

Penerapan Data Mining Menggunakan Metode Clustering Untuk Mengetahui Kelompok Kepatuhan Wajib Pajak Bumi dan Bangunan

Medina Aprilia Putri¹, Nining Rahaningsih², Fadhil M Basysyar³, Odi Nurdiawan⁴

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

³Sistem Informasi, STMIK IKMI Cirebon, Indonesia

⁴Manajemen Informatika, STMIK IKMI Cirebon, Indonesia

medinaaprilias@gmail.com

Info Artikel

Sejarah artikel :

Diterima Maret 2022

Direvisi Agustus 2022

Disetujui September 2022

Diterbitkan September 2022

ABSTRACT

Local taxes are taxes set by local governments whose collection authority and tax proceeds are used to fund regional expenditures. One of the taxes included in the authority of local governments is the Land and Building Tax (PBB). Land and Building Tax is one of the taxes that can be paid through the village government. With the increasing number of taxpayers in the village, the data on payment of tax contributions that go directly to the state treasury causes the Kendal Village government, Astanajapura District, Cirebon Regency not to know how many taxpayers are obedient and disobedient. This study uses data mining techniques namely the Clustering Method using the K-Means method. This study uses the Knowledge Discovery in Database (KDD) stage with the amount of data used as much as 1,159 in the form of taxpayer data for the Kendal Village community in 2021. The results of the RapidMiner test using the Davies Bouldin Index calculation obtained a cluster determination value with a value of 4 (0.862). Cluster 0 contains members who have a low level of compliance in paying PBB, Cluster 1 contains members of taxpayers with a moderate level of compliance in paying PBB, Cluster 2 has a high level of taxpayer compliance and Cluster 3 is a cluster with a very high level of taxpayer compliance. By having the most dominant average price determined by PBB in each cluster is Rp. 18.000,-.

Keywords : Clustering; Data Mining; K-Means; Land and Building Tax; Tax Compliance.

ABSTRAK

Pajak Daerah merupakan pajak yang ditetapkan oleh pemerintah Daerah yang wewenang pemungutan dan hasil pajaknya digunakan untuk mendanai pengeluaran daerah tersebut. Salah satu pajak yang termasuk wewenang pemerintah daerah adalah Pajak Bumi dan Bangunan (PBB). Pajak Bumi dan Bangunan merupakan salah satu pajak yang pembayarannya dapat melalui pemerintah Desa. Dengan semakin banyaknya wajib pajak yang ada di Desa, data pembayaran iuran Pajak yang langsung masuk ke kas negara menyebabkan pihak pemerintah Desa Kendal Kecamatan Astanajapura Kabupaten Cirebon tidak mengetahui seberapa banyak masyarakat wajib pajak yang patuh dan tidak patuh. Pada penelitian ini menggunakan Teknik *data mining* yaitu Metode *Clustering* dengan menggunakan metode *K-Means*. Penelitian ini menggunakan tahapan *Knowledge Discovery in Database (KDD)* dengan jumlah data yang digunakan sebanyak 1.159 yang berupa data Wajib Pajak masyarakat Desa Kendal Tahun 2021. Hasil dari pengujian RapidMiner dengan menggunakan perhitungan *Davies Bouldin Index* diperoleh nilai penentuan klaster dengan nilai 4 (0.862). *Cluster 0* berisikan anggota yang memiliki tingkat kepatuhan membayar PBB yang rendah, *Cluster 1* berisikan anggota Wajib pajak dengan tingkat kepatuhan

membayar PBB yang sedang, *Cluster 2* memiliki tingkat kepatuhan Wajib Pajak yang tinggi dan *Cluster 3* adalah cluster dengan tingkat kepatuhan Wajib Pajak yang sangat tinggi. Dengan memiliki rata-rata harga ketetapan PBB yang paling mendominasi pada setiap cluster adalah Rp. 18.000,-.

Kata Kunci : *Clustering; Data Mining; K-Means; Kepatuhan Pajak; Pajak Bumi dan Bangunan.*

PENDAHULUAN

Perpajakan merupakan salah satu bagian terpenting dalam Anggaran Pendapatan dan Belanja Negara (APBN). Hal ini terbukti dari besarnya kontribusi penerimaan pajak dalam mendanai APBN setiap tahunnya. Tercatat seluruh komponen penerimaan yang bersumber dari pajak hingga akhir September 2021 telah mencapai Rp 850,06 triliun atau 69,3 persen terhadap target APBN 2021. Pajak dikenakan kepada seluruh masyarakat sesuai dengan undang-undang yang berlaku. Berdasarkan Undang-Undang Nomor 23 tahun 2014 tentang Pemerintahan Daerah, Pemerintah Daerah berhak mengatur dan mengurus urusan pemerintahan dan kepentingan masyarakatnya sendiri. Salah satu sumber yang dimiliki dan dikelola oleh Pemerintah Daerah adalah Pendapatan Asli Daerah (PAD). Pendapatan Asli Daerah merupakan pendapatan yang diperoleh dari hasil pemungutan pajak kepada wajib pajak yang sesuai dengan peraturan perundang-undangan dan bersifat memaksa. Salah satu pajak yang termasuk PAD adalah Pajak Bumi dan Bangunan. Menurut undang-undang nomor 12 tahun 1985, pajak Bumi dan Bangunan adalah pajak yang bersifat kebendaan. Artinya, besarnya tanggungan pajak ditentukan oleh keadaan objek yaitu bumi/tanah dan bangunan bukan dari keadaan subjek (pelaku wajib pajak) itu sendiri.

