

Penerapan Data Mining Menggunakan Metode Teknik Classification Untuk Melihat Potensi Kepatuhan Wajib Pajak Bumi Dan Bangunan

¹Arif Rahmat Shaumi, ²Muhammad Faridz Ali, ³M Tsaqif Al Mutawakkil Simbolon

^{1,2,3}Program Studi Ilmu Komputer, Universitas Islam Negeri Sumatera Utara, Medan

¹arifrahmat220@gmail.com, ²farisali0305@gmail.com, ³mhdtsaqif11@gmail.com

Abstract: *The government implements development in Indonesia, requires substantial funds. The entry of cash from the Land and Building Tax is the most important part for the development of a region, with the results that have been obtained by the regional government can increase regional development with various infrastructures that help the community to carry out various activities and make the area more advanced. One type of tax is the Land and Building Tax (PBB). With the increasing number of taxpayers and data paying contributions directly into the treasury of state finances, the UPT BPPRD Subdistrict of Medan city did not know how many obedient and non-compliant taxpayers. In this study using data mining techniques, namely classification by applying the Naive Bayes algorithm and getting from the number of taxpayers as many as 1,647 taxpayers with an accuracy of 99.33% which have the potential not to be timely are found in the Medan sandpaper sub-district of 0.437 and the Medan Area sub-district with data of 0.229*

Keywords: *United Nations, Data Mining, taxpayers*

Abstrak: *Pemerintah melaksanakan pembangunan di Indonesia ini tentu memerlukan dana yang cukup besar. Pemasukan kas dari Pajak Bumi dan/atau Bangunan merupakan bagian terpenting untuk pembangunan suatu daerah, dengan hasil yang telah didapatkan pemerintah daerah dapat meningkatkan pembangunan daerah dengan berbagai infrastruktur yang membantu masyarakat dalam melakukan berbagai aktivitas serta menjadikan daerah tersebut menjadi lebih maju. Salah satu jenis pajak adalah Pajak Bumi dan Bangunan (PBB). Dengan semakin banyaknya wajib pajak dan data yang membayar iuran langsung masuk ke kas keuangan negara menyebabkan pihak UPT BPPRD kota Medan tidak mengetahui seberapa banyak wajib pajak yang patuh dan tidak patuh. Pada penelitian ini menggunakan teknik data mining yaitu classification dengan menerapkan algoritma naive bayes dan mendapatkan dari jumlah wajib pajak sebanyak 1.647 wajib pajak dengan akurasi sebesar 99.33% yang berpotensi tidak tepat waktu terdapat pada Kecamatan medan amplas sebesar 0.437 dan kecamatan medan area dengan data sebesar 0.229*

Kata kunci: *PBB, Data Mining, wajib pajak*

PENDAHULUAN

Pemerintah melaksanakan pembangunan di Indonesia ini tentu memerlukan dana yang cukup besar. Pemasukan kas dari Pajak Bumi dan/atau Bangunan merupakan bagian terpenting untuk pembangunan suatu daerah, dengan hasil yang telah didapatkan pemerintah daerah dapat meningkatkan pembangunan daerah dengan berbagai infrastruktur yang membantu masyarakat dalam melakukan berbagai aktivitas serta menjadikan daerah tersebut menjadi lebih maju. Pajak Bumi dan Bangunan (PBB) adalah pajak negara yang dikenakan terhadap bumi dan bangunan berdasarkan undang-undang nomor 12 tahun 1985 tentang pajak bumi dan bangunan (PBB) adalah pajak yang bersifat kebendaan dalam arti besarnya pajak terutang ditentukan oleh keadaan objek yaitu bumi/tanah dan atau bangunan. Keadaan subjek (siapa yang membayar) tidak ikut menentukan besarnya pajak (www.pajak.go.id).

Kepatuhan yang tinggi dari wajib pajak adalah suatu keadaan dimana wajib pajak memenuhi semua kewajibannya dalam membayar pajak, namun dalam realitanya tidak semua bahkan rendahnya tingkat wajib pajak patuh dalam membayar pajak. Dinas Pendapatan Daerah adalah suatu instansi yang bertujuan untuk mengumpulkan penerimaan Pendapatan Asli Daerah (PAD) yang



bersumber dari pajak daerah lainnya untuk pembangunan daerah. UPT Penyuluhan mempunyai tugas pokok melaksanakan sebagian fungsi Badan di Bidang Penyuluhan dan kehumasan dibidang pajak dan retribusi daerah pada lingkup Badan. UPT Penyuluhan menyelenggarakan fungsi. Penyelenggaraan pengkajian bahan petunjuk teknis Penyuluhan dan penyebarluasan informasi di Bidang Pajak dan retribusi daerah. Penyelenggaraan penyuluhan dan penyebarluasan informasi di bidang pajak dan retribusi daerah

Adapun Visi Badan Pengelola Pajak dan Retribusi Daerah Kota Medan adalah Dengan Terwujudnya Pendapatan Daerah Sebagai Andalan Pembiayaan Pembangunan Daerah. Serta memiliki Misi Badan Pengelola Pajak dan Retribusi Daerah Kota Medan adalah

- a. Meningkatkan kualitas pelayanan terhadap sumber dan pengelola Pendapatan Daerah,
- b. Meningkatkan sarana dan prasarana Badan.
- c. Intensifikasi dan ekstensifikasi subyek dan obyek Pendapatan Daerah,
- d. Meningkatkan penegakan hukum, meningkatkan kesadaran wajib pajak terhadap kewajibannya dalam membayar pajak daerah.

Dengan melihat kemajuan infrastruktur kota Medan yang semakin maju menjadikan pajak bumi dan bangunan (PBB) sebagai pendapatan asli daerah (PAD) yang paling dominan, namun yang menjadi masalahnya dengan menjadikan PBB sebagai salah satu penyumbang terbesar PAD yang menyebabkan peningkatan jumlah wajib pajak bumi dan bangunan semakin terus mengalami kenaikan, sehingga secara otomatis membuat masyarakat menjadi hilang kesadaran dan membuat menurunnya kepatuhan wajib pajak dalam membayar iuran tersebut.

Dengan semakin meningkatnya jumlah wajib pajak dan sistem pembayaran melalui bank dengan iuran yang langsung masuk ke data kas keuangan negara menyebabkan pihak UPT BPPRD Kota Medan tidak mengetahui seberapa banyak wajib pajaknya yang patuh dan seberapa banyak wajib pajak yang tidak patuh dalam membayar iuran pajak bumi dan bangunan (PBB) tersebut. Dari hal tersebut dilihat manfaat data mining adalah untuk menangani banyaknya data. Untuk melihat kepatuhan WP digunakan teknik klasifikasi untuk menentukan potensi kepatuhan wajib pajak (WP). Data Mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat pada basis data. Data mining terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery Database (KDD)* (Vulandari, 2017).

