

# Implementasi Naive Bayes pada Sistem Asesmen Program Kemitraan dan Bina Lingkungan PT. Pos Indonesia

Fatia Amalia Maresti<sup>1</sup>, Kiki Mustaqim<sup>1</sup>, dan Nava Gia Ginasta<sup>2</sup>

<sup>1</sup>Program Studi S1 Sains Data, Universitas Logistik dan Bisnis Internasional, Bandung

<sup>2</sup>Program Studi S1 Bisnis Digital, Universitas Logistik dan Bisnis Internasional, Bandung

Jln. Sariasih No. 54, Kota Bandung, 40151, Indonesia

E-mail : fatiaamalia@ulbi.ac.id, kiki@ulbi.ac.id, navagia@ulbi.ac.id

## Abstrak

Selama operasional Program Kemitraan dan Bina Lingkungan (PKBL) PT. Pos Indonesia terjadi kredit macet yang berdampak pada arus sirkulasi pendapatan. Salah satu penyebabnya adalah proses asesmen awal data yang kurang mendalam dengan melihat aspek kriteria calon mitra dari catatan kredit sebelumnya. Tujuan penelitian adalah memberikan rekomendasi sebagai dasar pertimbangan kelayakan calon mitra usaha. Melalui analisis data profil usaha, berupa sektor usaha, provinsi, status usaha, tahun pendirian, jumlah modal, hasil penjualan, total kekayaan bersih, usia pemilik, penghasilan pemilik, jumlah pinjaman, dan label, menggunakan metode klasifikasi Naive Bayes karena kecepatan dan akurasi. Hasil evaluasi dari 20 data test, terdapat 18 (90%) data terklasifikasi benar, yaitu 6 (33%) data terkategori layak dan 12 (67%) data terkategori tidak layak mendapat pinjaman. Algoritma Naive Bayes memberikan hasil akurasi sebesar 90%, dengan presisi dan recall masing-masing sebesar 92%. Hasil ini menunjukkan bahwa algoritma Naive Bayes layak digunakan untuk penentuan kemitraan usaha PKBL PT. Pos Indonesia.

**Kata kunci** : Klasifikasi, Naive Bayes, PKBL, PT. Pos Indonesia, Asesmen.

## Pendahuluan

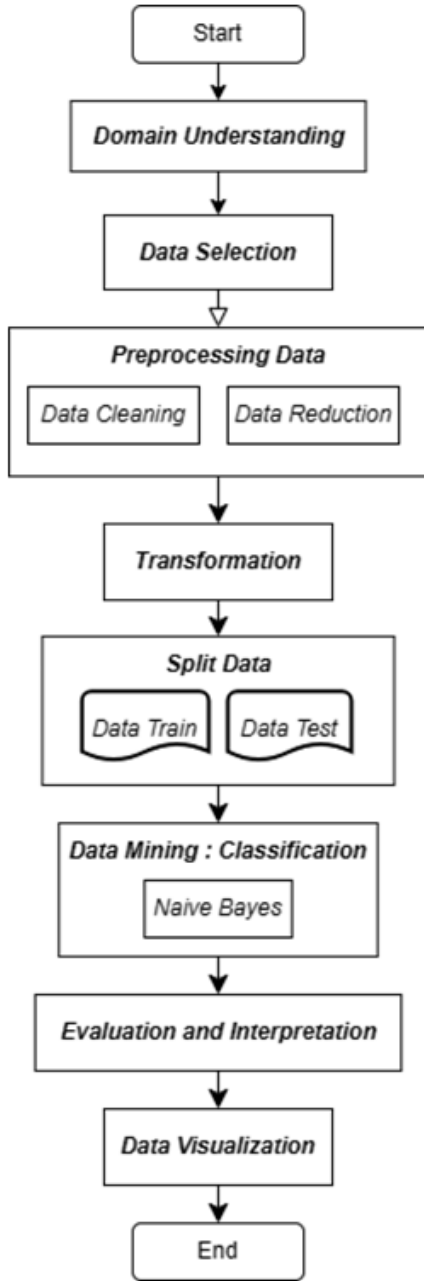
Tanggung Jawab Lingkungan atau yang dikenal sebagai *Corporate Social Responsibility* (CSR) merupakan salah satu kewajiban perusahaan dalam memperkuat kemampuannya dalam beradaptasi dengan lingkungan, pengampu kepentingan dan komunitas yang terikat dengannya. Praktik bisnis perusahaan harus fokus pada tiga hal penting untuk memperkuat citranya dalam masyarakat yaitu ekonomi, sosial dan lingkungan [1]. Program CSR diimplementasikan oleh BUMN melalui Program Kemitraan dan Bina Lingkungan. Pada dasarnya PKBL terdiri Program Kemitraan dengan Usaha Kecil dan Program Bina Lingkungan yang terdiri dari Bantuan Bencana Alam, Bidang Kesehatan Masyarakat, Bidang Pendidikan dan Pelatihan, Bidang Agama, Bidang Sarana dan Prasarana serta Bidang Pelestarian Alam [2].

Salah satu BUMN yang melaksanakan PKBL adalah PT Pos Indonesia yang diatur dalam Peraturan Menteri BUMN nomor Per-02/MBU/7/2017 tentang Perubahan Kedua atas Peraturan Menteri BUMN nomor Per-09/MBU/7/2015 tentang Program Kemitraan dan Bina Lingkungan Badan Us-

aha Milik Negara. Dalam Program Kemitraan PT. Pos Indonesia, Mitra Binaan adalah usaha kecil yang merupakan kegiatan ekonomi rakyat berskala kecil dan memenuhi kriteria kekayaan bersih serta kepemilikan sebagaimana diatur dalam Perundangan Menteri. Program Kemitraan PT. Pos Indonesia bertujuan memperkuat kemampuan usaha kecil agar kuat dan tangguh. Namun, selama operasional PKBL, khususnya dalam proses pinjaman Mitra Binaan, terjadi kredit macet yang disebabkan oleh kesulitan Mitra Binaan dalam mengembangkan usahanya. Hal tersebut tentunya berdampak pada arus sirkulasi kas maupun pendapatan PT. Pos Indonesia. Salah satu penyebab lainnya adalah proses asesmen awal data yang kurang mendalam dengan melihat aspek kriteria calon mitra maupun histori dari catatan kredit macet sebelumnya.

Penelitian ini bertujuan untuk meminimalisir peningkatan kredit macet dan menjadi dasar pertimbangan kelayakan asesmen calon mitra dengan mengimplementasikan Algoritma Naive Bayes dalam pengambilan keputusannya. *Naive Bayes* merupakan salah satu teknik klasifikasi dalam pencarian kesimpulan dengan menggunakan probabilitas [3]. Kriteria Mitra Binaan yang dapat

mengikuti Program Kemitraan dan laporan data kredit Mitra Binaan PKBL PT. Pos Indonesia akan dijadikan dasar dalam memberikan rekomendasi asesmen calon mitra. Dalam mengimplementasikan *Naive Bayes* juga memperhatikan aspek fungsional untuk memudahkan pengguna menggunakan sistem tersebut.



Gambar 1: Diagram Alir Tahap Penelitian

### Penelitian yang Terkait

Penelitian terkait sistem asesmen Mitra Binaan PKBL PT. Pos Indonesia masih sangat minim. Adapun penelitian terdahulu yang berkaitan tentang sistem asesmen dan PKBL BUMN adalah sebagai berikut :

1. Indra G. T. I. melakukan penelitian pada tahun 2021 tentang pembuatan aplikasi ases-

men calon debitur dengan mengimplementasikan algoritma Naive Bayes guna memberikan rekomendasi kelayakan calon debitur yang memiliki potensi kredit macet atau tidak. Hasil pengujian dari aplikasi asesmen calon debitur didapatkan nilai tingkat akurasi sebesar 86% [4].

2. M. D. A. Pangestu, dkk melakukan penelitian pada tahun 2022 tentang implementasi PKBL PT. Telekomunikasi Kandatel Jember dan dampak dari pelaksanaan program. Metode dalam penelitian ini menggunakan pendekatan kualitatif dengan sumber data primer dan sekunder, teknik analisis data meliputi reduksi, penyajian data, dan penarikan kesimpulan [2].
3. Tahun 2021 R. Dewisari, dkk melakukan penelitian dengan metode studi kepustakaan, menggunakan data sekunder yang bersumber dari laporan keuangan dan laporan PKBL perusahaan untuk mengetahui hubungan antara kinerja keuangan perusahaan dengan penyaluran dana PKBL tahun 2015-2019. Pengukuran kinerja keuangan perusahaan dilakukan pada BUMN yang kegiatan usahanya melibatkan penggunaan dan/atau pengelolaan SDA serta menyediakan informasi PKBL yang memadai, yaitu PT. Brantas Abipraya, PT. Jasa Marga, dan PT. Semen Indonesia, menggunakan perhitungan rasio profitabilitas berupa return on asset [5].
4. D.A. Kurniawan, dkk pada 2018 melakukan penelitian pada Koperasi Simpan Pinjam Baitul Maal wa Tamwil (KSPPS BMT) “Arta Jiwa Mandiri” Wonogiri yang bergerak di bidang kredit syariah dan simpan pinjam. Dibangun sebuah aplikasi yang dapat memprediksi calon anggota yang layak mendapatkan pinjaman dari koperasi dengan teknik data mining algoritma Naive Bayes. Hasil dari penelitian ini didapatkan nilai akurasi sebesar 80% [6].
5. Penelitian yang dilakukan oleh R. Rachman dan R.N. Handayani (2021) adalah menentukan prediksi kelancaran pembayaran sewa teras UMKM. Data yang digunakan berjumlah 504 data, dengan atribut kelas yang dibutuhkan Area Manager, Kota, Jenis Kelamin Penyewa, Rata-rata Umur Penyewa, dan Status Pembayaran. Algoritma klasifikasi naive bayes menghasilkan persentasi Accuracy 81.81%, Precision 66.66%, Recall 100% dan AUC 0.800 untuk keakuratan prediksi [7].

