

Algoritmos de agrupamiento no supervisado *k-means* con python

Síntesis: Aprendizaje no supervisado *K-means*



Como se ha visto en el desarrollo de este ejercicio es necesario que antes de aplicar cualquier algoritmo de **machine learning** se realice la exploración de los datos y detectar patrones que luego permitirán identificar cuál es el algoritmo de aprendizaje que se puede usar, el entendimiento de los datos es lo que más debería consumir tiempo.



El algoritmo de aprendizaje no supervisado, específicamente *K-means* está basado en la identificación de patrones como *clúster* y los datos toman una gran importancia en el aprendizaje automático, conceptualizando la importancia de la ciencia de datos, *bigdata*, tipos de variables y códigos fuente en Python.

Con el algoritmo *K-means* se pueden revisar patrones en diferentes exploraciones o en una muestra determinada obtenida aleatoriamente y en nuevas muestras diferentes al *daset* inicial. Con Python se hace la exploración de datos usando anaconda y jupyter *lab*, donde se toman las estadísticas descriptivas del *dataset* y se pueden realizar gráficos como **boxplots**, gráficos de dispersión y gráficos de pares o **paiplot**.

Así pues, un resumen de lo visto en el presente componente podrá ser visualizado en el siguiente mapa conceptual:

