



PENERAPAN ALGORITMA K-MEANS UNTUK PENGELOMPOKAN WAJIB PAJAK BUMI DAN BANGUNAN (PBB) BERMASALAH DI KELURAHAN BINJAI ESTATE

Raudho Khoirani ^{1*}, Yani Maulita ², Anton Sihombing ³

^{1,2,3} Program Studi Sistem Informasi, STMIK Kaputama

E-mail: raudhokhairani@gmail.com

ABSTRACT

The number of Land and Building Tax (PBB) taxpayers in the Binjai Estate Urban Village increases annually however, existing data has historically been used solely for routine recording and administration without further analysis. This study aims to apply the K-Means algorithm to categorize data on delinquent PBB taxpayers based on the variables of occupation, payment status, and tax assessment amount. The clustering process follows the Knowledge Discovery in Databases (KDD) stages data selection, preprocessing, transformation, data mining, and evaluation utilizing MATLAB software with an interface developed via MATLAB GUIDE. The algorithm employs Euclidean Distance measurements to determine the proximity of data points to cluster centroids. The study utilizes data from 600 PBB taxpayers covering the 2021-2025 period, processed through the interface using two cluster configurations: K=2 and K=3. Results indicate that the most dominant and consistent group across both configurations consists of taxpayers who are farmers, characterized by uncertain payment status and tax assessments ranging from IDR 51,000 to IDR 150,000. The developed interface displays the number of members per cluster, centroid values, clustering results, and 3D graphs, thereby assisting the Binjai Estate Urban Village in improving the efficiency of PBB collection.

Keywords: Clustering, K-Means, Land_and_Building_Tax, MATLAB, Problematic_Taxpayers

ABSTRAK

Wajib pajak Pajak Bumi dan Bangunan (PBB) di Kelurahan Binjai Estate jumlahnya terus bertambah setiap tahun, namun data yang ada selama ini hanya digunakan untuk pencatatan dan administrasi rutin tanpa diolah lebih lanjut. Penelitian ini bertujuan menerapkan algoritma *K-Means* untuk mengelompokkan data wajib pajak PBB bermasalah berdasarkan variabel pekerjaan, status pembayaran, dan ketetapan PBB. Proses pengelompokan dilakukan melalui tahapan *Knowledge Discovery in Database*

Article History

Received: Agustus 2026

Reviewed: Agustus 2026

Published: Agustus 2026

Plagiarism Checker No
237.u873247.fh8374

Copyright : Author
Publish by : Kohesi



This work is licensed
under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



(KDD), yaitu data *selection*, *preprocessing*, *transformation*, data *mining*, dan *evaluation*, dengan bantuan aplikasi *MATLAB* melalui *interface* yang dibangun menggunakan *MATLAB GUIDE*. Proses perhitungan dalam algoritma ini menggunakan pengukuran *Euclidean Distance* untuk menentukan tingkat kedekatan antar data terhadap titik pusat *cluster*. Data yang digunakan sebanyak 600 data wajib pajak PBB periode 2021-2025, yang diolah menggunakan *interface* yang telah diterapkan dengan dua jumlah *cluster*, yaitu $k=2$ dan $k=3$. Hasil penelitian menunjukkan kelompok yang paling dominan dan konsisten muncul pada kedua jumlah *cluster* adalah wajib pajak yang memiliki pekerjaan petani, dengan kondisi pembayaran diragukan serta ketetapan PBB berkisar Rp51.000 hingga Rp150.000. *Interface* yang dibangun mampu menampilkan jumlah anggota tiap *cluster*, nilai *centroid*, hasil pengelompokan, dan grafik 3 dimensi, sehingga dapat membantu Kelurahan Binjai Estate melakukan efisiensi penagihan PBB.

Kata Kunci: *Clustering*, *K-Means*, *MATLAB*, PBB, Wajib_Pajak_Bermasalah

1. PENDAHULUAN

Pendahuluan: Pajak Bumi dan Bangunan (PBB) merupakan salah satu sumber pendapatan daerah yang berperan strategis dalam mendukung pembiayaan pembangunan. Tingkat kepatuhan wajib pajak dalam membayar PBB menjadi faktor penentu tercapainya target penerimaan daerah [1].

