

## Analisis Data Mining Terhadap Data Faktor Perceraian Di Sumatera Utara Dengan Metode K-Means Clustering

Ayu Ofta Sari<sup>1</sup>, Muhammad Iqbal<sup>2</sup>

<sup>1,2</sup>Program Studi Magister Teknologi Informasi, Universitas Pembangunan Pancabudi, Medan, Indonesia

Email: <sup>1</sup>ayuofta2849@gmail.com, <sup>2</sup>muhammadiqbal@dosen.pancabudi.ac.id

Email Penulis Korespondensi: ayuofta2849@gmail.com

### Abstrak

Tingginya angka perceraian di Provinsi Sumatera Utara terus meningkat setiap tahunnya. Penelitian ini bertujuan untuk mengelompokkan data perceraian berdasarkan kabupaten dan faktor penyebabnya dengan menerapkan metode K-Means Clustering menggunakan RapidMiner. Data yang dianalisis mencakup 34 kabupaten/kota di Sumatera Utara, dengan indikator utama seperti faktor ekonomi, kekerasan dalam rumah tangga, dan faktor lain yang berkontribusi terhadap perceraian. Hasil analisis menunjukkan bahwa data dapat dikelompokkan menjadi tiga klaster dengan karakteristik yang berbeda. Klaster 0, yang mencakup 26 wilayah, menunjukkan tingkat perceraian tinggi yang dipengaruhi oleh faktor ekonomi dan sosial yang kompleks. Klaster 1, dengan 3 wilayah, mencerminkan kondisi unik yang memerlukan kajian lebih mendalam untuk mengidentifikasi penyebab utama. Sementara itu, Klaster 2, dengan 4 wilayah, mencerminkan tingkat perceraian yang relatif lebih rendah dengan pengaruh faktor penyebab yang lebih sederhana. Hasil penelitian ini memberikan wawasan strategis bagi instansi terkait untuk merumuskan kebijakan yang lebih efektif dalam menekan angka perceraian, seperti program edukasi pasangan, peningkatan ekonomi, dan intervensi sosial yang tepat sasaran. Dengan memanfaatkan pendekatan berbasis data ini, diharapkan pemerintah dan lembaga terkait dapat mengambil langkah-langkah preventif yang lebih baik untuk meningkatkan kesejahteraan masyarakat.

**Kata Kunci :** Faktor Perceraian, Rapidminer, K-Means Clustering

### Abstract

The high divorce rate in North Sumatra Province continues to increase every year. This research aims to group divorce data based on districts and causal factors by applying the K-Means Clustering method using RapidMiner. The data analyzed covers 34 districts/cities in North Sumatra, with main indicators such as economic factors, domestic violence and other factors that contribute to divorce. The analysis results show that the data can be grouped into three clusters with different characteristics. Cluster 0, which includes 26 regions, shows a high divorce rate influenced by complex economic and social factors. Cluster 1, with 3 regions, reflects unique conditions that require more in-depth study to identify the main causes. Meanwhile, Cluster 2, with 4 regions, reflects a relatively lower divorce rate with the influence of simpler causal factors. The results of this research provide strategic insight for relevant agencies to formulate more effective policies in reducing the divorce rate, such as couple education programs, economic improvement, and targeted social interventions. By utilizing this data-based approach, it is hoped that the government and related institutions can take better preventive steps to improve community welfare.

