especifican las condiciones para la realización del pronóstico sobre determinados atributos. Se presenta un esquema para el diseño del experimento, desde la preparación de los datos hasta la determinación de los parámetros a utilizar en cada método a emplear.

Capítulo III Análisis de los resultados.

3.1 Introducción.

En este capítulo se presentan los principales resultados de la experimentación, y el análisis que de ellos se derivan. Se selecciona el mejor método para los datos.

3.2 Selección de rasgos

La selección de rasgos consiste en encontrar el subconjunto de atributos del conjunto de datos original que mejor describe los objetos del dominio; tiene como meta reducir la dimensión del conjunto de rasgos a través de la selección del subconjunto de rasgos de mejor desempeño bajo algún criterio de clasificación, durante este proceso se eliminan rasgos que podrían resultar irrelevantes y redundantes proporcionando así una mejor representación de la información original reduciendo significativamente el costo computacional y contribuyendo a una mejor generalización del algoritmo de aprendizaje.

Una vez que las bases de datos se han convertido a formato **arff**, con fines de trabajarlas en WEKA están listas para realizar una selección de rasgos.

Este paso es de gran importancia ya que puede mejorar el rendimiento de las técnicas a emplear, permitiendo exceptuar o suprimir atributos no determinantes en el proceso de clasificación, dejando solo aquellos estrictamente necesarios.

Este proceso de selección de rasgo se les aplicó a los modelos:

- ◆ Modelo 1er año.
- ◆ Modelo 2do año.
- ◆ Modelo 1er año y 2do año.

Sobre cada uno de los modelos se aplicó más de un evaluador, más de un método de búsqueda y dos modos de evaluación, tal como se describe en las siguientes tablas. Para cada modelo se especifica la cantidad de atributos que resultaron del método y se referencia en la última columna el anexo donde puede verse el resultado en detalle. Todos los anexos de este epígrafe se agrupan bajo el título **Anexo B**.

Modelo 1er Año.

Bases de Datos	Evaluador de atributo	Método de búsqueda	Modo Eval.	Cantidad de atributos seleccionados	Anexo
Modelo 1er año. Arff	CfsSubsetEval	BestFirst -D 1 -N 5	10 Folds CV	15	B11
			Training set	4 (1,3,13,14)	B12
		FCBFSearch	10 Folds CV	15	B13
			Training set	15	B14
	InfoGainAttribute Eval	Ranker	10 Folds CV	15	B15
			Training set	15	B16

Tabla 4 Métodos aplicados al modelo 1er año.

De la aplicación de la selección de rasgos en el Modelo de 1er Año se obtiene:

- Con diferentes evaluadores y métodos de búsqueda, que solo en uno de los seis casos se eliminan atributos. En el resto de los casos resultan seleccionados los 15 atributos. El resultado de la aplicación de los modelos puede ser consultado en el Anexo B (del B11 al B16).
- La combinación de CfsSubsetEval con BestFirst y modo de evaluación Training Set (ver Anexo B12) selecciona solo 4 atributos: Mat I, FInf, MD y Mat II. Ciertamente, estos cuatro atributos tienen importancia en el primer año de la carrera pero no son los únicos, de los 15 atributos a discriminar quedan fuera de esta selección atributos como IP y DPOO que sin lugar a dudas, por la experiencia acumulada en el colectivo de año y por valoración de los profesores de estas asignaturas, así como por entrevistas a los estudiantes tienen mucho mayor peso que el que este método les confiere. Se compara este resultado con los ofrecidos por los algoritmos InfoGainAttributeEval con Ranker en ambos modos de evaluación y se comprueba lo aportado por los expertos y las entrevistas realizadas.(Ver Anexo B15 y B16)

Se decide por el equipo investigador lo siguiente:

- **Primero:** aceptar el **Modelo de 1er Año** con todos los atributos, teniendo en cuenta el resultado dado por 5 de los 6 métodos aplicados, y
- **Segundo:** probar con un segundo modelo para el 1er Año que incluya los atributos de mayor peso (según el InfoGainAttributeEval con Ranker con ambos modos de evaluación – anexos B15 y B16) al que se denomina **Modelo_1er Año_sel_rasg** que contiene los 10 atributos más importantes según el InfoGainAttributeEval con Ranker y modo de evaluación cross validation con 10 folds. Resultan seleccionados para este modelo los atributos: **Mat I, Mat II, FyS, MD, IP, DPOO, FInf, IdEx II, Alg y GA e IGS**, listados aquí según el orden que les confiere este modelo. –ver Anexo B15)

Modelo 2do Año.

Bases de Datos	Evaluador de atributo	Método de búsqueda	Modo Eval.	Cantidad de atributos seleccionados	Anexo
Modelo 2do año. Arff	CfsSubsetEval	BestFirst -D 1 -N 5	10 Folds CV	17	B21
			Training set	6 (1,3,4,10,11,12)	B22
		FCBFSearch	10 Folds CV	9	B23
			Training set	9	B24
	InfoGainAttributeEval	Ranker	10 Folds CV	17	B25
			Training set	17	B26

Tabla 5 Métodos aplicados al modelo 2do año.

El resultado de la aplicación de los modelos puede ser consultado en el Anexo B (del B21 al B26).

De la aplicación de la selección de rasgos en el Modelo de 2do Año se obtiene; con diferentes evaluadores y métodos de búsqueda, que:

- Cuando se utiliza CfsSubsetEval con BestFirst los resultados entre el modo de evaluación cross validation y training set difieren, obteniéndose en el primer caso el total de los atributos y en el segundo solo seis de ellos, siendo esta selección sometida a criterio de expertos. Se observa que en esta selección es eliminado el rasgo **Mat IV** que es considerado por todos como importante, tampoco incluye la **PProf**. (ver Anexos B21 y B22)
- Cuando se utiliza CfsSubsetEval con FCBFSearch los resultados convergen con ambos modos de evaluación, resultando seleccionados nueve rasgos en total, que incluye además de los seis del método anterior los rasgos **Mat IV**, **PProf** y **aprueba** que es el rasgo objetivo. (ver Anexos B23 y B24)
- Cuando se utiliza InfoGainAttributeEval con Ranker, se obtienen todos los rasgos ordenados por su importancia, para ambos modos de evaluación. Coincidentemente en este caso los primeros rasgos ordenados por este método coinciden con el resultado que ofrece el método anterior. (ver Anexos B25 y B26)

Teniendo en cuenta lo explicado antes, se decide:

- Seleccionar como rasgos para **Modelo_2do_año_sel_rasgos** los siguientes: **mat iii**, **ed1**, **iia**, **mat iv**, **Acomp**, **ed2**, **bd**, **PProf** además del rasgo objetivo **Aprueba**. Estos son los rasgos que aporta el método CfsSubsetEval con FCBFSearch en ambos modos de evaluación.
- De los rasgos seleccionados antes se observa que los atributos eliminados por los algoritmos usados son los correspondientes a las asignaturas **Dnac**, **epcap**, **idex3**, **segNac**, **TSP**, **EF 3**, **EF iv**, **idex 4** y **elec1**. Estos rasgos corresponden con las asignaturas del ciclo de humanidades y cultura física que contemplan todos los planes de estudio de todas las carreras universitarias en el país, por lo que se decide con fines comparativos utilizar ambos modelos, el modelo original con todas las asignaturas del año (**Modelo 2do Año**), y el modelo que se obtiene como resultado de la selección de rasgos.

Modelo 1er y 2do Año.

Bases de Datos	Evaluador de atributo	Método de búsqueda	Modo Eval.	Cantidad de atributos seleccionados	Anexo
Modelo 1er y 2do año. Arff	CfsSubsetEval	BestFirst -D 1 - N 5	10 Folds CV	33	B31
			Training set	15	B32
		FCBFSearch	10 Folds CV	33	B33
			Training set	33	B34
	InfoGainAttributeEval	Ranker	10 Folds CV	33	B35
			Training set	33	B36

Tabla 6 Métodos aplicados al modelo 2do año.

El resultado de la aplicación de los modelos puede ser consultado en el Anexo B (del B31 al B36).

De la aplicación de la selección de rasgos en el Modelo conjunto de 1er y 2do Año se obtiene; con diferentes evaluadores y métodos de búsqueda, que:

- Con diferentes evaluadores y métodos de búsqueda, solo en uno de los seis casos se eliminan atributos. En el resto de los casos resultan seleccionados los 33 atributos. El resultado de la aplicación de los modelos puede ser consultado en el Anexo B (del B31 al B36).
- La combinación de CfsSubsetEval con BestFirst y modo de evaluación Training Set (ver Anexo B32) selecciona solo 15

atributos de este modelo combinado, estos atributos tienen importancia durante los dos años de la carrera, pero llama la atención que incluye en esta selección a **idex 3 y EF 3**. Ciertamente no son considerados por los expertos como entre los significativos.

- Los algoritmos InfoGainAttributeEval con Ranker en ambos modos de evaluación aceptan los 33 atributos del modelo, ordenados por su grado de importancia dentro del conjunto de rasgos.
- La comparación de los resultados ofrecidos por estos modelos permite observar que ambos coinciden en los primeros 11 rasgos del ordenamiento que da InfoGainAttributeEval con Ranker. El rasgo número 12 de este propio algoritmo es **Mat iv**, por lo que se decide tomar en total 13 rasgos, eliminando **idex 3 y EF 3**, y sustituyendo **FInf** por **Mat iv** por considerarse no solo por el Ranker sino también por los expertos de mucha más importancia.

Se decide por el equipo investigador lo siguiente:

- **Primero:** entrenar el **Modelo de 1er y 2do Año** con todos los atributos, dado el resultados que se obtiene con la mayoría – 5 de 6 - de los métodos empleados.
- **Segundo:** probar con un segundo modelo para el 1er y 2do Año que incluya los atributos de mayor peso (según el InfoGainAttributeEval con Ranker con ambos modos de evaluación – anexos B35 y B36) al que se denomina **Modelo_1er y 2do Año_sel_rasg** que contiene los 13 atributos más importantes según el InfoGainAttributeEval con Ranker y modo de evaluación cross validation con 10 folds. Resultan seleccionados para este modelo los atributos: **ed1, iia, FyS , BD, FyS , MD, Mat III, ED2, Mat I, IP, Acomp, mat iv y Mat II**, listados aquí según el orden que les confiere este modelo. –ver Anexo B35)

Una vez terminada la selección de rasgos, se tienen 6 conjuntos de datos para explorar, tanto con redes neuronales como con árboles de decisión. Quedan entonces, los 3 conjuntos originales y los 3 que resultan de la selección de rasgos:

- Modelo 1er año_sel_rag.
- Modelo 2do año_sel_rag.
- Modelo 1er y 2do año_sel_rag.

Por la relación que guardan con los primeros, estos modelos pueden ser consultados en el **Anexo A, parte 2**.

3.3 Entrenamiento de la red neuronal.

Según los resultados alcanzados en el epígrafe anterior se procede al entrenamiento de la red neuronal, tal como se diseñó en el capítulo anterior. Se utiliza un Perceptrón Multicapa. Se cuenta al final del proceso de selección de rasgos con seis modelos para la experimentación, los 3 originales y los 3 que se obtienen del proceso de selección de rasgos. Se entrenan todos los modelos y se comparan resultados entre los pares de modelo, ya sea que se entrene con todos los rasgos o con la selección.

Como para cada modelo se prueban las 7 configuraciones de la red, resultan 42 modelos diferentes. (Ver Anexo C- Modelos y sus salidas)

Un resumen de los principales resultados se presentan a continuación, organizados por Modelo:

Modelo 1er año.

Todos atributos			Selección de rasgos			Mejor Variante
variante	%BC	Kappa	variante	%BC	Kappa	
V1	80,7018	0,5536	V1	85,0877	0,6523	
V2	83,3333	0,5921	V2	82,4561	0,5942	
V3	82,4561	0,5877	V3	86,8421	0,6883	X
V4	79,8246	0,5062	V4	83,3333	0,5921	
V5	80,7018	0,5605	V5	85,0877	0,6523	
V6	78,9474	0,5278	V6	85,0877	0,6523	
V7	79,8246	0,5440	V7	85,0877	0,6523	

Tabla 7 Resultados de los modelos de 1er año

Como se observa en la tabla anterior se obtienen mejores resultados (mayor % de bien clasificados) cuando se entrena la red para el modelo reducido. El índice de Kappa de todas las variantes de selección de rasgos exceptuando la V2 y la V4 en el modelo de SR son buenos ($>0,60$), en contraste con el modelo que usa todos los atributos, donde los coeficientes de Kappa dan valores dentro del rango de Moderado. Es de destacar que cuatro de las siete variantes dan resultados similares.

De acuerdo con lo anterior, se concluye:

- En primer año, la mejor variante es el modelo reducido: **Modelo 1er Año_sel_rasg**.

Modelo 2do año.

Todos atributos			Selección de rasgos			Mejor Variante
variante	%BC	Kappa	variante	%BC	Kappa	
V1	83,33	0,6615	V1	86,67	0,7205	
V2	83,33	0,6615	V2	87,50	0,7371	
V3	82,50	0,6412	V3	88,33	0,7523	X
V4	81,66	0,6253	V4	85,00	0,6794	
V5	80,83	0,6095	V5	86,67	0,7205	
V6	80,00	0,5913	V6	88,33	0,7555	X
V7	83,33	0,6615	V7	85,83	0,7002	

Tabla 8 Resultados de los modelos de 2do año

Como se observa en la tabla anterior se obtienen mejores resultados de % de bien clasificados cuando se entrena la red para el modelo reducido. También el índice de Kappa de todas las variantes de selección de rasgos es mejor en este modelo, aunque TODOS los coeficientes de Kappa dan valores en el rango de Bueno, por lo que se puede afirmar que aunque el modelo de selección de rasgos entrena mejor, el que utiliza todos los atributos también es bueno pues todos los valores están por encima de 0,60.

De acuerdo con lo anterior, se concluye:

- En segundo año, la mejor variante es el modelo reducido: **Modelo 2do Año_sel_rasg.**

Modelo 1er y 2do año.

Mejor variante	Todos atributos			Selección de rasgos		
	variante	%BC	Kappa	variante	%BC	Kappa
	V1	85,8209	0,7164	V1	81,3433	0,6269
	V2	84,3284	0,6866	V2	81,3433	0,6269
X	V3	86,7652	0,7313	V3	81,3433	0,6269
	V4	85,8209	0,7164	V4	82,8358	0,6567
	V5	85,0746	0,7015	V5	82,0896	0,6418
	V6	85,0746	0,7015	V6	82,8358	0,6567
	V7	85,8209	0,7164	V7	82,8358	0,6567

Tabla 9 Resultados de los modelos 1er y 2do año

Como se observa en la tabla anterior se obtienen mejores resultados de % de bien clasificados cuando se entrena la red para el modelo completo, siendo la mejor de todas la variante no 3, aunque como en el caso anterior el índice de Kappa habla a favor de todas las variantes pues está en el rango de Bueno. También el índice de Kappa de todas las variantes de selección de rasgos es superior a 0,60 por lo que la clasificación se considera en el rango de Buena, aunque ligeramente inferior a las variantes del otro grupo.

En este modelo varias variantes coinciden, obsérvese las V1, V4 y V7 así como las V5 y V6 por el lado del modelo completo. En el modelo con selección de rasgos las V1, V2 y V3 así como las V4, V6 y V7 resultan similares entre sí y si se comparan entre todas hay muy poca variación entre ellas.

De acuerdo con lo anterior, se concluye:

- En el modelo combinado, los mejores resultados se obtienen en el modelo con todos los rasgos: **Modelo 1er y 2do Año.**

3.3.1 Prueba de la red entrenada.

Según los resultados alcanzados en el entrenamiento de la red en el epígrafe anterior se procede a la prueba de la red entrenada, utilizando conjuntos de pruebas. Los conjuntos de prueba se confeccionan independientemente del conjunto de entrenamiento, y para ellos se incluyen un grupo de datos nuevos, con instancias completamente nuevas para la red, y algunas que pueden estar ya presentes en los modelos entrenados. Para consultar los conjuntos de prueba, vea **Anexos D.**

Se prueban todos los modelos y se comparan resultados entre los pares de modelo, ya sea que se pruebe con todos los rasgos o con la selección.

Como para cada modelo se prueban las 7 configuraciones de la red, resultan 42 modelos diferentes. (**Ver Anexo E- Modelos de prueba y sus salidas**)

Un resumen de los principales resultados se presentan a continuación, organizados por Modelo:

Modelo 1er año.

Todos los atributos						
Entrenamiento			Prueba			Mejor Variante
variante	%BC	Kappa	variante	%BC	Kappa	
V1	80,7018	0,5536	V1	92,85	0,8511	X
V2	83,3333	0,5921	V2	85,71	0,6957	
V3	82,4561	0,5877	V3	92,85	0,8511	X
V4	79,8246	0,5062	V4	92,85	0,8511	X
V5	80,7018	0,5605	V5	92,85	0,8511	X
V6	78,9474	0,5278	V6	92,85	0,8511	X
V7	79,8246	0,5440	V7	92,85	0,8511	X

Tabla 10 Resultados del modelo 1er año con entrenamiento y conjunto de prueba

Cuando se prueban las variantes para el modelo con todos los atributos, todas las variantes resultan muy buenas según % de bien clasificadas y los índices de Kappa. La V2 resulta la peor, lo que contrasta con que fue la que mejor entrenó la red. El índice de Kappa indica un grado de concordancia Muy

bueno. En la peor variante tampoco es malo, pues cae dentro del rango de Bueno, aunque las restantes son mucho mejores.

Selección de rasgos						
Entrenamiento			Prueba			Mejor Variante
variante	%BC	Kappa	variante	%BC	Kappa	
V1	85,0877	0,6523	V1	92,30	0,8434	X
V2	82,4561	0,5942	V2	84,61	0,6829	
V3	86,8421	0,6883	V3	76,92	0,5185	
V4	83,3333	0,5921	V4	92,30	0,8434	X
V5	85,0877	0,6523	V5	76,92	0,5185	
V6	85,0877	0,6523	V6	76,92	0,5185	
V7	85,0877	0,6523	V7	76,92	0,5185	

Tabla 11 Resultados del modelo 1er año selrasg con entrenamiento y conjunto de prueba

En este modelo de selección de rasgos, las mejores variantes resultan la V1 y la V4, con Muy buen índice de concordancia según Kappa. La V2 resulta también Buena mientras las restantes dan resultados moderados.

Una vez realizado el entrenamiento con conjunto de prueba, se concluye que:

- Tanto el modelo con todos los atributos como el modelo reducido entrenan bien con conjunto de prueba. Las variantes V1 y V4 resultan entre las mejores en ambos modelos, por lo que resultan candidatas a ser las seleccionadas.
- Como del entrenamiento de la red se tenía que el modelo reducido entrena mejor, entonces se valora como mejores para el uso de redes neuronales en primer año el modelo reducido con las variantes 1 ó 4.

Modelo 2do año.

Todos los atributos						
Entrenamiento			Prueba			Mejor Variante
variante	%BC	Kappa	variante	%BC	Kappa	
V1	83,33	0,6615	V1	100,00	1	X
V2	83,33	0,6615	V2	100,00	1	X
V3	82,5	0,6412	V3	100,00	1	X
V4	81,66	0,6253	V4	100,00	1	X
V5	80,83	0,6095	V5	100,00	1	X
V6	80,00	0,5913	V6	100,00	1	X
V7	83,33	0,6615	V7	100,00	1	X

Tabla 12 Resultados del modelo 2do año con entrenamiento y conjunto de prueba

Cuando se prueban las variantes para el modelo con todos los atributos, todas las variantes resultan Buenas, con un alto 100 % de bien clasificados, y además son equivalentes, no hay ninguna mejor que la otra en el caso de la

prueba. Con respecto a la red entrenada, todas se superan cuando hacen predicciones.

Selección de rasgos						
Entrenamiento			Prueba			Mejor Variante
variante	%BC	Kappa	variante	%BC	Kappa	
V1	86,67	0,7205	V1	86,66	0,5833	X
V2	87,50	0,7371	V2	80,00	0,4444	
V3	88,33	0,7523	V3	86,66	0,5833	X
V4	85,00	0,6794	V4	86,66	0,5833	X
V5	86,67	0,7205	V5	80,00	0,4444	
V6	88,33	0,7555	V6	80,00	0,4444	
V7	85,83	0,7002	V7	86,66	0,5833	X

Tabla 13 Resultados del modelo 2do año selrasg con entrenamiento y conjunto de prueba

En este modelo de selección de rasgos, cuatro de las siete variantes son Moderadas en la predicción según el índice de Kappa, con un alto % de bien clasificados. Si se comparan con la red entrenada, las mejores son la V1, V4 y la V7, pues superan el % de bien clasificados en la prueba de la red.

Una vez realizado el entrenamiento con conjunto de prueba, se concluye que:

- Tanto el modelo con todos los atributos como el modelo reducido entrenan bien con conjunto de prueba. Las variantes V1, V4 y V7 resultan las mejores para el modelo reducido, por lo que resultan adecuadas.
- Como del entrenamiento de la red se tenía que el modelo reducido entrena mejor, entonces se valora como mejores para el uso de redes neuronales en segundo año, el modelo reducido con las variantes 1, 4 ó 7.

Modelo de 1er y 2do año.

Todos los atributos						
Entrenamiento			Prueba			Mejor Variante
variante	%BC	Kappa	variante	%BC	Kappa	
V1	85,82	0,7164	V1	100	1	X
V2	84,32	0,6866	V2	100	1	X
V3	86,76	0,7313	V3	100	1	X
V4	85,82	0,7164	V4	100	1	X
V5	85,07	0,7015	V5	100	1	X
V6	85,07	0,7015	V6	100	1	X
V7	85,89	0,7164	V7	100	1	X

Tabla 14 Resultados del modelo 1er y 2do año con entrenamiento y conjunto de prueba

Cuando se prueban las variantes para el modelo con todos los atributos, todas las variantes resultan Muy buenas, con un 100 % de bien clasificados, y no hay ninguna variante mejor que la otra, sino que todas resultan equivalentes. Con respecto a la red entrenada, todas se superan cuando hacen predicciones.

Selección de rasgos						
Entrenamiento			Prueba			Mejor Variante
variante	%BC	Kappa	variante	%BC	Kappa	
V1	81,34	0,6269	V1	91,66	0,8222	
V2	81,34	0,6269	V2	95,83	0,9091	X
V3	81,34	0,6269	V3	87,50	0,7391	
V4	82,83	0,6567	V4	91,66	0,8222	
V5	82,08	0,6418	V5	87,50	0,7391	
V6	82,83	0,6567	V6	87,50	0,7391	
V7	82,83	0,6567	V7	83,33	0,6896	

Tabla 15 Resultados del modelo 1er y 2do año con entrenamiento y conjunto de prueba

En este modelo de selección de rasgos, todas las variantes son Muy buenas (Kappa > 0,80) o Buenas (Kappa > 0,60) en la predicción, con un muy alto % de bien clasificados, que en tres de las variantes supera el 90%. Aquí es llamativo que aunque todas las variantes entrenan bien, y son prácticamente similares, la V2 es mucho mejor en todo.

Una vez realizado el entrenamiento con conjunto de prueba, se concluye que:

- En este modelo resulta mejor el comportamiento del modelo completo cuando se prueba la red entrenada aunque el comportamiento del modelo reducido es muy bueno también, en particular la V2.

3.4 Entrenamiento del árbol de decisión.

Con los parámetros especificados en el capítulo anterior se entrena el árbol de decisión obteniéndose los resultados que se resumen a continuación. Los modelos obtenidos y los árboles generados pueden verse en el Anexo F.

Al ser seis modelos de datos y obtenerse dos modelos para cada uno, se generan 12 modelos de salidas y 12 árboles. (Ver Anexos F1-F12)

Modelo 1er Año.

Todos atributos			Selección de rasgos		
variante	%BC	Kappa	variante	%BC	Kappa
J48cv	92,9825	0,8324	J48cv	92,9825	0,8324
J48ts	96,4912	0,9162	J48ts	96,4912	0,91,62

Tabla 16 Comparación de los resultados de los árboles en los modelos de 1er año

En el caso del modelo de 1er año el árbol se comporta de idéntica forma para ambos modelos, con todos los rasgos y con la selección, en ambos modos de evaluación, tanto cuando se corre con cross validation como con training set. Ver Anexo F1 al F4.

El árbol obtenido para 1er año es el siguiente:

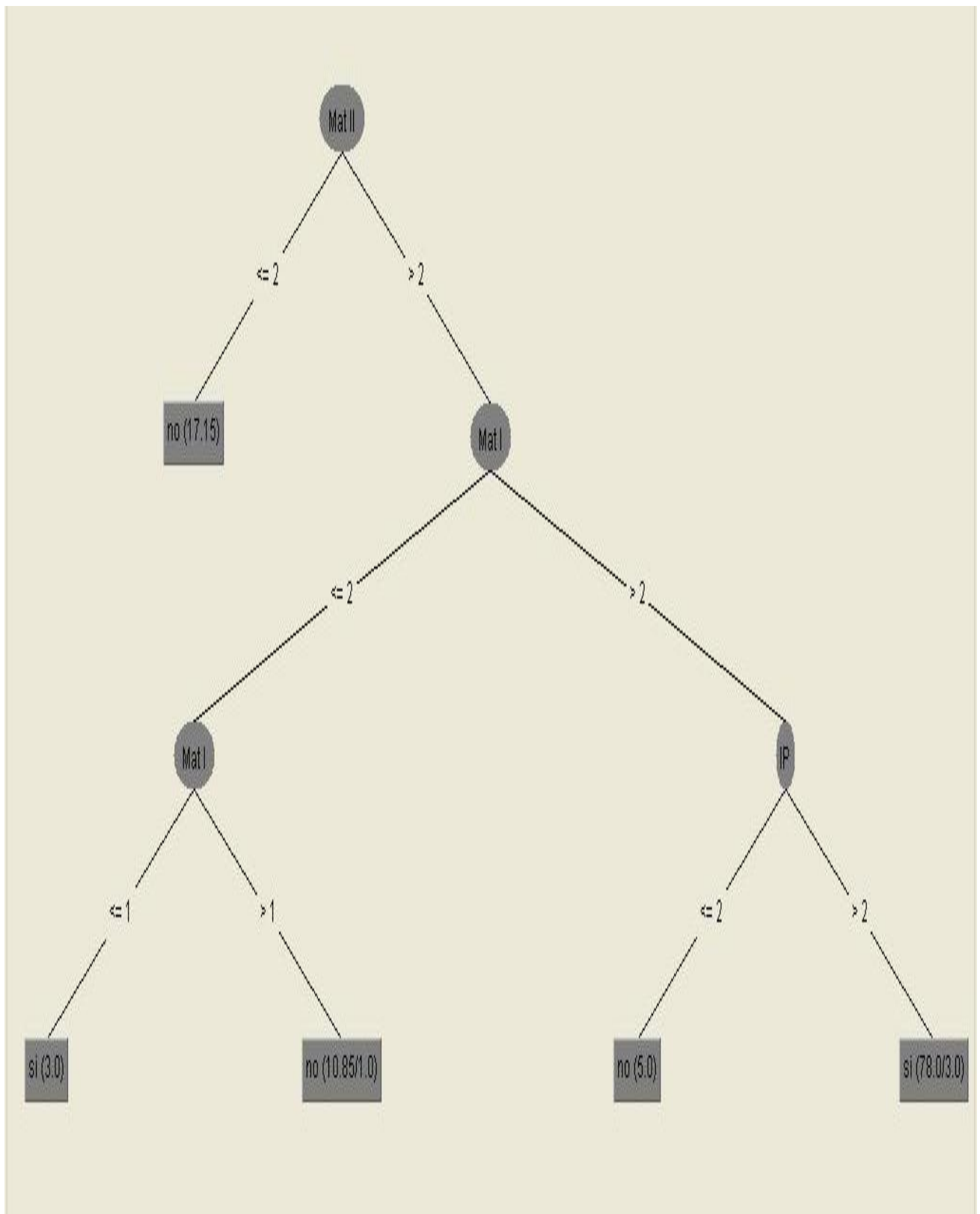


Figura 1 Árbol de decisión para el modelo 1er año

Modelo 2do año.

Todos atributos			Selección de rasgos		
variante	%BC	Kappa	variante	%BC	Kappa
J48cv	85,83	0,7021	J48cv	87,50	0,7354
J48ts	90,83	0,8060	J48ts	92,50	0,8423

Tabla 17 Comparación de los resultados de los árboles en los modelos de 2do año

En el caso del 2do año el árbol se comporta siempre mejor cuando se emplea el modelo de selección de rasgos, y al comparar los modos de evaluación el cross validation produce siempre menos instancias bien clasificadas que el training set, aunque el árbol que ambos generan es el mismo. Ver Anexo F5 al F8.

El árbol obtenido para 2do año es el siguiente:

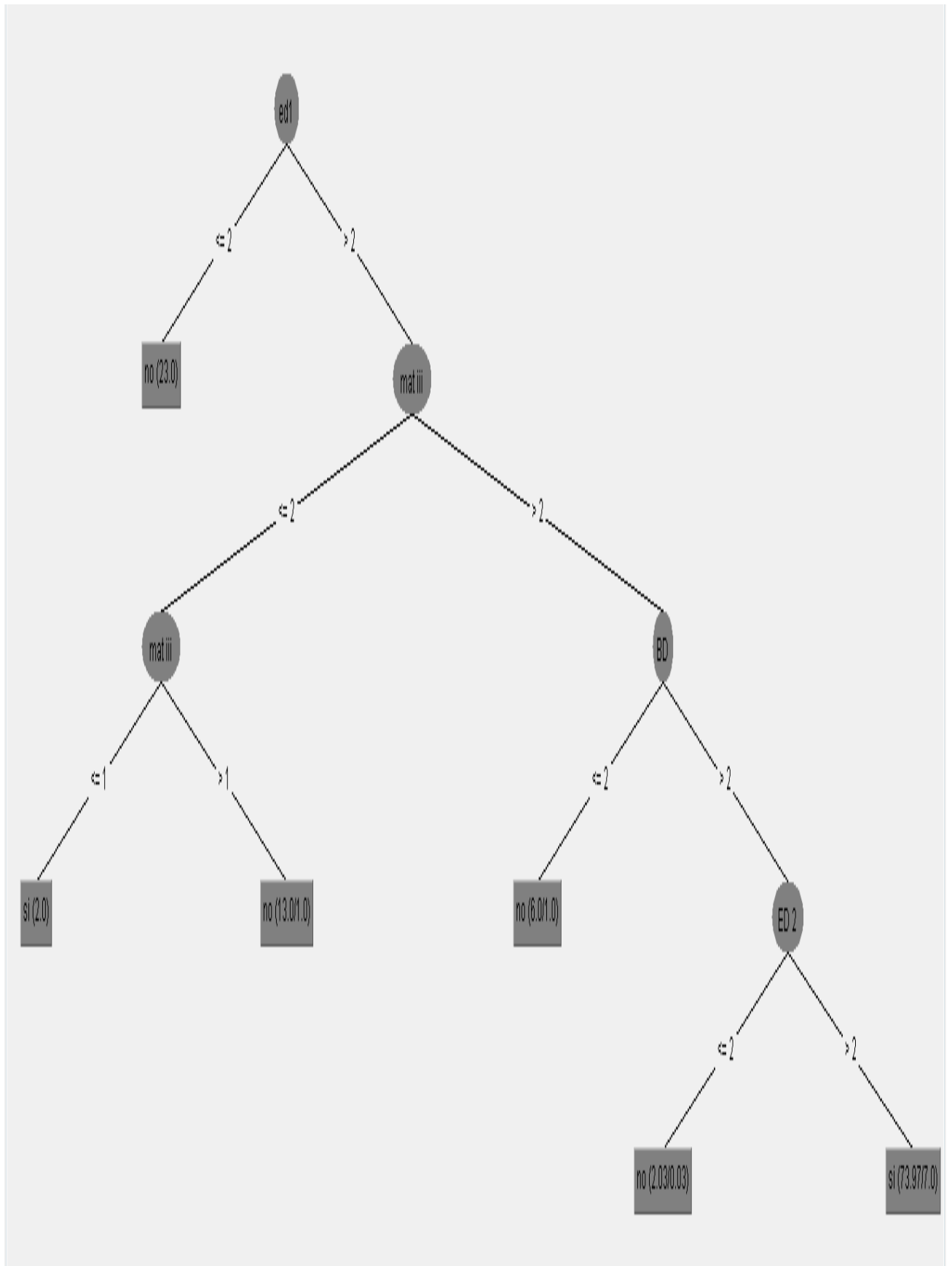


Figura 2 Árbol de decisión para el modelo 2do año

Modelo 1er Año y 2do año.

Todos atributos			Selección de rasgos		
variante	%BC	Kappa	variante	%BC	Kappa
J48cv	78,3582	0,5672	J48cv	76,8656	0,5373
J48ts	94,7761	0,8955	J48ts	88,0597	0,7612

Tabla 18 Comparación de los resultados de los árboles en los modelos de 1er y 2do año

En el caso de este modelo las dos mejores variantes son las del modelo con todos los atributos, y la que evalúa cross validation resulta Moderada mientras en contraste la del training set resulta Muy bueno. Ver Anexo F9 al F12.

En conclusión:

- El árbol de decisión resulta el mismo, para el caso de primer año, no se producen diferencias en lo cuantitativo cuando se emplean ambos modelos.
- En segundo año el mejor resulta el que se obtiene con el modelo reducido, el de selección de rasgos.
- En el caso del modelo combinado, resulta mejor el que se obtiene para todos los atributos.
- El entrenamiento con training set, aunque no es el mejor modo de evaluación, resulta útil pues su valor es un indicador del valor máximo a esperar cuando se entrena el árbol con un conjunto de prueba y con esa intención – y ninguna otra- ha sido usado en esta parte.

El árbol para el modelo de 1er y 2do año es el siguiente:

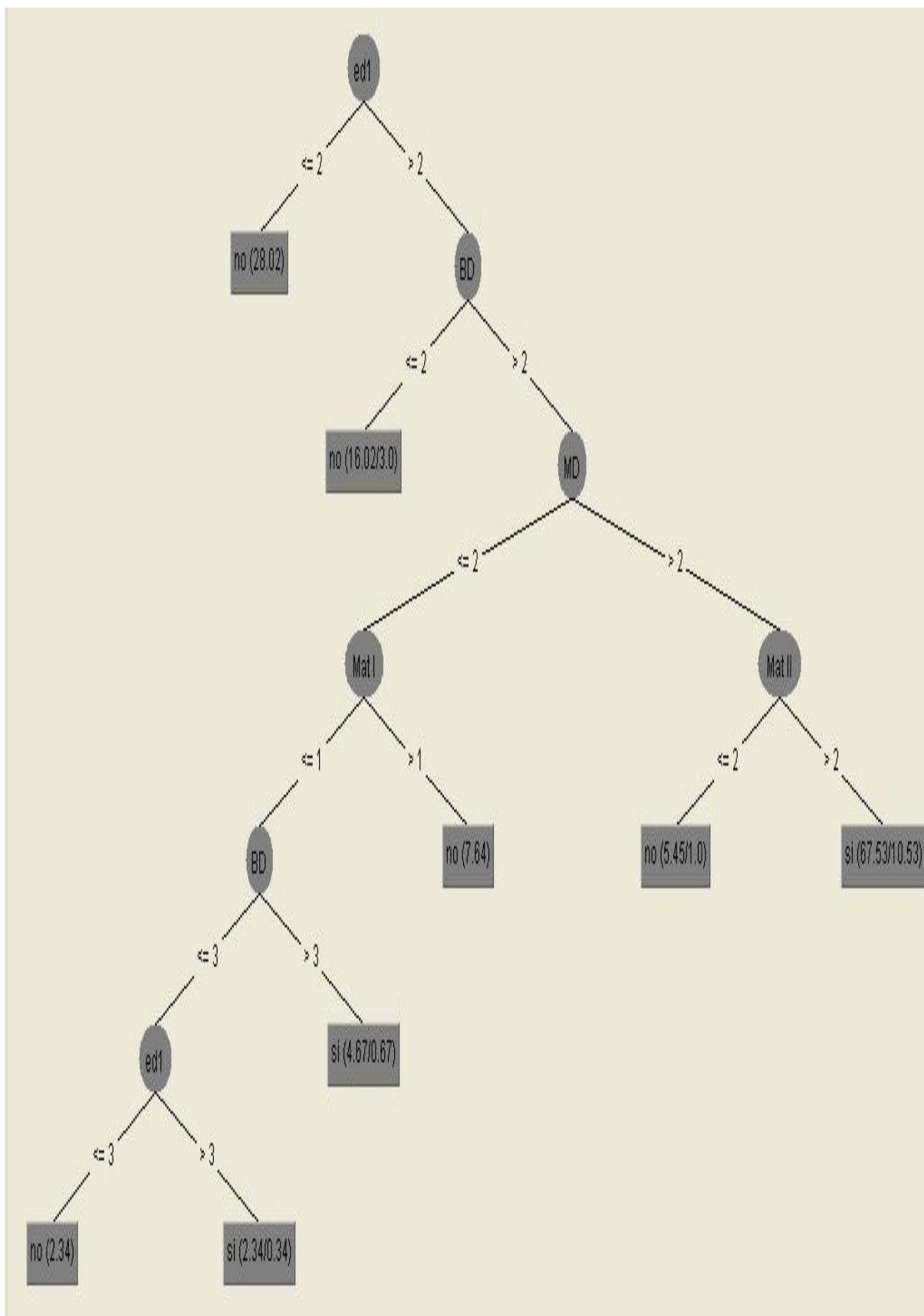


Figura 3 Árboles de decisión para el modelo 1er y 2do año

3.4.1 Prueba de los modelos de árboles.

Para la prueba de los modelos dados por los árboles, se utilizan los mismos conjuntos de entrenamiento empleados con las redes. Ver Anexo D.

En la prueba de los modelos entrenados, se obtienen los resultados que se resumen a continuación. Los modelos obtenidos pueden verse en el Anexo G.

Modelo 1er Año.

Todos atributos					
variante	%BC	Kappa	variante	%BC	Kappa
J48cv	92,9825	0,8324	J48cv	80,00	0,5238
J48ts	96,4912	0,9162			

Tabla 19 Resultados de los árboles con conjunto de prueba para el modelo 1er año

Selección de rasgos					
variante	%BC	Kappa	variante	%BC	Kappa
J48cv	92,9825	0,8324	J48cv	81,25	0,6250
J48ts	96,4912	0,9162			

Tabla 20 Resultados de los árboles con conjunto de prueba para el modelo 1er año selrasg

Aunque en el entrenamiento los árboles de decisión de ambos modelos fueron equivalentes, en la prueba del modelo se obtienen mejores resultados en el modelo de Selección de rasgos, dado que el índice de Kappa pasa de ser Moderado en el modelo completo a Bueno para el modelo reducido. El % de bien clasificados también es ligeramente mejor.

Modelo 2do año.

Todos atributos					
variante	%BC	Kappa	variante	%BC	Kappa
J48cv	85,83	0,7021	J48cv	86,66	0,6591
J48ts	90,83	0,8060			

Tabla 21 Resultados de los árboles con conjunto de prueba para el modelo 2do año

Selección de rasgos					
variante	%BC	Kappa	variante	%BC	Kappa
J48cv	87,50	0,7354	J48cv	87,50	0,7354
J48ts	92,50	0,8423			

Tabla 22 Resultados de los árboles con conjunto de prueba para el modelo 2do año selrasg

En el caso de segundo año, el modelo reducido había sido mejor en el entrenamiento. En la prueba de ambos modelos se obtienen también los mejores resultados en el modelo de Selección de rasgos, aunque el índice de

Kappa es Bueno en ambos casos, el % de bien clasificados también es ligeramente mejor en el modelo de selección de rasgos.

Modelo 1er Año y 2do año.

Todos atributos					
variante	%BC	Kappa	variante	%BC	Kappa
J48cv	78,3582	0,5672	J48cv	83,33	0,6596
J48ts	94,7761	0,8955			

Tabla 23 Resultados de los árboles con conjunto de prueba para el modelo 1er y 2do año

Selección de rasgos					
variante	%BC	Kappa	variante	%BC	Kappa
J48cv	76,8656	0,5373	J48cv	83,33	0,6596
J48ts	88,0597	0,7612			

Tabla 24 Resultados de los árboles con conjunto de prueba para el modelo 1er y 2do año selrasg

En el caso de este modelo, ambos, tanto el modelo con todos los atributos como el reducido, dan idénticos resultados a la hora de la prueba, con un índice de Kappa en el rango de Bueno, y un % de bien clasificados de 83,33.