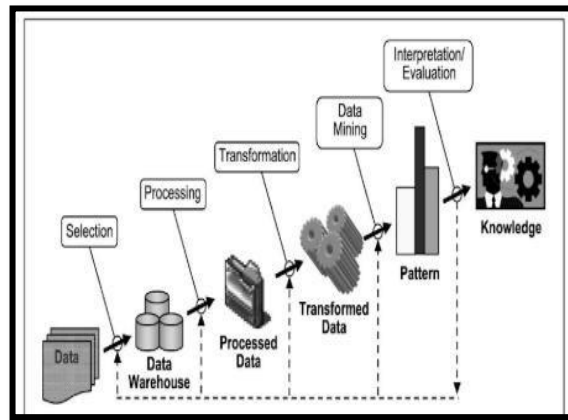
METODOLOGI PENELITIAN

Metode yang digunakan dalam penelitian ini adalah metode deskriptif. Metode deskriptif adalah suatu metode dalam meneliti suatu objek, suatu metode yang mengemukakan masalah dengan mengumpulkan data-data yang bertujuan untuk mendapatkan gambaran yang jelas mengenai suatu keadaan dengan cara menyajikan. penelitian deskriptif yaitu, penelitian yang dilakukan untuk mengetahui nilai variabel mandiri, baik satu variabel atau lebih (independen) tanpa membuat perbandingan, atau menghubungkan dengan variabel yang lain (Sugiyono, 2012).

Pada metode analisis akan dimining dengan melalui tahapan *knowledge discovery in database (KDD)*. Proses KDD adalah proses menggunakan data mining untuk mengekstrak pengetahuan apa yang dianggap sesuai dengan spesifikasi ukuran dan batas, menggunakan *database* bersama dengan *preprocessing* yang diperlukan, pengambilan sampel dan transformasi dari *database*.

Adapun untuk menganalisis data dalam penerapan data mining ini menggunakan tahapan *Knowledge Discover in Database (KDD)*. istilah data mining dan *knowledge discovery in database (KDD)* sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi yang tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan dengan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah data mining (Nofriansyah, 2014).





Gambar 1. Proses KDD

Berdasarkan gambar di atas, proses KDD secara garis besar dijelaskan sebagai berikut:

a. *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, kata memilih data-data seperti apa saja yang kita butuhkan untuk proses lebih lanjut dan kemudian data disimpan dalam suatu berkas, terpisah dari basis data operasional sehingga memberikan kemudahan untuk penggunaan berikutnya.

b. *Pre-processing (Cleaning)*

Pada umumnya data yang diperoleh, baik dari database suatu perusahaan maupun eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut data yang tidak relevan itu juga lebih baik dibuang karena keberadaannya bisa mengurangi mutu atau akurasi dari hasil data mining nantinya.

c. *Data Transformation*

Beberapa teknik data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa teknik standar seperti analisis asosiasi dan klustering hanya bisa menerima *input* kategorikal. Karenanya data berupa angka *numerik* yang berlanjut bisa dibagi-bagi menjadi beberapa interval. Proses ini sering disebut *binning*. Disini juga dilakukan pemilihan data yang diperlukan oleh teknik data mining yang dipakai. Transformasi dan pemilihan data ini juga menemukan kualitas dari hasil data mining nantinya karena ada beberapa karakteristik dari teknik-teknik data mining tertentu tergantung pada tahap ini.

d. *Data Mining*

Data mining adalah proses menjadi pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat tergantung pada tujuan dan proses KDD secara keseluruhan. Data yang digunakan untuk bisa menjadi sebuah model yang baik idealnya mencukupi sebagai data riset. Semakin banyak data dan semakin sedikit kesalahan (*error*) maka semakin bagus model yang dijadikan patokan.

e. *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 disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta pola atau informasi yang ada sebelumnya.

Data yang didapatkan dari UPT BPPRD kota Medan yaitu data PBB tahun 2017 pada kecamatan Medan Amplas Palembang dengan keseluruhan data yang telah membayar iuran PBB sebanyak 5.284 WP. Sebelum dilakukannya proses KDD data akan dihapuskan jumlah tagihan dibawah Rp.100.000 sesuai dengan peraturan pemerintah kota Medan tahun 2018 dengan di tidakkannya iuran untuk biaya dibawah Rp.100.000. Proses penghapusan data tersebut akan

dilakukan menggunakan *software pentaho* yang akan dijelaskan selanjutnya, setelah data tidak mengandung tagihan dibawah Rp. 100.000 data tersebut akan diproses pada tahapan KDD dengan tahapan yaitu seleksi data, *pre-processing* data, transformasi data, data mining dan evaluasi.

HASIL DAN PEMBAHASAN

Adapun untuk menganalisis data dalam penerapan data mining ini menggunakan tahapan *Knowledge Discover in Database (KDD)* yang sebelumnya akan dipisahkan menggunakan *tools Pentaho*, dengan tahapan sebagai berikut:

- 1) *Remove Data* \leq Rp. 100.000 menggunakan *Pentaho*
- 2) *Data Selection*
- 3) *Pre-Processing(Data Cleaning)*
- 4) *Transformation Data*
- 5) *Data Mining*
- 6) *Evaluation*

3.1 Remove Data Tagihan

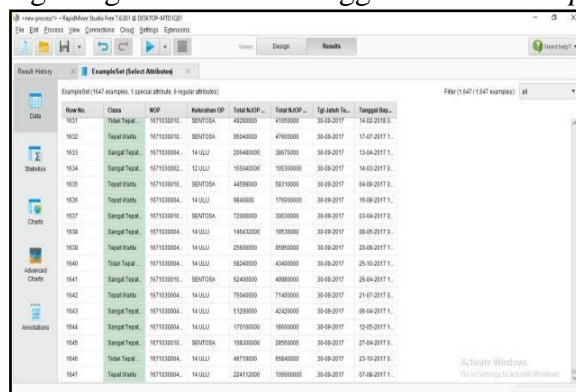
Data WP yang telah ada dipisahkan menggunakan *tools Pentaho*. *Pentaho* adalah kumpulan aplikasi *business intelligence (BI)* dan *data warehouse free open source software* yang berjalan diatas *platform java* (Putra, 2015).