**Keyword :** Divorce Factor, Rapidminer, K-Means Clustering

## 1. PENDAHULUAN

Perceraian merupakan isu sosial yang kompleks dan menjadi perhatian penting dalam masyarakat. Di Indonesia, jumlah kasus perceraian terus meningkat dalam beberapa tahun terakhir. Pada tahun 2022, jumlah kasus perceraian di Indonesia mencapai 516.334, meningkat 15,31% dibandingkan dengan 447.743 kasus pada tahun sebelumnya. Angka ini menjadi yang tertinggi dalam kurun waktu enam tahun terakhir. Sebagian besar kasus perceraian yang tercatat pada tahun tersebut adalah cerai gugat, yaitu permohonan cerai yang diajukan oleh pihak istri dan telah diputuskan oleh pengadilan. Sepanjang tahun 2022, tercatat sebanyak 388.358 kasus cerai gugat, atau setara dengan 75,21% dari total kasus perceraian. Sementara itu, cerai talak, yakni perceraian yang diajukan oleh pihak suami dan disahkan melalui putusan pengadilan, menyumbang 127.986 kasus atau sekitar 24,78% dari total kasus perceraian., dengan sebagian besar kasus terjadi di Provinsi Sumatera Utara. [1][2][3],[4][5] Fenomena ini memicu perhatian para peneliti dan pembuat kebijakan untuk menganalisis penyebab perceraian sekaligus mencari solusi guna menurunkan angka tersebut.

Beragam faktor memengaruhi perceraian, seperti masalah ekonomi, kekerasan dalam rumah tangga, ketidakcocokan pasangan, hingga perselisihan yang berulang. [6][7][8]. Penelitian oleh Lubis W dan Muktarruddin M (2023) menyebutkan bahwa kondisi ekonomi yang sulit dapat memicu ketegangan dalam rumah tangga, yang akhirnya berujung pada perceraian [9]. Oleh karena itu, diperlukan analisis data mendalam untuk mengidentifikasi pola dan tren perceraian di wilayah ini.

Dalam era digital, penggunaan teknologi analitik seperti RapidMiner dapat mempermudah pengolahan dan analisis data dalam skala besar. RapidMiner adalah platform analisis yang memungkinkan pengguna melakukan data mining, machine learning, dan analisis statistik tanpa memerlukan pengkodean yang rumit. Salah satu metode yang relevan adalah K-Means Clustering, yang dapat digunakan untuk mengelompokkan data perceraian berdasarkan kabupaten dan faktor penyebabnya. [10][11][12] Metode ini membantu peneliti mengenali kelompok populasi yang berisiko tinggi terhadap perceraian.

Penelitian ini bertujuan menerapkan metode K-Means Clustering menggunakan RapidMiner untuk menganalisis data perceraian di Sumatera Utara. Dengan pendekatan ini, diharapkan dapat ditemukan pola dan faktor dominan yang menjadi

pemicu perceraian. Hasil penelitian diharapkan memberikan wawasan baru kepada pemerintah dan lembaga terkait untuk merancang kebijakan yang lebih efektif dalam mengurangi angka perceraian. Selain itu, penelitian ini juga dapat menjadi referensi bagi studi-studi mendatang di bidang sosial dan ekonomi.

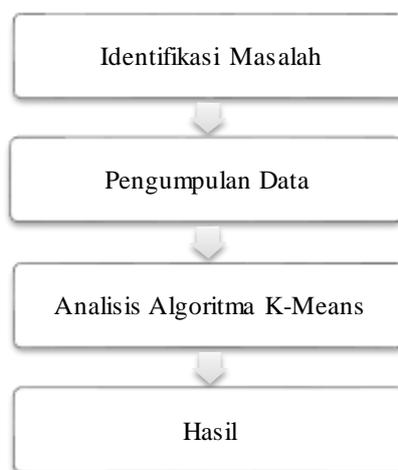
Melalui penelitian ini, diharapkan dapat tercipta pemahaman yang lebih baik mengenai perceraian di Sumatera Utara serta solusi berbasis data yang lebih tepat sasaran. Pendekatan ini diharapkan tidak hanya mengurangi angka perceraian, tetapi juga memberikan dampak positif bagi masyarakat.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining K-Means Clustering untuk menganalisis data perceraian berdasarkan kabupaten/kota di Provinsi Sumatera Utara. Adapun langkah-langkah yang akan dilakukan dalam penelitian ini dijelaskan sebagai berikut:

Kerangka kerja yang akan digunakan dalam penelitian ini dapat dilihat pada Gambar 1



