

# Minería de Datos

## Lab. #6

Ivan Saavedra, Ph.D.

[saavedrai@uninorte.edu.co](mailto:saavedrai@uninorte.edu.co)

Universidad del Norte  
División de Ingenierías  
Dpto. Ingeniería de Sistemas



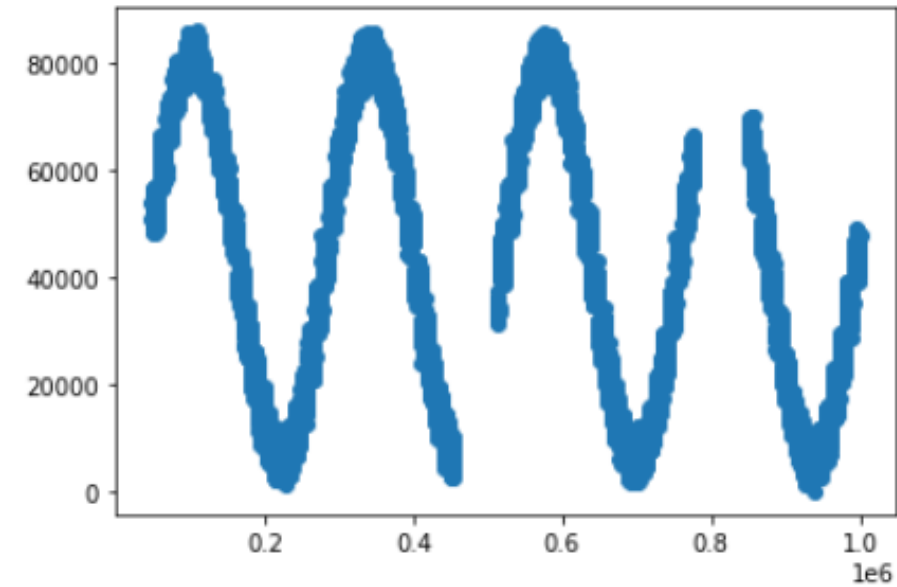
202030

# 1. Clustering Sin Signals

Aplique las técnicas de clustering K-means y DBSCAN para encontrar grupos de datos en el dataset suministrado (Sin.xlsx fuente: <http://cs.joensuu.fi/sipu/datasets/> ).

## Metodología

- Definir parámetros de los modelos
- Ajustar el modelo
- Compare los resultados de las dos técnicas y mencione cual seria la mejor para este caso



## 2. Clustering Fifa Players

Aplique K-means clustering en el dataset suministrado de jugadores Fifa con el objetivo de encontrar grupos de jugadores que presenten similitudes y así proveer múltiples opciones a la hora de obtener nuevos jugadores en un equipo de futbol.

### Metodología

- Preparar el dataset
- Definir parámetros del modelo
- Ajustar el modelo
- Evaluar el modelo
- Realizar este ejercicio
  1. Utilizando los siguientes features (Age and Overall)
  2. Utilizando los siguientes features (Age, Overall, Value, International Reputation, Weak Foot, y Skill Moves)



<https://www.transfermarkt.com/>

# Clustering

## **Entregable:**

- Archivos de Excel con los datos de entrada
- Un archivo de Jupyter Notebook con el desarrollo del análisis
- Las conclusiones y respuestas al objetivo del análisis deben ser contestadas en el mismo notebook.
- Se sugiere que comente las secciones de manera adecuada para una mejor interpretación de su análisis.
- La fecha de entrega es el día 15 de Noviembre del 2020 vía catalogo web enlace de laboratorios.