

Instituto Tecnológico y de Estudios Superiores de Occidente

Reconocimiento de validez oficial de estudios de nivel superior según acuerdo secretarial
15018, publicado en el Diario Oficial de la Federación del 29 de noviembre de 1976.

Departamento de Matemáticas y Física
Maestría en Ciencia de Datos



Recuerdos y redes: Detectando el Alzheimer antes del olvido

TRABAJO RECEPCIONAL que para obtener el **GRADO** de
Maestro en Ciencia de Datos

Presenta:
Héctor A. Galván Blanco

Director:
Jaime Emmanuel Alcalá Temores

Tlaquepaque, Jalisco, 23 de junio de 2025

Recuerdos y redes: Detectando el Alzheimer antes del olvido

Héctor A. Galván Blanco

Resumen

El diagnóstico temprano de enfermedades neurodegenerativas como el Alzheimer es fundamental para mejorar la calidad de vida de los pacientes y facilitar una intervención médica oportuna. Este trabajo explora el uso de modelos de aprendizaje supervisado para clasificar el estado cognitivo de individuos: Cognitivamente Normal (CN), Deterioro Cognitivo Leve (MCI) y Alzheimer (AD), a partir de variables clínicas y demográficas estructuradas. Se implementaron y compararon dos enfoques: un modelo de regresión logística multinomial y una red neuronal profunda, desarrollada tras múltiples pruebas con arquitecturas de diferente complejidad. El proceso incluyó limpieza, codificación, normalización y selección de variables. La red neuronal, optimizada con múltiples capas densas y regularización, logró un desempeño superior, particularmente en la detección de la etapa MCI (recall = 81 %), considerada crítica por su relevancia clínica. Estos resultados refuerzan el potencial de las redes neuronales como herramienta de apoyo al diagnóstico temprano en contextos clínicos.

Palabras clave: Alzheimer, deterioro cognitivo, redes neuronales, regresión logística, diagnóstico temprano.

Memories and Networks: Detecting Alzheimer Before the Fade

Héctor A. Galván Blanco

Abstract

Early diagnosis of neurodegenerative diseases such as Alzheimer's is crucial for improving patients' quality of life and enabling timely medical intervention. This study explores the use of supervised learning models to classify individuals' cognitive states—Cognitively Normal (CN), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD)—based on structured clinical and demographic data. Two approaches were implemented and compared: a multinomial logistic regression model and a deep neural network, developed through extensive testing with varying architectural complexity. The process involved data cleaning, encoding, normalization, and feature selection. The neural network, optimized with multiple dense layers and regularization, achieved superior performance, particularly in identifying the MCI stage (recall = 81%), which is considered clinically critical. These results highlight the potential of neural networks as a support tool for early diagnosis in clinical settings.

Keywords: Alzheimer's, cognitive impairment, neural networks, logistic regression, early diagnosis.

Tabla de Contenidos

| | Página |
|---|---------------|
| 1 Introducción | 13 |
| 1.1. Contexto | 13 |
| 1.2. Problema | 14 |
| 1.3. Objetivos | 14 |
| 1.3.1. Objetivo general | 15 |
| 1.3.2. Objetivos específicos | 15 |
| 2 Metodología | 17 |
| 2.1. Descripción de los datos | 17 |
| 2.2. Análisis exploratorio | 19 |
| 2.3. Descripción de los modelos | 21 |
| 2.4. Descripción de las métricas | 24 |
| 2.5. Descripción de los experimentos o simulaciones | 25 |
| 3 Resultados y discusión. | 27 |
| 3.1. Resultados y discusión | 27 |
| 4 Conclusiones y trabajo futuro. | 33 |
| 4.1. Conclusiones | 33 |
| 4.2. Trabajo futuro | 33 |
| Bibliografía | 35 |

Índice de figuras

| | Página |
|---|---------------|
| 2.1. Distribución de variables | 21 |
| 2.2. Número de pacientes acorde al sexo | 22 |
| 2.3. Diagnóstico clínico según sexo | 23 |
| 2.4. Diagnóstico clínico según nivel educativo | 23 |
| 3.1. Matriz de confusión de modelo de Regresión Logística . | 28 |
| 3.2. Curvas ROC para el modelo de Regresión Logística . . . | 29 |
| 3.3. Pérdida de la Red Neuronal por época | 31 |
| 3.4. Precisión de la Red Neuronal por época | 31 |

Índice de tablas

| | Página |
|--|---------------|
| 2.1. Variables no incluidas en los modelos | 19 |
| 2.2. Variables utilizadas en los modelos | 20 |
| 3.1. Métricas por clase | 27 |
| 3.2. Resultados Modelo de Red Neuronal | 30 |

Dedicado a...

- *Mariely: por todo tu amor, apoyo y contagioso entusiasmo. Som i serem.*
- *Familia Galván-Blanco: Por su soporte y siempre fomentar mi curiosidad.*
- *A mis compañer@s de clase por hacer más entendible y ameno este camino.*
- *Emmanuel Alcalá Temores, por su guía y dedicación en el desarrollo de este proyecto.*
- *Instituto Tecnológico de Estudios Superiores de Occidente (ITESO) por la oportunidad de aprender a través de sus profesores, y al mismo tiempo hacerlo yo, en mi rol como docente.*

1 Introducción

En este capítulo se presenta el contexto del objeto de estudio, la justificación del objeto de estudio, la definición del problema y los objetivos generales y específicos.

1.1 Contexto

En la actualidad, los sistemas de salud de todo el mundo enfrentan un reto cada vez más urgente: el envejecimiento de la población y el incremento asociado de enfermedades neurodegenerativas como la enfermedad de Alzheimer. De acuerdo con información de la Organización Mundial de la Salud (OMS), más de 55 millones de personas en el mundo viven con algún tipo de deterioro cognitivo severo, y se pronostica que esta cifra se duplicará para el 2050 [1]. Este padecimiento representa la causa más frecuente de estos trastornos, con un impacto significativo tanto en los pacientes como en sus familias y cuidadores, así como en los sistemas sanitarios.

