Facultad de Ingeniería

Carrera de Ingeniería Informática

"Modelo basado en aprendizaje automático para el pronóstico al egreso de pacientes con trauma grave de cráneo"

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

Autor: Roberto Suárez Surí

Tutores: MSc. Oscar Luis Muñoz González

MSc. Florencio Pons Moscoso

Consultante: Msc. Pedro Roberto Suárez Surí

Cienfuegos

Curso 2012-2013

"Año 55 de la Revolución"

PENSAMIENTO

Así como perfeccionamos las ciencias, debemos perfeccionar la moralidad, sin la cual el saber se destruye

J. Newton

DEDICATORIA

Especialmente dedico éste trabajo de diploma a mis padres María Luisa y Pedro Roberto a los cuales amo mucho; por la educación, cariño, dedicación, y amor que me han dado; complaciéndome siempre. A ellos por apoyarme y ayudarme siempre que lo he necesitado. También a mi hermano, a mis abuelos. A todos mis familiares, por lo cual le doy muchas gracias a Dios por haberme permitido formar parte de ellos y también a todos mis amigos.

AGRADECIMIENTOS

✓ *A mis padres por todos los sacrificios que han tenido que hacer para cumplir mis sueños, por creer y confiar siempre en mí, estar presentes cuando los necesito y por el amor incondicional que me brindan.*

✓ *A hermano por su apoyo, ayuda y confianza*.

✓ *A mis abuelos, tías, primo(a)s, a todos en general a todos lo que de una forma u otra han sido especiales conmigo.*

✓ *A mis tutores Oscar, Pons y mi padre, los cuales me han ayudado muchísimo con la tesis y han trabajado junto conmigo (se los recomiendo).*

✓ *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.*

✓ *Al team ALASKA (Jorge, Shinno, Luis, Pablo), que los considero como mis hermanos.*

Resumen

En esta investigación se aborda el problema del pronóstico de eventos probabilístico desde la perspectiva de las técnicas de inteligencia artificial. En específico se aplican las técnicas de aprendizaje automático: redes neuronales artificiales, árbol de decisión C4.5 y el IBK en el pronóstico del estado al egreso de pacientes con traumas graves de cráneo ingresados en la sala de cuidados intensivos del Hospital Clínico Quirúrgico Docente "Gustavo Aldereguía Lima" de la provincia de Cienfuegos. En la preparación de los datos se aplicó la selección de rasgos por los métodos *CfsSubsetEval y componentes principales.* En la investigación se utilizó el enfoque de la computación granular con el empleo de dos criterios: del especialista médico y de granulación óptima a través del Detector automático de interacciones Chi-cuadrado (CHAID). Se logró establecer un modelo de aprendizaje automático que garantiza como mínimo una mejora de un 12 % en la clasificación respecto a la realizada tradicionalmente.

Índice

Introducción 1
Capítulo 1 Marco teórico 5
1.1 Medicina en Cuba 5
1.2 La inteligencia artificial 6
1.2.1 Aprendizaje automático 9
1.3 Técnicas de clasificación. Caracterización general. 10
1.3.1 Redes neuronales artificiales 10
1.3.2 Árboles de decisión. 16
1.3.3 IBK: EL vecino más cercano. 18
1.4 Selección de Rasgos 20
1.4.1 CfsSubsetEval 20
1.4.2 PrincipalComponents (Componentes principales) 21
1.5 Computación granular 21
1.5.1 Granularidad 22
1.6 Conclusiones parciales del capítulo 23
Capítulo 2 Preprocesamiento de los datos 24
2.1 Conjuntos de datos 24
2.2 Herramientas a utilizar. 31
2.2.1 Weka 31
2.3 Selección de rasgos 34
2.4 Configuración de las técnicas de clasificación. 38
2.4.1 Parámetros considerados para el entrenamiento de la red neuronal MLP. 38
2.4.2 Parámetros a considerar para la utilización del J48. 40
2.4.3 Parámetros a considerar en la utilización del IBK 41
2.5 Indicadores para comparar resultados del experimento. 42
2.6 Conclusiones parciales del capítulo. 44
Capítulo 3 Análisis de los resultados. 46
3.1 Resultados del experimento. 46
3.2.1 Redes neuronales artificiales: MLP 46
3.2.2 Modelo al ingreso 46
3.2.2 Árbol de decisión j48 48
3.2.3 IBK (vecino más cercano) 49
3.3 Comparación de resultados de modelos al ingreso 51
3.4 Modelo a las 24 horas 52
3.4.1 MLP 52
3.4.2 Árbol de decisión J48 56
3.4.3 IBK (Vecino más cercano) 59
3.5 Comparación de los modelos a las 24 horas 61
3.6 Modelo a las 48 horas del ingreso 63
3.6.1 MPL 63
3.6.2 Árbol de decisión J48 66
3.6.3 IBK (Vecino más cercano) 69
3.7 Conclusiones de los resultados en el modelo a las 48 horas 71
3.8 Modelo para el pronóstico a las 72 horas del ingreso 73
3.8.1 MLP 73
3.8.2 Árbol de decisión J48 76
3.8.3 IBK (Vecino más cercano) 78
3.9 Comparación de modelos a las 72 horas 81

Conclusiones 83
Recomendaciones 85
Referencias Bibliográficas 86
Bibliografía 89
Anexos 93

Índice de tablas

Tabla 1Variables a utilizar en cada conjunto de datos 26
Tabla 2 Selección de rasgos en los conjuntos a las 24 horas. 35
Tabla 3 Selección de rasgos en los conjuntos a las 48 horas. 36
Tabla 4 Selección de rasgos en los conjuntos a las 72 horas. 37
Tabla 5 Define las capas ocultas 38
Tabla 6 Indicadores a comparar 42
Tabla 7 Coeficiente de Kappa 44
Tabla 8 Configuración de la red neuronal 46
Tabla 9 Percent_correct 46
Tabla 10 Kappa_statistic 47
Tabla 11True_positive_rate 47
Tabla 12True_negative_rate 47
Tabla 13Percent_correct 48
Tabla 14Kappa_statistic 48
Tabla 15 True_positive_rate 49
Tabla 16 True_negative_rate 49
Tabla 17 Percent_correct 49
Tabla 18 Kappa_statistic 50
Tabla 19 True_positive_rate 50
Tabla 20 True_negative_rate 50
Tabla 21 Comparación de las técnicas al ingreso 52
Tabla 22 Percent_correct 52
Tabla 23 Kappa_statistic 53
Tabla 24 True_positive_rate 54
Tabla 25 True_negative_rate 54
Tabla 26 False_positive_rate 55
Tabla 27 False_negative_rate 55
Tabla 28 Percent_correct 56
Tabla 29 Kappa_statistic 57
Tabla 30 True_positive_rate 57
Tabla 31 True_negative_rate 58
Tabla 32 False_positive_rate 58
Tabla 33 False_negative_rate 58
Tabla 34 Percent_correct 59
Tabla 35 Kappa_statistic 59
Tabla 36 True_positive_rate 60
Tabla 37 True_negative_rate 60
Tabla 38 False_positive_rate 61
Tabla 39 False_negative_rate 61
Tabla 40 Comparación de las técnicas a las 24 horas 62
Tabla 41 Percent_correct 63
Tabla 42 Kappa_statistic 64
Tabla 43 True_positive_rate 64
Tabla 44 True_negative_rate 65
Tabla 45 False_positive_rate 65
Tabla 46 False_negative_rate 65
Tabla 47 Percent_correct 66
Tabla 48 Kappa_statistic 66

Tabla 49 True_positive_rate 67
Tabla 50 True_negative_rate 67
Tabla 51 False_positive_rate 68
Tabla 52 False_negative_rate 68
Tabla 53 68
Tabla 54 Percent_correct 69
Tabla 55 Kappa_statistic 69
Tabla 56 True_positive_rate 70
Tabla 57 True_negative_rate 70
Tabla 58 False_positive_rate 71
Tabla 59 False_negative_rate 71
Tabla 60 Comparación de las técnicas a las 48 horas 72
Tabla 61 Percent_correct 73
Tabla 62 Kappa_statistic 74
Tabla 63 74
Tabla 64 True_positive_rate 74
Tabla 65 True_negative_rate 74
Tabla 66 False_positive_rate 75
Tabla 67 False_negative_rate 75
Tabla 68 Percent_correct 76
Tabla 69 Kappa_statistic 76
Tabla 70 True_positive_rate 77
Tabla 71 False_negative_rate 77
Tabla 72 False_positive_rate 77
Tabla 73 True_negative_rate 78
Tabla 74 Percent_correct 78
Tabla 75 Kappa_statistic 79
Tabla 76 True_positive_rate 79
Tabla 77 False_negative_rate 80
Tabla 78 False_positive_rate 80
Tabla 79 True_negative_rate 80
Tabla 80 Comparación de las técnicas a las 72 horas 82

Índice de figuras.

Figura 1 similitud de un modelo de red biológica y una artificial.[25] 11
Figura 2 Ejemplos de arquitectura de redes neuronal [26] 12
Figura 3 Red neuronal 13
Figura 4 Interfaz inicial del Weka 31
Figura 5 Estructura de un archivo.arff 32
Figura 6 Ejemplo de un archivo.arff 32
Figura 7 Interfaz de explorer del Weka 33

Índice de gráficos

Gráfico 1 Ejemplo de un árbol 17
Gráfico 2 Ejemplo de búsqueda del IBK 19
Gráfico 3 Comparación de resultados al ingreso 51
Gráfico 4 Comparación de resultados a las 24 horas 62
Gráfico 5 Comparación de resultados a las 48 horas 72
Gráfico 6 Comparación de resultados a las 72 horas 81

Introducción

El hombre en los diversos campos de su actividad se ha enfrentado a la necesidad de realizar pronósticos de eventos de carácter probabilístico, para lo cual se ha basado tradicionalmente en los métodos estadísticos los cuales se han enriquecidos considerablemente con el desarrollo de la informática. A partir de la segunda mitad del siglo XX junto al auge de las soluciones estadísticas comienza el uso de otros enfoques del problema del pronóstico tales como las técnicas de Inteligencia Artificial.

La experiencia científica acumulada en estos campos pronto se comenzó a transferir al campo de la medicina donde se han obtenido resultados en problemas tales como: procesamiento de imágenes, sistemas expertos para el diagnóstico de enfermedades, etc.

Entre los resultados alcanzados en la realización de pronósticos con el empleo de procedimientos estadísticos en el campo de la medicina se pueden mencionar:

-El pronóstico del estado al egreso de pacientes con enfermedades cerebro vasculares ingresados en terapia intermedia en el hospital clínico quirúrgico Gustavo Aldereguía Lima de Cienfuegos[1].

-Outcome prediction in mild traumatic brain injury: Age and clinical variables are stronger predictors than CT abnormalities[2].

-Pronóstico de supervivencia de infarto cerebral aterotrombótico usando aprendizaje automático[3]. En esta investigación se realiza la predicción del estado de los pacientes con lesión cerebral traumática apacible a través del empleo de una regresión logística binaria.

Estos resultados, considerados satisfactorios por la calidad de los pronósticos que proveen tienen la deficiencia que constituyen modelos estáticos que se hacen obsoletos con el paso del tiempo a consecuencia de los cambios medioambientales y su influencia en el sistema de relaciones entre las variables estudiadas, y carecen de la capacidad para actualizarse continuamente con los datos de los nuevos casos, lo cual hace necesario recalcularlos con datos recientes. Además están sometidos a supuestos muy restrictivos, que en ocasiones no se pueden garantizar.

Estos inconvenientes desaparecen con el uso se las técnicas de Inteligencia artificial (I.A.) debido al uso procesos de aprendizaje automático que se actualiza con el acceso a la nueva información.

"El desarrollo de tecnologías asociadas con las técnicas de inteligencia artificial, aplicadas a la medicina, representa una novedosa perspectiva, que puede reducir los costos, el tiempo de los tratamientos, los errores médicos; así como potenciar el uso de los recursos humanos en las ramas médicas con mayores requerimientos" [4].

En este sentido los investigadores trabajan en la búsqueda de modelos que proporcionen pronósticos lo más precisos posibles, lo que representa un campo de investigación inagotable por la diversidad de enfermedades a las que se enfrenta la ciencia médica.

Los resultados que se han obtenido en este campo de la investigación no siempre han tenido la aceptación adecuada, por una parte por el apego a las prácticas tradicionales y por otra por la complejidad que representa la implementación de su uso en el desempeño hospitalario.

"La posibilidad de una recreación del intelecto humano a escala artificial en el campo de la salud, tropieza con dificultades que es preciso tratar desde una sana postura crítica. Con frecuencia, los pronunciamientos al respecto, caracterizados por un gran optimismo con relación al futuro promisorio de la IA con fines médicos, tratan la realidad de forma fragmentada y estudian sólo ciertos aspectos de la relación médico-paciente"[4].

En las salas de terapia de los hospitales es tratada una gran cantidad de pacientes homogéneos según el tipo de enfermedad o traumas que poseen. Las políticas de tratamiento, además del diagnóstico del estado de los pacientes al ingreso dependen también de la perspectiva de evolución que se establece en cada caso[1]. También es una verdad garantizada que: *"La predicción exacta de resultado es deseable para el tratamiento óptimo"* [2].

El pronóstico de la evolución de estos pacientes es un problema complejo que en la actualidad se realiza solamente a partir de la experiencia de los especialistas que los atienden [5] y del uso de escalas que los ayudan a la toma de decisiones*. "El modelo de razonamiento, ante un problema o dificultad de orden clínico, suele seguir un proceso que no es sistemático, ni exhaustivo, ni crítico con la información; en él, se piensa primero en la experiencia o el*

propio conocimiento acumulado o se consulta con un colega, cuyo conocimiento es igualmente empírico. Es sorprendente e incluso alarmante que según diversas estimaciones, un alto porcentaje de las decisiones clínicas no tienen un fundamento científico sólido y sólo el 20% de la práctica médica se basa en actos de rigurosa efectividad"[5].

Se conoce que la predicción exacta no es posible lograrla, y aunque los porcentajes de clasificación correcta que en la actualidad se logran son aceptables, se espera que sea una actividad susceptible de mejora.

Según lo planteado anteriormente existe la necesidad y la posibilidad de apoyar al personal médico en el proceso de toma de decisiones, mediante los resultados de la aplicación de técnicas que permitan hacer un uso más completo de la gran gama de información que se acumula sobre los pacientes que han sido atendidos según un diagnóstico dado y que por su cuantía no es factible su explotación máxima en el momento de adoptar una decisión; de donde se deriva el siguiente problema:

PROBLEMA

¿Cómo mejorar el pronóstico del estado al egreso de los pacientes con traumas graves de cráneo?

IDEA A DEFENDER

Con el empleo de un modelo basado en técnicas de inteligencia artificial, se mejorará el pronóstico del estado al egreso de los pacientes con traumas craneales graves.

OBJETIVO GENERAL

Elaborar un modelo basado en técnicas de inteligencia artificial que permita realizar de forma fiable el pronóstico del estado al egreso de los pacientes con traumas craneales graves.

TAREAS

-Entrevistas y encuentros con los especialistas de traumas de cráneo.

-Estudio del estado del arte de la actividad de pronóstico en la medicina.

-Sistematización de los elementos teóricos de las técnicas de inteligencia artificial que se aplicarán para la realización de pronósticos.
-Elaboración del marco teórico del estudio que se realiza.
-Preparación de los conjuntos de datos para la aplicación de las técnicas de IA.
-Experimentación de las técnicas de IA con el conjunto de datos y validación de los resultados.
-Comparación de los resultados experimentales obtenidos entre las diversas técnicas aplicadas para seleccionar la técnica de IA más eficiente.
-Elaboración del documento final de la investigación.

APORTE PRÁCTICO

Un modelo elaborado para el pronóstico que apoya a los especialistas en la toma de decisiones durante la atención a los pacientes.
Se aplica la técnica de minería de datos Detector automático de interacciones Chi-cuadrado para el establecimiento de categorías que optimicen la relación entre las variables independientes y la dependiente en el proceso de granulación de los datos.

El trabajo estará estructurado de la siguiente forma:

Introducción, tres capítulos, conclusiones, recomendaciones, referencias bibliográficas y bibliografía. Los capítulos tendrán la información que a continuación se describe:
Capítulo 1: descripción de las técnicas inteligentes MLP, árbol de decisión j48 y IB1 (Vecino más cercano). Además incluye elementos esenciales sobre la selección de rasgos y granularidad de los datos.
Capítulo 2: diseño del experimento, preparación de los datos y aplicación de las técnicas del MLP, árbol de decisión j48 e IBK.
Capítulo 3: análisis de los resultados determinando el modelo para el pronóstico.

Capítulo 1 Marco teórico

En el presente capítulo se resume el desarrollo alcanzado por el sistema de salud en Cuba y la problemática de salud que representan los traumas de cráneo por accidentes, además se hace una reseña histórica del desarrollo de la inteligencia artificial y su uso creciente en las investigaciones de los últimos años. Se define el aprendizaje como elemento fundamental de un sistema inteligente y se describen tres técnicas de inteligencia artificial.: Perceptrón Multicapas (MLP), árbol de decisión j48 y un IBK (Vecino más cercano) con aprendizaje supervisado. También se describe el uso de la selección de rasgos para la ejecución de las técnicas IA y la técnica de granularidad como una posibilidad para la preparación de los datos

1.1 Medicina en Cuba

La Revolución cubana desde sus inicios dedicó grandes recursos para garantizar la salud a toda la población. Hoy se cuenta con un personal altamente calificado y una infraestructura que permite una atención de calidad a toda la población. La actividad investigativa en el campo de la salud ha permitido un gran desarrollo en la medicina y posibilitado la introducción de las más modernas técnicas existentes en todas las áreas de la práctica médica, convirtiendo al País en una potencia médica.

La accidentabilidad por las más diversas causas produce una cantidad considerable de lesionados con traumas craneales que reciben una atención especializada en las salas de cuidados intensivos. El pronóstico de la perspectiva del estado de esos pacientes ingresados en las instituciones de salud es de gran importancia para la toma de decisiones en cuanto a políticas terapéuticas por parte del personal especializado. Los traumas de cráneo encefálico son producidos debido a caídas, accidentes automovilísticos, golpes y otros. A nivel mundial cada siete segundos ocurre un accidente y cada cinco minutos fallece un individuo con trauma de cráneo; constituyendo la segunda causa de muerte en la población mundial después del cáncer. Entre estos accidentados el grupo de edad de 35 a 45 años es el de mayor letalidad.

En la provincia de Cienfuegos en el año 2012 la primera causa de muerte en menores de 45 años de edad fue la de los traumas de cráneo graves.

El pronóstico de la posible evolución de estos pacientes a partir de los datos clínicos iníciales, pruebas de laboratorio y de escalas diagnósticas, es un problema multifactorial complejo que puede ser apoyado por sistemas informáticos a partir de los resultados de la aplicación de métodos de inteligencia artificial.

En todos los hospitales del sistema nacional de salud se atienden en las salas de cuidados intensivos a los pacientes con traumas craneales. En la sala de cuidados intensivos del Hospital Gustavo Aldereguía Lima se atienden los pacientes con traumas craneales, donde desde hace varios años se lleva el registro de un grupo de variables de interés, conformando un conjunto de datos de gran utilidad para la investigación. Este registro incluye hasta la fecha 83 casos.

1.2 La inteligencia artificial

El desarrollo de las tecnologías en el mundo ha posibilitado a los investigadores enfrentar problemas que requieren investigar el sistema complejo de relaciones existentes entre la diversidad de variables que caracterizan el problema, dando lugar a formas de aprendizaje automatizado que se convierten en un recurso que aumenta considerablemente la posibilidad de análisis de la información existente. Para ello dedica grandes esfuerzos en la creación de algoritmos, métodos y técnicas que realicen procedimientos para agilizar el razonamiento inteligente en la construcción de robots máquinas con acciones físicas e inteligentes muy parecidas o iguales a las humanas. Todas estas acciones que simulan a los seres humanos se encuentran en el área de la inteligencia artificial que se ha constituido en una rama de las Ciencia de la Computación.

En el año 1903 se inventaron las válvulas a vacío y posteriormente se inventa el computador ENIAC que contenía 18000 de estas válvulas, de la cual se dice que es la primera máquina inteligente diseñada por el hombre[6]. En el 1937 el

matemático inglés Alan Mathison Turing realizó un estudio de gran relevancia donde introdujo la Máquina de Turing, además formalizó el concepto de algoritmo. Con esta máquina pudo probar que existían problemas sin solución, de los que ningún computador será capaz de dar un resultado. También se le considera el padre de la Inteligencia Artificial, por su famosa Prueba de Turing, que permitiría comprobar si un programa de ordenador puede ser tan inteligente como un ser humano, además Alan Turing tenía la premisa de que si durante el intercambio entre una computadora y el usuario, este último cree que está intercambiando con otro humano, entonces se dice que el sistema es inteligente[7]. En el año 1956 se acuñó el término "inteligencia artificial" en Dartmouth, durante una conferencia convocada por John McCarthy. En este campo de investigación se prosiguió trabajando y en 1980 los japoneses tuvieron un gran auge con los sistemas expertos.[8]

El estudio y la investigación en este campo ha alcanzado un avance considerable en cuanto al desarrollo y aplicación de técnicas y algoritmos inteligentes, etc. El objetivo principal de la I.A es que máquinas inteligentes puedan simular las habilidades y acciones físicas de los seres humanos no para su sustitución sino para su empleo en áreas de trabajo, máquinas capaces de obtener conocimiento, resolver problemas, predecir eventos a suceder, clasificar, evaluar información, entender símbolos, entre otros. Así de esta forma las acciones son más rápidas, precisas y sin riesgos humanos si el trabajo que se realiza es peligroso.

En la inteligencia artificial confluyen varios campos entre los que se pueden nombrar los citados por[9][10][11]:

- Aprendizaje Automático (Machine Learning).
- Ingeniería del conocimiento (Knowledge Engineering)
- Lógica difusa (Fuzzy Logic)
- Redes neuronales artificiales (Artificial Neural Networks)
- Sistemas reactivos (Reactive Systems)
- Sistemas Multiagente (Multi-Agent Systems)
- Sistemas basados en reglas (Rule-Based Systems)
- Razonamiento basado en casos (Case-Based Reasoning)

- Sistemas expertos (Expert Systems)
- Redes Bayesianas (Bayesian Networks)
- Computación evolutiva (Evolutionary Computation)
- Estrategias evolutivas
- Algoritmos genéticos (Genetic Algorithms)
- Técnicas de Representación de Conocimiento
- Redes semánticas (Semantic Networks
- Frames
- Visión artificial
- Audición artificial
- Lingüística computacional
- Procesamiento del lenguaje natural (Natural Language Processing)
- Minería de datos (Data Mining)

Como se dijo, la Inteligencia Artificial es utilizada en la actualidad en disímiles campos, convirtiéndose en los últimos años en una rama emergente de las Ciencias de la Computación, sobre la cual se dirigen grandes esfuerzos de la comunidad científica del mundo. Su aplicación penetra las más diversas esferas de la actividad social como en la robótica, industria mecánica, astronomía, medicina, entre otras; las cuales han reportando beneficios en las solución de estos problemas donde es necesaria principalmente la manipulación de información; además se ocupa de la representación, adquisición y procesamiento de conocimientos de forma automatizada.[9][10]

Las técnicas de inteligencia artificial se asemejan al comportamiento inteligente; es decir cuando se aplican para resolver problemas de toma de decisiones similares a las que toman los seres humanos.

Pero formalmente ¿Qué es Inteligencia Artificial? existen numerosas definiciones para el término pero en realidad no existe aún ninguna definición única que resuma adecuadamente todos los aspectos que el término quiere representar. Según [9] se define como "…*el estudio de cómo lograr que las computadoras hagan cosas que por el momento, las personas hacen mejor. Una computadora encuentra las raíces de una ecuación mucho más rápido y*

con mayor exactitud que un hombre; sin embargo, el hombre reconoce mucho mejor un conjunto de caracteres". Otra definición es la propuesta por [11] que la define *"como el estudio de las computaciones que permiten percibir, razonar y actuar".*

1.2.1 Aprendizaje automático

Como antes se mencionó en la I.A existen varios campos de investigación. De ellos resalta gran número de investigaciones realizadas con aprendizaje automático, conocido en inglés por *machine learning*.

¿Qué es aprendizaje automático? Se define el aprendizaje como elemento fundamental de cualquier sistema que se considere inteligente y su clasificación puedes ser: supervisado y no supervisado[12]. El aprendizaje supervisado parte de un conjunto de entrenamiento para el aprendizaje que consta de entradas y salidas (atributos y clases) y el no supervisado no requiere conocer la clase a predecir (salidas), el propio sistema hace un procesamiento organizativo para identificar grupos de datos que bajo características comunes (atributos) se agrupan en una misma clase.

Este campo de investigación comprende métodos, mecanismos y reglas, mediantes las cuales los computadores pueden aprender, razonar con gran rapidez semejándose a las de un ser humano a la hora de reconocimiento de patrones, imágenes, interpretación del lenguaje natural, toma de decisiones. Esta última permite a través de técnicas de clasificación predecir, pronosticar y clasificar según criterios y métodos grupos de objetos, animales, personas etc. que pertenecen a distintas categorías o clases especificas.

Cada uno de los grupos clasificados poseen características, es decir atributos, los cuales ofrecen la información que los particulariza y diferencia de otros grupos; todos estas características permiten categorizarlos en diferentes clases, quedando así una conjunto de datos.

Cada vez más, el aprendizaje automático está siendo utilizado en la solución de problemas prácticos de diferentes áreas. En el caso particular de la medicina su utilización es bastante frecuente. Un nicho significativo lo ha

constituido la problemática del diagnóstico médico en el cual varias soluciones basadas en aprendizaje automático han sido presentadas[13–22].

1.3 Técnicas de clasificación. Caracterización general.

Según la literatura consultada existen varias técnicas de clasificación y alguna de ellas con buenos resultados en el pronóstico médico[3].

Teniendo en cuenta la literatura consultada y los resultados preliminares alcanzados en la investigación se describen las 3 técnicas de clasificación que tuvieron mejor desempeño en problemas similares al tratado en esta investigación. Estas técnicas son: Perceptrón Multicapas (MLP), Árboles de decisión j48 e IBK (K-ésimo vecino más cercano).

1.3.1 Redes neuronales artificiales

Las redes neuronales (RNAs) o sistemas conexionistas como también se les llama, se han desarrollado en analogía con el funcionamiento de las redes biológicas del cerebro humano. Se constituyen en técnicas de inteligencia artificial que se han empleado en un amplio número de investigaciones de gran impacto en la actualidad.

Las redes neuronales, vistas como un área de la ingeniería del conocimiento, pueden situarse dentro de esta forma de tratar el problema de modelado. Son una herramienta utilizada en la resolución de problemas reales de gran complejidad y de diferentes áreas, como por ejemplo, el reconocimiento de patrones, diagnóstico médico, modelado de datos financieros, predicción de datos de crecimiento microbiano, etc .

La primera red neuronal artificial fue desarrollada por Warren McCulloch y Walter Pitss en el año 1943. Estos sistemas de procesamiento de información están conformados por neuronas conectadas con otras mediante enlaces de comunicación; ya que la primera neurona emite una salida hasta la conexión de otras neuronas[23]. Estos enlaces tienen un peso que contiene la información de influencia de una neurona con otra. Si este peso tiene un valor positivo corresponde a una relación de excitación fuerte y si el valor del peso es

negativo corresponde a una separación débil. Cada neurona ajusta las influencias de los enlaces recibidos de las entradas en una influencia total, todo esto se hace continuamente según una función de activación, esta pasa estrictamente la suma de los valores de las entradas a través de una función umbral para establecer la salida de las neuronas.[24]

Los factores de peso en los enlaces entre neuronas pueden ser adaptables o fijos, los adaptables recurren a un orden de enseñanza para convenir el valor de la fuerza de la conexión de las neuronas; además son básicos si no se conoce inicialmente cual deberá de ser su valor correcto. Si las neuronas emplean un peso fijo, su trabajo deberá estar anticipadamente definido. Los pesos serán definidos a partir de una representación completa del problema[24].

La figura 1 muestra la similitud de un modelo de red biológica y una artificial.