### Metode Penelitian

Penelitian ini memiliki enam tahapan utama yaitu *Domain Understanding*, *Data Selection*, *Prepro-*

cessing Data, Transformation, Split Data, Classification dan Evaluation. Gambar 1 merupakan diagram alir tahap penelitian yang dilakukan.

### Domain Understanding

Pada tahap awal peneliti melakukan analisis masalah yang ada dalam penelitian seperti melakukan research terhadap identifikasi masalah, tujuan penelitian, serta manfaat penelitian. Seiring observasi berlangsung, dilakukan juga kajian literatur yang berkaitan dengan topik penelitian. Dalam tahap ini akan dilakukan pengambilan data langsung pada bagian Program Kemitraan dan Bina Lingkungan PT. Pos Indonesia di Graha Pos Indonesia, Kota Bandung. Data berupa profil identitas usaha dan pemilik usaha Mitra Binaan PKBL PT. Pos Indonesia tahun 2023 pada provinsi tertentu.

### Data Selection

Selanjutnya memilah data apa saja yang akan digunakan. Secara keseluruhan data profil usaha terdiri dari 20 variabel, namun data diseleksi menjadi 11 variabel saja yang terdiri dari Sektor Usaha, Provinsi, Status Usaha, Tahun Pendirian, Jumlah Modal, Hasil Penjualan, Total Kekayaan Bersih, Usia Pemilik, Penghasilan Pemilik, Jumlah Pinjaman, dan Label. Variabel tersebut dipilih karena paling relevan untuk menganalisis pengajuan pinjaman usaha Mitra Binaan PKBL PT. Pos Indonesia.

### Preprocessing Data

Pada tahap ini dilakukan membersihkan, mengorganisasi, dan mempersiapkan data yang bertujuan untuk menghilangkan data error, seperti redudansi data dan missing value sehingga dapat mengeluarkan hasil lebih baik dalam menganalisis pengajuan pinjaman Mitra Binaan PKBL PT. Pos Indonesia.

### Transformation

Transformasi data dilakukan untuk memastikan bahwa data yang digunakan dalam analisis algoritma Naive Bayes sesuai dengan asumsi model dan dapat memberikan hasil yang akurat. Proses dalam transformasi data diantaranya encoding data kategorikal menjadi biner agar dapat dilakukan pengolahan dengan algoritma Naive Bayes.

### Split Data

Data dibagi menjadi data latih (*data train*) dan data uji (*data test*) dengan perbandingan 80:20. Data latih digunakan untuk pembentukan model klasifikasi Naive Bayes. Selanjutnya, model klasifikasi Naive Bayes yang sudah dilatih akan diuji dengan data test sehingga akan diperoleh performa model klasifikasi.

## Naive Bayes Classification

Algoritma Naive Bayes merupakan pengklasifikasian statistik yang berdasar pada teorema Bayes dan digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Naive Bayes memiliki kemampuan mengklasifikasi dan terbukti memiliki akurasi dan kecepatan yang tinggi saat diterapkan dalam data yang besar. Keunggulan lain dari Naive Bayes yaitu proses komputasi yang cepat, mudah diterapkan dengan struktur yang sederhana, dan efektif [8]. Dalam mengklasifikasi teks, metode ini mampu mendapatkan nilai akurasi yang tinggi dan kompleksitas run time yang baik dengan jumlah data yang besar [9]. Tujuan metode ini adalah untuk memperhatikan perkiraan probabilitas yang ada berdasarkan dari kategori data latih. Perumusan nilai peluang Naive Bayes yaitu sebagai berikut

