

Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM

Rizal Rachman¹, Rissa Nurfitriana Handayani²

^{1,2}Universitas Adhirajasa Reswara Sanjaya
Jl. Sekolah Internasional No 1-6 Antapani, Bandung, Indonesia

e-mail: ¹rizalrachman@ars.ac.id, ²rissa@ars.ac.id

Informasi Artikel

Diterima: 16-04-2021

Direvisi: 04-06-2021

Disetujui: 11-06-2021

Abstrak

Untuk meminimalisir jumlah penyewa teras untuk UMKM di depan toko Indomaret yang terlambat (*out standing*) setiap tahunnya dan untuk mengoptimalkan penyewa yang lancar setiap tahunnya. Salah satu cara yang bisa dilakukan manajemen PT. Indomarco PrismaTama untuk membantu menentukan prediksi kelancaran pembayaran sewa teras UMKM adalah dengan melakukan pengolahan data histori dari penyewa dengan memanfaatkan teknik *data mining* menggunakan algoritma *naïve bayes*. *Naïve Bayes* adalah suatu algoritma *data Mining* yang berfungsi memprediksi banyaknya penyewa yang terlambat pembayaran. Model yang dapat digunakan yaitu CRISP-DM, melalui pengenalan proses bisnis, pengenalan data, persiapan data, pemodelan, pengujian dan pengembangan. Algoritma yang digunakan untuk memprediksi probabilitas yaitu algoritma *Naïve Bayes*. Data yang digunakan berjumlah 504 data, dengan atribut yang sudah di klasifikasikan berdasarkan kelas yang dibutuhkan Area Manager, Kota, Jenis Kelamin Penyewa, Rata-rata Umur Penyewa, dan Status Pembayaran. Terdapat beberapa Probabilitas yang akan mengitung ke akuratan prediksinya, dan setelah di uji dengan menggunakan algoritma *naïve bayes*, maka diperoleh hasil persentasi *Accuracy* 81.81%, *Precision* 66.66%, *Recall* 100% dan *AUC* 0.800 Untuk keakuratan prediksi.

Kata Kunci: *Data Mining, Naïve Bayes, UMKM*

Abstract

To minimize the number of UMKM terrace tenants in front of the Indomaret store whose payments are late every year and to optimize tenants who have smoothness payments every year. One of the ways that the management of PT.Indomarco PrismaTama to help determine the prediction of smoothness payment of UMKM terrace rental is by processing historical data from tenants by utilizing data mining techniques using naïve bayes algorithm. Naïve Bayes is a data mining algorithm that function to predict the number of tenants who are late in payment. The method used is CRISP-DM, through the process business understanding, data understanding, data preparation, modeling, evaluation and deployment. The algorithm used to predict probability is the naïve bayes algorithm. The data used amount to 504 data, with attributes that have been classified based on the class required by the area manager, city, gender of the tenant, average age of the tenant, and payment status. There are several probabilities that will calculate the accuracy of the predictions, and after being tested using the naïve bayes algorithm, then the percentage results are obtained Accuracy 81.81%, Precision 66.66%, Recall 100% dan AUC 0.800 for prediction accuracy.

Keywords: *Data Mining, Naïve Bayes, UMKM*

1. Pendahuluan

Melihat perkembangan Koperasi dan Usaha Mikro Kecil dan Menengah (UMKM) masih butuh dukungan aturan-aturan *afirmasi* untuk Koperasi dan Usaha Mikro Kecil dan Menengah (UMKM) terhadap respon perubahan perekonomian pasar yang sangat dinamis. Koperasi dan Usaha

Mikro Kecil dan Menengah (UMKM) membutuhkan dukungan supaya bisa dimanfaatkan dalam memperbaiki susunan pengguna usaha di tingkat nasional sehingga menjadikan seimbang dan kuat, baik terhadap bidang usaha, strata atau sektoral (BPS, 2016).



Dalam perkembangan Usaha Mikro Kecil dan Menengah (UMKM) Indonesia, Pemerintah memberikan perhatian yang sangat besar. Dan sebagian bisnis yang lumayan berkembang untuk menghadapi kuatnya saingan pasar yaitu para penggiat UMKM (Sulistyawati, et al., 2018). Sebagian besar Usaha Mikro Kecil dan Menengah mempunyai hambatan dalam laporan keuangan dan administrasi maka dari itu laporan pembukuan masih belum rapih dan baik. Hal itu dikarena perkembangan Usaha Mikro Kecil dan Menengah diawali dari usaha pridiadi, selanjutnya berkembang serta membentuk suatu badan yaitu dengan berskala kecil dan menengah. Mengenai hal tersebut banyak dari mereka belum mempunyai cukup pengetahuan pada bidang admistrasi, pembukuan, dan pengetahuan dalam hal perpajakan (Ningrum, et al., 2016).

Tahapan model statistika yang bisa mengerjakan pengkategorian yaitu klasifikasi. Banyak pilihan metode klasifikasi dan model penelitian ini dapat menggunakan metode *naïve bayes*. Metode *naïve bayes* sudah banyak digunakan pada penelitian tentang *text mining*, salah satu kelebihan *naïve bayes* yaitu algoritma sederhana tetapi mempunyai nilai akurasi yang cukup tinggi. Nilai yang diperoleh saat tahap pengujian pada setiap pengukuran performa nilai akurasi, *precision*, *recall*, dan *F-Measure* sebesar 82,20%, 83,90%, 82,20%, dan 82,40%. Penggunaan data pengujian diperoleh untuk setiap pengukuran masing-masing performa akurasi, *precision*, *recall*, dan *F-Measure* sebesar 88,10%, 89,10%, 88,10%, dan 88,30% (ariadi, et al., 2015).

