**Optimización de hiperparametros | Hyperparameter Optimization**

Familiarizados con el concepto de Cross Validation vamos a utilizar este mismo principio de fondo para lograr automatizar un poco la selección y optimización de nuestros modelos.

**Problema**: Parece que encontramos un modelo de aprendizaje que parece funcionar, pero esto puede implicar que ahora tenemos que **encontrar la optimización de cada uno de los parámetros de este modelo**, encontrar el que mejor se ajuste y el que mejor resultado nos de.

1. Es facil perderse entre los conceptos de tantos parámetros. Tenemos flexibilidad para algoritmos básicos de Machine Learning, pero facil perderse.
2. Es difícil medir la sensibilidad de los mismos manualmente.
3. Es COSTOSO, en tiempo humano y computacionalmente.

Scikit Learn nos ofrece enfoques para automatizar el proceso de optimización paramétrica. Existen 3 enfoques principales, estos son:

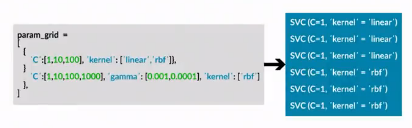
1. Optimización manual
2. Optimizacion por grilla de parámetros | GridSearchCV
3. Optimizacion por búsqueda aleatorizada |  
   .  
   .  
   **Optimización manual**
4. Escoger el modelo que queremos ajustar.
5. Buscar en la documentación de Scikit-Learn
6. Identificar parámetros y ajustes. Parámetros que vamos a necesitar y cuáles son los posibles ajustes que vamos a requerir para cada uno de estos parámetros.
7. Probar combinaciones una por una iterando a través de listas.  
   .  
   .  
   **Optimizacion por grilla de parámetros | GridSearchCV**

Es una forma organizada, exhaustiva y sistematica de probar todos los parametros que le digamos que tenga que probar, con los respectivos rangos de valores que le aportemos.

1. Definir una o varias métricas que queremos optimizar.
2. Identificar los posibles valores que pueden tener los parámetros.
3. Crear un diccionario de parámetros.
4. Usar Cross Validation.
5. Entrenar el modelo (e ir por un café)

La grilla de parámetros nos define GRUPOS DE PARÁMETROS que serán probados en todas sus combinaciones (Un grupo a la vez)

Ejemplo:

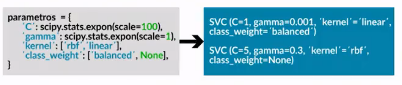
  
.  
.

**Optimizacion por búsqueda aleatorizada | RandomizedSearchCV**

Si no tenemos tanto tiempo para una prueba tan exhaustiva o queremos combinaciones aleatorias usaremos este metodo. Es lo mismo que el caso anterior, pero busca de forma aleatoria los parametros y Scikit Learn selecciona los mejores de las combinaciones aleatorias que se hicieron.

En este método, definimos escalas de valores para cada uno de los parámetros seleccionados, el sistema probará varias iteraciones (Configurables según los recursos) y mostrará la mejor combinación encontrada.

Ejemplo:

  
.  
.

**GridSearchCV vs RandomizedSearchCV**

* **GridSearchCV**
  + Cuando se quiera realizar un estudio a fondo sobre las implicaciones de los parámetros.
  + Se tenga tiempo.
  + Se tenga poder de procesamiento.
* **RandomizedSearchCV**
  + Cuando se quiera explorar posibles optimizaciones.
  + Haya poco tiempo.
  + Haya poco poder de procesamiento.