En general la prueba de los modelos de los árboles da como resultado:

- Resultan apropiados para el modelos de 1er año y el modelo de 2do año, el modelo reducido o de selección de rasgos.
- En el caso del modelo combinado de 1er y 2do años, la prueba no aporta información por tanto se selecciona el modelo completo, tal como se había obtenido de la etapa de entrenamiento.

3.5 Redes neuronales versus Árboles.

El entrenamiento de la red neuronal y del árbol de decisión, resultan adecuados para el problema planteado. No obstante si se tienen en cuenta los resultados obtenidos hasta ahora, aparentemente el árbol de decisión proporciona mejor resultados tanto en el entrenamiento como en la prueba. Atendiendo a que se ha realizado una sola iteración del experimento, antes de decidir por el árbol como modelo más apropiado a nuestro caso, realizamos un último experimento, donde comparamos el árbol de decisión versus todas las variantes de la red.

El objetivo será determinar cuál de los 8 métodos se comporta mejor, para cada par de modelos considerado.

Por tal motivo, el experimento lo dividimos en tres fases:

- 1ra fase: Modelo para 1er año.
- 2da fase: Modelo para 2do año.

- 3ra fase: Modelo para 1er y 2do año.

La diferencia entre el experimento actual y el anterior, viene dado porque ahora se realizan tres iteraciones con cada modelo, con lo cual, deben obtenerse resultados más precisos.

1ra fase: Modelo para 1er año.

Los parámetros especificados para este experimento son:

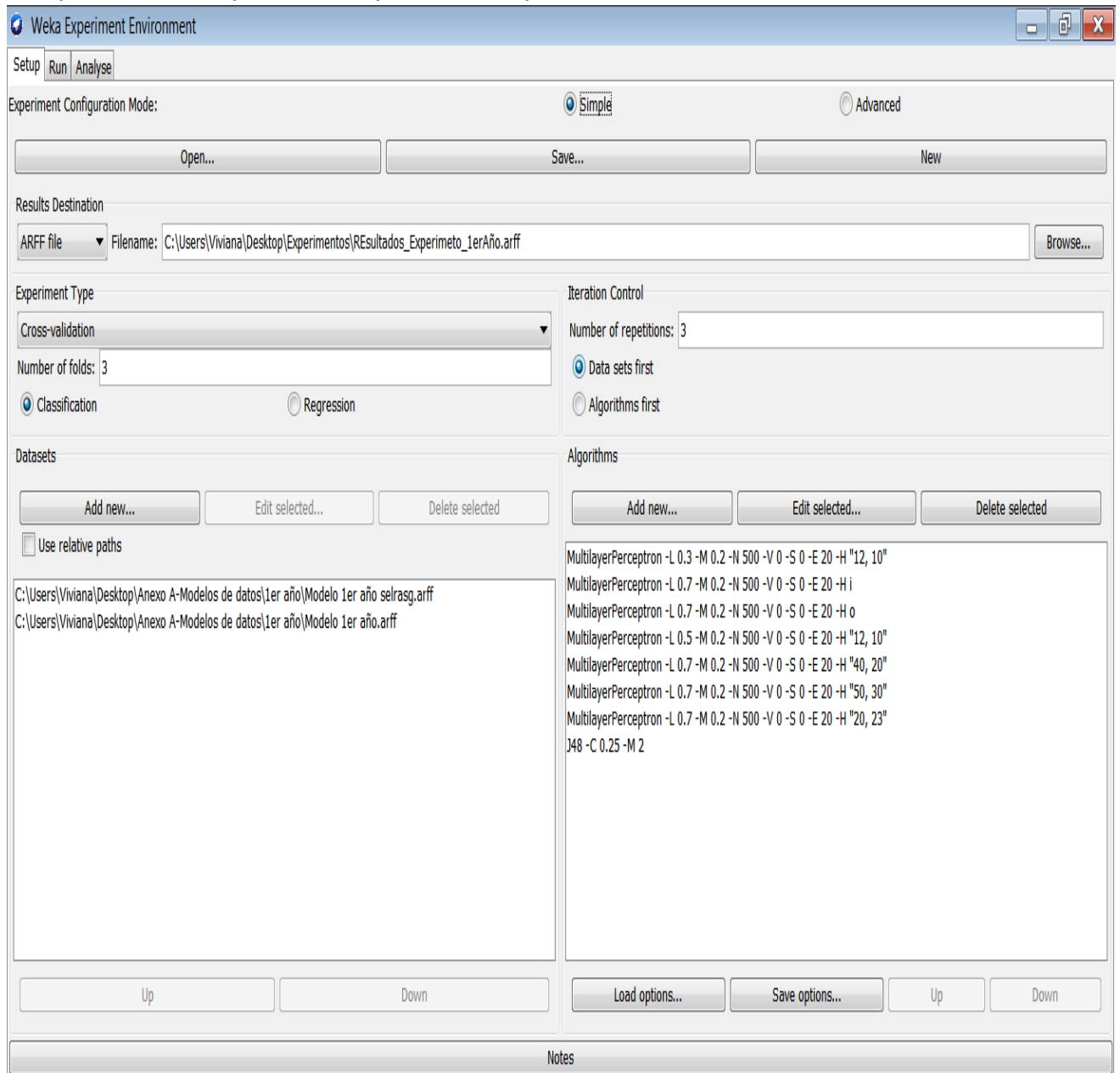


Figura 4 Parámetros especificados para 1er año

Un resumen de los resultados obtenidos se da en la siguiente tabla:

Por ciento de bien clasificados								
Modelo	J48	V1	V2	V3	V4	V5	V6	V7
1er año selrasg	92.11	83.92	81.29	86.26	85.09	85.09	84.50	84.21
1er año	92.11	81.29	83.04	82.75	80.70	81.29	79.53	81.29

Tabla 25 Por ciento bien clasificados por variantes en los modelos de 1er año

- Como es posible observar el J48 es superior a cualquiera de las variantes de la red neuronal en cualquiera de los dos modelos de datos (1ro y 1ro sel_rasg). Los índices de Kappa caen en el rango de Muy buena.
- Si nos fijamos solo en las variantes de la red, esta entrena y prueba mejor en seis de las siete variantes del modelo con selección de rasgos. En este caso los índices de Kappa son Buenos, mientras que para el modelo con todos los rasgos todos caen en el rango de Moderado.

Más información relacionada verse en el Anexo H1.

Se concluye que: para primer año, el algoritmo J48 es mejor que las variantes de la red neuronal.

2da fase: Modelo para 2do año.

Los parámetros especificados para este experimento son:

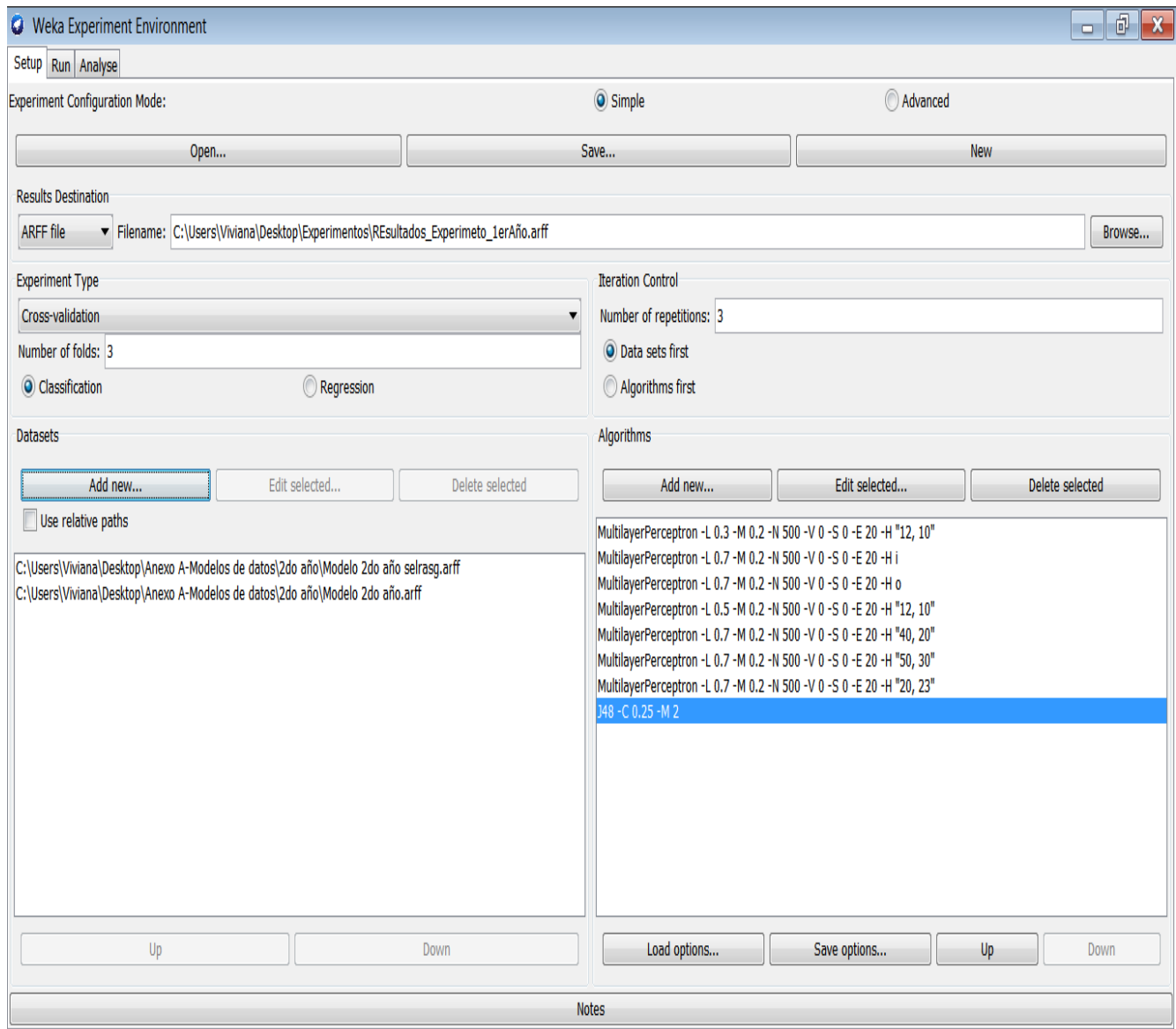


Figura 5 Parámetros especificados para 2do año

Un resumen de los resultados obtenidos se da en la siguiente tabla:

Porcentaje de bien clasificados								
Modelo	J48	V1	V2	V3	V4	V5	V6	V7
2do año selrasg	87.22	85.56	86.11	86.94	85.83	85.83	85.28	85.00
2do año	86.11	81.39	80.83	81.39	81.11	80.83	80.28	81.39

Tabla 26 Porcentaje bien clasificados por variantes en los modelos de 2do año

- Aquí en este caso, también es posible observar que el J48 es superior a cualquiera de las variantes de la red neuronal en cualquiera de los dos modelos de datos (2do y 2do sel_rasg). Los índices de Kappa caen en el rango de Muy buena.

- Si nos fijamos exclusivamente en las variantes de la red, esta entrena y prueba mejor en todas las variantes del modelo con selección de rasgos. Todos los índices de Kappa en este modelo caen en el rango de Buena, mientras en el modelo con todos los rasgos exceptuando dos el resto caen en el rango de Moderado.

Más información relacionada verse en el Anexo H2.

Se concluye que: para segundo año, el algoritmo J48 es mejor que las variantes de la red neuronal.

3ra fase: Modelo para 1ro y 2do año.

Los parámetros especificados para este experimento son:

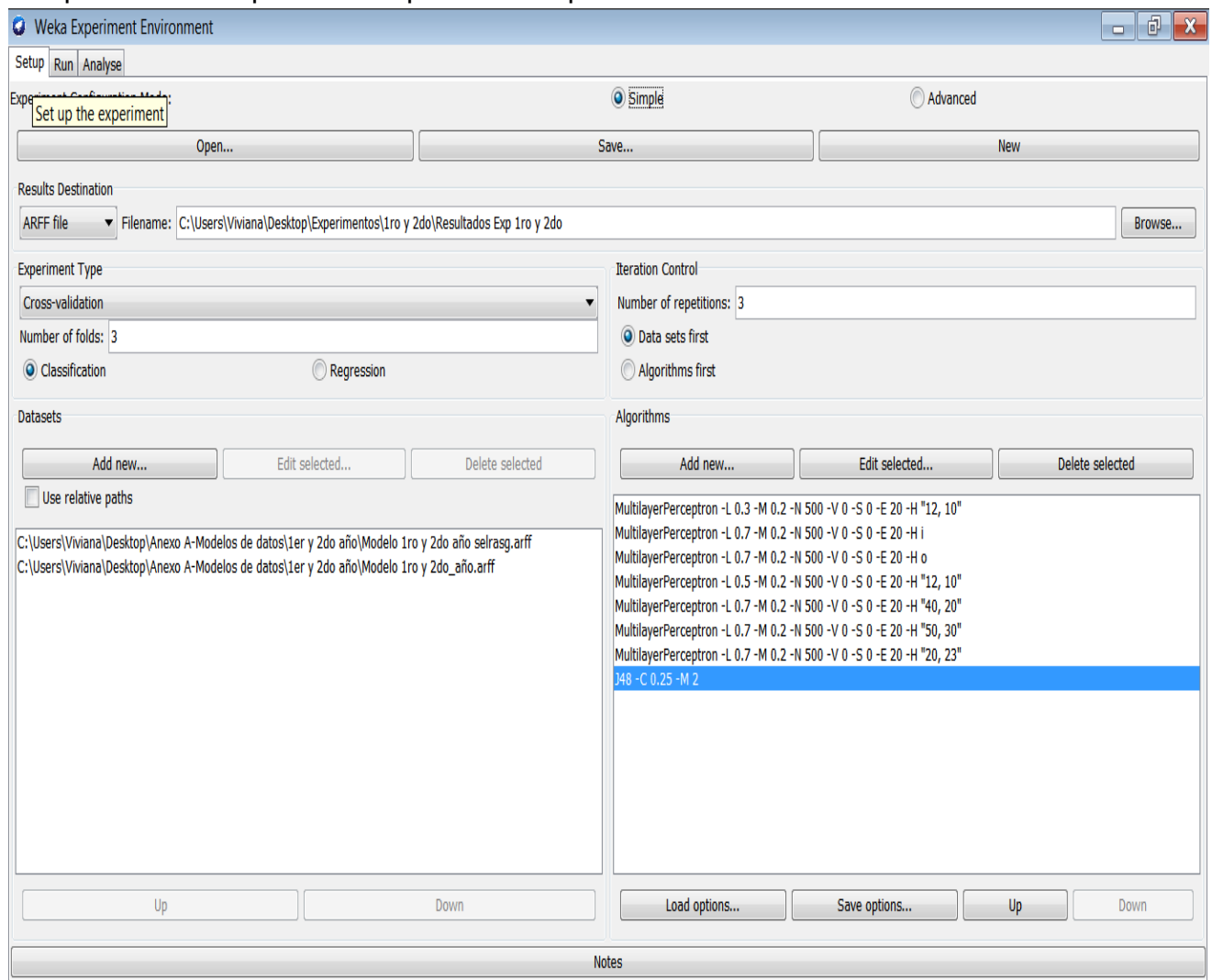


Figura 6 Parámetros especificados para 1er y 2do año

Un resumen de los resultados obtenidos se da en la siguiente tabla:

Porcentaje de bien clasificados								
Modelo	J48	V1	V2	V3	V4	V5	V6	V7
1ro y 2do año selrasg	71.40	79.84	79.83	78.84	79.84	79.34	80.07	79.08
1ro y 2do año	72.65	84.30	83.56	84.30	84.30	84.55	84.55	84.55

Tabla 27 Porcentaje bien clasificados por variantes en los modelos de 1er y 2do año

- Como es posible observar, en este caso el mejor entrenamiento viene dado por la red en sus variantes V5, V6 y V7 para el caso del modelo sin reducir, el J48 resulta el peor de todos en ambos modelos.
- Todos los índices de Kappa para el modelo 1ro y 2do año para las variantes de red neuronal caen en el intervalo de Bueno, mientras que para el modelo con selección de rasgos están en Moderado. El árbol de decisión es el peor de todos como ya se dijo y sus índices de Kappa aunque en el intervalo de Moderado está muy próximo a ser Débil.

Más información relacionada verse en el Anexo H3.

Se concluye que: para primer y segundo año, modelo combinado, el mejor algoritmo es la red neuronal, con el modelo sin reducción de rasgos: V5, V6 o V7 y para el modelo con selección de rasgos las variantes V1 o V4.

3.6 Pronósticos.

Del resultado del entrenamiento de las redes y el árbol de decisión se obtiene que los mejores resultados se encuentran en: Modelo 1ro sel-rasg y Modelo 2do sel_rasg. En ambos casos con el algoritmo J48.

Por tal motivo, utilizando los árboles de decisión que genera este algoritmo y los conjuntos de datos correspondientes a estos modelos, se realiza un pronóstico para las variables más influyentes en la variable objetivo de nuestro problema.

Modelo	Modelo 1er año_sel rasg	Modelo 2do año_sel rasg
Rasgos decisores	Mat II	ed1
	Mat I	mat iii
	IP	BD

Tabla 28 Rasgos decisores en los modelos

Se utiliza para ello el algoritmo SMOreg, este implementa las máquinas de vector soporte para regresión. Los parámetros pueden ser aprendidos usando

varios algoritmos. El algoritmo más popular (RegSMOImproved) es debido a Shevade, Keerthi [66] y es el que se usa por defecto. Este es el algoritmo para maquinas vector soporte que incluye Weka.

Pronóstico para primer Año.

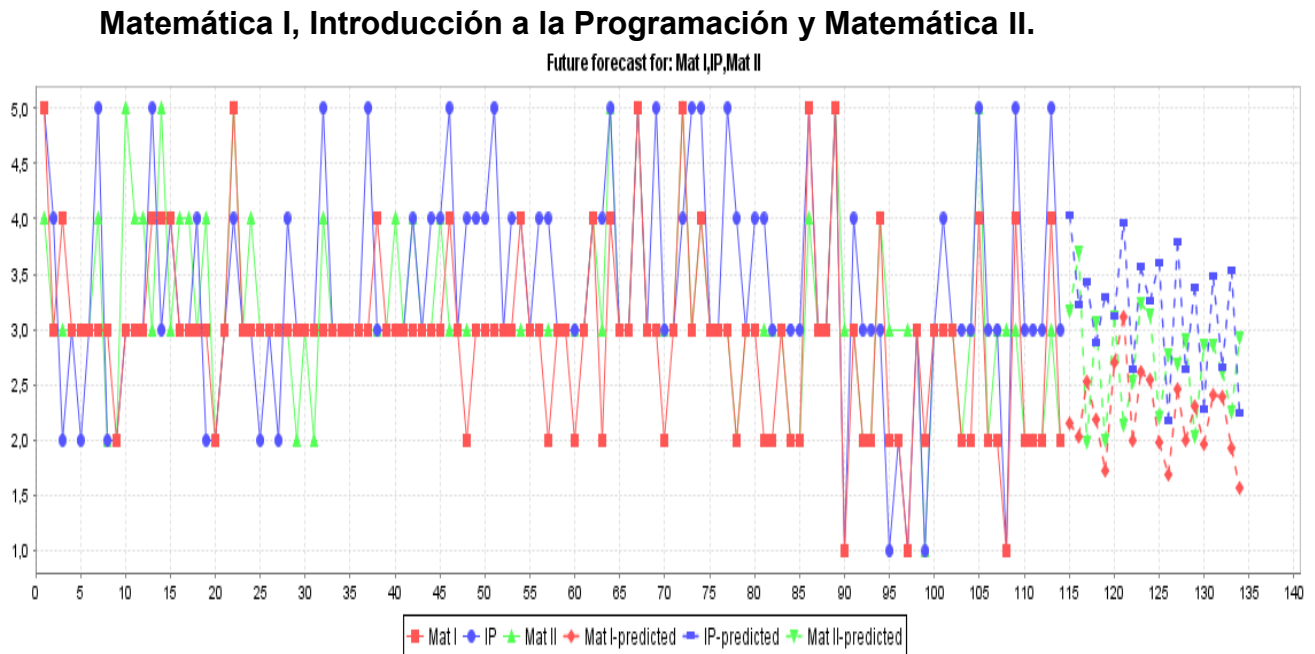


Figura 7 Pronóstico para Matemática I, Introducción a la Programación y Matemática II

Para interpretar la gráfica se debe conocer que los puntos unidos por líneas constituyen el conjunto de entrenamiento mientras que los puntos unidos por líneas discontinuas son los “nuevas” instancias que se obtienen sobre la base del modelo de pronóstico. Este valor en este caso es de 20, aunque puede ser variado.

La observación de la gráfica permite afirmar que según el pronóstico para 20 estudiantes, en Introducción a la Programación (ver puntos azules), por ejemplo, habría que esperar que un 35% desaprueba, un 55 % aprueba con nota de 3 y un 10% aprueba con nota de 4. Nadie obtiene 5. Como todo pronóstico tiene un porcentaje de certeza, y bastaría que se cumpliera en buena medida para considerarlo bueno. En este caso el profesor principal de la asignatura considera que es un buen pronóstico.

En Matemática I habría de esperar un 5 % de aprobado, y en Matemática II un 25 %.

Obsérvese como coinciden en el pronóstico, que varios de los alumnos que desaprueban Matemática I coinciden en Matemática II o Introducción a la Programación, en ocasiones en las tres asignaturas.

Información adicional del modelo obtenido puede hallarse en el Anexo I1.

Pronóstico para segundo Año.

Matemática III, Estructura de datos I y Bases de datos.

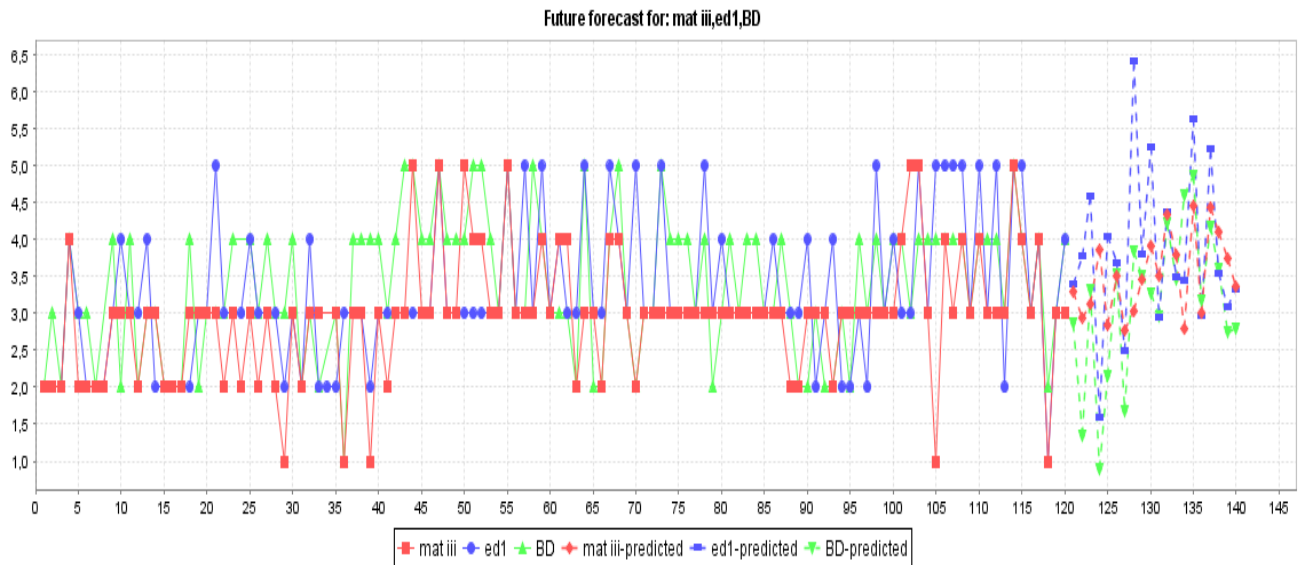


Figura 8 Pronóstico para Matemática III, Base de Datos y Estructura de Datos I

En segundo año para la asignatura Bases de Datos (**puntos verdes**) se prevé un 35% de desaprobados con un 45% de aprobados con 3, y un 20 % de aprobados con 4, es decir con calidad. Nadie obtiene 5. En Estructura de Datos se pronostica un 25 % de desaprobados y de los aprobados un 35 % con 4. Matemática III aporta un 20 % de desaprobados y solo un 20 % de calidad- uno de cada cinco obtiene 4 puntos-. Es interesante observar como en el pronóstico varios de los desaprobados lo están en más de una asignatura.

Información adicional del modelo obtenido puede hallarse en el Anexo I2.

Pronóstico combinado de 1ro y 2do año.

En este caso, para realizar el pronóstico en el modelo combinado se seleccionan las asignaturas que resultaron tener mayor influencia según el árbol de decisión J48.

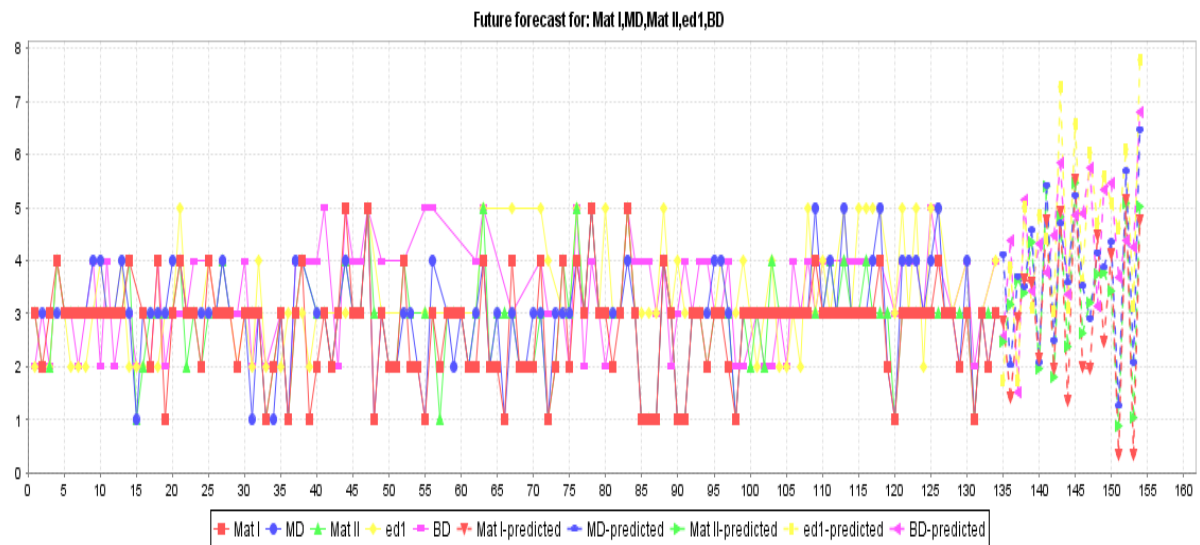


Figura 9 Pronóstico combinado para 1er y 2do año

Cuando en el modelo de 1er y 2do años, con selección de rasgos se aplica pronóstico se observa que los que desaproveban Matemática I coinciden con Matemática Discreta y/o con Matemática II.

Este modelo ofrece información a más largo plazo, aportando datos de la relación que se establece entre rasgos significativos en dos cursos académicos, obsérvese como Matemática I y II son más importantes en el aprobado después de dos cursos, que la Matemática III y IV, la Estructura de Datos I pesa más que la II en este período, y la Matemática Discreta que en primer año no quedó entre las de mayor peso, cuando es vista en un período mayor, adquiere importancia. Bases de datos pesa tanto en su año como en el período.

Información adicional del modelo obtenido puede hallarse en el Anexo I3.

3.7 Conclusiones parciales.

El análisis de los resultados realiza en este capítulo nos permite afirmar que:

- la selección de rasgos mejora los resultados que se obtienen en los modelos aplicados.
- las redes neuronales y los árboles de decisión una vez probados resultan métodos de aprendizaje automático adecuados a la predicción,
- que de la información relevante que aportan los métodos anteriores y de la posibilidad que nos brinda el algoritmo SMOreg se obtiene nueva información útil a la carrera.

Conclusiones

Una vez concluido el trabajo se arriba a las siguientes conclusiones:

- De la selección de rasgos:
 - La combinación de evaluadores y métodos de búsqueda mejora los resultados que se obtienen de utilizar uno solo de ellos, la mejor combinación en este problema estuvo dada por el evaluador CfsSubsetEval (con FCBFSearch ó BestFirst como métodos de búsqueda) unido al InfoGainAttributeEval con Ranker.
 - Se obtuvieron tres nuevos conjuntos de datos que resultaron muy útiles al entrenamiento de la red y árbol de decisión. En cada modelo se eliminaron un grupo de rasgos:
 - Primer año: 5
 - Segundo año: 9
 - Modelo conjunto: 19.
- Del entrenamiento de la red neuronal con 7 configuraciones diferentes se obtiene:
 - Que el algoritmo MultiLayerPerceptron arroja buenos resultados para el problema, siendo ligeramente superior cuando se entrenan los modelos reducidos por selección de rasgos. Porcentaje de bien clasificados para cada caso queda:
 - modelos completos: intervalo [78,90; 86,76],
 - modelos reducidos: intervalo [81,34; 88,33]
 - Los modelos de red reevaluados con un conjunto de prueba con datos nuevos, mejoran los resultados del entrenamiento (mejor % de clasificados) en el 84 % de los casos, considerándose muy positivo.
- Del entrenamiento del árbol de decisión se obtiene:
 - Excelentes resultados para los modelos de 1er y 2do año ([85,83; 92,98] % de BC), en contraste con los que se obtienen para el modelo combinado, que oscilan entre [76,86; 78,35].
 - Los modelos del árbol reevaluados con un conjunto de datos nuevos supera al entrenamiento en un 66 % de los casos, considerándose muy bueno, y oscilando entre el [81,25; 87,50] % de bien clasificados.
- Los experimentos comparativos entre pares de modelos arroja que:
 - El árbol de decisión J48 es mejor para los modelos de 1er año, y modelos de 2do año,

- Para el modelo combinado resulta mejor la red, para el caso del modelo reducido las variantes V1 o V4, y para el modelo completo, las variantes V5, V6 o V7.
- Los rasgos más influyentes en la variable objetivo según los modelos obtenidos son:
 - Para primer año: Matemática I, Matemática II e Introducción a la Programación.
 - Para segundo año: Estructura de datos I, Matemática III y Bases de datos.
 - Para el período de dos cursos: Matemática I, Matemática II, Matemática Discreta, Estructura de datos I y Bases de datos.
- Los resultados del pronóstico realizado a los rasgos característicos obtenidos de los modelos de árbol de decisión en los tres conjuntos de datos reducidos aporta información valiosa desde el punto de vista docente metodológico para las asignaturas, disciplinas, colectivos de año y carrera en general.

Recomendaciones

Al llegar al final del trabajo quedan como recomendaciones para trabajos futuros:

- Desarrollar en Java un sistema de predicción que pueda predecir el comportamiento de un alumno partiendo de los modelos de red y árbol de decisión obtenidos.
- Ampliar el conjunto de entrenamiento utilizando una muestra mayor de estudiantes considerando más cursos académicos.
- Incluir atributos no académicos para medir el peso que puedan tener en la predicción del valor de la variable objetivo.
- Continuar trabajando en la carrera con otros años pertenecientes al ciclo especializado.

Referencias Bibliográficas

- [1] Felder, R. and L. Silverman, *Learning and Teaching Styles in Engineering Education Application, in Engr. Education*. 1988.
- [2] Durán, E. and R. Costagauta, *Minería de datos para descubrir estilos de aprendizaje*. Revista Iberoamericana de Educación, 2007.
- [3] Quiñones, Yessica, "Aplicación de técnicas de Análisis de Cluster a la exploración de los Estilos de Aprendizaje en estudiantes de Ingeniería Informática," Universidad Carlos Rafael Rodríguez, Cienfuegos, 2011.
- [4] Felder, R. and R. Brent, *Learning and teaching styles in engineering education*. *Engineering Education Journal*. 2005.
- [5] José Alberto Chang García, "Determinación de los Estilos de Enseñanza predominantes en profesores de Ingeniería Informática aplicando técnicas de Análisis de Clúster," Universidad de Cienfuegos, Cuba, 2012.
- [6] Ian H. Witten and Eibe Frank, *Data Mining Practical Machine Learning Tools and Techniques*, Segunda. 2005.
- [7] Damián Jorge Matich, "Redes Neuronales: Conceptos Básicos y Aplicaciones.," Universidad Tecnológica Nacional – Facultad Regional Rosario, Departamento de Ingeniería Química, Argentina, 2001, p. 55.
- [8] Edgar Pérez Rivero, "PREDICCIÓN DE RESULTADOS ACADÉMICOS EN LA ASIGNATURA SO HACIENDO USO DE REDES NEURONALES," Instituto Superior Politécnico "José Antonio Echeverría," 2014.
- [9] Jorge Álvarez Blanco, "Predicción de resultados académicos en las asignaturas Estructuras de Datos 1 y 2," Instituto Superior Politécnico José Antonio Echeverría, 2014.
- [10] De Toro and Gisbert M, *Pequeño Larousse Ilustrado*. librería Larousse, Paris, Francia, 1926.
- [11] Stanton WJ, Etzel MJ, and Welker BJ, "Fundamentos del marketing." McGraw-Hil, 2004.
- [12] Carranza O and Sabrià F, *Logística, mejores prácticas en Latinoamérica*. Thompson Mexico, 2005.
- [13] Chase R, Jacobs R, and Aquilano N, *Administración de operaciones. Producción y cadena de suministros*. Editorial Mc Graw Hill, 2010.
- [14] Carabaña, J, "Origen social, inteligencia y rendimiento académico al final de la EGB," p. 2014.
- [15] Pérez Serrano. G, "Crítica al concepto de rendimiento académico," *Rev. Esp. Pedagog.*, vol. 174, pp. 521–534.
- [16] Vázquez Alonso, A, "La enseñanza de la Física (y la Química) en la E.G.B." *Revista Española de Física*, 1989.
- [17] "Tes de pronóstico académico." 2014.
- [18] Rodríguez Espinar, "Factores de rendimiento escolar," *Oikos-Tau*, 1982.
- [19] Pérez Serrano. G, "Origen social y rendimiento escolar Madrid," *Cent. Investig. Sociológicas*, 1981.
- [20] Gómez Molina, J. R, "Estudio descriptivo de las evaluaciones de B.U.P. y C.O.U. en el país valenciano." Bardón, 1982.
- [21] Ministerio de Educación y Ciencia (C.I.D.E.), "Avance datos diferencias alumnos control/experimentación en la Reforma de las Enseñanzas Medias." Manuscrito no publicado, 1986.
- [22] Reparaz, C., Touron, J., and Villanueva, C, "La predicción del rendimiento académico en el Curso de Orientación Universitaria." *Revista Española de Pedagogía.*, 1987.
- [23] Vázquez Alonso, A., "Rendimiento en Bachillerato: Aptitudes y atribución causal. Análisis del rendimiento objetivo en Física y Química," Universidad de las Islas Baleares, 1989.

- [24] Ed. Félix Varela, MES, *Reglamento de Organización Docente, Resolución no. 120/2010*, Ed. 2010, p. 22.
- [25] Sancho, F.V.N., "Estilo de Aprendizaje del Alumno y Técnicas para mejorar su aprendizaje," Jun. 2009.
- [26] Weiss, S. M and N. Indurkha, *Predictive data mining: A practical guide*. San Francisco, EEUU, 1998.
- [27] Abhishek Tiwaria and Arvind K.T. Sekhar, *Workflow based framework for life science informatics*, vol. 31. 2007.
- [28] "KNIME 2.1.0 lanzado." .
- [29] "KEEL 2." .
- [30] Derrac, "KEEL Una herramienta de software para el análisis de sistemas." 2010.
- [31] "KEEL. es." 2010.
- [32] J. Alcalá-Fdez, L. Sánchez, S. García, M.J. del Jesus, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Bacardit, V.M. Rivas, J.C. Fernández, and F. Herrera, "Documento de descripción de KEEL." 2009.
- [33] "Weka - EcuRed," 05-Feb-2015. [Online]. Available: <http://www.ecured.cu/index.php/Weka>. [Accessed: 05-Feb-2015].
- [34] Diego García Morate, "Manual de Weka." 2012.
- [35] Valadez, E.H, "Algoritmo de clustering basado en entropía para descubrir grupos en atributos de tipo mixto," Sección de Computación, Centro de Investigación y de Avanzados del Instituto Politécnico Nacional, DF, México, 2006.
- [36] J. Hernández, "Introducción a la Minería de Datos." Educación Pearson, Madrid, España, 2004.
- [37] Lucila L. Chiarvetto Peralta, Fernando A. Rey Saravia, and Nélida B. Brignole, "APLICACIÓN DE REDES NEURONALES ARTIFICIALES PARA LA PREDICCIÓN DE CALIDAD DE AIRE," vol. 27, Nov. 2008.
- [38] MARIA AUXILIADORA PEREZ, "APLICACIÓN DE MINERIA DE DATOS PARA LA PREDICCIÓN DEL RENDIMIENTO DE LA CAÑA DE AZÚCAR," UNIVERSIDAD CENTROCCIDENTAL "LISANDRO ALVARADO", Barquisimeto, Venezuela, 2005.
- [39] Sotolongo, G., Guzmán, and Maria Victoria, "Aplicaciones de las redes neuronales. El caso de la Bibliometría." 2001.
- [40] Carlos Serrano Cinca and Bonifacio Martin del Brio, "Predicción de la quiebra bancaria mediante el empleo de las redes neuronales artificiales," vol. 23, no. 74, 1993.
- [41] Roberto Suárez Surí, "Modelo basado en aprendizaje automático para el pronóstico al egreso de pacientes con trauma grave de cráneo," Universidad de Cienfuegos, Cuba, 2013.
- [42] Xabier Basogain Olabe, "Redes Neuronales Artificiales y sus aplicaciones." 2000.
- [43] Castellano, "Introducción a las redes neuronales aplicadas." 2013.
- [44] Luis Llano and MSc.(c) Andrés Hoyos, "Comparación del Desempeño de Funciones de Activación en Redes Feedforward para aproximar Funciones de Datos con y sin Ruido," p. 10, Jul. 2007.
- [45] P.A. Castillo, J.G. Castellano, J.J. Merelo, and A. Prieto, "Diseño de redes neuronales artificiales mediante algoritmos evolutivo," vol. 5, no. 14, 2001.
- [46] ESCOM, "Diseño y Entrenamiento de Redes Neuronales Artificiales," 2015. [Online]. Available: <http://es.slideshare.net/mentelibre/diseo-y-entrenamiento-de-redes-neuronales-artificiales>. [Accessed: 27-Jan-2015].
- [47] "Funciones en el modelo de Neurona Artificial," 2014. [Online]. Available: <http://www.redes-neuronales.com.es/tutorial-redes-neuronales/funciones-de-las-neuronas-artificiales.htm>. [Accessed: 05-Feb-2015].
- [48] María E. Valesani, Osvaldo P. Quintana, and Oscar A. Vallejos, "Imputación de datos con redes neuronales." 07-Sep-1999.

