

IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKKAN WILAYAH PELANGGARAN LALU LINTAS MENGGUNAKAN METODE K-MEANS PADA POLRES BENGKULU

Devi Sartika¹, Juju Jumadi²

Universitas Dehasen, Bengkulu

e-mail: ¹devisartika@unived.ac.id, ²juju.jumadi@unived.ac.id

Abstract: Numbers of congestion and especially traffic violations occur everywhere in the city of Bengkulu, especially at city junctions. The large number of road users who do not comply with traffic regulations properly, every day can increase the rate of accidents and violations of traffic rules in driving in the Bengkulu City area. Due to the large piles of data on traffic violations in the jurisdiction of the Bengkulu City Police, each year only this data is processed into Microsoft Excel, making the data irregular. The process of extracting useful information in order to obtain violation data can be carried out using data mining methods, one of which is k-means clustering. The implementation of the system uses the Visual Basic 2010 programming language. From the results of the tests conducted, it can be concluded that the areas with the most violations are the Tanah Patah and Betungan areas, for areas with violations in the moderate group there are Sawah Lebar, Pintu Batu, Padang Harapan, Kebun Tebeng and Simpang Nakau and for areas with low violations are the Swamp Makmur and Sungai Hitam areas.

Keywords: K-Means, Clustering, Offenses

Abstrak: Angka kemacetan dan terutama pelanggaran lalu lintas terjadi dimana-mana di daerah Kota Bengkulu terutama pada simpang-simpang kota. Banyaknya pengguna jalan yang tidak mematuhi peraturan berlalu lintas dengan baik, setiap harinya dapat menambah tingkat kecelakaan dan pelanggaran tata tertib lalu lintas dalam berkendara pada wilayah Kota Bengkulu. Dikarenakan banyaknya tumpukan data pelanggaran lalu lintas di wilayah hukum Polres Kota Bengkulu, setiap tahunnya yang hanya data tersebut diproses ke dalam Microsoft Excel sehingga membuat data tidak beraturan. Proses ekstraksi informasi yang berguna dalam rangka memperoleh data pelanggaran dapat dilakukan dengan menggunakan metode data mining yang salah satunya adalah k-means clustering. Implementasi sistem menggunakan bahasa pemrograman Visual Basic 2010. Dari hasil pengujian yang dilakukan dapat disimpulkan bahwa pengelompokan wilayah pelanggaran terbanyak adalah wilayah Tanah Patah dan Betungan, untuk wilayah dengan pelanggaran di kelompok sedang ada Sawah Lebar, Pintu Batu, Padang Harapan, Kebun Tebeng dan Simpang Nakau dan untuk wilayah dengan pelanggaran rendah adalah wilayah Rawa Makmur dan Sungai Hitam.

Kata kunci: K-Means, Clustering, Pelanggaran

PENDAHULUAN

Istilah data mining memiliki beberapa pandangan, seperti knowledge discovery ataupun pattern recognition. Istilah knowledge discovery atau penemuan pengetahuan tepat digunakan karena tujuan utama dari data mining memang untuk mendapatkan pengetahuan

yang masih tersembunyi di dalam bongkahan data. Sedangkan istilah untuk pattern recognition atau pengenalan pola tepat untuk digunakan karena guna menemukan pola yang tersembunyi di dalam bongkahan data (Nabila, Isnain, et al, 2021).

Data mining adalah proses yang mempekerjakan satu atau lebih teknik

pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis. Beberapa teknik dan sifat data mining yaitu Classification (Predictive), Clustering (Descriptive), Association Rule Discovery (Descriptive), Regression (Predictive), Deviation Detection (Predictive) (Silwattananusarn & Tuamsuk, 2012).

Data mining dibagi menjadi beberapa kelompok yaitu, deskripsi, estimasi, prediksi, klasifikasi, pengklasteran, dan asosiasi (Fajrin & Algifanri, 2018).

Clustering mengacu pada pengelompokan seperti record, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Cluster adalah kumpulan dari record yang memiliki kemiripan satu sama lain, dan berbeda dengan record di klaster lain. Clustering mencoba untuk membagi seluruh kumpulan data menjadi kelompok-kelompok yang relatif memiliki kemiripan, di mana kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal (Nabila, et al, 2021)

Clustering merupakan klasifikasi tanpa pengawasan dan kinerja merupakan proses partisi sekumpulan objek data dari satu set menjadi beberapa kelas. Hal ini dapat dilakukan dengan menerapkan berbagai persamaan dan langkah-langkah mengenai jarak algoritma, yaitu dengan Euclidean Distance (Hariani, Sarjan, & Syarli, 2021)

Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster dan data yang memiliki karakteristik berbeda dikelompokkan dengan cluster yang lain sehingga data yang berada dalam satu cluster memiliki tingkat variasi yang kecil (Reisandi, et al, 2021).

