

**PENERAPAN ALGORITMA NAIVE BAYES UNTUK MELIHAT
POTENSI KEPATUHAN PESERTA BPJS DALAM MEMBAYAR
TAGIHAN**

Rika Dinda Lestari¹, Raissa Amanda Putri²

Universitas Islam Negeri Sumatera Utara

E-mail: rkdnstar@gmail.com¹, raissa.ap@uinsu.ac.id²

Abstrak

Penelitian ini bertujuan untuk melihat potensi kepatuhan peserta BPJS dalam membayar tagihan dengan menggunakan algoritma naïve bayes. Jaminan Kesehatan Nasional (JKN) yang diselenggarakan oleh Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan merupakan jaminan sosial yang diwajibkan bagi seluruh warga negara Indonesia berdasarkan Undang-Undang Nomor 40 Tahun 2004 tentang Sistem Jaminan Sosial Nasional (SJSN). Dengan data sampel 1.499 kita dapat mengetahui peserta BPJS yang membayar tagihan dengan sangat tepat waktu, tepat waktu dan tidak tepat waktu dengan teknik klasifikasi dan menerapkan algoritma naïve bayes pada tools rapidminer diperoleh hasil akurasi sebesar 99.10% dengan nilai probability sangat tepat waktu 0.334, tepat waktu dengan nilai probability 0.182 dan tidak tepat waktu dengan nilai 0.484.

Kata Kunci — BPJS, Klasifikasi, Naive Bayes, Rapidminer.

1. PENDAHULUAN

Jaminan Kesehatan Nasional (JKN) yang diselenggarakan oleh Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan merupakan jaminan sosial yang diwajibkan bagi seluruh warga negara Indonesia berdasarkan Undang-Undang Nomor 40 Tahun 2004 tentang Sistem Jaminan Sosial Nasional (SJSN). BPJS tersebut dibentuk sesuai dengan amanat UU nomor 24 tahun 2011. (Novita, 2020) Undang-undang dasar yang dibuat untuk mengawasi sistem jaminan kesehatan disebut BPJS Kesehatan. Mengenai jaminan kesehatan, berlaku Peraturan Presiden Nomor 82 Tahun 2018. Sebuah organisasi hukum yang disebut Badan Penyelenggara Jaminan Sosial (BPJS) didirikan untuk mengawasi program jaminan sosial dengan tujuan memastikan terpenuhinya kebutuhan dari warga Indonesia yang membutuhkan sebuah asuransi kesehatan yang terjamin.

Asuransi kesehatan memberikan jaminan perlindungan kepada peserta dalam bentuk pelayanan kesehatan sehingga peserta dapat memanfaatkan perlindungan pelayanan kesehatan dalam memenuhi kebutuhannya. Setiap anggota masyarakat yang memberikan kontribusi secara mandiri atau didanai pemerintah berhak mendapatkan pelayanan kesehatan dasar. (Haq et al., 2022)

Tanggal pembayaran tagihan BPJS jatuh pada tanggal 10 setiap bulannya. Peserta BPJS yang melakukan pembayaran tagihan sebelum tanggal 10 dapat disebut dengan peserta BPJS yang sangat tepat waktu akan melakukan kewajiban tersebut. Kepatuhan

dalam membayar tagihan yang melakukan pembayaran setiap bulannya ditanggal 10 dikatakan seorang peserta BPJS yang tepat waktu dan bertanggung jawab akan kewajibannya. Seorang peserta BPJS dikatakan tidak tepat waktu apabila ia membayar tagihan lewat dari tanggal 10. Ketidakpatuhan dalam pembayaran tagihan BPJS Kesehatan sering terjadi jika peserta tersebut tidak memahami siklus BPJS Kesehatan. Untuk itu kita perlu memiliki pemahaman mengenai BPJS Kesehatan agar dapat mengapresiasi penuh dengan sikap yang tepat waktu.(Fajrini et al., 2021)

Peserta BPJS Kesehatan yang tidak melakukan pembayaran tepat waktu kemungkinan tidak akan dikenakan denda keterlambatan, namun kepesertaan BPJS akan dibekukan dan peserta BPJS tidak bisa mendapatkan pelayanan Kesehatan dari fasilitas Kesehatan yang bekerja sama dengan BPJS Kesehatan sampai melunasi tunggakan tersebut. Dengan semakin banyaknya peserta BPJS dan data-data pembayaran tagihan menyebabkan pihak dari BPJS Kesehatan belum membuat pola untuk mengklasifikasikan seberapa banyak peserta BPJS yang sangat tepat waktu, tepat waktu dan tidak tepat waktu berdasarkan wilayah masing-masing peserta tersebut.

Pada penelitian ini dalam mengklasifikasikan data peserta BPJS dalam membayar tagihan yang sangat tepat waktu, tepat waktu dan tidak tepat waktu berdasarkan wilayah dari masing-masing peserta BPJS diperoleh data sebanyak 1500 data di 3 bulan 2023 yakni januari, februari dan maret. Dan setelah dilakukannya tahapan KDD kini data tersebut sebanyak 1499 data, karena sebelumnya terdapat data *missing* didalamnya. Untuk masing masing wilayah dari kantor BPJS Kesehatan wilayah 1 cabang kota medan yaitu ada kota medan, kota binjai dan kabupaten langkat. Untuk menghasilkan Tingkat akurasi dari pengklasifikasian penelitian dengan algoritma naïve bayes maka diperlukan untuk data latih dan data uji sebanyak 70% untuk data latih dan 30% untuk data uji.

