

Plan Formativo: Ciencia de Datos	Nivel de Dificultad
Módulo: Aprendizaje no Supervisado	Medio
Tema: K-means	
Intención del aprendizaje o aprendizaje esperado:	
<ul style="list-style-type: none">Elaborar un modelo predictivo aplicando el algoritmo K-Means utilizando lenguaje Python para resolver un problema de clusterización	
Ejercicios planteados	
Clasificación de Clientes según su personalidad.	
<p>Luego de haber cerrado su cápsula de aprendizaje de clasificación, se le ha determinado que averigüe los intereses que pueden tener los clientes de la empresa KEPPLER, esto es necesario, para que se entregue la información al departamento de marketing y puedan realizar campañas para fidelizar a sus clientes.</p> <p>En el archivo clientes.csv encontrará el registro de cada uno de ellos, además el archivo contiene 9 categorías -actividades laborales- que son:</p> <ul style="list-style-type: none">Actor/actrizCantanteModelo	

- Tv, series
 - Radio
 - Tecnología
 - Deportes
 - Política
 - Escritor
1. Realice un resumen estadístico de sus datos.
 2. Entregue gráficos que acompañen los hallazgos encontrados en el punto 1.
 3. Encuentre una relación o grafique las variables: **op, ex y ag, además de las categorías.**
 4. Determine cuántos k o clúster son adecuados para esta consulta.
 5. Realice un modelo de k means con los cluster definidos en el punto 4.
 6. Entregue una representación gráfica de los cluster.
 7. A qué conclusión llega y cuál es su recomendación como experto.

Caso

Preguntas guía

- Clasificación
- Agrupación

- Métodos del código
- Gráficas 3D

Recursos Bibliográficos:

Referencias

[1] K-means

https://www.uniovi.es/computo/laboratorios_py/kmeans/kmeans.html

[2] Elección del valor de K

<https://jarroba.com/seleccion-del-numero-optimo-clusters/>

[3] Qué es el Tradeoff Bias – Variance

<https://www.themachinelearners.com/tradeoff-bias-variance/>