**TALLER REDES NEURONALES**

Diplomado en Big Data y Data Science

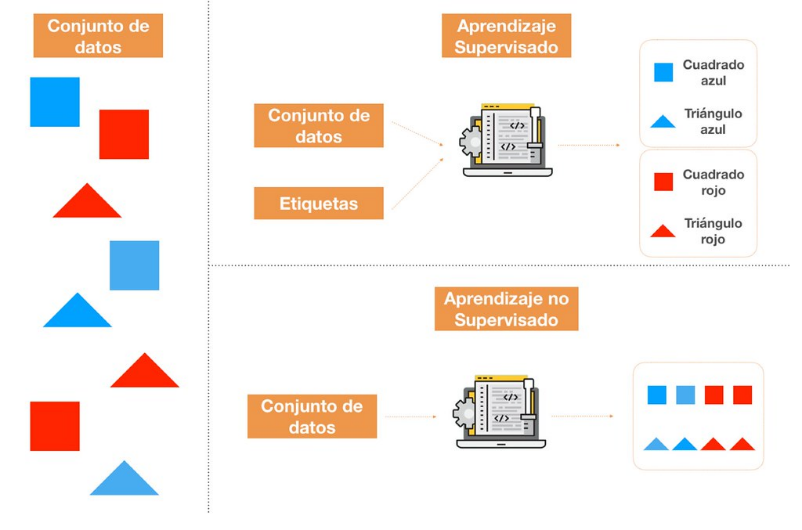
Docente: Nazly R. Hincapié

Estudiante: Natalia Santamaria Macias

1. ¿Qué es el aprendizaje supervisado y cuál es la diferencia fundamental con el aprendizaje no supervisado?

El aprendizaje supervisado, también conocido como machine learning supervisado, es una subcategoría de machine learning y la inteligencia artificial. Se define por su uso de conjuntos de datos etiquetados para entrenar algoritmos que clasifican datos o prevén resultados con precisión.

El aprendizaje supervisado se utiliza cuando se conocen las etiquetas de los datos, como en la clasificación o la regresión, mientras que el aprendizaje no supervisado se utiliza para explorar y descubrir patrones en los datos sin tener etiquetas.



1. ¿Qué significa el término clasificación en el contexto de aprendizaje de máquinas y cómo se diferencia de la regresión?

En el aprendizaje de máquinas, la clasificación se refiere al entrenamiento de un modelo en un conjunto de datos etiquetados para asignar puntos de datos a clases o para clasificar nuevos puntos de datos. En cambio, la regresión busca realizar la predicción de los valores numéricos a partir de los datos que representan el objeto.

1. ¿Cómo se define un modelo en el contexto del aprendizaje de máquinas?

Un modelo de aprendizaje automático es un archivo que se ha entrenado para reconocer determinados tipos de patrones. Puede entrenar un modelo con un conjunto de datos, y proporcionarle un algoritmo que puede usar para averiguar y obtener información de esos datos.

1. Explica la diferencia entre un conjunto de entrenamiento y un conjunto de prueba en el proceso de modelado

El conjunto de entrenamiento se refiere a un conjunto de datos que se utiliza para entrenar un modelo o algoritmo de aprendizaje automático.

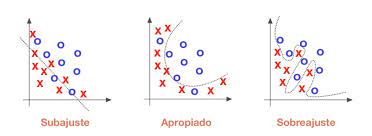
El conjunto de datos de prueba (o de test) es un conjunto de datos sobre el que vamos a aplicar nuestro modelo una vez ha sido entrenado para confirmar su rendimiento real.

1. ¿Qué es la función perdida y cuál es su objetivo en el aprendizaje supervisado?

Es un método para evaluar qué tan bien un algoritmo específico modela los datos otorgados. La función de pérdida, también denominada función objetivo durante el proceso de entrenamiento de la red, nos indica lo lejos que está en un momento dado, lo que la red nos ofrece como salida y el resultado que nosotros consideramos que es el correcto o deseado.