Data mining adalah proses yang menggabungkan kecerdasan buatan, metode statistik, matematika, dan pembelajaran mesin untuk menemukan dan mengekstrak informasi dan pengetahuan terkait dari database besar.(Kurniawan et al., 2022) Intinya, data mining adalah proses extraction untuk mendapatkan informasi penting yang implisit dan belum diketahui. (Prasetiyo et al., 2019) Dalam penelitian ini, metode klasifikasi data mining digunakan untuk melibatkan kelompok sasaran secara akurat. Tujuan dari metode klasifikasi ini adalah untuk mengelompokkan nilai suatu variabel yang tidak diketahui dari variabel lain yang telah ditentukan.(Riany & Testiana, 2023) Klasifikasi adalah Persiapan untuk mencari model yang mendeskripsikan data dan mengklasifikasikannya ke dalam kelas-kelas yang ada .(Undamayanti et al., 2022) Target dan proses *Knowledge Discovery in Database* (KDD) sangat dipengaruhi oleh pemilihan metode atau algoritma yang efektif. (Yunita, 2018) Kata lain yang digunakan dalam data mining adalah *Knowledge Discovery in Database*(KDD), yang memiliki arti serupa dengan data mining.(Widayati, 2019)

Dalam metode klasifikasi data mining, ada banyak algoritma, salah satunya adalah algorithm Naive Bayes..(Alfani, 2019) Algoritma diperlukan untuk mengatur komputer untuk mengambil langkah-langkah tertentu dalam pemecahan masalah. (Dianta, 2021) Salah satu teknologi klasifikasi algoritma yang mudah digunakan dan cepat diproses adalah metode klasifikasi Naive Bayes. (Rahman, 2023) Berdasarkan teorema Bayes, algoritma naïve Bayes merupakan metode klasifikasi statistik yang dapat digunakan untuk memprediksi probabilitas pengumpulan dari suatu kelas. Ia mampu mengklasifikasikan seperti pohon keputusan dan jaringan saraf.(Furqan et al., 2022) Naive Bayes menemukan peluang terbesar dari kemungkinan klasifikasi dengan melihat frekuensi tiap klasifikasi pada data pelatihan.(Hermanto et al., 2020) Berikut ini ialah rumus dari perhitungan manual naïve bayes :

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}$$

Keterangan :

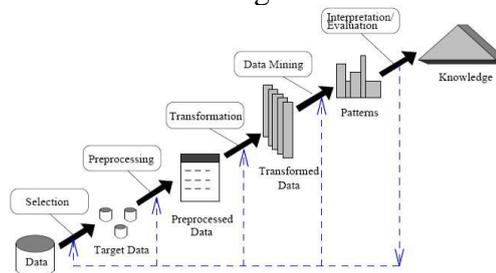
$P(H|E)$ = Probabilitas

H = Hipotesis

E = Bukti

2. METODE PENELITIAN

Metode penelitian adalah metode yang diciptakan peneliti untuk mengumpulkan data dan informasi selama penelitian yang dilakukannya. Dengan data-data yang mereka kumpulkan diharapkan mampu menghasilkan penelitian dengan kualitas terbaik. (Yuda, 2022) Metode pada penelitian ini adalah metode kuantitatif, yang bertujuan untuk menguji hipotesis dengan menggunakan teori-teori yang sudah ada. Metode kuantitatif yang merupakan penelitian yang menguji teori dengan meneliti hubungan antara variabel. Pada penelitian yang akan dilakukan ialah menerapkan algoritma naïve bayes dengan rapidminer untuk melihat potensi kepatuhan peserta BPJS dalam membayar tagihan. Sebelum dilakukannya penerapan dan pengujian, *Knowledge Discover in Database* (KDD) digunakan untuk menganalisis data pada penerapan data mining. Tahapan-tahapan ini sering digunakan bersama-sama untuk menjelaskan proses pengumpulan informasi tersembunyi dalam data berukuran besar. Berikut gambaran dari tahapan KDD:



Gambar 1 Tahapan KDD

3. HASIL DAN PEMBAHASAN

Adapun sampel data untuk penelitian yang akan dilakukan ialah sebanyak 1.500 data sebelum dilakukannya analisis data dengan tahapan KDD. Berikut data mentah yang diperoleh dari Kantor BPJS Kesehatan Kedeputusan Wilayah 1 Cabang Kota Medan:

Tabel 1 Dataset Peserta BPJS Kesehatan

No BPJS	Wilayah	Alamat	Tgl Jth Tempo	Tgl Bayar
1719622923	KOTA MEDAN	JL HALAT GG RA	10/02/2023	01/02/2023
2738206168	KOTA MEDAN	JL TITIPAPAN GC	10/02/2023	01/02/2023
7334111	KOTA MEDAN	JL AMALIUN NO	10/02/2023	01/02/2023
1150556163	KOTA MEDAN	JL SETIABUDI PS	10/02/2023	01/02/2023
2478230728	KOTA MEDAN	LINGK07	10/02/2023	01/02/2023
2498929435	KOTA MEDAN	JL ALFAKA 6 LK	10/02/2023	01/02/2023
7336945	KOTA MEDAN	JL BALAM NO 4	10/02/2023	01/02/2023
7337542	KOTA MEDAN	JL BALAM NO 4	10/02/2023	01/02/2023
7339544	KOTA MEDAN	jl a s muthalib peru	10/02/2023	01/02/2023
7340152	KOTA MEDAN	KOMP GRAHA P/	10/02/2023	01/02/2023
1853972842	KOTA MEDAN	JL KARYA BAKT	10/02/2023	01/02/2023
7343054	KOTA MEDAN	JL MENTENG II G	10/02/2023	01/02/2023
7343144	KOTA MEDAN	JL MADIO UTOM	10/02/2023	01/02/2023
7346845	KOTA MEDAN	JL GARU I NO 166	10/02/2023	01/02/2023
2474285073	KOTA MEDAN	JL BUNGA NCOL	10/02/2023	01/02/2023
7348228	KOTA MEDAN	JL DR. WAHIDIN	10/02/2023	01/02/2023
7350232	KOTA MEDAN	JL MARELAN III I	10/02/2023	01/02/2023
7350377	KOTA MEDAN	JL PEMBANGUN/	10/02/2023	01/02/2023