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \quad (1)$$

dengan

**X** : data dengan class yang belum diketahui

**C<sub>i</sub>** : hipotesis data X merupakan suatu class spesifik

**P(C<sub>i</sub>|X)** : probabilitas hipotesis C<sub>i</sub> berdasarkan kondisi X (posteriori probability)

**P(C<sub>i</sub>)** : probabilitas hipotesis C<sub>i</sub> (prior probability)

**P(X|C<sub>i</sub>)** : probabilitas X berdasarkan kondisi pada hipotesis C<sub>i</sub>

**P(X)** : probabilitas dari X

Klasifikasi Naive Bayes mengasumsikan bahwa keberadaan sebuah variabel tidak ada kaitannya dengan keberadaan variabel yang lain karena asumsi variabel tidak saling terkait (conditionally independent), ditulis dengan rumus [10] :

$$P(C_i) = \sum_{k=1}^n P(X_k|C_i) \quad (2)$$

### Evaluation and Interpretation

Untuk mengetahui performa metode klasifikasi Naive Bayes dilakukan tahap Evaluation. Teknik yang digunakan untuk mengevaluasi hasil yaitu *confusion matrix*. *Confusion Matrix* merupakan alat bantu dalam mengevaluasi performa dari berbagai algoritma dan matrix ini menghasilkan beberapa istilah dari kelas aktual dan kelas prediksi [11]. *Confusion matrix* digunakan untuk menghitung akurasi, presisi, dan recall. Adapun perhitungan akurasi menggunakan hasil pada confusion matrix yaitu

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

Dari persamaan (3) terdapat nilai TP (True Positive) yaitu jumlah data berlabel Layak dan diprediksi dengan benar, TN (True Negative) adalah jumlah data berlabel Tidak Layak dan diprediksi dengan benar, FP (False Positive) yaitu jumlah data berlabel Tidak Layak yang salah diklasifikasikan sebagai label Layak, dan FN (False Negative) adalah jumlah data berlabel Layak yang salah diklasifikasikan sebagai label Tidak Layak [12].

## Hasil dan Pembahasan

### Data Selection

Data terdiri dari 100 data profil usaha dengan 20 variabel, namun dalam penelitian ini hanya digunakan 11 variabel yaitu Sektor Usaha, Provinsi, Status Usaha, Tahun Pendirian, Jumlah Modal, Hasil Penjualan, Total Kekayaan Bersih, Usia Pemilik, Penghasilan Pemilik, Jumlah Pinjaman, dan Label.

### Preprocessing Data

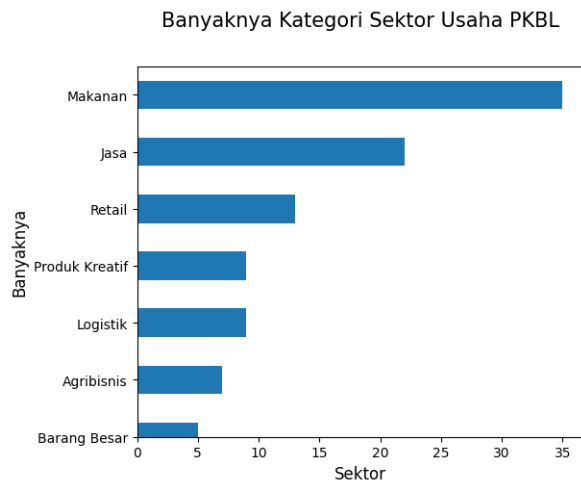
Data profil usaha berupa jenis usaha pada variabel Sektor disederhanakan menjadi 7 kategori yaitu sektor Makanan, Jasa, Retail, Agribisnis, Barang Besar, Logistik, dan Produk Kreatif. Untuk variabel lokasi usaha dituliskan berupa variabel Provinsi yang hanya mencakup 7 provinsi seperti yang tercantum pada Tabel 1.