Con esta problemática en mente, han sido desarrolladas varias iniciativas de investigación a nivel internacional. Fruto de ellas, han sido generadas bases de datos públicas que permiten analizar múltiples dimensiones de la enfermedad, desde factores genéticos hasta variables cognitivas y clínicas. Bases de datos como ADNI (Alzheimer's Disease Neuroimaging Initiative), AIBL (Australian Imaging, Biomarkers and Lifestyle Study of Ageing) y NACC (National Alzheimer's Coordinating Center) han abierto oportunidades para aplicar herramientas avanzadas de ciencia de datos en el estudio de la progresión de este padecimiento [2, 3].

Particularmente, el desarrollo de modelos de predicción del estado clínico de los pacientes a partir de información estructurada puede facilitar la detección temprana de alteraciones cognitivas y apoyar la toma de decisiones médicas. Justo en eso, se enfoca este trabajo.

La enfermedad de Alzheimer representa un desafío creciente para los sistemas de salud a nivel mundial. Su carácter progresivo y el impacto que genera en la calidad de vida de los pacientes y sus familias la convierten en una de las condiciones neurológicas más relevantes

del siglo XXI. Actualmente, más de 55 millones de personas viven con deterioro cognitivo en el mundo, y se espera que esta cifra se eleve a 139 millones para 2050 [1].

En México se estima que más de 1.3 millones de personas padecen esta enfermedad, y que este número podría triplicarse en los próximos 25 años [4]; [5]). Además, alrededor del 60% de los casos en el país no están diagnosticados, debido a la limitada disponibilidad de herramientas clínicas eficaces, personal capacitado y políticas públicas centradas en la detección temprana [6].

A nivel regional, América Latina y el Caribe experimentan una situación similar. Se proyecta que el número de personas con Alzheimer aumentará un 300% para 2050, pasando de aproximadamente 7.8 millones en 2021 a más de 32 millones [7]. Esta tendencia se ve agravada por el rápido envejecimiento de la población y la escasez de recursos destinados a su atención.

Bajo esta premisa,, el uso de herramientas de ciencia de datos ofrece una oportunidad para anticipar el deterioro cognitivo y apoyar la toma de decisiones médicas, especialmente en países con recursos limitados. Esta investigación tiene valor práctico y académico, ya que por un lado permite explorar y comparar el comportamiento de distintos modelos de clasificación en un problema de salud pública relevante, y por otro, fomenta la aplicación de técnicas avanzadas de aprendizaje automático en contextos reales, utilizando datos clínicos reales.

1.2 *Problema*

Existen herramientas clínicas para detectar el deterioro cognitivo relacionado con la enfermedad de Alzheimer, sin embargo, su aplicación puede ser limitada por factores como el acceso a especialistas, la variabilidad en los criterios de evaluación y la subjetividad en la interpretación de los resultados.

Por ello, resulta interesante explorar alternativas que permitan predecir el estado clínico de los pacientes a partir de variables cuantitativas disponibles en bases de datos estructuradas. En este sentido, persiste una pregunta clave que guía esta investigación:

¿Es posible predecir de manera precisa el estado clínico de pacientes con posible enfermedad de Alzheimer utilizando modelos de aprendizaje supervisado, y cómo se compara el desempeño de una regresión logística con el de una red neuronal?

1.3 *Objetivos*

1.3.1 *Objetivo general*

Explorar el potencial de los modelos de aprendizaje supervisado para predecir el estado clínico de pacientes con signos de deterioro cognitivo, utilizando datos clínicos estructurados, con el fin de comparar el desempeño de un enfoque clásico basado en regresión logística con uno más complejo basado en redes neuronales.

1.3.2 *Objetivos específicos*

1. Seleccionar y preparar un conjunto de datos clínicos estructurados para el desarrollo de modelos predictivos del estado clínico en pacientes con signos de deterioro cognitivo.
2. Implementar y evaluar dos modelos de clasificación supervisada —regresión logística y red neuronal— utilizando métricas estándar de desempeño.
3. Comparar los resultados obtenidos con ambos enfoques, identificando fortalezas, limitaciones y posibles aplicaciones en contextos clínicos.

2 Metodología

2.1 Descripción de los datos

Dado el carácter sensible y confidencial de los datos médicos, se llevó a cabo un proceso de solicitud formal a distintas instituciones especializadas en la recopilación y resguardo de información relacionada con enfermedades neurodegenerativas. Este proceso incluyó el compromiso explícito de utilizar los datos con fines estrictamente académicos. Se contactaron repositorios internacionales y centros de investigación enfocados en el estudio del Alzheimer y otros padecimientos relacionados.

Se concedió el acceso al repositorio IDA@ADNI (Image & Data Archive del Alzheimer's Disease Neuroimaging Initiative) de la Universidad del Sur de California (USC por sus siglas en inglés). Este repositorio proporciona una base de datos robusta, rica en variables clínicas, demográficas y genéticas. La muestra utilizada en este trabajo incluye pacientes clasificados clínicamente como con diagnóstico de Alzheimer, deterioro cognitivo leve (MCI, por sus siglas en inglés) y controles sanos (CN, cognitively normal).

La base de datos contiene información relevante como edad, sexo, nivel educativo, antecedentes familiares, resultados de pruebas cognitivas como el MMSE, genotipo ApoE, presencia de marcadores como TAU y beta-amiloide, así como datos sobre síntomas depresivos y antecedentes de diagnóstico clínico.

Para realizar el modelado predictivo, se seleccionaron un conjunto de variables relevantes tanto desde el punto de vista clínico como científico. La elección fue fundamentada en literatura especializada sobre factores asociados al desarrollo de la enfermedad de Alzheimer y en su potencial contribución al desempeño del modelo. Se describe la relevancia de cada una a continuación:

- **Genotipo (Apolipoproteína E – ApoE4):** La presencia del alelo ApoE4 es uno de los principales factores de riesgo genético asociados al desarrollo de Alzheimer. La herencia de una o dos copias de este alelo (ya sea por herencia materna o paterna), se ha vinculado con un aumento significativo en la probabilidad de

desarrollar la enfermedad. Se decidió incorporar esta variable como el aspecto genético dentro del modelo (descartando así otras variables redundantes).