Berdasarkan perwali (peraturan walikota) Medan yang baru bahwa tagihan pajak bumi dan bangunan (PBB) dibawah Rp. 100.000 akan dibebaskan pembayarannya, maka dari itu tagihan yang dibawah Rp. 100.000 akan dihapuskan dari sumber data yang ada menggunakan *software Pentaho*. Setelah dihilangkan setelah dihilangkan menggunakan *pentaho* menjadi sebanyak 1.647 data wajib pajak.

3.2 Data Selection

Classification adalah proses penemuanmodel atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya belum diketahui. (Vulandari, 2017)

Pada tahap ini *field* dari sumber data akan dipisahkan antara *field* yang tidak digunakan dan yang akan digunakan unyuk penggalian data pada tahap *pre-processing* selanjutnya untuk mendapatkan *knowledge* yang diinginkan denan menggunakan *tools rapidminer*.



Row No.	Class	NOP	Kelurahan OP	Total NJOP	Total NJOP	Tgl. Jatuh Tg.	Tanggal Bayar
1621	Tanah Perk.	1671033014	0207TSA	4820000	4100000	30-09-2017	14-10-2018
1622	Tapak Perk.	1671033014	0207TSA	3504000	4700000	30-09-2017	17-07-2017
1623	Sangat Perk.	1671033004	14 ULU	20948000	38070000	30-09-2017	13-04-2017
1624	Sangat Perk.	1671033002	12 ULU	15504000	16530000	30-09-2017	14-03-2017
1625	Tapak Perk.	1671033014	0207TSA	4428000	5010000	30-09-2017	04-08-2017
1626	Tapak Perk.	1671033004	14 ULU	9840000	17000000	30-09-2017	16-08-2017
1627	Sangat Perk.	1671033004	0207TSA	7200000	3000000	30-09-2017	03-04-2017
1628	Sangat Perk.	1671033004	14 ULU	14632000	19530000	30-09-2017	08-05-2017
1629	Tapak Perk.	1671033004	14 ULU	2580000	8990000	30-09-2017	20-09-2017
1630	Tapak Perk.	1671033004	14 ULU	5824000	4340000	30-09-2017	25-10-2017
1631	Sangat Perk.	1671033014	0207TSA	5240000	4998000	30-09-2017	28-04-2017
1632	Tapak Perk.	1671033004	14 ULU	7564000	7140000	30-09-2017	21-07-2017
1633	Sangat Perk.	1671033004	14 ULU	5120000	4242000	30-09-2017	06-04-2017
1634	Sangat Perk.	1671033004	14 ULU	17016000	1800000	30-09-2017	12-05-2017
1635	Sangat Perk.	1671033014	0207TSA	19820000	2050000	30-09-2017	27-04-2017
1636	Tanah Perk.	1671033004	14 ULU	4077800	8944000	30-09-2017	23-10-2017
1637	Tapak Perk.	1671033004	14 ULU	22412000	16000000	30-09-2017	07-08-2017

Gambar 2. Hasil Seleksi Data

Selection berfungsi untuk memilih data yang sesuai kebutuhan untuk mendapatkan pengetahuan, pada penelitian ini data yang sesuai untuk mendapatkan hasil yang diinginkan dan sesuai dengan prosesnya adalah NOP, kelurahan OP, total NJOP bumi, total NJOP bangunan, tanggal jatuh tempo, dan tanggal bayar serta ada tambahan *field* data yaitu estimasi dan class sebagai akurasi untuk prediksi potensi kepatuhan.

3.3 Pre-Processing(Data Cleaning)

Data yang telah diseleksi pada tahapan sebelumnya masih memiliki data yang kosong, ganda dan *missing*. Selain itu juga, data yang tidak relevan itu juga lebih baik dibuang karena keberadaannya bisa mengurangi akurasi dari hasil data mining nantinya. Dengan proses data yang telah diseleksi sebelumnya akan dilihat apakah pada data tersebut terdapat data yang kosong dan terdapat nilai *missing*, jika tidak terdapat data kosong dan *missing* data tersebut akan langsung diproses pada tahap tranformasi data, sedangkan jika terdapat data yang kosong dan *missing* data tersebut akan dibersihkan terlebih dahulu dengan menggunakan sebelum masuk pada tahap transformasi data.

Row No.	Class	NOP	Relevance OP	Total NJOP	Total NJOP	Tgl. Jmlh Tg.	Tanggal Bld.
1531	Tepat	187120010	30070000	41000000	41000000	30-09-2017	14-03-2017 E.
1532	Tepat	187120010	30070000	41000000	41000000	30-09-2017	17-01-2017 E.
1533	Sangat Tepat	187120004	141411	20840000	38070000	30-09-2017	13-04-2017 E.
1534	Sangat Tepat	187120002	121411	10540000	19030000	30-09-2017	14-03-2017 E.
1535	Tepat	187120010	30070000	44000000	50700000	30-09-2017	04-09-2017 E.
1536	Tepat	187120004	141411	8660000	17060000	30-09-2017	16-08-2017 E.
1537	Sangat Tepat	187120010	30070000	72000000	30030000	30-09-2017	03-04-2017 E.
1538	Sangat Tepat	187120004	141411	146432000	19530000	30-09-2017	06-05-2017 E.
1539	Tepat	187120004	141411	25000000	39000000	30-09-2017	20-09-2017 E.
1540	Tepat	187120004	141411	86240000	43400000	30-09-2017	26-10-2017 E.
1541	Sangat Tepat	187120010	30070000	62400000	48800000	30-09-2017	20-04-2017 E.
1542	Tepat	187120004	141411	75040000	71800000	30-09-2017	21-01-2017 E.
1543	Sangat Tepat	187120004	141411	51000000	42420000	30-09-2017	06-04-2017 E.
1544	Sangat Tepat	187120004	141411	171100000	18800000	30-09-2017	12-05-2017 E.
1545	Sangat Tepat	187120010	30070000	108200000	20500000	30-09-2017	27-08-2017 E.
1546	Tepat	187120004	141411	40700000	10440000	30-09-2017	23-10-2017 E.
1547	Tepat	187120004	141411	22412000	10200000	30-09-2017	07-06-2017 E.

Gambar 3. Hasil Pembersihan Data Kosong dan Missing

Gambar 3 merupakan hasil dari pembersihan data terhadap data kosong dan *missing* ternyata pada data PBB tidak terdapat data yang *missing* dan kosong maka data wajib pajak tetap sebanyak 1.647 data.

3.4 Transformation Data

Pada tahapan ini data yang telah dipilih dan dibersihkan selanjutnya akan ditransformasikan kedalam bentuk-bentuk yang cocok dan sesuai dengan algoritma yang digunakan untuk *procedure* penggalian data dengan cara melakukan normalisasi data yang akan digunakan *operator normalize* untuk mengubah data awal menjadi bentuk yang cocok untuk dilakukan penggalian data menggunakan *algoritma naive bayes*, tahap ini perlu dilakukan untuk memperbaiki nilai akurasi yang didapat pada tahap akhir nantinya, jika tidak dilakukan *normalize* data maka hasil akurasi yang akan didapat akan menjadi kecil.

