

Práctica 2 – Arquitectura DNN

Integrantes: Gonzalo Del Priore, Noelia Blanco, María Eugenia Puchkariov

Fecha de entrega: 15-08-2025

Introducción

La predicción de enfermedades a partir de información clínica constituye un desafío relevante en el ámbito de la salud. En este trabajo se aborda el problema de detección de diabetes utilizando el conocido dataset de Pima Indians Diabetes, que contiene variables clínicas de pacientes y una etiqueta binaria que indica la presencia o ausencia de la enfermedad. En la práctica anterior se implementó un Perceptrón Multicapa (MLP) como modelo base, el cual ofreció un rendimiento aceptable pero con limitaciones en la capacidad de detección de casos positivos. En esta nueva práctica se busca superar dicho modelo implementando una Red Neuronal Densa (DNN), más adecuada para datos tabulares y con la posibilidad de modelar relaciones no lineales entre variables, mejorando la capacidad predictiva.

Selección y Justificación de la Arquitectura

El dataset en cuestión está compuesto por variables tabulares independientes, sin estructura espacial ni temporal. Esto descarta el uso de redes convolucionales (CNN), más adecuadas para imágenes, y de redes recurrentes (RNN) o Transformers, enfocadas en secuencias. La DNN surge como la opción más apropiada al poder manejar variables continuas y categóricas, capturar relaciones complejas entre ellas y adaptarse a un número limitado de registros sin sobrecargar el entrenamiento. Además, permite introducir mecanismos de regularización como Batch Normalization y Dropout, que estabilizan y reducen el sobreajuste. En consecuencia, la elección de una DNN responde tanto a criterios técnicos como al objetivo del problema clínico planteado.

Preparación del Dataset

El dataset original contiene 768 registros y 8 variables predictoras, junto con la variable objetivo (Outcome). Durante la preparación se realizaron los siguientes pasos: 1) Reemplazo de ceros fisiológicamente imposibles en variables como Glucose, BloodPressure, SkinThickness, Insulin y BMI. 2) Imputación de los valores faltantes utilizando la mediana, lo que permitió conservar todos los registros sin introducir sesgos. 3) Estandarización de las variables mediante StandardScaler, para garantizar estabilidad numérica y eficiencia en el entrenamiento. 4) División estratificada en entrenamiento (80%) y prueba (20%), manteniendo la proporción de casos positivos (~35%) en ambas

particiones. Este proceso asegura que los datos estén limpios, comparables y en condiciones adecuadas para entrenar un modelo robusto.

Implementación y Entrenamiento del Modelo

El modelo implementado es una DNN de tres capas ocultas con 64, 32 y 16 neuronas respectivamente. Cada capa está seguida de Batch Normalization y Dropout, lo que aporta estabilidad y control del sobreajuste. La función de activación seleccionada fue ReLU en las capas ocultas y sigmoide en la capa de salida, apropiada para problemas binarios. Se utilizó el optimizador Adam con tasa de aprendizaje inicial de 0.001 y función de pérdida binary crossentropy. El entrenamiento incluyó callbacks como EarlyStopping, que detiene el proceso al no mejorar la validación, y ReduceLROnPlateau, que reduce la tasa de aprendizaje automáticamente. Estos mecanismos mejoran la eficiencia del proceso de ajuste, evitando sobreentrenamiento y logrando convergencia más estable. El entrenamiento se llevó a cabo durante un máximo de 150 épocas, con lotes de 32 observaciones, alcanzando un balance entre precisión de actualización y eficiencia computacional.

Evaluación del Modelo

El modelo fue evaluado en el conjunto de prueba con los siguientes resultados: - Accuracy: 0.73 - Precision (PPV): 0.62 - Recall (Sensibilidad): 0.59 - F1-score: 0.61 - Especificidad: 0.79 - AUC-ROC: 0.82 Estos valores reflejan una mejora respecto al MLP base, especialmente en precisión y especificidad. La matriz de confusión mostró 33 verdaderos positivos y 79 verdaderos negativos, con 21 falsos negativos y 21 falsos positivos, lo que confirma la necesidad de ajustar el umbral de decisión para priorizar la sensibilidad. Se evaluaron métricas clínicas como el F2-score y el F5-score, que otorgan mayor peso al recall, para demostrar la importancia de minimizar falsos negativos en diagnósticos médicos.

Comparación con MLP

El Perceptrón Multicapa de la práctica 1 obtuvo un accuracy de 0.72 y un F1 de 0.59. La DNN logró mejorar en precisión positiva (0.63 vs. 0.60) y en especificidad (0.85 vs. 0.80), aunque en el umbral estándar mostró un recall algo menor. Sin embargo, al ajustar el umbral a 0.155, la DNN alcanzó un recall de 0.93 con solo 4 falsos negativos, lo cual representa un gran avance en términos clínicos. Esto refleja que la DNN no solo mejora el rendimiento global, sino que también ofrece flexibilidad para adaptarse a criterios clínicos distintos mediante la selección adecuada de métricas y umbrales.

Visualizaciones

Se generaron gráficos de la evolución del accuracy y la pérdida en entrenamiento y validación, donde se observó un comportamiento estable y sin sobreajuste pronunciado. También se presentaron matrices de confusión comparativas para distintos umbrales de decisión, mostrando cómo varía el balance entre sensibilidad y especificidad. Finalmente, el gráfico de F2 y F5 en función del umbral permitió visualizar la relación entre recall y precisión, destacando que en entornos clínicos es preferible priorizar la sensibilidad, aun cuando aumenten los falsos positivos.

Conclusiones

El estudio realizado demuestra que la Red Neuronal Densa es una alternativa más robusta que el MLP base para este problema, logrando un mejor rendimiento en la mayoría de las métricas y mayor capacidad de generalización. La principal ventaja se encuentra en la posibilidad de ajustar el umbral de decisión, lo que permite priorizar la detección de casos positivos y reducir significativamente los falsos negativos, un aspecto clave en contextos médicos. Como limitaciones, se identificó la persistencia de falsos positivos en configuraciones con alta sensibilidad. Esto podría ser abordado en trabajos futuros aplicando técnicas de balanceo de clases como SMOTE, ajuste de pesos por clase o mediante una búsqueda más exhaustiva de hiperparámetros. En conclusión, la DNN no solo supera al modelo base en rendimiento global, sino que también ofrece la flexibilidad necesaria para adaptarse a criterios clínicos específicos, constituyendo una herramienta eficaz para la predicción de diabetes a partir de datos tabulares.