



Implementasi Algoritma Nave Bayes Untuk Memprediksi Kelayakan Kredit Nasabah

Aniq Astofa^{1*}, Eko Sutono²

^{1,2}Program Studi Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia

Email: ^{1*}dosen02360@unpam.ac.id

(* : coressponding author)

Abstrak - Dalam memberikan kredit ada resiko yang dihadapi oleh, yaitu terlambatnya pembayaran pengembalian bahkan kegagalan pembayaran kredit. Masalah seperti ini terjadi karena kurang akuratnya pihak pemberi kredit dalam penilaian terhadap kemampuan nasabah, sehingga mengakibatkan kesalahan dalam keputusan pemberian kredit yang berujung pada kemacetan kredit. Oleh karena itu untuk mengatasi hal tersebut, pihak dalam memberikan pinjaman perlu memprediksi kelayakan pengajuan kredit terlebih dahulu supaya resiko yang timbul dari pemberian kredit kepada calon debitur tidak terlalu besar. Cara untuk menentukan prediksi kelayakan pengajuan kredit yaitu dengan menggunakan algoritma *data mining*. Salah satu fungsi dalam *data mining* yang dapat digunakan untuk menyelesaikan masalah yang dihadapi pihak untuk memprediksi kelayakan pengajuan kredit adalah klasifikasi. Metode yang diusulkan adalah naïve bayes. Algoritma naïve bayes bertujuan untuk melakukan klasifikasi data pada kelas tertentu. Tujuan penelitian ini adalah untuk melihat pola prediksi dari setiap atribut-atribut yang terdapat pada data set dengan menggunakan algoritma naïve bayes dan melakukan pengujian data training terhadap data testing untuk melihat pemodelan data sudah baik atau belum. Dalam membangun aplikasi ini penulis menggunakan bahasa pemrograman visual basic 2008 serta mysql sebagai basis data. Dari hasil implementasi pengujian model data *training* terhadap data *testing* sebesar 75%.

Kata Kunci: Kredit, Data Mining, Klasifikasi, Naïve Bayes, Aplikasi

Abstract - In giving the credit risk faced by cooperatives, namely the delayed a payment returns payment failure even credit. A problem like this occurs because the lender parties less accurate in the assessment of the ability of the customer, resulting in errors in the decision granting credit that led to the credit bottlenecks. Therefore to overcome this, the cooperative in providing the loan need to predict the feasibility of filing in advance so that the credit risk arising from granting loans to prospective borrowers is not too large. How to determine eligibility for credit submission that is prediction using data mining algorithms. One of the functions in the data mining can be used to resolve the issues facing the cooperative parties to predict the feasibility of credit submission is classification. The proposed method is naive bayes. Naive bayes algorithm aims to conduct classification data on a particular class. The purpose of this research was to look at the patterns of prediction of any attributes in the dataset by using the algorithm of naive bayes and do testing data training data testing to look at data modeling been good or not. In building this application writers use visual basic 2008 programming language and mysql as the database. From the results of testing data model implementation training against testing data of 75%.

Keywords: Credit, Data Mining, Classification, Naïve Bayes, Applications

1 PENDAHULUAN

1.1 Latar Belakang

adalah salah satu lembaga yang bergerak dalam bidang jasa keuangan. Adapun KSU Rukun Artha Santosa ini merupakan sebuah bidang usaha yang bergerak di bidang serba usaha, dibentuk untuk mengatasi permasalahan ekonomi yang dialami oleh kelompok usaha perikanan tangkap di Juwana. Salah satu unit kegiatan utama KSU Rukun Artha Santosa adalah memberikan kredit simpan pinjam. Padatapan evaluasi dalam proses pemberian kredit menjadi masalah yang sangat penting bagi RAS, karena pihak masih kesulitan untuk menentukan pemohon yang layak mendapatkan fasilitas kredit dan tidak beresiko menyebabkan kredit macet. Masalah seperti ini terjadi karena kurang akuratnya pihak pemberi kredit dalam penilaian terhadap kemampuan nasabah, sehingga mengakibatkan kesalahan dalam keputusan pemberian kredit yang berujung pada kemacetan kredit. Oleh karena itu untuk mengatasi dan mengantisipasi hal tersebut, pihak dalam memberikan pinjaman perlu penerapan prinsip kehati-hatian supaya resiko yang timbul dari pemberian kredit kepada calon debitur tidak terlalu besar. Dalam menerapkan prinsip kehati-hatian