Dalam Undang-Undang nomor 28 tahun 2009 tentang Pajak Daerah dan Retribusi Daerah [1], Pajak bumi dan Bangunan (PBB) Perdesaan dan Perkotaan adalah pajak atas bumi dan/atau bangunan yang dimiliki, dikuasai, dan/atau dimanfaatkan oleh orang pribadi atau badan, kecuali kawasan yang digunakan untuk kegiatan usaha perkebunan, perhutanan dan pertambangan. Yang dimaksud bumi dalam pasal 28 ini adalah permukaan bumi yang meliputi tanah dan perairan pedalaman serta laut wilayah kabupaten/kota. Sedangkan Bangunan adalah konstruksi teknik yang ditanam atau diletakkan secara tetap pada tanah dan/atau perairan pedalaman dan/atau laut.

Pajak memiliki dua fungsi [2][3] yaitu, fungsi *budgetair* dan fungsi mengatur/regulasi. Fungsi *budgetair* adalah pajak yang berfungsi sebagai alat pemerintah untuk mendapatkan penghasilan dari masyarakat yang digunakan untuk berbagai kepentingan negara. Berdasarkan fungsi *budgetair* ini, adanya kesadaran dan kedisiplinan masyarakat untuk mematuhi kewajiban sebagai pelaku Wajib pajak sangat dibutuhkan. Hal ini yang akan menentukan tinggi atau rendahnya tingkat kepatuhan wajib pajak pada suatu daerah. Dengan semakin bertambahnya wajib pajak pada suatu daerah, maka data yang tersimpan juga akan semakin banyak.

Pada umumnya, kepatuhan pajak terbagi menjadi dua, yaitu kepatuhan formal dan kepatuhan material. Kepatuhan formal adalah keadaan dimana wajib pajak memenuhi kewajiban secara formal sesuai dengan ketentuan dalam Undang-Undang perpajakan. Maksudnya, kepatuhan pajak secara formal ini adalah kepatuhan yang mencakup sejauh mana wajib pajak patuh terhadap persyaratan

prosedural dan administrasi pajak, termasuk mengenai syarat pelaporan serta waktu untuk menyampaikan dan membayar pajak. Sedangkan arti dari kepatuhan material adalah keadaan dimana wajib pajak secara substantif atau hakikatnya memenuhi semua ketentuan material perpajakan yakni sesuai isi dan jiwa Undang-Undang. Wajib pajak yang memenuhi kepatuhan material adalah wajib pajak yang mengisi formulir Surat Pemberitahuan Objek Pajak (SPOP). SPOP tersebut diisi dengan jelas, benar dan lengkap sesuai dengan objek pajak dan wajib pajak membayar pajak yang terutang disertai dengan bukti SPPT. [4]

Menurut M. nanda Variesta Waworuntu dkk [5] Penerapan Metode *K-Means* pemetaan calon penerima JAMKESDA memberikan 2 kelompok yaitu kelompok mampu (cluster 1) yang memiliki jumlah penghasilan tinggi dan jumlah anggota keluarga sedikit. Sedangkan kelompok 2 yaitu kelompok tidak mampu (cluster 0) memiliki jumlah anggota sebanyak 334 anggota. sedangkan menurut [6] Penerapan Metode *K-Means* dalam pengelompokkan wilayah menurut intensitas kejadian bencana alam di Indonesia Tahun 2013-2018 dimana pengelompokkan ini bertujuan untuk mengetahui wilayah yang rawan terjadinya bencana alam. Hasil dari pengelompokan ini menghasilkan 2 cluster optimal, cluster pertama merupakan wilayah yang rawan terjadinya bencana alam dengan memiliki 3 anggota dari 34 Provinsi. Sedangkan cluster kedua berisikan wilayah yang bukan daerah rawan terjadinya bencana alam terdiri dari 31 anggota.

Desa Kendal Kecamatan Astanajapura merupakan salah satu Instansi Pemerintah yang memiliki wewenang untuk melakukan pemungutan Pajak Bumi dan Bangunan kepada masyarakat di wilayahnya sendiri. Desa Kendal juga mempunyai tanggung jawab terhadap laporan hasil penerimaan PBB sekaligus memperkirakan pemasukan pajak di wilayahnya. Pada tahun 2021, sebanyak 1.159 data wajib pajak Bumi dan Bangunan yang ada di Desa Kendal. Dengan data yang sebanyak ini, Sekretaris Desa selaku kordinator penanganan pajak Bumi dan Bangunan mengalami beberapa kesulitan untuk mengetahui kelompok kepatuhan Wajib Pajak dan mengetahui Blok mana yang tingkat kepatuhan dalam membayar pajak PBB yang tinggi dan yang rendah. Maka dari itu, dibuatlah sebuah pemodelan data agar petugas yang memiliki wewenang dalam PBB dapat mengetahui kelompok Wajib Pajak PBB dan Blok yang memiliki tingkat kepatuhan pajak yang tinggi dan rendah.

