

## Klasifikasi Data Pinjaman Koperasi Menggunakan Algoritma Naive Bayes

Tri Widiya Putra<sup>1</sup>, Eki Sriazi Lisnanda Sodikin.<sup>2</sup>, Muhamad Deden Ruhayat<sup>3</sup>  
Agus Yulianto<sup>4</sup>

<sup>1,2,3,4</sup>Program Studi Sistem Informasi, Universitas Nusa Mandiri Jakarta

e-mail: <sup>1</sup>[putra.amstrong@gmail.com](mailto:putra.amstrong@gmail.com), <sup>2</sup>[ekisriazilisnandas@gmail.com](mailto:ekisriazilisnandas@gmail.com),

<sup>3</sup>[dedencieie@gmail.com](mailto:dedencieie@gmail.com) <sup>4</sup>[agus.aag@nusamandiri.ac.id](mailto:agus.aag@nusamandiri.ac.id)

**Abstrak** - Koperasi sebagai badan usaha milik bersama yang berfungsi untuk menumbuhkan, mengembangkan dan mensejahterakan ekonomi anggota. pendapatan utama koperasi sangat bergantung pada besarnya jumlah angsuran yang dibayar oleh nasabah. Namun pada kenyataannya masalah yang terjadi justru timbul dari ketidak lancar proses pembayaran pinjaman yang mengganggu perputaran uang bahkan sampai menyebabkan kerugian bagi koperasi. Untuk membantu pihak koperasi karyawan PT. Tungmung Textille Bintang menganalisa kelayakan pemberian pinjaman kepada calon nasabah, penelitian ini menggunakan *Data mining* untuk menghitung nilai akurasi dari data histori peminjaman. Dengan menggunakan Metode *Naive Bayes*, nantinya data akan di klasifikasikan kedalam dua kategori yaitu lancar dan tidak lancar. Berdasarkan hasil pengujian mengenai penerapan metode *naive Bayes* untuk mengklasifikasi data historis pinjaman dengan menggunakan aplikasi *Rapid Miner* mendapatkan akurasi sebesar 72,99%, *precision* yaitu 84,25 % *class* lancar dan 35.97 % untuk *class* tidak lancar, *Recall* 81,22% pada *class* lancar dan 40,98% pada kelas tidak lancar, *Area Under Curve (AUC)* model algoritma *naive Bayes* adalah 0,681 hal ini menunjukkan model algoritma *naive Bayes* mencapai klasifikasi yang baik.

**Kata Kunci:** Koperasi, Data Mining, Naive Bayes

**Abstract** - Cooperatives as jointly owned enterprises that function to grow, develop and prosper members' economies. The main income of cooperatives is very dependent on the amount of installments paid by customers. But in reality the problems that occur actually arise from the non-fluency of the loan payment process which interferes with the circulation of money and even causes losses for the cooperative. To help the cooperative employees of PT. Tungmung Textille Bintang analyzes the feasibility of providing loans to prospective customers, this study uses *Data mining* to calculate the accuracy value of the historical data of borrowing. By using the method *Naive Bayes*, the data will be classified into two categories, namely current and non-current. Based on test results regarding the implementation of the method *naive Bayes* to classify the historical data of loans by comparing the data of testing using the application *Rapid Miner* get an accuracy of 72,99%, *precision* is 84,25% *class* smoothly and 35.97% for the *class* is not smooth, *Recall* 81,22% in *class* smoothly and 40,98% in the non-fluent *class*, the *Area Under Curve (AUC)* model of the algorithm *naive Bayes* is 0.681, this shows that the algorithm model *Naive Bayes* achieves a good classification.

**Keywords :** Cooperative, Data mining, Naive Bayes

### PENDAHULUAN

Koperasi ialah suatu organisasi badan usaha milik bersama yang memiliki peran sangat penting dalam menumbuhkan, mengembangkan potensi ekonomi, dan mensejahterakan rakyat (Habibulloh et al., 2021). Adanya koperasi menjadi salah satu cara anggotanya mendapatkan pinjaman dana untuk modal usaha sehingga diharapkan mampu memperbaiki kehidupan

perekonomian mereka. Koperasi karyawan PT. Tungmung Textile Bintang merupakan salah satu koperasi yang menyediakan jasa peminjaman pada anggotanya menggunakan sistem pembayaran bisa dicicil sesuai jangka saat yang sudah disepakati.