ini perlu melakukan manajemen resiko dengan mengidentifikasi dan memprediksi kelayakan nasabah dengan baik sebelum memberikan pinjaman dengan cara memperhatikan data historis pinjamannasabah. Cara untuk menentukan prediksi kelayakan nasabah dalam membayar pinjaman kepada yaitu dengan menggunakan sebuah algoritma *data mining* dan untuk menghitungnya diperlukan keahlian khusus yang mengerti *data mining*. *Data mining* dianggap mampu dalam proses pencarian informasi, mendapatkan sebuah solusi dalam pengambilan keputusan dan menemukan pola tersembunyi dari suatu data kemudian diolah sehingga menjadi sebuah informasi yang berguna bagi seseorang atau organisasi.

Dalam data mining terdapat beberapa algoritma klasifikasi yang dapat digunakan untuk analisa kredit, salah satunya adalah dengan algoritma *Naïve Bayes*. Klasifikasi *Naïve Bayes* berjalan sangat baik dibandingkan dengan model klasifikasi lainnya. Kejadian ini dibuktikan dengan penelitian yang dilakukan oleh Xhemali, J.Hinde dan G.Stone (2009) yang berjudul "*Naïve Bayes vs. Decision Tree vs. Neural Networks in the Clasification of Training Web Pages*", dalam penelitian tersebut menjelaskan bahwa "*Naïve Bayes Classifier* mempunyai tingkat akurasi yang lebih baik dibanding model *classifier* lainnya", dengan tingkat akurasi 93,20% dan nilai F-Measure lebih 95%. Penelitian lain yang dilakukan oleh Evaristus dan Aryuni (2014) berjudul "*Comparative Study of Data Mining Model for Credit Card Application Scoring In Bank*", melaporkan klasifikasi *Naïve Bayes* memiliki tingkat akurasi yang lebih baik dari pada klasifikasi *ID3*, yaitu dengan tingkat keakuratan *NBC* adalah 82% dan *ID3* adalah 76% .

Penelitian lainnya yang dilakukan oleh Suryati (2013) berjudul "*Aplikasi Pendukung Keputusan Pengajuan Kredit Sepeda Motor*", metode *Naïve Bayes* telah dapat digunakan untuk menentukan kelayakan kredit sepeda motor. Maka dari itu penggunaan metode *Naïve Bayes* ini akan sangat tepat untuk membangun sebuah aplikasi *data mining* untuk memprediksi penentuan kelayakan pengajuan kredit pada pemohon yang akan melakukan pengajuan kredit, sehingga pihak atau bagian kredit akan tepat dalam memperkirakan diterima atau ditolaknya pemohon yang akan melakukan pengajuan kredit.

1.2 Tujuan

Tujuan yang ingin dicapai dalam penulisan tugas akhir ini adalah merancang dan membangun sebuah aplikasi *data mining* untuk memprediksi kelayakan pemohon kredit dengan menggunakan algoritma *naïve bayes*. Diharapkan aplikasi ini dapat membantu pihak Rukun Artha Santosa untuk mengetahui apakah calon nasabah yang memohon kredit nantinya akan lancar atau tidak lancar dalam membayar piutang.

1.3 Batasan Masalah

1. Bagaimana pembuatan aplikasi *data mining* dengan menggunakan data historis debitur untuk memprediksi kelayakan pengajuan kredit apakah pemohon diterima atau ditolak dalam pengajuan kredit.
2. Data set yang digunakan dalam penelitian adalah data rekapitulasi peminjam yang berjumlah 163 record dari KSU "Rukun Artha Santosa".
3. Menggunakan aplikasi IDE Visual Basic 2008 untuk pembuatan tampilan antar muka dan MySQL sebagai basis data untuk menyimpan data nasabah.

1. TINJAUAN PUSTAKA

2.1 Data Mining

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer untuk menganalisis dan mengekstraksi pengetahuan secara otomatis. *Data mining* adalah pencarian informasi yang tersembunyi, dimana informasi tersebut sebelumnya belum dikenal dan berpotensi bermanfaat. Secara teknis *data mining* merupakan proses iteratif dan interaktif untuk menemukan pola atau model baru yang dapat digeneralisasi untuk masa yang akan datang, dan dapat dimengerti dalam suatu *database* yang besar.



