



# PENGELOMPOKAN DATA PBB MASYARAKAT BERDASARKAN JUMLAH PBB TERHUTANG PADA KANTOR DESA PADANG BRAHRANG MENGGUNAKAN METODE CLUSTERING

<sup>1</sup>Jenni Mardini

<sup>1</sup>Sistem Informasi, Sekolah Tinggi Ilmu Komputer KAPUTAMA  
Jl. Veteran No. 4A-9A, Binjai, Sumatera Utara, Indonesia

e-mail : jennimardini@gmail.com

Received: 2023-07-22

Revised: 2023-08-25

Accepted: 2023-09-28

Page : 54-61

**Abstrak** : Pajak Bumi dan Bangunan (PBB) perdesaan dan perkotaan adalah pajak atas bumi dan/bangunan yang dimiliki, dikuasai, dan/dimanfaatkan oleh orang pribadi atau badan, kecuali Kawasan yang digunakan untuk usaha perkebunan, perhutanan, dan pertambangan. Penelitian ini dapat membantu Pemerintah Desa Padang Brahrang dalam meningkatkan penerimaan Pendapatan Bagi Hasil (PBH) Pajak didesa Padang Brahrang. Penulisan proposal ini menggunakan metode *clustering* yang merupakan salah satu teknik data *Mining* untuk pengelompokan data PBB Masyarakat Desa Padang Brahrang Dengan menggunakan metode *clustering* algoritma *k-means*. Berdasarkan hasil proses *Clustering* menggunakan *MATLAB*, maka dapat dilihat pada keterangan : Dapat diketahui bahwasannya jumlah data terbanyak terdapat pada *cluster* 1 dengan jumlah data pada data pbb masyarakat yaitu : wilayah objek pajak (dusun) adalah dusun mawar dengan luas tanah anantara  $\leq 100M-2.000M$  dengan kisaran jumlah pajak terhutang yaitu  $\leq Rp. 10.000-Rp. 2.000.000$ .

**Kata kunci:** Algoritma\_K-Means, Clustering, Data\_Mining, Data PBB Masyarakat, Matlab, Gui

**Abstract** : Land and Building Tax (PBB) in rural and urban areas is a tax on land and/or buildings that are owned, controlled and/used by individuals or entities, except for areas used for plantation, forestry and mining businesses. This research can assist the Padang Brahrang Village Government in increasing tax revenue sharing (PBH) revenue in Padang Brahrang village. Writing this proposal uses the clustering method which is a data mining technique for grouping PBB data for the Padang Brahrang Village Community using the *k-means* algorithm clustering method.



*Based on the results of the Clustering process using MATLAB, it can be seen in the description: It can be seen that the largest amount of data is in cluster 1 with the amount of data in the UN community data, namely: the tax object area (hamlet) is Dusun Mawar with a land area between  $\leq 100M-2,000M$  with a range of tax payable, namely  $\leq Rp. 10,000-Rp. 2,000,000$ .*

**Keywords:** *K-Means\_Algorithm, Clustering, Data\_Mining, Community PBB Data, Matlab, Gui*



Journal of Matematics and Technology (MATECH) This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

## 1 Pendahuluan (or Introduction)

Pengelompokkan merupakan salah satu alternatif yang dapat digunakan untuk menggali suatu informasi agar menjadi pengetahuan baru bagi setiap orang yang ingin memperolehnya. dan dalam kasus ini informasi yang ingin digali ialah tentang data pajak bumi dan bangunan masyarakat yang ada di Desa Padang Brahrang Kecamatan Selesai. Penggalan ini dibuat untuk memperoleh informasi tentang data masyarakat wajib pajak terhutang yang ada di Desa Padang Brahrang berdasarkan wilayah dusunnya dan jumlah besaran nilai pajak terhutangnya agar memudahkan pemerintah Desa dalam mengetahui wilayah mana yang paling banyak jumlah target pajaknya dan wilayah dusun mana yang paling banyak masyarakat tidak memenuhi wajib pajaknya serta dapat membantu pihak Pemerintah Desa Padang Brahrang dalam meningkatkan penerimaan Pendapatan Bagi Hasil (PBH) Pajak di Desa Padang Brahrang.

