

LENGUAJE R MACHINE LEARNING

Objetivos del curso:

Aplicar técnicas de análisis y visualización de datos en un conjunto de datos complejo para problemas de machine learning.

El curso está diseñado para que una persona con conocimientos básicos de programación alcance un nivel de conocimiento que le permita ser funcional a la hora de abordar proyectos de machine learning utilizando Python y Lenguaje R a través de la "ide" "Spyder".

Duración, Fechas y Horario:

El curso de 30 horas. Pendiente de fijar fechas y horario.

Formato:

Presencial, en las instalaciones del centro de formación C2B en la calle Pérez Galdós 22 bajo 48010 Bilbao

Programa:

Outliers:

- Definición.
- Problemas de su no tratamiento.
- Detección de outliers:
 - Gráficamente.
 - Extracción de los outliers.
 - Determinación de las observaciones que incluyen outliers.
 - Extracción de los valores de los outliers.
- Tratamiento de los outliers:
 - Eliminación de las observaciones.
 - Transformación en NAs.
 - Creación de nuevas variables dummy de forma masiva.
 - Tratamiento individualizado.

Valores perdidos:

- Cuantificación de los valores perdidos:
 - Total.
 - Por variables.
 - Por observaciones.
- Análisis gráfico de los valores perdidos.
- Eliminación de los valores perdidos variable por variables.
- Eliminación masiva de los valores perdidos.
- Eliminación de variables con alto porcentaje de valores perdidos.
- Imputación por valores estadísticos (media, moda, mediana)
- Métodos de imputación múltiple.
- Creación de modelos para la imputación.

Estadísticos básicos: Cálculo e interpretación de la:

- Media.
- Mediana.
- Moda
- Varianza.



CAMINOS EUSKADI

- Covarianza.
- Correlación.

Regresión como base de la modelización:

- Regresión lineal simple.
- Regresión lineal general.
- Regresión polinómica.
- Regresión logística.
- Error Cuadrático Medio y Error Absoluto.
- Estimación de la recta de regresión.
- Análisis de la significatividad global del modelo.
- Análisis de la significatividad individual de las variables.
- Cálculo del Coeficiente de Determinación.
- Dummy Variables.

Machine Learning

- Definición.
- Mapa conceptual.
- Modelos supervisados
- Modelos no supervisados
- Series temporales univariantes