



TIC



▶ TALENTO
TECH
REGIÓN 3
CAUCA – NARIÑO
BOOTCAMP
ANÁLISIS DE DATOS





TIC



GUÍA INICIAL – INTEGRADOR UNIDAD 3

Contextualización de mis aprendizajes



TIC

Una de las habilidades principales de un analista de datos y un científico de datos es la capacidad de explorar los datos mediante técnicas estadísticas avanzadas que permiten encontrar relaciones entre datos similares, crear grupos de datos que comparten ciertas características y separar cada punto de datos según el tipo. En esta unidad se estudiarán los conceptos relacionados al aprendizaje de máquina clásico, teniendo en cuenta un enfoque centrado en los datos. Se explorarán los conceptos de regresión, clasificación, etiquetado de datos, preparación de la información y la creación y evaluación de modelos de aprendizaje de máquina. Se explorarán conceptos como los algoritmos de optimización que se emplean en el aprendizaje de máquina y se hará un taller para aplicar los conocimientos adquiridos.



Objetivo general



UNIDAD III

- **Objetivo general:** Al finalizar esta unidad el estudiante estará en capacidad de:
- **Comprender e implementar algoritmos de regresión**
- **Comprender e implementar algoritmos de clasificación**
- **Utilizar herramientas de machine learning para crear modelos**
- **Utilizar conjuntos de datos para entrenar modelos de regresión y clasificación**

Competencias a desarrollar

- Capacidad de abstracción, análisis y síntesis.
- Capacidad de investigación.
- Emplear técnicas de análisis exploratorio de datos.
- Aprendizaje de máquina

Capacidad de abstracción análisis y síntesis: Reconocer y entender los conceptos demostrados en la unidad.

Capacidad de investigación: Analizar con respecto a referencias externas de la información, capacidad de complementar utilizando herramientas disponibles.

Emplear técnicas de análisis exploratorio de datos: Utiliza las técnicas de análisis exploratorio para comprender el contenido de una base de datos.

Aprendizaje de máquina: Aplica conceptos estadísticos para crear modelos que expliquen a profundidad los detalles de los datos.

Activación de saberes previos

PLANTEAMIENTO DE LA SESIÓN

La solución del problema de regresión en machine learning consiste en encontrar un modelo que mejor se ajuste a los datos sin que el modelo se sobreentrene. Esto es, optimizar los parámetros para minimizar el error dadas ciertas restricciones. En esta lección se estudiarán los modelos de regresión y se implementarán utilizando el lenguaje Python.

MATERIALES

Guía para docentes del cuaderno de jupyter de regresión (archivo regresión.ipynb)

https://drive.google.com/file/d/li8iyRPZevnAqLsvKMLX0AnNQxSiGJSP/view?usp=drive_link