Di Kelurahan Binjai Estate, data wajib pajak PBB selama ini hanya digunakan untuk keperluan pencatatan dan administrasi rutin setiap tahunnya, tanpa diolah atau dianalisis lebih lanjut. Petugas kelurahan hanya mengetahui daftar wajib pajak beserta jumlah ketetapan PBB yang bervariasi, tanpa mengetahui gambaran wajib pajak yang bermasalah. Bertambahnya jumlah wajib pajak setiap tahun membuat volume data terus membesar sehingga identifikasi wajib pajak bermasalah secara manual menjadi semakin tidak praktis.

Belum adanya mekanisme pengelompokan pada data wajib pajak PBB bermasalah membuat pihak kelurahan kesulitan menentukan prioritas penagihan yang tepat. Dengan mengelompokkan wajib pajak berdasarkan ketetapan PBB, status pembayaran, dan pekerjaan, pihak kelurahan dapat mengetahui jenis pekerjaan yang paling banyak berkontribusi terhadap tunggakan, sehingga kebijakan penagihan dapat disesuaikan dengan kemampuan finansial wajib pajak, baik melalui pemotongan langsung dari penghasilan maupun penjadwalan ulang pembayaran. Tanpa pengelompokan tersebut, potensi kehilangan Pendapatan Asli Daerah (PAD) akibat tunggakan yang tidak tertangani secara terstruktur akan semakin besar [2].

Dengan menerapkan teknik data mining menggunakan metode K-Means Clustering, data wajib pajak PBB di Kelurahan Binjai Estate dapat diolah berdasarkan ketetapan PBB, status pembayaran, dan pekerjaan untuk menghasilkan pengelompokan yang mencerminkan



karakteristik data. Algoritma K-Means dipilih karena kemudahannya dalam memproses data numerik berskala besar serta menghasilkan pengelompokan yang dapat langsung digunakan sebagai dasar pengambilan keputusan. Proses pengelompokan diimplementasikan menggunakan aplikasi MATLAB, yang dipilih karena kemampuan komputasi numerik yang kuat serta dukungan visualisasi hasil clustering secara grafis [3].

Penelitian terdahulu telah membuktikan efektivitas K-Means dalam mengelompokkan data wajib pajak. Cantika dkk. [4] menerapkan metode K-Means Clustering berbasis MATLAB pada data wajib pajak PBB untuk mengetahui tingkat kepatuhan pembayaran. Penelitian lain menerapkan K-Means untuk mengelompokkan data PBB berdasarkan kelurahan guna mengidentifikasi pola kepatuhan antarwilayah [5], serta berdasarkan tingkat kepatuhan pembayaran menggunakan pengolahan berbasis MATLAB [6]. Studi-studi tersebut menegaskan bahwa K-Means Clustering yang diimplementasikan dengan MATLAB mampu menghasilkan pengelompokan data wajib pajak yang jelas dan sistematis.

Berdasarkan uraian tersebut, penelitian ini menerapkan algoritma K-Means untuk mengelompokkan data wajib pajak PBB bermasalah di Kelurahan Binjai Estate, dengan tujuan memberikan gambaran yang lebih jelas mengenai kelompok wajib pajak bermasalah berdasarkan karakteristiknya, sehingga langkah penagihan yang diambil dapat lebih terencana dan berdampak nyata terhadap peningkatan kepatuhan pembayaran PBB.

2. METODE PENELITIAN

2.1 PBB dan Variabel Penelitian

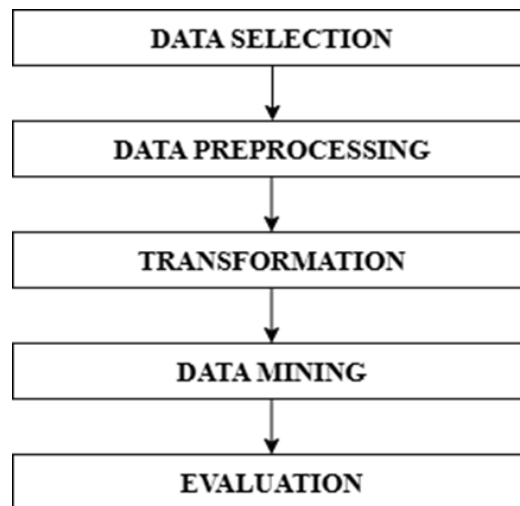
Pajak Bumi dan Bangunan (PBB) merupakan pajak yang berkaitan dengan kepemilikan, penguasaan, atau pemanfaatan tanah dan/atau bangunan. Dalam penelitian ini, data wajib pajak PBB di Kelurahan Binjai Estate digunakan sebagai objek pengelompokan untuk mengidentifikasi karakteristik wajib pajak bermasalah.