Pada Kantor Kepala Desa Padang Brahrang data PBB Masyarakatnya masih tergabung dalam satu buku besar dengan isi nama-nama wajib pajaknya, terkadang hanya dengan membagikan SPPT pajaknya saja tanpa di catat dan data piutang pajak hanya akan didapat dari pihak kecamatan apabila tanggal waktu target pajak telah selesai sehingga prosesnya kurang efisien. Dengan pemanfaatan data mining menggunakan data masukan berupa data piutang wajib pajak masyarakat pada Desa Padang Brahrang, data tersebut akan diolah dengan metode clustering menggunakan algoritma k-means dengan menggunakan variabel wilayah objek pajak (dusun), variabel luas tanah dan variabel jumlah pajak terhutang (piutang) pada data daftar himpunan ketetapan Pajak Bumi dan Bangunan (DHKP) dan Daftar Piutang Pajak dari kecamatan yang ada di Desa Padang Brahrang. Berdasarkan hasil pengelompokan tersebut, pengolahan data yang dilakukan dapat menghasilkan informasi yang berguna terhadap pihak Desa Padang Brahrang terutama dalam hal meningkatkan Pembagian pendapatan bagi hasil (PBH) pajak di Desa Padang Brahrang.

## 2 Tinjauan Literatur (or Literature Review)

Dalam penelitian ini, penulis sedikit banyak terinspirasi dan mereferensi dari penelitian-penelitian sebelumnya yang berkaitan dengan latar belakang masalah pada penelitian ini yaitu:



Menurut penelitian (Medina Aprilia Putri et al.2022) dengan judul “Penerapan Data Mining Menggunakan Metode Clustering Untuk Mengetahui Kelompok Kepatuhan Wajib Pajak Bumi dan Bangunan” menyimpulkan bahwa *Data Mining* dapat menghasilkan informasi yang bermanfaat, dimana dapat dibuktikan bahwa hasil dari penelitian yang dilakukan dengan metode *clustering* dapat membantu mengelompokkan data untuk tingkat kepatuhan PBB masyarakat dan dapat mengetahui data wajib pajak yang paling banyak menunggak[1]

Dan menurut (Noviana et al.2019) dengan judul “ Penerapan *Data Mining* Menentukan Kelompok Prioritas Penerima Bantuan Beras Rastra Dengan Clustering *K-Means*”. Menyatakan bahwa *Data Mining* dapat membantu untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dalam pengolahan data penerima bantuan beras sejahtera (rastra) dan dapat membantu dalam menentukan penduduk yang menjadi prioritas utama untuk mendapatkan bantuan di tengah banyaknya data penduduk di kecamatan muara papalik. *Software* yang digunakan yaitu *weka*. [2]

Menurut Mardi (Amna Wahyudin, S et al.2023 hal:9) *Data Mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan yaitu:

1. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *Data Mining* disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. *Pre-procesing/Cleaning*

Sebelum proses *Data Mining* dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi focus KDD. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang konsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (*tipografi*). Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

3. *Transformation*

*Coding* adalah proses tranformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *Data Mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data Mining*

*Data Mining* yaitu proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. *Interpretation Evaluation*

Pola informasi yang dihasilkan dari proses *Data Mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari Proses KDD yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.[3]

Menurut (Dasril Aldo, S.Kom et al.2021 hal:64) *Data Mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan yaitu:

1. Deskripsi

Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit di dukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.



2. Estimasi  
Estimasi hampir sama dengan klasifikasi, kecuali variable target lebih kearah *numeric* dari pada kearah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variable target sebagai nilai prediksi.
3. Prediksi  
Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan datang di masa mendatang.
4. Klasifikasi  
Dalam klasifikasi, terdapat target variable kategori. Sebagai contoh penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.
5. Pengklusteran  
Pengklusteran merupakan peneglompokan *record*, pengamatan ataupun memperhatikan dan membentuk kelas objek-objek yang dimiliki kemiripan. *Cluster* adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam *cluster* lain.
6. Asosiasi

Tugas asosiasi dalam *Data Mining* adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.[4]