Sebelum menerapkan dan menguji algoritma naïve bayes, Untuk menganalisis data pada penelitian ini menggunakan tahapan *Knowledge Discover in Database (KDD)* yaitu sebagai berikut:

Input

Input data merupakan tahapan awal untuk menganalisis data. Data yang akan diinput ialah data mentah yang diberikan oleh pihak BPJS Kesehatan. pada data mentah yang diberikan ada data yang tidak perlu dipakai, untuk itu perlu di selection terlebih dahulu.

	A	B	C	D	E	F	G
1	No BPJS	Wilayah	Alamat	Tgl Jth Tempo	Tgl Bayar	Estimasi (Wa...	Estimasi
2	1719622923...	KOTA MEDAN	JL HALAT GG ...	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
3	2738206168....	KOTA MEDAN	JL TITIPAPAN...	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
4	7334111.000	KOTA MEDAN	JL AMALIUN ...	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
5	1150556163....	KOTA MEDAN	JL SETIABUD...	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
6	2478230728....	KOTA MEDAN	LINGK07	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
7	2498929435....	KOTA MEDAN	JL ALFAKA 6 ...	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
8	7336945.000	KOTA MEDAN	JL BALAM NO 4	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
9	7337542.000	KOTA MEDAN	JL BALAM NO 4	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
10	7339544.000	KOTA MEDAN	jl a s muthalib...	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
11	7340152.000	KOTA MEDAN	KOMP GRAH...	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
12	1853972842....	KOTA MEDAN	JL KARYA BA...	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
13	7343054.000	KOTA MEDAN	JL MENTENG...	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...
14	7343144.000	KOTA MEDAN	JL MADIO UT...	Feb 10, 2023 ...	Feb 1, 2023 1...	9.000	Sangat Tepat ...

Gambar 2 Input Data Uji

	A	B	C	D	E	F	G
1	No BPJS	Wilayah	Alamat	Tgl Jth Tempo	Tgl Bayar	Estimasi (Wa...	Estimasi
2	832015.000	KOTA MEDAN	JL MANCANG...	Jan 10, 2023 ...	Jan 19, 2023 ...	-9.000	Tidak Tepat ...
3	1030252.000	KOTA MEDAN	JL SELAMAT ...	Jan 10, 2023 ...	Jan 6, 2023 1...	4.000	Sangat Tepat ...
4	1030948.000	KOTA MEDAN	JL KLAMBIR V...	Jan 10, 2023 ...	Jan 29, 2023 ...	-19.000	Tidak Tepat ...
5	1031229.000	KOTA MEDAN	JL ENGSEL LI...	Jan 10, 2023 ...	Jan 10, 2023 ...	0.000	Tepat Waktu
6	1180247.000	KOTA MEDAN	JLGARUDA N...	Jan 10, 2023 ...	Jan 10, 2023 ...	0.000	Tepat Waktu
7	1826755108....	KOTA MEDAN	JLN. SATRIA ...	Jan 10, 2023 ...	Jan 10, 2023 ...	0.000	Tepat Waktu
8	1721417.000	KOTA MEDAN	JL LAMPU GG...	Jan 10, 2023 ...	Jan 10, 2023 ...	0.000	Tepat Waktu
9	1721417.000	KOTA MEDAN	DUSUN RAM...	Jan 10, 2023 ...	Jan 10, 2023 ...	0.000	Tepat Waktu
10	1792991.000	KOTA MEDAN	JL GEDUNG ...	Jan 10, 2023 ...	Jan 11, 2023 ...	-1.000	Tidak Tepat ...
11	1899347.000	KOTA MEDAN	DUSUN KANT...	Jan 10, 2023 ...	Jan 10, 2023 ...	0.000	Tepat Waktu
12	2045158.000	KOTA MEDAN	DUSUN SUM...	Jan 10, 2023 ...	Jan 10, 2023 ...	0.000	Tepat Waktu
13	2131345.000	KOTA MEDAN	DUSUN PAN...	Jan 10, 2023 ...	Jan 10, 2023 ...	0.000	Tepat Waktu
14	2142573.000	KOTA MEDAN	JL PARANG IV...	Jan 10, 2023 ...	Jan 10, 2023 ...	0.000	Tepat Waktu

Gambar 1 Input Data Latih

Data Selection

Pada tahap ini fungsinya untuk menyeleksi data sesuai kebutuhan. Dalam penelitian ini diperlukan data-data yang sesuai seperti Nomor BPJS, Wilayah, Tanggal Jatuh Tempo, dan Tanggal Pembayaran agar diperoleh hasil yang sesuai dengan prosesnya. Selain itu, terdapat tambahan kolom data yaitu estimasi dan estimasi (hari) untuk akurasi dalam memperkirakan potensi kepatuhan.