Variabel yang digunakan dalam penelitian ini terdiri atas pekerjaan, status pembayaran, dan ketetapan PBB. Pekerjaan digunakan untuk menggambarkan karakteristik wajib pajak berdasarkan jenis pekerjaan yang dimiliki. Status pembayaran menggambarkan kondisi pembayaran atau tunggakan PBB. Berdasarkan Permendagri Nomor 73 Tahun 2015, status pembayaran dapat dikelompokkan menjadi lancar sampai dengan 1 tahun, kurang lancar lebih dari 1 sampai 2 tahun, diragukan lebih dari 2 sampai 3 tahun, dan macet lebih dari 3 tahun.

Ketetapan PBB merupakan jumlah pajak yang harus dibayarkan oleh wajib pajak sebagaimana ditetapkan oleh pemerintah daerah dan tercantum dalam Surat Pemberitahuan Pajak Terutang (SPPT), dengan penetapannya berkaitan dengan Nilai Jual Objek Pajak (NJOP). Ketiga variabel tersebut kemudian digunakan sebagai dasar dalam proses pengelompokan data menggunakan algoritma K-Means.

2.2 Metode Knowledge Discovery in Database (KDD)

Penelitian ini menerapkan metode Knowledge Discovery in Database (KDD), yaitu serangkaian proses yang digunakan untuk memperoleh informasi atau pengetahuan dari sekumpulan data. Tahapan KDD yang digunakan meliputi data selection, data preprocessing, transformation, data mining, dan evaluation, sebagaimana diilustrasikan pada Gambar 1.



Gambar 1. Struktur Metodologi Penelitian (Tahapan KDD)

Berdasarkan gambar diatas dapat diketahui bahwa ada beberapa tahapan yang digunakan dalam penelitian ini sebagai berikut :

a. *Data Selection*

Tahap data selection merupakan tahap pemilihan data yang sesuai dengan tujuan penelitian. Data yang digunakan adalah data wajib pajak PBB bermasalah di Kelurahan Binjai Estate periode 2021-2025, meliputi variabel pekerjaan, status pembayaran, dan ketetapan PBB.

b. *Data Preprocessing*

Tahap data preprocessing dilakukan untuk mengubah data mentah menjadi data yang lebih bersih dan layak digunakan, meliputi pemeriksaan inkonsistensi dan kelengkapan data, sehingga dataset yang dihasilkan lebih terstruktur dan siap diolah.

c. *Transformation*

Tahap transformation mengubah atribut pekerjaan, status pembayaran, dan ketetapan PBB ke dalam bentuk kode numerik kategorikal, sehingga jarak antar data dapat dihitung.

d. *Data Mining*

Data yang telah disiapkan dianalisis menggunakan algoritma K-Means Clustering, yang mengelompokkan data ke dalam sejumlah cluster berdasarkan kemiripan karakteristiknya.

e. *Evaluation*

Tahap evaluation menilai apakah hasil pengelompokan yang diperoleh telah sesuai dengan tujuan penelitian, termasuk memeriksa apakah hasil cluster tetap stabil pada pengujian yang dilakukan berulang.

2.3 Algoritma K-Means Clustering

K-Means merupakan metode clustering non-hierarki yang membagi kumpulan data menjadi dua kelompok atau lebih berdasarkan kemiripan data. Prosedur clustering yang digunakan terdiri atas langkah-langkah berikut: (1) menentukan jumlah cluster (K) dan pusat cluster (centroid) awal; (2) menghitung jarak setiap data terhadap setiap centroid menggunakan Euclidean distance; (3) mengelompokkan setiap data ke cluster dengan centroid terdekat; (4) menghitung ulang nilai centroid berdasarkan rata-rata data pada masing-masing



cluster; dan (5) mengulang langkah beberapa kali hingga tidak terjadi lagi perubahan anggota cluster.

