

Introducción al aprendizaje automático

Visto el tipo de sistemas expertos que queremos conseguir tras aplicar alguna/s técnica/s de ML, tenemos que ver cómo aprenden estos sistemas para obtener esa función de clasificación o regresión, que en adelante la denominaremos “hipótesis”.

Para que los sistemas aprendan, se ha de tener un conjunto de datos (o data set) de aprendizaje o entrenamiento que son utilizados para obtener una hipótesis (modelo o función) que generalice esos datos adecuadamente. Cuando se habla de “generalizar”, se habla de predecir la salida a partir de nuevos datos de entrada (datos de test) distintos a los datos de entrenamiento.

Este proceso de aprendizaje para la obtención de una hipótesis, lo podemos ver esquematizado en la siguiente imagen. En primer lugar contamos con un conjunto de datos que los utilizaremos para entrenar o enseñar al sistema. Aplicando alguna de las técnicas de ML, obtendremos una hipótesis que a priori debería ser la mejor función que se ajusta y generaliza los datos de entrenamiento. Para entender este esquema, supongamos que estamos en el caso de un aprendizaje supervisado (que explicamos en el punto de Tipos de Aprendizaje), en el que para cada entrada de los datos de entrenamiento, conocemos cuál va a ser su salida. Por tanto; para obtener la hipótesis en este tipo de aprendizaje, tenemos que obtener una función que se ajuste a esos datos de entrenamiento; o dicho de otra manera, que minimice el “error empírico” que es el error medido tras aplicar la hipótesis a los datos de entrenamiento. Ciertamente lo que interesa es que la hipótesis obtenida tenga el menor error posible con los datos de test, pero se asume que los datos de entrenamiento son una muestra lo suficientemente representativa como para que el error cometido con los datos de test sea similar al error empírico.

En resumen el objetivo de estas técnicas es encontrar aquella función o modelo (hipótesis); dentro de todo el conjunto de funciones o modelos, que mejor se ajusta y generaliza los datos de entrenamiento para aplicarlo a los datos de test y así obtener una predicción, recomendación, clasificación, etc