Sebagai forum finansial, penghasilan primer yang didapat oleh koperasi dipengaruhi total angsuran yang dibayarkan nasabah (Lestari & Badrul, 2020). Namun



pada kenyataannya masalah yang terjadi justru timbul dari ketidak lancar proses pembayaran pinjaman yang dilakukan anggota koperasi sehingga mengganggu perputaran uang pada koperasi. Tak jarang juga sampai ada nasabah yang kabur dan tidak melunaskan angsuran yang akhirnya menyebabkan kerugian bagi koperasi. Masalah-masalah itu bisa terjadi karena pertimbangan koperasi yang kurang matang dalam memberikan persetujuan pinjaman kepada calon nasabah. Tidak dapat memprediksi mana yang nantinya akan lancar proses kreditnya dan atau bermasalah. Masalah tunggakan pinjaman bisa diminimalisir tergantung asal performa *Credit Analyst* lebih optimal (Lestari & Badrul, 2020).

Guna membantu pihak koperasi menganalisa kelayakan pemberian pinjaman pada calon nasabah, penelitian ini menggunakan *Data Mining* buat menghitung nilai akurasi dari data histori peminjaman koperasi. *Data mining* didefinisikan sebagai suatu proses buat mencari pola berasal sekumpulan data yang terdapat di dalam *database* lalu dianalisis sehingga menghasilkan suatu informasi eksklusif untuk dimanfaatkan pada proses selanjutnya (Sartika et al., 2017). memakai Metode *Naïve bayes*, nantinya data akan di klasifikasikan kedalam dua kategori yaitu lancar serta tidak lancar. *Naïve Bayes* artinya pengelompokan data statistik yang bisa dipakai untk memprediksi peluang kelompok dalam suatu kelas (Borman et al., 2020). Penggunaan metode NBC (*Naïve Bayes Clasifier*) karena mempunyai tingkat keakuratan serta kecepatan yang tinggi ketika diterapkan ke dalam database yang besar (Susanto & Indriyani, 2019).

## METODOLOGI PENELITIAN

### 1. Tahapan Penelitian

Berikut ini adalah Langkah-langkah yang penulis gunakan dalam melakukan penelitian :



Gambar 1 Tahapan Penelitian

Sumber : (Putra et al., 2021)

#### a. Perumusan Masalah

Pada tahap awal penelitian, menentukan masalah apa yang terjadi pada koperasi PT. Tunggung Textille Bintan. Masih adanya pinjaman yang pembayarannya tidak lancar menyulitkan koperasi untuk mengelola informasi keuangan adalah masalah utama yang terjadi. Tidak adanya analisis lebih lanjut mengenai permasalahan yang dihadapi juga menjadi salah satu faktor kesulitan dalam menentukan pemberian pinjaman kepada calon peminjan.

#### b. Studi Literatur

Setelah mengerti dan memahami permasalahan apa yang sedang terjadi, langkah selanjutnya yaitu mencari referensi ataupun teori yang relevan beserta penelitian yang dilakukan. Pada hal ini peneliti menggunakan karya ilmiah dari beberapa jurnal dan juga buku sebagai bahan referensi.

#### c. Pemilihan Metode

Berdasarkan hasil literatur mengenai permasalahan yang ada yaitu tentang kredit macet ataupun tunggakan pembayaran pinjaman pada koperasi. *Naïve Bayes Clasifier (NBC)* adalah metode yang dipilih untuk mengukur dan juga menganalisa seberapa akurat tingkat kelancaran dan kemacetan pada data peminjaman di koperasi PT. Tunggung Textille Bintan.

#### d. Pengumpulan Data

Proses pendataan dilakukan melalui observasi dan wawancara kepada koperasi untuk mengetahui data yang dibutuhkan dalam penelitian ini. Hasil yang diperoleh berupa data pinjaman nasabah koperasi tahun 2020, maka nantinya data tersebut akan menjadi bahan pengujian dalam penelitian ini.