- [49] Félix de Moya Anegón, Víctor Herrero Solana, and Vicente Guerrero Bote, "La aplicación de Redes Neuronales Artificiales (RNA): a la recuperación de la información." 2005.
- [50] Carlos Serrano Cinca and José Luis Gallizo Larraz, "Las redes neuronales artificiales en el tratamiento de la información financiera," 05-Feb-2015. [Online]. Available: <http://ciberconta.unizar.es/biblioteca/0004/SerGall96.html>. [Accessed: 05-Feb-2015].
- [51] Alfonso Palmer Pol and Juan José Montañó Moreno, "REDES NEURONALES ARTIFICIALES APLICADAS AL ANÁLISIS DE SUPERVIVENCIA: ANÁLISIS COMPARATIVO CON EL MODELO DE REGRESIÓN DE COX EN SU ASPECTO PREDICTIVO," 05-Feb-2015. [Online]. Available: <http://www.psicothema.com/psicothema.asp?id=775>. [Accessed: 05-Feb-2015].
- [52] Ramírez Q. Juan A. and Chacón M. Mario I., "Redes neuronales artificiales para el procesamiento de imágenes, una revisión de la última década," vol. 9, Jul. 2001.
- [53] "Árbol de decisión - Descuadrando." [Online]. Available: http://descuadrando.com/%C3%81rbol_de_decisi%C3%B3n. [Accessed: 07-May-2015].
- [54] Jorge Enrique Albelo Mengana, Ramón Quiza Sardiñas, and Marcelino Rivas Santana, "Elementos de minería de datos." Nov-2007.
- [55] "Redes neuronales ventajas y desventajas," 16-Dec-2014. [Online]. Available: [.http://egkafati.bligoo.com/content/view/184582/Redes-neuronales-ventajas-y-desventajas.html](http://egkafati.bligoo.com/content/view/184582/Redes-neuronales-ventajas-y-desventajas.html). [Accessed: 16-Dec-2014].
- [56] McClelland, J and Rumelhart D. E., *The PDP Research Group. Parallel Distributed Processing*, vol. 1 and 2. MITPress, 1986.
- [57] Corchado, Emilio and Fyfe Colin, *Introducción a las Minería de Datos, Capítulo 13*. Prentice Hall, Pearson, 2009.
- [58] Breiman, L, Friedman, J.H., Olshen, RA., and Stone, e.J., *Classification and Regression Trees*. Wadsworth and Brooks/Cole, Monterey, 1984.
- [59] Quinlan, J.R, Michalski, R; Carbonell, and J.; Mitchell, T, *Learning Efficient Classification Procedures and Their Application to Chess End Games*. Machine Learning: An Artificial Intelligence Approach. Morgan Kaufmann, San Mateo, CA, 1983.
- [60] Quinlan, J.R, *Induction of Decision Trees*, vol. 1. Machine Learning, 1986.
- [61] "C4.5. Programs for Machine Learning." Morgan Kaufmann, San Francisco, EEUU, 1993.
- [62] Cestnik, G., Kononenko, I., Bratko, I., and Lavrac, N, "Assistant 86: A Knowledge Acquisition Tool for Sophisticated Users." *Progress in Machine Learning*, Sigma Press, 1987.
- [63] Buntine, W., "Learning Classification Trees Statistics and Computing." 1992.
- [64] Brodley, e. E. and Utgoff, P. E, "Multivariate decision trees," *Mach. Learn.*, vol. 19, pp. 45– 77, 1995.
- [65] Rafiuzaman, M., *Forecasting Chaotic Stock Market Data using Time Series Data Mining*, vol. 101. International Journal of Computer Applications (0975 – 8887), 2014.
- [66] S.K. Shevade, S.S. Keerthi, C. Bhattacharyya, and K.R.K. Murthy, "Improvements to the SMO algorithm for SVM regression." In *IEEE Transactions on Neural Networks*, 1999.

Bibliografía

- [1] Chase R, Jacobs R, and Aquilano N, *Administración de operaciones. Producción y cadena de suministros*. Editorial Mc Graw Hill, 2010.
- [2] Valadez, E.H, "Algoritmo de clustering basado en entropía para descubrir grupos en atributos de tipo mixto," Sección de Computación, Centro de Investigación y de Avanzados del Instituto Politécnico Nacional, DF, México, 2006.
- [3] Ing. Yadira Robles Aranda, MSc. Anthony Rafael Sotolongo, and Ing. Edgar Rojas, "Algoritmos de minería de datos: Árboles de decisión y reglas de inducción integrados a PostgreSQL." 2012.
- [4] PaolaAndreaCardonaHernández, "Aplicación de árboles de decisión en modelos de riesgo crediticio," 07-May-2015. [Online]. Available: <http://www.redalyc.org/articulo.oa?id=89927204>. [Accessed: 07-May-2015].
- [5] MARIA AUXILIADORA PEREZ, "APLICACIÓN DE MINERÍA DE DATOS PARA LA PREDICCIÓN DEL RENDIMIENTO DE LA CAÑA DE AZÚCAR," UNIVERSIDAD CENTROCCIDENTAL "LISANDRO ALVARADO", Barquisimeto, Venezuela, 2005.
- [6] Lucila L. Chiarvetto Peralta, Fernando A. Rey Saravia, and Nélide B. Brignole, "APLICACIÓN DE REDES NEURONALES ARTIFICIALES PARA LA PREDICCIÓN DE CALIDAD DE AIRE," vol. 27, Nov. 2008.
- [7] Quiñones, Yessica, "Aplicación de técnicas de Análisis de Cluster a la exploración de los Estilos de Aprendizaje en estudiantes de Ingeniería Informática," Universidad Carlos Rafael Rodríguez, Cienfuegos, 2011.
- [8] Sotolongo, G., Guzmán, and Maria Victoria, "Aplicaciones de las redes neuronales. El caso de la Bibliometría." 2001.
- [9] "Árbol de decisión - Descuadrando," 07-May-2015. [Online]. Available: http://descuadrando.com/%C3%81rbol_de_decisi%C3%B3n. [Accessed: 07-May-2015].
- [10] "Árbol de decisión - EcuRed," 07-May-2015. [Online]. Available: http://www.ecured.cu/index.php/%C3%81rbol_de_decisi%C3%B3n. [Accessed: 07-May-2015].
- [11] "Árbol de decisión, una herramienta para decidir bien | Alto Nivel," 07-May-2015. [Online]. Available: <http://www.altonivel.com.mx/36690-arbol-de-decision-una-herramienta-para-decidir-correctamente.html>. [Accessed: 07-May-2015].
- [12] nebrija, "Árboles de decisión," 2014.
- [13] SECURED, "Árboles de decisión definición." 2014.
- [14] "árboles De Decisión Weka," 07-May-2015. [Online]. Available: <http://es.slideshare.net/lecastillox/rboles-de-decisin-weka>. [Accessed: 07-May-2015].
- [15] Cestnik, G., Kononenko, I., Bratko, I., and Lavrac, N, "Assistant 86: A Knowledge Acquisition Tool for Sophisticated Users." Progress in Machine Learning, Sigma Press, 1987.
- [16] Ministerio de Educación y Ciencia (C.I.D.E.), "Avance datos diferencias alumnos control/experimentación en la Reforma de las Enseñanzas Medias." Manuscrito no publicado, 1986.
- [17] "C4.5. Programs for Machine Learning." Morgan Kaufmann, San Francisco, EEUU, 1993.
- [18] Breiman, L, Friedman, J.H., Olshen, R.A., and Stone, C.J., *Classification and Regression Trees*. Wadsworth and Brooks/Cole, Monterey, 1984.

- [19] Luis Llano and MSc.(c) Andrés Hoyos, "Comparación del Desempeño de Funciones de Activación en Redes Feedforward para aproximar Funciones de Datos con y sin Ruido," p. 10, Jul. 2007.
- [20] Pérez Serrano. G, "Crítica al concepto de rendimiento académico," *Revista Española de Pedagogía*, vol. 174, pp. 521–534.
- [21] Ian H. Witten and Eibe Frank, *Data Mining Practical Machine Learning Tools and Techniques*, Segunda. 2005.
- [22] "Definición de inteligencia artificial - Qué es, Significado y Concepto," 2015. [Online]. Available: <http://definicion.de/inteligencia-artificial/>. [Accessed: 05-Feb-2015].
- [23] José Alberto Chang García, "Determinación de los Estilos de Enseñanza predominantes en profesores de Ingeniería Informática aplicando técnicas de Análisis de Clúster," Universidad de Cienfuegos, Cuba, 2012.
- [24] P.A. Castillo, J.G. Castellano, J.J. Merelo, and A. Prieto, "Diseño de redes neuronales artificiales mediante algoritmos evolutivo," vol. 5, no. 14, 2001.
- [25] ESCOM, "Diseño y Entrenamiento de Redes Neuronales Artificiales," 2015. [Online]. Available: <http://es.slideshare.net/mentelibre/diseo-y-entrenamiento-de-redes-neuronales-artificiales>. [Accessed: 27-Jan-2015].
- [26] J. Alcalá-Fdez, L. Sánchez, S. García, M.J. del Jesus, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Bacardit, V.M. Rivas, J.C. Fernández, and F. Herrera, "Documento de descripción de KEEL." 2009.
- [27] Jorge Enrique Albelo Mengana, Ramón Quiza Sardiñas, and Marcelino Rivas Santana, "Elementos de minería de datos." Nov-2007.
- [28] "ELEMENTOS DE UNA RED NEURONAL," 2014. [Online]. Available: <http://www.hugo-inc.com/RNA/Unidad%201/1.6.html>. [Accessed: 05-Feb-2015].
- [29] Sancho, F.V.N., "Estilo de Aprendizaje del Alumno y Técnicas para mejorar su aprendizaje," Jun. 2009.
- [30] Gómez Molina, J. R, "Estudio descriptivo de las evaluaciones de B.U.P. y C.O.U. en el país valenciano." Bardón, 1982.
- [31] Rodríguez Espinar, "Factores de rendimiento escolar," *Oikos-Tau*, 1982.
- [32] Rafiuzaman, M., *Forecasting Chaotic Stock Market Data using Time Series Data Mining*, vol. 101. International Journal of Computer Applications (0975 – 8887), 2014.
- [33] "Funciones en el modelo de Neurona Artificial," 2014. [Online]. Available: <http://www.redes-neuronales.com.es/tutorial-redes-neuronales/funciones-de-las-neuronas-artificiales.htm>. [Accessed: 05-Feb-2015].
- [34] Stanton WJ, Etzel MJ, and Welker BJ, "Fundamentos del marketing." McGraw-Hil, 2004.
- [35] J.L. CUBERO, F. BERZAL, and F. HERRERA, "FUNDAMENTOS DE MINERÍA DE DATOS." 2005.
- [36] A LIRIO MAYORGA DELGADO and HECTOR JULIO PEDROZA, "HERRAMIENTA DE SOFTWARE PARA ENTRENAR Y SIMULAR REDES NEURONALES ARTIFICIALES," UNIVERSIDAD FRANCISCO DE PAULA SANTANDER, SAN JOSE DE CUCUTA, COLOMBIA, 2003.
- [37] "Historia de las redes neuronales," 2012. [Online]. Available: <http://ingenieria.udea.edu.co/investigacion/mecatronica/mecatronica/redes.htm>. [Accessed: 05-Feb-2015].
- [38] S.K. Shevade, S.S. Keerthi, C. Bhattacharyya, and K.R.K. Murthy, "Improvements to the SMO algorithm for SVM regression." In *IEEE Transactions on Neural Networks*, 1999.

- [39] María E. Valesani, Osvaldo P. Quintana, and Oscar A. Vallejos, "Imputación de datos con redes neuronales." 07-Sep-1999.
- [40] Quinlan, J.R, *Induction of Decision Trees*, vol. 1. Machine Learning, 1986.
- [41] "INFOFICS N°32," 16-Dec-2014. [Online]. Available: <http://www.usmp.edu.pe/publicaciones/boletin/fia/info32/pag4.htm>. [Accessed: 16-Dec-2014].
- [42] "Inteligencia artificial - EcuRed," 2015. [Online]. Available: http://www.ecured.cu/index.php/Inteligencia_artificial. [Accessed: 05-Feb-2015].
- [43] Limance, "Inteligencia artificial - Monografias.com," 2014. [Online]. Available: <http://www.monografias.com/trabajos16/la-inteligencia-artificial/la-inteligencia-artificial.shtml>. [Accessed: 05-Feb-2015].
- [44] Rubén Cárdenes Almeida, "Inteligencia Artificial Practica 2: Redes Neuronales." 2011.
- [45] Ing. Edgar Rojas, "Introducción a árboles de decisión," 2014.
- [46] J. Hernández, "Introducción a la Minería de Datos." Educación Pearson, Madrid, España, 2004.
- [47] Corchado, Emilio and Fyfe Colin, *Introducción a las Minería de Datos, Capítulo 13*. Prentice Hall, Pearson, 2009.
- [48] Castellano, "Introducción a las redes neuronales aplicadas." 2013.
- [49] "Introducción a las redes neuronales artificiales," 2013. [Online]. Available: http://magomar.webs.upv.es/rna/tutorial/RNA_intro.html. [Accessed: 05-Feb-2015].
- [50] Patricio García Báez, "Introducción a las Redes Neuronales y su aplicación a la Investigación Astrofísica," Instituto de astrofísica de Canarias, España, 2014.
- [51] "KEEL 2." .
- [52] "KEEL. es." 2010.
- [53] Derrac, "KEEL Una herramienta de software para el análisis de sistemas." 2010.
- [54] "KNIME 2.1.0 lanzado." .
- [55] Félix de Moya Anegón, Víctor Herrero Solana, and Vicente Guerrero Bote, "La aplicación de Redes Neuronales Artificiales (RNA): a la recuperación de la información." 2005.
- [56] "La educación en Cuba," 07-May-2015. [Online]. Available: <http://www.granma.cu/granmad/secciones/ref-fidel/art33.html>. [Accessed: 07-May-2015].
- [57] Vázquez Alonso, A, "La enseñanza de la Física (y la Química) en la E.G.B." Revista Española de Física, 1989.
- [58] Reparaz, C., Touron, J., and Villanueva, C, "La predicción del rendimiento académico en el Curso de Orientación Universitaria." Revista Eyraila de Pedagogía., 1987.
- [59] Carlos Serrano Cinca and José Luis Gallizo Larraz, "Las redes neuronales artificiales en el tratamiento de la información financiera," 05-Feb-2015. [Online]. Available: <http://ciberconta.unizar.es/biblioteca/0004/SerGall96.html>. [Accessed: 05-Feb-2015].
- [60] Felder, R. and L. Silverman, *Learning and Teaching Styles in Engineering Education Application, in Engr. Education*. 1988.
- [61] Felder, R. and R. Brent, *Learning and teaching styles in engineering education. Engineering Education Journal*. 2005.
- [62] Buntine, W., "Learning Classification Trees Statistics and Computing." 1992.
- [63] Quinlan, J.R, Michalski, R; Carbonell, and J.; Mitchell, T, *Learning Efficient Classification Procedures and Their Application to Chess End Games*. Machine

- Learning: An Artificial Intelligence Approach. Morgan Kaufmann, San Mateo, CA, 1983.
- [64] Carranza O and Sabrià F, *Logística, mejores prácticas en Latinoamérica*. Thompson Mexico, 2005.
- [65] Diego García Morate, "Manual de Weka." 2012.
- [66] "MATLAB - EcuRed," 2015. [Online]. Available: <http://www.ecured.cu/index.php/MATLAB>. [Accessed: 05-Feb-2015].
- [67] Durán, E. and R. Costagauta, *Minería de datos para descubrir estilos de aprendizaje*. Revista Iberoamericana de Educación, 2007.
- [68] Roberto Suárez Surí, "Modelo basado en aprendizaje automático para el pronóstico al egreso de pacientes con trauma grave de cráneo," Universidad de Cienfuegos, Cuba, 2013.
- [69] Brodley, e. E. and Utgoff, P. E, "Multivariate decision trees," *Machine Learning*, vol. 19, pp. 45– 77, 1995.
- [70] Prof. Andrés E Reyes Polanco., "NOCIONES DE INTELIGENCIA ARTIFICIAL." 2013.
- [71] Carabaña, J, "Origen social, inteligencia y rendimiento académico al final de la EGB," p. 2014.
- [72] Pérez Serrano. G, "Origen social y rendimiento escolar Madrid," *Centro de Investigaciones Sociológicas*, 1981.
- [73] De Toro and Gisbert M, *Pequeño Larousse Ilustrado*. librería Larousse, Paris, Francia, 1926.
- [74] "Portal de la Educación Cubana," 07-May-2015. [Online]. Available: <http://www.rimed.cu/>. [Accessed: 07-May-2015].
- [75] "PrecisionTree para análisis de decisión con árboles de decisión - Palisade," 07-May-2015. [Online]. Available: <http://www.palisade-lta.com/precisiontree/>. [Accessed: 07-May-2015].
- [76] Carlos Serrano Cinca and Bonifacio Martin del Brio, "Predicción de la quiebra bancaria mediante el empleo de las redes neuronales artificiales," vol. 23, no. 74, 1993.
- [77] Edgar Pérez Rivero, "PREDICCIÓN DE RESULTADOS ACADÉMICOS EN LA ASIGNATURA SO HACIENDO USO DE REDES NEURONALES," Instituto Superior Politécnico "José Antonio Echeverría," 2014.
- [78] Jorge Álvarez Blanco, "Predicción de resultados académicos en las asignaturas Estructuras de Datos 1 y 2," Instituto Superior Politécnico José Antonio Echeverría, 2014.
- [79] Weiss, S. M and N. Indurkha, *Predictive data mining: A practical guide*. San Francisco, EEUU, 1998.
- [80] Softonic, "Pythia," 2014. [Online]. Available: <http://pythia.softonic.com/>. [Accessed: 05-Feb-2015].
- [81] Roger Loadiza, "¿Qué es la inteligencia artificial?," 2015. [Online]. Available: http://bvs.sld.cu/revistas/san/vol2_2_98/san15298.htm. [Accessed: 05-Feb-2015].
- [82] "¿Qué son las redes neuronales," 2002. [Online]. Available: http://info.fisica.uson.mx/arnulfo.castellanos/archivos_html/quesonredneu.htm. [Accessed: 16-Dec-2014].
- [83] Avellano, "Redes Neuronales," 2014. [Online]. Available: <http://avellano.usal.es/~lalonso/RNA/>. [Accessed: 27-Jan-2015].
- [84] Daniel Ochoa, "Redes Neuronales," 01-Jun-2009. [Online]. Available: <http://es.slideshare.net/lennydan11/redes-neuronales-1514948>. [Accessed: 16-Dec-2014].

- [85] Fernando Izarrieta and Carlos Zaavedra, "Redes Neuronales Artificiales," Universidad de Concepcion, Chile, 2012.
- [86] Alfonso Palmer Pol and Juan José Montaña Moreno, "REDES NEURONALES ARTIFICIALES APLICADAS AL ANÁLISIS DE SUPERVIVENCIA: ANÁLISIS COMPARATIVO CON EL MODELO DE REGRESIÓN DE COX EN SU ASPECTO PREDICTIVO," 05-Feb-2015. [Online]. Available: <http://www.psicothema.com/psicothema.asp?id=775>. [Accessed: 05-Feb-2015].
- [87] "Redes neuronales artificiales Fundamentos, modelos y aplicaciones - Monografias.com," 2015. [Online]. Available: <http://www.monografias.com/trabajos12/redneur/redneur.shtml>. [Accessed: 05-Feb-2015].
- [88] Juan A. Botía, "Redes neuronales artificiales para aprendizaje automático y minería de datos," Universidad de Murcia, España, 10-Apr-2007.
- [89] Ramírez Q. Juan A. and Chacón M. Mario I., "Redes neuronales artificiales para el procesamiento de imágenes, una revisión de la última década," vol. 9, Jul. 2001.
- [90] Xabier Basogain Olabe, "Redes Neuronales Artificiales y sus aplicaciones." 2000.
- [91] Damián Jorge Matich, "Redes Neuronales: Conceptos Básicos y Aplicaciones.," Universidad Tecnológica Nacional – Facultad Regional Rosario, Departamento de Ingeniería Química, Argentina, 2001, p. 55.
- [92] "Redes neuronales - EcuRed," 2014. [Online]. Available: http://www.ecured.cu/index.php/Redes_neuronales. [Accessed: 27-Jan-2015].
- [93] "Redes neuronales ventajas y desventajas," 16-Dec-2014. [Online]. Available: <http://egkafati.bligoo.com/content/view/184582/Redes-neuronales-ventajas-y-desventajas.html>. [Accessed: 16-Dec-2014].
- [94] Ed. Félix Varela, MES, *Reglamento de Organización Docente, Resolución no. 120/2010*, Ed. 2010, p. 22.
- [95] Vázquez Alonso, A., "Rendimiento en Bachillerato: Aptitudes y atribución causal. Análisis del rendimiento objetivo en Física y Química," Universidad de las Islas Baleares, 1989.
- [96] María Itati Araujo, Santiago Iván Ferrari Alve, and Sonia I. Mariño, "Revisión de algoritmos de Redes Neuronales en dos herramientas de Minería de Datos," *CyTA*, vol. 11, no. 4, Jun. 2012.
- [97] Alfonso Palmer Pol, "SENSITIVITY NEURAL NETWORK." 2014.
- [98] "Sensitivity Neural Network 1.0 (SNN); programa informático para la simulación del comportamiento de las Redes Neuronales Artificiales (RNA)," 2015. [Online]. Available: http://tecnologiauib.com/es/advanced_knowledge/show/sensitivity-neural-network-1-0-snn-programa-informatico-para-la-simulacion-del-comportamiento-de-las-redes-neuronales-artificiales-rna/23. [Accessed: 05-Feb-2015].
- [99] Sergio Moriello, "SISTEMAS COMPLEJOS, CAOS Y VIDA ARTIFICIAL." 2009.
- [100] Florentino Fdez-Riverola and Juan M. Corchado, "Sistemas híbridos neuro-simbólicos: una revisión." 2012.
- [101] Universidad de Oviedo, "Sistemas inteligentes Árboles de decisión," 2014.
- [102] Javier De Andrés Suárez, "TÉCNICAS DE INTELIGENCIA ARTIFICIAL APLICADAS AL ANÁLISIS DE LA SOLVENCIA EMPRESARIAL." 2013.
- [103] "Temas de investigación (1, educativa, Madrid Servido de Publicaciones M.E.C.," *MEC-1NCIE*, pp. 29–72, 1979.
- [104] "Tes de pronóstico académico." 2014.
- [105] McClelland, J and Rumelhart D. E., *The PDP Research Group. Parallel Distributed Processing*, vol. 1 and 2. MITPress, 1986.

- [106] Natalia L. Weinbach and Carlos I. Ches~nevar, "Un Primer Acercamiento al Aprendizaje del Codigo Genetico usando Redes Neuronales." 2012.
- [107] "Weka - EcuRed," 05-Feb-2015. [Online]. Available: <http://www.ecured.cu/index.php/Weka>. [Accessed: 05-Feb-2015].
- [108] Abhishek Tiwaria and Arvind K.T. Sekhar, *Workflow based framework for life science informatics*, vol. 31. 2007.

Glosario de Términos

Pronóstico.

Conjetura acerca de lo que puede suceder, señal por donde se conjetura una cosa futura.

Aprendizaje.

Que aprende adquiriendo conocimiento del entorno u otra fuente.

Automático.

Algo que hace algo por si solo.

Anexos.

Anexo A.

Anexo A1.1- Modelo 1er Año.

@relation 'Modelo_1er año'

@attribute 'Mat I' numeric
@attribute FyS numeric
@attribute FInf numeric
@attribute IP numeric
@attribute 'EF I ' numeric
@attribute 'Id0 I' numeric
@attribute 'Alg y GA' numeric
@attribute Hluba numeric
@attribute 'Id0 II' numeric
@attribute DPOO numeric
@attribute 'EF II' numeric
@attribute IGS numeric
@attribute MD numeric
@attribute 'Mat II' numeric
@attribute EPlap numeric
@attribute aprueba {si,no}

@data

5,5,4,5,5,5,4,3,4,5,5,4,4,4,5,si
3,4,4,4,5,5,3,4,4,5,5,4,3,3,3,si
4,3,4,2,5,4,3,5,3,3,5,3,3,3,4,no
3,4,4,3,5,4,4,4,3,5,5,3,3,3,3,si
3,4,4,2,5,4,3,4,3,3,5,3,3,3,4,no
3,4,4,3,5,5,3,4,3,3,5,3,3,3,4,no
3,5,5,5,5,5,4,5,3,4,5,4,3,4,4,si
3,4,4,2,5,5,3,3,3,2,5,3,3,2,4,no
2,3,4,2,5,4,3,3,3,2,5,3,3,2,4,no
3,5,5,3,5,4,3,5,3,4,5,4,3,5,4,si
3,5,5,3,5,5,4,4,4,3,5,3,3,4,4,si
3,4,4,3,5,4,3,4,3,3,5,4,3,4,3,no
4,4,5,5,5,5,3,4,5,5,5,4,4,3,3,si
4,4,4,3,5,5,3,4,5,3,5,3,3,5,3,si
4,4,4,4,5,4,3,3,3,4,5,3,3,3,3,si
3,4,4,3,5,5,3,5,4,3,5,3,3,4,4,si
3,4,4,3,5,4,3,4,3,3,5,3,3,4,4,si
3,4,4,4,5,5,4,5,4,3,5,3,3,3,4,si
3,5,4,2,5,4,3,4,4,3,5,3,3,4,3,no
2,3,4,2,5,4,2,3,3,2,5,4,2,2,3,no
3,3,5,3,5,4,3,4,3,3,5,3,3,3,3,si
5,5,4,4,5,5,4,4,3,5,5,4,4,5,4,si
3,3,4,3,5,5,3,5,3,3,5,4,3,3,4,si
3,4,5,3,5,5,4,5,4,5,5,3,3,4,4,si
3,5,4,2,5,4,4,5,3,3,5,3,3,3,5,no
3,4,4,3,5,5,3,5,3,3,5,4,2,3,4,si
3,3,4,2,5,4,2,3,3,3,5,3,3,3,3,no

3,4,5,4,5,3,3,3,3,3,5,4,3,3,3,si
3,0,4,3,3,4,3,5,4,4,5,4,3,2,0,no
3,4,4,3,5,4,3,4,4,3,5,5,3,3,4,si
3,0,4,3,5,4,3,5,4,4,5,4,3,2,0,no
3,4,5,5,3,5,3,4,4,4,5,3,3,4,3,si
3,4,5,3,5,3,3,5,3,3,5,5,3,3,4,si
3,0,4,3,5,5,3,5,5,4,5,5,3,3,0,si
3,5,5,3,5,4,3,5,4,4,5,5,3,3,4,si
3,4,4,3,5,3,3,3,4,3,5,3,3,3,4,si
3,5,5,5,5,5,3,4,5,5,5,5,3,3,3,si
4,0,5,3,5,0,4,5,0,4,5,5,5,3,0,si
3,4,4,3,5,3,3,4,3,3,5,4,3,3,3,si
3,5,5,3,5,3,4,5,4,5,5,3,4,4,5,si
3,4,5,3,5,3,4,3,3,3,5,3,3,3,3,si
3,5,5,4,5,4,5,5,4,5,5,5,5,4,5,si
3,4,4,3,4,3,3,3,3,4,5,3,3,3,4,si
3,5,5,4,5,4,3,4,4,3,5,4,3,3,4,si
3,4,5,4,5,4,4,4,4,4,5,4,3,4,4,si
4,0,5,5,5,5,4,5,5,5,5,5,5,3,0,si
3,4,4,3,?,4,3,4,4,4,5,3,4,3,3,si
2,4,5,4,5,3,3,5,3,3,5,4,2,3,3,si
3,4,5,4,5,4,4,4,4,5,5,5,4,3,3,si
3,4,5,4,4,3,3,4,3,4,5,4,4,3,4,si
3,5,5,5,4,3,3,4,3,5,5,4,4,3,4,si
3,5,5,3,5,4,3,4,4,3,5,4,3,3,4,si
3,5,5,4,5,4,4,5,4,5,5,5,4,3,4,si
4,5,5,4,5,4,5,5,4,5,5,4,5,3,4,si
3,4,5,3,5,4,3,5,4,3,5,3,3,3,4,si
3,4,5,4,5,4,3,4,4,3,5,3,3,3,3,si
2,4,5,4,5,4,3,5,4,3,5,4,2,3,4,no
3,4,5,3,5,4,4,4,3,3,5,3,4,3,4,si
3,4,4,3,5,4,3,4,4,4,5,4,3,3,4,si
2,4,5,3,3,3,3,4,3,3,5,3,2,3,3,no
3,4,5,3,4,4,3,4,4,4,5,5,3,3,4,si
4,3,5,4,4,4,3,4,4,5,5,3,3,4,5,si
2,4,5,4,5,4,3,4,3,4,5,3,3,3,5,no
4,5,5,5,5,4,4,5,4,5,5,5,4,5,5,si
3,4,4,3,5,4,3,5,4,4,5,3,3,3,5,si
3,3,4,3,4,4,3,4,4,3,5,4,3,3,3,si
5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,si
3,?,4,3,5,4,3,4,4,4,5,4,3,3,4,si
3,4,5,5,5,3,3,5,4,5,5,5,3,3,5,si
2,4,4,3,5,3,3,4,3,3,5,4,3,3,5,no
3,3,4,3,5,4,3,4,3,3,5,3,3,3,3,si
5,5,5,4,5,4,4,4,4,5,5,4,4,5,5,si
3,3,5,5,5,4,3,4,4,4,5,5,3,3,4,si
4,5,5,5,5,4,3,4,4,5,5,5,4,4,3,si
3,3,3,3,5,3,3,4,3,4,5,4,3,3,4,si
3,3,4,3,5,3,3,4,3,4,5,5,3,3,4,si
3,4,5,5,5,5,3,5,4,4,5,4,3,3,5,si
2,3,5,4,5,5,3,4,4,3,5,3,3,2,5,no
3,4,4,3,5,3,3,5,3,4,5,5,4,3,5,si
3,5,5,4,5,5,3,5,5,4,5,5,4,3,5,si
2,4,5,4,5,3,3,5,3,5,5,5,3,3,3,no
2,4,4,3,4,4,3,4,4,3,5,5,3,3,4,no

3,5,4,3,5,4,3,5,4,4,5,5,3,3,5,si
2,3,4,3,5,3,3,4,3,3,5,4,2,2,5,no
2,3,4,3,4,4,3,4,4,3,5,4,2,2,3,no
5,5,5,5,5,5,5,5,5,5,5,5,4,4,5,si
3,0,3,3,5,3,3,3,5,3,5,5,3,3,0,si
3,0,3,3,4,3,3,3,4,3,5,3,3,3,0,si
5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,si
1,3,5,1,1,4,1,5,4,1,1,5,1,3,5,si
3,3,5,4,5,4,3,5,4,5,5,5,3,3,5,si
2,4,5,3,5,3,3,4,3,3,5,4,2,2,4,no
2,3,4,3,5,3,3,4,3,3,5,4,2,2,3,no
4,0,3,3,5,4,4,4,5,4,5,4,3,4,0,si
2,1,1,1,1,1,1,1,1,1,1,1,3,3,1,no
2,2,2,2,?,?,2,?,?,?,?,?,?,no
1,3,5,1,1,5,1,5,5,3,1,4,1,3,4,si
3,4,4,3,5,5,4,4,5,4,5,5,4,3,4,si
2,1,1,1,1,1,1,1,1,1,1,1,?,1,1,no
3,4,5,3,5,4,3,5,5,5,5,5,3,3,5,si
3,3,4,4,5,4,3,4,4,4,5,4,2,3,4,no
3,3,4,3,5,3,3,3,3,3,5,4,3,3,4,si
2,2,5,3,5,3,3,3,2,4,5,4,2,2,5,no
2,4,5,3,5,?,3,5,4,4,5,4,3,3,4,no
4,4,3,5,5,4,4,4,5,5,5,5,4,5,4,si
2,2,3,3,5,4,3,4,3,3,5,4,2,2,4,no
2,3,4,3,3,4,3,4,4,4,5,4,3,3,3,no
1,3,5,1,1,4,1,4,4,?,1,3,1,3,5,si
4,5,5,5,5,4,3,5,5,5,5,5,3,3,5,si
2,3,4,3,5,3,3,4,3,3,5,4,2,2,3,no
2,2,3,3,5,3,3,4,3,3,5,4,2,2,3,no
2,2,4,3,4,4,3,4,4,3,5,4,3,2,3,no
4,5,5,5,5,5,3,5,5,5,5,3,3,3,4,si
2,4,5,3,5,4,3,4,4,3,5,4,3,2,4,no

Anexo A1.2- Modelo 2do Año.

```
@relation 'Modelo 2do año'

@attribute 'Mat III' numeric
@attribute Epcap numeric
@attribute 'ED I' numeric
@attribute IIA numeric
@attribute 'IdEx III' numeric
@attribute SegNac numeric
@attribute 'EF 3' numeric
@attribute 'elec ' numeric
@attribute 'Mat IV' numeric
@attribute Argomp numeric
@attribute 'ED 2' numeric
@attribute BD numeric
@attribute TSP numeric
@attribute 'IdEx IV' numeric
@attribute DeNac numeric
@attribute 'EF iv' numeric
@attribute Pprof numeric
@attribute aprueba {no,si}

@data
2,3,2,2,3,2,4,3,3,3,2,2,3,3,2,4,2,no
2,3,2,2,3,3,5,4,3,2,2,3,4,4,3,5,2,no
2,2,2,2,4,3,5,3,3,2,2,2,3,4,3,5,2,no
4,3,4,3,4,4,4,5,5,3,5,4,5,4,5,5,5,si
2,3,3,3,3,3,5,4,3,3,3,3,3,4,4,5,4,si
2,3,2,2,3,3,5,4,3,2,3,3,3,3,4,5,4,no
2,2,2,2,3,3,5,3,2,2,2,2,2,4,2,5,2,no
2,4,2,2,4,4,4,4,?, ?, ?, ?, ?, ?, ?, 5, ?, no
3,2,3,3,3,3,4,4,3,3,3,4,3,4,3,4,5,no
3,3,4,3,3,5,3,4,2,2,3,2,3,3,3,3,2,no
3,3,3,3,4,4,4,4,3,3,4,4,5,4,3,5,4,si
2,3,3,3,3,3,5,4,3,3,3,2,3,4,3,4,5,no
3,3,4,3,3,3,2,3,3,4,3,3,4,4,3,3,5,no
3,5,2,3,3,5,4,4,3,3,2,3,3,3,4,4,2,no
2,2,2,2,5,3,4,3,2,2,2,2,2,5,3,5,5,no
2,2,2,2,4,5,4,3,?, ?, ?, ?, ?, ?, ?, 5, ?, no
2,0,2,2,3,0,3,5,2,2,2,2,0,?, 0,3,2,no
3,4,2,3,3,3,3,4,3,3,2,4,3,3,3,3,4,no
3,3,3,3,5,4,4,3,3,3,3,2,3,5,3,0,5,si
3,3,3,3,4,3,4,3,3,3,3,3,4,4,4,4,5,si
3,4,5,3,3,3,5,5,4,5,4,3,3,4,3,4,5,si
2,3,3,3,4,4,4,4,3,3,3,3,3,4,4,3,5,no
3,3,3,3,3,3,5,4,3,3,5,4,3,4,4,5,5,si
2,2,3,2,3,2,5,3,?, ?, ?, ?, ?, ?, ?, ?, no
3,3,4,3,4,3,4,3,3,3,4,4,3,4,2,4,5,no
2,3,3,3,4,5,5,4,3,3,3,3,3,4,4,4,4,no
3,3,3,3,4,4,5,4,3,3,3,4,3,4,5,5,5,si
2,3,3,2,3,3,5,4,3,2,3,3,3,3,3,5,4,no
```

1,4,2,2,3,1,1,1,3,2,2,3,4,4,1,3,4,no
3,4,3,3,3,3,5,3,3,3,2,4,3,4,4,5,4,no
2,5,2,2,4,3,3,4,3,3,2,2,5,4,4,4,5,no
3,3,4,3,4,3,5,4,3,3,4,3,3,4,4,5,4,si
3,1,2,3,1,3,1,4,3,3,2,2,1,1,4,1,5,no
?,4,2,2,4,?,5,3,?,?,?,?,?,?,no
3,3,2,3,3,0,4,5,3,2,2,3,3,?,0,4,3,no
1,1,3,3,1,1,1,1,1,1,3,1,5,1,1,1,1,si
3,3,3,3,3,4,5,4,3,3,3,4,3,4,4,4,4,si
3,3,3,3,4,3,5,4,3,3,3,4,3,4,5,5,5,si
1,1,2,3,1,1,1,1,1,3,3,4,3,1,1,1,1,no
3,3,3,4,3,2,4,3,3,3,3,4,3,4,2,4,3,no
2,3,3,3,3,3,5,4,3,4,3,3,4,4,3,5,3,no
3,5,3,3,5,4,5,4,4,4,3,4,4,4,5,5,3,si
3,4,3,4,4,4,5,5,4,3,3,5,4,5,5,5,5,si
5,5,3,4,4,5,5,5,5,4,4,5,5,5,5,5,5,si
3,0,3,3,4,0,5,4,4,4,4,4,0,4,0,5,5,si
3,0,3,3,3,0,5,4,4,3,3,4,0,3,0,5,4,si
5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,si
3,4,3,3,4,5,5,5,4,3,3,4,5,5,5,5,5,si
3,3,3,2,5,3,4,3,3,3,2,4,4,5,3,5,4,no
5,0,3,3,4,0,5,4,4,3,4,4,0,4,0,5,4,si
4,4,3,3,5,5,5,5,5,4,3,5,5,5,5,5,5,si
4,4,3,4,5,4,5,4,4,3,3,5,4,5,5,5,5,si
3,4,3,3,5,3,5,4,4,3,3,4,4,4,5,5,5,si
3,3,3,3,4,4,4,4,3,3,3,3,3,4,4,3,5,si
5,5,5,4,5,3,5,4,5,5,5,5,5,5,5,5,5,si
3,3,3,3,3,3,5,4,3,4,3,3,3,3,3,5,4,si
3,5,5,3,5,5,5,5,4,3,5,3,5,5,5,5,4,si
3,5,3,3,5,3,5,4,3,3,3,5,4,5,4,5,5,si
4,5,5,4,5,5,5,4,4,3,4,4,4,5,5,5,4,si
3,0,3,3,3,0,4,5,3,4,3,3,3,3,0,4,3,si
4,4,4,4,4,3,4,?,3,?,5,3,4,4,3,5,3,si
4,3,3,3,5,4,5,4,3,3,3,3,4,5,4,5,5,si
2,4,3,3,5,4,5,4,3,3,3,2,4,4,4,5,3,no
3,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,si
3,3,3,3,5,3,?,3,3,3,3,2,3,5,4,5,3,no
2,5,3,3,4,4,5,4,3,3,3,2,5,4,4,5,3,no
4,5,5,5,5,5,5,5,5,5,5,4,5,5,5,5,5,si
4,3,4,3,3,5,3,4,4,3,?,5,3,3,3,3,4,si
3,4,3,3,5,4,5,3,3,3,3,3,5,5,4,5,4,si
2,4,5,3,4,3,?,3,3,4,4,2,3,4,4,5,5,no
3,5,3,3,4,5,5,3,5,4,3,3,5,5,5,5,4,si
3,4,3,3,4,3,?,3,3,3,3,3,3,4,4,5,5,si
3,5,5,5,5,5,5,4,5,5,5,5,4,5,5,5,4,si
3,4,3,3,5,4,5,3,4,4,4,4,5,4,4,5,4,si
3,5,3,3,4,5,5,5,5,3,4,4,5,5,5,5,5,si
3,4,3,3,5,3,5,5,4,3,3,4,1,4,5,5,?,si
3,4,3,3,4,4,4,4,4,4,3,3,2,4,4,5,5,no
3,4,5,4,5,4,5,5,4,4,5,4,5,5,4,5,4,si
3,5,3,3,5,3,5,3,3,3,3,2,5,5,4,5,3,no
3,4,4,4,3,4,4,5,3,3,4,3,4,3,4,3,5,si
3,3,3,?,4,5,5,4,3,3,3,4,1,3,1,5,3,si
3,4,3,3,4,4,5,3,4,3,3,3,5,4,4,5,3,si
3,5,3,3,5,4,5,4,5,4,3,4,5,5,4,5,5,si

3,4,3,3,5,3,5,3,3,3,3,4,3,5,4,5,4,si
3,4,3,3,5,5,5,3,4,3,3,3,5,5,5,5,4,si
3,5,4,4,5,5,5,5,5,4,3,3,5,5,5,5,4,si
3,4,3,3,5,5,5,4,4,4,3,4,4,5,4,5,5,si
2,4,3,3,4,3,5,3,3,2,3,3,3,4,3,5,5,no
2,4,3,3,4,3,5,4,3,3,3,2,3,4,4,5,4,no
3,3,4,3,3,5,5,4,3,3,3,2,3,2,4,5,4,no
3,3,2,3,5,5,5,5,3,3,2,3,4,3,4,5,4,no
3,0,3,2,4,0,5,5,2,2,2,2,0,4,0,5,3,no
2,2,4,2,5,5,5,5,2,3,3,2,4,4,4,5,4,no
3,3,2,3,5,5,5,5,3,3,3,3,4,3,4,5,2,no
3,0,2,3,5,0,5,5,3,3,2,2,0,5,0,5,5,no
3,4,3,3,5,5,5,5,4,4,3,4,4,4,4,5,4,si
3,3,2,3,3,5,5,5,3,3,3,3,4,3,4,5,4,no
3,3,5,4,5,5,5,5,3,4,4,4,3,5,4,5,4,si
3,4,3,3,3,5,5,4,4,3,3,3,4,3,5,5,4,si
3,3,4,3,4,1,5,4,3,3,3,4,3,3,1,5,4,si
4,4,3,4,5,5,5,5,4,3,4,4,3,4,4,5,5,si
5,0,3,3,5,0,5,5,3,3,4,3,0,4,0,5,5,si
5,4,5,4,5,5,5,5,4,4,4,4,4,4,4,5,5,si
3,3,3,3,5,5,5,4,3,3,3,4,4,2,4,5,4,no
1,3,5,4,5,1,5,1,4,1,4,4,5,4,1,5,5,si
4,3,5,3,5,5,5,5,3,4,3,4,4,4,4,5,5,si
3,3,5,3,3,5,5,5,3,4,4,4,3,3,4,5,4,si
4,0,5,4,5,0,5,5,4,4,3,4,0,5,0,5,?,si
3,3,3,3,2,5,5,5,3,3,3,3,3,2,4,4,2,no
4,3,5,4,5,5,5,5,3,4,4,4,3,4,4,5,5,si
3,3,3,3,4,5,5,5,3,5,3,4,4,4,4,5,4,si
3,4,5,4,5,5,5,5,3,3,4,4,4,4,4,5,5,si
3,4,2,2,5,5,5,5,3,3,4,3,4,4,4,5,4,no
5,4,5,3,5,5,5,5,3,4,4,5,4,4,4,5,4,si
4,3,5,3,4,5,5,5,3,4,4,4,3,4,4,5,4,si
3,3,3,3,4,5,5,5,3,4,3,3,4,4,4,5,4,si
4,4,4,3,4,5,5,5,3,4,4,4,4,4,4,2,4,si
1,3,1,2,2,5,5,5,1,2,1,2,3,2,4,1,2,no
3,3,3,3,4,5,5,4,3,3,3,3,4,3,4,5,3,si
3,4,4,3,5,5,5,5,3,5,3,4,4,4,4,5,5,si

