

Laboratorio #2

<https://github.com/andresquez/DeLe-Lab2>

Objetivo de la Red:

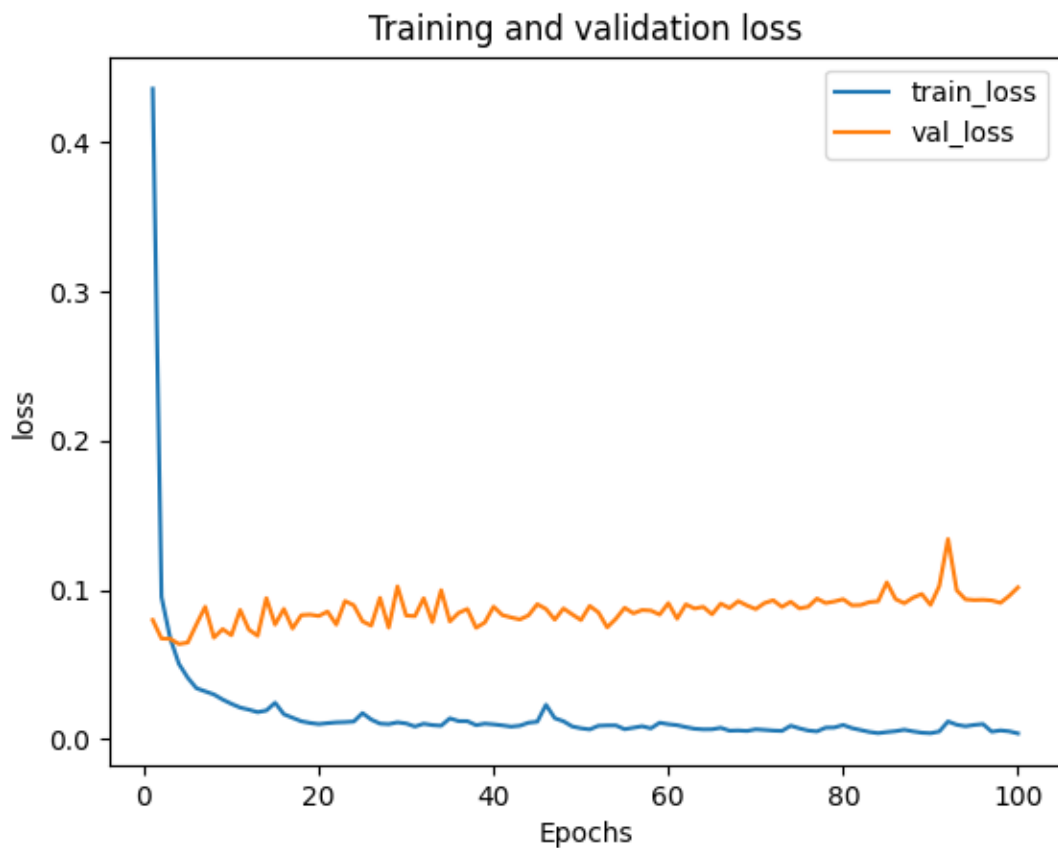
El objetivo de la red neuronal es predecir la recaudación mundial de películas (Worldwide gross \$) en función de diversas características de las películas, como la duración, el presupuesto de producción, las calificaciones y el número de votos. La elección de utilizar una red neuronal es por la capacidad de estos modelos para capturar relaciones no lineales complejas entre las variables de entrada y la variable objetivo. La red neuronal puede manejar múltiples características de entrada y aprender patrones complejos a partir de datos grandes, lo que la hace adecuada para tareas de regresión en conjuntos de datos con múltiples características. En este caso tenemos variables con valores diversos y varias características lo cual hace que la implementación sea un buen caso de uso para las redes neuronales.

Composición y Resultados de cada Red Neuronal:

Todas fueron probadas con un estándar de 100 épocas, una distribución de entrenamiento vs prueba de 80-20 y también 80-20 para el set de validación.

