

PENGELOMPOKAN DATA PENERIMAAN PAJAK BUMI DAN BANGUNAN BERDASARKAN KELURAHAN DI KOTA TASIKMALAYA MENGGUNAKAN ALGORITMA K-MEANS

¹ Rini Nursaniah, ² Nining Rahaningsih, ³ Irfan Ali ⁴ Nisa Dienwati Nuris

^{1,2}Komputerisasi Akuntansi, STMIK IKMI Cirebon

³Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

⁴Teknik Informatika, STMIK IKMI Cirebon

Sekolah Tinggi Manajemen Informatika dan Komputer IKMI Cirebon,
Jl. Perjuangan No. 10B, Karyamulya, Kesambi, Kota Cirebon, Jawa Barat 45135
rnursaniah@gmail.com

ABSTRAK

Pajak adalah aspek penting dalam ekonomi sebuah negara, karena digunakan untuk meningkatkan kesejahteraan masyarakat. Pajak Bumi dan Bangunan (PBB) merupakan salah satu jenis pajak daerah yang dipungut oleh Pemerintah Daerah. Masalah yang terjadi adalah target dan penerimaan realisasi PBB-P2 pada tahun 2021 masih belum memenuhi target yang telah ditetapkan. Hal ini menunjukkan perbedaan dalam tingkat kepatuhan pajak disetiap kelurahan di kota Tasikmalaya yang dapat mempengaruhi jumlah penerimaan pajak. Akar masalah penelitian ini adalah belum diketahui kelompok penerimaan PBB terendah dan tertinggi, sehingga memerlukan penerapan algoritma *K-Means Clustering* dalam mengelompokkan data untuk mencari solusinya. Penelitian ini menggunakan tahapan KDD dan data yang digunakan yaitu data jumlah penerimaan PBB tahun 2021 oleh Badan Pendapatan Daerah bersumber dari *website* Open Data Kota Tasikmalaya dengan jumlah data sebanyak 69 data. Hasil yang diperoleh yaitu nilai *Davies Bouldin Index* sebesar 0,370 merupakan nilai paling optimal dengan 6 *cluster*. *cluster* 0 dengan jumlah data 8 kelurahan, *cluster* 1 dengan jumlah data 30 kelurahan, *cluster* 2 dengan jumlah data 3 kelurahan, *cluster* 3 dengan jumlah data 19 kelurahan, *cluster* 4 dengan jumlah data 8 kelurahan, *cluster* 5 dengan jumlah data 1 kelurahan. Penerimaan PBB tertinggi terletak pada *cluster* 2 serta penerimaan PBB terendah terletak pada *cluster* 1.

Kata kunci : *K-Means Clustering*, Kota Tasikmalaya, Pajak, Pajak Bumi dan Bangunan

1. PENDAHULUAN

Pajak adalah suatu kewajiban yang disetorkan oleh wajib pajak orang pribadi ataupun badan kepada pemerintah berupa uang dan digunakan untuk meningkatkan kesejahteraan masyarakat. Pajak Daerah adalah kewajiban dari masyarakat suatu daerah untuk diberikan kepada pemerintah daerah agar digunakan untuk kepentingan umum suatu daerah. Pajak yang dibayarkan oleh masyarakat tersebut digunakan untuk bantuan sosial, pembangunan fasilitas dan infrastruktur bagi masyarakat itu sendiri. Salah satu jenis pajak yang termasuk pajak daerah dan dipungut oleh pemerintah daerah yaitu Pajak Bumi dan Bangunan (PBB). Pajak Bumi dan Bangunan (PBB) adalah sebuah kewajiban pembayaran berupa uang yang terjadi karena adanya tanah dan bangunan yang memberikan manfaat berupa status sosial dan ekonomi bagi wajib pajak orang pribadi maupun badan. Jika ditinjau dari sifatnya, Pajak Bumi dan Bangunan memiliki sifat berwujud dengan ciri khusus pada objek berupa tanah dan bangunan oleh karena itu, penentuan besaran pajak berdasarkan jenis harta atau bangunan yang bersangkutan[1]

Klasterisasi adalah mengelompokkan data berdasarkan kemiripan data, dilihat dari nilai atribut kemiripannya atau dengan cara pengukuran jarak datanya, sehingga data yang mirip dikelompokkan. *K-*

means Clustering merupakan salah satu metode algoritma yang sifatnya *unsupervised learning* (berfokus pada analisis data tanpa adanya label) untuk mengelompokkan suatu data kedalam beberapa klaster [2]. Data yang digunakan cocok untuk diolah menggunakan metode ini karena untuk mengetahui kelompok kelurahan penerimaan pajak bumi bangunan tertinggi dengan mengelompokkan data menjadi beberapa klaster.

Salah satu penelitian terdahulu yang telah dilakukan dengan menggunakan algoritma *Clustering* yaitu oleh (Irwan et al, 2022) yang berjudul "Pengelompokan Jenis Penerimaan Pajak di Kota Makassar Menggunakan Fuzzy Clustering". Penelitian ini memfokuskan pada penerapan metode algoritma *Fuzzy C-Means* (FCM) untuk mengelompokkan jenis penerimaan pajak yang paling tinggi di tiap kecamatan di kota Makassar sehingga pemerintah dapat menilai kesejahteraan sosial secara ekonomi dengan lebih akurat. Variabel yang digunakan adalah jenis - jenis penerimaan pajak yaitu Pajak Penghasilan (PPh), Pajak Pertambahan Nilai (PPN) serta Pajak Bumi dan Bangunan (PBB). Hasil dari penelitian ini menunjukkan bahwa cluster yang optimal adalah 7 cluster dengan nilai index PC 0,457. Dalam analisis ini, setiap kecamatan dikelompokkan dalam klaster tertentu, dengan potensi penerimaan pajak tertinggi pada klaster ketujuh dan

potensi penerimaan pajak terendah pada kluster keenam.