Menurut ( Cahyo Prianto et al. 2020 hal:14) *clustering* adalah “suatu metode mempartisi data-set menjadi beberapa sub-set atau kelompok sedemikian rupa sehingga elemen-elemen dari suatu kelompok tertentu memiliki *set property* yang *disshare* bersama, dengan tingkat similaritas yang tinggi dalam suatu kelompok dan tingkat similaritas antar kelompok yang rendah”. Sehingga *cluster* tidak harus sama akan tetapi pengelompokannya berdasarkan pada kedekatan dari suatu karakteristik sampel yang ada salah satunya dengan menggunakan rumus jarak *eclidean*. Tujuan utama dari metode clustering adalah mengelompokkan sejumlah data/obyek kedalam cluster (*group*) sehingga dalam setiap cluster akan berisi data yang semirip mungkin [5].

Menurut James et al (Amna Wahyudin S et al.2023 hal:66) menyatakan bahwa “ *K-Means* merupakan salah satu metode pengelompokan data non hierarki (sekatan) yang berusaha mempartisi data ke dalam *cluster* atau kelompok sehingga data yang memiliki karakteristik yang sama akan dikelompokkan ke dalam satu *cluster* yang sama dan data yang memiliki karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. [6]

Menurut (Indah Werdiningsih.2022 hal:55) menyatakan bahwa *Matlab* merupakan singkatan dari Matrix Laboratory. *Matlab* merupakan sebuah platfor pemogramana dan komputasi numeric yang digunakan untuk menganalisis data, mengembangkan algoritma dan membuat model. Leh karena itu *matlab* banyak digunakan beberapa engineer dan ilmuwan di berbagai bidang riset yang memerlukan komputasi numeric yang kompleks [7].

Menurut (I Made Andik Setiawan.2022 hal:1) *Graphical User Interface (GUI)* *Matlab* adalah faslitas yang disediakan oleh *matlab* untuk programing berbasis object [8] .

Menurut Hughes dan Kapoor (Martha Rianty., S.E., MSi.2021 hal:12) usaha atau bisnis adalah

Menurut (S. Purwaningsih.2020 hal:1) Masyarakat merupakan sekumpulan manusia yang hidup secara berdampingan dengan segala kebudayaan dan kepribadiannya [9].

### 3 Metode Penelitian (or Research Method)

Metodologi penelitian ini dilakukan untuk mencari sesuatu secara sistematis dengan menggunakan metode ilmiah serta sumber yang berlaku. Dengan adanya proses ini, dapat memberikan hasil penelitian yang baik dan tepat. Atas dasar metodologi penelitian yang digunakan, maka dapat dibuat suatu alur kegiatan seperti berikut :



1. Persiapan

Tahap ini adalah kegiatan awal dari penelitian, yaitu dengan penentuan dari latar belakang masalah kemudian dilakukan batasan masalah dan selanjutnya dilakukan penentuan tujuan serta manfaat yang dilakukan dalam penyusunan proses kerja sistem. Setelah itu, penulis menentukan kebutuhan data penelitian diantaranya menentukan metode yang akan digunakan, kriteria dari data usaha masyarakat. Setelah itu penulis membuat surat izin penelitian yang ditujukan kepada Kantor Kepala Desa Tanjung Merahe untuk meminta data dan mengumpulkan data yang dibutuhkan dalam penelitian ini.

2. Kajian Teori

Dalam tahap ini penulis mengumpulkan berbagai teori baik dari buku yang dipinjam dari perpustakaan, jurnal maupun internet untuk mendukung penelitian yang akan dilakukan. Teori yang dikumpulkan antara lain mengenai, Data Mining, Metode *Clustering*, algoritma *K-Means*, Usaha Masyarakat, *Matlab*, *Gui*, *Flowchart* dan teori yang mendukung lainnya.

3. Pengumpulan Data

Tahap ini merupakan pengumpulan data yang diperlukan dalam pembuatan skripsi, yaitu Data Usaha Masyarakat yang diperoleh Dari Kantor Kepala Desa Tanjung Merahe.