rumus jarak *Euclidean* :

$$D_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2} \dots \dots \dots (1)$$

Nilai centroid juga diperbarui dengan menghitung rata-rata dari seluruh data dalam kluster tersebut menggunakan rumus Euclidean:

$$\frac{x_1 + x_2 + x_3 + \dots + x_n}{\sum x}$$

2.4 Data Penelitian

Data yang digunakan terdiri atas 600 data wajib pajak PBB bermasalah di Kelurahan Binjai Estate periode 2021-2025, meliputi pekerjaan, status pembayaran, dan ketetapan PBB. Tabel 1 menyajikan contoh data.

No	NOP	Nama	Pekerjaan	Status Pembayaran	Ketetapan PBB (Rp)
1	12.76.***.***.***.***.0	D*****o	Pensiunan	4 Tahun Menunggak	80.030
2	12.76.***.***.***.***.0	J***N	Buruh	5 Tahun Menunggak	43.200
3	12.76.***.***.***.***.0	U***N Y***N	Pegawai Swasta	1 Tahun Menunggak	270.160
4	12.76.***.***.***.***.0	H***N T**E	Wiraswasta	3 Tahun Menunggak	31.244
5	12.76.***.***.***.***.0	N****L	Petani	3 Tahun Menunggak	305.864

Tabel 1. Contoh Data Penelitian (2021-2025)

3. HASIL DAN PEMBAHASAN

Hasil dan Pembahasan:

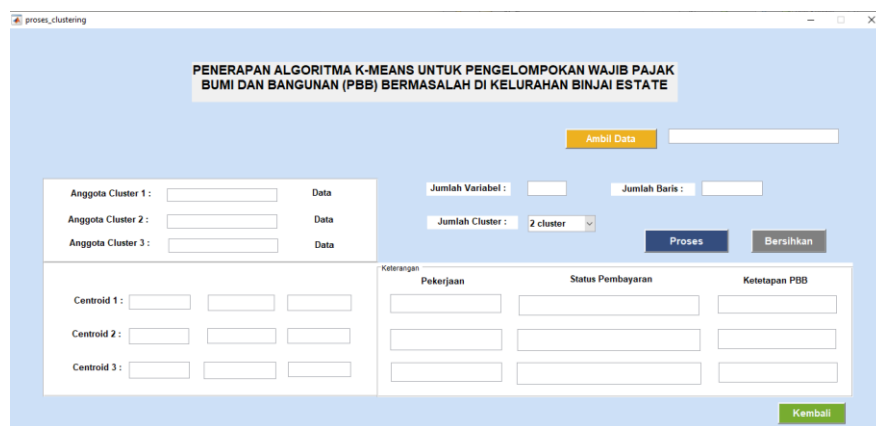
3.1 Antarmuka Aplikasi

Proses clustering diterapkan sebagai aplikasi berbasis *MATLAB GUIDE* yang terdiri atas dua halaman. Halaman pertama yaitu menampilkan menu utama:



Gambar 2. Halaman Menu Utama

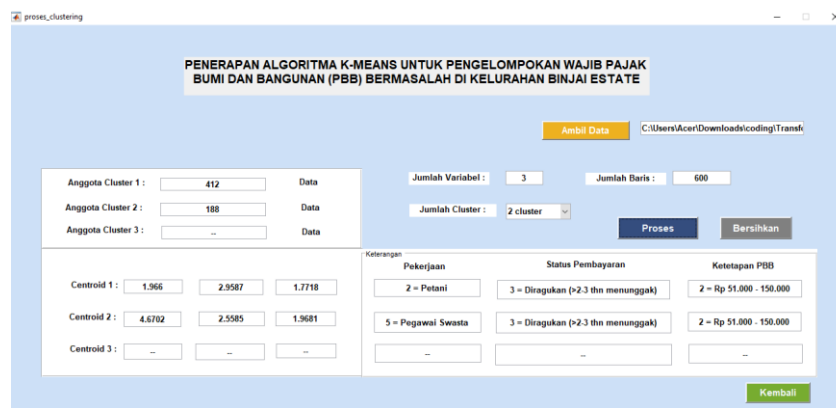
Selanjutnya halaman proses clustering memungkinkan pengguna menambahkan dataset wajib pajak dari berkas Excel yang sudah ditransformasikan Dalam bentuk numerik, menampilkan jumlah variabel dan baris data yang terdeteksi, serta menyediakan menu pop-up untuk memilih jumlah cluster (dua atau tiga cluster). Setelah tombol Proses ditekan, aplikasi menghitung K-Means Clustering dan menampilkan jumlah anggota tiap cluster, nilai centroid, hasil tiap cluster (pekerjaan, status pembayaran, dan kategori ketetapan PBB), serta grafik tiga dimensi hasil clustering. Tombol Clear mengembalikan tampilan ke kondisi awal agar antarmuka dapat digunakan kembali untuk pengujian baru.