Data yang akan digunakan dalam penelitian ini berisikan 6 (enam) atribut, yaitu Nomor Objek Pajak (NOP), Nama Wajib Pajak, Ketentuan PBB dan Alamat Objek Pajak. Sedangkan untuk atribut status dan atribut SPPT merupakan atribut tambahan yang ditambahkan oleh peneliti. Atribut SPPT menjelaskan keberadaan SPPT apakah SPPT terhutang pajak berada di Kantor Pemerintah Desa Kendal atau tidak. Sedangkan Atribut Status berisi apakah wajib pajak sudah membayar iuran PBB atau belum ditentukan oleh keberadaan Surat Pemberitahuan Pajak Terutang (SPPT). Jika wajib pajak sudah membayar iuran PBB, maka SPPT akan diserahkan kepada wajib pajak itu sendiri. Sedangkan jika wajib Pajak belum melakukan pembayaran PBB, SPPT masih disimpan di Kantor Pemerintah Desa Kendal. Hal ini yang menentukan pelabelan pada kolom STATUS dengan menggunakan kriteria 'LUNAS' bagi wajib pajak yang sudah membayar PBB dan kriteria 'MENUNGGAK' bagi wajib pajak yang belum melakukan pembayaran PBB.

Tabel 1. Data Wajib Pajak PBB

NOP	NAMA WAJIB PAJAK	KETETAPAN PBB 2021	ALAMAT OBJEK PAJAK	STATUS	SPPT
32.11.080.006.001.0001.0	Tuni'ah	18.000	Pon	Lunas	Tidak Ada
32.11.080.006.001.0148.0	Sudira	23.534	Manis	Lunas	Tidak Ada
32.11.080.006.001.0184.0	Tarimah	18.000	Manis	Menunggak	Ada
32.11.080.006.001.0479.0	Mualimin, H	52.682	Wage	Lunas	Tidak Ada
32.11.080.006.001.0480.0	Sukardi	25.460	Wage	Menunggak	Ada
32.11.080.006.001.0048.0	Ru'ati	18.000	Pon	Menunggak	Ada
32.11.080.006.001.0403.0	Rubaiah	99.044	Pahing	Lunas	Tidak Ada

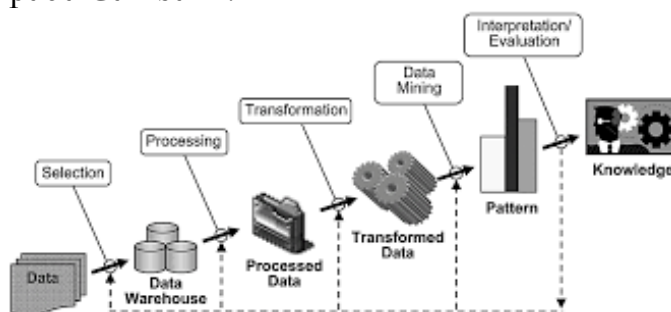
Sumber: Kantor Pemerintah Desa Kendal

METODE

Adapun teknik pengumpulan data dalam penelitian ini yaitu :

1. Observasi; objek penelitian ini adalah Pemerintah Desa Kendal. Jenis data yang digunakan adalah data primer, yang dimaksud data primer adalah data yang didapat langsung oleh kantor Pemerintah Desa Kendal dengan menggunakan metode wawancara, studi dokumentasi dan Observasi.
2. Wawancara; pada metode penelitian ini peneliti melakukan wawancara dengan Sekretaris Desa selaku Kordinator Pajak Bumi dan Bangunan Desa Kendal.
3. Studi Dokumentasi; adapun penelitian ini dokumen yang digunakan yaitu buku, jurnal terkreditasi dan data pendukung.

Dalam melakukan analisis dan pembahasan terhadap masalah yang terjadi, maka penelitian ini menggunakan data kuantitatif. Sumber data yang digunakan adalah data sekunder yaitu data yang berasal dari kantor Pemerintah Desa Kendal Kecamatan Astanajapura Kabupaten Cirebon. Proses pengambilan data yang akan digunakan dalam *data mining* dilakukan secara bertahap. Setelah mendapatkan data mentah kemudian akan dilakukan seleksi dan diolah menjadi sebuah informasi atau benang merah. Metode *Knowledge discovery in database* (KDD) adalah salah satu proses yang menggunakan *data mining* sebagai alat untuk mengekstrak pengetahuan yang dianggap sesuai dengan spesifikasi ukuran dan batas, menggunakan *database* secara bersamaan dengan *preprocessing* yang diperlukan dan pengambilan sampel serta informasi dari *database* [7]. Berikut tahapan proses KDD yang akan tunjukan pada Gambar 1.



