

INF-393 Máquinas de Aprendizaje

PAUTA Quiz 0.

SE RUEGA AL LECTOR ESTUDIOSO QUE COMUNIQUE AL PROFESOR CUALQUIER
ERROR QUE DESCUBRA EN ESTE DOCUMENTO.

1. *Según el autor, ¿Cuáles son los 3 elementos principales de un algoritmo de Machine Learning? Describálos brevemente.*

Los 3 componentes críticos son:

- (a) **Representación:** Es decir, una forma de representar los datos y el modelo que permita manipularlos computacionalmente. La elección de una representación para los datos implica la elección de las variables o atributos que se usarán para aprender la relación entre el patrón de entrada y la salida que se quiere aproximar. La elección de una representación para el modelo implica la elección de un espacio de hipótesis y su eventual parametrización.
- (b) **Evaluación:** Un funcional que permita distinguir entre modelos buenos y malos.
- (c) **Optimización:** Un algoritmo que permita buscar el mejor modelo con respecto a la función objetivo elegida.

2. *Explique qué se entiende por sobre-ajuste.*

Se dice que un algoritmo de aprendizaje cae en sobre-ajuste cuando exhibe un error de entrenamiento mucho más pequeño que su error de predicción real (test).

3. *¿Por qué el autor afirma que un algoritmo de entrenamiento que encuentra un óptimo local podría ser mejor que uno que encuentra el óptimo global?*

Porque la función objetivo utilizada durante el entrenamiento es solamente un proxy (un estimador) del verdadero objetivo (generalizar).

4. *Explique el problema denominado “maldición de la dimensionalidad”.*

De acuerdo al autor, en espacios altamente dimensionales, nuestra intuición geométrica (habituada a una dimensionalidad reducida) no es necesariamente correcta (de hecho la mayoría de las veces no lo es). Como consecuencia, un método que “parece funcionar” en espacios de baja dimensionalidad, no necesariamente funcionará en espacios altamente dimensionales.

Otra expresión del problema, consiste en que el número de datos necesarios para cubrir uniformemente el espacio de características aumenta exponencialmente en la dimensionalidad. Como consecuencia, si las distribuciones que necesitamos aprender (condicionales en el caso de aprendizaje supervisado) fuesen uniformes o de amplio soporte, podríamos necesitar un número de datos de entrenamiento exponencialmente grande en la dimensionalidad del espacio característico.

5. *Según el autor, ¿Cuál es la clave de los proyectos exitosos en aprendizaje automático? ¿Qué se puede hacer si los datos no viene "preparados para el aprendizaje"?*

El factor más relevante es una buena ingeniería de las características que se usan para representar a los datos. Si tenemos muchos atributos independientes y correlacionados con la salida que deseamos aproximar el aprendizaje será más fácil. Por el contrario, si la respuesta es una función muy compleja de las características o es independiente de ellas, el aprendizaje resultará más difícil. Si los datos no vienen directamente en las condiciones previamente descritas, podemos construir características que si lo estén.s