**Gambar 1.** Tahapan Penelitian

### 2.2 Identifikasi Masalah

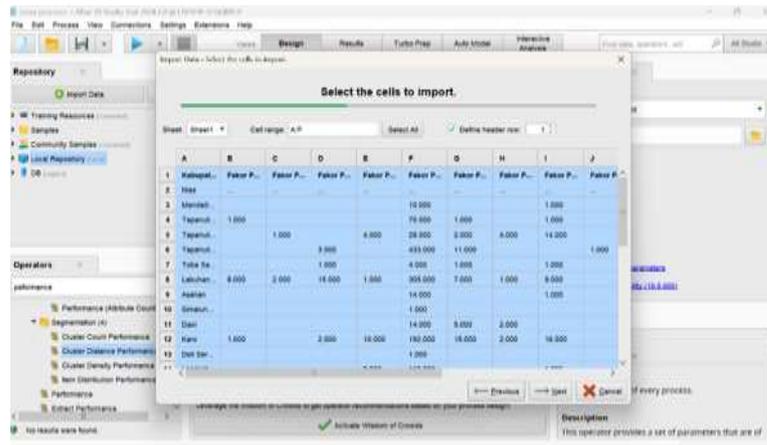
Tingginya angka perceraian ini tidak hanya mencerminkan masalah dalam hubungan antar pasangan tetapi juga mencerminkan kondisi sosial ekonomi di wilayah tersebut. Sayangnya, belum banyak penelitian yang mengelompokkan wilayah berdasarkan karakteristik perceraian. Oleh karena itu, diperlukan analisis mendalam menggunakan metode yang mampu mengidentifikasi pola tersembunyi dalam data perceraian untuk memberikan solusi yang tepat sasaran.

### 2.3 Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari instansi resmi, seperti Badan Pusat Statistik (BPS) dan Pengadilan Agama di Provinsi Sumatera Utara. Data tersebut mencakup jumlah kasus perceraian dan faktor-faktor penyebabnya di setiap kabupaten/kota pada tahun tertentu, misalnya tahun 2022. Data yang dikumpulkan mencakup:

- a. Total kasus perceraian per kabupaten/kota.
- b. Faktor penyebab perceraian, seperti perselisihan, faktor ekonomi, meninggalkan salah satu pihak, dan lainnya

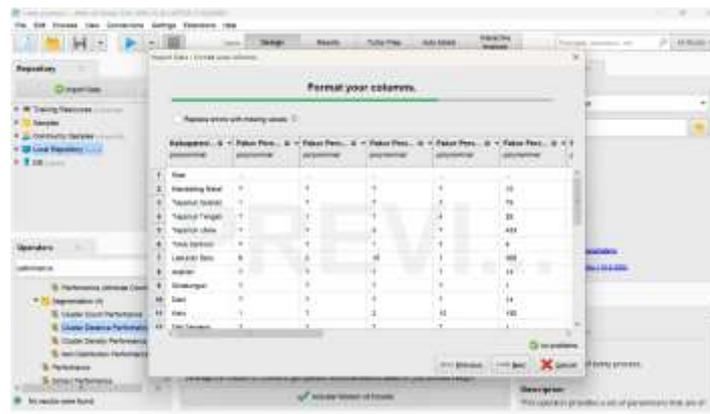




Gambar 3. Selection

### 3.2 Tahap *Preprocessing* data

Tahapan berikutnya adalah *preprocessing* data, yang mencakup beberapa langkah penting, seperti menghapus data duplikat, memperbaiki kesalahan pada data, dan langkah-langkah lain yang relevan. Tujuan utama dari tahap ini adalah untuk menghasilkan atribut data yang bersih dan berkualitas, sehingga data tersebut siap digunakan untuk proses selanjutnya, yaitu transformasi data