4. Analisa Data

Tahapan ini berupa merupakan tahapan mengelola dan menganalisa data yang telah diperoleh yaitu Data Usaha Masyarakat kemudian data tersebut dilakukan transformasi untuk dapat dilakukan analisa yang kemudian diproses dengan menggunakan metode *Clustering* untuk mendapatkan sebuah informasi yang baru atau dikelompokan sesuai dengan variabel yang telah ditentukan.

5. Pengujian dan implementasi

Tahap ini merupakan tahapan yang melakukan pengujian validasi dan implementasi data yang telah dianalisa sebelumnya serta penyusunan program. Pada tahap ini juga dilakukan pengujian dengan memproses data yang telah diproses dengan menggunakan *software matlab* sebagai implementasi proses pengelompokan data usaha masyarakat berdasarkan jenis usahanya.

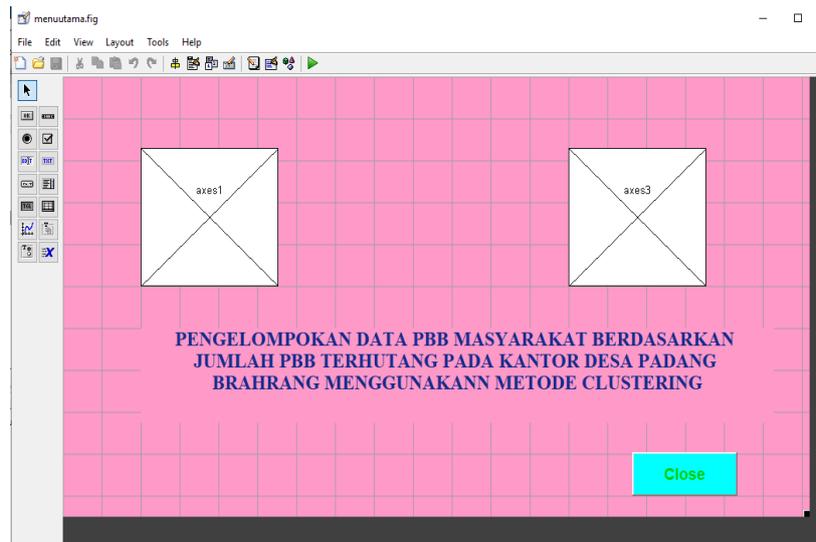
6. Tahap Akhir

Tahap ini merupakan tahap pengambilan kesimpulan dan saran yang dapat dilakukan dalam penyusunan penelitian ini. Dengan adanya kesimpulan maka akan diketahui hasil dari keseluruhan penelitian dan diharapkan dengan saran akan ada perbaikan dan manfaat bagi yang lain.

## 4 Hasil dan Pembahasan (or Results and Analysis)

### Tampilan Halaman Utama

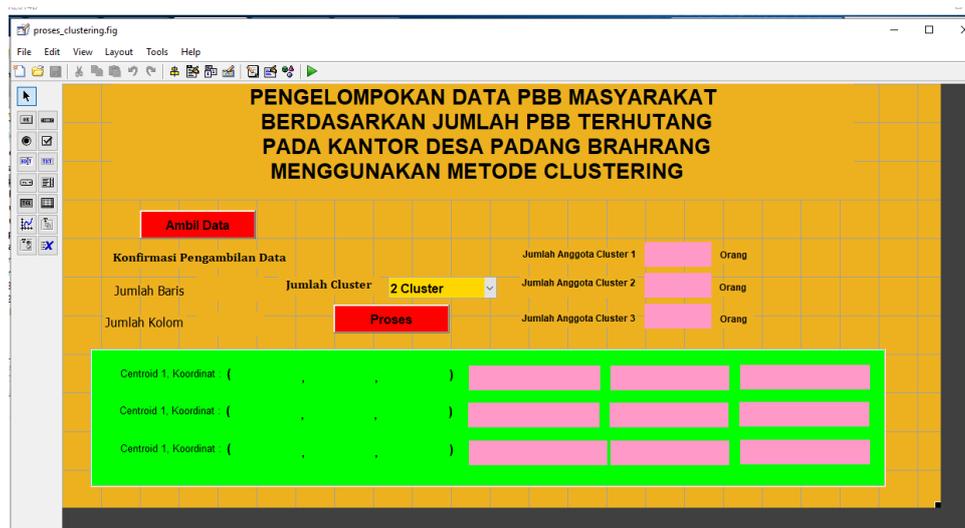
Halaman ini tampil ketika pertama kali membuka aplikasi *Matlab* untuk data mining yang akan dirancang , pada halaman ini akan terlihat semua *link-link* yang digunakan dalam aplikasi ini. Di sini hanya menampilkan gambar, tulisan, logo, judul aplikasi dan *buttonclose*. Dengan desain *form* sebagai berikut:



