

Trabajo de Fin de Grado

Análisis comparativo de la Red Ontogénica SuperGNG en el diagnóstico temprano de la Enfermedad de Alzheimer

TITULACIÓN: Grado en Ingeniería Informática

AUTOR: Fernando Pellicer Sarmiento

TUTORIZADO POR:

Carmen Paz Suárez Araujo y Pablo Carmelo Fernández López

Fecha: Julio/2024

Agradecimientos

En primer lugar, quiero agradecer a mis tutores, Carmen Paz Suárez Araujo y Pablo Carmelo Fernández López, por guiarme y brindarme apoyo en la elaboración de este proyecto.

A mi familia y amigos, especialmente a mis padres y a mi hermano, quienes han hecho posible esta etapa y la han hecho más amena y agradable.

A Ylermi Cabrera León, por sus valiosos consejos y por su fundamental asistencia en el desarrollo del proyecto, además del trato cordial a lo largo de este periodo.

Muchas gracias a todos.

Resumen

El envejecimiento poblacional ha incrementado las enfermedades asociadas al envejecimiento, destacando las demencias, especialmente la enfermedad de Alzheimer (EA) como la más prevalente. Esto ha impulsado la investigación en nuevos métodos diagnósticos. Los métodos computacionales, basados en aprendizaje automático (AA), han proporcionado una importante aportación. Este trabajo se centra en los basados en computación neuronal, evaluando la red neuronal ontogénica SuperGNG para el diagnóstico temprano de la EA, mostrando resultados altamente prometedores, $AUC \geq 0.99$, en comparación con otros métodos de AA (MLP, SVM, etc.), para todos los problemas abordados. En los desarrollos realizados se han usado dos conjuntos de datos de ADNI, uno conteniendo criterios diagnósticos invasivos y otro no invasivos; evaluándose también la influencia del factor de riesgo edad en el diagnóstico.

Abstract

The aging population has led to an increase in age-related diseases, with dementias, particularly Alzheimer's (AD), being the most prevalent. This has driven research into new diagnostic methods. Computational methods based on machine learning (ML) have made significant contributions. This work focuses on those based on neural computing, evaluating the ontogenic neural network SuperRGNG for the early diagnosis of AD, showing highly promising results, $AUC \geq 0.99$, compared to other ML methods (MLP, SVM, etc.) for all addressed problems. The developments have used two ADNI datasets, one containing invasive diagnostic criteria and the other non-invasive; the influence of the risk factor age on diagnosis has also been evaluated.

Contenido

Capítulo 1: Introducción	1
1.1. Motivación.....	1
1.2. Estado actual y objetivos	2
1.3. Competencias específicas y aportaciones del trabajo	2
1.3.1. Competencias específicas.....	2
1.3.2. Aportación del trabajo en el entorno socioeconómico, técnico o científico	3
1.4. Conceptos básicos clínicos relacionados con el estudio.	4
1.4.1. Demencia.....	4
1.4.2. Enfermedad del Alzheimer	5
1.4.3. Deterioro Cognitivo Leve	6
1.5. Análisis de la base de datos ADNI.....	7
1.6. Métodos computacionales	8
1.6.1. Métodos No Neuronales	9
1.6.2. Métodos Neuronales	13
1.7. Coeficientes de calidad de clúster	18
1.7.1. Coeficiente de Silhouette	20
1.7.2. Coeficiente Davies-Bouldin.....	21
1.7.3. Coeficiente Calinski–Harabasz.....	22
Capítulo 2: Desarrollo Del Sistema Software.....	24
2.1. Metodología Software	24
2.2. Arquitectura del sistema software	25
2.2.1. Requisitos funcionales.....	25
2.2.2. Diseño y desarrollo del sistema software.....	26
2.2.3. Infraestructura software y hardware.....	29
2.2.4. Sistema Software	30
Capítulo 3: Análisis de la SuperGNG para la detección de la Enfermedad del Alzheimer.....	32
3.1. Estudio paramétrico de la SuperGNG usando conjuntos de datos sintéticos	32
3.1.1. Selección de los conjuntos de datos sintéticos	33
3.1.2. Implementación de los conjuntos de datos sintéticos	36
3.1.3. Resultados en conjuntos de datos sintéticos	37
3.2 Evaluación y Pre-procesado de los Conjuntos de Datos.....	39

3.2.1. Estudio y análisis de los conjuntos de datos	39
3.2.2. Estudio estadístico de los datos.	42
3.2.3. Pre-procesado de los datos según los coeficientes de clustering	44
3.3 Estudio comparativo de la red neuronal SuperGNG con otros métodos de aprendizaje automático	47
3.3.1. Elección de parámetros de los clasificadores.....	48
3.3.2. Comparación entre clasificadores frente a conjuntos de datos no invasivos	50
3.3.3. Comparación de clasificadores frente a conjunto de datos con pruebas invasivas	61
3.3.4. Comparación global de los clasificadores.....	73
3.4. Discusión.....	74
Capítulo 4: Conclusiones y trabajos futuros.....	77
4.1. Conclusiones.....	77
4.2. Trabajos Futuros	78
Capítulo 5: Bibliografía	79

Índice de figuras

Figura 1: Comparación entre un cerebro sano y un cerebro afectado por la EA. Obtenido de: [16]	6
Figura 2: Esquema Diagnóstico del DCL Obtenido de: [22]	7
Figura 3: Problema no Lineal Obtenido de [34]	10
Figura 4: Separación de clases con un hiperplano en 3 dimensiones, del problema no lineal de la Figura 3. Obtenido de [34]	11
Figura 5: Estructura básica de un árbol de decisión Obtenido de: [36]	11
Figura 6: División en una variable predictora continua X_i utilizando el punto de división c .	13
Figura 7: Arquitectura de una red neuronal Perceptrón Multicapa. Obtenido de: [45]	14
Figura 8: Comportamiento de una red de gas neuronal creciente, empezando por sus estados iniciales (fila superior) hasta sus estados finales (fila inferior). Obtenido de [50].	16
Figura 9: Estructura de MyGNG. Obtenido de: [53]	17
Figura 10: Estructura de SuperGNG, donde x_i es el componente i -th del vector de entrada; w_{li} el peso de la neurona l_i y u_l la neurona del módulo GNG. Fuente: [52].	18
Figura 11: Intra-Clúster e Inter-Clúster. Obtenido de: [54].	19
Figura 12: Cálculo del coeficiente de silhouette. Obtenido de [57]	20
Figura 13: Cálculo del coeficiente de DB. Obtenido de [57]	22
Figura 14: Cálculo del coeficiente de CH. Obtenido de [66].	23
Figura 15: Diagrama del diseño del sistema	27
Figura 16: Diagrama UML del sistema software	28
Figura 17: Flujos de trabajo Implementados en el Sistema Software Final	30
Figura 18: Código del Flujo de trabajo 1	31
Figura 19: Conjunto de datos blobs con diferente grado de solapamiento	33
Figura 20: Blobs desbalanceados	35
Figura 21: Representación del conjunto de datos Two Concentric Squares	36
Figura 22: Implementación en el código de los conjuntos de datos sintéticos	37
Figura 23: A la izquierda el conjunto de datos blobs 62% OVL y a la derecha los resultados de clasificación de la red SuperGNG.	39
Figura 24: Gráfico de los p-valores de las características conjunto de datos sin pruebas invasivas	43
Figura 25: Gráfico de los p-valores del conjunto de datos con pruebas invasivas	44
Figura 26: Cálculo de métricas de rendimiento de los clasificadores	48
Figura 27: Relación entre componentes y las características del conjunto de datos usadas Power Transformer Yeo-Johnson - Neighborhood CA	50
Figura 28: Representación de los datos mediante el preprocesado Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)	51
Figura 29: Curvas ROC de los clasificadores - Power Transformer Yeo-Johnson - Neighborhood CA	51
Figura 30: Relación entre componentes y las características del conjunto de datos usadas MaxAbsScaler - Neighborhood CA	52

Figura 31: : Representación de los datos mediante el preprocesado MaxAbsScaler - Neighborhood CA (PC1, PC2).....	53
Figura 32: Curvas ROC de los clasificadores - MaxAbsScaler - Neighborhood CA (PC1, PC2).....	53
Figura 33: Relación entre componentes y las características del conjunto de datos usadas Robust Scaler - Neighborhood CA	54
Figura 34: : Representación de los datos mediante el preprocesado Robust Scaler - Neighborhood CA (PC1, PC2).....	55
Figura 35: Curvas ROC de los clasificadores - Robust Scaler - Neighborhood CA (PC1, PC2).....	55
Figura 36: Relación entre componentes y las características del conjunto de datos usadas Power Transformer Yeo-Johnson - Neighborhood CA	56
Figura 37: : Representación de los datos mediante el preprocesado Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2).....	57
Figura 38: Curvas ROC de los clasificadores - Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)	57
Figura 39: Relación entre componentes y las características del conjunto de datos usadas MaxAbsScaler - Neighborhood CA	58
Figura 40: : Representación de los datos mediante el preprocesado MaxAbsScaler - Neighborhood CA (PC1, PC2).....	59
Figura 41: Curvas ROC de los clasificadores - MaxAbsScaler - Neighborhood CA (PC1, PC2).....	59
Figura 42: Relación entre componentes y las características del conjunto de datos usadas Robust Scaler - Neighborhood CA	60
Figura 43: : Representación de los datos mediante el preprocesado Robust Scaler - Neighborhood CA (PC1, PC2).....	60
Figura 44: Curvas ROC de los clasificadores - Robust Scaler - Neighborhood CA (PC1, PC2).....	61
Figura 45: Relación entre componentes y las características del conjunto de datos usadas Power Transformer Yeo-Johnson - Neighborhood CA	62
Figura 46: : Representación de los datos mediante el preprocesado Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2).....	62
Figura 47: Curvas ROC de los clasificadores - Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)	63
Figura 48: Relación entre componentes y las características del conjunto de datos usadas MaxAbsScaler - Neighborhood CA	64
Figura 49: : Representación de los datos mediante el preprocesado MaxAbsScaler - Neighborhood CA (PC1, PC2).....	64
Figura 50: Curvas ROC de los clasificadores - MaxAbsScaler - Neighborhood CA (PC1, PC2).....	65
Figura 51: Relación entre componentes y las características del conjunto de datos usadas Robust Scaler - Neighborhood CA	66
Figura 52: : Representación de los datos mediante el preprocesado Robust Scaler - Neighborhood CA (PC1, PC2).....	66
Figura 53: Curvas ROC de los clasificadores - Robust Scaler - Neighborhood CA (PC1, PC2).....	67
Figura 54: Relación entre componentes y las características del conjunto de datos usadas Power Transformer Yeo-Johnson - Neighborhood CA	68

Figura 55: : Representación de los datos mediante el preprocesado Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)	68
Figura 56: Curvas ROC de los clasificadores - Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)	69
Figura 57: : Relación entre componentes y las características del conjunto de datos usadas MaxAbsScaler - Neighborhood CA	70
Figura 58: Representación de los datos mediante el preprocesado MaxAbsScaler - Neighborhood CA (PC1, PC2).....	70
Figura 59: Curvas ROC de los clasificadores - MaxAbsScaler - Neighborhood CA (PC1, PC2).....	70
Figura 60: Relación entre componentes y las características del conjunto de datos usadas Robust Scaler – Neighborhood CA.....	71
Figura 61: : Representación de los datos mediante el preprocesado Robust Scaler - Neighborhood CA (PC1, PC2).....	72
Figura 62: Curvas ROC de los clasificadores - Robust Scaler - Neighborhood CA (PC1, PC2).....	72

Índice de tablas

Tabla 1: Competencias específicas y motivo de cumplimiento.....	3
Tabla 2: Características de cada fase ADNI. Datos obtenidos de: [23]	8
Tabla 3: Características de los equipos utilizados	29
Tabla 4: Flujos de trabajo en el sistema final.....	31
Tabla 5: Valores de clustering en blobs según el solapamiento existente	34
Tabla 6: Valores de clustering en blobs desbalanceados	35
Tabla 7: Valores de clustering en Two Concentric Squares	36
Tabla 8: Resultado del estudio paramétrico de la SuperGNG con datos sintéticos.....	38
Tabla 9: Características de cada fase después de preprocesar los datos	40
Tabla 10: Definiciones y códigos de los criterios diagnósticos que conforman del conjunto de datos sin pruebas invasivas [23].....	41
Tabla 11: Definiciones y códigos de los criterios diagnósticos que conforman el conjunto de datos con pruebas invasivas [23].	41
Tabla 12: Estadísticos del conjunto de datos sin pruebas invasivas	42
Tabla 13: Estadísticos del conjunto de datos con pruebas invasivas.....	43
Tabla 14: TOP 4 mejores combinaciones de escalado, proyectado y componentes de cada métrica en los conjuntos de datos.....	46
Tabla 15: Combinación de escalado y proyectado más comunes en los conjuntos de datos, según los resultados de las coeficientes	47
Tabla 16: Parámetros de las 5 mejores configuraciones de SuperGNG comparadas con la configuración semilla de "blobs"	49
Tabla 17: Parámetros de configuración de los clasificadores	49
Tabla 18: Métricas de Rendimiento Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)	52
Tabla 19: Métricas de rendimiento ROC MaxAbsScaler - Neighborhood CA (PC1, PC2).....	54
Tabla 20: Métricas de rendimiento Robust Scaler - Neighborhood CA (PC1, PC2)	55
Tabla 21: Métricas de Rendimiento Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)	58
Tabla 22: Métricas de Rendimiento MaxAbsScaler - Neighborhood CA (PC1, PC2).....	60
Tabla 23: Métricas de Rendimiento Robust Scaler - Neighborhood CA (PC1, PC2)	61
Tabla 24: Métricas de Rendimiento Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)	63
Tabla 25: Métricas de Rendimiento MaxAbsScaler - Neighborhood CA (PC1, PC2).....	65
Tabla 26: Métricas de Rendimiento Robust Scaler - Neighborhood CA (PC1, PC2)	67
Tabla 27: Métricas de rendimiento Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)	69
Tabla 28: Métricas de Rendimiento MaxAbsScaler - Neighborhood CA (PC1, PC2).....	71
Tabla 29: Métricas de Rendimiento Robust Scaler - Neighborhood CA (PC1, PC2)	73
Tabla 30: Valores de AUC obtenidos por los clasificadores en los conjuntos de datos Con edad..	73
Tabla 31: Valores de AUC obtenidos por los clasificadores en los conjunto de datos Sin edad.....	74

Índice de Algoritmos

Algoritmo 1: Algoritmo GNG	15
Algoritmo 2: Algoritmo Red Ontogénica SuperGNG.....	18

Capítulo 1: Introducción

Se va a analizar una arquitectura neuronal para el diagnóstico temprano de la Enfermedad del Alzheimer (EA) utilizando una base de datos de sujetos con enfermedades neurodegenerativas. Esta base de datos contiene una amplia gama de datos clínicos asociados a sujetos en diferentes etapas de la enfermedad.

Para el análisis, disponemos de cuatro conjuntos de datos. Los primeros dos conjuntos incluyen criterios clínicos de Alzheimer, así como el factor de riesgo asociado a la edad. Los otros dos conjuntos son versiones modificadas de los primeros, en los que se ha eliminado el factor de riesgo de la edad, manteniendo los otros criterios diagnósticos.

Para preprocesar los datos, se ha propuesto un método para la elección de los procesos de proyección, escalado y componentes utilizando coeficientes de clustering. Esta propuesta garantiza la imparcialidad entre los diferentes conjuntos de datos a la hora de ser clasificados.

Para optimizar la arquitectura, primero se realizan pruebas utilizando conjuntos de datos sintéticos. Estos conjuntos de datos permiten ajustar los parámetros y evaluar el rendimiento del modelo en un entorno controlado. Esto permitirá determinar los mejores valores de los parámetros que conforman la arquitectura neuronal en estudio, Supervised Reconfigurable Growing Neural Gas (SuperGNG), para optimizar el rendimiento y desempeño de la misma tanto en entornos sintéticos como en ámbitos clínicos reales.

En este primer capítulo se presentan los conceptos clínicos necesarios para entender el estudio, una descripción detallada de la base de datos “Alzheimer's Disease Neuroimaging Initiative” (ADNI), y una explicación de los clasificadores y coeficientes de calidad de clustering utilizados en el estudio.

El resto del documento se compone de cuatro capítulos más. El segundo aborda el desarrollo del sistema software. El tercer capítulo presenta el estudio comparativo. El cuarto está compuesto de las conclusiones derivadas del estudio y propuestas de trabajos futuros. El último capítulo contiene la bibliografía utilizada en el trabajo.

1.1. Motivación

La motivación del trabajo viene dada por la propuesta de TFT de la Dra. Carmen Paz Suárez Araujo y el Dr. Pablo Carmelo Fernández López, junto a mi interés de colaborar en proyectos que aporten valor a estudios relacionados con el diagnóstico temprano de la EA.

La red ontogénica SuperGNG es una implementación del estudiante de doctorado Ylmeri León Cabrera, que ha realizado una nueva arquitectura neuronal y surgió la necesidad de analizar su desempeño en la clasificación de sujetos con Alzheimer. Para ello se realizan primero estudios con datos sintéticos que han permitido estudiar los nuevos parámetros de la red y su comportamiento. Posteriormente se realizó un estudio comparativo, del desempeño de esta, utilizando datos reales de sujetos con la EA con otros clasificadores ampliamente utilizados.

1.2. Estado actual y objetivos

Actualmente, en el ámbito del diagnóstico temprano de la EA, se están desarrollando biomarcadores específicos en sangre y líquido cefalorraquídeo, así como técnicas avanzadas de neuroimagen como tomografías por emisión de positrones (PET) y resonancias magnéticas (RM) para detectar la enfermedad en sus etapas iniciales [1]. Además de estas innovaciones, se utilizan otros criterios diagnósticos como los tests neuropsicológicos, que evalúan funciones cognitivas específicas y pueden revelar cambios en etapas iniciales de la enfermedad [2].

El uso de algoritmos de inteligencia artificial y el amplio despliegue de herramientas computacionales están facilitando el procesamiento y análisis de grandes volúmenes de datos clínicos y de imagen, proporcionando un soporte integral para el diagnóstico temprano y mejorando la precisión en la predicción del riesgo y la identificación de la EA [3].

Los objetivos de este trabajo son los siguientes:

- Desarrollo de un Sistema Software:
 - Diseñar y desarrollar el sistema software para realizar el estudio comparativo.
 - Desarrollo de componentes software que sean capaces de manejar el conjunto de datos y preprocesarlos.
- Estudio y Optimización de SuperGNG para ayuda en la Detección de la EA:
 - Evaluar el rendimiento de la arquitectura objetivo en el diagnóstico temprano del Alzheimer.
 - Estudio Comparativo de la SuperGNG y otros Métodos Paradigmáticos de Aprendizaje Automático (AA) para la Detección de la EA.

1.3. Competencias específicas y aportaciones del trabajo

1.3.1. Competencias específicas

En este trabajo, a medida que se ha desarrollado se han utilizado distintas herramientas y conocimientos pertenecientes a las ramas de la Computación Inteligente, Aprendizaje Automático e Ingeniería del Software. Dentro de las competencias específicas del Grado

de Ingeniería Informática, las más ligadas a este estudio son las que encontramos en la Tabla 1, donde se incluyen sus respectivas justificaciones.

Competencia específica	Descripción	Motivo de cumplimiento
CI1	Capacidad para diseñar, desarrollar, seleccionar y evaluar aplicaciones y sistemas informáticos, asegurando su fiabilidad, seguridad y calidad, conforme a principios éticos y a la legislación y normativa vigente.	En este trabajo se ha diseñado y desarrollado un sistema software para realizar las pruebas de cara a la comparación.
CI6	Conocimiento y aplicación de los procedimientos algorítmicos básicos de las tecnologías informáticas para diseñar soluciones a problemas, analizando la idoneidad y complejidad de los algoritmos propuestos.	La integración de los módulos utilizados en el sistema software ha llevado al uso de algoritmos y soluciones a los problemas a medidas que se presentaban y eran necesarios.
CI15	Conocimiento y aplicación de los principios fundamentales y técnicas básicas de los sistemas inteligentes y su aplicación práctica.	Esta competencia se cumple ya que ha sido necesario comprender y manejar, distintos métodos y técnicas de AA, así como modelos de computación neuronal hasta implementar arquitecturas neuronales determinadas.

Tabla 1: Competencias específicas y motivo de cumplimiento

1.3.2. Aportación del trabajo en el entorno socioeconómico, técnico o científico

En primer lugar, el sistema software desarrollado podrá ser utilizado en futuros estudios, lo que facilitará su realización. Desde el punto de vista científico-técnico, esta contribución se destaca por su capacidad para mejorar la eficiencia en la realización de estudios futuros.

Desde el punto de vista científico-técnico, se ha hecho una aportación de conjuntos de datos sintéticos para el análisis de clasificadores, los cuales podrán ser utilizados para evaluar el rendimiento de redes en investigaciones posteriores.

Una de las contribuciones más destacadas de este trabajo es su aportación al diagnóstico temprano del Alzheimer, comparando conjuntos de datos con pruebas invasivas y no invasivas. Desde el punto de vista socioeconómico, este estudio abre nuevas posibilidades al demostrar que es posible clasificar a los sujetos utilizando métodos no dañinos, incómodos o que requieran mucho tiempo. Además, se ha corroborado que la edad puede facilitar significativamente la clasificación de sujetos, siendo un factor de riesgo con un impacto positivo en los resultados.

Por otra parte, se ha propuesto un método para estudiar cuál combinación de escalado y proyección es la más adecuada para un conjunto de datos, utilizando coeficientes de clustering. Esta metodología promueve la equidad al no favorecer un conjunto de datos sobre otros, fortaleciendo así el rigor científico en la selección de técnicas de análisis.

Finalmente, desde el punto de vista socioeconómico, se valida un sistema inteligente de ayuda al diagnóstico temprano de la EA de uso en cualquier ámbito clínico, pero muy especialmente en atención primaria, que es donde primero reporta el sujeto síntomas de esta posible patología no comunicable asociada al envejecimiento.

1.4. Conceptos básicos clínicos relacionados con el estudio.

En este apartado se aborda el objetivo de adquirir conocimientos acerca de los conceptos de demencia, la EA y el deterioro cognitivo leve (DCL).

Por otro lado, la base de datos ADNI, debe ser entendida y estudiada (nomenclatura, estructura, formatos...) de tal manera que pueda utilizarse en futuras actividades.

Llegados a este punto, debemos comprender cual es el interés del estudio de la demencia. Para ello, se presenta una citación realizada por el Ministerio de Sanidad Consumo y Bienestar Social, que muestra la importancia que tiene la demencia en la sociedad actual:

“El abordaje de las demencias es una prioridad de salud pública y un problema social y sanitario de primer orden. Se ha convertido en la gran epidemia silenciosa del siglo XXI y en el gran reto para la sostenibilidad de cualquier sistema social y sanitario” [4].

1.4.1. Demencia

Demencia, (*del inglés “Dementia”*) se define como un síndrome donde existe un deterioro cognitivo persistente, que impide al individuo que lo padece realizar sus actividades sociales o laborales de forma habitual [5]. La demencia puede ser causada por un gran número de enfermedades, y generalmente se caracteriza por una evolución lenta, progresiva e irreversible. “Un estimado de 47.5 millones de personas viven actualmente con demencia. Alrededor del 60% de esta enfermedad recae en países de ingresos bajos y medios, que son los que tienen menor capacidad para hacer frente a la situación” [6].

Las enfermedades degenerativas son de difícil diagnóstico, aunque un estudio del perfil neuropsicopatológico predominante y los exámenes de neuroimagen entre otros biomarcadores permiten tener una seguridad razonable sobre el diagnóstico [7].

En cuanto a los criterios de diagnóstico de demencia más utilizados, cabe destacar los del Manual Diagnóstico y Estadístico de los Trastornos Mentales de la Asociación Americana de Psiquiatría, entre los que se encuentran deterioro de la memoria, la presencia de déficits cognitivos (afasia, apraxia, agnosia o deterioro de la función ejecutiva), entre otros criterios [8] [7].

Dentro de las distintas enfermedades que engloba la demencia, la EA es la demencia más prevalente, seguido de la demencia Vascular, las degeneraciones frontotemporales y la demencia con cuerpos de Lewy [7] [9].

La Organización Mundial de la Salud (OMS) realizó una conferencia en 2015 donde su directora en aquel momento (Margaret Chan) destacó un dato preocupante acerca de la demencia: “A medida que el envejecimiento de la población continúa acelerándose, se espera que el número de casos de demencia se duplique cada 20 años” [6].

1.4.2. Enfermedad del Alzheimer

Es un trastorno neurológico de causa incierta que afecta principalmente a individuos mayores de 65 años ocasionado generalmente por una pérdida progresiva de memoria. Se caracteriza por un inicio insidioso y una evolución lenta y progresiva [10]. En la actualidad no existe cura para la EA, por ello se ha incrementado en gran medida la búsqueda de métodos para la predicción temprana y la búsqueda de tratamientos efectivos [11].

Existen varias hipótesis sobre las causas de la EA, las más destacadas son:

- Hipótesis genéticas
 - Hipótesis de las proteínas β -amiloide: esta hipótesis afirma que la EA está relacionada con la acumulación de las proteínas β -amiloide [12]. Esta hipótesis es el modelo dominante de patogénesis de la EA y es la que guía el desarrollo de posibles tratamientos [12].
 - Hipótesis de las proteínas tau: se basa en que la acumulación de las proteínas tau se encuentra comúnmente en la EA, que se agrega dentro de los cuerpos de las células nerviosas. [12].
 - Hipótesis del sueño: Se basa en la hipótesis de las proteínas β -amiloide. Según esta hipótesis, la evidencia acumulada muestra que la alteración del sueño (común en adultos mayores de 60 años) contribuye al deterioro cognitivo y es el causante del aumento de la proteína β -amiloide [13].
- Hipótesis ambientales
 - Hipótesis de Toxicidad por metales: Afirma que la EA viene dada por factores ambientales producido por metales. Principalmente se hace referencia al aluminio, que es un metal que está muy integrado en nuestra vida cotidiana (alimentación, cosmética, medicina, etc.) [14]. Afirma que el aluminio está unido a la transferrina plasmática y a las moléculas de citrato que pueden mediar la transferencia de aluminio al cerebro [14]. Hay estudios que demuestran que se acumula en las áreas de la corteza, hipocampo y cerebelo [14].
 - Hipótesis de factores médicos: Estas hipótesis afirman que el causante de la EA son enfermedades como la obesidad, la diabetes y enfermedades cardiovasculares [14].