Gambar 4. *Preprocessing* data

### 3.3 Transformasi Data

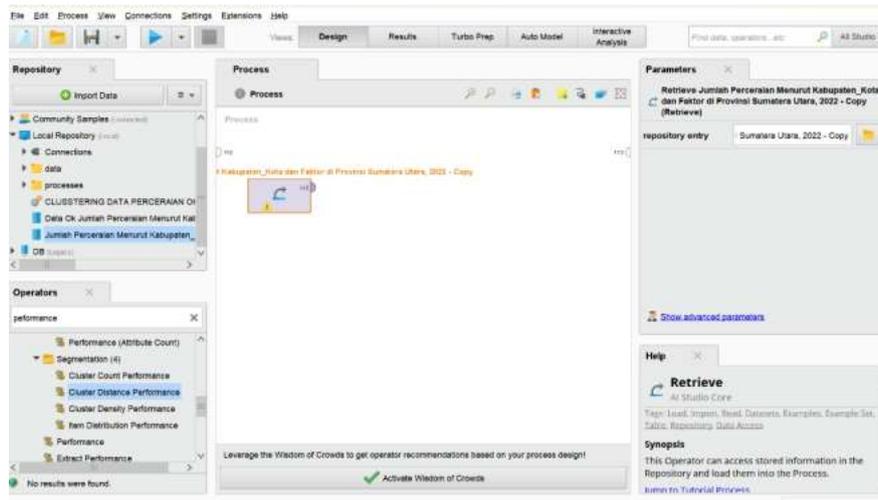
Tahap transformasi adalah proses di mana data yang telah melalui tahap sebelumnya diubah menjadi format atau bentuk yang sesuai dengan kebutuhan peneliti. Pada tahap ini, data dipersiapkan untuk dapat digunakan secara efektif dalam langkah-langkah berikutnya dalam proses penelitian



Gambar 5. Transformasi data

### 3.4 Clustering data dengan menggunakan RapidMiner Studio

Tahapan berikutnya adalah meminta operator untuk mengambil data dari lokasi yang telah diimpor dan memastikan data tersebut siap untuk diproses ke langkah selanjutnya



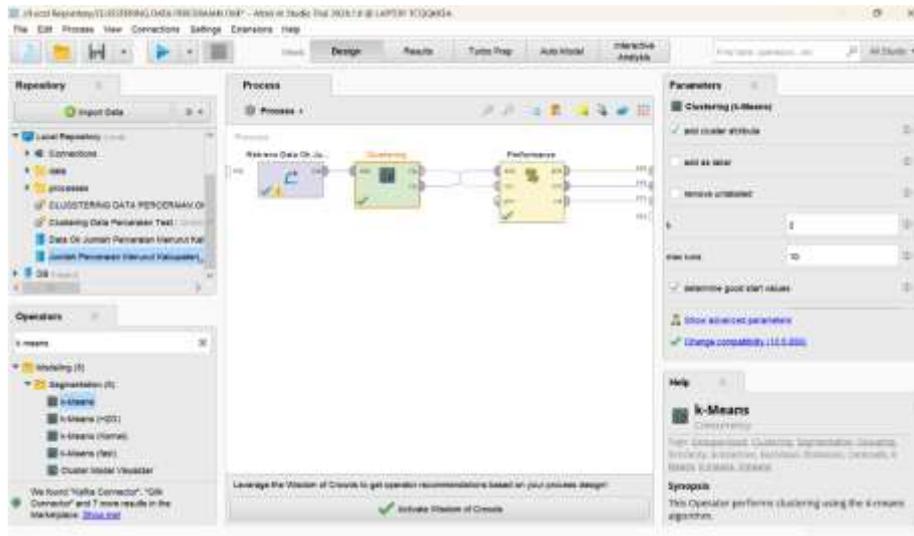
Gambar 6. Operator Retrieve Data

Tahap berikutnya adalah membangun model clustering menggunakan metode k-means, yang akan menghasilkan sejumlah kluster sesuai dengan kebutuhan analisis. Model ini bertujuan untuk mengelompokkan data berdasarkan kemiripan atau karakteristik tertentu, sehingga dapat mendukung proses interpretasi dan pengambilan keputusan selanjutnya



Gambar 7. Operator K-Means

Operator yang terakhir digunakan adalah *cluster distance performance* digunakan untuk mengoptimalkan nilai *Index Davies-Bouldin (DBI)*. *Davies Bouldin Index (DBI)* adalah metode yang digunakan untuk mengukur validitas kluster dalam analisis data menggunakan metode pengelompokan apa pun. Operator ini bertujuan untuk menurunkan nilai DBI, karena DBI yang lebih kecil menunjukkan klusterisasi yang lebih ideal. Dengan demikian, proses ini membantu menentukan jumlah kluster yang paling optimal berdasarkan ukuran dan distribusi data. [20]

