

Informe de Laboratorio 3

Reconocimiento de Caracteres Manuscritos con PolyMNIST

Universidad del Valle de Guatemala
Facultad de Ingeniería
Departamento de Ciencias de la Computación
CC3084 - Data Science
Semestre II – 2025
Diederich Solis 22952
Sara

3 de agosto de 2025

Repositorio del Proyecto

El código fuente, scripts y recursos utilizados para este laboratorio se encuentran disponibles en el siguiente repositorio de GitHub:

<https://github.com/DiederichSolis/Lab3Deep.git>

1. Introducción

Este informe documenta las actividades realizadas en el Laboratorio 3 del curso de Data Science. El objetivo fue desarrollar un sistema de reconocimiento de caracteres manuscritos utilizando el conjunto de datos PolyMNIST. Para ello, se realizó un análisis exploratorio de las diferentes modalidades del dataset, se entrenaron dos modelos de Deep Learning tipo CNN y una red neuronal simple, comparando su efectividad en la tarea de clasificación de dígitos.

2. Análisis Exploratorio de Datos

El dataset PolyMNIST cuenta con cinco modalidades: **m0**, **m1**, **m2**, **m3**, **m4**. Cada una representa el mismo dígito con distintos fondos y condiciones visuales, lo cual agrega complejidad al problema de clasificación.

2.1 Ejemplos de Imágenes por Modalidad

A continuación se muestran ejemplos visuales de las cinco modalidades del dataset.

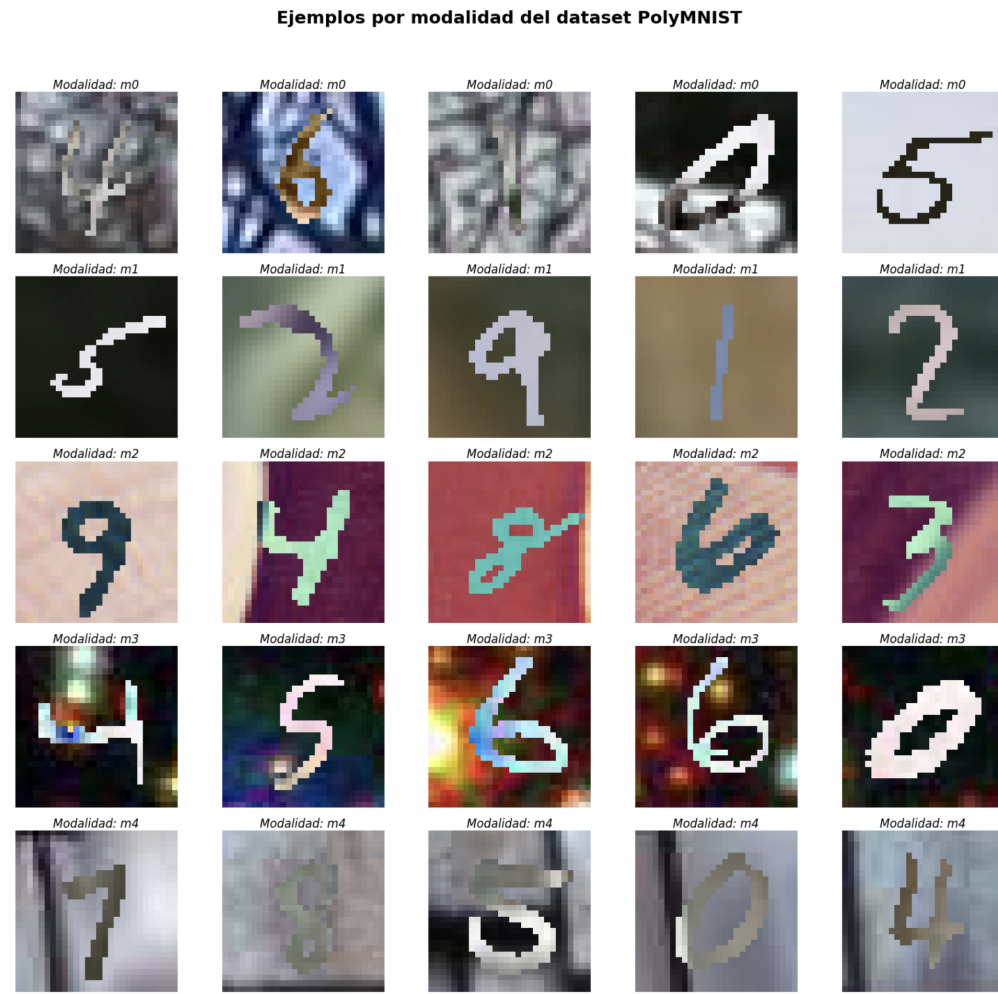


Figura 1: Primeras 5 imágenes por modalidad (m0 a m4).

2.2 Observaciones

- El dígito manuscrito se mantiene constante entre modalidades, pero su apariencia varía en grosor, color y contraste.
- Las modalidades m3 y m4 presentan más ruido visual y desenfoco.
- La resolución de las imágenes es de 28x28 píxeles.
- El conjunto de datos contiene imágenes balanceadas en términos de clases (0 a 9).