En cuanto al diagnóstico de la EA, este puede ser realizado a partir de neuroimágenes, debido a que un cerebro de una persona cognitivamente normal (CN) es diferenciable del de un paciente que padece EA [15], podemos ver esta diferencia en la Figura 1. Este

tipo de diagnóstico permite estudiar al individuo de manera poco invasiva comparado con otras metodologías como las biopsias cerebrales [15].

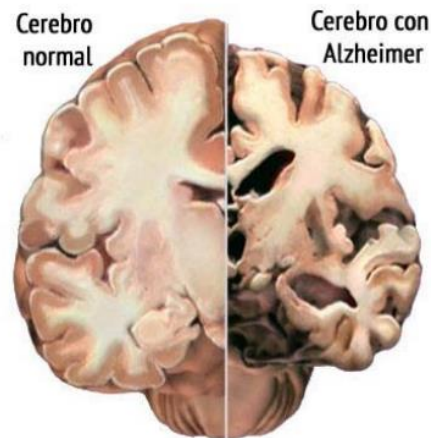


Figura 1: Comparación entre un cerebro sano y un cerebro afectado por la EA. Obtenido de: [16]

Para el diagnóstico de la EA, las medidas neuropsicológicas (p. ej., pruebas cognitivas), neuroimagen y biomarcadores se combinan con frecuencia para detectar la EA en etapa temprana y predecir la conversión del DCL a la EA [17]. En la práctica clínica, las pruebas neuropsicológicas son los métodos predominantes utilizados para identificar la EA, debido a que estos son menos invasivos y tienen un mejor coste [17].

Una prueba comúnmente utilizada es la denominada ADAS (Alzheimer's Disease Assessment Scale), que consiste en una evaluación psicológica para detectar la gravedad de la EA en los sujetos [15]. Otras pruebas también utilizadas son MMSE (Mini-Mental State Examination), MoCA (Montreal Cognitive Assessment) y FAQ (Functional Activities Questionnaire) [15].

Cuando la enfermedad se encuentra en etapas avanzadas, suele generar en el individuo demencia y posteriormente la muerte [18]. El tiempo medio de supervivencia una vez diagnosticada la EA se estima que oscila los 10 años, aunque puede ser menor en caso de personas de edad avanzada [9].

Entre los factores de riesgo de la EA se encuentran en primer lugar la edad (envejecimiento normal de las personas) y, en segundo lugar, la posible influencia de la genética familiar, dado que se ha observado que aproximadamente el 40% de las personas que la padecen tienen antecedentes familiares de la enfermedad [19].

1.4.3. Deterioro Cognitivo Leve

Se define como un síndrome que manifiesta un déficit cognitivo de un grado mayor al esperado para la edad y el nivel cultural del individuo. Aunque no interfiere de manera considerable en las actividades diarias del individuo [20].

Algunos de los criterios diagnósticos del DCL son: Que el paciente presente deterioro cognitivo evidenciable, se descarta la presencia de demencia, las alteraciones son

informadas por paciente o por un familiar y se puede diagnosticar mediante pruebas neuropsicológicas [21].

Podemos diferenciar varios subtipos de deterioro cognitivo dependiendo del dominio cognitivo afectado: amnésico (exclusivamente a la memoria), amnésico multidominio (a la memoria y otros dominios), no amnésico (se altera la función ejecutiva o el lenguaje) y no amnésico multidominio (se alteran 2 o más dominios diferentes a la memoria) [21].

A continuación, se muestra un esquema de diagnóstico del DCL en la Figura 2, que muestra lo mencionado en los párrafos anteriores.

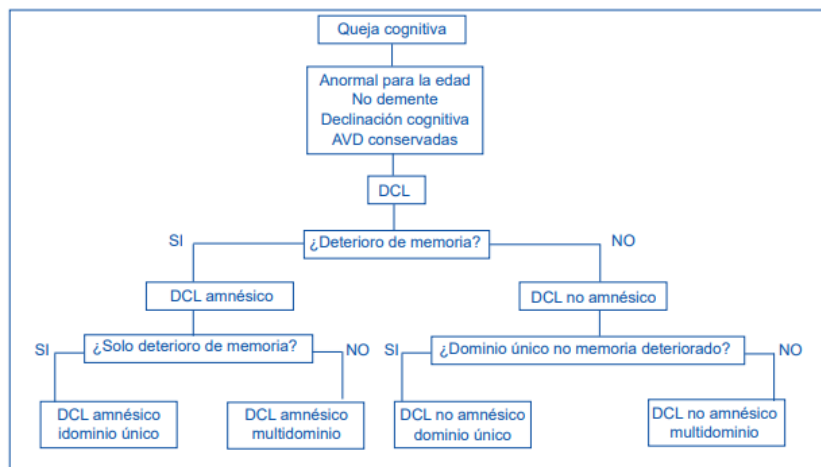


Figura 2: Esquema Diagnóstico del DCL Obtenido de: [22]

1.5. Análisis de la base de datos ADNI

Es un estudio de investigación global enfocado a la investigación y desarrollo de tratamientos que retrasen o detengan la progresión de la EA [23]. Los investigadores participantes rastrean y estudian el estado y la progresión de la EA en el cerebro humano a partir de marcadores biológicos, clínicos y genéticos de los individuos que forman esta base de datos. Están diagnosticados en distintos estados de deterioro cognitivo como: “Mild Cognitive Impairment” (MCI), “Early Mild Cognitive Impairment” (EMCI), “Late Mild Cognitive Impairment” (LMCI), “Alzheimer’s Disease” (AD), así como sujetos de control, “Cognitively Normal” (CN), [23].

Los objetivos más destacados del proyecto ADNI son: Desarrollar métodos para predecir la EA, la búsqueda de tratamientos o medidas de prevención y la construcción de la base de datos sobre la enfermedad [24]. ADNI se divide en 4 fases, ADNI-1, ADNI-GO, ADNI-2, ADNI-3, creadas en el orden de mención.

Las muestras utilizadas para las distintas fases se organizan en la Tabla 2 indicando el número de muestras y tipos de sujetos que participaron, además de características como los objetivos, financiación y fechas de cada una. Es importante señalar que los mejores criterios médicos perduran a lo largo de las distintas fases del estudio ADNI, manteniéndose consistentes con el tiempo gracias a su aportación en el diagnóstico.

Características del Estudio	ADNI-1	ADNI-GO	ADNI-2	ADNI-3
Objetivo Principal	Desarrollar biomarcadores como medidas de resultado para ensayos clínicos	Examinar biomarcadores en etapas más tempranas de la enfermedad	Desarrollar biomarcadores como predictores del deterioro cognitivo y como medidas de resultado	Estudiar el uso de PET con tau y técnicas de imágenes funcionales en ensayos clínicos
Financiación	40 millones federales (NIA), \$27 millones industria y fundaciones	\$24 millones de fondos de la Ley de Recuperación Estadounidense	\$40 millones federales (NIA), \$27 millones industrias y fundaciones	\$40 millones federales (NIA), hasta \$20 millones industrias y fundaciones
Duración/fecha de inicio	5 años/octubre 2004	2 años/septiembre 2009	5 años/septiembre 2011	5 años/septiembre 2016
Cohorte	200 personas de edad avanzada 400 MCI 200 AD	Existentes ADNI-1 + 200 EMCI	Existentes ADNI-1 y ADNI-GO + 150 personas de edad avanzada 100 EMCI 150 LMCI 150 AD	Existentes en ADNI-1, ADNI-GO, ADNI-2 + 133 personas de edad avanzada 151 MCI 87 AD

Tabla 2: Características de cada fase ADNI. Datos obtenidos de: [23]

Hay que comentar la dificultad que existe en la obtención de datos, debido a que los sujetos se someten a una serie de pruebas alargadas en el tiempo, y muchos de ellos no presentan (por distintos motivos) algunas de las pruebas, sobre todo las finales [25]. Además, los sujetos que se someten a las pruebas tienen que cumplir unas características muy concretas para poder ser seleccionado como muestra válida [25].

1.6. Métodos computacionales

Siendo el "Método Computacional" una de las principales metodologías seguidas en el desarrollo de este trabajo, estando focalizado el mismo en el desarrollo y análisis de soluciones computacionales para la ayuda al diagnóstico temprano de la EA, es de obligado cumplimiento el dedicarles un espacio a los métodos computacionales utilizados en nuestro estudio.

Nos dedicaremos a métodos de AA, centrándonos con más detenimiento en las redes neuronales artificiales superficiales, que es el enfoque computacional base de la solución computacional inteligente propuesta y analizada en este TFG.

Empezaremos con la definición de aprendizaje: "El aprendizaje se refiere, (...), a un amplio espectro de situaciones en las cuales el aprendiz incrementa su conocimiento o sus habilidades para cumplir una tarea" [26].

A lo largo de los últimos años, se ha incorporado en nuestra sociedad el concepto de AA. Los métodos computacionales supervisados y no supervisados pertenecen a este tipo de aprendizaje, y tienen numerosas aplicaciones en distintos campos como la obtención de patrones o diagnóstico médico [27].

Por otra parte, la inteligencia artificial (IA) era definida por John McCarthy como: “La ciencia y la ingeniería para crear máquinas inteligentes, especialmente programas informáticos inteligentes. Está relacionada con la tarea similar de utilizar ordenadores para comprender la inteligencia humana, pero la IA no tiene por qué limitarse a métodos que sean biológicamente observables” [28].

En el campo de la inteligencia artificial, una de las ramas más destacadas son las Redes Neuronales Artificiales (RNAs), que se definen como aquellos sistemas de procesamiento de información paralelo y distribuido que contienen elementos procesadores de información, que definen el comportamiento del conjunto del sistema software a partir de sus interacciones locales y de reglas de aprendizaje [29].

Además, en el AA nos encontramos con una taxonomía cuya base, podríamos decir, es la inspiración biológica o no, de la misma. Así tenemos los denominados métodos neuronales y los métodos no neuronales. Tratan de emular el comportamiento del cerebro humano, que se caracteriza por el aprendizaje conocimientos a partir de conjuntos de datos. Mientras que los métodos no neuronales están más dentro de la IA clásica o simbólica [29].

1.6.1. Métodos No Neuronales

Clasificador Naïve Bayes

Es un clasificador que utiliza como base el teorema de Naïve Bayes (NB). Este clasificador simplifica el aprendizaje al asumir que las características son independientes dada la clase. Aunque la suposición de independencia que asumen estos clasificadores generalmente es pobre, en la práctica este tipo de clasificadores son capaces de competir con clasificadores más sofisticados [30].

Como se indicaba anteriormente, los clasificadores bayesianos utilizan como base el teorema de bayes, que establece lo siguiente: “Sea $\{A_1, A_2, \dots, A_i, \dots, A_n\}$ un conjunto de sucesos mutuamente excluyentes y cuya unión es el total o sea 1, y tales que la probabilidad de cada uno de ellos es distinta de cero. Sea B un suceso cualquiera del que se conocen las probabilidades condicionales $P(B|A_i)$. Entonces la probabilidad $P(A_i|B)$ viene dada por la expresión:” [31]. Ecuación 1.

$$P(A_i|B) = \frac{P(B|A_i) P(A_i)}{P(B)} = \frac{P(B|A_i) P(A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)}$$

Ecuación 1

Dónde:

$P(A_i)$: Probabilidades a priori.

$P(B|A_i)$: Probabilidad de B en la hipótesis de A_i .

$P(A_i|B)$: Probabilidades a posteriori.

Este método destaca por consumir menos recursos computacionales comparado con algoritmos parecidos (Sequential Minimum Optimization, Chunking...), además de que tiene mayor facilidad a la hora de ser programado y no requiere cálculos complejos para resolver problema de programación cuadrática [32].

Máquinas de Vectores Soporte

Las máquinas de vectores de soporte (“Support vector machines” en inglés) (SVM) se definen como técnicas de clasificación de 2 clases, basadas en la transformación del espacio de datos de entrada en otro con una dimensión mayor, para posteriormente resolver el problema con el uso de un hiperplano [33].

Para entender la esencia de SVM debemos comprender 4 conceptos:

- Hiperplano de separación: Dado un conjunto de datos, diremos que los datos son linealmente separables si se puede encontrar un hiperplano a partir de una función lineal que sea capaz de posicionarse de manera que, si tenemos dos clases, una de ellas se encuentre en una parte del hiperplano y por consiguiente la otra clase en el otro lado [34].
- Hiperplano con margen duro: Si los datos de separación son perfectamente separables con una función lineal estamos hablando de margen duro [34].
- Hiperplano con margen blando: Normalmente, a la hora de separar datos existen casos en los que se la función lineal es incapaz de separar perfectamente las clases (y por lo tanto existen errores). Este tipo de hiperplanos que tienen estos errores se llaman hiperplanos con margen blando [34].
- Función Kernel: Cuando los datos no pueden ser separados por funciones lineales surgen las funciones kernel. Estas funciones proyectan los datos de entrada en una dimensión superior donde si puedan ser separados [34]. Podemos ver un ejemplo donde la necesidad de la función kernel es clave para encontrar un hiperplano que separa los datos de ambas clases en la Figura 3 y Figura 4.

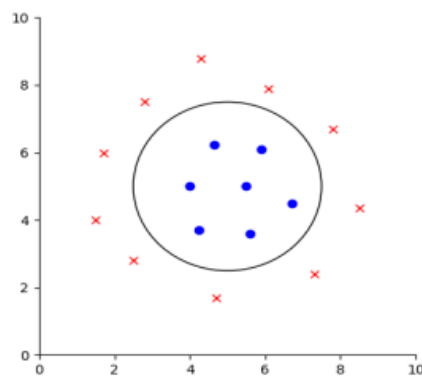


Figura 3: Problema no Lineal Obtenido de [34]

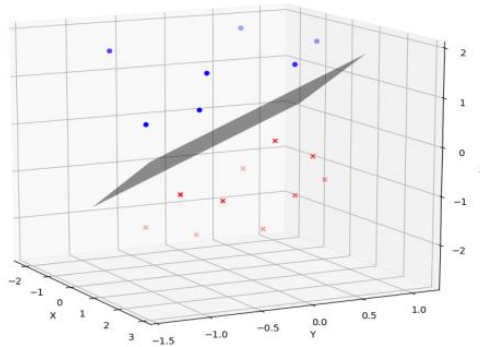


Figura 4: Separación de clases con un hiperplano en 3 dimensiones, del problema no lineal de la Figura 3. Obtenido de [34]

Las máquinas de vectores de soporte empezaron con una limitación significativa: no todos los problemas se pueden resolver de forma lineal, lo cual limitaba enormemente este tipo de algoritmos en las aplicaciones prácticas. Este problema fue solucionado con los métodos kernel, los cuales permiten, como se comentó anteriormente, clasificar datos con estructuras no lineales [35].

Árbol de decisión

Un árbol es una estructura discreta formada por una colección de nodos y aristas que se distribuyen en una estructura jerárquica. Cuando hablamos de un árbol de decisión (del inglés “Decision Tree”) (DT), nos referimos a un modelo que tiene como objetivo el aprendizaje inductivo mediante observaciones y construcciones lógicas. Generalmente hablamos de árboles binarios, aunque no existe limitación respecto a esta característica [36].

Cada uno de los nodos del árbol representa una decisión, menos los nodos hoja que son resultados del árbol [36]. Podemos observar la estructura de un árbol de decisión en la Figura 5.

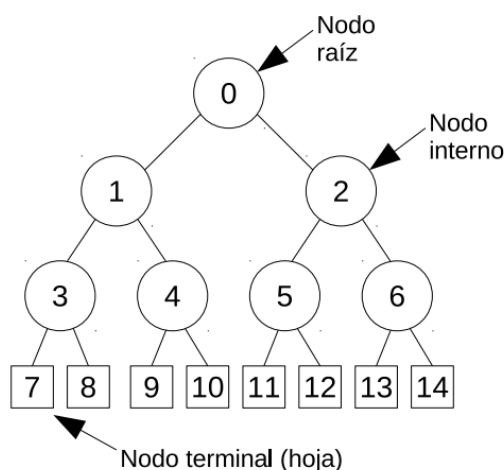


Figura 5: Estructura básica de un árbol de decisión Obtenido de: [36]

Como norma general, a la hora de seleccionar el mejor atributo de un nodo (que incluye la homogeneidad o pureza de los conjuntos), se utilizan dos métodos, la ganancia de información y la impureza de Gini [37].

En la primera metodología de ganancia de información (Ecuación 3), se calcula la probabilidad cuál de las opciones que tenemos es más acertada desde el punto de vista probabilístico [38].

$$G(S, A) = H(S) - H(S, A)$$

Ecuación 2

Donde $H(S)$ es la entropía de S y $H(S, A)$ es el valor esperado de la entropía después de que S se particiona teniendo en cuenta el atributo A [38].

Para comprender la Ecuación 2, es fundamental definir el concepto de entropía. La entropía es una medida de la incertidumbre o impureza de los valores en un sistema, y sus valores típicamente oscilan entre 0 y 1. La fórmula que define la entropía viene dada por la Ecuación 3 [38].

$$H(S) = -Pp \log_2(Pp) - Pn \log_2(Pn)$$

Ecuación 3

Donde Pp y Pn son las probabilidades de acertar y fallar respectivamente.

La impureza de Gini, se refiere a una medida utilizada en clasificación que evalúa qué tan frecuentemente se clasifica erróneamente una elección al hacerla aleatoriamente. Esta medida se calcula mediante la Ecuación 4 [39].

$$Gini(t) = 1 - \sum_{c \in C} p_{t,c}^2$$

Ecuación 4

Donde C es el conjunto de clases y $p_{t,c}$ es la probabilidad de ocurrencia de la clase c en el nodo t [39].

DT se define por su fácil implementación y su flexibilidad. Aunque consta de inconvenientes como son el costo de entrenamiento y que es propenso al sobreajuste cuando se vuelve complejo [40].

Bosque Aleatorio

Del inglés "Random Forest" (RF), consiste en una combinación de árboles predictivos, donde cada árbol tiene como entrada un vector con valores aleatorios de la muestra y comparten la misma distribución que los demás árboles que constituyen el bosque. Este algoritmo es un refinamiento adicional de los distintos modelos de árbol de clasificación y regresión [37].

Los resultados que se obtienen de los distintos árboles del bosque son combinados de manera que quede una única decisión como resultado. El método comúnmente utilizado es el voto de la mayoría o promedio de los resultados obtenidos [36].

Los árboles que se utilizan en el algoritmo se basan en árboles de partición recursivos binarios (generalmente). Todos ellos dividen el espacio predictor utilizando particiones binarias en variables individuales. El nodo raíz comprende todo el espacio predictor [41]. Podemos ver un ejemplo de la toma de decisión de las variables del árbol en la Figura 6.

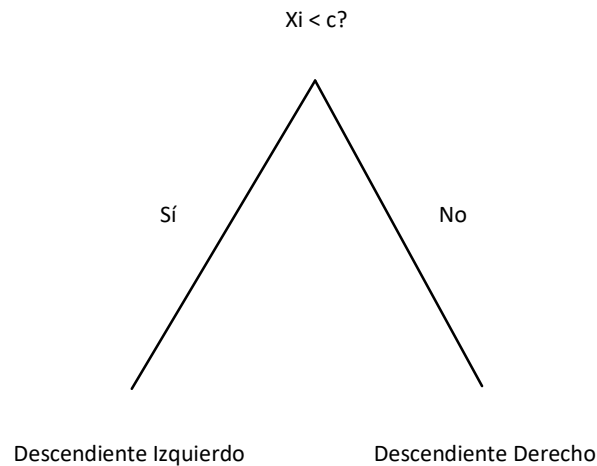


Figura 6: División en una variable predictora continua X_i utilizando el punto de división c .

La tasa de error de un bosque, por lo general, es la media de tasa de errores de los árboles individuales. Cuando hablamos en casos de clasificación es simple obtener esta tasa, debido a que podemos diferenciar sin mayor problema entre las distintas clases. Por otra parte, cuando hablamos de regresión, se complica debido a que una predicción puntual para un valor continuo (matemáticamente) es siempre incorrecta [42].

1.6.2. Métodos Neuronales

Perceptrón Multicapa

Antes de explicar la red neuronal, es importante entender qué es un perceptrón simple: Consiste en una función de decisión que pondera las entradas y las combina linealmente, seguida de una función de activación que produce una salida binaria. Sin embargo, el perceptrón tiene limitaciones: no puede resolver problemas no lineales como, por ejemplo, la operación 'XOR' [43].

“Multilayer Perceptron” (MLP) es una arquitectura de red de las RNAs supervisada y unidireccional, capaces de resolver problemas no lineales. [44]

Podemos ver la estructura de este tipo de red en la Figura 7. En general, está formada por una capa de entrada, varias capas ocultas y una capa de salida. Se conectan entre sí de tal manera que la entrada de las capas ocultas recibe la entrada de la capa previa a ellas. No existe conexión entre las neuronas de la misma capa [45] [44].

El aprendizaje de MLP se lleva a cabo de manera supervisada, utilizando un conjunto de datos que incluye tanto ejemplos de entrada como las salidas deseadas correspondientes. Este proceso se enfoca en minimizar los errores entre las salidas generadas por la red y las salidas esperadas, mediante un método conocido como

retropropagación del error o "backpropagation" [46]. La retropropagación ajusta los pesos de la red de manera iterativa, propagando el error hacia atrás a través de las capas. Esto se logra calculando el gradiente del error con respecto a cada peso y ajustando los pesos en la dirección que reduce el error, mejorando así la precisión de las predicciones de la red [46].

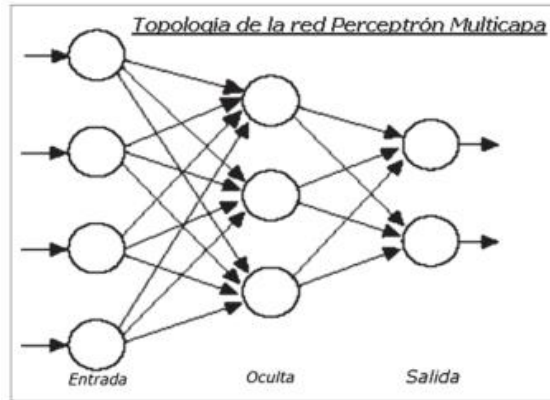


Figura 7: Arquitectura de una red neuronal Perceptrón Multicapa. Obtenido de: [45]

Las neuronas están conectadas a partir de pesos sinápticos los cuáles se ajustan durante el entrenamiento de la red y constituyen el conocimiento de la red neuronal. [45]

La operación que realizan los perceptrones viene dada por la ecuación 6 [47].

$$z_k = \sum_{j=1}^q w'_{kj} y_j - \theta'_k = \sum_{j=1}^q w'_{kj} f \left(\sum_{i=1}^n w_{ji} x_i - \theta_j \right) - \theta'_k$$

Ecuación 6

Donde:

x_i entradas de la red

y_j salidas, w_{ij} pesos y θ_j umbrales de la capa oculta

z_k salidas, w_{kj} pesos y θ'_k umbrales de la capa de salida.

f es la función de activación

Para todo $i = 1, \dots, n$; $j = 1, \dots, q$; $k = 1, \dots, m$.

MLP destaca por su capacidad para descubrir patrones sobre el conjunto de datos de entrenamiento, sin supuestos estadísticos. Pero a su vez, al igual que la mayoría, una de las desventajas es el costo computacional y temporal a la hora de ser entrenado y el problema del sobreentrenamiento, aunque es una de las arquitecturas neuronales más utilizadas en todas las aplicaciones reales. [45].

Gas Neuronal Creciente

La RNA "Growing Neural Gas" (GNG) es un modelo de aprendizaje autoorganizativo dinámico, es decir, que es capaz de resolver problemas solamente con datos de entrada, sin información adicional de los resultados esperados. Su función es la cuantificación espacial de datos [48].

Esta red está diseñada de tal manera que permite que el número de neuronas que contiene aumente o disminuya según sea necesario para mantener una medida del error que acumulan cada una de las neuronas de la red. Este tipo de red se basa en un modelo ontogénico, que es un enfoque en el cual el sistema puede modificar su estructura interna y el número de sus componentes (en este caso, neuronas) en respuesta a las necesidades del entorno o la tarea específica. Cuando se han llevado a cabo un cierto número de épocas (λ) se realiza el proceso de neurogénesis: se añade una neurona artificial (a lo largo del documento se usará "neurona") entre la neurona con mayor error acumulado y su vecina con mayor error acumulado. En caso de que, al finalizar una iteración, alguna neurona no tenga conexiones, esta es eliminada de la red teniendo lugar una apoptosis neuronal [49].

La GNG consta de [50]:

- Un conjunto de A unidades (o nodos). Cada unidad $c \in A$ tiene asociado un vector de referencia $w_c \in \mathbb{R}^n$. Los vectores de referencia pueden considerarse como posiciones en el espacio de entrada de las unidades correspondientes.
- Un conjunto de N conexiones (o aristas) entre pares de unidades. Estas conexiones no están ponderadas. El propósito de estas es la definición de estructura topológica.

Además, se trabajará con un espacio de datos de entrada o entorno de información, que obedecen a alguna función de densidad de probabilidad desconocida $P(\varepsilon)$ [50].

El algoritmo de GNG consiste en lo siguiente [50]:

Algoritmo 1: Algoritmo GNG

1. Comienza con 2 unidades a y b en posiciones aleatorias W_a y W_b en el espacio de entrada \mathbb{R}^n .
2. Genera una señal de entrada de acuerdo con $P(\varepsilon)$.
3. Encuentra la unidad más cercana s_1 y la segunda unidad más cercana s_2 .
4. Incrementa la edad de todas las aristas que derivan de s_1 .
5. Agrega la distancia al cuadrado entre la señal de entrada y la unidad más cercana en el espacio de entrada a una variable de contador local:

$$\Delta error(s_1) = ||w_{s_1} - \varepsilon||^2$$

6. Mueve s_1 y sus vecinos topológicos directos hacia $P(\varepsilon)$ por factores E_b y E_n , respectivamente, de la distancia total:

$$\Delta w_{s_1} = \epsilon_b(\varepsilon - w_{s_1})$$

$$\Delta w_n = \epsilon_n(\varepsilon - w_n) \text{ para todos los vecinos } n \text{ de } s_1$$

7. Si s_1 y s_2 están conectados por una arista, establece la edad de esta arista en cero. Si no existe tal arista, créala.

