



ALGORITMA FUZZY C-MEANS DENGAN METODE ELBOW UNTUK MENGELOMPOKKAN KEPOLISIAN PROVINSI DI INDONESIA BERDASARKAN JUMLAH KASUS (*CRIME TOTAL*, *CRIME CLEARED*, *CRIME RATE*, *CLEARANCE RATE*)

Anggar Wahyu Hadiyatullah¹, Hardian Oktavianto², Habibatul Azizah Al Faruq³
Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember
Email : anggarwahyu87@gmail.com

ABSTRAK

Crime Total (jumlah kejahatan yang dilaporkan) adalah peristiwa yang dilaporkan kepada kepolisian dari laporan masyarakat atau peristiwa yang pelakunya di tangkap langsung oleh polisi, *Crime Clearance Rate* (angka penyelesaian tindak kejahatan) adalah presentase penyelesaian tindak kejahatan oleh polisi *Crime Rate* (angka kejahatan) adalah angka kejahatan atau resiko terkena kejahatan setahun dibagi dengan jumlah penduduk. Menurut data Badan Pusat Statistika Republik Indonesia, jumlah kasus *Crime Total*, *Crime Cleared*, *Crime Rate Clearance Rate*, di 34 Provinsi di Indonesia periode 2016 - 2018. mengalami peningkatan yang tidak signifikan dan relatif stabil. Berdasarkan penelitian sebelumnya, maka penelitian ini dilakukan untuk mengelompokkan provinsi di Indonesia berdasarkan jumlah kasus *Crime Total*, *Crime Cleared*, *Crime Rate Clearance Rate* dengan menggunakan metode *clustering* yaitu algoritma *Fuzzy C-Means*. Untuk pengukuran *cluster* optimum dalam menentukan *cluster* terbaik, metode yang digunakan adalah metode *Elbow*. Data yang digunakan yaitu jumlah kasus *Crime Total*, *Crime Cleared*, *Crime Rate Clearance Rate* di 34 provinsi di Indonesia periode tahun 2016 - 2018. Dari serangkaian pengujian mulai dari 2 cluster sampai 10 *cluster*, dihasilkan *cluster* optimum berada pada 2 *cluster* berdasarkan jarak SSE (*Sum of Squares Error*) pada metode *Elbow*. Pada *cluster* 1 terdiri dari 4 anggota provinsi dan *cluster* 2 terdiri dari 30 anggota provinsi. Berdasarkan hasil karakteristik data tahun 2018, *cluster* 1 memiliki jumlah kasus *Crime Total*, *Crime Cleared*, *Crime Rate Clearance Rate* lebih rendah dibandingkan dengan *cluster* 2.

Kata kunci: *Crime Total*, *Crime Cleared*, *Crime Rate Clearance Rate*, *clustering*, *fuzzy c-means*, *elbow*.

ABSTRACT

Crime Total (Total Crime Reported) is an incident reported to the police from public reports or an incident where the perpetrator was arrested directly by the police, the *Crime Clearance Rate* is the percentage of crime by the police. The *Crime Rate* is crime affected by crime is above the population. According to data from the Central Statistics Agency of the Republic of Indonesia, the number of total criminal cases, settlement crimes, crime exemption rates, in 34 provinces in Indonesia for the 2016-2018 period has increased insignificantly and is relatively stable. Based on previous research, this study was conducted to group provinces in Indonesia based on the number cases of *Crime Total*, *Crime Cleared Rates*, and *Crime Rate Clearance Rates* using the clustering method, namely the *Fuzzy C-Means* algorithm. The optimal cluster measurement in determining the best cluster, the method used is the *Elbow* method. The data used are the number of *Crime Total*, *Criminal Actions of Settlement*, the *Clearance Level of Crime* in 34 provinces in Indonesia for the 2018 period. from a series of tests ranging from 2 clusters to 10 cluster, the resulting optimum cluster is in 2 clusters based on the distance SSE (*Sum of Squares Error*) on the *Elbow* method. *Cluster 1* consists of 4 provincial members and *cluster 2* consists of 30 provincial members. Based on the results of data comparisons in 2018, *cluster 1* has a lower number of *Crime Total*, *Crime Cleared*, *Crime Rate Clearance Rate* cases than *cluster 2*.