#### e. Pengolahan Data

Ditahap ini, data yang didapatkan dari koperasi karyawan PT. Tunggung Textille Bintan kemudian diolah menggunakan metode *Naïve Bayes Clasifier*. Dimulai dari pembersihan data, penentuan atribut, mengklasifikasikan data pada kategori lancar dan tidak lancar, hingga didapatkanlah data utuh yang siap diujikan.

#### f. Analisa Hasil

Pada tahap ini, penulis melakukan analisis hasil pengolahan data menggunakan perhitungan manual yang kemudian dicocokkan juga dengan pengujian menggunakan Aplikasi *Rapid Miner*. Fungsi penggunaan *rapid miner* ini guna mengukur dan mengetahui seberapa besar nilai akurasi data untuk menentukan kelancaran atau tidaknya proses pinjaman di koperasi PT. Tunggung Textille Bintan.

#### g. Kesimpulan

dijadikan sebagai bahan pertimbangan pemberian pinjaman kepada nasabah yang akan meminjam Perhitungan dengan menggunakan metode *naïve bayes* menghasilkan kriteria nasabah seperti apa dan bagaimana yang proses peminjamannya lancar dan tidak lancar. Hasil tersebut juga dikuatkan dengan tingkat akurasi yang diukur dari aplikasi *Rapid Miner*, yang nantinya dapat Kesimpulan dijadikan sebagai

bahan pertimbangan pemberian pinjaman kepada nasabah yang akan meminjam

## HASIL DAN PEMBAHASAN

Dalam menentukan data anggota koperasi yang lancar dan tidak lancar ada beberapa kriteria atau atribut yang peneliti gunakan. Kriteria yang peneliti gunakan sebagai penelitian dalam menentukan anggota yang lancar dan tidak lancar sebagai berikut:

### 1. Jenis kelamin

Dalam atribut jenis kelamin terdapat dua jenis kelamin yaitu Laki-laki dan Perempuan.

### 2. Status Nikah

Status pernikahan menjadi salah atribut yang dapat mempengaruhi hasil prediksi dengan klasifikasi Naive Bayes.

### 3. Penghasilan

Adapun jumlah penghasilan yang terdapat di Koperasi PT. Tung Mung Textile Bintang yang ada pada data peminjaman tersebut yaitu : Rp. 4.600.000 , Rp.4.800.000, Rp.5.000.000, Rp.5.500.000.

### 4. Jenis Pinjaman

Jenis Pinjaman yang ada di Koperasi PT. Tung Mung Textile Bintang diantaranya: Emergency, Siaga, Taktis UPDB.

### 5. Jumlah Pinjaman

Jumlah Pinjaman yang diberikan Koperasi PT. Tung Mung Textile Bintang diantaranya berjumlah: Rp.1.000.000, Rp.2.000.000, Rp.8.000.000, Rp.15.000.000, Rp.20.000.000.

### 6. Lama Cicilan

Pada data peminjaman di Koperasi PT. Tung Mung Textile Bintang terdapat atribut lama cicilan, bulanan hingga tahunan diantaranya mulai dari 4 bulan, hingga 48 bulan (4 tahun).

### 7. Cicilan

Dalam data peminjaman ini terdapat atribut cicilan, dimana dalam cicilan ini mulai dari Rp.132500, hingga Rp.1351200.

## 1. Perhitungan Naive Bayes

Menghitung Jumlah kelas atau Label

Tahap pertama perhitungan pencarian kelancaran dengan metode *naive bayes* adalah dengan mencari probabilitas dari masing – masing class, didalam data peminjaman terdapat dua class yaitu class lancar dan tidak lancar.

