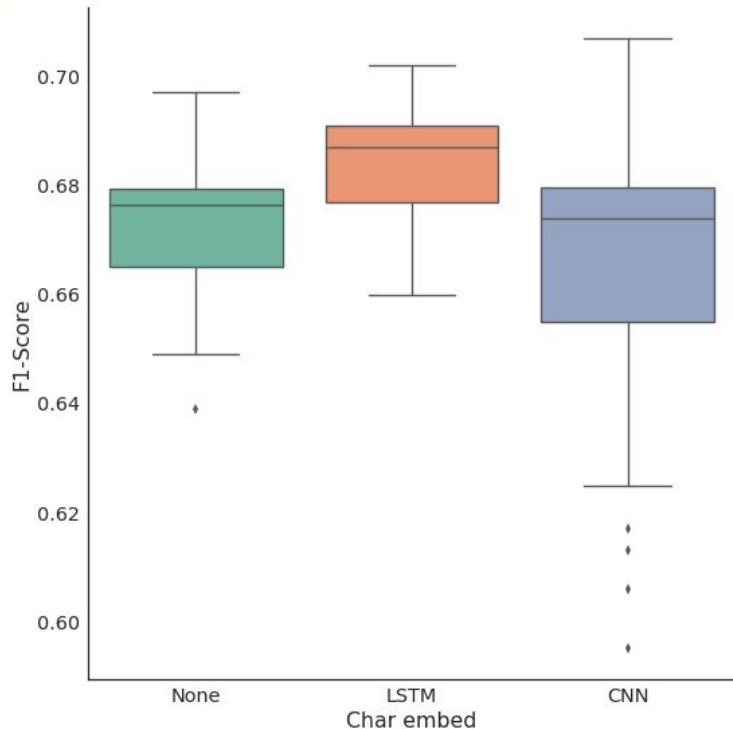


Entrenamiento de Redes Neuronales

Diplomatura en Ciencia de Datos, Aprendizaje
Automático y sus Aplicaciones

Variación de modelos



¿Cuál de los tres modelos es mejor?



Variación de modelos

Why Comparing Single Performance Scores Does Not Allow to Draw Conclusions About Machine Learning Approaches, Neimers and Gureyvich, CoRR 2017



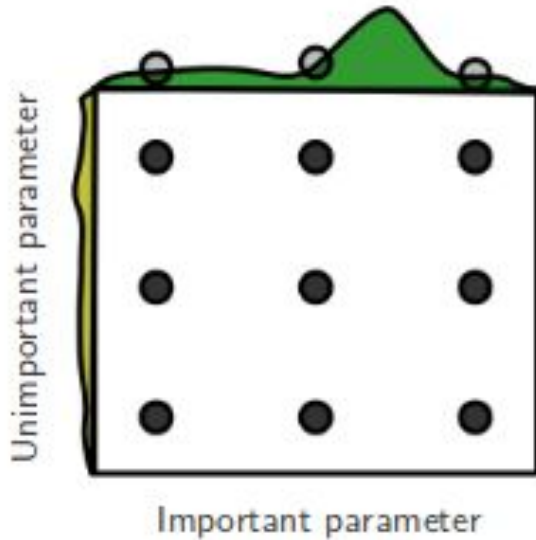
Exploración random

[Random Search for Hyper-Parameter Optimization](#), J. Begstra and Y. Bengio, JMLR 2016

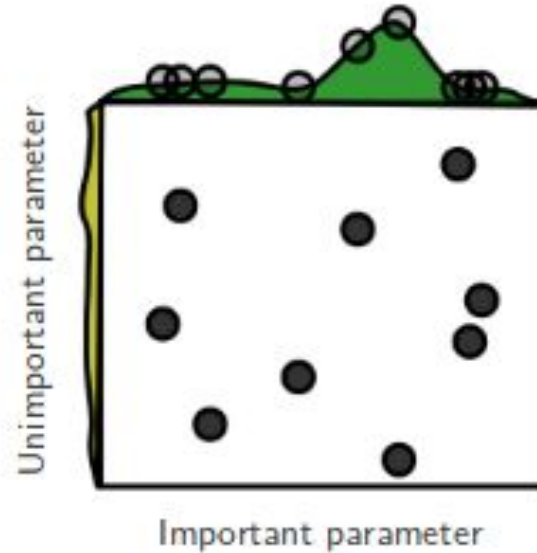
- *only a few of the hyper-parameters really matter*
- *random search es mejor que grid search*

Exploración random

Grid Layout



Random Layout





Paciencia

Si el modelo no mejora la performance en X iteraciones, terminar en entrenamiento.

Utilizar una partición distinta para determinar el corte que para evaluar!



Síntomas de error

- Siempre predice el mismo valor
- Correr el mismo modelos con inicializaciones distintas da resultados muy dispares
- Diferencia grande entre train y dev



Síntomas de error

- Los pesos no cambian luego del entrenamiento
- Todos los gradientes tienen valores iguales a 0
- No activación de las neuronas
- Loss no se corresponde con métricas
- Hay NaN!