

Checkpoint 4 - Grupo 23

Introducción

Para este capítulo utilizamos el dataset resultante del capítulo 1, con exactamente el mismo procesamiento de datos. Además, utilizamos la técnica de min max scaling para escalar los datos de entrada ya que esto nos trae grandes beneficios para nuestro modelo. Entre tantos beneficios, podemos recalcar que mejora la convergencia de la red neuronal, iguala la influencia de las características y además evita problemas numéricos debido a que algunas funciones de activación no trabajan muy bien si los valores de entrada son muy grandes.

Construcción del modelo

Detallar como mínimo los siguientes puntos para el “mejor” modelo obtenido:

- ¿Cuál fue la arquitectura escogida?
 - Capa entrada
 - Neuronas de entrada: 256
 - Función de activación: relu
 - Tipo de conexión: Densa
 - Capas ocultas
 - 1. dropout de 0.2
 - 2. neuronas: 128, función: relu, tipo: densa
 - 3. neuronas: 64, función: relu, tipo: densa
 - Capa de salida
 - cantidad de neuronas: 1
 - función de activación: sigmoidea
- Hiperparametros optimizados: kernel_initializer, cantidad de neuronas
- Optimizador utilizado: Adam
- Regularización: L1 y L2 en las diferentes capas
- Ciclos de entrenamiento: 60

Cuadro de Resultados

Realizar un cuadro de resultados comparando los modelos que entrenaron (entre ellos debe figurar cuál es el que seleccionaron como mejor predictor). Confeccionar el siguiente cuadro con esta información:

Medidas de rendimiento en el conjunto de TEST:

- F1
- Precision
- Recall
- Metrica X
- Resultado obtenido en Kaggle.

Para la métrica X vamos a utilizar el F1 score con los datos de entrenamiento, tal como hicimos en los capítulos anteriores

Modelo	F1-Test	Precision Test	Recall Test	Metrica X	Kaggle
modelo_1	0,863	0,845	0,881	0,873	0,848
modelo_2	0,843	0,819	0,867	0,843	
modelo_3	0,852	0,833	0,871	0,854	
modelo_4	0,865	0,860	0,870	0,877	0,855

Nota: indicar brevemente en qué consiste cada modelo de la tabla y detallar el caso del mejor modelo.

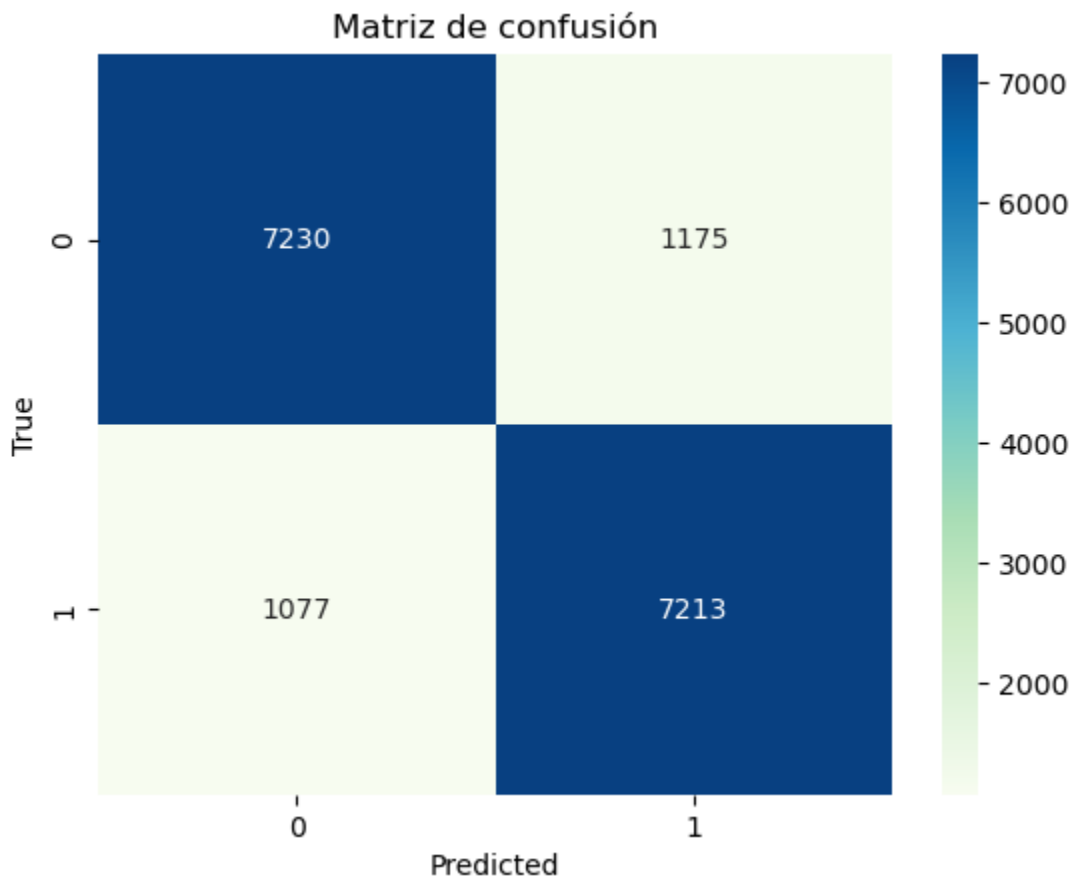
Modelo 1: modelo básico, una capa de entrada densa con 128 neuronas con activación relu y una de salida con activación sigmoidea

Modelo 2: modelo un poco más avanzado, una capa de entrada densa con 128 neuronas con activación relu, un dropout de 0.2 una de salida con activación sigmoidea. Además utiliza regularizaciones l1 y l2.

Modelo 3: es muy parecido al modelo 2, pero se le suma el learning rate y el batch normalization, los cuales evitan notablemente el overfit.

Modelo 4: último modelo, varias capas ocultas, dropouts y regularizaciones para evitar el overfit.

Matriz de Confusion



Como podemos observar el modelo está bastante balanceado, predice correctamente casi la misma cantidad de cada clase. Sin embargo se puede apreciar que tiende a clasificar erróneamente prediciendo más la clase 1 sobre la clase 0

Tareas Realizadas

Dado que somos solo dos integrantes, decidimos ir tarea a tarea, tomando decisiones en conjunto y dividiéndonos la carga de cada parte a la mitad.

Integrante	Tarea
Mateo Vroonland	Todo
Juan Pablo Carosi Warburg	Todo