Application of the K-Means Clustering Algorithm in Mapping the Regional Voter Strategy for the Legislative Candidates for the DPR RI

by Jurnal Komitek

Submission date: 01-Feb-2022 01:07AM (UTC+0900)
Submission ID: 1752021070
File name: 34._dhika-alfatah-stia.doc (494.5K)
Word count: 3682
Character count: 24008

Penerbit ADM
RUMAH PENGETAHUAN
Jurnal KOMITEK
Jurnal Komputer, informasi Dan Teknologi

Application of the K-Means Clustering Algorithm in Mapping the Regional Voter Strategy for the Legislative Candidates for the DPR RI

Penerapan Algoritma K-Means Clustering dalam Memetakan Strategi Daerah Pemilih pada Calon Legislatif DPR RI

Dhika Alfatah
Sekolah Tinggi Ilmu Administrasi Bengkulu
Email: [1)] dhikaalfatah8@gmail.com

How to Cite :
Alfatah, D. (2021). Application of the K-Means Clustering Algorithm in Mapping the Regional Voter Strategy for the Legislative Candidates for the DPR RI. JURNAL Komitek, 1(2). DOI: https://doi.org/10.53697/jkomitek.v1i2

ARTICLE HISTORY
Received [29 November 2021]
Revised [10 Desember 2021]
Accepted [26 Desember2021]

KEYWORDS
Application of the K-Means Clustering Algorithm, Mapping the Regional, Strategy for the Legislative Candidates

ABSTRAK

Penelitian ini menerapkan Data Mining dengan menggunakan metode Clustering untuk memetakan strategi pemilihan dearah pemilih pada Calon Legislatif Kota Bengkulu Algoritma yang digunakan yaitu K-Means Clustering, di mana data dikelompokkan berdasarkan karakteristik yang sama akan dimasukkan ke dalam kelompok yang sama dan set data yang dimasukkan ke dalam kelompok tidak tumpang tindih. Informasi yang ditampilkan berupa kelompok – kelompok data produk berdasarkan tingkat suara pemilih disetiap desa, sehingga diketahui daerah/desa mana yang tingkat pemilih tinggi, sedang ataupun rendah. Pengujian dilakukan dengan aplikasi RapidMiner 5.3, sehingga menghasilkan cluster-cluster yang tingkat pemilihnya tinggi, sedang atau pun rendah

ABSTRACT

This study applies Data Mining by using the Clustering method to map the electoral strategy for the Legislative Candidates in Bengkulu City. The algorithm used is K-Means Clustering, where data grouped based on the same characteristics will be included in the same group and data sets are entered into the same group. in non-overlapping groups. The information displayed is in the form of product data groups based on the voter vote level in each village, so that it is known which regions/villages have high, medium or low voter levels. The test was carried out with the RapidMiner 5.3 application, resulting in clusters with high, medium or low voter rates..

PENDAHULUAN

Strategi dalam menghadapi pemilihan legislatif daerah merupakan perencanaan yang cermat yang disusun dan dilaksanakan oleh tim kampanye yang memiliki tujuan mencapai kemenangan atas sasaran yang ditentukan dalam pilkada. Sasaran merupakan apa yang ingin dicapai oleh kandidat dan tim kampanye dalam hal ini adalah target dukungan pemilihan yang diwujudkan dalam pemberian suara kepada kandidat tersebut.

Ruang lingkup Strategi dalam menghadapi pemilihan legislaif daerah merupakan perencanaan yang cermat yang disusun dan dilaksanakan oleh tim kampanye yang memiliki tujuan mencapai kemenangan atas sasaran yang ditentukan dalam pilkada. Sasaran merupakan apa yang ingin dicapai oleh kandidat dan tim kampanye dalam hal ini adalah target dukungan pemilihan yang diwujudkan dalam pemberian suara kepada kandidat tersebut. Ruang lingkup pembahasan strategi tak sebatas pada tatanan konsep atau rencana, namun yang terpenting adalah bagaimana kandidat dan tim kampanye tersebut mengimplementasikannya dilapangan.

Dalam persaingan menjadi anggota DPR RI tidaklah mudah, perlu adanya strategi masing – masing antara Caleg DPR RI untuk mendapatkan kursi jabatannya. Para calon legislatif bersaing dengan berbagai cara apapun. Namun jika dalam pemetaan wilayah basis suara pemilihnya yang tidak diketahui.