Gambar 1 Tampilan Menu Utama

### Tampilan Proses Clustreing

Pada halaman ini akan terlihat keseluruhan proses *data mining* sampai pada pemunculan grafik dan keterangan *centroid* sebagai hasil dari perhitungan dengan metode *clustering* menggunakan algoritma *k-means*. Dengan desain *inform* sebagai berikut :



Gambar 2 Tampilan Menu Proses Clustering

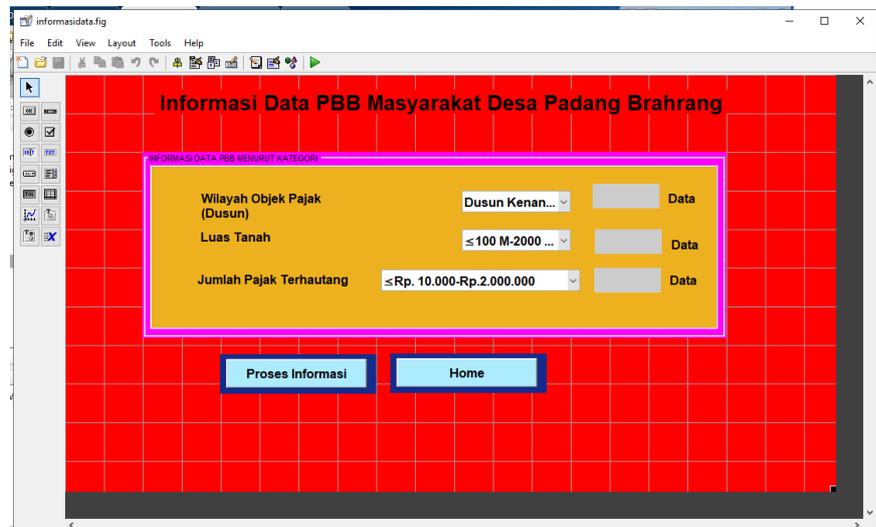
Keterangan Gambar :

1. *Button* Ambil Data berfungsi untuk mengambil data/import file dari *microsoft excel* yang akan dilakukan proses *clustering*.
2. Jumlah baris dan jumlah kolom untuk menampilkan jumlah baris dan kolom sesuai dengan data yang ada pada *microsoft excel* yang diimport untuk selanjutnya diproses *clustering*.

3. Jumlah *cluster* untuk menentukan jumlah pusat *cluster* yang akan diproses *clustering* dengan memilih angka *cluster* dimenu *pop up* yang disediakan.
4. *Button* proses *cluster* digunakan untuk memproses perhitungan *clustering* dengan *algoritama k-means* dalam proses *data mining* tersebut, kemudian akan muncul hasil *clustering* pada masing-masing *centroid* dan juga muncul grafik hasil *clustering* berikut keterangan dari *centroid* tersebut.
5. *Button* Menu utama untuk kembali kemenu utama.

### Halaman Informasi Data Usaha Masyarakat

Pada halaman Informasi Data Usaha Masyarakat akan muncul data Usaha Masyarakat menurut kategori yang tersimpan di Microsoft *Excel* yang telah terkoneksi dengan *Matlab*, di mana data tersebut mengalami proses seleksi berdasarkan Umur Usaha, Jenis Bidang Usaha dan Pendapatan Rata-rata Usaha Per-bulanya.tampilanya seperti dibawah ini :



Gambar 3 Tampilan Menu Informasi Usaha Masyarakat

## 5 Kesimpulan (or Conclusion)

Setelah melakukan peneitian ini, dapat disimpulkan bahwa :

