

Overfitting y Underfitting

El aprendizaje supervisado se entiende mejor como la aproximación de una función target (f) que mapea variables de entrada (X) a una salida (Y).

$$Y = f(X)$$

Esta caracterización describe el rango de los problemas de clasificación o predicción y los algoritmos de machine learning que resuelven mejor cada uno de ellos.

Una consideración importante al aprender la función target de los datos de entrenamiento es, qué tan bien el modelo puede generalizar nuevos datos. La generalización es importante porque los datos que tenemos para el entrenamiento son sólo una muestra, están incompletos y sucios ó ruidosos.

Generalización en Machine Learning

En machine learning el aprendizaje de una función target desde los datos de entrenamiento se describe como aprendizaje inductivo.

Inductivo se refiere a aprender conceptos generales a través de ejemplos específicos que es exactamente el problema que el aprendizaje supervisado pretende resolver. Esto es diferente de la deducción que es el otro extremo y busca aprender conceptos específicos de reglas generales.

La generalización se refiere a qué tan bien aplican los conceptos aprendidos por un modelo de machine learning a ejemplos que no fueron utilizados para entrenar o para que el modelo aprendiera.

La meta de un buen modelo de machine learning es generalizar bien desde los datos de entrenamiento a cualquier dato nuevo que encaje en el mismo. Esto nos permite hacer predicciones en el futuro de datos que el modelo nunca ha visto.

Hay una terminología que se usa en machine learning cuando hablamos de qué tan bien aprende y generaliza los nuevos datos un modelo de machine learning, llamada overfitting y underfitting.

Overfitting y underfitting son las dos mayores causas de mal desempeño en algoritmos de machine learning.

Fit estadístico

En estadística, el fit o entrenamiento se refiere a que tan bien se aproxima la función target.

Es bueno usar estos términos en machine learning porque los algoritmos de aprendizaje supervisado buscan aproximar la función para las variables de salida dadas las variables de entrada.

La estadística a menudo describe la calidad del entrenamiento, lo que involucra medidas usadas para estimar que tan bien la aproximación hace juego con la función target.

Algunos de estos métodos son útiles en machine learning, pero algunas de estas técnicas asumen que sabemos la forma de la función target a la que estamos aproximando, lo que no es el caso en machine learning.

Si supiéramos la forma de la función target, la usaríamos directamente para hacer predicciones, en lugar de intentar aprender una aproximación de las muestras de nuestros datos de entrenamiento.

Underfitting & Overfitting Machine Learning

Underfitting

Entreno al modelo con
1 sola raza de perro



Muestra nueva:
¿Es perro?



La máquina fallará en reconocer al perro por falta de suficientes muestras. No puede generalizar el conocimiento.

Overfitting

Entreno al modelo con
10 razas de perro color marrón



Muestra nueva:
¿Es perro?



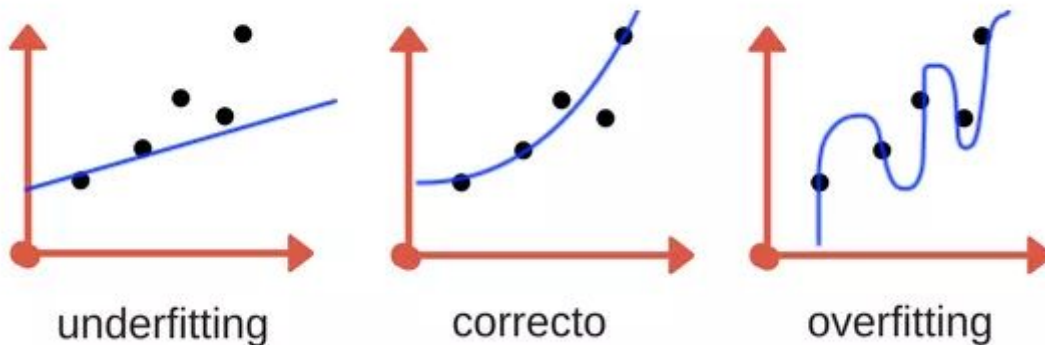
La máquina fallará en reconocer un perro nuevo porque no tiene estrictamente los mismos valores de las muestras de entrenamiento.

Overfitting en Machine Learning

Overfitting se refiere a un modelo que se ajusta a los datos demasiado bien.

El Overfitting sucede cuando un modelo aprende el detalle y ruido en los datos de entrenamiento en la medida que esto impacta negativamente el desempeño de un modelo cuando recibe datos nuevos.

Esto significa que el ruido o las fluctuaciones aleatorias en los datos de entrenamiento es adquirido y aprendido como conceptos por el modelo. El problema es que estos conceptos no aplican en los datos nuevos e impactan negativamente la habilidad del modelo para generalizar.



Underfitting en Machine Learning

El Underfitting se refiere a un modelo que no puede modelar los datos de entrenamiento ni datos nuevos.

Un modelo de machine learning con underfit no será un modelo adecuado y será obvio ya que tendrá mal desempeño en los datos de entrenamiento.

Para resolverlo se puede intentar usar otro algoritmo o utilizar más muestras.

Obteniendo un buen fin en Machine Learning

Para intentar que estos problemas nos afecten lo menos posible, podemos llevar a cabo diversas acciones:

- Cantidad mínima de muestras tanto para entrenar el modelo como para validarlo.
- Clases variadas y equilibradas en cantidad: En caso de aprendizaje supervisado y suponiendo que tenemos que clasificar diversas clases o categorías, es importante

que los datos de entrenamiento estén balanceados. Supongamos que tenemos que diferenciar entre manzanas, peras y bananas, debemos tener muchas fotos de las 3 frutas y en cantidades similares. Si tenemos muy pocas fotos de peras, esto afectará en el aprendizaje de nuestro algoritmo para identificar esa fruta.

- Conjunto de validación de datos. Siempre subdividir nuestro conjunto de datos y mantener una porción del mismo “oculto” a nuestra máquina entrenada. Esto nos permitirá obtener una valoración de aciertos/fallos real del modelo y también nos permitirá detectar fácilmente efectos del overfitting /underfitting.
- Parameter Tunning o Ajuste de Parámetros: deberemos experimentar sobre todo dando más/menos “tiempo/iteraciones” al entrenamiento y su aprendizaje hasta encontrar el equilibrio.
- Cantidad excesiva de Dimensiones (features), con muchas variantes distintas, sin suficientes muestras. A veces conviene eliminar o reducir la cantidad de características que utilizaremos para entrenar el modelo. Una herramienta útil para hacerlo es PCA.
- Si nuestro modelo es una red neuronal artificial –deep learning-, podemos caer en overfitting si usamos capas ocultas en exceso, ya que haríamos que el modelo memorice las posibles salidas, en vez de ser flexible y adecuar las activaciones a las entradas nuevas.
- Si el modelo entrenado con el conjunto de test tiene un 90% de aciertos y con el conjunto de validación tiene un porcentaje muy bajo, esto señala claramente un problema de overfitting.
- Si en el conjunto de validación sólo se acierta un tipo de clase (por ejemplo “peras”) o el único resultado que se obtiene es siempre el mismo valor será que se produjo un problema de underfitting.