Maka tentu akan berpengaruh besar dalam strategi pemenangan. Tim pemenangan Calon Legistalif harus dapat memetakan situasi dan kondisi dilapangan bisa dengan dibantu dengan hasil survey, data kongkret dan lain - lain. Tentu kondisi tersebut berbeda – beda. Oleh karena itu langkah yang dilakukan bervariasi. Untuk mengetahui kondisi tersebut salah satunya adalah melakukan pemanfaatan data perolehan suara dari data periode yang sebelumnya.

Ketersediaan data yang banyak dan kebutuhan akan informasi atau pengetahuan sebagai pendukung pengambilan keputusan untuk memetakan daerah pilihannya dan dukungan infrastruktur dibidang teknik informatika merupakan cikal bakal dari lahirnya teknologi data mining. Penggunaan teknik Data Mining diharapkan dapat membantu mempercepat proses pengambilan keputusan, memungkinkan perusahaan untuk mengelola informasi yang terkandung didalam data hasil perolehan suara dengan variabel yang mempengaruhi perolehan suara sehingga menjadi sebuah pengetahuan (knowledge) yang baru. Melalui pengetahuan yang didapat, maka tim Calon Legislatif akan lebih mudah memetakan daerah pilihan dengan mengetahui kondisi suara yang didapat pada periode lalu.

Data Mining sudah banyak diterapkan dalam berbagai bidang. Salah satunya metode clustering khususnya metode K-Means, antara lain oleh Oyelade (2010) menjelaskan tentang Metode K-Mean dapat dimplementasikan untuk menganalisis kinerja Mahasiswa akademik. Kemudian penelitian yang dilakukan oleh Bhatia (2013) menyatakan bahwa K- Means clustering merupakan algoritma yang mudah dan sederhana. Algoritma Clustering memiliki daya tarik yang luas dan digunakan untuk analisis data eksplorasi. Nikhil (2013) mengemukakan tentang peningkatan algoritma Clustering K-Means menggunakan waktu yang lebih baik dan akurat. Analisis cluster merupakan pengelompokan satu set objek dengan benda yang sama yang lebih mirip satu sama lain. Sunila (2013) melakukan analisis dengan 4 (empat) algoritma clustering yaitu K-Means Clustering, Farther first clustering, Density Based Clustering, Filtered clusterer. Analisis keempat teknik ini diimplementasikan dengan software WEKA. Bondu (2013) menjelaskan tentang peningkatan algoritma K-Means clustering dapat diimplementasikan untuk informasi donor darah. Sehingga dengan adanya peningkatan K-Means clustering ini, dapat menghasilkan cluster yang akurat untuk menemukan informasi donor. Kemudian Rajashree (2010) mengimplementasikan Algoritma K- Mean clustering untuk meningkatkan efisiensi dan akurasi tugas pertambangan dengan data yang memiliki dimensi tinggi. Ramesh (2012) mengimplementasikan Data Mining dengan algoritma K-Means untuk Multicore Heterogen Compute Cluster. Kemudian penelitian yang dilakukan oleh Ramzi (2015) melakukan peningkatan algoritma K-Means clustering untuk penemuan pola dalam data kesehatan. Dengan Demikian maka tim pemenangan Caleg tersebut akan bisa mengambil inisiatif yang diperlukan. Algoritma k-maens bekerja dengan mengclustering unsur – unsur yang dapat membantu memetakan kondisi daerah pemilih..

LANDASAN TEORI

Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) merupakan bagian dari Data Mining. Dimana Knowledge Discovery in Database (KDD) itu sendiri merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Dimana hasil dari proses tersebut akan membentuk pola-pola dari kumpulan data, yang sering disebut dengan pengenalan pola (pattern recognition). Dalam penelitian ini fokus pada metode Data Mining dengan kasus pengelompokan data (clustering). Adanya data dalam skala besar memungkinkan metode data mining dengan teknik clustering yang dapat mengelompokkan data ke dalam beberapa kelompok yang diinginkan (Deka, et al, 2014). Goldie (2012), memaparkan Knowledge Discovery in Database (KDD) terdiri dari tiga proses utama, yaitu :