Gambar 3. Halaman Proses Clustering

3.2 Pengujian Dua Cluster

Pengujian dua cluster diuji berulang hingga nilai centroid stabil, yang tercapai mulai pengujian keempat. Gambar 4 menunjukkan hasil stabil pada halaman proses clustering, diperoleh pada pengujian kelima.



PENERAPAN ALGORITMA K-MEANS UNTUK PENGELOMPOKAN WAJIB PAJAK BUMI DAN BANGUNAN (PBB) BERMASALAH DI KELURAHAN BINJAI ESTATE

Ambil Data C:\Users\Acer\Downloads\coding\Transf...

Anggota Cluster 1 : 412 Data
 Anggota Cluster 2 : 188 Data
 Anggota Cluster 3 : -- Data

Jumlah Variabel : 3 Jumlah Baris : 600
 Jumlah Cluster : 2 cluster

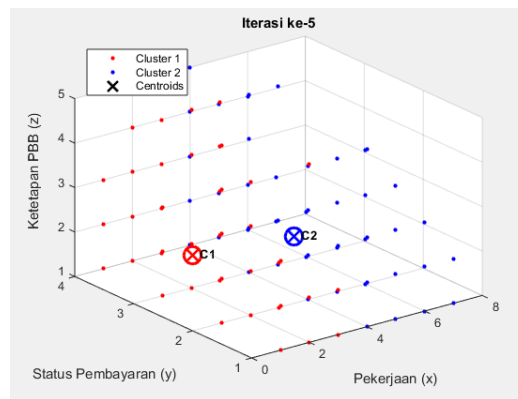
Proses Bersihkan

Centroid 1 : 1.966 2.9587 1.7718
 Centroid 2 : 4.6702 2.5585 1.9681
 Centroid 3 : -- -- --

Keterangan	Pekerjaan	Status Pembayaran	Ketetapan PBB
2 = Petani	3 = Diragukan (>2.3 thn menunggak)	2 = Rp 51.000 - 150.000	
5 = Pegawai Swasta	3 = Diragukan (>2.3 thn menunggak)	2 = Rp 51.000 - 150.000	
--	--	--	

Kembali

Gambar 4. Hasil Dua Cluster (Pengujian ke-5)



Gambar 5. Grafik Tiga Dimensi Hasil Dua Cluster

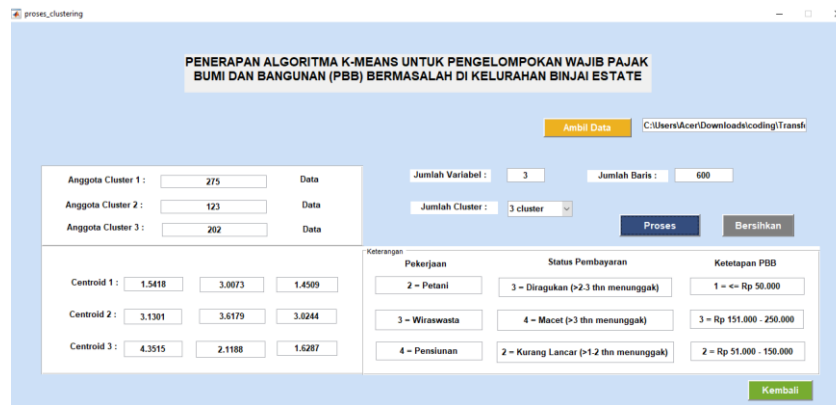
Sebagaimana ditunjukkan pada Gambar 4, Cluster 1 beranggotakan 412 data dengan centroid (1,966; 2,9587; 1,7718), didominasi wajib pajak berpekerjaan petani dengan status pembayaran diragukan (2-3 tahun menunggak) dan ketetapan PBB antara Rp51.000-Rp150.000. Cluster 2 beranggotakan 188 data dengan centroid (4,6702; 2,5585; 1,9681), didominasi wajib pajak berpekerjaan pegawai swasta dengan status pembayaran dan kisaran ketetapan yang sama. Tabel 2 merangkum hasil akhir clustering ini.