Gambar 1. Tahapan Penelitian [7]

Adapun penjelasan tahapan penelitian yaitu :

1. *Data Selection* (pemilihan data), seringkali data yang diperoleh tidak semua digunakan, oleh karena itu hanya data yang sesuai dengan kebutuhan analisis yang akan diambil dari database.
2. *Data Cleaning*, pada tahap awal *data mining* ini, data mentah yang telah diperoleh akan dibersihkan dari kesalahan atau tidak lengkapnya kriteria data.
3. *Transformasi Data* (transformasi data), tahap ini dilakukan dengan menempatkan data yang relevan ke dalam prosedur *data mining*.
4. *Data Mining*, tahap utama dari proses pengumpulan data adalah *data mining*, yang merupakan identifikasi persyaratan umum yang disepakati untuk mengambil teknik atau metode tertentu.
5. Interpretasi / Evaluasi, model informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

Data mining menurut Han Jiawei dan M. Kamber adalah proses mengekstraksi pola-pola yang menarik (implisit, tidak diketahui sebelumnya dan berpotensi untuk dimanfaatkan) dari data yang berukuran besar [8]. Proses *data mining* adalah salah satu alat yang digunakan untuk menemukan pola dan hubungan antar data. Pola dan hubungan ini sebelumnya belum diketahui karena berada dalam kumpulan data yang amat besar. Alat-alat ini yang nantinya dapat menggabungkan model statistik, teknik machine learning, dan algoritma.

Clustering merupakan salah satu metode *data mining* yang bersifat tanpa arahan (*unsupervised*). Terdapat dua jenis data *clustering* yang sering digunakan dalam proses pengelompokan data yaitu *hierarchical* (hirarki) data *clustering* dan *non-hierarchical* (non hirarki) data *clustering* [9].

K-Means clustering merupakan metode yang dikembangkan oleh Mac Queen pada tahun 1967. *K-means* adalah metode pengklasteran secara *partitioning* yang memisahkan data ke dalam kelompok yang berbeda. Dengan *partitioning* secara iterasi, *K-means* mampu meminimalkan rata-rata jarak setiap data ke klasternya [5] Berikut Langkah-langkah yang terdapat pada algoritma *K-means* [10] :

1. Tentukan jumlah kluster (k), tetapkan pusat kluster secara acak
2. Hitung jarak setiap data ke pusat cluster
3. Kelompokkan data ke dalam kluster dengan jarak yang paling pendek
4. Hitung pusat kluster baru
5. Ulangi Langkah 2 (dua) sampai 4 (empat) sehingga tidak ada lagi data yang pindah ke cluster lain.

HASIL DAN PEMBAHASAN

Data Selection

Pada tahap awal, sumber data yang telah diperoleh akan dipisahkan antara data yang akan digunakan dan data yang tidak digunakan. Data yang digunakan hanya data yang sesuai dengan variabel perhitungan yang digunakan untuk dianalisis.

	A	B	C	D	E
1	NOP	KETETAPAN PBB 2021	ALAMAT OBJEK PAJAK	STATUS	SPPT
2	32.11.080.006.001.0001.0	18.000	PON	LUNAS	TIDAK ADA
3	32.11.080.006.001.0002.0	19.254	PON	LUNAS	TIDAK ADA
4	32.11.080.006.001.0003.0	18.000	PON	LUNAS	TIDAK ADA
5	32.11.080.006.001.0004.0	18.000	PON	LUNAS	TIDAK ADA
6	32.11.080.006.001.0005.0	18.000	PON	LUNAS	TIDAK ADA
7	32.11.080.006.001.0006.0	18.000	PON	LUNAS	TIDAK ADA
8	32.11.080.006.001.0007.0	31.472	PON	LUNAS	TIDAK ADA
9	32.11.080.006.001.0008.0	18.000	PON	LUNAS	TIDAK ADA
10	32.11.080.006.001.0009.0	24.320	PON	LUNAS	TIDAK ADA
11	32.11.080.006.001.0010.0	18.000	PON	LUNAS	TIDAK ADA
12	32.11.080.006.001.0011.0	18.000	PON	LUNAS	TIDAK ADA
13	32.11.080.006.001.0012.0	18.000	PON	MENUNGGAK	ADA
14	32.11.080.006.001.0013.0	25.834	PON	LUNAS	TIDAK ADA
15	32.11.080.006.001.0014.0	24.734	PON	LUNAS	TIDAK ADA
16	32.11.080.006.001.0015.0	18.000	PON	LUNAS	TIDAK ADA
17	32.11.080.006.001.0016.0	27.238	PON	LUNAS	TIDAK ADA
18	32.11.080.006.001.0017.0	18.000	PON	LUNAS	TIDAK ADA
19	32.11.080.006.001.0018.0	18.000	PON	LUNAS	TIDAK ADA
20	32.11.080.006.001.0020.0	71.694	PON	MENUNGGAK	ADA

Gambar 2. Data yang akan digunakan

Pada penelitian ini, atribut yang sesuai dengan proses yang akan dilakukan hanyalah Nomor Objek Pajak (NOP) dan ketentuan PBB 2021. Sedangkan Alamat objek pajak, keberadaan SPPT, serta atribut Status yang digunakan sebagai akurasi untuk prediksi potensi nanti adalah atribut yang ditambahkan oleh peneliti untuk menunjang pemodelan data nanti.