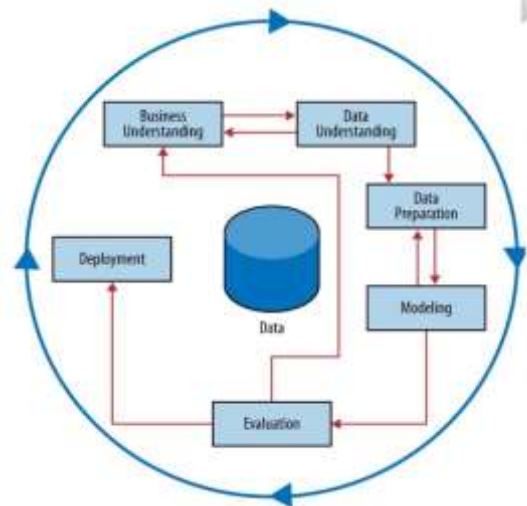
Saat ini jumlah gerai di Indomaret berjumlah kurang lebih 16.000 gerai, meliputi 40% gerai yang dimiliki waralaba dan juga 60% gerai yang dimiliki oleh perusahaan. Pemasok barang dagangan yang disalurkan ke setiap gerai sebagian besar berasal dari pusat distribusi Indomaret yang berjumlah 33 sudah menyediakan kurang lebih 5.000 tipe produk. Sekarang Indomaret didukung dengan kedatangan sewa teras mitra Usaha Mikro Kecil dan Menengah di depan seluruh gerai (Indomaret.co.id, 2018)

Biaya sewa teras merupakan komponen penting dalam memperhitungkan profit. Bilamana biaya sewa teras terlalu tinggi untuk ukuran usahanya, maka penyewa tidak bisa menggarap ladang tersebut dengan keuntungan, sekalipun disana ada emasnya. Untuk itulah di perlukan negosiasi yang baik dengan pihak lahan. Pada dasarnya pihak penyewa menginginkan lahannya berhasil digarap, banyak pengunjung yang datang, sehingga mendapatkan keuntungan finansial dan popularitas (Maitimu, 2017).

Penggunaan sesuatu dengan uang yang dibayarkan atau membayar uang karena meminjamkan atau memakai sesuatu. Sewa menyewa merupakan suatu perjanjian bersama dengan seseorang atau pihak yang berjanji untuk menyerahkan barang dalam waktu yang sudah dijanjikan dengan membayarkan suatu tempat dan harga oleh seseorang atau pihak terakhir sudah disanggupi suatu pembayarannya (Pradana, Waspada, 2019).

Data mining adalah suatu proses semi otomatis yang memakai suatu teknik matematika, statistik, *machine learning*, dan kecerdasan buatan untuk mengidentifikasi dan mengekstraksi suatu informasi pengetahuan yang terkait dari macam-macam basis data yang sangat besar (Maharani, et al., 2017).

Kerangka kerja CRISP-DM (*Cross-Industri Standar for Data mining*) adalah salah satu metodologi *data mining* yang sudah tersusun oleh beberapa Anggota Komisi Eropa pada tahun 1996 dan sudah ditetapkan sebagai suatu standarisasi proses dalam *data mining*. *Data mining* memiliki enam tahapan CRISP-DM yaitu fase pemahaman bisnis, pemahaman data, pengolahan data, pemodelan, evaluasi, penyebaran (Setiawan, 2016).



Sumber : (Setiawan, 2016).

Gambar 1. Fase CRISP-DM

Naive Bayes adalah suatu klasifikasi kemungkinan sederhana yang dapat menghitung seluruh kemungkinan dengan menggabungkan sejumlah kombinasi dan frekuensi suatu nilai dari basis data yang didapatkan. Suatu algoritma memanfaatkan *teorema bayes* dan memperkirakan seluruh atribut yang bebas dan saling lepas yang dapat diberikan oleh suatu nilai pada kelas variabel. *naive bayes* adalah klasifikasi dengan suatu metode kemungkinan dan perhitungan yang ditemukan oleh seseorang ilmuwan dari Inggris

yaitu Thomas Bayes menghasilkan prediksi peluang yang akan datang berdasarkan suatu pengalaman sebelumnya (Saleh, 2015).

Teorema Bayes memiliki bentuk umum sebagai berikut :

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

Di mana :

- X = Kelas data yang belum diketahui
- H = Hipotesa data X adalah kelas spesifik
- P(H|X) = Kemungkinan Hipotesa H berdasarkan keadaan X (posteriori prob.)
- P(H) = Kemungkinan Hipotesa H (prior prob.)
- P(X|H) = Kemungkinan X berdasarkan keadaan tersebut
- P(X) = Kemungkinan dari X

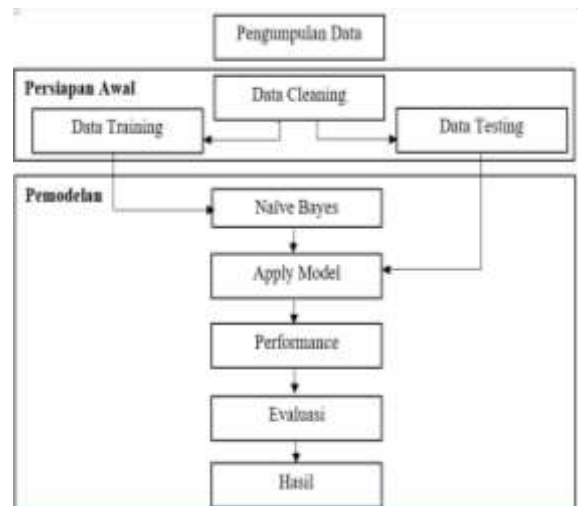
Metode algoritma *Naïve Bayes* dan *support Vektor Machine (SVM)* untuk melakukan pengujian dalam memprediksi suatu keberhasilan pada pengobatan imunoterapi agar penyakit kutil dengan memanfaatkan bahasa pemrograman R pada R Studio, Hasil pengujian tersebut menunjukkan bahwa metode *Naïve Bayes* menghasilkan nilai yang lebih unggul dibandingkan dengan metode *Support Vektor Machine (SVM)* (Supriyatna, Mustika, 2018). Untuk perbandingan penetapan status penanganan kecelakaan kerja. Dari hasil kinerja pengklasifikasian yang baik dengan memanfaatkan metode *Naïve Bayes* dan *Support Vektor Machine (SVM)*, menghasilkan nilai akurasi yang lebih baik (Puridewi, Nugraha, 2018).