Metode perhitungannya yaitu dengan mencari sejumlah data lancar dan tidak lancar dari keseluruhan data training kemudian membaginya dengan keseluruhan data. Hasil dari perhitungannya tertera pada tabel dibawah ini :

Tabel 1. Menghitung Jumlah Class atau Label

| Label               | Jumlah data |                     | Hasil |
|---------------------|-------------|---------------------|-------|
|                     | Jumlah data | Jumlah seluruh data |       |
| P (y=lancar )       | 474         | 596                 | 0.80  |
| P (y=tidak lancar ) | 122         | 596                 | 0.20  |

Sumber : (Putra et al., 2021)

## A. Menghitung probabilitas masing-masing atribut.

### 1.) Atribut Jenis Kelamin

Tabel 2. Probabilitas atribut jenis kelamin

|                    | Jenis Kelamin |              |           |                 |
|--------------------|---------------|--------------|-----------|-----------------|
|                    | Lancar        | Tidak lancar | P(lancar) | P(tidak lancar) |
| <b>Laki - laki</b> | 26            | 8            | 0.05      | 0.07            |
| <b>Perempuan</b>   | 448           | 114          | 0.95      | 0.93            |
| <b>Total</b>       | 474           | 122          |           |                 |

Sumber : (Putra et al., 2021)

### 2.) Status Perkawinan

Tabel 3. Probabilitas atribut Status Perkawinan

|                      | Status Perkawinan |              |           |                 |
|----------------------|-------------------|--------------|-----------|-----------------|
|                      | Lancar            | Tidak lancar | P(lancar) | P(tidak lancar) |
| <b>Menikah</b>       | 271               | 70           | 0,57      | 0,57            |
| <b>Belum Menikah</b> | 203               | 52           | 0,43      | 0,43            |
| <b>Total</b>         | 474               | 122          |           |                 |

Sumber : (Putra et al., 2021)

### 3.) Atribut Penghasilan

Tabel 4. Probabilitas atribut Penghasilan

|                  | Penghasilan |              |           |                 |
|------------------|-------------|--------------|-----------|-----------------|
|                  | Lancar      | Tidak lancar | P(lancar) | P(tidak lancar) |
| <b>4.600.000</b> | 108         | 24           | 0,23      | 0,20            |
| <b>4.800.000</b> | 144         | 21           | 0,30      | 0,17            |
| <b>5.000.000</b> | 110         | 33           | 0,23      | 0,27            |
| <b>5.500.000</b> | 112         | 44           | 0,24      | 0,36            |
| <b>Total</b>     | 474         | 122          |           |                 |

Sumber : (Putra et al., 2021)

### 4.) Atribut Jenis Pinjaman

Tabel 5. Probabilitas atribut Jenis Pinjaman

|                  | Jenis Pinjaman |              |           |                 |
|------------------|----------------|--------------|-----------|-----------------|
|                  | Lancar         | Tidak lancar | P(lancar) | P(tidak lancar) |
| <b>Emergency</b> | 3              | 7            | 0,01      | 0,06            |
| <b>Siaga</b>     | 108            | 55           | 0,23      | 0,45            |
| <b>Taktis</b>    | 363            | 60           | 0,77      | 0,49            |
| <b>UPDB</b>      | 474            | 122          |           |                 |
| <b>Total</b>     |                |              |           |                 |

Sumber : (Putra et al., 2021)

5.) Atribut Jumlah Pinjaman

Tabel 6. Probabilitas atribut Jumlah Pinjaman

|          | Jumlah pinjaman |              |           |                 |
|----------|-----------------|--------------|-----------|-----------------|
|          | Lancar          | Tidak lancar | P(lancar) | P(tidak lancar) |
| 1000000  | 35              | 15           | 0,07      | 0,12            |
| 2000000  | 0               | 2            | 0,00      | 0,02            |
| 8000000  | 76              | 45           | 0,16      | 0,37            |
| 15000000 | 34              | 2            | 0,07      | 0,02            |
| 20000000 | 329             | 58           | 0,69      | 0,48            |
| total    | 474             | 122          |           |                 |

Sumber : (Putra et al., 2021)

6.) Atribut Cicilan

Tabel 7. Probabilitas Atribut Cicilan

|         | Cicilan |              |           |                 |
|---------|---------|--------------|-----------|-----------------|
|         | Lancar  | Tidak lancar | P(lancar) | P(tidak lancar) |
| 132500  | 32      | 10           | 0.07      | 0.08            |
| 260000  | 3       | 7            | 0.01      | 0.06            |
| 460000  | 67      | 42           | 0.14      | 0.34            |
| 506700  | 34      | 2            | 0.07      | 0.02            |
| 536700  | 124     | 14           | 0.26      | 0.11            |
| 675600  | 154     | 32           | 0.32      | 0.26            |
| 920000  | 9       | 3            | 0.02      | 0.02            |
| 1073400 | 10      | 2            | 0.02      | 0.02            |
| 1351200 | 41      | 10           | 0.09      | 0.08            |
| Total   | 474     | 122          |           |                 |