- **Edad:** Se le considera el principal factor de riesgo no modificable en esta enfermedad. Entre más edad se alcanza, mayor es la probabilidad de presentar deterioro cognitivo.
- **Género:** Existe evidencia de que el riesgo y la evolución del Alzheimer pueden variar entre hombres y mujeres. Esta diferencia puede explicarse por factores tanto biológicos como sociales. Incluir esta variable permite al modelo capturar posibles sesgos o patrones relacionados con el sexo biológico.
- **Nivel educativo (PTEDUCAT):** Las personas con mayor nivel educativo suelen tener una mayor resistencia al deterioro clínico, aun en presencia de patología cerebral. Esta variable se asocia con la noción de “reserva cognitiva”, entendida como la capacidad del cerebro para compensar el daño neurodegenerativo. Esta variable permite ajustar por diferencias individuales importantes en el pronóstico.
- **TAU:** La acumulación anómala de la proteína TAU es uno de los principales biomarcadores asociados a los procesos degenerativos del Alzheimer. Su medición fortalece la capacidad explicativa del modelo al tomar en cuenta el grado de afectación cerebral del paciente.
- **MMSE Score (MMSCORE):** El Mini-Mental State Examination (MMSE) es una herramienta estandarizada para evaluar el estado cognitivo general del paciente. Su puntuación sirve como un indicador confiable del grado de afectación cognitiva, siendo útil tanto para el diagnóstico como para el seguimiento clínico. Su inclusión aporta una dimensión objetiva del funcionamiento mental del individuo.
- **TOTALMOD:** Esta variable corresponde al puntaje total obtenido en una batería de pruebas cognitivas complementarias al MMSE. Su inclusión permite ampliar el espectro de evaluación cognitiva, considerando dominios específicos adicionales que enriquecen la caracterización clínica del paciente.
- **Diagnóstico (variable objetivo):** Representa la categoría clínica asignada al paciente: Alzheimer (Dementia), deterioro cognitivo leve (MCI) o control sano (CN). Constituye la variable dependiente del modelo, cuyo objetivo es clasificar correctamente a los individuos en una de estas tres clases diagnósticas.

Se dejaron fuera algunas variables que se consideraron redundantes (más información en el siguiente apartado) y otras que incluían o hacían alusión a imágenes cerebrales, ya que uno de los retos de la investigación consistía precisamente en verificar qué tan necesarios eran estudios en el diagnóstico de Alzheimer, ya que pueden resultar costosos y/o incómodos para el paciente.

2.2 *Análisis exploratorio*

Volviendo a las variables seleccionadas, en IDA@ADNI se cuenta con información detallada de 4,888 pacientes, de los cuales se extrajo una muestra de 1,271 individuos que contaban con presencia del alelo ApoE4, ya fuera del lado materno, del paterno o de ambos. Primeramente se realizó un análisis de calidad de los datos. Basándose en la teoría sobre Alzheimer, se decidió dejar fuera las siguientes variables del dataset:

| Variable eliminada | Significado | Motivo de eliminación |
|--------------------|---|---|
| MOTHAD | Antecedentes de Alzheimer por parte de la madre | Demasiados datos faltantes. Además esta información ya está presente en la variable "Genotipo" (Genotype) |
| FATHAD | Antecedentes de Alzheimer por parte de la madre | Demasiados datos faltantes. Además esta información ya está presente en la variable "Genotipo" (Genotype) |
| BCFAQ | Empeoramiento clínicamente relevante en las actividades de la vida diaria | Redundante con MMSCORE. Demasiados datos faltantes. |
| BCPREDX | Diagnóstico en la visita previa. | Alta correlación con la variable objetivo: Diagnóstico (Diagnosis) |
| CONSENSRES | Presencia de amiloide en el cerebro. | Alta correlación con la variable objetivo: Diagnóstico (Diagnosis) |
| SIBAD | Presencia de Alzheimer en hermanos. | Redundante al ya contar con la variable "Genotipo" (Genotype) |
| BCDEPRES | El paciente tiene antecedentes de depresión clínica | Demasiados datos faltantes. En varios casos podría ser redundante con la variable objetivo. |

Tabla 2.1: Variables no incluidas en los modelos

Una vez eliminada la información relativa a las variables de la tabla anterior, se descartaron las líneas relativas a los pacientes que no contaban con información de la variable objetivo (Diagnosis). Se explican a continuación los valores posibles que podrían ser tomados en cada una de las variables que sí fueron consideradas en la conformación del modelo predictivo.

| Atributo | Descripción |
|---|---|
| Genotipo (Genotype): Genotipo del padre /Genotipo de la madre. | 2/2 , 2/3, 2/4, 3/3, 3/4, 4/4 (Valores 4 son más propensos a presentar Alzheimer) |
| Edad (subject_age): Edad del paciente. | Variable discreta en enteros. |
| Sexo del paciente (PTGENDER) | 1 = Masculino. 2 = Femenino |
| Nivel educativo del paciente (PTEDUCAT): Años con educación escolar cursados. | Variable discreta en enteros. |
| TAU: Niveles de acumulación de esta proteína en el cerebro. | Variable discreta en enteros. |
| MMSE Score (MMSCORE): Resultado en el Mini-Mental State Examination, el cual es una herramienta estandarizada para evaluar el estado cognitivo general del paciente. | Variable discreta con valores posibles de 0 a 30 |
| TOTALMOD: Resultados en pruebas complementarias al MMSE | Variabla discreta con valores posibles de 0 a 85 |
| Diagnóstico (Diagnosis): Es la variable objetivo, diagnóstico del paciente. | 1: CN (Control Sano) , 2: MCI (Deterioro cognitivo leve), 3: Demencia |

Tabla 2.2: Variables utilizadas en los modelos

Como parte del análisis exploratorio, se examinó la distribución de las variables numéricas seleccionadas: `subject_age`, `TAU`, `MMSCORE` y `TOTALMOD`, con el objetivo de identificar posibles sesgos y evaluar estrategias de imputación adecuadas.