Data Cleaning

Dengan kemungkinan masih adanya data yang kosong, ganda dan missing, dilakukannya proses *cleaning* masih sangat dibutuhkan. Tujuannya agar data yang tidak relevan itu bisa terhapus karena keberadaannya dapat mengurangi tingkat akurasi dari hasil *data mining* nantinya.

Row No.	NOP	KETETAPAN	ALAMAT OB.	STATUS	SPPT
1	32.11.080.00...	18000	PON	LUNAS	TIDAK ADA
2	32.11.080.00...	19254	PON	LUNAS	TIDAK ADA
3	32.11.080.00...	18000	PON	LUNAS	TIDAK ADA
4	32.11.080.00...	18000	PON	LUNAS	TIDAK ADA
5	32.11.080.00...	18000	PON	LUNAS	TIDAK ADA
6	32.11.080.00...	18000	PON	LUNAS	TIDAK ADA
7	32.11.080.00...	31472	PON	LUNAS	TIDAK ADA
8	32.11.080.00...	18000	PON	LUNAS	TIDAK ADA
9	32.11.080.00...	24320	PON	LUNAS	TIDAK ADA
10	32.11.080.00...	18000	PON	LUNAS	TIDAK ADA
11	32.11.080.00...	18000	PON	LUNAS	TIDAK ADA
12	32.11.080.00...	18000	PON	MENUNGGAK	ADA
13	32.11.080.00...	25834	PON	LUNAS	TIDAK ADA
14	32.11.080.00...	24734	PON	LUNAS	TIDAK ADA
15	32.11.080.00...	18000	PON	LUNAS	TIDAK ADA
16	32.11.080.00...	27238	PON	LUNAS	TIDAK ADA
17	32.11.080.00...	18000	PON	LUNAS	TIDAK ADA

Gambar 3. Hasil data cleaning

Gambar 3 merupakan hasil dari pembersihan data terhadap data kosong dan data *missing*. hal ini dihasilkan dari mengaktifkannya *Replace errors with missing values* pada saat proses memasukan data kedalam *rapidminer*. Pada kenyataannya, data PBB Kendal tahun 2021 tidak memiliki data yang kosong dan *missing*, sehingga data wajib pajak tetap sebanyak 1.159 data.

Data Transformation

Pada tahapan ini, data yang terpilih dan dibersihkan selanjutnya akan ditransformasikan kedalam bentuk-bentuk yang sesuai dengan algoritma yang digunakan. Prosedur perubahan data yang dilakukan pada penelitian ini memiliki dua tahapan. Tahap pertama adalah mengubah variabel yang digunakan menjadi numerik dengan menggunakan operator *Nominal to Numerical* karena bentuk data yang di dapat masih bervariasi nominal. Tahap ini merupakan tahap yang penting, karena *K-Means* adalah algoritma yang hanya bisa menggunakan data yang bervariasi numerik.

Row No.	NOP	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.
1	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
2	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
3	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
4	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
5	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
6	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
7	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
8	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
9	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
10	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
11	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
12	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
13	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
14	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
15	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
16	32.11.080.00..	1	0	0	0	0	0	0	0	0	0

Gambar 4. Hasil transformasi data

Tahap kedua adalah normalisasi data dengan menggunakan operator *Normalize*. Tahap ini bertujuan untuk memperkecil jarak antar nominal ketetapan PBB, Hal ini dilakukan karena besarnya selisih jumlah ketetapan PBB sehingga jika jarak tersebut terlalu besar, hasil dari proses pemodelan ini akan menjadi kurang optimal. Pada proses normalisasi menggunakan *rapidminer* ini, metode transformasi yang digunakan adalah metode *range transformation* yaitu metode dimana nilai jarak yang diinginkan bisa ditentukan. Pada penelitian ini parameter yang digunakan adalah nilai *min* = 0.0 dan nilai *max* = 1.0.

Row No.	NOP	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.	ALAMAT OB.
1	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
2	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
3	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
4	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
5	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
6	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
7	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
8	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
9	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
10	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
11	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
12	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
13	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
14	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
15	32.11.080.00..	1	0	0	0	0	0	0	0	0	0
16	32.11.080.00..	1	0	0	0	0	0	0	0	0	0

Gambar 5. Hasil normalisasi

Data Mining

Pada tahapan ini dilakukannya pemodelan data menggunakan hasil dari transformasi yang sudah dilakukan. Adapun metode yang digunakan dalam penelitian ini adalah metode *clustering* dengan menerapkan algoritma *k-means* dimana diketahui nilai *k* = 4, *max runs* sebanyak 60 kali putaran dan jenis *measure*

types nya adalah NumericalMeasures dengan menggunakan metode perhitungan EuclideanDistance dan max optimization steps nya berjumlah 100.