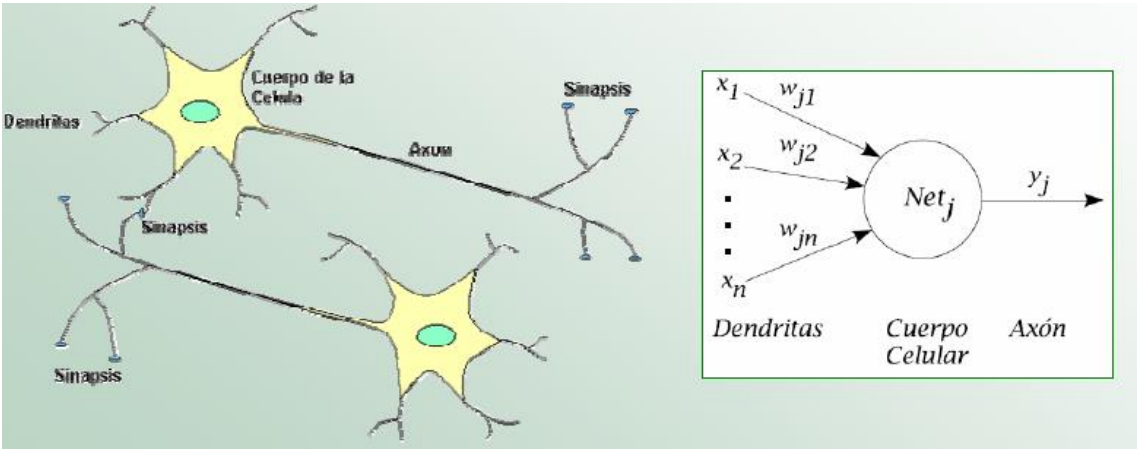

Figura 1 similitud de un modelo de red biológica y una artificial.[25]

Las redes neuronales se congregan en bloques estructurales los cuales se nombran capas. Estas se dividen en tres tipos de capas, la primera es la capa de entrada la cual conforma el primer bloque; el segundo bloque esta compuesto por una o varias capas ocultas y el tercer bloque contiene la capa de salida. La capa de entrada esta formada por las neuronas que tienen la tarea de introducir los patrones de entrada en la red, la cual toman de los datos o indicaciones procedentes del entorno del problema a resolver. Además en este primer bloque no se efectúa ningún procedimiento.

Las capas ocultas están formadas por aquellas neuronas cuyas entradas provienen de capas anteriores y cuyas salidas pasan a neuronas de capas posteriores.

La capa de salida es aquella que está formada por neuronas cuyos valores de salida corresponden al valor del resultado calculado por toda la red.

Como conclusión se puede decir que la estructura de una red neuronal de tipo multilayer perceptron esta formada por tres tipos de capas.

Las RNAs presentan distintos tipos de arquitectura en correspondencia con la estructura de las capas. Pueden ser monocapas y multicapas. Las redes monocapas son aquellas formadas por una sola capa y las multicapas su formación esta dada por varias capas. Atendiendo al flujo de la información (datos) podemos comentar de redes recurrentes (feedback) o realimentadas donde la información puede fluir entre las capas en cualquier sentido; y las redes unidireccionales (feedforward) las cuales la información fluye en un único sentido de las neuronas de entrada a las neuronas de salida.

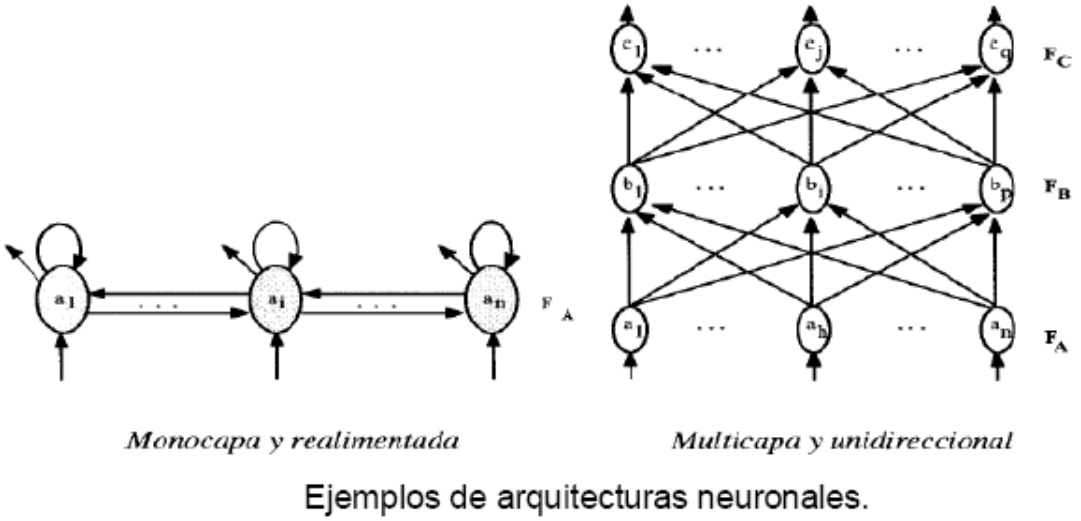

Figura 2 Ejemplos de arquitectura de redes neuronal [26]

La topología de las redes neuronales presenta diferentes tipos, los cuales pueden ser de acuerdo a sus enlaces y a la distribución espacial entre las neuronas. En un sistema de RNAs las conexiones o enlaces son direccionales, la información solamente puede propagarse en un solo sentido.

A continuación se muestra en la siguiente imagen el diseño de una red neuronal de tres capas:

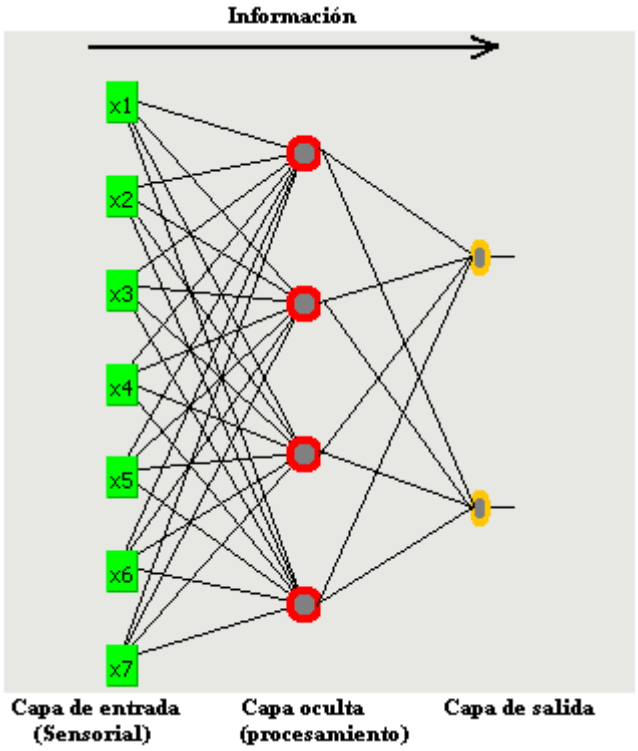

Figura 3 Red neuronal

Una vez desarrollado la arquitectura y composición de una red neuronal artificial veremos el proceso de aprendizaje.

En este proceso se le proporciona a la red una muestra de casos aleatorios del conjunto de datos disponibles, y se ajustan los pesos de manera que cada estímulo (entrada) produzca la salida deseada.

Utilizando aprendizaje supervisado, la exactitud de los resultados depende del conocimiento por la red de la información de una muestra de patrones, de la salida u objetivo que se quiere, además de la precisión en los pesos hasta que su salida sea la esperada. De esta forma la red es competente de precisar las entradas y salidas sin necesidad de plantear una representación funcional de partida.

En el aprendizaje no supervisado se presentan a la red un conjunto de casos sin asociar a la respuesta que se desea. La red, por medio de la regla de aprendizaje, estima la función: densidad de probabilidad, a partir de lo cual puede reconocerse las regularidades en el conjunto de entradas, extraer rasgos, o agrupar patrones según sus similitudes (clustering).

Sin embargo las respuestas finales de las redes neuronales dependen de las condiciones iniciales del entrenamiento de la red, ya que si el entrenamiento es hecho con datos errados, o con escasos juegos de datos, los resultados finales pueden no ser correctos, ni los esperados para darle una solución precisa y eficiente al problema propuesto.

Se conocen alrededor de cincuenta tipos de modelos de redes neuronales artificiales: Perceptrón, Perceptrón Multicapa, Mapa de Kohonen, ADALINE, MADALINE, Hopfield, Máquina de Boltzmann, entre otras que han sido utilizadas en varias soluciones prácticas como las publicadas en [27–37]

1.3.1.1 Modelo de red neuronal artificial. Perceptrón Multicapas.

Uno de los primeros modelos de redes neuronales fue el Perceptrón el cual surge en el año 1962 su creación se debe a Rosemblatt. 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[9].

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 divulgaron un estudio realizado al respecto. Todo estos trabajos investigativos permitieron llegar a la idea de adicionarle a la red capas ocultas, surgiendo así un nuevo modelo de red neuronal artificial, la arquitectura del perceptrón multicapa (MLP). Esta red se caracteriza porque cada neurona oculta o de salida esta diseñada para realizar dos cálculos: la dispersión de la señal hacia delante (forward pass) y el vector gradiente necesario para la retro-propagación *(*backward pass*)*[24]. El proceso de

entrenamiento del MLP se desarrolla mediante un algoritmo de retropropagación (backpropagation) de errores o variantes, este es el algoritmo que resuelve el problema de adiestrar las neuronas de las capas ocultas. El algoritmo permite a la red aprender de un juego de datos predefinido de pares de entradas y salidas proporcionadas como ejemplo. En el caso de de la aplicación de un patrón de entrada como estímulo para la primera capa de la red esta información se va propagando por todas las capas hasta generar la salida. La señal de salida se compara entonces con la salida deseada y se determina el error para cada neurona de salida. Si existiera más de una capa oculta este cálculo de error no se puede hacer directamente en las neuronas de la capas ocultas, solo se puede hacer si hay una sola capa intermedia (capa oculta). Lo que se hace en caso de haber más de una capa intermedia es propagar hacia atrás el error, llamado este proceso retropropagación de errores. Estos errores se envían hacia las neuronas de las otras capas, partiendo de la capa de salida. Cada neurona que tuvo algún error, propaga hacia todos las neuronas de la capa oculta, que aporten directamente a su salida, una fracción del error ocurrido durante el proceso. Esto no sólo se hace para una neurona, sino que se realiza la misma operación para todas las neuronas de la capa de salida. De esta manera, cada neurona de las capas intermedias recibirá una fracción de todos los errores que ocasionó su salida en la propagación anterior. Todo este proceso se repite en cada capa hasta que todas las neuronas de la red hayan recibido un valor de error.

Esto permite apoyarse en el valor de error percibido, la actualización de los pesos de conexión de cada neurona, y que la red pueda codificar todas las tramas de entrenamiento.

Este algoritmo presenta una deficiencia a la hora de minimizar la búsqueda de la función de error, pues puede obtener un mínimo local o algún punto estacionario perturbado o encontrar el mínimo global de la función de error. Aunque necesariamente no siempre tiene que alcanzarse el mínimo global en todas las aplicaciones, se puede preestablecer un error mínimo.

De la aplicación y experimentación de la retropropagación en diferentes problemas resultó que este es capaz de enfrentar problemas de clasificación de gran envergadura, de manera eficaz y relativamente con facilidad. De forma

global se ha establecido que el MLP es una red con grandes ventajas si el entrenamiento se desarrolla con rigurosidad y tratando en todo momento del proceso de reducir al mínimo el error; llevando todo esto a obtener un buen entrenamiento que posibilite una clasificación fiable y deseada.

Para muchos investigadores el MLP es la red neuronal más idónea para solucionar problemas de clasificación supervisada.

1.3.2 Árboles de decisión.

Los procesos de clasificación de casos, fenómenos naturales, enfermedades, entre otros, han sido abordados con el empleo de técnicas estadísticas o modelos probabilísticos y más recientemente en la perspectiva de la inteligencia artificial con el uso de los árboles de decisión. A continuación nos centraremos en esta técnica de clasificación dentro de la I.A.

La propuesta de estas técnicas fueron hechas en sus inicios por Hunt, Marin y Stone [40]. Este campo fue tomado por otros investigadores realizando estudios notables que se enriquecieron de nuevas teorías y definiciones en las que se destaca Quinlan, Breiman, Friedman, Olshen [40].

Los árboles de decisión son sistemas de aprendizajes inductivos como también se les conoce. Se entiende como aprendizaje inductivo: *"…es aquel aprendizaje que parte de casos particulares (experiencias) y obtiene casos generales (modelos o reglas)"* [40].
Los sistemas de aprendizaje inductivos son aplicados para extraer información, clasificar y predecir, posibilitando la ayuda a los expertos en el análisis de conocimiento acumulado en conjuntos de datos.

EL algoritmo ID3 es el algoritmo básico para la obtención del árbol de decisión[12]. Mediante este algoritmo, la conformación de árboles de decisión se realiza mediante un proceso recursivo. Su estructura es un grafo que empieza con el nodo raíz y termina en las hojas. El nodo raíz tiene nodos descendientes, de los cuales pueden descender otros nodos. Todos los nodos que no tengan descendientes son las hojas, las cuales portan el valor o

nombre de las clases. La estructura del árbol se comienza con la búsqueda del nodo raíz el cual es atributo más revelador de información de la conjunto de datos con respecto a las clases. Esto se determina mediante el cálculo de entropía cuanto más pequeño sea este valor menor será el grado de incertidumbre que aporta el atributo:

$$G(Q,A) = Entropía(Q) - \sum_{v \in Valores(A)}^{n} - \frac{|Qv|}{|Q|} Entropía(Qv)$$

Entropía indica el valor de entropía y se calcula por la siguiente fórmula

$$Entropía(Q) - \sum_{i=1}^{n} - p_i \log_2 p_i$$

n: cantidad de clases

Q: conjunto de casos.

p_i : cantidad de casos de Q que pertenecen a la clase i

A: atributo

valores(A): todos los posibles valores que puede tomar el atributo A

Q_v: subconjunto de Q cuyas instancias tienen el valor v en el atributo A

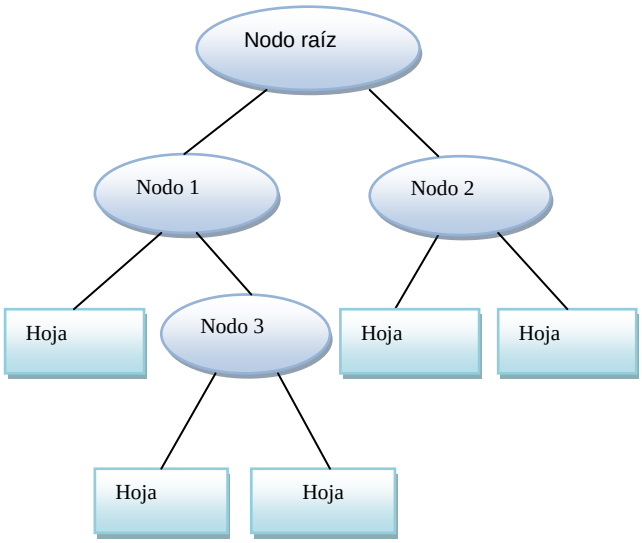

Gráfico 1 Ejemplo de un árbol

Una vez conformado un árbol de decisión se extraen las reglas del mismo, las cuales pueden ser empleadas para clasificar nuevos casos en el sistema de clases establecido.

Posteriores modificaciones al ID3 han generado variantes más sofisticadas de algoritmos que generan árboles de decisión. Entre las más conocidas está la variante C4.5[12]

1.3.2.1 Algoritmo C4.5

EL algoritmo C4.5 fue propuesto por Ross Quinlan en Programs for Machine Learning. Morgan Kaufmann Publishers, 1993. Este algoritmo basado en el ID3 presenta un mecanismo de post-poda el cual permite la formación de un árbol más simple para el entendimiento. Es capaz de extraer información en grandes conjuntos de datos compuestas por atributos continuos y categóricos [40]. La estrategia de búsqueda que utiliza es primero en profundidad (depth-first).

Entre las mejoras que incluye con respecto a su antecesor se encuentran:

- Permite utilizarse tanto con atributos continuos como con atributos discretos. C4.5 crea un umbral y luego se divide la lista en aquellos cuyo valor de atributo es superior al umbral y los que son menores o iguales a él.
- Puede utilizarse aun cuando hayan valores de atributos faltantes - C4.5 permite valores de los atributos para ser marcado como para faltantes. Los valores faltantes de los atributos simplemente no se usa en los cálculos de la ganancia y la entropía. [40]

Una de las implementaciones más utilizadas de este algoritmo es la conocida como J48. Esta es una implementación open source en lenguaje de programación Java en la herramienta Weka que se describe en este documento.

1.3.3 IBK: EL vecino más cercano.

En la actualidad existe un numeroso grupo de clasificadores, entre los que se encuentran los *Lazy***,** que en español significa holgazán o perezoso. Este grupo esta compuesto por varias técnicas, dentro de las cuales encontramos

de IB1, IBK, kStart entre otras. Nos centraremos en el IBK, esta técnica se conoce también como el vecino más cercano. Su proceso de clasificación se basa en clasificar un nuevo caso comparándolo con los casos más cercanos mediante la distancia euclidiana.

$$d_E(i,j) = \sqrt{(X_{i1} - X_{j1})^2 (X_{i2} - X_{j2})^2}$$

En el funcionamiento del algoritmo de clasificación de vecinos más cercanos se parte de un conjunto de datos de X casos, los cuales pertenecen a clases disjuntas. Se tiene un nuevo caso P para clasificar. Se determinan los K vecinos más cercanos al nuevo caso a clasificar mediante la distancia euclidiana, siendo el número K ajustable a partir de pruebas preliminares que permitan determinar cuál es el valor de k más apto para la clasificación.

Se le asignará el nuevo caso (P) a la clase que más se repita entre los k vecinos más cercanos.

Ejemplo:

Clases: puntos verde ( ) y puntos azules ( )

K = 5

X = 15

P = ?

P es el punto Rojo ( )

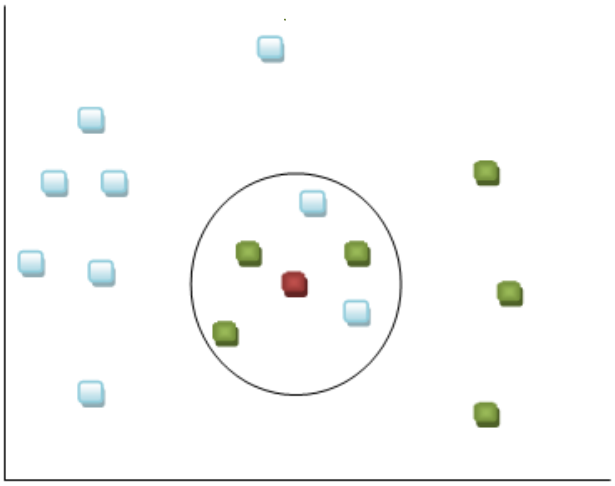

Gráfico 2 Ejemplo de búsqueda del IBK

Como se nota en la figura, los K vecinos de P son 2 azules y 3 verdes, siendo estos los más predominantes; por lo que P se clasifica como verde. Si

ocurriese el caso que el número de vecinos fuera par y existiera la misma cantidad de casos en cada clase, entonces hay que aplicar una regla heurística para el rompimiento de este equilibrio. En este caso se puede hacer la elección de el vecino más cercano o también escoger la clase con las distancia media menor [41].

1.4 Selección de Rasgos

Cuando un proceso está caracterizado por una gran cantidad de atributos entre los cuales pueden existir algunos que aportan información redundante o que no tienen una incidencia significativa sobre la variable dependiente del proceso, la selección de los rasgos esenciales aporta beneficios en el análisis de la información con el propósito de clasificar a los individuos en las clases definidas por las categorías de la variable dependiente. Este proceso puede incidir en el ahorro de tiempo y recursos por medición de variables irrelevantes o redundantes, además de eliminar variables de confusión que pueden afectar la calidad de la clasificación.

La selección de rasgos incluye dos componentes fundamentales: el método de búsqueda y la función de evaluación.

La función de evaluación permite calcular la cantidad de subconjuntos de rasgos, mientras que el método de búsqueda, por lo general heurístico, es el encargado de generar los subconjuntos de rasgos.

Entre las funciones de evaluación utilizadas en los métodos de selección de rasgos dos de las más utilizadas son la implementación *CfsSubsetEval* y el de componentes principales.

1.4.1 CfsSubsetEval

La función de evaluación *CfsSubsetEval* se basa en calcular la correlación de la clase con cada atributo existente en el juego de datos, y eliminar los atributos que presentan una correlación muy alta como atributos redundantes. Según este método los subconjuntos elegidos son aquellos

altamente correlacionados con el atributo que precisa las clases y con poca correlación entre ellos[15].

1.4.2 PrincipalComponents (Componentes principales)

La función de evaluación de componentes principales consiste en la transformación de p variables en p combinaciones lineales, llamadas componentes principales, que se caracterizan porque la primera componente explica la mayor parte de la variabilidad de conjunto de datos, la segunda componente explica la mayor parte de la variabilidad restante y así sucesivamente hasta la componente p-ésima. Estas componentes constituyen variables artificiales, además de ser ortogonales entre sí. Las componentes se obtienen por transformaciones algebraicas a partir de la diagonalización de la matriz de correlaciones de las variables originales. Las p componentes explican el 100 % de la variación total del conjunto de datos. En múltiples aplicaciones se utilizan solo las primeras componentes que logran explicar un porcentaje alto de variabilidad total.

El software Weka para realizar una selección de rasgos tiene la opción del empleo de una función de evaluación, un método de búsqueda o de manera personalizada por el especialista.

1.5 Computación granular

La computación granular, conocida por su acrónimo en inglés como GrC, se definió en el resumen de la conferencia IEEE-GrC'2006 como una teoría de cómputo general para el uso efectivo de gránulos, tales como clases, clústeres, subconjuntos, grupos e intervalos para construir un modelo computacional eficiente para aplicaciones complejas con una gran cantidades de datos, información y conocimiento[42].

La Computación granular se centra en la solución de problemas basado en conceptos de sentido común de gránulos, vista granulada, granularidad y jerarquía[43].

Cada vez más, la computación granular juega un rol muy importante en la bioinformática, los negocios electrónicos, seguridad, aprendizaje automático, minería de datos, entre otras áreas en términos de eficiencia, efectividad, robustez e incertidumbre[42].

Esencialmente la computación granular es un estudio interdisciplinar y multidisciplinar que emerge de muchas disciplinas desde la filosofía y la psicología hasta la inteligencia artificial. (Granular computing, past, present and future).

1.5.1 Granularidad

La experiencia en la actividad investigativa que se basa en el empleo de relaciones entre variables, permite establecer que en muchas ocasiones pequeñas diferencias dadas por las mediciones son irrelevantes para obtener una buena clasificación y que la granulación conveniente de los atributos puede expresar el contenido esencial del sistema de relaciones, con beneficios en el proceso de clasificación.
La granularidad incluye la construcción de 3 componentes básicos: Los gránulos, la vista granular y la jerarquía.

Un gránulo es interpretado como una de las numerosas partículas pequeñas que forman una unidad mayor.

La vista granular, también conocida como nivel consiste de entidades (gránulos) cuyas propiedades caracterizan y describen al elemento en estudio.

La jerarquía se define como el orden en que se establecen los niveles o vistas granulares. De modo que en un sentido se va de un nivel con mayor granularidad a otro de menor granularidad[43].

Además, la granularidad también incluye dos operaciones básicas que son la combinación (unión de gránulos para conformar un gránulo a otro nivel) y la

descomposición (la separación de unidades de un gránulo para formar nuevos gránulos a otro nivel).

Desde esta perspectiva es necesario establecer los criterios y métodos de granulación según el problema a tratar.

Ejemplos de investigaciones que se basan en la computación granular obteniendo resultados satisfactorios son [44–52]

1.6 Conclusiones parciales del capítulo

El avance de la medicina en Cuba es una verdad innegable que se consolida cada vez más por su capital humano y la utilización de nuevas tecnologías. La inteligencia artificial y más específicamente el aprendizaje automático se muestran como alternativas factibles y eficientes en la solución de problemas de clasificación que han tenido gran notoriedad en el pronóstico en áreas tan sensibles como la salud.

Según la bibliografía consultada, las técnicas de clasificación más utilizadas en problemas similares al tratado en esta investigación son las redes neuronales artifíciales, particularmente el perceptrón multicapa, los árboles de decisión, en particular el C4.5 y las llamadas técnicas “lazy”, particularmente la del k-ésimo vecino más cercano.

La selección de rasgos aporta una premisa de valor en la preparación de los datos para la aplicación de las técnicas de clasificación. Esto toma mayor connotación desde la perspectiva de la computación granular al permitir diferentes niveles de granularidad de los datos que influyen en la clasificación.

Capítulo 2 Preprocesamiento de los datos

En este capítulo se presentan las herramientas empleadas en el desarrollo de la investigación. Se precisan los resultados de las selecciones de rasgos realizadas a las variables originales y de la granulación para la aplicación posterior de las técnicas de aprendizaje automático: MLP, árbol de decisión C4.5 (con su implementación J48) y el KNN (con su implementación IBK). Además se ajustan los parámetros de cada una de estas técnicas.

2.1 Conjuntos de datos

En esta investigación se utiliza la compilación de datos de pacientes ingresados con traumas graves de cráneo en la sala de cuidados intensivos del hospital clínico quirúrgico Gustavo Aldereguía Lima de Cienfuegos. Esta compilación cuenta con 40 variables recogidas durante varios años. A partir de la misma se conformaron varios conjuntos de entrenamientos. Estos conjuntos están construidos para obtener el pronóstico en cada uno de cuatro momentos del tiempo en que se actualiza la información del comportamiento de las variables que caracterizan el estado de los pacientes, es decir al ingreso, 24 horas, 48 horas y 72 horas.

Las variables registradas son:

- Edad,
- Sexo,
- TAC(Escala de coma de marshall)
- ECG(Escala de Coma de Glasgow),
- Alteraciones pupilares,
- TAS(Presión arterial sistólica),
- Hipoxemia,
- PIC (Presión intra craneal),
- PPC (Presión de perfusión cerebral),
- IPR (Índice del pulso del líquido cefalorraquídeo sobre respiración),
- Glicemia,
- Osmolaridad,
- Egreso.

En el caso de las variables Edad, Sexo y TAC estás tienen una única medición para cada instancia, lo mismo sucede para la variable egreso que en este caso será la variable dependiente y que define las clases en la clasificación. El resto de las variables (9 en total) se miden en cada uno de los cuatros momentos ya indicados. Esto hace que realmente sean 36 variables que junto a las 3 de única medición y la variable que representa las clases completan las 40 variables.

Para nombrar las variables iniciales que se transforman en 4 variables según los momentos de medición, se le agrega a estas, según sea el caso, la letra I (Ingreso) o los números 24, 48 o 72 para las horas posteriores al ingreso. Por ejemplo: PICI, PIC24, PIC48 y PIC72 representan la presión intracraneal al ingreso y a las 24 horas, 48 horas y a las 72 horas de ingresado el paciente.

A continuación se detallan los conjuntos de datos conformados para la experimentación:

- Al Ingreso: es un conjunto constituido solo por variables medidas al ingresar el paciente. (13 variables)

- A las 24 horas: este conjunto está conformado por las variables medidas al ingreso y también por las medidas a las 24 horas de estar el paciente en el hospital. (22 variables)

- A las 48 horas: este conjunto está formado por las variables de ingreso, las variables medidas a las 24 horas y las medidas a las 48 horas. (31 variables)

- A las 72 horas: este último conjunto está conformado por todas las variables medidas en los cuatro momentos que se han considerado. (40 variables)

A continuación se presentan las variables seleccionadas en cada uno de los conjuntos de datos:

Tabla 1Variables a utilizar en cada conjunto de datos

Al ingreso	A las 24 horas	A las 48 horas	A las 72 horas
1. Edad 2. Sexo 3. ECGI 4. Alt. PupilaresI 5. TASI 6. HipoxemiaI 7. PICI 8. PPCI 9. IPRI 10. GlicemiaI 11. OsmolaridadI 12. TAC 13. Egreso	1. Edad 2. Sexo 3. ECGI 4. ECG24 5. Alt. PupilaresI 6. Alt. Pupilares24 7. TASI 8. TAS24 9. HipoxemiaI 10. Hipoxemia24 11. PICI 12. PIC24 13. PPCI 14. PPC24 15. IPRI 16. IPR24 17. GlicemiaI 18. Glicemia24 19. OsmolaridadI 20. Osmolaridad24 21. TAC 22. Egreso	1. Edad 2. Sexo 3. ECGI 4. ECG24 5. ECG48 6. Alt. PupilaresI 7. Alt. Pupilares24 8. Alt. Pupilares48 9. TASI 10. TAS24 11. TAS48 12. HipoxemiaI 13. Hipoxemia24 14. Hipoxemia48 15. PICI 16. PIC24 17. PIC48 18. PPCI 19. PPC24 20. PPC48 21. IPRI 22. IPR24 23. IPR48 24. GlicemiaI 25. Glicemia24 26. Glicemia48 27. OsmolaridadI 28. Osmolaridad24 29. Osmolaridad48 30. TAC 31. Egreso	1. Edad 2. Sexo 3. ECGI 4. ECG24 5. ECG48 6. ECG72 7. Alt. PupilaresI 8. Alt. Pupilares24 9. Alt. Pupilares48 10. Alt. Pupilares72 11. TASI 12. TAS24 13. TAS48 14. TAS72 15. HipoxemiaI 16. Hipoxemia24 17. Hipoxemia48 18. Hipoxemia72 19. PICI 20. PIC24 21. PIC48 22. PIC72 23. PPCI 24. PPC24 25. PPC48 26. PPC72 27. IPRI 28. IPR24 29. IPR48 30. IPR72 31. GlicemiaI 32. Glicemia24 33. Glicemia48 34. Glicemia72 35. OsmolaridadI 36. Osmolaridad24 37. Osmolaridad48 38. Osmolaridad72 39. TAC 40. Egreso

Inicialmente los conjuntos de datos establecidos fueron sometidos a un análisis exploratorio de datos utilizando el paquete estadístico SPSS, con la intención de realizar una limpieza de datos buscando casos anómalos o casos

perdidos. Luego de realizado el análisis se confirmó que no existían casos anómalos. En el caso de los conjuntos de datos para las 48 y 72 horas se detectaron valores perdidos. Estos valores perdidos se justifican por ser casos que en la variable Egreso en momentos de medición anteriores al del conjunto de datos tuvieron como valor fallecido por lo que en los conjuntos de datos posteriores no aparecen mediciones. En tales casos se eliminaron las instancias que contenían estos valores perdidos.

