

T.M.FAJAR PRAMUDYA

1301172735

IFIK-41-03

## LATAR BELAKANG

Data mining merupakan salah satu teknik yang digunakan untuk mendapatkan pengetahuan baru dengan memanfaatkan jumlah data yang sangat besar. Teknik data mining digunakan untuk memeriksa basis data berukuran besar sebagai cara untuk menemukan pola yang baru dan berguna. Tidak semua pekerjaan pencarian informasi dinyatakan sebagai data mining.

Langkah-langkah yang terlibat dalam preprocessing data meliputi menggabungkan data dari berbagai sumber, membersihkan (cleaning) data untuk membuang noise dan observasi duplikat, dan menyeleksi record dan fitur yang relevan untuk pekerjaan data mining.

Karena terdapat banyak cara mengumpulkan dan menyimpan data, tahapan preprocessing data merupakan langkah yang banyak menghabiskan waktu dalam knowledge discovery in databases (KDD).

Hasil dari data mining sering kali diintegrasikan dengan decision support system (DSS). Sebagai contoh, dalam aplikasi bisnis informasi yang dihasilkan oleh data mining dapat diintegrasikan dengan tool manajemen kampanye produk sehingga promosi pemasaran yang efektif yang dilaksanakan dan dapat diuji.

Terdapat beberapa teknik di dalam data mining diantaranya yaitu : Clustering dan Classification.

Clustering adalah metode penganalisaan data, yang sering dimasukkan sebagai salah satu metode Data Mining, yang memiliki tujuan untuk mengelompokkan data dengan karakteristik yang sama ke suatu wilayah yang sama dan data dengan karakteristik yang berbeda ke wilayah yang lain.

Classification adalah sebuah metode dari Data Mining yang digunakan untuk memprediksi katagori atau kelas dari suatu data instance berdasarkan sekumpulan atribut-atribut dari data tersebut.

## METODE

Metode yang digunakan adalah K-mens.

K-Means Clustering adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan tanpa supervisi (unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi.

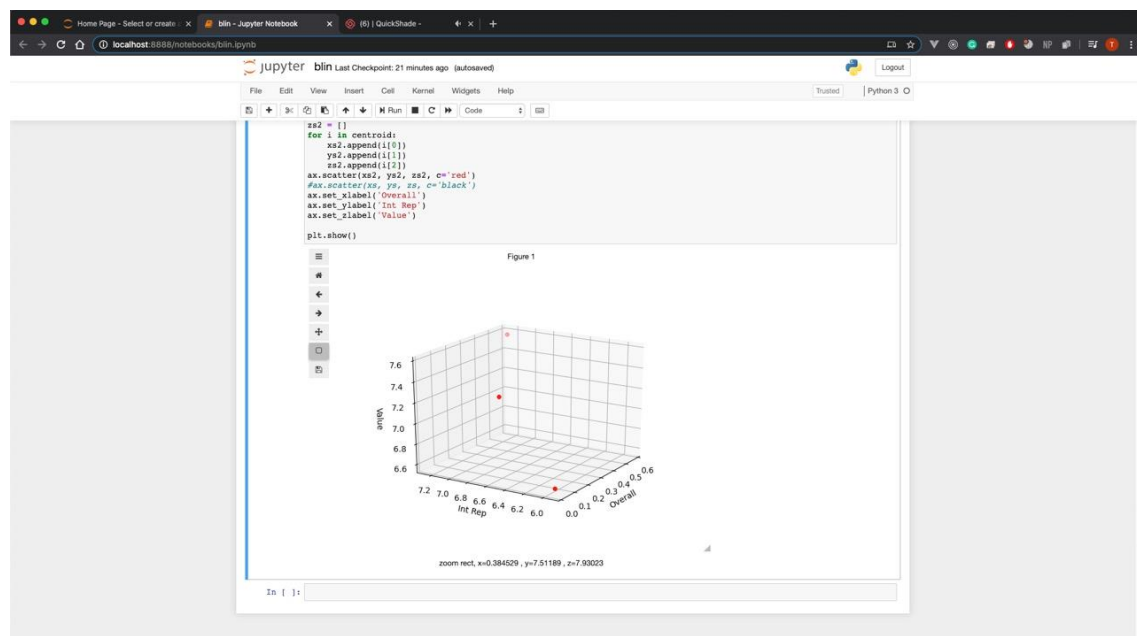
Terdapat dua jenis data clustering yang sering dipergunakan dalam proses pengelompokan data yaitu Hierarchical dan Non-Hierarchical, dan K-Means merupakan salah satu metode data clustering non-hierarchical atau Partitional Clustering.

Metode K-Means Clustering berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain.

Metode K-Means Clustering bertujuan untuk meminimalisasikan objective function yang diset dalam proses clustering dengan cara meminimalkan variasi antar data yang ada di dalam suatu cluster dan memaksimalkan variasi dengan data yang ada di cluster lainnya.

## KESIMPULAN

Untuk kedepan mungkin bisa menggunakan metode lain untuk Teknik clustering dan classification.



Link video :

<https://www.loom.com/share/cfdd46e7fd3e4cf4908cdc745158410d>