

TFG

Aarón Casado Monge¹, Juan José Cuadrado Gallego¹

University of Alcalá, Polytechnic School, Computer Science Department, Scientific and Technological Campus, Polytechnic Building. Office: 0243, 28805, Alcalá de Henares, Madrid, Spain

Abstract

Clusterización (qué es y para qué sirve) - Técnicas (para qué sirven) - Paquetes

Keywords: BigData, Statistics, Biostatistics, Data Science, Clustering

1. Introducción

2. Clustering

Cluster Analysis o **Clustering**, adaptado al español como **Clusterización**, Agrupamiento o
5 Análisis de Grupos es un método de clasificación no supervisada, la cual busca definir, para una característica determinada o Suceso Elemental (SE) ¹, un conjunto de grupos de observaciones (suceso) ² con valores cercanos. Estos grupos son los denominados clusters o grupos y permiten a partir de los
10 diferentes sucesos elementales que configuran dicho suceso, asignar dicho SE al mismo. Clustering nos permite definir los valores de cada cluster durante el proceso de clasificación [5].

15 Este procedimiento es uno de los muchos que ofrece **Data Mining**, una rama de **Data Science** (Ciencia de los Datos), que se ha posicionado rápidamente como una de las disciplinas más influyentes en plena era de la información, donde
20 el volumen de datos generado diariamente y que se almacenan en bases de datos es inmenso, pero para poder darle uso, es necesario organizarlos y agruparlos de manera adecuada (Data Warehousing), analizarlos para obtener información y
25 conocimiento de los mismos (Data Mining) y presentar los resultados de manera apropiada (Visualización). Y aunque Data Science es todavía joven y carece de una definición oficial, podríamos definirla

como la unión de la Estadística, la Inteligencia Artificial y la Programación aplicados a ese gran volumen de datos (Big Data ³) permitiendo extraer conocimiento a partir de ellos [4].

Email addresses: aaron.casado@uah.es (Aarón Casado Monge), jcg@uah.es (Juan José Cuadrado Gallego)

¹Definir suceso elemental

²Definir Suceso

³Definir Big Data

3. Referencias

- [1] Ostasiewicz, W. (2014). The Emergence of Statistical Science. Silesian Statistical Review / Slaski Przegląd Statystyczny, 12, 75–81.
- [2] https://en.wikipedia.org/wiki/History_of_statistics
- [3] Everyman's Science
- [4] Apuntes JJ Data Science
- [5] Apuntes JJ clustering
<https://doi.org/10.15611/sps.2014.12.04>
<https://en.wikipedia.org/wiki/Biostatistics>
<https://en.wikipedia.org/wiki/Statistics>
<https://journal.emwa.org/statistics/history-of-biostatistics/> (PDF samename)
- Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques (3rd ed., p. 740). 225 Wyman Street, Waltham, MA 02451, USA: Morgan Kaufmann Publishers, Elsevier.
- <https://normas-apa.org/referencias/citar-diccionario/>
<https://www.scribbr.es/detector-de-plagio/>
<https://tablesgenerator.com/>
elsevier dos páginas latex