Teniendo en cuenta pruebas preliminares con los conjuntos de datos conformados y lo significativo de la computación granular en los problemas de clasificación, se decidió conformar otros conjuntos de datos similares pero con otros niveles de granularidad. Para la conformación de estos conjuntos de datos se establecieron 2 criterios de granularidad. El primer criterio está basado en el conocimiento médico. El segundo criterio está basado en técnicas estadísticas. A continuación se exponen los gránulos del nivel establecido según cada criterio de granularidad de las variables:

Criterio 1: Granulación de los datos a partir del conocimiento médico:
Para cada variante se definieron los gránulos (subgrupos) que puede asumir la variable a partir de la opinión del experto (especialista médico). .

Variable: Edad.
Subgrupo 1: edad **<** 60 años
Subgrupo 2: edad **≥** 60 años

Variable: Escala de Coma de Glasgow:
Subgrupo 1: 3 **≤** ECG **<** 6 puntos
Subgrupo 2: ECG **≥** 6 puntos

Variable: Alteraciones pupilares:
Subgrupo 1: si
Subgrupo 2: no

Variable: Presión arterial sistólica (TAS)

Subgrupo 1: TAS **≤** 90 mmHg

Subgrupo 2: TAS **>** 90 mmHg

Variable: Hipoxemia

Subgrupo 1: PaO_2 **<** 60

Subgrupo 2: PaO_2 **≥** 60

Variable: Presión intracraneal (PIC)

Subgrupo 1: 0 **≤** PIC **≤** 15

Subgrupo 2: 15 **>** PIC **<** 20

Subgrupo 3: PIC **≥** 20

Variable: Presión de perfusión cerebral (PPC)

Subgrupo 1: PPC **<** 70

Subgrupo 2: 70 **≤** PPC **≥** 90

Subgrupo 3: PPC **>** 90

Variable: Índice de pulso del líquido cefalorraquídeo sobre respiración (IPR)

Subgrupo 1: normal

Subgrupo 2: IPR **>** 2,5

Variable: Glicemia

Subgrupo 1: Glicemia **≤** 5,5

Subgrupo 2: Glicemia **>** 5,5

Variable: Osmolaridad

Subgrupo 1: normal

Subgrupo 2: Osmolaridad **>** 310

Variable: Escala de coma de marshall (TAC)

Subgrupo 1: LD I

Subgrupo 2: LD II

Subgrupo 3: LD III

Subgrupo 4: LD IV

Subgrupo 5: LFNE

Subgrupo 6: LFE

Una vez conformada los 4 conjuntos de datos siguiendo este criterio de granularidad se les nombró Modelos Especialistas y se diferencian según el momento de la medición al igual que los conjuntos de datos iníciales. Así

Criterio 2: Granulación basada en técnicas estadísticas. Este criterio consiste en granular las variables originales de modo que maximicen la relación con la variable dependiente (estado al egreso) lo cual se logra mediante la aplicación de la técnica de minería de datos detector automático de interacciones Chi-cuadrado (CHAID) con el empleo del paquete estadístico SPSS en su versión 18. Según este criterio se conforman nuevos conjuntos de datos cuyas variables son granuladas de la siguiente manera::

Variable: Edad se dividió en dos subgrupos:

Subgrupo 1: edad **≤** 34 años

Subgrupo 2: edad **>** 34 años

Variable: Escala de Coma de Glasgow se dividió en dos subgrupos:

Subgrupo 1: ECG **≤** 4 puntos

Subgrupo 2: ECG **>** 4 puntos

Variable: Alteraciones pupilares se dividió en dos subgrupos:

Subgrupo 1: si

Subgrupo 2: no

Variable: Presión arterial sistólica (TAS)

Subgrupo 1: TAS **≤** 115 mmHg

Subgrupo2: TAS **>** 115 mmHg

Variable: Hipoxemia

Subgrupo 1: PaO_2 **<** 60

Subgrupo 2: PaO_2 **≥** 60

Variable: Presión intracraneal (PIC)

Subgrupo 1: PIC **≤** 28

Subgrupo 2: PIC **>** 28

Variable: Presión de perfusión cerebral (PPC)

Subgrupo 1: PPC **≤** 72

Subgrupo 2: PPC **>** 72

Variable: Índice de pulso del líquido cefalorraquídeo sobre respiración (IPR)

Subgrupo 1: Normal

Subgrupo 2: IPR > 2,5

Variable: Glicemia

Subgrupo 1: Glicemia **≤** 5

Subgrupo 2: 5 **<** Glicemia **≤** 9,8

Subgrupo 3: Glicemia **>** 9,8

Variable: Osmolaridad

Subgrupo 1: Osmolaridad **≤** 298

Subgrupo 2: Osmolaridad **>** 298

Variable: Escala de coma de marshall (TAC)

Subgrupo 1: LD I y LFE

Subgrupo 2: LD II y LD III

Subgrupo 3: LD IV y LFNE

Una vez conformados estos conjuntos de datos se les nombró Modelos Óptimos.

De esta manera quedan construidos los 12 conjuntos de datos con los que se trabajará en la investigación.

2.2 Herramientas a utilizar.

Para el desarrollo de la investigación se utilizaron, esencialmente, 2 herramientas para el procesamiento de datos: Weka 3.52 y SPSS Versión 18.0.

2.2.1 Weka

Weka (*Waikato Environment for Knowledge Analysis*) incluye un conjunto de librerías JAVA que facilita el manejo de la información existente para el proceso de análisis y extracción de conocimientos contenidos en las conjuntos de datos, a través del empleo de una gran cantidad de técnicas de inteligencia artificial. Es un software de desarrollo libre, de código abierto que puede ser modificado. Además puede tener acceso a conjuntos de datos SQL mediante una conexión JDBC (Java Database Connectivity).

Weka es multiplataforma y de fácil manejo por parte del usuario a la hora de su utilización.

Su interfaz brinda varias opciones entre las que resaltan la posibilidad de exploración de técnicas con conjuntos de datos y la experimentación para la comparación entre las técnicas y entre los conjuntos de datos.

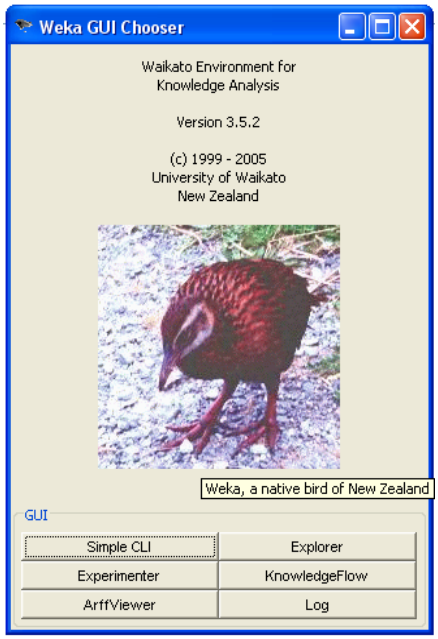

Figura 4 Interfaz inicial del Weka

Para procesar los conjuntos de datos en Weka, estos deben ser editados en el formato que soporta la herramienta (archivos .arff). Estos archivos tienen su propia estructura, tal y como se muestra en la figura 5.

```
%COMNETARIOS
@relation NOMBRE_DEL_CONJUNTO_DE_DATOS

@attribute variable1 real
@attribute variable2 real
...
@attribute valor1 numeric
@attribute valor2 numeric
...
@attribute num1 integer
@attribute num2 integer
...
@attribute Class1{X1,X2...,Xn}
@attribute Class2{S1,S2...,Sn}

@data
%conjunto de casos
variable1, variable2, valor1, valo2,num1,num2,X1,S1
.
.
.
variable1, variable2, valor1, valo2,num1,num2,X2,S2
```

Figura 5 Estructura de un archivo.arff

A continuación se muestra, a modo de ejemplo, en la figura 6 como quedaría conformada la información en el archivo Modelo_Ingreso_Normales.arff que contiene el conjunto de datos con la granularidad inicial y las variables medidas al ingreso.

```
@relation Modelo_Ingreso_Normales

@attribute Edad numeric
@attribute Sexo numeric
@attribute ECG_Ingreso numeric
@attribute Alteracion_Pupila_Ingreso numeric
@attribute TAS_Ingreso numeric
@attribute PaO2_Ingreso numeric
@attribute PIC_Ingreso numeric
@attribute PPC_Ingreso numeric
@attribute IPR_Ingreso numeric
@attribute Glicemia_Ingreso numeric
@attribute Osmolaridad_Ingreso numeric
@attribute TAC_Ingreso numeric
@attribute Class_Egreso {0,1}

@data
30,0,8,0,110,2,2,86,1,5.6,279,1,0
26,0,3,0,85,2,4,62,1,7.3,279,1,1
40,1,6,0,140,2,13,64,1,4.3,288,2,0
64,1,4,0,150,2,12,78,1,8.3,276,2,0
52,1,5,0,140,2,6,90,1,7.2,286,2,0
```

Figura 6 Ejemplo de un archivo.arff

La opción de exploración del Weka presenta varias opciones como: preprocesamiento, clasificación, agrupamiento (Clúster), asociación, selección de atributos y visualización.

Como se puede apreciar la figura 7, una vez abierto un conjunto de datos en WEKA, se obtiene una información detallada de los casos que presenta, precisando instancias, atributos y cantidad de instancias pertenecientes a cada clase a clasificar.

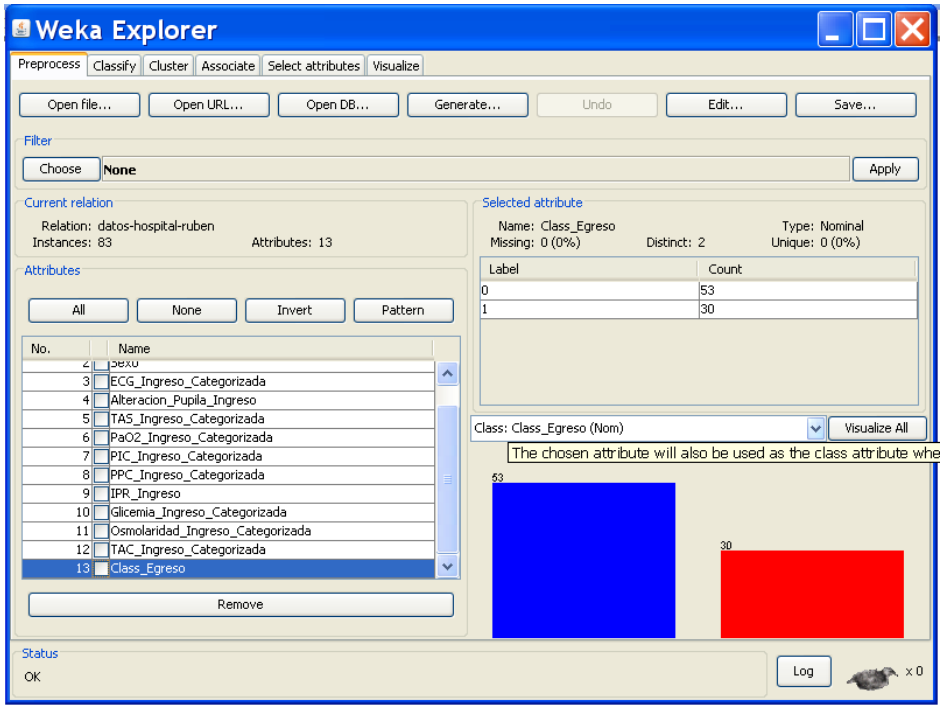

Figura 7 Interfaz de explorer del Weka

Cantidad de instancias: 83

Cantidad de atributos: 13

Cantidad de instancias por clases: 53 vivos y 30 fallecidos

Para los 3 conjuntos de 72 horas las cantidades de instancia, atributos y cantidad de instancias por clases es la misma. Solo se diferencian en el nivel de granularidad. En los otros modelos se presenta una situación similar, aunque la cantidad de casos difiere en los modelos según el tiempo de

medición ya que como se explicó con anterioridad se eliminaron casos perdidos en los conjuntos de 48 y 72 horas.

2.3 Selección de rasgos

Teniendo en cuenta lo significativo de la selección de rasgos en los problemas de clasificación se realizó un estudio para prescindir, en cada conjunto de datos, de los rasgos no determinantes en el proceso de clasificación, dejando solo los necesarios. Para este estudio se utilizó las facilidades del Weka en cuanto a la selección de rasgos.

El proceso de selección de rasgo se les aplicó a los conjuntos de 24, 48 y 72 horas. No se realizó a los conjuntos de ingreso por tener pocos atributos y la eliminación de uno o algunos de ellos no repercute significativamente.

Tabla 2 Selección de rasgos en los conjuntos a las 24 horas.

Conjunto de datos	Evaluador de atributo	Método de búsqueda	Cantidad atributos escogidos	Atributos escogidos	Nombre del nuevo conjunto
Modelo 24h Normales.arff	CfsSubsetEval	BestFirst –D 1 –N 5	10	4,6,8,10,11,14,16,18,20,21	Modelo 24h Normales RASGO 1.arff
	PrincipalComponents-R0.95	Ranker-T-1.7976931348623157E308-N-1	15	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15	Modelo 24h Normales RASGO 2.arff
Modelo 24h Especialista.arff	CfsSubsetEval	BestFirst-D1-N5	9	4,6,8,10,14,16,18,19,21	Modelo 24h Especialista RASGOS 1.arff
	PrincipalComponents-R0.95	Ranker-T-1.7976931348623157E308-N-1	16	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16	Modelo 24h Especialista RASGOS 2.arff
Modelo 24h Óptimos.arff	CfsSubsetEval	BestFirst –D 1 –N 5	9	4,6,8,10,12,14,16,18,21	Modelo 24h Óptimos RASGOS 1.arff
	PrincipalComponents-R0.95	Ranker-T-1.7976931348623157E308-N-1	16	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16	Modelo 24h Óptimos RASGOS 2.arff

Tabla 3 Selección de rasgos en los conjuntos a las 48 horas.

Conjunto de datos	Evaluador de atributo	Método de búsqueda	Cantidad atributos escogidos	Atributos escogidos	Nombre del nuevo conjunto
Modelo 48h Normales.arff	CfsSubsetEval	BestFirst-D1 –N 5	12	4,5,7,10,13,15, 20,22,25,26, 28,30	Modelo 48h Normales RASGO 1.arff
	PrincipalComponents –R 0.95	Ranker –T -1.7976931348623157E308 –N -1	17	1,2,3,4,5,6,7,8,9, 10,11,12,13,14,15,16,17	Modelo 48h Normales RASGO 2.arff
Modelo 48h Especialista.arff	CfsSubsetEval	BestFirst –D1-N 5	10	4,5,7,10,13, 17,20,22,26, 30	Modelo 48h Especialista RASGOS 1.arff
	PrincipalComponents –R 0.95	Ranker –T -1.7976931348623157E308 –N -1	19	1,2,3,4,5,6,7,8,9, 10,11,12,13,14, 15,16,17,18,19	Modelo 48h Especialista RASGOS 2.arff
Modelo 48h Óptimos.arff	CfsSubsetEval	BestFirst –D 1 –N 5	15	4,5,7,8,10,13,14, 15,19,20,22,24, 25,26,30	Modelo 48h Óptimos RASGOS 1.arff
	PrincipalComponents –R 0.95	Ranker –T -1.7976931348623157E308 –N -1	20	1,2,3,4,5,6,7,8,9, 10,11,12,13,14,15,16,17,18,19,20	Modelo 48h Óptimos RASGOS 2.arff

Tabla 4 Selección de rasgos en los conjuntos a las 72 horas.

Conjunto de datos	Evaluador de atributo	Método de búsqueda	Cantidad atributos escogidos	Atributos escogidos	Nombre del nuevo conjunto
Modelo 72h Normales.arff	CfsSubsetEval	BestFirst –D 1 –N 5	16	5,6,7,12,14,15,18,21,26,28, 32,33,34,36, 38,39	Modelo 72h Normales RASGO 1.arff
	PrincipalComponents –R 0.95	Ranker –T -1.7976931348623157E308 –N -1	19	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18, 19	Modelo 72h Normales RASGO 2.arff
Modelo 72h Especialista.arff	CfsSubsetEval	BestFirst –D 1 –N 5	11	5,6,10,12,15,18,21,26, 28,33,39	Modelo 72h Especialista RASGOS 1.arff
	PrincipalComponents –R 0.95	Ranker –T -1.7976931348623157E308 –N -1	20	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18, 19,20	Modelo 72h Especialista RASGOS 2.arff
Modelo 72h Optimos.arff	CfsSubsetEval	BestFirst –D 1 –N 5	15	5,6,10,12,16,18,20,24, 25,28,32,33,34,38,39	Modelo 72h Optimos RASGOS 1.arff
	PrincipalComponents –R 0.95	Ranker –T -1.7976931348623157E308 –N -1	22	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18, 19,20,21,22	Modelo 72h Optimos RASGOS 2.arff

Luego de la selección de rasgos realizada, se conformaron nuevos conjuntos a partir de los previamente conformados para las 24, 48 y 72 horas obteniéndose, como se pudo ver en las tablas 2, 3 y 4. Seis nuevos conjuntos por cada uno de los 3 momentos de medición. Finalmente se realiza el experimento con los 30 conjuntos de datos conformados.

2.4 Configuración de las técnicas de clasificación.

En estudios preliminares apoyados en la bibliografía consultada y en las facilidades de la opción Explorer del Weka, se probaron varias técnicas de clasificación con varias configuraciones de sus parámetros. Este estudio permitió identificar cuáles técnicas proporcionaban los mejores resultados, para definitivamente incluirlas en el experimento final a realizar. Estas técnicas fueron: La red neuronal artificial MLP, el árbol de decisión C4.5 y de las llamadas técnicas Lazy el k-ésimo vecino más cercano. Estas técnicas fueron explicadas en el capítulo anterior.

2.4.1 Parámetros considerados para el entrenamiento de la red neuronal MLP.

hiddenLayers: Define las capas ocultas de la red neuronal. Se define por medio de una lista de números enteros positivos, uno por cada capa oculta, y separados por coma que representan la cantidad de neuronas en la capa. Pueden utilizarse comodines para la cantidad de neuronas en cada capa:

Tabla 5 Define las capas ocultas

Comodín	Significado
i	cantidad de neuronas = cantidad de atributos
o	cantidad de neuronas = cantidad de clases
t	cantidad de neuronas = cantidad de atributos + cantidad de clases
a	cantidad de neuronas = (cantidad de atributos + cantidad de clases)/2

LearningRate: Define el valor de la tasa de aprendizaje. Es un valor numérico.

Momentum: impulso aplicado a los pesos durante la actualización.

nominalToBinaryFilter: procesa las instancias con filtro. Ayuda a mejorar el rendimiento de la red si hay atributos nominales en los datos. Se establece en true o false. Por defecto es true

normalizeAttributes: normaliza los atributos. Ayuda a mejorar el rendimiento de la red. Para esto la clase no tiene que ser numérica, también normaliza atributos nominales (estableciendo los valores entre -1 y 1). Se establece en true o false. Por defecto es true

normalizeNumericClass: normaliza la clase si ésta es numérica. Ayuda a mejorar el rendimiento de la red. Normaliza la clase entre -1 y 1. Este es un proceso interno, los valores mostrados son dados en los rangos originales. Se establece en true o false. Por defecto es true

randomSeed: inicializa un número aleatorio el cual es usado para ubicar los pesos iniciales en las conexiones entre los nodos.

trainingTime: establece la cantidad de épocas (tiempos) durante el entrenamiento. El valor es numérico. Por defecto tiene valor 500.

A partir de los estudios preliminares se determinó que la configuración de parámetros más adecuada para los conjuntos de datos a procesar es la siguiente:

learningRate: 0.3

momentum: 0.2

nominalToBinaryFilter: true

normalizeAttributes: true

normalizeNumericClass: true

randomSeed: 0

trainingTime: 500

En el caso del parámetro hiddenLayers se probaron varias configuraciones de redes neuronales variando la cantidad de capas y de neuronas. Los mejores resultados se centraron en 3 topologías que se describen más adelante.

2.4.2 Parámetros a considerar para la utilización del J48.

binarySplits: Indica si se utiliza separación binaria en atributos nominales cuando se construye el árbol. Toma como valor true o false. El valor por defecto es true.

confidenceFactor: Establece el factor de confianza utilizado para la poda del árbol (valores más pequeños incurren en más poda). Es un valor numérico. EL valor por defecto es 0.25

minNumObj: Indica el número mínimo de instancias por hoja. Es un valor numérico. El valor por defecto es 2.

nunFolds: Determina la cantidad de datos utilizados para la poda de reducción de error. Uno se utiliza para la poda y el resto para el crecimiento del árbol. Es un valor numérico. El valor por defecto es 3.

reducedErrorPruning: Establece si se utiliza poda de reducción de error o la poda del C4.5. Toma valor true (poda de reducción de error) o false (poda del C4.5). El valor por defecto es false.

saveInstanceData: Establece si se guardan los datos del entrenamiento para la visualización. Toma valor true o false. El valor por defecto es false.

Seed: Indica la semilla usada para la aleatoriedad de los datos cuando se utiliza la poda de reducción del error. Valor numérico. El valor por defecto es 1.

subtreeRaising: Establece si se considera la operación *raising* del sub-árbol cuando se está podando. Toma valor true o false. El valor por defecto es true.

Unpruned: Indica si la poda se llevará a cabo. O no. Toma valor true o false. El valor por defecto es false

A partir del estudio preliminar se estableció que la configuración que logra mejores resultados es la siguiente:

La configuración de los parámetros se estableció como se presenta a continuación.

binarySplits: false

confidenceFactor: 0.25

minNumObj: 2

nunFolds: 3

reducedErrorPruning: false

saveInstanceData: false

seed: 1

subtreeRaising: true

unpruned: false

2.4.3 Parámetros a considerar en la utilización del IBK

KNN: Permite establecer el número k de vecinos más cercano a tener en cuenta al clasificar. Es un valor entero positivo mayor que 0. El valor por defecto es 1.

crossValidate: Indica si se utilizará la técnica de validación cruzada hold-one-out para seleccionar el mejor valor de k. Toma valor true o false. El valor por defecto es false.

Distance Weighting: Establece el método de peso de distancia a utilizar. Toma los siguientes valores: No distance weighting (sin peso de distancia), weight by 1/distancia (Peso: 1/distancia) y weight by 1-distancia (Peso: 1-distancia). El valor por defecto es No distance weighting

mean Squared: Indica si será utilizado el error cuadrático medio o el error cuadrático medio absoluto. Toma valor true (será utilizado el error medio cuadrático) o false (se utilizará el error medio cuadrático absoluto).

nearestNeighbourSearchAlgorithm: Indica el algoritmo de búsqueda del vecino más cercano a utilizar. Los valores posibles son: KDTree (Método k-d tree) y LinearNN (Método linear para el vecino más cercano)

KNN: 1

crossValidate: False

distaceWeighting: No distance weighting

meanSquared: False

nearestNeighbourSearchAlgorithm: LinearNN

2.5 Indicadores para comparar resultados del experimento.

Varios son los indicadores que se utilizan en el análisis de los experimentos. La decisión de utilizarlos está supeditada a varios elementos como puede ser la naturaleza de lo que se desea demostrar.

Teniendo en cuenta la bibliografía consultada y la importancia que tiene la clasificación en el pronóstico médico por las consecuencias que de esta se derivan se utilizarán varios indicadores de los resultados de los modelos que se experimentan. Para una mejor comprensión de los mismos, se incluye una breve explicación:

Sea una tabla de contingencia de 2x2 donde se registra la clasificación de una muestra de **n** individuos en dos categorías: Negativo y Positivo (negativo**:** para pacientes pronosticados vivos y positivo: para pacientes pronosticados muertos) según la clasificación real y la clasificación que se realiza empleando un procedimiento determinado.

Tabla 6 Indicadores a comparar

Realidad	Pronóstico		Total
	Negativo (Vivo = 0)	Positivo (Fallecido = 1)	
Negativo (Vivo = 0)	a	c	$a+c=f_1$
Positivo (Fallecido = 1)	d	b	$d+b=f_2$
Total	$a+d=c_1$	$c+b=c_2$	n

Partiendo de estas clasificaciones tabuladas, se pueden definir las siguientes frecuencias absolutas:

a = **Número de verdaderos positivos (**son casos positivos que son clasificados como positivos**)**

c = **Número de falsos positivos (**son los casos negativos que son calificados como positivos**)**

a = **Número de verdaderos negativos (**son los caso negativos clasificados como negativos**)**

d = **Número de falsos negativos (**son los casos positivos clasificados como negativos**)**

Donde n = a+b+c+d

A partir de la cual se obtienen las frecuencias relativas de algunos indicadores que constituyen índices de la calidad de la clasificación, que se definen a continuación:

Frecuencia relativa de falso positivos (FRFP):

$$FRFP = \frac{c}{a+c}$$

Frecuencia relativa de falso negativo (FRFN):

$$FRFN = \frac{d}{d+b}$$

Frecuencia relativa verdaderos positivos (FRVP):

$$FRVP = \frac{b}{b+d}$$

Frecuencia relativa verdaderos negativo (FRVN):

$$FRVN = \frac{a}{a+c}$$

Porcentaje de clasificación correcta (%CC):

$$\%CC = \frac{a+b}{n} * 100\,\%$$

Coeficiente Kappa

El coeficiente Kappa de Cohen es una medida del acuerdo entre dos jueces o procedimientos, en este caso la clasificación real y la clasificación pronosticada según la técnica de I.A. empleada. Es un índice más adecuado que el porcentaje de buena clasificación debido que excluye a las coincidencias entre los jueces producidas por azar.

Según la tabla 6 la cantidad de casos concordantes entre realidad y pronóstico son la suma de los elementos en la diagonal principal, a+d.

Se define como la proporción observada de casos concordantes a $p_0 = \frac{a+d}{n}$.

La concordancia esperada por azar entre realidad y pronóstico es la que se obtiene considerando que realidad y pronóstico son independientes, es decir los casos que se deberían observar en la diagonal principal bajo el supuesto de independencia, a la cual definimos como frecuencia esperada. El cálculo de la

frecuencia esperada en la diagonal principal (cantidad de acuerdos por azar) se calcula sumando las frecuencias esperadas en cada casilla de la diagonal principal. La frecuencia esperada en una casilla de la diagonal principal se obtiene multiplicando el total marginal de la fila correspondiente por el total marginal de la columna correspondiente y dividiendo el producto por el total de la muestra. Según las notaciones de la tabla 6 la frecuencia esperada en la primera casilla de la diagonal principal es $\frac{f_1 \cdot c_1}{n}$ y en la segunda casilla de la diagonal principal es $\frac{f_2 \cdot c_2}{n}$. La suma de las frecuencias esperadas en la diagonal principal es $\frac{f_1 \cdot c_1}{n} + \frac{f_2 \cdot c_2}{n} = \frac{f_1 \cdot c_1 + f_2 \cdot c_2}{n}$, la cual constituye la frecuencia esperada por azar. Si se define a $p_e = \frac{\frac{f_1 \cdot c_1 + f_2 \cdot c_2}{n}}{n} = \frac{f_1 \cdot c_1 + f_2 \cdot c_2}{n^2}$

El coeficiente o índice Kappa se define como $K = \frac{p_0 - p_e}{1 - p_e}$, que cumple la propiedad que $0 \leq K \leq 1$.

Escala para la valoración de índice de Kappa:

Tabla 7 Coeficiente de Kappa

Valor de K	Grado de concordancia
< 0.20	Pobre
0.21 – 0.40	Débil
0.41 – 0.60	Moderada
0.61 – 0.80	Buena
0.81 – 1.00	Muy buena

2.6 Conclusiones parciales del capítulo.

Luego del análisis de la información a procesar se decidió conformar varios conjuntos de datos atendiendo a los 4 momentos de medición de los valores de las variables y los niveles de granularidad en que se presentan estos valores. Realizado un estudio para selección de rasgos se decidió descartar rasgos en los conjuntos de datos de las 24, 48 y 72 horas que en pruebas preliminares presentaron mejoras en los resultados. De estas pruebas preliminares se definieron también las técnicas a emplear en el experimento.

Para el procesamiento de los datos se trabajó con las herramientas Weka y SPSS versión 18.0. Estas herramientas poseen una adecuada interfaz e implementación de los métodos utilizados en el procesamiento de los datos incluyendo reportes del análisis estadístico de los resultados alcanzados a partir de los indicadores seleccionados.

Capítulo 3 Análisis de los resultados.