Tahapan proses dalam penggunaan *data mining* yang merupakan proses *Knowledge Discovery in Database* secara garisbesar dapat dijelaskan sebagai berikut :

1. *Pre-processing*

Sebelum proses *data mining* dapat dijelaskan, perlu dilakukan proses pembersihan data pada data yang menjadi fokus KDD. Proses pembersihan data mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak. Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

2. *Data Selection*

Pemilihan data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.

3. *Transformation*

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. Proses Mining

Proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma yang tepatsangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. *Interpretation* (Evaluasi)

Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Pola untuk melihat apakah ada sesuatu yang baru dan menarik dan dilakukan iterasi jika diperlukan. Tahap ini akan diketahui apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang adasebelumnya.

2.2 Klasifikasi

Berdasarkan tugasnya data mining dibagi menjadi beberapa kelompok, yaitu: [10] Deskripsi, Estimasi, Prediksi, Klasifikasi, Pengklusteran, dan Asosiasi. Klasifikasi merupakan bagian dari algoritma data mining, algoritma klasifikasi ini adalah algoritma yang menggunakan data dengan target (label) yang berupa nilai kategorikal / nominal.

2.3 Naïve Bayes

Naïve Bayes merupakan salah satu algoritma data mining dengan metode klasifikasi. Naïve Bayes adalah pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. Teorema tersebut dikombinasikan dengan *Naïve* dimana diasumsikan kondisi antar atribut saling bebas.

Persamaan dari teorema *Bayes* adalah:

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

2.4 Distribusi Frekuensi Kelompok

Data yang didapatkan peneliti yang masih berupa data random dapat dibuat menjadi data yang berkelompok, yaitu data yang telah disusun ke dalam kelas-kelas tertentu. Distribusi frekuensi adalah susunan data menurut kelas interval tertentu atau menurut kategori tertentu dalam sebuah daftar.

Penyusunan suatu tabel distribusi frekuensi perlu dilakukantahapan penyusunan data atau dilakukan dengan rumus:

Banyak Kelas Interval:

$$P_i = \frac{R}{n} = 1 + 3.3 \log n$$

Keterangan:

P_i = Panjang Kelas Interval

n = Banyaknya Kelas Interval

R = Rentang

2.5 Database MySQL

Menurut [19] *database* merupakan suatu kumpulan data yang terintegrasi dan diatur sedemikian rupa sehingga data tersebut dapat dikelola, dicari, dan diambil secara cepat dan tepat.

2.6 Struktur Navigasi

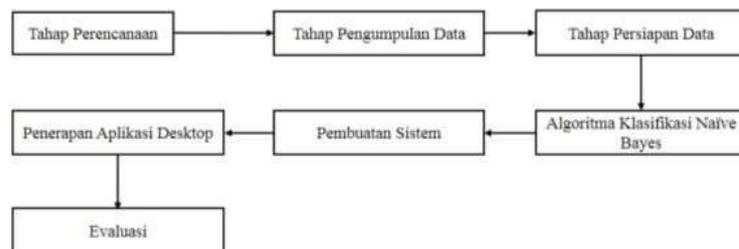
Struktur navigasi merupakan sebuah alur yang dipakai dalam pembuatan aplikasi. Sebelum menyusun sebuah aplikasi multimedia kedalam sebuah *software*, kita terlebih dahulu menentukan alur apa yang akan digunakan dalam membuat aplikasi.

2.7 UML

Unified Modeling Language (UML) merupakan bahasa spesifikasi standar untuk mendokumentasikan, menspesifikasikan, dan membangun sistem perangkat lunak. UML dikembangkan sebagai suatu alat untuk analisis dan desain berorientasi objek oleh Grady Booch, James Rumbaugh, dan Ivar Jacobson.

2. METODE PENELITIAN

Sebagai proses awal analisa data dilakukan pada awal penelitian ini dan untuk memberikan kemudahan dibuat suatu tahapan penelitian yang dapat dilihat pada gambar 1.