Sumber : (Putra et al., 2021)

7.) Atribut Lama Cicilan

Tabel 8. Probabilitas atribut Lama Cicilan

|       | Lama cicilan |              |           |                 |
|-------|--------------|--------------|-----------|-----------------|
|       | Lancar       | Tidak lancar | P(lancar) | P(tidak lancar) |
| 4     | 3            | 5            | 0,01      | 0,04            |
| 8     | 32           | 12           | 0,07      | 0,10            |
| 10    | 9            | 3            | 0,02      | 0,02            |
| 18    | 41           | 10           | 0,09      | 0,08            |
| 20    | 67           | 42           | 0,14      | 0,34            |
| 24    | 10           | 2            | 0,02      | 0,02            |
| 36    | 188          | 34           | 0,40      | 0,28            |
| 48    | 124          | 14           | 0,26      | 0,11            |
| Total | 474          | 122          |           |                 |

Sumber : (Putra et al., 2021)

B. Confusion Matrix perhitungan manual

Agar mempermudah perhitungan *naïve bayes* secara manual maka dibuat pengukuran kinerja klasifikasi dengan menggunakan Confusion Matrix berdasarkan *accuracy*, *error*, *sensitivity/Recall*, *Precision*, dan *specificity* yang ditunjukkan oleh Tabel IV.8 hasil penghitungan dalam laporan tabel sebagai berikut:

Tabel 9. Confusion Matrix

|                       | True Lancar | True Tidak lancar | Class Precision |
|-----------------------|-------------|-------------------|-----------------|
| Prediksi Lancar       | 385         | 72                | 84.25%          |
| Prediksi Tidak Lancar | 89          | 50                | 35.97%          |
| Class Recall          | 81.22%      | 40.98%            |                 |

Sumber : (Putra et al., 2021)

Jumlah keseluruhan data yang diuji 596 record dengan pengklasifikasian sebagai berikut :

True Positif (TP) : 385  
False Positif (FP) : 72  
False Negatif (FN) : 89  
True Negatif (TN) : 50

Hasil Perhitungan untuk mencari nilai *accuracy*, *error*, *sensitivity/Recall*, *Precision*, *specificity*, pada tabel .10 ini:

Tabel 10. Hasil Confusion Matrix

| Parameter   | Nilai  | Presentase |
|-------------|--------|------------|
| Akurasi     | 0,7299 | 72.99%     |
| Error       | 0,2701 | 27.01%     |
| Recall      | 0,8122 | 81.22%     |
| Sensitifity | 0,8425 | 84.25%     |
| Precison    | 0,4098 | 40.98%     |
| Spesificity |        |            |

Sumber : (Putra et al., 2021)

Nilai *confusion matrix* menunjukkan hasil pengukuran manual dengan tingkat akurasi sebesar 72,99%. Hasil perhitungan tersebut nantinya juga akan dibandingkan menggunakan perhitungan aplikasi *Rapid Miner*, apakah hasilnya sama atau berbeda. Dari perbandingan kedua cara perhitungan tersebut nantinya dapat diketahui bahwa algoritma *naïve Bayes* termasuk dalam tingkat klasifikasi baik atau tidak untuk pengolahan data pinjaman koperasi PT. Tungmung Textill Bintan.