En este capítulo se presentan los resultados alcanzados en la aplicación de las técnicas de clasificación de inteligencia artificial: red neuronal artificial MLP en tres variantes, el árbol de decisión J48 y el IBK, en los conjuntos de datos construidos en los cuatro momentos del tiempo del paciente ingresado en que se actualizan las mediciones de las variables: al ingreso, las 24 horas, 48 horas y 72 horas; estructurados en tres variantes, los llamados: Modelo_Ingreso_Normales, Modelo_Ingreso_Especialista y Modelo_Ingreso_Óptimos. También se realiza la comparación de estos resultados y se concluye con la selección de los modelos que mejor pronostican el estado de los pacientes en el egreso.

3.1 Resultados del experimento.

3.2.1 Redes neuronales artificiales: MLP

Para la realización del experimento del Multilayer Perceptrón se establecieron tres MLP diferentes en cuanto a la configuración de cantidad de capas ocultas y la cantidad de neuronas por cada capa oculta, a continuación se detallan:

Tabla 8 Configuración de la red neuronal

MLP	Capas ocultas	Cantidad de neuronas por capas		
(1) MLP	2	3	2	
(2) MLP	2	6	3	
(3) MLP	3	14	10	6

3.2.2 Modelo al ingreso

Empleando las tres configuraciones de MLP contra las tres variantes de conjunto de datos en el modelo de ingreso, se obtuvo como resultado:

Tabla 9 Percent_correct

```
Analysing:  Percent_correct

Dataset                      (1) function | (2) funct (3) funct
----------------------------------------------------------------
Modelo_Ingreso_Normales  (100)     84.54  |   84.57     83.22
Modelo_Ingreso_Especialis(100)     80.10  |   79.51     80.10
Modelo_Ingreso_Óptimos   (100)     86.38  |   87.54     87.72
----------------------------------------------------------------
                                 (v/ /*) |   (0/3/0)   (0/3/0)
(1) functions.MultilayerPerceptron '-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E
20 -H \"3, 2\"'
```

```
(2) functions.MultilayerPerceptron '-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E
20 -H \"6, 3\"'
(3) functions.MultilayerPerceptron '-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E
20 -H \"14, 10, 6\"'
```

Los mejores modelos, según el porcentaje de casos clasificados correctamente, son los realizados con el conjunto de datos Modelo_Ingreso_Óptimos y especialmente el obtenido a partir de la función (3) con un 87.72 % de aciertos.

Nota: los promedios que se presentan en los resultados de verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos son obtenidos a partir de cada repetición y de la validación cruzada.

Tabla 10 Kappa_statistic

```
Analysing:  Kappa_statistic

Dataset                   (1) function | (2) func (3) func
----------------------------------------------------------
Modelo_Ingreso_Normales  (100)  0.67    |   0.66     0.63
Modelo_Ingreso_Especialis(100)  0.56    |   0.55     0.56
Modelo_Ingreso_Óptimos   (100)  0.70    |   0.73     0.73
----------------------------------------------------------
                              (v/ /*) |  (0/3/0)  (0/3/0)
```

El valor del coeficiente Kappa permite determinar que los mejores resultados en la clasificación se logran con el conjunto de datos Modelo_Ingreso_Óptimos y especialmente los obtenidos a partir de las configuraciones de la red neuronal definidas por las funciones 2 y 3 con un 73 % de aciertos.

Tabla 11True_positive_rate

```
Analysing:  True_positive_rate

Dataset                    (1) functio | (2) func (3) func
----------------------------------------------------------
Modelo_Ingreso_Normales  (100)    0.87 |   0.88     0.86
Modelo_Ingreso_Especialis(100)    0.86 |   0.85     0.86
Modelo_Ingreso_Óptimos   (100)    0.92 |   0.91     0.92
----------------------------------------------------------
                              (v/ /*) |  (0/3/0)  (0/3/0)
```

La frecuencia relativa de verdaderos positivos no muestra diferencias significativas entre las tres configuraciones de la red neuronal empleada, pero si se observan diferencias entre los resultados obtenidos con cada conjunto de datos, siendo la más efectiva la llamada Modelo_Ingreso_Óptimos con un 92 % de verdaderos positivos.

Tabla 12True_negative_rate

```
Analysing:  True_negative_rate

Dataset                   (1) functio | (2) func (3) func
----------------------------------------------------------
```

```
Modelo_Ingreso_Normales  (100)   0.80 |   0.78     0.77
Modelo_Ingreso_Especialis(100)   0.70 |   0.70     0.70
Modelo_Ingreso_Óptimos   (100)   0.77 |   0.81     0.80
-----------------------------------------------------
                              (v/ /*) |  (0/3/0)  (0/3/0)
```

Según la frecuencia de verdaderos negativos no se aprecian grandes diferencias en los resultados de las tres configuraciones de la red neuronal ni entre los tres conjuntos de datos empleados, siendo los resultados más altos los obtenidos con el conjunto de datos modelo ingresos óptimos con las funciones 2 y 3, y el conjunto de datos modelo ingreso normales con la función 1.

Se puede resumir que en la aplicación de la técnica MLP en estos 3 tres conjuntos de datos al ingreso que la configuración de la red según la cantidad de capas y neuronas por capa no incide sustancialmente en la clasificación, sin embargo se aprecia cierta ventaja en la configuración definida como función 2. La configuración de los conjuntos de datos que contribuyen mejor al desempeño de la red neuronal es la definida según el criterio de granulación óptima obtenido con la aplicación de la técnica estadística de minería de datos CHAID.

3.2.2 Árbol de decisión j48

Tabla 13Percent_correct

```
Analysing:  Percent_correct

Dataset                   (1) trees.J48
----------------------------------------
Modelo_Ingreso_Normales  (100)  78.32 |
Modelo_Ingreso_Especialis(100)  79.01 |
Modelo_Ingreso_Óptimos   (100)  83.69 |
----------------------------------------
(v/ /*)                               |
```

Según el porcentaje de casos clasificados correctamente el árbol de decisión J48 con el conjunto de datos Modelo_Ingreso_Óptimos fue la de mejor resultado con un 83.69 % de aciertos.

Tabla 14Kappa_statistic

```
Analysing:  Kappa_statistic

Dataset                   (1) trees.J48
----------------------------------------
Modelo_Ingreso_Normales   (100)   0.51 |
Modelo_Ingreso_Especialis (100)   0.53 |
```

```
Modelo_Ingreso_Óptimos    (100)    0.63 |
----------------------------------------
(v/ /*)                                 |
```

El coeficiente Kappa de 0.63 muestra que el árbol J48 con datos modelos ingreso óptimos fue el de mejores resultados.

Tabla 15 True_positive_rate

```
Analysing:  True_positive_rate

Dataset                    (1) trees.J48
----------------------------------------
Modelo_Ingreso_Normales  (100)    0.86 |
Modelo_Ingreso_Especialis(100)    0.85 |
Modelo_Ingreso_Óptimos   (100)    0.90 |
----------------------------------------
(v/ /*)                                 |
```

El 90 % de verdaderos positivos obtenido con el árbol J48 a partir de los datos modelos ingreso óptimos fue el mejor resultado según este indicador. Igual comportamiento muestra el porcentaje de verdaderos negativos, con un 73 %. (Tablas 15 y 16)

Tabla 16 True_negative_rate

```
Analysing:  True_negative_rate

Dataset                    (1) trees.J48
----------------------------------------
Modelo_Ingreso_Normales  (100)    0.64 |
Modelo_Ingreso_Especialis(100)    0.68 |
Modelo_Ingreso_Óptimos   (100)    0.73 |
----------------------------------------
(v/ /*)                                 |
```

En resumen se puede plantear que el árbol de decisión J48 alcanza los mejores resultados en la clasificación con el conjunto de datos configurada según el criterio de granulación óptima.

3.2.3 IBK (vecino más cercano)

Tabla 17 Percent_correct

```
Analysing:  Percent_correct

Dataset                    (1) lazy.IBk
----------------------------------------
Modelo_Ingreso_Normales  (100)    81.60 |
Modelo_Ingreso_Especialis(100)    82.00 |
```

```
Modelo_Ingreso_Óptimos   (100)    92.18 |
----------------------------------------
(v/ /*)                                 |
```

Según el porcentaje de casos clasificados correctamente el IBK con el conjunto de datos Modelo_Ingreso_Óptimos fue el mejor resultado con un 92.18 % de aciertos.

Tabla 18 Kappa_statistic

```
Analysing:  Kappa_statistic

Dataset                    (1) lazy.IBk
----------------------------------------
Modelo_Ingreso_Normales  (100)     0.57 |
Modelo_Ingreso_Especialis(100)     0.58 |
Modelo_Ingreso_Óptimos   (100)     0.81 |
----------------------------------------
(v/ /*)                                 |
```

El coeficiente Kappa obtenido por el IBK con el conjunto de datos Modelo_Ingreso_Óptimos de 0.81 es el mejor de los tres modelos aplicados.
El 99 % de verdaderos positivos y el 79 % de verdaderos negativos obtenidos al aplicar el IBK en el conjunto de datos Modelo_Ingreso_Óptimos muestran que esta configuración de el conjunto de datos es la de mejores resultados. (Tablas 19 y 20)

En resumen se puede concluir que el conjunto de datos Modelo_Ingreso_Óptimos ha garantizado el mejor desempeño a la técnica de I.A. en la clasificación del estado al egreso de los pacientes en el momento del ingreso.

Tabla 19 True_positive_rate

```
Analysing:  True_positive_rate

Dataset                    (1) lazy.IBk
----------------------------------------
Modelo_Ingreso_Normales  (100)    0.92 |
Modelo_Ingreso_Especialis(100)    0.92 |
Modelo_Ingreso_Óptimos   (100)    0.99 |
----------------------------------------
(v/ /*)                                 |
```

Tabla 20 True_negative_rate

```
Analysing:  True_negative_rate

Dataset                    (1) lazy.IBk
----------------------------------------
Modelo_Ingreso_Normales  (100)    0.63 |
Modelo_Ingreso_Especialis(100)    0.64 |
Modelo_Ingreso_Óptimos   (100)    0.79 |
----------------------------------------
(v/ /*)                                |
```

Nota: además de los indicadores analizados en los resultados; se presentan otros indicadores, que pueden ser consultados en el anexo1 los cuales son: true_negative_rate, num_true_negative, false_negative_rate y num_false_negative,

3.3 Comparación de resultados de modelos al ingreso

En el gráfico 3 se muestra comparativamente el desempeño de cada uno de los cinco modelos de aprendizaje automático aplicados al momento del ingreso respecto a los indicadores porcentaje de clasificación correcta y coeficiente Kappa ampliado a una escala de 1 a100, donde se aprecia que es la técnica IBK fue la de mejor desempeño.

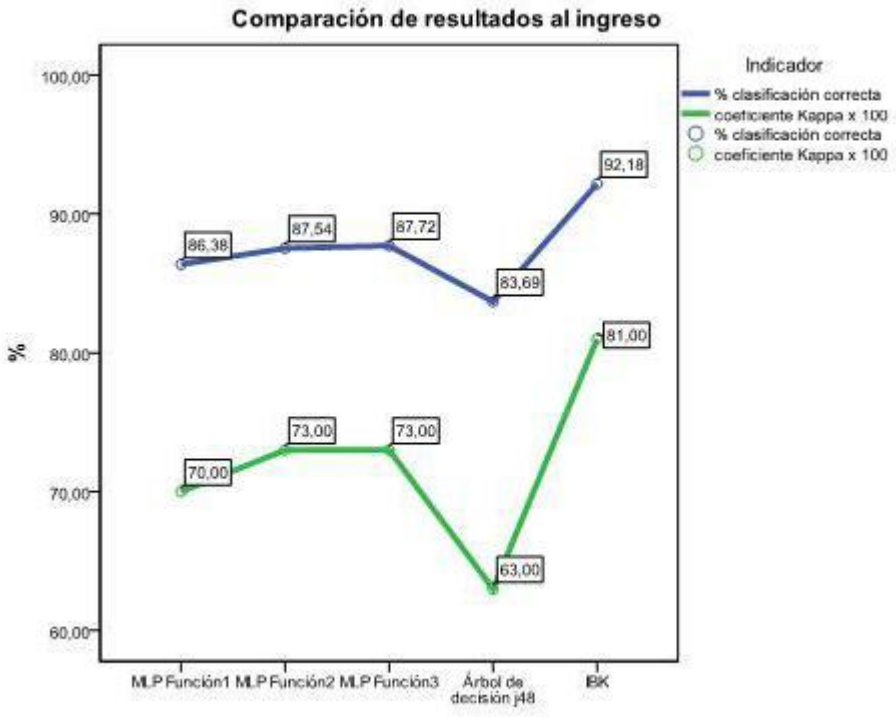

Gráfico 3 Comparación de resultados al ingreso

En la tabla 21 se presentan los resultados de las técnicas que mostraron los resultados próximos a la seleccionada como la de mejor desempeño junto al

conjunto de datos más ventajoso. En el caso de la red neuronal MLP se presenta solo la configuración de mejor desempeño.

Tabla 21 Comparación de las técnicas al ingreso

Indicadores	(3) MLP	Árbol de Decisión J48	IBK (vecino más cercano)
Conjunto de dato de mejor desempeño	Modelo_Ingreso_Óptimos	Modelo_Ingreso_Óptimos	Modelo_Ingreso_Óptimos
Porcent_correct	87.72	83.69	92.18
Kappa_statistc	0.73	0.63	0.81
True_positive_rate	0.92	0.90	0.99
False_positive_rate	0.20	0.27	0.21
True_negative_rate	0.80	0.73	0.79
False_Negative_rate	0.08	0.10	0.01

De acuerdo a los indicadores de calidad de la clasificación utilizados se puede concluir que la técnica de inteligencia artificial IBK obtiene generalmente los mejores indicadores en la calidad de la clasificación y que el conjunto de datos estructurada según el criterio de granulación óptima es la que más contribuye al desempeño de las técnicas de IA al momento del ingreso.

3.4 Modelo a las 24 horas

3.4.1 MLP

Para el pronóstico a las 24 horas del ingreso se experimentó con los tres MLP configurados como en la tabla 8 y con las nueve configuraciones del conjunto de datos para este momento del ingreso, lo cual generó la corrida de 27 modelos cuyos resultados se exponen a continuación:

Tabla 22 Percent_correct

```
Analysing:  Percent_correct

Dataset                    (1) function | (2) funct (3) funct
-------------------------------------------------------------
Modelo_24H_Normales        (100)  89.88 |   90.90     90.81
Modelo_24h_Normales_RASGO1 (100)  93.00 |   94.07     93.86
Modelo_24h_Normales_RASGO2 (100)  90.10 |   90.10     90.81
Modelo_24H_Especialista    (100)  90.65 |   90.64     90.68
Modelo_24h_Especialista_R1 (100)  92.72 |   92.72     92.60
Modelo_24h_Especialista_R2 (100)  91.49 |   91.49     91.74
Modelo_24H_Óptimo          (100)  90.69 |   90.92     91.17
Modelo_24h_Óptimos_RASGOS 1(100)  92.64 |   93.01     93.47
Modelo_24h_Óptimos_RASGOS 2(100)  90.96 |   90.25     90.36
-------------------------------------------------------------
                                (v/ /*) |   (0/9/0)   (0/9/0)
```

Según el porcentaje de casos clasificados correctamente fue el modelo MLP función2 con el conjunto de datosModelo_24h_Normales_RASGO1 la de mejor desempeño, con un 94.07 % de aciertos. Aunque con resultados próximos a estos se comportaron los modelos: MLP función1 con el conjunto de datos Modelo_24h_Normales_RASGO1, MLP función3 con el conjunto de datos Modelo_24h_Normales_RASGO1, MLP función2 con el conjunto de datosModelo_24h_óptimos_RASGO1 y MLP función3 con el conjunto de datosModelo_24h_óptimos_RASGO1. (Tablas 22)

De estos resultados se puede inferir además que los conjuntos de datos que mejor contribuyen al desempeño de la técnica MLP fueron Modelo_24h_Normales_RASGO1 y Modelo_24h_óptimos_RASGO1.
Se puede afirmar que la selección de rasgos 1 es la que más contribuye al desempeño de la red neuronal; mientras que no se observa diferencia sustancial entre las formas de configurar los datos según: estructura original, granulados según criterio de especialista y según criterio óptimo en cuanto al porcentaje de clasificación correcta.

Tabla 23 Kappa_statistic

Analysing: Kappa_statistic

Dataset	(1) functio	(2) func	(3) func
Modelo_24H_Normales	(100) 0.78	0.80	0.80
Modelo_24h_Normales_RASGO1	(100) 0.85	0.87	0.87
Modelo_24h_Normales_RASGO2	(100) 0.78	0.78	0.80
Modelo_24H_Especialista	(100) 0.79	0.79	0.79
Modelo_24h_Especialista_R1	(100) 0.85	0.85	0.84
Modelo_24h_Especialista_R2	(100) 0.82	0.82	0.82
Modelo_24H_Óptimo	(100) 0.80	0.80	0.81
Modelo_24h_Óptimos_RASGOS1	(100) 0.84	0.85	0.86
Modelo_24h_Óptimos_RASGOS2	(100) 0.80	0.79	0.79
	(v/ /*)	(0/9/0)	(0/9/0)

La comparación de los 27 modelos corridos respecto al indicador coeficiente Kappa (Tablas 23) permitió establecer que todos muestran un desempeño en la clasificación de bueno o muy bueno ; siendo los de mejor desempeño el modelo MLP función2 con el conjunto de datos Modelo_24h_Normales_RASGO1 y modelo MLP función3 con el conjunto de datos Modelo_24h_Normales_RASGO1. Aunque los modelos: modelo MLP función1 con el conjunto de datos Modelo_24h_Normales_RASGO1, modelo

MLP función1 con el conjunto de datos Modelo_24h_Especialista_RASGO1, modelo MLP función2 con el conjunto de datosModelo_24h_Especialista_RASGO1, modelo MLP función3 con el conjunto de datos Modelo_24h_Especialista_RASGO1, modelo MLP función1 con el conjunto de datos Modelo_24h_Óptimos_RASGO1, modelo MLP función2 con el conjunto de datos Modelo_24h_Óptimos_RASGO1 y modelo MLP función3 con el conjunto de datos Modelo_24h_Óptimos_RASGO1 muestran un desempeño muy próximo a los de mejor desempeño.

Se puede afirmar además que el método de selección de rasgos1 es el que mejor contribución hace al desempeño de la red neuronal experimentada; mientras que la forma de configurar los datos según: estructura original, granulados según criterio de especialista y según criterio óptimo no muestran diferencias significativas en cuanto a este indicador.

Tabla 24 True_positive_rate

Analysing: True_positive_rate

Dataset	(1) functio	(2) func	(3) func
Modelo_24H_Normales	(100) 0.93	0.93	0.93
Modelo_24h_Normales_RASGO1	(100) 0.96	0.96	0.95
Modelo_24h_Normales_RASGO2	(100) 0.93	0.93	0.93
Modelo_24H_Especialista	(100) 0.94	0.94	0.94
Modelo_24h_Especialista_R1	(100) 0.93	0.93	0.93
Modelo_24h_Especialista_R2	(100) 0.93	0.93	0.93
Modelo_24H_Óptimo	(100) 0.93	0.93	0.93
Modelo_24h_Óptimos_RASGOS1	(100) 0.95	0.95	0.95
Modelo_24h_Óptimos_RASGOS2	(100) 0.93	0.92	0.92
	(v/ /*)	(0/9/0)	(0/9/0)

En el indicador razón de verdaderos positivos (Tablas 24) muestra los resultados más prominentes el modelo MLP en sus tres configuraciones con el conjunto de datos Modelo_24h_Normales_RASGO1 y el modelo MLP en sus tres configuraciones con el conjunto de datos Modelo_24h_Óptimos_RASGO1.

En este indicador el método de selección de rasgos1 es el que mejor contribución hace al desempeño de la red neuronal experimentada.

Tabla 25 True_negative_rate

Analysing: True_negative_rate

Dataset	(1) functio	(2) func	(3) func
Modelo_24H_Normales	(100) 0.85	0.88	0.87
Modelo_24h_Normales_RASGO1	(100) 0.88	0.91	0.91
Modelo_24h_Normales_RASGO2	(100) 0.84	0.85	0.87

Modelo_24H_Especialista	(100)	0.85	0.85	0.85
Modelo_24h_Especialista_R1	(100)	0.92	0.92	0.92
Modelo_24h_Especialista_R2	(100)	0.89	0.89	0.90
Modelo_24H_Óptimo	(100)	0.86	0.86	0.87
Modelo_24h_Óptimos_RASGOS1	(100)	0.89	0.90	0.90
Modelo_24h_Óptimos_RASGOS2	(100)	0.87	0.87	0.87
		(v/ /*)	(0/9/0)	(0/9/0)

En el indicador razón de verdaderos negativos (Tablas 25) los resultados más prominentes los muestra el modelo MLP en sus tres configuraciones con el conjunto de datos Modelo_24h_Especialista_RASGO1 y el modelo MLP en sus configuraciones 2 y 3 con el conjunto de datos Modelo_24h_Normales_RASGO1.

En este indicador el método de selección de rasgos1 es el que mejor contribución hace al desempeño de la red neuronal experimentada.

Tabla 26 False_positive_rate

Analysing: False_positive_rate

Dataset		(1) functio	(2) func	(3) func
Modelo_24H_Normales	(100)	0.15	0.12	0.13
Modelo_24h_Normales_RASGO1	(100)	0.12	0.09	0.09
Modelo_24h_Normales_RASGO2	(100)	0.16	0.15	0.13
Modelo_24H_Especialista	(100)	0.15	0.15	0.15
Modelo_24h_Especialista_R1	(100)	0.08	0.08	0.08
Modelo_24h_Especialista_R2	(100)	0.11	0.11	0.10
Modelo_24H_Óptimo	(100)	0.14	0.14	0.13
Modelo_24h_Óptimos_RASGOS1	(100)	0.11	0.10	0.10
Modelo_24h_Óptimos_RASGOS2	(100)	0.13	0.13	0.13
		(v/ /*)	(0/9/0)	(0/9/0)

En cuanto a la razón de de falsos positivos (Tablas 26) el modelo de mejor desempeño resultó ser modelo MLP en sus tres configuraciones con el conjunto de datos Modelo_24h_Especialista_RASGO1 con un 8 % de falsos positivos.

Tabla 27 False_negative_rate

Analysing: False_negative_rate

Dataset		(1) functio	(2) func	(3) func
Modelo_24H_Normales	(100)	0.07	0.07	0.07
Modelo_24h_Normales_RASGO1	(100)	0.04	0.04	0.05
Modelo_24h_Normales_RASGO2	(100)	0.07	0.07	0.07
Modelo_24H_Especialista	(100)	0.06	0.06	0.06
Modelo_24h_Especialista_R1	(100)	0.07	0.07	0.07
Modelo_24h_Especialista_R2	(100)	0.07	0.07	0.07

```
Modelo_24H_Óptimo          (100)   0.07 |    0.07      0.07
Modelo_24h_Óptimos_RASGOS1(100)    0.05 |    0.05      0.05
Modelo_24h_Óptimos_RASGOS2(100)   0.07 |    0.08      0.08
-----------------------------------------------------
                              (v/ /*) |   (0/9/0)   (0/9/0)
```

En cuanto a la razón de de falsos negativos (Tablas 27) los modelos de mejor desempeño resultaron ser el modelo MLP en sus tres configuraciones: 1,2 y3 con el conjunto de datos Modelo_24h_Normales_RASGO1 con un 4 %, 4 % y 5% respectivamente de falsos negativos y el modelo MLP en sus tres configuraciones con el conjunto de datos Modelo_24h_ÓPTIMOS_RASGO1 con un 5 % de falsos negativos.

El análisis de estos indicadores permite concluir en el presente estudio que la configuración de la red neuronal según capas y neuronas no influye significativamente en la calidad de los indicadores analizados, mientras que la selección de rasgos 1 es la que más contribuye al desempeño de la red neuronal.

3.4.2 Árbol de decisión J48

Se experimentó con nueve modelos del tipo árbol de decisión j48 generados por la distintas configuraciones de la conjunto de datos.

Resultados:

Tabla 28 Percent_correct

```
Analysing:  Percent_correct

Dataset                    (1) trees.J48
----------------------------------------
Modelo_24H_Normales      (100)    88.74 |
Modelo_24h_Normales_RASGO1(100)   89.96 |
Modelo_24h_Normales_RASGO2(100)   88.99 |
Modelo_24H_Especialista  (100)    84.92 |
Modelo_24h_Especialista_R1(100)   86.63 |
Modelo_24h_Especialista_R2(100)   93.08 |
Modelo_24H_Óptimo        (100)    88.11 |
Modelo_24h_Óptimos_RASGOS1(100)   91.14 |
Modelo_24h_Óptimos_RASGOS2(100)   96.08 |
----------------------------------------
(v/ /*)                                 |
```

Los modelos de mejor resultado según el porcentaje de casos clasificados correctamente (Tablas 28) fueron los corridos con el conjunto de datos Modelo_24h_Óptimos_RASGOS2 con un 96.08 % de aciertos y el corrido con

el conjunto de datos Modelo_24h_especialista_RASGOS2 con un 96.08 % de aciertos.

En este indicador la selección de rasgos 2 es la que mayormente mejor contribuye al desempeño de la técnica árbol de decisión j48.

Tabla 29 Kappa_statistic

```
Analysing:  Kappa_statistic

Dataset                    (1) trees.J48
----------------------------------------
Modelo_24H_Normales      (100)   0.75 |
Modelo_24h_Normales_RASGO1(100)  0.78 |
Modelo_24h_Normales_RASGO2(100)  0.77 |
Modelo_24H_Especialista  (100)   0.67 |
Modelo_24h_Especialista_R1(100)  0.71 |
Modelo_24h_Especialista_R2(100)  0.85 |
Modelo_24H_Óptimo        (100)   0.75 |
Modelo_24h_Óptimos_RASGOS1(100)  0.81 |
Modelo_24h_Óptimos_RASGOS2(100)  0.92 |
----------------------------------------
(v/ /*)                                |
```

Un coeficiente Kappa es de 0.92 (Tablas 29) obtenido por el árbol de decisión j48 con el conjunto de datos Modelo_24h_Óptimos_RASGOS2, indica una diferencia sustancial con los demás modelos experimentados.

Tabla 30 True_positive_rate

```
Analysing:  True_positive_rate

Dataset                    (1) trees.J48
----------------------------------------
Modelo_24H_Normales      (100)   0.93 |
Modelo_24h_Normales_RASGO1(100)  0.94 |
Modelo_24h_Normales_RASGO2(100)  0.90 |
Modelo_24H_Especialista  (100)   0.88 |
Modelo_24h_Especialista_R1(100)  0.90 |
Modelo_24h_Especialista_R2(100)  0.93 |
Modelo_24H_Óptimo        (100)   0.90 |
Modelo_24h_Óptimos_RASGOS1(100)  0.94 |
Modelo_24h_Óptimos_RASGOS2(100)  0.96 |
----------------------------------------
(v/ /*)                                |
```

Los indicadores de razón de verdaderos positivos, razón de verdaderos negativos, razón de falsos positivos y razón de falsos negativos (Tablas 30, 31,32 y 33) muestra como resultado más relevante el obtenido por el modelo árbol de decisión j48 con el conjunto de datos Modelo_24h_Óptimos_RASGOS2. Resaltadas en verde claro aparecen estas

razones en este modelo y en otros modelos que muestran razones significativas por su magnitud.

Tabla 31 True_negative_rate

Analysing: True_negative_rate

Dataset	(1) trees.J48
Modelo_24H_Normales (100)	0.81
Modelo_24h_Normales_RASGO1 (100)	0.82
Modelo_24h_Normales_RASGO2 (100)	0.87
Modelo_24H_Especialista (100)	0.79
Modelo_24h_Especialista_R1 (100)	0.80
Modelo_24h_Especialista_R2 (100)	0.93
Modelo_24H_Óptimo (100)	0.84
Modelo_24h_Óptimos_RASGOS1 (100)	0.86
Modelo_24h_Óptimos_RASGOS2 (100)	0.97

(v/ /*)

Tabla 32 False_positive_rate

Analysing: False_positive_rate

Dataset	(1) trees.J48
Modelo_24H_Normales (100)	0.19
Modelo_24h_Normales_RASGO1 (100)	0.18
Modelo_24h_Normales_RASGO2 (100)	0.13
Modelo_24H_Especialista (100)	0.21
Modelo_24h_Especialista_R1 (100)	0.20
Modelo_24h_Especialista_R2 (100)	0.07
Modelo_24H_Óptimo (100)	0.16
Modelo_24h_Óptimos_RASGOS1 (100)	0.14
Modelo_24h_Óptimos_RASGOS2 (100)	0.03

(v/ /*)

Tabla 33 False_negative_rate

Analysing: False_negative_rate

Dataset	(1) trees.J48
Modelo_24H_Normales (100)	0.07
Modelo_24h_Normales_RASGO1 (100)	0.06
Modelo_24h_Normales_RASGO2 (100)	0.10
Modelo_24H_Especialista (100)	0.12
Modelo_24h_Especialista_R1 (100)	0.10
Modelo_24h_Especialista_R2 (100)	0.07
Modelo_24H_Óptimo (100)	0.10
Modelo_24h_Óptimos_RASGOS1 (100)	0.06
Modelo_24h_Óptimos_RASGOS2 (100)	0.04

(v/ /*)

A modo de resumen se puede plantear que el modelo de árbol de decisión j48 de mejor desempeño fue el corrido sobre el conjunto de datos Modelo_24h_Óptimos_RASGOS2.