Masalah yang perlu diselesaikan pada penelitian ini yaitu penerimaan pajak PBB-P2 oleh Badan Pendapatan Daerah Kota Tasikmalaya pada tahun 2021 masih belum memenuhi target yang telah ditetapkan. Hal ini menunjukkan adanya perbedaan dalam tingkat kepatuhan pajak di setiap kelurahan di kota Tasikmalaya yang dapat mempengaruhi jumlah penerimaan pajak seperti kurangnya kesadaran masyarakat dalam membayar pajak, kemampuan ekonomi masyarakat yang menjadi berkurang ataupun pelayanan publik yang kurang memadai dari pemerintah [3].

Tujuan dari penelitian ini adalah untuk mengetahui jumlah kluster yang optimal untuk digunakan dalam mengelompokkan penerimaan pajak berdasarkan kelurahan menggunakan algoritma *K-Means Clustering* sehingga dapat membantu dalam mengembangkan kebijakan pajak yang lebih adil agar dapat memberikan manfaat bagi masyarakat itu sendiri.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data mining merupakan suatu proses analisis untuk menemukan pola dari kumpulan data yang besar, sehingga menjadi informasi berbentuk pola yang mempunyai arti dalam menunjang pengambilan keputusan [4]. *Data mining* melibatkan penggunaan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk menganalisis data dalam jumlah besar dan mengekstrak informasi yang bermakna [5]. [6] Pembagian *data mining* berdasarkan tugas yang dapat dilaksanakan ada 6 (enam) yaitu : deskripsi, estimasi, prediksi, klasifikasi, pengklusteran dan asosiasi.

2.2. RapidMiner

Merupakan sebuah perangkat lunak yang bersifat terbuka (*open source*) dan memiliki banyak kegunaan dalam analisis data, yang berfokus pada fitur-fitur analisis *data mining*, *text mining* dan analisis prediksi. Memberikan pengguna pemahaman dan mendukung dalam pengambilan keputusan. Operator tersebut memberikan pengguna kemampuan yang sangat luas dalam menjalankan berbagai tugas analisis data [7].

2.3. Clustering

Clustering adalah teknik dalam *data mining* dengan tujuan untuk menguraikan data ke dalam kluster atau kelompok tertentu. Data yang dikelompokkan berdasarkan kemiripan setiap entitas data. Data dapat disusun secara efisien kedalam kelompok yang memiliki kesamaan karakteristik sehingga suatu data dapat dianalisis dengan mudah dan dapat dimengerti oleh penggunaannya. Dalam proses *clustering*, tidak diperlukan label kelas untuk setiap data yang sedang diproses. Label baru hanya

akan diberikan setelah terbentuknya kluster dan biasa dikenal dengan istilah '*unsupervised learning*' [8] Metode *clustering* dibagi menjadi dua jenis yaitu [9]:

a. Hierarchical Clustering

Pendekatan hierarki atau biasa disebut dengan *hierarchical clustering* merupakan salah satu pendekatan yang diterapkan untuk mengelompokkan fitur produk dengan cara yang terstruktur, dengan tujuan untuk menghitung ukuran karakteristik yang serupa dan kedekatan antar fungsi produk. Arah pengelompokan memiliki dua sifat yaitu *divide (top to down)* dan *agglomerative (bottom up)*. Beberapa metode dalam pendekatan ini yaitu tautan tunggal (*single linkage*), tautan penuh (*complete linkage*), tautan rata-rata (*average linkage*) dan variasi minimum (*ward's minimum variance*).

b. Partitional Clustering

Pendekatan partisi atau non hierarkis juga biasa disebut dengan *partitional clustering* merupakan salah satu pendekatan dengan cara menghasilkan partisi kemudian mengevaluasi dengan kriteria tertentu. Beberapa metode dalam *partitional clustering* yaitu : *K-Means*, *Fuzzy K-Means* dan *Mixture Modelling*.

2.4. K-Means

K-means merupakan algoritma yang digunakan untuk mengelompokkan suatu data menjadi (k) kluster dimana setiap data dengan karakteristik yang sama dikelompokkan kedalam kelompok yang sama dan data dengan karakteristik yang berbeda dikelompokkan kedalam kelompok berbeda lainnya. Algoritma ini digunakan untuk pengelompokan data dan pengenalan pola. *K-means* dikenal dengan kesederhanaannya, efisiensi dalam memproses dataset yang besar, dan stabilitas dalam memberikan hasil dengan jumlah iterasi yang memadai [10] Berikut adalah langkah-langkah dalam menjelaskan proses *K-Means Clustering* [8]

- Pilih jumlah kluster (k) yang diinginkan, kemudian tetapkan pusat kluster secara acak.
- Hitung jarak antara setiap data dengan setiap titik pusat kluster.
- Kategorikan data kedalam kluster dengan jarak paling dekat.
- Hitung pusat kluster baru dengan cara mengambil rata-rata dari semua data pada kluster tersebut.
- Ulangi langkah b sampai dengan d secara berulang sehingga tidak ada data yang beralih ke kluster lain.

