**Validación de los datos.**

1. En ML La última palabra la tienen los datos

Nuestras intuiciones no tienen nada que hacer frente a lo que digan nuestros datos y las matemáticas que aplicamos sobre estos datos.

1. Necesitamos mentalidad de testeo

Necesitamos mentalizarnos como si fuésemos testers continuos de nuestra aplicación. No se trata solo de probar al principio o al final, sino que tendremos que probar constantemente durante todo el proceso varias formas con varios conjuntos de datos, con varios configuraciones de parámetros y con varias distribuciones de nuestros datos para poder encontrar la solución óptima de nuestro problema que tenemos pendiente.

1. *“Todos los modelos son malos, solamente que algunos resultan útiles”* : Nunca nuestros modelos van a corresponder con la realidad al 100%, lo que sucede es que si realizamos ciertas configuraciones, y con cierta habilidad y destreza, podremos llegar a un punto en el que nuestro modelo va a ser útil para el caso específico con el que estamos trabajando.

**Tipos de validación.**

1. Dividir nuestros datos en Entrenamiento/ Prueba (Hold- On).

*“train\_test\_split”*

Recomendable cuando tenemos lo siguiente:

Necesitamos hacer un prototipado rápido.

Cuando no se tiene mucho conocimiento: Cuando recién estamos iniciando en ML.

No se cuenta con abundante poder de cómputo

1. Usar Validación cruzada (K-Folds).

Recomendable cuando tenemos lo siguiente:

“ Recomendable en la mayoría de casos ”

Se cuenta con un equipo suficiente para desarrollar ML

Se requiere la integración con técnicas de optimización paramétrica.

Se tiene más tiempo para las pruebas



1. Validación Cruzada (LOOCV).

Recomendable cuando tenemos lo siguiente:

Se tiene un gran poder de cómputo.

Se cuentan con pocos datos para dividir por Training/Test.

Personas que sufren de TOC y quieren probar todos los casos posibles.

Nota: LOOCV modelo deja un dato por fuera y entrena con el resto de datos. Este proceso lo realiza con todas las combinaciones posibles de nuestro dataset.