8. Elimina las aristas con una edad mayor que a la edad máxima. Si esto resulta en puntos sin aristas derivadas, elimínalos también.
9. Si el número de señales de entrada generadas hasta ahora es un múltiplo entero de un parámetro λ , inserta una nueva unidad de la siguiente manera:
 - Determina la unidad q con el máximo error acumulado.
 - Inserta una nueva unidad r a mitad de camino entre q y su vecino f con la variable de error más grande:
$$w_r = 0.5 (w_q - w_f)$$
 - Inserta aristas que conecten la nueva unidad r con las unidades q y f , y elimina la arista original entre q y f .
10. Disminuye todas las variables de error multiplicándolas por una constante d .
11. Si no se cumple el criterio de parada se continúa con el paso 2.

Podemos ver el comportamiento de GNG en la Figura 8. GNG empieza con un número reducido de neuronas e incrementa a medida que lo necesita [50].

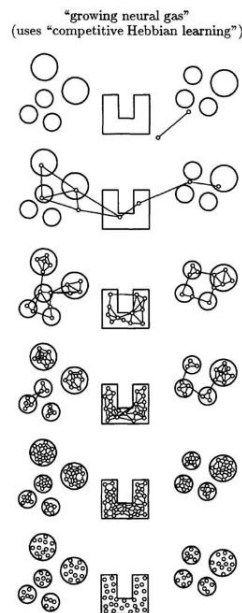


Figura 8: Comportamiento de una red de gas neuronal creciente, empezando por sus estados iniciales (fila superior) hasta sus estados finales (fila inferior). Obtenido de [50].

GNG tiene un gran potencial no solo por no necesitar conocer entradas entrada, sino también por ser una red no supervisada que no requiere conocer la salida deseada. Además, posee la capacidad de adaptar dinámicamente el espacio de entrada, mostrando un alto carácter ontogénico. Por ello, son ideales para problemas en los que desconocemos total o parcialmente la distribución de entrada, tamaño de la red u otras características [51].

Por otra parte, una desventaja notoria es que una inadecuada combinación de hiperparámetros puede derivar en dos problemas: que la red no crezca lo suficiente y por lo tanto no represente correctamente los datos, y la creación de pequeños clústeres que representen de manera errónea el espacio de datos [52].

Valiéndose de la bondad de GNG, e intentando superar carencias y mejorar el desempeño y rendimiento de arquitecturas neuronales supervisadas paradigmáticas, surgen nuevas arquitecturas como “Modular Hybrid Growing Neural Gas” (MyGNG) y SuperGNG.

- MyGNG se trata de una arquitectura neuronal modular e híbrida formada por una primera capa de entrada, un módulo autoorganizado a partir de GNG y una última capa que se basa en el algoritmo del perceptrón (algoritmo supervisado). El objetivo del desarrollo de esta nueva arquitectura neuronal ontogénica híbrida ha sido aportar un sistema inteligente de mayor capacidad y efectividad en el diagnóstico diferencial de la EA [53]. Podemos apreciar la estructura de la red MyGNG en la Figura 9.

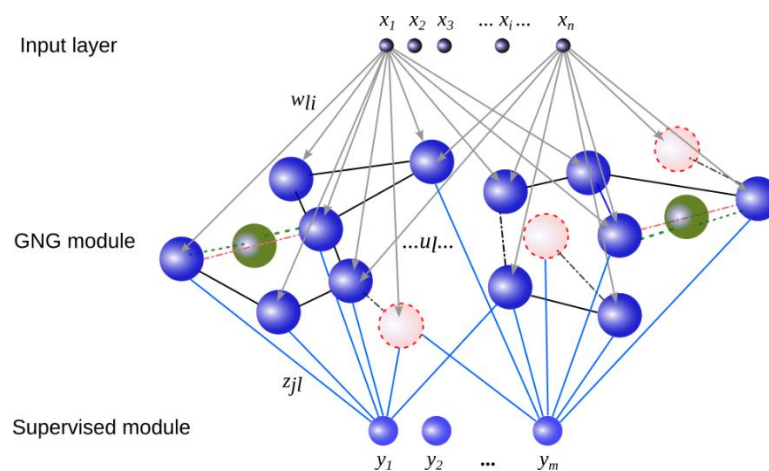


Figura 9: Estructura de MyGNG. Obtenido de: [53]

- SuperGNG es una versión supervisada y mejorada de GNG, debido a que posee una característica reconfigurable dentro del algoritmo GNG, que le hace capaz de resolver problemas y muy especialmente en el ámbito del diagnóstico de la EA [53]. Esta arquitectura neuronal ontogénica es capaz de cambiar sus conexiones y adaptar automáticamente su topología durante el aprendizaje basándose en las clases de los datos, a diferencia de otras arquitecturas [52]. Podemos ver la estructura de esta red neuronal en la Figura 10.

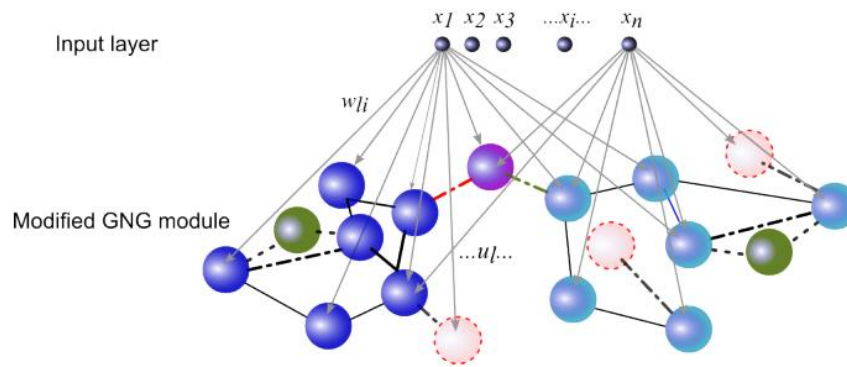


Figura 10: Estructura de SuperGNG, donde x_i es el componente i -th del vector de entrada; w_{li} el peso de la neurona li y u_i la neurona del módulo GNG. Fuente: [52].

Algoritmo 2: Algoritmo Red Ontogénica SuperGNG

Esta nueva implementación basada en GNG introduce un paso adicional entre los pasos 9 y 10 del algoritmo GNG mencionado previamente. Este nuevo paso, consta de las siguientes operaciones:

- Separar los clústeres/neuronas que están conectadas, pero no deberían porque pertenecen a distintas clases.
- Conectar los clústeres/neuronas que no están conectadas, pero deberían porque detectan entradas pertenecientes a la misma clase.

Para realizar estos pasos, se añaden dos parámetros a la red:

- El primero indica a partir de qué número de épocas se empiezan a realizar estas operaciones. Este parámetro se añade debido a que, si se empezara desde épocas muy tempranas a realizar estas operaciones, la red podría no crecer.
- El segundo indica la frecuencia para realizar las operaciones dentro de cada época, es decir, cada cuantas iteraciones se llevan a cabo. Esto depende del tamaño del conjunto de datos que se esté evaluando.

Como se ha observado en las explicaciones anteriores, estos cambios se realizan mientras la red se entrena, lo cual podría resultar ventajoso de cara a mejorar el rendimiento de GNG.

1.7. Coeficientes de calidad de clúster

Uno de los pasos importantes cuando trabajamos con data driven methods (métodos basados en datos) son los correspondientes a determinados métodos de data mining que nos ayudan a preparar adecuadamente los datos para optimizar también, la eficacia del método de procesado a usar. Es decir, pre-procesamos los datos antes de constituir los vectores de características que alimentarán el sistema basado en redes neuronales, con el que se va a trabajar. Es por ello tan importante identificar un buen coeficiente para obtener la calidad del clúster. Es esto lo que justifica esta última sección del capítulo.

Así, los objetivos de este apartado consisten en identificar el coeficiente de *clustering* más adecuados para comparar diversas combinaciones de proyección, escalado y componentes, con el fin de encontrar cuáles de ellas tienen una mejor separación entre las diferentes clases, para facilitar su posterior clasificación.

Por lo general, para el estudio de validación de *clustering* se emplean tres coeficientes: Silhouette, Davies-Bouldin (DB) y Calinski-Harabasz (CH). En este trabajo se propone emplearlos para evaluar la calidad de clústeres antes de su clasificación.

Antes de empezar con los coeficientes debemos tener claro los conceptos de cohesión y compacidad en el contexto de los clústeres. Mientras que una alta cohesión significa que los puntos dentro de un clúster están más cerca de su centroide, una alta compacidad significa que los puntos dentro de un clúster están más cerca unos de otros, lo que indica que el clúster es más compacto.

Por otra parte, también hay que comentar los conceptos de inter-clúster e intra-clúster. Así, el intra-clúster es la distancia entre puntos de datos dentro del mismo clúster y el inter-clúster hace referencia a la distancia entre muestras de diferentes clústeres de un conjunto. Observamos estos conceptos en la Figura 11.

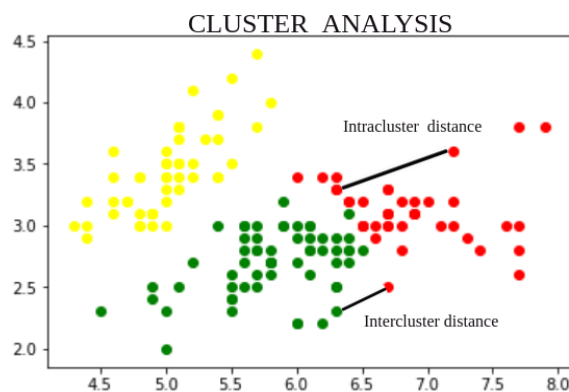


Figura 11: Intra-Clúster e Inter-Clúster. Obtenido de: [54]

Con respecto a los métodos de pre-procesado que hemos usado en nuestro estudio, han sido los de proyección siguientes: Unprojected, PCA, Incremental PCA, FastICA, Kernel PCA rbf, Kernel PCA linear, Kernel PCA poly, KernelPCA_sigmoid, KernelPCA cosine, Neighborhood CA, Dictionary Learning, Truncated SVD, NMF, Factor Analysis, Linear Discriminant Analysis.

Por otro lado, los escalados fueron: Max Abs Scaler, Min Max Scaler, Normalizer, Power Transformer Yeo-Johnson, Quantile Transformer Gaussian, Quantile Transformer Uniform, Robust Scaler, Estándar Scaler, Unscaled.

Los tres coeficientes se utilizan como indicadores de validación interna. Estos evalúan características inherentes de los datos, como su compacidad, separación y cohesión. Su evaluación no depende de información externa.

1.7.1. Coeficiente de Silhouette

Este coeficiente permite cuantificar cuán buena es cada una de las asignaciones de las muestras a un clúster comparando su similitud con el resto de las muestras del mismo clúster. [55]

Para cada muestra i , el coeficiente de silhouette (s_i) es obtenido de la siguiente manera [56] [55]:

1. Determina la media de las distancias (a_i) entre la observación i y las demás observaciones dentro del mismo clúster. Cuanto menor sea a_i , más similar será con las otras observaciones en su clúster.
2. Calcula la distancia promedio entre la observación i y los clústeres restantes. Aquí, la distancia promedio entre i y un clúster C se define como la media de las distancias entre i y todas las observaciones en el clúster C .
3. Identifica b_i como la menor de las distancias promedio entre i y los otros clústeres, es decir, la distancia al clúster más cercano.
4. Calcular el valor de silhouette como:

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)}$$

Ecuación 5: Fórmula del coeficiente de silhouette

Un ejemplo de los elementos involucrados en el cálculo del coeficiente de silhouette, se puede ver en la Figura 2, donde la muestra i pertenece al clúster a , donde vemos que solo se escoge el clúster más cercano para evaluar la puntuación de cada una de las muestras [56].

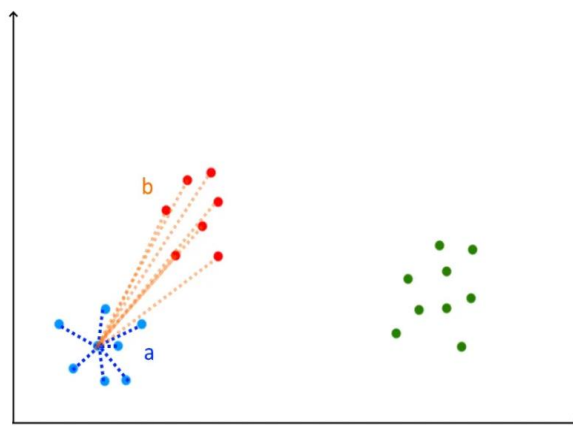


Figura 12: Cálculo del coeficiente de silhouette. Obtenido de [57]

Su valor oscila entre -1 y +1, donde un valor alto indica una buena concordancia del objeto con su propio cúmulo y una mala concordancia con los cúmulos vecinos [56] [58].

- Un valor del coeficiente cercano a +1 indica que el punto de datos está en el clúster correcto [59].

- Un valor del coeficiente cercano al 0 significa que los clústeres están superpuestos (está entre dos clústeres) [59] [60].
- Un valor del coeficiente cercano a -1 significa que el punto de datos está en un clúster incorrecto [59].

El valor promedio de los coeficientes de silhouette de todas las muestras es el que nos indica la calidad del clúster [61].

El coeficiente de Silhouette produce una mayor precisión en los resultados que el coeficiente DB. Sin embargo, la complejidad computacional de Silhouette es mucho mayor que DB [58].

1.7.2. Coeficiente Davies-Bouldin

Cuando este coeficiente tiene valores cercanos al 0, se considerará una mejor calidad de la agrupación. Esto señala que las muestras están más compactas dentro de sus grupos, al tiempo que están bien separadas del resto [62].

Este coeficiente [63] se define como:

$$S_i = \frac{1}{k} \sum_{i=1, i \neq j}^k \left(\frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right)$$

Ecuación 6: Fórmula de DB

Donde:

- k es el número de clústeres
- σ_i es la distancia promedio entre cada punto en el clúster i y el centroide del clúster
- σ_j es la distancia promedio entre cada punto del clúster j y el centroide del clúster
- $d(c_i, c_j)$ es la distancia entre los centroides de los 2 clústeres.

Valores pequeños para el coeficiente DB indica clústeres compactos, y cuyos centros están bien separados los unos de los otros [60].

En la Figura 3 podemos observar los componentes que están involucrados en el cálculo del coeficiente de DB.

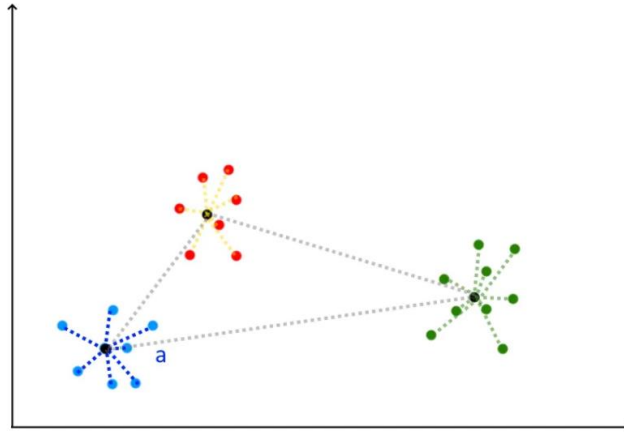


Figura 13: Cálculo del coeficiente de DB. Obtenido de [57]

Debido a su naturaleza de coeficiente de representante único, es más efectivo cuando se trata de conglomerados de forma convexa [64].

1.7.3. Coeficiente Calinski–Harabasz

Esta métrica posibilita la comparación de dos agrupamientos para determinar cuál presenta una definición más precisa de sus grupos. Cuanto mayor sea este indicador, menor será la dispersión de las muestras dentro de los grupos y, a su vez, mayor será la separación entre los grupos [62].

La fórmula de CH consiste en una ecuación donde el numerador es la suma de cuadrados entre grupos (Between-Group Sum of Squares) (BGSS) dividida por sus grados de libertad, y el denominador es la suma de cuadrados dentro del grupo (Within-Group Sum of Squares) (WGSS) dividida por sus grados de libertad [65].

$$CH = \frac{BGSS}{WGSS} \times \frac{n_e - k}{k - 1}$$

Ecuación 7: Fórmula de CH

Donde:

- n_e es el número de puntos de datos.
- k es el número de clústeres.
- $BGSS$ es la suma ponderada de distancias euclidianas al cuadrado entre cada centroide de conglomerado (media del conglomerado) y el centroide de datos general (media de todos los datos).

$$BGSS = \sum_{q \in k} n_q (c_q - c_e)(c_q - c_e)^T$$

Ecuación 8: Suma de cuadrados entre conglomerados

- Donde n_q es el número de puntos en el clúster q , c_q es el centroide de q y c_e es el centroide general de los datos.
- $WGSS$ es la suma de distancias euclidianas al cuadrado entre los puntos de datos y sus respectivos centroides del grupo:

$$WGSS = \sum_{q \in k} \sum_{x \in \text{cluster } q} (x - c_q)(x - c_q)^T$$

Ecuación 9: Suma de cuadrados dentro del grupo

- Donde n_q es el número de puntos en el clúster q , c_q es el centroide de q y c_e es el centroide general de los datos.

En la Figura 4 podemos ver un ejemplo de los componentes que influyen en el cálculo del coeficiente de CH.

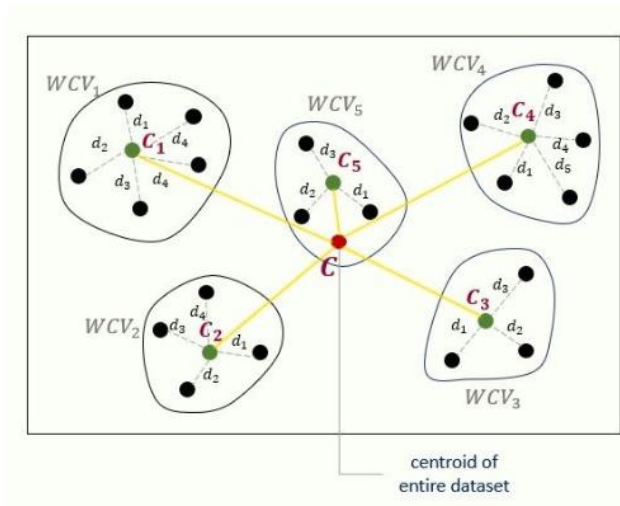


Figura 14: Cálculo del coeficiente de CH. Obtenido de [66]

Capítulo 2: Desarrollo Del Sistema Software

En este capítulo se trata el sistema software de ayuda al diagnóstico temprano de la EA que se ha desarrollado para realizar el estudio objetivo de este TFT.

2.1. Metodología Software

La metodología utilizada en este trabajo ha sido una metodología en cascada que consta de tres fases, cada una con un enfoque específico y una serie de actividades diseñadas para asegurar el éxito del desarrollo del sistema software. A continuación, se describen estas fases en detalle:

Fase de Declaración Explícita del Problema

- **Objetivo y Requisitos Funcionales:** En esta fase se plantea claramente el objetivo del sistema y se definen los requisitos funcionales que debe cumplir los problemas específicos que el sistema debe resolver.

Fase de Diseño del Sistema

- **Diseño del Sistema:** Se lleva a cabo el diseño detallado del sistema, que incluye la integración de la red SuperGNG y otros clasificadores similares que serán utilizados para comparaciones posteriores.

Fase de Construcción y Realización de Pruebas del Sistema

- **Implementación:** En esta fase, se pone en práctica el diseño realizado en la fase anterior. Esto implica codificar el sistema.
- **Pruebas del Sistema:** Se realizan pruebas tanto de las funciones utilizadas, como de los resultados que se esperan de cada uno de los flujos de trabajo, para verificar el correcto funcionamiento del sistema.

El desarrollo de este sistema fue un esfuerzo colaborativo llevado a cabo junto con los integrantes del grupo de investigación COMCIENCIA. Se realizaron reuniones semanales para discutir dudas y resolver problemas surgidos en el código, fomentando un entorno de trabajo en equipo y asegurando que el proyecto sigue un desarrollo esperado.

Dado que la red SuperGNG incluye 18 parámetros y presenta una gran complejidad, recibí asistencia del grupo de investigación, especialmente de Ylermi Cabrera León, quien propuso la red y posee un conocimiento profundo de su funcionamiento. Esta colaboración fue esencial para el análisis y ajuste de los nuevos parámetros, permitiéndome desglosar los resultados y realizar un estudio exhaustivo. Gracias a esta

ayuda, fue posible llevar a cabo un análisis detallado y una evaluación comparativa efectiva de los clasificadores en diversos conjuntos de datos.

2.2. Arquitectura del sistema software

El objetivo es la implementación de un sistema software, a partir de la creación o integración de componentes de software, para llevar a cabo un análisis comparativo que permita evaluar el desempeño de la red ontogénica SuperGNG en el diagnóstico temprano de enfermedades neurodegenerativas.

2.2.1. Requisitos funcionales

Los requisitos funcionales de la arquitectura del sistema software a desarrollar son los siguientes:

- Carga / Creación de Datos:
 - Implementar funciones para leer conjuntos de datos desde archivos en formatos comunes como CSV, Excel o archivos de texto.
 - Desarrollar capacidades para generar conjuntos de datos sintéticos según las especificaciones requeridas.
- Preprocesamiento de Datos:
 - Integrar/Implementar las funciones de preprocesamiento incluidas tareas propias de data mining tales como: limpieza, normalización, transformación, escalado y proyección de datos.
 - Incorporar funciones para particionar datos para su entrenamiento y manejo de valores faltantes, características irrelevantes y variables categóricas.
 - Permitir obtener los coeficientes de clustering de conjuntos de datos.
- Integración de clasificadores
 - El sistema tendrá la integración de los clasificadores MLP, SVM, NB, DT y RF.
 - El sistema debe integrar la arquitectura neuronal ontogénica SuperGNG.
- Entrenamiento de los clasificadores:
 - El sistema debe proporcionar las funcionalidades necesarias para el entrenamiento de los clasificadores.
 - Debe permitir ajustar parámetros de distintos clasificadores que tenga integrados que le permita realizar procesos de comparación, testeo y evaluación de los mismos.
- Validación y Evaluación del Modelo:
 - Se debe implementar un módulo de validación cruzada para evaluar el rendimiento del modelo entrenado.
 - Debe incluir métricas de evaluación como precisión, sensibilidad, especificidad y área bajo la curva ROC.
 - El sistema será capaz de crear gráficas en 3 dimensiones para estudiar las métricas de rendimiento.
- Análisis estadístico

- Se debe implementar un módulo capaz de calcular el p-valor y otros descriptivos estadísticos de los datos de entrada.
- Debe incluirse el cálculo del p-valor para las tablas de resultados de los clasificadores previo a la comparación de resultados (estudio de significancia de los resultados a la hora de comparar).
- Exportación de Resultados:
 - El sistema debe permitir exportar los resultados del análisis, incluidas las métricas de rendimiento del modelo en el formato que se le indique.

2.2.2. Diseño y desarrollo del sistema software

El trabajo realizado consistió en implementar un sistema utilizando una serie de módulos y bibliotecas ya existentes, creando además los módulos necesarios para el correcto funcionamiento del mismo. Se optimizó y adaptó el código preexistente para satisfacer las necesidades específicas del proyecto. Se seleccionaron aquellos módulos y bibliotecas que se alineaban con los requerimientos del proyecto. Además, se desarrollaron nuevos módulos de código, asegurando en todo momento no romper ni alterar el funcionamiento del código existente.

El diseño del sistema se basa en tres flujos de trabajo, los cuáles podemos observar en la Figura 15, en este diagrama se indican los procesos que incluyen cada uno de los flujos de trabajo a partir de recuadros de color que se le ha asignado a cada uno. Si nos fijamos en detalle, dos de los bloques que se muestran (carga/creación de los datos y el preprocesado) son comunes a todos los flujos de trabajo, y es por ello por lo que el sistema se encargará de gestionar el uso de estos bloques acorde a la necesidad que exija cada uno de ellos.

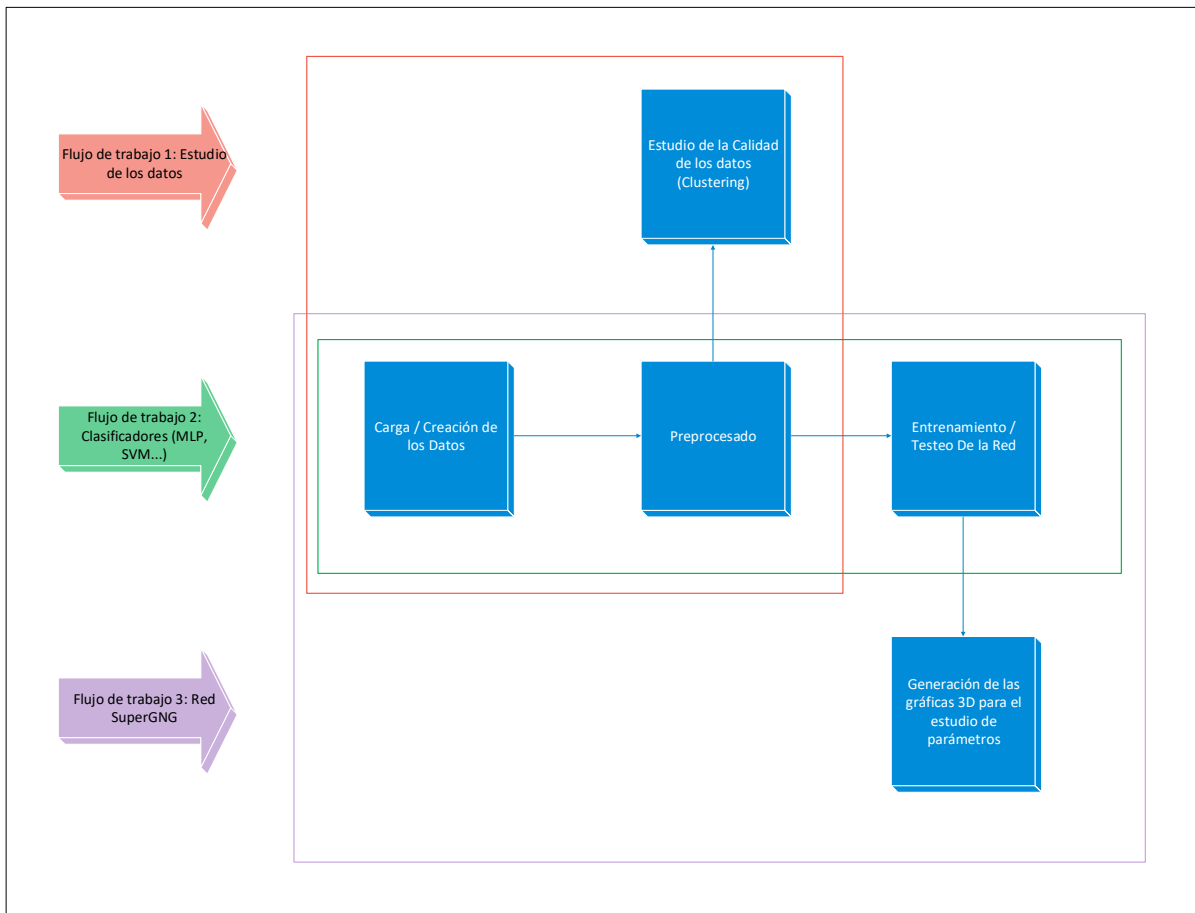


Figura 15: Diagrama del diseño del sistema

Teniendo en cuenta los módulos ya existentes, en algunos de los cuales han sido necesarias modificaciones para poder ser integrados en el sistema, se ha realizado un diagrama de componentes UML (Figura 16), explicando las distintas relaciones entre cada uno de ellos. Todo el sistema se centra en la función `manageWorkFlow()` que es la responsable de gestionar el correcto funcionamiento de los distintos flujos de trabajo.

