

Redes Neuronales

Sebastian Makkos, Rodolfo Albornoz, Valeria Brzoza

Organización de datos, 1C23

Para este checkpoint, tuvimos que construir distintas redes neuronales de clasificación, variando sus parámetros, la optimización, su arquitectura y posteriormente elegir el mejor modelo entre ellos para poder analizarlo con sus métricas y su matriz de confusión.

En el primer modelo, utilizamos activación sigmoid, una densidad de entrada de 16, un aprendizaje de 0,01. Para la una función de perdida utilizamos binary cross entropy y utilizamos épocas y baches con un valor de 100. Para este modelo, obtuvimos una métrica de f1 score de 0.7575969450868607 como el mejor modelo hasta ahora.

En el segundo modelo, añadimos un kernel inicializador uniforme a la capa de entrada. Aumentamos la cantidad de neuronas de la capa oculta y disminuimos el aprendizaje para que el aprendizaje sea más lento y profundo a un valor de 0,0001. Para este modelo, obtuvimos una métrica de f1 score de 0.7458151188826101 como el segundo mejor modelo.

Y para el último modelo, utilizamos GridSearch cross validation, una activación softmax, distinto a los demás, un learning rate de 0,0001 y una función de pérdida utilizamos categorical cross entropy. Para este modelo, no probamos con parámetros más grandes por el tiempo de aprendizaje, pero fue algo que lo tuvimos en cuenta en el desarrollo del modelo mismo. Para este modelo, obtuvimos una métrica de f1 score de 0.6812831457767429 como la peor de las tres.

En general, también probamos otras cosas como optimizadores (adam, nesterov, momentum) y otras funciones de activación como linear, pero no vimos mejoras significativas en la performance de los modelos. De hecho, por ejemplo, al usar linear, empeoro bastante la performance.

Para concluir, podemos decir que nuestro mejor modelo fue el primero, en el cual no fue optimizado y el aprendizaje no es el más bajo comparado con los demás. Y analizando sus métricas y matriz de confusión, concluimos que no estaba overfitted.