Algoritma K-Means merupakan salah satu algoritma Clustering yang masuk ke dalam kelompok Unsupervised learning yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok dengan

sistem partisi (Nabila, et al, 2021). Algoritma K-Means merupakan teknik clustering berbasis jarak yang membagi data ke dalam beberapa cluster dan algoritma ini hanya bekerja pada atribut angka atau numerik.

. Setelah iterasi K-Means berhenti, setiap objek dalam dataset menjadi anggota dari suatu cluster. Nilai cluster ditentukan dengan mencari seluruh objek untuk menemukan cluster dengan jarak terdekat ke objek. Algoritma K Means akan mengelompokkan item data dalam suatu dataset ke suatu cluster berdasarkan jarak terdekat (Baradwaj et al., 2017).

Nilai centroid awal yang dipilih secara acak yang menjadi titik pusat awal, akan dihitung jarak dengan semua data menggunakan rumus Euclidean Distance. Data yang memiliki jarak pendek terhadap centroid akan membuat sebuah cluster. Proses ini berkelanjutan sampai tidak terjadi perubahan pada setiap kelompok (Agrawal & Gupta, 2017).

Versi baru dari Visual Studio inversi terbaru dibuat lebih sederhana untuk mempermudah pengguna dalam mempelajarinya dan memenuhi kebutuhan para Programme (Yesputra, 2017).

Sistem basis data merupakan sekumpulan basis data dengan para pemakai yang menggunakan basis data secara bersama-sama, personil yang merancang dan mengelola basis data, teknik-teknik untuk merancang dan mengelola basis data, serta sistem komputer yang mendukungnya. Komponen utama penyusun sistem basis data adalah perangkat keras, sistem operasi, basis data, sistem pengelola basis data (DBMS), pemakai (Programmer, User mahir, user umum, user khusus) (Pamungkas, 2017)

Basis data merupakan gabungan file data yang dibentuk dengan hubungan/relasi yang logis dan dapat diungkapkan dengan catatan serta bersifat independen. Adapun basis data adalah tempat berkumpulnya data yang saling berhubungan dalam suatu wadah (organisasi/perusahaan) bertujuan agar dapat mempermudah dan

mempercepat untuk pemanggilan atau pemanfaatan kembali data tersebut (Lubis, 2016).

ERD (Entity Relationship Diagram) adalah model teknik pendekatan yang menyatakan atau menggambarkan hubungan suatu model. Didalam hubungan ini tersebut dinyatakan yang utama dari ERD adalah menunjukkan objek data (Entity) dan hubungan (Relationship), yang ada pada Entity berikutnya. (Fridayanthie & Mahdiati, 2016).

Flowchart dapat diartikan sebagai suatu alat atau sarana yang menunjukkan langkah-langkah yang harus dilaksanakan dalam menyelesaikan suatu permasalahan untuk komputasi dengan cara mengekspresikannya ke dalam serangkaian simbol-simbol grafis khusus. Manfaat yang akan diperoleh bila menggunakan flowchart dalam pemecahan masalah komputasi: Terbiasa berfikir secara sistematis dan terstruktur, Mudah mengecek dan menemukan bagian-bagian prosedur yang tidak valid dan bertele-tele. Prosedur akan mudah dikembangkan (Nuraini, 2015).

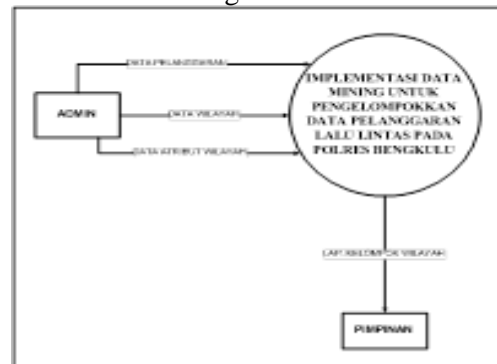
METODE

Dalam penelitian ini, metodologi penelitian yang digunakan adalah metode System Development Life Cycle (SDLC). SDLC ini berfungsi untuk menggambarkan tahapan-tahapan utama dan langkah-langkah dari setiap tahapan yang secara garis besar terbagi dalam 3 (tiga) kegiatan.