En el diagrama de UML de componentes, Figura 17, se marcan en verde los módulos desarrollados para este trabajo, incluyendo la lógica de flujos de trabajo, el cálculo de estadísticas y la elección de gráficas 3D, así como la gráfica surface 3D.

En naranja, el módulo que estaba previamente desarrollado pero ha sido necesario desarrollar y refactorizar nuevas funcionalidades para el sistema, como puede ser la clasificación de 3 clases o la inclusión de la Red SuperGNG para graficar junto a los otros clasificadores resultados a través de las medidas de rendimiento de los distintos clasificadores, como la curva ROC.

Los módulos relacionados con el desarrollo de la Red Ontogénica SuperGNG se destacan en rojo.

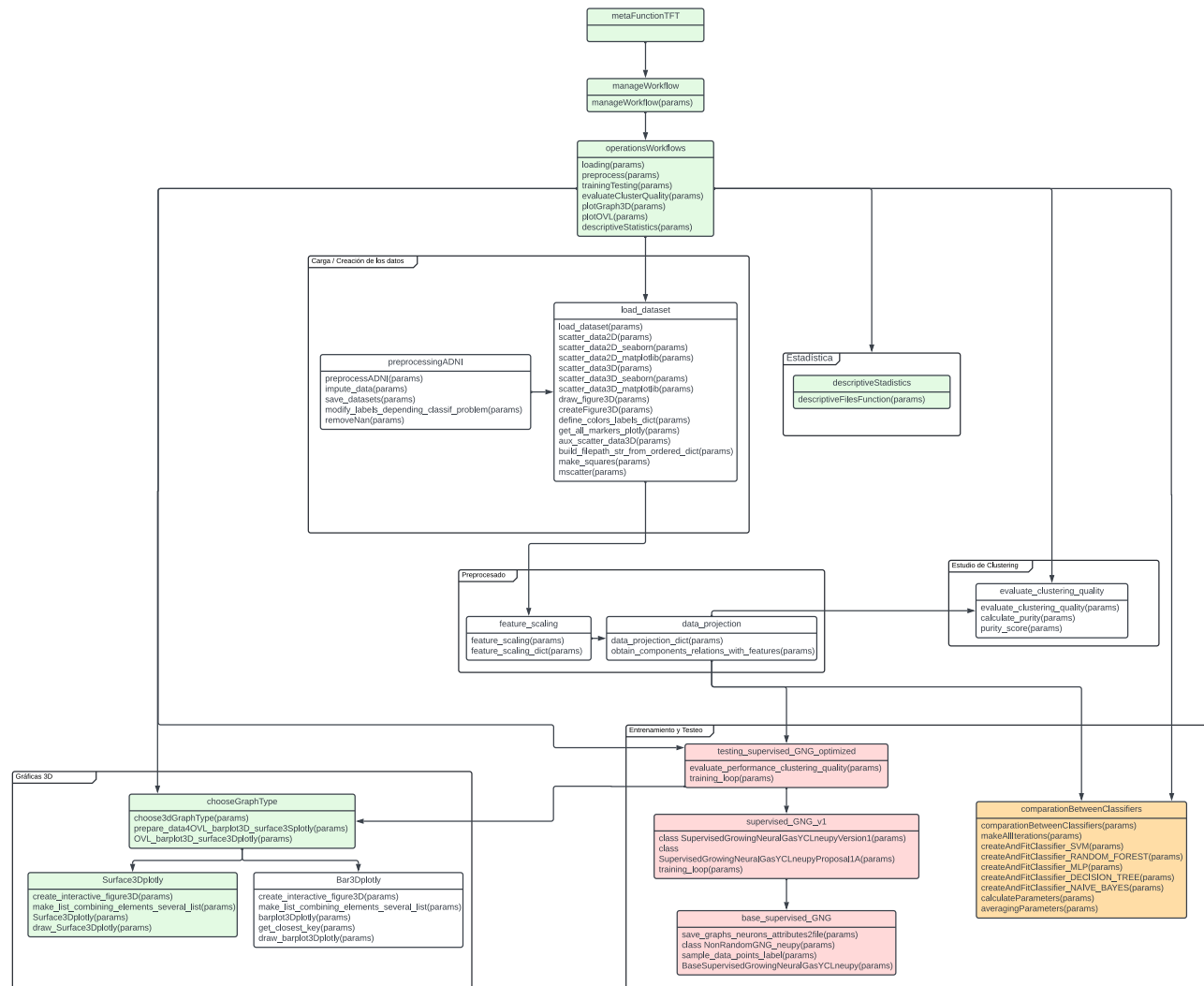


Figura 16: Diagrama UML del sistema software

Para mayor simplicidad, este diagrama presenta sólo los módulos principales del sistema, omitiendo los módulos auxiliares como el guardado de archivos o la creación de formatos de datos.

En el módulo "metaFunctionTFT.py" se encuentran los parámetros necesarios para realizar la llamada a los distintos flujos de trabajo. El módulo "manageWorkFlow" realiza llamadas a las funciones del módulo "operationsWorkFlow.py" para ejecutar las operaciones de los flujos de trabajo.

El módulo "descriptiveStatistics.py" incluye tanto el cálculo de estadísticas descriptivas para datos de entrada (media, mediana, etc.) como el cálculo del p-valor al realizar comparaciones entre clasificadores.

Otro módulo a destacar es "comparationBetweenClassifiers.py". Inicialmente, este módulo clasificaba problemas binarios, pero se ha ampliado para incluir la clasificación de clases ternarias y la posibilidad de añadir la red SuperGNG a los resultados finales utilizando sus predicciones. Este módulo puede ser utilizado tanto para una sola ejecución como para una ejecución particionada, calculando posteriormente la media y exportando los datos resultantes.

2.2.3. Infraestructura software y hardware

Para este trabajo se han utilizado dos equipos informáticos uno para la ejecución de los clasificadores y otro para la implementación del sistema y de los conjuntos de datos sintéticos. Las características de estos equipos se pueden ver en la Tabla 3.

Equipo	HP ENVY Desktop TE02-0xxx	Victus by HP Gaming Laptop 16-r0xxx
Procesador	12 th Gen Intel® Core™ i7-12700, 2100 Mhz, 12 procesadores principales, 20 procesadores lógicos	13th Gen Intel® Core™ i7-13700H, 2400Mhz, 14 procesadores principales, 20 procesadores lógicos
Gráfica	NVIDIA GeForce RTX 3060, 12 GB	NVIDIA GeForce RTX 4060 Laptop GPU, 8GB
Sistema Operativo	Windows 11 Home, versión 23H2	Windows 11 Home, versión 23H2
Memoria RAM	32 GB	16 GB
Memoria de Almacenamiento	1 TB	1 TB

Tabla 3: Características de los equipos utilizados

Para la realización de este proyecto se utiliza Python 3.6.8. El uso de esta versión se a una restricción en la versión de TensorFlow requerida por una de las bibliotecas (neupy) que utiliza la implementación de la red ontogénica SuperGNG.

Las principales librerías usadas han sido las siguientes:

- Neupy (versión 0.8.2): Se utiliza para RNAs y aprendizaje profundo. Enfocado a este proyecto, contiene la red GNG. Enlace página principal: <http://neupy.com>.
- Sklearn (versión 0.22.1): Contiene un conjunto de módulos para AA y la minería de datos. Enlace página principal: <https://pypi.python.org/pypi/scikit-learn/>.
- Pandas (versión 0.25.3): Librería que contiene potentes estructuras de datos para el análisis de datos, las series temporales y las estadísticas. Enlace página principal: <http://pandas.pydata.org>.
- Tensorflow (versión 1.15.0): Una biblioteca de código abierto para la computación numérica y el AA a gran escala. Facilita la implementación de redes neuronales profundas y otras aplicaciones de aprendizaje automático. Enlace página principal: <https://www.tensorflow.org/>.
- Numpy (versión 1.19.5): Proporciona soporte para matrices grandes y multidimensionales junto con una colección de funciones matemáticas de alto nivel para operar con ellas. Enlace página principal: <https://numpy.org/>.

2.2.4. Sistema Software

El sistema software fue mejorado en funcionalidad de forma dinámica, después de haber realizado operaciones de testeo del mismo, donde se detectó la necesidad de incorporar la implementación de nuevos flujos de trabajo que se adaptan mejor a las necesidades del estudio. Para la elección de los distintos flujos de trabajo en la ejecución del sistema se utiliza la función “selectWorkflow()” que vemos en la Figura 17.

```
def selectWorkflow():
    if(workflow == 1):
        workflow1(loadingDictInput, preprocessDict, trainingDict, graphDict)
    elif(workflow == 2):
        workflow2(loadingDictInput, preprocessDict, trainingDict, graphDict)
    elif(workflow == 3):
        workflow3(loadingDictInput, preprocessDict, clusteringDict)
    elif(workflow == 4):
        workflow4(statisticsDict)
    elif(workflow == "3D" or workflow == "3d"):
        workflow3D(graphDict)
    else:
        print("Error: The workflow is not valid")
        exit()
```

Figura 17: Flujos de trabajo Implementados en el Sistema Software Final

Los nuevos flujos de trabajo que se han implementado se ven en la Tabla 4, donde se indica, además, cuáles son las funciones que realiza cada uno. Los flujos de trabajo 1, 2 y 3 son los expuestos en el diseño del sistema (Figura 15). El número 4 es el cálculo de estadísticas tanto de los datos de entrada como de comparativa entre clasificadores.

El flujo de trabajo “3D” es auxiliar y se encarga de graficar las métricas de rendimiento de los distintos clasificadores, partir de resultados.

	Carga	Preprocesado	Entrenamiento y Testeo	Métricas 3D	Estudio Clustering	Obtención de Estadísticos
Flujo de trabajo 1 (SupeRGNG)	X	X	X	X		
Flujo de trabajo 2 (MLP, SVM...)	X	X	X			
Flujo de trabajo 3 (Clustering)	X	X			X	
Flujo de trabajo 4 (Estadística)						X
3D				X		

Tabla 4: Flujos de trabajo en el sistema final

Un ejemplo de función de flujo de trabajo se puede ver en la Figura 18, donde se observa que se encuentran bien definidas las 3 funciones de las que se hace uso, carga, preprocesado y entrenamiento de la red.

```
def workflow1(loadingDictInput, preprocessDict, trainingDict, graphDict):
    data, labels_true, features_names, saving_path0, title_str_d, labels_column_name, save_model_data_labels2binary_file_path, \
    color_list, use_centers_str, all_configs_dict, = \
        loading(loadingDictInput)

    scaled_projected_data_dict, keys_where_valueerror_happened_s_p= \
        preprocess(data, labels_true, features_names, saving_path0, title_str_d, preprocessDict)

    trainingTesting(saving_path0, scaled_projected_data_dict, keys_where_valueerror_happened_s_p, title_str_d, labels_true, color_list,
        features_names, labels_column_name, all_configs_dict, use_centers_str, save_model_data_labels2binary_file_path,
        trainingDict, graphDict)
```

Figura 18: Código del Flujo de trabajo 1

Capítulo 3: Análisis de la SuperGNG para la detección de la Enfermedad del Alzheimer

La solución computacional para la ayuda al diagnóstico de la EA que proponemos en este trabajo es un sistema inteligente basado en arquitecturas neuronales ontogénicas híbridas, concretamente la SuperGNG, la cual fue descrita en el primer capítulo de esta memoria.

Nuestra propuesta consiste en la detección de la EA frente a sujetos sanos, desarrollando un sistema que sea de uso en atención primaria.

Determinaremos también, un conjunto mínimo de criterios diagnósticos óptimos que puedan convertirse en un conjunto de criterios diagnósticos estándar de uso clínico.

Este estudio y desarrollo será el contenido de este capítulo, distribuido de la siguiente forma: Optimizaremos la arquitectura SuperGNG para lo cual realizaremos un estudio paramétrico de la misma usando datos sintéticos. Posteriormente se realiza el estudio de preprocesado con los coeficientes de clustering, seguido del estudio con datos reales de Alzheimer y finalmente, la discusión de los resultados obtenidos.

3.1. Estudio paramétrico de la SuperGNG usando conjuntos de datos sintéticos

Para evaluar los parámetros de la red, se han escogido tres conjuntos de datos sintéticos distintos, los cuales presentan características de distribución y estructura que reflejen condiciones propias de los datos clínicos reales para el diagnóstico de la EA, a los cuales se tendrá que enfrentar la red neuronal SuperGNG.

Así los conjuntos se deben a que presentan los siguientes retos a la hora de clasificar las clases:

- Blobs: Desarrollar clasificadores con capacidad de tratar y resolver problemas con clases superpuestas entre sí.
- Muestras no balanceadas (Unbalanced Blobs): Desarrollar clasificadores con clases no balanceadas.
- Dos cuadrados concéntricos (Two Concentric Squares): Desarrollar clasificadores con capacidad para clasificar clases no linealmente separables.

A continuación, detallaremos las características de cada uno de los conjuntos de datos mencionados.

3.1.1. Selección de los conjuntos de datos sintéticos

Seleccionaremos para este estudio, conjuntos de datos que manifiesten las características propias de los datos clínicos reales a los que se enfrentará la SuperGNG en la ayuda al diagnóstico de la EA.

Blobs

Este conjunto de datos está formado por 999 elementos, igualmente repartidos entre 3 clases con distribución Gaussiana. Con este conjunto de datos analizaremos paramétricamente la SuperGNG frente al solapamiento de los datos. Este es un estudio fundamental para realizar, ya que en los datos clínicos el solapamiento entre las clases diagnósticas es una constante, siendo una fuerte característica en las demencias.

El número total de distribuciones diferentes de blobs será 11, denominadas variantes. En cada una de las cuales habrá un diferente solapamiento entre clases, aumentando el mismo en cada una de ellas, desde el no solapamiento, 0% de solapamiento, hasta un 62% de solapamiento.

Podemos ver la variación del conjunto de datos blobs en la Figura 19.

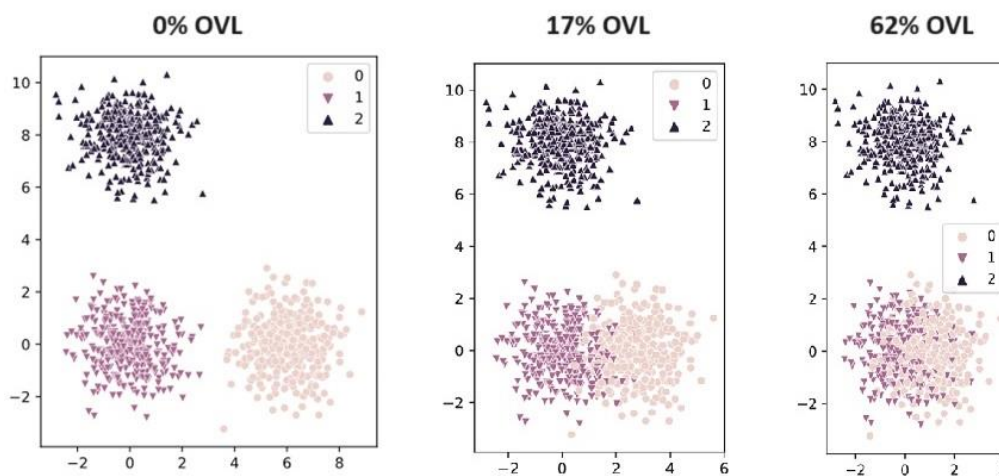


Figura 19: Conjunto de datos blobs con diferente grado de solapamiento

Además, con el fin de obtener más información, se han obtenido los valores de los 3 coeficientes de calidad de clustering que se explicaron anteriormente. Podemos ver estos resultados en la Tabla 5, donde claramente la calidad se reduce a medida que aumenta el porcentaje de solapamiento.

Blobs OVL (%)	silhouette	davies_bouldin	calinski_harabasz
0	0,90846617	0,37490792	5928,15268
1	0,89518183	0,39960389	5587,09477
2	0,8668889	0,44628394	5130,64937
5	0,82415727	0,51023	4740,37375
10	0,76066945	0,60319511	4416,2679
17	0,7039125	0,69278683	4236,95832
21	0,67063619	0,75073005	4158,33182
26	0,63426318	0,82124002	4087,05752
32	0,59488979	0,90890209	4023,13542
45	0,50921082	1,16876301	3917,34781
62	0,42146269	1,67482084	3840,96899

Tabla 5: Valores de clustering en blobs según el solapamiento existente

Unbalanced Blobs

Este conjunto de datos está compuesto de 999 elementos, de forma no balanceada entre tres clases con una distribución Gaussiana. Consiste en 9 variantes donde se aumenta el número de muestras en una de las clases, y este porcentaje de aumento disminuye en otra de las clases. Las distintas variantes se construyen aumentando un 10% las muestras.

Este conjunto de datos se ha seleccionado debido a que se estudia una característica que se muy presente en los datos médicos, el no balanceo de los mismos, lo que afecta tremendamente el proceso de clasificación y por ende de diagnóstico. Así habrá clases bien representadas y otras clases con bastante más baja representación lo que complica su detección.

El desbalanceo existente de este conjunto se puede observar en la Figura 20, se aumenta el porcentaje de muestras en una clase mientras que en otras disminuye.

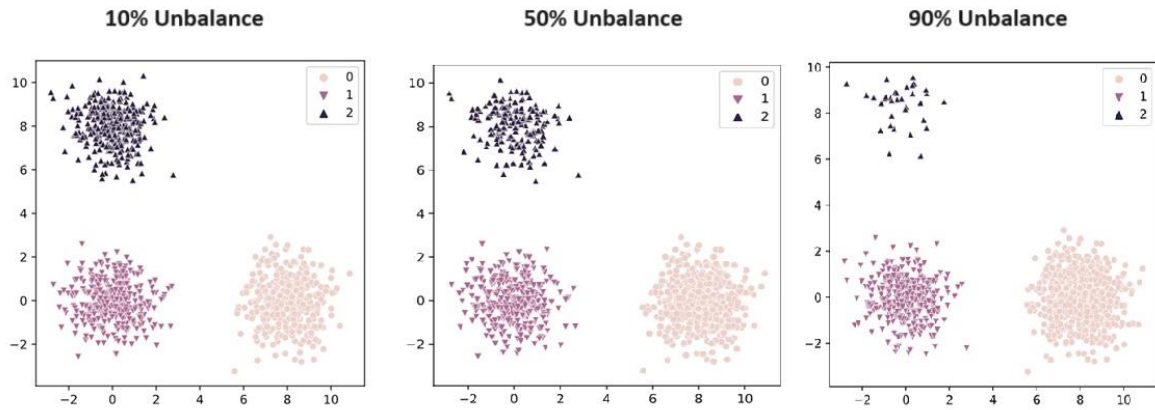


Figura 20: Blobs desbalanceados

Al igual que con blobs, se han obtenido valores de los coeficientes de clustering para cada una de las variaciones. En este caso, vemos que los coeficientes DB y CH no sufren cambios tan significativos como para el caso de blobs, aunque vemos que si existe una disminución de la calidad a medida que una de las clases tiene cada vez menos muestras. Por otra parte, la métrica Silhouette, que evalúa la separación entre muestras, mantiene un buen resultado en todas las variantes, ya que no se produce un solapamiento. Vemos esto en la Tabla 6.

Unbalanced-Blobs Unbalance (%)	silhouette	davies_bouldin	calinski_harabasz
10	0,937147624	0,306850209	7537,965964
20	0,93723561	0,304551423	7410,224893
30	0,937594435	0,302694294	7258,681007
40	0,938163292	0,300207747	7036,544523
50	0,938572695	0,298866266	6725,538122
60	0,938340527	0,297390066	6277,938948
70	0,938885557	0,295684943	5831,74911
80	0,938160611	0,295003546	5211,846999
90	0,938581165	0,289202044	4576,019058

Tabla 6: Valores de clustering en blobs desbalanceados

Two Concentric Squares

Por último, tenemos el conjunto de datos “Two Concentric Squares” con una distribución concéntrica y no linealmente separable compuesto de 1000 muestras, divididas en dos clases. Podemos ver la distribución de este conjunto de datos en la Figura 21.

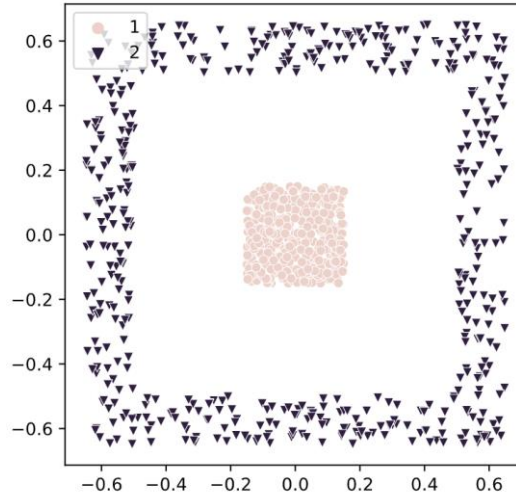


Figura 21: Representación del conjunto de datos Two Concentric Squares

Además, se ha calculado el coeficiente de cada uno de los coeficientes para este conjunto de datos. Vemos que, aunque las clases tienen una separación significativa, los valores de los coeficientes no indican este hecho debido a sus limitaciones. En este caso se remarca que en los conjuntos de datos que tienen formas determinadas, no funcionan correctamente estos coeficientes. Observamos estos resultados en la Tabla 7.

	silhouette	davies_bouldin	calinski_harabasz
Two Concentric Squares	0,22183175	129,347853	1661,2257

Tabla 7: Valores de clustering en Two Concentric Squares

3.1.2. Implementación de los conjuntos de datos sintéticos

Para la implementación de los conjuntos de datos sintéticos se ha utilizado la función “makeblobs()” de sklearn y se han ido cambiando las parámetros para obtener el comportamiento que se espera en cada caso. Estas funciones se utilizan para crear tanto “blobs” como “unbalanced blobs”. Podemos ver el fragmento de código de esta implementación en la Figura 22.

```

elif dataset_name in ('blobs', 'gaussian'):
    #https://scikit-learn.org/stable/modules/generated/sklearn.datasets.make_blobs.html

    data, labels = datasets.make_blobs(n_samples=param_dict['num_samples'],
                                       centers=param_dict['centers'],
                                       cluster_std=param_dict['cluster_std'],
                                       random_state=param_dict['random_seed']) #random_state=8
    labels_names = tuple(np.unique(labels)) # As there can be any number of classes, named 0, 1, 2...
    features_names = ('X', 'Y')

elif dataset_name in ('unbalanced_blobs', ):
    #https://scikit-learn.org/stable/modules/generated/sklearn.datasets.make_blobs.html

    data, labels = datasets.make_blobs([n_samples=param_dict['num_samples'],
                                       centers=param_dict['centers'],
                                       cluster_std=param_dict['cluster_std'],
                                       random_state=param_dict['random_seed']] #random_state=8
    labels_names = tuple(np.unique(labels)) # As there can be any number of classes, named 0, 1, 2...
    features_names = ('X', 'Y')

```

Figura 22: Implementación en el código de los conjuntos de datos sintéticos

Por otra parte, el conjunto de datos no gaussiano “Two Concentric Squares” se encuentra en un fichero CSV, por lo que en este caso lo que se hace es cargar los datos y entrenarlos, sin necesidad de crearlos.

3.1.3. Resultados en conjuntos de datos sintéticos

Se realiza el estudio paramétrico de la red SuperRGNG con todos los conjuntos descritos en la sección anterior. En este estudio se determinan los intervalos de valores óptimos para los hiperparámetros más destacados de la SuperRGNG para alcanzar rendimientos de la red de máximo valor, que vemos en la Tabla 8. Los conjuntos de datos 'Two Concentric Squares', 'unbalanced_blobs' y 'blobs' han sido evaluados en términos de su impacto en la red neuronal. En el caso de 'Two Concentric Squares' y 'unbalanced_blobs' los resultados se han mostrado estables sin problemas significativos, mientras que el conjunto 'blobs', con su superposición de clases, ha proporcionado valiosa información sobre el comportamiento de la red.

Los parámetros que se analizan de la red SuperRGNG son: “epochs”, “winner_step”, “neighbor_step”, “max_nodes”, “n_iter_before_neuron_added”, “max_edge_age”, “initial_n_epochs_w_m_i_b” y “n_iter_before_m_i_b_a_i_n_e” [67].

Parámetros	Valores Óptimos	Observaciones de Comportamiento
epochs	(10, 1000)	Un valor muy alto implica tiempos de entrenamiento mayores y no necesariamente mejor rendimiento.
winner_step (e_b)	(0.1, 0.9)	Aumentar para incrementar el número de conexiones. Pero se debe notar e_n.
neighbor_step (e_n)	[0.002, 0.4]	Otros autores: 0.06, 0.012 Otros autores: preferiblemente menor que e_b.
max_nodes	Cuanto más pequeño, mejor. PERO no tan pequeño que los datos de entrada no estén bien representados (valores $\leq 1\%$ del tamaño de los datos de entrada son desaconejados fuertemente).	Cuando las clases están muy cerca, reduce el número de neuronas.
n_iter_before_neuron_added. (λ)	[25, 700]	Generalmente, es difícil ajustarlo, por lo que se recomienda probar varios valores.
max_edge_age (a_max)	[25, 700]	Generalmente, es difícil ajustarlo, por lo que se recomienda probar varios valores.
initial_n_epochs_w_m_i_b	Mejor no por encima del número de épocas. Generalmente, los valores cercanos a las épocas son mejores.	Valores por debajo de 100 aumentaron el tiempo de entrenamiento.
n_iter_before_m_i_b_a_i_n_e	Valores cercanos a la longitud de los datos de entrenamiento, pero mejor no por encima de ella.	Valores por debajo de 100 aumentaron el tiempo de entrenamiento.