Row No.	Estimasi	No BPJS	Wilayah	Tgl Jth Tempo	Tgl Bayar	Estimasi (W...
986	Tidak Tepat ...	2082027543	KOTA BINJAI	Mar 10, 2023	Mar 14, 2023	-4
987	Tidak Tepat ...	2081473841	KOTA BINJAI	Mar 10, 2023	Mar 14, 2023	-4
988	Tidak Tepat ...	2081107034	KOTA BINJAI	Mar 10, 2023	Mar 14, 2023	-4
989	Tidak Tepat ...	2084051272	KOTA BINJAI	Mar 10, 2023	Mar 14, 2023	-4
990	Tidak Tepat ...	2085051431	KOTA BINJAI	Mar 10, 2023	Mar 14, 2023	-4
991	Tidak Tepat ...	2085033701	KOTA BINJAI	Mar 10, 2023	Mar 14, 2023	-4
992	Tidak Tepat ...	2084420777	KOTA BINJAI	Mar 10, 2023	Mar 14, 2023	-4
993	Tepat Waktu	2093557217	KOTA BINJAI	Mar 10, 2023	Mar 10, 2023	0
994	Sangat Tepat...	2199547811	KOTA BINJAI	Mar 10, 2023	Mar 7, 2023	3
995	Sangat Tepat...	2199086076	KOTA BINJAI	Mar 10, 2023	Mar 1, 2023	9
996	Sangat Tepat...	1265066087	KOTA BINJAI	Mar 10, 2023	Mar 5, 2023	5
997	Sangat Tepat...	2200634752	KOTA BINJAI	Mar 10, 2023	Mar 7, 2023	3
998	Sangat Tepat...	2201932124	KOTA BINJAI	Mar 10, 2023	Mar 9, 2023	1
999	Tepat Waktu	2201085415	KOTA BINJAI	Mar 10, 2023	Mar 10, 2023	0
1000	Tidak Tepat ...	2200818993	KOTA BINJAI	Mar 10, 2023	Mar 11, 2023	-1

ExampleSet (1,000 examples, 1 special attribute, 5 regular attributes)

Gambar 4 Hasil Selection Data Uji

Row No.	Estimasi	No BPJS	Wilayah	Tgl Jth Tempo	Tgl Bayar	Estimasi (W...
485	Tidak Tepat ...	7313/14	KOTA MEDAN	Feb 10, 2023	Feb 19, 2023	-9
487	Tidak Tepat ...	7314243	KOTA MEDAN	Feb 10, 2023	Feb 19, 2023	-9
488	Sangat Tepat...	7316853	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
489	Sangat Tepat...	7318517	KOTA MEDAN	Feb 10, 2023	Feb 2, 2023	8
490	Sangat Tepat...	7321465	KOTA MEDAN	Feb 10, 2023	Feb 3, 2023	7
491	Sangat Tepat...	7321904	KOTA MEDAN	Feb 10, 2023	Feb 4, 2023	6
492	Sangat Tepat...	7322264	KOTA MEDAN	Feb 10, 2023	Feb 5, 2023	5
493	Sangat Tepat...	7322657	KOTA MEDAN	Feb 10, 2023	Feb 6, 2023	4
494	Sangat Tepat...	7324413	KOTA MEDAN	Feb 10, 2023	Feb 7, 2023	3
495	Sangat Tepat...	1872268942	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
496	Sangat Tepat...	7325381	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
497	Sangat Tepat...	7326426	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
498	Sangat Tepat...	7328812	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
499	Sangat Tepat...	7333604	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
500	Sangat Tepat...	764256	?	Feb 11, 2023	Feb 2, 2023	9

ExampleSet (500 examples, 1 special attribute, 5 regular attributes)

Gambar 2 Hasil Selection Data Latih

Preprocessing Data

Jika data yang telah dipilih pada tahap sebelumnya masih memiliki data kosong/missing, kita dapat memeriksanya dengan memprosesnya. Jika tidak ada, data akan segera diproses pada tahap transformasi, sementara jika ada, data akan dibersihkan

Label	Estimasi	No BPJS	Wilayah	Tgl Jth Tempo	Tgl Bayar	Estimasi (W...
▼	Estimasi	Nominal	0	Least: Sangat Tepat Waktu (1...)	Most: Tidak Tepat Waktu (242)	Values: Tidak Tepat Waktu (242), Tepat Waktu (133), ...[1
▼	No BPJS	Real	0	Min: 764256	Max: 3047540736	Average: 293891153.862
▼	Wilayah	Nominal	1	Least: KOTA MEDAN (499)	Most: KOTA MEDAN (499)	Values: KOTA MEDAN (499)
▼	Tgl Jth Tempo	Date-time	0	Earliest date: Jan 10, 2023	Latest date: Feb 11, 2023	Duration: 32 days
▼	Tgl Bayar	Date-time	0	Earliest date: Jan 1, 2023	Latest date: Feb 25, 2023	Duration: 55 days
▼	Estimasi (Waktu)	Integer	0	Min: -21	Max: 9	Average: -2.766

Gambar 3 Data Missing Pada Data Latih

Terlihat pada gambar diatas terdapat 1 data missing pada data latih, yang perlu untuk dihapuskan dari data agar lebih akurat.