Row No.	NOP	cluster	ALAMAT OB...	ALAMAT OB...	ALAMAT OB...	ALAMAT OB...	ALAMAT OB...	ALAMAT OB...	ALAMAT OB...	ALAMAT OB...	ALAMAT OB...
1	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
2	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
3	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
4	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
5	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
6	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
7	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
8	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
9	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
10	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
11	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
12	32.11.080.00...	cluster_0	1	0	0	0	0	0	0	0	0
13	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
14	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
15	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0
16	32.11.080.00...	cluster_3	1	0	0	0	0	0	0	0	0

Gambar 6. Hasil Clustering

Interpretasi

Pada tahap ini pola informasi yang telah dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

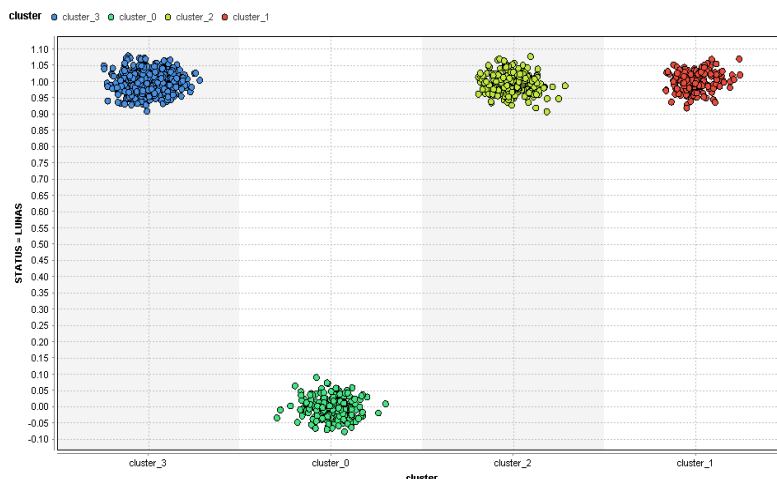
Tabel 2. anggota cluster berdasarkan penyebutan Tingkat Kepatuhan

	Menunggak	Lunas	Tingkat Kepatuhan
Cluster_0	222	146	rendah
Cluster_1		154	Sedang
Cluster_2		246	tinggi
Cluster_3		391	sangat tinggi

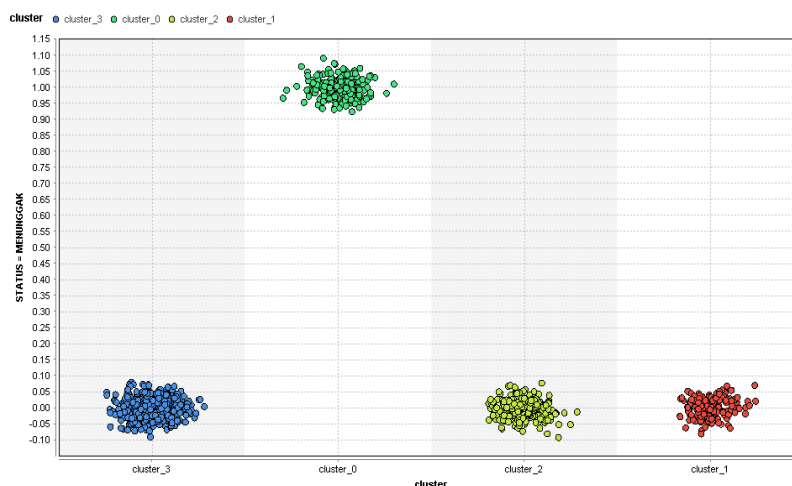
Setelah melakukan 4 (empat) percobaan untuk menentukan nilai k, Pada Percobaan ketiga dengan menggunakan nilai k yang ditentukan adalah 4 menghasilkan nilai Performance Vektor (Davies Bouldin Index) yang lebih optimal dibandingkan dengan percobaan pertama dan kedua. Nilai yang dihasilkan yaitu nilai yang mendekati 0 dengan nilai -0.862.

Tabel 3. Percobaan dalam menentukan nilai k

Nilai k	Average (avg)	Davies Bouldin Index (DBI)
2	-0.843	-0.899
3	-0.710	-0.872
4	-0.534	-0.862
5	-0.461	-1.028