Keywords: *Crime Total*, *Crime Cleared*, *Crime Rate Clearance Rate*, *clustering*, *fuzzy c-means*, *elbow*.

1. PENDAHULUAN

Menurut Maslow (1943) bahwa setiap manusia mempunyai 5 kebutuhan dasar ialah fisiologis, keamanan, cinta, harga diri serta aktualisasi diri. Fungsi dan wewenang negara adalah memberikan rasa aman kepada masyarakatnya karena pada saat ini bisa dikatakan bahwa rasa aman merupakan salah satu kebutuhan utama yang harus didapatkan oleh setiap manusia dalam menjalankan seluruh aktivitasnya. Kewajiban pemerintah dan negara Indonesia sudah tercantum di dalam Pembukaan UUD 1945 yang berbunyi: “..... Pemerintah dan Negara Indonesia yang melindungi segenap bangsa Indonesia dan seluruh tumpah darah Indonesia” karena menurut Maslow (1943) dalam teori hierarki kebutuhan manusia bahwa rasa aman berada pada tingkatan ke dua di bawah kebutuhan dasar seperti sandang, pangan dan papan.

Kasus kejahatan sebagai suatu problem nyaris di segala dunia. Menurut data Badan Pusat Statistika (BPS), jumlah kejahatan di Indonesia pada tahun 2016 sebanyak 357.197 kasus pada tahun 2017 sebanyak 336.652 kasus, pada tahun 2018 terdapat 294.281 kasus. Dari banyaknya daerah yang tingkat kejahatannya bervariasi, maka dilakukanlah pengolahan data-data dengan metode data mining, penelitian sebelumnya yang dilakukan oleh Rinawati (2020) dengan studi kasus “*K-Means* algoritma data mining pada Kasus Daerah Pelaku Kejahatan Pencurian Berdasarkan Provinsi”. Dataset yang digunakan dari tahun 2016 sampai 2018. Pada riset tersebut, tidak menggunakan pengukuran *cluster* optimum untuk memastikan suatu *cluster* terbaik. Jadi penulis ingin melakukan penelitian data dari jumlah kasus *Crime Total*, *Crime Cleared*, *Crime Rate*, *Clearance Rate* menurut kepolisian Provinsi di Indonesia pada tahun 2016 di 34 Provinsi di Indonesia. Algoritma yang akan digunakan oleh penulis merupakan Algoritma *Fuzzy C-Means* serta memakai metode *Elbow* untuk memastikan *cluster* optimum dalam penentuan *cluster* terbaik. Maka dari itu, riset ini dibuat dengan judul “Algoritma *Fuzzy C-Means* dengan metode *Elbow* untuk mengelompokkan Kepolisian Provinsi di Indonesia berdasarkan jumlah *Crime Total*, *Crime Cleared*, *Crime Rate*, *Clearance Rate* di Kepolisian Provinsi di Indonesia”.

2. PENELITIAN TERKAIT

A. Kepolisian Daerah di Indonesia

Menurut kepala kepolisian NKRI, satuan kewilayahan kepolisian daerah (polda) tersebar di 32 daerah dan kepolisian resort (polres) di 400an satuan. Masing-masing mempunyai kondisi geososial yang luas dan berbeda. Polda bertugas menyelenggarakan tugas polri pada wilayah tingkat I, yaitu provinsi. Polda dipimpin oleh kepala kepolisian negara republik indonesia daerah (kapolda) yang bertanggung jawab kepada kapolri

B. Kejahatan

Kejahatan menurut kamus besar bahasa Indonesia merupakan sikap yang berlawanan dengan nilai-nilai serta norma-norma yang berlaku serta sudah disahkan oleh hukum tertulis (hukum pidana) Menurut Quinney (1970) kejahatan merupakan sesuatu rumusan tentang sikap manusia yang diciptakan oleh yang berwenang dalam sesuatu penduduk yang secara politisi terorganisir. Kriminologi di Indonesia memandang kejahatan selaku sikap yang sudah diutus oleh majelis hukum.