Data Mining

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai database besar. Dan Data Mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Di mana hasil dari proses penggalian tersebut akan membentuk pola-pola dari kumpulan data, yang sering disebut dengan pengenalan pola (Deka, et al, 2014). Data Mining adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data warehouse, atau penyimpanan informasi lainnya. Data mining berkaitan dengan

bidang ilmu–ilmu lain, seperti database system, data warehousing, statistik, machine learning, information retrieval, dan komputasi tingkat tinggi. Selain itu, Data Mining didukung oleh ilmu lain seperti neural network, pengenalan pola, spatial data analysis, image database, signal processing.

Data mining didefinisikan sebagai proses menemukan pola-pola dalam data. Proses ini otomatis atau seringnya semiotomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi (Budanis dan Nofi, 2014). Data Mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Lindawati, 2008). Data Mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode Data Mining ini dapat digunakan untuk mengambil keputusan di masa depan. Data Mining ini juga dikenal dengan istilah pattern recognition yang merupakan metode pengolahan data berskala besar oleh karena itu Data Mining ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi. Secara umum kajian Data Mining membahas metode-metode seperti, clustering, klasifikasi, regresi, seleksi variable, dan market basket analisis (Johan, 2013). Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Afrisawati, 2013). Data Mining sudah banyak diterapkan dalam berbagai bidang. Salah satunya metode clustering khususnya metode K-Means, antara lain oleh Oyelade (2010) menjelaskan tentang Metode K-Mean dapat dimplementasikan untuk menganalisis kinerja Mahasiswa akademik. Kemampuan untuk memantau perkembangan prestasi akademik siswa adalah masalah penting untuk akademik pendidikan tinggi. Sebuah sistem yang digunakan untuk menganalisis siswa. Hasil yang didapatkan berdasarkan analisis cluster dengan menggunakan standar statistik algoritma untuk mengatur data yang nilai sesuai dengan tingkat kinerja.

Tujuan penelitian ini adalah untuk menentukan keputusan yang efektif oleh akademik. Kemudian Ramesh (2012) mengimplementasikan Data Mining dengan algoritma K-Means untuk Multicore Heterogen Compute Cluster. Dalam penelitian ini, kinerja algoritma K-Mean Data diimplementasikan pada heterogen untuk menghitung cluster dengan pemrograman multi core. Program multicore diimplementasikan untuk komputasi paralel dan memanfaatkan kekuatan komputasi maksimum cluster heterogen. Dalam penelitian ini juga menganalisis efisiensi dan kinerja algoritma Data Mining K-Mean untuk dataset yang besar seperi chess.txt. Dataset ini dibagi menjadi jumlah core dan inti menghitung dataset independen dan membuat cluster data dari dataset yang sama pada setiap inti prosesor.

Tahapan–Tahapan Data Mining

Karena Data Mining adalah suatu rangkaian proses. Data Mining dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantaraan knowledge base (Lindawati, 2008).

Menurut Budanis dan Nofi (2014), Data Mining memiliki tahapan–tahapan antara lain:

1. Pembersihan data (data cleaning). Pembersihan data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan.
2. Integrasi data (data integration). Integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru.
3. Seleksi Data (Data Selection). Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database.
4. Transformasi data (Data Transformation). Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam Data Mining.
5. Proses mining. Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.
6. Evaluasi pola (pattern evaluation). Untuk mengidentifikasi pola-pola menarik kedalam knowledge based yang ditemukan.

Presentasi pengetahuan (knowledge presentation), Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna

K-Means Clustering

Menurut Tutik (2014) K-Means adalah salah satu algoritma yang menggunakan metode partisi. K-Means adalah algoritma clustering yang membagi masing-masing item data ke dalam satu cluster. K-Means adalah suatu teknik pengelompokan data yang mana keberadaan tiap-tiap titik data dalam suatu