1. Dengan menggunakan metode *clustering* maka akan mendapatkan hasil berupa informasi baru dengan menggunakan 561 data yang telah diolah dan juga diuji oleh sistem, sehingga dapat diketahui wilayah objek pajak (dusun) adalah dusun mawar dengan luas tanah antara  $\leq 100M-2.000M$  dengan kisaran jumlah pajak terhutang yaitu  $\leq Rp. 10.000-Rp. 2.000.000$
2. Berdasarkan hasil proses *Clustering* menggunakan *MATLAB*, maka dapat dilihat pada keterangan : Dapat diketahui bahwasannya jumlah data terbanyak terdapat pada *cluster* 1 dengan jumlah data pada data pbb masyarakat yaitu : wilayah objek pajak (dusun) adalah dusun mawar dengan luas tanah antara  $\leq 100M-2.000M$  dengan kisaran jumlah pajak terhutang yaitu  $\leq Rp. 10.000-Rp. 2.000.000$ .



3. Dengan dilakukan implementasi menggunakan *matlab* maka telah mendapatkan hasil kelompok data pbb masyarakat yang tidak patuh membayar pajak sesuai dengan variabel yang telah ditentukan.

### Referensi (Reference)

- [1] M. A. Putri, N. Rahaningsih, F. M. Basysyar, and O. Nurdiawan, "Penerapan Data Mining Menggunakan Metode Clustering Untuk Mengetahui Kelompok Kepatuhan Wajib Pajak Bumi dan Bangunan," *J. Account. Inf. Syst.*, vol. 5, no. 2, pp. 145–156, 2022, [Online]. Available: <https://jurnal.masoemiversity.ac.id/index.php/aims/article/view/496>
- [2] T. Noviana, J. Jasmir, and Y. Novianto, "Penerapan Data Mining Menentukan Kelompok Prioritas Penerima Bantuan Beras Rastra Dengan Clustering K-Means," *Progr. Stud. Tek. Inform. Stikom Din. Bangsa*, pp. 159–174, 2019.
- [3] L. W. S. Amna Wahyuddin.s, I Gede Iwan Sudipa, Tri Andi EPutra, Ahmad Jumaidi Wahyuddin, Wara Alfa Syukrila, Anindya Krisna Wardani, Nono Keryana, Tutuk Indriyani, *DATA MINING*. Padang, Sumatera Barat: PT. Global Eksekutif Teknologi, 2023.
- [4] S. K. M. K. Dasril Aldo, S. P. M. K. Afriosa Syawitri, S. K. M. K. Alwendi, S. K. M. K. Darmansah, and S. K. M. K. Khairunnisa Samosir, *DATA MINING*. Insan Cendekia Mandiri, 2021. [Online]. Available: <https://books.google.co.id/books?id=zWgtEAAAQBAJ>
- [5] S. B. Cahyo Prianto, *PEMBUATAN APLIKASI CLUSTERING GANGGUAN JARINGAN MENGGUNAKAN METODE KMEANS CLUSTERING*. Bandung: Kreatif Industri Nusantara, 2020.
- [6] D. Aldo, "Data Mining Sales of Skin Care Products Using the K-Means Method," *Sinkron*, vol. 8, no. 1, pp. 295–304, 2023, doi: 10.33395/sinkron.v8i1.12007.
- [7] I. Werdiningsih, D. C. R. Novitasari, and D. Z. Haq, *Pengelolaan Data Mining dengan Pemrograman Matlab*. Airlangga University Press, 2022. [Online]. Available: <https://books.google.co.id/books?id=CgOdEAAAQBAJ>
- [8] M. A. Setiawan, *Graphical User Interface (GUI) Matlab dan Interface dengan Arduino*. Bangka Belitung: Politeknik Manufaktur Negeri Bangka Belitung, 2022. [Online]. Available: <https://books.google.co.id/books?id=FeR3EAAAQBAJ>
- [9] S. Purwaningsih, *Pranata Sosial dalam Kehidupan Masyarakat*. Alprin, 2020. [Online]. Available: <https://books.google.co.id/books?id=uAYBEAAAQBAJ>