3.4.3 IBK (Vecino más cercano)

El experimentó se realizó con nueve modelos del tipo IBK generados por la distintas configuraciones de la conjunto de datos.

Resultados:

Tabla 34 Percent_correct

```
Analysing:  Percent_correct

Dataset                    (1) lazy.IBk
---------------------------------------
Modelo_24H_Normales      (100)   89.15 |
Modelo_24h_Normales_RASGO1(100)  92.14 |
Modelo_24h_Normales_RASGO2(100)  88.58 |
Modelo_24H_Especialista  (100)   91.13 |
Modelo_24h_Especialista_R1(100)  92.90 |
Modelo_24h_Especialista_R2(100)  86.24 |
Modelo_24H_Óptimo        (100)   94.08 |
Modelo_24h_Óptimos_RASGOS1(100)  90.65 |
Modelo_24h_Óptimos_RASGOS2(100)  90.54 |
---------------------------------------
(v/ /*)                                |
```

Los modelos de mejor desempeño según el porcentaje de casos clasificados correctamente (Tablas 34) fueron los corridos con el conjunto de datos Modelo_24h_Óptimos con un 94.08 % de aciertos seguido por los modelos corridos con los conjuntos de datos Modelo_24h_especialista_RASGOS1 y Modelo_24h_Normales_RASGOS1 con un 92.90 % y 92,14% de aciertos respectivamente.

Tabla 35 Kappa_statistic

```
Analysing:  Kappa_statistic

Dataset                    (1) lazy.IBk
---------------------------------------
Modelo_24H_Normales      (100)   0.75 |
Modelo_24h_Normales_RASGO1(100)  0.83 |
Modelo_24h_Normales_RASGO2(100)  0.73 |
Modelo_24H_Especialista  (100)   0.79 |
Modelo_24h_Especialista_R1(100)  0.84 |
Modelo_24h_Especialista_R2(100)  0.67 |
Modelo_24H_Óptimo        (100)   0.86 |
Modelo_24h_Óptimos_RASGOS1(100)  0.78 |
Modelo_24h_Óptimos_RASGOS2(100)  0.78 |
---------------------------------------
(v/ /*)                                |
```

El coeficiente Kappa (Tablas 35) muestra al modelo IBK con el conjunto de datos Modelo_24h_Óptimo sin selección de rasgos como la de mejor

desempeño, con un coeficiente Kappa de 0,86, seguido por los modelos con los conjuntos de datos Modelo_24h_especialista_RASGOS1 y Modelo_24h_Normales_RASGOS1 con coeficientes Kappa de 0,84 y 0,83 respectivamente. Estos coeficientes clasifican la calidad de los pronósticos realizados en la categoría de muy bueno.

Tabla 36 True_positive_rate

Analysing: True_positive_rate

Dataset	(1)	lazy.IBk
Modelo_24H_Normales	(100)	0.96
Modelo_24h_Normales_RASGO1	(100)	0.95
Modelo_24h_Normales_RASGO2	(100)	0.98
Modelo_24H_Especialista	(100)	0.98
Modelo_24h_Especialista_R1	(100)	0.96
Modelo_24h_Especialista_R2	(100)	0.98
Modelo_24H_Óptimo	(100)	0.98
Modelo_24h_Óptimos_RASGOS1	(100)	0.97
Modelo_24h_Óptimos_RASGOS2	(100)	0.98

(v/ /*)

Los indicadores de razón de verdaderos positivos y razón de falsos negativos (Tablas 36 y 39) muestran en general resultados significativamente positivos siendo los más relevantes los obtenido por el modelo IBK con las conjunto de datos Modelo_24h_Normales_RASGOS2, Modelo_24h_Especialista, Modelo_24h_Especialista_RASGOS2, Modelo_24h_Óptimos y Modelo_24h_Óptimos_RASGOS2 .

Las razones de verdaderos negativos y de falsos positivos (Tablas 37 y 38) muestra los resultados más relevantes en los modelos Modelo_24h_Normales_RASGOS1, Modelo_24h_Especialista _RASGOS1 y Modelo_24h_Óptimos

Tabla 37 True_negative_rate

Analysing: True_negative_rate

Dataset	(1)	lazy.IBk
Modelo_24H_Normales	(100)	0.77
Modelo_24h_Normales_RASGO1	(100)	0.86
Modelo_24h_Normales_RASGO2	(100)	0.72
Modelo_24H_Especialista	(100)	0.79
Modelo_24h_Especialista_R1	(100)	0.87
Modelo_24h_Especialista_R2	(100)	0.65
Modelo_24H_Óptimo	(100)	0.87
Modelo_24h_Óptimos_RASGOS1	(100)	0.80
Modelo_24h_Óptimos_RASGOS2	(100)	0.77

```
(v/ /*)                                |
```

Tabla 38 False_positive_rate

```
Analysing:  False_positive_rate

Dataset                    (1) lazy.IBk
----------------------------------------
Modelo_24H_Normales       (100)    0.23 |
Modelo_24h_Normales_RASGO1(100)    0.14 |
Modelo_24h_Normales_RASGO2(100)    0.28 |
Modelo_24H_Especialista   (100)    0.21 |
Modelo_24h_Especialista_R1(100)    0.13 |
Modelo_24h_Especialista_R2(100)    0.35 |
Modelo_24H_Óptimo         (100)    0.13 |
Modelo_24h_Óptimos_RASGOS1(100)    0.20 |
Modelo_24h_Óptimos_RASGOS2(100)    0.23 |
----------------------------------------
(v/ /*)                                |
```

Tabla 39 False_negative_rate

```
Analysing:  False_negative_rate

Dataset                    (1) lazy.IBk
----------------------------------------
Modelo_24H_Normales       (100)    0.04 |
Modelo_24h_Normales_RASGO1(100)    0.05 |
Modelo_24h_Normales_RASGO2(100)    0.02 |
Modelo_24H_Especialista   (100)    0.02 |
Modelo_24h_Especialista_R1(100)    0.04 |
Modelo_24h_Especialista_R2(100)    0.02 |
Modelo_24H_Óptimo         (100)    0.02 |
Modelo_24h_Óptimos_RASGOS1(100)    0.03 |
Modelo_24h_Óptimos_RASGOS2(100)    0.02 |
----------------------------------------
(v/ /*)                                |
```

Según el porcentaje de clasificación correcta y el índice Kappa los modelos de mejor desempeño son los modelos árbol de decisión j48 con los conjuntos de datos Modelo_24h_Normales_RASGOS1, Modelo_24h_Especialista _RASGOS1 y Modelo_24h_Óptimos, siendo el más relevante este último.

Nota: además de los indicadores analizados en los resultados; se presentan otros indicadores, que pueden ser consultados en el anexo2 los cuales son: true_negative_rate, num_true_negative, false_negative_rate y num_false_negative,

3.5 Comparación de los modelos a las 24 horas

En el gráfico 4 se muestra comparativamente el desempeño de cada uno de los cinco modelos de aprendizaje automático aplicados a las 24 horas del ingreso respecto a los indicadores porcentaje de clasificación correcta y

coeficiente Kappa ampliado a una escala de 1 a 100, donde se aprecia que es la técnica árbol de decisión la de mejor desempeño.

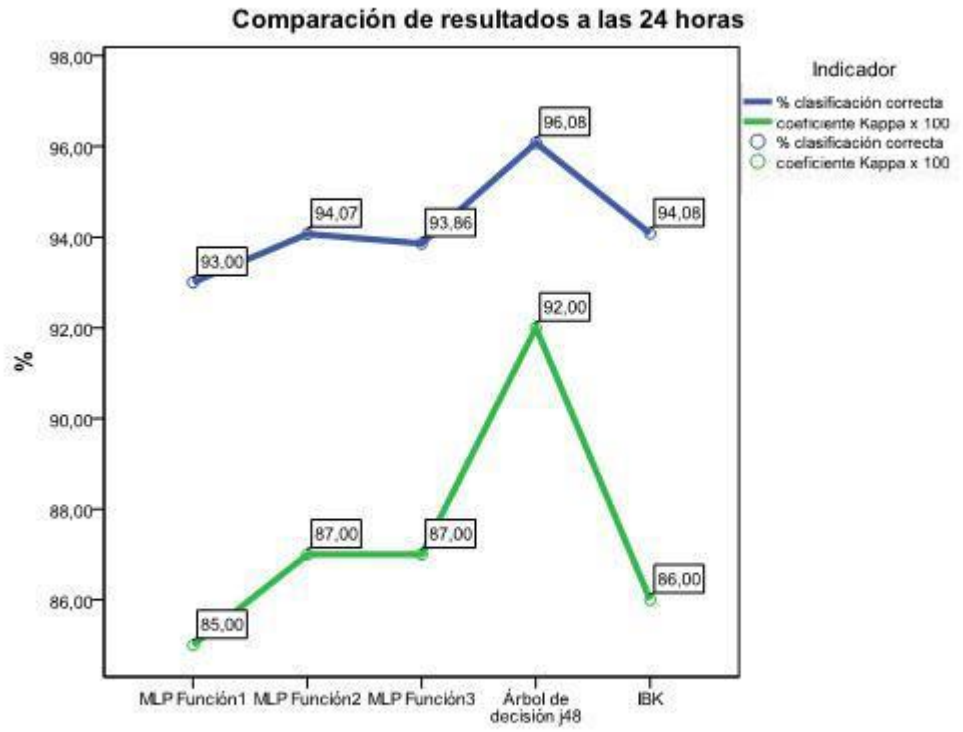

Gráfico 4 Comparación de resultados a las 24 horas

En la tabla 40 se presentan los resultados de los mejores modelos corridos en cada una de las técnicas aplicadas, así como la base de dato con la cual obtienen esos indicadores a las 24 horas del ingreso.

Tabla 40 Comparación de las técnicas a las 24 horas

Indicadores	(2) MLP	Árbol de Decisión J48	IBK (vecino más cercano)
Conjunto de dato de mejor desempeño	Modelo_24h_Normales_RASGO1 Función 2	Modelo_24H_Óptimo_ RASGOS2	Modelo_24H_Óptimo
Porcent_correct	94.07	96.08	94.08
Kappa_statistc	0.87	0.92	0.86
True_positive_rate	0.96	0.96	0.98
False_positive_rate	0.09	0.03	0.13
True_negative_rate	0.91	0.97	0.87
False_Negative_rate	0.04	0.04	0.02

El modelo de mejor desempeño en la mayoría de los indicadores resultó ser el árbol de decisión J48 con una selección de rasgos según el procedimiento 2 y con datos granulados según el criterio óptimo, el cual muestra indicadores altamente positivos según todos los indicadores considerados en esta investigación.

3.6 Modelo a las 48 horas del ingreso

3.6.1 MPL

Para el pronóstico a las 48 horas del ingreso se experimentó con los tres MLP configurados como en la tabla 8 y con las nueve configuraciones del conjunto de datos para este momento del ingreso, lo cual generó la corrida de 27 modelos cuyos resultados se exponen a continuación:

Tabla 41 Percent_correct

Analysing: Percent_correct

Dataset	(1) function	(2) funct	(3) funct
Modelo_48h_Normales (100)	91.77	91.63	91.25
Modelo_48h_Normales_RASGO1 (100)	93.93	93.93	94.05
Modelo_48h_Normales_RASGO2 (100)	89.34	89.34	89.46
Modelo_48h_Espacialista (100)	94.79	94.91	95.55
Modelo_48h_Especialista_R1 (100)	95.80	95.68	95.68
Modelo_48h_Especialista_R2 (100)	95.18	94.91	94.91
Modelo_48h_Óptimos (100)	92.16	92.41	92.16
Modelo_48h_Óptimos_RASGOS1 (100)	93.07	93.86	93.11
Modelo_48h_Óptimos_RASGOS2 (100)	92.02	92.02	91.77
(v/ /*)		(0/9/0)	(0/9/0)

Según el porcentaje de casos clasificados correctamente (Tabla 41) los modelos MLP en sus tres configuraciones con el conjunto de datos Modelo_48h_Especialista_RASGO1, modelos MLP Función3 con el conjunto de datos Modelo_48h_Especialista y el modelo MLP Función1 con el conjunto de datos Modelo_48h_Especialista_RASGO2 son los de mejor desempeño, siendo el primero de ellos el de más altos indicadores.

De estos resultados se puede inferir además que los conjuntos de datos que mejor contribuyen al desempeño de la técnica MLP fueron las que se obtuvieron con la granulación de los datos empleando el criterio del especialista.

Se aprecia cierta tendencia favorable a la selección de rasgos 1 en la mejora del desempeño de la red neuronal.

Tabla 42 Kappa_statistic

Analysing: Kappa_statistic

Dataset		(1) functio	(2) func	(3) func
Modelo_48h_Normales	(100)	0.79	0.79	0.78
Modelo_48h_Normales_RASGO1	(100)	0.85	0.85	0.85
Modelo_48h_Normales_RASGO2	(100)	0.74	0.74	0.74
Modelo_48h_Espacialista	(100)	0.88	0.88	0.89
Modelo_48h_Especialista_R1	(100)	0.90	0.90	0.90
Modelo_48h_Especialista_R2	(100)	0.89	0.89	0.89
Modelo_48h_Óptimos	(100)	0.80	0.80	0.80
Modelo_48h_Óptimos_RASGOS1	(100)	0.83	0.85	0.83
Modelo_48h_Óptimos_RASGOS2	(100)	0.80	0.80	0.79
		(v/ /*)	(0/9/0)	(0/9/0)

El coeficiente Kappa (Tabla 42) muestra que los modelos MLP en sus tres configuraciones con el conjunto de datos Modelo_48h_Especialista_RASGO1, es la de mejor desempeño, seguida por modelos MLP Función3 con el conjunto de datos Modelo_48h_Especialista y el modelo MLP Función1 con el conjunto de datos Modelo_48h_Especialista_RASGO2 .

De estos resultados se puede inferir además que el conjunto de datos que mejor contribuyen al desempeño de la técnica MLP fueron las que se obtuvieron con la granulación de los datos empleando el criterio del especialista.

Se aprecia además en este indicador cierta tendencia, favorable a la selección de rasgos 1, en la mejora del desempeño de la red neuronal.

Tabla 43 True_positive_rate

Analysing: True_positive_rate

Dataset		(1) functio	(2) func	(3) func
Modelo_48h_Normales	(100)	0.96	0.96	0.96
Modelo_48h_Normales_RASGO	(100)	0.98	0.98	0.98
Modelo_48h_Normales_RASGO	(100)	0.94	0.94	0.94
Modelo_48h_Espacialista	(100)	0.96	0.96	0.97
Modelo_48h_Especialista_R	(100)	0.97	0.97	0.97
Modelo_48h_Especialista_R	(100)	0.95	0.95	0.95
Modelo_48h_Óptimos	(100)	0.98	0.98	0.98
Modelo_48h_Óptimos_RASGOS	(100)	0.97	0.98	0.97
Modelo_48h_Óptimos_RASGOS	(100)	0.97	0.96	0.96
		(v/ /*)	(0/9/0)	(0/9/0)

En todos los modelos se obtienen altos porcentajes de verdaderos positivos y de verdaderos negativos. (Tabla 43 y 44)

Tabla 44 True_negative_rate

Analysing: True_negative_rate

Dataset		(1) functio	(2) func	(3) func
Modelo_48h_Normales	(100)	0.83	0.82	0.82
Modelo_48h_Normales_RASGO1	(100)	0.86	0.86	0.86
Modelo_48h_Normales_RASGO2	(100)	0.80	0.80	0.80
Modelo_48h_Espacialista	(100)	0.92	0.92	0.92
Modelo_48h_Especialista_R1	(100)	0.93	0.92	0.92
Modelo_48h_Especialista_R2	(100)	0.95	0.95	0.95
Modelo_48h_Óptimos	(100)	0.80	0.81	0.80
Modelo_48h_Óptimos_RASGOS1	(100)	0.86	0.86	0.86
Modelo_48h_Óptimos_RASGOS2	(100)	0.82	0.83	0.82
		(v/ /*)	(0/9/0)	(0/9/0)

Tabla 45 False_positive_rate

Analysing: False_positive_rate

Dataset		(1) functio	(2) func	(3) func
Modelo_48h_Normales	(100)	0.17	0.18	0.18
Modelo_48h_Normales_RASGO	(100)	0.14	0.14	0.14
Modelo_48h_Normales_RASGO	(100)	0.20	0.20	0.20
Modelo_48h_Espacialista	(100)	0.08	0.08	0.08
Modelo_48h_Especialista_R	(100)	0.07	0.08	0.08
Modelo_48h_Especialista_R	(100)	0.05	0.05	0.05
Modelo_48h_Óptimos	(100)	0.20	0.19	0.20
Modelo_48h_Óptimos_RASGOS	(100)	0.14	0.14	0.14
Modelo_48h_Óptimos_RASGOS	(100)	0.18	0.17	0.18
		(v/ /*)	(0/9/0)	(0/9/0)

El modelo MPL en sus tres configuraciones con el conjunto de datos de Modelo_48h_Especialista_RASGO2 es la de valores más favorables en el comportamiento de la razón de falsos positivos. (Tabla 45)

Tabla 46 False_negative_rate

Analysing: False_negative_rate

Dataset		(1) functio	(2) func	(3) func
Modelo_48h_Normales	(100)	0.04	0.04	0.04
Modelo_48h_Normales_RASGO1	(100)	0.02	0.02	0.02
Modelo_48h_Normales_RASGO2	(100)	0.06	0.06	0.06
Modelo_48h_Espacialista	(100)	0.04	0.04	0.03
Modelo_48h_Especialista_R1	(100)	0.03	0.03	0.03
Modelo_48h_Especialista_R2	(100)	0.05	0.05	0.05
Modelo_48h_Óptimos	(100)	0.02	0.02	0.02
Modelo_48h_Óptimos_RASGOS1	(100)	0.03	0.02	0.03
Modelo_48h_Óptimos_RASGOS2	(100)	0.03	0.04	0.04

```
                                (v/ /*) |   (0/9/0)   (0/9/0)
```

El mejor comportamiento según la razón de falsos negativos (Tabla 46) se obtiene por los modelos MPL en sus tres configuraciones con las conjunto de datos Modelo_48h_Normales_RASGO1 y Modelo_48h_Óptimos, además del modelo MPL Función 2 con las conjunto de datos Modelo_48h__ Óptimos_ RASGO1.

En resumen el modelo MPL en sus tres configuraciones con el conjunto de datos Modelo_48h_especialista_RASGO1 es el de mejor desempeño.
La granulación de datos por criterio del especialista es la que más contribuye al desempeño de la red neuronal.

3.6.2 Árbol de decisión J48

El experimentó se realizó con nueve modelos con el empleo de la técnica j48 generados por la distintas configuraciones de la conjunto de datos.

Tabla 47 Percent_correct

```
Analysing:  Percent_correct

Dataset                    (1) trees.J48
----------------------------------------
Modelo_48h_Normales      (100)    92.46 |
Modelo_48h_Normales_RASGO1(100)   93.50 |
Modelo_48h_Normales_RASGO2(100)   91.23 |
Modelo_48h_Espacialista  (100)    96.71 |
Modelo_48h_Especialista_R1(100)   96.71 |
Modelo_48h_Especialista_R2(100)   96.16 |
Modelo_48h_Óptimos       (100)    88.41 |
Modelo_48h_Óptimos_RASGOS1(100)   90.04 |
Modelo_48h_Óptimos_RASGOS2(100)   90.04 |
----------------------------------------
(v/ /*)                                 |
```

Según el porcentaje de casos clasificados correctamente (Tabla 47) son los modelos árbol de decisión J48 con los conjuntos de datos Modelo_48h_Especialista, Modelo_48h_Especialista_RASGOS1 y Modelo_48h_Especialista_RASGOS2 los que muestran un mejor desempeño con porcentajes de 96.71, 96.16 respectivamente.

Tabla 48 Kappa_statistic

```
Analysing:  Kappa_statistic

Dataset                    (1) trees.J48
```

Modelo_48h_Normales	(100)	0.83
Modelo_48h_Normales_RASGO1	(100)	0.85
Modelo_48h_Normales_RASGO2	(100)	0.80
Modelo_48h_Espacialista	(100)	0.92
Modelo_48h_Especialista_R1	(100)	0.92
Modelo_48h_Especialista_R2	(100)	0.91
Modelo_48h_Óptimos	(100)	0.72
Modelo_48h_Óptimos_RASGOS1	(100)	0.76
Modelo_48h_Óptimos_RASGOS2	(100)	0.76

(v/ /*)

Respecto al coeficiente Kappa (Tabla 43) muestran un mejor desempeño los modelos J48 con los conjuntos de datos Modelo_48h_Especialista, Modelo_48h_Especialista_RASGOS1 y Modelo_48h_Especialista_RASGOS2.

Tabla 49 True_positive_rate

Analysing: True_positive_rate

Dataset	(1)	trees.J48
Modelo_48h_Normales	(100)	0.94
Modelo_48h_Normales_RASGO1	(100)	0.94
Modelo_48h_Normales_RASGO2	(100)	0.93
Modelo_48h_Espacialista	(100)	0.97
Modelo_48h_Especialista_R1	(100)	0.97
Modelo_48h_Especialista_R2	(100)	0.96
Modelo_48h_Óptimos	(100)	0.93
Modelo_48h_Óptimos_RASGOS1	(100)	0.94
Modelo_48h_Óptimos_RASGOS2	(100)	0.93

(v/ /*)

Tabla 49

Las mayores razones de verdaderos positivos (Tabla 49) de la técnica j48 se logran con los conjuntos de datos Modelo_48h_Especialista y Modelo_48h_Especialista_RASGOS1, siendo también muy próximos a estos la del modelo con el conjunto de datos Modelo_48h_Especialista_RASGOS2.

Tabla 50 True_negative_rate

Analysing: True_negative_rate

Dataset	(1)	trees.J48
Modelo_48h_Normales	(100)	0.90
Modelo_48h_Normales_RASGO	(100)	0.93
Modelo_48h_Normales_RASGO	(100)	0.88
Modelo_48h_Espacialista	(100)	0.95
Modelo_48h_Especialista_R	(100)	0.95
Modelo_48h_Especialista_R	(100)	0.96
Modelo_48h_Óptimos	(100)	0.77
Modelo_48h_Óptimos_RASGOS	(100)	0.82
Modelo_48h_Óptimos_RASGOS	(100)	0.84

```
---------------------------------------
(v/ /*)                                |
```

El modelo j48 con el conjunto de datos Modelo_48h_Especialista_RASGOS2 muestra el mayor porcentaje de verdaderos negativos, 96 % , (Tabla 50) mientras los modelos con los conjuntos de datos Modelo_48h_Especialista y Modelo_48h_Especialista_RASGOS muestran resultados muy próximos a este.

Tabla 51 False_positive_rate

```
Analysing:  False_positive_rate
Dataset                    (1) trees.J48
---------------------------------------
Modelo_48h_Normales      (100)    0.10 |
Modelo_48h_Normales_RASGO1(100)   0.07 |
Modelo_48h_Normales_RASGO2(100)   0.12 |
Modelo_48h_Espacialista  (100)    0.05 |
Modelo_48h_Especialista_R1(100)   0.05 |
Modelo_48h_Especialista_R2(100)   0.04 |
Modelo_48h_Óptimos       (100)    0.23 |
Modelo_48h_Óptimos_RASGOS1(100)   0.18 |
Modelo_48h_Óptimos_RASGOS2(100)   0.16 |
---------------------------------------
(v/ /*)                                |
```

El modelo j48 con el conjunto de datos Modelo_48h_Especialista_RASGOS2 muestra el menor porcentaje de verdaderos negativos, 4 %, (Tabla 51) mientras los modelos con los conjuntos de datos Modelo_48h_Especialista y Modelo_48h_Especialista_RASGOS muestran resultados muy próximos a este.

Tabla 52 False_negative_rate

Tabla 53

```
Analysing:  False_negative_rate

Dataset                    (1) trees.J48
---------------------------------------
Modelo_48h_Normales      (100)    0.06 |
Modelo_48h_Normales_RASGO1(100)   0.06 |
Modelo_48h_Normales_RASGO2(100)   0.07 |
Modelo_48h_Espacialista  (100)    0.03 |
Modelo_48h_Especialista_R1(100)   0.03 |
Modelo_48h_Especialista_R2(100)   0.04 |
Modelo_48h_Óptimos       (100)    0.07 |
Modelo_48h_Óptimos_RASGOS1(100)   0.06 |
Modelo_48h_Óptimos_RASGOS2(100)   0.07 |
---------------------------------------
(v/ /*)                                |
```

Las menores razones de falsos negativos (Tabla 52) del j48 se logran con los conjuntos de datos Modelo_48h_Especialista y

Modelo_48h_Especialista_RASGOS1, siendo también muy próximos a estas la del modelo con el conjunto de datos Modelo_48h_Especialista_RASGOS2.

En resumen se puede concluir planteando que los modelos de mejor desempeño a las 48 horas son los que se corren con los conjuntos de datos granulados con el criterio del especialista, particularmente sin selección de rasgos y con selección de rasgos según el criterio rasgos2.

3.6.3 IBK (Vecino más cercano)

El experimentó se realizó con nueve modelos IBK generados por la distintas configuraciones de la conjunto de datos.

Tabla 54 Percent_correct

Analysing: Percent_correct

Dataset	(1) lazy.IBk
Modelo_48h_Normales (100)	89.52
Modelo_48h_Normales_RASGO1 (100)	92.34
Modelo_48h_Normales_RASGO2 (100)	88.07
Modelo_48h_Espacialista (100)	90.02
Modelo_48h_Especialista_R1 (100)	94.89
Modelo_48h_Especialista_R2 (100)	88.50
Modelo_48h_Óptimos (100)	88.70
Modelo_48h_Óptimos_RASGOS1 (100)	93.54
Modelo_48h_Óptimos_RASGOS2 (100)	85.80

(v/ /*)

El modelo IBK con mayor porcentaje de clasificación correcta (Tabla 53) es el corrido con el conjunto de datos Modelo_48h_Especialista_RASGOS1, con un desempeño próximo a este se encuentran los modelos corridos con los conjuntos de datos Modelo_48h_Normales_RASGOS1 y Modelo_48h_Óptimos_RASGOS1

Tabla 55 Kappa_statistic

Analysing: Kappa_statistic

Dataset	(1) lazy.IBk
Modelo_48h_Normales (100)	0.74
Modelo_48h_Normales_RASGO1 (100)	0.82
Modelo_48h_Normales_RASGO2 (100)	0.68
Modelo_48h_Espacialista (100)	0.75
Modelo_48h_Especialista_R1 (100)	0.88
Modelo_48h_Especialista_R2 (100)	0.70
Modelo_48h_Óptimos (100)	0.71
Modelo_48h_Óptimos_RASGOS1 (100)	0.84

```
Modelo_48h_Óptimos_RASGOS2(100)   0.61 |
-------------------------------------
(v/ /*)                              |
```

El mejor desempeño según el coeficiente Kappa (Tabla 54) fue 0.88, logrado con el conjunto de datos Modelo_48h_Especialista_RASGOS1.

Tabla 56 True_positive_rate

```
Analysing:  True_positive_rate

Dataset                    (1) lazy.IBk
-------------------------------------
Modelo_48h_Normales       (100)   0.96 |
Modelo_48h_Normales_RASGO1(100)   0.96 |
Modelo_48h_Normales_RASGO2(100)   0.98 |
Modelo_48h_Espacialista   (100)   0.96 |
Modelo_48h_Especialista_R1(100)   0.96 |
Modelo_48h_Especialista_R2(100)   0.97 |
Modelo_48h_Óptimos        (100)   0.96 |
Modelo_48h_Óptimos_RASGOS1(100)   0.98 |
Modelo_48h_Óptimos_RASGOS2(100)   0.98 |
-------------------------------------
(v/ /*)                              |
```

En cuanto a razón de verdaderos positivos y falsos negativos (Tabla 55 y 58) los mejores resultados se alcanzan en las corridas con los conjuntos de datos Modelo_48h_Normales_RASGOS2, Modelo_48h_Óptimos_RASGOS1 y Modelo_48h_Óptimos_RASGOS2.

