

## **Deep Learning**

### Laboratorio 2

---

### **Objetivo de la red**

El objetivo de la red fue reducir el índice de aprobación de las películas (approval\_Index) utilizando características como el año y mes de producción, duración, promedio de calificación, número de votos, presupuesto de producción, ingresos domésticos e internacionales, y géneros.

Las redes neuronales son capaces de capturar relaciones no lineales complejas entre las características de entrada y la variable objetivo, algo que modelos lineales no pueden hacer de manera efectiva. Además, las redes neuronales pueden ajustarse y adaptarse a diferentes tipos de datos y patrones, proporcionando una gran flexibilidad en su implementación.

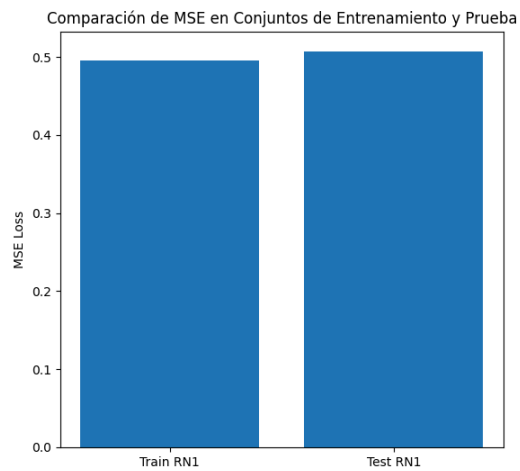
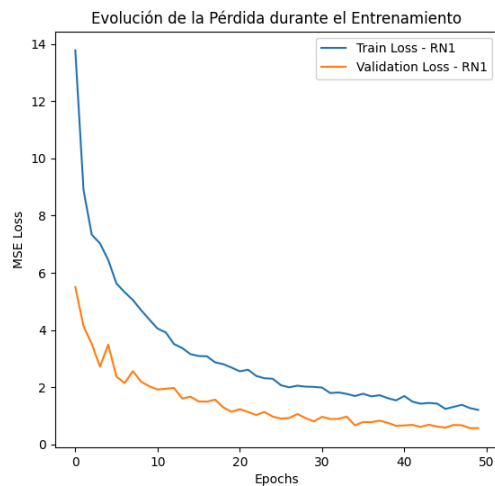
En muchos casos, las redes neuronales pueden ofrecer un rendimiento superior en comparación con otros modelos de aprendizaje automático, especialmente cuando se dispone de un conjunto de datos suficientemente grande y diverso.

### **Composición y resultados**

#### **Red neuronal 1:**

La red neuronal cuenta con tres capas densas con 64, 32 y 16 neuronas, respectivamente, en donde su función de activación es ReLU en todas las capas ocultas. ReLU es conocida por su eficiencia en entrenamiento y su capacidad para mitigar el problema del gradiente desvaneciente. Esta red neuronal se regularizó con el método de dropout con una tasa de 0.5 para prevenir el sobreajuste al apagar aleatoriamente neuronas durante el entrenamiento.

La función de pérdida utilizada fue la del MSE y se puede visualizar con la siguiente gráfica:



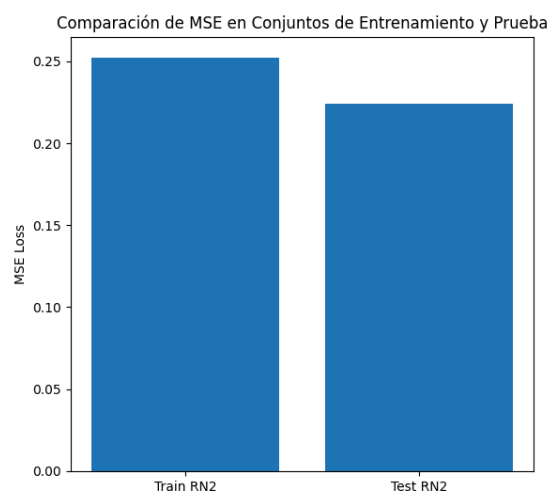
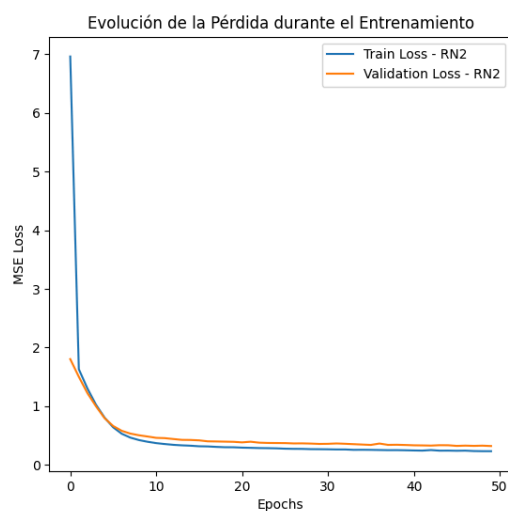
En donde la función de pérdida fue la siguiente:

- MSE en conjunto de entrenamiento: 0.496
- MSE en conjunto de prueba: 0.507

## Red neuronal 2:

La red neuronal cuenta con dos capas densas con 128 y 64 neuronas, en donde su función de activación es Sigmoid en todas las capas ocultas. Aunque Sigmoid puede sufrir saturación en las neuronas, es efectiva en ciertos contextos de clasificación binaria y permite el modelado de relaciones complejas. Esta red neuronal se regularizó con el método de L2 con una tasa de 0.001 para reducir el sobreajuste penalizando grandes coeficientes.