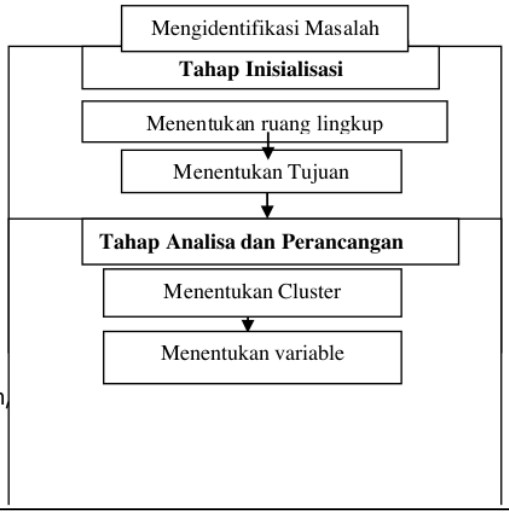
cluster ditentukan oleh derajat keanggotaan. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981 (Deka Dwinavinta et al, 2014). Menurut Johan (2013) K-means clustering merupakan salah satu metode data clustering non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih cluster/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster/kelompok dan data yang memiliki karakteristik yang berbeda dikkelompokan dengan cluster/kelompok yang lain sehingga data yang berada dalam satu cluster/kelompok memiliki tingkat variasi yang kecil.

K-Means merupakan salah satu metode pengelompokan data nonhirarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok (Afrisawati, 2013). K-Means clustering sudah banyak digunakan dalam penelitian diantaranya oleh Bhatia (2013) yang menyatakan bahwa K- Means clustering merupakan algoritma yang mudah dan sederhana. Algoritma Clustering memiliki daya tarik yang luas dan digunakan untuk analisis data eksplorasi. Dalam penelitiannya mengunakan Algoritma Clustering K-Mean untuk menganalisis data eksplorasi. Sehingga menghasilkan eksperimental pendekatan yang berbeda untuk k cluster. Dan membandingkan pada dataset yang berbeda dengan menggunakan K-Mean. Algoritma ini diimplementasikan menggunakan MATLAB R2009b. Kemudian penelitian yang dilakukan oleh Nikhil (2013) yang mengemukakan tentang peningkatan algoritma Clustering K-Means menggunakan waktu yang lebih baik dan akurat. Analisis cluster merupakan pengelompokan satu set objek dengan benda yang sama yang lebih mirip satu sama lain. Sunila (2013) melakukan analisis dengan 4 (empat) algoritma clustering yaitu K-Means Clustering, Farther first clustering, Density Based Clustering, Filtered clusterer. Analisis keempat teknik ini diimplementasikan dengan software WEKA. Keempat cluster ini dianalisa dan dibandingkan. Kesimpulan dalam penelitian ini dalah kinerja algoritma K-Means lebih baik dari Dansity. Dan semua algoritma memiliki beberapa ambiguitas sendiri. Kualitas semua algoritma menjadi sangat baik ketika menggunakan dataset besar. Algoritma K-Means lebih cepat melakukan pengelompokan dibandingkan dengan algoritma lainnya. Bondu (2013) menjelaskan tentang peningkatan algoritma K-Means clustering dapat diimplementasikan untuk informasi donor darah. Penelitian ini di latar belakangi oleh jumlah kecelakaan dan penyakit yang meningkat dan mengkhawatirkan. Sehingga mengakibatkan peningkatan besar dalam permintaan untuk darah. Dalam penelitian ini peningkatan algoritma K- Means clustering ini, dapat menghasilkan cluster yang akurat untuk menemukan informasi donor. Kemudian Rajashree (2010) mengimplementasikan Algoritma K- Mean clustering untuk meningkatkan efisiensi dan akurasi tugas pertambangan dengan data yang memiliki dimensi tinggi.

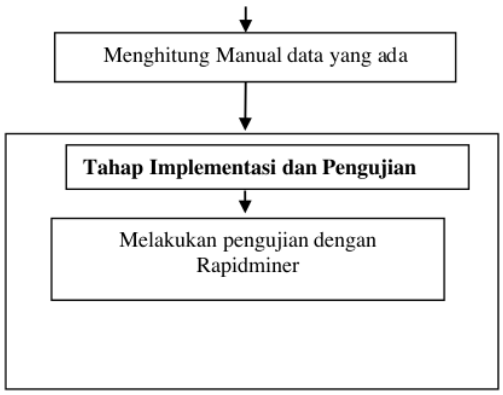
METODE PENELITIAN

Penelitian ini memiliki beberapa tahapan dalam pelaksanaan kegiatan yang tertuang pada kerangka kerja penelitian yaitu identifikasi masalah, analisa masalah, menentukan tujuan, mempelajari literature, mengumpulkan data, analisa Algoritma *K-Means Clustering*, pengolahan data, implementasi, dan pengujian urutan kerangka kerja yang harus diikuti, urutan kerangka kerja ini merupakan langkah–langkah yang dilakukan dalam penulisan. Adapun kerangka kerja yang digunakan dalam penulisan ini adalah seperti terlihat pada gambar 1.

Kerangka kerja yang peneliti lakukan dalam penelitian dapat dilihat pada gambar berikut :



Penerbit ADM
RUMAH PENGETAHUAN
Jurnal KOMITEK
Jurnal Komputer, informasi Dan Teknologi



Gambar 1 Kerangka Kerja (Frame Work Penelitian)

HASIL DAN PEMBAHASAN

Hasil

Dari hasil penelitian dalam memetakan daerah strategi pemilih, pada calon legislatif Anggota DPR RI dengan algoritma K-means, maka data yang didapat diolah menjadi data yang sesuai dengan kebutuhan dalam perhitungan algoritma, sehingga dapat dilakukan perhitungan.

Hasil analisa yang akan diperoleh adalah dapat mengetahui kondisi desa, mulai dari kategori tingkat pemilih tinggi, tingkat pemilih sedang, tingkat pemilih rendah, sehingga tim pemenangan Calon Legislatif dapat memetakan strategi daerah mana yang perlu dilakukan serangan terhadap konsidi sekarang. Adapun data yang akan diuji yaitu diambil dalam wilayah DAPIL (Daerah Pemilihan) Bengkulu Selatan. Teridiri dari 11 kecamatan terdiri dari 158 desa/kelurahan. Namun untuk pengujian secara manual, maka data diambil secara random dengan dua Kecamatan diantaranya Kecamatan 29 kelurahan.

Dalam memetakan strategi daerah pemilih, maka peneliti menggunakan data hasil penelitian dari Tim Pemenangan salah satu caleg terpilih yaitu data hasil perolehan suara, dan Kontribusi (yang pernah dilakukan) pada masyarakat yaitu jumlah kunjungan dan bantuan. Sehingga jumlah persentase suara yang didapat pada periode lalu, dan kontribusi adalah menjadi penentu dalam memetakan wilayah pemilih.

Variabel yang digunakan sangat berpengaruh dalam memetakan strategi pemilihan adalah Persentase hasil suara dalam masing – masing desa pada periode lalu, kemudian dengan kontribusi yang telah dilakukan sebelumnya, dan rekapan data kunjungan ke tiap Desa. Variabel ini disusun sesuai dari Tim Pemenangan.

Dalam kebutuhan data yang akan diolah, maka ada beberapa langkah rangkaian yang harus dilakukan dalam pengolahan data dengan melakukan beberapa tahapan *praproses*.

Gambar 2 Flowchart PraProses Data Mining

Dari gambar 2 diatas merupakan tahapan – tahapan yang harus dilakukan sebelum melakukan proses clustering dengan algoritma k-means.

1. **Data Cleaning (Pembersihan Data).** Melakukan pembersihan data dari data yang tidak relevan dan tidak konsisten.
2. **Integrasi Data.** Merupakan penggabungan dari data – data pengdukung pada pengolahan data dari beberapa data digabung dalam satu data baru. Pada penelitian ini data yang akan digabung ialah Daftar Pemilih Tetap (DPT), Data Bantuan (Yang pernah dilakukan) dalam tiap daerah,dan Rekapan Data Kunjungan.
3. **Seleksi Data.** Dari data yang didapat maka dilakukan seleksi data , dapil Bengkulu Selatan memiliki 158 data desa namun karena tidak semua data – data tersebut dipakai, dan akan disesuaikan dengan analisis yang akan diolah. Maka diambil sampel sebanyak dua kecamatan Manna dan Kota Manna dengan 29 data desa.
4. **Tranformasi Data.** Setelah dilakukan seleksi data maka mentranformasikan atau mengubah data kedalam format yang sesuai untuk diproses dalam datamining. Tranformasi dilakukan sehingga data siap diolah.
5. **Proses Mining.** Melakukan Proses data mining dengan menerapkan algoritma k-means untuk menemukan pengetahuan dan tersembunyi dari data.

Pembahasan

Dalam memetakan suara wilayah pemilih dengan algoritma K-Means maka mengikuti langkah – langkah dari algoritma K-Means berdasarkan data yang didapat kemudian dilakukan pengolahan data sehingga sesuai dengan aturan dari algoritma K-Means clustering,

Langkah –Langkah Algoritma K-Means Clustering

Data yang digunakan dalam pengelompokan atau clustering adalah data perolehan suara, data Kontribusi terhadap daerah pemilih (Bantuan dan Kunjungan) pada satu kecamatan pada periode 2014. Untuk mengelompokan data diatas dengan menjadi beberapa cluster, maka dilakukan beberapa langkah sebagai berikut :

1. **Menentukan Jumlah Cluster**

   Pada penelitian ini data – data akan di kelompokan menjadi tiga cluster. Yaitu cluster 0 merupakan kelompok dengan kondisi daerah pemilih dengan tingkat sedang, , cluster 1 kelompok suara pemilih dengan tingkatan tinggi, dan cluster 2 kelompok produk dengan suara pemilih berpotensi rendah.
2. **Menentukan titik pusat *cluster* (*centroid*).**

Penerbit ADM RUMAH PENGETAHUAN
Jurnal KOMITEK
Jurnal Komputer, informasi Dan Teknologi

Dalam penelitian ini titik pusat awal ditentukan secara *random* atau acak. Maka didapat titik pusat dari setiap *cluster* pada tabel 4.3 berikut:

Tabel 1 Titik Pusat Awal Setiap Cluster

Data Ke	Nama Desa/ Kelurahan	Persentase Suara	Bantuan	Kunjungan
1	Gunung Kembang	7.48	1	0
36	Manggul	14.05	1	5
153	Padang Niur	2.27	1	1

Menghitung Obyek Ke Pusat

Setelah Menentukan Centroid awal, maka sesiap data mendapatkan centroid terdekatnya yaitu dengan menghitung jarak setiap data ke masing- masing centriod menggunakan rumus korelasi antar dua obyek yaitu *Euclidean Distance*. Adapun penhitungan *centroida* awal secara manual dilakukan penghitungan 10 sampel. Seperti pada perhitungan dibawah ini :

Persamaan :

$$D(C0,1) = \sqrt{(x\text{-}1i - x1j)^2 + (x\text{-}2i - x2j)^2 + \ ...+ (x\text{-}ki - xkj)^2} \quad \ldots (1)$$

Dimana :
D(I,j) = Jarak data ke I ke pusat cluster j
Xki = Data ke I Pada atribut data ke k
Xkj = Titik Pusat ke j pada atribut k

Cluster	Kecamatan	Desa	Kondisi
Cluster 2	Kota manna	Padang niur	Tingkat pemilih tinggi
	Kota manna	Tebat kubu	

Dari tabel kesimpulan di atas dapat diketahui bahwa daerah pemilih sedang , tingkat pemilih tinggi, dan tingkat pemilih rendah sehingga bagi tim kampanye sudah memiliki gambaran secara lebih detail untuk dapat melakukan beberapa strategi khusus untuk menambah jumlah pemilih pada daerah yang rendah yaitu cluster 2. Sedangkan daerah pemilih yang dalam keadaan tertinggi atau tingkat aman yaitu di cluster 1. Sedangkan kondisi dalam keadaan sedang pada cluster 0.

KESIMPULAN DAN SARAN

Kesimpulan

1. Dalam Memetakan strategi pemilih pada calon legislatif harus mengetahui kondisi setiap daerah pilihan, konsidi tersebut sangat membantu bagi tim pemenangan dalam memetakan stratetgi. Harus optimis bahwasanya harus dikuasai dengan baik, mendapatkan data – data yang kongkret, dan selanjutnya kita harus mengetahui kebutuhan – kebutuhan pada daerah pilihan, maka dengan begitu , jalan untuk memenangkan pemilu akan lebih mudah.
2. Untuk menerapkan data mining dalam memetakan strategi pemilih pada calon legislatif dengan algotritma k-means clustering yaitu ada dua tahapan.
   a. Pra proses. Tahapan praproses ini adalah proses pengolahan data – data yang akan diolah,pertama yang dilakukan, pembersihan data (data cleaning), Integrasi data (data integration), seleksi data (data Selection), tranformasi data (data traformation), dan terakhir dilakukan proses mining.
   b. Proses. Tahapan menerapkan algoritma kedalam kasus yang ada, langkahnya menetukan jumlah kluster, yang ditentukan 3 cluster yaitu dengan tingkatan pemilih tinggi, sedang, rendah. Setelah itu menentukan centroid dari data yang ada, selanjutnya dilakukan penghitungan jarak minimumnya, dan mengahsilkan cluster baru, dan dilakukan secara berulang sampai cluster tidak berpindah lagi.
3. Implementasi dengan rapidminer dilakukan untuk menerapkan bahwa algoritma k-means dengan data yang ada dapat diuji dengan baik sehingga menghasilkan pengetahuan baru. Pengujian ini

dengan rapidminer berfungsi untuk menganalisa kembali dan memastikan data yang diolah adalah benar.

4. Bahwa algoritma k-means ini dapat diterapkan dalam kasus memetakan strategi daerah pemilih sehingga hasil penelitian ini menghasilkan suatu model pengetahuan yang berguna bagi tim pemenangan dalam konteks pemilu.

Saran

1. Algoritma k-means sangat mendukung prediksi suatu keputusan, mungkin dalam kasus hal yang lain juga, bisa menilai kinerja karyawan, kategori beasiswa dan lainya yang dapat diteliti bagi peneliti selanjutnya
2. Dalam memetakan strategi pemilu tidaklah mudah, data –data yang lengkap tentu akan sangat menbantu peneliti bagi menentukan strategi pemenangan Calon Legislatif
3. Sebelum menentukan kasus agar dipahami terlebih dahulu kasus yang akan diteliti sehingga memudahkan peneliti untuk pengolah data sehingga menjadi pengaetahuan baru.

.

DAFTAR PUSTAKA

Afrisawati. (2013). *"Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means."* Ed. *Pelita Informatika Budi Darma, Volume : V,No. 3.*

Bondu Venkateswarlu and Prof G.S.V.Prasad Raju. (2013). *"Mine Blood Donors Information through Improved K-Means Clustering." International Journal of Computational Science and Information Technology (IJCSITY) Vol.1, No.3.*

Budanis Dwi Meilani dan Nofi Susanti, (2014), *"Aplikasi Data Mining Untuk Menghasilkan Pola Kelulusan Siswa Dengan Metode Naïve Bayes."* Ed. *Jurnal LINK Vol 21/No.2.*

Deka Dwinavinta,et.al. (2014). *"Klasterisasi Judul Buku dengan Menggunakan Metode K-Means."* Ed. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI).*

Dr. M.P.S Bhatia and Deepika Khurana. (2013). *"Experimental study of Data clustering using k-Means and modified algorithms." International Journal of Data Mining & Knowlegde Management Process (IJDKP). Vol. 3 No.3.*

Er. Nikhil Chaturvedi and Er. Anand Rajavat. (2013). *"An Improvement in K-mean Clustering Algorithm Using Better Time and Accuracy." International Journal of Programming Language and Aplications (IJPLA) Vol. 3 No.4.*

Fadlina. (2014). *"Data Mining untuk Analisa Tingkat Kejahatan Jalanan Dengan Algoritma Association Rule Metode Apriori (Studi Kasus Di Polsekta Medan Sunggal)." Voumel.III No.1.*

Goldie Gunadi dan Dana Indra Sensuse, (2012). *"Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori dan Frequent Pattern Growth (FP-Grouth) : Studi Kasus Percetakan PT. Gramedia."* Ed. *Jurnal Telematika MKom Vol.4 No.1.*

Godara, Sunila. (2013). *"Analysis Of Various Clustering Algorithms".*Ed Journal International IJITEE, Vol 3 Issue 1.