2.5. Pajak

Pajak adalah pembayaran iuran kepada negara oleh rakyat yang bersifat memaksa tanpa adanya balas jasa dan diatur oleh undang-undang untuk digunakan bagi kepentingan rakyat. Pajak memiliki dua fungsi yaitu sumber keuangan negara (*budgetair*)

artinya pajak berfungsi sebagai sumber pendapatan yang dimanfaatkan untuk membayarkan pengeluaran negara serta fungsi regularend (pengatur) artinya pajak dimanfaatkan untuk menata pelaksanaan kebijakan pemerintah dalam bidang sosial ekonomi agar mencapai suatu tujuan misalnya pengenaan pajak atas barang mewah (PPnBM) dari masyarakat yang memakai barang mewah, semakin mewah suatu barang maka pajak yang dibayarkan akan semakin tinggi [11].

2.6. Pajak Daerah

Pajak Daerah merupakan salah satu sumber dari Pendapatan Asli Daerah (PAD). Pendapatan asli daerah adalah penerimaan yang bersumber dari pajak daerah, retribusi daerah dan lain sebagainya yang didapatkan oleh Daerah berdasarkan peraturan daerah sesuai dengan undang-undang yang berlaku. Pajak Daerah adalah jenis pajak yang dipungut oleh pemerintah daerah terhadap wajib pajak orang pribadi atau badan. Jenis pajak daerah yang pungutannya dilakukan oleh Daerah (Kabupaten/Kota) yaitu Pajak Hotel, Pajak Restoran, Pajak Hiburan, Pajak Reklame, Pajak atas mineral bukan logam dan batuan, Pajak Parkir, Pajak Air Tanah, Pajak Sarang Burung Walet, Pajak Bumi dan Bangunan Pedesaan dan Perkotaan dan Pajak BPHTB.

2.7. Pajak Bumi dan Bangunan

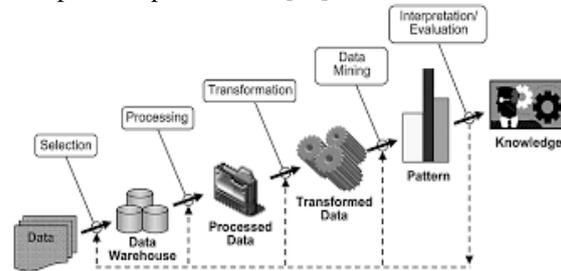
Menurut Undang-Undang Perpajakan dan Retribusi Daerah No 28 Tahun 2009 Pajak Bumi dan Bangunan (PBB) merupakan pajak yang dikenakan atas bangunan diatas tanah atau bangunan yang dimiliki, dikuasai atau dimanfaatkan oleh orang pribadi maupun badan kecuali perkebunan, perhutanan dan pertambangan. Pajak Bumi dan Bangunan (PBB) merupakan salah satu sumber pendapatan dari pajak daerah yang relatif besar karena penghasilan tersebut digunakan untuk kegiatan pembangunan suatu daerah, pembangunan infrastruktur dan fasilitas umum bagi masyarakat. PBB diatur dalam Undang-Undang Nomor 12 Tahun 1994 tentang Pajak Bumi dan Bangunan. Undang-undang tersebut mengatur berbagai aspek seperti objek pajak, subjek pajak, tarif pajak, ketentuan pembebasan pajak, kewajiban pelaporan dan sanksi yang berlaku bagi wajib pajak yang tidak memenuhi kewajibannya.

Pengelolaan PBB-P2 mencakup Pajak Bumi dan Bangunan Pedesaan serta Pajak Bumi dan Bangunan Perkotaan yang merupakan tanggung jawab dari Pemerintah Daerah. Tarif Pajak Bumi dan Bangunan ditentukan oleh beberapa faktor yaitu jenis bangunan, luas tanah dan nilai tanah. Pajak ini merupakan kewajiban masyarakat yang harus dibayarkan setiap tahunnya kepada pemerintah daerah [12]

3. METODE PENELITIAN

Tahapan Metode KDD (*Knowledge Discovery in Database*) merupakan metode atau tahapan yang

digunakan untuk menemukan, menyusun ulang dan menyempurnakan data serta pola yang memiliki makna dari basis data mentah. Berikut ini adalah tahapan dari proses KDD[13]:



Gambar 1. Tahapan KDD

Keterangan dari gambar 1 adalah sebagai berikut :