Tabla 57 True_negative_rate

```
Analysing:  True_negative_rate

Dataset                    (1) lazy.IBk
-------------------------------------
Modelo_48h_Normales       (100)   0.76 |
Modelo_48h_Normales_RASGO1(100)   0.85 |
Modelo_48h_Normales_RASGO2(100)   0.67 |
Modelo_48h_Espacialista   (100)   0.77 |
Modelo_48h_Especialista_R1(100)   0.93 |
Modelo_48h_Especialista_R2(100)   0.71 |
Modelo_48h_Óptimos        (100)   0.72 |
Modelo_48h_Óptimos_RASGOS1(100)   0.84 |
Modelo_48h_Óptimos_RASGOS2(100)   0.60 |
-------------------------------------
(v/ /*)                              |
```

El porcentaje de verdaderos negativos y falsos positivos (Tabla 56 y 57) más adecuados corresponde a el conjunto de datos Modelo_48h_Especialista_RASGOS1

Tabla 58 False_positive_rate

```
Analysing:  False_positive_rate
Dataset                    (1) lazy.IBk
-------------------------------------
Modelo_48h_Normales       (100)   0.24 |
Modelo_48h_Normales_RASGO1(100)   0.15 |
Modelo_48h_Normales_RASGO2(100)   0.33 |
Modelo_48h_Espacialista   (100)   0.23 |
Modelo_48h_Especialista_R1(100)   0.07 |
Modelo_48h_Especialista_R2(100)   0.29 |
Modelo_48h_Óptimos        (100)   0.28 |
Modelo_48h_Óptimos_RASGOS1(100)   0.16 |
Modelo_48h_Óptimos_RASGOS2(100)   0.40 |
-------------------------------------
(v/ /*)                              |
```

Tabla 59 False_negative_rate

```
Analysing:  False_negative_rate

Dataset                    (1) lazy.IBk
-------------------------------------
Modelo_48h_Normales       (100)   0.04 |
Modelo_48h_Normales_RASGO1(100)   0.04 |
Modelo_48h_Normales_RASGO2(100)   0.02 |
Modelo_48h_Espacialista   (100)   0.04 |
Modelo_48h_Especialista_R1(100)   0.04 |
Modelo_48h_Especialista_R2(100)   0.03 |
Modelo_48h_Óptimos        (100)   0.04 |
Modelo_48h_Óptimos_RASGOS1(100)   0.02 |
Modelo_48h_Óptimos_RASGOS2(100)   0.02 |
-------------------------------------
(v/ /*)
```

Nota: además de los indicadores analizados en los resultados; se presentan otros indicadores, que pueden ser consultados en el anexo3 los cuales son: true_negative_rate, num_true_negative, false_negative_rate y num_false_negative,

3.7 Conclusiones de los resultados en el modelo a las 48 horas

El gráfico 5 muestra comparativamente el desempeño de cada uno de los cinco modelos de aprendizaje automático aplicados a las 48 horas del ingreso respecto a los indicadores porcentaje de clasificación correcta y coeficiente Kappa ampliado a una escala de 1 a100, donde se aprecia que es la técnica árbol de decisión la de mejor desempeño.

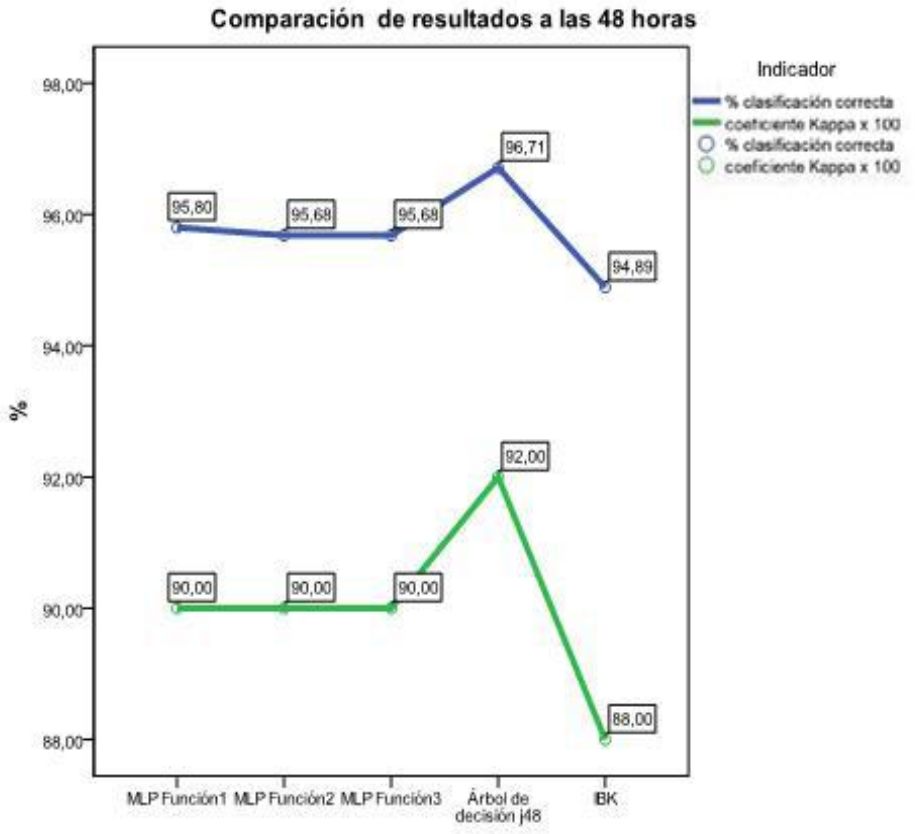

Gráfico 5 Comparación de resultados a las 48 horas

Se concluye además que el conjunto de datos Modelo_48h_Especialista_RASGOS1 es el de mejor desempeño respecto a todos los indicadores.

En la tabla 59 se presentan los mejores resultados de cada una de las técnicas aplicadas.

Tabla 60 Comparación de las técnicas a las 48 horas

Indicadores	(1) MLP	Árbol de Decisión J48	IBK (vecino más cercano)
Base de dato	Modelo_48h_Especialista_ RASGOS1	-Modelo_48h_Especialista -Modelo_48h_Especialista_ RASGOS1	Modelo_48h_Especialista_ RASGOS1
Porcent_correct	95.80	96.71	94.89
Kappa_statistc	0.90	0.92	0.88
True_positive_rate	0.97	0.97	0.96

False_positive_rate	0.07	0.05	0.07
True_negative_rate	0.93	0.95	0.93
False_Negative_rate	0.03	0.03	0.04

De ella se concluye:

- los modelos de Árbol de decisión j48 con los datos Modelo_48h_Especialista y Modelo_48h_Especialista_RASGOS1 son los de mejor desempeño para el pronóstico a 48 horas del ingreso.

-La selección de rasgos por el criterio RASGOS 1 es el que mejor contribuye al desempeño del árbol de decisión j48.

-La granulación de los datos según criterio del especialista es la que mejor contribuye al desempeño del árbol de decisión j48.

3.8 Modelo para el pronóstico a las 72 horas del ingreso

3.8.1 MLP

Para el pronóstico a las 72 horas del ingreso se experimentó con los tres modelos MLP configurados en la tabla 8 y con las nueve configuraciones del conjunto de datos para este momento del ingreso, lo cual generó la corrida de 27 modelos cuyos resultados se exponen a continuación:

Tabla 61 Percent_correct

Analysing: Percent_correct

Dataset	(1) function	(2) funct	(3) funct
Modelo_72h_Normales (100)	95.83	95.83	94.98
Modelo_72h_Normales_RASGO1 (100)	97.12	97.12	97.12
Modelo_72h_Normales_RASGO2 (100)	95.69	95.69	95.83
Modelo_72h_Especialista (100)	94.83	94.98	94.98
Modelo_72h_Especialista_R1 (100)	97.43	97.43	97.57
Modelo_72h_Especialista_R2 (100)	94.38	94.38	94.52
Modelo_72h_Óptimos (100)	90.67	90.24	90.10
Modelo_72h_Óptimos_RASGOS1 (100)	95.55	95.55	95.55
Modelo_72h_Óptimos_RASGOS2 (100)	90.55	91.10	90.67
	(v/ /*)	(0/9/0)	(0/9/0)

Según el porcentaje de casos clasificados correctamente (Tabla 60) el modelo MLP de mejor desempeño es el obtenido a partir de la función (3) con el conjunto de datos Modelo_72h_Especialista_RASGOS1, aunque todos los

modelos independiente de la cantidad de capas y neuronas cuando se corren en los conjuntos de datos Modelo_72h_Normales_RASGOS1 y Modelo_72h_Especialista_RASGOS1 muestran un desempeño próximo al mejor.

Se destaca la selección de rasgos1 como aquel que más aporta al desempeño de la técnica MLP.

Tabla 62 Kappa_statistic

Tabla 63

Analysing: Kappa_statistic

Dataset	(1) functio	(2) func	(3) func
Modelo_72h_Normales	(100) 0.88	0.88	0.86
Modelo_72h_Normales_RASGO	(100) 0.92	0.92	0.92
Modelo_72h_Normales_RASGO	(100) 0.87	0.87	0.88
Modelo_72h_Especialista	(100) 0.85	0.85	0.85
Modelo_72h_Especialista_R	(100) 0.94	0.94	0.94
Modelo_72h_Especialista_R	(100) 0.83	0.83	0.83
Modelo_72h_Óptimos	(100) 0.71	0.70	0.69
Modelo_72h_Óptimos_RASGOS	(100) 0.87	0.87	0.87
Modelo_72h_Óptimos_RASGOS	(100) 0.71	0.73	0.72
	(v/ /*)	(0/9/0)	(0/9/0)

El coeficiente Kappa es de 0.94, (Tabla 61) resultado clasificado como muy bueno.

Tabla 64 True_positive_rate

Analysing: True_positive_rate

Dataset	(1) functio	(2) func	(3) func
Modelo_72h_Normales	(100) 0.97	0.97	0.97
Modelo_72h_Normales_RASGO1	(100) 0.98	0.98	0.98
Modelo_72h_Normales_RASGO2	(100) 0.98	0.98	0.98
Modelo_72h_Especialista	(100) 0.97	0.97	0.97
Modelo_72h_Especialista_R1	(100) 0.97	0.97	0.98
Modelo_72h_Especialista_R2	(100) 0.97	0.97	0.97
Modelo_72h_Óptimos	(100) 0.97	0.96	0.96
Modelo_72h_Óptimos_RASGOS1	(100) 0.98	0.98	0.98
Modelo_72h_Óptimos_RASGOS2	(100) 0.97	0.97	0.97
	(v/ /*)	(0/9/0)	(0/9/0)

Se obtiene en verdaderos positivos (Tabla 62) el mejor resultado de 98 % con el conjunto de datos Modelo_72h_Óptimos_RASGOS1 .

Tabla 65 True_negative_rate

Analysing: True_negative_rate

Dataset	(1) functio	(2) func	(3) func
Modelo_72h_Normales	(100) 0.92	0.92	0.89
Modelo_72h_Normales_RASGO1	(100) 0.94	0.94	0.94
Modelo_72h_Normales_RASGO2	(100) 0.90	0.90	0.91
Modelo_72h_Especialista	(100) 0.88	0.88	0.88
Modelo_72h_Especialista_R1	(100) 0.98	0.98	0.98
Modelo_72h_Especialista_R2	(100) 0.85	0.85	0.85
Modelo_72h_Óptimos	(100) 0.73	0.73	0.72
Modelo_72h_Óptimos_RASGOS1	(100) 0.89	0.89	0.89
Modelo_72h_Óptimos_RASGOS2	(100) 0.73	0.75	0.74
	(v/ /*)	(0/9/0)	(0/9/0)

Un 98% de verdaderos negativos (Tabla 63) y 3 % de falsos positivos (Tabla 64) son los resultados más sobresalientes, obtenidos con el conjunto de datos Modelo_72h_Especialista_RASGOS1 .

Tabla 66 False_positive_rate

Analysing: False_positive_rate

Dataset	(1) functio	(2) func	(3) func
Modelo_72h_Normales	(100) 0.09	0.09	0.12
Modelo_72h_Normales_RASGO	(100) 0.06	0.06	0.06
Modelo_72h_Normales_RASGO	(100) 0.10	0.10	0.10
Modelo_72h_Especialista	(100) 0.13	0.13	0.13
Modelo_72h_Especialista_R	(100) 0.03	0.03	0.03
Modelo_72h_Especialista_R	(100) 0.15	0.15	0.15
Modelo_72h_Óptimos	(100) 0.27	0.28	0.28
Modelo_72h_Óptimos_RASGOS	(100) 0.12	0.12	0.12
Modelo_72h_Óptimos_RASGOS	(100) 0.27	0.26	0.26
	(v/ /*)	(0/9/0)	(0/9/0)

Tabla 67 False_negative_rate

Analysing: False_negative_rate

Dataset	(1) functio	(2) func	(3) func
Modelo_72h_Normales	(100) 0.03	0.03	0.03
Modelo_72h_Normales_RASGO	(100) 0.02	0.02	0.02
Modelo_72h_Normales_RASGO	(100) 0.02	0.02	0.02
Modelo_72h_Especialista	(100) 0.03	0.03	0.03
Modelo_72h_Especialista_R	(100) 0.03	0.03	0.02
Modelo_72h_Especialista_R	(100) 0.03	0.03	0.03
Modelo_72h_Óptimos	(100) 0.03	0.04	0.04
Modelo_72h_Óptimos_RASGOS	(100) 0.02	0.02	0.02
Modelo_72h_Óptimos_RASGOS	(100) 0.03	0.03	0.03
	(v/ /*)	(0/9/0)	(0/9/0)

En el indicador de falsos negativos (Tabla 65) muestran el mejor resultado, 2%, el modelo MLP en sus tres configuraciones y con los conjuntos de datos:

Modelo_72h_Normales_RASGOS1, Modelo_72h_Normales_RASGOS2 y Modelo_72h_Óptimos_RASGOS1.

Se concluye que el modelo de mejor desempeño es MLP obtenido a partir de la función (3) con el conjunto de datos Modelo_72h_Especialista_RASGOS1, con las demás configuraciones y con el conjunto de datos Modelo_72h_Normales_RASGOS1 se obtiene resultados muy próximos a este.

3.8.2 Árbol de decisión J48

El experimentó se realizó con nueve modelos con el empleo de la técnica j48 generados por la distintas configuraciones de la conjunto de datos.

Tabla 68 Percent_correct

Analysing: Percent_correct

Dataset	(1) trees.J48
Modelo_72h_Normales (100)	88.48
Modelo_72h_Normales_RASGO1 (100)	89.33
Modelo_72h_Normales_RASGO2 (100)	90.50
Modelo_72h_Especialista (100)	90.64
Modelo_72h_Especialista_R1 (100)	91.07
Modelo_72h_Especialista_R2 (100)	88.88
Modelo_72h_Óptimos (100)	88.67
Modelo_72h_Óptimos_RASGOS1 (100)	90.10
Modelo_72h_Óptimos_RASGOS2 (100)	90.40
(v/ /*)	

Según el porcentaje de casos clasificados correctamente (Tabla 66) el árbol de decisión J48 con el conjunto de datos Modelo_72h_Especialista_RASGOS1 fue el de mejor desempeño con un 91.07 % de aciertos. Aunque este modelo con los conjuntos de datos Modelo_72h_Normales_RASGOS2, Modelo_72h_Especialista y Modelo_72h_Óptimos_RASGOS1 alcanzan porcentajes de clasificación correcta muy próximos al de mejor desempeño.

Tabla 69 Kappa_statistic

Analysing: Kappa_statistic

Dataset	(1) trees.J48
Modelo_72h_Normales (100)	0.67
Modelo_72h_Normales_RASGO1 (100)	0.69
Modelo_72h_Normales_RASGO2 (100)	0.75
Modelo_72h_Especialista (100)	0.75
Modelo_72h_Especialista_R1 (100)	0.76
Modelo_72h_Especialista_R2 (100)	0.70
Modelo_72h_Óptimos (100)	0.67
Modelo_72h_Óptimos_RASGOS1 (100)	0.72

```
Modelo_72h_Óptimos_RASGOS2(100)  0.74 |
----------------------------------------
(v/ /*)                                |
```

Según el coeficiente Kappa (Tabla 66) el árbol de decisión J48 con el conjunto de datos Modelo_72h_Especialista_RASGOS1 fue el de mejor desempeño con un índice de 0,76. Aunque este modelo con los conjuntos de datos Modelo_72h_Normales_RASGOS2 y Modelo_72h_Especialista alcanzan índices de Kappa muy próximos al obtenido por el modelo de mejor desempeño.

Tabla 70 True_positive_rate

```
Analysing:  True_positive_rate

Dataset                    (1) trees.J48
----------------------------------------
Modelo_72h_Normales      (100)   0.94 |
Modelo_72h_Normales_RASGO1(100)  0.95 |
Modelo_72h_Normales_RASGO2(100)  0.93 |
Modelo_72h_Especialista  (100)   0.94 |
Modelo_72h_Especialista_R1(100)  0.94 |
Modelo_72h_Especialista_R2(100)  0.92 |
Modelo_72h_Óptimos       (100)   0.95 |
Modelo_72h_Óptimos_RASGOS1(100)  0.94 |
Modelo_72h_Óptimos_RASGOS2(100)  0.94 |
----------------------------------------
(v/ /*)                                |
```

Las razones de verdaderos positivos y de falsos negativos (Tabla 68 y 69) más favorables son alcanzadas por el árbol de decisión j48 con los conjuntos de datos Modelo_72h_Normales_RASGOS1 y Modelo_72h_Óptimos con 0,95 % y 5 % respectivamente.

Tabla 71 False_negative_rate

```
Analysing:  False_negative_rate

Dataset                    (1) trees.J48
----------------------------------------
Modelo_72h_Normales      (100)  0.06 |
Modelo_72h_Normales_RASGO(100)  0.05 |
Modelo_72h_Normales_RASGO(100)  0.07 |
Modelo_72h_Especialista  (100)  0.06 |
Modelo_72h_Especialista_R(100)  0.06 |
Modelo_72h_Especialista_R(100)  0.08 |
Modelo_72h_Óptimos       (100)  0.05 |
Modelo_72h_Óptimos_RASGOS(100)  0.06 |
Modelo_72h_Óptimos_RASGOS(100)  0.06 |
----------------------------------------
(v/ /*)                                |
```

Tabla 72 False_positive_rate

```
Analysing:  False_positive_rate

Dataset                    (1) trees.J48
---------------------------------------
Modelo_72h_Normales      (100)  0.27 |
Modelo_72h_Normales_RASGO(100)  0.27 |
Modelo_72h_Normales_RASGO(100)  0.16 |
Modelo_72h_Especialista  (100)  0.17 |
Modelo_72h_Especialista_R(100)  0.16 |
Modelo_72h_Especialista_R(100)  0.21 |
Modelo_72h_Óptimos       (100)  0.29 |
Modelo_72h_Óptimos_RASGOS(100)  0.21 |
Modelo_72h_Óptimos_RASGOS(100)  0.19 |
---------------------------------------
(v/ /*)                              |
```

Las razones de falsos positivos y de verdaderos negativos (Tabla 70 y 71) más favorables son alcanzadas por el árbol de decisión j48 con los conjuntos de datos Modelo_72h_Normales_RASGOS2 y Modelo_72h_Especialista_RASGOS1 con 16 % y 84 % respectivamente.

Tabla 73 True_negative_rate

```
Analysing:  True_negative_rate

Dataset                    (1) trees.J48
---------------------------------------
Modelo_72h_Normales      (100)  0.73 |
Modelo_72h_Normales_RASGO(100)  0.74 |
Modelo_72h_Normales_RASGO(100)  0.84 |
Modelo_72h_Especialista  (100)  0.83 |
Modelo_72h_Especialista_R(100)  0.84 |
Modelo_72h_Especialista_R(100)  0.79 |
Modelo_72h_Óptimos       (100)  0.71 |
Modelo_72h_Óptimos_RASGOS(100)  0.80 |
Modelo_72h_Óptimos_RASGOS(100)  0.81 |
---------------------------------------
(v/ /*)
```

Se concluye que el modelo de mejor desempeño en árbol de decisión j48 es el corrido con el conjunto de datos Modelo_72h_Especialista_RASGOS1.

3.8.3 IBK (Vecino más cercano)

El experimentó con el empleo del IBK se realizó con nueve modelos con el generados por la distintas configuraciones de la conjunto de datos.

Tabla 74 Percent_correct

```
Analysing:  Percent_correct

Dataset                    (1) lazy.IBk
```

```
---------------------------------------
Modelo_72h_Normales      (100)   87.60 |
Modelo_72h_Normales_RASGO1(100)  94.24 |
Modelo_72h_Normales_RASGO2(100)  85.62 |
Modelo_72h_Especialista  (100)   92.50 |
Modelo_72h_Especialista_R1(100)  98.00 |
Modelo_72h_Especialista_R2(100)  90.90 |
Modelo_72h_Óptimos       (100)   90.17 |
Modelo_72h_Óptimos_RASGOS1(100)  90.48 |
Modelo_72h_Óptimos_RASGOS2(100)  85.86 |
---------------------------------------
(v/ /*)                                |
```

Según el porcentaje de casos clasificados correctamente (Tabla 72 y 73) el modelo IBK de mejor desempeño fue el corrido con el conjunto de datos Modelo_72h_Especialista_RASGOS1 con un 98.00 % de aciertos, que también alcanza el mayor coeficiente de Kappa con un índice de 0,95.

Tabla 75 Kappa_statistic

```
Analysing:  Kappa_statistic

Dataset                   (1) lazy.IBk
---------------------------------------
Modelo_72h_Normales      (100)   0.63 |
Modelo_72h_Normales_RASGO1(100)  0.83 |
Modelo_72h_Normales_RASGO2(100)  0.51 |
Modelo_72h_Especialista  (100)   0.81 |
Modelo_72h_Especialista_R1(100)  0.95 |
Modelo_72h_Especialista_R2(100)  0.68 |
Modelo_72h_Óptimos       (100)   0.68 |
Modelo_72h_Óptimos_RASGOS1(100)  0.71 |
Modelo_72h_Óptimos_RASGOS2(100)  0.49 |
---------------------------------------
```

Tabla 76 True_positive_rate

```
Analysing:  True_positive_rate

Dataset                   (1) lazy.IBk
---------------------------------------
Modelo_72h_Normales      (100)   0.94 |
Modelo_72h_Normales_RASGO1(100)  0.96 |
Modelo_72h_Normales_RASGO2(100)  0.98 |
Modelo_72h_Especialista  (100)   0.94 |
Modelo_72h_Especialista_R1(100)  0.98 |
Modelo_72h_Especialista_R2(100)  1.00 |
Modelo_72h_Óptimos       (100)   0.98 |
Modelo_72h_Óptimos_RASGOS1(100)  0.96 |
Modelo_72h_Óptimos_RASGOS2(100)  1.00 |
---------------------------------------
(v/ /*)                                |
```

Los modelos IBK de mejor desempeño según los indicadores verdaderos positivos y falsos negativos (Tabla 74 y 75) fueron los corridos con los

conjuntos de datos Modelo_72h_Normales_RASGOS2, Modelo_72h_Especialista_RASGOS1 y Modelo_72h_Óptimos,con un 98 % de verdaderospositivos y un 2 % de falsos negativos.

Tabla 77 False_negative_rate

Analysing: False_negative_rate

Dataset	(1)	lazy.IBk
Modelo_72h_Normales	(100)	0.06
Modelo_72h_Normales_RASGO1	(100)	0.04
Modelo_72h_Normales_RASGO2	(100)	0.02
Modelo_72h_Especialista	(100)	0.06
Modelo_72h_Especialista_R1	(100)	0.02
Modelo_72h_Especialista_R2	(100)	0.00
Modelo_72h_Óptimos	(100)	0.02
Modelo_72h_Óptimos_RASGOS1	(100)	0.04
Modelo_72h_Óptimos_RASGOS2	(100)	0.00

(v/ /*)

Tabla 78 False_positive_rate

Analysing: False_positive_rate

Dataset	(1)	lazy.IBk
Modelo_72h_Normales	(100)	0.32
Modelo_72h_Normales_RASGO1	(100)	0.13
Modelo_72h_Normales_RASGO2	(100)	0.51
Modelo_72h_Especialista	(100)	0.12
Modelo_72h_Especialista_R1	(100)	0.02
Modelo_72h_Especialista_R2	(100)	0.36
Modelo_72h_Óptimos	(100)	0.32
Modelo_72h_Óptimos_RASGOS1	(100)	0.26
Modelo_72h_Óptimos_RASGOS2	(100)	0.54

(v/ /*)

Los modelos IBK de mejor desempeño según los indicadores falsos positivos y verdaderos negativos (Tabla 75 y 76) fue corrido con el conjunto de datos Modelo_72h_Especialista_RASGOS1, con un 2 % de falsos positivos y un 98 % de verdaderos negativos.

Tabla 79 True_negative_rate

Analysing: True_negative_rate

Dataset	(1)	lazy.IBk
Modelo_72h_Normales	(100)	0.69
Modelo_72h_Normales_RASGO1	(100)	0.88
Modelo_72h_Normales_RASGO2	(100)	0.50
Modelo_72h_Especialista	(100)	0.88
Modelo_72h_Especialista_R1	(100)	0.98
Modelo_72h_Especialista_R2	(100)	0.65
Modelo_72h_Óptimos	(100)	0.69

```
Modelo_72h_Óptimos_RASGOS1(100)   0.74 |
Modelo_72h_Óptimos_RASGOS2(100)   0.46 |
----------------------------------------
(v/ /*)                                |
```

Nota: además de los indicadores analizados en los resultados; se presentan otros indicadores, que pueden ser consultados en el anexo4 los cuales son: true_negative_rate, num_true_negative, false_negative_rate y num_false_negative,

3.9 Comparación de modelos a las 72 horas

El gráfico 6 muestra comparativamente el desempeño de cada uno de los cinco modelos de aprendizaje automático aplicados a las 72 horas del ingreso respecto a los indicadores porcentaje de clasificación correcta y coeficiente Kappa ampliado a una escala de 1 a100, donde se aprecia que es la técnica IBK la de mejor desempeño.

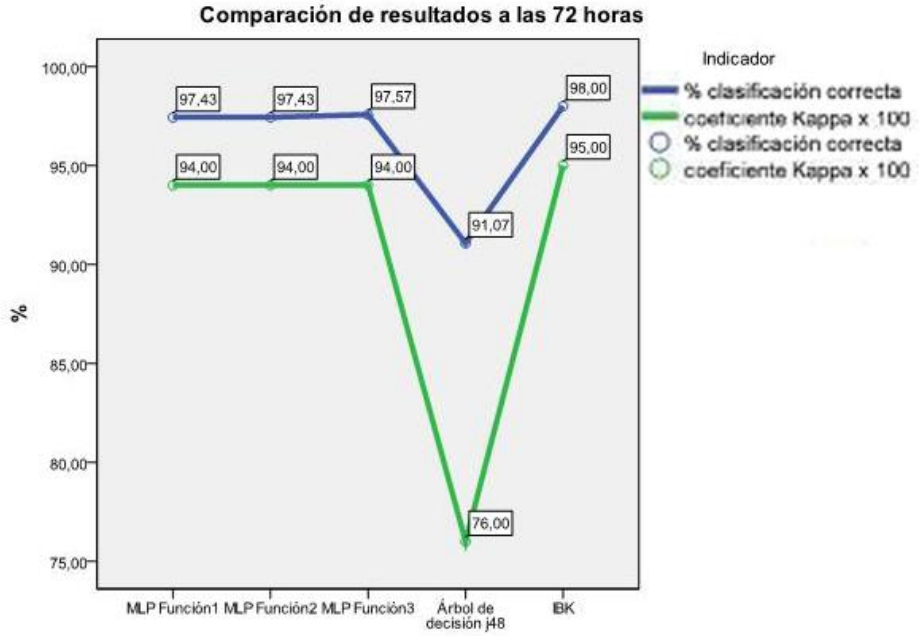

Gráfico 6 Comparación de resultados a las 72 horas

Se concluye además que el conjunto de datos Modelo_72h_Especialista_RASGOS1 fue la de mejor contribución al desempeño de las técnicas aplicadas.

En la tabla 77 se presentan los mejores resultados de cada técnica aplicada, así como el conjunto de datos con la que obtuvo el mejor desempeño de las técnicas de aprendizaje automático en la clasificación en el modelo de pronóstico a las 72 horas del ingreso.

Tabla 80 Comparación de las técnicas a las 72 horas

Indicadores	(3) MLP	Árbol de Decisión J48	IBK (vecino más cercano)
Conjunto de dato de mejor desempeño	Modelo_72h_Especialista_ RASGOS1	Modelo_72h_Especialista_ RASGOS1	Modelo_72h_Especialista_ RASGOS1
Porcent_correct	97.54	91.07	98.00
Kappa_statistc	0.94	0.76	0.95
True_positive_rate	0.98	0.94	0.98
False_positive_rate	0.03	0.16	0.02
True_negative_rate	0.98	0.84	0.98
False_Negative_rate	0.02	0.06	0.02

Las comparaciones realizadas permiten concluir que:

-Todos los modelos experimentados mostraron resultados positivos destacándose en la mayoría de los indicadores los modelos MLP e IBK.

-La preparación de los datos con el criterio de selección de rasgos RASGOS1 fue el que más aportó al desempeño de las técnicas de IA aplicadas.

- La preparación de los datos con el criterio de especialista fue el que más aportó al desempeño de las técnicas de IA aplicadas.

-El modelo de mejor desempeño en el pronóstico a las 72 horas fue IBK con el conjunto de datos Modelo_72h_Especialista_RASGOS1.