Dari 100 dataset yang dimiliki, variabel Label dibagi menjadi 2 kategori yaitu Label Tidak Layak [0] sebanyak 51 data dan Label Layak [1] sebanyak 49 data. Gambar 2 menunjukkan grafik batang variabel Sektor dengan hasil sektor usaha yang paling banyak menjadi mitra adalah sektor makanan diikuti oleh jasa. Untuk variabel Provinsi pada Gambar 3 diperoleh bahwa provinsi Kepulauan Riau yang paling banyak dan Provinsi Jawa

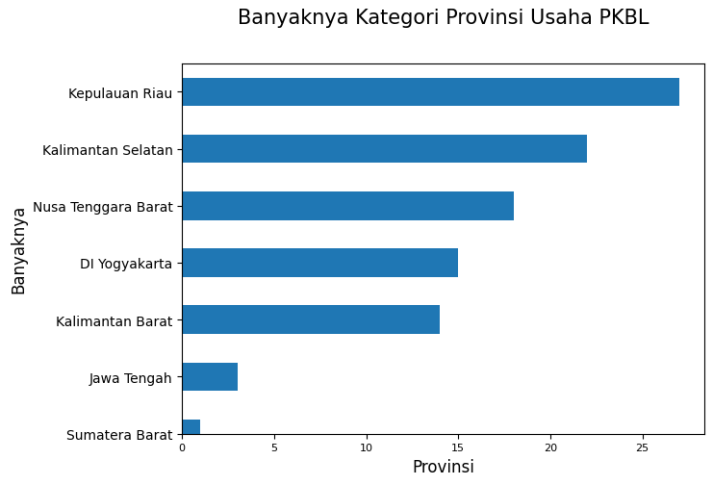
Tengah dan Sumatra Barat yang paling sedikit.

Tabel 1: . Kategori Setiap Variabel Data Profil Usaha PKBL

No	Variabel	Tipe Data	Isi Variabel
C1	Sektor	Kategorik	1) Makanan 2) Jasa 3) Retail 4) Agribisnis 5) Barang Besar 6) Logistik 7) Produk Kreatif
C2	Provinsi	Kategorik	1) Sumatera Barat 2) Kepulauan Riau 3) Jawa Tengah 4) DI Yogyakarta 5) Kalimantan Barat 6) Kalimantan Selatan 7) Nusa Tenggara Barat
C3	Status	Kategorik	1) Milik 2) Sewa
C4	Tahun Pendirian	Numerik	Tahun pendirian usaha mulai dari tahun 1985 sampai 2023
C5	Jumlah Modal	Numerik	Jumlah modal dalam setahun bernilai mulai dari Rp 450.000 sampai Rp 500.000.000
C6	Hasil Penjualan	Numerik	Hasil penjualan dalam setahun bernilai mulai dari Rp 2.500.000 sampai Rp 660.000.000
C7	Total Kekayaan Bersih	Numerik	Total kekayaan bersih dalam setahun bernilai mulai dari Rp 2.000.000 sampai Rp 2.071.851.000
C8	Usia Pemilik	Numerik	Usia pemilik usaha per September 2023 mulai dari 31 tahun sampai 64 tahun
C9	Penghasilan	Numerik	Penghasilan dalam setahun bernilai mulai dari Rp 2.050.000 sampai Rp 74.400.000
C10	Jumlah Pinjaman	Numerik	Jumlah modal dalam setahun bernilai mulai dari Rp 800.000 sampai Rp 100.000.000
C11	Label	Kategorik	0) Tidak Layak mendapat pinjaman 1) Layak mendapat pinjaman



Gambar 2: Diagram batang kategori Sektor data profil usaha



Gambar 3: Diagram batang kategori Provinsi data profil usaha

Statistika deskriptif data berupa rata-rata, standar deviasi, minimum, median, dan maksimum dari variabel numerikal yaitu Tahun.Berdiri, Jml.Modal, Hsl.Penjualan, Total.Kekayaan, Usia, Penghasilan, dan Pinjaman ditunjukkan pada Tabel 2.

Tabel 2: Sari Numerik Variabel Kategorik Data Profil Usaha

Var. Numerik	Rataan	Standar Dev
Tahun.Berdiri	2015,23	6,58
Jml.Modal	69.891.500	111.282.981
Hsl.Penjualan	94.002.900	131.154.126
Total.Kekayaan	236.308.510	342.346.735
Usia	47,61	7,92
Penghasilan	12.871.400	15.248.880
Pinjaman	30.471.500	27.604.476

Var. Numerik	Minimum	Median	Maksimum
Tahun.Berdiri	1985	2018	2023
Jml.Modal	450.000	35.000.000	500.000.000
Hsl.Penjualan	2.500.000	57.600.000	660.000.000
Total.Kekayaan	2.000.000	95.500.000	2071.851.000
Usia	31	49,01	63,44
Penghasilan	2.050.000	6.350.000	74.400.000
Pinjaman	800.000	30.000.000	100.000.000

**Transformation**

Variabel numerik pada data dikelompokkan menjadi dua kategori dengan pemisah pembulatan nilai rataannya. Tabel 3 menjabarkan isi dari variabel C4 sampai C10 data profil usaha.