La variable `subject_age` mostró una distribución aproximadamente normal, por lo que en caso de haber sido necesario se habrían utilizado técnicas de imputación basadas en la media sin comprometer la representación del conjunto de datos, sin embargo, no existieron datos faltantes en esta variable. Por otro lado, las variables `TAU` y `TOTALMOD` presentaron una ligera asimetría positiva (sesgo a la derecha), lo que podría indicar la presencia de valores atípicos elevados. Para estas variables, se consideró más apropiado aplicar una imputación basada en la mediana, ya que es más robusta frente a distribuciones no simétricas.

Finalmente, la variable `MMSCORE` mostró un sesgo marcado hacia la izquierda, indicando que la mayoría de los pacientes obtuvieron puntuaciones altas en esta prueba cognitiva. Dada esta fuerte asimetría negativase utilizó la mediana como medida de imputación.

Además del análisis de las variables numéricas, se exploraron las características sociodemográficas de la muestra. En primer lugar, se evaluó la distribución por género, observándose la proporción de hombres y mujeres sin desequilibrios considerables (Figura 2.2).

Posteriormente, se analizaron las relaciones entre el diagnóstico clínico y dos variables independientes: el sexo y los años de educación. Estas visualizaciones permitieron identificar patrones preliminares, tales como diferencias en la prevalencia de condiciones cognitivas entre hombres y mujeres, así como una posible asociación entre mayor nivel

| Atributo | Descripción |
|---|---|
| Genotipo (Genotype): Genotipo del padre /Genotipo de la madre. | 2/2 , 2/3, 2/4, 3/3, 3/4, 4/4 (Valores 4 son más propensos a presentar Alzheimer) |
| Edad (subject_age): Edad del paciente. | Variable discreta en enteros. |
| Sexo del paciente (PTGENDER) | 1 = Masculino. 2 = Femenino |
| Nivel educativo del paciente (PTEDUCAT): Años con educación escolar cursados. | Variable discreta en enteros. |
| TAU: Niveles de acumulación de esta proteína en el cerebro. | Variable discreta en enteros. |
| MMSE Score (MMSCORE): Resultado en el Mini-Mental State Examination, el cual es una herramienta estandarizada para evaluar el estado cognitivo general del paciente. | Variable discreta con valores posibles de 0 a 30 |
| TOTALMOD: Resultados en pruebas complementarias al MMSE | Variabla discreta con valores posibles de 0 a 85 |
| Diagnóstico (Diagnosis): Es la variable objetivo, diagnóstico del paciente. | 1: CN (Control Sano) , 2: MCI (Deterioro cognitivo leve), 3: Demencia |

Figura 2.1: Distribución de variables

educativo y menor frecuencia de diagnóstico de Alzheimer (3) o MCI (2) y control sano (1).

2.3 Descripción de los modelos

En este apartado se describen los dos modelos que serán utilizados para predecir el estado cognitivo de los pacientes (diagnóstico): la regresión logística multinomial y una red neuronal artificial. La regresión logística será utilizada como modelo de referencia (benchmark) para evaluar el rendimiento del modelo más complejo basado en redes neuronales.

Regresión logística multinomial

Técnica estadística utilizada para modelar la probabilidad de pertenencia de una observación a una clase categórica. En su forma binaria, estima la probabilidad de que la variable dependiente Y tome el valor 1, dado un vector de características $X = (x_1, x_2, \dots, x_n)$.

En el caso de la clasificación multiclase, como la utilizada en este documento, se emplea la regresión logística multinomial. Para K clases posibles (en este caso: CN, MCI y Alzheimer), el modelo estima las probabilidades de pertenencia a cada clase mediante una función *softmax*.

Las fórmulas relativas a la regresión logística son:

Regresión logística binaria

$$P(Y = 1 | X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)}} \quad (2.1)$$

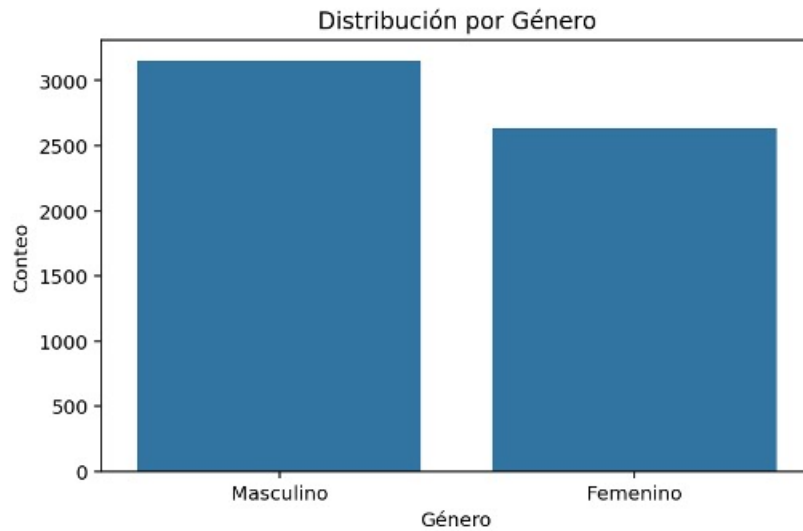


Figura 2.2: Número de pacientes acorde al sexo

Regresión logística multinomial

$$P(Y = k | X) = \frac{e^{\beta_{k0} + \beta_{k1}x_1 + \dots + \beta_{kn}x_n}}{\sum_{j=1}^K e^{\beta_{j0} + \beta_{j1}x_1 + \dots + \beta_{jn}x_n}} \quad (2.2)$$

Donde:

- β_{kj} : Coeficientes estimados para la clase k y la variable j .
- k : Clase específica para la cual se calcula la probabilidad.
- K : Número total de clases posibles.
- La clase de referencia se elige normalmente como aquella para la cual todos los parámetros son cero.

