

Trabajo Práctico 1 — Reservas de Hotel

Organizacion de Datos
Curso Rodriguez
Primer cuatrimestre de 2023

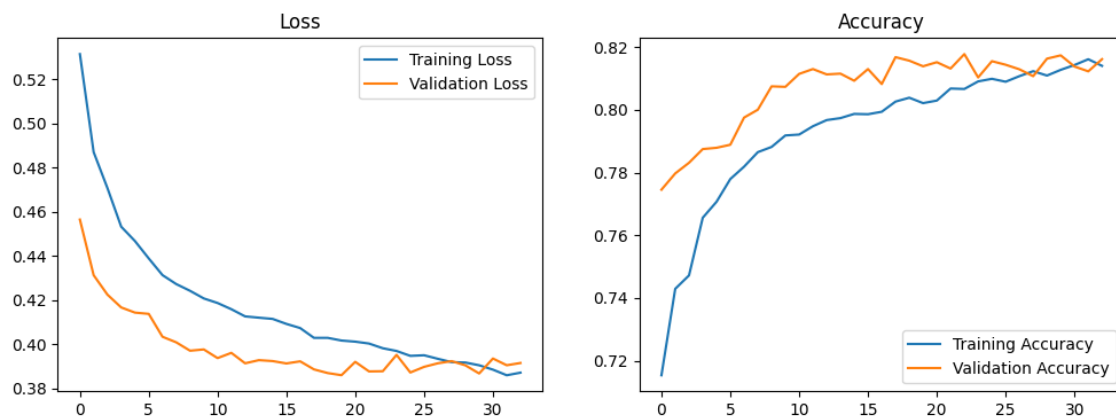
Alumno	Padrón	gitHub
Camila Gonzalez	105661	c-gonzalez-a
Eduardo Martín Bocanegra	106028	martinboca
Mateo Cabrera	108118	m-cabrerar

1. Introducción

En esta última entrega, el objetivo fue entrenar una red neuronal para hacer nuestras predicciones del estado de una reserva de hotel, utilizando herramientas como back propagation y early stopping para evitar el sobreajuste del modelo a los datos de entrenamiento.

2. Entrenamiento

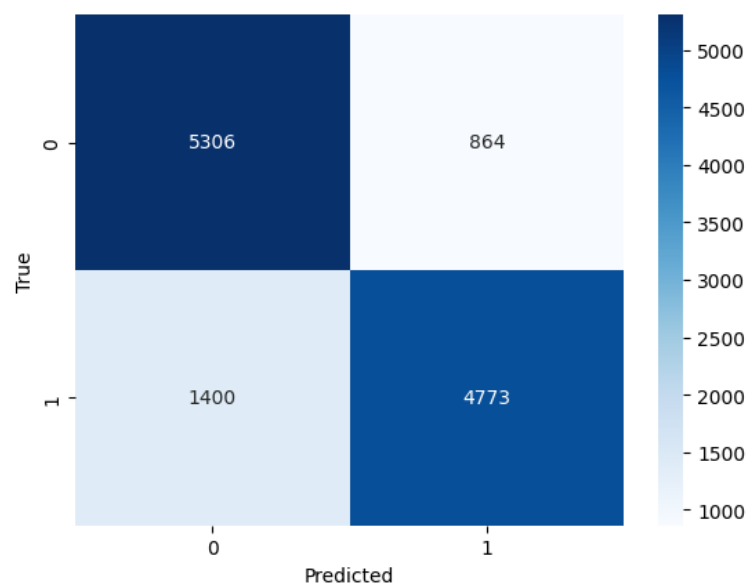
Al entrenar el modelo, observamos un incremento logarítmico del accuracy del modelo que se asienta al rededor de 0.81 luego de unas 25 iteraciones. De igual manera, el loss tiende a disminuir a 0.39 hacia el final de nuestro entrenamiento.



Al rededor de las 30 iteraciones el entrenamiento terminó con el uso de early stopping para evitar sobre ajustar el modelo.

3. Resultados

Obtuvimos un f1 score de 0.82, que es bastante bueno, pero no llega a ser tan bueno como el modelo de XGBoost implementado en la entrega anterior. La matriz de confusión quedó así:



4. Conclusiones

Luego de haber probado distintos modelos de predicciones, árboles de decisión, Random Forest, KNN, Boosting, Voting, Stacking, etc. pudimos observar que los que tuvieron mejor performance fueron siempre los modelos de ensamble, aquellos que utilizan varios modelos para realizar las predicciones. Dentro de los modelos híbridos, con los que mejor resultados obtuvimos fueron los híbridos. Voting y Stacking obtuvieron un f1 de 0.84321 y 0.85402 respectivamente, pero todos estuvieron relativamente cerca. Estos scores fue aceptables, pero es definitivamente mejorable. Con mejor y más profundo feature engineering, quizás probando otros modelos, etc.