Row No.	Class	NOP	Total NJOP	Total NJOP	Relevance OP	Tgl. Jmlh Tg.	Tanggal Bld.
1	Sangat Tepat	0	-0.204	-0.125	101411	30-09-2017	20-03-2017 E.
2	Sangat Tepat	0	-0.202	-0.222	101411	30-09-2017	07-08-2017 E.
3	Sangat Tepat	0	-0.211	-0.202	101411	30-09-2017	22-08-2017 E.
4	Tepat	0	-0.168	-0.279	101411	30-09-2017	28-09-2017 E.
5	Tepat	0	-0.207	-0.046	101411	30-09-2017	14-09-2017 E.
6	Sangat Tepat	0	-0.209	-0.208	10655174	30-09-2017	16-08-2017 E.
7	Tepat	0	-0.117	-0.104	101411	30-09-2017	20-04-2017 E.
8	Sangat Tepat	0	-0.168	-0.073	101411	30-09-2017	08-04-2017 E.
9	Tepat	0	-0.201	-0.276	101411	30-09-2017	25-08-2017 E.
10	Tepat	0	-0.091	-0.108	10655174	30-09-2017	16-08-2017 E.
11	Tepat	0	-0.275	-0.107	101411	30-09-2017	13-09-2017 E.
12	Sangat Tepat	0	-0.202	-0.122	101411	30-09-2017	09-09-2017 E.
13	Tepat	0	-0.228	-0.106	101411	30-09-2017	24-01-2017 E.
14	Tepat	0	-0.176	-0.226	10655174	30-09-2017	16-08-2017 E.
15	Sangat Tepat	0	-0.174	-0.108	101411	30-09-2017	08-04-2017 E.
16	Tepat	0	-0.228	-0.108	101411	30-09-2017	14-01-2017 E.
17	Tepat	0	1.372	-0.227	10655174	30-09-2017	27-08-2017 E.

Gambar 4. Hasil transformasi data

Pada gambar 4 dapat dilihat data yang telah dilakukan transformasi menjadi bentuk yang sesuai untuk dilakukan penggalian data pada tahapan selanjutnya, terdapat perubahan pada *record* dari *field* NOP, Total NJOP Bumi, dan Total NJOP Bangunan diubah menjadi bentuk yang cocok untuk dimining.

3.5 Data Mining

Pada tahap ini dilakukan 2 pengujian yaitu dengan melakukan perhitungan manual dan perhitungan dari *rapidminer* untuk mendapatkan hasil yang akurat.

1) Perhitungan Manual Algoritma *naive bayes*



Berikut perhitungan manual dari algoritma *Naive Bayes* dengan menggunakan dataset training sebanyak 200 record dengan diambil acak dari data PBB kecamatan Medan Amplas yang telah dilakukan pembuangan data dengan tagihan kurang dari Rp. 100.000. Perhitungan manual algoritma *naive bayes* dilakukan menggunakan rumus sebagai berikut:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \dots\dots\dots(1)$$

Keterangan:

P (H|X) = Probabilitas

H = Hipotesis

X = Bukti .

Sebelum dilakukan perhitungan manual harus ditentukan dulu estimasi dari setiap tanggal bayar dan tanggal jatuh tempo dengan melakukan perhitungan dan membuat peraturan untuk mendapatkan estimasi, maka aturan atau *rule* yang terbentuk adalah sebagai berikut:

- a. *If* tanggal bayar = januari 2017 *Then* Estimasi = 4
- b. *If* tanggal bayar = februari 2017 *Then* Estimasi = 5
- c. *If* tanggal bayar = maret 2017 *Then* Estimasi = 6
- d. *If* tanggal bayar = april 2017 *Then* Estimasi = 7
- e. *If* tanggal bayar = mey 2017 *Then* Estimasi = 8
- f. *If* tanggal bayar = juni 2017 *Then* Estimasi = 9
- g. *If* tanggal bayar = juli 2017 *Then* Estimasi = 10
- h. *If* tanggal bayar = agustus 2017 *Then* Estimasi = 11
- i. *If* tanggal bayar = september 2017 *Then* Estimasi = 12
- j. *If* tanggal bayar = oktober 2017 *Then* Estimasi = 13
- k. *If* tanggal bayar = november 2017 *Then* Estimasi = 14
- l. *If* tanggal bayar = desember 2017 *Then* Estimasi = 15
- m. *If* tanggal bayar = januari 2018 *Then* Estimasi = 16
- n. *If* tanggal bayar = februari 2018 *Then* Estimasi = 17
- o. *If* tanggal bayar = maret 2018 *Then* Estimasi = 18
- p. *If* tanggal bayar = april 2018 *Then* Estimasi = 19
- q. *If* tanggal bayar = mey 2018 *Then* Estimasi = 20
- r. *If* tanggal bayar = juni 2018 *Then* Estimasi = 21
- s. *If* estimasi ≤ 9 *Then* estimasi(class) = SangatTepat Waktu
- t. *If* estimasi = 10 – 12 *Then* estimasi(class) = Tepat Waktu
- u. *If* estimasi > 12 *Then* estimasi(class) = TidakTepat Waktu

Adapun dataset untuk melakukan perhitungan manual terdiri dari 200 record yang diambil acak dari data PBB kecamatan ulu yang telah dihapuskan data yang dibawah Rp. 100.000, dengan contoh data seperti padatabel 1.