Hasil penerapan C4.5 dan *Naïve Bayes* dalam metode klasifikasi penerimaan beasiswa dalam penelitian ini menunjukkan perbandingan kinerja model algoritma *Naïve Bayes* menghasilkan nilai lebih baik dan tingkat nilai akurasi metode algoritma *Naïve Bayes* berjumlah 95.11% dalam waktu 0s. (Anam, Santoso, 2018). Untuk mengetahui fitur seleksi hubungan variabel nilai dan latar pendidikan. Hal ini dapat di tunjukan dengan adanya kenaikan performa dari pengujian algoritma *naïve bayes* waktu 1s dengan akurasi 64,77% dan algoritma *naïve bayes – forward selection* waktu 6s dengan akurasi 78.08%. Kenaikan performa yang didapat sebesar 13.31% (Astuti, et al., 2018). Dalam menentukan seleksi penerima beasiswa dengan metode *naïve bayes classifier*. Pengujian kelayakan berdasarkan dari uji testing 1 menghasilkan sebesar 96,56% dan uji testing 2 dengan presentase sebesar 90,33%, mencakup pada kategori yang layak. Sistem kelayakan yang digunakan oleh suatu lembaga pendidikan supaya dapat meningkatkan kemanfaatan dan kemudahan dengan nilai akurasi berdasarkan klasifikasi yang cukup tinggi (Rahman, Suryanto, 2017).

Nasabah potensial pada AJB Bumi Putra menggunakan algoritma *Naïve Bayes*. Model algoritma *Naïve Bayes* mampu membantu banyak pihak khususnya Bumi Putra agar menggunakan data dengan baik seperti untuk pengklasifikasian basis data nasabah potensial, pengklasifikasian mampu digunakan data yang perorangan atau data yang berkelompok. (Pakpahan, Irawan, 2017). Perkiraan tingkat kredit yang dilakukan oleh suatu bank sudah meraih keberhasilan yang sudah dicapai pada prediksi tingkat kemudahn dalam pembayaran kredit, metode algoritma *Naïve Bayes* dengan jenis *Forward Selection* dapat memperkirakan kemudahan dalam pembayaran kredit selanjutnya yang dapat dibuktikan dengan nilai akurasi *Naïve Bayes* model *Forward Selection* dapat mencapai akurasi 71.97% (Hasan, 2017).

2. Metode Penelitian

Data penelitian ini menggunakan data primer yang diperoleh langsung saat studi kasus di PT.Indomarco Prismatama. Data yang digunakan berupa data sewa teras UMKM di halaman toko Indomaret cabang Bandung pada tahun 2018. Dan tahapan yang akan disajikan mengenai metodologi penelitian keseluruhan yang didasari pada *Cross Industry Standard Proses For Data Mining (CRISP-DM)*. Berikut adalah tahapan yang akan dilakukan dalam penelitian ini seperti yang terlihat pada gambar 2.



Sumber : (Hasil Analisa)
 Gambar 2 : Desain Penelitian

Berdasarkan Gambar 2 tersebut mengenai model *Cross Industry Standard Proses For Data Mining (CRISP-DM)* memiliki 6 tahapan yaitu proses bisnis yang ada di PT.Indomarco Prismatama, pemahaman dan pengumpulan data, persiapan data untuk seleksi dan pembersihan, pemodelan dengan algoritma *Naïve Bayes*, *Apply Model* dan *Performance*,

evaluasi dengan pengujian data dan yang terakhir pengembangan hasil.

3. Hasil dan Pembahasan

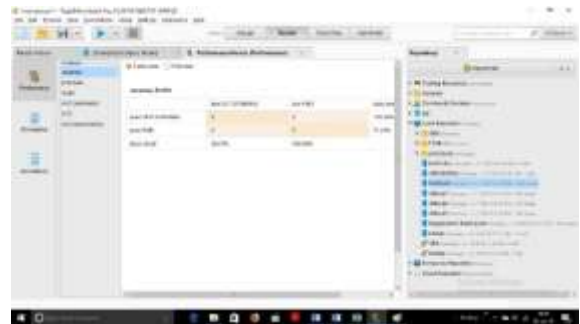
Dalam pengumpulan data akan digunakan dalam *Training* dan *Testing* untuk penelitian ini adalah data *privat* yang belum pernah di analisa. Data ini mempunyai 504 *Record* dengan 13 atribut (yang akan dipakai dengan mengklasifikasikan menjadi 6 atribut). Dan parameter yang menentukan kelancaran dalam pembayaran sewa teras UMKM adalah AM (Area Manager), kota, jenis kelamin penyewa, rata-rata umur penyewa, rata-rata biaya sewa, dan status pembayaran (*paid* dan *outstanding*). *Sample* data yang digunakan untuk penelitian ini adalah data penyewa sewa teras UMKM di halaman teras toko Indomaret cabang Bandung sebanyak 504 data yang terdiri dari 13 atribut dan yang telah di klasifikasi menjadi 6 atribut.

