

Prueba sorpresa

1. ¿En qué consiste la generalización en Machine Learning?

- La generalización en el Machine Learning hace referencia a la capacidad de un modelo de generar outputs útiles, generales respecto a datos diferentes a aquellos en un dataset (de entrenamiento).

2. ¿En qué consiste el compromiso **bias-variance**?

- Este término aparece cuando hablamos del overfitting de un modelo, para corregirlo buscamos regularizar el modelo, forzar el algoritmo a crear un modelo menos complejo; esto suele incrementar el bias pero reducir la varianza, y a eso se le conoce como el bias-variance tradeoff.

3. ¿Cuál es la diferencia entre el conjunto de entrenamiento, validación, evaluación?

- Estos conjuntos de datos suelen venir de el dataset con el cual trabaja el modelo, suelen tener una distribución de 70%, 30% entre entrenamiento y validación normalmente y se usan para evitar que el modelo se “acostumbre” a los datos del conjunto de entrenamiento, testear que esté haciendo el trabajo correctamente. El conjunto de evaluación puede hacer parte de nuestro dataset inicial (no necesariamente) pero se usa para evaluar la eficiencia de nuestro modelo para la tarea que se le está dando.

4. Explique la diferencia entre **regularización L1 y L2**.

- Si bien son 2 formas de regularizar un modelo, la diferencia está en la norma que usan y los efectos que esta termina teniendo en el modelo, L1 se usa para facilitar la interpretabilidad del modelo, L2 suele ser mejor a la hora de maximizar el rendimiento del mismo.

5. ¿En qué consiste la validación cruzada y para qué sirve?

- Creo que esto hacía referencia al “shift” a la hora de validar el modelo e implica en vez de alimentar el modelo con el orden que tiene el dataset alimentarlo con los datos de validación pero en un orden aleatorio y así asegurar que el modelo no está aprendiendo el orden del dataset si no una representación útil.

6. ¿Cuál es la diferencia entre **probabilidad y verosimilitud**?

- Probabilidad es la frecuencia con la que un evento ocurre en unas determinadas repeticiones, verosimilitud hace referencia a una función con la cual podemos inferir un valor a partir de unas observaciones.

7. ¿Cómo se puede garantizar que un modelo no sobre-aprenda?

- El término “sobre-aprender” hace referencia a un modelo que ha aprendido tanto un dataset que ahora sus outputs son réplicas del dataset y no se puede desplegar correctamente con datos nuevos, para esto existen herramientas que detienen el aprendizaje del modelo antes de tiempo, cuando el dataset de validación pierde efecto se para el aprendizaje.

8. ¿Cuál es la diferencia entre causalidad y correlación?

- La diferencia es que en causalidad un evento implica/modifica el segundo o viceversa y la correlación sólo implica que existe una relación entre 2 o más variables, sin embargo el resultado de estas puede estar ligada a alguna otra u no, para tener causalidad necesitamos correlación, pero la correlación no implica causalidad.

9. ¿Cómo se puede **determinar cuál es el algoritmo de Machine Learning** que aplica para un conjunto de datos en particular?

- Pues en general hay muchos algoritmos para muchas cosas, pero supongo que la distribución de los datos y/o como se interpretan es lo que define que herramientas

usar en los algoritmos, yo sé que las convoluciones son muy útiles a nivel de imágenes, y que hay algoritmos que tienen restricciones a nivel de inputs (solo aceptan positivos, son muy útiles para números entre 0 y 1), etc. Pero realmente desconozco si hay un algoritmo como tal para determinar algoritmos de Machine Learning según conjuntos de datos.

10. ¿En qué consiste el aprendizaje por refuerzo?

- El aprendizaje por refuerzo es un campo del Machine Learning que consiste en entrenar modelos que reciben estímulos de un “ambiente” a estos estímulos les corresponderá respuestas y estas respuestas tienen una recompensa, a partir de estas recompensas se entrena el modelo para que tenga la mejor respuesta ante un estímulo estipulado.

11. ¿Cuál es la principal diferencia entre Machine Learning y Deep Learning?

- Una diferencia podría ser que el Deep Learning es un campo del Machine Learning, pero realmente la diferencia más grande diría que es la interpretabilidad de estos 2 campos, el machine learning brinda muchas herramientas para lograr que una máquina aprenda, algunos más o menos interpretables, sin embargo el deep learning es una black box en la mayoría de los casos, es poco interpretable pero encuentra muy buenas representaciones de los datos, el problema es que son muy abstractas.

12. ¿Explique en que consiste el algoritmo de gradiente en descenso?

- Básicamente tomar una función, elegir un punto aleatorio y calcular su derivada, si esta es positiva, tomar un punto menor al elegido inicialmente, si la derivada es negativa, tomar un punto mayor al elegido inicialmente, este algoritmo busca encontrar los mínimos de una función, puntos donde la derivada es 0 o muy cercano a 0.

13. ¿Cuál es la diferencia entre **gradiente estocástico y gradiente en descenso?**

14. ¿Qué es una **matriz de confusión y para qué sirve?**

15. Defina precisión y **recall**

- Precisión es una métrica respecto a que tan similar son los outputs del modelo y los outputs esperados, ya sea que manejemos un algoritmo de ML supervisado, no-supervisado, recall es una métrica de cuántas veces nuestro algoritmo de ML predice nuestros verdaderos positivos, es decir, nuestra clase objetivo

16. ¿En qué consiste la **maldición de la dimensionalidad?**

- Esto era relacionado a variedades (?) “As the number of features or dimensions grow the amount of data we need to generalize accurately grows exponentially” -Att: Wikipedia

17. ¿Qué es el **shattering de un conjunto de puntos?**

18. ¿Qué es la **dimensión VC de un espacio de hipótesis?**

- Yo creo recordar que el espacio hipótesis era un espacio de gran dimensión donde presumimos existían nuestros inputs, qué es la dimensión VC no recuerdo.

19. ¿Cuál es la **dimensión VC de las redes neuronales profundas?**

20. ¿Qué relación existe en la dimensión VC de un espacio de hipótesis y el **número de muestras requeridas para que un modelo generalice?**

21. ¿En qué consiste el fenómeno del **doblo descenso?**

22. ¿Cuáles son las garantías de convergencia del **gradiente estocástico en un modelo de redes neuronales profundas?**