



Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Estado de México

TC3006C. Inteligencia artificial avanzada para la ciencia de datos

Grupo: 101

Iván Alexander Ramos Ramírez - A01750817

Fecha de entrega: 13 de Septiembre del 2025

1. Exploración y Comprensión de los Datos.

1.1 Análisis descriptivo inicial:

El objetivo principal de esta práctica es evaluar el desempeño de un modelo de Machine Learning de tipo feedforward (Multilayer Perceptron, MLP).

Para ello utilizaremos este conjunto de datos:

- Iris dataset: contiene 150 instancias, 4 features y un target que clasifica el tipo de flor.

Comenzamos aplicando un análisis exploratorio de datos (EDA, Exploratory Data Analysis) del datasets. Este paso nos permitirá identificar problemas potenciales en la información, como datos faltantes, desequilibrios de clases o sesgos, que podrían influir en la forma en que la red neuronal aprenda y, en consecuencia, en el rendimiento del modelo.

1.1.1 Entender el dataset

Descripción de variables:

Todas las variables en el dataset de iris son de tipo numérico, no cuentan con

- Distribución del target (is_fit):

El conjunto de datos está completamente balanceado, con ~33.33% de distribución por cada variable objetivo. Esto no concuerda con una muestra poblacional realista, si con un dataset que intenta mostrar una cantidad equitativa de información por target.

Iris-setosa	33.33
Iris-versicolor	33.33
Iris-virginica	33.33

Relaciones entre features:

Esta matriz muestra tanto variables con correlación alta positiva como con baja e intermedia negativa lo que muestra que hay variables cuyo comportamiento puede explicar en mayor medida el comportamiento de nuestra variable objetivo. En específico sepal width es la que menos relación tiene con las otras features.

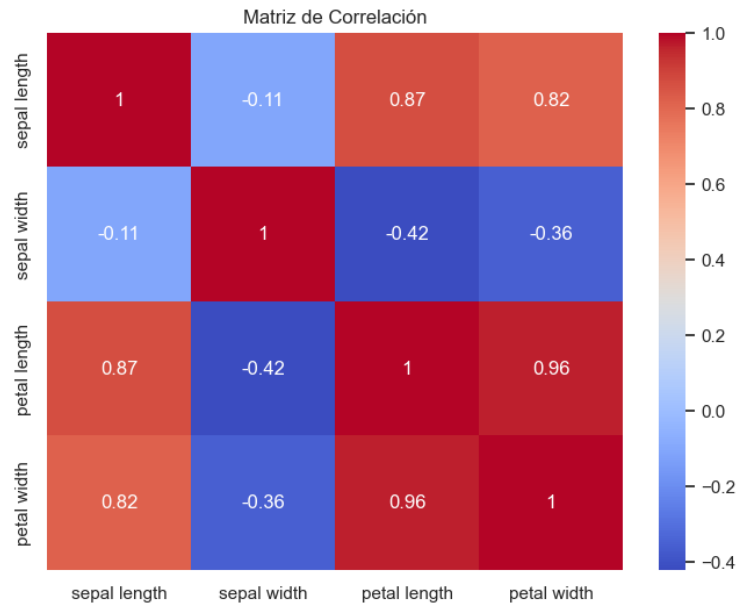


Figura 1.