Gambar 1. Tahapan Penelitian

Tahap perencanaan yaitu dengan mengidentifikasi masalah dan menentukan tujuan. Proses identifikasi masalah dilakukan dengan melihat, tidak beresiko menyebabkan kredit macet. Atribut yang digunakan adalah sebagai berikut:

Tabel 1. Atribut

Atribut	Keterangan
Jenis Kelamin	Jenis Kelamin nasabah
Umur	Umur nasabah



Status	Status Pernikahannasabah
Pekerjaan	Pekerjaan nasabah
Pinjaman	Jumlah Pinjaman nasabah
Jangka Waktu	Jangka waktu pelunasan
Agunan	Jaminan
Kolekbilitas	Kualitas kredit nasabah berupa lancaratau tidak lancar

Data yang digunakan untuk mengamati, memahami dan menilai situasi bisnis saat ini. Kemudian tahap pengumpulan data Pada tahap ini, pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan dalam rangka mencapai tujuan penelitian. Peneliti menggunakan beberapa teknik untuk mengumpulandata, yaitu *observasi* dan *study literature*. Kemudian tahap persiapan data, persiapan data dilakukan agar data sesuai dengan kebutuhan *Data Mining*. Pada tahap persiapan data ini meliputi pemilihan atribut, *data training*, *data testing*, dan pengelompokkan data. Datayang dipakai pada penelitian ini sebanyak 263 data yang terdiri dari 9 atribut.

3. HASIL DAN PEMBAHASAN

3.1 Model dan Variabel

Persiapan data dilakukan berdasarkan penemuan masalah yang terdapat pada sistem pembiayaan. Permasalahan yang ditemukan pada penelitian ini yaitu sulitnya bagian kredit dalam tahap evaluasi, untuk menentukan pemohon yang layak mendapatkan fasilitas kredit dan yang menguji model sebanyak 114 sebagai data training dan 49 sebagai data testing. Untuk mempermudah proses klasifikasi dilakukan pengelompokkan nilai masing-masing atribut yang digunakan dalam penelitian. Pengelompokkan didasarkan pada nilai masing-masing atribut yang digunakan dalam penelitian. Sebagian variabel datayang ada dalam bentuk angka, untuk penggunaan model algoritma naïve bayes data angka harus ditransformasikan ke dalam klasifikasi atau kelompok berdasar interval. Pilihankelas interval kelas tergantung pada jumlah kelas untuk ditribusi tertentu danukuran data. Berikut adalah rumus untuk menentukan ukuran interval kelas

Banyak Kelas Interval

$$P_i = \frac{R}{n} = 1 + 3.3 \log n$$

Keterangan:

P_i = Panjang Kelas Interval

n = Banyaknya Kelas Interval

R = Rentang

Diketahui

Min = 1,000,000

Max = 160,000,000

Rentang = (max - min) = (160,000,000 – 1,000,000) = 159,000,000

Banyak kelas interval = $1+3.3 \log 163 = 8$

Panjang kelas interval = (rentang / banyak kelas interval) = (159,000,000 / 8) = 19,875,000

a. Kelas Pertama

Panjang kelas pertama = (min + panjang kelas interval) - 1 = (1,000,000 + 19,875,000) - 1 = 20,874,999

Maka interval kelas pertama adalah (1,000,000 - 20,874,999).

b. Kelas Kedua

Batas bawah kelas kedua dimulai dari setelah batas terakhir pada kelas pertama maka batas bawahnya adalah 20,875,000

Panjang kelas kedua = (20,875,000 + panjang kelas interval) - 1 = (20,875,000 + 19,875,000) - 1 = 40,749,999

Maka kelas interval kedua adalah (20,875,000 - 40,749,999). Untuk kelas-kelas berikutnya cukup mengikuti langkah kelas kedua sampai batas angka max pada nilai atribut pinjaman.

Pengelompokkan nilai untuk atribut Jangka Waktu kredit, digunakan untuk menerangkan jangka waktu pelunasan kredit nasabah, menurut Gunarto Suhardi (2003) [30] dalam penggolongan jangka waktu kredit dibagi menjadi 3 kategori yaitu:

- 1 – 12 bulan = Jangka Pendek
- 13 – 36 bulan = Jangka Menengah
- > 36 = Jangka Panjang

