

PEMETAAN DAERAH RAWAN KECELAKAAN MENGGUNAKAN METODE FUZZY C-MEANS

Dwi Puspitasari¹, Yan Watequlis Syaifudin², Resqy Dwi Nofyandi³

^{1,2,3} Program Studi Teknik Informatika, Jurusan Teknologi Informasi, Politeknik Negeri Malang
¹dwi.puspitasari@polinema.ac.id, ²qulis@polinema.ac.id, ³resqynofyandi@gmail.com

Abstrak

Kecelakaan lalu lintas adalah suatu peristiwa di jalan yang tidak diduga dan tidak disengaja melibatkan kendaraan dengan atau tanpa pengguna jalan lain yang mengakibatkan korban manusia dan kerugian harta benda. Angka kecelakaan hingga menyebabkan korban meninggal dunia meningkat tajam di wilayah hukum Polres Pasuruan Kota. Data kecelakaan lalu lintas sangat berhubungan erat dengan data yang bersifat spasial temporal, yaitu memiliki informasi utama berupa lokasi dan waktu kejadian. Namun, pemanfaatan data kecelakaan lalu lintas masih kurang optimal karena hanya disajikan dalam bentuk grafik statistika sehingga potensi informasi dari data tersebut tidak tersampaikan secara maksimal.

Dengan memanfaatkan teknik data mining, data mengenai kecelakaan lalu lintas dapat digali untuk mengelompokkan ruas-ruas jalan berdasarkan faktor kesamaan karakteristik yang melekat pada data. Salah satu metode klusterisasi yang digunakan untuk mencapai tujuan ini adalah algoritma Fuzzy C-Means. Hasil pengelompokan data kecelakaan tadi divisualisasikan menggunakan peta yang menggambarkan pemetaan daerah hasil clustering di wilayah hukum Polres Pasuruan Kota.

Kata kunci : *clustering, fuzzy c-means, kecelakaan lalu lintas, data mining*

1. Pendahuluan

Menurut Undang-undang Republik Indonesia Nomor 22 Tahun 2009 Tentang Lalu Lintas dan Angkutan Jalan, Kecelakaan Lalu Lintas adalah suatu peristiwa di Jalan yang tidak diduga dan tidak disengaja melibatkan Kendaraan dengan atau tanpa Pengguna Jalan lain yang mengakibatkan korban manusia dan/atau kerugian harta benda. Data Korps Lalu Lintas Kepolisian Republik Indonesia menyebutkan bahwa setiap tahun ada 28.000-38.000 orang meninggal akibat kecelakaan lalu lintas di Indonesia. Jumlah tersebut membuat Indonesia berada di peringkat pertama negara dengan rasio tertinggi kematian akibat kecelakaan lalu lintas di dunia. Menurut Jusri Pulubuhu (Pendiri Jakarta Defensive Driving Consulting), posisi Indonesia sebagai negara dengan rasio tertinggi kematian akibat kecelakaan lalu lintas sesuai fakta di lapangan. Sebab, saat ini sangat banyak pengguna kendaraan di Indonesia tidak mengerti peraturan lalu lintas. Saking banyaknya jumlah orang yang tidak tahu aturan ini berimbas terhadap kacaunya lalu lintas secara keseluruhan.

Pada wilayah hukum Polres Pasuruan Kota terhimpun data kejadian kecelakaan lalu lintas tahun 2017 sejumlah 428 kasus. Dengan korban meninggal 110 jiwa, korban mengalami luka berat 47 jiwa dan korban mengalami luka ringan 441 jiwa. Sedangkan pada tahun 2016 sejumlah 311 kasus, terdapat kenaikan $\pm 37\%$. Kejadian kecelakaan lalu lintas di wilayah hukum Polres Pasuruan Kota disebabkan