Berdasarkan observasi yang dilakukan, bahwa berdasarkan uraian latar belakang yang telah dijelaskan pada bab sebelumnya, maka dapat diketahui bahwa data pelanggaran lalu lintas di wilayah hukum Polres Kota Bengkulu, setiap tahunnya yang hanya data tersebut diproses ke dalam Microsoft Excel sehingga membuat data tidak beraturan. Data yang sudah masuk hanya diproses secara manual sehingga menyulitkan pihak kepolisian lalu lintas untuk mengetahui wilayah mana saja dan jenis

pelanggaran berkendaraan yang sering melakukan pelanggaran. Maka dari itu di perlukan suatu pengelompokan pelanggaran berkendaraan dengan data-data pelanggaran berkendaraan tersebut yang akan menjadi variabel adalah wilayah dan jenis pelanggaran apa saja yang dilakukan.

Dalam perancangan sistem baru ini penulis membuat suatu sistem yang dapat membantu meng-cluster data pelanggaran lalu lintas. Untuk itulah penulis berkeinginan untuk merancang dan membangun sistem data mining untuk data pelanggaran lalu lintas pada Sat Lantas Polres Bengkulu.



Gambar. Diagram Konteks

HASIL DAN PEMBAHASAN

Aplikasi penerapan data mining untuk pengelompokan pelanggaran lalu lintas pada Polres Bengkulu telah dibuat sesuai dengan rancangan diberi nama App.Clustering_Pelanggaran.Exe.

Dimana aplikasi ini telah selesai dibuat dan dapat dipergunakan untuk membantu proses pengelompokan wilayah atau daerah dengan jumlah pelanggaran diprovinsi Bengkulu. Aplikasi ini telah berhasil dibangun sesuai dengan kebutuhan yang diperlukan oleh sistem, sehingga aplikasi ini diharapkan mampu menunjang dan membantu pihak yang terkait dalam proses pengolahan data dan menghasilkan informasi yang cepat, efektif dan akurat.

Pada form ini akan mengelola data wilayah beserta data jenis pelanggaran serta jumlah pelanggaran yang terjadi

pada suatu wilayah yang sudah dimasukkan ke sistem. Seperti yang terlihat pada gambar berikut ini :

Kode	Nama	Nama Pelanggaran	Jenis Pelanggaran	Jumlah
0727151605	Tanah Patah	K-01	Tidak Menggunakan Helm	230
0727151605	Tanah Patah	K-02	Sabuk Pengaman	139
0727151605	Tanah Patah	K-03	Melanggar Lalu L.	130
0727151605	Tanah Patah	K-04	Tidak Membawa...	110
0727151605	Tanah Patah	K-05	Kelebihan Muatan	103
0727151605	Tanah Patah	K-06	Melawan Arah	235
0727151605	Tanah Patah	K-07	Kelengkapan Ka...	852
0658351608	Sawah Lebar	K-01	Tidak Mengguna...	774
0658351608	Sawah Lebar	K-02	Sabuk Pengaman	672
0658351608	Sawah Lebar	K-03	Melanggar Lalu L.	161
0658351608	Sawah Lebar	K-04	Tidak Membawa...	439
0658351608	Sawah Lebar	K-05	Kelebihan Muatan	852
0658351608	Sawah Lebar	K-06	Melawan Arah	115
0658351608	Sawah Lebar	K-07	Kelengkapan Ka...	290
0677151904	Pintu Batu	K-01	Tidak Mengguna...	119
0677151904	Pintu Batu	K-02	Sabuk Pengaman	243

Gambar Form Input Jumlah Pelanggaran

Form proses pengelompokan ini merupakan yang akan digunakan untuk memproses data wilayah dan jumlah jenis pelanggaran menggunakan metode K-Means Clustering dengan terlebih dahulu menentukan titik awal centroid dan jumlah cluster yang diinginkan sehingga menghasilkan kelompok untuk setiap masing-masing wilayah. Adapun tampilan dan Form Proses K-Means Clustering dapat dilihat pada gambar di bawah ini.

Kode	Nama	KodeKelompok
0727151605	Tanah Patah	C-1
0658351608	Sawah Lebar	C-2
0677151904	Pintu Batu	C-2
0672715201	Padang Harapan	C-2
06724151921	Kebun Tebing	C-2
215724151442	Kawasan Wisata	C-3
07726951470	Rawa Makmur	C-3
010724151951	Simpang Nakau	C-2
06726151952	Betungan	C-1
025927154529	Sungai Hitam	C-3

Kode	Jumlah-Member	Variance
C-1	2	514.0537423266
C-2	5	519.9597116677
C-3	3	844.2298290635