Row No.	Estimasi	No BPJS	Wilayah	Tgl Jth Tempo	Tgl Bayar	Estimasi (W...
485	Tidak Tepat...	7313409	KOTA MEDAN	Feb 10, 2023	Feb 19, 2023	-9
486	Tidak Tepat...	7313714	KOTA MEDAN	Feb 10, 2023	Feb 19, 2023	-9
487	Tidak Tepat...	7314243	KOTA MEDAN	Feb 10, 2023	Feb 19, 2023	-9
488	Sangat Tepat...	7316853	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
489	Sangat Tepat...	7318517	KOTA MEDAN	Feb 10, 2023	Feb 2, 2023	8
490	Sangat Tepat...	7321465	KOTA MEDAN	Feb 10, 2023	Feb 3, 2023	7
491	Sangat Tepat...	7321904	KOTA MEDAN	Feb 10, 2023	Feb 4, 2023	6
492	Sangat Tepat...	7322264	KOTA MEDAN	Feb 10, 2023	Feb 5, 2023	5
493	Sangat Tepat...	7322657	KOTA MEDAN	Feb 10, 2023	Feb 6, 2023	4
494	Sangat Tepat...	7324413	KOTA MEDAN	Feb 10, 2023	Feb 7, 2023	3
495	Sangat Tepat...	1872268942	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
496	Sangat Tepat...	7325381	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
497	Sangat Tepat...	7326426	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
498	Sangat Tepat...	7328812	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9
499	Sangat Tepat...	7333604	KOTA MEDAN	Feb 10, 2023	Feb 1, 2023	9

ExampleSet (499 examples, 1 special attribute, 5 regular attributes)

Gambar 7 Hasil Preprocessing Data

Gambar diatas menunjukkan hasil dari preprocessing data. Kini data tersebut tersisa sebanyak 499 data dan tidak lagi ada data kosong/missing.

Data Transformation

Setelah memberesihkan data yang missing, maka di tahap ini kita akan menormalisasikan data nya dengan panel normalize pada rapidminer agar data yang digunakan lebih sesuai untuk diterapkandengan algoritma naïve bayes:

Z-Transformation

```
Normalize 2 attributes to mean 0 and variance 1.
Using
No BPJS --> mean: 7.40934261038E8, variance: 1.00094961623943424E18
Estimasi ( Waktu ) --> mean: -1.623, variance: 42.905776776776
```

Z-Transformation

```
Normalize 2 attributes to mean 0 and variance 1.
Using
No BPJS --> mean: 2.93891153862E8, variance: 5.2911777622010426E17
Estimasi ( Waktu ) --> mean: -2.766, variance: 55.97920240480962
```

Gambar 4 Hasil Transformasi pada Data Uji dan Data Latih

Data yang telah diubah menjadi bentuk yang akurat dan sesuai untuk tahap selanjutnya dari data mining ditunjukkan pada gambar di atas.

Data Mining

Untuk mendapatkan hasil yang tepat, maka akan dilakukan 2 pengujian yaitu pengujian perhitungan manual dan pengujian dengan tools rapidminer. Untuk melakukan perhitungan manual harus dibuat dulu estimasi dari setiap tanggal bayar dan tanggal jatuh tempo. Dataset untuk melakukan perhitungan manual terdiri dari 1000 data uji dan 499 data latih setelah melakukan transformasi data, karena terdapat data yang kosng/missing. Berikut rumus perhitungan manual algoritma naïve bayes:

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}$$

Keterangan :

$P(H|E)$ = Probabilitas

H = Hipotesis

E = Bukti

$P1$ = Sangat tepat waktu

$P2$ = Tepat waktu

P3 = Tidak tepat waktu

Hitung jumlah kelas estimasi, diman jumlah kelas untuk setiap estimasi dibagi dengan total data.

- a. P1 Kelas = “sangat tepat waktu”= jumlah P1
= $334/1000$
= 0.334
- b. P2 Kelas = “tepat waktu” = jumlah P2
= $182/1000$
= 0.182
- c. P3 Kelas = “tidak tepat waktu” = jumlah P3
= $484/1000$
= 0.484

Selanjutnya, hitung jumlah kelas untuk setiap prediksi (sangat tepat waktu, tepat waktu, atau tidak tepat waktu) berdasarkan wilayah yang menjadi bagian dari cabang Medan seperti kota medan, kabupaten langkat dan kota binjai. Kemudian, bagi jumlah atribut total untuk tiap region.

- a. $P(\text{wilayah} = \text{“kota medan”} \mid \text{Kelas} = \text{“sangat tepat waktu”})$
= $138/301$
= 0.45
- b. $P(\text{wilayah} = \text{“kota medan”} \mid \text{Kelas} = \text{“tepat waktu”})$
= $29/301$
= 0.09
- c. $P(\text{wilayah} = \text{“kota medan”} \mid \text{Kelas} = \text{“tidak tepat waktu”})$
= $134/301$
= 0.44
- d. $P(\text{wilayah} = \text{“kabupaten langkat”} \mid \text{Kelas} = \text{“sangat tepat waktu”})$
= $60/245$
= 0.24
- e. $P(\text{wilayah} = \text{“kabupaten langkat”} \mid \text{Kelas} = \text{“tepat waktu”})$
= $21/245$
= 0.08
- f. $P(\text{wilayah} = \text{“kabupaten langkat”} \mid \text{Kelas} = \text{“tidak tepat waktu”})$
= $164/245$
= 0.66
- g. $P(\text{wilayah} = \text{“kota binjai”} \mid \text{Kelas} = \text{“sangat tepat waktu”})$
= $136/454$
= 0.29
- h. $P(\text{wilayah} = \text{“kota binjai”} \mid \text{Kelas} = \text{“tepat waktu”})$
= $132/454$
= 0.29
- i. $P(\text{wilayah} = \text{“kota binjai”} \mid \text{Kelas} = \text{“tidak tepat waktu”})$
= $186/454$
= 0.40

Selanjutnya kalikan Semua Hasil yang diperoleh dari hasil $P(\text{wilayah} \mid \text{Kelas Estimasi})$ dengan sesama kelas estimasinya (P1, P2, dan P3) , adalah sebagai berikut :