Tabel 1. Data Training

Area Manager	Kota	Jenis Kelamin	Rata-Rata Umur Penyewa	Rata-Rata Biaya Sewa	Status Pembayaran
Adp	Kab Bandung	Pria	Umur < 40	Sewa > 500000	Out Standing
Adp	Kab Bandung	Pria	Umur < 40	Sewa = 500000	Paid
Adp	Kab Bandung	Pria	Umur < 40	Sewa < 500000	Paid
Adp	Kota Cimahi	Wanita	Umur < 40	Sewa = 500000	Out Standing
Adp	Kota Cimahi	Pria	Umur > 40	Sewa > 500000	Paid
Ags	Kab Bandung	Pria	Umur < 40	Sewa > 500000	Out Standing
Ags	Kab Bandung	Wanita	Umur < 40	Sewa > 500000	Out Standing
Ags	Kab Bandung	Pria	Umur > 40	Sewa = 500000	Paid
Ags	Kab Bandung	Wanita	Umur > 40	Sewa > 500000	Paid
Bcp	Kab Bandung	Pria	Umur > 40	Sewa = 500000	Out Standing
Bcp	Kab Bandung	Pria	Umur > 40	Sewa > 500000	Out Standing

Sumber : Data sewa teras di toko Indomaret

Dalam tahap ini memperlihatkan dimana pengujian model menggunakan *Software RapidMiner Version 9.2* terhadap algoritma *Naïve Bayes* dengan menggunakan metode *validasi Apply model* dan *Performance*. Untuk pengujian ini menggunakan *dataset* primer yang didapatkan dalam studi kasus di PT.Indomarco Prismatama untuk di implementasikan pada proses pengujian model kemudiann *dataset* diujikan pada algoritma *Naïve Bayes* dengan metode *validasi Apply Model* dan *Performance*. Setelah dilakukan uji model maka diperoleh hasil *Accuracy* sebesar 81.82%, *Precision* sebesar 100%, *recall* sebesar 66.67%. Ditunjukkan dengan gambar 3 di bawah ini.



Gambar 3. Hasil *Accuracy* menggunakan Algoritma *Naïve Bayes*

Dari *Confusion Matrix* pada gambar 3. dapat diketahui probabilitas atau peluang menggunakan rumus *Theorema Bayes* di perlihatkan di tabel 2 .sebagai berikut:

Tabel 2. *Confusion Matrix*

	True Out Standing	True Paid	Class Precision
Pred. Out Standing	4 (True Positive= TP)	0 (False Negative= FN)	100%
Pred. Paid	2 (False Positive= FP)	5 (True Negative= TN)	71.43%
Class Recall	66.67%	100%	

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} * 100$$

$$= \frac{4+5}{4+0+2+5} * 100$$

$$= \frac{9}{11} * 100$$

$$= 0.8181 * 100$$

$$= 81.81 \%$$

$$\text{Precision} = \frac{TP}{TP+FP} * 100$$

$$= \frac{4}{4+2} * 100$$

$$= \frac{4}{6} * 100$$

$$= 0.66666 * 100$$

$$= 66.66 \%$$

$$\text{Recall} = \frac{TP}{TP+FN} * 100$$

$$= \frac{4}{4+0} * 100$$

$$= \frac{4}{4} * 100$$

$$= 1 * 100$$

$$= 100 \%$$

Keterangan :

TP = Prediksi positif yang positif

FN = Prediksi positif yang negative

FP = Prediksi negative yang positif

TN = Prediksi negative yang negative



Gambar 1. Hasil AUC dengan menggunakan Algoritma *Naïve Bayes*

Pada gambar 4. menunjukan dari hasil akurasi untuk klasifikasi prediksi *dataset* sewa teras UMKM di toko Indomaret dapat diperoleh dengan penerapan metode algoritma *naïve*

bayes. Pada penelitian ini diperoleh hasil *Accuracy* 81.81%, *Precision* 66.66 %, dan *Recall* 100%. Pada gambar 4.2 menghasilkan nilai AUC sebesar 0.88-1.00 yang termasuk kedalam kategori *Excellent Classification Prediction*.

Untuk mencari nilai akurasi pada data testing dalam pemodelan *Naïve Bayes* dilakukan dengan cara perhitungan Teorema Bayes. Dimana *dataset* atau data testing di klasifikasikan probabilitas tiap variabel kelasnya. Data yang digunakan sebanyak 504 data sewa teras UMKM di toko Indomaret.

Berikut adalah contoh klasifikasi dengan menggunakan perhitungan manual untuk salah satu testing data dengan menggunakan rumus Teorema Bayes. Ditunjukkan pada tebal 3. Berikut ini.

Tabel 3
Data Testing yang belum di Klasifikasikan

NO	AM	TOKO	RUPIAH SEWA/BULAN			PEMBAYARAN
			LAHAN	GEROKBANK	TOTAL SEWA	
1	YWA	TGU3	500,00	-	500,00	OUTSTANDING
2	EH O	FK42	550,00	-	550,00	PAID
3	HY O	FN0K	450,00	-	450,00	OUTSTANDING
4	PCA	T355	350,00	-	350,00	PAID
5	PCA	T355	350,00	-	350,00	PAID
6	PCA	T355	350,00	-	350,00	PAID
7	TFI	TH97	750,00	-	750,00	OUTSTANDING
8	HFI	TH14	550,00	-	550,00	PAID
9	FRY	TRX4	1,600,000	-	1,600,000	OUTSTANDING
10	IRD	FULN	1,400,000	-	1,400,000	OUTSTANDING

Tabel 1.
Data Yang Sudah di Klasifikasikan (Data Uji)

Area Mana ger	Kota	Jenis Kelamin	Rata - Rata Umur Penyewa	Rata-Rata Biaya Sewa	Status Pembayaran
Adp	Kab Bandung	Pria	Umur < 40	Sewa > 500000	Out Standing
Adp	Kab Bandung	Pria	Umur < 40	Sewa = 500000	Paid
Adp	Kab Bandung	Pria	Umur < 40	Sewa < 500000	Paid
Adp	Kota Cima hi	Wanita	Umur < 40	Sewa = 500000	Out Standing
Adp	Kota Cima hi	Pria	Umur > 40	Sewa > 500000	Paid
Aggs	Kab Bandung	Pria	Umur < 40	Sewa > 500000	Out Standing
Aggs	Kab Bandung	Wanita	Umur < 40	Sewa > 500000	Out Standing
Aggs	Kab Bandung	Pria	Umur > 40	Sewa = 500000	Paid
Aggs	Kab Bandung	Wanita	Umur > 40	Sewa > 500000	Paid
Bcp	Kab Bandung	Pria	Umur > 40	Sewa = 500000	Out Standing
Bcp	Kab Bandung	Pria	Umur > 40	Sewa > 500000	Out Standing

Setelah data uji diklasifikasikan dan dibuatkan data probabilitas setiap atribut supaya bisa menghasilkan nilai yang akan menjadi bahan perhitungan nilai *Accuracy*, *Precision*, dan *Recall*. Berikut data probabilitas.