Conclusiones

Del análisis de los resultados del experimento realizado con el modelo planteado se derivan las conclusiones siguientes:

- El modelo propuesto se apoya en más de una técnica de clasificación atendiendo al momento de la misma
    - Al momento del ingreso, la técnica más eficiente en precisión al clasificar fue la **IBK,** utilizando el conjunto de datos **Modelo_Ingreso_Óptimos**.
    - A las 24 horas de ingresado, la técnica más eficiente en precisión al clasificar fue el **Árbol de decisión J48**, utilizando el conjunto de datos **Modelo_24H_Óptimos_RASGOS2**.
    - A las 48 horas de ingresado el paciente, la técnica más eficiente en precisión al clasificar fue el **Árbol de decisión J48** con el conjunto de datos **Modelo_48H_Especialista_RASGOS1**.
    - Para el modelo a las 72 horas la técnica y conjunto de datos más eficiente en precisión a la hora de clasificar en cuanto a los indicadores que se midieron es el **IBK** y el conjunto de datos **Modelo_72H_Especialista_RASGOS1**.
- Los modelos aplicados mejoraron en un porcentaje significativo el pronóstico con respecto al que se obtiene por la clasificación convencional realizada por los especialistas.
- Se utilizó la técnica de minería de datos Detector automático de interacciones Chi-cuadrado (CHAID) para la granulación de variables independientes como criterio que optimiza la relación con la variable dependiente, nombrando esta como granulación óptima.
- La granulación de los datos influyó positivamente en la calidad de clasificación.
- La granulación óptima es la que mejor contribuye al desempeño de las técnicas de aprendizaje automático en la clasificación al ingreso y a las 24 horas.

- La granulación basada en el criterio del especialista es la que mejor contribuye al desempeño de las técnicas de aprendizaje automático en la clasificación a las 48 y 72 horas.
- Entre los métodos de selección de rasgos el de mejor desempeño fue CfsSubsetEval.

Recomendaciones

- Realizar en el futuro un estudio similar al hecho en esta investigación que permita mejorar o confirmar el pronóstico con el modelo obtenido utilizando una mayor cantidad de casos.
- Elaborar un sistema informático para el pronóstico de pacientes con traumas graves de cráneo, utilizando el modelo propuesto.

Referencias Bibliográficas

[1] R. Bembibre Taboada, «Creación y validación de un instrumento para el seguimiento de pacientes con Enfermedad Cerebrovascular», Doctorado, 2003.

[2] Jacobs, Beems, Stulemeijer, Van Vugt, Borm, y Van Der Vliet, «Outcome prediction in mild traumatic brain injury: Age and clinical variables are stronger predictors than CT abnormalities», 2010.

[3] S. T. Dánel, «PRONÓSTICO DE SUPERVIVENCIA DE INFARTO CEREBRAL ATEROTROMBÓTICO USANDO APRENDIZAJE ATOMATIZADO», presented at the Temática: Gestión del Conocimiento y la Información para la Salud (Minería de datos).

[4] R. Á. Á. María del Carmen Expósito Gallardo, «Aplicaciones de la inteligencia artificial en la Medicina: perspectivas y problemas».

[5] L. Boucourt Rivera, «Su excelencia: la medicina basada en evidencias», may 2003.

[6] S. McCartney, *ENIAC: the triumphs and tragedies of the world's first computer*. Berkley Books, 2001.

[7] C. Teuscher, *Alan Turing: Life and Legacy of a Great Thinker*. Springer, 2004.

[8] H. Books, *Articles on History of Artificial Intelligence, Including: Alan Turing, Eliza, Herbert Simon, LISP Machine, Seymour Papert, Warren Sturgis McCulloch,*. Hephaestus Books, 2011.

[9] E. Rich y K. Knight, *Artificial intelligence*. McGraw-Hill, 1991.

[10] S. J. Russell, P. Norvig, E. Davis, S. J. Russell, y S. J. Russell, *Artificial intelligence: a modern approach*, vol. 2. Prentice hall Englewood Cliffs, 2010.

[11] P. H. Winston, *Artificial intelligence*. Addison-Wesley, 1984.

[12] R. S. Michalski, J. G. Carbonell, y T. M. Mitchell, *Machine Learning: An Artificial Intelligence Approach*. Tioga Publishing Company, 1994.

[13] R. L. Linscheid, J. H. Dobyns, J. W. Beabout, y R. S. Bryan, «Traumatic Instability of the Wrist Diagnosis, Classification, and Pathomechanics», *The Journal of Bone & Joint Surgery*, vol. 84, n°. 1, pp. 142–142, 2002.

[14] C. A. Pena-Reyes y M. Sipper, «A fuzzy-genetic approach to breast cancer diagnosis», *Artificial intelligence in medicine*, vol. 17, n°. 2, pp. 131–155, 1999.

[15] K. Polat, S. Sahan, H. Kodaz, y S. Günes, «A new classification method for breast cancer diagnosis: feature selection artificial immune recognition system (FS-AIRS)», in *Advances in Natural Computation*, Springer, 2005, pp. 830–838.

[16] P. Lucas, «Bayesian analysis, pattern analysis, and data mining in health care», *Current opinion in critical care*, vol. 10, n°. 5, pp. 399–403, 2004.

[17] E. K. Viikki, «Decision tree induction in the diagnosis of otoneurological diseases», *Informatics for Health and Social Care*, vol. 24, n°. 4, pp. 277–289, 1999.

[18] N. Webinar, Q. E. Roundtable, y B. W. Part, «Electronic Medical Records for Clinical Research: Application to the Identification of Heart Failure-Page 3», *Am J Manag Care*, vol. 13, n°. part 1, pp. 281–288, 2007.

[19] J. Dem\vsar, B. Zupan, N. Aoki, M. J. Wall, T. H. Granchi, y J. Robert Beck, «Feature mining and predictive model construction from severe trauma patient's data», *International journal of medical informatics*, vol. 63, n°. 1, pp. 41–50, 2001.

[20] L. Zhang, D. Samaras, D. Tomasi, N. Volkow, y R. Goldstein, «Machine learning for clinical diagnosis from functional magnetic resonance imaging», in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, vol. 1, pp. 1211–1217.

[21] D. West y V. West, «Model selection for a medical diagnostic decision support system: a breast cancer detection case», *Artificial Intelligence in medicine*, vol. 20, nº. 3, pp. 183–204, 2000.
[22] P. Cunningham, J. Carney, y S. Jacob, «Stability problems with artificial neural networks and the ensemble solution», *Artificial Intelligence in Medicine*, vol. 20, nº. 3, pp. 217–225, 2000.
[23] B. Müller y M. T. Strickland, *Neural Networks: An Introduction*. Springer, 1995.
[24] P. Peretto, *An Introduction to the Modeling of Neural Networks*. Cambridge University Press, 1992.
[25] H. César y G. P., U. de Córdoba Pedro A, «Predicción de mapas de cosecha de girasol infestado de Predicción de mapas de cosecha de girasol infestado de Ridolfia Ridolfia segetum segetum mediante Redes Neuronales Evolutivas de Unidad mediante Redes Neuronales Evolutivas de Unidad Producto», Universiada Central de Marta Abreus las Villas, abr-2008.
[26] «Clasificación de las RNA». [Online]. Available: http://www.ibiblio.org/pub/linux/docs/LuCaS/Presentaciones/200304curso-glisa/redes_neuronales/curso-glisa-redes_neuronales-html/x84.html. [Accessed: 03-mar-2013].
[27] G. Deboeck y T. Kohonen, *Visual explorations in finance: with self-organizing maps*, vol. 2. Springer London, 1998.
[28] K. Y. Lee, A. Sode-Yome, y J. H. Park, «Adaptive Hopfield neural networks for economic load dispatch», *Power Systems, IEEE Transactions on*, vol. 13, nº. 2, pp. 519–526, 1998.
[29] B. Widrow y M. A. Lehr, «Adaptive neural networks and their applications», *International journal of intelligent systems*, vol. 8, nº. 4, pp. 453–507, 1993.
[30] C. Christodoulou y M. Georgiopoulos, *Applications of neural networks in electromagnetics*. Artech House, Inc., 2000.
[31] J. J. Hopfield, «Neural networks and physical systems with emergent collective computational abilities», *Proceedings of the national academy of sciences*, vol. 79, nº. 8, pp. 2554–2558, 1982.
[32] K. A. Smith, «Neural networks for combinatorial optimization: a review of more than a decade of research», *INFORMS Journal on Computing*, vol. 11, nº. 1, pp. 15–34, 1999.
[33] T. Y. Lin y C. H. Tseng, «Optimum design for artificial neural networks: an example in a bicycle derailleur system», *Engineering Applications of Artificial Intelligence*, vol. 13, nº. 1, pp. 3–14, 2000.
[34] A. Ultsch, «Self-organizing neural networks for visualisation and classification», in *Information and classification*, Springer, 1993, pp. 307–313.
[35] T. Kohonen, G. Barna, y R. Chrisley, «Statistical pattern recognition with neural networks: Benchmarking studies», in *Neural Networks, 1988., IEEE International Conference on*, 1988, pp. 61–68.
[36] F. Y.-P. Simon y Y. Takefuji, «Stochastic neural networks for solving job-shop scheduling. II. architecture and simulations», in *Neural Networks, 1988., IEEE International Conference on*, 1988, pp. 283–290.
[37] T. Fukuda y T. Shibata, «Theory and applications of neural networks for industrial control systems», *Industrial Electronics, IEEE Transactions on*, vol. 39, nº. 6, pp. 472–489, 1992.
[40] D. C. Á. JOSÉ, «NUEVOS ENFOQUES EN APRENDIZAJE INCREMENTAL», Universidad de Málaga, 2007.

[41] M. Abdelmalik, I. Iñaki, y L. Pedro, «Clasificadores K-NN». .
[42] A. Bargiela y W. Pedrycz, «The roots of granular computing», in *Proceedings of 2006 IEEE International Conference on Granular Computing*, 2006, pp. 806–809.
[43] Y. Yao y N. Zhong, «Granular computing», *Wiley Encyclopedia of Computer Science and Engineering*, 2009.
[44] J. Stepaniuk, *Rough–granular computing in knowledge discovery and data mining*, vol. 152. Springer Verlag, 2008.
[45] Y. H. Chen y Y. Y. Yao, «Multiview intelligent data analysis based on granular computing», in *Proceedings of 2006 IEEE International Conference on Granular Computing*, 2006, pp. 281–286.
[46] T. Y. Lin y E. Louie, «Modeling the real world for data mining: granular computing approach», in *IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th*, 2001, pp. 3044–3049.
[47] D.-W. Wang, C.-J. Liau, y T. Hsu, «Medical privacy protection based on granular computing», *Artificial intelligence in medicine*, vol. 32, n°. 2, pp. 137–149, 2004.
[48] T. Y. Lin, «Granular computing on binary relations analysis of conflict and Chinese wall security policy», in *Rough Sets and Current Trends in Computing*, 2002, pp. 296–299.
[49] S. Calegari y D. Ciucci, «Granular computing applied to ontologies», *International journal of approximate reasoning*, vol. 51, n°. 4, pp. 391–409, 2010.
[50] W.-Z. Wu, Y. Leung, y J.-S. Mi, «Granular computing and knowledge reduction in formal contexts», *Knowledge and Data Engineering, IEEE Transactions on*, vol. 21, n°. 10, pp. 1461–1474, 2009.
[51] A. C. Zaily, «Aplicación de Técnicas de Clasificación de Inteligencia Artificial en el Pronóstico de Temperaturas en el Centro Meteorológico de Cienfuegos», Cienfuegos, 2010.
[52] T. Y. Lin, X. Hu, y E. Louie, «A fast association rule algorithm based on bitmap and granular computing», in *Fuzzy Systems, 2003. FUZZ'03. The 12th IEEE International Conference on*, 2003, vol. 1, pp. 678–683.

Bibliografía

[1] T. Y. Lin, X. Hu, y E. Louie, «A fast association rule algorithm based on bitmap and granular computing», in *Fuzzy Systems, 2003. FUZZ'03. The 12th IEEE International Conference on*, 2003, vol. 1, pp. 678–683.

[2] C. A. Pena-Reyes y M. Sipper, «A fuzzy-genetic approach to breast cancer diagnosis», *Artificial intelligence in medicine*, vol. 17, nº. 2, pp. 131–155, 1999.

[3] M. Gabor, «A Lazy Model-Based Approach to On-Line Classication», 1998.

[4] K. Polat, S. Sahan, H. Kodaz, y S. Günes, «A new classification method for breast cancer diagnosis: feature selection artificial immune recognition system (FS-AIRS)», in *Advances in Natural Computation*, Springer, 2005, pp. 830–838.

[5] K. Y. Lee, A. Sode-Yome, y J. H. Park, «Adaptive Hopfield neural networks for economic load dispatch», *Power Systems, IEEE Transactions on*, vol. 13, nº. 2, pp. 519–526, 1998.

[6] B. Widrow y M. A. Lehr, «Adaptive neural networks and their applications», *International journal of intelligent systems*, vol. 8, nº. 4, pp. 453–507, 1993.

[7] C. Teuscher, *Alan Turing: Life and Legacy of a Great Thinker*. Springer, 2004.

[8] R. Quinlan, «An Artificial Intelligence Approach», 1983.

[9] P. Peretto, *An Introduction to the Modeling of Neural Networks*. Cambridge University Press, 1992.

[10] A. M. Oscar J., «Análisis Detallado del Artículo "Tabu Search"».

[11] A. C. Zaily, «Aplicación de Técnicas de Clasificación de Inteligencia Artificial en el Pronóstico de Temperaturas en el Centro Meteorológico de Cienfuegos», Cienfuegos, 2010.

[12] R. Á. Á. María del Carmen Expósito Gallardo, «Aplicaciones de la inteligencia artificial en la Medicina: perspectivas y problemas».

[13] C. Christodoulou y M. Georgiopoulos, *Applications of neural networks in electromagnetics*. Artech House, Inc., 2000.

[14] H. Books, *Articles on History of Artificial Intelligence, Including: Alan Turing, Eliza, Herbert Simon, LISP Machine, Seymour Papert, Warren Sturgis McCulloch,*. Hephaestus Books, 2011.

[15] E. Rich y K. Knight, *Artificial intelligence*. McGraw-Hill, 1991.

[16] P. H. Winston, *Artificial intelligence*. Addison-Wesley, 1984.

[17] S. J. Russell, P. Norvig, E. Davis, S. J. Russell, y S. J. Russell, *Artificial intelligence: a modern approach*, vol. 2. Prentice hall Englewood Cliffs, 2010.

[18] G. F. Luger, *Artificial intelligence: Structures and strategies for complex problem solving*. Addison Wesley Longman, 2005.

[19] P. Lucas, «Bayesian analysis, pattern analysis, and data mining in health care», *Current opinion in critical care*, vol. 10, nº. 5, pp. 399–403, 2004.

[20] H. Kitano, «Building complex systems using developmental process: An engineering approach», in *Evolvable Systems: From Biology to Hardware*, Springer, 1998, pp. 218–229.

[21] «Clasificación de las RNA». [Online]. Available: http://www.ibiblio.org/pub/linux/docs/LuCaS/Presentaciones/200304curso-glisa/redes_neuronales/curso-glisa-redes_neuronales-html/x84.html. [Accessed: 03-mar-2013].

[22] M. Abdelmalik, I. Iñaki, y L. Pedro, «Clasificadores K-NN». .

[23] R. Bembibre Taboada, «Creación y validación de un instrumento para el seguimiento de pacientes con Enfermedad Cerebrovascular», Doctorado, 2003.

[24] W. Ian H, Eibe Frank, y H. Mark A, *Data Mining Practical Machine Learning Tools and Techniques*. .

[25] I. H. Witten, E. Frank, y M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques*. Elsevier, 2011.

[26] E. K. Viikki, «Decision tree induction in the diagnosis of otoneurological diseases», *Informatics for Health and Social Care*, vol. 24, nº. 4, pp. 277–289, 1999.

[27] «Educación Médica Superior - El método clínico: muerte y resurrección», 01-ene-2002. [Online]. Available: http://scielo.sld.cu/scielo.php?pid=S0864-21412000000200001&script=sci_arttext. [Accessed: 01-ene-2002].

[28] N. Webinar, Q. E. Roundtable, y B. W. Part, «Electronic Medical Records for Clinical Research: Application to the Identification of Heart Failure-Page 3», *Am J Manag Care*, vol. 13, nº. part 1, pp. 281–288, 2007.

[29] S. McCartney, *ENIAC: the triumphs and tragedies of the world's first computer*. Berkley Books, 2001.

[30] J. Dem\vsar, B. Zupan, N. Aoki, M. J. Wall, T. H. Granchi, y J. Robert Beck, «Feature mining and predictive model construction from severe trauma patient's data», *International journal of medical informatics*, vol. 63, nº. 1, pp. 41–50, 2001.

[31] Y. Yao y N. Zhong, «Granular computing», *Wiley Encyclopedia of Computer Science and Engineering*, 2009.

[32] W.-Z. Wu, Y. Leung, y J.-S. Mi, «Granular computing and knowledge reduction in formal contexts», *Knowledge and Data Engineering, IEEE Transactions on*, vol. 21, nº. 10, pp. 1461–1474, 2009.

[33] S. Calegari y D. Ciucci, «Granular computing applied to ontologies», *International journal of approximate reasoning*, vol. 51, nº. 4, pp. 391–409, 2010.

[34] T. Y. Lin, «Granular computing on binary relations analysis of conflict and Chinese wall security policy», in *Rough Sets and Current Trends in Computing*, 2002, pp. 296–299.

[35] Center for History and New Media, «Guía rápida». [Online]. Available: http://zotero.org/support/quick_start_guide.

[36] *IA ElaineRich*. .

[37] «Inteligencia artificial - EcuRed», 02-abr-2013. [Online]. Available: file:///G:/TRABAJO%20%20TESIS%205%20A%C3%91O/2013/Inteligencia_artificial.htm. [Accessed: 02-abr-2013].

[38] M. Tom M., *Machine Learning*. .

[39] L. Zhang, D. Samaras, D. Tomasi, N. Volkow, y R. Goldstein, «Machine learning for clinical diagnosis from functional magnetic resonance imaging», in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, vol. 1, pp. 1211–1217.

[40] D. Michie, D.J. Spiegelhalter, y C.C. Taylor, *Machine Learning, Neural and Statistical Classification*. 1994.

[41] R. S. Michalski, R. S. Michalski, G. Tecuci, J. G. Carbonell, y T. M. Mitchell, *Machine Learning: A Multistrategy Approach*. Morgan Kaufmann, 1994.

[42] R. S. Michalski, J. G. Carbonell, y T. M. Mitchell, *Machine Learning: An Artificial Intelligence Approach*. Tioga Publishing Company, 1994.

[43] G. M. Diego, *MANUAL DE WEKA*. .

[44] D.-W. Wang, C.-J. Liau, y T. Hsu, «Medical privacy protection based on granular computing», *Artificial intelligence in medicine*, vol. 32, nº. 2, pp. 137–149, 2004.

[45] D. West y V. West, «Model selection for a medical diagnostic decision support

system: a breast cancer detection case», *Artificial Intelligence in medicine*, vol. 20, nº. 3, pp. 183–204, 2000.

[46] T. Y. Lin y E. Louie, «Modeling the real world for data mining: granular computing approach», in *IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th*, 2001, pp. 3044–3049.

[47] S. L. Nimali, «Modelo para el pronóstico del consumo eléctrico en el Hotel Jagua», Cienfuegos, 2010.

[48] M. E. ALFONSO CARLOS, «Modelos de Regresión Basados en Redes Neuronales de Unidades Producto Diseñadas y Entrenadas Mediante Algoritmos de Optimización Híbrida. Aplicaciones», UNIVERSIDAD DE GRANADA E.T.S. de Ingeniería Informática, 2005.

[49] Y. H. Chen y Y. Y. Yao, «Multiview intelligent data analysis based on granular computing», in *Proceedings of 2006 IEEE International Conference on Granular Computing*, 2006, pp. 281–286.

[50] J. J. Hopfield, «Neural networks and physical systems with emergent collective computational abilities», *Proceedings of the national academy of sciences*, vol. 79, nº. 8, pp. 2554–2558, 1982.

[51] K. A. Smith, «Neural networks for combinatorial optimization: a review of more than a decade of research», *INFORMS Journal on Computing*, vol. 11, nº. 1, pp. 15–34, 1999.

[52] B. Müller y M. T. Strickland, *Neural Networks: An Introduction*. Springer, 1995.

[53] D. C. Á. JOSÉ, «NUEVOS ENFOQUES EN APRENDIZAJE INCREMENTAL», Universidad de Málaga, 2007.

[54] T. Y. Lin y C. H. Tseng, «Optimum design for artificial neural networks: an example in a bicycle derailleur system», *Engineering Applications of Artificial Intelligence*, vol. 13, nº. 1, pp. 3–14, 2000.

[55] Jacobs, Beems, Stulemeijer, Van Vugt, Borm, y Van Der Vliet, «Outcome prediction in mild traumatic brain injury: Age and clinical variables are stronger predictors than CT abnormalities», 2010.

[56] R. Epstein, G. Roberts, y G. Beber, *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Springer, 2008.

[57] A. M. Tyrrell, E. Sanchez, D. Floreano, G. Tempesti, D. Mange, J.-M. Moreno, J. Rosenberg, y A. E. Villa, «Poetic tissue: An integrated architecture for bio-inspired hardware», in *Evolvable Systems: From Biology to Hardware*, Springer, 2003, pp. 129–140.

[58] H. César y G. P., U. de Córdoba Pedro A, «Predicción de mapas de cosecha de girasol infestado de Ridolfia Ridolfia segetum segetum mediante Redes Neuronales Evolutivas de Unidad del Producto», Universiada Central de Marta Abreus las Villas, may-2008.

[59] S. T. Dánel, «PRONÓSTICO DE SUPERVIVENCIA DE INFARTO CEREBRAL ATEROTROMBÓTICO USANDO APRENDIZAJE ATOMATIZADO», presented at the Temática: Gestión del Conocimiento y la Información para la Salud (Minería de datos).

[60] H. M. D. Janny, «Propuesta de técnica de Inteligencia Artificial para la detección de anomalías en la red de datos de la Universidad de Cienfuegos.», Carlos Rafael Rodríguez, 2009.

[61] «Revista Cubana de Medicina Militar - Comportamiento de factores pronósticos clínicos y demográficos relacionados con el traumatismo craneoencefálico», 02-abr-2013. [Online]. Available: http://scielo.sld.cu/scielo.php?pid=S0138-65572010000200004&script=sci_arttext. [Accessed: 02-abr-2013].

[62] H. Vafaie y K. De Jong, «Robust feature selection algorithms», in *Tools with Artificial Intelligence, 1993. TAI'93. Proceedings., Fifth International Conference on*, 1993, pp. 356–363.

[63] J. Stepaniuk, *Rough–granular computing in knowledge discovery and data mining*, vol. 152. Springer Verlag, 2008.

[64] C. O. Angel, R. B. Rocioo, y A. D. Yurlenisz, «SELECCIÓN DE ATRIBUTOS PREDICTIVOS DEL RENDIMIENTO ACADÉMICO DE ESTUDIANTES EN UN MODELO DE B--- LEARNING», semtiebre 2011.

[65] A. Ultsch, «Self-organizing neural networks for visualisation and classification», in *Information and classification*, Springer, 1993, pp. 307–313.

[66] P. Cunningham, J. Carney, y S. Jacob, «Stability problems with artificial neural networks and the ensemble solution», *Artificial Intelligence in Medicine*, vol. 20, nº. 3, pp. 217–225, 2000.

[67] T. Kohonen, G. Barna, y R. Chrisley, «Statistical pattern recognition with neural networks: Benchmarking studies», in *Neural Networks, 1988., IEEE International Conference on*, 1988, pp. 61–68.

[68] F. Y.-P. Simon y Y. Takefuji, «Stochastic neural networks for solving job-shop scheduling. II. architecture and simulations», in *Neural Networks, 1988., IEEE International Conference on*, 1988, pp. 283–290.

[69] L. Boucourt Rivera, «Su excelencia: la medicina basada en evidencias», may 2003.

[70] R. Rojas y U. Hashagen, *The First Computers: History and Architectures*. MIT Press, 2002.

[71] A. Bargiela y W. Pedrycz, «The roots of granular computing», 2006, pp. 806–809.

[72] T. Fukuda y T. Shibata, «Theory and applications of neural networks for industrial control systems», *Industrial Electronics, IEEE Transactions on*, vol. 39, nº. 6, pp. 472–489, 1992.

[73] R. L. Linscheid, J. H. Dobyns, J. W. Beabout, y R. S. Bryan, «Traumatic Instability of the Wrist Diagnosis, Classification, and Pathomechanics», *The Journal of Bone & Joint Surgery*, vol. 84, nº. 1, pp. 142–142, 2002.

[74] «UPV-EHU - Facultad de Informática - Minería de Datos», 02-abr-2013. [Online]. Available: file:///G:/TRABAJO%20%20TESIS%205%20A%C3%91O/19.12/practicas1213.htm. [Accessed: 02-abr-2013].

[75] G. Deboeck y T. Kohonen, *Visual explorations in finance: with self-organizing maps*, vol. 2. Springer London, 1998.

Anexos

Anexo1 Indicadores al momento del Ingreso

MLP

```
Analysing:  Num_true_positives

Dataset                    (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_Ingreso_Normales  (100)   4.63 |   4.67     4.59
Modelo_Ingreso_Especialis(100)   4.55 |   4.50     4.55
Modelo_Ingreso_Óptimos   (100)   4.86 |   4.85     4.88
-----------------------------------------------------------
                              (v/ /*) |  (0/3/0)  (0/3/0)


Analysing:  False_positive_rate

Dataset                    (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_Ingreso_Normales  (100)   0.20 |   0.22     0.23
Modelo_Ingreso_Especialis(100)   0.30 |   0.30     0.30
Modelo_Ingreso_Óptimos   (100)   0.23 |   0.19     0.20
-----------------------------------------------------------
                              (v/ /*) |  (0/3/0)  (0/3/0)
Analysing:  Num_false_positives

Dataset                     (1) functio | (2) func (3) func
------------------------------------------------------------
Modelo_Ingreso_Normales   (100)   0.61 |   0.65     0.68
Modelo_Ingreso_Especialis (100)   0.90 |   0.90     0.90
Modelo_Ingreso_Óptimos    (100)   0.69 |   0.58     0.59
------------------------------------------------------------
                               (v/ /*) |  (0/3/0)  (0/3/0)



.Analysing:  Num_true_negatives

Dataset                     (1) functio | (2) func (3) func
------------------------------------------------------------
Modelo_Ingreso_Normales  (100)    2.39 |   2.35     2.32
Modelo_Ingreso_Especialis(100)    2.10 |   2.10     2.10
Modelo_Ingreso_Óptimos   (100)    2.31 |   2.42     2.41
------------------------------------------------------------
                               (v/ /*) |  (0/3/0)  (0/3/0)

Analysing:  False_negative_rate

Dataset                     (1) functio | (2) func (3) func
------------------------------------------------------------
Modelo_Ingreso_Normales  (100)    0.13 |   0.12     0.14
Modelo_Ingreso_Especialis(100)    0.14 |   0.15     0.14
Modelo_Ingreso_Óptimos   (100)    0.08 |   0.09     0.08
------------------------------------------------------------
                               (v/ /*) |  (0/3/0)  (0/3/0)
```

```
Analysing:  Num_false_negatives

Dataset                    (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_Ingreso_Normales  (100)   0.67 |   0.63     0.71
Modelo_Ingreso_Especialis(100)   0.75 |   0.80     0.75
Modelo_Ingreso_Óptimos   (100)   0.44 |   0.45     0.42
-----------------------------------------------------------
                              (v/ /*) |  (0/3/0)  (0/3/0)
```

J48

```
Analysing:  Num_true_positives

Dataset                    (1) trees.J48
----------------------------------------
Modelo_Ingreso_Normales  (100)   4.58 |
Modelo_Ingreso_Especialis(100)   4.51 |
Modelo_Ingreso_Óptimos   (100)   4.77 |
----------------------------------------
(v/ /*)                               |

Analysing:  False_positive_rate

Dataset                    (1) trees.J48
----------------------------------------
Modelo_Ingreso_Normales  (100)   0.36 |
Modelo_Ingreso_Especialis(100)   0.32 |
Modelo_Ingreso_Óptimos   (100)   0.27 |
----------------------------------------
(v/ /*)                               |

Analysing:  Num_false_positives

Dataset                    (1) trees.J48
----------------------------------------
Modelo_Ingreso_Normales  (100)    1.08 |
Modelo_Ingreso_Especialis(100)    0.95 |
Modelo_Ingreso_Óptimos   (100)    0.82 |
----------------------------------------
(v/ /*)                                |


Analysing:  Num_true_negatives

Dataset                    (1) trees.J48
----------------------------------------
Modelo_Ingreso_Normales  (100)   1.92 |
Modelo_Ingreso_Especialis(100)   2.05 |
Modelo_Ingreso_Óptimos   (100)   2.18 |
----------------------------------------
(v/ /*)                               |

Analysing:  False_negative_rate

Dataset                    (1) trees.J48
----------------------------------------
Modelo_Ingreso_Normales  (100)   0.14 |
```

```
Modelo_Ingreso_Especialis(100)   0.15 |
Modelo_Ingreso_Óptimos   (100)   0.10 |
----------------------------------------
(v/ /*)                               |

Analysing:  Num_false_negatives

Dataset                   (1) trees.J48
----------------------------------------
Modelo_Ingreso_Normales  (100)   0.72 |
Modelo_Ingreso_Especialis(100)   0.79 |
Modelo_Ingreso_Óptimos   (100)   0.53 |
----------------------------------------
(v/ /*)                               |
```