Gambar 7. Scatter Plot status Lunas

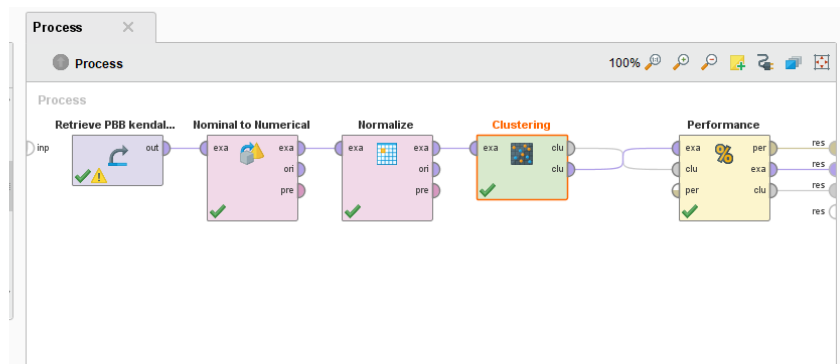


Gambar 8. Scatter Plot status Menunggak

Setelah melakukan pemodelan *clustering*, maka tahap selanjutnya adalah melakukan validasi *clustering* dengan menggunakan operator *cluster distance performance* agar diketahui nilai DBI (*Davies Bouldin Index*). Hal ini memiliki tujuan untuk memaksimalkan pengukuran jarak *cluster* satu dengan *cluster* yang lain. Pada operator ini parameter *main criterion* nya diubah menjadi *Davies Bouldin Index*. *Davies Bouldin Index* (DBI) adalah metode yang digunakan untuk mengevaluasi *cluster* yang dikenalkan oleh David L. Davies dan Donald W. Bouldin. DBI adalah suatu metode yang digunakan untuk mengukur validitas *cluster* dalam metode *clustering*. Pengukuran dalam DBI dapat memaksimalkan jarak antara *cluster* C_i dan C_j dan pada saat yang sama mencoba untuk meminimalkan jarak antara titik-titik dalam *cluster* [10].

Keanggotaan *cluster* dikelompokkan berdasarkan kemiripannya, percobaan yang digunakan hanya percobaan dengan nilai 4 *cluster*. Dari pengamatan hasil *clustering*, diperoleh bahwa wajib pajak yang menempati *cluster* 0 memiliki kepatuhan dalam membayar pajak yang rendah dengan memiliki 222 Wajib Pajak yang menunggak dan 146 wajib pajak yang lunas, *cluster* 0 ini didominasi oleh blok pahing dengan 68 wajib pajak yang menunggak. Tingkat kesadaran wajib pajak pada *cluster* 1 bisa dikatakan sedang, artinya jumlah wajib pajak yang Lunas dalam

cluster ini mengalami kenaikan dalam hal wajib pajak yang Lunas dan tidak memiliki wajib pajak yang menunggak, anggota dari cluster 1 ini didominasi oleh blok wage dengan jumlah wajib pajak yang sudah membayar pajak sebanyak 154 wajib pajak.



Gambar 9. Proses Pengujian

Cluster 2 bisa disebut sebagai cluster yang memiliki tingkat kepatuhan pajak yang tinggi. Hal ini disebabkan oleh meningkatnya status lunas wajib pajak dari cluster sebelumnya yaitu sebanyak 246 anggota. Cluster ini didominasi oleh blok manis dengan tidak memiliki anggota cluster yang menunggak. Cluster terakhir yang dihasilkan adalah cluster 3. Cluster ini bisa disebut sebagai cluster dengan tingkat kepatuhannya sangat tinggi, hal ini disebabkan karena status Lunas terhadap pembayaran iuran PBB pada cluster ini memiliki 391 anggota dan tidak memiliki anggota menunggak dengan didominasi oleh Blok Pahing sebanyak 151 anggota.

Untuk mengetahui kisaran harga ketetapan PBB pada setiap klaster, peneliti menggunakan tools dari Microsoft excel yaitu Filter kemudian menggunakan rumus persentase berikut ini:

$$\text{Persentase (\%)} = \frac{\text{Jumlah bagian}}{\text{Jumlah total}} \times 100\%$$

Tabel 4. jumlah anggota cluster

	Cluster_0	Cluster_1	Cluster_2	Cluster_3
PON	37	-	-	142
MANIS	17	-	246	-
PAHING	69	-	-	151
WAGE	38	154	-	-
TRIKEM	46	-	-	106
PUTAT	5	-	-	27
SILAWUNG	6	-	-	29
SIMUBAL	3	-	-	34
GEMPOL	4	-	-	19
CIBUK	1	-	-	25
TOTAL =	226	154	246	533

Setelah melakukan perhitungan, dapat ditemukan bahwa kisaran harga ketetapan PBB pada cluster 0 adalah Rp. 18.000 dengan memiliki 174 anggota dari

226 anggota *cluster 0* serta jumlah persentasenya sebesar 77%. Pada *cluster 1* memiliki 117 dari 154 Wajib Pajak dengan harga ketetapan PBB nya adalah Rp. 18.000,- dan persentase mencapai 76%. Dari 246 anggota *cluster 2*, 187 anggota lainnya adalah Wajib Pajak yang memiliki nilai ketetapan PBB Rp.18.000,- yang kemudian memperoleh persentase sebesar 76%. Pada *cluster 3* juga ditemukan bahwa Rp. 18.000,- adalah harga ketetapan PBB yang paling banyak dimiliki Wajib Pajak, pada *cluster* ini persentase yang dihasilkan adalah 44,3%. Maka, dapat disimpulkan bahwa baik *cluster 0*, *cluster 1*, *cluster 2* maupun *cluster 3* memiliki kesamaan dalam hal dominasi harga ketetapan PBB 2021 yaitu Rp. 18.000,-.