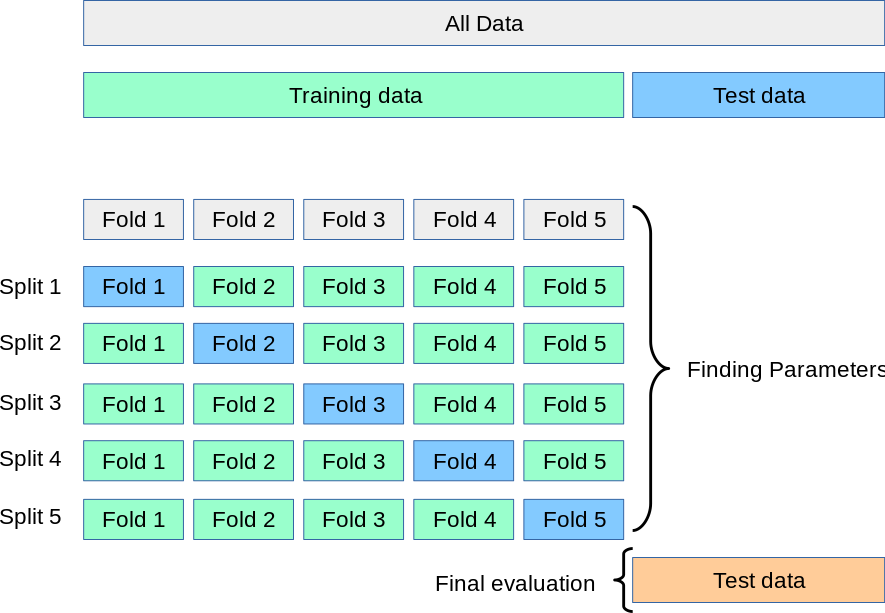
1. Describe qué es sobre ajuste y Cómo puede afectar la capacidad predictiva de un modelo.

El sobreajuste es un comportamiento de aprendizaje automático no deseado que se produce cuando el modelo de aprendizaje automático proporciona predicciones precisas para los datos de entrenamiento, pero no para los datos nuevos.



1. ¿Qué es la validación Cruzada y cuál es su utilidad en la evaluación de modelos?

La validación cruzada es una técnica para evaluar modelos de ML mediante el entrenamiento de varios modelos de ML en subconjuntos de los datos de entrada disponibles y evaluarlos con el subconjunto complementario de los datos. Ayuda a garantizar la exactitud de las predicciones del modelo del aprendizaje automático que se está utilizando.



1. Explica qué es la selección de características y Por qué es importante en el proceso de construcción de modelos

La Selección de Características es el proceso de seleccionar las más importante y/o relevante características de un conjunto de datos, con el objetivo de mejorar el rendimiento de predicción de los predictores, proporcionar predictores más rápidos y más rentables y proporcionar una mejor comprensión del proceso.

1. ¿Qué es la matriz de confusión y cómo se utiliza para evaluar el rendimiento de un modelo de clasificación?

La matriz de confusión es una herramienta muy útil para valorar cómo de bueno es un modelo clasificación basado en aprendizaje automático.

Ahora es el momento de descifrar y extraer información valiosa de esta representación tabular. La interpretación de la matriz de confusión parte de la base de entender las métricas: precisión, especificidad, sensibilidad, y valor F1.

1. ¿Qué es el sesgo y la varianza en el contexto de modelos de aprendizaje y cómo afectan el rendimiento del modelo?

El sesgo en los modelos de IA se refiere a la tendencia de los modelos a producir resultados que no son precisos o justos. Este sesgo puede provenir de muchos factores, incluyendo la calidad de los datos de entrenamiento, los algoritmos utilizados y los supuestos subyacentes en el modelo.

La varianza de un estimador es cuánto varía la predicción según los datos que utilicemos para el entrenamiento.