Anexo A1.3-Modelo combinado 1er y 2do año.

```
@relation 'Modelo 1ro y 2do_año'

@attribute 'Mat I' numeric
@attribute FyS numeric
@attribute FInf numeric
@attribute IP numeric
@attribute 'EF I ' numeric
@attribute 'Id0 I' numeric
@attribute 'Alg y GA' numeric
@attribute Hluba numeric
@attribute 'Id0 II' numeric
@attribute DPOO numeric
@attribute 'EF II' numeric
@attribute IGS numeric
@attribute MD numeric
@attribute 'Mat II' numeric
@attribute EPlap numeric
@attribute 'mat iii' numeric
@attribute eplap numeric
@attribute edl numeric
@attribute iia numeric
@attribute 'id0 3' numeric
@attribute segNa1 numeric
@attribute 'EF 3' numeric
@attribute 'ele1 1' numeric
@attribute 'mat iv' numeric
@attribute Alomp numeric
@attribute 'ED 2' numeric
@attribute BD numeric
@attribute TSP numeric
@attribute 'id0 4' numeric
@attribute Dna1 numeric
@attribute 'EF iv' numeric
@attribute Pprof numeric
@attribute aprueba {no,si}

@data
3,3,4,3,5,5,3,4,3,3,5,4,3,3,4,2,3,2,2,3,2,4,3,3,3,2,2,3,3,2,4,2,no
2,3,4,3,5,5,3,4,3,4,5,4,3,2,5,2,3,2,2,3,3,5,4,3,2,2,3,4,4,3,5,2,no
3,3,3,3,5,5,3,4,3,3,5,4,3,2,4,2,2,2,2,4,3,5,3,3,2,2,2,3,4,3,5,2,no
4,4,4,3,5,5,4,5,4,5,5,5,3,4,5,4,3,4,3,4,4,4,5,5,3,5,4,5,4,5,5,5,si
3,3,3,3,5,5,3,4,3,3,5,4,3,3,5,2,3,3,3,3,3,5,4,3,3,3,3,3,4,4,5,4,si
3,3,3,3,5,5,3,3,3,4,5,4,3,3,5,2,3,2,2,3,3,5,4,3,2,3,3,3,3,4,5,4,no
3,3,1,3,1,5,3,1,3,3,5,4,3,3,4,2,2,2,2,3,3,5,3,2,2,2,2,2,4,2,5,2,no
3,3,4,3,5,5,3,4,4,3,5,4,3,3,5,2,4,2,2,4,4,4,4,?,?,?,?,?,?,?,5,?,no
3,3,4,3,5,5,3,3,4,3,5,4,4,3,3,3,2,3,3,3,3,4,4,3,3,3,4,3,4,3,4,5,no
3,3,4,3,3,5,3,5,3,3,5,4,4,3,4,3,3,4,3,3,5,3,4,2,2,3,2,3,3,3,2,no
3,4,5,3,5,5,4,5,3,3,5,4,3,3,5,3,3,3,3,4,4,4,4,3,3,4,4,5,4,3,5,4,si
3,3,3,3,5,5,4,3,3,5,5,3,3,5,2,3,3,3,3,3,5,4,3,3,3,2,3,4,3,4,5,no
3,3,5,4,3,5,4,3,3,3,5,4,4,3,5,3,3,4,3,3,3,2,3,3,4,3,3,4,4,3,3,5,no
4,4,3,3,5,5,3,5,4,3,5,4,3,4,5,3,5,2,3,3,5,4,4,3,3,2,3,3,3,4,4,2,no
```

?,1,1,1,1,1,1,1,1,1,1,5,4,1,1,1,2,2,2,2,5,3,4,3,2,2,2,2,2,5,3,5,5,no
3,3,4,3,5,5,3,3,4,3,5,4,3,2,5,2,2,2,2,4,5,4,3,?,?,?,?,?,5,?,no
2,0,4,3,5,5,3,4,4,3,5,4,3,3,0,2,0,2,2,3,0,3,5,2,2,2,2,0,?,0,3,2,no
4,3,4,3,5,5,3,5,4,3,5,4,3,3,4,3,4,2,3,3,3,3,4,3,3,2,4,3,3,3,4,no
1,1,4,1,1,1,1,3,1,1,5,5,3,3,5,3,3,3,3,5,4,4,3,3,3,3,2,3,5,3,0,5,si
3,3,3,3,5,5,3,4,3,3,5,4,4,3,4,3,3,3,3,4,3,4,3,3,3,3,3,4,4,4,4,5,si
4,3,4,3,5,5,4,4,4,3,5,4,4,4,4,3,4,5,3,3,3,5,5,4,5,4,3,3,4,3,4,5,si
3,4,4,3,3,5,3,4,4,4,5,5,3,2,4,2,3,3,3,4,4,4,4,3,3,3,3,3,4,4,3,5,no
3,4,4,3,5,5,3,5,3,4,5,4,3,3,4,3,3,3,3,3,5,4,3,3,5,4,3,4,4,5,5,si
2,3,4,3,5,5,3,4,3,4,5,4,3,2,4,2,2,3,2,3,2,5,3,?,?,?,?,?,?,?,no
4,1,4,3,1,1,1,3,3,3,5,4,3,3,1,3,3,4,3,4,3,4,3,3,3,4,4,3,4,2,4,5,si
3,3,4,3,5,5,3,3,3,3,5,4,3,3,5,2,3,3,3,4,5,5,4,3,3,3,3,3,4,4,4,4,no
3,4,4,3,5,5,3,4,3,5,5,5,4,4,4,3,3,3,3,4,4,5,4,3,3,3,4,3,4,5,5,5,si
3,3,4,3,5,5,3,4,3,3,5,4,3,3,4,2,3,3,2,3,3,5,4,3,2,3,3,3,3,3,5,4,no
2,3,3,3,5,5,2,5,?,?,?,?,?,?,?,1,4,2,2,3,1,1,1,3,2,2,3,4,4,1,3,4,no
3,3,3,3,5,5,3,3,3,3,5,4,3,3,4,3,4,3,4,3,3,3,3,5,3,3,3,2,4,3,4,5,4,no
3,1,1,1,1,1,3,1,1,1,1,1,1,1,3,1,2,5,2,2,4,3,3,4,3,3,2,2,5,4,4,4,5,no
3,3,4,3,5,5,3,5,4,3,5,4,3,3,4,3,3,4,3,4,3,5,4,3,3,4,3,3,4,4,5,4,si
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,3,1,2,3,1,3,1,4,3,3,2,2,1,1,4,1,5,no
2,1,3,1,1,1,1,5,1,3,5,4,1,2,4,?,4,2,?,4,?,5,3,?,?,?,?,?,?,?,no
3,4,3,3,5,5,3,3,3,3,5,4,3,3,4,3,3,2,3,3,0,4,5,3,2,2,3,3,?,0,4,3,no
1,1,1,1,5,1,1,1,1,1,1,1,1,1,1,1,1,3,3,1,1,1,1,1,1,3,1,5,1,1,1,1,si
3,4,4,3,5,5,3,5,3,4,5,4,4,3,4,3,3,3,3,3,4,5,4,3,3,3,4,3,4,4,4,4,si
4,4,5,3,5,5,4,4,4,3,5,4,4,4,4,3,3,3,3,4,3,5,4,3,3,3,4,3,4,5,5,5,si
1,1,1,1,1,1,1,1,1,1,?,?,?,?,?,1,1,2,3,1,1,1,1,1,3,3,4,3,1,1,1,1,no
2,4,4,3,4,4,3,4,4,3,5,5,3,3,4,3,5,3,3,5,4,5,4,4,4,3,4,4,4,5,5,3,si
3,5,4,3,5,4,3,5,4,4,5,5,3,3,5,3,4,3,4,4,4,5,5,4,3,3,5,4,5,5,5,5,si
2,3,4,3,5,3,3,4,3,3,5,4,2,2,5,?,?,?,?,?,?,?,?,?,?,?,?,?,no
3,3,4,3,4,4,3,4,4,3,5,4,3,3,3,2,5,3,3,4,4,5,4,3,3,3,2,5,4,4,5,3,no
5,5,5,5,5,5,5,5,5,5,5,5,5,4,4,5,5,5,3,4,4,5,5,5,5,4,4,5,5,5,5,si
3,0,3,3,5,3,3,3,5,3,5,5,3,3,0,3,0,3,3,4,0,5,4,4,4,4,4,0,4,0,5,5,si
3,0,3,3,4,3,3,3,4,3,5,3,3,3,0,3,0,3,3,0,5,4,4,3,3,4,0,3,0,5,4,si
5,si
1,3,5,1,1,4,1,5,4,1,1,5,1,3,5,?,?,?,?,?,?,?,?,?,?,?,?,no
3,3,5,4,5,4,3,5,4,5,5,5,3,3,5,3,4,3,3,4,5,5,5,4,3,3,4,5,5,5,5,si
2,4,5,3,5,3,3,4,3,3,5,4,2,2,4,?,?,?,?,?,?,?,?,?,?,?,?,no
2,3,4,3,5,3,3,4,3,3,5,4,2,2,3,?,?,?,?,?,?,?,?,?,?,?,no
4,0,3,3,5,4,4,4,5,4,5,4,3,4,0,5,0,3,3,4,0,5,4,4,3,4,4,0,4,0,5,4,si
2,1,1,1,1,1,1,1,1,1,1,1,1,3,3,1,?,?,?,?,?,?,?,?,?,?,?,no
2,2,2,2,?,?,2,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,no
1,3,5,1,1,5,1,5,5,3,1,4,1,3,4,4,4,3,3,5,5,5,5,5,4,3,5,5,5,5,5,si
3,4,4,3,5,5,4,4,5,4,5,5,4,3,4,4,4,3,4,5,4,5,4,4,3,3,5,4,5,5,5,5,si
2,1,1,1,1,1,1,1,1,1,1,1,1,?,1,1,?,?,?,?,?,?,?,?,?,?,no
3,4,5,3,5,4,3,5,5,5,5,5,3,3,5,?,?,?,?,?,?,?,?,?,?,?,no
3,3,4,4,5,4,3,4,4,4,5,4,2,3,4,?,?,?,?,?,?,?,?,?,?,?,no
3,3,4,3,5,3,3,3,3,3,5,4,3,3,4,?,?,?,?,?,?,?,?,?,?,?,no
2,2,5,3,5,3,3,3,2,4,5,4,2,2,5,?,?,?,?,?,?,?,?,?,?,no
2,4,5,3,5,?,3,5,4,4,5,4,3,3,4,3,4,3,3,5,3,5,4,4,3,3,4,4,4,5,5,5,si
4,4,3,5,5,4,4,4,5,5,5,5,4,5,4,5,5,5,4,5,3,5,4,5,5,5,5,5,5,5,si
2,2,3,3,5,4,3,4,3,3,5,4,2,2,4,?,?,?,?,?,?,?,?,?,?,no
2,3,4,3,3,4,3,4,4,4,5,4,3,3,3,?,?,?,?,?,?,?,?,?,?,no
1,3,5,1,1,4,1,4,4,?,1,3,1,3,5,?,?,?,?,?,?,?,?,?,?,no
4,5,5,5,5,4,3,5,5,5,5,5,3,3,5,3,5,5,3,5,5,5,5,4,3,5,3,5,5,5,4,si
2,3,4,3,5,3,3,4,3,3,5,4,2,2,3,?,?,?,?,?,?,?,?,?,?,no
2,2,3,3,5,3,3,4,3,3,5,4,2,2,3,?,?,?,?,?,?,?,?,?,?,no
2,2,4,3,4,4,3,4,4,3,5,4,3,2,3,?,?,?,?,?,?,?,?,?,?,no
4,5,5,5,5,3,5,5,5,5,3,3,3,4,4,5,5,4,5,5,5,4,4,3,4,4,4,5,5,5,4,si
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,4,4,4,4,4,3,4,?,3,?,5,3,4,4,3,5,3,si
2,4,5,3,5,4,3,4,4,3,5,4,3,2,4,?,?,?,?,?,?,?,?,?,?,no
4,3,5,4,4,4,3,4,4,5,5,3,3,4,5,4,3,3,3,5,4,5,4,3,3,3,3,4,5,4,5,5,si
2,4,5,4,5,4,3,4,3,4,5,3,3,3,5,2,4,3,3,5,4,5,4,3,3,3,2,4,4,4,5,3,no

4,5,5,5,5,4,4,5,4,5,5,5,4,5,5,3,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,si
3,4,4,3,5,4,3,5,4,4,5,3,3,3,5,3,3,3,5,3,?,3,3,3,3,2,3,5,4,5,3,no
5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,4,5,5,5,5,5,5,5,5,5,4,5,5,5,5,5,si
3,?,4,3,5,4,3,4,4,4,5,4,3,3,4,3,4,3,3,5,4,5,3,3,3,3,3,5,5,4,5,4,si
3,4,5,5,5,3,3,5,4,5,5,5,3,3,5,2,4,5,3,4,3,?,3,3,4,4,2,3,4,4,5,5,no
2,4,4,3,5,3,3,4,3,3,5,4,3,3,5,3,5,3,3,4,5,5,3,5,4,3,3,5,5,5,5,4,si
3,3,4,3,5,4,3,4,3,3,5,3,3,3,3,3,4,3,3,4,3,?,3,3,3,3,3,3,4,4,5,5,si
5,5,5,4,5,4,4,4,4,5,5,4,4,5,5,3,5,5,5,5,5,5,5,4,5,5,5,5,4,5,5,5,4,si
3,3,5,5,5,4,3,4,4,4,5,5,3,3,4,3,4,3,3,5,4,5,3,4,4,4,4,5,4,4,5,4,si
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,3,5,3,3,4,5,5,5,5,3,4,4,5,5,5,5,5,si
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,3,4,3,3,5,3,5,5,4,3,3,4,1,4,5,5,?,si
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,3,4,3,3,4,4,4,4,4,4,3,3,2,4,4,5,5,no
4,5,5,5,5,4,3,4,4,5,5,5,4,4,3,3,4,5,4,5,4,5,5,4,4,5,4,5,5,4,5,4,si
3,3,3,3,5,3,3,4,3,4,5,4,3,3,4,3,5,3,3,5,3,5,3,3,3,3,2,5,5,4,5,3,si
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,3,4,4,4,4,3,4,4,5,3,3,4,3,4,3,5,si
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,3,3,4,3,?,4,5,5,4,3,3,3,4,1,3,1,5,3,si
3,3,4,3,5,3,3,4,3,4,5,5,3,3,4,3,4,3,3,4,4,5,3,4,3,3,3,5,4,4,5,3,si
3,4,5,5,5,5,3,5,4,4,5,4,3,3,5,3,5,3,3,5,4,5,4,5,4,3,4,5,5,4,5,5,si
2,3,5,4,5,5,3,4,4,3,5,3,3,2,5,3,4,3,3,5,3,5,3,3,3,3,4,3,5,4,5,4,si
3,4,4,3,5,3,3,5,3,4,5,5,4,3,5,3,4,3,3,5,5,5,3,4,3,3,3,5,5,5,5,4,si
3,5,5,4,5,5,3,5,5,4,5,5,4,3,5,3,5,4,4,5,5,5,5,5,4,3,3,5,5,5,5,4,si
2,4,5,4,5,3,3,5,3,5,5,5,3,3,3,3,4,3,3,5,5,5,4,4,4,3,4,4,5,4,5,5,si
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,2,4,3,3,4,3,5,4,3,3,3,2,3,4,4,5,4,no
3,4,5,4,5,3,3,3,3,3,5,4,3,3,3,3,3,4,3,3,5,5,4,3,3,3,2,3,2,4,5,4,no
3,0,4,3,3,4,3,5,4,4,5,4,3,2,0,?,?,?,?,?,?,?,?,?,?,?,?,?,?,no
3,4,4,3,5,4,3,4,4,3,5,5,3,3,4,3,3,2,3,5,5,5,5,3,3,2,3,4,3,4,5,4,no
3,0,4,3,5,4,3,5,4,4,5,4,3,2,0,3,0,3,2,4,0,5,5,2,2,2,2,0,4,0,5,3,no
3,4,5,5,3,5,3,4,4,4,5,3,3,4,3,2,2,4,2,5,5,5,5,2,3,3,2,4,4,4,5,4,no
3,4,5,3,5,3,3,5,3,3,5,5,3,3,4,3,3,2,3,5,5,5,5,3,3,3,3,4,3,4,5,2,no
3,0,4,3,5,5,3,5,5,4,5,5,3,3,0,3,0,2,3,5,0,5,5,3,3,2,2,0,5,0,5,5,no
3,5,5,3,5,4,3,5,4,4,5,5,3,3,4,3,4,3,3,5,5,5,5,4,4,3,4,4,4,4,5,4,si
3,4,4,3,5,3,3,3,4,3,5,3,3,3,4,3,2,3,3,5,5,5,3,3,3,3,4,3,4,5,4,no
3,5,5,5,5,5,3,4,5,5,5,5,3,3,3,3,2,3,4,5,5,5,5,3,4,4,4,3,5,4,5,4,si
4,0,5,3,5,0,4,5,0,4,5,5,5,3,0,5,0,3,3,5,0,5,5,3,3,4,3,0,4,0,5,5,si
3,4,4,3,5,3,3,4,3,3,5,4,3,3,3,3,3,4,3,4,1,5,4,3,3,3,4,3,3,1,5,4,si
3,5,5,3,5,3,4,5,4,5,5,3,4,4,5,4,4,3,4,5,5,5,5,4,3,4,4,3,4,4,5,5,si
3,4,5,3,5,3,4,3,3,3,5,3,3,3,3,?,?,?,?,?,?,?,?,?,?,?,?,?,?,no
3,5,5,4,5,4,5,5,4,5,5,5,5,4,5,5,4,5,4,5,5,5,5,4,4,4,4,4,4,4,5,5,si
3,4,4,3,4,3,3,3,3,4,5,3,3,3,4,3,3,3,3,5,5,5,4,3,3,3,4,4,2,4,5,4,no
3,5,5,4,5,4,3,4,4,3,5,4,3,3,4,1,3,5,4,5,1,5,1,4,1,4,4,5,4,1,5,5,si
3,4,5,4,5,4,4,4,4,4,5,4,3,4,4,4,3,5,3,5,5,5,5,3,4,3,4,4,4,4,5,5,si
3,4,4,3,?,4,3,4,4,4,5,3,4,3,3,3,3,5,3,3,5,5,5,3,4,4,4,3,3,4,5,4,si
4,0,5,5,5,5,4,5,5,5,5,5,5,3,0,4,0,5,4,5,0,5,5,4,4,3,4,0,5,0,5,?,si
2,4,5,4,5,3,3,5,3,3,5,4,2,3,3,?,?,?,?,?,?,?,?,?,?,?,?,?,?,no
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,3,3,3,3,2,5,5,5,3,3,3,3,3,2,4,4,2,no
3,4,5,4,5,4,4,4,4,4,5,5,5,4,3,3,4,3,5,4,5,5,5,5,3,4,4,4,3,4,4,5,5,si
3,4,5,4,4,3,3,4,3,4,5,4,4,3,4,3,3,3,3,4,5,5,5,3,5,3,4,4,4,4,5,4,si
3,5,5,5,4,3,3,4,3,5,5,4,4,3,4,3,4,5,4,5,5,5,5,3,3,4,4,4,4,4,5,5,si
3,5,5,3,5,4,3,4,4,3,5,4,3,3,4,3,4,2,2,5,5,5,5,3,3,4,3,4,4,4,5,4,no
3,5,5,4,5,4,4,5,4,5,5,4,3,4,5,4,5,3,5,5,5,5,3,4,4,5,4,4,5,4,si
4,5,5,4,5,4,5,5,4,5,5,4,5,3,4,4,3,5,3,4,5,5,5,3,4,4,4,3,4,4,5,4,si
3,4,5,3,5,4,3,5,4,3,5,3,3,3,4,?,?,?,?,?,?,?,?,?,?,?,?,?,?,no
3,4,5,4,5,4,3,4,4,3,5,3,3,3,3,3,3,3,3,4,5,5,5,3,4,3,3,4,4,4,5,4,si
2,4,5,4,5,4,3,5,4,3,5,4,2,3,4,?,?,?,?,?,?,?,?,?,?,?,?,?,?,no
3,4,5,3,5,4,4,4,3,3,5,3,4,3,4,4,4,4,3,4,5,5,5,3,4,4,4,4,4,2,4,si
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,3,1,2,2,5,5,5,1,2,1,2,3,2,4,1,2,no
3,4,4,3,5,4,3,4,4,4,5,4,3,3,4,3,3,3,3,4,5,5,4,3,3,3,3,4,3,4,5,3,si
2,4,5,3,3,3,3,4,3,3,5,3,2,3,3,?,?,?,?,?,?,?,?,?,?,?,?,?,?,no
3,4,5,3,4,4,3,4,4,4,5,5,3,3,4,3,4,4,3,5,5,5,5,3,5,3,4,4,4,4,5,5,si

Anexo A2.1- Modelo 1er año con selección de rasgos.

```
@relation 'Modelo_1er año selrasg'
```

```
@attribute 'Mat I' numeric
@attribute FyS numeric
@attribute FInf numeric
@attribute IP numeric
@attribute 'Alg y GA' numeric
@attribute 'Id0 II' numeric
@attribute DPOO numeric
@attribute IGS numeric
@attribute MD numeric
@attribute 'Mat II' numeric
@attribute aprueba {si,no}
```

```
@data
5,5,4,5,4,4,5,4,4,4,si
3,4,4,4,3,4,5,4,3,3,si
4,3,4,2,3,3,3,3,3,3,no
3,4,4,3,4,3,5,3,3,3,si
3,4,4,2,3,3,3,3,3,3,no
3,4,4,3,3,3,3,3,3,3,no
3,5,5,5,4,3,4,4,3,4,si
3,4,4,2,3,3,2,3,3,2,no
2,3,4,2,3,3,2,3,3,2,no
3,5,5,3,3,3,4,4,3,5,si
3,5,5,3,4,4,3,3,3,4,si
3,4,4,3,3,3,3,4,3,4,no
4,4,5,5,3,5,5,4,4,3,si
4,4,4,3,3,5,3,3,3,5,si
4,4,4,4,3,3,4,3,3,3,si
3,4,4,3,3,4,3,3,3,4,si
3,4,4,3,3,3,3,3,3,4,si
3,4,4,4,4,4,3,3,3,3,si
3,5,4,2,3,4,3,3,3,4,no
2,3,4,2,2,3,2,4,2,2,no
3,3,5,3,3,3,3,3,3,3,si
5,5,4,4,4,3,5,4,4,5,si
3,3,4,3,3,3,3,4,3,3,si
3,4,5,3,4,4,5,3,3,4,si
3,5,4,2,4,3,3,3,3,3,no
3,4,4,3,3,3,3,4,2,3,si
3,3,4,2,2,3,3,3,3,3,no
3,4,5,4,3,3,3,4,3,3,si
3,0,4,3,3,4,4,4,3,2,no
3,4,4,3,3,4,3,5,3,3,si
3,0,4,3,3,4,4,4,3,2,no
3,4,5,5,3,4,4,3,3,4,si
```

3,4,5,3,3,3,3,5,3,3,si
3,0,4,3,3,5,4,5,3,3,si
3,5,5,3,3,4,4,5,3,3,si
3,4,4,3,3,4,3,3,3,3,si
3,5,5,5,3,5,5,5,3,3,si
4,0,5,3,4,0,4,5,5,3,si
3,4,4,3,3,3,3,4,3,3,si
3,5,5,3,4,4,5,3,4,4,si
3,4,5,3,4,3,3,3,3,3,si
3,5,5,4,5,4,5,5,5,4,si
3,4,4,3,3,3,4,3,3,3,si
3,5,5,4,3,4,3,4,3,3,si
3,4,5,4,4,4,4,4,3,4,si
4,0,5,5,4,5,5,5,5,3,si
3,4,4,3,3,4,4,3,4,3,si
2,4,5,4,3,3,3,4,2,3,si
3,4,5,4,4,4,5,5,4,3,si
3,4,5,4,3,3,4,4,4,3,si
3,5,5,5,3,3,5,4,4,3,si
3,5,5,3,3,4,3,4,3,3,si
3,5,5,4,4,4,5,5,4,3,si
4,5,5,4,5,4,5,4,5,3,si
3,4,5,3,3,4,3,3,3,3,si
3,4,5,4,3,4,3,3,3,3,si
2,4,5,4,3,4,3,4,2,3,no
3,4,5,3,4,3,3,3,4,3,si
3,4,4,3,3,4,4,4,3,3,si
2,4,5,3,3,3,3,3,2,3,no
3,4,5,3,3,4,4,5,3,3,si
4,3,5,4,3,4,5,3,3,4,si
2,4,5,4,3,3,4,3,3,3,no
4,5,5,5,4,4,5,5,4,5,si
3,4,4,3,3,4,4,3,3,3,si
3,3,4,3,3,4,3,4,3,3,si
5,5,5,5,5,5,5,5,5,5,si
3,?,4,3,3,4,4,4,3,3,si
3,4,5,5,3,4,5,5,3,3,si
2,4,4,3,3,3,3,4,3,3,no
3,3,4,3,3,3,3,3,3,3,si
5,5,5,4,4,4,5,4,4,5,si
3,3,5,5,3,4,4,5,3,3,si
4,5,5,5,3,4,5,5,4,4,si
3,3,3,3,3,3,4,4,3,3,si
3,3,4,3,3,3,4,5,3,3,si
3,4,5,5,3,4,4,4,3,3,si
2,3,5,4,3,4,3,3,3,2,no
3,4,4,3,3,3,4,5,4,3,si
3,5,5,4,3,5,4,5,4,3,si
2,4,5,4,3,3,5,5,3,3,no
2,4,4,3,3,4,3,5,3,3,no
3,5,4,3,3,4,4,5,3,3,si
2,3,4,3,3,3,3,4,2,2,no
2,3,4,3,3,4,3,4,2,2,no
5,5,5,5,5,5,5,5,4,4,si
3,0,3,3,3,5,3,5,3,3,si

3,0,3,3,3,4,3,3,3,3,si
5,5,5,5,5,5,5,5,5,5,si
1,3,5,1,1,4,1,5,1,3,si
3,3,5,4,3,4,5,5,3,3,si
2,4,5,3,3,3,3,4,2,2,no
2,3,4,3,3,3,3,4,2,2,no
4,0,3,3,4,5,4,4,3,4,si
2,1,1,1,1,1,1,1,3,3,no
2,2,2,2,2,?,?,?,?,no
1,3,5,1,1,5,3,4,1,3,si
3,4,4,3,4,5,4,5,4,3,si
2,1,1,1,1,1,1,1,?,1,no
3,4,5,3,3,5,5,5,3,3,si
3,3,4,4,3,4,4,4,2,3,no
3,3,4,3,3,3,3,4,3,3,si
2,2,5,3,3,2,4,4,2,2,no
2,4,5,3,3,4,4,4,3,3,no
4,4,3,5,4,5,5,5,4,5,si
2,2,3,3,3,3,3,4,2,2,no
2,3,4,3,3,4,4,4,3,3,no
1,3,5,1,1,4,?,3,1,3,si
4,5,5,5,3,5,5,5,3,3,si
2,3,4,3,3,3,3,4,2,2,no
2,2,3,3,3,3,3,4,2,2,no
2,2,4,3,3,4,3,4,3,2,no
4,5,5,5,3,5,5,3,3,3,si
2,4,5,3,3,4,3,4,3,2,no

Anexo A2.2.-Modelo 2do año con selección de rasgos.

```
@relation 'Modelo 2do año selrasg'
```

```
@attribute 'mat iii' numeric
@attribute ed1 numeric
@attribute iia numeric
@attribute 'mat iv' numeric
@attribute Alomp numeric
@attribute 'ED 2' numeric
@attribute BD numeric
@attribute Pprof numeric
@attribute aprueba {no,si}
```

```
@data
2,2,2,3,3,2,2,2,no
2,2,2,3,2,2,3,2,no
2,2,2,3,2,2,2,2,no
4,4,3,5,3,5,4,5,si
2,3,3,3,3,3,3,4,si
2,2,2,3,2,3,3,4,no
2,2,2,2,2,2,2,2,no
2,2,2,?,?,?,?,no
3,3,3,3,3,3,4,5,no
3,4,3,2,2,3,2,2,no
3,3,3,3,3,4,4,4,si
2,3,3,3,3,3,2,5,no
3,4,3,3,4,3,3,5,no
3,2,3,3,3,2,3,2,no
2,2,2,2,2,2,2,5,no
2,2,2,?,?,?,?,no
2,2,2,2,2,2,2,2,no
3,2,3,3,3,2,4,4,no
3,3,3,3,3,3,2,5,si
3,3,3,3,3,3,3,5,si
3,5,3,4,5,4,3,5,si
2,3,3,3,3,3,3,5,no
3,3,3,3,3,5,4,5,si
2,3,2,?,?,?,?,no
3,4,3,3,3,4,4,5,no
2,3,3,3,3,3,3,4,no
3,3,3,3,3,3,4,5,si
2,3,2,3,2,3,3,4,no
1,2,2,3,2,2,3,4,no
3,3,3,3,3,2,4,4,no
2,2,2,3,3,2,2,5,no
3,4,3,3,3,4,3,4,si
3,2,3,3,3,2,2,5,no
?,2,2,?,?,?,?,no
```

3,2,3,3,2,2,3,3,no
1,3,3,1,1,3,1,1,si
3,3,3,3,3,3,4,4,si
3,3,3,3,3,3,4,5,si
1,2,3,1,3,3,4,1,no
3,3,4,3,3,3,4,3,no
2,3,3,3,4,3,3,3,no
3,3,3,4,4,3,4,3,si
3,3,4,4,3,3,5,5,si
5,3,4,5,4,4,5,5,si
3,3,3,4,4,4,4,5,si
3,3,3,4,3,3,4,4,si
5,5,5,5,5,5,5,5,si
3,3,3,4,3,3,4,5,si
3,3,2,3,3,2,4,4,no
5,3,3,4,3,4,4,4,si
4,3,3,5,4,3,5,5,si
4,3,4,4,3,3,5,5,si
3,3,3,4,3,3,4,5,si
3,3,3,3,3,3,3,5,si
5,5,4,5,5,5,5,5,si
3,3,3,3,4,3,3,4,si
3,5,3,4,3,5,3,4,si
3,3,3,3,3,3,5,5,si
4,5,4,4,3,4,4,4,si
3,3,3,3,4,3,3,3,si
4,4,4,3,?,5,3,3,si
4,3,3,3,3,3,3,5,si
2,3,3,3,3,3,2,3,no
3,5,5,5,5,5,5,5,si
3,3,3,3,3,3,2,3,no
2,3,3,3,3,3,2,3,no
4,5,5,5,5,5,4,5,si
4,4,3,4,3,?,5,4,si
3,3,3,3,3,3,3,4,si
2,5,3,3,4,4,2,5,no
3,3,3,5,4,3,3,4,si
3,3,3,3,3,3,3,5,si
3,5,5,5,5,5,5,4,si
3,3,3,4,4,4,4,4,si
3,3,3,5,3,4,4,5,si
3,3,3,4,3,3,4,?,si
3,3,3,4,4,3,3,5,no
3,5,4,4,4,5,4,4,si
3,3,3,3,3,3,2,3,no
3,4,4,3,3,4,3,5,si
3,3,?,3,3,3,4,3,si
3,3,3,4,3,3,3,3,si
3,3,3,5,4,3,4,5,si
3,3,3,3,3,3,4,4,si
3,3,3,4,3,3,3,4,si
3,4,4,5,4,3,3,4,si
3,3,3,4,4,3,4,5,si
2,3,3,3,2,3,3,5,no
2,3,3,3,3,3,2,4,no

3,4,3,3,3,3,2,4,no
3,2,3,3,3,2,3,4,no
3,3,2,2,2,2,2,3,no
2,4,2,2,3,3,2,4,no
3,2,3,3,3,3,3,2,no
3,2,3,3,3,2,2,5,no
3,3,3,4,4,3,4,4,si
3,2,3,3,3,3,3,4,no
3,5,4,3,4,4,4,4,si
3,3,3,4,3,3,3,4,si
3,4,3,3,3,3,4,4,si
4,3,4,4,3,4,4,5,si
5,3,3,3,3,4,3,5,si
5,5,4,4,4,4,4,5,si
3,3,3,3,3,3,4,4,no
1,5,4,4,1,4,4,5,si
4,5,3,3,4,3,4,5,si
3,5,3,3,4,4,4,4,si
4,5,4,4,4,3,4,?,si
3,3,3,3,3,3,3,2,no
4,5,4,3,4,4,4,5,si
3,3,3,3,5,3,4,4,si
3,5,4,3,3,4,4,5,si
3,2,2,3,3,4,3,4,no
5,5,3,3,4,4,5,4,si
4,5,3,3,4,4,4,4,si
3,3,3,3,4,3,3,4,si
4,4,3,3,4,4,4,4,si
1,1,2,1,2,1,2,2,no
3,3,3,3,3,3,3,3,si
3,4,3,3,5,3,4,5,si

Anexo A23.-Modelo combinado 1er y 2do año, con selección de rasgos.

```
@relation 'Modelo 1ro y 2do año sel_rasg'
```

```
@attribute 'Mat I' numeric
@attribute FyS numeric
@attribute IP numeric
@attribute DPOO numeric
@attribute MD numeric
@attribute 'Mat II' numeric
@attribute 'mat iii' numeric
@attribute edl numeric
@attribute iia numeric
@attribute 'mat iv' numeric
@attribute Alomp numeric
@attribute 'ED 2' numeric
@attribute BD numeric
@attribute aprueba {no,si}
```

```
@data
```

```
3,3,3,3,3,3,2,2,2,3,3,2,2,no
2,3,3,4,3,2,2,2,2,3,2,2,3,no
3,3,3,3,3,2,2,2,2,3,2,2,2,no
4,4,3,5,3,4,4,4,3,5,3,5,4,si
3,3,3,3,3,3,2,3,3,3,3,3,3,si
3,3,3,4,3,3,2,2,2,3,2,3,3,no
3,3,3,3,3,3,2,2,2,2,2,2,2,no
3,3,3,3,3,3,2,2,2,?,?,?,no
3,3,3,3,4,3,3,3,3,3,3,3,4,no
3,3,3,3,4,3,3,4,3,2,2,3,2,no
3,4,3,3,3,3,3,3,3,3,3,4,4,si
3,3,3,3,3,3,2,3,3,3,3,3,2,no
3,3,4,3,4,3,3,4,3,3,4,3,3,no
4,4,3,3,3,4,3,2,3,3,3,2,3,no
?,1,1,1,1,1,2,2,2,2,2,2,2,no
3,3,3,3,3,2,2,2,2,?,?,?,no
2,0,3,3,3,3,2,2,2,2,2,2,2,no
4,3,3,3,3,3,3,2,3,3,3,2,4,no
1,1,1,1,3,3,3,3,3,3,3,3,2,si
3,3,3,3,4,3,3,3,3,3,3,3,3,si
4,3,3,3,4,4,3,5,3,4,5,4,3,si
3,4,3,4,3,2,2,3,3,3,3,3,3,no
3,4,3,4,3,3,3,3,3,3,3,5,4,si
2,3,3,4,3,2,2,3,2,?,?,?,no
4,1,3,3,3,3,3,4,3,3,3,4,4,si
3,3,3,3,3,3,2,3,3,3,3,3,3,no
3,4,3,5,4,4,3,3,3,3,3,3,4,si
3,3,3,3,3,3,2,3,2,3,2,3,3,no
```

2,3,3,?,?,?,1,2,2,3,2,2,3,no
3,3,3,3,3,3,3,3,3,3,2,4,no
3,1,1,1,1,3,2,2,2,3,3,2,2,no
3,3,3,3,3,3,3,4,3,3,3,4,3,si
1,1,1,1,1,1,3,2,3,3,3,2,2,no
2,1,1,3,1,2,?,2,?,?,?,?,no
3,4,3,3,3,3,3,2,3,3,2,2,3,no
1,1,1,1,1,1,1,3,3,1,1,3,1,si
3,4,3,4,4,3,3,3,3,3,3,3,4,si
4,4,3,3,4,4,3,3,3,3,3,3,4,si
1,1,1,1,?,?,1,2,3,1,3,3,4,no
2,4,3,3,3,3,3,3,3,4,4,3,4,si
3,5,3,4,3,3,3,3,4,4,3,3,5,si
2,3,3,3,2,2,?,?,?,?,?,no
3,3,3,3,3,3,2,3,3,3,3,3,2,no
5,5,5,5,4,4,5,3,4,5,4,4,5,si
3,0,3,3,3,3,3,3,3,4,4,4,4,si
3,0,3,3,3,3,3,3,3,4,3,3,4,si
5,5,5,5,5,5,5,5,5,5,5,5,5,si
1,3,1,1,1,3,?,?,?,?,?,no
3,3,4,5,3,3,3,3,3,4,3,3,4,si
2,4,3,3,2,2,?,?,?,?,?,no
2,3,3,3,2,2,?,?,?,?,?,no
4,0,3,4,3,4,5,3,3,4,3,4,4,si
2,1,1,1,3,3,?,?,?,?,?,no
2,2,2,?,?,?,?,?,?,?,?,no
1,3,1,3,1,3,4,3,3,5,4,3,5,si
3,4,3,4,4,3,4,3,4,4,3,3,5,si
2,1,1,1,?,1,?,?,?,?,?,no
3,4,3,5,3,3,?,?,?,?,?,no
3,3,4,4,2,3,?,?,?,?,?,no
3,3,3,3,3,3,?,?,?,?,?,no
2,2,3,4,2,2,?,?,?,?,?,no
2,4,3,4,3,3,3,3,3,4,3,3,4,si
4,4,5,5,4,5,5,5,4,5,5,5,5,si
2,2,3,3,2,2,?,?,?,?,?,no
2,3,3,4,3,3,?,?,?,?,?,no
1,3,1,?,1,3,?,?,?,?,?,no
4,5,5,5,3,3,3,5,3,4,3,5,3,si
2,3,3,3,2,2,?,?,?,?,?,no
2,2,3,3,2,2,?,?,?,?,?,no
2,2,3,3,3,2,?,?,?,?,?,no
4,5,5,5,3,3,4,5,4,4,3,4,4,si
1,1,1,1,1,1,4,4,4,3,?,5,3,si
2,4,3,3,3,2,?,?,?,?,?,no
4,3,4,5,3,4,4,3,3,3,3,3,3,si
2,4,4,4,3,3,2,3,3,3,3,3,2,no
4,5,5,5,4,5,3,5,5,5,5,5,5,si
3,4,3,4,3,3,3,3,3,3,3,3,2,no
5,5,5,5,5,5,4,5,5,5,5,5,4,si
3,?,3,4,3,3,3,3,3,3,3,3,3,si
3,4,5,5,3,3,2,5,3,3,4,4,2,no
2,4,3,3,3,3,3,3,3,5,4,3,3,si
3,3,3,3,3,3,3,3,3,3,3,3,3,si
5,5,4,5,4,5,3,5,5,5,5,5,5,si

3,3,5,4,3,3,3,3,3,4,4,4,4,si
1,1,1,1,1,1,3,3,3,5,3,4,4,si
1,1,1,1,1,1,3,3,3,4,3,3,4,si
1,1,1,1,1,1,3,3,3,4,4,3,3,no
4,5,5,5,4,4,3,5,4,4,4,5,4,si
3,3,3,4,3,3,3,3,3,3,3,2,si
1,1,1,1,1,1,3,4,4,3,3,4,3,si
1,1,1,1,1,1,3,3,?,3,3,3,4,si
3,3,3,4,3,3,3,3,3,4,3,3,3,si
3,4,5,4,3,3,3,3,3,5,4,3,4,si
2,3,4,3,3,2,3,3,3,3,3,3,4,si
3,4,3,4,4,3,3,3,3,4,3,3,3,si
3,5,4,4,4,3,3,4,4,5,4,3,3,si
2,4,4,5,3,3,3,3,3,4,4,3,4,si
1,1,1,1,1,1,2,3,3,3,3,3,2,no
3,4,4,3,3,3,3,4,3,3,3,3,2,no
3,0,3,4,3,2,?,?,?,?,no
3,4,3,3,3,3,3,2,3,3,3,2,3,no
3,0,3,4,3,2,3,3,2,2,2,2,2,no
3,4,5,4,3,4,2,4,2,2,3,3,2,no
3,4,3,3,3,3,3,2,3,3,3,3,3,no
3,0,3,4,3,3,3,2,3,3,3,2,2,no
3,5,3,4,3,3,3,3,3,4,4,3,4,si
3,4,3,3,3,3,3,2,3,3,3,3,3,no
3,5,5,5,3,3,3,5,4,3,4,4,4,si
4,0,3,4,5,3,5,3,3,3,3,4,3,si
3,4,3,3,3,3,3,4,3,3,3,3,4,si
3,5,3,5,4,4,4,3,4,4,3,4,4,si
3,4,3,3,3,3,?,?,?,?,no
3,5,4,5,5,4,5,5,4,4,4,4,4,si
3,4,3,4,3,3,3,3,3,3,3,3,4,no
3,5,4,3,3,3,1,5,4,4,1,4,4,si
3,4,4,4,3,4,4,5,3,3,4,3,4,si
3,4,3,4,4,3,3,5,3,3,4,4,4,si
4,0,5,5,5,3,4,5,4,4,4,3,4,si
2,4,4,3,2,3,?,?,?,?,no
1,1,1,1,1,1,3,3,3,3,3,3,3,no
3,4,4,5,4,3,4,5,4,3,4,4,4,si
3,4,4,4,4,3,3,3,3,3,5,3,4,si
3,5,5,5,4,3,3,5,4,3,3,4,4,si
3,5,3,3,3,3,3,2,2,3,3,4,3,no
3,5,4,5,4,3,5,5,3,3,4,4,5,si
4,5,4,5,5,3,4,5,3,3,4,4,4,si
3,4,3,3,3,3,?,?,?,?,no
3,4,4,3,3,3,3,3,3,3,4,3,3,si
2,4,4,3,2,3,?,?,?,?,no
3,4,3,3,4,3,4,4,3,3,4,4,4,si
1,1,1,1,1,1,1,1,2,1,2,1,2,no
3,4,3,4,3,3,3,3,3,3,3,3,3,si
2,4,3,3,2,3,?,?,?,?,no
3,4,3,4,3,3,3,4,3,3,5,3,4,si

Anexo B.