- Seleksi data (*Data Selection*)**
Proses seleksi atau pemilihan data, bertujuan untuk memilih data yang akan diterima untuk dianalisis lebih lanjut.
- Pemilihan data (*Preprocessing*)**
Proses ini bertujuan untuk mengidentifikasi dan menghilangkan data-data yang tidak lengkap, tidak akurat, tidak relevan, atau terkontaminasi.
- Transformasi data (*Transformation*)**
Setelah data dibersihkan, langkah selanjutnya dalam proses KDD adalah transformasi data. Proses transformasi ini melibatkan konversi data mentah ke dalam format yang lebih sesuai untuk proses data mining
- Data mining***
Hal ini merupakan inti dari proses KDD, yaitu penerapan berbagai teknik data mining seperti pengelompokkan, klasifikasi, regresi untuk mengekstrak pola, pengetahuan dan informasi dari data.
- Evaluasi Pola (*Evaluation*)**
Setelah pola diidentifikasi, pola tersebut perlu dievaluasi kualitas dan potensi kegunaannya terhadap tujuan bisnis atau penelitian. Dalam penelitian ini menggunakan metode *Davies Bouldin Index* dengan tujuan untuk menentukan jumlah kluster yang optimal, nilai *Davies Bouldin Index (DBI)* yang semakin mendekati nilai 0 menunjukkan semakin baik kluster yang diperoleh [14]. Pada tahun 1979 David L. Davies dan Donald W. Bouldin memperkenalkan metode *Davies-Bouldin Index (DBI)*. Merupakan metode yang memiliki tujuan untuk melakukan evaluasi pada metode pengelompokan (*clustering*) yang menghasilkan seberapa baik nilai dari tiap cluster terhadap data yang digunakan agar mendapatkan kelompok atau cluster yang optimal berdasarkan kemiripan antar data dari tiap cluster[15]

4. HASIL DAN PEMBAHASAN

4.1. Seleksi Data (*Data Selection*)

Pada tahap ini data yang digunakan yaitu data penerimaan Pajak Bumi dan Bangunan Kota Tasikmalaya tahun 2021 dengan jumlah sebanyak 69

data sebagai bahan penelitian. Tahap pertama adalah menggunakan operator *Read CSV*. Operator *Read CSV* digunakan untuk mengelola data dengan cara membaca data dalam bentuk format *csv (comma separated values)*. Kemudian dilakukan pemilihan data yang akan digunakan. Data yang sudah diproses dapat ditampilkan pada menu *Results* ditampilkan pada gambar berikut ini.

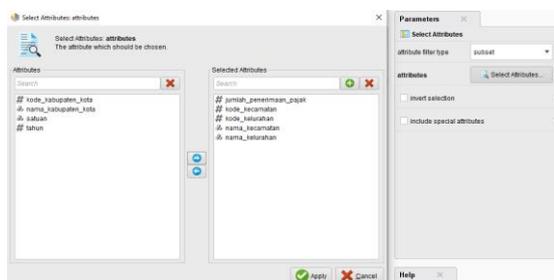
Row No.	nama_kelur...	kode_keca...	nama_keca...	kode_kelura...	jumlah_pen...	tahun
1	LEUWILIANG	3278010	KAWALI	327801001	8907882	2021
2	URUG	3278010	KAWALI	327801002	278212732	2021
3	GUNUNGTA...	3278010	KAWALI	327801003	253882107	2021
4	GUNUNGGE...	3278010	KAWALI	327801004	263859480	2021
5	TALAGASARI	3278010	KAWALI	327801005	131348841	2021
6	TANJUNG	3278010	KAWALI	327801006	105775479	2021
7	CIBELUTI	3278010	KAWALI	327801007	173255899	2021
8	KARANGANY...	3278010	KAWALI	327801008	184035898	2021
9	CILAMAJANG	3278010	KAWALI	327801009	163831627	2021
10	KARSAMENAK	3278010	KAWALI	327801010	712045285	2021
11	SETIAMULYA	3278020	TAMANSARI	327802001	170124479	2021
12	SETIAWARGI	3278020	TAMANSARI	327802002	257313479	2021
13	TAMANSARI	3278020	TAMANSARI	327802003	265559058	2021

Gambar 2. Tampilan data

Pada gambar diatas merupakan hasil dari proses import data yang menampilkan data dari operator *Read CSV* pada menu *Results* sehingga proses dapat dilanjutkan pada tahap berikutnya.

4.2. Pemilihan data (*Preprocessing*)

Tahapan ini merupakan cara untuk mengolah data dari sebelumnya data mentah menjadi data yang siap untuk diproses lebih lanjut. Tahapan dari *Preprocessing* yang dilakukan yaitu data *reduction*. *Data reduction* adalah pengurangan dari suatu data yang diambil tetapi hal tersebut tidak akan mengubah hasil analisis dari data itu sendiri. Atribut yang tidak dibutuhkan pada data penerimaan Pajak Bumi dan Bangunan ini ada 4 atribut yaitu atribut kode kabupaten kota, nama kabupaten kota, satuan dan tahun. Sehingga atribut yang digunakan untuk dianalisis lebih lanjut memiliki 5 atribut yaitu kode kecamatan, nama kecamatan, kode kelurahan, nama kelurahan dan jumlah penerimaan pajak. Operator *Select Attribute* akan digunakan dalam proses ini.



Gambar 3. Proses pemilihan data

Gambar diatas merupakan proses seleksi atribut, pada parameter *attribute filter type* pilih *subset* yaitu memilih atribut tertentu dari sebuah *ExampleSet* dengan membuang 4 atribut yang tidak dibutuhkan dengan tujuan untuk fokus pada atribut yang relevan dengan data yang digunakan.