IBK

```
Analysing:  Num_true_positives

Dataset                   (1) lazy.IBk
----------------------------------------
Modelo_Ingreso_Normales  (100)   4.89 |
Modelo_Ingreso_Especialis(100)   4.88 |
Modelo_Ingreso_Óptimos   (100)   5.27 |
----------------------------------------
(v/ /*)                               |

Analysing:  False_positive_rate

Dataset                   (1) lazy.IBk
----------------------------------------
Modelo_Ingreso_Normales  (100)   0.37 |
Modelo_Ingreso_Especialis(100)   0.36 |
Modelo_Ingreso_Óptimos   (100)   0.21 |
----------------------------------------
(v/ /*)                               |

Analysing:  Num_false_positives

Dataset                   (1) lazy.IBk
----------------------------------------
Modelo_Ingreso_Normales  (100)   1.11 |
Modelo_Ingreso_Especialis(100)   1.07 |
Modelo_Ingreso_Óptimos   (100)   0.62 |
----------------------------------------
(v/ /*)                               |


Analysing:  Num_true_negatives

Dataset                   (1) lazy.IBk
----------------------------------------
Modelo_Ingreso_Normales  (100)   1.89 |
Modelo_Ingreso_Especialis(100)   1.93 |
Modelo_Ingreso_Óptimos   (100)   2.38 |
----------------------------------------
(v/ /*)                               |

Analysing:  False_negative_rate
```

```
Dataset                   (1) lazy.IBk
----------------------------------------
Modelo_Ingreso_Normales  (100)   0.08 |
Modelo_Ingreso_Especialis(100)   0.08 |
Modelo_Ingreso_Óptimos   (100)   0.01 |
----------------------------------------
(v/ /*)                               |

Analysing:  Num_false_negatives

Dataset                   (1) lazy.IBk
----------------------------------------
Modelo_Ingreso_Normales  (100)   0.41 |
Modelo_Ingreso_Especialis(100)   0.42 |
Modelo_Ingreso_Óptimos   (100)   0.03 |
----------------------------------------
(v/ /*)                               |
```

Anexo2 Indicadores a las 24 horas de ingresado

MLP

```
Analysing:  Num_false_positives

Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_24H_Normales      (100)  0.46 |   0.37     0.38
Modelo_24h_Normales_RASGO(100)  0.35 |   0.26     0.27
Modelo_24h_Normales_RASGO(100)  0.47 |   0.45     0.39
Modelo_24H_Especialista  (100)  0.46 |   0.46     0.45
Modelo_24h_Especialista_R(100)  0.24 |   0.24     0.25
Modelo_24h_Especialista_R(100)  0.33 |   0.33     0.30
Modelo_24H_Óptimo        (100)  0.43 |   0.41     0.39
Modelo_24h_Óptimos_RASGOS(100)  0.34 |   0.31     0.29
Modelo_24h_Óptimos_RASGOS(100)  0.40 |   0.40     0.40
-----------------------------------------------------------
                            (v/ /*) |  (0/9/0)  (0/9/0)

Analysing:  Num_true_positives
Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_24H_Normales      (100)  4.92 |   4.92     4.92
Modelo_24h_Normales_RASGO(100)  5.07 |   5.07     5.06
Modelo_24h_Normales_RASGO(100)  4.95 |   4.93     4.93
Modelo_24H_Especialista  (100)  4.99 |   4.99     4.98
Modelo_24h_Especialista_R(100)  4.94 |   4.94     4.94
Modelo_24h_Especialista_R(100)  4.93 |   4.93     4.92
Modelo_24H_Óptimo        (100)  4.96 |   4.96     4.96
Modelo_24h_Óptimos_RASGOS(100)  5.03 |   5.03     5.05
Modelo_24h_Óptimos_RASGOS(100)  4.95 |   4.89     4.90
-----------------------------------------------------------
                            (v/ /*) |  (0/9/0)  (0/9/0)

Analysing:  Num_true_negatives

Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_24H_Normales      (100)  2.54 |   2.63     2.62
Modelo_24h_Normales_RASGO(100)  2.65 |   2.74     2.73
Modelo_24h_Normales_RASGO(100)  2.53 |   2.55     2.61
Modelo_24H_Especialista  (100)  2.54 |   2.54     2.55
Modelo_24h_Especialista_R(100)  2.76 |   2.76     2.75
Modelo_24h_Especialista_R(100)  2.67 |   2.67     2.70
Modelo_24H_Óptimo        (100)  2.57 |   2.59     2.61
Modelo_24h_Óptimos_RASGOS(100)  2.66 |   2.69     2.71
Modelo_24h_Óptimos_RASGOS(100)  2.60 |   2.60     2.60
-----------------------------------------------------------
                            (v/ /*) |  (0/9/0)  (0/9/0)
```

El número de casos de vivos clasificados; y que son vivo es de 2.74 este número es un promedio.

```
Analysing:  Num_false_negatives

Dataset                  (1) functio | (2) func (3) func
```

```
-------------------------------------------------------------
Modelo_24H_Normales      (100)  0.38 |   0.38     0.38
Modelo_24h_Normales_RASGO(100)  0.23 |   0.23     0.24
Modelo_24h_Normales_RASGO(100)  0.35 |   0.37     0.37
Modelo_24H_Especialista  (100)  0.31 |   0.31     0.32
Modelo_24h_Especialista_R(100)  0.36 |   0.36     0.36
Modelo_24h_Especialista_R(100)  0.37 |   0.37     0.38
Modelo_24H_Óptimo        (100)  0.34 |   0.34     0.34
Modelo_24h_Óptimos_RASGOS(100)  0.27 |   0.27     0.25
Modelo_24h_Óptimos_RASGOS(100)  0.35 |   0.41     0.40
-------------------------------------------------------------
                           (v/ /*) |  (0/9/0)  (0/9/0)
```

J84

```
Analysing:  Num_true_positives

Dataset                   (1) trees.J48
----------------------------------------
Modelo_24H_Normales      (100)   4.95 |
Modelo_24h_Normales_RASGO(100)   5.00 |
Modelo_24h_Normales_RASGO(100)   4.78 |
Modelo_24H_Especialista  (100)   4.69 |
Modelo_24h_Especialista_R(100)   4.79 |
Modelo_24h_Especialista_R(100)   4.94 |
Modelo_24H_Óptimo        (100)   4.80 |
Modelo_24h_Óptimos_RASGOS(100)   5.00 |
Modelo_24h_Óptimos_RASGOS(100)   5.08 |
----------------------------------------
(v/ /*)                                |


Analysing:  Num_false_positives
Dataset                   (1) trees.J48
----------------------------------------
Modelo_24H_Normales      (100)   0.58 |
Modelo_24h_Normales_RASGO(100)   0.53 |
Modelo_24h_Normales_RASGO(100)   0.39 |
Modelo_24H_Especialista  (100)   0.64 |
Modelo_24h_Especialista_R(100)   0.60 |
Modelo_24h_Especialista_R(100)   0.21 |
Modelo_24H_Óptimo        (100)   0.48 |
Modelo_24h_Óptimos_RASGOS(100)   0.43 |
Modelo_24h_Óptimos_RASGOS(100)   0.10 |
----------------------------------------
(v/ /*)                                |


Analysing:  Num_true_negatives

Dataset                   (1) trees.J48
----------------------------------------
Modelo_24H_Normales      (100)   2.42 |
Modelo_24h_Normales_RASGO(100)   2.47 |
Modelo_24h_Normales_RASGO(100)   2.61 |
Modelo_24H_Especialista  (100)   2.36 |
Modelo_24h_Especialista_R(100)   2.40 |
Modelo_24h_Especialista_R(100)   2.79 |
Modelo_24H_Óptimo        (100)   2.52 |
Modelo_24h_Óptimos_RASGOS(100)   2.57 |
```

```
Modelo_24h_Óptimos_RASGOS(100)   2.90 |
----------------------------------------
(v/ /*)                                |


Analysing:  Num_false_negatives

Dataset                   (1) trees.J48
----------------------------------------
Modelo_24H_Normales      (100)   0.35 |
Modelo_24h_Normales_RASGO(100)   0.30 |
Modelo_24h_Normales_RASGO(100)   0.52 |
Modelo_24H_Especialista  (100)   0.61 |
Modelo_24h_Especialista_R(100)   0.51 |
Modelo_24h_Especialista_R(100)   0.36 |
Modelo_24H_Óptimo        (100)   0.50 |
Modelo_24h_Óptimos_RASGOS(100)   0.30 |
Modelo_24h_Óptimos_RASGOS(100)   0.22 |
----------------------------------------
(v/ /*)                                |
```

IBK

```
Analysing:  Num_true_positives

Dataset                   (1) lazy.IBk
----------------------------------------
Modelo_24H_Normales      (100)  5.08 |
Modelo_24h_Normales_RASGO(100)  5.06 |
Modelo_24h_Normales_RASGO(100)  5.18 |
Modelo_24H_Especialista  (100)  5.19 |
Modelo_24h_Especialista_R(100)  5.10 |
Modelo_24h_Especialista_R(100)  5.21 |
Modelo_24H_Óptimo        (100)  5.20 |
Modelo_24h_Óptimos_RASGOS(100)  5.12 |
Modelo_24h_Óptimos_RASGOS(100)  5.22 |
----------------------------------------
(v/ /*)                                |


Analysing:  Num_false_positives

Dataset                   (1) lazy.IBk
----------------------------------------
Modelo_24H_Normales      (100)  0.68 |
Modelo_24h_Normales_RASGO(100)  0.41 |
Modelo_24h_Normales_RASGO(100)  0.83 |
Modelo_24H_Especialista  (100)  0.63 |
Modelo_24h_Especialista_R(100)  0.38 |
Modelo_24h_Especialista_R(100)  1.05 |
Modelo_24H_Óptimo        (100)  0.39 |
Modelo_24h_Óptimos_RASGOS(100)  0.60 |
Modelo_24h_Óptimos_RASGOS(100)  0.70 |
----------------------------------------
(v/ /*)                                |


Analysing:  Num_true_negatives

Dataset                   (1) lazy.IBk
```

```
---------------------------------------
Modelo_24H_Normales      (100)  2.32 |
Modelo_24h_Normales_RASGO(100)  2.59 |
Modelo_24h_Normales_RASGO(100)  2.17 |
Modelo_24H_Especialista  (100)  2.37 |
Modelo_24h_Especialista_R(100)  2.62 |
Modelo_24h_Especialista_R(100)  1.95 |
Modelo_24H_Óptimo        (100)  2.61 |
Modelo_24h_Óptimos_RASGOS(100)  2.40 |
Modelo_24h_Óptimos_RASGOS(100)  2.30 |
---------------------------------------
(v/ /*)                               |


Analysing:  Num_false_negatives

Dataset                   (1) lazy.IBk
---------------------------------------
Modelo_24H_Normales      (100)  0.22 |
Modelo_24h_Normales_RASGO(100)  0.24 |
Modelo_24h_Normales_RASGO(100)  0.12 |
Modelo_24H_Especialista  (100)  0.11 |
Modelo_24h_Especialista_R(100)  0.20 |
Modelo_24h_Especialista_R(100)  0.09 |
Modelo_24H_Óptimo        (100)  0.10 |
Modelo_24h_Óptimos_RASGOS(100)  0.18 |
Modelo_24h_Óptimos_RASGOS(100)  0.08 |
---------------------------------------
(v/ /*)                               |
```

Anexo3 Indicadores a las 48 horas de ingresado

MLP

```
Analysing:  Num_false_positives

Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_48h_Normales      (100)  0.44 |   0.45     0.46
Modelo_48h_Normales_RASGO(100)  0.37 |   0.37     0.36
Modelo_48h_Normales_RASGO(100)  0.50 |   0.50     0.49
Modelo_48h_Espacialista  (100)  0.20 |   0.20     0.20
Modelo_48h_Especialista_R(100)  0.19 |   0.20     0.20
Modelo_48h_Especialista_R(100)  0.14 |   0.13     0.13
Modelo_48h_Óptimos       (100)  0.50 |   0.48     0.49
Modelo_48h_Óptimos_RASGOS(100)  0.37 |   0.36     0.38
Modelo_48h_Óptimos_RASGOS(100)  0.44 |   0.43     0.45
-----------------------------------------------------------
                             (v/ /*) |  (0/9/0)  (0/9/0)

Analysing:  Num_true_positives

Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_48h_Normales      (100)  5.10 |   5.10     5.08
Modelo_48h_Normales_RASGO(100)  5.20 |   5.20     5.20
Modelo_48h_Normales_RASGO(100)  4.97 |   4.97     4.97
Modelo_48h_Espacialista  (100)  5.09 |   5.10     5.15
Modelo_48h_Especialista_R(100)  5.16 |   5.16     5.16
Modelo_48h_Especialista_R(100)  5.06 |   5.03     5.03
Modelo_48h_Óptimos       (100)  5.19 |   5.19     5.18
Modelo_48h_Óptimos_RASGOS(100)  5.13 |   5.18     5.14
Modelo_48h_Óptimos_RASGOS(100)  5.12 |   5.11     5.11
-----------------------------------------------------------
                             (v/ /*) |  (0/9/0)  (0/9/0)


Analysing:  Num_true_negatives

Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_48h_Normales      (100)  2.06 |   2.05     2.04
Modelo_48h_Normales_RASGO(100)  2.13 |   2.13     2.14
Modelo_48h_Normales_RASGO(100)  2.00 |   2.00     2.01
Modelo_48h_Espacialista  (100)  2.30 |   2.30     2.30
Modelo_48h_Especialista_R(100)  2.31 |   2.30     2.30
Modelo_48h_Especialista_R(100)  2.36 |   2.37     2.37
Modelo_48h_Óptimos       (100)  2.00 |   2.02     2.01
Modelo_48h_Óptimos_RASGOS(100)  2.13 |   2.14     2.12
Modelo_48h_Óptimos_RASGOS(100)  2.06 |   2.07     2.05
-----------------------------------------------------------
                             (v/ /*) |  (0/9/0)  (0/9/0)


Analysing:  Num_false_negatives

Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_48h_Normales      (100)  0.20 |   0.20     0.22
```

```
Modelo_48h_Normales_RASGO(100)  0.10 |   0.10     0.10
Modelo_48h_Normales_RASGO(100)  0.33 |   0.33     0.33
Modelo_48h_Espacialista  (100)  0.21 |   0.20     0.15
Modelo_48h_Especialista_R(100)  0.14 |   0.14     0.14
Modelo_48h_Especialista_R(100)  0.24 |   0.27     0.27
Modelo_48h_Óptimos       (100)  0.11 |   0.11     0.12
Modelo_48h_Óptimos_RASGOS(100)  0.17 |   0.12     0.16
Modelo_48h_Óptimos_RASGOS(100)  0.18 |   0.19     0.19
-----------------------------------------------------------
                             (v/ /*) |  (0/9/0)  (0/9/0)
```

J48

```
Analysing:  Num_true_positives

Dataset                   (1) trees.J48
---------------------------------------
Modelo_48h_Normales      (100)  4.98 |
Modelo_48h_Normales_RASGO(100)  4.98 |
Modelo_48h_Normales_RASGO(100)  4.94 |
Modelo_48h_Espacialista  (100)  5.16 |
Modelo_48h_Especialista_R(100)  5.16 |
Modelo_48h_Especialista_R(100)  5.10 |
Modelo_48h_Óptimos       (100)  4.95 |
Modelo_48h_Óptimos_RASGOS(100)  4.98 |
Modelo_48h_Óptimos_RASGOS(100)  4.93 |
---------------------------------------
(v/ /*)                               |


Analysing:  Num_false_positives

Dataset                   (1) trees.J48
---------------------------------------
Modelo_48h_Normales      (100)  0.27 |
Modelo_48h_Normales_RASGO(100)  0.19 |
Modelo_48h_Normales_RASGO(100)  0.33 |
Modelo_48h_Espacialista  (100)  0.12 |
Modelo_48h_Especialista_R(100)  0.12 |
Modelo_48h_Especialista_R(100)  0.10 |
Modelo_48h_Óptimos       (100)  0.56 |
Modelo_48h_Óptimos_RASGOS(100)  0.46 |
Modelo_48h_Óptimos_RASGOS(100)  0.41 |
---------------------------------------
(v/ /*)                               |


Analysing:  Num_true_negatives

Dataset                   (1) trees.J48
---------------------------------------
Modelo_48h_Normales      (100)  2.23 |
Modelo_48h_Normales_RASGO(100)  2.31 |
Modelo_48h_Normales_RASGO(100)  2.17 |
Modelo_48h_Espacialista  (100)  2.38 |
Modelo_48h_Especialista_R(100)  2.38 |
Modelo_48h_Especialista_R(100)  2.40 |
Modelo_48h_Óptimos       (100)  1.94 |
Modelo_48h_Óptimos_RASGOS(100)  2.04 |
Modelo_48h_Óptimos_RASGOS(100)  2.09 |
```

```
---------------------------------------
(v/ /*)                                |

Analysing:  Num_false_negatives

Dataset                    (1) trees.J48
---------------------------------------
Modelo_48h_Normales      (100)  0.32 |
Modelo_48h_Normales_RASGO(100)  0.32 |
Modelo_48h_Normales_RASGO(100)  0.36 |
Modelo_48h_Espacialista  (100)  0.14 |
Modelo_48h_Especialista_R(100)  0.14 |
Modelo_48h_Especialista_R(100)  0.20 |
Modelo_48h_Óptimos       (100)  0.35 |
Modelo_48h_Óptimos_RASGOS(100)  0.32 |
Modelo_48h_Óptimos_RASGOS(100)  0.37 |
---------------------------------------
(v/ /*)                                |
```

IBK

```
Analysing:  Num_true_positives

Dataset                    (1) lazy.IBk
---------------------------------------
Modelo_48h_Normales      (100)  5.10 |
Modelo_48h_Normales_RASGO(100)  5.10 |
Modelo_48h_Normales_RASGO(100)  5.19 |
Modelo_48h_Espacialista  (100)  5.10 |
Modelo_48h_Especialista_R(100)  5.10 |
Modelo_48h_Especialista_R(100)  5.13 |
Modelo_48h_Óptimos       (100)  5.11 |
Modelo_48h_Óptimos_RASGOS(100)  5.20 |
Modelo_48h_Óptimos_RASGOS(100)  5.20 |
---------------------------------------
(v/ /*)                                |

Analysing:  Num_false_negatives

Dataset                    (1) lazy.IBk
---------------------------------------
Modelo_48h_Normales      (100)  0.20 |
Modelo_48h_Normales_RASGO(100)  0.20 |
Modelo_48h_Normales_RASGO(100)  0.11 |
Modelo_48h_Espacialista  (100)  0.20 |
Modelo_48h_Especialista_R(100)  0.20 |
Modelo_48h_Especialista_R(100)  0.17 |
Modelo_48h_Óptimos       (100)  0.19 |
Modelo_48h_Óptimos_RASGOS(100)  0.10 |
Modelo_48h_Óptimos_RASGOS(100)  0.10 |
---------------------------------------
(v/ /*)                                |

Analysing:  Num_false_positives

Dataset                    (1) lazy.IBk
---------------------------------------
Modelo_48h_Normales      (100)  0.62 |
Modelo_48h_Normales_RASGO(100)  0.40 |
Modelo_48h_Normales_RASGO(100)  0.82 |
Modelo_48h_Espacialista  (100)  0.58 |
Modelo_48h_Especialista_R(100)  0.20 |
```

```
Modelo_48h_Especialista_R(100)  0.73 |
Modelo_48h_Óptimos       (100)  0.69 |
Modelo_48h_Óptimos_RASGOS(100)  0.40 |
Modelo_48h_Óptimos_RASGOS(100)  1.01 |
---------------------------------------
(v/ /*)                               |
```

```
Analysing:  Num_true_negatives

Dataset                   (1) lazy.IBk
---------------------------------------
Modelo_48h_Normales      (100)  1.88 |
Modelo_48h_Normales_RASGO(100)  2.10 |
Modelo_48h_Normales_RASGO(100)  1.68 |
Modelo_48h_Espacialista  (100)  1.92 |
Modelo_48h_Especialista_R(100)  2.30 |
Modelo_48h_Especialista_R(100)  1.77 |
Modelo_48h_Óptimos       (100)  1.81 |
Modelo_48h_Óptimos_RASGOS(100)  2.10 |
Modelo_48h_Óptimos_RASGOS(100)  1.49 |
---------------------------------------
(v/ /*)                               |
```

Anexo4 Indicadores a las 72 horas de ingresado

MPL

```
Analysing:  Num_true_positives

Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_72h_Normales      (100)  4.96 |   4.96     4.96
Modelo_72h_Normales_RASGO(100)  5.00 |   5.00     5.00
Modelo_72h_Normales_RASGO(100)  4.98 |   4.98     4.98
Modelo_72h_Especialista  (100)  4.96 |   4.97     4.97
Modelo_72h_Especialista_R(100)  4.97 |   4.97     4.98
Modelo_72h_Especialista_R(100)  4.96 |   4.96     4.97
Modelo_72h_Óptimos       (100)  4.93 |   4.91     4.91
Modelo_72h_Óptimos_RASGOS(100)  5.00 |   5.00     5.00
Modelo_72h_Óptimos_RASGOS(100)  4.93 |   4.94     4.93
-----------------------------------------------------------
                             (v/ /*) |  (0/9/0)  (0/9/0)


Analysing:  Num_false_positives

Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_72h_Normales      (100)  0.15 |   0.15     0.21
Modelo_72h_Normales_RASGO(100)  0.10 |   0.10     0.10
Modelo_72h_Normales_RASGO(100)  0.18 |   0.18     0.17
Modelo_72h_Especialista  (100)  0.22 |   0.22     0.22
Modelo_72h_Especialista_R(100)  0.05 |   0.05     0.05
Modelo_72h_Especialista_R(100)  0.25 |   0.25     0.25
Modelo_72h_Óptimos       (100)  0.48 |   0.49     0.50
Modelo_72h_Óptimos_RASGOS(100)  0.21 |   0.21     0.21
Modelo_72h_Óptimos_RASGOS(100)  0.49 |   0.46     0.48
-----------------------------------------------------------
                             (v/ /*) |  (0/9/0)  (0/9/0)

Analysing:  Num_true_negatives

Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
Modelo_72h_Normales      (100)  1.65 |   1.65     1.59
Modelo_72h_Normales_RASGO(100)  1.70 |   1.70     1.70
Modelo_72h_Normales_RASGO(100)  1.62 |   1.62     1.63
Modelo_72h_Especialista  (100)  1.58 |   1.58     1.58
Modelo_72h_Especialista_R(100)  1.75 |   1.75     1.75
Modelo_72h_Especialista_R(100)  1.55 |   1.55     1.55
Modelo_72h_Óptimos       (100)  1.32 |   1.31     1.30
Modelo_72h_Óptimos_RASGOS(100)  1.59 |   1.59     1.59
Modelo_72h_Óptimos_RASGOS(100)  1.31 |   1.34     1.32
-----------------------------------------------------------
                             (v/ /*) |  (0/9/0)  (0/9/0)


Analysing:  Num_false_negatives

Dataset                  (1) functio | (2) func (3) func
-----------------------------------------------------------
```

```
Modelo_72h_Normales      (100)  0.14 |   0.14     0.14
Modelo_72h_Normales_RASGO(100)  0.10 |   0.10     0.10
Modelo_72h_Normales_RASGO(100)  0.12 |   0.12     0.12
Modelo_72h_Especialista  (100)  0.14 |   0.13     0.13
Modelo_72h_Especialista_R(100)  0.13 |   0.13     0.12
Modelo_72h_Especialista_R(100)  0.14 |   0.14     0.13
Modelo_72h_Óptimos       (100)  0.17 |   0.19     0.19
Modelo_72h_Óptimos_RASGOS(100)  0.10 |   0.10     0.10
Modelo_72h_Óptimos_RASGOS(100)  0.17 |   0.16     0.17
----------------------------------------------------------
                           (v/ /*) |  (0/9/0)   (0/9/0)
```

J48

```
Llevar loamarillo para anexo
Analysing:  Num_false_positives

Dataset                   (1) trees.J48
---------------------------------------
Modelo_72h_Normales      (100)  0.49 |
Modelo_72h_Normales_RASGO(100)  0.48 |
Modelo_72h_Normales_RASGO(100)  0.29 |
Modelo_72h_Especialista  (100)  0.32 |
Modelo_72h_Especialista_R(100)  0.30 |
Modelo_72h_Especialista_R(100)  0.37 |
Modelo_72h_Óptimos       (100)  0.54 |
Modelo_72h_Óptimos_RASGOS(100)  0.38 |
Modelo_72h_Óptimos_RASGOS(100)  0.35 |
---------------------------------------
(v/ /*)                               |


Analysing:  Num_true_negatives

Dataset                   (1) trees.J48
---------------------------------------
Modelo_72h_Normales      (100)   1.31 |
Modelo_72h_Normales_RASGO(100)   1.32 |
Modelo_72h_Normales_RASGO(100)   1.51 |
Modelo_72h_Especialista  (100)   1.48 |
Modelo_72h_Especialista_R(100)   1.50 |
Modelo_72h_Especialista_R(100)   1.43 |
Modelo_72h_Óptimos       (100)   1.26 |
Modelo_72h_Óptimos_RASGOS(100)  1.42 |
Modelo_72h_Óptimos_RASGOS(100)  1.45 |
---------------------------------------
(v/ /*)                               |

Analysing:  Num_true_positives

Dataset                   (1) trees.J48
---------------------------------------
Modelo_72h_Normales      (100)  4.79 |
Modelo_72h_Normales_RASGO(100)  4.84 |
Modelo_72h_Normales_RASGO(100)  4.73 |
Modelo_72h_Especialista  (100)  4.77 |
Modelo_72h_Especialista_R(100)  4.78 |
Modelo_72h_Especialista_R(100)  4.70 |
Modelo_72h_Óptimos       (100)  4.85 |
Modelo_72h_Óptimos_RASGOS(100)  4.79 |
Modelo_72h_Óptimos_RASGOS(100)  4.78 |
---------------------------------------
```

```
(v/ /*)                                |

Analysing:  Num_false_negatives

Dataset                    (1) trees.J48
---------------------------------------
Modelo_72h_Normales      (100)  0.31 |
Modelo_72h_Normales_RASGO(100)  0.26 |
Modelo_72h_Normales_RASGO(100)  0.37 |
Modelo_72h_Especialista  (100)  0.33 |
Modelo_72h_Especialista_R(100)  0.32 |
Modelo_72h_Especialista_R(100)  0.40 |
Modelo_72h_Óptimos       (100)  0.25 |
Modelo_72h_Óptimos_RASGOS(100)  0.31 |
Modelo_72h_Óptimos_RASGOS(100)  0.32 |
---------------------------------------
(v/ /*)                                |
```

IBK

Llevar lo amarillo para anexos

```
Analysing:  Num_true_positives

Dataset                    (1) lazy.IBk
---------------------------------------
Modelo_72h_Normales      (100)  4.81 |
Modelo_72h_Normales_RASGO(100)  4.90 |
Modelo_72h_Normales_RASGO(100)  5.02 |
Modelo_72h_Especialista  (100)  4.80 |
Modelo_72h_Especialista_R(100)  5.00 |
Modelo_72h_Especialista_R(100)  5.10 |
Modelo_72h_Óptimos       (100)  4.99 |
Modelo_72h_Óptimos_RASGOS(100)  4.91 |
Modelo_72h_Óptimos_RASGOS(100)  5.08 |
---------------------------------------
(v/ /*)                                |

Analysing:  Num_false_positives

Dataset                    (1) lazy.IBk
---------------------------------------
Modelo_72h_Normales      (100)  0.57 |
Modelo_72h_Normales_RASGO(100)  0.20 |
Modelo_72h_Normales_RASGO(100)  0.92 |
Modelo_72h_Especialista  (100)  0.22 |
Modelo_72h_Especialista_R(100)  0.04 |
Modelo_72h_Especialista_R(100)  0.63 |
Modelo_72h_Óptimos       (100)  0.57 |
Modelo_72h_Óptimos_RASGOS(100)  0.47 |
Modelo_72h_Óptimos_RASGOS(100)  0.96 |
---------------------------------------
(v/ /*)                                |

Dataset                    (1) lazy.IBk
---------------------------------------
Modelo_72h_Normales      (100)  1.23 |
Modelo_72h_Normales_RASGO(100)  1.60 |
Modelo_72h_Normales_RASGO(100)  0.88 |
```

```
Modelo_72h_Especialista  (100)  1.58 |
Modelo_72h_Especialista_R(100)  1.76 |
Modelo_72h_Especialista_R(100)  1.17 |
Modelo_72h_Óptimos       (100)  1.23 |
Modelo_72h_Óptimos_RASGOS(100)  1.33 |
Modelo_72h_Óptimos_RASGOS(100)  0.84 |
---------------------------------------
(v/ /*)                               |

Analysing:  Num_false_negatives

Dataset                   (1) lazy.IBk
---------------------------------------
Modelo_72h_Normales      (100)  0.29 |
Modelo_72h_Normales_RASGO(100)  0.20 |
Modelo_72h_Normales_RASGO(100)  0.08 |
Modelo_72h_Especialista  (100)  0.30 |
Modelo_72h_Especialista_R(100)  0.10 |
Modelo_72h_Especialista_R(100)  0.00 |
Modelo_72h_Óptimos       (100)  0.11 |
Modelo_72h_Óptimos_RASGOS(100)  0.19 |
Modelo_72h_Óptimos_RASGOS(100)  0.02 |
---------------------------------------
(v/ /*)                               |
```