- a. Untuk semua atribut Kelas Estimasi = “sangat tepat waktu” $\times P(E \mid \text{Kelas Estimasi} = \text{“sangat tepat waktu”})$
= $0.45 \times 0.24 \times 0.29$
= 0.03

- b. Untuk semua atribut Kelas Estimasi = “tepat waktu” $\times P(E|Kelas\ Estimasi= “tepat\ waktu”)$
 $= 0.09 \times 0.08 \times 0.29$
 $= 0.002$
- c. Untuk semua atribut Kelas Estimasi = “tidak tepat waktu” $\times P(E|Kelas\ Estimasi= “tidak\ tepat\ waktu”)$
 $= 0.44 \times 0.66 \times 0.40$
 $= 0.11$

Tahap selanjutnya ialah kalikan hasil yang didapat dari $P(E|Kelas\ estimasi)$ dengan hasil yang didapat dari hasil perkalian semua variable ($P1, P2$ dan $P3$) diatas :

- a. Perkalian prior probabilitas dengan semua atribut Kelas Estimasi = “sangat tepat waktu” $P1 | Kelas\ Estimasi = “sangat\ tepat\ waktu” \times P(E|Kelas\ Estimasi = “sangat\ tepat\ waktu”)$
 $= 0.334 \times 0.03$
 $= 0.01$
- b. Perkalian prior probabilitas dengan semua atribut Kelas Estimasi = “tepat waktu” $P2 | Kelas\ Estimasi = “tepat\ waktu” \times P(E|Kelas\ Estimasi = “tepat\ waktu”)$
 $= 0.182 \times 0.002$
 $= 0.0003$
- c. Perkalian prior probabilitas dengan semua atribut Kelas Estimasi = “tidak tepat waktu” $P3 | Kelas\ Estimasi = “tidak\ tepat\ waktu” \times P(E|Kelas\ Estimasi = “tidak\ tepat\ waktu”)$
 $= 0.484 \times 0.11$
 $= 0.05$

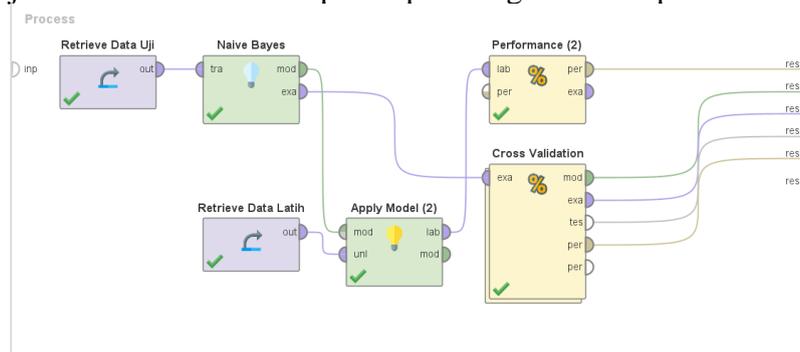
Dapat disimpulkan bahwa data tidak tepat waktu lebih unggul diantara data sangat tepat waktu dan tepat waktu.

Evaluation

Pada tahap ini dihasilkan pola- pola untuk model klasifikasi yang dievaluasi untuk mengetahui apakah hipotesis yang ada memang tercapai. Proses dari tahap ini akan dilihat pada hasil penerapan dan pengujian berikut:

Penerapan

Untuk mengetahui potensi kepatuhan peserta BPJS Kesehatan dalam membayar tagihan digunakan tools rapidminer dengan menggunakan algoritma naïve bayes, gambar berikut menunjukkan model hasil dari penerapan dengan tools rapidminer:



Gambar 5 Model Penerapan Naive Bayes dengan Rapidminer

Pengujian

A. Akurasi Prediksi

Dalam percobaan yang dilakukan dengan algoritma naïve bayes melalui tools rapidminer diperoleh waktu operasinya sebesar 0 second dan 1 second yang menunjukkan bahwa operasinya sangat cepat, dapat dilihat pada gambar berikut:



Gambar 10 Hasil Operasi Pada Rapidminer

Dengan Tingkat akurasi sebesar 99.10%, algoritma naïve bayes memiliki Tingkat kekuatan yang sangat tinggi. Hasil akurasi tersebut menunjukkan bahwa pola tersebut dapat digunakan untuk mengklasifikasikan kepatuhan peserta BPJS Kesehatan. Hasil akurasi yang tinggi ini juga disebabkan oleh keakuratan data uji dan data latih yang berarti polanya dapat diprediksi dengan akurat. Adapun hasil akurasi dapat dilihat pada gambar berikut.

accuracy: 99.10% +/- 1.29% (micro average: 99.10%)

	true Sangat Tepat Waktu	true Tidak Tepat Waktu	true Tepat Waktu	class precision
pred. Sangat Tepat Waktu	325	0	0	100.00%
pred. Tidak Tepat Waktu	9	484	0	98.17%
pred. Tepat Waktu	0	0	182	100.00%
class recall	97.31%	100.00%	100.00%	

Gambar 6 Hasil Akurasi

SimpleDistribution

Distribution model for label attribute Estimasi

```

Class Sangat Tepat Waktu (0.334)
5 distributions

Class Tidak Tepat Waktu (0.484)
5 distributions

Class Tepat Waktu (0.182)
5 distributions

```

Gambar 7 Tabel Distribusi

Dari gambar diatas dapat dihasilkan bahwa data BPJS Kesehatan Kedepuitan Wilayah 1 Cabang Kota Medan, kelas tidak tepat waktu dengan nilai sebesar 0.484 lebih besar dari kelas sangat tepat waktu dan kelas tepat waktu Adapun model distribusi untuk label atribut kelas estimasi adalah sebagai berikut :