Tabel 1. Dataset

NOP	KELURAHAN OP	TGL JATUH TEMPO	TGL BAYAR	ESTIMASI
167103000500430030	16 ULU	30-09-2017	05-04-2017 09:43:08	SangatTepatWaktu
167103000300200570	13 ULU	30-09-2017	30-03-2017 12:17:31	SangatTepatWaktu
167103000300201200	13 ULU	30-09-2017	13-03-2017 09:13:13	SangatTepatWaktu
167103001000202380	SENTOSA	30-09-2017	31-03-2017 11:07:01	SangatTepatWaktu
167103001000203230	SENTOSA	30-09-2017	22-05-2017 08:50:58	SangatTepatWaktu
167103000501311450	16 ULU	30-09-2017	04-10-2017 12:14:56	TidakTepatWaktu
167103000400401220	14 ULU	30-09-2017	10-05-2017 11:26:25	SangatTepatWaktu
167103000500902900	16 ULU	30-09-2017	07-06-2017 11:08:53	SangatTepatWaktu
167103000500201740	16 ULU	30-09-2017	22-06-2017 10:59:30	SangatTepatWaktu
167103000500900070	16 ULU	30-09-2017	26-09-2017 10:27:49	TepatWaktu
167103000500203440	16 ULU	30-09-2017	14-08-2017 09:51:20	TepatWaktu
167103001000102170	SENTOSA	30-09-2017	03-04-2018 08:05:57	SangatTepatWaktu
167103000500201380	16 ULU	30-09-2017	12-05-2017 10:21:59	SangatTepatWaktu
167103000900601290	TANGGA TAKAT	30-09-2017	05-06-2017 12:59:23	SangatTepatWaktu
167103000901003540	TANGGA TAKAT	30-09-2017	19-05-2017 10:54:34	SangatTepatWaktu
167103000900402780	TANGGA TAKAT	30-09-2017	Belum Bayar	TidakTepatWaktu
167103000501107680	16 ULU	30-09-2017	Belum Bayar	TidakTepatWaktu
167103000501306870	16 ULU	30-09-2017	Belum Bayar	TidakTepatWaktu
167103000500500920	16 ULU	30-09-2017	Belum Bayar	TidakTepatWaktu
167103000501101320	16 ULU	30-09-2017	Belum Bayar	TidakTepatWaktu



Menghitung jumlah kelas dari estimasi berdasarkan klasifikasi yang terbentuk (*prior probability*), dimana jumlah tiap kelas masing- masing estimasi yang ada dibagi dengan jumlah seluruh data. C1 untuk kelas estimasi sangat tepat waktu, C2 untuk kelas tepat waktu dan C3 untuk kelas tidak tepat waktu:

- a. C1(Class Estimasi = „Sangat Tepat Waktu“) =jumlah “sangat tepat waktu” pada kolom Estimasi = $97/200 = 0,48$
- b. C2 (Class Estimasi = “Tepat Waktu”) = jumlah “tidak tepat waktu” pada kolom Estimasi = $48/200 = 0,24$
- c. C3 (Class Estimasi = “Tidak Tepat Waktu”) = jumlah “tidak tepat waktu” pada kolom Estimasi = $55/200 = 0,27$

Kemudian Menghitung jumlah kasus yang sama pada setiap atribut dari kelas Estimasi (sangat tepat waktu , tepat waktu/tidak tepat waktu) berdasarkan masing-masing kelurahan OP (13 ULU, 14 ULU, 16 ULU TANGGA TAKAT dan SENTOSA) kemudian dibagi dengan seluruh jumlah atribut kelurahan OP masing-masing.

- a. $P(\text{Kelurahan} = \text{“11 ULU”} \mid \text{Class Estimasi} = \text{“sangat tepat waktu”}) = 7/8 = 0,87$
- b. $P(\text{Kelurahan} = \text{“11 ULU”} \mid \text{Class Estimasi} = \text{“tepat waktu”}) = 1/8 = 0,12$
- c. $P(\text{Kelurahan} = \text{“11 ULU”} \mid \text{Class Estimasi} = \text{“tidak tepat waktu”}) = 0/8 = 0$
- d. $P(\text{Kelurahan} = \text{“12 ULU”} \mid \text{Class Estimasi} = \text{“sangat tepat waktu”}) = 7/12 = 0,58$
- e. $P(\text{Kelurahan} = \text{“12 ULU”} \mid \text{Class Estimasi} = \text{“tepat waktu”}) = 3/12 = 0,25$
- f. $P(\text{Kelurahan} = \text{“12 ULU”} \mid \text{Class Estimasi} = \text{“tidak tepat waktu”}) = 2/12 = 0,16$
- g. $P(\text{Kelurahan} = \text{“13 ULU”} \mid \text{Class Estimasi} = \text{“sangat tepat waktu”}) = 18/28 = 0,64$
- h. $P(\text{Kelurahan} = \text{“13 ULU”} \mid \text{Class Estimasi} = \text{“tepat waktu”}) = 6/28 = 0,21$
- i. $P(\text{Kelurahan} = \text{“13 ULU”} \mid \text{Class Estimasi} = \text{“tidak tepat waktu”}) = 4/28 = 0,14$
- j. $P(\text{Kelurahan} = \text{“14 ULU”} \mid \text{Class Estimasi} = \text{“sangat tepat waktu”}) = 6/15 = 0,4$
- k. $P(\text{Kelurahan} = \text{“14 ULU”} \mid \text{Class Estimasi} = \text{“tepat waktu”}) = 6/15 = 0,4$
- l. $P(\text{Kelurahan} = \text{“14 ULU”} \mid \text{Class Estimasi} = \text{“tidak tepat waktu”}) = 3/15 = 0,2$
- m. $P(\text{Kelurahan} = \text{“16 ULU”} \mid \text{Class Estimasi} = \text{“sangat tepat waktu”}) = 26/70 = 0,37$
- n. $P(\text{Kelurahan} = \text{“16 ULU”} \mid \text{Class Estimasi} = \text{“tepat waktu”}) = 16/70 = 0,22$
- o. $P(\text{Kelurahan} = \text{“16 ULU”} \mid \text{Class Estimasi} = \text{“tidak tepat waktu”}) = 28/70 = 0,4$
- p. $P(\text{Kelurahan} = \text{“Tangga Takat”} \mid \text{Class Estimasi} = \text{“sangat tepat waktu”}) = 21/37 = 0,56$
- q. $P(\text{Kelurahan} = \text{“Tangga Takat”} \mid \text{Class Estimasi} = \text{“tepat waktu”}) = 10/37 = 0,27$
- r. $P(\text{Kelurahan} = \text{“Tangga Takat”} \mid \text{Class Estimasi} = \text{“tidak tepat waktu”}) = 6/37 = 0,16$
- s. $P(\text{Kelurahan} = \text{“Sentosa”} \mid \text{Class Estimasi} = \text{“sangat tepat waktu”}) = 12/30 = 0,4$
- t. $P(\text{Kelurahan} = \text{“Sentosa”} \mid \text{Class Estimasi} = \text{“tepat waktu”}) = 6/30 = 0,2$
- u. $P(\text{Kelurahan} = \text{“Sentosa”} \mid \text{Class Estimasi} = \text{“tidak tepat waktu”}) = 12/30 = 0,4$

Selanjutnya kalikan Semua Hasil *Variable* yang didapat dari hasil P(KELURAHAN OP|Class Estimasi) dengan sesama kelas estimasinya (C1, C2 dan C3) , adalah sebagai berikut :