Resultados de los algoritmos aplicados a la Selección de Rasgos.

Anexo B1.1

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.BestFirst -D 1 -N 5
Relation: Modelo_1ro año-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R16
Instances: 114
Attributes: 16
Mat I
FyS
FInf
IP
EF I
IdEx I
Alg y GA
HCuba
IdEx II
DPOO
EF II
IGS
MD
Mat II
EPCap
aprueba
Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

number of folds (%) attribute
10(100 %) 1 Mat I
6(60 %) 2 FyS
10(100 %) 3 FInf
2(20 %) 4 IP
0(0 %) 5 EF I
0(0 %) 6 IdEx I
0(0 %) 7 Alg y GA
1(10 %) 8 HCuba
0(0 %) 9 IdEx II
3(30 %) 10 DPOO
0(0 %) 11 EF II
0(0 %) 12 IGS
10(100 %) 13 MD
10(100 %) 14 Mat II
0(0 %) 15 EPCap

Anexo B1.2

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.BestFirst -D 1 -N 5
Relation: Modelo_1ro año-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R16
Instances: 114
Attributes: 16
Mat I
FyS
FInf
IP
EF I
IdEx I
Alg y GA
HCuba
IdEx II
DPOO
EF II
IGS
MD
Mat II
EPCap
aprueba
Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method:
Best first.
Start set: no attributes
Search direction: forward
Stale search after 5 node expansions
Total number of subsets evaluated: 103
Merit of best subset found: 0.414

Attribute Subset Evaluator (supervised, Class (nominal): 16 aprueba):
CFS Subset Evaluator
Including locally predictive attributes
Selected attributes: 1,3,13,14 : 4
Mat I
FInf
MD
Mat II

Anexo B1.3

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.FCBFSearch -D
false -T -1.7976931348623157E308 -N -1
Relation: Modelo_1ro año-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R16
Instances: 114
Attributes: 16
Mat I
FyS
FInf
IP
EF I
IdEx I
Alg y GA
HCuba
IdEx II
DPOO
EF II
IGS
MD
Mat II
EPCap
aprueba
Evaluation mode: 10-fold cross-validation

Anexo B1.5

=== Run information ===

Evaluator:
weka.attributeSelection.InfoGainAttributeEval
Search: weka.attributeSelection.Ranker -T -
1.7976931348623157E308 -N -1
Relation: Modelo_1ro año-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R16
Instances: 114
Attributes: 16
Mat I
FyS
FInf
IP
EF I
IdEx I
Alg y GA
HCuba

Anexo B1.4

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.FCBFSearch -D
false -T -1.7976931348623157E308 -N -1
Relation: Modelo_1ro año-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R16
Instances: 114
Attributes: 16
Mat I
FyS
FInf
IP
EF I
IdEx I
Alg y GA
HCuba
IdEx II
DPOO
EF II
IGS
MD
Mat II
EPCap
aprueba
Evaluation mode: evaluate on all training data

Anexo B1.6

=== Run information ===

Evaluator:
weka.attributeSelection.InfoGainAttributeEval
Search: weka.attributeSelection.Ranker -T -
1.7976931348623157E308 -N -1
Relation: Modelo_1ro año-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R1-
weka.filters.unsupervised.attribute.Remove-R16
Instances: 114
Attributes: 16
Mat I
FyS
FInf
IP
EF I
IdEx I
Alg y GA
HCuba

IdEx II
DPOO
EF II
IGS
MD
Mat II
EPCap
aprueba
Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

average merit	average rank	attribute
0.531 +- 0.028	1 +- 0	1 Mat I
0.438 +- 0.031	2 +- 0	14 Mat II
0.356 +- 0.031	3.4 +- 0.49	2 FyS
0.325 +- 0.027	4.1 +- 1.14	13 MD
0.292 +- 0.024	5 +- 0.77	4 IP
0.281 +- 0.024	5.7 +- 0.78	10 DPOO
0.239 +- 0.019	7 +- 0.63	3 Flnf
0.203 +- 0.022	8.5 +- 1.12	9 IdEx II
0.191 +- 0.019	9 +- 0.89	7 Alg y GA
0.166 +- 0.015	10.1 +- 0.83	12 IGS
0.171 +- 0.019	10.2 +- 0.98	8 HCuba
0.119 +- 0.018	12.2 +- 0.6	6 IdEx I
0.095 +- 0.013	12.9 +- 0.3	15 EPCap
0.054 +- 0.015	13.9 +- 0.3	5 EF I
0.014 +- 0.004	15 +- 0	11 EF II

IdEx II
DPOO
EF II
IGS
MD
Mat II
EPCap
aprueba
Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method:
Attribute ranking.

Attribute Evaluator (supervised, Class (nominal):
16 aprueba):
Information Gain Ranking Filter

Ranked attributes:
0.5274 1 Mat I
0.4347 14 Mat II
0.3479 2 FyS
0.3213 13 MD
0.2845 4 IP
0.2743 10 DPOO
0.2337 3 Flnf
0.198 9 IdEx II
0.1852 7 Alg y GA
0.1657 8 HCuba
0.1603 12 IGS
0.112 6 IdEx I
0.0861 15 EPCap
0.0444 5 EF I
0.0102 11 EF II

Selected attributes:
1,14,2,13,4,10,3,9,7,8,12,6,15,5,11 : 15

Anexo B2.1

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.BestFirst -D 1 - N 5
Relation: Modelo 2do año-
weka.filters.unsupervised.attribute.Remove-R1-4,22
Instances: 120
Attributes: 18
mat iii
epcap
ed1
iia

Anexo B2.2

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.BestFirst -D 1 - N 5
Relation: Modelo 2do año-
weka.filters.unsupervised.attribute.Remove-R1-4,22
Instances: 120
Attributes: 18
mat iii
epcap
ed1
iia

idex 3
segNac
EF 3
elec 1
mat iv
Acomp
ED 2
BD
TSP
idex 4
Dnac
EF iv
Pprof
aprueba
Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

number of folds (%) attribute
10(100 %) 1 mat iii
1(10 %) 2 epcap
10(100 %) 3 ed1
10(100 %) 4 iia
0(0 %) 5 idex 3
0(0 %) 6 segNac
4(40 %) 7 EF 3
0(0 %) 8 elec 1
4(40 %) 9 mat iv
5(50 %) 10 Acomp
10(100 %) 11 ED 2
10(100 %) 12 BD
0(0 %) 13 TSP
0(0 %) 14 idex 4
0(0 %) 15 Dnac
0(0 %) 16 EF iv
3(30 %) 17 Pprof

Anexo B2.3

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.FCBFSearch -D true -T -1.7976931348623157E308 -N -1
Relation: Modelo 2do año-2clases-
weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16
Instances: 120
Attributes: 9
mat iii
ed1
iia
mat iv
A1omp
ED 2

idex 3
segNac
EF 3
elec 1
mat iv
Acomp
ED 2
BD
TSP
idex 4
Dnac
EF iv
Pprof
aprueba
Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method:
Best first.
Start set: no attributes
Search direction: forward
Stale search after 5 node expansions
Total number of subsets evaluated: 136
Merit of best subset found: 0.412

Attribute Subset Evaluator (supervised, Class (nominal): 18 aprueba):
CFS Subset Evaluator
Including locally predictive attributes

Selected attributes: 1,3,4,10,11,12 : 6
mat iii
ed1
iia
Acomp
ED 2
BD

Anexo B2.4

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.FCBFSearch -D true -T -1.7976931348623157E308 -N -1
Relation: Modelo 2do año-2clases-
weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16
Instances: 120
Attributes: 9
mat iii
ed1
iia
mat iv
A1omp
ED 2

BD
Pprof
aprueba2
Evaluation mode: 10-fold cross-validation

Anexo B2.5

=== Run information ===

Evaluator:
weka.attributeSelection.InfoGainAttributeEval
Search: weka.attributeSelection.Ranker -T -
1.7976931348623157E308 -N -1
Relation: Modelo 2do año-
weka.filters.unsupervised.attribute.Remove-R1-
4,22
Instances: 120
Attributes: 18
mat iii
epcap
ed1
iia
idex 3
segNac
EF 3
elec 1
mat iv
Acomp
ED 2
BD
TSP
idex 4
Dnac
EF iv
Pprof
aprueba
Evaluation mode: 10-fold cross-validation
=== Attribute selection 10 fold cross-validation
(stratified), seed: 1 ===

average merit	average rank	attribute
0.475 +- 0.028	1.1 +- 0.3	3 ed1
0.403 +- 0.028	2.6 +- 1.11	1 mat iii
0.4 +- 0.032	2.9 +- 0.7	12 BD
0.383 +- 0.027	3.6 +- 0.66	11 ED 2
0.329 +- 0.019	5.7 +- 1	9 mat iv
0.315 +- 0.038	6.4 +- 1.69	4 iia
0.314 +- 0.025	6.9 +- 1.14	10 Acomp
0.299 +- 0.015	7.5 +- 1.12	15 Dnac
0.291 +- 0.027	8.3 +- 0.9	2 epcap
0.23 +- 0.029	10.8 +- 0.98	17 Pprof
0.226 +- 0.028	10.9 +- 0.94	5 idex 3
0.198 +- 0.021	12.8 +- 1.33	6 segNac
0.196 +- 0.028	13.2 +- 1.17	13 TSP
0.178 +- 0.03	13.7 +- 2.28	7 EF 3
0.175 +- 0.026	14.4 +- 0.92	16 EF iv
0.156 +- 0.014	15.6 +- 0.8	14 idex 4

BD
Pprof
aprueba2
Evaluation mode: evaluate on all training data

Anexo B2.6

=== Run information ===

Evaluator:
weka.attributeSelection.InfoGainAttributeEval
Search: weka.attributeSelection.Ranker -T -
1.7976931348623157E308 -N -1
Relation: Modelo 2do año-
weka.filters.unsupervised.attribute.Remove-R1-
4,22
Instances: 120
Attributes: 18
mat iii
epcap
ed1
iia
idex 3
segNac
EF 3
elec 1
mat iv
Acomp
ED 2
BD
TSP
idex 4
Dnac
EF iv
Pprof
aprueba
Evaluation mode: evaluate on all training data
=== Attribute Selection on all input data ===
Search Method:Attribute ranking.

Attribute Evaluator (supervised, Class (nominal):
18 aprueba):
Information Gain Ranking Filter
Ranked attributes:
0.468 3 ed1
0.396 12 BD
0.394 1 mat iii
0.38 11 ED 2
0.323 9 mat iv
0.312 10 Acomp
0.303 4 iia
0.288 15 Dnac
0.279 2 epcap
0.222 17 Pprof
0.215 5 idex 3
0.186 13 TSP
0.185 6 segNac

0.133 +- 0.014 16.6 +- 0.92 8 elec 1

0.17 7 EF 3
0.167 16 EF iv
0.148 14 idex 4
0.124 8 elec 1
Selected attributes:
3,12,1,11,9,10,4,15,2,17,5,13,6,7,16,14,8 : 17

Anexo B3.1

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.BestFirst -D 1 - N 5
Relation: Modelo 1ro y 2do anyo-
weka.filters.unsupervised.attribute.Remove-R1,34
Instances: 134
Attributes: 33
Mat I
FyS
FInf
IP
EF I
IdEx I
Alg y GA
HCuba
IdEx II
DPOO
EF II
IGS
MD
Mat II
EPCap
mat iii
epcap
ed1
iia
idex 3
segNac
EF 3
elec 1
mat iv
Acomp
ED 2
BD
TSP
idex 4
Dnac
EF iv
Pprof
aprueba

Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

number of folds (%) attribute

Anexo B3.2

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.BestFirst -D 1 - N 5
Relation: Modelo 1ro y 2do anyo-
weka.filters.unsupervised.attribute.Remove-R1,34
Instances: 134
Attributes: 33
Mat I
FyS
FInf
IP
EF I
IdEx I
Alg y GA
HCuba
IdEx II
DPOO
EF II
IGS
MD
Mat II
EPCap
mat iii
epcap
ed1
iia
idex 3
segNac
EF 3
elec 1
mat iv
Acomp
ED 2
BD
TSP
idex 4
Dnac
EF iv
Pprof
aprueba

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method:
Best first.

10(100 %) 1 Mat I
 10(100 %) 2 FyS
 4(40 %) 3 Flnf
 10(100 %) 4 IP
 2(20 %) 5 EF I
 3(30 %) 6 IdEx I
 3(30 %) 7 Alg y GA
 1(10 %) 8 HCuba
 0(0 %) 9 IdEx II
 10(100 %) 10 DPOO
 0(0 %) 11 EF II
 0(0 %) 12 IGS
 10(100 %) 13 MD
 10(100 %) 14 Mat II
 0(0 %) 15 EPCap
 10(100 %) 16 mat iii
 1(10 %) 17 epcap
 10(100 %) 18 ed1
 10(100 %) 19 iia
 6(60 %) 20 idex 3
 0(0 %) 21 segNac
 10(100 %) 22 EF 3
 0(0 %) 23 elec 1
 6(60 %) 24 mat iv
 6(60 %) 25 Acomp
 10(100 %) 26 ED 2
 10(100 %) 27 BD
 0(0 %) 28 TSP
 0(0 %) 29 idex 4
 0(0 %) 30 Dnac
 1(10 %) 31 EF iv
 1(10 %) 32 Pprof

Start set: no attributes
 Search direction: forward
 Stale search after 5 node expansions
 Total number of subsets evaluated: 480
 Merit of best subset found: 0.203

Attribute Subset Evaluator (supervised, Class
 (nominal): 33 aprueba):
 CFS Subset Evaluator
 Including locally predictive attributes

Selected attributes:
 1,2,3,4,10,13,14,16,18,19,20,22,25,26,27 : 15

Mat I
 FyS
 Flnf
 IP
 DPOO
 MD
 Mat II
 mat iii
 ed1
 iia
 idex 3
 EF 3
 Acomp
 ED 2
 BD

Anexo B3.3

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
 Search: weka.attributeSelection.FCBFSearch -D
 false -T -1.7976931348623157E308 -N -1
 Relation: Modelo 1ro y 2do anyo-
 weka.filters.unsupervised.attribute.Remove-R1,34
 Instances: 134
 Attributes: 33

Mat I
 FyS
 Flnf
 IP
 EF I
 IdEx I
 Alg y GA
 HCuba
 IdEx II
 DPOO
 EF II
 IGS
 MD
 Mat II

Anexo B3.4

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval
 Search: weka.attributeSelection.FCBFSearch -D
 false -T -1.7976931348623157E308 -N -1
 Relation: Modelo 1ro y 2do anyo-
 weka.filters.unsupervised.attribute.Remove-R1,34
 Instances: 134
 Attributes: 33

Mat I
 FyS
 Flnf
 IP
 EF I
 IdEx I
 Alg y GA
 HCuba
 IdEx II
 DPOO
 EF II
 IGS
 MD
 Mat II

EPCap	EPCap
mat iii	mat iii
epcap	epcap
ed1	ed1
iia	iia
idex 3	idex 3
segNac	segNac
EF 3	EF 3
elec 1	elec 1
mat iv	mat iv
Acomp	Acomp
ED 2	ED 2
BD	BD
TSP	TSP
idex 4	idex 4
Dnac	Dnac
EF iv	EF iv
Pprof	Pprof
aprueba	aprueba
Evaluation mode: 10-fold cross-validation	Evaluation mode: evaluate on all training data

Anexo B3.5

=== Run information ===

Evaluator:
weka.attributeSelection.InfoGainAttributeEval
Search: weka.attributeSelection.Ranker -T -
1.7976931348623157E308 -N -1
Relation: Modelo 1ro y 2do anyo-
weka.filters.unsupervised.attribute.Remove-R1,34
Instances: 134
Attributes: 33
Mat I
FyS
Flnf
IP
EF I
IdEx I
Alg y GA
HCuba
IdEx II
DPOO
EF II
IGS
MD
Mat II
EPCap
mat iii
epcap
ed1
iia
idex 3
segNac
EF 3
elec 1
mat iv

Anexo B3.6

=== Run information ===

Evaluator:
weka.attributeSelection.InfoGainAttributeEval
Search: weka.attributeSelection.Ranker -T -
1.7976931348623157E308 -N -1
Relation: Modelo 1ro y 2do anyo-
weka.filters.unsupervised.attribute.Remove-R1,34
Instances: 134
Attributes: 33
Mat I
FyS
Flnf
IP
EF I
IdEx I
Alg y GA
HCuba
IdEx II
DPOO
EF II
IGS
MD
Mat II
EPCap
mat iii
epcap
ed1
iia
idex 3
segNac
EF 3
elec 1
mat iv

Acomp
ED 2
BD
TSP
idex 4
Dnac
EF iv
Pprof
aprueba
Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation
(stratified), seed: 1 ===

average merit	average rank	attribute
0.188 +- 0.014	1.6 +- 1.28	18 ed1
0.165 +- 0.014	4 +- 2.05	19 iia
0.16 +- 0.014	5.3 +- 2.49	2 FyS
0.158 +- 0.016	5.4 +- 2.65	27 BD
0.158 +- 0.016	5.8 +- 2.79	10 DPOO
0.159 +- 0.01	5.9 +- 2.26	13 MD
0.158 +- 0.015	6.1 +- 2.59	16 mat iii
0.156 +- 0.014	6.2 +- 2.52	26 ED 2
0.159 +- 0.013	6.6 +- 3.2	1 Mat I
0.14 +- 0.012	10.4 +- 2.84	4 IP
0.132 +- 0.015	11.6 +- 2.91	25 Acomp
0.13 +- 0.013	12.1 +- 1.58	24 mat iv
0.121 +- 0.007	13.2 +- 1.89	14 Mat II
0.119 +- 0.009	13.7 +- 2	17 epcap
0.121 +- 0.007	14.3 +- 1.49	30 Dnac
0.119 +- 0.014	14.7 +- 2.57	7 Alg y GA
0.1 +- 0.009	18 +- 1.95	20 idex 3
0.095 +- 0.009	19.3 +- 2	28 TSP
0.094 +- 0.011	19.6 +- 2.01	22 EF 3
0.09 +- 0.01	20.4 +- 1.74	31 EF iv
0.09 +- 0.011	20.7 +- 2.05	32 Pprof
0.087 +- 0.012	21.1 +- 3.59	3 Flnf
0.077 +- 0.007	24.1 +- 2.91	29 idex 4
0.075 +- 0.01	24.5 +- 2.29	9 IdEx II
0.073 +- 0.011	24.9 +- 3.01	6 IdEx I
0.074 +- 0.008	25.2 +- 2.68	21 segNac
0.066 +- 0.011	26.4 +- 2.91	5 EF I
0.068 +- 0.009	26.7 +- 1.19	12 IGS
0.062 +- 0.01	28.5 +- 1.69	8 HCuba
0.06 +- 0.008	28.7 +- 1.49	15 EPCap
0.044 +- 0.007	31.4 +- 0.49	23 elec 1
0.042 +- 0.009	31.6 +- 0.49	11 EF II

Acomp
ED 2
BD
TSP
idex 4
Dnac
EF iv
Pprof
aprueba
Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method:
Attribute ranking.

Attribute Evaluator (supervised, Class (nominal):
33 aprueba):
Information Gain Ranking Filter

Ranked attributes:

0.1853	18 ed1
0.1638	19 iia
0.1582	2 FyS
0.1575	13 MD
0.1565	27 BD
0.1563	1 Mat I
0.1562	10 DPOO
0.155	16 mat iii
0.1545	26 ED 2
0.1366	4 IP
0.1298	25 Acomp
0.1276	24 mat iv
0.1198	14 Mat II
0.1169	7 Alg y GA
0.1164	30 Dnac
0.1155	17 epcap
0.0971	20 idex 3
0.0917	28 TSP
0.091	22 EF 3
0.087	31 EF iv
0.0869	32 Pprof
0.0855	3 Flnf
0.0732	29 idex 4
0.0729	9 IdEx II
0.071	6 IdEx I
0.069	21 segNac
0.0664	12 IGS
0.0637	5 EF I
0.0597	8 HCuba
0.0581	15 EPCap
0.0416	11 EF II
0.0409	23 elec 1

Selected attributes:
18,19,2,13,27,1,10,16,26,4,25,24,14,7,30,17,20,28,
22,31,32,3,29,9,6,21,12,5,8,15,11,23 : 32

Anexo C.

Modelos resultantes del entrenamiento de la Red Neuronal para las diferentes variantes.

Anexo C1 - Modelo 1er año- v1 MLP

```
=== Run information ===
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:    Modelo_1er año
Instances:   114
Attributes:  16
             Mat I
             FyS
             FInf
             IP
             EF I
             Id0 I
             Alg y GA
             H1uba
             Id0 II
             DPOO
             EF II
             IGS
             MD
             Mat II
             EP1ap
             aprueba
Test mode:   3-fold cross-validation
=== Classifier model (full training set) ===
=== Summary ===

Correctly Classified Instances      92           80.7018 %
Incorrectly Classified Instances    22           19.2982 %
Kappa statistic                    0.5536
Mean absolute error                 0.1906
Root mean squared error            0.4068
Relative absolute error            44.5764 %
Root relative squared error        88.1465 %
Coverage of cases (0.95 level)    87.7193 %
Mean rel. region size (0.95 level) 57.0175 %
Total Number of Instances         114

=== Detailed Accuracy By Class ===

TP Rate   FP Rate   Precision   Recall   F-Measure   ROC Area   Class
0.848     0.286     0.87       0.848    0.859       0.86       si
0.714     0.152     0.676     0.714    0.694       0.86       no
Weighted Avg.
0.807     0.245     0.81       0.807    0.808       0.86

=== Confusion Matrix ===
  a  b   <-- classified as
67 12 |  a = si
10 25 |  b = no
```

Anexo C2 Modelo 1er año- v2 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation:     Modelo_1er año
Instances:    114
Attributes:   16
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPOO
              EF II
              IGS
              MD
              Mat II
              EP1ap
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===
Time taken to build model: 1.72 seconds
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	95	83.3333 %
Incorrectly Classified Instances	19	16.6667 %
Kappa statistic	0.5921	
Mean absolute error	0.1675	
Root mean squared error	0.3682	
Relative absolute error	39.186 %	
Root relative squared error	79.7826 %	
Coverage of cases (0.95 level)	90.3509 %	
Mean rel. region size (0.95 level)	57.8947 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class	
0.911	0.343	0.857	0.911	0.883	0.868	si	0.657
0.089	0.767	0.657	0.708	0.868	no		
Weighted Avg.							
0.833	0.265	0.829	0.833	0.829	0.868		

=== Confusion Matrix ===

```
a  b  <-- classified as
72  7  |  a = si
12 23  |  b = no
```

Anexo C3 Modelo 1er año- v3 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2

-N 500 -V 0 -S 0 -E 20 -H 0

Relation: Modelo_1er año

Instances: 114

Attributes: 16

Mat I

FyS

FInf

IP

EF I

Id0 I

Alg y GA

H1uba

Id0 II

DPOO

EF II

IGS

MD

Mat II

EP1ap

aprueba

Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 0.36 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	94	82.4561 %
Incorrectly Classified Instances	20	17.5439 %
Kappa statistic	0.5877	
Mean absolute error	0.1903	
Root mean squared error	0.3955	
Relative absolute error	44.5115 %	
Root relative squared error	85.6979 %	
Coverage of cases (0.95 level)	88.5965 %	
Mean rel. region size (0.95 level)	60.9649 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.873	0.286	0.873	0.873	0.873	0.843	si
0.714	0.127	0.714	0.714	0.714	0.843	no
Weighted Avg.						
0.825	0.237	0.825	0.825	0.825	0.843	

=== Confusion Matrix ===

a	b	<-- classified as
69	10	a = si
10	25	b = no

Anexo C4 Modelo 1er año- v4 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:     Modelo_1er año
Instances:    114
Attributes:   16
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPOO
              EF II
              IGS
              MD
              Mat II
              EP1ap
              aprueba
Test mode:    5-fold cross-validation
```

```
=== Classifier model (full training set) ===
Time taken to build model: 2.53 seconds
=== Stratified cross-validation ===
=== Summary ===
```

Correctly Classified Instances	91	79.8246 %
Incorrectly Classified Instances	23	20.1754 %
Kappa statistic	0.5062	
Mean absolute error	0.2113	
Root mean squared error	0.4371	
Relative absolute error	49.4645 %	
Root relative squared error	94.7655 %	
Coverage of cases (0.95 level)	84.2105 %	
Mean rel. region size (0.95 level)	57.0175 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.886	0.4	0.833	0.886	0.859	0.829	si
0.6	0.114	0.7	0.6	0.646	0.829	no
Weighted Avg.						
0.798	0.312	0.792	0.798	0.794	0.829	

=== Confusion Matrix ===

```
  a  b  <-- classified as
70  9  |  a = si
14 21  |  b = no
```

Anexo C5 Modelo 1er año- v5 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation:     Modelo_1er año
Instances:    114
Attributes:   16
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPOO
              EF II
              IGS
              MD
              Mat II
              EP1ap
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===
Time taken to build model: 9.02 seconds
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	92	80.7018 %
Incorrectly Classified Instances	22	19.2982 %
Kappa statistic	0.5605	
Mean absolute error	0.2034	
Root mean squared error	0.4101	
Relative absolute error	47.581 %	
Root relative squared error	88.871 %	
Coverage of cases (0.95 level)	89.4737 %	
Mean rel. region size (0.95 level)	61.8421 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.835	0.257	0.88	0.835	0.857	0.841	si
0.743	0.165	0.667	0.743	0.703	0.841	no
Weighted Avg.						
0.807	0.229	0.815	0.807	0.81	0.841	

=== Confusion Matrix ===

```
a  b  <-- classified as
66 13 | a = si
 9 26 | b = no
```

Anexo C6 Modelo 1er año- v6 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation:     Modelo_1er año
Instances:    114
Attributes:   16
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPOO
              EF II
              IGS
              MD
              Mat II
              EP1ap
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===
Time taken to build model: 14.06 seconds
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	90	78.9474 %
Incorrectly Classified Instances	24	21.0526 %
Kappa statistic	0.5278	
Mean absolute error	0.1997	
Root mean squared error	0.4209	
Relative absolute error	46.7133 %	
Root relative squared error	91.1952 %	
Coverage of cases (0.95 level)	86.8421 %	
Mean rel. region size (0.95 level)	55.7018 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.81	0.257	0.877	0.81	0.842	0.845	si
0.743	0.19	0.634	0.743	0.684	0.845	no
Weighted Avg.						
0.789	0.236	0.802	0.789	0.794	0.845	

=== Confusion Matrix ===

```
a  b  <-- classified as
64 15 | a = si
 9 26 | b = no
```

Anexo C7 Modelo 1er año- v7 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "20, 23"
Relation:     Modelo_1er año
Instances:    114
Attributes:   16
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPOO
              EF II
              IGS
              MD
              Mat II
              EP1ap
              aprueba
Test mode:    3-fold cross-validation
```

```
=== Classifier model (full training set) ===
Time taken to build model: 5.64 seconds
=== Stratified cross-validation ===
=== Summary ===
```

Correctly Classified Instances	91	79.8246 %
Incorrectly Classified Instances	23	20.1754 %
Kappa statistic	0.544	
Mean absolute error	0.2037	
Root mean squared error	0.4108	
Relative absolute error	47.6313 %	
Root relative squared error	89.0182 %	
Coverage of cases (0.95 level)	89.4737 %	
Mean rel. region size (0.95 level)	61.8421 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.823	0.257	0.878	0.823	0.85	0.841	si
0.743	0.177	0.65	0.743	0.693	0.841	no
Weighted Avg.						
0.798	0.233	0.808	0.798	0.802	0.841	

=== Confusion Matrix ===

```
a  b  <-- classified as
65 14 | a = si
 9 26 | b = no
```

Anexo C8 Modelo 1er año selrasg- v1 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:     Modelo_1er año selrasg
Instances:    114
Attributes:   11
              Mat I
              FyS
              FInf
              IP
              Alg y GA
              Id0 II
              DPOO
              IGS
              MD
              Mat II
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 2.05 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	97	85.0877 %
Incorrectly Classified Instances	17	14.9123 %
Kappa statistic	0.6523	
Mean absolute error	0.1654	
Root mean squared error	0.3575	
Relative absolute error	38.6788 %	
Root relative squared error	77.4566 %	
Coverage of cases (0.95 level)	90.3509 %	
Mean rel. region size (0.95 level)	60.5263 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.886	0.229	0.897	0.886	0.892	0.86	si
0.771	0.114	0.75	0.771	0.761	0.86	no
Weighted Avg.						
0.851	0.193	0.852	0.851	0.851	0.86	

=== Confusion Matrix ===

```
a  b  <-- classified as
70  9  |  a = si
 8 27  |  b = no
```

Anexo C9 Modelo 1er año selrasg- v2 MLP

=== Run information ===

```
Scheme:          weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation:        Modelo_1er año selrasg
Instances:       114
Attributes:      11
                  Mat I
                  FyS
                  FInf
                  IP
                  Alg y GA
                  Id0 II
                  DPOO
                  IGS
                  MD
                  Mat II
                  prueba
Test mode:       3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 0.95 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	94	82.4561 %
Incorrectly Classified Instances	20	17.5439 %
Kappa statistic	0.5942	
Mean absolute error	0.1741	
Root mean squared error	0.3836	
Relative absolute error	40.724 %	
Root relative squared error	83.1126 %	
Coverage of cases (0.95 level)	89.4737 %	
Mean rel. region size (0.95 level)	56.5789 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.861	0.257	0.883	0.861	0.872	0.824	si
0.743	0.139	0.703	0.743	0.722	0.824	no
Weighted Avg.						
0.825	0.221	0.828	0.825	0.826	0.824	

=== Confusion Matrix ===

```
a  b  <-- classified as
68 11 | a = si
 9 26 | b = no
```

Anexo C10 Modelo 1er año selrasg- v3 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H 0
Relation: Modelo_1er año selrasg
Instances: 114
Attributes: 11
Mat I
FyS
FInf
IP
Alg y GA
Id0 II
DPOO
IGS
MD
Mat II
aprueba
Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 0.31 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	99	86.8421 %
Incorrectly Classified Instances	15	13.1579 %
Kappa statistic	0.6883	
Mean absolute error	0.1587	
Root mean squared error	0.3526	
Relative absolute error	37.1187 %	
Root relative squared error	76.403 %	
Coverage of cases (0.95 level)	90.3509 %	
Mean rel. region size (0.95 level)	60.0877 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.911	0.229	0.9	0.911	0.906	0.863	si
0.771	0.089	0.794	0.771	0.783	0.863	no
Weighted Avg.						
0.868	0.186	0.867	0.868	0.868	0.863	

=== Confusion Matrix ===

a	b	<-- classified as
72	7	a = si
8	27	b = no

Anexo C11 Modelo 1er año selrasg- v4 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:     Modelo_1er año selrasg
Instances:    114
Attributes:   11
              Mat I
              FyS
              FInf
              IP
              Alg y GA
              Id0 II
              DPOO
              IGS
              MD
              Mat II
              prueba
Test mode:    5-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 1.92 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	95	83.3333 %
Incorrectly Classified Instances	19	16.6667 %
Kappa statistic	0.5921	
Mean absolute error	0.1805	
Root mean squared error	0.3906	
Relative absolute error	42.2604 %	
Root relative squared error	84.6684 %	
Coverage of cases (0.95 level)	87.7193 %	
Mean rel. region size (0.95 level)	59.2105 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.911	0.343	0.857	0.911	0.883	0.807	si
0.657	0.089	0.767	0.657	0.708	0.807	no
Weighted Avg.						
0.833	0.265	0.829	0.833	0.829	0.807	

=== Confusion Matrix ===

```
  a  b  <-- classified as
72  7  |  a = si
12 23  |  b = no
```

Anexo C12 Modelo 1er año selrasg- v5 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation:     Modelo_1er año selrasg
Instances:    114
Attributes:   11
              Mat I
              FyS
              FInf
              IP
              Alg y GA
              Id0 II
              DPOO
              IGS
              MD
              Mat II
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 8.17 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	97	85.0877 %
Incorrectly Classified Instances	17	14.9123 %
Kappa statistic	0.6523	
Mean absolute error	0.1712	
Root mean squared error	0.3709	
Relative absolute error	40.0416 %	
Root relative squared error	80.3715 %	
Coverage of cases (0.95 level)	90.3509 %	
Mean rel. region size (0.95 level)	60.0877 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.886	0.229	0.897	0.886	0.892	0.867	si
0.771	0.114	0.75	0.771	0.761	0.867	no
Weighted Avg.						
0.851	0.193	0.852	0.851	0.851	0.867	

=== Confusion Matrix ===

```
a  b  <-- classified as
70  9  |  a = si
 8 27  |  b = no
```

Anexo C13 Modelo 1er año selrasg- v6 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation: Modelo_1er año selrasg
Instances: 114
Attributes: 11
Mat I
FyS
FInf
IP
Alg y GA
Id0 II
DPOO
IGS
MD
Mat II
aprueba
Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 12.95 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	97	85.0877 %
Incorrectly Classified Instances	17	14.9123 %
Kappa statistic	0.6523	
Mean absolute error	0.1703	
Root mean squared error	0.3705	
Relative absolute error	39.8326 %	
Root relative squared error	80.2821 %	
Coverage of cases (0.95 level)	90.3509 %	
Mean rel. region size (0.95 level)	59.6491 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.886	0.229	0.897	0.886	0.892	0.864	si
0.771	0.114	0.75	0.771	0.761	0.864	no
Weighted Avg.						
0.851	0.193	0.852	0.851	0.851	0.864	

=== Confusion Matrix ===

a	b	<-- classified as
70	9	a = si
8	27	b = no

Anexo C14 Modelo 1er año selrasg- v7 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "20, 23"
Relation: Modelo_1er año selrasg
Instances: 114
Attributes: 11
Mat I
FyS
FInf
IP
Alg y GA
Id0 II
DPOO
IGS
MD
Mat II
aprueba
Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 4.89 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	97	85.0877 %
Incorrectly Classified Instances	17	14.9123 %
Kappa statistic	0.6523	
Mean absolute error	0.1693	
Root mean squared error	0.3684	
Relative absolute error	39.5842 %	
Root relative squared error	79.8194 %	
Coverage of cases (0.95 level)	90.3509 %	
Mean rel. region size (0.95 level)	60.0877 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.886	0.229	0.897	0.886	0.892	0.869	si
0.771	0.114	0.75	0.771	0.761	0.869	no
Weighted Avg.						
0.851	0.193	0.852	0.851	0.851	0.869	

=== Confusion Matrix ===

a	b	<-- classified as
70	9	a = si
8	27	b = no

Anexo C15 Modelo 2do año- v1 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"

Relation: Modelo 2do año-2clases

Instances: 120

Attributes: 18
mat iii
eplap
ed1
iaa
id0 3
segNa1
EF 3
ele1 1
mat iv
A1omp
ED 2
BD
TSP
id0 4
Dna1
EF iv
Pprof
aprueba2

Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 4.47 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	100	83.3333 %
Incorrectly Classified Instances	20	16.6667 %
Kappa statistic	0.6615	
Mean absolute error	0.182	
Root mean squared error	0.4011	
Relative absolute error	37.6268 %	
Root relative squared error	81.57 %	
Coverage of cases (0.95 level)	85.8333 %	
Mean rel. region size (0.95 level)	55 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.857	0.183	0.764	0.857	0.808	0.665	0.884	0.784	no
0.817	0.143	0.892	0.817	0.853	0.665	0.884	0.93	si
Weighted Avg.								
0.833	0.159	0.84	0.833	0.834	0.665	0.884	0.87	

=== Confusion Matrix ===

a	b	<-- classified as
42	7	a = no
13	58	b = si

Anexo C16 Modelo 2do año- v2 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2

-N 500 -V 0 -S 0 -E 20 -H i

Relation: Modelo 2do año-2clases

Instances: 120

Attributes: 18

mat iii

eplap

ed1

iaa

id0 3

segNa1

EF 3

ele1 1

mat iv

A1omp

ED 2

BD

TSP

id0 4

Dna1

EF iv

Pprof

aprueba2

Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 4.04 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	100	83.3333 %
Incorrectly Classified Instances	20	16.6667 %
Kappa statistic	0.6615	
Mean absolute error	0.1716	
Root mean squared error	0.3966	
Relative absolute error	35.4774 %	
Root relative squared error	80.664 %	
Coverage of cases (0.95 level)	86.6667 %	
Mean rel. region size (0.95 level)	54.5833 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.857	0.183	0.764	0.857	0.808	0.665	0.886	0.816	no
0.817	0.143	0.892	0.817	0.853	0.665	0.886	0.924	si
Weighted Avg.								
0.833	0.159	0.84	0.833	0.834	0.665	0.886	0.88	

=== Confusion Matrix ===

a	b	<-- classified as
42	7	a = no
13	58	b = si

Anexo C17 Modelo 2do año- v3 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2

-N 500 -V 0 -S 0 -E 20 -H 0

Relation: Modelo 2do año-2clases

Instances: 120

Attributes: 18

mat iii

eplap

ed1

iaa

id0 3

segNa1

EF 3

ele1 1

mat iv

A1omp

ED 2

BD

TSP

id0 4

Dna1

EF iv

Pprof

aprueba2

Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 0.97 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	99	82.5	%
Incorrectly Classified Instances	21	17.5	%
Kappa statistic	0.6412		
Mean absolute error	0.1904		
Root mean squared error	0.4004		
Relative absolute error	39.3666	%	
Root relative squared error	81.4329	%	
Coverage of cases (0.95 level)	89.1667	%	
Mean rel. region size (0.95 level)	57.5	%	
Total Number of Instances	120		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.816	0.169	0.769	0.816	0.792	0.642	0.877	0.781	no
0.831	0.184	0.868	0.831	0.849	0.642	0.877	0.921	si
Weighted Avg.								
0.825	0.178	0.827	0.825	0.826	0.642	0.877	0.864	

=== Confusion Matrix ===

a	b	<-- classified as
40	9	a = no
12	59	b = si

Anexo C18 Modelo 2do año- v4 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2

-N 500 -V 0 -S 0 -E 20 -H "12, 10"

Relation: Modelo 2do año-2clases

Instances: 120

Attributes: 18
mat iii
eplap
ed1
iaa
id0 3
segNa1
EF 3
ele1 1
mat iv
A1omp
ED 2
BD
TSP
id0 4
Dna1
EF iv
Pprof
aprueba2

Test mode: 5-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 4.42 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	98	81.6667 %
Incorrectly Classified Instances	22	18.3333 %
Kappa statistic	0.6253	
Mean absolute error	0.1849	
Root mean squared error	0.4088	
Relative absolute error	38.2357 %	
Root relative squared error	83.1363 %	
Coverage of cases (0.95 level)	89.1667 %	
Mean rel. region size (0.95 level)	55.8333 %	
Total Number of Instances	120	

== Detailed Accuracy By Class ==

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.816	0.183	0.755	0.816	0.784	0.627	0.888	0.801	no
0.817	0.184	0.866	0.817	0.841	0.627	0.888	0.926	si
Weighted Avg.								
0.817	0.183	0.82	0.817	0.818	0.627	0.888	0.875	

=== Confusion Matrix ===

a	b	<-- classified as
40	9	a = no
13	58	b = si

Anexo C19 Modelo 2do año- v5 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"

Relation: Modelo 2do año-2clases

Instances: 120

Attributes: 18
mat iii
eplap
ed1
iaa
id0 3
segNa1
EF 3
ele1 1
mat iv
A1omp
ED 2
BD
TSP
id0 4
Dna1
EF iv
Pprof
aprueba2

Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 18.75 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	97	80.8333 %
Incorrectly Classified Instances	23	19.1667 %
Kappa statistic	0.6095	
Mean absolute error	0.1987	
Root mean squared error	0.4248	
Relative absolute error	41.0818 %	
Root relative squared error	86.3906 %	
Coverage of cases (0.95 level)	85 %	
Mean rel. region size (0.95 level)	55.4167 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.816	0.197	0.741	0.816	0.777	0.612	0.878	0.771	no
0.803	0.184	0.864	0.803	0.832	0.612	0.878	0.927	si
Weighted Avg.								
0.808	0.189	0.813	0.808	0.809	0.612	0.878	0.863	

=== Confusion Matrix ===

a	b	<-- classified as
40	9	a = no
14	57	b = si

Anexo C20 Modelo 2do año- v6 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2

-N 500 -V 0 -S 0 -E 20 -H "50, 30"

Relation: Modelo 2do año-2clases

Instances: 120

Attributes: 18
mat iii
eplap
ed1
iaa
id0 3
segNa1
EF 3
ele1 1
mat iv
A1omp
ED 2
BD
TSP
id0 4
Dna1
EF iv
Pprof
aprueba2

Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 29.09 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	96	80	%
Incorrectly Classified Instances	24	20	%
Kappa statistic	0.5913		
Mean absolute error	0.1988		
Root mean squared error	0.4287		
Relative absolute error	41.0985 %		
Root relative squared error	87.1933 %		
Coverage of cases (0.95 level)	85 %		
Mean rel. region size (0.95 level)	55 %		
Total Number of Instances	120		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.796	0.197	0.736	0.796	0.765	0.593	0.883	0.781	no
0.803	0.204	0.851	0.803	0.826	0.593	0.883	0.93	si
Weighted Avg.								
0.8	0.201	0.804	0.8	0.801	0.593	0.883	0.869	

=== Confusion Matrix ===