PENUTUP

Berdasarkan hasil analisis dan pembahasan diatas, dapat disimpulkan bahwa dengan menggunakan metode *K-Means*, data Wajib Pajak Bumi dan Bangunan Desa Kendal pada tahun 2021 dapat dikelompokkan menjadi 4 *cluster*. Blok yang memiliki tingkat kepatuhan PBB yang rendah adalah Blok Pahing yang beranggotakan 68 Wajib Pajak yang menunggak. Sedangkan rata-rata harga ketetapan PBB yang paling mendominasi pada setiap *cluster* adalah Rp. 18.000,-, dengan persentase harga ketetapan PBB pada *cluster 0* sebesar 77%, *cluster 1* 76%, *cluster 2* 76% dan *cluster 3* sebesar 44,3%. Dengan bantuan *software* RapidMiner yang digunakan selama penelitian ini, maka didapat sebuah knowledge:

1. *Cluster 0* adalah *cluster* dengan tingkat kepatuhan pajaknya masih rendah. Pada *cluster* ini jumlah anggota Wajib Pajak yang menunggak lebih banyak dibandingkan dengan jumlah Wajib Pajak yang lunas yaitu 4 anggota Wajib Pajak yang Lunas dan 222 anggota Wajib Pajak yang Menunggak.
2. *Cluster 1* adalah *cluster* yang memiliki tingkat kepatuhan Pajak Bumi dan Bangunan yang sedang, karena pada *cluster* ini memiliki 154 Wajib Pajak yang Lunas dan tidak memiliki anggota Wajib Pajak yang menunggak.
3. *Cluster 2* merupakan *cluster* dengan tingkat kepatuhan yang tinggi, hal ini dikarenakan jumlah Wajib Pajak yang Lunas mencapai 246 anggota.
4. Sedangkan *cluster 3* adalah *cluster* dengan tingkat kepatuhan PBB yang paling tinggi dengan memiliki 533 anggota.

Adapun saran yang dapat diberikan setelah melakukan penelitian adalah:

1. Penelitian selanjutnya diharapkan dapat menambahkan variable penentu lainnya sehingga wajib pajak dapat dikategorikan sebagai wajib pajak yang lunas dalam membayar iuran PBB.
2. Untuk selanjutnya diharapkan penelitian ini dapat diterapkan dengan menggunakan algoritma yang lain agar dapat mengembangkan penelitian dimasa yang akan datang.
3. Bagi Pemerintah Desa Kendal diharapkan dapat meningkatkan sosialisasi tentang pentingnya membayar PBB kepada masyarakatnya sehingga masyarakat Desa Kendal tersebut mengetahui waktu pembayaran pajak serta dapat meningkatkan kesadaran untuk membayar pajak khususnya kepada blok yang tingkat kepatuhan pajaknya masih rendah.

DAFTAR PUSTAKA

- [1] Undang-Undang No 28 Tahun 2009 tentang Pajak Daerah dan Retribusi Daerah.
- [2] Mardiasmo, *Perpajakan*, Edisi Terbaru 2018. Yogyakarta, 2018.

- [3] A. Rohendi, "Fungsi Budgeter dan Fungsi Regulasi dalam Ketentuan Perpajakan Indonesia," Apr. 2014.
- [4] R. A. Kamaroellah, "Analisis Kepatuhan Wajib Pajak Bumi dan Bangunan Berdasarkan Realisasi Penerimaan Pajak Bumi dan Bangunan (PBB) Pada Dinas Pendapatan Daerah Kabupaten Pamekasan," *Jurnal Ekonomi dan Perbankan Syariah*, vol. 4 No. 1 Juni, no. 1, 2017.
- [5] M. Nanda Variestha Waworuntu and M. Faisal Amin, "Penerapan metode K-Means Pemetaan calon Penerima JAMKESDA," vol. 05, p. 190, 2018.
- [6] M. S. Yana *et al.*, "Penerapan Metode K-Means dalam Pengelompokan Wilayah Menurut Intensitas Kejadian Bencana Alam di Indonesia Tahun 2013-2018," 2018.
- [7] Q. Widayati, "Penerapan Data Mining Menggunakan Metode Teknik Classification... ..(Qoriani Widayati)," *Jurnal Ilmiah MATRIk*, vol. Vol.20 No.2, no. PBB, Data Mining, 2018, [Online]. Available: www.pajak.go.id
- [8] A. Sani, J. Teknologika, and E. Muningsih, "Penerapan Metode K-Means Clustering pada Perusahaan."
- [9] R. Rosmini, A. Fadlil, and S. Sunardi, "Implementasi Metode K-Means Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah," *IT Journal Research Development*, vol. 3, no. 1, pp. 22-31, Aug. 2018, doi: 10.25299/itjrd.2018.vol3(1).1773.
- [10] F. Farahdinna, I. Nurdiansyah, A. Suryani, and A. Wibowo, "Perbandingan Algoritma K-Means dan K-Medoids dalam Klasterisasi Produk Asuransi Perusahaan Nasional," *Jurnal Ilmiah FIFO*, vol. 11, no. 2, p. 208, Nov. 2019, doi: 10.22441/fifo.2019.v11i2.010.