- a. Untuk semua atribut *Class Estimasi* = “sangat tepat waktu”
 $P(X|\text{Class Estimasi} = \text{“sangat tepat waktu”}) = 0,87 \times 0,58 \times 0,64 \times 0,4 \times 0,37 \times 0,56 \times 0,4 = 0,01$
- b. Untuk semua atribut *Class Estimasi* = “tepat waktu” $P(X|\text{Class Estimasi} = \text{“tepat waktu”}) = 0,12 \times 0,25 \times 0,21 \times 0,4 \times 0,22 \times 0,27 \times 0,2 = 0,04$
- c. Untuk semua atribut *Class Estimasi* = “tidak tepat waktu” $P(X|\text{Class Estimasi} = \text{“tidak tepat waktu”}) = 0 \times 0,16 \times 0,14 \times 0,2 \times 0,4 \times 0,16 \times 0,4 = 0$

Lalu kalikan hasil yang didapat dari P (NOP|kelas estimasi) dengan hasil yang didapat dari hasil perkalian semua *variable* (C1, C2 dan C3) diatas :

- a. Perkalian *priot probability* dengan semuaatribut *Class Estimasi* = “sangat tepat waktu” $P (Ci) \mid \text{Class Estimasi} = \text{“sangat tepat waktu”}) \times P(X|\text{Class Estimasi} = \text{“sangat tepat waktu”}) = 0,48 \times 0,01 = 0,0048$
- b. Perkalian *priot probability* dengan semuaatribut *Class Estimasi* = “tepat waktu”
 $P (Ci) \mid \text{Class Estimasi} = \text{“tepat waktu”}) \times P(X|\text{Class Estimasi} = \text{“tepat waktu”})$



$$= 0,24 \times 0,04$$

$$= 0,0096$$

- c. Perkalian *prior probability* dengan semua atribut *Class Estimasi* = “tidak tepat waktu”
 $P(C_i) | \text{Class Estimasi} = \text{“tidak tepat waktu”} \times P(X | \text{Class Estimasi} = \text{“tidak tepat waktu”})$
 $= 0,27 \times 0 = 0$

Dari hasil perhitungan $P(X|C_i) \times P(C_i)$ didapat perbandingan dimana $P(C_i | \text{Class Estimasi} = \text{“Tepat Waktu”}) \times P(X | \text{Class Estimasi} = \text{“tepat waktu”}) > P(C_i | \text{Class Estimasi} = \text{“Sangat Tepat Waktu”}) \times P(X | \text{Class Estimasi} = \text{“sangat tepat waktu”})$ & $P(C_i | \text{Class Estimasi} = \text{“tidak tepat waktu”}) \times P(X | \text{Class Estimasi} = \text{“tidak tepat waktu”})$

Kesimpulan :

Perhitungan antara perkalian *Class Estimasi* “ sangat tepat waktu”, *Class Estimasi* “tepat waktu”, dan *Class Estimasi* “tidak tepat waktu” menunjukkan bahwa nilai *Class Estimasi* = “tepat waktu” lebih besar dibandingkan *class Estimasi* “sangat tepat waktu” dan *class estimasi* “tidak tepat waktu”

1) Penerapan Algoritma Naive Bayes Pada RapidMiner

Tahapan implementasi algoritma *naive bayes* menggunakan *tools rapidminer* untuk menentukan potensi kepatuhan wajib pajak. Gambar 5 berikut merupakan hasil dari potensi kepatuhan wajib pajak dengan ditambahkan nilai *confidence* sebagai hasil prediksi dari *tools rapidminer*.

ID	KAP	Class	confidence/Sangat tepat Waktu	confidence/tepat Waktu	confidence/Sangat tepat Waktu	tepat Waktu	tepat Waktu
1	0	tepat Waktu	0,001	0,002	0,000	0,100	-0,220
2	0	tepat Waktu	0,001	0,002	0,007	-0,070	-0,050
3	0	Sangat tepat Waktu	0,000	0,002	0,000	-0,100	-0,070
4	0	tepat Waktu	0,001	0,002	0,000	-0,270	-0,070
5	0	Sangat tepat Waktu	0,002	0,002	0,000	0,200	-0,070
6	0	Sangat tepat Waktu	0,001	0,002	0,007	-0,050	0,000
7	0	tepat Waktu	0,000	0,000	0,000	0,170	-0,070
8	0	Sangat tepat Waktu	0,002	0,000	0,000	0,007	-0,007
9	0	tepat Waktu	0,002	0,001	0,007	-0,100	0,000
10	0	Sangat tepat Waktu	0,000	0,002	0,000	-0,140	-0,000
11	0	tepat Waktu	0,000	0,000	1,000	0,000	-0,110
12	0	tepat Waktu	0,001	0,000	0,000	0,150	0,110
13	0	Sangat tepat Waktu	0,000	0,002	0,000	0,167	-0,000
14	0	Sangat tepat Waktu	0,002	0,002	0,000	0,200	-0,000
15	0	tepat Waktu	0,001	0,000	0,000	0,200	-0,200
16	0	tepat Waktu	0,002	0,000	0,000	-0,000	-0,100

Gambar 5. Hasil Potensi

3.6 Evaluation

Pada tahap ini pola-pola hasil akhir dari proses data mining akan dipresentasikan dalam bentuk visual untuk membantu *user* dalam mengerti pola tersebut.

1) Akurasi Prediksi

Pada percobaan dengan algoritma *naive bayes* dengan menggunakan *tools rapidminer* diperoleh waktu eksekusi sebesar 1second dapat diartikan eksekusi berjalan cukup cepat. Hasil eksekusi dapat dilihat pada gambar 6 berikut.

Process	Results	Completed	Execution Time
Process (1 results)	Process (1 results)	Completed Aug 31, 2018 8:30:16 AM	0 s
Process (1 results)	Process (1 results)	Completed Aug 31, 2018 8:34:13 AM	0 s
Process (1 results)	Process (1 results)	Completed Aug 31, 2018 8:35:37 AM	0 s
Process (1 results)	Process (1 results)	Completed Aug 31, 2018 8:37:16 AM	0 s
Process (1 results)	Process (1 results)	Completed Aug 31, 2018 8:41:38 AM	0 s
Process (2 results)	Process (2 results)	Completed Aug 31, 2018 8:51:14 AM	1 s
Process (3 results)	Process (3 results)	Completed Aug 31, 2018 8:52:35 AM	0 s