Tabla 8: Resultado del estudio paramétrico de la SuperGNG con datos sintéticos

En la Figura 23 podemos ver el comportamiento de la red en Blobs con una superposición del 62%, donde se observa la alta capacidad de clasificación que muestra la SuperGNG, aun trabajando en entornos de información con un alto grado de solapamiento entre sus datos. Para ser un conjunto de datos con una complejidad elevada, la red muestra una eficiencia altamente significativa a la hora de clasificar.

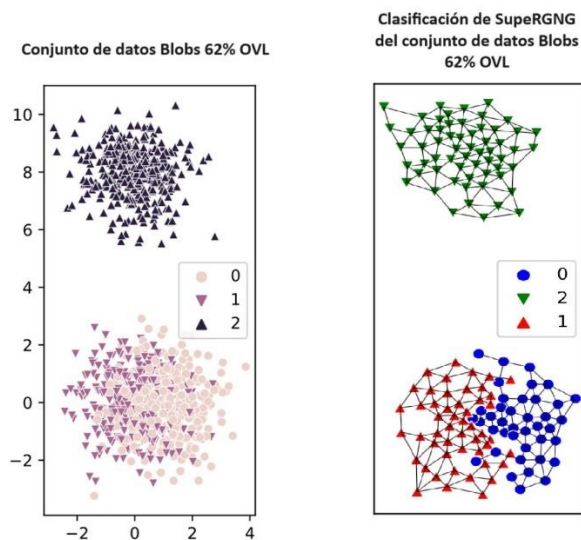


Figura 23: A la izquierda el conjunto de datos blobs 62% OVL y a la derecha los resultados de clasificación de la red SuperGNG

En base a lo reflejado en la Tabla 8, los cuatro hiperparámetros de interés para el estudio son los siguientes:

- "initial_n_epochs_w_m_i_b": indica la época inicial para establecer los límites entre clases.
- "n_iter_before_m_i_b_a_i_n_e": Indica con qué frecuencia los límites entre clases se establecerán después de las épocas iniciales (indicado por la propiedad anterior).
- "max_nodes": Indica el máximo número de nodos de la red.
- "n_iter_before_neuron_added": Indica cada cuantas iteraciones se añaden nuevas neuronas.

3.2 Evaluación y Pre-procesado de los Conjuntos de Datos

3.2.1. Estudio y análisis de los conjuntos de datos

Al determinar los criterios diagnósticos que formarán parte de nuestro conjunto de datos se deben seleccionar aquellos que demuestren un mejor desempeño en el diagnóstico temprano, para lo cual se ha considerado la opinión de expertos médicos, neurólogos y geriatras, así como un estudio de ranqueo de características (criterios diagnóstico-realizados con distintos métodos como Extreme Gradient Boosting (XGBOOST) y boosting (entre otras, Adaptive Boosting (AdaBoost)) [68].

Después de filtrar y limpiar los datos de ADNI, obtenemos los siguientes resultados que podemos observar en la Tabla 9.

	ADNI 1		ADNI 2		ADNI 3	
	Criterios diagnósticos	Sujetos	Criterios diagnósticos	Sujetos	Criterios diagnósticos	Sujetos
CN	136	229	194	188	195	83
AD	136	188	194	150	195	17

Tabla 9: Características de cada fase después de preprocesar los datos

Esta diferencia de valores respecto a lo comentado anteriormente se debe a falta de valores en criterios diagnósticos importantes de los sujetos.

Las clases que integrarán el conjunto de datos objetivo del estudio serán la EA y CN.

Para el estudio que se realiza en este trabajo, que está centrado en la ayuda al diagnóstico de la EA, es decir a resolver el problema de clasificación, se incluirán las clases CN - EA, y se utilizará la base de datos ADNI 1. Esta decisión se debe a que en esta fase de ADNI después de preprocesar los datos, tiene el mayor número de CN (229) y de EA (188).

Para este estudio, se utilizarán dos conjuntos de datos distintos con criterios diagnósticos de ADNI 1. El primero estará compuesto por criterios diagnósticos no invasivos para los sujetos. El segundo conjunto incluye características que implican pruebas invasivas para la persona. Esto significa que puedan resultar incómodas o dolorosas, requiriendo, en la mayoría de los casos, una infraestructura médica compleja, específica y de alto coste. Esto hace que solo se pueda encontrar en centros especializados y/u hospitales, aunque tiene la ventaja de que algunas proporcionan biomarcadores más sensibles que las no invasivas.

En la Tabla 10 podemos observar los características que incluye el conjunto de datos sin pruebas invasivas, junto a la definición de cada una, así como el rango de valores que poseen y a qué tipo de pruebas pertenecen.

Termino	Definición	Código	Case Report Form (CRF)
AGE	Edad inicial	-	Tablas clave de ADNI fusionadas en una sola tabla
FAQREM	Recordar citas, ocasiones familiares, vacaciones, medicamentos.	0=Normal 1=Nunca lo hice, pero podría hacerlo ahora	Cuestionario de evaluación funcional
FAQFORM	Reunir registros de impuestos, asuntos comerciales u otros documentos.	2=Nunca lo hice, ahora tendría dificultades	
FAQTRAVL	Viajar fuera del vecindario, conducir o hacer arreglos para tomar el transporte público.	3=Tiene dificultades, pero lo hace solo 4=Requiere ayuda 5=Dependiente	
MMSCORE	Puntuación total MMSE	0..30	Mini examen del estado mental
MMFLAGDL	20. Marcador (alt: Penny)	1=Correcto; 2=Incorrecto	

Tabla 10: Definiciones y códigos de los criterios diagnósticos que conforman del conjunto de datos sin pruebas invasivas [23].

Por otro lado, tenemos el conjunto de datos con pruebas invasivas, más precisamente las pruebas invasivas o de alto coste que se incluyen son “ABETA” y “VENTRICLES”. Las definiciones y otras características de este conjunto se encuentran en la Tabla 11.

Termino	Definición	Código	Case Report Form (CRF)
AGE	Edad inicial	-	Tablas clave de ADNI fusionadas en una sola tabla
MMSCORE	Puntuación total MMSE	0..30	Mini examen del estado mental
FAQTOTAL	Puntuación total	0..30	Cuestionario de evaluación funcional
ABETA	PTAU/ABETA	-	Datos de biomarcadores longitudinales de UPENN (4 años)
VENTRICLES	Ventrículos UCSF	-	Tablas clave de ADNI fusionadas en una sola tabla

Tabla 11: Definiciones y códigos de los criterios diagnósticos que conforman el conjunto de datos con pruebas invasivas [23].

3.2.2. Estudio estadístico de los datos.

Una vez desglosados los dos conjuntos de datos que se utilizaran para este trabajo, realizaremos un estudio estadístico de los mismos.

En la Tabla 12 observamos los valores estadísticos del conjunto de datos sin pruebas invasivas para ambas clases.

Crterios Diagnósticos	Estadísticos	CN (229)	EA (188)
AGE	Mean (SD)	75,87 (5,02)	75,26 (7,53)
	Median	75,6	75,65
	Range	59,9-89,6	55,1-90,9
	P-value	0,369	
FAQREM	Mean (SD)	0.12 (0.61)	3.65 (1.36)
	Median	0	4
	Range	0-4	0-5
	P-value	5.01E-86	
FAQFORM	Mean (SD)	0.12 (0.53)	3.65 (1.57)
	Median	0	4
	Range	0-5	0-5
	P-value	2.5E-78	
FAQTRAVL	Mean (SD)	0.03 (0.28)	3.16 (1.89)
	Median	0	4
	Range	0-3	0-5
	P-value	6.33E-69	
MMSCORE	Mean (SD)	29.11 (0.1)	23.28 (2.04)
	Median	29	23
	Range	25-30	18-27
	P-value	4.98E-72	
MMFLAGDL	Mean (SD)	1.14 (0.34)	1.84 (0.36)
	Median	1	2
	Range	1-2	1-2
	P-value	1.04E-51	

Tabla 12: Estadísticos del conjunto de datos sin pruebas invasivas

El p-valor que se ha obtenido ha sido con un Alpha de 0.05. Los resultados que hemos visto en la Tabla 12, todas los criterios diagnósticos son estadísticamente a excepción de la edad , lo cual es lógico, puesto que esta no es un criterio diagnóstico, sino un posible factor de riesgo. Se da una representación gráfica para mostrar las diferencias entre los distintos p-valores en la Figura 24.

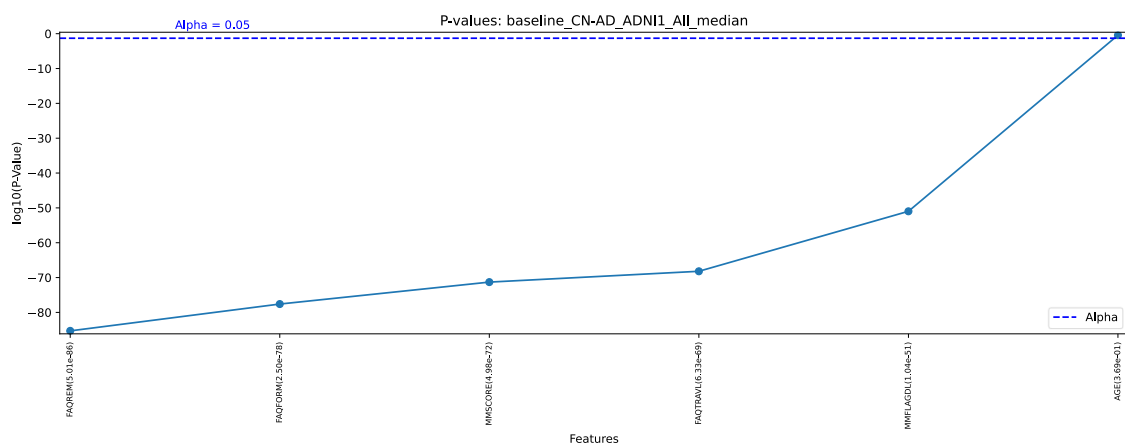


Figura 24: Gráfico de los p-valores de las características conjunto de datos sin pruebas invasivas

Por otra parte, en la Tabla 13 presentamos los datos estadísticos del conjunto de datos con pruebas invasivas. Como era de esperar, la edad sigue sin ser significativa, debido a que los datos de esta característica son los mismo que el anterior conjunto de datos.

Crterios Diagnósticos	Estadísticos	CN (229)	EA (188)
AGE	Mean (SD)	75,87 (5,02)	75,26 (7,53)
	Median	75,6	75,65
	Range	59,9-89,6	55,1-90,9
	P-value	0,369	
MMSCORE	Mean (SD)	29,11 (0,1)	23,28 (2,04)
	Median	29	23
	Range	25-29	18-27
	P-value	4,98E-72	
FAQTOTAL	Mean (SD)	0,14 (0,6)	13,14 (6,84)
	Median	0	12
	Range	0-6	0-30
	P-value	4,34E-89	
ABETA	Mean (SD)	1234,71 (454,74)	607,64 (275,11)
	Median	1232,5	552,5
	Range	200-3592	212,3-2568
	P-value	7,03E-48	
VENTRICLES	Mean (SD)	35785,35 (20332,6)	51290,72 (26784,4)
	Median	32271	45391
	Range	5834-118875	9166-147064
	P-value	2,87E-09	

Tabla 13: Estadísticos del conjunto de datos con pruebas invasivas

Se muestra también gráficamente los p-valores obtenidos en la Figura 25. Al igual que la anterior prueba, el Alpha tiene un valor de 0,05.

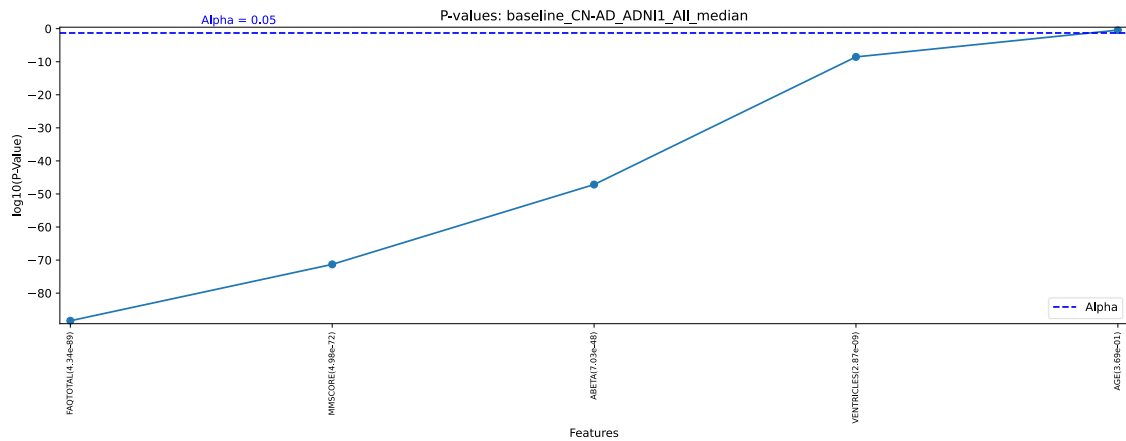


Figura 25: Gráfico de los p-valores del conjunto de datos con pruebas invasivas

Como observamos en la Figura 24 y la Figura 25, los valores estadísticos de las características de ambos conjuntos de datos son generalmente muy positivos. Como se ha comentado anteriormente, las personas con edad avanzada son más propensas a tener Alzheimer.

Es por este motivo que se incluye aun no siendo estadísticamente significativa. Aunque la estadística añade más información de valor a los datos, hay que tener en cuenta que el p-valor es una frecuencia relativa que indica si es cierta una determinada hipótesis [69]. Pero como sabemos, aceptar una hipótesis como posible no implica que esta sea cierta [69].

3.2.3. Pre-procesado de los datos según los coeficientes de clustering

Para alcanzar la mejor representación de la información que va a ser tratada por los métodos de clasificación para la detección de la EA, se debe realizar un buen y adecuado preprocesado del espacio de entrada. Usaremos las siguientes tres métricas de clustering, Silhouette, DB y CH, que fueron analizadas y usadas en el estudio paramétrico de la SuperGNG en el capítulo anterior. Así, seleccionaremos la combinación óptima de los procesos de escalado, proyectado y componentes, siguiendo la metodología utilizada en el estudio de calidad de clústeres en [70].

Al llevar a cabo un estudio comparativo, estos coeficientes sabemos que proporcionan una guía sólida para la toma de decisiones. Dado que cada métrica tiene sus propias maneras de realizar el estudio de calidad, esta combinación permitirá que los coeficientes se complementen mutuamente, brindando una evaluación más completa y robusta del rendimiento del *clustering*.

La metodología empleada consiste en seleccionar una combinación específica de técnicas de data mining para el preprocesado de los datos (escalado, proyección y componentes) que presentan el mejor comportamiento en los conjuntos de datos utilizados. Este enfoque es necesario para obtener una combinación unificada que permita comparar de manera efectiva el desempeño de diferentes clasificadores. Identificamos las cuatro mejores combinaciones para cada métrica y conjunto de datos, y calculamos la frecuencia de aparición de cada combinación en el TOP 4 obtenido.

Es por ello por lo que se utilizarán las 3 combinaciones más usuales. De esta manera podremos evaluar si los coeficientes nos aportan valor al estudio y además realizar una comparación entre las combinaciones que se obtengan como más comunes.

Los resultados de los coeficientes se muestran en la Tabla 14.

	Con Edad						Sin Edad					
	Conjunto de datos Invasivo			Conjunto de datos No Invasivo			Conjunto de datos Invasivo			Conjunto de datos No Invasivo		
	Escalado	Proyectado	Dimensi ón	Escalado	Proyectado	Dimensi ón	Escalado	Proyectado	Dimensi ón	Escalado	Proyectado	Dimensi ón
Shilouette	PowerTransformer_Yeo-Johnson	NeighborhoodCA	2	PowerTransformer_Yeo-Johnson	NeighborhoodCA	2	PowerTransformer_Yeo-Johnson	NeighborhoodCA	3	PowerTransformer_Yeo-Johnson	FactorAnalysis	3
	PowerTransformer_Yeo-Johnson	NeighborhoodCA	3	StandardScaler	NeighborhoodCA	2	PowerTransformer_Yeo-Johnson	NeighborhoodCA	2	PowerTransformer_Yeo-Johnson	FactorAnalysis	2
	RobustScaler	NeighborhoodCA	2	MaxAbsScaler	NeighborhoodCA	2	MaxAbsScaler	NeighborhoodCA	2	MaxAbsScaler	NeighborhoodCA	3
	StandardScaler	NeighborhoodCA	3	MinMaxScaler	NeighborhoodCA	2	RobustScaler	NeighborhoodCA	2	MaxAbsScaler	NeighborhoodCA	2
DB	PowerTransformer_Yeo-Johnson	Unprojected	2	PowerTransformer_Yeo-Johnson	NeighborhoodCA	2	PowerTransformer_Yeo-Johnson	Unprojected	2	PowerTransformer_Yeo-Johnson	FactorAnalysis	3
	MaxAbsScaler	TruncatedSVD	2	MinMaxScaler	NeighborhoodCA	2	Normalizer	TruncatedSVD	2	MaxAbsScaler	NeighborhoodCA	2
	RobustScaler	NeighborhoodCA	2	MaxAbsScaler	NeighborhoodCA	2	PowerTransformer_Yeo-Johnson	NeighborhoodCA	3	MinMaxScaler	NeighborhoodCA	2
	StandardScaler	NeighborhoodCA	3	StandardScaler	NeighborhoodCA	2	RobustScaler	NeighborhoodCA	2	MaxAbsScaler	NeighborhoodCA	3
CH	PowerTransformer_Yeo-Johnson	NeighborhoodCA	2	PowerTransformer_Yeo-Johnson	NeighborhoodCA	2	PowerTransformer_Yeo-Johnson	NeighborhoodCA	3	PowerTransformer_Yeo-Johnson	FactorAnalysis	3
	RobustScaler	NeighborhoodCA	2	MaxAbsScaler	NeighborhoodCA	2	PowerTransformer_Yeo-Johnson	NeighborhoodCA	2	MaxAbsScaler	NeighborhoodCA	2
	PowerTransformer_Yeo-Johnson	NeighborhoodCA	3	MinMaxScaler	NeighborhoodCA	2	RobustScaler	NeighborhoodCA	2	MinMaxScaler	NeighborhoodCA	2
	StandardScaler	NeighborhoodCA	3	MinMaxScaler	NeighborhoodCA	3	MaxAbsScaler	NeighborhoodCA	2	MaxAbsScaler	NeighborhoodCA	3

Tabla 14: TOP 4 mejores combinaciones de escalado, proyectado y componentes de cada métrica en los conjuntos de datos

Una vez analizados los resultados de los coeficientes, se muestra en la Tabla 15 las tres combinaciones más usuales, y las que serán utilizadas para el estudio.

Escalado	Proyectado	Dimensión	Apariciones
PowerTransformer_Yeo-Johnson	NeighborhoodCA	2	8
MaxAbsScaler	NeighborhoodCA	2	7
RobustScaler	NeighborhoodCA	2	6

Tabla 15: Combinación de escalado y proyectado más comunes en los conjuntos de datos, según los resultados de las coeficientes

Entre los 3 coeficientes para cada uno de los conjuntos de datos, tenemos que las combinaciones antes mencionadas aparecen en el 50% de las opciones posibles, lo que indica que pueden ser una buena representación para la comparación posterior.

La primera combinación de preprocesado, vemos en la Tabla 14 que tiene un mayor número de apariciones en los conjuntos de datos con edad, mientras que la segunda aparece más en los conjuntos de datos sin edad. La tercera combinación aparece de igual manera en los conjuntos de datos con y sin edad.

Estas tres combinaciones permiten realizar una comparación en 3 perspectivas, lo cual impide que los resultados que obtengamos beneficien a un solo conjunto de datos.

Hay que destacar que el proyectado Neighborhood CA aparece con mucha frecuencia junto a distintos escalados, lo que parece indicar que este proyectado obtiene buenos resultados a la hora de proyectar datos de nuestro espacio de entrada.

3.3 Estudio comparativo de la red neuronal SuperGNG con otros métodos de aprendizaje automático

Abordamos la detección de la EA resolviendo el problema de clasificación binaria CN-EA mediante la arquitectura neuronal SuperGNG, y demostraremos, no solo a través de su resultado, sino también a través de un estudio comparativo con un conjunto de métodos de AA al uso, tanto neuronales como no neuronales, la alta eficacia y el mejor desempeño del problema clínico real planteado, por la solución computacional propuesta, la red neuronal ontogénica SuperGNG. Comenzaremos este desarrollo con el particionado “Stratified K-Fold” con 5 “folds” en cada caso.

Para la obtención de las métricas se hace uso de la función “calculateMetrics()” cuya implementación se puede ver en la Figura 26. La implementación realizada contempla la posibilidad del cálculo de métricas de rendimiento del clasificador no solo para clasificaciones binarias, sino ternarias o n-arias. El AUC se obtiene con la función “roc_curve” de la biblioteca sklearn, mientras que, si son 3 clases, se obtienen los distintos AUC de cada clase por separado y se utiliza la dinámica “micro” para obtener el nuevo valor de AUC.

```

def calculateMetrics(y_test, y_test_pred, num_clases):
    if num_clases == 2:
        fpr, tpr, thresholds = metrics.roc_curve(y_test, y_test_pred)
        AUC = metrics.auc(fpr, tpr)
        tn, fp, fn, tp = metrics.confusion_matrix(np rint(y_test), np rint(y_test_pred)).ravel()
    else:
        fpr_macro, tpr_macro, roc_auc_macro, fpr_micro, tpr_micro, roc_auc_micro = calculate_macro_micro_roc(y_test, y_test_pred, 3)
        fpr = fpr_micro
        tpr = tpr_micro

        # Calcular el AUC promedio
        AUC = roc_auc_micro

        conf_matrix = metrics.confusion_matrix(np rint(y_test), np rint(y_test_pred))

        tp = [conf_matrix[i, i] for i in range(len(conf_matrix))]
        fp = [sum(conf_matrix[:, i]) - conf_matrix[i, i] for i in range(len(conf_matrix))]
        fn = [sum(conf_matrix[i, :]) - conf_matrix[i, i] for i in range(len(conf_matrix))]

        tn = []
        for i in range(len(conf_matrix)):
            correct_predictions_other_classes = sum(sum(conf_matrix)) - sum(conf_matrix[:, i]) - sum(conf_matrix[i, :]) + conf_matrix[i, i]
            tn.append(correct_predictions_other_classes)

        tn = sum(tn)
        tp = sum(tp)
        fp = sum(fp)
        fn = sum(fn)

        ACCURACY = (tp + tn) / (tp + tn + fp + fn)
        PRECISION = tp / (tp + fp)
        RECALL = tp / (tp + fn)
        SPECIFICITY = tn / (tn + fp)
        ppv_cuiPositive = tp / (tp + fp)
        CUIPOSITIVE = PRECISION * ppv_cuiPositive
        npv_cuiNegative = tn / (tn + fn)
        CUINEGATIVE = SPECIFICITY * npv_cuiNegative

    return fpr, tpr, AUC, ACCURACY, PRECISION, RECALL, SPECIFICITY, CUIPOSITIVE, CUINEGATIVE

```

Figura 26: Cálculo de métricas de rendimiento de los clasificadores

Una vez las se calculan las métricas, se hace un promedio de los 5 K-Folds, obteniendo así los resultados de los clasificadores en el estudio.

Finalmente se calculan las métricas adicionales, incluyendo la métrica “Clinical Utility Index” (CUI). Esta métrica de eficacia clínica real permite determinar que el grado de medicina traslacional es útil en la práctica clínica, lo cual da un alto valor añadido al producto desarrollado [71].

3.3.1. Elección de parámetros de los clasificadores

La optimización de la arquitectura SuperGNG para llevar a cabo la clasificación CN-EA se hace conjuntos de datos sintéticos, utilizando una configuración que tuvo buenos resultados en el conjunto de datos “blobs” (que llamaremos semilla), adaptado a los conjuntos de datos ADNI, debido a que pasamos de 999 muestra a 417.

Una vez evaluada la configuración, se realizaron pruebas sobre los conjuntos de datos donde se obtuvieron resultados de AUC superiores a 0.99 para todos los conjuntos de datos estudiados, con pruebas invasivas y no invasivas, considerando la edad como una característica más del espacio de entrada, y excluyendo la misma, como tal, lo cual nos llevó al análisis de la edad como factor de riesgo para la EA. Pero debido a que nos encontramos con clases que tienen en algunos

casos distribuciones distintas a las estudiadas con los datos sintéticos, se han realizado 50 configuraciones con los parámetros más significativos de SuperGNG.

En la Tabla 16 podemos ver las 5 configuraciones que obtuvieron mejores resultados, donde se muestran los parámetros de cada una de ellas.

Config	n_iter_before_neuro_n_added	max_edge_age	max_nodos	initial_n_epochs_w_m_i_b	n_iter_before_m_i_b_a_i_n_e
7	150	200	45	300	300
38	250	150	50	300	200
6	150	200	45	300	200
5	150	150	60	300	300
19	200	150	45	300	300

Tabla 16: Parámetros de las 5 mejores configuraciones de SuperGNG comparadas con la configuración semilla de "blobs"

Vemos que la mejor configuración reduce aún más los nodos utilizados, que es una de las características que se comentaron anteriormente.

Por otra parte, las configuraciones de los clasificadores que se han utilizado para la comparación son las que se presentan en la Tabla 17.