Gambar Form Proses Pengelompokan

Untuk menampilkan hasil atau laporan dari proses pengelompokan wilayah berdasarkan jumlah pelanggaran yang telah dilakukan pada form proses K-Means Clustering dilakukan dengan cara meng-klik menu laporan pada menu utama seperti gambar dibawah ini :

No	Nama	Jenis	Jumlah
1	0727151605	Tanah Patah	230
2	0658351608	Sawah Lebar	139
3	0677151904	Pintu Batu	119
4	0672715201	Padang Harapan	774
5	06724151921	Kebun Tebing	103
6	215724151442	Kawasan Wisata	439
7	07726951470	Rawa Makmur	852
8	010724151951	Simpang Nakau	290
9	06726151952	Betungan	115
10	025927154529	Sungai Hitam	243

Gambar Laporan Kelompok Wilayah Pelanggaran

Pengujian yang dilakukan pada aplikasi ini adalah dengan menggunakan teknik black box, seperti yang telah dijelaskan pada Bab III sebelumnya. Tujuan dari pengujian ini adalah untuk menjamin bahwa perangkat lunak yang dibangun memiliki kualitas yang handal, yaitu mampu mempresentasikan kajian pokok dari spesifikasi analisis, perancangan dan pengkodean dari perangkat lunak itu sendiri. Berikut tabel pengujian black box.

Pengujian sistem dilakukan untuk mengetahui bagaimana jalannya kerja sistem dalam melakukan proses pengelompokan wilayah berdasarkan jenis pelanggaran dengan menggunakan metode K-Means Clustering

SIMPULAN

Dari hasil analisis, implementasi, dan pengujian yang telah dilakukan, maka diperoleh kesimpulan sebagai bahwa Hasil implementasi pengelompokan wilayah pengguna alat kontrasepsi dengan menggunakan metode K-Means Clustering, dimana Algoritma K-mean sangat sensitive pada pembangkitan titik pusat awal secara random, hasil pengelompokan bersifat tidak unik (selalu berubah-ubah) dan proses pengerjaannya cepat tetapi keakuratannya tidak dijamin. Teknik clustering K-Means dapat digunakan oleh pihak Polres Bengkulu dalam pengelompokan wilayah berdasarkan pelanggaran menjadi 3 kelompok (C1= pelanggaran terbanyak),

(C2=Pelanggaran Sedang), (C3=Pelanggaran Sedikit) untuk menganalisis untuk melakukan operasi patuh sehingga dapat menciptakan masyarakat patuh dan sadar hukum.

DAFTAR PUSTAKA

- Eska, J. (2016). Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4.5. *Jurnal Teknologi dan Sistem Informasi (Jurteks)* , Vol.2 No.2 ISSN : 2407-1811.
- Fajrin, A. A., & Algifanri, M. (2018). Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma FP-Growth Pada Data Transaksi Penjualan Spare Part Motor. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, Vol.05 No.01. ISSN:2406-7857.
- Fridayanthie, E. W., & Mahdiati, T. (2016). RANCANG BANGUN SISTEM INFORMASI PERMINTAAN ATK BERBASIS INTRANET (STUDI KASUS: KEJAKSAAN NEGERI RANGKASBITUNG). *JURNAL KHATULISTIWA INFORMATIKA*, VOL. IV, NO.2 , 126-138.
- Hariani, Sarjan, M., & Syarli. (2021). SISTEM INFORMASI PENILAIAN KINERJA PELAYANAN POLI GIGI PADA PUSKESMAS MENGGUNAKAN ALGORITMA K-MEANS BERBASIS WEB.

Journal Pegguruang: Conference Series, 188-192.

- Mardalius. (2017). IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN KELAS KELOMPOK BIMBINGAN BELAJAR TAMBAHAN (STUDI KASUS : SISWA SMA NEGERI 1 RANAH PESISIR). *Proceeding SEMILOKA ROYAL 2017 "Teknologi Mobile"*, 105.
- Nabila, Z., Isnain, A., Permata, & Abidin, Z. (2021). ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS. *Jurnal Teknologi dan Sistem Informasi (JTSI)*, 100-108.
- Nuraini, R. (2015). DESAIN ALGORITMA OPERASI PERKALIAN MATRIKS MENGGUNAKAN METODE FLOWCHART. *JURNAL TEKNIK KOMPUTER*, 144 -151.
- Silwattananusarn, T., & Tuamsuk, K. (2012). *Data Mining and Its Applications for Knowledge Management : A Literature Review from 2007 to 2012*. International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.2, No.5,.
- Yesputra, R. (2017). *BELAJAR VISUAL BASIC. NET DENGAN VISUAL STUDIO 2010*. Medan: Royal Asahan Press.