C. Data Mining

Data mining adalah sebuah proses untuk mencari atau menggali nilai dari suatu data berupa objek yang belum diketahui secara manual menurut Pramudiono (2006).

D. Clustering

Clustering merupakan proses dalam datamining yang digunakan dalam mencerna suatu informasi yang tidak dikenal objeknya. *Clustering* diketahui dengan sebutan *cluster analysis*. Teknik bekerja *clustering* merupakan yaitu mengelompokkan informasi ke dalam sebagian kelas ataupun yang biasa disebut dengan *cluster*. Objek-objek yang terdapat pada kelas tersebut mempunyai tingkatan kemiripan yang besar bila dibandingkan dengan objek yang lainya yang terdapat di kelas tersebut, namun mempunyai tingkatan kemiripan yang rendah bila dibanding dengan objek yang terdapat di kelas/*cluster* lain (Han serta Kamber, 2006)

E. Fuzzy C-Means

Menurut Kusumadewi dan Purnomo (2004), langkah perhitungan algoritma *Fuzzy C-Means* yakni sebagai berikut:

1. Input data yang akan di cluster X , berbentuk matriks berdimensi $n \times m$ (n = jumlah ilustrasi, data, m = atribut tiap data). X_{ij} = data ilustrasi ke- i ($i=1,2,\dots,n$),atribut ke- j ($j = 1,2,\dots,m$).
2. Memastikan :
 - Jumlah cluster = c ;
 - Pangkat bobot = w ;
 - Maksimum iterasi = MaxIter
 - Error terkecil yang diinginkan = ξ
 - Fungsi objektif awal = $P_0 = 0$
 - Iterasi awal = $t = 1$
3. Bilangan random μ_{ik} , $i = 1, 2, 3, \dots,n$; $k = 1, 2, 3, \dots,n$; sebagai elemen matriks partisi awal U (derajat keanggotaan dalam cluster).
 Perhitungan tiap kolom atribut:

$$Q_j = \sum_{k=1}^c \mu_{ik}$$

Q_j adalah jumlah nilai dengan derajat keanggotaan per kolom = 1
 Dengan $j = 1,2,\dots,m$.

4. Menghitung pusat cluster ke- k : V_{kj} dengan $k=1, 2, 3, \dots,c$; dan $j=1, 2, 3, \dots,m$

$$V_{kj} = \frac{\sum_{i=1}^n (\mu_{ik} \times X_{ij})}{\sum_{i=1}^n (\mu_{ik})} \quad w$$

V_{kj} = pusat cluster
 i = iterasi

μ_{ik} = perubahan matrix partisi
 X_{ij} = atribut

5. Menghitung fungsi objektif pada iterasi ke-t, P_t ;

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right) (\mu_{ik})^w \quad (3)$$

P_t = Fungsi objektif
 $\sum_{i=1}^n$ = jumlah data cluster
 $\sum_{i=1}^n$ = jumlah awal perhitungan cluster

6. Menghitung pergantian matrix partisi:

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}$$

μ_{ik}

Dengan $i=1,2,\dots,n$; dan $k=1,2,\dots,c$.

Cek keadaan menyudahi :

Bila $(|P_t - P_{t-1}| < \xi)$ atau $(t > \text{MaxIter})$ hingga menyudahi.

Bila tidak : $t = t + 1$, mengulang ke langkah empat.

F. Metode Elbow

Metode elbow merupakan sesuatu tata cara serta analisis yang dipakai untuk pemilihan jumlah cluster ataupun kelompok terbaik untuk memperoleh suatu data dengan memandang hasil presentase serta perbandingan antara jumlah cluster yang hendak membentuk sesuatu siku pada sesuatu titik, Berikut ini merupakan rumus dari SSE:

$$SSE = \sum_{k=1}^k \sum_{x_i \in S_k} \|X_i - C_k\|^2$$

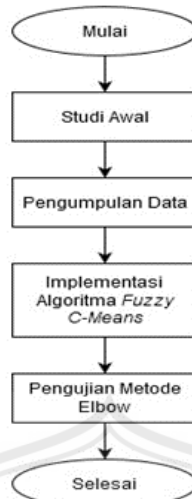
X_i = nilai atribut dari data ke-i

C_k = nilai atribut titik pusat cluster ke-i

G. Rstudio

RStudio adalah aplikasi yang di bangun oleh JJAllarie yang digunakan dalam bidang analisis data pengolahan informasi R. R ditulis oleh Ross Ihaka dan Robert Gentleman di Universitas Auckland, Selandia Baru. R adalah suatu bahasa komputer dan merupakan lingkungan pemograman interaktif untuk analisis data dan grafik.

H. Diagram Alur Penelitian



Gambar 1. Diagram Alur Penelitian

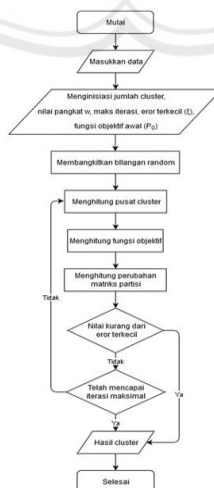
I. Studi Awal

Langkah awal dari penelitian ini adalah mempelajari serta mencari kasus yang akan diteliti serta diuji, kemudian memastikan ruang lingkup permasalahan, latar belakang, serta memahami sebagian literatur antara lain merupakan buletin, paper, dan lain-lain yang berkaitan dengan kasus serta mencari pemecahan dari permasalahan tersebut

J. Pengumpulan Data

Tipe data yang digunakan dalam penelitian ini adalah data sekunder, yang diperoleh dari sumber yang sudah ada yaitu dari Badan Pusat Statistika (BPS). Pengumpulan data diambil dari Statistik Kriminal 2019 yang diakses pada link <https://www.bps.go.id/publication/2019/12/12/66c0114edb7517a33063871f/s-tatistik-kriminal-2019.html>

K. Implementasi Algoritma Fuzzy C-Means



Gambar 2. Flowchart Fuzzy C-Means

3. HASIL DAN PEMBAHASAN

Hasil yang diperoleh dari perhitungan yang sudah dilaksanakan, data yang sudah didapat hendak di *cluster* memakai algoritma *Fuzzy C- Means* setelah itu diolah untuk menciptakan hasil *cluster* optimum ataupun *cluster* terbaik dengan memakai tata cara *Elbow*. Data yang digunakan merupakan data jumlah permasalahan *Crime Total*, *Crime Cleared*, *Crime Rate Clearance Rate*, di 34 Provinsi di Indonesia periode 2018.

A. *Fuzzy C-Means* pada RStudio

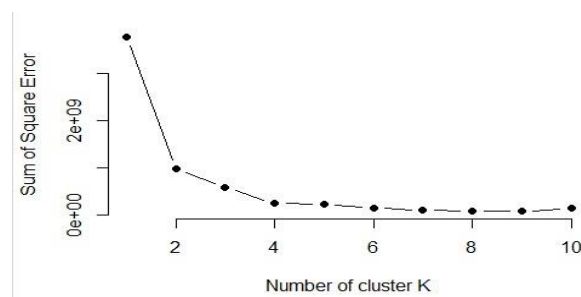
Pengolahan data menggunakan RStudio yang di *cluster* menggunakan algoritma *Fuzzy C- Means* dari 2 *cluster* hingga 10 *cluster*. Dari hasil yang eksekusi pada RStudio merupakan jumlah iterasi, pusat *cluster*, fungsi objektif, serta derajat keanggotaan setiap objek terhadap masing- masing *cluster*, ditampilkan ke dalam *plot* dari tiap *cluster* yang tercipta