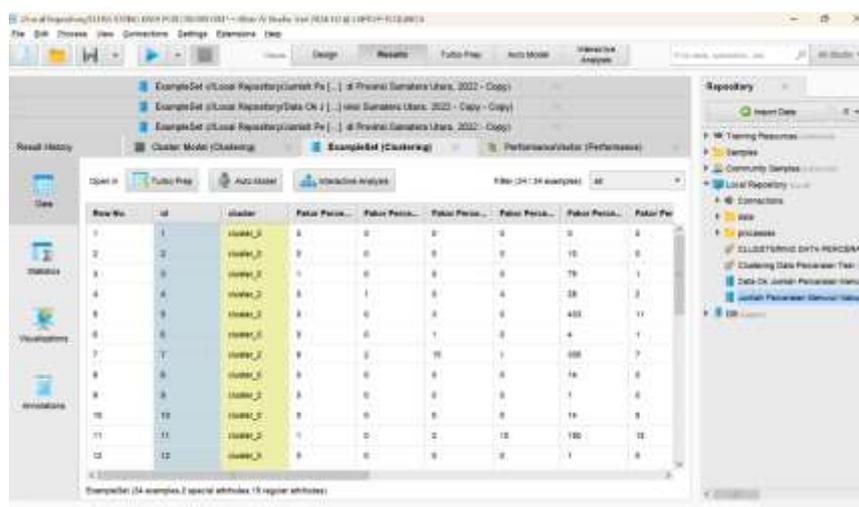


**Gambar 8.** Operator *K-Means*

Setelah menerapkan teknik data mining dan mendapatkan hasil clustering, langkah berikutnya adalah tahap evaluasi. Tahap ini bertujuan untuk menilai sejauh mana model yang telah dibangun berhasil memenuhi tujuan dan kriteria yang ditetapkan. Evaluasi dilakukan dengan memeriksa performa model serta kualitas hasil clustering yang telah dihasilkan, guna memastikan model tersebut efektif dan sesuai dengan kebutuhan analisis. Dalam model clustering yang menggunakan metode *K-means*, jumlah kluster yang diuji, 3,4, dan 5 kluster. Nilai DBI untuk setiap kluster dianalisis untuk menentukan hasil terbaik, dengan kluster 3 dipilih sebagai hasil optimal karena memiliki nilai DBI sebesar **-0,382**, yang menunjukkan validitas kluster yang baik.

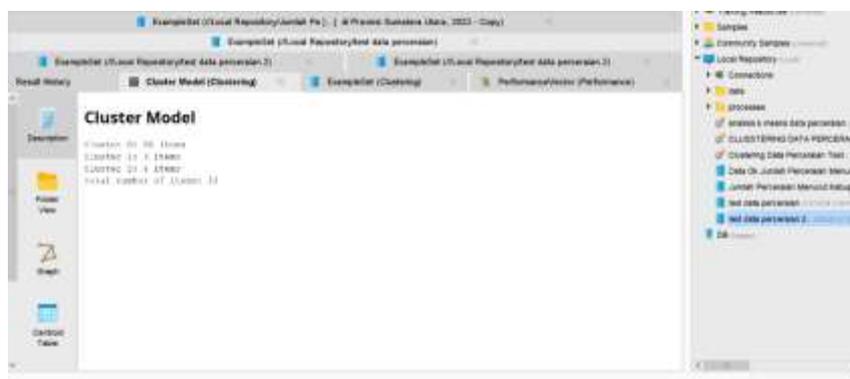
Tabel 1. Nilai DBI dalam setiap cluster

Cluster	Nilai DBI
3	-0,382
4	-0,530
5	-0,482



**Gambar 9.** Operator *Performance*

Selanjutnya, tahap ini menjelaskan hasil dari Operator Cluster Model, yang menghasilkan pembagian ke dalam 3 kluster. Dari total 34 dataset yang dianalisis, masing-masing kluster memiliki jumlah anggota yang berbeda: Cluster 0 terdiri dari 26 anggota, Cluster 1 berisi 3 anggota, dan Cluster 2 memiliki 4 anggota. Hasil ini memberikan gambaran mengenai distribusi data pada setiap kluster berdasarkan karakteristik yang telah dianalisis sebelumnya.