Tabel 5.
Probabilitas Jenis Kelamin Penyewa

Jenis Kelamin	Paid	Out Standing
Pria	74%	74%
Wanita	26%	26%
Total	100%	100%

Berdasarkan tabel 5 yang sudah diklasifikasi berdasarkan kelas masing-masing maka untuk mencari nilainya berapa persen disetiap probabilitasnya. Perhitungan mencari persentase peluang status pembayaran di setiap

probabilitas. Contohnya pada probabilitas jenis kelamin. Maka untuk mencari berapa persen pada probabilitas jenis kelamin pria dan wanita yang status pembayarannya *paid* dan *out standing* sebagai berikut.

Keterangan :

A = Jumlah jenis kelamin laki-laki = 375

B = Jumlah jenis kelamin wanita = 129

C = Jumlah status paid = 277

D = Jumlah status Out Standing = 227

E = Jumlah jenis kelamin laki-laki, status paid = 206

F = Jumlah jenis kelamin laki-laki, status Out Standing = 169

G = Jumlah jenis kelamin wanita, status paid = 71

H = Jumlah jenis kelamin wanita, status Out Standing = 58

$$P. \text{ Pria, Paid} = \frac{E}{C} * 100 = \frac{206}{277} * 100$$

$$= 0.74 * 100$$

$$= 74 \%$$

$$P. \text{ Pria, Out Standing} = \frac{F}{D} * 100 = \frac{169}{227} * 100$$

$$= 0.74 * 100$$

$$= 74 \%$$

$$P. \text{ Wanita, Paid} = \frac{G}{C} * 100 = \frac{71}{277} * 100$$

$$= 0.256 * 100$$

$$= 25.6 \%$$

$$P. \text{ Wanita, Out Standing} = \frac{H}{D} * 100 = \frac{58}{227} * 100$$

$$= 0.255 * 100$$

$$= 25.5 \%$$

Tabel 6.
Probabilitas Rata-rata Umur Penyewa

Rata-Rata Umur Penyewa	Paid	Out Standing
Umur < 40	68%	78%
Umur > 40	32%	22%
TOTAL	100%	100%

Berdasarkan tabel 6. di atas yang sudah diklasifikasi berdasarkan kelas masing-masing maka untuk mencari nilainya berapa persen disetiap probabilitasnya. Perhitungan mencari persentase peluang status pembayaran di setiap probabilitas. Contoh pada probabilitas rata-rata umur penyewa. Maka untuk mencari berapa

persen pada probabilitas rata-rata umur <40 tahun dan >40 tahun yang status pembayarannya *paid* dan *out standing* sebagai berikut.

Keterangan :

- A = Jumlah umur <40 tahun = 365
- B = Jumlah umur >40 tahun = 139
- C = Jumlah status paid = 277
- D = Jumlah status Out Standing = 227
- E = Jumlah umur <40 th, status paid = 189
- F = Jumlah umur <40 th, status Out Standing = 36
- G = Jumlah umur >40 th, status paid = 88
- H = Jumlah umur >40 th, status Out Standing = 51

$$\begin{aligned} \text{P. umur <40 th, Paid} &= \frac{E}{C} * 100 \\ &= \frac{189}{277} * 100 \\ &= 0.68 * 100 \\ &= 68 \% \end{aligned}$$

$$\begin{aligned} \text{P. umur <40 th, Out Standing} &= \frac{F}{D} * 100 = \frac{36}{227} * 100 \\ &= 0.158 * 100 \\ &= 15.8 \% \end{aligned}$$

$$\begin{aligned} \text{P. umur >40 th, Paid} &= \frac{G}{C} * 100 = \frac{88}{277} * 100 \\ &= 0.317 * 100 \\ &= 31.7 \% \end{aligned}$$

$$\begin{aligned} \text{P. Wanita, Out Standing} &= \frac{H}{D} * 100 = \frac{51}{227} * 100 \\ &= 0.224 * 100 \\ &= 22.4 \% \end{aligned}$$

Tabel 7.
 Probabilitas Rata-rata Biaya Sewa

Rata-Rata Sewa	Biaya	Paid	Out Standing
SEWA < 500000		9%	3%
SEWA > 500000		78%	87%
SEWA = 500000		13%	10%
TOTAL		100%	100%

Berdasarkan tabel 7. di atas yang sudah diklasifikasi berdasarkan kelas masing-masing maka untuk mencari nilainya berapa persen disetiap probabilitasnya. Perhitungan mencari persentase peluang status pembayaran di setiap probabilitas. Contoh pada probabilitas rata-rata biaya sewa. Maka untuk mencari berapa persen pada probabilitas rata-rata sewa <500.000, =500.000 dan >500.000 yang status

pembayarannya *paid* dan *out standing* sebagai berikut.

Keterangan :

- A = Jumlah rata-rata < 500.000 = 365
- B = Jumlah rata-rata = 500.000 = 139
- C = Jumlah rata-rata > 500.000 = 139
- D = Jumlah status paid = 277
- E = Jumlah status Out Standing = 227
- F = Jumlah rata-rata < 500.000, status paid = 25
- G = Jumlah rata-rata < 500.000, status Out Standing = 7
- H = Jumlah rata-rata = 500.000, status paid = 35
- I = Jumlah rata-rata = 500.000, status Out Standing = 23
- J = Jumlah rata-rata > 500.000, status paid = 217
- K = Jumlah rata-rata > 500.000, status Out Standing = 197

$$\begin{aligned} \text{P. rata-rata < 500.000, Paid} &= \frac{F}{D} * 100 \\ &= \frac{25}{277} * 100 \\ &= 0.09 * 100 \\ &= 9,02 \% \end{aligned}$$

$$\begin{aligned} \text{P. rata-rata < 500.000, Out Standing} &= \frac{G}{E} * 100 \\ &= \frac{7}{227} * 100 \\ &= 0.030 * 100 \\ &= 3.08 \% \end{aligned}$$