- Sangat tepat waktu : 5 distributions
- Tepat waktu : 5 distributions
- Tidak tepat waktu : 5 distributions

B. Hasil Pengujian

Setelah melakukan penelitian, diperoleh informasi baru dari proses data mining untuk mengetahui apakah peserta BPJS melakukan pembayaran tagihannya sangat tepat waktu, tepat waktu, atau tidak tepat waktu. Berdasarkan hasil penelitian, kita dapat mengetahui informasi dan pengetahuan tentang proses data mining untuk menentukan bagian mana yang sangat tepat waktu, tepat waktu, atau tidak tepat waktu. Setelah menyelesaikan tahapan KDD (*Knowledge Discovery In Database*) dari data peserta BPJS Kesehatan Kedepuitan Wilayah 1 Cabang Kota Medan tahun 2023, penelitian ini menemukan hasil yang sejalan dengan tujuan untuk menghasilkan informasi kepatuhan peserta BPJS Kesehatan yang mungkin berpotensi membayar tidak tepat pada waktunya menjadi tepat waktu.

Dari proses perhitungan data mining menggunakan algortima naive bayes dengan menggunakan teknik klasifikasi dan tingkat keakurasiannya, dihasilkan informasi berdasarkan perhitungan data peserta BPJS Kesehatan kedepuitan wilayah 1 cabang

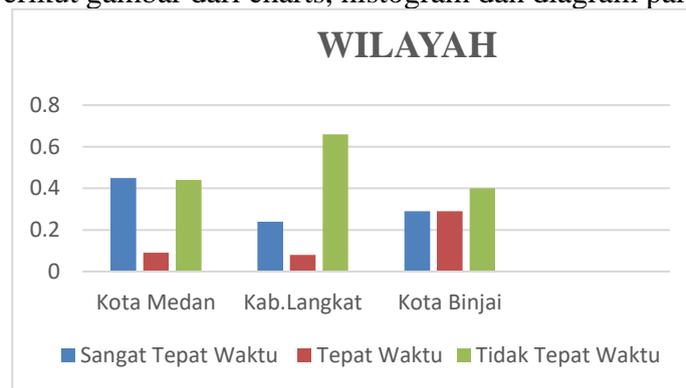
medan menunjukkan kelas sangat tepat waktu dengan perkalian prior probabilitas bernilai 0.334, kelas tepat waktu bernilai 0.182 dan kelas tidak tepat waktu bernilai 0.484 serta menurut perhitungan dari tools rapidminer wilayah Kab.Langkat dan Kota Binjai yang berpotensi tidak tepat waktu dengan akurasi sebesar 99.10%. salah satu hal yang mungkin membuat peserta BPJS membayar dengan tidak tepat waktu ialah besarnya tagihan BPJS Kesehatan tersebut dan juga faktor ekonomi yang mengakibatkan peserta BPJS membayar dengan tidak tepat waktu. Adapun data peserta BPJS dengan estimasi kepatuhannya sangat tepat waktu dapat dilihat pada tabel 1:

Tabel 1 Hasil

Estimasi	wilayah	Jumlah	Prior probability
Sangat tepat waktu	Kota medan	138	0.45
	Kabupaten langkat	60	0.24
	Kota binjai	136	0.29
Tepat waktu	Kota medan	29	0.09
	Kabupaten langkat	21	0.08
	Kota binjai	132	0.29
Tidak tepat waktu	Kota medan	134	0.44
	Kabupaten langkat	164	0.66
	Kota binjai	186	0.40

C. Knowledge Representation

Ini adalah langkah terakhir, di mana informasi yang telah diketahui ditunjukkan secara visual kepada pengguna melalui berbagai visualisasi, seperti charts, histograms, dan diagram Pareto. Berikut gambar dari charts, histogram dan diagram pareto:



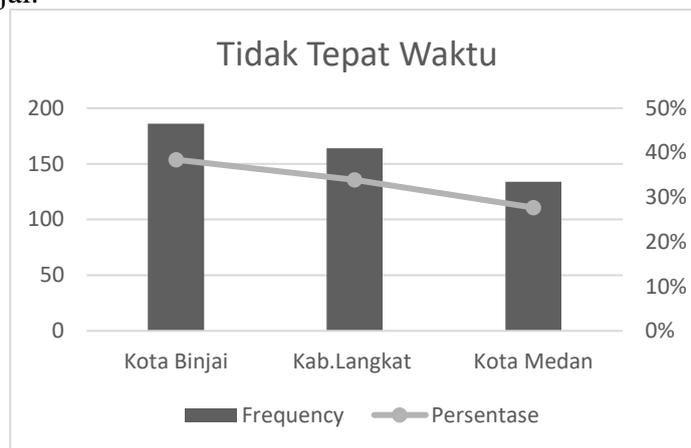
Gambar 1 Grafik 1

Grafik 1 memperlihatkan data wajib pajak dengan nilai Prior Probabilitas sangat tepat waktu, tepat waktu dan tidak tepat waktu.



Gambar 2 Grafik 2

Gambar diatas adalah grafik 2 yang menghasilkan sangat tepat waktu, tepat waktu dan tidak tepat waktu. Wilayah yang berpotensi tinggi untuk tidak tepat waktu yaitu wilayah Kota Binjai.



Gambar 3 Grafik 3

Gambar diatas adalah grafik 3 yang dihasilkan Kelas tidak tepat waktu terdapat pada wilayah kota binjai.