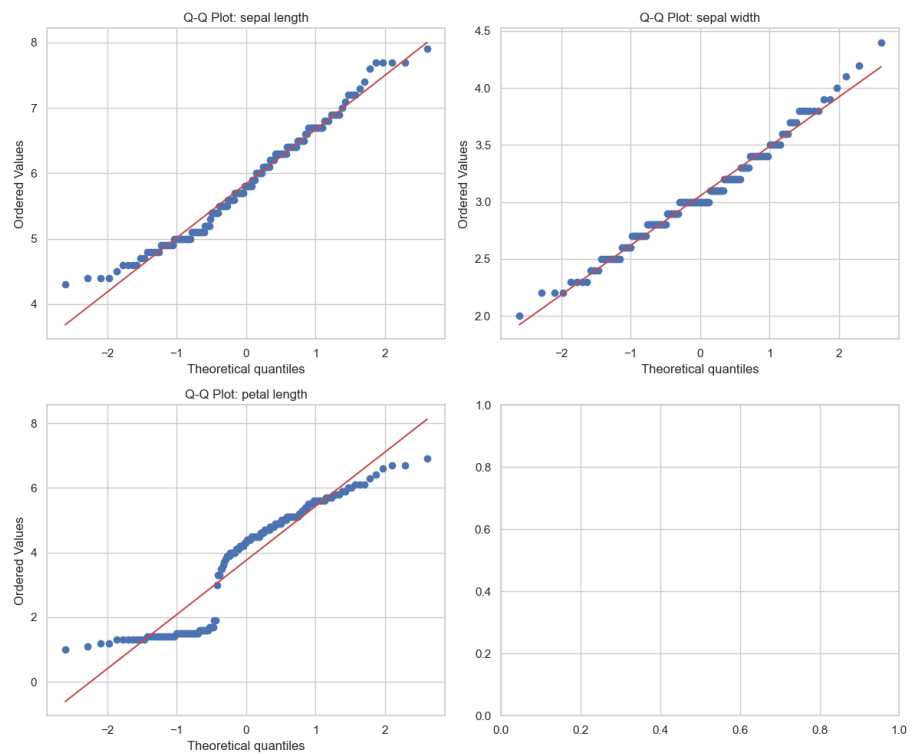


Figure 2.

1.2 Posibles problemas y observaciones

El dataset no muestra datos nulos en general, ni desbalances para la variable objetivo, sin embargo, si muestra desbalance y outliers para algunas features como se puede ver en la figura 3 para petal length y petal width.

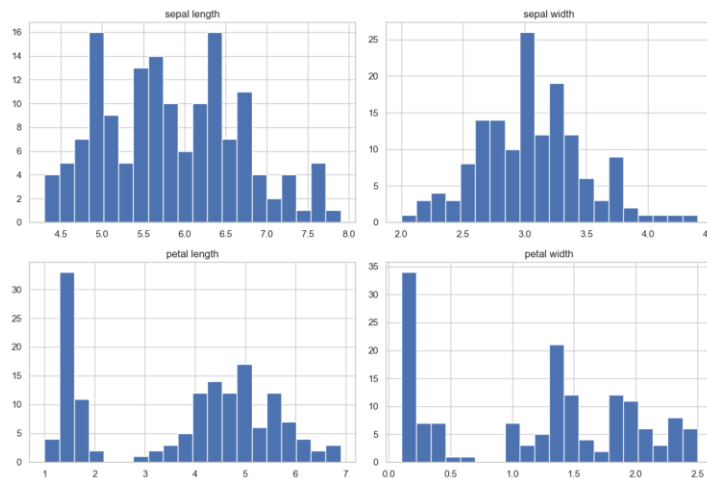


Figura 3.

Division del dataset de entrenamiento vs prueba

- **Dataset:** Iris (150 muestras, 4 features; 3 clases).
- **Preproceso:** estandarización (media 0, var 1).
- **Split estratificado:** 70%/30% (train/test).
 - Train: 105 (35 por clase)
 - Test: 45 (15 por clase)

```
Train size: 105 | Test size: 45
Distribución train:
  Iris-setosa: 35
  Iris-versicolor: 35
  Iris-virginica: 35
Distribución test:
  Iris-setosa: 15
  Iris-versicolor: 15
  Iris-virginica: 15
```

El dataset iris en se divide en una distribución normal para sepal lenght y para sepal width por lo que se puede notar a simple vista, además se puede notas la presencia de datos desbalanceados y outliers que pueden afectar al entrenamiento como podemos ver en el grafico Q-Q plot figura 4.