C. Hasil Perhitungan menggunakan Aplikasi *Rapid Miner*.

1.) *Accuracy*

|                   | True Lancar | True Tidak Lancar | Class Precision |
|-------------------|-------------|-------------------|-----------------|
| pred LANCAR       | 385         | 72                | 84.25%          |
| pred TIDAK LANCAR | 89          | 50                | 35.97%          |
| class recall      | 81.22%      | 40.98%            |                 |

Sumber : (Putra et al., 2021)

Gambar 2. *Accuracy*

2.) *Recall atau Sensitivity*

|                   | True Lancar | True Tidak Lancar | Class Precision |
|-------------------|-------------|-------------------|-----------------|
| pred LANCAR       | 385         | 72                | 84.25%          |
| pred TIDAK LANCAR | 89          | 50                | 35.97%          |
| class recall      | 81.22%      | 40.98%            |                 |

Sumber : (Putra et al., 2021)

Gambar 3. *Recall atau Sensitivity*

### 3.) Precision

|              | tol LANCAR | tol TOLAK | class precision |
|--------------|------------|-----------|-----------------|
| pred LANCAR  | 72         | 7         | 84.2%           |
| pred TOLAK   | 8          | 59        | 88.1%           |
| class recall | 81.2%      | 48.9%     |                 |

Sumber : (Putra et al., 2021)

Gambar 4. Precision

## KESIMPULAN

Sesuai hasil penelitian yang sudah dilakukan mengenai penerapan metode naïve bayes untuk mengklasifikasi data pinjaman pada PT TUNG MUNG TEXTIE BINTAN dengan membandingkan data training dan data testing menggunakan software rapid miner mendapatkan akurasi sebanyak 72,99% sehingga penerapan metode naïve bayes ini baik untuk diterapkan pada data peminjaman.

Dengan adanya hasil akurasi yang cukup, bisa dijadikan bahan pertimbangan bagi koperasi untuk menyampaikan kepercayaan kepada karyawan yang akan mengajukan karena hasil penelitian potensi lancar lebih besar dibandingkan yang tidak lancar.

## REFERENSI

- Borman, R. I., Wati, M., Teknik, F., Ftik, K., Indonesia, U. T., Zainal, J., Pagar, A., No, A., & Lampung, B. (2020). *Penerapan Data Mining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandarlampung Dengan Algoritma Naïve Bayes*. 9.
- Habibulloh, W. M., Topiq, S., Adhirajasa, U., Sanjaya, R., Adhirajasa, U., & Sanjaya, R. (2021). *Klasifikasi Kelayakan Kredit Menggunakan Algoritma Naive Bayes Pada KSP Mekar Jaya Maleber*. 3(1), 92–99.
- Lestari, S., & Badrul, M. (2020). *IMPLEMENTASI KLASIFIKASI NAIVE BAYES UNTUK PREDIKSI*. 7(1), 8–16.
- Putra, T. widi, Sodikin, E. S. L., & Ruhayat, M. D. (2021). *LAPORAN AKHIR PENELITIAN : Klasifikasi Data Pinjaman Koperasi Menggunakan Algoritma Naive Bayes*.
- Sartika, D., Sensuse, D. I., Indo, U., Mandiri, G., & Komputer, F. I. (2017). *Perbandingan Algoritma Klasifikasi Naive Bayes , Nearest Neighbour , dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian*. 1(2), 151–161.
- Susanto, W., & Indriyani, L. (2019). *Analisis*

*Penerapan Naïve Bayes Untuk Memprediksi Resiko Kredit Anggota Koperasi Keluarga Guru*. 6(2), 262–270.

## PROFIL PENULIS

PROFIL PENULIS Eki Sriazi Lisnanda Sodikin., Tahun 2014 dari SMKN 9 Jurusan Teknik Komputer Jaringan, Argo Subur Tangerang. Tahun 2017 lulus dari SMKN 9 Jurusan Teknik Komputer Jaringan, Argo Subur Tangerang. Saat ini menjadi Mahasiswa Universitas Nusa Mandiri (S1).

PROFIL PENULIS Muhamad Deden Ruhayat, Pendidikan MA NEGERI 2 TANGERANG Jurusan Keagamaan Tahun Ajaran 2013 – 2016. Saat ini menjadi Mahasiswa Universitas Nusa Mandiri Fakultas Ilmu Komputer Jurusan Sistem Informasi (S1).

PROFIL PENULIS Tri Widiya Putra, Pendidikan SMK 9 JAKARTA Jurusan Pemasaran Tahun Ajaran 2014 – 2017. Saat ini menjadi Mahasiswa Universitas Nusa Mandiri Fakultas Ilmu Komputer Jurusan Sistem Informasi (S1).