```
a b <-- classified as
39 10 | a = no
14 57 | b = si
```

Anexo C21 Modelo 2do año- v7 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2

-N 500 -V 0 -S 0 -E 20 -H "20, 23"

Relation: Modelo 2do año-2clases

Instances: 120

Attributes: 18
mat iii
eplap
ed1
iaa
id0 3
segNa1
EF 3
ele1 1
mat iv
A1omp
ED 2
BD
TSP
id0 4
Dna1
EF iv
Pprof
aprueba2

Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 10.62 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	100	83.3333 %
Incorrectly Classified Instances	20	16.6667 %
Kappa statistic	0.6615	
Mean absolute error	0.1771	
Root mean squared error	0.4012	
Relative absolute error	36.6075 %	
Root relative squared error	81.5954 %	
Coverage of cases (0.95 level)	85 %	
Mean rel. region size (0.95 level)	54.5833 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.857	0.183	0.764	0.857	0.808	0.665	0.881	0.768	no
0.817	0.143	0.892	0.817	0.853	0.665	0.881	0.929	si
Weighted Avg.								
0.833	0.159	0.84	0.833	0.834	0.665	0.881	0.863	

=== Confusion Matrix ===

```
a b <-- classified as
42 7 | a = no
13 58 | b = si
```

Anexo C22 Modelo 2do año selrasg- v1 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"

Relation: Modelo 2do año-2clases-

weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16

Instances: 120

Attributes: 9

mat iii

ed1

iaa

mat iv

Alomp

ED 2

BD

Pprof

aprueba2

Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 3.42 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	104	86.6667 %
Incorrectly Classified Instances	16	13.3333 %
Kappa statistic	0.7205	
Mean absolute error	0.1483	
Root mean squared error	0.351	
Relative absolute error	30.6568 %	
Root relative squared error	71.3878 %	
Coverage of cases (0.95 level)	90.8333 %	
Mean rel. region size (0.95 level)	57.5 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.796	0.085	0.867	0.796	0.83	0.722	0.89	0.863	no
0.915	0.204	0.867	0.915	0.89	0.722	0.89	0.909	si
Weighted Avg.								
0.867	0.155	0.867	0.867	0.866	0.722	0.89	0.891	

=== Confusion Matrix ===

a	b	<-- classified as
39	10	a = no
6	65	b = si

Anexo C23 Modelo 2do año selrasg- v2 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation:     Modelo 2do año-2clases-
weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16
Instances:    120
Attributes:    9
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               Pprof
               aprueba2
Test mode:     3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 1.14 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	105	87.5	%
Incorrectly Classified Instances	15	12.5	%
Kappa statistic	0.7371		
Mean absolute error	0.1394		
Root mean squared error	0.3439		
Relative absolute error	28.8109	%	
Root relative squared error	69.9323	%	
Coverage of cases (0.95 level)	89.1667	%	
Mean rel. region size (0.95 level)	55.8333	%	
Total Number of Instances	120		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.796	0.07	0.886	0.796	0.839	0.74	0.887	0.87	no
0.93	0.204	0.868	0.93	0.898	0.74	0.887	0.909	si
Weighted Avg.								
0.875	0.15	0.876	0.875	0.874	0.74	0.887	0.893	

=== Confusion Matrix ===

```
a  b  <-- classified as
39 10 | a = no
 5 66 | b = si
```

Anexo C24 Modelo 2do año selrasg- v3 MLP

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2

-N 500 -V 0 -S 0 -E 20 -H o

Relation: Modelo 2do año-2clases-

weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16

Instances: 120

Attributes: 9

mat iii

ed1

iaa

mat iv

Alomp

ED 2

BD

Pprof

aprueba2

Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 0.41 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	106	88.3333 %
Incorrectly Classified Instances	14	11.6667 %
Kappa statistic	0.7523	
Mean absolute error	0.1493	
Root mean squared error	0.3255	
Relative absolute error	30.8626 %	
Root relative squared error	66.2065 %	
Coverage of cases (0.95 level)	95 %	
Mean rel. region size (0.95 level)	68.3333 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.776	0.042	0.927	0.776	0.844	0.76	0.902	0.857	no
0.958	0.224	0.861	0.958	0.907	0.76	0.902	0.912	si
Weighted Avg.								
0.883	0.15	0.888	0.883	0.881	0.76	0.902	0.89	

=== Confusion Matrix ===

a	b	<-- classified as
38	11	a = no
3	68	b = si

Anexo C25 Modelo 2do año selrasg- v4 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:     Modelo 2do año-2clases-
weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16
Instances:    120
Attributes:    9
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               Pprof
               aprueba2
Test mode:     5-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 3.47 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	102	85	%
Incorrectly Classified Instances	18	15	%
Kappa statistic	0.6794		
Mean absolute error	0.158		
Root mean squared error	0.3777		
Relative absolute error	32.6627 %		
Root relative squared error	76.8186 %		
Coverage of cases (0.95 level)	87.5 %		
Mean rel. region size (0.95 level)	54.5833 %		
Total Number of Instances	120		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.714	0.056	0.897	0.714	0.795	0.69	0.857	0.807	no
0.944	0.286	0.827	0.944	0.882	0.69	0.857	0.89	si
Weighted Avg.								
0.85	0.192	0.856	0.85	0.846	0.69	0.857	0.856	

=== Confusion Matrix ===

```
a  b  <-- classified as
35 14 | a = no
 4 67 | b = si
```

Anexo C26 Modelo 2do año selrasg- v5 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation:     Modelo 2do año-2clases-
weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16
Instances:    120
Attributes:    9
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               Pprof
               aprueba2
Test mode:     3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 15.08 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	104	86.6667 %
Incorrectly Classified Instances	16	13.3333 %
Kappa statistic	0.7205	
Mean absolute error	0.1655	
Root mean squared error	0.3573	
Relative absolute error	34.2168 %	
Root relative squared error	72.6675 %	
Coverage of cases (0.95 level)	90.8333 %	
Mean rel. region size (0.95 level)	61.6667 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.796	0.085	0.867	0.796	0.83	0.722	0.894	0.88	no
0.915	0.204	0.867	0.915	0.89	0.722	0.894	0.917	si
Weighted Avg.								
0.867	0.155	0.867	0.867	0.866	0.722	0.894	0.901	

=== Confusion Matrix ===

```
a  b  <-- classified as
39 10 | a = no
6  65 | b = si
```

Anexo C27 Modelo 2do año selrasg- v6 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation:     Modelo 2do año-2clases-
weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16
Instances:    120
Attributes:    9
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               Pprof
               aprueba2
Test mode:     3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 25.99 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	106	88.3333 %
Incorrectly Classified Instances	14	11.6667 %
Kappa statistic	0.7555	
Mean absolute error	0.1526	
Root mean squared error	0.3454	
Relative absolute error	31.5431 %	
Root relative squared error	70.2542 %	
Coverage of cases (0.95 level)	90.8333 %	
Mean rel. region size (0.95 level)	60 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.816	0.07	0.889	0.816	0.851	0.757	0.896	0.9	no
0.93	0.184	0.88	0.93	0.904	0.757	0.896	0.91	si
Weighted Avg.								
0.883	0.137	0.884	0.883	0.882	0.757	0.896	0.906	

=== Confusion Matrix ===

```
a  b  <-- classified as
40  9  |  a = no
 5 66  |  b = si
```

Anexo C28 Modelo 2do año selrasg- v7 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "20, 23"
Relation:     Modelo 2do año-2clases-
weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16
Instances:    120
Attributes:    9
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               Pprof
               aprueba2
Test mode:     3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 8.91 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	103	85.8333 %
Incorrectly Classified Instances	17	14.1667 %
Kappa statistic	0.7002	
Mean absolute error	0.152	
Root mean squared error	0.3581	
Relative absolute error	31.4313 %	
Root relative squared error	72.8356 %	
Coverage of cases (0.95 level)	91.6667 %	
Mean rel. region size (0.95 level)	58.3333 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.755	0.07	0.881	0.755	0.813	0.706	0.886	0.883	no
0.93	0.245	0.846	0.93	0.886	0.706	0.886	0.9	si
Weighted Avg.								
0.858	0.174	0.86	0.858	0.856	0.706	0.886	0.893	

=== Confusion Matrix ===

```
a  b  <-- classified as
37 12 | a = no
 5 66 | b = si
```

Anexo C29 Modelo 1er y 2do año- v1 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:     Modelo 1ro y 2do_año
Instances:    134
Attributes:   33
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPOO
              EF II
              IGS
              MD
              Mat II
              EP1ap
              mat iii
              eplap
              ed1
              iia
              id0 3
              segNa1
              EF 3
              ele1 1
              mat iv
              A1omp
              ED 2
              BD
              TSP
              id0 4
              Dna1
              EF iv
              Pprof
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 4.28 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	115	85.8209 %
Incorrectly Classified Instances	19	14.1791 %
Kappa statistic	0.7164	
Mean absolute error	0.1384	
Root mean squared error	0.3498	
Relative absolute error	27.675 %	
Root relative squared error	69.9468 %	
Coverage of cases (0.95 level)	91.0448 %	
Mean rel. region size (0.95 level)	54.4776 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	ROCArea	Class
0.851	0.134	0.864	0.851	0.857	0.919	no
0.866	0.149	0.853	0.866	0.859	0.919	si
Weighted Avg.						
0.858	0.142	0.858	0.858	0.858	0.919	

=== Confusion Matrix ===

```
a  b  <-- classified as
57 10 | a = no
 9 58 | b = si
```

Anexo C30 Modelo 1er y 2do año- v2 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation:     Modelo 1ro y 2do_año
Instances:    134
Attributes:   33
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPO0
              EF II
              IGS
              MD
              Mat II
              EP1ap
              mat iii
              ep1ap
              ed1
              iia
              id0 3
              segNa1
              EF 3
              ele1 1
              mat iv
              A1omp
              ED 2
              BD
              TSP
              id0 4
              Dna1
              EF iv
              Pprof
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 7.5 seconds

```
=== Stratified cross-validation ===  
=== Summary ===
```

Correctly Classified Instances	113	84.3284 %
Incorrectly Classified Instances	21	15.6716 %
Kappa statistic	0.6866	
Mean absolute error	0.1451	
Root mean squared error	0.3576	
Relative absolute error	29.0105 %	
Root relative squared error	71.5033 %	
Coverage of cases (0.95 level)	92.5373 %	
Mean rel. region size (0.95 level)	55.597 %	
Total Number of Instances	134	

```
=== Detailed Accuracy By Class ===
```

TPRate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.851	0.164	0.838	0.851	0.844	0.93	no
0.836	0.149	0.848	0.836	0.842	0.93	si
Weighted Avg.						
0.843	0.157	0.843	0.843	0.843	0.93	

```
=== Confusion Matrix ===
```

```
  a  b  <-- classified as  
57 10 |  a = no  
11 56 |  b = si
```

Anexo C31 Modelo 1er y 2do año- v3 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H o
Relation:     Modelo 1ro y 2do_año
Instances:    134
Attributes:   33
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPOO
              EF II
              IGS
              MD
              Mat II
              EP1ap
              mat iii
              eplap
              ed1
              iia
              id0 3
              segNa1
              EF 3
              ele1 1
              mat iv
              A1omp
              ED 2
              BD
              TSP
              id0 4
              Dna1
              EF iv
              Pprof
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 0.69 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	116	86.5672 %
Incorrectly Classified Instances	18	13.4328 %
Kappa statistic	0.7313	
Mean absolute error	0.1384	
Root mean squared error	0.3414	
Relative absolute error	27.6738 %	
Root relative squared error	68.2757 %	
Coverage of cases (0.95 level)	91.791 %	
Mean rel. region size (0.95 level)	55.597 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.881	0.149	0.855	0.881	0.868	0.934	no
0.851	0.119	0.877	0.851	0.864	0.934	si
Weighted Avg.						
0.866	0.134	0.866	0.866	0.866	0.934	

=== Confusion Matrix ===

```
a  b  <-- classified as
59  8 | a = no
10 57 | b = si
```

Anexo C32 Modelo 1er y 2do año- v4 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:     Modelo 1ro y 2do_año
Instances:    134
Attributes:   33
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPO0
              EF II
              IGS
              MD
              Mat II
              EP1ap
              mat iii
              ep1ap
              ed1
              iia
              id0 3
              segNa1
              EF 3
              ele1 1
              mat iv
              A1omp
              ED 2
              BD
              TSP
              id0 4
              Dna1
              EF iv
              Pprof
              aprueba
Test mode:    5-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 4.11 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	115	85.8209 %
Incorrectly Classified Instances	19	14.1791 %
Kappa statistic	0.7164	
Mean absolute error	0.1492	
Root mean squared error	0.3598	
Relative absolute error	29.8399 %	
Root relative squared error	71.9371 %	
Coverage of cases (0.95 level)	89.5522 %	
Mean rel. region size (0.95 level)	55.2239 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.836	0.119	0.875	0.836	0.855	0.911	no
0.881	0.164	0.843	0.881	0.861	0.911	si
Weighted Avg.						
0.858	0.142	0.859	0.858	0.858	0.911	

=== Confusion Matrix ===

```
a  b  <-- classified as
56 11 | a = no
 8 59 | b = si
```

Anexo C33 Modelo 1er y 2do año- v5 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation:     Modelo 1ro y 2do_año
Instances:    134
Attributes:   33
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPO0
              EF II
              IGS
              MD
              Mat II
              EP1ap
              mat iii
              ep1ap
              ed1
              iia
              id0 3
              segNa1
              EF 3
              ele1 1
              mat iv
              A1omp
              ED 2
              BD
              TSP
              id0 4
              Dna1
              EF iv
              Pprof
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 14.7 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	114	85.0746 %
Incorrectly Classified Instances	20	14.9254 %
Kappa statistic	0.7015	
Mean absolute error	0.1415	
Root mean squared error	0.3502	
Relative absolute error	28.2967 %	
Root relative squared error	70.024 %	
Coverage of cases (0.95 level)	92.5373 %	
Mean rel. region size (0.95 level)	55.2239 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.836	0.134	0.862	0.836	0.848	0.929	no
0.866	0.164	0.841	0.866	0.853	0.929	si
Weighted Avg.						
0.851	0.149	0.851	0.851	0.851	0.929	

=== Confusion Matrix ===

```
a  b  <-- classified as
56 11 | a = no
 9 58 | b = si
```

Anexo C34 Modelo 1er y 2do año- v6 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation:     Modelo 1ro y 2do_año
Instances:    134
Attributes:   33
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPO0
              EF II
              IGS
              MD
              Mat II
              EP1ap
              mat iii
              ep1ap
              ed1
              iia
              id0 3
              segNa1
              EF 3
              ele1 1
              mat iv
              A1omp
              ED 2
              BD
              TSP
              id0 4
              Dna1
              EF iv
              Pprof
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 21.73 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	114	85.0746 %
Incorrectly Classified Instances	20	14.9254 %
Kappa statistic	0.7015	
Mean absolute error	0.1396	
Root mean squared error	0.3473	
Relative absolute error	27.9182 %	
Root relative squared error	69.4389 %	
Coverage of cases (0.95 level)	92.5373 %	
Mean rel. region size (0.95 level)	55.2239 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.851	0.149	0.851	0.851	0.851	0.93	no
0.851	0.149	0.851	0.851	0.851	0.93	si
Weighted Avg.						
0.851	0.149	0.851	0.851	0.851	0.93	

=== Confusion Matrix ===

```
a  b  <-- classified as
57 10 | a = no
10 57 | b = si
```

Anexo C35 Modelo 1er y 2do año- v7 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "20, 23"
Relation:     Modelo 1ro y 2do_año
Instances:    134
Attributes:   33
              Mat I
              FyS
              FInf
              IP
              EF I
              Id0 I
              Alg y GA
              H1uba
              Id0 II
              DPO0
              EF II
              IGS
              MD
              Mat II
              EP1ap
              mat iii
              ep1ap
              ed1
              iia
              id0 3
              segNa1
              EF 3
              ele1 1
              mat iv
              A1omp
              ED 2
              BD
              TSP
              id0 4
              Dna1
              EF iv
              Pprof
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 8.66 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	115	85.8209 %
Incorrectly Classified Instances	19	14.1791 %
Kappa statistic	0.7164	
Mean absolute error	0.1337	
Root mean squared error	0.3356	
Relative absolute error	26.745 %	
Root relative squared error	67.0997 %	
Coverage of cases (0.95 level)	92.5373 %	
Mean rel. region size (0.95 level)	55.2239 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.851	0.134	0.864	0.851	0.857	0.93	no
0.866	0.149	0.853	0.866	0.859	0.93	si
Weighted Avg.						
0.858	0.142	0.858	0.858	0.858	0.93	

=== Confusion Matrix ===

```
a  b  <-- classified as
57 10 | a = no
 9 58 | b = si
```

Anexo C36 Modelo 1er y 2do año selrasg- v1 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:     Modelo 1ro y 2do año sel_rasg
Instances:    134
Attributes:    14
               Mat I
               FyS
               IP
               DPOO
               MD
               Mat II
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               aprueba
Test mode:     3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 2.42 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	109	81.3433 %
Incorrectly Classified Instances	25	18.6567 %
Kappa statistic	0.6269	
Mean absolute error	0.1812	
Root mean squared error	0.4112	
Relative absolute error	36.2258 %	
Root relative squared error	82.2327 %	
Coverage of cases (0.95 level)	84.3284 %	
Mean rel. region size (0.95 level)	52.2388 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.821	0.194	0.809	0.821	0.815	0.87	no
0.806	0.179	0.818	0.806	0.812	0.87	si
Weighted Avg.						
0.813	0.187	0.814	0.813	0.813	0.87	

=== Confusion Matrix ===

```
a  b  <-- classified as
55 12 | a = no
13 54 | b = si
```

Anexo C37 Modelo 1er y 2do año selrasg

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation:     Modelo 1ro y 2do año sel_rasg
Instances:    134
Attributes:   14
              Mat I
              FyS
              IP
              DPOO
              MD
              Mat II
              mat iii
              ed1
              iia
              mat iv
              A1omp
              ED 2
              BD
              aprueba
Test mode:    3-fold cross-validation
```

```
=== Classifier model (full training set) ===
Time taken to build model: 1.92 seconds
=== Stratified cross-validation ===
=== Summary ===
```

Correctly Classified Instances	109	81.3433 %
Incorrectly Classified Instances	25	18.6567 %
Kappa statistic	0.6269	
Mean absolute error	0.1822	
Root mean squared error	0.4032	
Relative absolute error	36.4375 %	
Root relative squared error	80.62 %	
Coverage of cases (0.95 level)	86.5672 %	
Mean rel. region size (0.95 level)	55.597 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.836	0.209	0.8	0.836	0.818	0.88	no
0.791	0.164	0.828	0.791	0.809	0.88	si
Weighted Avg.						
0.813	0.187	0.814	0.813	0.813	0.88	

=== Confusion Matrix ===

```
  a  b  <-- classified as
56 11 |  a = no
14 53 |  b = si
```

Anexo C38 Modelo 1er y 2do año selrasg- v3 MLP

=== Run information ===

```
Scheme:          weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H o
Relation:        Modelo 1ro y 2do año sel_rasg
Instances:       134
Attributes:      14
                  Mat I
                  FyS
                  IP
                  DPOO
                  MD
                  Mat II
                  mat iii
                  ed1
                  iia
                  mat iv
                  Alomp
                  ED 2
                  BD
                  aprueba
Test mode:       3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 0.41 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	109	81.3433 %
Incorrectly Classified Instances	25	18.6567 %
Kappa statistic	0.6269	
Mean absolute error	0.1883	
Root mean squared error	0.4194	
Relative absolute error	37.6578 %	
Root relative squared error	83.8668 %	
Coverage of cases (0.95 level)	83.5821 %	
Mean rel. region size (0.95 level)	52.6119 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.806	0.179	0.818	0.806	0.812	0.864	no
0.821	0.194	0.809	0.821	0.815	0.864	si
Weighted Avg.						
0.813	0.187	0.814	0.813	0.813	0.864	

=== Confusion Matrix ===

```
  a  b  <-- classified as
54 13 | a = no
12 55 | b = si
```

Anexo C39 Modelo 1er y 2do año selrasg- v4 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:    Modelo 1ro y 2do año sel_rasg
Instances:   134
Attributes:  14
              Mat I
              FyS
              IP
              DPOO
              MD
              Mat II
              mat iii
              ed1
              iia
              mat iv
              Alomp
              ED 2
              BD
              aprueba
Test mode:   5-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 2.78 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	111	82.8358 %
Incorrectly Classified Instances	23	17.1642 %
Kappa statistic	0.6567	
Mean absolute error	0.2004	
Root mean squared error	0.4053	
Relative absolute error	40.0759 %	
Root relative squared error	81.0401 %	
Coverage of cases (0.95 level)	88.806 %	
Mean rel. region size (0.95 level)	64.1791 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.866	0.209	0.806	0.866	0.835	0.855	no
0.791	0.134	0.855	0.791	0.822	0.855	si
Weighted Avg.						
0.828	0.172	0.83	0.828	0.828	0.855	

=== Confusion Matrix ===

```
  a  b  <-- classified as
58  9  |  a = no
14 53  |  b = si
```

Anexo C40 Modelo 1er y 2do año selrasg- v5 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation:     Modelo 1ro y 2do año sel_rasg
Instances:    134
Attributes:   14
              Mat I
              FyS
              IP
              DPOO
              MD
              Mat II
              mat iii
              ed1
              iia
              mat iv
              A1omp
              ED 2
              BD
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 10.05 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	110	82.0896 %
Incorrectly Classified Instances	24	17.9104 %
Kappa statistic	0.6418	
Mean absolute error	0.1754	
Root mean squared error	0.4053	
Relative absolute error	35.0742 %	
Root relative squared error	81.034 %	
Coverage of cases (0.95 level)	85.0746 %	
Mean rel. region size (0.95 level)	52.2388 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.851	0.209	0.803	0.851	0.826	0.862	no
0.791	0.149	0.841	0.791	0.815	0.862	si
Weighted Avg.						
0.821	0.179	0.822	0.821	0.821	0.862	

=== Confusion Matrix ===

```
  a  b  <-- classified as
57 10 |  a = no
14 53 |  b = si
```

Anexo C41 Modelo 1er y 2do año selrasg- v6 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation:     Modelo 1ro y 2do año sel_rasg
Instances:    134
Attributes:   14
              Mat I
              FyS
              IP
              DPOO
              MD
              Mat II
              mat iii
              ed1
              iia
              mat iv
              Alomp
              ED 2
              BD
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 15.95 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	111	82.8358 %
Incorrectly Classified Instances	23	17.1642 %
Kappa statistic	0.6567	
Mean absolute error	0.1705	
Root mean squared error	0.4023	
Relative absolute error	34.0876 %	
Root relative squared error	80.4499 %	
Coverage of cases (0.95 level)	85.0746 %	
Mean rel. region size (0.95 level)	52.2388 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.866	0.209	0.806	0.866	0.835	0.872	no
0.791	0.134	0.855	0.791	0.822	0.872	si
Weighted Avg.						
0.828	0.172	0.83	0.828	0.828	0.872	

=== Confusion Matrix ===

```
  a  b  <-- classified as
58  9  |  a = no
14 53  |  b = si
```

Anexo C42 Modelo 1er y 2do año selrasg- v7 MLP

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "20, 23"
Relation:     Modelo 1ro y 2do año sel_rasg
Instances:    134
Attributes:   14
              Mat I
              FyS
              IP
              DPOO
              MD
              Mat II
              mat iii
              ed1
              iia
              mat iv
              Alomp
              ED 2
              BD
              aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

Time taken to build model: 6.33 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	111	82.8358 %
Incorrectly Classified Instances	23	17.1642 %
Kappa statistic	0.6567	
Mean absolute error	0.1723	
Root mean squared error	0.4017	
Relative absolute error	34.4517 %	
Root relative squared error	80.3274 %	
Coverage of cases (0.95 level)	85.0746 %	
Mean rel. region size (0.95 level)	52.6119 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.881	0.224	0.797	0.881	0.837	0.869	no
0.776	0.119	0.867	0.776	0.819	0.869	si
Weighted Avg.						
0.828	0.172	0.832	0.828	0.828	0.869	

=== Confusion Matrix ===

```
  a  b  <-- classified as
59  8  |  a = no
15 52  |  b = si
```

Anexo D.

Conjuntos de pruebas para la evaluación de los modelos resultantes del entrenamiento.

Anexo D 1.1.

Conjunto de prueba Modelo 1er Año.

Relation 'CP-Modelo_1er año.csv'

Mat I,FyS,FInf,IP,EF I ,Id0 I,Alg y GA,H1uba,Id0 II,DPOO,EF II,IGS,MD,Mat II,EP1ap,aprueba
5,5,4,5,5,5,4,3,4,5,5,4,4,5,si
3,4,4,4,5,5,3,4,4,5,5,4,3,3,3,si
4,3,4,2,5,4,3,5,3,3,5,3,3,3,4,no
3,4,4,3,5,4,4,4,3,5,5,3,3,3,3,si
3,4,4,2,5,4,3,4,3,3,5,3,3,3,4,no
3,4,4,3,5,5,3,4,3,3,5,3,3,3,4,no
3,5,5,5,5,5,4,5,3,4,5,4,3,4,4,si
3,4,4,2,5,5,3,3,3,2,5,3,3,2,4,no
2,3,4,2,5,4,3,3,3,2,5,3,3,2,4,no
3,5,5,3,5,4,3,5,3,4,5,4,3,5,4,si
3,5,5,3,5,5,4,4,4,3,5,3,3,4,4,si
3,4,4,3,5,4,3,4,3,3,5,4,3,4,3,no
4,4,5,5,5,5,3,4,5,5,5,4,3,3,si
4,4,4,3,5,5,3,4,5,3,5,3,3,5,3,si

Anexo D 1.2.
Conjunto de prueba Modelo 2do Año.

```
@relation 'Modelo 2do año'

@attribute 'Mat III' numeric
@attribute Epcap numeric
@attribute 'ED I' numeric
@attribute IIA numeric
@attribute 'IdEx III' numeric
@attribute SegNac numeric
@attribute 'EF 3' numeric
@attribute 'elec ' numeric
@attribute 'Mat IV' numeric
@attribute Argomp numeric
@attribute 'ED 2' numeric
@attribute BD numeric
@attribute TSP numeric
@attribute 'IdEx IV' numeric
@attribute DeNac numeric
@attribute 'EF iv' numeric
@attribute Pprof numeric
@attribute aprueba {no,si}

@data
2,3,2,2,3,2,4,3,3,3,2,2,3,3,2,4,2,no
2,3,2,2,3,3,5,4,3,2,2,3,4,4,3,5,2,no
2,2,2,2,4,3,5,3,3,2,2,2,3,4,3,5,2,no
4,3,4,3,4,4,4,5,5,3,5,4,5,4,5,5,si
2,3,3,3,3,3,5,4,3,3,3,3,3,4,4,5,4,si
2,3,2,2,3,3,5,4,3,2,3,3,3,3,4,5,4,no
2,2,2,2,3,3,5,3,2,2,2,2,2,4,2,5,2,no
2,4,2,2,4,4,4,4,?,?,?, ?, ?, ?, ?,5,?,no
3,2,3,3,3,3,4,4,3,3,3,4,3,4,3,4,5,no
3,3,4,3,3,5,3,4,2,2,3,2,3,3,3,3,2,no
3,3,3,3,4,4,4,4,3,3,4,4,5,4,3,5,4,si
2,3,3,3,3,3,5,4,3,3,3,2,3,4,3,4,5,no
1,3,5,4,5,1,5,1,4,1,4,4,5,4,1,5,5,si
4,3,5,3,5,5,5,5,3,4,3,4,4,4,4,5,5,si
3,3,5,3,3,5,5,5,3,4,4,4,3,3,4,5,4,si
4,0,5,4,5,0,5,5,4,4,3,4,0,5,0,5,?,si
3,3,3,3,2,5,5,5,3,3,3,3,3,2,4,4,2,no
4,3,5,4,5,5,5,5,3,4,4,4,3,4,4,5,5,si
3,3,3,3,4,5,5,5,3,5,3,4,4,4,4,5,4,si
3,4,5,4,5,5,5,5,3,3,4,4,4,4,4,5,5,si
3,4,4,3,5,5,5,5,3,5,3,4,4,4,4,5,5,si
```

Anexo D1.3

Conjunto de prueba Modelo combinado 1er y 2do año.

```
@relation 'Modelo 1ro y 2do_año'

@attribute 'Mat I' numeric
@attribute FyS numeric
@attribute FInf numeric
@attribute IP numeric
@attribute 'EF I ' numeric
@attribute 'Id0 I' numeric
@attribute 'Alg y GA' numeric
@attribute Hluba numeric
@attribute 'Id0 II' numeric
@attribute DPOO numeric
@attribute 'EF II' numeric
@attribute IGS numeric
@attribute MD numeric
@attribute 'Mat II' numeric
@attribute EPlap numeric
@attribute 'mat iii' numeric
@attribute eplap numeric
@attribute edl numeric
@attribute iia numeric
@attribute 'id0 3' numeric
@attribute segNa1 numeric
@attribute 'EF 3' numeric
@attribute 'ele1 1' numeric
@attribute 'mat iv' numeric
@attribute Alomp numeric
@attribute 'ED 2' numeric
@attribute BD numeric
@attribute TSP numeric
@attribute 'id0 4' numeric
@attribute Dna1 numeric
@attribute 'EF iv' numeric
@attribute Pprof numeric
@attribute aprueba {no,si}

@data
3,3,4,3,5,5,3,4,3,3,5,4,3,3,4,2,3,2,2,3,2,4,3,3,3,2,2,3,3,2,4,2,no
2,3,4,3,5,5,3,4,3,4,5,4,3,2,5,2,3,2,2,3,3,5,4,3,2,2,3,4,4,3,5,2,no
3,3,3,3,5,5,3,4,3,3,5,4,3,2,4,2,2,2,2,4,3,5,3,3,2,2,2,3,4,3,5,2,no
4,4,4,3,5,5,4,5,4,5,5,5,3,4,5,4,3,4,3,4,4,4,5,5,3,5,4,5,4,5,5,5,si
3,3,3,3,5,5,3,4,3,3,5,4,3,3,5,2,3,3,3,3,3,5,4,3,3,3,3,3,4,4,5,4,si
3,3,3,3,5,5,3,3,3,4,5,4,3,3,5,2,3,2,2,3,3,5,4,3,2,3,3,3,3,4,5,4,no
2,2,2,2,?, ?,2,?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?,no
1,3,5,1,1,5,1,5,5,3,1,4,1,3,4,4,4,3,3,5,5,5,5,5,4,3,5,5,5,5,5,si
3,4,4,3,5,5,4,4,5,4,5,5,4,3,4,4,4,3,4,5,4,5,4,4,3,3,5,4,5,5,5,5,si
2,1,1,1,1,1,1,1,1,1,1,1,1,?,1,1,?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?,no
3,4,5,3,5,4,3,5,5,5,5,5,3,3,5,?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?,no
```

3,3,4,4,5,4,3,4,4,4,5,4,2,3,4,?,?,?,?,?,?,?,?,?,?,?,?,?,no
3,3,4,3,5,3,3,3,3,3,5,4,3,3,4,?,?,?,?,?,?,?,?,?,?,?,?,?,no
2,2,5,3,5,3,3,3,2,4,5,4,2,2,5,?,?,?,?,?,?,?,?,?,?,?,?,?,no
2,4,5,3,5,?,3,5,4,4,5,4,3,3,4,3,4,3,3,5,3,5,4,4,3,3,4,4,4,5,5,5,si
4,4,3,5,5,4,4,4,5,5,5,5,4,5,4,5,5,5,4,5,3,5,4,5,5,5,5,5,5,5,si
3,4,5,5,3,5,3,4,4,4,5,3,3,4,3,2,2,4,2,5,5,5,5,2,3,3,2,4,4,4,5,4,no
3,4,5,3,5,3,3,5,3,3,5,5,3,3,4,3,3,2,3,5,5,5,5,3,3,3,3,4,3,4,5,2,no
3,0,4,3,5,5,3,5,5,4,5,5,3,3,0,3,0,2,3,5,0,5,5,3,3,2,2,0,5,0,5,5,no
3,5,5,3,5,4,3,5,4,4,5,5,3,3,4,3,4,3,3,5,5,5,5,4,4,3,4,4,4,4,5,4,si
3,4,4,3,5,3,3,3,4,3,5,3,3,3,4,3,3,2,3,3,5,5,5,3,3,3,3,4,3,4,5,4,no
3,5,5,5,5,5,3,4,5,5,5,5,3,3,3,3,3,5,4,5,5,5,5,3,4,4,4,3,5,4,5,4,si
4,0,5,3,5,0,4,5,0,4,5,5,5,3,0,5,0,3,3,5,0,5,5,3,3,4,3,0,4,0,5,5,si
3,4,4,3,5,3,3,4,3,3,5,4,3,3,3,3,3,4,3,4,1,5,4,3,3,3,4,3,3,1,5,4,si

Anexo D2.1.

Conjunto de prueba Modelo 1er año con selección de rasgos.

Relation:CP-Modelo_1er año selrasg.csv

Mat I,FyS,FInf,IP,Alg y GA,Id0 II,DPOO,IGS,MD,Mat II,aprueba
5,5,4,5,4,4,5,4,4,4,si
3,4,4,4,3,4,5,4,3,3,si
4,3,4,2,3,3,3,3,3,3,no
3,4,4,3,4,3,5,3,3,3,si
3,4,4,2,3,3,3,3,3,3,no
3,4,4,3,3,3,3,3,3,3,no
3,5,5,5,4,3,4,4,3,4,si
3,4,4,2,3,3,2,3,3,2,no
2,3,4,2,3,3,2,3,3,2,no
3,5,5,3,3,3,4,4,3,5,si
3,5,5,3,4,4,3,3,3,4,si
3,4,4,3,3,3,3,4,3,4,no
4,4,5,5,3,5,5,4,4,3,si

Anexo D2.2

Modelo 2do año con selección de rasgos.

Relation: Modelo 2do año selrasg.csv

'matiii',ed1,iaa,'mativ',A1omp,'ED2',BD,Pprof,aprueba2
2,2,2,3,3,2,2,2
2,2,2,3,2,2,3,2
2,2,2,3,2,2,2,2
3,3,3,3,3,4,4,4
2,3,3,3,3,3,2,5
2,2,2,2,2,2,2,2
3,2,3,3,3,2,4,4
2,3,3,3,3,3,3,5
3,3,3,3,3,5,4,5
2,3,2,?,?,?,?,?
3,4,3,3,3,4,4,5
2,3,3,3,3,3,3,4
3,3,3,3,3,3,4,5
3,4,3,3,3,4,3,4
3,3,3,3,3,3,4,4
3,3,3,3,3,3,4,5
3,3,3,4,4,3,4,3
3,3,4,4,3,3,5,5
5,3,4,5,4,4,5,5
3,3,3,4,4,4,4,5
3,3,3,4,3,3,4,4

Anexo D2.3.-Modelo combinado, con selección de rasgos.

