

Trabajo Práctico 1 - Checkpoint 4

Grupo 21 - 'df' is not defined

[7506/9558] Organizacion de Datos
Primer cuatrimestre de 2023

Alumno	Padron	Email
Martin Pata Fraile de Manterola	106226	mpata@fi.uba.ar
Tobias Emilio Serpa	108266	tserpa@fi.uba.ar
Juan Francisco Cuevas	107963	jcuevas@fi.uba.ar

En esta parte pudimos implementar las redes neuronales, algunas simples con menos capas y otras mas complejas con mas capas, las cuales claramente son mas complejas y mas costosas pero tambien dan resultados mejores. Lo que pudimos notar es que no hubo redes neuronales las cuales destaquen a la hora de hacer las predicciones, todos los modelos que hayamos variando sus hyperparametros intentando conseguir los mejores para la misma, en la mayoria de los casos obtuvimos resultados muy parecidos.

Cuando arrancamos con las primeras redes ahi teniamos modelos no tan buenos, pero a medida que fuimos probando y agregando capas estos modelos mejoraron pero no lo suficiente como para poder dar la mejor prediccion final, ya que esta la conseguimos con los ensamblados (con XGBoost). Lo que pudimos notar es que la hora de usar SGD (optimizador basico de las redes neuronales) los modelos eran mas lentos y ademas predecian de una peor manera, esto mejoro cuando implementamos el optimizador ADAM el cual es mucho mas efectivo y rapido que SGD. Tambien notamos que con 100 epochs era mas que suficiente para conseguir los mejores modelos (esto se puede ver en el grafico tambien, como alrededor de 100 epochs es el pico maximo del accuracy).

Aparte podemos concluir que el accuracy con los datos de training era muy bueno (casi un 0,93) pero a la hora de predecir nuevos datos este disminuia gradualmente dando un accuracy muy bajo respecto al que daba con los datos de training (bajando hasta 0,84 aproximadamente, lo cual no es valor malo pero no es tan bueno haciendo la comparacion).

Por lo que podriamos decir que redes neuronales no es el mejor clasificador para el problema propuesto en este caso, apesar de ser buenos clasificadores depende mucho del problema que se proponga y que se intente predecir ya que no con todo va a funcionar igual de bien.