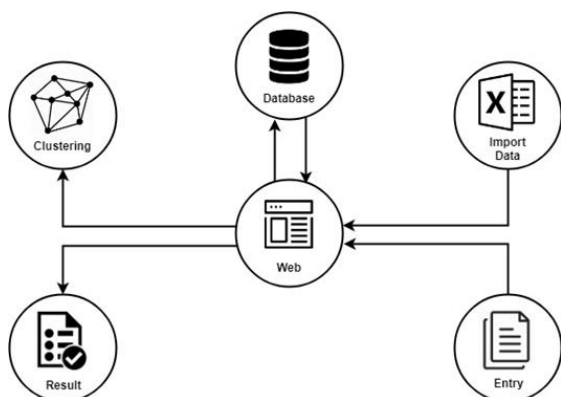
oleh beragam faktor, diantaranya pengendara tidak memperhatikan markah jalan, melaju dengan kecepatan tinggi, saat menyalip kurang memperhatikan kendaraan di depannya serta mengantuk saat mengemudi. Hal tersebut patut menjadi perhatian dan penanganan dari pihak Polres Pasuruan Kota untuk menentukan kebijakan demi mengurangi terjadinya kecelakaan lalu lintas di wilayah hukum Polres Pasuruan Kota. Diperlukan pengelompokan ruas jalan berdasarkan intensitas terjadinya kecelakaan lalu lintas, sebagai informasi tambahan bagi pihak Polres Pasuruan Kota dalam menentukan kebijakan agar kebijakan yang dibuat dapat berdampak maksimal dalam mengurangi kecelakaan lalu lintas pada ruas jalan tersebut.

Dari rekaman data kecelakaan lalu lintas dapat dilakukan *clustering* untuk mengelompokkan ruas jalan berdasarkan intensitas terjadinya kecelakaan lalu lintas. *Clustering* ditujukan untuk mengelompokkan ruas jalan di wilayah hukum Polres Pasuruan Kota berdasarkan faktor kesamaan karakteristik yang ada di dataset yaitu jumlah kecelakaan, jumlah kendaraan yang terlibat, serta jumlah korban (meninggal dunia, luka berat, dan luka ringan) akibat kecelakaan yang terjadi dalam suatu rentan waktu tertentu. Dengan menggunakan *clustering*, dapat diperoleh tingkat kerawanan suatu wilayah terhadap kecelakaan lalu lintas melalui *clustering* data kecelakaan di wilayah hukum Polres Pasuruan Kota. Salah satu metode pada teknik *clustering* adalah *Fuzzy C-Means*, metode *Fuzzy C-Means* memberikan hasil pengelompokan yang halus

atau tidak banyak menggeser pusat kluster. Hal ini disebabkan karena setiap data dilengkapi dengan fungsi keanggotaan (*membership function*) untuk menjadi anggota kluster yang ditemukan. Untuk memudahkan penyampaian informasi tentang daerah rawan kecelakaan dari proses *data mining*, informasi tersebut divisualisasikan dalam bentuk peta.

2. Analisis dan Perancangan

2.1 Analisis Sistem



Gambar 17. Gambaran Umum Sistem

Sistem akan memetakan daerah rawan kecelakaan berdasarkan hasil klusterisasi dari data rekaman kecelakaan lalu lintas. Masukan sistem berupa data kecelakaan lalu lintas pada seluruh ruas jalan yang tercatat pihak Polres Pasuruan Kota. Masukan tersebut akan diolah menggunakan teknik klusterisasi dengan metode Fuzzy C-Means dan akan menghasilkan keluaran berupa peta dengan keterangan berupa garis pada daerah (ruas jalan) dengan warna yang berbeda-beda sesuai dengan kluster yang terbentuk.

2.2 Analisis Kebutuhan

Analisis kebutuhan merupakan tahap untuk menganalisis segala sesuatu yang dibutuhkan untuk mengembangkan sistem, dalam hal ini terdapat beberapa hal yang harus dipenuhi antara lain meliputi dataset.