La función de pérdida utilizada fue la del MSE y se puede visualizar con la siguiente gráfica:



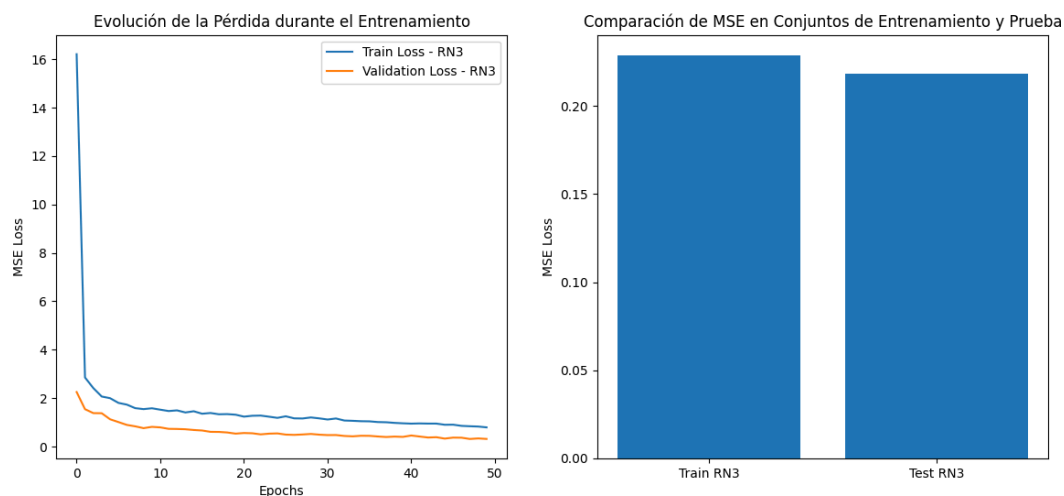
En donde la función de pérdida fue la siguiente:

- MSE en conjunto de entrenamiento: 0.251
- MSE en conjunto de prueba: 0.224

### Red neuronal 3:

La red neuronal cuenta con cuatro capas densas con 128, 64, 32 y 16 neuronas, en donde la función de activación utilizada es Tanh en todas las capas ocultas. Tanh puede modelar relaciones no lineales más efectivamente que Sigmoid, ya que su salida está centrada en cero, lo que ayuda con la convergencia durante el entrenamiento. La regularización utilizada es una combinación de Dropout con una tasa de 0.3 y L2 con una tasa de 0.0001, proporcionando una robusta prevención del sobreajuste.

La función de pérdida utilizada fue la del MSE y se puede visualizar con la siguiente gráfica:



En donde la función de pérdida fue la siguiente:

- MSE en conjunto de entrenamiento: 0.229
- MSE en conjunto de prueba: 0.219

## Discusión

La Red Neuronal 1, con una estructura de tres capas y funciones de activación ReLU, mostró un mayor MSE en comparación con las otras redes, indicando una menor precisión. La Red Neuronal 2, con funciones de activación Sigmoid y regularización L2, mejoró significativamente, pero mostró una ligera tendencia al sobreajuste. Finalmente, la Red Neuronal 3, que combina funciones de activación Tanh con Dropout y L2, obtuvo el menor MSE en el conjunto de prueba y una pequeña diferencia entre el MSE de entrenamiento y de prueba, lo que sugiere una excelente capacidad de generalización y precisión, haciendo de ella la opción más robusta y efectiva para predecir el índice de aprobación de las películas.

## Selección de red neuronal óptima

La Red Neuronal 3 es seleccionada como el modelo óptimo debido a su superior rendimiento, con el menor MSE en el conjunto de prueba (0.219), y su pequeña diferencia entre el MSE de entrenamiento (0.229) y de prueba, lo que sugiere una excelente capacidad de generalización sin sobreajuste. La combinación de funciones de activación Tanh, Dropout y L2 regularización permite a esta red capturar relaciones complejas en los datos de manera robusta y flexible, proporcionando una mayor precisión y fiabilidad en la predicción del índice de aprobación de las películas.

Red Neuronal 1 no fue seleccionada debido a su mayor MSE en comparación con las otras redes, lo que indica una menor precisión en sus predicciones. A pesar de utilizar ReLU y Dropout para prevenir el sobreajuste, su rendimiento general fue inferior.

Red Neuronal 2, aunque mostró una mejora significativa respecto a la Red Neuronal 1 con un menor MSE, todavía presentó una diferencia considerable entre el MSE de entrenamiento y de prueba, sugiriendo un posible sobreajuste el cual se puede visualizar en los resultados presentados anteriormente. La función de activación Sigmoid y la regularización L2 no fueron suficientes para superar el rendimiento de la Red Neuronal 3.