Row No.	nama_kelur...	kode_keca...	nama_keca...	kode_kelura...	jumlah_pen...
1	LEUWILIANG	3278010	KAWALI	327801001	8907882
2	URUG	3278010	KAWALI	327801002	278212732
3	GUNUNGTA...	3278010	KAWALI	327801003	253882107
4	GUNUNGGE...	3278010	KAWALI	327801004	263859480
5	TALAGASARI	3278010	KAWALI	327801005	131348841
6	TANJUNG	3278010	KAWALI	327801006	105775479
7	CIBELUTI	3278010	KAWALI	327801007	173255899
8	KARANGANY...	3278010	KAWALI	327801008	184035898
9	CILAMAJANG	3278010	KAWALI	327801009	163831627
10	KARSAMENAK	3278010	KAWALI	327801010	712045285
11	SETIAMULYA	3278020	TAMANSARI	327802001	170124479
12	SETIAWARGI	3278020	TAMANSARI	327802002	257313479
13	TAMANSARI	3278020	TAMANSARI	327802003	265559058

Gambar 4. Hasil dari data reduction

Menu *Result* ditampilkan pada gambar 4 memuat hasil setelah diproses menggunakan operator *Select Attribute* agar memperkecil kesalahan data serta dalam melakukan model *K-Means* lebih cepat.

4.3. Transformasi Data

Transformasi data adalah proses mengubah dataset dengan tujuan agar data lebih sesuai dengan proses pemodelan *data mining*. Dalam proses ini menggunakan algoritma *K-Means Clustering* dimana atribut dari suatu data harus berupa *numerik*. Oleh karena itu, data yang memiliki atribut *nominal* harus diubah menjadi *numerik* menggunakan operator *Nominal to Numerical*.

Row No.	nama_kelur...	nama_keca...	kode_keca...	kode_kelura...	jumlah_pen...
1	LEUWILIANG	0	3278010	327801001	8907882
2	URUG	0	3278010	327801002	278212732
3	GUNUNGTA...	0	3278010	327801003	253882107
4	GUNUNGGE...	0	3278010	327801004	263859480
5	TALAGASARI	0	3278010	327801005	131348841
6	TANJUNG	0	3278010	327801006	105775479
7	CIBELUTI	0	3278010	327801007	173255899
8	KARANGANY...	0	3278010	327801008	184035898
9	CILAMAJANG	0	3278010	327801009	163831627
10	KARSAMENAK	0	3278010	327801010	712045285
11	SETIAMULYA	1	3278020	327802001	170124479
12	SETIAWARGI	1	3278020	327802002	257313479
13	TAMANSARI	1	3278020	327802003	265559058

Gambar 5. Hasil Transformasi Data

Pada gambar 5 merupakan hasil dari transformasi data pada atribut nama kecamatan kota dari sebelumnya bentuk nominal menjadi numerik, kemudian pada parameter *unique integer* dicentang (✓) sehingga setiap nilai yang terdapat pada atribut tersebut diubah menjadi nilai numerik yang unik atau berbeda yaitu 0,1, 2 dan seterusnya agar dapat diproses ke tahap berikutnya.

4.4. Data Mining

Tahap selanjutnya yaitu pemodelan data, metode yang digunakan untuk penelitian ini adalah *clustering* dengan menerapkan algoritma *K-Means* menggunakan operator *Clustering (K-Means)* dengan fungsi untuk mengolah data dengan metode *K-Means* untuk mengelompokkan data penerimaan Pajak Bumi dan Bangunan menjadi 6 klaster.

Row No.	nama_kelur...	cluster	nama_keca...	kode_keca...	kode_kelura...	jumlah_pen...
1	LEUWILIANG	cluster_1	0	3278010	3278010001	89070862
2	LURUG	cluster_3	0	3278010	3278010002	278212732
3	GUNUNGTAL	cluster_3	0	3278010	3278010003	253882107
4	GUNUNGSE...	cluster_3	0	3278010	3278010004	263898480
5	TALAGASARI	cluster_1	0	3278010	3278010005	131348941
6	TANJUNG	cluster_1	0	3278010	3278010006	105775479
7	CIBEUTI	cluster_1	0	3278010	3278010007	173255899
8	KARANGANY...	cluster_1	0	3278010	3278010008	184035898
9	CILAMAJANG	cluster_1	0	3278010	3278010009	163831627
10	KARSAMENAK	cluster_0	0	3278010	3278010010	712045285
11	SETHAMULYA	cluster_1	1	3278020	3278020001	170124479
12	SETIAWARGI	cluster_3	1	3278020	3278020002	257313479
13	TAMUNGSI	cluster_3	1	3278020	3278020003	265559058

Gambar 6. Hasil Clustering

Pada gambar diatas menampilkan hasil clustering dengan diketahui $k = 6$, $max\ runs$ sebanyak 10 kali putaran dan $max\ optimization\ step$ berjumlah 100.

4.5. Evaluasi Pola (Evaluation)

Setelah proses pemodelan menggunakan metode K -Means dibuat, selanjutnya adalah hasil yang diperoleh setelah melakukan eksekusi data menggunakan aplikasi *RapidMiner*. Operator yang digunakan yaitu operator *Performance (Cluster Distance Performance)* berfungsi untuk evaluasi kinerja dari metode *clustering*. Parameter *performance* yaitu *maximize* dicentang (✓) supaya hasil *result* pada *performance* tidak minus.