Este modelo ofrece interpretabilidad directa: los signos y magnitudes de los coeficientes permiten analizar la dirección e intensidad del efecto de cada variable sobre la probabilidad de pertenecer a una clase.

Redes neuronales En este estudio, se utilizó un modelo secuencial con dos capas ocultas, cada una con activación *ReLU*, y una capa de salida con activación *softmax* para clasificación multiclase (CN, MCI, Alzheimer).

Este enfoque es consistente con lo reportado por varios autores [3], [8], quienes destacan que los modelos de redes neuronales profundas han mostrado buenos resultados en tareas de diagnóstico

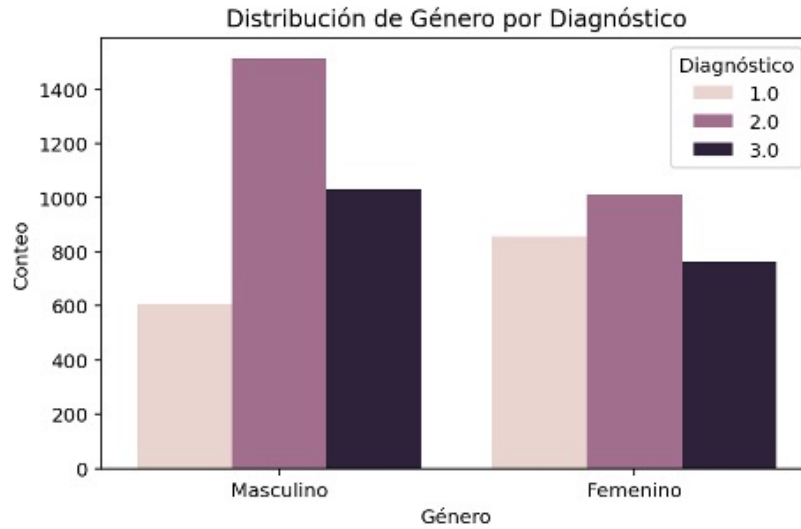


Figura 2.3: Diagnóstico clínico según sexo

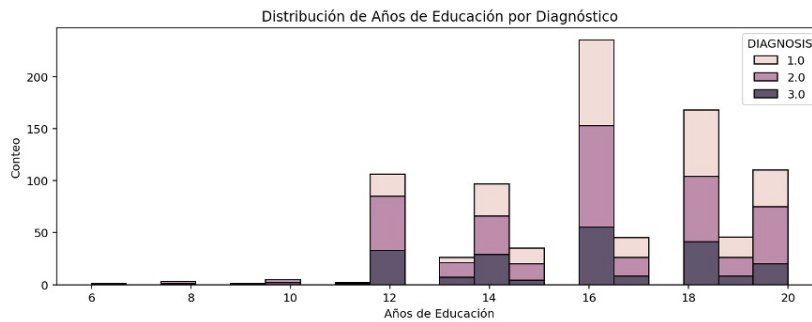


Figura 2.4: Diagnóstico clínico según nivel educativo

El modelo se entrena para minimizar la función de pérdida de entropía cruzada categórica:

$$L = - \sum_{i=1}^N \sum_{k=1}^K y_{ik} \log(\hat{y}_{ik}) \quad (2.3)$$

Donde:

- y_{ik} : Clase real para el individuo i , codificada en *one-hot encoding*.
- \hat{y}_{ik} : Probabilidad predicha por la red neuronal de que el individuo i pertenezca a la clase k .
- N : Número total de observaciones.
- K : Número total de clases.

El algoritmo de entrenamiento utilizado fue *Adam*, un optimizador

basado en gradiente descendente adaptativo, eficiente para problemas con muchos parámetros.

Además, los datos fueron normalizados previamente mediante *standard scaling*, y las variables categóricas fueron transformadas adecuadamente. Por ejemplo, la variable GENOTYPE fue dividida en Father_ApoE4 y Mother_ApoE4, mientras que DIAGNOSIS fue codificada como variable categórica multiclase.

2.4 Descripción de las métricas

Para evaluar el rendimiento de ambos modelos: Regresión Logística Multinomial y la Red Neuronal, se utilizaron métricas de clasificación multiclase. Estas métricas permiten identificar tanto el rendimiento global de los modelos como su desempeño al clasificar específicamente cada grupo: cognitivamente normal (CN), deterioro cognitivo leve (MCI) y demencia.

Exactitud (Accuracy)

Mide la proporción de predicciones correctas en relación con el total de observaciones. Es una métrica global del modelo:

$$\text{Exactitud} = \frac{\text{Numero de predicciones correctas}}{\text{Numero total de observaciones}}$$

Una alta exactitud sugiere que el modelo acierta con frecuencia en sus predicciones. Para poder utilizar esta métrica y obtener resultados relevantes, es necesario contar con clases balanceadas.

Precisión (Precision)

Evalúa la proporción de casos correctamente clasificados como positivos en relación con todos los casos que el modelo predijo como positivos:

$$\text{Precision} = \frac{\text{Verdaderos positivos}}{\text{Verdaderos positivos} + \text{Falsos positivos}}$$

Esta métrica es útil para entender qué tan confiables son las predicciones positivas. En contextos clínicos, como el presente estudio, una alta precisión implica que cuando el modelo predice una condición como deterioro cognitivo, es muy probable que sea correcta.

Sensibilidad (Recall)

La sensibilidad (también llamada *recall*) mide la proporción de verdaderos positivos identificados correctamente por el modelo. Es decir, qué tan bien el modelo detecta todos los casos reales de una clase

específica:

$$\text{Sensibilidad} = \frac{\text{Verdaderos positivos}}{\text{Verdaderos positivos} + \text{Falsos negativos}}$$

Una alta sensibilidad indica que el modelo tiene una baja tasa de falsos negativos, lo que es clave para el diagnóstico temprano, ya que permite capturar la mayor cantidad posible de casos reales.