**Split Data**

Variabel bebas X terdiri dari 10 variabel yaitu Sektor, Provinsi, Status, Tahun.Berdiri, Jml.Modal,

Hsl.Penjualan, Total.Kekayaan, Usia, Penghasilan, dan Pinjaman, kemudian variabel tak bebas y yaitu variabel Label. Setiap variabel X dan y dibagi menjadi data latih (data train) dan data uji (data test) dengan perbandingan 80:20. Sehingga ukuran data\_train X adalah (80,10) data dan data\_test X sebanyak (20,10) data. Sedangkan data\_train y adalah (80,1) data dan data\_test y sebanyak (20,1) data.

Tabel 3: Transformasi Variabel Numerik Menjadi Kategori

No	Variabel	Isi Variabel
C4	Tahun Pendirian	1. T1 : Kurang dari 2015 2. T2 : Lebih dan sama dengan 2015
C5	Jumlah Modal	1. M1 : Kurang dari 70.000.000 2. M2 : Lebih dan sama dengan 70.000.000
C6	Hasil Penjualan	1. J1 : Kurang dari 94.000.000 2. J2 : Lebih dan sama dengan 94.000.000
C7	Total Kekayaan Bersih	1. K1 : Kurang dari 240.000.000 2. K2 : Lebih dan sama dengan 240.000.000
C8	Usia Pemilik	1. U1 : 20 – 47 tahun 2. U2 : 47 – 60 tahun
C9	Penghasilan	1. Ph1 : Kurang dari 13.000.000 2. Ph2 : Lebih dan sama dengan 13.000.000
C10	Jumlah Pinjaman	1. Pn1 : Kurang dari 30.000.000 2. Pn2 : Lebih dan sama dengan 30.000.000

**Naïve Bayes Classification**

Naive Bayes merupakan sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut [13]. Klasifikasi Naive Bayes menggunakan python dengan package GaussionNB.

Dalam mengimplementasikan Naive Bayes, dilakukan perhitungan peluang dari masing-masing C1 hingga C10 terhadap peluang C11 (Label). Langkah pertama adalah membuat peluang 0 dan 1 dari C11:

$$P(C11=0)=51/100 =0,51$$

$$P(C11=1)=49/100=0,49$$

Berikutnya adalah menghitung peluang C1 hingga C10 terhadap C11, yang disebut dengan X. Lakukan transformasi data numerik dalam kategori. Misalnya akan dilakukan prediksi label Layak atau Tidak Layak terhadap data uji berikut:

1. Sektor = Makanan → C1=Makanan
2. Provinsi = Nusa Tenggara Barat → C2=NTB
3. Status Usaha = Milik → C3=Milik
4. Tahun Pendirian = 2020 → C4=T2
5. Jumlah Modal = Rp 50.000.000 → C5=M1
6. Hasil Penjualan = Rp 40.600.000 → C6=J1
7. T. Kekayaan Bersih = Rp 65.000.000 → C7=K1
8. Usia Pemilik = 50 → C8=U2
9. Penghasilan = Rp 12.000.000 → C9=Ph1
10. Jumlah Pinjaman = Rp 41.000.000 → C10=Pn2

Langkah berikutnya adalah menghitung jumlah data dari masing-masing kategori C1 hingga C10 terhadap C11. Misal dari 100 data, kategori C1 = 'Makanan' ada sebanyak 35 data dengan kategori label C11 = 0 yang bersesuaian sebanyak 14 data dan C11 = 1 sebanyak 21. Langkah yang sama dilakukan untuk menghitung C2 hingga C10. Setelah dilakukan perhitungan jumlah data dari masing-masing kategori, selanjutnya dihitung peluangnya dengan penjabaran  $P(X|C11=0)$  sebagai berikut:

- $P(C1=Makanan | C11=0) = 14/51 = 0,27$
- $P(C2=NTB | C11=0) = 3/51 = 0,06$
- $P(C3=Milik | C11=0) = 37/51 = 0,73$
- $P(C4=T2 | C11=0) = 40/51 = 0,78$
- $P(C5=M1 | C11=0) = 41/51 = 0,80$
- $P(C6=J1 | C11=0) = 33/51 = 0,65$
- $P(C7=K1 | C11=0) = 51/51 = 1$
- $P(C8=U2 | C11=0) = 29/51 = 0,57$
- $P(C9=Ph1 | C11=0) = 34/51 = 0,67$
- $P(C10=Pn2 | C11=0) = 24/51 = 0,47$