$$\begin{aligned} \text{P. rata-rata = 500.000, Paid} &= \frac{H}{D} * 100 \\ &= \frac{35}{277} * 100 \\ &= 0.126 * 100 \\ &= 12.6 \% \end{aligned}$$

$$\begin{aligned} \text{P. rata-rata = 500.000, Out Standing} &= \frac{I}{E} * 100 \\ &= \frac{23}{227} * 100 \\ &= 0.101 * 100 \\ &= 10,1 \% \end{aligned}$$

$$\begin{aligned} \text{P. rata-rata > 500.000, Paid} &= \frac{J}{D} * 100 \\ &= \frac{217}{277} * 100 \\ &= 0.783 * 100 \\ &= 78,3 \% \end{aligned}$$

$$\text{P. rata-rata > 500.000, Out Standing}$$

$$= \frac{K}{E} * 100$$

$$= \frac{197}{227} * 100$$

$$= 0.867 * 100$$

$$= 86.7 \%$$

Tabel 8. Probabilitas Kota

Kota	Paid	Out Standing
Kab Bandung	33%	38%
Kodya Bandung	50%	47%
Kab Garut	6%	12%
Kab Tasikmalaya	1%	2%
Kota Cimahi	10%	1%
Total	100%	100%

Berdasarkan tabel 8 di atas yang sudah diklasifikasi berdasarkan kelas masing-masing maka untuk mencari nilainya berapa persen disetiap probabilitasnya. Perhitungan mencari persentase peluang status pembayaran di setiap probabilitas. Contoh pada probabilitas kota. Maka untuk mencari berapa persen pada probabilitas Kab Bandung, Kodya Bandung, Kab Garut, Kab Tasikmalaya, dan Kota Cimahi yang status pembayarannya *paid* dan *out standing* sebagai berikut.

Keterangan :

A = Jumlah Kab.Bandung = 178

B = Jumlah Kodya Bandung = 244

C = Jumlah Kab.Garut = 46

D = Jumlah Kab.Tasikmalaya = 6

E = Jumlah Kota Cimahi = 30

F = Jumlah status paid = 277

G = Jumlah status Out Standing = 227

H = Jumlah Kab.Bandung, status paid = 93

I = Jumlah Kab.Bandung, status Out Standing = 86

J = Jumlah Kodya Bandung, status paid = 138

K = Jumlah Kodya Bandung, status Out Standing = 106

L = Jumlah Kab.Garut, status paid = 18

M = Jumlah Kab.Garut, status Out Standing = 28

N = Jumlah Kab.Tasikmalaya, status paid = 2

O = Jumlah Kab.Tasikmalaya, status Out Standing = 4

P = Jumlah Kota Cimahi, status paid = 27

Q = Jumlah Kota Cimahi, status Out Standing = 3

P. Kab.Bandung,Paid

$$= \frac{H}{F} * 100$$

$$= \frac{93}{277} * 100$$

$$= 0.335 * 100$$

$$= 33.5 \%$$

P. Kab.Bandung, Out Standing

$$= \frac{I}{G} * 100$$

$$= \frac{86}{227} * 100$$

$$= 0.378 * 100$$

$$= 37.8 \%$$

P. Kodya Bandung,Paid

$$= \frac{J}{F} * 100$$

$$= \frac{138}{277} * 100$$

$$= 0.498 * 100$$

$$= 49.8 \%$$

P. Kodya Bandung, Out Standing

$$= \frac{K}{G} * 100$$

$$= \frac{106}{227} * 100$$

$$= 0.466 * 100$$

$$= 46,6 \%$$

P. Jumlah Kab.Garut,Paid

$$= \frac{L}{F} * 100$$

$$= \frac{18}{277} * 100$$

$$= 0.064 * 100$$

$$= 6.49 \%$$

P. Jumlah Kab.Garut, Out Standing

$$= \frac{M}{G} * 100$$

$$= \frac{28}{227} * 100$$

$$= 0.123 * 100$$

$$= 12.33 \%$$

P. Jumlah Kab.Tasikmalaya,Paid

$$= \frac{N}{F} * 100$$

$$= \frac{2}{277} * 100$$

$$= 0.007 * 100$$

$$= 0.72 \%$$

P. Jumlah Kab.Tasikmalaya, Out Standing

$$= \frac{O}{G} * 100$$

$$= \frac{4}{227} * 100$$

$$= 0.017 * 100$$

$$= 1.76 \%$$

P. Jumlah Kota Cimahi, Paid

$$\begin{aligned} &= \frac{P}{F} * 100 \\ &= \frac{27}{277} * 100 \\ &= 0.097 * 100 \\ &= 9,7 \% \end{aligned}$$

P. Jumlah Kota Cimahi, Out Standing

$$\begin{aligned} &= \frac{Q}{G} * 100 \\ &= \frac{3}{227} * 100 \\ &= 0.013 * 100 \\ &= 1.32 \% \end{aligned}$$

Tabel 9. Probabilitas Area Manager

Area Manager	Paid	Out Standing
ADP	5%	1%
AGS	6%	6%
BCP	5%	9%
EHO	7%	7%
FRY	5%	8%
HES	3%	4%
HFI	11%	1%
HYO	3%	4%
ICF	13%	15%
IRD	4%	12%
PCA	12%	3%
SGP	4%	10%
TFI	11%	5%
THS	4%	8%
YWA	8%	6%
TOTAL	100%	100%

Berdasarkan tabel 9. di atas yang sudah diklasifikasi berdasarkan kelas masing-masing maka untuk mencari nilainya berapa persen disetiap probabilitasnya. Perhitungan mencari persentase peluang status pembayaran di setiap probabilitas. Contoh pada probabilitas area manager. Maka untuk mencari berapa persen pada probabilitas ADP, AGS, BCP, EHO, FRY, HES, HFI, HYO, ICF, IRD, PCA, SGP, TFI, THS, dan YWA yang status pembayarannya *paid* dan *out standing* sebagai berikut.