F1-Score

Es una media combinada de precisión y sensibilidad. Permite balancear la evaluación cuando hay un *trade-off* entre estas dos métricas:

$$F_1 = 2 \cdot \frac{\text{Precisión} \cdot \text{Recall}}{\text{Precisión} + \text{Recall}}$$

Esta métrica es especialmente valiosa en problemas donde no hay un equilibrio entre clases o donde tanto los falsos positivos como los falsos negativos tienen consecuencias significativas, como en entornos médicos.

En el marco de este proyecto, si bien todas las métricas mencionadas aportan información importante, no todas tienen el mismo peso. En aspectos de salud, la sensibilidad (recall) suele ser prioritaria, ya que representa la capacidad del modelo para identificar correctamente a los pacientes que realmente presentan una condición médica. Fallar en detectar un paciente afectado (falso negativo) puede tener consecuencias clínicas graves, como retrasar el diagnóstico y el inicio de tratamiento.

También la precisión tiene su dosis de relevancia, ya que evita etiquetar erróneamente a personas sanas como enfermas. Esta confusión podría generar ansiedad y llevar a intervenciones innecesarias.

En lo que respecta al F1-score, su utilidad radica en balancear precisión y sensibilidad, siendo una métrica robusta para evaluar el modelo cuando ambas dimensiones son importantes.

2.5 Descripción de los experimentos o simulaciones

A continuación se describe en detalle la configuración de los modelos predictivos utilizados, así como las decisiones de diseño adoptadas en cada etapa del proceso. Como ya se ha mencionado previamente, el objetivo fue evaluar la capacidad de ambos modelos para clasificar correctamente a los individuos en una de tres categorías: Cognitivamente normal (CN), Deterioro Cognitivo Leve (MCI) o Demencia (AD).

Regresión logística multinomial

Fue empleada como modelo base (benchmark) por su interpretabilidad y así establecer un punto de referencia. Se optó por una versión multinomial, ya que en el problema de clasificación se incluyen más de dos clases mutuamente excluyentes. Se utilizó la función `LogisticRegression` del módulo `sklearn.linear_model`, configurada con `multi_class="multinomial"` y `solver="lbfgs"`, por su estabilidad para problemas multiclase con conjuntos de datos de tamaño moderado.

Parámetros principales:

- **Regularización:** Se mantuvo el valor por defecto $C = 1.0$, tras probar también con valores como 0.1 y 10 sin mejoras significativas.
- **Estandarización:** Se utilizó `StandardScaler` para normalizar las variables predictoras, buscando así evitar que variables con mayor magnitud dominaran el modelo.
- **División de datos:** El conjunto de entrenamiento y prueba se dividió en 80% - 20%.

Red neuronal multicapa

El segundo modelo fue una red neuronal feedforward. Esta red incluyó lo siguiente:

- Una capa de entrada con el número de nodos correspondiente al número de variables predictoras.
- Cuatro capas ocultas con 128, 64, 32 y 16 neuronas respectivamente, todas con activación ReLU.
- Una capa de salida con tres neuronas y activación softmax.
- El modelo se entrenó con la función de pérdida `categorical_crossentropy` y el optimizador Adam.

Previamente se había experimentado con arquitecturas más simples; sin embargo, no se lograba una mejora significativa en la clasificación de la clase MCI, que clínicamente es la más relevante para los objetivos de este trabajo por lo que se decidió ajustar la complejidad de los modelos, ya que según algunos estudios [2], [9], esto mejora la clasificación en tareas clínicas específicas

3 Resultados y discusión

3.1 Resultados y discusión

Regresión logística

Desempeño general: El modelo alcanzó una exactitud global del 68%. Aunque este valor es razonable para un modelo base, se consideró importante analizar el comportamiento por clase, primordialmente por el hecho de que, al estar evaluando un contexto de salud, ciertos errores pueden ser más costosos que otros. Estos resultados se muestran en la tabla 3.1

| Clase | Precisión | Recall | F1-score | Soporte |
|-----------|-----------|--------|----------|---------|
| CN | 0.69 | 0.47 | 0.56 | 291 |
| MCI | 0.63 | 0.76 | 0.69 | 523 |
| Alzheimer | 0.78 | 0.75 | 0.76 | 343 |

Tabla 3.1: Métricas por clase

Analizando a detalle, en lo relativo a la clase 1 (Cognitivamente normal): El modelo tuvo una precisión del 69% pero sólo un 47% de los casos fueron correctamente identificados como sanos (recall). Esto implica una tasa alta de falsos positivos, es decir, personas sanas clasificadas erróneamente con algún grado de deterioro.

En cuanto a la clase 2 (MCI), se obtuvo el mayor recall (76%), lo que indica una buena capacidad para detectar casos con deterioro cognitivo leve. Sin embargo, la precisión fue más baja, lo que sugiere cierta confusión con otras clases, especialmente Alzheimer.

Finalmente, en la clase 3 (Alzheimer): El modelo mostró su mejor desempeño general, con una precisión del 78% y un recall del 75%. Esto indica una alta efectividad al identificar correctamente casos con el diagnóstico de Alzheimer.

Se generó también una matriz de confusión (Figura 3.1) que reveló que el mayor número de errores se dio entre las clases MCI y Alzheimer, lo que es consistente con la superposición clínica que existe entre esas dos condiciones. Además, se observaron errores en la clasificación de

individuos sanos como MCI, lo que permite observar que el modelo tiende a sobreclasificar hacia condiciones de enfermedad.