Sedangkan untuk  $P(X|C11=1)$  sebagai berikut :

- $P(C1=Makanan | C11=1) = 21/49 = 0,43$
- $P(C2=NTB | C11=1) = 15/49 = 0,31$
- $P(C3=Milik | C11=1) = 33/49 = 0,67$

- $P(C4=T2 | C11=1) = 33/49 = 0,67$
- $P(C5=M1 | C11=1) = 32/49 = 0,65$
- $P(C6=J1 | C11=1) = 42/49 = 0,86$
- $P(C7=K1 | C11=1) = 22/49 = 0,45$
- $P(C8=U2 | C11=1) = 27/49 = 0,55$
- $P(C9=Ph1 | C11=1) = 41/49 = 0,84$
- $P(C10=Pn2 | C11=1) = 24/49 = 0,49$

Selanjutnya adalah mengalikan antara  $P(X|C11)$  dengan  $P(C11)$ , sehingga diperoleh hasil sebagai berikut:

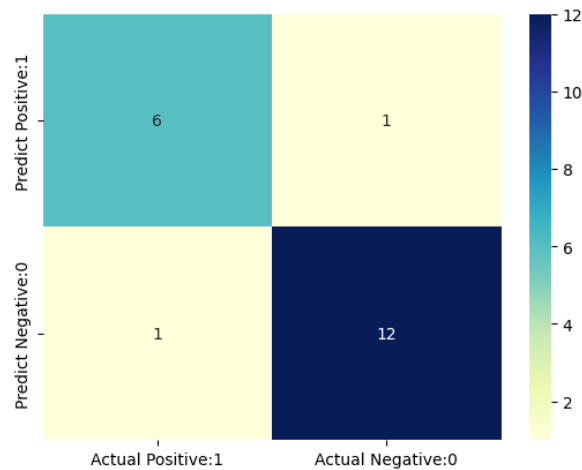
$$\begin{aligned} \text{Hasil hitung data uji dengan nilai } C11 = 0 \\ P(X \cap C11=0) &= P(X|C11=0) \cdot P(C11=0) \\ &= (0,000853) \cdot (0,51) \\ &= 0,000435 \\ \text{Hasil hitung data uji dengan nilai } C11 = 1 \\ P(X \cap C11=1) &= P(X|C11=1) \cdot P(C11=1) \\ &= (0,003377) \cdot (0,49) \\ &= 0,001655 \end{aligned}$$

Dari hasil di atas diperoleh bahwa peluang data uji dengan C11=1, yaitu 0,001655 lebih besar dari pada data uji dengan C11=0, yaitu 0,000435. Hasil bermakna bahwa data uji tersebut menghasilkan peluang memperoleh Label=1, yaitu Layak, lebih besar dibandingkan dengan peluang mendapatkan Label=0 yaitu Tidak Layak. Sehingga data uji dengan nilai kriteria seperti contoh yang diberikan di atas diprediksi masuk dalam kategori Layak mendapatkan pinjaman program mitra binaan PKBL PT. Pos Indonesia.

### ***Evaluation***

Proses penentuan hasil dari true positive dan true negative pada metode Naive Bayes berdasarkan hasil peluangnya. Hasil evaluasi dari 20 data test, terdapat 18 (90%) data yang terklasifikasi dengan benar, yaitu 6 (33%) data terkategori layak mendapat pinjaman, dan 12 (67%) data terkategori tidak layak mendapat pinjaman, ditunjukkan pada grafik confusion matrix pada Gambar 4.

Presisi untuk kelas data yang tidak layak mendapat pinjaman yaitu 86% dan presisi kelas yang layak mendapat pinjaman adalah 92%. Nilai Recall untuk kelas data tidak layak sebesar 86%, sedangkan recall kelas data yang layak mendapat pinjaman adalah 92%. Serta diperoleh nilai akurasi algoritma Naive Bayes sebagai metode klasifikasi dengan membandingkan data hasil prediksi dengan data test yaitu 90%. Hal ini menunjukkan bahwa algoritma Naive Bayes layak digunakan untuk proses penentuan kemitraan usaha PKBL PT. Pos Indonesia.