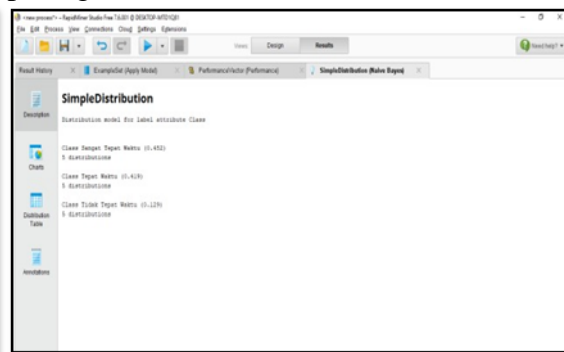
Gambar 6. Hasil eksekusi



Hasil akurasi model *naive bayes* menunjukkan bahwa nilai tingkat akurasi yang dihasilkan oleh algoritma *naive bayes* memiliki tingkat kekuatan yang sangat tinggi. Hal ini dibuktikan dengan hasil akurasi perhitungan mencapai 99.33%, nilai tersebut membuktikan bahwa model yang dibangun dapat digunakan untuk melakukan klasifikasi potensi kepatuhan wajib pajak, nilai yang tinggi tersebut juga disebabkan oleh kompleksnya data yang mengakibatkan model dapat memprediksi dengan akurat. Adapun hasil akurasi dapat dilihat pada gambar 7.

	Hasil Sangat Tepat Waktu	Hasil Tepat Waktu	Hasil Tidak Tepat Waktu	Class Precision
pred Sangat Tepat Waktu	336	3	0	99.12%
pred Tepat Waktu	2	316	0	99.36%
pred Tidak Tepat Waktu	0	36	36	100.00%
class recall	88.41%	89.64%	100.00%	

Gambar 7. Hasil Akurasi



Gambar 8 Model Distribusi

Dari gambar 8 dapat ditarik kesimpulan bahwa data PBB kecamatan seberang ulu, kelas sangat tepat waktu dengan nilaisebesar 0.382 lebih besar dari kelas tepat waktu dan kelas tidak tepat waktu. Adapun model distribusi untuk label atribut kelas estimasiadalah sebagai berikut:

- Kelas sangat tepat waktu : 5 *distribution*
- Kelas tepat waktu : 5 *distribution*
- Kelas tidak tepat waktu : 5 *distribution*

2) Hasil pengujian

Berdasarkan penelitian yang telah dilakukan, telah didapatkan suatu informasi dan pengetahuan baru dari proses data mining dalam menentukan wajib pajak di kelurahan mana yang berpotensi tidak patuh atau tidak tepat waktu, dan kebanyakan wajib pajak mana yang patuh membayar iuran PBB yang didapat setelah dilakukan tahapan *knowledge discovery in database (KDD)* dari data PBB kecamatan seberang ulu 2 tahun 2017. Dari penelitian tersebut didapatkan hasil yang sesuai dengan tujuan untuk menghasilkan informasi yang baru tentang kepatuhan wajib pajak yang berpotensi tidak tepat waktu.

Dari proses perhitungan data mining menggunakan algoritma *naive bayes* dengan menerapkan teknik *classification* dan tingkat keakurasiannya, dihasilkan informasi berdasarkan perhitungan data PBB kecamatan ulu pada tahun 2017 menunjukkan kelas sangat tepat waktu dengan perkalian *prior probability* senilai 0,48, sedangkan kelas tidak tepat waktu dengan *prior probability* senilai 0,27 dan kelas tepat waktu dengan *prior probability* senilai 0,24 serta menurut perhitungan dari *tools rapidminer* kelurahan yang tidak tepat waktu ada di kelurahan 16 ulu dari sebagian wajib pajak dengan data sebanyak 0.437 dan berpotensi tidak tepat waktu terdapat pada kelurahan tangga takat dengan data sebanyak 0.229 dengan akurasi sebesar 99.33% yang diakibatkan berpotensi tidak tepat waktu karena besarnya total NJOP bumi dan bangunan yang mengakibatkan tingginya jumlah tagihan menyebabkan sulitnya masyarakat untuk membayar iuran PBB pertahunnya

3) Hasil Analisis

Dari proses perhitungan data mining secara manual dan dengan implementasi pada *tools rapidminer* didapatkan hasil wajib pajak di setiap kelurahan yang ada pada kecamatan seberang ulu patuh dalam membayar iuran PBB, didapatkan kelurahan yang berpotensi tidak tepat waktu terdapat pada sebagian wajib pajak kelurahan 16 ulu dengan data sebanyak 42 wajib pajak dan kelurahan tangga takat dengan data sebanyak 22 wajib pajak. Dengan hasil penelitian ini diharapkan dapat menjadi bahan referensi instansi pemerintah dalam memberikan arahan terhadap kelurahan tersebut agar menjadi patuh dalam membayar iuran PBB.

Adapun data wajib pajak dengan estimasi kepatuhannya sangat tepat waktu dapat dilihat pada tabel 2. Jumlah pada tabel 2 ini didapat dari jumlah seluruh kelas estimasi sangat tepat waktu pada



dataset yang digunakan.

Tabel 2. Hasil Pengujian

Estimasi	Kelurahan	Jumlah	Prior Probability
Sangat Tepat Waktu	16 ULU	0.405	0.361
	TANGGA TAKAT	0.246	0.142
	SENTOSA	0.121	0.139
	14 ULU	0.109	0.156
	13 ULU	0.077	0.125
	12 ULU	0.024	0.028
	11 ULU	0.018	0.049

Tabel 3. Hasil Tepat Waktu

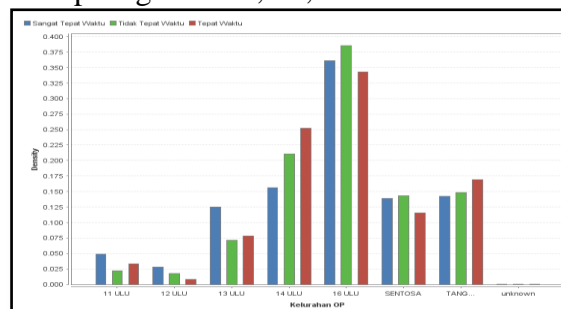
Estimasi	Kelurahan	Jumlah	Prior Probability
Tepat Waktu	16 ULU	0.434	0.386
	TANGGA TAKAT	0.173	0.148
	14 ULU	0.150	0.211
	SENTOSA	0.141	0.143
	13 ULU	0.073	0.072
	11 ULU	0.022	0.022
	12 ULU	0.006	0.0018

Tabel 4. Hasil Tidak Tepat Waktu

Estimasi	Kelurahan	Jumlah	Prior Probability
Tidak Tepat Waktu	16 ULU	0.437	0.343
	TANGGA TAKAT	0.229	0.169
	SENTOSA	0.146	0.116
	14 ULU	0.104	0.252
	13 ULU	0.073	0.079
	11 ULU	0.01	0.033
	12 ULU	0	0.008

4) Knowledge Representation

Ini merupakan tahapan akhir dimana pengetahuan yang telah ditemukan secara visual ditampilkan kepada user dengan berbagai visualisasi seperti *chart*, *histogram*, dan *pareto*. Adapun visualisasi tersebut dapat dilihat pada gambar 9, 10, dan 11.