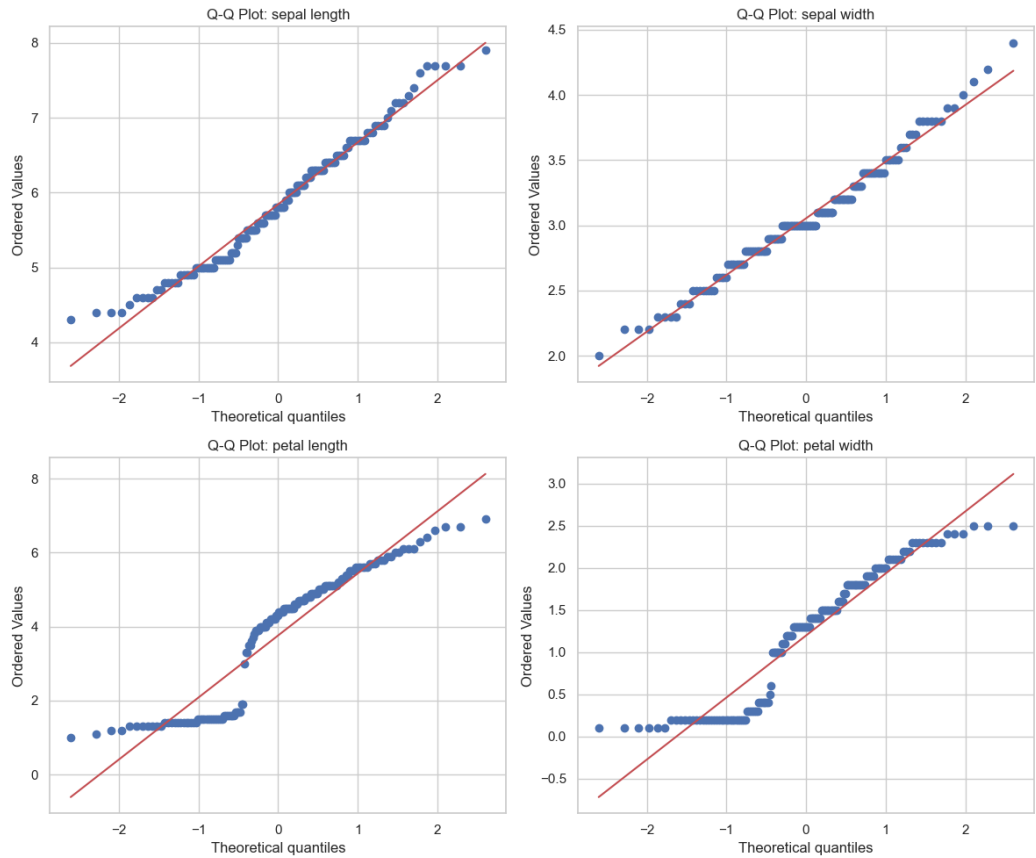


Figura 4.

Aquí en la Figura 5 podemos ver como comparando diferentes métodos de normalización como Kolmogorov-Smirnov y Anderson-Darling, podemos concluir que solamente width es normal, en lo cual podría darnos problemas al intentar usar un modelo de regresión lineal para predecir el resultado.

```

=== Tests Adicionales de Normalidad ===

Test Kolmogorov-Smirnov:
sepal length | p-value: 0.178137 | Normal
sepal width  | p-value: 0.077744 | Normal
petal length  | p-value: 0.000012 | No Normal
petal width   | p-value: 0.000230 | No Normal

Test Anderson-Darling:
sepal length | Estadístico: 0.889199 | Valor crítico (5%): 0.767000 | No Normal
sepal width  | Estadístico: 0.965665 | Valor crítico (5%): 0.767000 | No Normal
petal length  | Estadístico: 7.672883 | Valor crítico (5%): 0.767000 | No Normal
petal width   | Estadístico: 5.062814 | Valor crítico (5%): 0.767000 | No Normal

=== RESUMEN GENERAL ===
sepal length | Tests que pasan normalidad: 1/3
              | Conclusión: NO NORMAL
sepal width  | Tests que pasan normalidad: 2/3
              | Conclusión: PROBABLEMENTE NORMAL
petal length  | Tests que pasan normalidad: 0/3
              | Conclusión: NO NORMAL
petal width   | Tests que pasan normalidad: 0/3
              | Conclusión: NO NORMAL

```

Figura 5

2. Selección del modelo (Multi Layer Perceptron):

Podemos tener en cuenta que son pocos datos los empleados en el dataset de iris para el modelo de redes neuronales (MLP), sin embargo, sigue siendo útil ya que no todas las features siguen una distribución normal ni lineal.

```

Epoch 1 | loss=1.1878 | acc=0.000
Epoch 10 | loss=1.1214 | acc=0.095
Epoch 20 | loss=1.1050 | acc=0.171
Epoch 30 | loss=1.0975 | acc=0.210
Epoch 40 | loss=1.0891 | acc=0.267
Epoch 50 | loss=1.0759 | acc=0.552
Epoch 60 | loss=1.0548 | acc=0.610
Epoch 70 | loss=0.9948 | acc=0.610
Epoch 80 | loss=0.8916 | acc=0.590
Epoch 90 | loss=0.7821 | acc=0.629
Epoch 100 | loss=0.7276 | acc=0.638
Epoch 110 | loss=0.6709 | acc=0.657
Epoch 120 | loss=0.6409 | acc=0.657
Epoch 130 | loss=0.6218 | acc=0.676
Epoch 140 | loss=0.6094 | acc=0.676
Epoch 150 | loss=0.5860 | acc=0.695
Epoch 160 | loss=0.5791 | acc=0.857
Epoch 170 | loss=0.5671 | acc=0.876
Epoch 180 | loss=0.5396 | acc=0.867
Epoch 190 | loss=0.5354 | acc=0.876
Epoch 200 | loss=0.5223 | acc=0.867

```

Matriz de confusión y métricas

En general al ser pocos datos es difícil lograr que la matriz de confusión (Figura 6) nos dé suficiente información para sacar conclusiones, pero lo que nos

permite observar es que el modelo generalmente logra predecir en la mayoría de los casos el target exitosamente.