Cluster	Avg. within centroid distance	Davies Bouldin Index (DBI)
cluster_0	2131609434504373.000	
cluster_1	2296059460221230.000	
cluster_2	121991536483294.200	
cluster_3	8608132449097819.000	
cluster_4	2147707422247342.000	
cluster_5	3155548469420430.000	
cluster_6	0.000	0.370

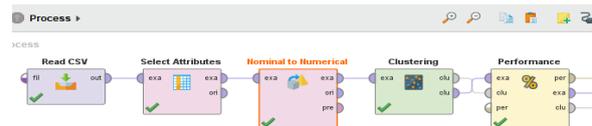
Gambar 7. Hasil dari PerformanceVector

Berdasarkan gambar diatas fungsi dari operator *performance* yaitu untuk mengetahui nilai dari *Avg.within centroid distance* dan *Davies Bouldin Index (DBI)*. Data dikelompokkan berdasarkan jarak terdekat menggunakan algoritma k -Means. Tabel percobaan dalam menentukan nilai k adalah sebagai berikut.

Tabel 1. Hasil percobaan dalam menentukan nilai k

Nilai K	DBI
2	0,497
3	0,405
4	0,441
5	0,431
6	0,370
7	0,406
8	0,450
9	0,374
10	0,409

Setelah melakukan pengujian sebanyak 10 (sepuluh) kali untuk menentukan nilai k , pada percobaan kelima dengan nilai $k = 6$ hasilnya adalah nilai dari *Davies Bouldin Index (DBI)* yang diperoleh lebih optimal dibandingkan dengan pengujian sebelumnya. Dengan nilai yang dihasilkan adalah nilai yang paling baik yaitu mendekati nilai nol (0) pada 0,370. Setelah proses evaluasi dilakukan menggunakan metode *DBI* dan mendapatkan nilai kluster yang optimal yaitu 6 cluster. Pemodelan pada algoritma k -means ditampilkan pada gambar dibawah ini :



Gambar 8. Model dari algoritma k -means

Pada gambar 8 menampilkan proses pengolahan data menggunakan metode K -Means. Pengolahan data menggunakan algoritma K -Means *Clustering* pada tools *RapidMiner* menggunakan 5 operator yaitu *Read CSV* berfungsi untuk membaca data dengan format *csv* untuk diproses lebih lanjut, operator *Select Attribute* berfungsi untuk mengurangi atau menghapus data yang tidak dibutuhkan, Operator *Nominal to Numerical* berfungsi untuk mengubah data dengan atribut berupa nominal menjadi numerik, Operator *Clustering (K-Means)* berfungsi untuk mengelompokkan data menjadi 6 kluster dengan $max\ runs$ sebanyak 10 kali putaran, Operator *Performance (Cluster Distance Performance)* berfungsi untuk mengevaluasi kinerja dari k -means *clustering* dalam mengelompokkan data penerimaan Pajak Bumi dan Bangunan.

Jumlah data pada tiap kluster ditunjukkan pada gambar dibawah ini.

Cluster Model

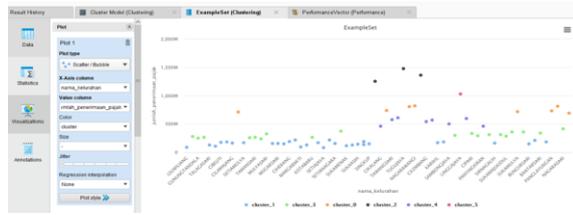
```

Cluster 0: 8 items
Cluster 1: 30 items
Cluster 2: 3 items
Cluster 3: 19 items
Cluster 4: 8 items
Cluster 5: 1 items
Total number of items: 69
  
```

Gambar 9. Jumlah data tiap cluster

Dari gambar diatas dapat dijelaskan bahwa untuk kluster 0 memiliki jumlah data 8 kelurahan, kluster 1 memiliki jumlah data 30 kelurahan, kluster 2 memiliki jumlah data 3 kelurahan, kluster 3 memiliki jumlah data 19 kelurahan, kluster 4 memiliki jumlah data 8 kelurahan dan kluster 5 memiliki jumlah data 1 kelurahan dengan total keseluruhan 69 kelurahan.

Tampilan hasil cluster pada data penerimaan Pajak Bumi dan Bangunan adalah sebagai berikut .



Gambar 10. Visualisasi data

Pada gambar diatas merupakan visualisasi data menggunakan plot *Scatter/Bubble*, dapat disimpulkan bahwa dalam pengelompokan data penerimaan pajak Bumi dan Bangunan pada 69 kelurahan di kota Tasikmalaya terbagi menjadi 6 *cluster*. Penerimaan Pajak Bumi dan Bangunan tertinggi terletak pada *cluster 2* dengan jumlah data 3 kelurahan serta penerimaan Pajak Bumi dan Bangunan terendah pada *cluster 1* dengan jumlah data 30 kelurahan.

Berikut adalah hasil klaster yang ditampilkan pada tabel 2.