Data yang digunakan adalah data rekaman kecelakaan lalu lintas dari Polres Pasuruan Kota tahun 2017. Adapun atribut dari data rekaman kecelakaan lalu lintas yang akan dilibatkan dalam proses klusterisasi adalah sebagai berikut :

Tabel 2. Data Proses Klusterisasi

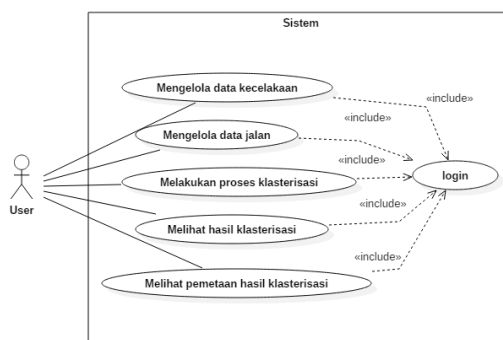
No	Data	Keterangan
1.	Jumlah kecelakaan	Jumlah kecelakaan yang terjadi pada ruas jalan tertentu
2.	Jumlah kendaraan yang terlibat	Jumlah kendaraan yang terlibat kecelakaan lalu lintas
3.	Jumlah korban meninggal dunia	Jumlah korban kecelakaan lalu lintas yang mengakibatkan meninggal dunia
4.	Jumlah korban luka berat	Jumlah korban kecelakaan lalu lintas yang mengakibatkan luka berat
5.	Jumlah korban luka ringan	Jumlah korban kecelakaan lalu lintas yang mengakibatkan luka ringan

2.3 Perancangan Sistem

Pada tahap ini yaitu berupa perancangan sistem yang digunakan untuk memberikan gambaran umum tentang sistem yang akan dibangun.

Use Case Diagram

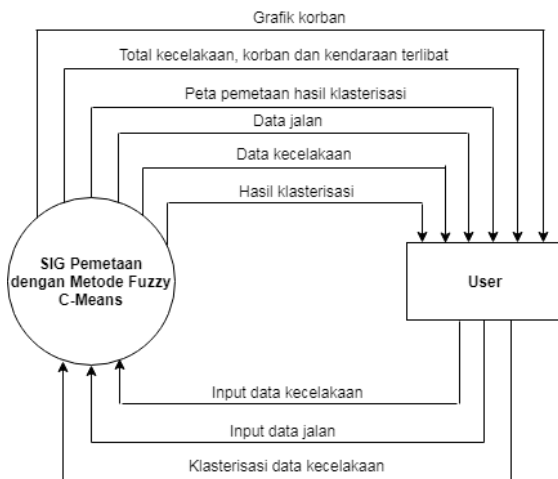
Pengguna dalam sistem ini hanya ada satu, yang diinisialisasikan dengan “user”. Interaksi yang dapat dilakukan pengguna dengan sistem ini adalah mengelola data kecelakaan, melakukan proses klusterisasi, melihat hasil klusterisasi dan melihat pemetaan hasil klusterisasi.



Gambar 18. Use Case Diagram

2.4 Context Diagram

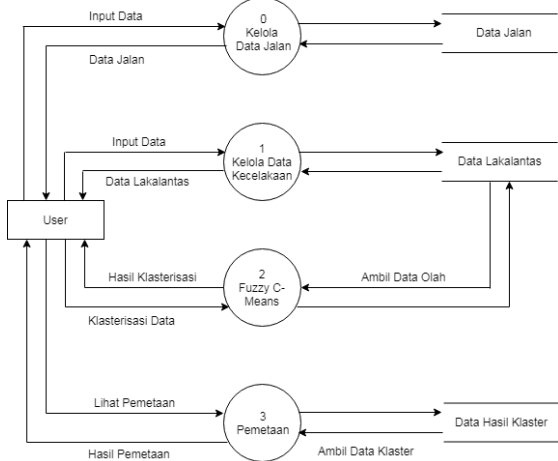
Context diagram adalah diagram yang terdiri dari suatu proses dan menggambarkan ruang lingkup suatu sistem. Context diagram dari sistem yang akan dibuat dapat dilihat pada gambar dibawah.



