

**PELABELAN DATA TRAINING ALGORITMA NAËVE BAYES
MENGUNAKAN ALGORITMA K-MEANS UNTUK KLASIFIKASI
RESIKO KECELAKAAN KENDARAAN BERMOTOR DI KOTA
SEMARANG**

AZIZATUL MUNAWAROH

(Pembimbing : Aris Marjuni, S.Si, M.Kom)

Teknik Informatika - S1, FIK, Universitas Dian Nuswantoro

www.dinus.ac.id

Email : 111201207271@mhs.dinus.ac.id

ABSTRAK

Jumlah kecelakaan lalu lintas di kota Semarang, beberapa tahun belakangan ini menunjukkan angka kenaikan yang cukup signifikan. Hal ini diakibatkan oleh banyaknya pengemudi kendaraan yang kurang bahkan tidak sadar sama sekali akan keselamatan diri sendiri dalam mengemudi seperti melanggar peraturan lalu lintas, tidak menggunakan standart keselamatan yang memadai, dan lain sebagainya. Tujuan dilaksanakannya penelitian ini adalah untuk mengetahui dimana lokasi yang sering terjadi kecelakaan, khususnya di kota semarang. Dalam penelitian ini, akan dilakukan proses clustering dan classification untuk mengolah data. Sedangkan untuk algoritma yang digunakan dalam proses clustering adalah algoritma k-means dan untuk proses classification menggunakan algoritma naËve bayes. NaËve Bayes merupakan teknik prediksi berbasis probabilistik berdasar penerapan teorema atau aturan bayes, Sehingga merupakan algoritma klasifikasi sederhana, cepat namun memiliki akurasi tinggi. Clustering akan di bagi menjadi 3 cluster yaitu tidak rawan, rawan, dan sangat rawan. Atribut yang digunakan adalah kendaraan korban, waktu kejadian, wilayah kejadian, dan factor penyebab kecelakaan. Proses pengujian menggunakan Rapid Miner dengan klasifikasi menghasilkan akurasi global sebesar 99.59%. Hasil dari penelitian ini digunakan untuk menentukan daerah di Kota Semarang yang rawan terhadap kecelakaan lalu lintas. Algoritma K-Means sangat membantu dalam proses pelabelan data training Algoritma NaËve Bayes.

Kata Kunci : Kecelakaan Lalu Lintas, Clustering, Classification, K-Means, NaËve Bayes

TRAINING DATA LABELING OF NAIVE BAYES ALGORITHM USING K-MEANS ALGORITHM FOR RISK CLASSIFICATION OF TRAFFIC ACCIDENT AT SEMARANG

AZIZATUL MUNAWAROH

(Lecturer : Aris Marjuni, S.Si, M.Kom)

*Bachelor of Informatics Engineering - S1, Faculty of Computer
Science, DINUS University*

www.dinus.ac.id

Email : 111201207271@mhs.dinus.ac.id

ABSTRACT

The number of traffic accidents in the city of Semarang , for the last years, showed a significant increase . It is caused by a number of vehicle drivers who lack even not at all be aware of the safety of themselves in driving such a traffic violation , not using adequate safety standards , and etc. The purpose of this study is to determine where the location have a high-risk traffic accident, especially in the city of Semarang . In this study, will do clustering and classification process to processed the data . As for the algorithm used in the process of the clustering algorithm is k -means and to the process of classification is naive Bayes algorithm. Naive Bayes is a simple probabilistic-based prediction technique are based on the application of Bayes theorem or rule, so that is one of the classification algorithms is simple, fast but has a high accuracy. Clustering will be divided into three clusters, namely not prone, prone, and very vulnerable. The attribute used is the victim vehicle, the time of events, area events, and factor in the accident. The testing process using Rapid Miner classification generate global accuracy of 99.59%. The results of this research are used to determine the affluent in the city of Semarang prone to traffic accidents. K-Means algorithm is very helpful in the process of labeling the training data Naive Bayes algorithm.

Keyword : Traffic Accidents, Clustering, Classification, K-Means, Naive Bayes