Gambar 4: . Grafik Confusion Matrix sebagai tahap evaluasi klasifikasi Naive bayes

## Penutup

Penelitian ini memberikan kemudahan rekomendasi dalam menentukan calon mitra usaha yang dapat dikategorikan Layak atau Tidak Layak mendapatkan pinjaman melalui 10 parameter yang terdiri dari Sektor Usaha, Provinsi Usaha, Status Usaha, Tahun Pendirian, Jumlah Modal, Hasil Penjualan, Total Kekayaan Bersih, Usia Pemilik, Penghasilan Pemilik, dan Jumlah Pinjaman. Penelitian ini mengimplementasikan algoritma Naïve Bayes dalam membuat model analisis kelayakan calon mitra usaha PKBL PT. Pos Indonesia. Model klasifikasi Naïve Bayes menghasilkan akurasi sebesar 90%, presisi sebesar 92% dan recall sebesar 92%. Model klasifikasi Naive Bayes layak digunakan untuk proses penentuan kemitraan usaha PKBL PT. Pos Indonesia. Adapun saran untuk pengembangan selanjutnya adalah menambah jumlah data train sehingga kualitas rekomendasi yang dihasilkan lebih akurat serta mengkombinasikan metode klasifikasi lainnya sebagai pembanding dengan metode Naïve Bayes.

## Daftar Pustaka

- [1] M. Marthin, M. B. Salinding dan I. Akim, "Implementasi Prinsip Corporate Social Responsibility (Csr) Berdasarkan Undang-Undang Nomor 40 Tahun 2007 Tentang Perseroan Terbatas", *J. Priv. Commer. Law*, vol. 1, no. 1, pp. 111–132, doi: 10.15294/jpcl.v1i1.12358, 2018.
- [2] D. A. Pangestu, P. Suharso dan W. Hartanto, "Implementasi Program Kemitraan Dan Bina Lingkungan (PKBL) sebagai Tanggung Jawab Sosial Perusahaan PT. Telekomunikasi Kandat Jember", *J. Pendidik. Ekon.*, vol. 16, pp. 323–333, doi: 10.19184/jpe.v16i2.25741, 2022.
- [3] M. Sabransyah, Y. N. Nasution dan D. Tisna, "Aplikasi Metode Naive Bayes dalam Prediksi Risiko Penyakit Jantung", *J. EKSPONENSIAL*, vol. 8, no. 2, pp. 111–118, 2017.
- [4] I. G. T. Isa, "Aplikasi Asesmen Calon Debitur menggunakan Naive Bayes di Koperasi Mitra Sejahtera SMK Negeri 1 Kota Sukabumi", *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 10, no. 1, pp. 31–39, doi: 10.32736/sisfokom.v10i1.1013, 2021.
- [5] R. Dewisari dan R. S. Ubed, "Penyaluran Dana Program Kemitraan Dan Bina Lingkungan (Pkbl) Dan Kinerja Keuangan Bumh", *Indones. Rich J.*, vol. 2, no. 1, pp. 49–58, doi: 10.31092/irj.v2i1.22, 2021.
- [6] D. A. Kurniawan dan Y. I. Kurniawan, "Aplikasi Prediksi Kelayakan Calon Anggota Kredit Menggunakan Algoritma Naïve Bayes", *J. Teknol. dan Manaj. Inform.*, vol. 4, no. 1, doi: 10.26905/jtmi.v4i1.1831, 2018.
- [7] R. Rachman dan R. N. Handayani, "Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM", *J. Inform.*, vol. 8, no. 2, pp. 111–122, doi: 10.31294/ji.v8i2.10494, 2021.
- [8] Y. A. S. Ivan dan P. P. Adikara, "Classification of Indonesian Hate Speech on Twitter Using Naïve Bayes and Selection of Information Gain Feature with Word Normalization", *J. Inf. Technol. Dev. Comput. Sci.*, vol. 3, no. 5, pp. 4914–4922, 2019.
- [9] A. S. Fitriani, "Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur", *JTAM (Jurnal Teor. dan Apl. Mat.)*, vol. 3, no. 2, pp. 98–104, 2019.
- [10] H. F. Putro, R. T. Vlandari dan W. L. Y. Saptomo, "Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan", *J. Teknol. Inf. dan Komun.*, vol. 8, no. 2, doi: 10.30646/tikom-sin.v8i2.500, 2020.

- [11] R. Tesalonika dan E. Mailoa, "Implementasi algoritma naive bayes untuk analisis sentimen isu resesi ekonomi 2023 di indonesia pada platform twitter", vol. 9, no. 1, pp. 34–40, 2024.
- [12] D. Normawati dan S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter", J. Sains Komput. Inform., vol. 5, no. 2, pp. 697–711, 2021.
- [13] F. Liantoni dan H. Nugroho, "Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan Knearest Neighbor", J. Siman-tec, vol. 5, no. 1, pp. 9–16, 2015.