Cluster	Pekerjaan Dominan	Status Pembayaran	Ketetapan PBB (Rp)	Anggota
1	Petani	Diragukan (2-3 thn)	51.000-150.000	412
2	Pegawai Swasta	Diragukan (2-3 thn)	51.000-150.000	188

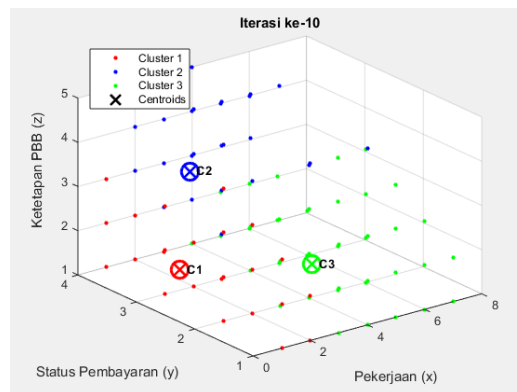
Tabel 2. Hasil Akhir Clustering - Dua Cluster

3.3 Pengujian Tiga Cluster

Pengujian tiga cluster diuji sebanyak sepuluh kali dan mencapai hasil stabil mulai pengujian kesembilan. Gambar 6 menunjukkan hasil stabil pada halaman proses clustering, diperoleh pada pengujian kesepuluh.



Gambar 6. Hasil Tiga Cluster (Pengujian ke-10)



Gambar 7. Grafik Tiga Dimensi Hasil Tiga Cluster

Sebagaimana ditunjukkan pada Gambar 6, Cluster 1 beranggotakan 275 data dengan centroid (1,5418; 3,0073; 1,4509), didominasi wajib pajak petani dengan status pembayaran diragukan dan ketetapan PBB hingga Rp50.000. Cluster 2 beranggotakan 123 data dengan centroid (3,1301; 3,6179; 3,0244), didominasi wiraswasta dengan status macet (lebih dari 3 tahun menunggak) dan ketetapan antara Rp151.000-Rp250.000. Cluster 3 beranggotakan 202 data dengan centroid (4,3515; 2,1188; 1,6287), didominasi pensiunan dengan status kurang lancar (1-2 tahun menunggak) dan ketetapan antara Rp51.000-Rp150.000. Tabel 3 merangkum hasil akhir clustering ini.

Cluster	Pekerjaan Dominan	Status Pembayaran	Ketetapan PBB (Rp)	Anggota
1	Petani	Diragukan (2-3 thn)	hingga 50.000	275
2	Wiraswasta	Macet (>3 thn)	151.000-250.000	123
3	Pensiunan	Kurang Lancar (1-2 thn)	51.000-150.000	202

Tabel 3. Hasil Akhir Clustering - Tiga Cluster

3.4 Pembahasan

Pada pengujian dua cluster maupun tiga cluster, kelompok yang paling dominan dan konsisten muncul adalah wajib pajak berpekerjaan petani dengan tunggakan 2-3 tahun, status pembayaran diragukan, dan ketetapan PBB antara Rp51.000-Rp150.000, yaitu 412 dari 600 data (68,7%) pada pengujian dua cluster dan 275 dari 600 data (45,8%) pada pengujian tiga cluster. Kelompok terbesar kedua pada pengujian dua cluster adalah wajib pajak pegawai swasta dengan status dan kisaran ketetapan yang sama, sedangkan pada pengujian tiga cluster sisa



data terbagi menjadi kelompok wiraswasta dengan tunggakan lebih dari tiga tahun berstatus macet, serta kelompok pensiunan dengan tunggakan 1-2 tahun berstatus kurang lancar.

