



SDC - Learning
Software Learning and Solutions

INGENIERÍA DE CARACTERÍSTICAS EN R



**MANUEL
SIGÜEÑAS**

Agile Data Scientist and
ScrumStudy Trainer



MARTES 17
6:00 PM [GMT-5]



FLOR MURGA
+51 960 781 999
Asesora Académica

¿Qué es?

Feature Engineering

Feature Selection

Feature Transformation

Feature Creation
(Encoding, Binning)

Feature Extraction
(Automated in Deep Learning)

¿Qué es?

Para Azure:

Es parte del Proceso de Ciencia de Datos en Equipo (TDSP). Considera dos componentes Feature engineering y Feature selection.

- **Ingeniería de características:** Intenta crear características relevantes adicionales a partir de las características sin procesar existentes en los datos.
- **Selección de características:** Selecciona un subconjunto relevante de características o variables en un intento de reducir la dimensionalidad del problema de entrenamiento.

<https://docs.microsoft.com/en-us/azure/architecture/data-science-process/select-features>

Para AWS:

Es un proceso para seleccionar y transformar variables al crear un modelo de Machine Learning. Generalmente incluye la creación de features, la transformación de features, la extracción de features y la selección de features. <https://docs.aws.amazon.com/wellarchitected/latest/machine-learning-lens/feature-engineering.html>

Requerimiento de los modelos

Características de clientes y desean realizar un proyecto de **segmentación de clientes**

K-means (No supervisado)

- ¿Cómo ingresamos las variables a este modelo? ¿Qué variables? ¿están en el formato apropiado? ¿relación de variables?
- Escalamiento: estandarización o normalización
- Está basado en distancias (euclideana, coseno) y que las variables estén en una distinta unidad de medida afectaría el proceso de aprendizaje

Requerimiento de los modelos

Naive Bayes (Supervisado)

- Los predictores deben ser discretos: por ejemplo, variables como edad deben ser discretizadas
- Probabilidad condicional necesitamos que los predictores estén discretizados.

Regresión (Supervisado)

- La variable target u objetivo no tiene comportamiento normal ¿qué hacer?
- Transformar la variable: SoftMax, MinMax, entre otras.



¡MUCHAS GRACIAS!



INICIO: 21 MAYO
SÁBADO: 9:30 AM a 1PM



FLOR MURGA
+51 960 781 999
Asesora Académica

SDC - Learning
Software Learning and Solutions