```
@relation 'Modelo 1ro y 2do año sel_rasg'
```

```
@attribute 'Mat I' numeric
@attribute FyS numeric
@attribute IP numeric
@attribute DPOO numeric
@attribute MD numeric
@attribute 'Mat II' numeric
@attribute 'mat iii' numeric
@attribute ed1 numeric
@attribute iia numeric
@attribute 'mat iv' numeric
@attribute Alomp numeric
@attribute 'ED 2' numeric
@attribute BD numeric
@attribute aprueba {no,si}
```

```
@data
3,3,3,3,3,3,2,2,2,3,3,2,2,no
2,3,3,4,3,2,2,2,2,3,2,2,3,no
3,3,3,3,3,2,2,2,2,3,2,2,2,no
4,4,3,5,3,4,4,4,3,5,3,5,4,si
3,3,3,3,3,3,2,3,3,3,3,3,3,si
3,3,3,4,3,3,2,2,2,3,2,3,3,no
3,3,3,3,3,3,2,2,2,2,2,2,2,no
3,3,3,3,3,3,2,2,2,?,?,?,no
3,3,3,3,4,3,3,3,3,3,3,3,4,no
3,3,3,3,4,3,3,4,3,2,2,3,2,no
3,4,3,3,3,3,3,3,3,3,3,4,4,si
3,3,3,3,3,3,2,3,3,3,3,3,2,no
3,3,4,3,4,3,3,4,3,3,4,3,3,no
4,4,3,3,3,4,3,2,3,3,3,2,3,no
?,1,1,1,1,1,2,2,2,2,2,2,2,no
3,3,3,3,3,2,2,2,2,?,?,?,no
2,0,3,3,3,3,2,2,2,2,2,2,2,no
4,3,3,3,3,3,3,2,3,3,3,2,4,no
1,1,1,1,3,3,3,3,3,3,3,3,2,si
3,3,3,3,4,3,3,3,3,3,3,3,3,si
4,3,3,3,4,4,3,5,3,4,5,4,3,si
3,4,3,4,3,2,2,3,3,3,3,3,3,no
3,4,3,4,3,3,3,3,3,3,3,5,4,si
2,3,3,4,3,2,2,3,2,?,?,?,no
```

Anexo E.

Resultados de la prueba de los Modelos entrenados por la red neuronal.

Anexo E1 Modelo 1er año- CP- V1

=== Model information ===

Filename: v1.model
Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation: Modelo_1er año
Attributes: 16
Mat I
FyS
FInf
IP
EF I
Id0 I
Alg y GA
H1uba
Id0 II
DPOO
EF II
IGS
MD
Mat II
EP1ap
aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP-1er_año
Instances: 14
Attributes: 16

=== Summary ===

Correctly Classified Instances	13	92.8571 %
Incorrectly Classified Instances	1	7.1429 %
Kappa statistic	0.8511	
Mean absolute error	0.0809	
Root mean squared error	0.2681	
Coverage of cases (0.95 level)	92.8571 %	
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.167	0.889	1	0.941	0.861	0.958	0.97	si
0.833	0	1	0.833	0.909	0.861	0.958	0.958	no
Weighted Avg.								
0.929	0.095	0.937	0.929	0.927	0.861	0.958	0.965	

=== Confusion Matrix ===

a b <-- classified as
8 0 | a = si
1 5 | b = no

Anexo E2 Modelo 1er año- CP- V2

=== Model information ===

```
Filename:      v2.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation:      Modelo_1er año
Attributes:    16
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DPOO
               EF II
               IGS
               MD
               Mat II
               EP1ap
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:      CP-1er_año
Instances:     14
Attributes:    16
```

=== Summary ===

Correctly Classified Instances	12	85.7143 %
Incorrectly Classified Instances	2	14.2857 %
Kappa statistic	0.6957	
Mean absolute error	0.1475	
Root mean squared error	0.3782	
Coverage of cases (0.95 level)	85.7143 %	
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
1	0.333	0.8	1	0.889	0.73	0.875	0.916	si
0.667	0	1	0.667	0.8	0.73	0.875	0.882	no
Weighted Avg.								
0.857	0.19	0.886	0.857	0.851	0.73	0.875	0.901	

=== Confusion Matrix ===

```
a b  <-- classified as
8 0 | a = si
2 4 | b = no
```

Anexo E3 Modelo 1er año- CP- V3

=== Model information ===

```
Filename:      v3.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H o
Relation:      Modelo_1er año
Attributes:    16
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DP00
               EF II
               IGS
               MD
               Mat II
               EP1ap
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:      CP-1er_año
Instances:     14
Attributes:    16
```

=== Summary ===

Correctly Classified Instances	13	92.8571 %
Incorrectly Classified Instances	1	7.1429 %
Kappa statistic	0.8511	
Mean absolute error	0.093	
Root mean squared error	0.2633	
Coverage of cases (0.95 level)	92.8571 %	
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
1	0.167	0.889	1	0.941	0.861	0.938	0.953	si
0.833	0	1	0.833	0.909	0.861	0.938	0.944	no
Weighted Avg.								
0.929	0.095	0.937	0.929	0.927	0.861	0.938	0.949	

=== Confusion Matrix ===

```
a b  <-- classified as
8 0 | a = si
1 5 | b = no
```

Anexo E4 Modelo 1er año- CP- V4

=== Model information ===

```
Filename:      v4.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:      Modelo_1er año
Attributes:    16
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DPOO
               EF II
               IGS
               MD
               Mat II
               EP1ap
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP-1er_año
Instances:     14
Attributes:    16
```

=== Summary ===

Correctly Classified Instances	13	92.8571 %
Incorrectly Classified Instances	1	7.1429 %
Kappa statistic	0.8511	
Mean absolute error	0.0809	
Root mean squared error	0.2681	
Coverage of cases (0.95 level)	92.8571 %	
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
1	0.167	0.889	1	0.941	0.861	0.958	0.97	si
0.833	0	1	0.833	0.909	0.861	0.958	0.958	no
Weighted Avg.								
0.929	0.095	0.937	0.929	0.927	0.861	0.958	0.965	

=== Confusion Matrix ===

```
a b  <-- classified as
8 0 | a = si
1 5 | b = no
```

Anexo E5 Modelo 1er año- CP- V5

== Model information ==

```
Filename:      v5.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation:      Modelo_1er año
Attributes:     16
                Mat I
                FyS
                FInf
                IP
                EF I
                Id0 I
                Alg y GA
                H1uba
                Id0 II
                DPOO
                EF II
                IGS
                MD
                Mat II
                EP1ap
                aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

Relation: CP-1er_año

Instances: 14

Attributes: 16

=== Summary ===

Correctly Classified Instances	13	92.8571 %
Incorrectly Classified Instances	1	7.1429 %
Kappa statistic	0.8511	
Mean absolute error	0.0782	
Root mean squared error	0.2677	
Coverage of cases (0.95 level)	92.8571 %	
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
1	0.167	0.889	1	0.941	0.861	0.938	0.953	si
0.833	0	1	0.833	0.909	0.861	0.938	0.944	no
Weighted Avg.								
0.929	0.095	0.937	0.929	0.927	0.861	0.938	0.949	

=== Confusion Matrix ===

```
a b  <-- classified as
8 0 | a = si
1 5 | b = no
```

Anexo E6 Modelo 1er año- CP- V6

=== Model information ===

```
Filename:      v6.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation:      Modelo_1er año
Attributes:    16
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DPOO
               EF II
               IGS
               MD
               Mat II
               EP1ap
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP-1er_año
Instances:     14
Attributes:    16
```

=== Summary ===

Correctly Classified Instances	13	92.8571 %
Incorrectly Classified Instances	1	7.1429 %
Kappa statistic	0.8511	
Mean absolute error	0.0759	
Root mean squared error	0.2675	
Coverage of cases (0.95 level)	92.8571 %	
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.167	0.889	1	0.941	0.861	0.938	0.953	si
0.833	0	1	0.833	0.909	0.861	0.938	0.944	no
Weighted Avg.								
0.929	0.095	0.937	0.929	0.927	0.861	0.938	0.949	

=== Confusion Matrix ===

```
a b  <-- classified as
8 0 | a = si
1 5 | b = no
```

Anexo E7 Modelo 1er año- CP- V7

=== Model information ===

```
Filename:      v7.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "20, 23"
Relation:      Modelo_1er año
Attributes:     16
                Mat I
                FyS
                FInf
                IP
                EF I
                Id0 I
                Alg y GA
                H1uba
                Id0 II
                DPOO
                EF II
                IGS
                MD
                Mat II
                EP1ap
                aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP-1er_año
Instances:     14
Attributes:     16
```

=== Summary ===

Correctly Classified Instances	13	92.8571 %
Incorrectly Classified Instances	1	7.1429 %
Kappa statistic	0.8511	
Mean absolute error	0.0773	
Root mean squared error	0.2676	
Coverage of cases (0.95 level)	92.8571 %	
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROC Area	PRCArea	Class
1	0.167	0.889	1	0.941	0.861	0.938	0.953	si
0.833	0	1	0.833	0.909	0.861	0.938	0.944	no
Weighted Avg.								
0.929	0.095	0.937	0.929	0.927	0.861	0.938	0.949	

=== Confusion Matrix ===

```
a b  <-- classified as
8 0 | a = si
1 5 | b = no
```

Anexo E8 Modelo 1er año selrasg- CP- V1

=== Model information ===

```
Filename:      v1.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:      Modelo_1er año selrasg
Attributes:    11
               Mat I
               FyS
               FInf
               IP
               Alg y GA
               Id0 II
               DPOO
               IGS
               MD
               Mat II
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:      CP-1er_año selrasg
Instances:     13
Attributes:    11
```

=== Summary ===

Correctly Classified Instances	12	92.3077 %
Incorrectly Classified Instances	1	7.6923 %
Kappa statistic	0.8434	
Mean absolute error	0.0936	
Root mean squared error	0.2813	
Coverage of cases (0.95 level)	92.3077 %	
Total Number of Instances	13	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.167	0.875	1	0.933	0.854	1	1	si
0.833	0	1	0.833	0.909	0.854	1	1	no
Weighted Avg.								
0.923	0.09	0.933	0.923	0.922	0.854	1	1	

=== Confusion Matrix ===

```
a b  <-- classified as
7 0 | a = si
1 5 | b = no
```

Anexo E9 Modelo 1er año selrasg- CP- V2

=== Model information ===

```
Filename:      v2.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation:      Modelo_1er año selrasg
Attributes:    11
               Mat I
               FyS
               FInf
               IP
               Alg y GA
               Id0 II
               DPOO
               IGS
               MD
               Mat II
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:      CP-1er_año selrasg
Instances:     13
Attributes:    11
```

=== Summary ===

Correctly Classified Instances	11	84.6154 %
Incorrectly Classified Instances	2	15.3846 %
Kappa statistic	0.6829	
Mean absolute error	0.1636	
Root mean squared error	0.3938	
Coverage of cases (0.95 level)	84.6154 %	
Total Number of Instances	13	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.333	0.778	1	0.875	0.72	0.762	0.787	si
0.667	0	1	0.667	0.8	0.72	0.762	0.841	no
Weighted Avg.								
0.846	0.179	0.88	0.846	0.84	0.72	0.762	0.812	

=== Confusion Matrix ===

```
a b  <-- classified as
7 0 | a = si
2 4 | b = no
```

Anexo E10 Modelo 1er año selrasg- CP- V3

=== Model information ===

Filename: v3.model
Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H o
Relation: Modelo_1er año selrasg
Attributes: 11
Mat I
FyS
FInf
IP
Alg y GA
Id0 II
DPOO
IGS
MD
Mat II
aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP-1er_año selrasg
Instances: 13
Attributes: 11

=== Summary ===

Correctly Classified Instances	10	76.9231 %
Incorrectly Classified Instances	3	23.0769 %
Kappa statistic	0.5185	
Mean absolute error	0.2447	
Root mean squared error	0.4767	
Coverage of cases (0.95 level)	76.9231 %	
Total Number of Instances	13	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.5	0.7	1	0.824	0.592	0.952	0.968	si
0.5	0	1	0.5	0.667	0.592	0.952	0.948	no
Weighted Avg.								
0.769	0.269	0.838	0.769	0.751	0.592	0.952	0.959	

=== Confusion Matrix ===

a b <-- classified as
7 0 | a = si
3 3 | b = no

Anexo E11 Modelo 1er año selrasg- CP- V4

=== Model information ===

```
Filename:      v4.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:      Modelo_1er año selrasg
Attributes:     11
                Mat I
                FyS
                FInf
                IP
                Alg y GA
                Id0 II
                DPOO
                IGS
                MD
                Mat II
                aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP-1er_año selrasg
Instances:     13
Attributes:     11
```

=== Summary ===

Correctly Classified Instances	12	92.3077 %
Incorrectly Classified Instances	1	7.6923 %
Kappa statistic	0.8434	
Mean absolute error	0.0936	
Root mean squared error	0.2813	
Coverage of cases (0.95 level)	92.3077 %	
Total Number of Instances	13	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.167	0.875	1	0.933	0.854	1	1	si
0.833	0	1	0.833	0.909	0.854	1	1	no
Weighted Avg.								
0.923	0.09	0.933	0.923	0.922	0.854	1	1	

=== Confusion Matrix ===

```
a b  <-- classified as
7 0 | a = si
1 5 | b = no
```

Anexo E12 Modelo 1er año selrasg- CP- V5

=== Model information ===

```
Filename:      v5.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation:      Modelo_1er año selrasg
Attributes:    11
               Mat I
               FyS
               FInf
               IP
               Alg y GA
               Id0 II
               DPOO
               IGS
               MD
               Mat II
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

Relation: CP-1er_año selrasg

Instances: 13

Attributes: 11

=== Summary ===

Correctly Classified Instances	10	76.9231 %
Incorrectly Classified Instances	3	23.0769 %
Kappa statistic	0.5185	
Mean absolute error	0.2344	
Root mean squared error	0.4805	
Coverage of cases (0.95 level)	76.9231 %	
Total Number of Instances	13	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.5	0.7	1	0.824	0.592	0.833	0.881	si
0.5	0	1	0.5	0.667	0.592	0.833	0.849	no
Weighted Avg.								
0.769	0.269	0.838	0.769	0.751	0.592	0.833	0.866	

=== Confusion Matrix ===

```
a b  <-- classified as
7 0 | a = si
3 3 | b = no
```

Anexo E13 Modelo 1er año selrasg- CP- V6

=== Model information ===

```
Filename:      v6.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
               -N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation:      Modelo_1er año selrasg
Attributes:    11
               Mat I
               FyS
               FInf
               IP
               Alg y GA
               Id0 II
               DPOO
               IGS
               MD
               Mat II
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP-1er_año selrasg
Instances:     13
Attributes:    11
```

=== Summary ===

Correctly Classified Instances	10	76.9231 %
Incorrectly Classified Instances	3	23.0769 %
Kappa statistic	0.5185	
Mean absolute error	0.2325	
Root mean squared error	0.4804	
Coverage of cases (0.95 level)	76.9231 %	
Total Number of Instances	13	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.5	0.7	1	0.824	0.592	0.81	0.869	si
0.5	0	1	0.5	0.667	0.592	0.81	0.826	no
Weighted Avg.								
0.769	0.269	0.838	0.769	0.751	0.592	0.81	0.849	

=== Confusion Matrix ===

```
a b  <-- classified as
7 0 | a = si
3 3 | b = no
```

Anexo E14 Modelo 1er año selrasg- CP- V7

=== Model information ===

```
Filename:      v7.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "20, 23"
Relation:      Modelo_1er año selrasg
Attributes:    11
               Mat I
               FyS
               FInf
               IP
               Alg y GA
               Id0 II
               DPOO
               IGS
               MD
               Mat II
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

Relation: CP-1er_año selrasg

Instances: 13

Attributes: 11

=== Summary ===

Correctly Classified Instances	10	76.9231 %
Incorrectly Classified Instances	3	23.0769 %
Kappa statistic	0.5185	
Mean absolute error	0.233	
Root mean squared error	0.4804	
Coverage of cases (0.95 level)	76.9231 %	
Total Number of Instances	13	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.5	0.7	1	0.824	0.592	0.833	0.881	si
0.5	0	1	0.5	0.667	0.592	0.833	0.849	no
Weighted Avg.								
0.769	0.269	0.838	0.769	0.751	0.592	0.833	0.866	

=== Confusion Matrix ===

```
a b  <-- classified as
7 0 | a = si
3 3 | b = no
```

Anexo E15 Modelo 2do año- CP- V1

=== Model information ===

Filename: v1.model
Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation: Modelo 2do año
Attributes: 18
Mat III
Epcap
ED I
IIA
IdEx III
SegNac
EF III
elec
Mat IV
Arqcomp
ED II
BD
TSP
IdEx IV
DeNac
EF IV
Pprof
Aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

Relation: CP- 2do año

Instances: 15

Attributes: 18

=== Summary ===

Correctly Classified Instances	15	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0066		
Root mean squared error	0.0147		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	ROCArea	Class
1	0	1	1	1	1	no
1	0	1	1	1	1	si
Weighted Avg.						
1	0	1	1	1	1	

=== Confusion Matrix ===

```
a b <-- classified as
11 0 | a = no
0 4 | b = si
```

Anexo E16 Modelo 2do año- CP- V2

=== Model information ===

Filename: v2.model
Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation: Modelo 2do año
Attributes: 18
Mat III
Epcap
ED I
IIA
IdEx III
SegNac
EF III
elec
Mat IV
Arqcomp
ED II
BD
TSP
IdEx IV
DeNac
EF IV
Pprof
Aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP- 2do año
Instances: 15
Attributes: 18

=== Summary ===

Correctly Classified Instances	15	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0028		
Root mean squared error	0.0108		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	no
1	0	1	1	1	1	si
Weighted Avg.						
1	0	1	1	1	1	

=== Confusion Matrix ===

a b <-- classified as
11 0 | a = no
0 4 | b = si

Anexo E17 Modelo 2do año- CP- V3

=== Model information ===

```

Filename:      v3.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H 0
Relation:      Modelo 2do año
Attributes:     18
                Mat III
                Epcap
                ED I
                IIA
                IdEx III
                SegNac
                EF III
                elec
                Mat IV
                Arqcomp
                ED II
                BD
                TSP
                IdEx IV
                DeNac
                EF IV
                Pprof
                Aprueba
    
```

=== Classifier model ===

=== Re-evaluation on test set ===

```

User supplied test set
Relation:      CP- 2do año
Instances:     15
Attributes:     18
    
```

=== Summary ===

Correctly Classified Instances	15	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0049		
Root mean squared error	0.0152		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	no
1	0	1	1	1	1	si
Weighted Avg.						
1	0	1	1	1	1	

=== Confusion Matrix ===

```

a  b  <-- classified as
11  0 | a = no
 0  4 | b = si
    
```

Anexo E18 Modelo 2do año- CP- V4

=== Model information ===

Filename: v4.model
Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation: Modelo 2do año
Attributes: 18
Mat III
Epcap
ED I
IIA
IdEx III
SegNac
EF III
elec
Mat IV
Arqcomp
ED II
BD
TSP
IdEx IV
DeNac
EF IV
Pprof
Aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP- 2do año
Instances: 15
Attributes: 18

=== Summary ===

Correctly Classified Instances	15	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0066		
Root mean squared error	0.0147		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	no
1	0	1	1	1	1	si
Weighted Avg.						
1	0	1	1	1	1	

=== Confusion Matrix ===

a	b	<-- classified as
11	0	a = no
0	4	b = si

Anexo E19 Modelo 2do año- CP- V5

=== Model information ===

Filename: v5.model
Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation: Modelo 2do año
Attributes: 18
Mat III
Epcap
ED I
IIA
IdEx III
SegNac
EF III
elec
Mat IV
Arqcomp
ED II
BD
TSP
IdEx IV
DeNac
EF IV
Pprof
Aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

Relation: CP- 2do año
Instances: 15
Attributes: 18

=== Summary ===

Correctly Classified Instances	15	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0035		
Root mean squared error	0.0085		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	no
1	0	1	1	1	1	si
Weighted Avg.						
1	0	1	1	1	1	

=== Confusion Matrix ===

a b <-- classified as
11 0 | a = no
0 4 | b = si

Anexo E20 Modelo 2do año- CP- V6

=== Model information ===

Filename: v6.model
Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation: Modelo 2do año
Attributes: 18
Mat III
Epcap
ED I
IIA
IdEx III
SegNac
EF III
elec
Mat IV
Arqcomp
ED II
BD
TSP
IdEx IV
DeNac
EF IV
Pprof
Aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP- 2do año
Instances: 15
Attributes: 18

=== Summary ===

Correctly Classified Instances	15	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0034		
Root mean squared error	0.0089		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	no
1	0	1	1	1	1	si
Weighted Avg.						
1	0	1	1	1	1	

=== Confusion Matrix ===

```
a b <-- classified as
11 0 | a = no
0 4 | b = si
```

Anexo E21 Modelo 2do año- CP- V7

=== Model information ===

Filename: v7.model
Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "20, 23"
Relation: Modelo 2do año
Attributes: 18
Mat III
Epcap
ED I
IIA
IdEx III
SegNac
EF III
elec
Mat IV
Arqcomp
ED II
BD
TSP
IdEx IV
DeNac
EF IV
Pprof
Aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

Relation: CP- 2do año
Instances: 15
Attributes: 18

=== Summary ===

Correctly Classified Instances	15	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0026		
Root mean squared error	0.0071		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	no
1	0	1	1	1	1	si
Weighted Avg.						
1	0	1	1	1	1	

=== Confusion Matrix ===

a b <-- classified as
11 0 | a = no
0 4 | b = si

Anexo E22 Modelo 2do año selrasg- CP- V1

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:     Modelo 2do año selrasg
Instances:    120
Attributes:    9
               Mat III
               ED I
               IIA
               Mat IV
               Arqcomp
               ED II
               BD
               Pprof
               Aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:     CP- 2do año selrasg.arff
Instances:    15
Attributes:    9
```

=== Summary ===

Correctly Classified Instances	13	86.6667 %
Incorrectly Classified Instances	2	13.3333 %
Kappa statistic	0.5833	
Mean absolute error	0.1442	
Root mean squared error	0.3139	
Coverage of cases (0.95 level)	100	%
Total Number of Instances	15	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.917	0.333	0.917	0.917	0.917	0.944	no
0.667	0.083	0.667	0.667	0.667	0.944	si
Weighted Avg.						
0.867	0.283	0.867	0.867	0.867	0.944	

=== Confusion Matrix ===

```
a  b  <-- classified as
11  1 | a = no
 1  2 | b = si
```

Anexo E23 Modelo 2do año selrasg- CP- V2

=== Run information ===

```
Scheme:      weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation:     Modelo 2do año selrasg
Instances:    120
Attributes:    9
               Mat III
               ED I
               IIA
               Mat IV
               Arqcomp
               ED II
               BD
               Pprof
               Aprueba
Test mode:    3-fold cross-validation
```

=== Classifier model (full training set) ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:     CP- 2do año selrasg.arff
Instances:    15
Attributes:    9
```

=== Summary ===

Correctly Classified Instances	12	80	%
Incorrectly Classified Instances	3	20	%
Kappa statistic	0.4444		
Mean absolute error	0.1514		
Root mean squared error	0.3361		
Coverage of cases (0.95 level)	93.3333 %		
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.833	0.333	0.909	0.833	0.87	0.917	no
0.667	0.167	0.5	0.667	0.571	0.917	si
Weighted Avg.						
0.8	0.3	0.827	0.8	0.81	0.917	

=== Confusion Matrix ===

```
a  b  <-- classified as
10  2 | a = no
 1  2 | b = si
```

Anexo E24 Modelo 2do año selrasg- CP- V3

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H 0
Relation: Modelo 2do año selrasg
Instances: 120
Attributes: 9
Mat III
ED I
IIA
Mat IV
Arqcomp
ED II
BD
Pprof
Aprueba
Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

=== Re-evaluation on test set ===

User supplied test set

Relation: CP- 2do año selrasg.arff
Instances: 15
Attributes: 9

=== Summary ===

Correctly Classified Instances	13	86.6667 %
Incorrectly Classified Instances	2	13.3333 %
Kappa statistic	0.5833	
Mean absolute error	0.1477	
Root mean squared error	0.3521	
Coverage of cases (0.95 level)	93.3333 %	
Total Number of Instances	15	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.917	0.333	0.917	0.917	0.917	0.917	no
0.667	0.083	0.667	0.667	0.667	0.917	si
Weighted Avg.						
0.867	0.283	0.867	0.867	0.867	0.917	

=== Confusion Matrix ===

```
a  b  <-- classified as
11  1 | a = no
 1  2 | b = si
```

Anexo E25 Modelo 2do año selrasg- CP- V4

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation: Modelo 2do año selrasg
Instances: 120
Attributes: 9
Mat III
ED I
IIA
Mat IV
Arqcomp
ED II
BD
Pprof
Aprueba
Test mode: 5-fold cross-validation

=== Classifier model (full training set) ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP- 2do año selrasg.arff
Instances: 15
Attributes: 9

=== Summary ===

Correctly Classified Instances	13	86.6667 %
Incorrectly Classified Instances	2	13.3333 %
Kappa statistic	0.5833	
Mean absolute error	0.1442	
Root mean squared error	0.3139	
Coverage of cases (0.95 level)	100	%
Total Number of Instances	15	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.917	0.333	0.917	0.917	0.917	0.944	no
0.667	0.083	0.667	0.667	0.667	0.944	si
Weighted Avg.						
0.867	0.283	0.867	0.867	0.867	0.944	

=== Confusion Matrix ===

a	b	<-- classified as
11	1	a = no
1	2	b = si

Anexo E26 Modelo 2do año selrasg- CP- V5

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation: Modelo 2do año selrasg
Instances: 120
Attributes: 9
Mat III
ED I
IIA
Mat IV
Arqcomp
ED II
BD
Pprof
Aprueba
Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

=== Re-evaluation on test set ===

User supplied test set

Relation: CP- 2do año selrasg.arff
Instances: 15
Attributes: 9

=== Summary ===

Correctly Classified Instances	12	80	%
Incorrectly Classified Instances	3	20	%
Kappa statistic	0.4444		
Mean absolute error	0.1675		
Root mean squared error	0.3629		
Coverage of cases (0.95 level)	93.3333 %		
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.833	0.333	0.909	0.833	0.87	0.917	no
0.667	0.167	0.5	0.667	0.571	0.917	si
Weighted Avg.						
0.8	0.3	0.827	0.8	0.81	0.917	

=== Confusion Matrix ===

a	b	<-- classified as
10	2	a = no
1	2	b = si

Anexo E27 Modelo 2do año selrasg- CP- V6

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation: Modelo 2do año selrasg
Instances: 120
Attributes: 9
Mat III
ED I
IIA
Mat IV
Arqcomp
ED II
BD
Pprof
Aprueba
Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

=== Re-evaluation on test set ===

User supplied test set

Relation: CP- 2do año selrasg.arff
Instances: 15
Attributes: 9

=== Summary ===

Correctly Classified Instances	12	80	%
Incorrectly Classified Instances	3	20	%
Kappa statistic	0.4444		
Mean absolute error	0.1639		
Root mean squared error	0.3602		
Coverage of cases (0.95 level)	93.3333	%	
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.833	0.333	0.909	0.833	0.87	0.917	no
0.667	0.167	0.5	0.667	0.571	0.917	si
Weighted Avg.						
0.8	0.3	0.827	0.8	0.81	0.917	

=== Confusion Matrix ===

a	b	<-- classified as
10	2	a = no
1	2	b = si

Anexo E28 Modelo 2do año selrasg- CP- V7

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "20, 23"
Relation: Modelo 2do año selrasg
Instances: 120
Attributes: 9
Mat III
ED I
IIA
Mat IV
Arqcomp
ED II
BD
Pprof
Aprueba
Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

=== Re-evaluation on test set ===

User supplied test set

Relation: CP- 2do año selrasg.arff
Instances: 15
Attributes: 9

=== Summary ===

Correctly Classified Instances	13	86.6667 %
Incorrectly Classified Instances	2	13.3333 %
Kappa statistic	0.5833	
Mean absolute error	0.1126	
Root mean squared error	0.2742	
Coverage of cases (0.95 level)	100	%
Total Number of Instances	15	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.917	0.333	0.917	0.917	0.917	0.972	no
0.667	0.083	0.667	0.667	0.667	0.972	si
Weighted Avg.						
0.867	0.283	0.867	0.867	0.867	0.972	

=== Confusion Matrix ===

a	b	<-- classified as
11	1	a = no
1	2	b = si

Anexo E29 Modelo 1er y 2do año- CP- V1

=== Model information ===

```
Filename:      v1.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:      Modelo 1ro y 2do_año
Attributes:    33
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DP00
               EF II
               IGS
               MD
               Mat II
               EP1ap
               mat iii
               ep1ap
               ed1
               iia
               id0 3
               segNa1
               EF 3
               ele1 1
               mat iv
               A1omp
               ED 2
               BD
               TSP
               id0 4
               Dna1
               EF iv
               Pprof
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP- 1ro y 2do_año
Instances:     24
Attributes:    33
```

=== Summary ===

Correctly Classified Instances	24	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0062		
Root mean squared error	0.0115		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0	1	1	1	1	1	1	no
1	0	1	1	1	1	1	1	si
Weighted Avg.								
1	0	1	1	1	1	1	1	

=== Confusion Matrix ===

```
a  b  <-- classified as
15  0 | a = no
  0  9 | b = si
```

Anexo E30 Modelo 1er y 2do año- CP- V2

=== Model information ===

```
Filename:      v2.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation:      Modelo 1ro y 2do_año
Attributes:    33
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DP00
               EF II
               IGS
               MD
               Mat II
               EP1ap
               mat iii
               ep1ap
               ed1
               iia
               id0 3
               segNa1
               EF 3
               ele1 1
               mat iv
               A1omp
               ED 2
               BD
               TSP
               id0 4
               Dna1
               EF iv
               Pprof
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP- 1ro y 2do_año
Instances:     24
Attributes:    33
```

=== Summary ===

Correctly Classified Instances	24	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0095		
Root mean squared error	0.0278		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0	1	1	1	1	1	1	no
1	0	1	1	1	1	1	1	si
Weighted Avg.								
1	0	1	1	1	1	1	1	

=== Confusion Matrix ===

```
a  b  <-- classified as
15  0 | a = no
  0  9 | b = si
```

Anexo E31 Modelo 1er y 2do año- CP- V3

=== Model information ===

```
Filename:      v3.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H o
Relation:      Modelo 1ro y 2do_año
Attributes:    33
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DP00
               EF II
               IGS
               MD
               Mat II
               EP1ap
               mat iii
               ep1ap
               ed1
               iia
               id0 3
               segNa1
               EF 3
               ele1 1
               mat iv
               A1omp
               ED 2
               BD
               TSP
               id0 4
               Dna1
               EF iv
               Pprof
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP- 1ro y 2do_año
Instances:     24
Attributes:    33
```

=== Summary ===

Correctly Classified Instances	24	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0105		
Root mean squared error	0.0179		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0	1	1	1	1	1	1	no
1	0	1	1	1	1	1	1	si
Weighted Avg.								
1	0	1	1	1	1	1	1	

=== Confusion Matrix ===

```
a  b  <-- classified as
15  0 | a = no
  0  9 | b = si
```

Anexo E32 Modelo 1er y 2do año- CP- V4

=== Model information ===

```
Filename:      v4.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:      Modelo 1ro y 2do_año
Attributes:    33
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DP00
               EF II
               IGS
               MD
               Mat II
               EP1ap
               mat iii
               ep1ap
               ed1
               iia
               id0 3
               segNa1
               EF 3
               ele1 1
               mat iv
               A1omp
               ED 2
               BD
               TSP
               id0 4
               Dna1
               EF iv
               Pprof
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP- 1ro y 2do_año
Instances:     24
Attributes:    33
```

=== Summary ===

Correctly Classified Instances	24	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0062		
Root mean squared error	0.0115		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0	1	1	1	1	1	1	no
1	0	1	1	1	1	1	1	si
Weighted Avg.								
1	0	1	1	1	1	1	1	

=== Confusion Matrix ===

```
a  b  <-- classified as
15  0 | a = no
  0  9 | b = si
```

Anexo E33 Modelo 1er y 2do año- CP- V5

=== Model information ===

```
Filename:      v5.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation:      Modelo 1ro y 2do_año
Attributes:    33
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DP00
               EF II
               IGS
               MD
               Mat II
               EP1ap
               mat iii
               ep1ap
               ed1
               iia
               id0 3
               segNa1
               EF 3
               ele1 1
               mat iv
               A1omp
               ED 2
               BD
               TSP
               id0 4
               Dna1
               EF iv
               Pprof
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP- 1ro y 2do_año
Instances:     24
Attributes:    33
```

=== Summary ===

Correctly Classified Instances	24	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0032		
Root mean squared error	0.0062		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0	1	1	1	1	1	1	no
1	0	1	1	1	1	1	1	si
Weighted Avg.								
1	0	1	1	1	1	1	1	

=== Confusion Matrix ===

```
a  b  <-- classified as
15  0 | a = no
  0  9 | b = si
```

Anexo E34 Modelo 1er y 2do año- CP- V6

=== Model information ===

```
Filename:      v6.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation:      Modelo 1ro y 2do_año
Attributes:    33
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DP00
               EF II
               IGS
               MD
               Mat II
               EP1ap
               mat iii
               ep1ap
               ed1
               iia
               id0 3
               segNa1
               EF 3
               ele1 1
               mat iv
               A1omp
               ED 2
               BD
               TSP
               id0 4
               Dna1
               EF iv
               Pprof
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP- 1ro y 2do_año
Instances:     24
Attributes:    33
```

=== Summary ===

Correctly Classified Instances	24	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0032		
Root mean squared error	0.0062		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0	1	1	1	1	1	1	no
1	0	1	1	1	1	1	1	si
Weighted Avg.								
1	0	1	1	1	1	1	1	

=== Confusion Matrix ===

```
a  b  <-- classified as
15  0 | a = no
  0  9 | b = si
```

Anexo E35 Modelo 1er y 2do año- CP- V7

=== Model information ===

```
Filename:      v7.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "20, 23"
Relation:      Modelo 1ro y 2do_año
Attributes:    33
               Mat I
               FyS
               FInf
               IP
               EF I
               Id0 I
               Alg y GA
               H1uba
               Id0 II
               DP00
               EF II
               IGS
               MD
               Mat II
               EP1ap
               mat iii
               ep1ap
               ed1
               iia
               id0 3
               segNa1
               EF 3
               ele1 1
               mat iv
               A1omp
               ED 2
               BD
               TSP
               id0 4
               Dna1
               EF iv
               Pprof
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP- 1ro y 2do_año
Instances:     24
Attributes:    33
```

=== Summary ===

Correctly Classified Instances	24	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0033		
Root mean squared error	0.0066		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0	1	1	1	1	1	1	no
1	0	1	1	1	1	1	1	si
Weighted Avg.								
1	0	1	1	1	1	1	1	

=== Confusion Matrix ===

```
a  b  <-- classified as
15  0 | a = no
  0  9 | b = si
```

Anexo E36 Modelo 1er y 2do año selrasg- CP- V1

=== Model information ===

```
Filename:      v1.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:      Modelo 1ro y 2do año sel_rasg
Attributes:    14
               Mat I
               FyS
               IP
               DP00
               MD
               Mat II
               mat iii
               ed1
               iia
               mat iv
               Alomp
               ED 2
               BD
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:      CP- 1er y 2do año selrasg
Instances:     24
Attributes:    14
```

=== Summary ===

Correctly Classified Instances	22	91.6667 %
Incorrectly Classified Instances	2	8.3333 %
Kappa statistic	0.8222	
Mean absolute error	0.0955	
Root mean squared error	0.2894	
Coverage of cases (0.95 level)	91.6667 %	
Total Number of Instances	24	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.933	0.111	0.933	0.933	0.933	0.822	0.922	0.946	no
0.889	0.067	0.889	0.889	0.889	0.822	0.922	0.906	si
Weighted Avg.								
0.917	0.094	0.917	0.917	0.917	0.822	0.922	0.931	

=== Confusion Matrix ===

```
  a  b  <-- classified as
14  1  |  a = no
 1  8  |  b = si
```

Anexo E37 Modelo 1er y 2do año selrasg- CP- V2

=== Model information ===

```
Filename:      v2.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H i
Relation:      Modelo 1ro y 2do año sel_rasg
Attributes:    14
               Mat I
               FyS
               IP
               DPOO
               MD
               Mat II
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:      CP- 1er y 2do año selrasg
Instances:     24
Attributes:    14
```

=== Summary ===

Correctly Classified Instances	23	95.8333 %
Incorrectly Classified Instances	1	4.1667 %
Kappa statistic	0.9091	
Mean absolute error	0.055	
Root mean squared error	0.1471	
Coverage of cases (0.95 level)	100 %	
Total Number of Instances	24	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.111	0.938	1	0.968	0.913	0.996	0.996	no
0.889	0	1	0.889	0.941	0.913	0.996	0.989	si
Weighted Avg.								
0.958	0.069	0.961	0.958	0.958	0.913	0.996	0.993	

=== Confusion Matrix ===

```
a  b  <-- classified as
15  0  |  a = no
 1  8  |  b = si
```

Anexo E38 Modelo 1er y 2do año selrasg- CP- V3

=== Model information ===

```
Filename:      v3.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H o
Relation:      Modelo 1ro y 2do año sel_rasg
Attributes:    14
               Mat I
               FyS
               IP
               DP00
               MD
               Mat II
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

```
User supplied test set
Relation:      CP- 1er y 2do año selrasg
Instances:     24
Attributes:    14
```

=== Summary ===

Correctly Classified Instances	21	87.5	%
Incorrectly Classified Instances	3	12.5	%
Kappa statistic	0.7391		
Mean absolute error	0.1414		
Root mean squared error	0.3324		
Coverage of cases (0.95 level)	95.8333	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.867	0.111	0.929	0.867	0.897	0.742	0.893	0.927	no
0.889	0.133	0.8	0.889	0.842	0.742	0.893	0.87	si
Weighted Avg.								
0.875	0.119	0.88	0.875	0.876	0.742	0.893	0.906	

=== Confusion Matrix ===

```
a  b  <-- classified as
13  2 | a = no
 1  8 | b = si
```

Anexo E39 Modelo 1er y 2do año selrasg- CP- V4

=== Model information ===

```
Filename:      v4.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "12, 10"
Relation:      Modelo 1ro y 2do año sel_rasg
Attributes:    14
               Mat I
               FyS
               IP
               DPOO
               MD
               Mat II
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:      CP- 1er y 2do año selrasg
Instances:     24
Attributes:    14
```

=== Summary ===

Correctly Classified Instances	22	91.6667 %
Incorrectly Classified Instances	2	8.3333 %
Kappa statistic	0.8222	
Mean absolute error	0.0955	
Root mean squared error	0.2894	
Coverage of cases (0.95 level)	91.6667 %	
Total Number of Instances	24	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.933	0.111	0.933	0.933	0.933	0.822	0.922	0.946	no
0.889	0.067	0.889	0.889	0.889	0.822	0.922	0.906	si
Weighted Avg.								
0.917	0.094	0.917	0.917	0.917	0.822	0.922	0.931	

=== Confusion Matrix ===

```
a  b  <-- classified as
14  1  |  a = no
 1  8  |  b = si
```

Anexo E40 Modelo 1er y 2do año selrasg- CP- V5

=== Model information ===

```
Filename:      v5.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "40, 20"
Relation:      Modelo 1ro y 2do año sel_rasg
Attributes:    14
               Mat I
               FyS
               IP
               DPOO
               MD
               Mat II
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:      CP- 1er y 2do año selrasg
Instances:     24
Attributes:    14
```

=== Summary ===

Correctly Classified Instances	21	87.5	%
Incorrectly Classified Instances	3	12.5	%
Kappa statistic	0.7391		
Mean absolute error	0.1268		
Root mean squared error	0.3535		
Coverage of cases (0.95 level)	87.5	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.867	0.111	0.929	0.867	0.897	0.742	0.9	0.944	no
0.889	0.133	0.8	0.889	0.842	0.742	0.9	0.849	si
Weighted Avg.								
0.875	0.119	0.88	0.875	0.876	0.742	0.9	0.909	

=== Confusion Matrix ===

```
a  b  <-- classified as
13  2 | a = no
 1  8 | b = si
```

Anexo E41 Modelo 1er y 2do año selrasg- CP- V6

=== Model information ===

```
Filename:      v6.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation:      Modelo 1ro y 2do año sel_rasg
Attributes:    14
               Mat I
               FyS
               IP
               DPOO
               MD
               Mat II
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:      CP- 1er y 2do año selrasg
Instances:     24
Attributes:    14
```

=== Summary ===

Correctly Classified Instances	21	87.5	%
Incorrectly Classified Instances	3	12.5	%
Kappa statistic	0.7391		
Mean absolute error	0.1279		
Root mean squared error	0.3538		
Coverage of cases (0.95 level)	87.5	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.867	0.111	0.929	0.867	0.897	0.742	0.893	0.942	no
0.889	0.133	0.8	0.889	0.842	0.742	0.893	0.827	si
Weighted Avg.								
0.875	0.119	0.88	0.875	0.876	0.742	0.893	0.899	

=== Confusion Matrix ===

```
a  b  <-- classified as
13  2 | a = no
 1  8 | b = si
```

Anexo E42 Modelo 1er y 2do año selrasg- CP- V7

=== Model information ===

```
Filename:      v6.model
Scheme:        weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.2
-N 500 -V 0 -S 0 -E 20 -H "50, 30"
Relation:      Modelo 1ro y 2do año sel_rasg
Attributes:    14
               Mat I
               FyS
               IP
               DPOO
               MD
               Mat II
               mat iii
               ed1
               iia
               mat iv
               A1omp
               ED 2
               BD
               aprueba
```

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set

```
Relation:      CP- 1er y 2do año selrasg
Instances:     24
Attributes:    14
```

=== Summary ===

Correctly Classified Instances	21	87.5	%
Incorrectly Classified Instances	3	12.5	%
Kappa statistic	0.7391		
Mean absolute error	0.1279		
Root mean squared error	0.3538		
Coverage of cases (0.95 level)	87.5	%	
Total Number of Instances	24		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.867	0.111	0.929	0.867	0.897	0.742	0.893	0.942	no
0.889	0.133	0.8	0.889	0.842	0.742	0.893	0.827	si
Weighted Avg.								
0.875	0.119	0.88	0.875	0.876	0.742	0.893	0.899	

=== Confusion Matrix ===