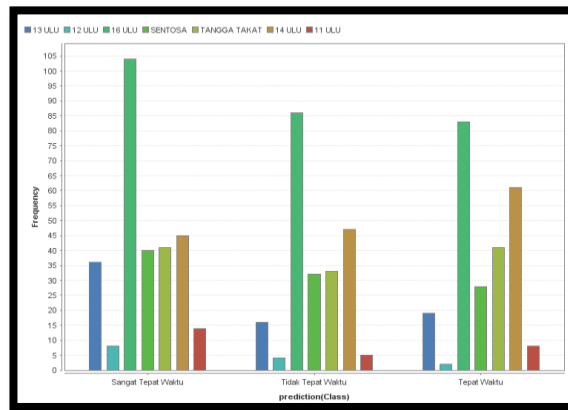


Gambar 9. Chart

Gambar 9 memperlihatkan data wajib pajak kelurahan 16 ulu dengan presentase diagram tertinggi yang memiliki data terbanyak diantara kelurahan lain dengan *class* estimasi tertinggi tidak

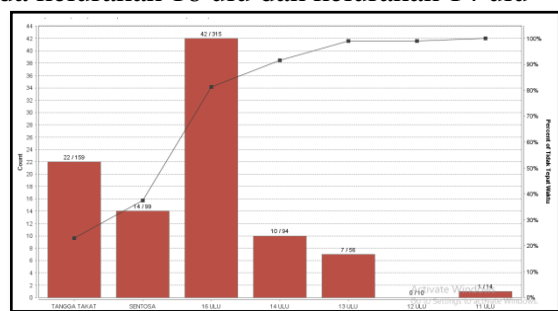


tepat waktu mencapai >0.375.



Gambar 10. Histogram

Pada gambar 10 memperlihatkan *class* estimasi sangat tepat waktu, tepat waktu dan tidak teapat waktu tertinggi terdapat pada kelurahan 16 ulu dan kelurahan 14 ulu



Gambar 11. Pareto

Pada gambar 11 memperlihatkan *class* tidak tepat waktu tertinggi terdapat pada kelurahan 16 ulu dan tangga takat.

KESIMPULAN

Berdasarkan perhitungan data mining menggunakan algoritma *naive bayes* pada teknik *classification* dapat ditarik kesimpulan bahwa kelas estimasi pembayaran wajib pajak sangattepat waktu lebih besar dari pada kelas estimasi tepat waktu dan tidak tepat waktu.

Dari hasil *observasi* terhadap data PBB kecamatan seberang ulu 2 melalui teknik *classification* menggunakan algoritma *naive bayes* didapatkan kelurahan yang berpotensi tidak tepat waktu terdapat pada sebagian wajib pajak kelurahan 16 ulu dengan data 0.437 dan kelurahan tangga takat dengan data sebanyak 0.229 yang diakibatkan oleh besarnya jumlah total NJOP bumi dan bangunan yang membuat semakin tingginya jumlah tagihan menyebabkan sulitnya masyarakat membayar iuran PBB pertahunnya.

Tools rapidminer berhasil menentukan potensi kepatuhan wajib pajak per kelurahan mana yang berpotensi tidak tepat waktu padatahun 2017 dengan akurasi sebesar 99.33% dari jumlah seluruh data wajib pajak PBB kecamatan seberang ulu 2 sebesar 1.467 wajib pajak setelah dilakukan penghapusan tagihan dibawah Rp. 100.000.

DAFTAR RUJUKAN

- [1] Y. Irawan, "Penerapan Data Mining Untuk Evaluasi Data Penjualan Menggunakan Metode Clustering Dan Algoritma Hirarki Divisive Di Perusahaan Media World Pekanbaru," J. Teknol. Inf. Univ. Lambung Mangkurat, vol. 4, no. 1, pp. 13–20, 2019.
- [2] N. L. W. S. R. Ginantra et al., Data mining dan penerapan algoritma. Yayasan Kita Menulis, 2021.
- [3] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," J. Media Inform. Budidarma, vol. 4, no. 2, pp. 437–444, 2020.



- [4] A. F. O. Pasaribu and others, “Analisis Pola Menggunakan Metode C4. 5 Untuk Peminatan Jurusan Siswa Berdasarkan Kurikulum (Studi Kasus: Sman 1 Natar),” *J. Teknol. Dan Sist. Inf.*, vol. 2, no. 1, pp. 80–85, 2021.
- [5] I. Istiningsih and others, “IMPLEMENTASI DATA MINING PENGELOMPOKKAN PELANGGAN MENGGUNAKAN RFM DAN K-MEANS CLUSTERING STUDI KASUS DI INDONESIA DIGITAL PRINTING YOGYAKARTA,” *STMIK AKAKOM YOGYAKARTA*, 2020.
- [6] R. N. Situmorang, “Klasifikasi Kesegaran Ikan Berdasarkan Ekstraksi Fitur Menggunakan Metode K Nearest Neighbor Dan Hue Saturation Value,” *Universitas Islam Negeri Sumatera Utara Medan*, 2021.
- [7] A. Sapitri, “PERBANDINGAN METODE K-NEAREST NEIGHBOR (K-NN) DAN METODE K MEANS DALAM MENENTUKAN TINGKAT PENJUALAN PRODUK PADA PERUSAHAAN KAYU ELANG PERKASA,” *Universitas Darma Persada*, 2021.
- [8] R. Y. Hayuningtyas, “Penerapan Algoritma Naive Bayes untuk Rekomendasi Pakaian Wanita,” *J. Inform.*, vol. 6, no. 1, pp. 18–22, 2019.
- [9] N. Pramesti, “Klasifikasi Persediaan Barang Menggunakan Naive Bayes,” *J. DATA Sci. & Inform.*, vol. 1, no. 2, pp. 53–57, 2021.
- [10] M. F. Rifai, H. Jatnika, and B. Valentino, “Penerapan Algoritma Naive Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (MOS),” 2019.
- [11] M. Misdrum, F. Syarifuddin, and A. A. Widodo, “Klasifikasi Data Set Virus Corona Menggunakan Metode Naive Bayes Classifier,” *SPIRIT*, vol. 12, no. 2, 2020

