

FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS STIKUBANK (UNISBANK) SEMARANG

Program Studi : Sistem Informasi
Tugas Akhir Sarjana Komputer
Semester Ganjil Tahun 2020/2021

***Clustering Kecelakaan Lalu Lintas Menggunakan
Algoritma K-Means***

Sofiani Maya Rodriques
NIM : 16.01.55.0039

ABSTRAK

Kecelakaan lalu lintas merupakan kejadian yang banyak terjadi di jalan raya, yang dapat mengakibatkan kerugian harta benda dan juga menimbulkan luka, sampai meninggal dunia. Kota Semarang adalah kota besar dengan banyak jalan raya sehingga kejadian kecelakaan lalu lintas dapat terjadi kapanpun dan tidak terduga. Secara statistik angka kecelakaan lalu lintas dari tahun ke tahun semakin meningkat, kejadian tersebut terekam oleh pihak Satlantas Kota Semarang yang dapat diolah agar menghasilkan informasi mengenai tingkat kecelakaan lalu lintas.

Maka permasalahan yang diteliti adalah bagaimana melakukan pengelompokan data lokasi kejadian kecelakaan lalu lintas berdasarkan data lokasi kejadian kecelakaan (kecamatan), jenis kecelakaan, karakteristik pelaku (usia, jenis kelamin, profesi), dan jenis kendaraan, dengan menerapkan metode teknik data mining yaitu *Clustering* dengan menggunakan kombinasi metode *Silhouette* dan *K-Means*. Metode *Silhouette* digunakan untuk mencari optimasi jumlah kluster terbaik menggunakan Bahasa R.

Hasil yang didapat dari proses *Silhouette* adalah 4 kluster, dimana nilai *average silhouette width* tertinggi pada grafik 0.42 yaitu menunjukan pada kluster ke 2. Algoritma *K-Means* digunakan untuk mengelompokan data kedalam 4 Kluster, dengan hasil rata-rata tertinggi ada pada Kluster4.

Kata Kunci : *Kecelakaan lalu lintas, Data Mining, Clustering, Silhouette*

Abstract

Traffic accidents are many incidents that occur on the highway, which can result in property losses and also cause injuries, until death. Semarang is a big city with many highways so that traffic accidents can happen at any time and unexpectedly. Statistically, the number of traffic accidents from year to year is increasing, the incident is recorded by the Satlantas of Semarang City which can be processed in order to produce information about the rate of traffic accidents.

So the problem examined is how to group the data of traffic accident locations based on data on the location of accidents (subdistricts), types of accidents, characteristics of perpetrators (age, gender, profession), and types of vehicles, by applying data mining techniques, namely Clustering using a combination of Silhouette and K-Means methods. Silhouette method is used to find the best cluster number optimization using R Language.

The results obtained from the Silhouette process are 4 clusters, where the highest average silhouette width value in the graph 0.42 is indicated in the 2nd cluster. The K-Means algorithm is used to group data into 4 Clusters, with the highest average results being cluster4.

Keywords: *Traffic Accident, Data Mining, Clustering, Silhouette*

Semarang, 20 Januari 2021

Pembimbing



(Kristophorus Hadiono, Ph.D)

NIDN: 0622027601