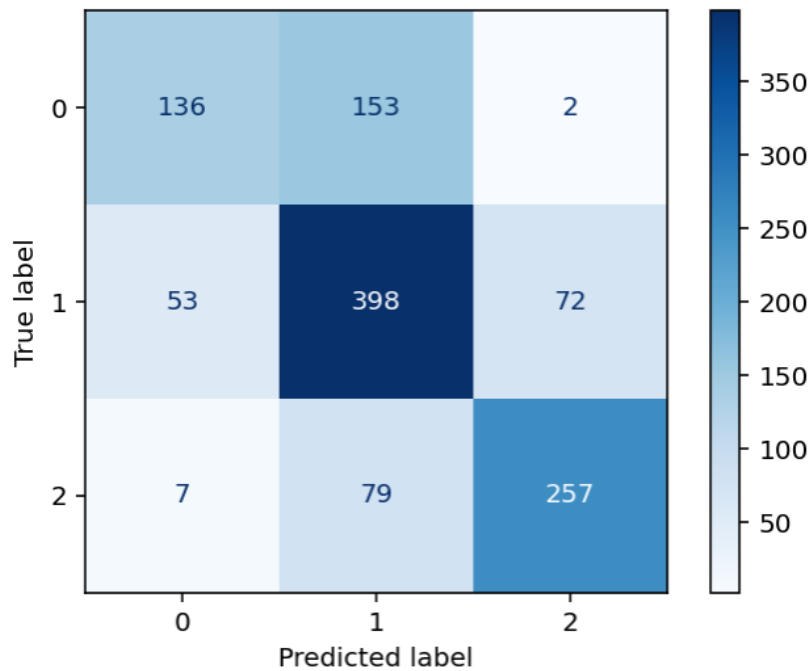


Figura 3.1: Matriz de confusión de modelo de Regresión Logística

Finalmente se generaron curvas ROC (Receiver Operating Characteristic), para analizar la capacidad del modelo para distinguir entre clases en un contexto de clasificación multiclase. Se calcula una curva ROC para cada clase, considerando un esquema de “uno contra todos”.

Como muestra la imagen 3.2, el resultado más bajo se obtuvo en la clase 2 (MCI). Esta es la clase en la que más exactitud se busca, ya que al ser la intermedia, es donde las acciones y tratamientos tienen un mayor peso, por lo que su diagnóstico acertado resulta más relevante.

Red neuronal

Durante el desarrollo del modelo basado en redes neuronales, se realizaron múltiples intentos iniciales utilizando arquitecturas más simples con una o dos capas ocultas y un número reducido de neuronas. Sin embargo, estas configuraciones ofrecieron resultados limitados, particularmente en el *recall* de la clase MCI.

De acuerdo con expertos del área clínica, esta etapa intermedia entre el deterioro cognitivo normal y el Alzheimer representa un punto crítico para una intervención temprana. Por ello, maximizar el *recall* en

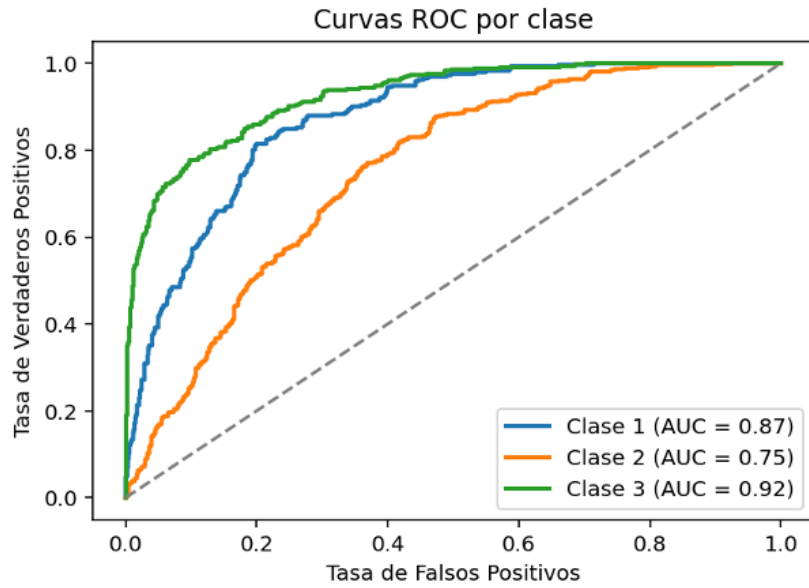


Figura 3.2: Curvas ROC para el modelo de Regresión Logística

esta categoría fue considerado prioritario, ya que permite detectar la mayor cantidad posible de casos verdaderos y actuar en consecuencia de forma oportuna.

Bajo estos criterios, se diseñó una red neuronal más profunda con las siguientes características:

- Una capa de entrada con el número de nodos correspondiente al número de variables predictoras utilizadas.
- Cuatro capas ocultas con 128, 64, 32 y 16 neuronas, respectivamente, todas activadas con la función ReLU.
- Una capa de salida con tres neuronas y activación *softmax*.
- El modelo se entrenó con la función de pérdida *categorical_crossentropy* y el optimizador *Adam*.

El modelo alcanzó una precisión general del 71 %, destacando el desempeño en la detección de casos de MCI, con un recall de 0.81, lo cual confirma que la ampliación de la red tuvo un impactopositivo en la capacidad del modelo para identificar correctamente esta clase prioritaria. También hubo buenos niveles de precisión y F1-score también para la clase AD, mientras que en la clase CN se presentó el desmepeño más bajo, sugiriendo que algunos individuos con CN fueron clasificados como MCI, un error menos costoso clínicamente que el inverso.

| Reporte de Clasificación: | | | | |
|---------------------------|-----------|--------|----------|---------|
| | precision | recall | f1-score | support |
| CN | 0.71 | 0.51 | 0.59 | 291 |
| MCI | 0.65 | 0.81 | 0.72 | 523 |
| AD | 0.86 | 0.73 | 0.79 | 343 |
| accuracy | | | 0.71 | 1157 |
| macro avg | 0.74 | 0.68 | 0.70 | 1157 |
| weighted avg | 0.73 | 0.71 | 0.71 | 1157 |

Tabla 3.2: Resultados Modelo de Red Neuronal

A continuación, se presentan dos gráficas que muestran la evolución del desempeño del modelo de red neuronal durante el proceso de entrenamiento. La primera gráfica ilustra cómo disminuyó la función de pérdida (*loss*) a lo largo de las épocas, indicando que el modelo fue ajustando sus parámetros para minimizar el error.