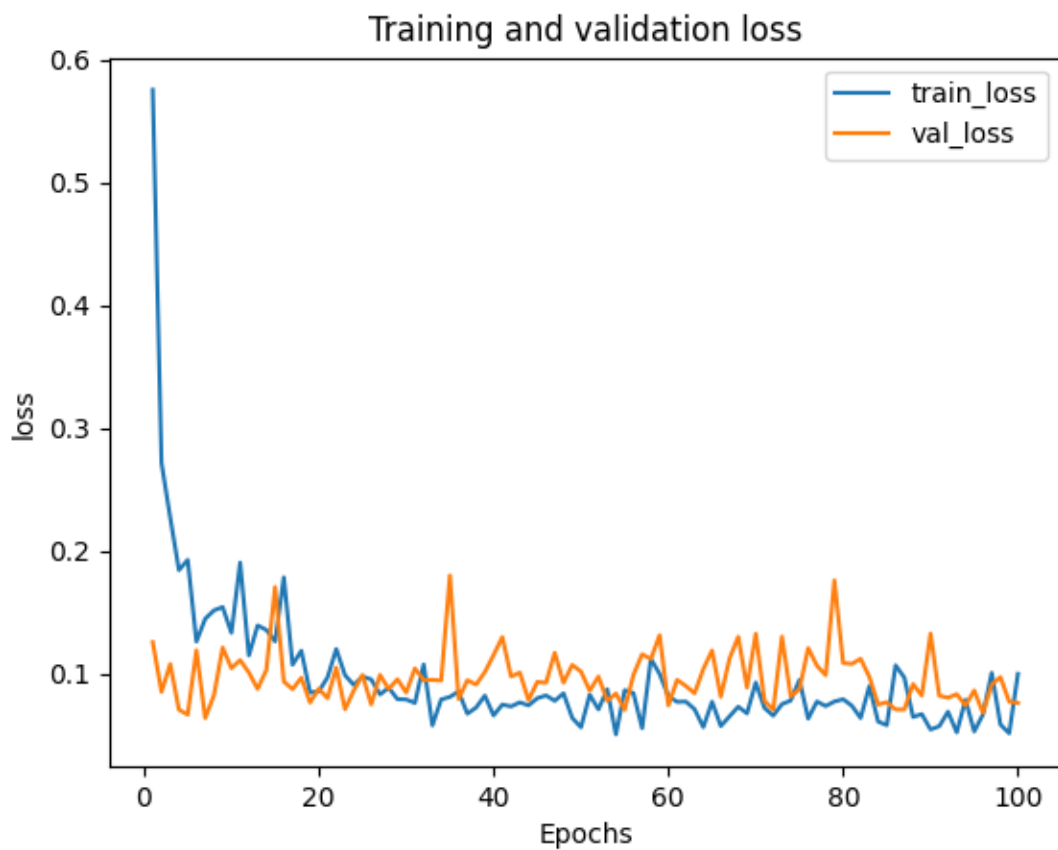
Red #1

- Configuración:
 - Sin Dropout
 - Sin regularización
 - Función de activación: ReLu
- Composición:
 - Tres capas densas con funciones de activación ReLU, con 64, 128, 64 neuronas respectivamente, y una capa de salida con una sola neurona para la regresión.
- Resultados:
 - Validation loss: 0.09289



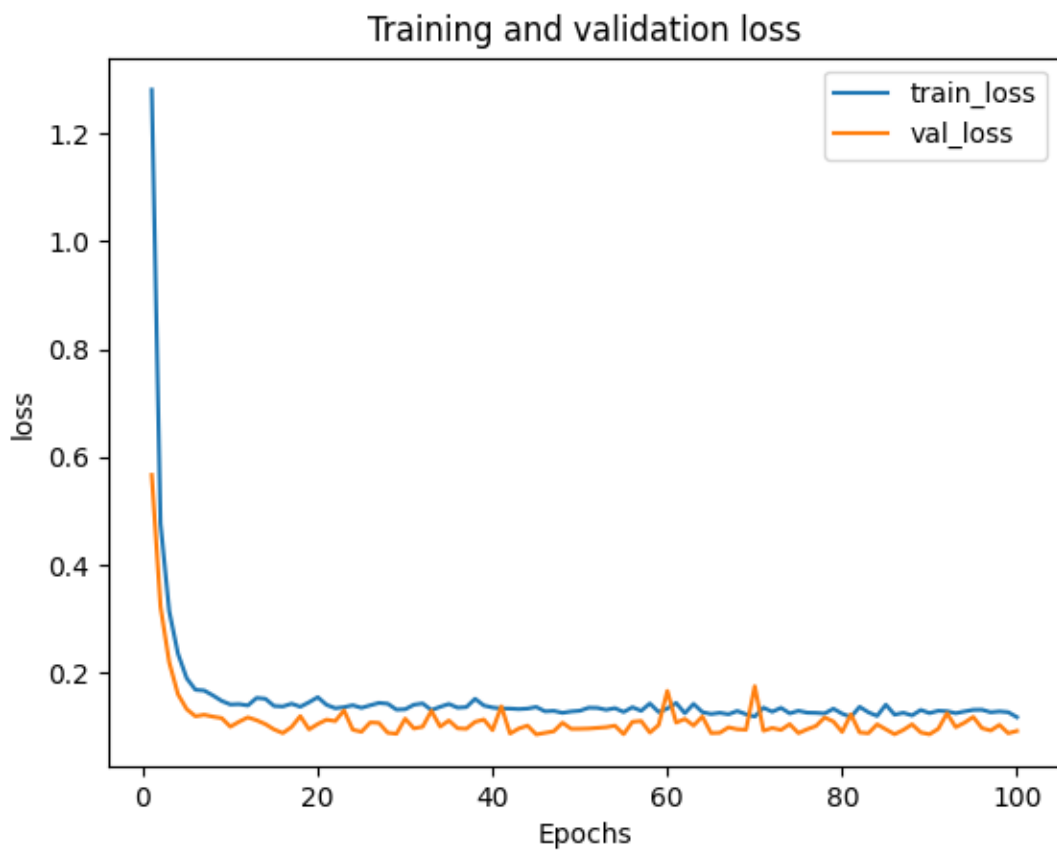
Red #2

- Configuración:
 - Con Dropout
 - Sin regularización
 - Función de activación: ReLu
- Composición:
 - Cuatro capas densas con funciones de activación ReLU, con 64, 128, 128, y 64 neuronas respectivamente, intercaladas con capas de Dropout con una tasa de 0.5 en cada una, y una capa de salida con una sola neurona para la regresión.
 -
- Resultados:
 - Validation loss: 0.15051