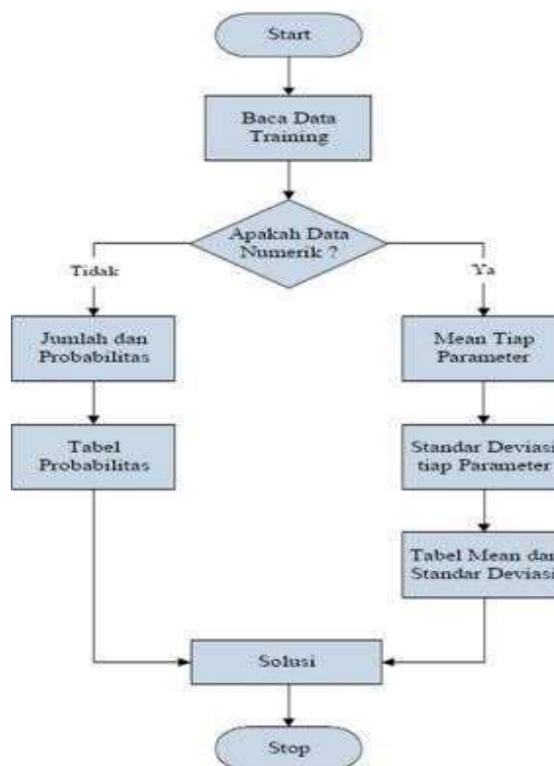
Tabel 2. Atribut dan Nilai Kategori

Atribut	Keterangan
Jenis Kelamin	<ul style="list-style-type: none"> • Laki-laki = L • Perempuan = P
Umur	<ul style="list-style-type: none"> • 20 – 24 • 25 – 29 • 30 – 34 • 35 – 39 • 40 – 44 • 45 – 49 • 50 – 54 • 55 – 59
Status	<ul style="list-style-type: none"> • Kawin • Belum Kawin
Pekerjaan	<ul style="list-style-type: none"> • Petani • Nelayan • Wiraswasta • PedagangPNS
Pinjaman	<ul style="list-style-type: none"> • 1,000,000 – 20,874,999 • 20,875,000 – 40,749,999 • 40,750,000 – 60,624,999 • 60,625,000 – 80,499,999 • 80,500,000 – 100,374,999 • 100,375,000 – 120,249,999 • 120,250,000 – 140,124,999 • 140,125,000 – 160,000,000

Jangka Waktu	<ul style="list-style-type: none"> • 1 – 12 bulan = Jangka Pendek • 13 – 36 bulan = Jangka Menengah • > 36 = Jangka Panjang
Sistem	<ul style="list-style-type: none"> • Flat = F • RC
Agunan	<ul style="list-style-type: none"> • BPKB Sepeda Motor = BPKB SPM • BPKB Mobil = BPKB MBL • Sertifikat Rumah = Sertifikat • Akta Kapal • Lain-Lain
Koleabilitas	<ul style="list-style-type: none"> • Lancar • Tidak Lancar

3.2 Algoritma Program

Algoritma yang digunakan yaitu algoritma naïve bayes. Algoritma klasifikasi naïve bayes adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Bentuk pemodelan algoritma naïve bayes dapat dilihat pada gambar 2.



Gambar 2. Flowchart Naïve Bayes

3.3 Implementasi Dengan Perhitungan Naïve Bayes

Klasifikasi data set apabila diberikandata testing sebagai berikut:

- Jenis Kelamin = L
- Umur = 35-39
- Status = Kawin



- d. Pekerjaan = PNS
- e. Pinjaman = 1.000.000 – 20.874.999
- f. Jangka Waktu = Jangka Menengah
- g. Sistem = F
- h. Agunan = BPKB_SPM

Apabila diberikan data testing, maka klasifikasi dapat ditentukan melalui langkah berikut:

1. Menghitung $P(C_i)$, jumlah *class* dari kolektibilitas berdasarkan klasifikasi yang dibentuk (*prior probability*).
 - a. *Prior* Lancar atau C1 (Class Kolektibilitas = “Lancar”) $P(Y=Lancar) = 7/13 = 0.538$
 - b. “Jumlah data lancar pada data training dibagi dengan jumlah keseluruhan data”
 - c. *Prior* Tidak Lancar atau C2 (Class Kolektibilitas = “Tidak Lancar”) $P(Y=Tidak Lancar) = 6/13 = 0.461$ “Jumlah data Tidak Lancar pada data training dibagi dengan jumlah keseluruhan data”
2. Menghitung $P(X|C_i), i = 1,2$, jumlah kasus yang sama pada setiap atribut dari kelas Kolektibilitas (Lancar / Tidak Lancar) berdasarkan data testing.