Tabel 1. Hasil *cluster*

Cluster	Kelurahan
0	Karsamenak, Empangsari, Tuguraja, Nagarawangi, Sukamulya, Panglayungan, Cipedes, Sukamanah
1	Leuwiliang, Talagasari, Tanjung, Cibeut, Karanganyar, Cimalajang, Setiamulya, Sukahurip, Mugasari, Sumelap, Cihayang, Ciakar, Margabakti, Awipari, Kersanagara, Setiajaya, Setiaratu, Setianagara, Sukamenak, Purbaratu, Sukaasih, Sukajaya, Singkup, Karikil, Cigantang, Sirnagalih, Bungursari, Sukajaya, Cibunigeulis, Sukalaksana
2	Kahuripan, Tugujaya, Yudanagara
3	Urug, Gunungtandala, Gununggede, Setiawargi, Tamansari, Tamanjaya, Mulyasari, Kotabaru, Sukanagara, Sambongpari, Cipari, Cipawitra, Parakannyasag, Indihiang, Sukamajukidul, Sukamajukaler, Sukarindik, Bantarsari, Nagarasari.
4	Cikalang, Tawangasari, Lengongsari, Cilembang, Argasari, Sambongjaya, Mangkubumi, Panyingkiran.
5	Linggajaya

Pada tabel 2 merupakan hasil klaster berisikan beberapa Kelurahan yang terbagi kedalam 6 *cluster* berdasarkan jumlah penerimaan Pajak Bumi dan Bangunan di kota Tasikmalaya.

Kelompok dengan jumlah penerimaan Pajak Bumi dan Bangunan tertinggi terletak pada *cluster 2* dan merupakan kelompok dengan penerimaan Pajak Bumi dan Bangunan terbesar dari kelompok *cluster* lainnya karena berdasarkan data dari Dinas Kependudukan dan Pencatatan Sipil jumlah penduduk dari Kelurahan Kahuripan Kota Tasikmalaya pada tahun 2021 sebanyak 21.972 jiwa yang menandakan jumlah penduduk pada kelurahan tersebut cukup banyak. Kelurahan Tugujaya dan Yudanagara termasuk Kecamatan Cihideung yang

merupakan pusat Kota Tasikmalaya yaitu dijadikan pusat pelayanan umum, perdagangan, jasa dan perkantoran. Memiliki fasilitas dalam perdagangan seperti minimarket sebanyak 23 unit, pasar dengan 4 unit permanen, pasar dengan 2 unit semi permanen serta pertokoan sebanyak 37 unit [16]. Menurut [17] jumlah penduduk yang banyak maka semakin banyak tempat tinggal masyarakat yang berdampak pada meningkatnya penerimaan Pajak Bumi dan Bangunan. Luas lahan bangunan berdampak baik terhadap penerimaan Pajak Bumi dan Bangunan, sebab semakin luas lahan bangunan maka jumlah pajak Pajak Bumi dan Bangunan yang terutang semakin tinggi kemudian wajib pajak memiliki kesadaran yang tinggi dalam kewajiban untuk membayar juga akan meningkatkan penerimaan pajak.

Kelompok dengan jumlah penerimaan Pajak Bumi dan Bangunan terendah terletak pada *cluster 1* karena adanya beberapa kelurahan termasuk Kecamatan Indihiang dan Bungursari dengan jumlah penduduk terendah daripada jumlah penduduk Kecamatan lain di Kota Tasikmalaya sehingga penerimaan Pajak Bumi dan Bangunan yang diperoleh juga rendah, faktor lain yaitu kemungkinan SPPT (surat pemberitahuan pajak terutang) yang tidak diterima oleh wajib pajak atau data wajib pajak yang tidak sesuai, rendahnya pengetahuan masyarakat mengenai kewajiban membayar pajak karena masyarakat lebih mengutamakan untuk kebutuhan sehari hari sehingga tidak patuh dalam membayar pajak[18].

5. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian menggunakan metode *k-means* adalah data jumlah penerimaan pajak bumi dan bangunan berdasarkan kelurahan di kota Tasikmalaya tahun 2021 dikelompokkan menjadi 6 klaster dengan hasil evaluasi menggunakan DBI (*Davies Bouldin Index*) dengan nilai optimal sebesar 0,370 yaitu Penerimaan Pajak Bumi dan Bangunan dikelompokkan menjadi 6 *cluster*. *cluster 0* dengan jumlah data 8 kelurahan, *cluster 1* dengan jumlah data 30 kelurahan, *cluster 2* dengan jumlah data 3 kelurahan, *cluster 3* dengan jumlah data 19 kelurahan, *cluster 4* dengan jumlah data 8 kelurahan serta *cluster 5* dengan jumlah data 1 kelurahan. Beberapa saran terhadap pengelompokan data penerimaan Pajak Bumi dan Bangunan berdasarkan kelurahan yaitu untuk penelitian selanjutnya adalah penambahan jumlah dataset agar mendapatkan hasil lebih akurat serta dalam pengelompokan data dapat menggunakan metode *clustering* yang lain untuk mengetahui perbedaan dari hasil yang didapatkan kemudian untuk Pemerintah Daerah diharapkan untuk menghimbau masyarakat agar taat dalam membayar pajak seperti melakukan sosialisasi kepada wajib pajak mengenai pentingnya membayar pajak agar meningkatnya pembangunan serta realisasi

Pajak Bumi dan Bangunan dapat mencapai target yang telah ditentukan.