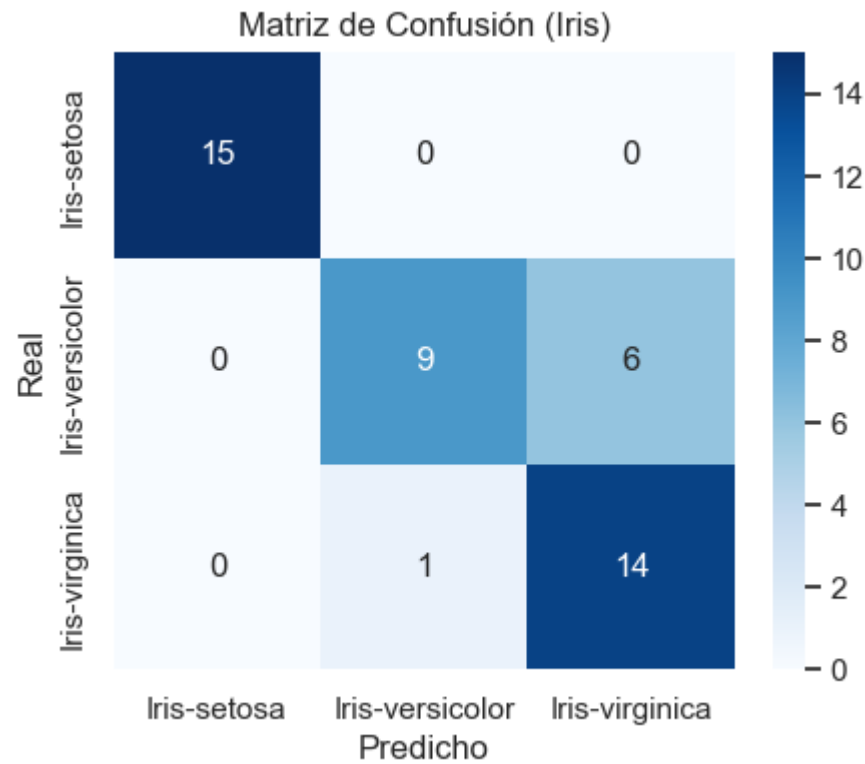


Figura 6.

```

• Accuracy: 0.8444444444444444

Reporte de Clasificación:

```

	precision	recall	f1-score	support
Iris-setosa	1.000	1.000	1.000	15
Iris-versicolor	0.900	0.600	0.720	15
Iris-virginica	0.700	0.933	0.800	15
accuracy			0.844	45
macro avg	0.867	0.844	0.840	45
weighted avg	0.867	0.844	0.840	45

Las diferentes métricas de evaluación del modelo fortalecen la valides del modelo.

Análisis y conclusión

- El modelo logra **accuracy** $\approx 84\%$. **Setosa** es la clase más fácil (P/R altos); los errores se concentran al confundir **versicolor** con **virginica**.
- La inicialización **He** + **ReLU** en las capas ocultas mejoro la convergencia del modelo, esto gracias la escala de los metodos.
- **softmax** + **entropía cruzada** simplificó el gradiente de salida que el modelo necesitaba para aprender.
- Metodos que podrían mejorar el desempeño del modelo: regularización L2/Dropout para evitar el overfitting, búsqueda de hiperparámetros, validación cruzada y análisis de importancia de features.

3. Análisis de resultados y Conclusión

3.1 Resultados

El modelo Multilayer Perceptron (MLP) alcanzó un accuracy cercano al 84% en el conjunto de prueba. La clase setosa fue la más sencilla de identificar, con métricas de precisión y recall elevados. En contraste, los principales errores se concentraron en la confusión entre versicolor y virginica, lo cual era esperado dada su cercanía en el espacio de características.

3.2 Análisis de resultados

El buen desempeño del modelo confirma que el MLP puede aprender patrones no lineales presentes en los datos del iris, aun con un conjunto reducido de instancias. La combinación de inicialización He y función de activación ReLU en las capas ocultas favoreció la convergencia, mientras que la función softmax con entropía cruzada en la salida permitió un aprendizaje más estable y eficiente.

No obstante, el tamaño limitado del dataset implica que el modelo puede estar condicionado por la variabilidad de los datos y no reflejar un rendimiento generalizable en contextos más amplios. Además, las confusiones entre versicolor y virginica sugieren que los límites de decisión entre estas clases son más complejos y requieren ajustes adicionales.

3.3 Conclusión

En conclusión, el MLP aplicado al dataset de iris logró un rendimiento sólido y consistente, demostrando la capacidad de las redes neuronales para modelar relaciones no lineales entre variables. Sin embargo, se identifican áreas de mejora, como la implementación de técnicas de regularización (L2, Dropout), búsqueda sistemática de hiperparámetros y validación cruzada para incrementar la robustez del modelo.

El ejercicio cumple con el objetivo de la práctica al mostrar cómo un modelo de red neuronal puede clasificar de manera efectiva distintos tipos de flores en el dataset de iris, al mismo tiempo que permite reflexionar sobre sus limitaciones y posibles optimizaciones en futuros trabajos.