$$P(\text{Jenis Kelamin=L} | Y=Lancar) = 3/7 = 0.428$$

$$P(\text{Jenis Kelamin=L} | Y=Tidak Lancar) = 4/6 = 0.666$$

$$P(\text{Umur=35-39} | Y=Lancar) = 1/7 = 0.142$$

$$P(\text{Umur=35-39} | Y=Tidak Lancar) = 1/6 = 0.166$$

$$P(\text{Status=Kawin} | Y=Lancar) = 6/7 = 0.857$$

$$P(\text{Status=Kawin} | Y=Tidak Lancar) = 4/6 = 0.666$$

$$P(\text{Pekerjaan=PNS} | Y=Lancar) = 1/7 = 0.142$$

$$P(\text{Pekerjaan=PNS} | Y=Tidak Lancar) = 3/6 = 0.5$$

$$P(\text{Pinjaman=1.000.000-20.874.999} | Y=Lancar) = 6/7 = 0.857$$

$$P(\text{Pinjaman=1.000.000-20.874.999} | Y=Tidak Lancar) = 2/6 = 0.333$$

$$P(\text{Jangka Waktu=Jangka Menengah} | Y=Lancar) = 1/7 = 0.142$$

$$P(\text{Jangka Waktu=Jangka Menengah} | Y=Tidak Lancar) = 4/6 = 0.666$$

$$P(\text{Sistem=F} | Y=Lancar) = 5/7 = 0.714$$

$$P(\text{Sistem=F} | Y=Tidak Lancar) = 5/6 = 0.833$$

$$P(\text{Agunan=BPKB_SPM} | Y=Lancar) = 5/7 = 0.714$$

$$P(\text{Agunan=BPKB_SPM} | Y=Tidak Lancar) = 1/6 = 0.166$$

3.4 Pengujian Naïve Bayes

Pengujian ini bertujuan untuk mengetahui proses algoritma naïve bayes dalam mengklasifikasikan data ke dalam kelas yang telah ditentukan. Pada uji coba ini diberikan data *training* untuk membentuk tabel probabilitas. Kemudian selanjutnya akan diberikan data *testing* untuk menguji tabel probabilitas yang sudah terbentuk. Jumlah data *training* adalah 114 *record* sedangkan data *testing* 49 *record*. Pengujian dilakukan untuk melihat tingkat akurasi dari pemodelan yang di dapat, untuk menghitung tingkat akurasi digunakan persamaan jumlah data yang benar dibagi dengan jumlah data yang di uji kemudian di kalikan dengan 100%. Berikut adalah gambar sampel hasil uji yang di dapat.

Kolektibilitas	Prediksi	Keterangan
Tidak Lancar	Lancar	Salah
Lancar	Lancar	Benar
Tidak Lancar	Lancar	Salah
Lancar	Lancar	Benar
Tidak Lancar	Lancar	Salah
Tidak Lancar	Tidak Lan...	Benar
Lancar	Lancar	Benar
Lancar	Lancar	Benar

Gambar 3. Sampel Hasil Uji

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan dari hasil analisis, perancangan, implementasi dan pengujian yang telah di bahas pada bab-bab sebelumnya, maka penulis dapat menarik beberapa kesimpulan, yaitu:

1. Aplikasi data mining dengan algoritma naïve bayes ini dapat digunakan pada Rukun Artha Santosa Juwana Pati.
2. Aplikasi data mining dengan algoritma naïve bayes ini dapat mempermudah pihak untuk melakukan evaluasi penentuan kelayakan pemohon dengan cepat.
3. Pengujian pada data rekapitulasi peminjam yang berjumlah 163 record dari Rukun Artha Santosa Juwana Pati dengan proses *mining* algoritma naïve bayes menghasilkan tingkat akurasi 71%, dimana dalam pengujian model data, data set dibagi menjadi dua bagian yaitu data *training* dan data *testing* dengan persentase 70% untuk data *training* dan sisanya 30% untuk data *testing*.
4. Penentuan data *training* dapat mempengaruhi hasil pengujian, karena pola dari data training tersebut akan dijadikan sebagai rule untuk menentukan kelas pada data *testing*. Sehingga besar atau kecil tingkat akurasi yang di dapat dipengaruhi oleh penentuan data *training*.

