

TP FINAL SIA

Contenido

Introducción:.....	2
Aclaraciones importantes.....	3

Introducción:

El laboratorio de Inteligencia Artificial de la empresa Ultralistic nos ha solicitado realizar nuestra propia implementación del algoritmo K-Medias, ya que posee un [dataset](#) sobre la calidad vitivinícola y pretende analizar cuales son las agrupaciones relevantes y “naturales” del mismo, más allá de las clases presentes en dicho dataset, y para ello desea utilizar el mencionado algoritmo.

Para la implementación se establecieron los siguientes requerimientos:

- a) Se puede escribir en el lenguaje de programación preferido por el grupo de trabajo.
- b) No se admiten implementaciones que posean la utilización de librerías o frameworks (o cualquier cosa parecida) que ya contengan el algoritmo solicitado.
- c) Debe realizarse una implementación del algoritmo sin “vectorizar” y otra “vectorizada”.
- d) Esta permitido utilizar librerías o frameworks (o cualquier cosa parecida) para soporte de los cálculos matemáticos, como pueden ser las operaciones de matrices, cálculos estadísticos, etc.
- e) Las implementaciones deben ser flexibles, es decir que deben de poder setearse la cantidad de clusters que se desean y además se debe poder manejar cualquier número de ejemplos y atributos del dataset correspondiente.
- f) Los tipos de datos que deben soportar las implementaciones propias de K-medias, por supuesto que deben ser normalizados de alguna forma para evitar inconvenientes.
- g) Se debe realizar una comparativa entre las implementaciones propias y la implementación de alguna librería o framework disponible en el mercado (dicha librería o framework puede ser que se encuentre escrita en otro lenguaje de programación).

Teniendo en cuenta dichos requerimientos se debe lograr:

- 1) Construir las implementaciones de K-medias solicitadas cumpliendo con lo establecido en los requerimientos.
- 2) Realizar una comparativa (puede ser gráfica) que contenga múltiples ejecuciones del dataset para diferentes números de clusters utilizando tanto las propias implementaciones de K-medias como las de terceros.
- 3) Se espera que las implementaciones propias de K-Medias posean una interfaz amigable y permitan “predecir” a qué grupo (o cluster) pertenecería un elemento (o registro) que se encuentre fuera del dataset .

Aclaraciones importantes

El TP puede ser realizado en grupos de hasta 2 integrantes y deben estar todos los integrantes en la mesa examinadora.

Se pueden solicitar al profesor, luego de que finalice el cursado, (por TEAMS) clases de consulta referidas al TP final.

Se destaca que en la mesa de examen final se espera:

1. Presentación de un informe en formato de documento PDF con la prolijidad acorde. (10%)
2. Explicación teórico/práctica del trabajo practico por parte de **todos** los integrantes del grupo (40%)
3. Construir una aplicación que corra sin presentar errores. (20%)
4. Que dicha aplicación cuente con una interfaz gráfica funcional de manera que permita interactuar con las implementaciones propias y cuente con alguna funcionalidad extra que los integrantes del grupo crean necesaria. (30%)