B. Menentukan Jumlah Cluster Optimum

Jika telah melewati proses *cluster* menggunakan *Fuzzy C- Means*, sehingga akan dicoba proses metode *Elbow* untuk mencari *cluster* optimum dalam memastikan *cluster* terbaik. Hasil perhitungan metode *Elbow* pada RStudio ditampilkan dalam nilai SSE (Sum of Squares Error) serta grafik yang terdiri dari sumbu x serta sumbu y. Nilai pada sumbu x merupakan jumlah *cluster* yaitu *cluster* 2 sampai *cluster* 10, sebaliknya sumbu y merupakan nilai SSE (Sum of Squares Error) yang dihasilkan dari pusat setiap *cluster*

C	SSE	Jarak	Keterangan
1	3756052074	-	-
2	975707614	2780344460	Jarak C1 ke C2
3	587145618	388561996	Jarak C2 ke C3
4	244689813	342455805	Jarak C3 ke C4
5	231754385	12935428	Jarak C4 ke C5
6	151155868	80598517	Jarak C5 ke C6
7	98790377	52365491	Jarak C6 ke C7
8	87581280	11209097	Jarak C7 ke C8
9	74645852	1293542	Jarak C8 ke C9
10	138233016	-63587164	Jarak C9 ke C10

Tabel 1. Hasil Nilai *Elbow*



diambil sebagai *cluster* terbaik yaitu titik yang membentuk suatu siku. Hasil pada titik yang membentuk siku yaitu pada titik yang terjalin penyusutan yang ekstrem antara 2 titik *cluster* serta setelah itu diiringi oleh nilai yang relatif konstan. Jadi, *cluster* optimum ada pada *cluster* 2.

C. Profiling Cluster Optimum

Proses clustering menggunakan algoritma Fuzzy C-Means untuk mencari cluster terbaik dengan metode Elbow, diketahui cluster optimum berada di cluster 2. Untuk mengetahui karakter dari masing-masing cluster, maka digunakan profiling cluster data jumlah kasus Crime Total, Crime Cleared, Crime Rate, Clearance Rate dari masing-masing provinsi yang masuk ke dalam kelompok C1 dan C2 melalui proses perhitungan yang telah dilakukan sebelumnya.

Pada cluster 1 terdiri dari 4 anggota kepolisian daerah dengan rentang Crime total 21.498 - 34.655 kasus, Crime cleared 11.521 - 30.577 kasus, Clearance Rate 58.74 - 88.23 kasus, Crime Rate 67-255 kasus.

Pada cluster 2 terdiri dari 30 anggota kepolisian daerah dengan rentang Crime Total 396 - 16.209 kasus, Crime Cleared 277 - 12.150 kasus, Clearance Rate 24,25 - 83,59 kasus, Crime Rate 27 - 416 kasus. Dari hasil karakteristik data tersebut, cluster 1 memiliki jumlah kasus lebih tinggi dibandingkan dengan cluster 2.

3. DAFTAR PUSTAKA

1. Kusumadewi, Sri dan Hari Purnomo. 2004. Aplikasi Logika Fuzzy untuk Pendukung Keputusan. Edisi 1. Yogyakarta; Graha Ilmu. (hal: 1,3,7-11,25-26,30,39-45).
2. Han, J. dan M. Kamber. 2006. Data Mining Concepts and Techniques Second Edition. San Francisco: Morgan Kaufmann.
3. H. Maslow. 1943. A Theory of Human Motivation, Originally Published in Psychological Review, 50, 370-396. York University, Toronto, Ontario ISSN 1492-3713.
4. Melvin M. Belli, Quinney .1971. The Social Reality of Crime, 69 MICH. L. REV. 978
5. Statistik Kriminal 2019, ISSN : 2089-5291 No. Publikasi: 04330.1903 Katalog: 4401002 Penerbit: © Badan Pusat Statistik.
6. Rinawati, R., Sihombing, E. G., Dewi, L. S., & Arisawati, E. 2020. Analisis Algoritma Datamining pada Kasus Daerah Pelaku Kejahatan Pencurian Berdasarkan Provinsi. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 4(1), 77-87