```
a  b  <-- classified as
13  2  |  a = no
 1  8  |  b = si
```

Anexo F.

Modelos resultantes del entrenamiento del árbol de decisión.

Anexo F1 Modelo 1er año- J48- cross-validation

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo_1er año
Instances: 114
Attributes: 16
Mat I
FyS
FInf
IP
EF I
Id0 I
Alg y GA
H1uba
Id0 II
DPOO
EF II
IGS
MD
Mat II
EP1ap
aprueba

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	106	92.9825 %
Incorrectly Classified Instances	8	7.0175 %
Kappa statistic	0.8324	
Mean absolute error	0.1062	
Root mean squared error	0.2573	
Relative absolute error	24.8662 %	
Root relative squared error	55.7446 %	
Coverage of cases (0.95 level)	95.614 %	
Mean rel. region size (0.95 level)	63.5965 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.962	0.143	0.938	0.962	0.95	0.906	si
0.857	0.038	0.909	0.857	0.882	0.906	no
Weighted Avg.						
0.93	0.111	0.929	0.93	0.929	0.906	

=== Confusion Matrix ===

```
a b <-- classified as
76 3 | a = si
5 30 | b = no
```

Anexo F2 Modelo 1er año- J48- transing set

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo_1er año
Instances: 114
Attributes: 16
Mat I
FyS
FInf
IP
EF I
Id0 I
Alg y GA
H1uba
Id0 II
DPOO
EF II
IGS
MD
Mat II
EP1ap
aprueba
Test mode: evaluate on training data

=== Classifier model (full training set) ===

Time taken to build model: 0 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	110	96.4912 %
Incorrectly Classified Instances	4	3.5088 %
Kappa statistic	0.9162	
Mean absolute error	0.0665	
Root mean squared error	0.1824	
Relative absolute error	15.5891 %	
Root relative squared error	39.5358 %	
Coverage of cases (0.95 level)	97.3684 %	
Mean rel. region size (0.95 level)	54.8246 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.987	0.086	0.963	0.987	0.975	0.957	si
0.914	0.013	0.97	0.914	0.941	0.957	no
Weighted Avg.						
0.965	0.063	0.965	0.965	0.965	0.957	

=== Confusion Matrix ===

a b <-- classified as
78 1 | a = si
3 32 | b = no

Anexo F3 Modelo 1er año selrasg- J48- cross-validation

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo_1er año selrasg
Instances: 114
Attributes: 11
Mat I
FyS
FInf
IP
Alg y GA
Id0 II
DPOO
IGS
MD
Mat II
aprueba
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	106	92.9825 %
Incorrectly Classified Instances	8	7.0175 %
Kappa statistic	0.8324	
Mean absolute error	0.1062	
Root mean squared error	0.2573	
Relative absolute error	24.8662 %	
Root relative squared error	55.7446 %	
Coverage of cases (0.95 level)	95.614 %	
Mean rel. region size (0.95 level)	63.5965 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.962	0.143	0.938	0.962	0.95	0.906	si
0.857	0.038	0.909	0.857	0.882	0.906	no
Weighted Avg.						
0.93	0.111	0.929	0.93	0.929	0.906	

=== Confusion Matrix ===

a	b	<-- classified as
76	3	a = si
5	30	b = no

Anexo F4 Modelo 1er año selrasg- J48- training set

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo_1er año selrasg
Instances: 114
Attributes: 11
Mat I
FyS
FInf
IP
Alg y GA
Id0 II
DPOO
IGS
MD
Mat II
aprueba
Test mode: evaluate on training data

=== Classifier model (full training set) ===

Time taken to build model: 0 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances	110	96.4912 %
Incorrectly Classified Instances	4	3.5088 %
Kappa statistic	0.9162	
Mean absolute error	0.0665	
Root mean squared error	0.1824	
Relative absolute error	15.5891 %	
Root relative squared error	39.5358 %	
Coverage of cases (0.95 level)	97.3684 %	
Mean rel. region size (0.95 level)	54.8246 %	
Total Number of Instances	114	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.987	0.086	0.963	0.987	0.975	0.957	si
0.914	0.013	0.97	0.914	0.941	0.957	no
Weighted Avg.						
0.965	0.063	0.965	0.965	0.965	0.957	

=== Confusion Matrix ===

```
a  b  <-- classified as
78  1  |  a = si
 3 32 |  b = no
```

Anexo F5 Modelo 2do año- J48- cross-validation

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo 2do año-2clases
Instances: 120
Attributes: 18
mat iii
ep1ap
ed1
iaa
id0 3
segNa1
EF 3
ele1 1
mat iv
A1omp
ED 2
BD
TSP
id0 4
Dna1
EF iv
Pprof
aprueba2
Test mode: 3-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 0.07 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	103	85.8333 %
Incorrectly Classified Instances	17	14.1667 %
Kappa statistic	0.7021	
Mean absolute error	0.1828	
Root mean squared error	0.3423	
Relative absolute error	37.7951 %	
Root relative squared error	69.6161 %	
Coverage of cases (0.95 level)	97.5 %	
Mean rel. region size (0.95 level)	79.5833 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.776	0.085	0.864	0.776	0.817	0.705	0.884	0.854	no
0.915	0.224	0.855	0.915	0.884	0.705	0.884	0.872	si
Weighted Avg.								
0.858	0.167	0.859	0.858	0.857	0.705	0.884	0.865	

=== Confusion Matrix ===

```
a  b  <-- classified as
38 11 | a = no
 6 65 | b = si
```

Anexo F6 Modelo 2do año- J48- training set

=== Run information ===

```
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    Modelo 2do año-2clases
Instances:   120
Attributes:  18
             mat iii
             eplap
             ed1
             iia
             id0 3
             segNa1
             EF 3
             ele1 1
             mat iv
             A1omp
             ED 2
             BD
             TSP
             id0 4
             Dna1
             EF iv
             Pprof
             aprueba2
Test mode:   evaluate on training data
```

=== Classifier model (full training set) ===

Time taken to build model: 0.03 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0 seconds

=== Summary ===

Correctly Classified Instances	109	90.8333 %
Incorrectly Classified Instances	11	9.1667 %
Kappa statistic	0.806	
Mean absolute error	0.1574	
Root mean squared error	0.2803	
Relative absolute error	32.564 %	
Root relative squared error	57.025 %	
Coverage of cases (0.95 level)	100 %	
Mean rel. region size (0.95 level)	85 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.816	0.028	0.952	0.816	0.879	0.812	0.907	0.888	no
0.972	0.184	0.885	0.972	0.926	0.812	0.907	0.886	si
Weighted Avg.								
0.908	0.12	0.912	0.908	0.907	0.812	0.907	0.887	

=== Confusion Matrix ===

```
a  b  <-- classified as
40  9  |  a = no
 2 69  |  b = si
```

Anexo F7 Modelo 2do año selrasg- J48- cross-validation

=== Run information ===

```
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    Modelo 2do año-2clases-
weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16
Instances:   120
Attributes:  9
             mat iii
             ed1
             iia
             mat iv
             Alomp
             ED 2
             BD
             Pprof
             aprueba2
Test mode:   10-fold cross-validation
```

=== Classifier model (full training set) ===
Time taken to build model: 0.09 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	105	87.5	%
Incorrectly Classified Instances	15	12.5	%
Kappa statistic	0.7354		
Mean absolute error	0.1978		
Root mean squared error	0.3352		
Relative absolute error	40.8907	%	
Root relative squared error	68.1766	%	
Coverage of cases (0.95 level)	100	%	
Mean rel. region size (0.95 level)	90.4167	%	
Total Number of Instances	120		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.776	0.056	0.905	0.776	0.835	0.741	0.861	0.865	no
0.944	0.224	0.859	0.944	0.899	0.741	0.861	0.839	si
Weighted Avg.								
0.875	0.156	0.878	0.875	0.873	0.741	0.861	0.849	

=== Confusion Matrix ===

```
a  b  <-- classified as
38 11 | a = no
 4 67 | b = si
```

Anexo F8 Modelo 2do año selrasg- J48- training set

=== Run information ===

```
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:     Modelo 2do año-2clases-
weka.filters.unsupervised.attribute.Remove-R2,5-8,13-16
Instances:    120
Attributes:    9
               mat iii
               ed1
               iia
               mat iv
               Alomp
               ED 2
               BD
               Pprof
               aprueba2
```

Test mode: evaluate on training data

=== Classifier model (full training set) ===

Time taken to build model: 0.01 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0.01 seconds

=== Summary ===

Correctly Classified Instances	111	92.5	%
Incorrectly Classified Instances	9	7.5	%
Kappa statistic	0.8423		
Mean absolute error	0.1353		
Root mean squared error	0.2598		
Relative absolute error	27.9932	%	
Root relative squared error	52.8561	%	
Coverage of cases (0.95 level)	100	%	
Mean rel. region size (0.95 level)	88.75	%	
Total Number of Instances	120		

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.857	0.028	0.955	0.857	0.903	0.846	0.924	0.905	no
0.972	0.143	0.908	0.972	0.939	0.846	0.924	0.906	si
Weighted Avg.								
0.925	0.096	0.927	0.925	0.924	0.846	0.924	0.906	

=== Confusion Matrix ===

```
a  b  <-- classified as
42  7  |  a = no
 2 69  |  b = si
```

Anexo F9 Modelo 1er y 2do año- J48- cross-validation

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo 1ro y 2do_año
Instances: 134
Attributes: 33
Mat I
FyS
FInf
IP
EF I
Id0 I
Alg y GA
H1uba
Id0 II
DPOO
EF II
IGS
MD
Mat II
EP1ap
mat iii
ep1ap
ed1
iia
id0 3
segNa1
EF 3
ele1 1
mat iv
A1omp
ED 2
BD
TSP
id0 4
Dna1
EF iv
Pprof
aprueba

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 0.03 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	105	78.3582 %
Incorrectly Classified Instances	29	21.6418 %
Kappa statistic	0.5672	
Mean absolute error	0.253	
Root mean squared error	0.3987	
Relative absolute error	50.577 %	
Root relative squared error	79.7068 %	

Coverage of cases (0.95 level)	97.7612 %
Mean rel. region size (0.95 level)	83.209 %
Total Number of Instances	134

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.716	0.149	0.828	0.716	0.768	0.853	no
0.851	0.284	0.75	0.851	0.797	0.853	si
Weighted Avg.						
0.784	0.216	0.789	0.784	0.783	0.853	

=== Confusion Matrix ===

```
  a  b  <-- classified as
48 19 |  a = no
10 57 |  b = si
```

Anexo F10 Modelo 1er y 2do año- J48- training set

=== Run information ===

```
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    Modelo 1ro y 2do_año
Instances:   134
Attributes:  33
             Mat I
             FyS
             FInf
             IP
             EF I
             Id0 I
             Alg y GA
             H1uba
             Id0 II
             DPOO
             EF II
             IGS
             MD
             Mat II
             EP1ap
             mat iii
             ep1ap
             ed1
             iia
             id0 3
             segNa1
             EF 3
             ele1 1
             mat iv
             A1omp
             ED 2
             BD
             TSP
             id0 4
             Dna1
             EF iv
             Pprof
             aprueba
Test mode:   evaluate on training data
```

=== Classifier model (full training set) ===

Time taken to build model: 0.03 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	127	94.7761 %
Incorrectly Classified Instances	7	5.2239 %
Kappa statistic	0.8955	
Mean absolute error	0.1725	
Root mean squared error	0.2698	
Relative absolute error	34.4969 %	
Root relative squared error	53.9632 %	
Coverage of cases (0.95 level)	100 %	
Mean rel. region size (0.95 level)	78.7313 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.925	0.03	0.969	0.925	0.947	0.966	no
0.97	0.075	0.929	0.97	0.949	0.966	si
Weighted Avg.						
0.948	0.052	0.949	0.948	0.948	0.966	

=== Confusion Matrix ===

a	b	<-- classified as
62	5	a = no
2	65	b = si

Anexo F11 Modelo 1er y 2do año selrasg- J48- cross-validation

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo 1ro y 2do año sel_rasg
Instances: 134
Attributes: 14
Mat I
FyS
IP
DP00
MD
Mat II
mat iii
ed1
iia
mat iv
A1omp
ED 2
BD
aprueba
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	103	76.8657 %
Incorrectly Classified Instances	31	23.1343 %
Kappa statistic	0.5373	
Mean absolute error	0.2734	
Root mean squared error	0.4061	
Relative absolute error	54.6656 %	
Root relative squared error	81.1861 %	
Coverage of cases (0.95 level)	97.0149 %	
Mean rel. region size (0.95 level)	84.3284 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.716	0.179	0.8	0.716	0.756	0.847	no
0.821	0.284	0.743	0.821	0.78	0.847	si
Weighted Avg.						
0.769	0.231	0.772	0.769	0.768	0.847	

=== Confusion Matrix ===

```
a b <-- classified as
48 19 | a = no
12 55 | b = si
```

Anexo F12 Modelo 1er y 2do año selrasg- J48- training set

=== Run information ===

```
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    Modelo 1ro y 2do año sel_rasg
Instances:   134
Attributes:  14
             Mat I
             FyS
             IP
             DPOO
             MD
             Mat II
             mat iii
             ed1
             iia
             mat iv
             A1omp
             ED 2
             BD
             aprueba
Test mode:   evaluate on training data
```

=== Classifier model (full training set) ===

Time taken to build model: 0.02 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	118	88.0597 %
Incorrectly Classified Instances	16	11.9403 %
Kappa statistic	0.7612	
Mean absolute error	0.1941	
Root mean squared error	0.2947	
Relative absolute error	38.8254 %	
Root relative squared error	58.9352 %	
Coverage of cases (0.95 level)	100 %	
Mean rel. region size (0.95 level)	86.5672 %	
Total Number of Instances	134	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.821	0.06	0.932	0.821	0.873	0.943	no
0.94	0.179	0.84	0.94	0.887	0.943	si
Weighted Avg.						
0.881	0.119	0.886	0.881	0.88	0.943	

=== Confusion Matrix ===

```
a  b  <-- classified as
55 12 | a = no
 4 63 | b = si
```

Anexo G.

Resultados de la prueba de los Modelos entrenados por el árbol de decisión.

Anexo G1 Modelo 1er año- J48- CP

=== Model information ===

Filename: j48.model
Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo_1er año
Attributes: 16
Mat I
FyS
FInf
IP
EF I
Id0 I
Alg y GA
H1uba
Id0 II
DPOO
EF II
IGS
MD
Mat II
EP1ap
aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP-1er_año
Instances: 14
Attributes: 16

=== Summary ===

Correctly Classified Instances	12	85.7143 %
Incorrectly Classified Instances	2	14.2857 %
Kappa statistic	0.6957	
Mean absolute error	0.1593	
Root mean squared error	0.3646	
Coverage of cases (0.95 level)	85.7143 %	
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.333	0.8	1	0.889	0.73	0.833	0.8	si
0.667	0	1	0.667	0.8	0.73	0.833	0.81	no
Weighted Avg.								
0.857	0.19	0.886	0.857	0.851	0.73	0.833	0.804	

=== Confusion Matrix ===

a b <-- classified as
8 0 | a = si
2 4 | b = no

Anexo G2 Modelo 1er año selrasg- J48- CP

=== Model information ===

Filename: j48.model
Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo_1er año selrasg
Attributes: 11
Mat I
FyS
FInf
IP
Alg y GA
Id0 II
DPOO
IGS
MD
Mat II
aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP-1er_año selrasg
Instances: 13
Attributes: 11

=== Summary ===

Correctly Classified Instances	11	84.6154 %
Incorrectly Classified Instances	2	15.3846 %
Kappa statistic	0.6829	
Mean absolute error	0.1686	
Root mean squared error	0.3782	
Coverage of cases (0.95 level)	84.6154 %	
Total Number of Instances	13	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
1	0.333	0.778	1	0.875	0.72	0.833	0.778	si
0.667	0	1	0.667	0.8	0.72	0.833	0.821	no
Weighted Avg.								
0.846	0.179	0.88	0.846	0.84	0.72	0.833	0.798	

=== Confusion Matrix ===

a b <-- classified as
7 0 | a = si
2 4 | b = no

Anexo G3 Modelo 2do año- J48- CP

=== Model information ===

Filename: j48.model
Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo 2do año
Attributes: 18
Mat III
Epcap
ED I
IIA
IdEx III
SegNac
EF III
elec
Mat IV
Arqcomp
ED II
BD
TSP
IdEx IV
DeNac
EF IV
Pprof
Aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP- 2do año
Instances: 15
Attributes: 18

=== Summary ===

Correctly Classified Instances	13	86.6667 %
Incorrectly Classified Instances	2	13.3333 %
Kappa statistic	0.6591	
Mean absolute error	0.1322	
Root mean squared error	0.3028	
Coverage of cases (0.95 level)	100 %	
Total Number of Instances	15	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.909	0.25	0.909	0.909	0.909	0.943	no
0.75	0.091	0.75	0.75	0.75	0.943	si
Weighted Avg.						
0.867	0.208	0.867	0.867	0.867	0.943	

=== Confusion Matrix ===

a	b	<-- classified as
10	1	a = no
1	3	b = si

Anexo G4 Modelo 2do año selrasg- J48- CP

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo 2do año selrasg
Instances: 120
Attributes: 9
Mat III
ED I
IIA
Mat IV
Arqcomp
ED II
BD
Pprof
Aprueba
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===
Time taken to build model: 0.03 seconds

=== Re-evaluation on test set ===
User supplied test set
Relation: CP- 2do año selrasg.arff
Instances: 15
Attributes: 9

=== Summary ===

Correctly Classified Instances	12	80	%
Incorrectly Classified Instances	3	20	%
Kappa statistic	0.4444		
Mean absolute error	0.2111		
Root mean squared error	0.4117		
Coverage of cases (0.95 level)	100	%	
Total Number of Instances	15		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.833	0.333	0.909	0.833	0.87	0.847	no
0.667	0.167	0.5	0.667	0.571	0.847	si
Weighted Avg.						
0.8	0.3	0.827	0.8	0.81	0.847	

=== Confusion Matrix ===

a	b	<-- classified as
10	2	a = no
1	2	b = si

Anexo G5 Modelo 1er y 2do año- J48- CP

=== Model information ===

Filename: j48.model
Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo 1ro y 2do_año
Attributes: 33
Mat I
FyS
FInf
IP
EF I
Id0 I
Alg y GA
H1uba
Id0 II
DPOO
EF II
IGS
MD
Mat II
EP1ap
mat iii
ep1ap
ed1
iia
id0 3
segNa1
EF 3
ele1 1
mat iv
A1omp
ED 2
BD
TSP
id0 4
Dna1
EF iv
Pprof
aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP- 1ro y 2do_año
Instances: 24
Attributes: 33

=== Summary ===

Correctly Classified Instances	20	83.3333 %
Incorrectly Classified Instances	4	16.6667 %
Kappa statistic	0.6596	
Mean absolute error	0.2111	
Root mean squared error	0.3538	
Coverage of cases (0.95 level)	100	%
Total Number of Instances	24	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.8	0.111	0.923	0.8	0.857	0.669	0.889	0.915	no
0.889	0.2	0.727	0.889	0.8	0.669	0.889	0.743	si
Weighted Avg.								
0.833	0.144	0.85	0.833	0.836	0.669	0.889	0.851	

=== Confusion Matrix ===

a	b	<-- classified as
12	3	a = no
1	8	b = si

Anexo G6 Modelo 1er y 2do año selrasg- J48- CP

=== Model information ===

Filename: j48.model
Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Modelo 1ro y 2do año sel_rasg
Attributes: 14
Mat I
FyS
IP
DP00
MD
Mat II
mat iii
ed1
iia
mat iv
Alomp
ED 2
BD
aprueba

=== Classifier model ===

=== Re-evaluation on test set ===

User supplied test set
Relation: CP- 1er y 2do año selrasg
Instances: 24
Attributes: 14

=== Summary ===

Correctly Classified Instances	20	83.3333 %
Incorrectly Classified Instances	4	16.6667 %
Kappa statistic	0.6596	
Mean absolute error	0.2146	
Root mean squared error	0.3592	
Coverage of cases (0.95 level)	100 %	
Total Number of Instances	24	

=== Detailed Accuracy By Class ===

TPRate	FPRate	Precision	Recall	F-Measure	MCC	ROCArea	PRCArea	Class
0.8	0.111	0.923	0.8	0.857	0.669	0.881	0.915	no
0.889	0.2	0.727	0.889	0.8	0.669	0.881	0.718	si
Weighted Avg.								
0.833	0.144	0.85	0.833	0.836	0.669	0.881	0.841	

=== Confusion Matrix ===

a	b	<-- classified as
12	3	a = no
1	8	b = si

Anexo H.
Experimentos Redes versus Árbol de Decisión.
Anexo H1.

Experimento para modelos de 1er año.

Tester: weka.experiment.PairedCorrectedTTester

Analysing: Percent_correct

Datasets: 2

Resultsets: 8

Confidence: 0.05 (two tailed)

Sorted by: -

Date: 25/05/15 10:20

Dataset	(8)	trees.J4		(1)	funcnt(2)	funcnt	(3)	funcnt	(4)	funcnt	(5)	funcnt	(6)	funcnt(7)	funcnt
Modelo_1er año selrasg	(9)	92.11		83.92	83.92	86.26	85.09	85.09	84.50	84.21					
Modelo_1er año	(9)	92.11		81.29	83.04 *	82.75	80.70	81.29	79.53 *	81.29					
		(v/	/*)		(0/2/0)	(0/1/1)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/1/1)	(0/2/0)			

Analysing: Kappa_statistic

Dataset	(8)	trees.J4		(1)	funcnt(2)	funcnt	(3)	funcnt	(4)	funcnt	(5)	funcnt	(6)	funcnt(7)	funcnt
Modelo_1er año selrasg	(9)	0.81		0.63	0.62	0.68	0.66	0.66	0.65	0.64					
Modelo_1er año	(9)	0.81		0.58	0.61 *	0.60	0.56	0.58	0.54 *	0.57					
		(v/	/*)		(0/2/0)	(0/1/1)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/1/1)	(0/2/0)			

Analysing: Mean_absolute_error

Dataset	(8)	trees.J		(1)	func	(2)	func	(3)	func	(4)	func	(5)	func	(6)	func	(7)	func
Modelo_1er año selrasg	(9)	0.12		0.17		0.16		0.15		0.17		0.17		0.17		0.17	
Modelo_1er año	(9)	0.12		0.19		0.17		0.18		0.19		0.20		0.20		0.19	
		(v/	/*)		(0/2/0)	(0/2/0)		(0/2/0)		(0/2/0)		(0/2/0)		(0/2/0)		(0/2/0)	

Key:

- (1) functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"12, 10\" - 5990607817048210779
- (2) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H i - 5990607817048210779
- (3) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H o - 5990607817048210779
- (4) functions.MultilayerPerceptron -L 0.5 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"12, 10\" - 5990607817048210779
- (5) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"40, 20\" - 5990607817048210779
- (6) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"50, 30\" - 5990607817048210779
- (7) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"20, 23\" - 5990607817048210779
- (8) trees.J48 -C 0.25 -M 2 -217733168393644444

Anexo H2.
Experimento para modelos de 2do año.

Tester: weka.experiment.PairedCorrectedTTester

Analysing: Percent_correct

Datasets: 2

Resultsets: 8

Confidence: 0.05 (two tailed)

Sorted by: -

Date: 25/05/15 11:21

Dataset	(8)	trees.J4		(1)	funct(2)	funct	(3)	funct	(4)	funct	(5)	funct	(6)	funct(7)	funct
Modelo 2do año selrasg	(9)	87.22		85.56	86.11	86.94	85.83	85.83	85.28	85.00					
Modelo 2do año	(9)	86.11		81.39	80.83	81.39	81.11	80.83	80.28	81.39					
		(v/ /*)		(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)

Analysing: Kappa_statistic

Dataset	(8)	trees.J4		(1)	funct(2)	funct	(3)	funct	(4)	funct	(5)	funct	(6)	funct(7)	funct
Modelo 2do año selrasg	(9)	0.73		0.70	0.71	0.73	0.70	0.70	0.69	0.68					
Modelo 2do año	(9)	0.71		0.62	0.60	0.61	0.61	0.60	0.59	0.62					
		(v/ /*)		(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)

Analysing: Mean_absolute_error

Dataset	(8)	trees.J4		(1)	funct(2)	funct	(3)	funct	(4)	funct	(5)	funct	(6)	funct(7)	funct
Modelo 2do año selrasg	(9)	0.19		0.16	0.17	0.16	0.16	0.16	0.16	0.16					
Modelo 2do año	(9)	0.18		0.20	0.19	0.20	0.19	0.20	0.20	0.19					
		(v/ /*)		(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)

Key:

(1) functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"12, 10\" - 5990607817048210779

(2) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H i - 5990607817048210779

(3) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H o - 5990607817048210779

(4) functions.MultilayerPerceptron -L 0.5 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"12, 10\" - 5990607817048210779

(5) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"40, 20\" - 5990607817048210779

(6) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"50, 30\" - 5990607817048210779

(7) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"20, 23\" - 5990607817048210779

(8) trees.J48 -C 0.25 -M 2 -217733168393644444

Anexo H3.
Experimento para modelos de 1er y 2do año.

Tester: weka.experiment.PairedCorrectedTTester

Analysing: Percent_correct

Datasets: 2

Resultsets: 8

Confidence: 0.05 (two tailed)

Sorted by: -

Date: 25/05/15 11:30

Dataset	(8) trees.J4	(1) funct	(2) funct	(3) funct	(4) funct	(5) funct	(6) funct	(7) funct
Modelo 1ro y 2do año sel	(9) 71.40	79.84	79.83	78.84	79.84	79.34	80.07	79.08
Modelo 1ro y 2do_año	(9) 72.65	84.30	83.56	84.30 v	84.30	84.55	84.55	84.55
	(v/ /*)	(0/2/0)	(0/2/0)	(1/1/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)

Analysing: Kappa_statistic

Dataset	(8) trees.J4	(1) funct	(2) funct	(3) funct	(4) funct	(5) funct	(6) funct	(7) funct
Modelo 1ro y 2do año sel	(9) 0.43	0.60	0.60	0.58	0.60	0.59	0.60	0.58
Modelo 1ro y 2do_año	(9) 0.45	0.69	0.67	0.69 v	0.69	0.69	0.69	0.69
	(v/ /*)	(0/2/0)	(0/2/0)	(1/1/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)

Analysing: Mean_absolute_error

Dataset	(8) trees.J4	(1) funct	(2) funct	(3) funct	(4) funct	(5) funct	(6) funct	(7) funct
Modelo 1ro y 2do año sel	(9) 0.30	0.21	0.21 *	0.21	0.21	0.21 *	0.21 *	0.21
Modelo 1ro y 2do_año	(9) 0.29	0.16 *	0.16 *	0.17 *	0.16 *	0.16 *	0.16 *	0.16 *
	(v/ /*)	(0/2/0)	(0/2/0)	(1/1/0)	(0/2/0)	(0/2/0)	(0/2/0)	(0/2/0)

Key:

(1) functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"12, 10\" - 5990607817048210779

(2) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H i - 5990607817048210779

(3) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H o - 5990607817048210779

(4) functions.MultilayerPerceptron -L 0.5 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"12, 10\" - 5990607817048210779

(5) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"40, 20\" - 5990607817048210779

(6) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"50, 30\" - 5990607817048210779

(7) functions.MultilayerPerceptron -L 0.7 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H \"20, 23\" - 5990607817048210779

(8) trees.J48 -C 0.25 -M 2 -217733168393644444

Anexo I.

Pronósticos de los rasgos más característicos.

Anexo I1.

Pronóstico Modelo_1er año selrasg

=== Run information ===

Scheme:

SMOreg -C 1.0 -N 0 -I "RegSMOImproved -L 0.0010 -W 1 -P 1.0E-12 -T 0.0010 -V" -K
"PolyKernel -C 250007 -E 1.0"

Lagged and derived variable options:

-F "[Mat I, IP, Mat II]" -L 1 -M 12

Relation: Modelo_1er año selrasg

Instances: 114

Attributes: 11

Mat I
FyS
FInf
IP
Alg y GA
Id0 II
DPOO
IGS
MD
Mat II
aprueba

Number of kernel evaluations: 6555 (99.566% cached)

=== Future predictions from end of training data ===

inst#	Mat I	IP	Mat II
1	5	5	4
2	3	4	3
3	4	2	3
4	3	3	3
5	3	2	3
6	3	3	3
7	3	5	4
8	3	2	2
9	2	2	2
10	3	3	5
11	3	3	4
12	3	3	4
13	4	5	3
14	4	3	5
15	4	4	3

16	3	3	4
17	3	3	4
18	3	4	3
19	3	2	4
20	2	2	2
21	3	3	3
22	5	4	5
23	3	3	3
24	3	3	4
25	3	2	3
26	3	3	3
27	3	2	3
28	3	4	3
29	3	3	2
30	3	3	3
31	3	3	2
32	3	5	4
33	3	3	3
34	3	3	3
35	3	3	3
36	3	3	3
37	3	5	3
38	4	3	3
39	3	3	3
40	3	3	4
41	3	3	3
42	3	4	4
43	3	3	3
44	3	4	3
45	3	4	4
46	4	5	3
47	3	3	3
48	2	4	3
49	3	4	3
50	3	4	3
51	3	5	3
52	3	3	3
53	3	4	3
54	4	4	3
55	3	3	3
56	3	4	3
57	2	4	3
58	3	3	3
59	3	3	3
60	2	3	3
61	3	3	3
62	4	4	4
63	2	4	3
64	4	5	5
65	3	3	3
66	3	3	3

67	5	5	5
68	3	3	3
69	3	5	3
70	2	3	3
71	3	3	3
72	5	4	5
73	3	5	3
74	4	5	4
75	3	3	3
76	3	3	3
77	3	5	3
78	2	4	2
79	3	3	3
80	3	4	3
81	2	4	3
82	2	3	3
83	3	3	3
84	2	3	2
85	2	3	2
86	5	5	4
87	3	3	3
88	3	3	3
89	5	5	5
90	1	1	3
91	3	4	3
92	2	3	2
93	2	3	2
94	4	3	4
95	2	1	3
96	2	2	?
97	1	1	3
98	3	3	3
99	2	1	1
100	3	3	3
101	3	4	3
102	3	3	3
103	2	3	2
104	2	3	3
105	4	5	5
106	2	3	2
107	2	3	3
108	1	1	3
109	4	5	3
110	2	3	2
111	2	3	2
112	2	3	2
113	4	5	3
114	2	3	2
115*	2.1504	4.0346	3.1602
116*	2.0304	3.2347	3.6985
117*	2.5267	3.4367	1.985

118*	2.1901	2.8814	3.0638
119*	1.7306	3.296	2.0043
120*	2.701	3.1273	3.1189
121*	3.1131	3.9678	2.13
122*	1.9946	2.6504	2.5369
123*	2.6241	3.5661	3.243
124*	2.5407	3.2553	3.1286
125*	1.9794	3.6111	2.2093
126*	1.6869	2.1759	2.7716
127*	2.4594	3.7962	2.6871
128*	2.0036	2.6403	2.9031
129*	2.3009	3.3861	2.0364
130*	1.9664	2.2867	2.8556
131*	2.4032	3.4762	2.8557
132*	2.3868	2.6634	2.6231
133*	1.9244	3.5419	2.2603
134*	1.5693	2.2467	2.9252

Anexo I2.

Pronóstico Modelo 2do año selrasg

=== Run information ===

Scheme:

SMOreg -C 1.0 -N 0 -I "RegSMOImproved -L 0.0010 -W 1 -P 1.0E-12 -T 0.0010 -V" -K
"PolyKernel -C 250007 -E 1.0"

Lagged and derived variable options:

-F "[mat iii, ed1, BD]" -L 1 -M 12

Relation: Modelo 2do año selrasg

Instances: 120

Attributes: 9

mat iii
ed1
iia
mat iv
A1omp
ED 2
BD
Pprof
Aprueba

Number of kernel evaluations: 7260 (99.194% cached)

=== Future predictions from end of training data ===

inst#	mat iii	ed1	BD
1	2	2	2
2	2	2	3
3	2	2	2
4	4	4	4
5	2	3	3
6	2	2	3
7	2	2	2
8	2	2	?
9	3	3	4
10	3	4	2
11	3	3	4
12	2	3	2
13	3	4	3
14	3	2	3
15	2	2	2
16	2	2	?
17	2	2	2
18	3	2	4
19	3	3	2
20	3	3	3
21	3	5	3

22	2	3	3
23	3	3	4
24	2	3	?
25	3	4	4
26	2	3	3
27	3	3	4
28	2	3	3
29	1	2	3
30	3	3	4
31	2	2	2
32	3	4	3
33	3	2	2
34	?	2	?
35	3	2	3
36	1	3	1
37	3	3	4
38	3	3	4
39	1	2	4
40	3	3	4
41	2	3	3
42	3	3	4
43	3	3	5
44	5	3	5
45	3	3	4
46	3	3	4
47	5	5	5
48	3	3	4
49	3	3	4
50	5	3	4
51	4	3	5
52	4	3	5
53	3	3	4
54	3	3	3
55	5	5	5
56	3	3	3
57	3	5	3
58	3	3	5
59	4	5	4
60	3	3	3
61	4	4	3
62	4	3	3
63	2	3	2
64	3	5	5
65	3	3	2
66	2	3	2
67	4	5	4
68	4	4	5
69	3	3	3
70	2	5	2
71	3	3	3
72	3	3	3

73	3	5	5
74	3	3	4
75	3	3	4
76	3	3	4
77	3	3	3
78	3	5	4
79	3	3	2
80	3	4	3
81	3	3	4
82	3	3	3
83	3	3	4
84	3	3	4
85	3	3	3
86	3	4	3
87	3	3	4
88	2	3	3
89	2	3	2
90	3	4	2
91	3	2	3
92	3	3	2
93	2	4	2
94	3	2	3
95	3	2	2
96	3	3	4
97	3	2	3
98	3	5	4
99	3	3	3
100	3	4	4
101	4	3	4
102	5	3	3
103	5	5	4
104	3	3	4
105	1	5	4
106	4	5	4
107	3	5	4
108	4	5	4
109	3	3	3
110	4	5	4
111	3	3	4
112	3	5	4
113	3	2	3
114	5	5	5
115	4	5	4
116	3	3	3
117	4	4	4
118	1	1	2
119	3	3	3
120	3	4	4
121*	3.3072	3.3999	2.8631
122*	2.9493	3.7719	1.3374
123*	3.1331	4.5995	3.3235

124*	3.8567	1.5912	0.8888
125*	2.8407	4.0556	2.1421
126*	3.5065	3.6981	3.5287
127*	2.7764	2.4916	1.6785
128*	3.0281	6.4301	3.8431
129*	3.4587	3.8008	3.5042
130*	3.9256	5.2602	3.2815
131*	3.5177	2.9479	2.9738
132*	4.3376	4.3704	4.2226
133*	3.7954	3.5086	3.461
134*	2.7908	3.4507	4.5976
135*	4.4534	5.6491	4.8664
136*	3.0113	2.9664	3.1761
137*	4.4479	5.2311	4.1735
138*	4.1077	3.5487	3.6043
139*	3.7383	3.0858	2.754
140*	3.3673	3.3387	2.7956

Anexo I3.

Pronóstico de Modelo 1ro y 2do sel_rasg.

=== Run information ===

Scheme:

SMOreg -C 1.0 -N 0 -I "RegSMOImproved -L 0.0010 -W 1 -P 1.0E-12 -T 0.0010 -V" -K
"PolyKernel -C 250007 -E 1.0"

Lagged and derived variable options:

-F "[Mat I, MD, Mat II, ed1, BD]" -L 1 -M 12

Relation: Modelo 1ro y 2do año sel_rasg

Instances: 134

Attributes: 14

Mat I
FyS
IP
DPOO
MD
Mat II
mat iii
ed1
iia
mat iv
A1omp
ED 2
BD
Aprueba

Number of kernel evaluations: 9045 (99.565% cached)

=== Future predictions from end of training data ===

i#	Mat I	MD	Mat II	ed1	BD
1	3	3	3	2	2
2	2	3	2	2	3
3	3	3	2	2	2
4	4	3	4	4	4
5	3	3	3	3	3
6	3	3	3	2	3
7	3	3	3	2	2
8	3	3	3	2	?
9	3	4	3	3	4
10	3	4	3	4	2
11	3	3	3	3	4
12	3	3	3	3	2
13	3	4	3	4	3
14	4	3	4	2	3

15	?	1	1	2	2
16	3	3	2	2	?
17	2	3	3	2	2
18	4	3	3	2	4
19	1	3	3	3	2
20	3	4	3	3	3
21	4	4	4	5	3
22	3	3	2	3	3
23	3	3	3	3	4
24	2	3	2	3	?
25	4	3	3	4	4
26	3	3	3	3	3
27	3	4	4	3	4
28	3	3	3	3	3
29	2	?	?	2	3
30	3	3	3	3	4
31	3	1	3	2	2
32	3	3	3	4	3
33	1	1	1	2	2
34	2	1	2	2	?
35	3	3	3	2	3
36	1	1	1	3	1
37	3	4	3	3	4
38	4	4	4	3	4
39	1	?	?	2	4
40	2	3	3	3	4
41	3	3	3	3	5
42	2	2	2	?	?
43	3	3	3	3	2
44	5	4	4	3	5
45	3	3	3	3	4
46	3	3	3	3	4
47	5	5	5	5	5
48	1	1	3	?	?
49	3	3	3	3	4
50	2	2	2	?	?
51	2	2	2	?	?
52	4	3	4	3	4
53	2	3	3	?	?
54	2	?	?	?	?
55	1	1	3	3	5
56	3	4	3	3	5
57	2	?	1	?	?
58	3	3	3	?	?
59	3	2	3	?	?
60	3	3	3	?	?
61	2	2	2	?	?
62	2	3	3	3	4
63	4	4	5	5	5
64	2	2	2	?	?
65	2	3	3	?	?

66	1	1	3	?	?
67	4	3	3	5	3
68	2	2	2	?	?
69	2	2	2	?	?
70	2	3	2	?	?
71	4	3	3	5	4
72	1	1	1	4	3
73	2	3	2	?	?
74	4	3	4	3	3
75	2	3	3	3	2
76	4	4	5	5	5
77	3	3	3	3	2
78	5	5	5	5	4
79	3	3	3	3	3
80	3	3	3	5	2
81	2	3	3	3	3
82	3	3	3	3	3
83	5	4	5	5	5
84	3	3	3	3	4
85	1	1	1	3	4
86	1	1	1	3	4
87	1	1	1	3	3
88	4	4	4	5	4
89	3	3	3	3	2
90	1	1	1	4	3
91	1	1	1	3	4
92	3	3	3	3	3
93	3	3	3	3	4
94	2	3	2	3	4
95	3	4	3	3	3
96	3	4	3	4	3
97	2	3	3	3	4
98	1	1	1	3	2
99	3	3	3	4	2
100	3	3	2	?	?
101	3	3	3	2	3
102	3	3	2	3	2
103	3	3	4	4	2
104	3	3	3	2	3
105	3	3	3	2	2
106	3	3	3	3	4
107	3	3	3	2	3
108	3	3	3	5	4
109	4	5	3	3	3
110	3	3	3	4	4
111	3	4	4	3	4
112	3	3	3	?	?
113	3	5	4	5	4
114	3	3	3	3	4
115	3	3	3	5	4
116	3	3	4	5	4

117	3	4	3	5	4
118	4	5	3	5	4
119	2	2	3	?	?
120	1	1	1	3	3
121	3	4	3	5	4
122	3	4	3	3	4
123	3	4	3	5	4
124	3	3	3	2	3
125	3	4	3	5	5
126	4	5	3	5	4
127	3	3	3	?	?
128	3	3	3	3	3
129	2	2	3	?	?
130	3	4	3	4	4
131	1	1	1	1	2
132	3	3	3	3	3
133	2	2	3	?	?
134	3	3	3	4	4
135*	2.8664	4.1128	2.4698	1.7554	2.5597
136*	1.4183	2.0372	3.1833	3.8785	4.3709
137*	2.9102	3.7035	3.6493	1.7431	1.5141
138*	3.6524	3.5146	3.3874	5.021	5.1454
139*	3.5803	4.5771	4.342	3.1046	3.4258
140*	2.1628	2.0817	1.9553	4.849	4.3173
141*	4.7593	5.4228	5.4147	4.4185	3.7714
142*	1.9812	2.5055	1.7974	3.0418	4.4765
143*	4.9192	4.7063	4.8137	7.2898	5.8412
144*	1.3759	3.5953	2.3898	3.0949	3.373
145*	5.5411	5.2314	5.474	6.5603	4.8449
146*	1.983	3.5346	2.6249	3.6145	4.879
147*	1.9562	2.9153	3.2056	6.0493	5.7461
148*	4.4918	4.1519	3.7309	4.6927	3.1331
149*	2.4595	3.8797	3.7659	5.6057	5.3453
150*	4.1153	4.3606	3.4212	5.0877	5.4678
151*	0.3497	1.2698	0.8871	4.6031	3.6885
152*	5.1492	5.7012	5.0687	6.0877	4.3822
153*	0.342	2.0899	1.0328	3.1431	4.249
154*	4.7791	6.4687	5.0243	7.7753	6.8096