Keterangan :

- A = Jumlah ADP = 18
- B = Jumlah AGS = 30
- C = Jumlah BCP = 34
- D = Jumlah EHO = 36
- E = Jumlah FRY = 33

F = Jumlah HES = 15

- G = Jumlah HFI = 33
- H = Jumlah HYO = 17
- I = Jumlah ICF = 70
- J = Jumlah IRD = 37
- K = Jumlah PCA = 39
- L = Jumlah SGP = 35
- M = Jumlah TFI = 43
- N = Jumlah THS = 30
- O = Jumlah YWA = 34

- P = Jumlah status paid = 277
- Q = Jumlah status Out Standing = 227
- R = Jumlah ADP, status paid = 15
- S = Jumlah ADP, status Out Standing = 3
- T = Jumlah AGS, status paid = 17
- U = Jumlah AGS, status Out Standing = 13
- V = Jumlah BCP, status paid = 13
- W = Jumlah BCP, status Out Standing = 21
- X = Jumlah EHO, status paid = 19
- Y = Jumlah EHO, status Out Standing = 17
- aa = Jumlah FRY, status paid = 14
- ab = Jumlah FRY, status Out Standing = 19
- ac = Jumlah HES, status paid = 7
- ad = Jumlah HES, status Out Standing = 8
- ae = Jumlah HFI, status paid = 30
- af = Jumlah HFI, status Out Standing = 3
- ag = Jumlah HYO, status paid = 8
- ah = Jumlah HYO, status Out Standing = 9
- aj = Jumlah ICF, status paid = 35
- ak = Jumlah ICF, status Out Standing = 35
- al = Jumlah IRD, status paid = 10
- am = Jumlah IRD, status Out Standing = 27
- an = Jumlah PCA, status paid = 33
- ap = Jumlah PCA, status Out Standing = 6
- aq = Jumlah SGP, status paid = 12
- ar = Jumlah SGP, status Out Standing = 23
- ba = Jumlah TFI, status paid = 31
- bs = Jumlah TFI, status Out Standing = 12
- as = Jumlah THS, status paid = 12
- at = Jumlah THS, status Out Standing = 18
- av = Jumlah YWA, status paid = 21
- ax = Jumlah YWA, status Out Standing = 13

$$\begin{aligned} \text{P. ADP, Paid} &= \frac{R}{P} * 100 \\ &= \frac{15}{277} * 100 \\ &= 0.054 * 100 \\ &= 5.41 \% \end{aligned}$$

$$\begin{aligned} \text{P. ADP, Out Standing} &= \frac{S}{Q} * 100 \\ &= \frac{3}{227} * 100 \\ &= 0.013 * 100 \\ &= 1.32 \% \end{aligned}$$

$$\begin{aligned} \text{P. AGS, Paid} &= \frac{T}{P} * 100 \\ &= \frac{17}{277} * 100 \end{aligned}$$

	$= 0.062 * 100$		$= \frac{8}{227} * 100$
	$= 6.13 \%$		$= 0.035 * 100$
P. AGS, Out Standing	$= \frac{U}{Q} * 100$		$= 3.52 \%$
	$= \frac{13}{227} * 100$	P. HFI, Paid	$= \frac{ac}{P} * 100$
	$= 0.057 * 100$		$= \frac{30}{277} * 100$
	$= 5,72 \%$		$= 0.108 * 100$
P. BCP, Paid	$= \frac{V}{P} * 100$		$= 10.83 \%$
	$= \frac{13}{277} * 100$	P. HFI, Out Standing	$= \frac{af}{Q} * 100$
	$= 0.046 * 100$		$= \frac{3}{227} * 100$
	$= 4.69 \%$		$= 0.013 * 100$
P. BCP, Out Standing	$= \frac{W}{Q} * 100$		$= 1,32 \%$
	$= \frac{21}{227} * 100$	P. HYO, Paid	$= \frac{ag}{P} * 100$
	$= 0.092 * 100$		$= \frac{8}{277} * 100$
	$= 9,25 \%$		$= 0.028 * 100$
P. EHO, Paid	$= \frac{X}{P} * 100$		$= 2.88 \%$
	$= \frac{19}{277} * 100$	P. HYO, Out Standing	$= \frac{ah}{Q} * 100$
	$= 0.068 * 100$		$= \frac{9}{227} * 100$
	$= 6.85 \%$		$= 0.039 * 100$
P. EHO, Out Standing	$= \frac{Y}{Q} * 100$		$= 3.96 \%$
	$= \frac{17}{227} * 100$	P. ICF, Paid	$= \frac{ai}{P} * 100$
	$= 0.074 * 100$		$= \frac{35}{277} * 100$
	$= 7.48 \%$		$= 0.126 * 100$
P. FRY, Paid	$= \frac{aa}{P} * 100$		$= 12.63 \%$
	$= \frac{14}{277} * 100$	P. ICF, Out Standing	$= \frac{ak}{Q} * 100$
	$= 0.050 * 100$		$= \frac{35}{227} * 100$
	$= 5.05 \%$		$= 0.154 * 100$
P. FRY, Out Standing	$= \frac{ab}{Q} * 100$		$= 15.41 \%$
	$= \frac{19}{227} * 100$	P. IRD, Paid	$= \frac{al}{P} * 100$
	$= 0.083 * 100$		$= \frac{10}{277} * 100$
	$= 8.37 \%$		$= 0.036 * 100$
P. HES, Paid	$= \frac{ac}{P} * 100$		$= 3.61 \%$
	$= \frac{7}{277} * 100$	P. IRD, Out Standing	$= \frac{am}{Q} * 100$
	$= 0.025 * 100$		$= \frac{27}{227} * 100$
	$= 2.52 \%$		$= 0.118 * 100$
P. HES, Out Standing	$= \frac{ad}{Q} * 100$		$= 11.89 \%$

$$\begin{aligned} \text{P. PCA, Paid} &= \frac{an}{P} * 100 \\ &= \frac{33}{277} * 100 \\ &= 0.119 * 100 \\ &= 11.91 \% \end{aligned}$$