La segunda gráfica muestra la evolución de la precisión (*accuracy*), evidenciando cómo el modelo fue mejorando su capacidad para clasificar correctamente las observaciones. Estas curvas permiten evaluar la estabilidad y convergencia del modelo, así como detectar posibles signos de sobreajuste o subajuste.

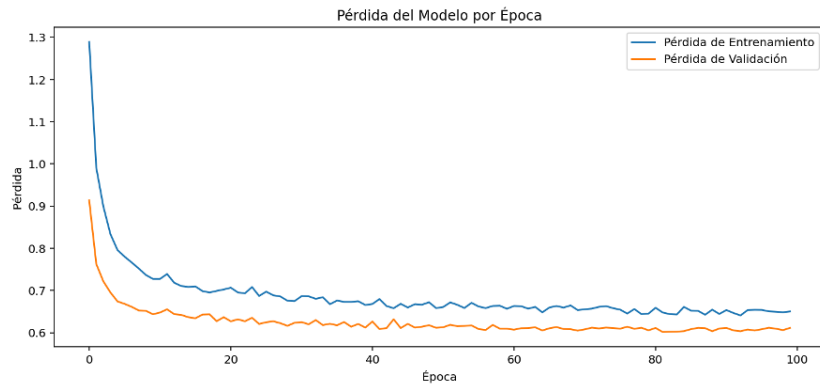


Figura 3.3: Pérdida de la Red Neuronal por época

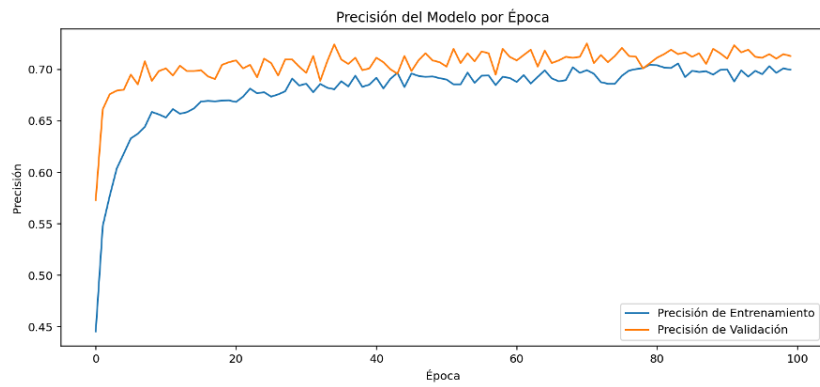


Figura 3.4: Precisión de la Red Neuronal por época

4 Conclusiones y trabajo futuro

4.1 Conclusiones

Este trabajo logró explorar el potencial de los modelos de aprendizaje supervisado para predecir el estado clínico de pacientes con signos de deterioro cognitivo, cumpliendo con los objetivos planteados. Se seleccionó y preparó un conjunto de datos clínicos estructurados, incluyendo variables demográficas y cognitivas relevantes, lo que permitió desarrollar modelos predictivos robustos.

Se implementaron dos enfoques: regresión logística como modelo de referencia y una red neuronal como alternativa. Si bien la regresión logística obtuvo resultados aceptables, la red neuronal mostró un mejor desempeño global, especialmente en la detección de la etapa MCI (recall = 81 %), considerada crítica para la intervención temprana.

La comparación entre ambos modelos permitió identificar fortalezas y limitaciones concretas. La red neuronal superó a la regresión en precisión y recall para la mayoría de las clases, aunque presentó dificultades en la identificación de sujetos con diagnóstico CN.

4.2 Trabajo futuro

Los resultados obtenidos sugieren la necesidad de explorar el uso de variables adicionales, como biomarcadores o imágenes médicas, para mejorar la discriminación entre clases en futuras investigaciones.

Asimismo, se sugiere investigar con modelos de distintas características a los explorados en este trabajo,

En resumen, este estudio demuestra el potencial de las redes neuronales para apoyar el diagnóstico temprano del deterioro cognitivo, pero también resalta la necesidad de seguir refinando estos modelos para lograr una mayor precisión y confiabilidad en entornos reales.

Bibliografía

- [1] World Health Organization (WHO), "Global status report on the public health response to dementia," 2021.
- [2] E. E. Bron, M. Smits, W. M. van der Flier, H. Vrenken, F. Barkhof, P. Scheltens, and B. van Ginneken, "Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural mri: The caddementia challenge," *NeuroImage*, vol. 111, pp. 562–579, 2015.
- [3] T. Pereira, C. Marques, F. Ferreira, F. M. Couto, and M. Costa, "Machine learning approaches to detect alzheimer's disease using structural mri: A review," *Current Alzheimer Research*, vol. 15, no. 2, pp. 114–142, 2018.
- [4] Instituto Nacional de Estadística y Geografía (INEGI), "Estadísticas a propósito del día mundial del alzheimer (21 de septiembre)," 2022.
- [5] Fundación Alzheimer México A.C., "Panorama de la enfermedad en méxico," 2023.
- [6] Secretaría de Salud (México), "Programa de acción específico para la atención de personas adultas mayores 2021–2024," 2021.
- [7] Organización Panamericana de la Salud / Organización Mundial de la Salud (OPS/OMS), "Plan de acción mundial sobre la respuesta de salud pública a la demencia 2017–2025," 2022.
- [8] S. Vieira, W. H. L. Pinaya, and A. Mechelli, "Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications," *Neuroscience & Biobehavioral Reviews*, vol. 74, pp. 58–75, 2017.
- [9] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. Setio, F. Ciompi, M. Ghafoorian, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.