**Gambar 10.** Cluster Model

#### 4. KESIMPULAN

Hasil analisis menunjukkan bahwa data perceraian di Sumatera Utara dapat dikelompokkan menjadi tiga kluster yang berbeda. Kluster 0, yang terdiri dari 26 wilayah, menunjukkan bahwa banyak daerah menghadapi tingkat perceraian yang tinggi dengan berbagai faktor penyebab, termasuk ekonomi dan kekerasan dalam rumah tangga. Hal ini menandakan perlunya perhatian lebih dari instansi terkait untuk menerapkan program intervensi yang efektif dan dukungan bagi pasangan di wilayah ini. Kluster 1, yang mencakup tiga wilayah (Tapanuli Tengah, Labuhan Batu, dan Karo), menunjukkan kondisi perceraian yang lebih moderat, sedangkan Kluster 2, yang terdiri dari empat wilayah (Tapanuli Utara, Simalungun, Langkat, dan Tanjung Balai), mungkin memiliki faktor-faktor perceraian yang berbeda dan memerlukan pendekatan yang spesifik. Oleh karena itu, hasil penelitian ini dapat menjadi dasar bagi instansi terkait untuk merumuskan kebijakan yang lebih tepat sasaran, termasuk peningkatan program edukasi, konseling, dan penelitian lebih lanjut untuk memahami faktor-faktor yang mempengaruhi perceraian di setiap kluster. Implementasi kebijakan yang responsif terhadap karakteristik masing-masing kluster diharapkan dapat meningkatkan kualitas hidup masyarakat dan mengurangi angka perceraian di Sumatera Utara.

#### DAFTAR PUSTAKA

- [1] N. Januari, "MENGKALI AKAR MASALAH: Analisis Kasus Perceraian di Indonesia," *AKADEMIK: Jurnal Mahasiswa Humanis*, vol. 3, no. 3, pp. 120–130, 2023, doi: 10.37481/jmh.v3i3.613.
- [2] I. Z. Fahurrochman, M. Fadhilah, S. F. Al-Jabar, and Y. Herdiyanto, "Pengaruh Pernikahan Dini Dan Perceraian Dalam Perspektif Hukum Dan Psikologi Di Desa Ciluncat," *Proceedings UIN Sunan Gunung Djati Bandung*, vol. 3, no. 3, pp. 316–331, 2021, [Online]. Available: <https://proceedings.uinsgd.ac.id/index.php/Proceedings>
- [3] D. Siregar *et al.*, "Studi hukum tentang tingkat perceraian dan efeknya terhadap anak," *Jurnal Derma Pengabdian Dosen Perguruan Tinggi (Jurnal DEPUTI)*, vol. 3, no. 2, pp. 178–185, 2023, doi: 10.54123/deputi.v3i2.276.
- [4] Asiva Noor Rachmayani, "Gambaran kasus korban kekerasan dalam rumah tangga di RSUD Dr. Pringadi Medan," p. 6, 2015.
- [5] R. H. Harahap, "Relevansi Bimbingan Perkawinan Pranikah Dengan Tingginya Angka Perceraian: Studi Analisis Terhadap Pelaksanaan Bimbingan Perkawinan Pranikah Di KUA Kota Medan," *Mizan: Journal of Islamic Law*, vol. 5, no. 3, p. 393, 2021, doi: 10.32507/mizan.v5i3.1098.
- [6] E. Fidyawati *et al.*, "Disharmoni Keluarga Pekerja Migran Indonesia ( PMI ) Tulungagung Dalam Perspektif Teori Struktural Fungsional Emile Durkheim PENDAHULUAN Pekerja Migran Indonesia ( PMI ) merupakan istilah pada warga negara Indonesia yang bekerja di luar negeri . 1 Para pe," vol. 8, no. 1, pp. 1140–1155, 2024.
- [7] J. Ilmu and P. Dan, "Pelatihan Pra-Nikah Untuk Menangani Pernikahan Usia Dini Di Desa Cijagang Cianjur," vol. 06, no. 4, pp. 115–127, 2024.
- [8] D. Harga, D. A. N. Kualitas, and P. Literature, "Vol. 13 No. 2 Desember 2022," vol. 13, no. 2, pp. 75–86, 2022.
- [9] W. G. Lubis and M. Muktaruddin, "Peran konseling pranikah dalam menurunkan angka perceraian di kota Tanjung Balai," *Jurnal EDUCATIO: Jurnal Pendidikan Indonesia*, vol. 9, no. 2, p. 995, 2023, doi: 10.29210/1202323413.
- [10] Normah, B. Rifai, S. Vambudi, and R. Maulana, "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE," *Jurnal Teknik Komputer AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.
- [11] W. Sudrajat, I. Cholid, and J. Petrus, "Penerapan Algoritma K-Means Clustering untuk Pengelompokan UMKM Menggunakan Rapidminer," *Jurnal JUPITER*, vol. 14, no. 1, pp. 27–36, 2022.