Johan Oscar Ong. (2013). *"Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing Presiden University."* Ed. *Jurnal Ilmiah Teknik Industri, Vol.12, No.1.*

Lindawati. (2008). *"Data Mining Dengan Teknik Clustering Dalam Pengklasifikasian Data Mahasiswa Studi Kasus Prediksi Lama Studi Mahasiswa Universitas Bina Nusantara."* Ed. *Seminar Nasional Informatika.*

Mujib Ridwan, et.al. (2013). *"Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier."* Ed. *Jurnal EECCIS Vol.7, No.1.*

Oyelade. O. J, dkk (2010). *"Application of k-Means Clustering algorithm for prediction of Students' Academic Performance." International Journal of Computer Science and Information Security. Vol. 7 No.1.*

Rajashree, dkk. (2010). *"A hybridized K-means clustering approach for high dimensional dataset." International Journal of Engineering, Science and Technology Vol. 2, No. 2.*

Ramesh Singh Yadava and P.K.Mishra. (2012). *"Performance Analysis of High Performance k-Mean Data Mining Algorithm for Multicore Heterogeneous Compute Cluster." Vol. 2 No.4.*

Ramzi A. Haraty, dkk. (2015*). "An Enhancedk-Means Clustering Algorithm for Pattern Discovery in Healthcare Data." Hindawi Publishing Corporation International Journal of Distributed Sensor Networks Volume 2015, Article ID 615740.*

Tutik Khotimah. (2014). *"Pengelompokan Surat Dalam Al Qur'an Menggunakan Algoritma K-Means."* Ed. *Jurnal Simetris, Vol 5 No 1.*

Application of the K-Means Clustering Algorithm in Mapping the Regional Voter Strategy for the Legislative Candidates for the DPR RI

ORIGINALITY REPORT

24%	25%	22%	16%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.jurnal.stmikroyal.ac.id Internet Source	2%
2	eprints.sinus.ac.id Internet Source	2%
3	es.slideshare.net Internet Source	2%
4	repository.uksw.edu Internet Source	2%
5	digilib.uin-suka.ac.id Internet Source	2%
6	journal-isi.org Internet Source	1%
7	repository.unpkediri.ac.id Internet Source	1%
8	www.stmik-time.ac.id Internet Source	1%

9	Ratih Yulia Hayuningtyas. "Penerapan Algoritma Naïve Bayes untuk Rekomendasi Pakaian Wanita", Jurnal Informatika, 2019 Publication	1 %
10	medium.com Internet Source	1 %
11	journal.pdmbengkulu.org Internet Source	1 %
12	nero.trunojoyo.ac.id Internet Source	1 %
13	repository.unmuhjember.ac.id Internet Source	1 %
14	www.ie.its.ac.id Internet Source	1 %
15	Frskila Parhusip, Agus Perdana Windarto, Irfan Sudahri Damanik, Eka Irawan, Ilham Syahputra Saragih. "KLASIFIKASI FAKTOR PENYEBAB RENDAHNYA MINAT MAHASISWA DALAM MENULIS ARTIKEL ILMIAH", Jurnal RESISTOR (Rekayasa Sistem Komputer), 2021 Publication	1 %
16	Triuli Novianti, Iwan Santosa. "PENENTUAN JADWAL KERJA BERDASARKAN KLASIFIKASI DATA KARYAWAN MENGGUNAKAN METODE DECISION TREE C4.5 (Studi Kasus Universitas Muhammadiyah Surabaya)", Jurnal Komunika	1 %

: Jurnal Komunikasi, Media dan Informatika, 2016
Publication

17	Deri Lianda, Niko Surya Atmaja. "Prediksi Data Buku Favorit Menggunakan Metode Naïve Bayes (Studi Kasus: Universitas Dehasen Bengkulu)", Pseudocode, 2021 Publication	1 %
18	Agus Nur Khomarudin, Supratman Zakir, Rina Novita, Endrawati, Mohd Zahiri bin Awang Mat, Efmi Maiyana. "K-Mean Clustering Algorithm in Grouping Prospective Scholarship Recipients", Journal of Physics: Conference Series, 2021 Publication	1 %
19	download.garuda.ristekdikti.go.id Internet Source	1 %

Exclude quotes	On	Exclude matches	< 25 words
Exclude bibliography	On		

Application of the K-Means Clustering Algorithm in Mapping the Regional Voter Strategy for the Legislative Candidates for the DPR RI

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9