4. KESIMPULAN

Kesimpulan yang dapat diambil pada penelitian ini ialah penerapan Algoritma Naïve Bayes untuk melihat potensi kepatuhan peserta BPJS dalam membayar tagihan telah berhasil dilakukan dengan bantuan dari tools rapidminer. Berdasarkan perhitungan manual dengan algoritma naive bayes pada teknik klasifikasi diperoleh hasil kelas estimasi pembayaran tagihan peserta BPJS Kesehatan tidak tepat waktu lebih besar dari pada kelas estimasi sangat tepat waktu dan tepat waktu. Dari hasil observasi terhadap data BPJS Kesehatan ke deputian wilayah 1 cabang medan melalui teknik klasifikasi dengan algoritma naive bayes didapatkan wilayah yang berpotensi tidak tepat waktu terdapat pada sebagian peserta BPJS Kesehatan wilaya kota binjai dengan nilai sebanyak 0.484 dengan Tingkat akurasi sebesar 99.10%. yang menyebabkan peserta BPJS Kesehatan sulit dalam membayar tagihan perbulannya ialah karena semakin besarnya jumlah total tagihan dan besarnya total tagihan menunggak

DAFTAR PUSTAKA

- Alfani, V. (2019). Data Mining Untuk Klasifikasi Pinjaman Kredit Pensiunan Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Pelita Informatika*, 18(April), 281–286.
- Dianta, I. A. (2021). *Logika dan Algoritma* (M. . Danang, S.Kom. (ed.)). Yayasan Prima Agus Teknik Redaksi.
- Fajrini, F., Latifah, N., Hermansyah, D., & Firda, N. (2021). Studi Ketidakpatuhan Membayar Iuran BPJS Bukan Penerima Upah di Kelurahan Cempaka Putih. *Public Health Journal*. <https://media.neliti.com/media/publications/227635-peran-umkm-dalam-pembangunan-dan-kesejah-7d176a2c.pdf%0Ahttp://journal.uinsgd.ac.id/index.php/jispo/article/view/1752>
- Furqan, M., Nasution, Y. R., & Fadillah, R. (2022). Klasifikasi Penyakit Kulit Menggunakan Algoritma Naïve Bayes Berdasarkan Tekstur Warna Berbasis Android. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 6(1), 12–20.
- Haq, M. S., Fachrin, S. A., & Alwi, M. K. (2022). Faktor Yang Berhubungan Dengan Kepatuhan Membayar Iuran BPJS Peserta Mandiri Di Puskesmas Rappokalling Kota Makassar. *Window of Public Health Journal*, 3(4), 615–623. <https://doi.org/10.33096/woph.v3i4.464>
- Hermanto, H., Mustopa, A., & Kuntoro, A. Y. (2020). Algoritma Klasifikasi Naive Bayes Dan Support Vector Machine Dalam Layanan Komplain Mahasiswa. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, 5(2), 211–220. <https://doi.org/10.33480/jitk.v5i2.1181>
- Kurniawan, R. A., Hasibuan, M. S., Piramida, P., & Ramadhan, R. S. (2022). Penerapan Algoritma K-Means Untuk Clustering Tempat Makan Di Batubara. *Journal of Computer Science and Informatics Engineering (CoSIE)*, 01(1), 10–18. <https://doi.org/10.55537/cosie.v1i1.27>
- Novita, M. (2020). Faktor-Faktor Yang Berhubungan Dengan Kepatuhan Membayar Iuran BPJS Kesehatan Mandiri pada Pasien Rawat Jalan di Puskesmas Lubuk Basung Kab. Agam. *Study Fenomenologi Pengalaman Keluarga Pasien Dalam penanganan Prehospital Pasca Deteksi Dini Stroke*, 11(2), 282–289.
- Oon Wira Yuda, Darmawan Tuti, Lim Sheih Yee, S. (2022). *sinta 4 (8).pdf: Vol. Vol.8, No.*
- Prasetyo, M. A. M. B., Mirqotussa'adah, E. L. H. M. A. J. H., & Nurzahputra, S. H. R. A. (2019). *Data Mining Algoritma C4.5*.
- Rahman, A. (2023). Klasifikasi Performa Akademik Siswa Menggunakan Metode Decision Tree dan Naive Bayes. *Jurnal SAINTEKOM*, 13(1), 22–31. <https://doi.org/10.33020/saintekom.v13i1.349>
- Riany, A. F., & Testiana, G. (2023). Penerapan Data Mining untuk Klasifikasi Penyakit Stroke Menggunakan Algoritma Naïve Bayes. *Jurnal SAINTEKOM*, 13(1), 42–54. <https://doi.org/10.33020/saintekom.v13i1.352>
- Undamayanti, E., Iman Hermanto, T., Kaniawulan, I., Studi, P., Informatika, T., Teknologi, S. T., & Purwakarta, W. (2022). Analisis Sentimen Menggunakan Metode Naive Bayes Berbasis Particle Swarm Optimization Terhadap Pelaksanaan Program Merdeka Belajar Kampus Merdeka. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 6(2), 916–930.
- Widayati, Q. (2019). Penerapan Data Mining Menggunakan Metode Teknik Classification Untuk Melihat Potensi Kepatuhan Wajib Pajak Bumi Dan Bangunan. *Jurnal Ilmiah Matrik*, 20(2), 157–168. <https://doi.org/10.33557/jurnalmatrik.v20i2.119>
- Yunita, F. (2018). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru. *Sistemasi*, 7(3), 238. <https://doi.org/10.32520/stmsi.v7i3.388>