Gambar 19. Context Diagram

2.5 Data Flow Diagram

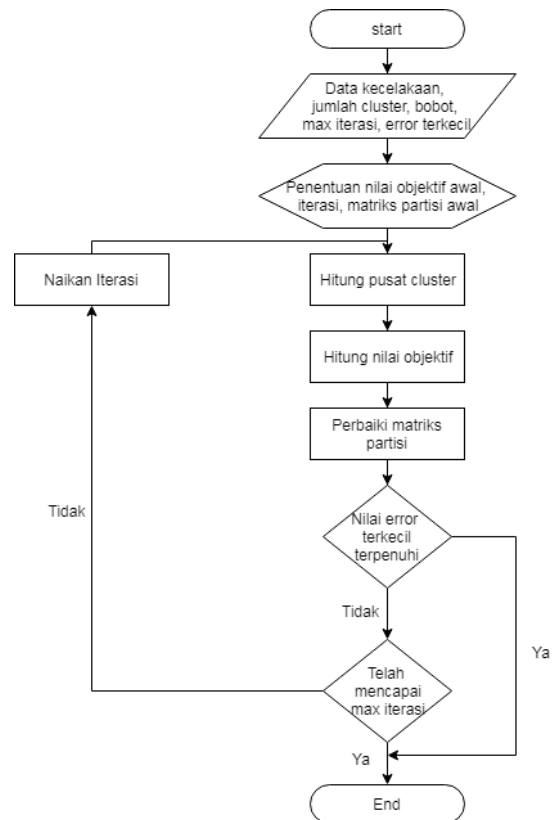
Data flow diagram merupakan model dari sistem untuk menggambarkan pembagian sistem yang lebih kecil. Maka dapat di gambarkan Data Flow Diagram (DFD) pada gambar dibawah.



Gambar 20. Data Flow Diagram

2.6 Metode Fuzzy C-Means

Pada Gambar 6, masukan untuk melakukan proses klusterisasi menggunakan metode fuzzy c-means antara lain: data yang akan diolah, jumlah kluster, nilai pangkat atau bobot, nilai maksimum iterasi dan error terkecil. Setelah itu hitung pusat kluster, nilai objektif dan hitung ulang matriks nilai random. Perhitungan berhenti apabila memenuhi salah satu kondisi, telah mencapai nilai error terkecil yang dimasukkan atau telah mencapai maksimum iterasi yang ditentukan.