Red #3

- Configuración:
 - Sin Dropout
 - Con regularización L2
 - Función de activación: Tanh
- Composición:
 - Dos capas densas con funciones de activación Tanh y regularización L2 con 128 y 64 neuronas respectivamente, y una capa de salida con una sola neurona para la regresión.
- Resultados:
 - Validation loss: 0.16758



Discusión de Resultados y Diferencias de Rendimiento

Red #1: Sin Dropout, Sin Regularización, Función de Activación ReLU

- Fortalezas:
 - Rendimiento: La Red #1 mostró el test loss y validation loss más bajos, indicando que logró el mejor ajuste en los datos de prueba y validación. Esto sugiere que el modelo tiene una excelente capacidad de generalización, lo cual es crucial para predecir datos no vistos con precisión.
 - Simplicidad y Eficiencia: La arquitectura más sencilla de la Red #1 (con tres capas densas) sin técnicas adicionales de regularización o Dropout, contribuye a una mayor eficiencia en el entrenamiento y menor complejidad computacional.
- Debilidades:
 - Sin Regularización: Aunque el modelo funcionó bien en este caso, la falta de técnicas de regularización puede ser un riesgo en situaciones donde el conjunto de datos es más propenso al sobreajuste. No obstante, en este escenario específico, el modelo no parece estar sobreajustado.

Red #2: Con Dropout, Sin Regularización, Función de Activación ReLU

- Fortalezas:
 - Prevención de Sobreajuste: La incorporación de Dropout en la Red #2 ayuda a prevenir el sobreajuste al reducir la dependencia entre neuronas durante el entrenamiento. Esto es útil en redes más profundas o en conjuntos de datos con riesgo de sobreajuste.
- Debilidades:
 - Rendimiento Inferior: Aunque el Dropout es una técnica efectiva de regularización, la Red #2 mostró un test loss y validation loss mayores que la Red #1. Esto sugiere que, en este caso, el Dropout pudo haber sido excesivo o que la adición de capas no contribuyó positivamente al rendimiento. La alta tasa de Dropout (0.5) y la mayor complejidad del modelo podrían haber llevado a una pérdida de información y una menor capacidad de ajuste.

Red #3: Sin Dropout, Con Regularización L2, Función de Activación Tanh

- Fortalezas:
 - Regularización L2: esta es efectiva para reducir el riesgo de sobreajuste al penalizar los pesos grandes, ayudando a mantener el modelo más generalizado y menos susceptible a datos ruidosos.
- Debilidades:
 - Rendimiento Menor: La Red #3 obtuvo el test loss y validation loss más altos, indicando que la combinación de Tanh y regularización L2 no resultó en un mejor ajuste en este caso. La función de activación Tanh puede llevar a problemas de saturación, lo que puede afectar negativamente la convergencia del modelo. Además, la regularización L2 puede no haber sido necesaria si el modelo ya estaba funcionando adecuadamente sin ella.

Modelo óptimo:

Luego de analizar los resultados obtenidos, determinamos que el modelo elegido es: Red #1. Esto debido a distintos factores, entre ellos:

- Rendimiento Superior: La Red #1 alcanzó el test loss y validation loss más bajos, indicando el mejor en ambos conjuntos. Esto demuestra que este modelo tiene la mejor capacidad de generalización y ajuste a los datos disponibles.
- Simplicidad y Eficiencia: La Red #1 es la más sencilla en términos de arquitectura, sin técnicas adicionales de regularización ni Dropout, lo que puede ser ventajoso si el modelo no está sobre ajustado y el conjunto de datos es lo suficientemente representativo.
- Estabilidad en Resultados: La simplicidad del modelo y la función de activación ReLU contribuyen a un rendimiento estable y eficiente, sin la necesidad de regularización compleja o técnicas que puedan complicar el entrenamiento y ajuste del modelo.

Referencias:

Sotaquirá, M. (2021) La Regresión Lineal en Keras. Recuperado de:

<https://www.codificandobits.com/curso/fundamentos-deep-learning-python/redes-neuronales-7-regresion-lineal-keras/>

Berenguer Triana, Juan (2023). Programación de una red neuronal y ajuste de sus parámetros por backpropagation. Proyecto Fin de Carrera / Trabajo Fin de Grado, E.T.S.I. Industriales (UPM)
Recuperado de: <https://oa.upm.es/view/institution/Industriales/>

Castillo, C. (2023) Función de error o costo de una red neuronal. Recuperado de:

<https://carloscb.com/blog/funcion-de-error-o-costo-de-una-red-neuronal>