

Checkpoint 4 - Grupo 24

Introducción

En este checkpoint se crearon distintas redes neuronales variando su regularización y se optimizó los hiperparamétricos del mejor modelo mediante GridSearch y Cross Validation.

Construcción del modelo

- Arquitectura escogida del mejor modelo:
 - Capa entrada
 - Neuronas de entrada: 128
 - Función de activación: Relu
 - Tipo de conexión: Dense (Fully Connected)
 - Capas ocultas
 - Neuronas por capa: 64
 - Función de activación: Relu
 - Tipo de conexión: Dense
 - Capa de salida
 - Cantidad de neuronas 2
 - función de activación: Sigmoid

- ¿Qué hiperparametros se optimizaron?

Los hiperparametros optimizados fueron Epochs y batches utilizando GridSearch.

- ¿Qué optimizador se utilizó?

El optimizador que utilizamos fue Adam con Learning Rate de 0.001, Beta_1 de 0.9 y Beta_2 de 0.999

- ¿Se utilizó alguna técnica de regularización? ¿Cuál?

Utilizamos 4 técnicas distintas de regularización buscando la mejor predicción, inicialmente usamos "Dropout", con una probabilidad de eliminación de 0,4. Luego probamos utilizar L1, L2 en modelos separados y por último cambiar L1 y L2.

- ¿Cuántos ciclos de entrenamiento utilizó?

Utilizamos 100 ciclos de entrenamiento.

Cuadro de Resultados

Medidas de rendimiento en el conjunto de TEST:

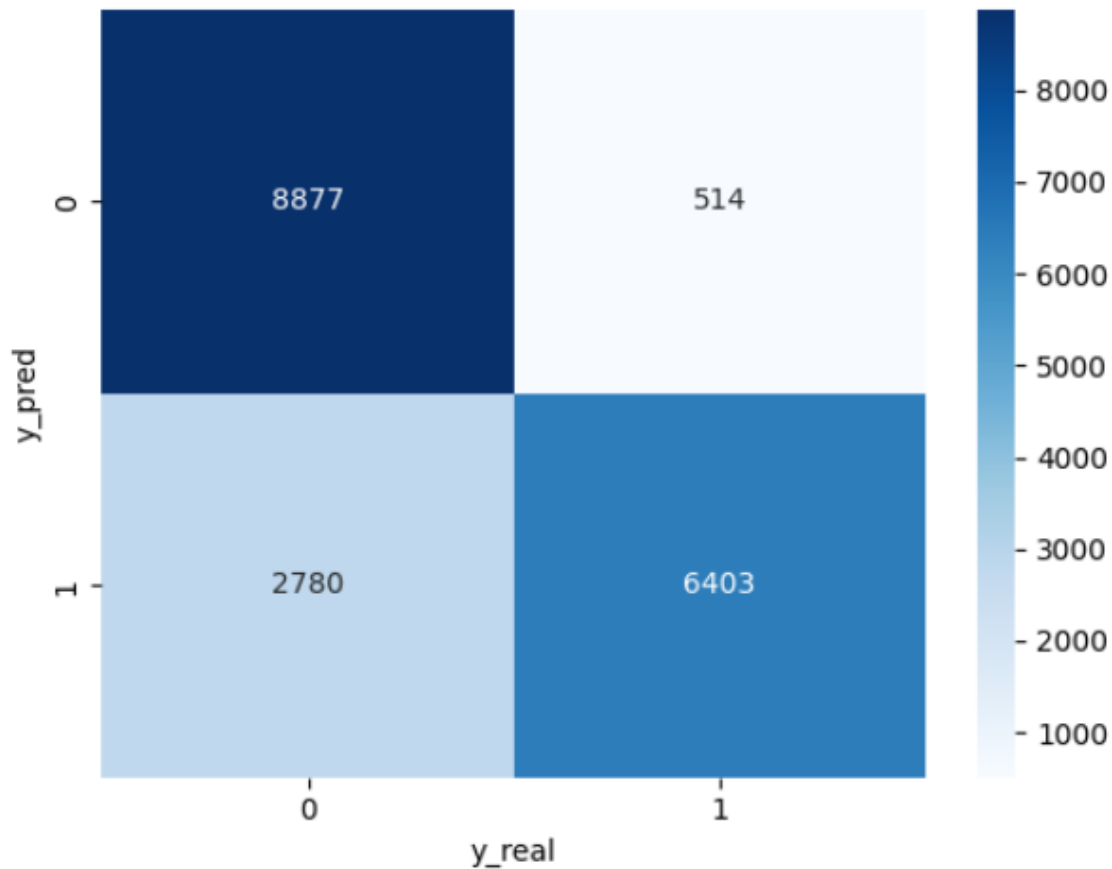
- F1
- Precision
- Recall
- Accuracy
- Resultado obtenido en Kaggle.

Modelo	F1-Test	Presicion Test	Recall Test	Accuracy	Kaggle
Modelo Dropout	0.7991	0.6923	0.9292	0.8225	0.8225
Modelo L2	0.7925	0.6828	0.9102	0.8121	0.8160
Modelo L1 y L2	0.7892	0.6801	0.9146	0.8152	0.8008
Modelo L1	0.7844	0.6887	0.9165	0.8136	0.7806

Los modelos que vemos en las tablas fueron creados con la regularización que su propio nombre indica y observamos mediante los resultados en kaggle que el Dropout supera a todos los otros modelos planteados.

El modelo con el Dropout consiste en fijar un parámetro entre 0 y 1 el cual funciona como una probabilidad de eliminar la neurona con el fin de disminuir el overfitting. En nuestro caso optamos por el número 0.4.

Matriz de Confusion



Podemos observar que con estos tipos de modelos se obtiene un valor de falsos positivos más elevado a comparación de los otros modelos planteados en checkpoints anteriores. Creemos que esto se debe a alguna alteración hecha en el CHP1 en la parte limpieza del dataset que genera este error a la hora de predecir.

Tareas Realizadas

Integrante	Tarea
Lautaro Torraca	Realizamos una tarea compartida a lo largo de este checkpoint.
Negrotti Gianluca	
Marco Tosi	