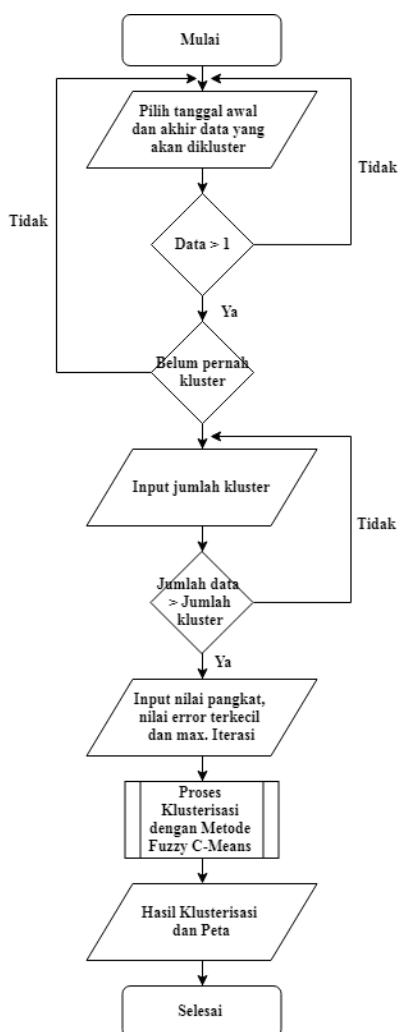


Gambar 21. Diagram Alir Fuzzy C-Means

2.7 Alur Proses Klusterisasi Sistem

Gambar 7 menjelaskan tentang proses klusterisasi pada sistem yang akan dibangun, user diminta memilih data yang akan digunakan dalam proses klusterisasi. Kemudian sistem akan memeriksa jumlah data yang dipilih, apabila jumlah data tidak memenuhi kondisi untuk dilakukan proses klusterisasi, user diminta memilih data yang lain. Jika data yang dipilih memenuhi kondisi untuk dilakukan proses klusterisasi, sistem akan memeriksa data yang dipilih apabila data pernah diklusterisasi sebelumnya maka user akan diminta untuk memilih data lain.

Apabila data yang dipilih telah memenuhi dua kondisi sebelumnya, sistem akan memeriksa jumlah kluster yang dimasukkan oleh user. Jumlah kluster yang dimasukkan harus tidak lebih besar dari jumlah data yang dipilih, jika jumlah kluster yang dimasukkan lebih besar dari jumlah data yang dipilih maka user diminta memasukkan jumlah kluster lain. Setelah kondisi tersebut telah dipenuhi, data dapat diklusterisasi menggunakan metode fuzzy c-means. Keluaran dari sistem berupa hasil klusterisasi dan ditampilkan dalam bentuk peta.



Gambar 22. Diagram Alir Sistem

2.8 Modified Partition Coefficient

Untuk pengujian hasil kluster, digunakan metode modified partition coefficient (MPC). MPC merupakan metode yang digunakan untuk menguji validitas jumlah kluster. MPC sendiri merupakan pengembangan dari metode *partition coefficient* (PC). *Partition coefficient* (PC) merupakan metode yang mengukur jumlah kluster yang mengalami *overlap*. Nilai PC berada dalam batas $\frac{1}{c} \leq PC(c) \leq 1$. Pada umumnya jumlah kluster yang paling optimal ditentukan dari nilai PC yang paling besar ($\max_{2 \leq c \leq n-1} PC(c)$). Berikut adalah algoritma metode PC: [9]

$$PC(c) = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^N (\mu_{ij}^2) \tag{1}$$

Dimana:

- c = jumlah kluster
- N = jumlah data
- μ_{ij} = derajat keanggotaan data ke- j pada kluster ke- i
- $PC(c)$ = nilai indeks PC pada kluster ke- c

Partition coefficient cenderung mengalami perubahan yang monoton terhadap beragam nilai c (jumlah kluster). Modifikasi dari indeks PC (*Modified*

Partition Coefficient / MPC) dapat mengurangi perubahan yang monoton tersebut. Nilai MPC berada dalam batas $0 \leq PC(c) \leq 1$. Pada umumnya jumlah kluster yang paling optimal ditentukan dari nilai MPC yang paling besar ($\max_{2 \leq c \leq n-1} PC(c)$). Berikut adalah algoritma metode MPC: [9]

$$MPC(c) = 1 - \frac{c}{c-1} (1 - PC(c)) \tag{2}$$

Dimana:

- c = jumlah kluster
- $MPC(c)$ = nilai indeks MPC pada kluster ke- c

3. Implementasi Dan Pengujian

3.1 Pengujian Metode

Pengujian dilakukan dengan cara melakukan proses klusterisasi pada sistem, perbandingan hasil uji sistem dengan hasil kluster melalui perhitungan manual dengan excel, pengujian validitas kluster menggunakan Modified Partition Coefficient. Data yang digunakan secara manual dan melalui sistem yakni data yang diambil dari pihak Polres Pasuruan Kota.

3.2 Pengujian Metode Fuzzy C-Means pada Sistem

Untuk mengetahui hasil dari metode yang telah diterapkan pada sistem maka dilakukan proses pengujian dengan keseluruhan data yaitu sejumlah 36 data kecelakaan pada jalan di kota Pasuruan. Nilai parameter yang digunakan antara lain : jumlah kluster = 3, pangkat = 2, epsilon = 0.000001 dan maksimum iterasi = 100. Hasil pengujian ditunjukkan pada gambar dibawah.

