

**Nama Kelompok :**

- Billy Aprillio (2017-104)
- Muhammad Sahedi N (2017-086)
- L Yasril Imam (2017-051)

**Judul Artikel :**

COMPLEX-IT: A Case-Based Modeling and Scenario Simulation Platform for Social Inquiry

**Metode :** Framework CBBS (case-based scenario simulation), K-means Clustering, ANN

Penelitian ini bertujuan untuk mensimulasikan setiap data yang di cluster, melihat setiap perbedaan major dan minor pada setiap cluster yang ada. dengan mengetahui perbedaan major dan minor pada setiap cluster yang ada maka dapat memberikan hasil analisis pada data tersebut dan mendapatkan variable atau atribut yang paling berpengaruh terhadap clustering pada data tersebut.

Pada penelitian ini, tahapan pertama yang dilakukan oleh peneliti adalah preprocessing pada data dengan menyeleksi data mana yang perlu digunakan untuk penelitian tersebut. Kemudian peneliti mulai mengelompokkan data untuk menganalisis data mana yang memiliki hubungan antara 1 data dengan data yang lainnya. Dengan menggunakan K-means clustering data akan di cluster menjadi  $n$ -cluster ( $n$  = jumlah cluster yang di inginkan). Selanjutnya melakukan train data dengan menggunakan metode SOM (Self Organizing Maps). Lalu melakukan simulasi mapping dengan menggunakan CBBS (case-based scenario simulation) untuk melihat major dan minor setiap cluster dan setiap simulasi yang dilakukan.

**Dataset :**

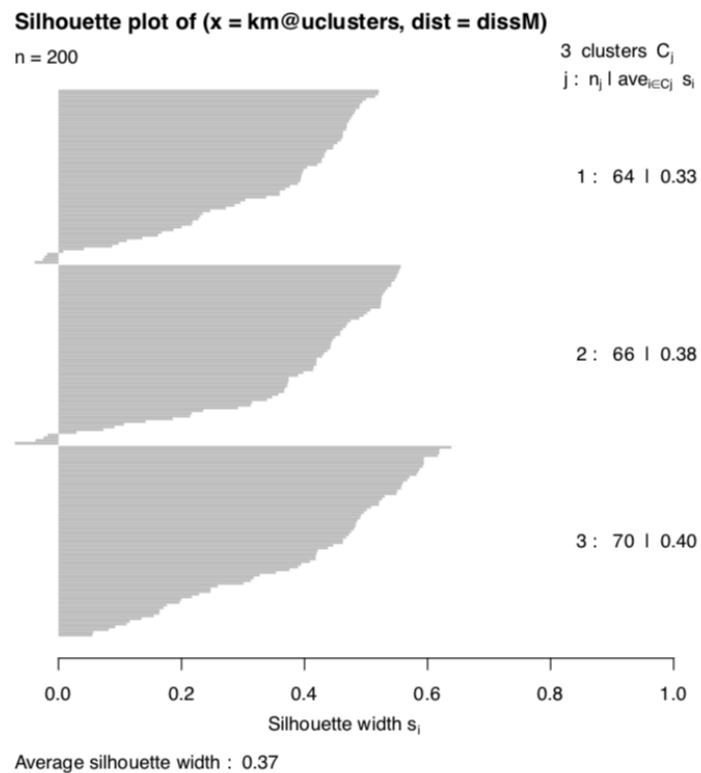
Dataset yang di gunakan adalah data mall customer. data tersebut memiliki 200 data. Dataset berisi tentang jenis kelamin, umur, pendapatan tahunan, nilai pengeluaran untuk mengetahui seperti apa target pelanggan mall agar lebih mudah dalam melakukan pemasaran.

**Perubahan yang di lakukan:**

Kami menggunakan dataset customer mall yang di dalamnya terindikasikan atribut yang memiliki nilai seberapa besar seorang customer menggunakan uangnya. Tujuan kami adalah melakukan clustering secara acak pada data tersebut lalu melakukan prediksi untuk mengetahui minor dan major terhadap variable tersebut dan mengetahui seberapa besar pengaruh variable tersebut. Data kami dapatkan dari Kaggle (<https://www.kaggle.com/vjchoudhary7/customer-segmentation-tutorial-in-python>). Analisis yang kami lakukan menggunakan metode yang sama, dengan menggunakan SOM untuk train data dan mapping data untuk koordinat x dan y, lalu mengcluster data tersebut menjadi beberapa cluster dan melakukan simulasi menggunakan CBBS pada tahap akhir untuk mendapatkan hasil analisa dari setiap quadrant (Banyaknya simulasi yang dilakukan). Kami mendapatkan hasil bahwa pada variable atau atribut Spending Score sangat paling berpengaruh pada data yang lain karena data tersebut memiliki banyak major yang berada pada setiap cluster dan setiap simulasi.

Quadrant 5	Quadrant 10	Quadrant 15	Quadrant 20	Quadrant 25
Quadrant 4	Quadrant 9	Quadrant 14	Quadrant 19	Quadrant 24
Quadrant 3	Quadrant 8	Quadrant 13	Quadrant 18	Quadrant 23
Quadrant 2	Quadrant 7	Quadrant 12	Quadrant 17	Quadrant 22
Quadrant 1	Quadrant 6	Quadrant 11	Quadrant 16	Quadrant 21

*Gambar 1 Analisis Quadrant*



*Gambar 2 Hasil Pseudo dan Silhoutte*

**Referensi :**

<https://arxiv.org/abs/2003.03099>