Clasificador	Parámetros de Configuración
MLP	<ul style="list-style-type: none"> - Capa oculta 1: 16 neuronas, activación 'relu' - Capa oculta 2: 8 neuronas, activación 'relu' - Capa de salida: 2 neuronas, activación 'softmax' - Optimizador: 'rmsprop' - Épocas: 100
RF	<ul style="list-style-type: none"> - Número de estimadores: 10 - Profundidad máxima: sin límite - Mínimo de muestras para dividir: 5
DT	<ul style="list-style-type: none"> - Mínimo de muestras por hoja: 2 - Mínimo de muestras para dividir: 5 - Profundidad máxima: sin límite
SVM	<ul style="list-style-type: none"> - Kernel: 'linear'
NB	-

Tabla 17: Parámetros de configuración de los clasificadores

3.3.2. Comparación entre clasificadores frente a conjuntos de datos no invasivos

Clasificación con AGE

Comenzamos con el conjunto de datos sin pruebas invasivas, añadiendo la edad como ayuda para la clasificación.

Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

En la Figura 28 podemos ver la relación que existe entre los distintos criterios pertenecientes al conjunto de datos y los componentes de la proyección. Se observa que el primer componente tiene relación inversa con MMSCORE y directa con los otros criterios. En cuanto al segundo componente, tiene una relación directa con FAQREM e indirecta con los otros criterios.

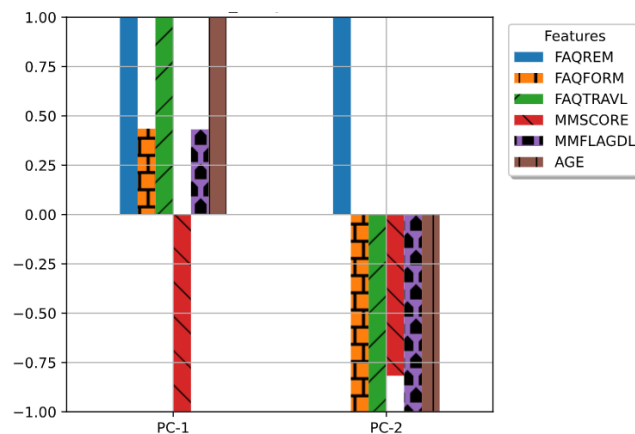


Figura 27: Relación entre componentes y las características del conjunto de datos usadas Power Transformer Yeo-Johnson - Neighborhood CA

La representación de los datos proyectados con este método de preprocesado es una muy buena representación de la información (Figura 28). Tiene las dos clases bastante bien separadas, aunque algunas muestras se encuentran bastante cerca de la clase opuesta.

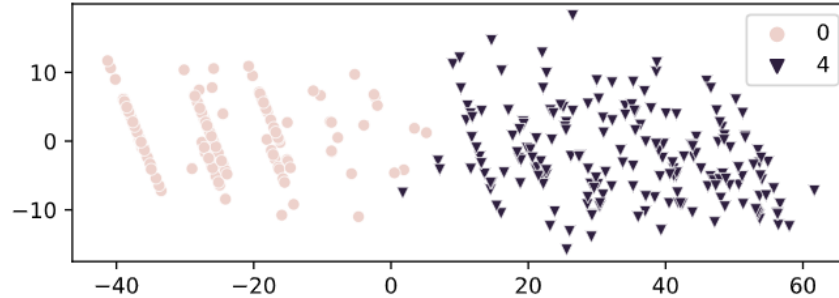


Figura 28: Representación de los datos mediante el preprocesado Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

Los resultados de los clasificadores han sido muy positivos, podemos ver esto tanto en el AUC (Figura 29) como las métricas de rendimiento adicionales (Tabla 18).

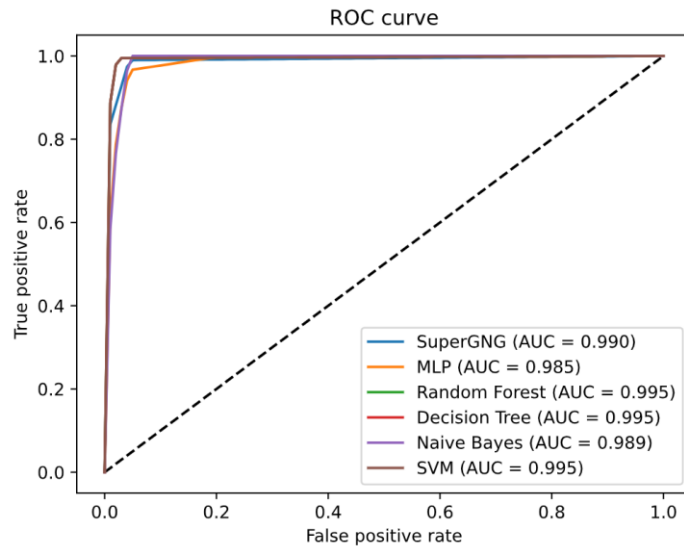


Figura 29: Curvas ROC de los clasificadores - Power Transformer Yeo-Johnson - Neighborhood CA

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SuperGNG	0,99	0,99	0,99	0,989	0,991	0,98	0,983
MLP	0,985	0,984	0,972	0,996	0,975	0,946	0,971
RF	0,995	0,995	0,995	0,995	0,996	0,99	0,991
DT	0,995	0,995	0,995	0,995	0,996	0,99	0,991
NB	0,989	0,988	0,975	1	0,978	0,951	0,978
SVM	0,995	0,995	0,995	0,995	0,996	0,99	0,991

Tabla 18: Métricas de Rendimiento Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

MaxAbsScaler - Neighborhood CA (PC1, PC2)

En la Figura 30 se observa que el primer componente se relaciona de manera inversa con los criterios MMFLAGDL y MMSCORE, mientras que mantiene una relación directa con los otros criterios. El segundo componente viene dado por una relación inversa de las pruebas MMSCORE y AGE, y directa con los otros criterios.

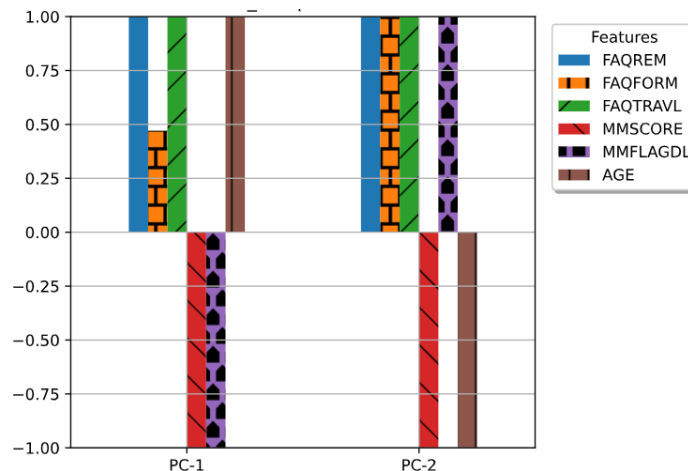


Figura 30: Relación entre componentes y las características del conjunto de datos usadas MaxAbsScaler - Neighborhood CA

Vemos además que los datos tienen una separación significativa entre las dos clases (Figura 31), con lo que estos métodos de preprocesado también proporcionan una representación de la información semánticamente relevante.

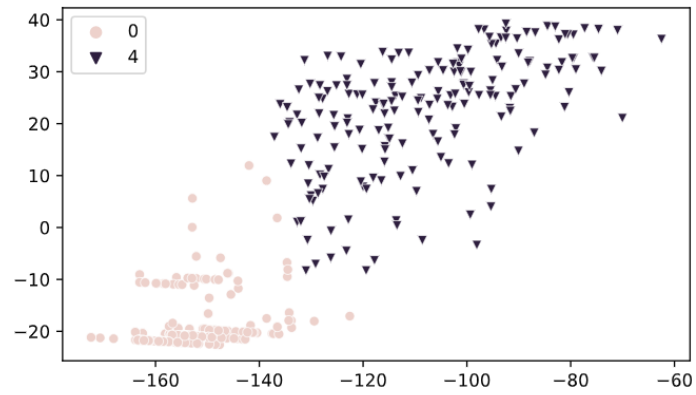


Figura 31: : Representación de los datos mediante el preprocesado MaxAbsScaler - Neighborhood CA (PC1, PC2)

Además, observamos que la curva ROC (Figura 32) y las métricas (Tabla 19) también han alcanzado valores muy elevados. MLP ha obtenido un rendimiento un poco menor comparado con los otros clasificadores.

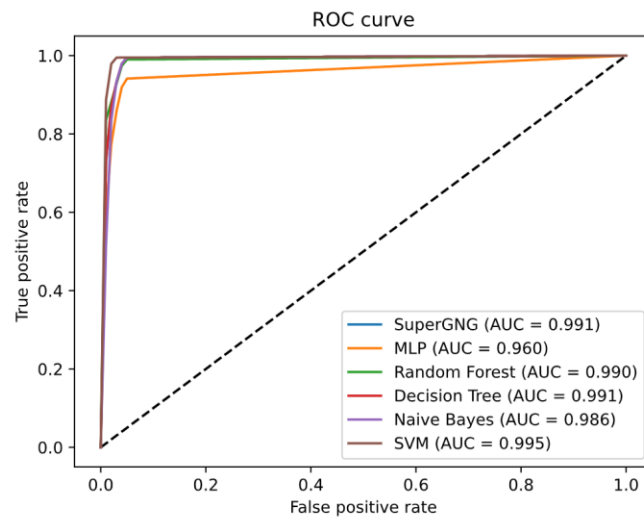


Figura 32: Curvas ROC de los clasificadores - MaxAbsScaler - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SuperGNG	0,991	0,99	0,985	0,995	0,987	0,97	0,983
MLP	0,96	0,963	0,977	0,939	0,982	0,955	0,935
RF	0,99	0,99	0,99	0,989	0,991	0,98	0,983
DT	0,991	0,99	0,984	0,995	0,987	0,97	0,983
NB	0,986	0,986	0,974	0,995	0,978	0,949	0,974
SVM	0,995	0,995	0,995	0,995	0,996	0,99	0,991

Tabla 19: Métricas de rendimiento ROC MaxAbsScaler - Neighborhood CA (PC1, PC2)

Robust Scaler - Neighborhood CA (PC1, PC2)

El último método de preprocesado estudiado ha sido Robust Scaler - Neighborhood CA. Vemos la relación entre criterios diagnósticos del conjunto de datos con la proyección Neighborhood CA en la Figura 33. El primer componente viene dado por una relación inversa con MMSCORE y directa con los criterios restantes. Por otra parte, el segundo componente se relaciona de manera inversa con MMFLAGDL, FAQTRAVL y FAQFORM (aunque esta última en menor medida), y el resto de los criterios con una relación directa.

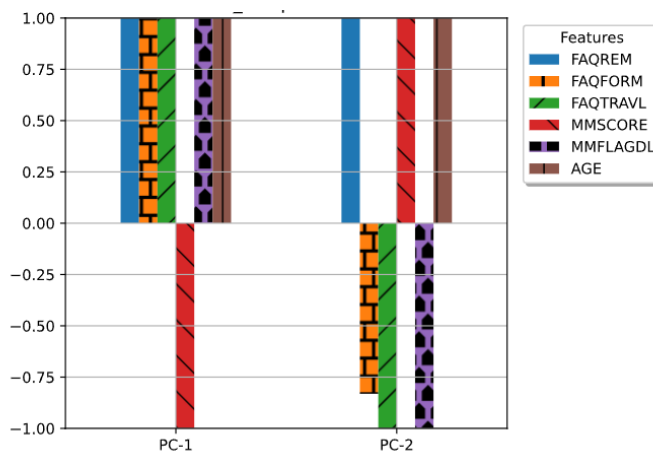


Figura 33: Relación entre componentes y las características del conjunto de datos usadas Robust Scaler - Neighborhood CA

La representación de los datos en este caso linealmente separable, Figura 34.

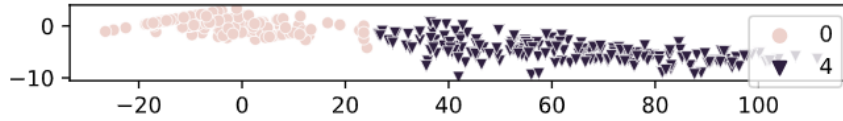


Figura 34: : Representación de los datos mediante el preprocesado Robust Scaler - Neighborhood CA (PC1, PC2)

Entre los resultados, se encuentran valores de AUC de 1 (Figura 35), de igual manera ocurre en las métricas de rendimiento adicionales (Tabla 20). Esto refuerza la alta calidad de la representación de la información.

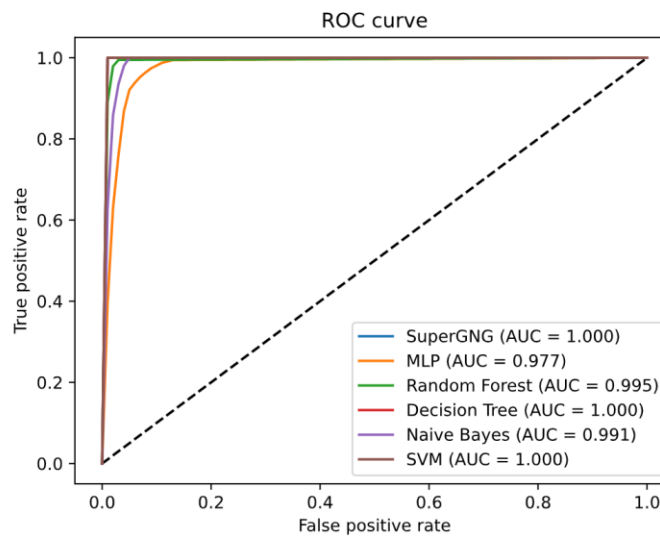


Figura 35: Curvas ROC de los clasificadores - Robust Scaler - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SuperGNG	1	1	1	1	1	1	1
MLP	0,977	0,976	0,955	0,994	0,961	0,914	0,956
RF	0,995	0,995	0,995	0,995	0,996	0,989	0,991
DT	1	1	1	1	1	1	1
NB	0,991	0,99	0,98	1	0,982	0,96	0,982
SVM	1	1	1	1	1	1	1

Tabla 20: Métricas de rendimiento Robust Scaler - Neighborhood CA (PC1, PC2)

Clasificación sin AGE

A continuación, repetiremos los experimentos de clasificación , pero excluyendo la edad del conjunto de datos.

Se realiza, a continuación, el estudio del rendimiento de los clasificadores para resolver la misma clasificación CN-EA con solo criterios diagnósticos, sin pruebas invasivas, pero sin considerar la característica AGE.

Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

A diferencia del apartado anterior, donde las relaciones entre las características eran de valores de 1 y -1, vemos en la Figura 36 valores inferiores. La relaciones en el primer componente vienen dadas por MMSCORE con una relación inversa y los otros criterios de manera directa, aunque con mayor aportación del criterio FAQREM. Por otra parte, el segundo componente tiene una relación directa con FAQREM, mientras MMFLAGDL tiene una relación inversa.

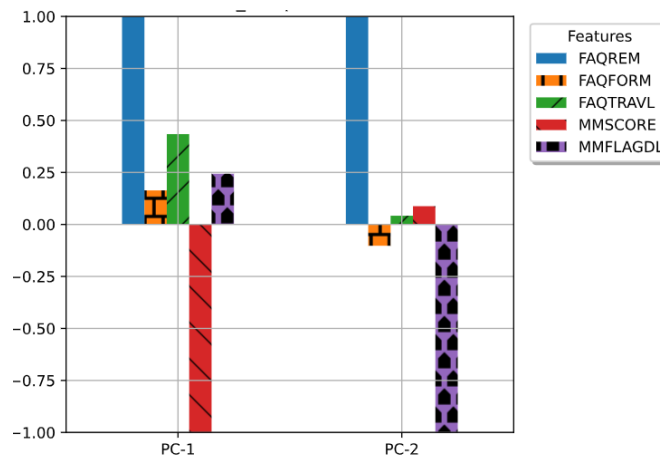


Figura 36: Relación entre componentes y las características del conjunto de datos usadas Power Transformer Yeo-Johnson - Neighborhood CA

Aunque vemos algunas muestras mezcladas (superposición), la separación entre las clases sigue siendo significativa. Podemos verlo en la Figura 37.

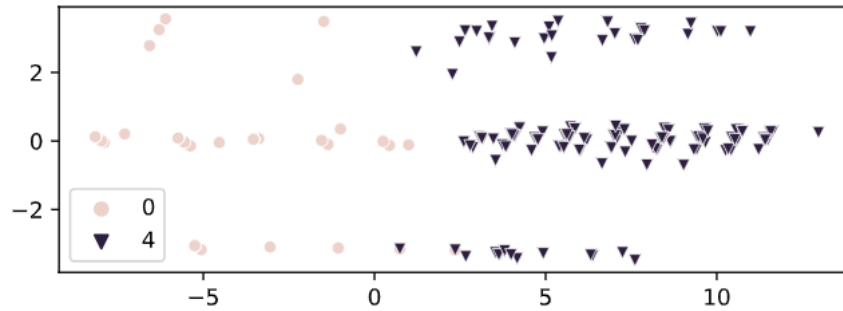


Figura 37: : Representación de los datos mediante el preprocesado Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

En este caso, como vimos un conjunto de datos con bastante separación de clases, observamos que los clasificadores obtienen valores muy elevados tanto en AUC (Figura 38) como en las métricas de rendimiento que se calcularon (Tabla 21).

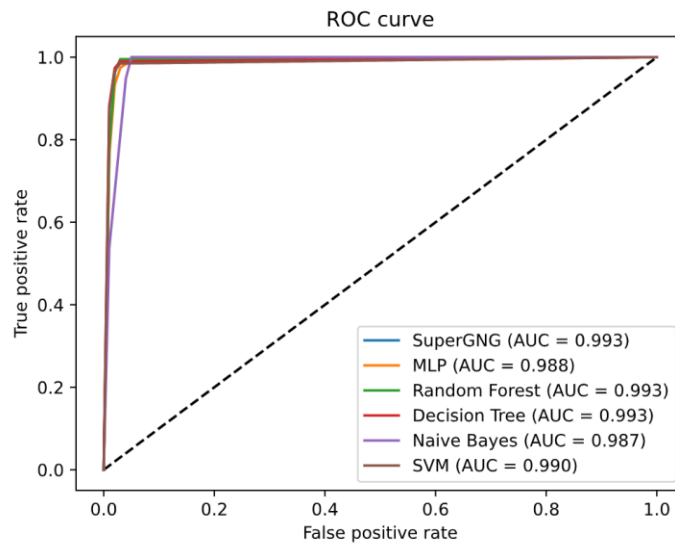


Figura 38: Curvas ROC de los clasificadores - Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SuperGNG	0,993	0,993	0,99	0,995	0,991	0,979	0,987
MLP	0,988	0,989	0,989	0,986	0,99	0,978	0,979
RF	0,993	0,993	0,99	0,995	0,991	0,979	0,987
DT	0,993	0,993	0,995	0,989	0,996	0,99	0,987
NB	0,987	0,986	0,97	1	0,974	0,941	0,974
SVM	0,99	0,99	0,995	0,984	0,996	0,99	0,983

Tabla 21: Métricas de Rendimiento Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

MaxAbsScaler - Neighborhood CA (Dim 2) (PC1, PC2)

Con este pre-procesado, la asociación entre las componentes proyectadas y las características del espacio de entrada, tiene una asociación significativa, con relación directa para el PC1 y las características FAQREM, FAQFORM y FAQTRAVL, mientras que los otros criterios tiene una relación inversa (Figura 39). En cuanto al segundo componente, la relación directa viene dada por FAQREM y MMFLAGDL, y la inversa por FAQFORM, MMSCORE y en menor medida por FAQTRAVL.

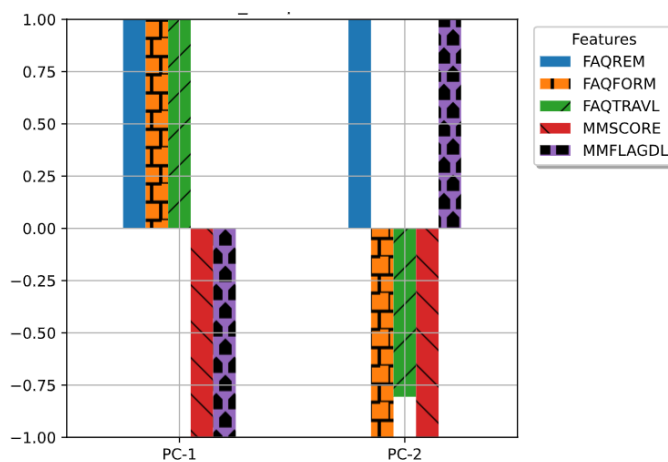


Figura 39: Relación entre componentes y las características del conjunto de datos usadas MaxAbsScaler - Neighborhood CA

Por otra parte, la representación de la información no es tan pura, existiendo algún pequeño solape entre las clases, Figura 40.

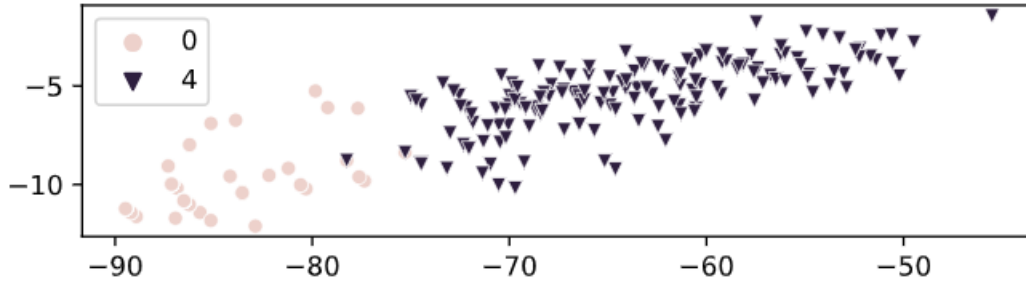


Figura 40: : Representación de los datos mediante el preprocesado MaxAbsScaler - Neighborhood CA (PC1, PC2)

Como era de esperar, los resultados de los clasificadores se han visto perjudicado por el solapamiento que comentamos anteriormente. Sobretudo podemos observarlo en el clasificador MLP, tanto en la Figura 41 y en la Tabla 22.

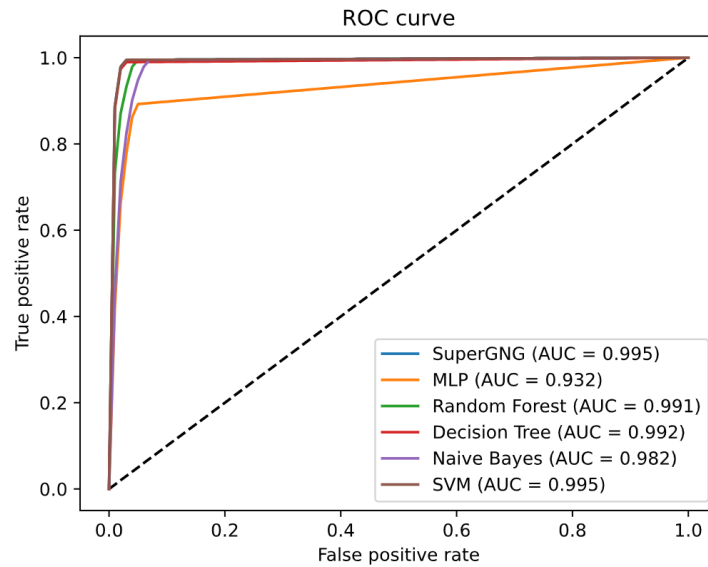


Figura 41: Curvas ROC de los clasificadores - MaxAbsScaler - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SuperGNG	0,995	0,995	0,995	0,995	0,996	0,99	0,991
MLP	0,932	0,936	0,966	0,89	0,974	0,933	0,892
RF	0,991	0,99	0,985	0,995	0,987	0,97	0,983
DT	0,992	0,993	0,995	0,989	0,996	0,99	0,987
NB	0,982	0,981	0,964	0,995	0,969	0,931	0,965
SVM	0,995	0,995	0,995	0,995	0,996	0,99	0,991

Tabla 22: Métricas de Rendimiento MaxAbsScaler - Neighborhood CA (PC1, PC2)

Robust Scaler - Neighborhood CA (Dim 2) (PC1, PC2)

Continuamos con la combinación Robust Scaler - Neighborhood CA, con relaciones entre componentes menos potentes en PC-2, siendo la relación con el FAQREM la realmente significativa (Figura 42). EL primer componente viene dado por una relación inversa con el criterio MMSCORE y directa con los otros criterios.

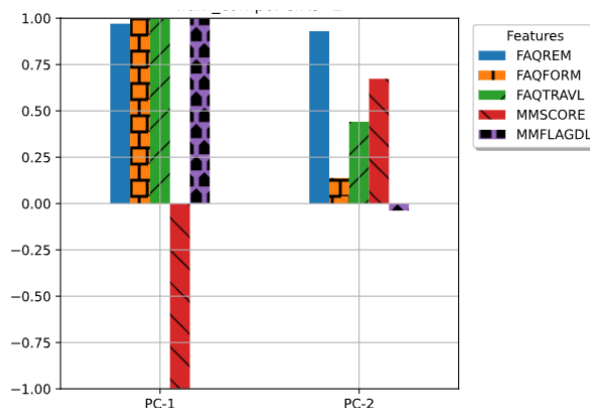


Figura 42: Relación entre componentes y las características del conjunto de datos usadas Robust Scaler - Neighborhood CA

En esta combinación de preprocesado la separación entre clases es bastante significativa, pero no completa, debido a un porcentaje pequeño de muestras que se superponen, Figura 43.



Figura 43: : Representación de los datos mediante el preprocesado Robust Scaler - Neighborhood CA (PC1, PC2)

En cualquier caso, el número de muestras superpuestas es inferior, lo que beneficia la clasificación como vemos en la Figura 44 y la Tabla 23.

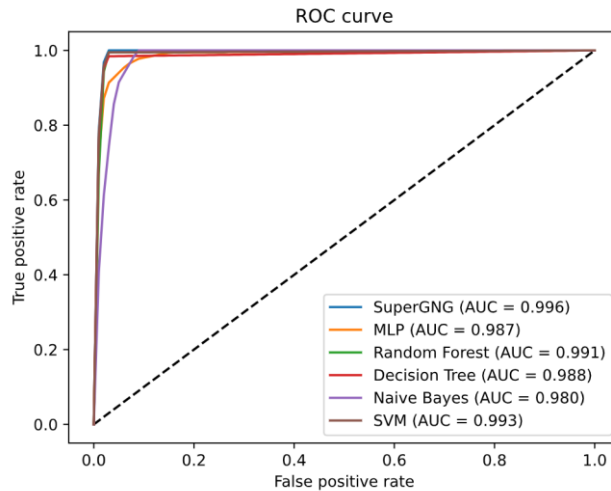


Figura 44: Curvas ROC de los clasificadores - Robust Scaler - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SuperGNG	0,996	0,995	0,99	1	0,991	0,979	0,991
MLP	0,987	0,987	0,979	0,994	0,981	0,959	0,976
RF	0,991	0,99	0,984	0,995	0,987	0,969	0,983
DT	0,988	0,988	0,99	0,984	0,991	0,979	0,979
NB	0,98	0,978	0,955	1	0,961	0,913	0,961
SVM	0,993	0,993	0,99	0,995	0,991	0,98	0,987

Tabla 23: Métricas de Rendimiento Robust Scaler - Neighborhood CA (PC1, PC2)

3.3.3. Comparación de clasificadores frente a conjunto de datos con pruebas invasivas

En este apartado se llevará a cabo un estudio comparativo del desempeño de clasificación con los conjuntos de datos con pruebas invasivas. Para ello se usará la misma metodología que se utilizó con el conjunto de datos con pruebas no invasivas.