4.2 Saran

Beberapa saran yang dapat diberikan penulis untuk perbaikan dan penelitian selanjutnya antara lain:

1. Dari sisi data yang digunakan diharapkan di masa mendatang jumlah record dan jumlah atribut yang digunakan dapat bertambah, dikarenakan algoritma naïve bayes untuk mendapatkan tingkat akurasi yang tinggi dibutuhkan data yang banyak dan juga atribut pendukung keputusan yang banyak.
2. Pada bagian program untuk melakukan pengujian ataupun prediksi masih dilakukan satu per satu tiap record, untuk pengembangan lebih lanjut diharapkan program ini dapat melakukan pengujian dan prediksi secara menyeluruh.

DAFTAR PUSTAKA

- [1] Vercellis, C., Business Intelligence, Data Mining and Optimization for Decision Making, England: Wiley, 2009.



- [2] N. Fitriani, "Perancangan dan Pembuatan Aplikasi Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Resiko Kredit Pada," Skripsi Teknik Informatika AMIKOM, Jogjakarta, 2014.
- [3] D. Indonesia, "Kamus Besar Bahasa Indonesia (KBBI)," Pusat Bahasa Departemen Pendidikan Nasional, Jakarta, 2008.
- [4] C. C. Ciptohartono, "Algoritma Klasifikasi Naive Bayes untuk Menilai Kelayakan Kredit," Skripsi Teknik Informatika Universitas Dian Nuswantoro, Semarang, 2014.
- [5] Kabir et al., "Credit Risk Assessment and Evaluation System for Industrial Project," International Journal of Trade, Economics and Finance, vol. 1, pp. 331-341, 2010.
- [6] E. S. Y. Pandie, "Implementasi Algoritma Data Mining K-Nearest Neighbour (K-NN) dalam Pengambilan Keputusan Pengajuan Kredit," Thesis Sistem Informasi Universitas Diponegoro, Semarang, 2012.
- [7] R. D. J. d. Noeryanti, "Aplikasi Metode K-Nearest Neighbor dan Analisis Diskriminan Untuk Analisis Resiko Kredit Pada Simpan Pinjam Di Kopinkra Sumber Rejeki," Jogjakarta, 2014.
- [8] R. Oktrivianto, "Aplikasi Data Mining untuk Memprediksi Kelas Resiko Pemberian Kredit Menggunakan Support Vector Machine (SVM)," Tesis Matematika Institut Teknologi Sepuluh Nopember, Surabaya, 2008.
- [9] M. Rizki, "Implementasi Data Mining untuk Memprediksi Data Nasabah Bank dalam Penawaran Deposito Berjangka Dengan Menggunakan Algoritma Klasifikasi Naive Bayes," Skripsi Teknik Informatika GUNADARMA, Jakarta, 2014.
- [10] M. Berry, Data Mining Techniques, John Wiley & Sons, 2004.
- [11] F. A. Hermawati, Data Mining, Yogyakarta: ANDI, 2013.
- [12] Bustami, "Penerapan Algoritma Naive Bayes untuk Mengklasifikasi Data Nasabah Asuransi," Jurnal Informatika, vol. 8, p. no.1, 2014.
- [13] C. J. a. R. G. Daniela Xhemali, "Naive Bayes vs. Decision Tree vs. Neural Networks in the Classification of Training Web Pages," IJCSI International Journal of Computer Science Issues, vol. 4, p. No.1, 2009.
- [14] E. D. M. a. M. Aryuni, "Comparative Study of Data Mining Model for Credit Card Application Scoring In Bank," Journal of Theoretical and Applied Information Technology, vol. 59, p. No. 2, 2014.
- [15] F. H. N. d. P. Suryati, "Aplikasi Sistem Pendukung Keputusan Pengajuan Kredit Sepeda Motor," in Seminar Nasional Teknologi Informasi & Komunikasi Terapan, Semarang, 2013.
- [16] K. d. E. T. Luthfi, Algoritma Data Mining, Yogyakarta: Andi, 2009.
- [17] M. I. Hasan, Pokok-pokok Materi Statistik I (Statistik Deskriptif), Jakarta: Bumi Aksara, 2001