DAFTAR PUSTAKA

- [1] Ratna Wulandari, "Pengaruh Kesadaran Wajib Pajak, Kualitas Pelayanan. Pengetahuan Perpajakan dan Sanksi Perpajakan Terhadap Kepatuhan Wajib Pajak PBB-P2 di Kabupaten Rembang," *J SMA (Jurnal Sains Manajemen dan Akuntansi)*, vol. 15, no. 1, pp. 86–103, Apr. 2023, doi: 10.37151/jsma.v15i1.120.
- [2] W. Xiang, N. Zhu, S. Ma, X. Meng, and M. An, "A dynamic shuffled differential evolution algorithm for data clustering," *Neurocomputing*, vol. 158, pp. 144–154, Jun. 2015, doi: 10.1016/j.neucom.2015.01.058.
- [3] A. Y. Supriatna, "PENDAPATAN DAERAH KOTA TASIKMALAYA," *Jurnal Administrasi dan Kebijakan Publik*, vol. 3, no. 2, 2022.
- [4] S. Sharyanto and D. Lestari, "Penerapan Data Mining Untuk Menentukan Segmentasi Pelanggan Dengan Menggunakan Algoritma K-Means dan Model RFM Pada E-Commerce," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 4, p. 866, Aug. 2022, doi: 10.30865/jurikom.v9i4.4525.
- [5] F. Ginting, E. Buulolo, and E. R. Siagian, "IMPLEMENTASI ALGORITMA REGRESI LINEAR SEDERHANA DALAM MEMREDIKSI BESARAN PENDAPATAN DAERAH (STUDI KASUS: DINAS PENDAPATAN KAB. DELI SERDANG)," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 3, no. 1, Nov. 2019, doi: 10.30865/komik.v3i1.1602.
- [6] R. R. Rerung, "Penerapan Data Mining dengan Memanfaatkan Metode Association Rule untuk Promosi Produk," *Jurnal Teknologi Rekayasa*, vol. 3, no. 1, p. 89, Jun. 2018, doi: 10.31544/jtera.v3.i1.2018.89-98.
- [7] D. Ramdhan, G. Dwilestari, R. D. Dana, A. Ajiz, and K. Kaslani, "Clustering Data Persediaan Barang Dengan Menggunakan Metode K-Means," *MEANS (Media Informasi Analisa dan Sistem)*, pp. 1–9, May 2022, doi: 10.54367/means.v7i1.1826.
- [8] F. Farahdinna, I. Nurdiansyah, A. Suryani, and A. Wibowo, "PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS DALAM KLASIFIKASI PRODUK ASURANSI PERUSAHAAN NASIONAL," *Jurnal Ilmiah FIFO*, vol. 11, no. 2, p. 208, Nov. 2019, doi: 10.22441/fifo.2019.v11i2.010.
- [9] G. E. I. Kambey, R. Sengkey, and A. Jacobus, "Penerapan Clustering pada Aplikasi Pendeteksi Kemiripan Dokumen Teks Bahasa Indonesia," *Teknik Informatika*, vol. 15, pp. 75–82, Apr. 2020.
- [10] R. K. Dinata, S. Safwandi, N. Hasdyna, and N. Azizah, "Analisis K-Means Clustering pada Data Sepeda Motor," *INFORMAL: Informatics Journal*, vol. 5, no. 1, p. 10, Apr. 2020, doi: 10.19184/isj.v5i1.17071.
- [11] A. Syarifudin, *BUKU AJAR PERPAJAKAN*. Kebumen: STIE Putra Bangsa, 2018.
- [12] M. N. Huda and G. Wicaksono, "Analisis Efektivitas Dan Kontribusi Penerimaan Pajak Bumi Dan Bangunan Perdesaan Dan Perkotaan Terhadap Pendapatan Asli Daerah Kota Yogyakarta," *Educoretax*, vol. 1, no. 4, pp. 284–290, Dec. 2021, doi: 10.54957/educoretax.v1i4.108.
- [13] Haris Kurniawan, Sarjon Defit, and Sumijan, "Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal," *Journal of Applied Computer Science and Technology*, vol. 1, no. 2, pp. 80–89, Dec. 2020, doi: 10.52158/jacost.v1i2.102.
- [14] D. A. I. C. Dewi and D. A. K. Pramita, "Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali," *Matrix: Jurnal Manajemen Teknologi dan Informatika*, vol. 9, no. 3, pp. 102–109, Nov. 2019, doi: 10.31940/matrix.v9i3.1662.
- [15] E. Muningsih, I. Maryani, and V. R. Handayani, "Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa," *Sains dan Manajemen*, vol. 9, Mar. 2021.
- [16] A. Hamid, *KECAMATAN CIHIDEUNG DALAM ANGKA 2022*. BPS Kota Tasikmalaya, 2022.
- [17] L. Lisnawati, "PENGARUH FAKTOR DEMOGRAFIS DAN PRODUK DOMESTIK REGIONAL BRUTO (PDRB) TERHADAP REALISASI PAJAK BUMI DAN BANGUNAN PERKOTAAN DAN PEDESAAN (PBB-P2) DAN DAMPAKNYA PADA PENERIMAAN PAJAK DAERAH DI 18 KABUPATEN DI JAWA BARAT," *LAND JOURNAL*, vol. 1, no. 2, pp. 203–219, Sep. 2021, doi: 10.47491/landjournal.v1i2.713.
- [18] R. A. Muslim, "Implementasi Pemungutan dan Penegakan Hukum Pajak Bumi Dan Bangunan Perdesaan dan Perkotaan di Kota Tasikmalaya (Berdasarkan Peraturan Daerah Kota Tasikmalaya Nomor 4 Tahun 2011 Tentang Pajak Daerah)," *Supremasi Hukum: Jurnal Kajian Ilmu Hukum*, vol. 7, no. 2, Nov. 2018, doi: 10.14421/sh.v7i2.2041.