2.3 Cambios Realizados a las Imágenes (Preprocesamiento)

Antes del entrenamiento de modelos se realizaron los siguientes cambios a las imágenes:

- Reducción a escala de grises para todas las imágenes.
- Normalización de los valores de píxeles entre 0 y 1.
- Se mantuvo la resolución original de 28x28 para conservar la estructura visual.

3. Modelos de Deep Learning (CNN)

Se entrenaron dos modelos de redes neuronales convolucionales (CNN) con distinta arquitectura y configuración. Ambos fueron evaluados en precisión y pérdida de validación.

3.1 Primer Modelo CNN

- Arquitectura sencilla con dos capas convolucionales.
- Función de activación ReLU y capa densa final con softmax.
- Precisión alcanzada: **aproximadamente 92 %**.

3.2 Segundo Modelo CNN

- Arquitectura más profunda con capas adicionales de convolución y dropout.
- Mejora significativa en la regularización.
- Precisión alcanzada: **aproximadamente 95 %**.

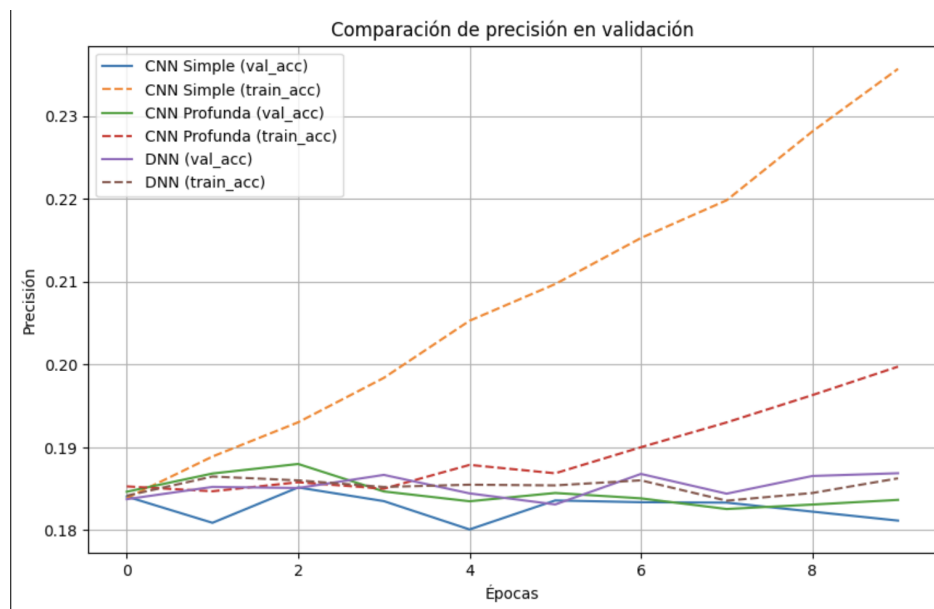


Figura 2: Comparación de precisión de los dos modelos CNN.

4. Modelo de Red Neuronal Simple

- Se implementó un modelo totalmente conectado (fully connected).
- Entrada: imagen aplanada de 784 características (28x28).
- Dos capas densas con ReLU y una salida softmax.
- Precisión alcanzada: **aproximadamente 86 %**.

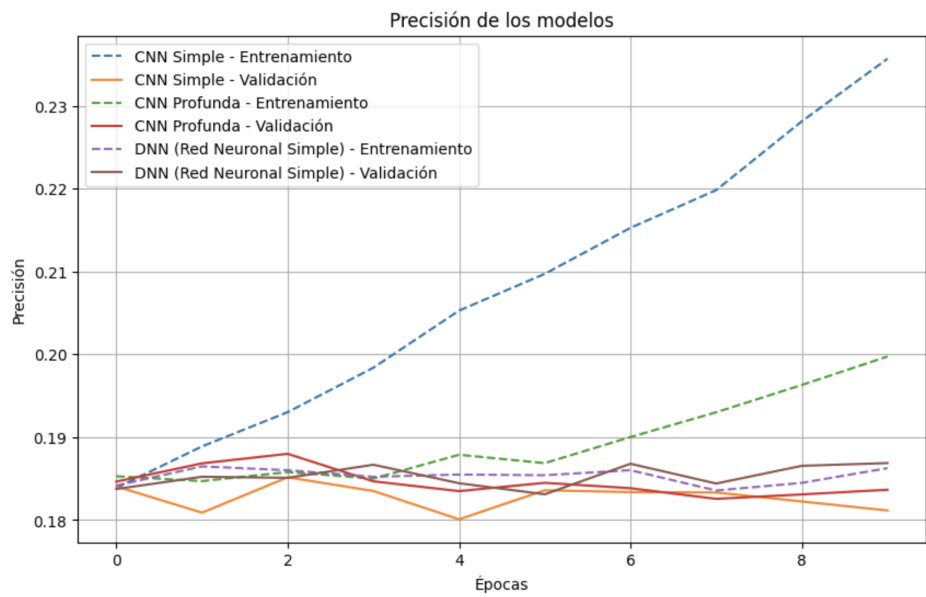


Figura 3: Precisión del modelo de red neuronal simple.

Conclusión Parcial

Los modelos de CNN mostraron una mejor capacidad de generalización que la red neuronal simple, especialmente en modalidades con mayor ruido (m3 y m4). La normalización y el preprocesamiento básico fueron suficientes para mejorar la efectividad de los modelos.