

Aprendizaje Profundo - Practico N°2

Integrantes:

- Buzzi, Sergio
- Diaz, Carolina
- Fabro, Juan

En este documento se presentan brevemente los resultados obtenidos al experimentar con distintas arquitecturas de redes neuronales sobre el dataset petfinder.

Dado que las posibles configuraciones son innumerables, se decide plantear una arquitectura base y estudiar que ocurre al modificar unos pocos hiperparámetros. En todos los casos se aplica sobre la descripción el embedding glove, empleando convoluciones 1D de 2, 3,5 y 10 palabras. Luego se combina la información de los distintos tipos de variables en una red fully connected unidireccional, en la cual se emplean capas cuya cantidad de neuronas van decreciendo progresivamente.

Como conocemos el dataset de una materia previa, sabemos que es muy difícil lograr una accuracy superior al 40% en el conjunto de test. Efectivamente eso es lo sucedió, pero en las primeras pruebas notamos que los modelos resultaban en un marcado sobreajuste en los datos de entrenamiento y development. Por lo tanto, en cada una de las capas ocultas de la neuronal artificial fully connected se aplican operaciones de dropout con la intención de reducir los problemas de overfitting.

En la tabla de la página siguiente, se muestran los resultados obtenidos al plantear diversos valores de los hiperparámetros: Tamaño de la capa de descripción, tamaño de la primera capa oculta y tasa de dropout. En el caso de este último hiperparámetro se probó con los valores 0,50 y 0,75; siendo este valor demasiado alto en comparación a lo que se emplea habitualmente. Es intrigante observar que aún con una estrategia de regularización tan agresiva no se consigue solucionar el problema de overfitting.

| Experi mento | Hiperparámetros | | | Métricas | | |
|-----------------|----------------------------|-----------------------|------------------|----------|--------------|---------------|
| | DESCRIPTION _LAYER_SIZE | HIDDEN_LAYER _SIZE | DROPOUT _RATE | Dev_loss | Dev_accuracy | Test_accuracy |
| 1 | 256 | 128 | 0.50 | 0.6266 | 0.7520 | 0.3182 |
| 2 | 256 | 128 | 0.75 | 0.5757 | 0.8016 | 0.3114 |
| 3 | 256 | 256 | 0.50 | 0.1526 | 0.9424 | 0.3144 |
| 4 | 256 | 256 | 0.75 | 0.1599 | 0.9381 | 0.3152 |
| 5 | 256 | 512 | 0.50 | 0.0918 | 0.9679 | 0.3265 |
| 6 | 256 | 512 | 0.75 | 0.0929 | 0.9688 | 0.3023 |
| 7 | 512 | 128 | 0.50 | 0.8148 | 0.6991 | 0.3182 |
| 8 | 512 | 128 | 0.75 | 0.6502 | 0.7643 | 0.3197 |
| 9 | 512 | 256 | 0.50 | 0.0989 | 0.9641 | 0.3288 |
| 10 | 512 | 256 | 0.75 | 0.1310 | 0.9589 | 0.3288 |
| 11 | 512 | 512 | 0.50 | 0.0595 | 0.9769 | 0.3394 |
| 12 | 512 | 512 | 0.75 | 0.0631 | 0.9759 | 0.3303 |
| 13 | 1024 | 128 | 0.50 | 0.7815 | 0.6750 | 0.3379 |
| 14 | 1024 | 128 | 0.75 | 0.8375 | 0.6504 | 0.3326 |
| 15 | 1024 | 256 | 0.50 | 0.1129 | 0.9570 | 0.3046 |
| 16 | 1024 | 256 | 0.75 | 0.1253 | 0.9608 | 0.3432 |
| 17 | 1024 | 512 | 0.50 | 0.0566 | 0.9750 | 0.3258 |
| 18 | 1024 | 512 | 0.75 | 1.2765 | 0.4454 | 0.3220 |

Aparentemente, al incrementar el tamaño de la capa de descripción se consigue un mejor ajuste en el conjunto de test. No son tan claros los resultados para el caso de las modificaciones en el tamaño de las capas ocultas y la tasa de dropout.