Las configuraciones de los clasificadores que se han utilizado para la clasificación son las que se presentan en la Tabla 16 y la Tabla 17.

Clasificación con AGE

Al igual que hicimos en el apartado anterior, empezaremos añadiendo la edad al conjunto de características usadas en este análisis comparativo.

Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

Regresamos a las relaciones entre características con valores extremos 1 y -1. Se puede observar este hecho en la Figura 45. El primer y segundo componente vienen dado por una relación directa por VENTRICLES, FAQTOTAL Y AGE, e inversa por MMSCORE.

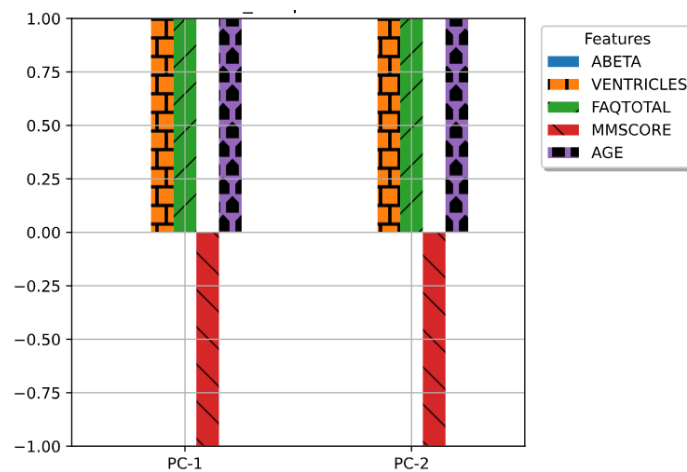


Figura 45: Relación entre componentes y las características del conjunto de datos usadas Power Transformer Yeo-Johnson - Neighborhood CA

La separación de las clases en la representación de información obtenida es altamente favorable para la clasificación a realizar (Figura 46).

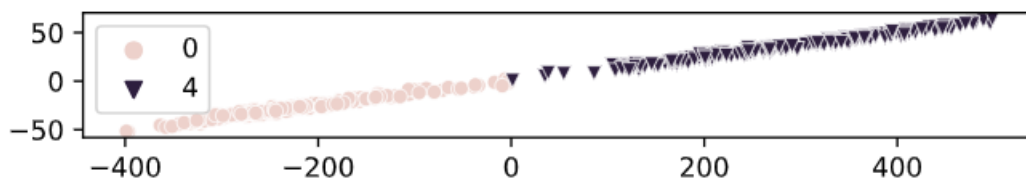


Figura 46: : Representación de los datos mediante el preprocesado Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

Los resultados en todos los casos de los clasificadores son valores muy próximos a 1, lo que indica que efectivamente esta combinación es adecuada para representar los datos. Estos resultados se presentan en las curvas ROC (Figura 47) y las métricas adicionales (Tabla 24).

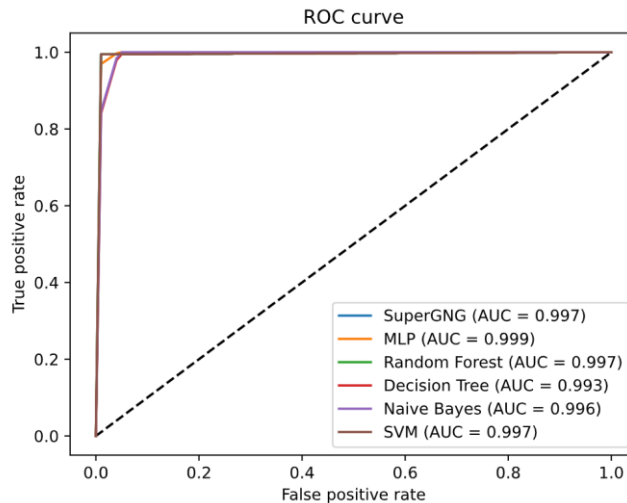


Figura 47: Curvas ROC de los clasificadores - Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SuperGNG	0,997	0,998	1	0,995	1	1	0,996
MLP	0,999	0,999	0,998	1	0,998	0,996	0,998
RF	0,997	0,998	1	0,995	1	1	0,996
DT	0,993	0,993	0,99	0,995	0,991	0,98	0,987
NB	0,996	0,995	0,99	1	0,991	0,98	0,991
SVM	0,997	0,998	1	0,995	1	1	0,996

Tabla 24: Métricas de Rendimiento Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

MaxAbsScaler - Neighborhood CA (PC1, PC2)

Las asociaciones de los componentes proyectados y los criterios se presentan en la Figura 48. El primer componente está asociado a una relación directa con VENTRICLES, FAQTOTAL Y AGE, mientras que los criterios restantes de manera inversa. El segundo componente se asocia por una relación directa con MMSCORE y ABETA, y el resto de los criterios de manera inversa.

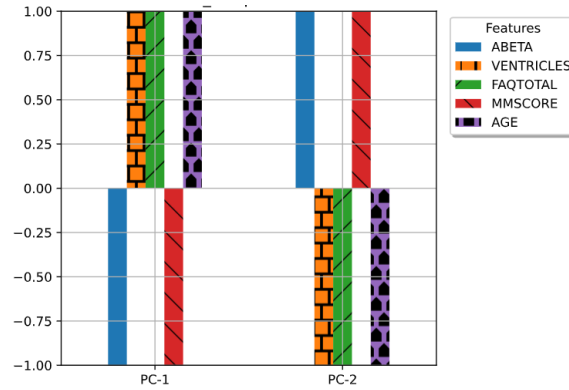


Figura 48: Relación entre componentes y las características del conjunto de datos usadas MaxAbsScaler - Neighborhood CA

En este caso tenemos la primera representación de la información donde las dos clases presentes en el conjunto de datos está perfectamente separada, lo que pronostica una clasificación altamente precisa, Figura 49.

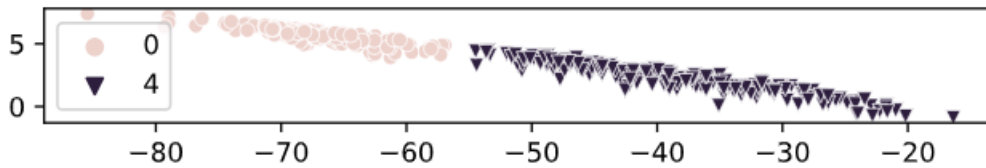


Figura 49: : Representación de los datos mediante el preprocesado MaxAbsScaler - Neighborhood CA (PC1, PC2)

Los resultados de clasificación, según las métricas de rendimiento usadas, de todos los clasificadores son de los mejores que se han obtenido , hasta el momento, en este estudio (Figura 50)(Tabla 25).

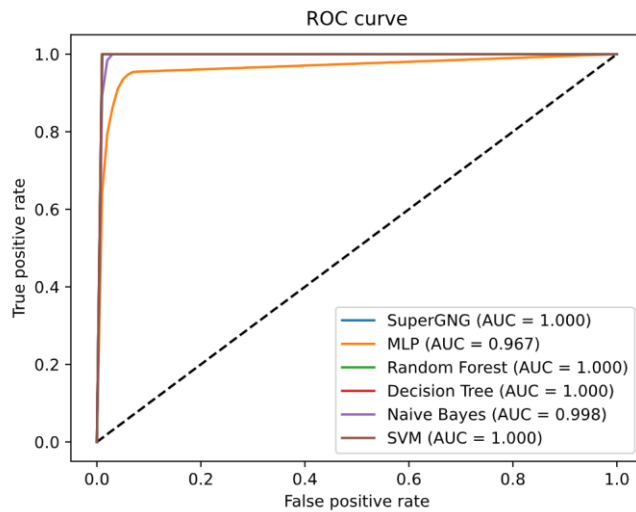


Figura 50: Curvas ROC de los clasificadores - MaxAbsScaler - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SuperGNG	1	1	1	1	1	1	1
MLP	0,967	0,968	0,979	0,951	0,983	0,959	0,945
RF	1	1	1	1	1	1	1
DT	1	1	1	1	1	1	1
NB	0,998	0,998	0,995	1	0,996	0,99	0,996
SVM	1	1	1	1	1	1	1

Tabla 25: Métricas de Rendimiento MaxAbsScaler - Neighborhood CA (PC1, PC2)

Robust Scaler - Neighborhood CA (PC1, PC2)

La relación entre características en algunos casos se siguen los mismos valores que en el apartado anterior. En la Figura 51 se muestran las que se usan en este apartado. Ambos componentes vienen dados por una relación directa por VENTRICLES (aunque con menor aportación en el primer

componete), FAQTOTAL y AGE. De manera inversa por ABETA (con menor aportación en el segundo componente) y MMSOCRE.

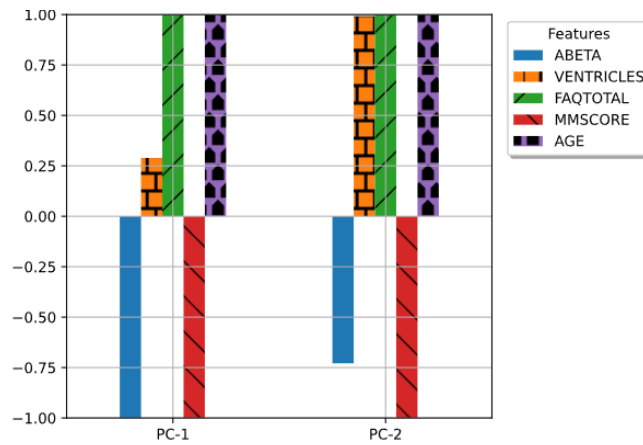


Figura 51: Relación entre componentes y las características del conjunto de datos usadas Robust Scaler - Neighborhood CA

Nuevamente esta representación de la información es de alto valor semántico, con una separación muy marcada entre clases (Figura 52).

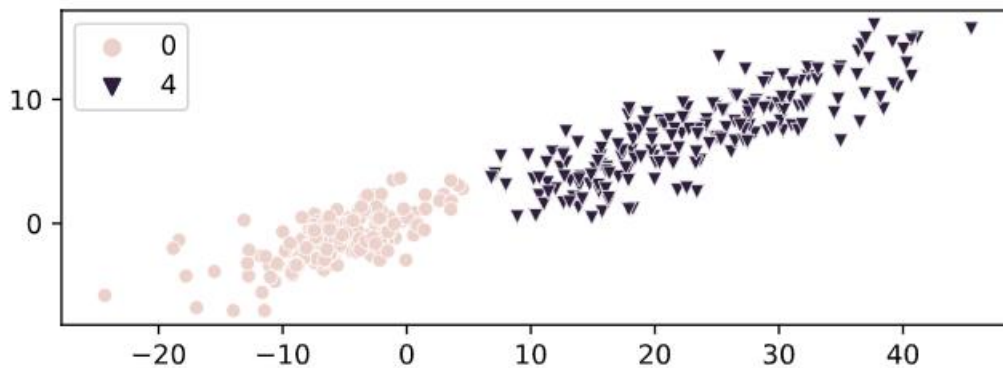


Figura 52: : Representación de los datos mediante el preprocesado Robust Scaler - Neighborhood CA (PC1, PC2)

Superando incluso preprocesado, inmediatamente anterior, con esta combinación de escalado-proyección, se alcanzan, los mejores resultados hasta ahora vistos. Cuatro de los clasificadores obtienen el valor máximo para la AUC =1 (Figura 52) igual que en el resto de las métricas de rendimiento analizadas (Tabla 26).

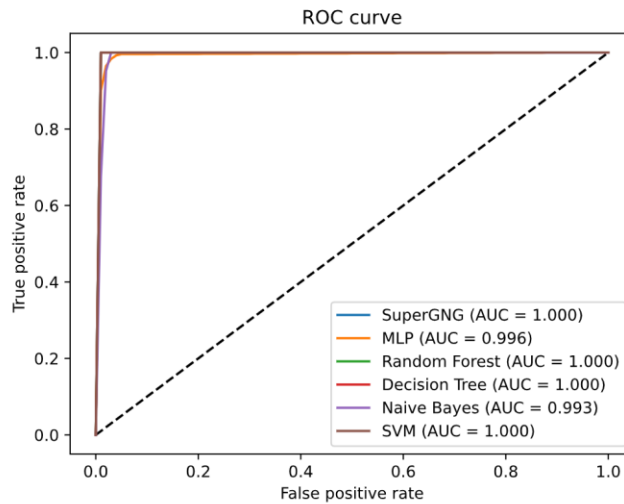


Figura 53: Curvas ROC de los clasificadores - Robust Scaler - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SuperGNG	1	1	1	1	1	1	1
MLP	0,996	0,996	0,995	0,996	0,996	0,99	0,992
RF	1	1	1	1	1	1	1
DT	1	1	1	1	1	1	1
NB	0,993	0,993	0,984	1	0,987	0,969	0,987
SVM	1	1	1	1	1	1	1

Tabla 26: Métricas de Rendimiento Robust Scaler - Neighborhood CA (PC1, PC2)

Clasificación sin AGE

A continuación, repetiremos los experimentos de clasificación, pero excluyendo la edad del conjunto de datos.

Se realiza, a continuación, el estudio del rendimiento de los clasificadores para resolver la misma clasificación CN-EA con solo criterios diagnósticos, con pruebas invasivas, pero sin considerar la característica AGE.

Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

Vemos que algunas relaciones entre características tienen valores mínimos o incluso nulos para la proyección (Figura 54). En esta relación no aparece el criterio ABETA en ninguno de los

componentes. El primer componente viene dado por una relación directa por VENTRICLES y FAQTOTAL, e inversa por MMSOCRE. Mientras que el segundo componente viene dado por FAQTOTAL de manera directa y el resto de los criterios inversa.

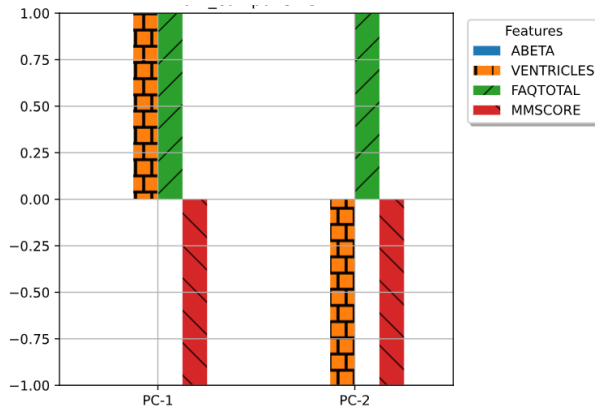


Figura 54: Relación entre componentes y las características del conjunto de datos usadas Power Transformer Yeo-Johnson - Neighborhood CA

La separación entre clases en este caso es muy significativa, pero hay varias muestras que se encuentran integradas entre las muestras de las clases opuestas.

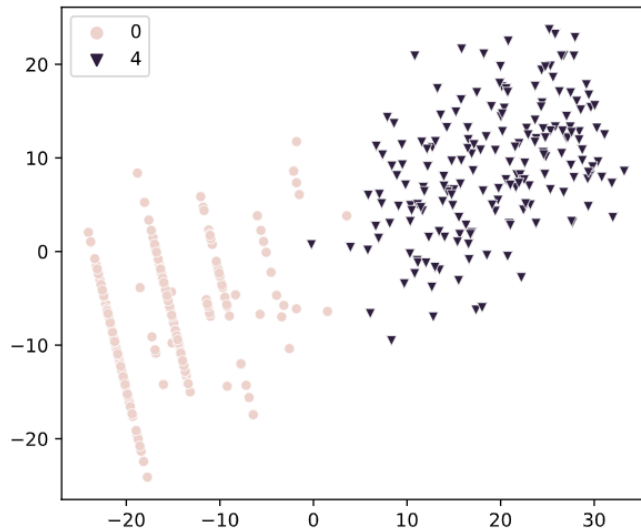


Figura 55: : Representación de los datos mediante el preprocesado Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

Aunque los resultados son muy positivos tanto en AUC (Figura 56) como en las métricas adicionales de rendimiento (Tabla 27), se nota la penalización de las muestras mezcladas en las clases y probablemente es el motivo por el cuál no se obtienen valores de 1 en todos los casos.

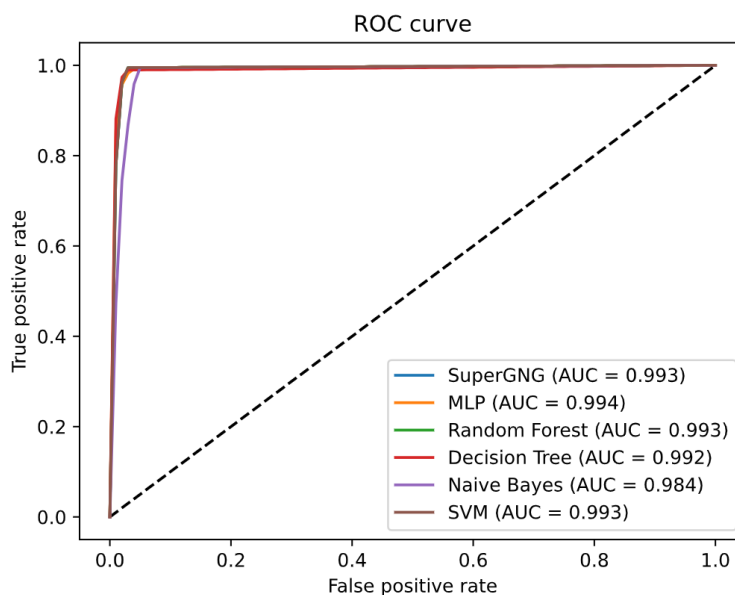


Figura 56: Curvas ROC de los clasificadores - Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SuperGNG	0,993	0,993	0,99	0,995	0,991	0,979	0,987
MLP	0,994	0,994	0,993	0,995	0,994	0,986	0,99
RF	0,993	0,993	0,99	0,995	0,991	0,979	0,987
DT	0,992	0,993	0,995	0,989	0,996	0,99	0,987
NB	0,984	0,983	0,969	0,995	0,974	0,94	0,969
SVM	0,993	0,993	0,99	0,995	0,991	0,979	0,987

Tabla 27: Métricas de rendimiento Power Transformer Yeo-Johnson - Neighborhood CA (PC1, PC2)

MaxAbsScaler - Neighborhood CA (PC1, PC2)

Volvemos a encontrarnos relaciones con valores en límites entre los criterios diagnósticos, observamos esto en la Figura 57. EL primer componente viene dado por FAQTOTAL con relación directa y el resto una relación inversa, aunque VENTRICLES en menor medida. El segundo componente está compuesto por una relación directa por ABETA y MMSCORE y el resto con una relación inversa.

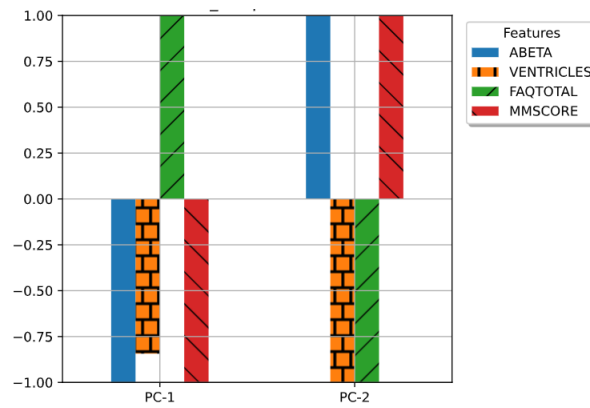


Figura 57: : Relación entre componentes y las características del conjunto de datos usadas MaxAbsScaler - Neighborhood CA

La representación de los datos es bastante buena, primera vez que vemos esta característica en conjuntos de datos con características invasivas.



Figura 58: Representación de los datos mediante el preprocesado MaxAbsScaler - Neighborhood CA (PC1, PC2)

Además, observamos que MLP cada vez que ha tenido valores menores de rendimiento ha sido con pruebas invasivas, y se repite este hecho como observamos en la Figura 59 y en la Tabla 28.

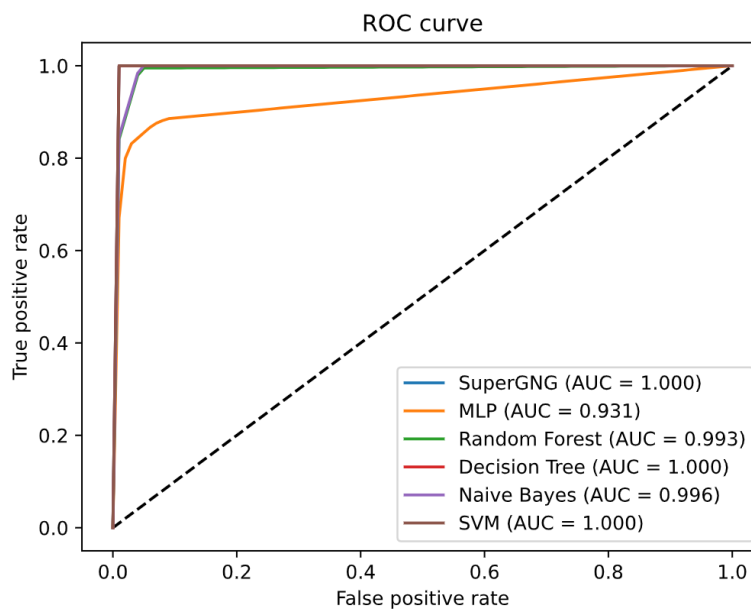


Figura 59: Curvas ROC de los clasificadores - MaxAbsScaler - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SupergNG	1	1	1	1	1	1	1
MLP	0,931	0,937	0,985	0,875	0,988	0,971	0,902
RF	0,993	0,993	0,99	0,995	0,991	0,98	0,987
DT	1	1	1	1	1	1	1
NB	0,996	0,995	0,99	1	0,991	0,98	0,991
SVM	1	1	1	1	1	1	1

Tabla 28: Métricas de Rendimiento MaxAbsScaler - Neighborhood CA (PC1, PC2)

Robust Scaler - Neighborhood CA (Dim 2) (PC1, PC2)

Robust Scaler y Neighborhood CA en este caso tiene variaciones muy significativas ente características como vemos en la Figura 60. El primer componente viene dado por VENTRICLES y FAQTOTAL de manera directa, y el resto inversa. El segundo componente se compone por VENTRICLES Y MMSCORE de manera directa y el resto de los criterios de manera inversa.

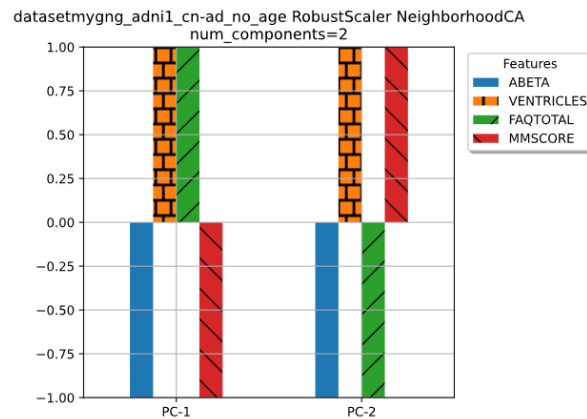


Figura 60: Relación entre componentes y las características del conjunto de datos usadas Robust Scaler – Neighborhood CA

Los datos, aunque dispersos, tienen una separación muy significativa entre clases, como se observa en la Figura 61.

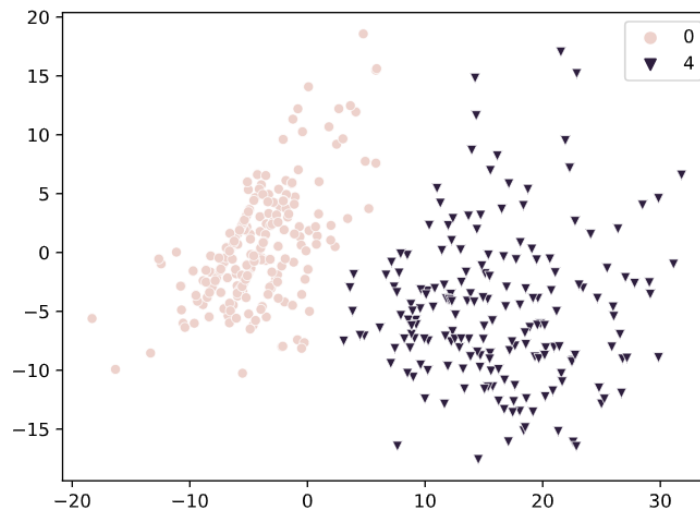


Figura 61: : Representación de los datos mediante el preprocesado Robust Scaler - Neighborhood CA (PC1, PC2)

Se ve reflejada la separación en los resultados de los clasificadores, donde todos y cada uno de ellos obtiene valores de AUC muy cercanos a 1 (Figura 62) y las otras métricas de rendimiento acorde a estos valores (Tabla 29).