- [12] D. A. Azhari, Y. Maulita, and S. Ramadani, “Pengelompokan Data Kriminal untuk Menentukan Pola Rawan Tindak Kriminal Menggunakan Algoritma K-Means ( Studi Kasus : Polsek Hamparan Perak ) judul penelitian yaitu “ Pengelompokan Data Kriminal Untuk Menentukan Pola Rawan Tindak Kriminal Menggunakan Alg.” vol. 2, no. 5, 2024.
- [13] J. F. Andry and H. Tannady, “Data Management Analysis for Predicting Stroke using RapidMiner,” vol. 2, no. 3, pp. 318–322, 2024.
- [14] Z. Fatah *et al.*, “IMPLEMENTASI RAPIDMINER PADA KLASTERISASI GEMPA BUMI DI INDONESIA BERDASARKAN KEDALAMAN MENGGUNAKAN K-MEANS,” vol. 1, no. 6, pp. 84–91, 2024.
- [15] S. Informasi, Y. I. Muasaroh, and Z. Fatah, “Jurnal Advance Research Informatika Implementasi RapidMiner dalam Optimasi Pembentukan Kelas Unggulan Menggunakan K-Means Clustering,” vol. 3, no. 1, pp. 66–72, 2024.
- [16] P. Rumah, S. Royal, A. Hidayah, D. Dulisep, and B. Angga, “Implementasi Algoritma K - Means Menggunakan RapidMiner untuk Klasterisasi Data Obat,” vol. 7, no. 2, pp. 200–211, 2024.
- [17] R. Adolph, “Penggunaan rapid miner untuk melihat pohon keputusan penilaian kinerja *account representative*” vol. 7, no. 2, pp. 1–23, 2016.
- [18] M. Rafi Nahjan, Nono Heryana, and Apriade Voutama, “Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 101–104, 2023, doi: 10.36040/jati.v7i1.6094.
- [19] K. Handoko, “Penerapan Data Mining Dalam Meningkatkan Mutu Pembelajaran Pada Instansi Perguruan Tinggi Menggunakan Metode K-Means Clustering (Studi Kasus Di Program Studi Tkj Akademi Komunitas Solok Selatan),” *Jurnal Teknologi dan Sistem Informasi*, vol. 02, no. 03, pp. 31–40, 2016, [Online]. Available: <http://teknosi.fti.unand.id/index.php/teknosi/article/view/70>
- [20] A. Saputra and R. Yusuf, “Perbandingan Algoritma DBSCAN dan K-MEANS dalam Segmentasi Pelanggan Pengguna Transportasi Publik Transjakarta Menggunakan Metode RFM,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 4, pp. 1346–1361, 2024, doi: 10.57152/malcom.v4i4.1516.