Kluster	Nama Jalan	Jumlah Kecelakaan	Kendaraan Tertibat	Korban MD	Korban LB	Korban LR
1	Jalan Dr. Pandjaitan	1	1	0	0	1
1	Jalan Letjen. R. Soeprapto	1	2	2	0	1
1	Jalan Diponegoro	1	2	1	0	0
1	Jalan Erlangga	1	2	0	0	1
1	Jalan Gajah Mada	4	8	0	0	5
1	Jalan Hasanuddin	1	2	0	0	1
1	Jalan Hayam Wuruk	2	4	0	1	1
1	Jalan Irian Jaya	1	2	0	0	1
1	Jalan KH. Kartini	1	2	0	0	2
1	Jalan KH. Abdul Hamid	2	4	0	0	2
1	Jalan KH. Agus Salim	2	4	0	0	4
1	Jalan KH. Wachid Hasyim	2	3	0	0	3
1	Jalan Lks. RE Martadinata	1	2	0	1	3
1	Jalan Pahlawan	4	7	2	0	6
1	Jalan Pattimura	3	5	1	0	3
1	Jalan Patimurus	2	3	2	0	1
1	Jalan S. Parman	1	1	0	0	1
1	Jalan KH. Samanhudi	1	2	0	1	0
1	Jalan Brigjend. Slamet Riyadi	2	3	0	1	1
1	Jalan Sumatera	1	2	0	1	0
1	Jalan Sunan Ampel	2	3	0	0	3
1	Jalan Trunojoyo	2	3	1	0	2
1	Jalan Untung Surapati	2	4	0	0	6
1	Jalan Urip Sumoharjo	2	4	0	0	2
1	Jalan Veteran	3	5	0	0	5
1	Jalan Balakota	6	12	3	0	6
1	Jalan Dr. Wahidin Sudirohusodo	9	15	2	0	11

Jalan Jend. Gatot Subroto	7	11	2	0	9
Jalan KH. Hasyim Aoy'ari	6	9	0	0	7
Jalan HOS. Cokroaminoto	8	15	1	2	12
Jalan KH. Ahmad Dahlan	7	14	1	0	11
Jalan KH. Mansyur	6	10	2	0	5
Jalan Panglima Sudirman	13	28	5	3	20
Jalan Soekarno Hatta	8	13	2	2	6
Jalan Ahmad Yani	25	43	11	6	28
Jalan Ir. H. Juanda	32	58	5	2	48

Gambar 23. Pengujian

Dari pengujian metode fuzzy c-means pada sistem diatas, menunjukkan bahwa sistem dapat melakukan klusterisasi data secara baik sesuai nilai parameter yang dimasukkan setelah iterasi ke-32 dengan nilai selisih fungsi objektif iterasi sebelumnya sebesar $9.8525686098583E-7$ dan waktu klusterisasi selama 0.13621211051941 detik.

3.3 Pengujian Terhadap Nilai Error Minimum (ξ)

Pengujian dilakukan untuk mendapatkan nilai error minimum (ξ). Pengujian dilakukan pada saat fungsi objektif telah mencapai kondisi konvergen. Nilai error diperoleh dengan menghitung selisih dari fungsi objektif yang didapatkan pada setiap iterasi. Fungsi objektif dikatakan telah konvergen apabila nilai yang dihasilkan sudah tidak mengalami perubahan lagi, sehingga kesalahan (error) yang dihasilkan bernilai 0. Pada pengujian ini, parameter jumlah cluster yang digunakan adalah sebanyak 3 dengan maksimum iterasi 100. Hasil pengujian nilai error terhadap fungsi objektif ditunjukkan pada Tabel 2.