$$\begin{aligned} \text{P. PCA, Out Standing} &= \frac{ap}{Q} * 100 \\ &= \frac{6}{227} * 100 \\ &= 0.026 * 100 \\ &= 2.64 \% \end{aligned}$$

$$\begin{aligned} \text{P. SGP, Paid} &= \frac{aq}{P} * 100 \\ &= \frac{12}{277} * 100 \\ &= 0.043 * 100 \\ &= 4.33 \% \end{aligned}$$

$$\begin{aligned} \text{P. SGP, Out Standing} &= \frac{ar}{Q} * 100 \\ &= \frac{23}{227} * 100 \\ &= 0.101 * 100 \\ &= 10.13 \% \end{aligned}$$

$$\begin{aligned} \text{P. TFI, Paid} &= \frac{ba}{P} * 100 \\ &= \frac{31}{277} * 100 \\ &= 0.111 * 100 \\ &= 11.19 \% \end{aligned}$$

$$\begin{aligned} \text{P. TFI, Out Standing} &= \frac{bs}{Q} * 100 \\ &= \frac{12}{227} * 100 \\ &= 0.05 * 100 \\ &= 5.28 \% \end{aligned}$$

$$\begin{aligned} \text{P. THS, Paid} &= \frac{as}{P} * 100 \\ &= \frac{12}{277} * 100 \\ &= 0.043 * 100 \\ &= 4.33 \% \end{aligned}$$

$$\begin{aligned} \text{P. THS, Out Standing} &= \frac{at}{Q} * 100 \\ &= \frac{18}{227} * 100 \\ &= 0.079 * 100 \\ &= 7.92 \% \end{aligned}$$

$$\begin{aligned} \text{P. YWA, Paid} &= \frac{ay}{P} * 100 \\ &= \frac{21}{277} * 100 \\ &= 0.075 * 100 \end{aligned}$$

$$\begin{aligned} \text{P. YWA, Out Standing} &= \frac{ax}{Q} * 100 \\ &= \frac{13}{227} * 100 \\ &= 0.057 * 100 \\ &= 5.72 \% \end{aligned}$$

4. Kesimpulan

Berdasarkan hasil pengolahan data dengan menggunakan aplikasi *Rapidminer* dengan menggunakan algoritma *Naïve Bayes* menghasilkan nilai *Accuracy* 81.81%, *Precision* 66.66%, dan *Recall* 100% dengan hasil AUC 0.800. Pengolahan data yang dilakukan menghasilkan informasi jumlah status pembayaran yang *out standing* berdasarkan beberapa kriteria dan lokasi, sehingga pihak manajemen bisa mengambil langkah untuk meminimalisir penyewa yang sering telat dalam pembayaran sewa teras UMKM. Untuk penelitian selanjutnya, bisa untuk menggunakan metode lain yang lebih akurat nilainya. Hasil dari penelitian ini dapat dijadikan sebagai salah satu referensi bagi pihak manajemen dalam mendukung keputusan untuk melakukan tindakan yang bisa meminimalisir terjadinya status pembayaran yang *out standing*.

Referensi

- Ariadi, et al. (2015). Klasifikasi Berita Indonesia Menggunakan Metode Naive Bayesian Classification Dan Support Vektor Machine Dengan Confix Srtipping Stemmer. *Jurnal Sains Dan Seni Its*.
- Astuti, et al. (2018). Algoritma Naive Bayes Dengan Fitur Seleksi Untuk Mengetahui Hubungan Variabel Nilai Dan Latar Belakang Pendidikan. *Jurnal Simetris*.
- BPS. (2016). *BPS*.
- Hasan. (2017). Prediksi Tingkat Kelancaran Pembayaran Kredit Bank Menggunakan Algoritma Naive Bayes Berbasis Forward Selection. *Ilkom Jurnal Ilmiah*.
- Indomaret.Co.Id. (2018).
- Maharani, et al. (2017). Implementasi Data Mining Untuk Pengaturan Layout Minimarket Dengan Menerapkan Association Rule. *Jurnal Riset Komputer (JURIKOM)*.
- Maitimu. (2017). In L. S. Maitimu. Jakarta.
- Ningrum, et al. (2016). Implementasi Peraturan Pemerintah Nomor 46 Tahun 2013 Tentang Tentang Penghasilan Yang Diperoleh Wajib Pajak Sektor Umkm (Studi Pada Kantor Pelayanan Pajak

- (Kkp) Pratama Sidoarjo Selatan). *Jurnal Perpajakan (Jejak)*.
- Pakpahan, Irawan. (2017). Penerapan Algoritma Naive Bayes Untuk Menentukan Nasabah Potensial Pada Ajb Bumiputera 1912. *Jurnal LPKIA*.
- Pradana, Waspada. (2019). Aplikasi Hybrid Pada Sistem Informasi Penyewaan Buku. *Jurnal Simetris*.
- Puridewi, Nugraha. (2018). Perbandingan Metode Naive Bayes, Support Vector Machine Dan Id3 Dalam Penetapan Status Penanganan Kecelakaan Kerja.
- Rahman, Suryanto. (2017). Implementasi Sistem Informasi Seleksi Penerima Beasiswa Dengan Metode Naive Bayes Classifier. *Jurnal Penelitian Pendidikan Indonesia (JPPI)*.
- Saleh. (2015). Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *Citec Journal*.
- Setiawan. (2016). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi (Studi Kasus : Politeknik Lp3i Jakarta). *Jurnal Lentera ICT*.
- Sulistiyawati, et al. (2018). Beberapa Faktor Yang Mempengaruhi Keunggulan Bersaing Pada UMKM Handycraft Di Semarang.
- Supriyatna, Mustika. (2018). Komparasi Algoritma Naive Bayes Dan SVM Untuk Memprediksi Keberhasilan Imunoterapi Pada Penyakit Kutil. *Jurnal Sains Komputer & Informatika (J-SAKTI)*.