Hasil ini menunjukkan bahwa petani merupakan kategori pekerjaan yang paling sering muncul dan berkaitan dengan permasalahan pembayaran PBB di Kelurahan Binjai Estate, Temuan ini sejalan dengan penelitian terdahulu yang menerapkan K-Means clustering pada data wajib pajak PBB, yang juga menunjukkan bahwa pengelompokan berdasarkan pekerjaan, status pembayaran, dan ketetapan pajak menghasilkan kelompok yang mendukung kebijakan penagihan yang lebih tepat sasaran [4], [5].

Antarmuka berbasis MATLAB GUIDE yang dikembangkan dalam penelitian ini mampu mengimpor data wajib pajak dari Excel, memvalidasi jumlah variabel dan baris data, menghitung jarak antar data dengan K-Means, serta menampilkan keanggotaan cluster, nilai centroid, keterangan, dan grafik tiga dimensi tanpa memerlukan perhitungan manual.

KESIMPULAN

Kesimpulan dan Saran: Berdasarkan hasil pengujian yang telah dilakukan terhadap 600 data wajib pajak PBB bermasalah di Kelurahan Binjai Estate menggunakan metode *K-Means clustering*, pengujian dilakukan sebanyak 5 kali pada 2 *cluster* dan 10 kali pada 3 *cluster*. Hasil dari seluruh proses pengujian pada *cluster* 2 maupun *cluster* 3, dapat disimpulkan bahwa wajib pajak dengan jenis pekerjaan Petani, yang telah menunggak selama lebih dari 2-3 tahun, dengan status pembayaran dalam kondisi diragukan, serta ketetapan PBB berkisar antara Rp51.000 hingga Rp150.000, merupakan kelompok yang paling dominan dan paling sering muncul secara konsisten pada seluruh pengujian. Antarmuka berbasis MATLAB GUIDE yang dikembangkan mampu mengimpor data, memvalidasi variabel dan baris data, menghitung jarak K-Means, serta menyajikan keanggotaan cluster, nilai centroid, interpretasi kategori, dan grafik tiga dimensi, sehingga membantu pihak kelurahan mengidentifikasi kelompok pekerjaan yang paling banyak bermasalah tanpa perhitungan manual. Penelitian selanjutnya disarankan menambahkan variabel lain di luar pekerjaan, status pembayaran, dan ketetapan PBB, serta membandingkan hasilnya dengan metode clustering lain guna memvalidasi lebih lanjut pengelompokan yang diperoleh.

DAFTAR PUSTAKA

- [1] Primadeni, F., Pradana, A. I., & Purwanto, E. (2024). Analisis Clustering untuk Pengelompokan Wilayah Berdasarkan Karakteristik PBB di Kabupaten Sragen. *Jurnal Pendidikan dan Teknologi Indonesia*, 4(11), 405-417. <https://doi.org/10.52436/1.jpti.500>
- [2] Sri Wahyuni & Sriani. (2025). Penerapan Algoritma K-Means untuk Pengelompokan Kepatuhan Wajib Pajak Bumi dan Bangunan di Kota Medan. *Journal of Computer Engineering, System and Science*, 10(1), 299-311.
- [3] Hakim, M. A., Prasetijo, A. B., & Eridani, D. (2023). Penerapan Data Mining dengan Algoritma K-Means Clustering Penyewaan Alat-Alat Event pada Studi Kasus CV. Dipo Rental Creativindo. 1(4), 148-155. <https://doi.org/10.14710/jtk.v1i4.37011>



- [4] Cantika, R., Fauzi, A., & Sihombing, A. (2024). Penerapan Metode Clustering untuk Mengetahui Kepatuhan Wajib Pajak Bumi dan Bangunan pada Desa Perkebunan Tanjung Keliling. (4).
- [5] Nursaniah, R., Rahaningsih, N., Ali, I., & Dienwati, N. (2024). Pengelompokan Data Penerimaan Pajak Bumi dan Bangunan Berdasarkan Kelurahan di Kota Tasikmalaya Menggunakan Algoritma K-Means. 8(2), 1477-1483.
- [6] Widiya, P., Rahaningsih, N., & Suarna, N. (2023). Pengelompokan Tingkat Kepatuhan Wajib Pajak Bumi dan Bangunan Menggunakan Algoritma K-Means. 7(2), 1358-1362.