Tabel 3. Nilai Error Minimum

Iterasi ke-	Fungsi Objektif	Nilai Error
1	4312.292029	4312.292029
2	2721.954099	1590.337931
3	2381.514584	340.4395147
4	1851.01251	530.5020739
5	1291.548428	559.4640825
6	933.6251226	357.9233051
7	806.9507085	126.6744141
8	783.0465058	23.90420265
9	777.6669349	5.379570913
10	776.3214606	1.34547432
11	775.9651117	0.356348874
12	775.8670588	0.098052922
13	775.8393794	0.02767939
14	775.8314349	0.007944526
15	775.8291306	0.002304257
16	775.8284579	0.000672677
17	775.8282608	0.00019715
18	775.8282029	0.0000579
19	775.8281858	0.0000171
20	775.8281808	0.000005
21	775.8281793	0.0000015
22	775.8281789	0.0000004
23	775.8281788	0.0000001
24	775.8281787	0.0000001
25	775.8281787	0

Berdasarkan Tabel 2, dapat diketahui bahwa nilai fungsi objektif menjadi konvergen ketika mencapai iterasi ke-25. Nilai error minimum yang didapat ketika nilai fungsi objektif menjadi konvergen adalah sebesar $0.0000001 (1 \times 10^{-7})$. Dengan demikian, nilai error minimum sebesar $0.0000001 (1 \times 10^{-7})$ merupakan nilai error minimum yang optimal untuk diterapkan.

3.4 Pengujian Terhadap Jumlah Klaster (c)

Pengujian terhadap jumlah klaster (c) dilakukan untuk mengetahui pengaruh banyaknya jumlah cluster terhadap nilai validitas MPC. Tujuan pengujian ini yaitu untuk mendapatkan nilai c terbaik yang memiliki nilai MPC paling baik untuk digunakan pada pengujian selanjutnya. Pengujian terhadap jumlah cluster dilakukan dengan membandingkan hasil nilai MPC dari banyaknya cluster yang berbeda. Pada pengujian ini, nilai c yang digunakan yaitu c=2 sampai dengan c=9 dengan nilai pembobot w=2. Hasil uji validitas terhadap variasi jumlah klaster ditunjukkan pada Tabel 3.

Tabel 4. Hasil Pengujian Klaster

Nilai Klaster (c)	Modified Partition Coefficient
2	0.91322627968556
3	0.81599660200797
4	0.83997965232639
5	0.79023728349124
6	0.64100357868794
7	0.61909890521743
8	0.6486521106883
9	0.59252048016113

Pada Tabel 3 juga terlihat bahwa nilai validitas MPC tertinggi diperoleh dengan nilai c=2. Dengan demikian, nilai c=2 merupakan jumlah cluster yang optimal untuk diterapkan.

3.5 Pengujian Terhadap Nilai Pangkat (w)

Pengujian juga dilakukan terhadap nilai pangkat (w) yang merupakan bobot pangkat pada perhitungan nilai keanggotaan metode fuzzy c-means. Nilai dari pembobot (w) bernilai $w > 1$. Tujuan dari pengujian ini untuk mendapatkan nilai w terbaik yang memiliki nilai MPC paling tinggi. Pada pengujian nilai w yang akan diuji adalah w=2 sampai dengan w=9 dengan jumlah klaster 2.

Tabel 5. Hasil Pengujian Pangkat

Nilai Pangkat (w)	Modified Partition Coefficient
2	0.91322627905116
3	0.33388459323707
4	3.3685498834757E-11
5	2.2204460492503E-16
6	4.4408920985006E-16
7	0
8	0
9	0

Pada Tabel 4 menunjukkan bahwa nilai MPC tertinggi didapatkan pada saat w bernilai 2, yaitu sebesar 0.91322627905116.