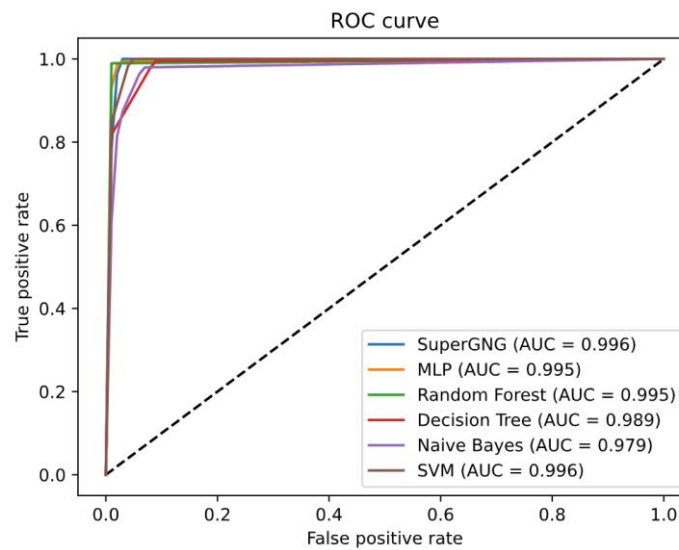


Figura 62: Curvas ROC de los clasificadores - Robust Scaler - Neighborhood CA (PC1, PC2)

Classifier	AUC	Accuracy	Precision	Recall	Specificity	CUI+	CUI-
SupeRGNG	0,996	0,995	0,99	1	0,991	0,98	0,991
MLP	0,995	0,996	0,997	0,994	0,997	0,994	0,992
RF	0,995	0,995	1	0,989	1	1	0,991
DT	0,989	0,988	0,98	0,995	0,983	0,963	0,978
NB	0,979	0,978	0,974	0,979	0,978	0,949	0,961
SVM	0,996	0,995	0,99	1	0,991	0,98	0,991

Tabla 29: Métricas de Rendimiento Robust Scaler - Neighborhood CA (PC1, PC2)

3.3.4. Comparación global de los clasificadores

Se realiza una comparación global de los distintos clasificadores utilizados para la detección de la EA a través de la clasificación binaria CN-EA, como hemos visto en las secciones anteriores. Esta comparación se lleva a cabo a través de la métrica de rendimiento de clasificadores más fiable, el AUC, como se puede ver en la Tabla 30 y la Tabla 31.

Con Edad						
	Invasivo			No Invasivo		
Clasificadores / Combinaciones	PowerTransformer_ / Yeo-Johnson / Neighbourhood CA/ 2	MaxAbsScaler / Neighbourhood CA/ 2	RobustScaler / Neighbourhood CA/ 2	PowerTransformer_ / Yeo-Johnson / Neighbourhood CA/ 2	MaxAbsScaler / Neighbourhood CA/ 2	RobustScaler / Neighbourhood CA/ 2
SupeRGNG	0,997	1	1	0,99	0,991	1
MLP	0,999	0,967	0,996	0,985	0,96	0,977
RF	0,997	1	1	0,995	0,99	0,995
DT	0,993	1	1	0,995	0,991	1
NB	0,996	0,998	0,993	0,989	0,986	0,991
SVM	0,997	1	1	0,995	0,995	1

Tabla 30: Valores de AUC obtenidos por los clasificadores en los conjuntos de datos Con edad

Sin Edad						
	Invasivo			No Invasivo		
Clasificadores / Combinaciones	PowerTransformer_ Yeo-Johnson / Neighbourhood CA/ 2	MaxAbsScaler / Neighbourhood CA/ 2	RobustScaler / Neighbourhood CA/ 2	PowerTransformer_ Yeo-Johnson / Neighbourhood CA/ 2	MaxAbsScaler / Neighbourhood CA/ 2	RobustScaler / Neighbourhood CA/ 2
SupergNG	0,993	1	0,996	0,993	0,995	0,996
MLP	0,994	0,931	0,995	0,988	0,932	0,987
RF	0,993	0,993	0,995	0,993	0,991	0,991
DT	0,992	1	0,989	0,993	0,992	0,988
NB	0,984	0,996	0,979	0,987	0,982	0,98
SVM	0,993	1	0,996	0,99	0,995	0,993

Tabla 31: Valores de AUC obtenidos por los clasificadores en los conjunto de datos Sin edad

Además, se ha pasado una prueba estadística para evaluar si podemos comparar los resultados entre los distintos clasificadores, teniendo en cuenta la distribución de los datos. Se ha obtenido un p-valor utilizando un Alpha de 0.05, el cual es significativo estadísticamente ($9,27611 \times 10^{-6} \leq 0,05$) utilizando el test estadístico de Friedman [72].

Se obtiene únicamente un p-valor porque este es el resultado de realizar la prueba uniendo la Tabla 30 y la Tabla 31. Debido a que nos encontramos con un estudio estadístico de múltiples clasificadores con múltiples conjuntos de datos distintos. Esta significancia nos indica que podemos rechazar la hipótesis de que todos los algoritmos se comportan de igual manera en distintos conjuntos de datos [73].

3.4. Discusión

Primero, es importante destacar la contribución del sistema desarrollado. Este sistema ha reducido significativamente el tiempo necesario para el entrenamiento y la clasificación de la SupergNG, así como para la ejecución de clasificadores como MLP y SVM, utilizando los mismos datos de entrada ya preprocesados para la comparación.

Además, se han incorporado herramientas para el análisis de datos, como gráficos 3D (por ejemplo, Surface) y herramientas estadísticas para analizar los datos y determinar la significancia en la comparación de clasificadores.

Para realizar la discusión, se han realizado la Tabla 30 y la Tabla 31 con todos los valores de AUC que se han obtenido a lo largo del documento.

Tras ver los resultados obtenidos en cada una de las combinaciones estudiadas, podemos afirmar que los coeficientes han aportado valor a la clasificación debido a que ningún clasificador ha obtenido menos de 0.931 de AUC en ninguno de los casos, obtenido en solo dos combinación con

el clasificador MLP, y siendo el siguiente valor más bajo de los otros clasificadores de 0.979. Además, generalmente la separación entre clases ha sido bastante significativa.

Hay que destacar que la combinación con la que hemos realizado la comparación del ranking de las métricas (Robust Scaler - Neighborhood CA) ha obtenido resultados muy positivos, a pesar de no tener tantas ocurrencias como las otras dos que hemos estudiado.

Como se comentó al principio, uno de los conjuntos de datos consta de pruebas invasivas, mientras que el otro consta de pruebas que no implican pruebas invasivas.

Si comparamos ambos conjuntos de datos con la característica de la edad, el conjunto de datos con pruebas no invasivas ha obtenido mejores resultados que las combinaciones del conjuntos de datos con pruebas invasivas como vemos si comparamos la Tabla 30 y la Tabla 31 cada clasificador entre sí. Además, hemos visto a lo largo del documento que la separación entre clases es mucho mayor al añadir la edad como factor de riesgo.

Como se mostró anteriormente, la edad es la única característica que obtuvo un p-valor superior al Alpha de 0.05, y por lo tanto no es significativo estadísticamente. Este suceso se debe a que es un posible factor de riesgo y no un criterio diagnóstico.

Por ello, generalmente se descartaría desde un principio. Pero en este estudio se ha demostrado que realmente, este factor de riesgo realmente es capaz de mejorar la clasificación respecto a los casos donde se ha descartado. Si nos fijamos en los valores de AUC de los clasificadores, generalmente predominan los valores elevados en los conjuntos de datos donde se incluye la edad.

Al poner en práctica este tipo de ejercicios, nos podemos dar cuenta que la estadística debe ser tomada en cuenta, aunque siempre recordando que la probabilidad no siempre acierta a la hora de definir las características que nos conviene usar [69].

Hay que destacar además que el conjunto de datos con pruebas no invasivas ha obtenido a su vez mejores resultados en la mayoría de los casos que si añadimos las pruebas invasivas. Esto sucede tanto incluyendo la edad como sin ella en el conjunto.

Esto es un punto muy importante, porque las pruebas invasivas pueden ser incómodas o incluso dolorosas para los sujetos, además de en algunos casos necesitar instrumental médico poco asequible para su realización. Es por ello por lo que el conjunto de datos con pruebas no invasivas debe ser tomado en cuenta como una buena opción para la clasificación y el diagnóstico temprano del Alzheimer.

Nos centramos ahora en evaluar el rendimiento de la red ontogénica SuperGNG en los conjuntos de datos de ADNI, donde los resultados son mejores de lo que se habría pensado antes de realizarlo.

De las 12 combinaciones evaluadas, los resultados de AUC de SuperGNG es el mejor obtenido entre los clasificadores, empatado en ocasiones con otros clasificadores, más concretamente en 9 de ellos. En los otros 4 la diferencia máxima que existe entre el mejor resultado de otros clasificadores es de 0.04 de AUC.

SuperGNG obtuvo valores de AUC mayores o iguales a 0.99 en todos los casos, lo que demuestra que la red es capaz de adaptarse correctamente a los distintos problemas de clasificación que se le han presentado. Además, obtiene un valor de AUC de 1 en 4 de las combinaciones de proyectado y escalado junto a DT y SVM, siendo estas las que mejores resultados han presentado.

Capítulo 4: Conclusiones y trabajos futuros

En este capítulo se exponen las conclusiones que se han sacado del estudio y los posibles trabajos futuros que este ha dejado abiertos y pueden ser realizados.

4.1. Conclusiones

- Los objetivos propuestos en el trabajo han sido alcanzados de forma completa y satisfactoria.
- La red ontogénica SuperGNG ha mostrado con su rendimiento que tiene utilidad en el diagnóstico temprano de la EA, alcanzando resultados altamente prometedores, con valores de AUC ≥ 0.99 en todos los casos aquí estudiados.
- Aunque los conjuntos de datos con pruebas invasivas mostraron mejores resultados, la diferencia no ha sido significativa. Por ello, se considera que las pruebas no invasivas tienen una alta relevancia para la ayuda al diagnóstico de la EA. Además de proporcionar accesibilidad, comodidad y seguridad a los sujetos.
- La edad se ha determinado factor de riesgo importante para la EA. Cuando este dato sociodemográfico se utiliza juntamente con los criterios diagnósticos indicados en la conclusión anterior, mejora la precisión del diagnóstico, pasando el valor de AUC desde un 0,96 a un 0,93 sin considerarla. Para reafirmar este hecho, observamos las métricas CUI+ y CUI-. El valor más bajo obtenido en los conjuntos de datos que incluyen la edad es de 0.914 (CUI+) y 0.935 (CUI-). En contraste, excluyendo la edad, estos valores disminuyen a 0.902 (CUI+) y 0.902 (CUI-).
- Los coeficientes de clustering utilizados aportan valor al estudio de los datos y de las distintas combinaciones de escalado y proyectado. Además, la combinación que ha mantenido unos resultados elevados en todos los conjuntos de datos ha sido el escalado “Power-Transformer” con la proyección “Neighborhood” con 2 componentes, por encima de las otras dos combinaciones.
- El sistema software desarrollado puede facilitar las tareas de clasificación y otras operaciones útiles en este tipo de estudios, como lo son la obtención de estadísticos o el graficado de métricas de rendimiento. Gracias a la automatización de procesos, se espera que tenga una contribución positiva en futuros estudios, además de reducir significativamente el tiempo necesario para su realización.
- El desarrollo del sistema se benefició significativamente de la metodología propuesta. También aseguró que el proyecto avanzara de acuerdo con lo planeado, demostrando la importancia de una metodología bien definida.

4.2. Trabajos Futuros

Los excelentes resultados obtenidos en la clasificación de la red SuperGNG abren la puerta a una herramienta potente que puede mejorar el diagnóstico y seguimiento de la EA. Una vez evaluado que permita a los profesionales de la salud utilizar esta herramienta para el diagnóstico temprano del Alzheimer, facilitando intervenciones farmacoterapéuticas personalizadas.

Seguir promoviendo, estableciendo y afianzando la colaboración con profesionales de la salud como neurólogos, geriatras y médicos de atención primaria, permitirá interpretar mejor los resultados y ajustar el modelo según las necesidades clínicas. Investigar las implicaciones éticas y sociales del uso de inteligencia artificial en el diagnóstico de enfermedades neurodegenerativas asegurará que las herramientas desarrolladas sean equitativas y respeten la privacidad de los sujetos, promoviendo su aceptación y uso generalizado.

Sería de interés también, evaluar SuperGNG como una herramienta de clasificación aplicado a otros entornos clínicos, como una solución computacional inteligente para la ayuda al diagnóstico de enfermedades de difícil detección y diagnóstico.

Estos trabajos futuros no solo mejorarán la precisión y aplicabilidad de la red SuperGNG, sino que también contribuirán a convertirla en una herramienta para la lucha contra el Alzheimer, mejorando el diagnóstico temprano y el tratamiento de esta enfermedad devastadora.

Capítulo 5: Bibliografía

- [1] L. Casado Lira, «Biomarcadores y biosensores para el diagnóstico del Alzheimer,» 2017.
- [2] R. Lanfranco, P. Manríquez-Navarro, L. Avello y A. Canales-Johnson, «Evaluación de la enfermedad de Alzheimer en etapa temprana: biomarcadores y pruebas neuropsicológicas,» *Revista médica de Chile*, 140(9), 1191-1200, 2012.
- [3] J. L. Piscoya Tirado, «Aplicación de la Inteligencia Artificial en el tratamiento de la enfermedad del Alzheimer: una revisión sistemática de la literatura,» 2023.
- [4] Ministerio de Sanidad, Consumo y Bienestar Social, «Plan Integral de Alzheimer y otras Demencias (2019-2023),» Madrid, 2020.
- [5] J. C. Mingote Adán, M. Gálvez Herrer, P. D. Pino Cuadrado y M. Gutiérrez García, «El paciente que padece un trastorno depresivo en el trabajo,» *Medicina y seguridad del trabajo*, 55(214), 41-63., 2009.
- [6] World Health Organization, «FIRST WHO MINISTERIAL CONFERENCE ON GLOBAL ACTION,» 2015.
- [7] R. Nitrini y S. M. D. Bruck, «Demencia: definición y clasificación,» *Revista neuropsicología, neuropsiquiatría y neurociencias*, 12(1), 75-98, 2012.
- [8] P. González-Martínez, J. Oltra-Cucarella, E. Sitges-Maciá y B. Bonete-López, «Revisión y actualización de los criterios de deterioro cognitivo objetivo y su implicación en el deterioro cognitivo leve y la demencia,» *Rev Neurol* 2021;72 (08):288-295, 2021.
- [9] M. J. Robles, I. Fort y F. Formiga, «Demencia, una enfermedad evolutiva: demencia severa. Identificación de demencia terminal,» *Revista Española de Geriátría y Gerontología*, Volume 44, Supplement 2, 2009.
- [10] C. Quinteros Alvarado y J. M. D. L. Á. Camey Quinteros, «Importancia del acompañamiento psicológico al cuidador primario de enfermos con demencia tipo alzheimer,» *Doctoral dissertation*, Universidad de San Carlos de Guatemala.
- [11] M. Herrera-Rivero, M. E. Hernández-Aguilar, J. Manzo y G. E. Aranda-Abreu, «Enfermedad de Alzheimer: inmunidad y diagnóstico,» *Revista de neurología*, 51(3), 153-164., 2010.
- [12] D. J Selkoe y J. Hardy, «The amyloid hypothesis of Alzheimer's disease at 25 years,» *EMBO Molecular Medicine*, pp. 595-608, 2016.
- [13] M. R. Irwin y M. V. Vitiello, «Implications of sleep disturbance and inflammation for Alzheimer's disease dementia,» *The Lancet Neurology* Volume 18, Issue 3, 2019.

- [14] K. Zeinab Breijyeh y R. , «Comprehensive Review on Alzheimer's Disease: Causes and Treatment,» *Molecules* 2020, 25(24), 5789, 2020.
- [15] Y. Cabrera-León, P. García Báez, P. Fernández-López y C. P. Suárez-Araujo, «Neural computation-based methods for the early,» *A systematic review*, 2023.
- [16] R. González, C. Martín, M. Martín, M. González, F. García de Blas y S. Riu, «Enfermedad de Alzheimer,» *Clínica, tratamiento y rehabilitación.*, 2000.
- [17] S. L. Warren, E. Reid, P. Whitfield, A. M. Helal, E. G. Abo Hamza y M. S. Hamid, «Cognitive and behavioral abnormalities in individuals with Alzheimer's disease, mild cognitive impairment, and subjective memory complaints,» *Current Psychology*, 43(1), 800-810., 2024.
- [18] M. Romano, M. D. Nissen, N. Del Huerto y C. Parquet, «Enfermedad de alzheimer,» *Revista de posgrado de la vía cátedra de medicina*, 75, 9-12., 2007.
- [19] J. C. Núria Setó-Salvia, «Genética en la enfermedad de Alzheimer,» *Revista de Neurología*, 50(6), 360-364., 2010.
- [20] L. A. M. Cadavid y A. M. Pérez-Acosta, «na aproximación de la literatura científica sobre la relación entre reconocimiento de emociones, envejecimiento, deterioro cognitivo y demencias.».
- [21] J. G. Rodríguez y G. G. Gutiérrez, «Definición y prevalencia del deterioro cognitivo leve,» *Revista española de Geriatria y gerontología*, 52, 3-6, 2017.
- [22] M. Pose y F. Manes, «Deterioro cognitivo leve,» *Acta Neurol Colomb*, 26(3), 7-12., 2010.
- [23] Alzheimer's Disease Neuroimaging Initiative, «ADNI,» 2024. [En línea]. Available: <https://adni.loni.usc.edu/>.
- [24] J. P. Heredia, J. T. Nalvaiz y S. O. Gasso, «Técnicas de aprendizaje supervisado en la enfermedad de Alzheimer,» *Aplicaciones al estudio ADNI*, 2014.
- [25] B. T. Wyman, D. J. Harvey, K. Crawford, M. A. Bernstein, O. Carmichael, P. E. Cole y A. D. N. I. , «Standardization of analysis sets for reporting results from ADNI MRI data,» *Alzheimer's & Dementia*, 9(3), 332-337, 2013.
- [26] A. Moreno, E. Armengol, J. Béjar Alonso, L. A. Belanche Muñoz, C. U. Cortés García, R. Gavaldà Mestre y R. Sánchez-Marrè, *Aprendizaje automático*, 1994.
- [27] G. Santafé, J. A. Lozano y P. Larranaga, *Aprendizaje discriminativo de clasificadores Bayesianos.*, *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*, 10(29), 39-47., 2006.
- [28] J. McCarthy, *WHAT IS ARTIFICIAL INTELLIGENCE?*, 2007.
- [29] R. F. López y J. M. F. Fernández, *Las redes neuronales artificiales.*, *Netbiblo*, 2008.

- [30] I. Rish, «An empirical study of the naive Bayes classifier.» In IJCAI 2001 workshop on empirical methods in artificial intelligence (Vol. 3, No. 22, pp. 41-46)., 2001.
- [31] R. Fernández Regalado, «El teorema de Bayes y su utilización en la interpretación de las pruebas diagnósticas en el laboratorio clínico.» Revista cubana de investigaciones biomédicas, 28(3), 158-165., 2009.
- [32] L. B. Martínez y A. M. Rodríguez, «Estudio e Implementación de una Máquina de Soporte Vectorial.»
- [33] P. L. Chinas, «SVM and ANN application to multivariate pattern recognition using scatter data.» IEEE Latin America Transactions, 13(5), 1633-1639., 2015.
- [34] G. Valenzuela González, «Aprendizaje Supervisado: Métodos, Propiedades y Aplicaciones.» 2022.
- [35] R. Burbidge y B. Buxton, «An introduction to support vector machines for data mining.» Keynote papers, young OR12, 3-15., 2001.
- [36] E. Fernández-Alvarado, Detección de manos en imágenes de profundidad mediante el uso de bosques de decisión aleatorios., 2015.
- [37] R. F. M. Merino y C. I. Ñ. Chacón, Bosques aleatorios como extensión de los árboles de clasificación con los programas R y Python., Interfases, (10), 165-189., 2017.
- [38] J. A. León Muñoz, J A y M. Á. Acevedo Franco, «Formulación para tratamientos adecuados a la fibrosis quística utilizando árboles de decisión.» 2015.
- [39] A. Fernández Khatiboun, «Machine learning en ciberseguridad.» 2019.
- [40] IBM, «IBM,» [En línea].
- [41] A. Cutler, D. R. Cutler y J. R. Stevens, «Random forests,» Ensemble machine learning: Methods and applications, 157-175., 2012.
- [42] S. J. Rigatti, Random forest, Journal of Insurance Medicine, 47(1), 31-39., 2017.
- [43] F. Murtagh, «Multilayer perceptrons for classification and regression,» Neurocomputing, Volume 2, Issues 5–6, 1991.
- [44] M. G. Longoni, E. Porcel, M. V. López y G. N. Dapozo, «Modelos de Redes Neuronales Perceptrón Multicapa y de Base Radial para la predicción del rendimiento académico de alumnos universitarios,» In XVI Congreso Argentino de Ciencias de la Computación., 2010.
- [45] G. A. Toro Bayona y I. A. Lizarazo Salcedo, «Evaluación de las redes neuronales artificiales Perceptron Multicapa y Fuzzy-Artmap en la clasificación de imágenes satelitales.» Ingeniería, 17(1), 61-72., 2012.

- [46] M. C. Popescu, V. E. Balas, L. Perescu-Popescu y N. Mastorakis, «Multilayer perceptron and neural networks.,» WSEAS Transactions on Circuits and Systems, 8(7), 579-588., 2009.
- [47] H. Vivas, «Optimización en entrenamiento del perceptrón multicapa.,» Universidad del Cauca, Popayán, 2014.
- [48] D. Fišer, F. Jan y K. Miroslav, «Growing neural gas efficiently,» Neurocomputing Volume 104, 2013.
- [49] Á. Serrano Musalem, «Redes neuronales artificiales auto-organizativas dinámicas.,» 2015.
- [50] B. Fritzke, «A Growing Neural Gas Network Learns Topologies,» Neural Information Processing Systems, Vol 7, 1995.
- [51] J. Holmström, «Growing neural gas,» Unpublished Master's Thesis, Uppsala University., 2002.
- [52] Y. Cabrera-León, P. G. Báez, P. Fernández-López y C. P. Suárez-Araujo, «Study on Mild Cognitive Impairment and Alzheimer's Disease Classification Using a New Ontogenic Neural Architecture, The Supervised Reconfigurable Growing Neural Gas,» 2023 Annual Modeling and Simulation Conference (ANNSIM), Hamilton, ON, Canada, 2023.
- [53] A. Sosa-Marrero y et al, «Detección de la enfermedad de Alzheimer frente al deterioro cognitivo leve mediante una nueva red neuronal híbrida modular.,» Rojas, I., Joya, G., Català, A. (eds) Avances en Inteligencia Computacional. IWANN 2021. Apuntes de Cátedra en Ciencias de la Computación(), vol 12862. Springer, Cham., 2021.
- [54] ACERVOLIMA, «Intra Cluster Distance and Inter Cluster Distance,» <https://acervolima.com/ml-distancia-intercluster-e-intracluster/>.
- [55] F. A. A. Plazas Niño, «Introducción al análisis clúster: una aplicación en la clasificación de campos petroleros.,» 2022. [En línea]. Available: <https://dspace.utpl.edu.ec/handle/20.500.11962/25558>.
- [56] P. J. Rousseeuw, «Silhouettes: A graphical aid to the interpretation and validation of cluster analysis,» Journal of Computational and Applied Mathematics, Volume 20, Pages 53-65, 1987.
- [57] N. Arriola Landa Cosio, «Cómo seleccionar el mejor valor de K en K-means,» medium.com, <https://medium.com/@nicolasarriola/c%C3%B3mo-seleccionar-el-mejor-valor-de-k-en-k-means-21121b604365#:~:text=Davies%20Bouldin%20index&text=Se%20usa%20para%20evaluar%20que,bajos%20nos%20indican%20mejores%20resultados.,> 2021.
- [58] S. Petrovic, «A comparison between the silhouette index and the davies-bouldin index in labelling ids clusters.,» In Proceedings of the 11th Nordic workshop of secure IT systems (Vol. 2006, pp. 53-64)., Citeseer., 2006.

- [59] sklearn, «sklearn.metrics.silhouette_score,» https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html, 2011.
- [60] L. F. Pastrán Ramírez y S. Gongora Aya, «Algoritmo de selección y validación del método de clusterización óptimo para datos no supervisados.,» 2021.
- [61] K. R. Shahapure y C. Nicholas, «Cluster quality analysis using silhouette score.,» In 2020 IEEE 7th international conference on data science and advanced analytics (DSAA) (pp. 747-748). IEEE., 2020.
- [62] Y. A. Wijaya, D. A. Kurniady, E. Setyanto, W. S. Tarihoran, D. Rusmana y R. Rahim, «Davies bouldin index algorithm for optimizing clustering case studies mapping school facilities.,» TEM J, 10(3), 1099-1103., 2021.
- [63] D. L. Davies y D. W. Bouldin, «A Cluster Separation Measure,» Pattern Analysis and Machine Intelligence, IEEE Transactions on. PAMI-1. 224 - 227., 1979.
- [64] O. M. Paola Favati, «An internal validity index for arbitrarily shaped clusters.,» Expert Systems with Applications, Volume 235, 2024.
- [65] T. Caliński y H. JA, «Dendrite Method for Cluster Analysis.,» Communications in Statistics - Theory and Methods. 3. 1-27., 1974.
- [66] D. Butvinik, «linkedin.com,» 2024. [En línea]. Available: https://www.linkedin.com/posts/danny-butvinik_artificialintelligence-machinelearning-activity-7157230907257454592-jaOl/.
- [67] Y. Cabrera-León, «Analysis of hyperparameters - Dynamic SuperRGNG,» Technical report, 6/6/24.
- [68] Y. Cabrera-León, P. García Báez, P. Fernández-López y C. P. Suárez-Araujo, «Neural Computation-Based Methods for the Early Diagnosis and Prognosis of Alzheimer’s Disease Not Using Neuroimaging Biomarkers: A Systematic Review,» J Alzheimers Dis. 2024; 98(3): 793–823., 2024.
- [69] C. Carazo-Díaz y L. Prieto-Valiente, «El valor de p del test no es un ‘índice matemático’, es simplemente una frecuencia relativa,» Rev Neurol 2024;78 (07):209-211, 2024.
- [70] F. Robledo Yagüe, «Clasificación de Solanum lycopersicum y parientes silvestres mediante técnicas de aprendizaje automático partiendo de datos genéticos,» Doctoral dissertation, Universitat Politècnica de València, 2019.
- [71] A. J Mitchell , «Quickly find the Qualitative and Quantitative Accuracy & Utility of Diagnostic Predictive and Screening Tests,» <https://psycho-oncology.info/cui.html>, 2012.
- [72] R. Eisinga, T. Heskes, B. Pelzer y M. Te Grotenhuis, «Exact p-values for pairwise comparison of Friedman rank sums, with application to comparing classifiers,» BMC bioinformatics, 18, 1-18, 2017.

[73] N. Japkowicz, «Evaluating learning algorithms: a classification perspective,» Cambridge University Press.