4. Kesimpulan

Berdasarkan hasil pengujian dan analisis penelitian terhadap pemetaan daerah rawan kecelakaan dengan menggunakan metode fuzzy c-means, maka dapat diambil beberapa kesimpulan yaitu:

1. Metode fuzzy c-means dapat digunakan untuk melakukan pengelompokan jalan di kota Pasuruan berdasarkan data rekaman kecelakaan lalu lintas pihak Polres Pasuruan Kota.
2. Klaster data dengan nilai pusat klaster tinggi, merupakan klaster dengan atribut data data kecelakaan yang tinggi begitu juga sebaliknya.
3. Berdasarkan pengujian yang telah dilakukan, nilai error minimum yang didapatkan pada saat fungsi objektif telah konvergen adalah sebesar 0.0000001 (1 x 10⁻⁷). Banyaknya klaster yang optimal untuk digunakan adalah sebanyak 2 klaster dengan nilai validitas MPC sebesar 0.91322627968556. Sedangkan nilai pangkat yang optimal untuk digunakan adalah yaitu 2 dengan nilai validitas MPC sebesar 0.91322627968556.

Daftar Pustaka :

Alvian, Wijaya Kusuma. "Implementasi Data Mining dengan Algoritma Fuzzy C-Means Studi Kasus Penjualan di UD Subur Baru." Skripsi, Fakultas Ilmu Komputer (2014).
 Anggraini, Fenny, and Sugeng Mingparwoto. "PENERAPAN METODE ALGORITMA BELLMAN-FORD DALAM APLIKASI

PENCARIAN LOKASI PERSEROAN TERBATAS DI PT. JAKARTA INDUSTRIAL ESTATE PULOGADUNG (PT. JIEP)." *Jurnal Teknologi* 7.1 (2015): 28-34.
 Asril, Elvira, Fana Wiza, and Yogi Yunefri. "Analisis Data Lulusan dengan Data Mining untuk Mendukung Strategi Promosi Universitas Lancang Kuning." *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi* 6.2 (2015): 24-32.
 Astuti, Femi Dwi. "Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Menggunakan Algoritma Hard C-Means." *Data Manajemen dan Teknologi Informasi (DASI)* 18.1 (2017): 64-69.
 Faisal, Radfan. (2017). "Angka Laka Lantas di Kota Pasuruan Meningkat, Ini Jumlah Korbannya". Didapat dari: <https://www.jawapos.com/radarbromo/read/2017/12/31/36985/angka-laka-lantas-di-kota-pasuruan-meningkat-ini-jumlah-korbannya>.
 Niam, A., Suprayogi, A., Awaluddin, M., & Putra Wijaya, A. (2014). *Aplikasi Openstreetmap Untuk Sistem Informasi Geografis Kantor Pelayanan Umum (Studi Kasus: Kota Salatiga)* (Doctoral dissertation, Universitas Diponegoro).
 Putri, Cahaya Eka. "Analisis Karakteristik Kecelakaan dan Faktor Penyebab Kecelakaan pada Lokasi Blackspot di Kota Kayu Agung." *Jurnal Teknik Sipil dan Lingkungan* 2.1 (2015): pp-154.
 Ramadhan, Aditya, Zuliar Efendi, and Mustakim Mustakim. "Perbandingan K-Means dan Fuzzy C-Means untuk Pengelompokan Data User Knowledge Modeling." *Seminar Nasional Teknologi Informasi Komunikasi dan Industri*. 2017.
 Republik Indonesia. 2002. *Undang-Undang Republik Indonesia Nomor 22 Tahun 2009 Tentang Lalu Lintas Dan Angkutan Jalan*.
 Rudi, Alsadad. (2017). "Kematian akibat Kecelakaan di Indonesia Tertinggi di Dunia". Didapat dari: <http://otomotif.kompas.com/read/2017/12/04/100400715/kematian-akibat-kecelakaan-di-indonesia-tertinggi-di-dunia>
 Tanjung, Priscilia Angeline. "Penerapan Fuzzy C-Means Clustering Pada Data Nasabah Bank". Skripsi, Fakultas Sains dan Teknologi (2016).