

KLASIFIKASI DATA DEBITUR UNTUK MENENTUKAN KELAYAKAN KREDIT DENGAN MENGGUNAKAN METODE NAIVE BAYES

Deny cahya Mahendra¹, Achmad Wahid Kurniawan²

Mahasiswa Program Studi Teknik Informatika-S1, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang
Jl. Imam Bonjol No. 207, Jl.Nakula No.5-11, Semarang, Kode Pos.50131, Telp.(024) 3517261
Email : deny.cahya76@gmail.com, 111201106051@mhs.dinus.ac.id

Abstrak

Kredit adalah peminjaman uang dengan pembayaran secara tidak langsung atau membayar dengan secara berkala, dengan batas jumlah pinjaman tertentu yang diizinkan oleh pihak bank. Kredit merupakan sumber utama penghasilan dalam sebuah bank dan resiko bisnis. Karena dalam proses pemberian kredit tidak jarang terjadi kredit macet, di mana akan terjadi tunggakan kredit dalam masa angsuran. Data mining merupakan teknik yang memanfaatkan data dengan jumlah yang besar untuk mendapatkan informasi atau data yang berharga untuk mengambil keputusan yang penting. Data mining juga telah terbukti digunakan dalam perbankan yang mengklasifikasikan data yang berguna dan berukuran besar. Dalam penelitian ini studikusus yang dilakukan pada data debitur Bank BRI dengan menggunakan model Naive Bayes (NBC). Dari hasil penelitian ini untuk menentukan prediksi kelayakan kredit lancar atau kredit macet, dan evaluasi performance naive bayes. Untuk hasil penelitian algoritma naive bayes ini, dapat di terapkan dalam menilai kelayakan kredit pada akurasi data awal sebesar 79,84%, sedangkan data yang sudah melalui tahap pengolahan data yaitu sebesar 88,61%, dan terakhir data yang sudah melalui pengolahan data dan konversi data memiliki accuracy paling tinggi yaitu 90,28%, data yang diolah dengan pre-processing lebih unggul dibandingkan data yang belum diolah, sekalipun algoritma naive bayes mampu menangani data yang hilang atau tidak lengkap.

Kata kunci: data mining, debitur, klasifikasi, algoritma naive bayes.

Abstract

Credit is lending money which payment method by using indirect payment or gradual payment method, with certain loan amount limits permitted by the bank. Credit is the main source of income of a bank and also a business risk. Because in a credit granting process, bad credit is not rare, where could occur loan arrears installment in the future. Data mining is a technique that utilizes large amounts of data to obtain valuable information or data to take important decisions. Data mining has also been shown to be used in banking to classify useful and large data. In this research, case study carried out at BRI Bank debtor data using the model Naive Bayes (NBC). From the results of this study is to determine the prediction of credit feasibility; good credit or bad credit, and performance evaluation of naive Bayes. For this naive Bayes algorithm research results can be applied in assessing the credit feasibility on the accuracy of the initial data by 79.84%, whereas the data that has been through a phase of data processing is equal to 88.61%, and the last data which has been through data processing and conversion data has the highest accuracy is 90.28%, data processed by pre-processing is superior compared to the data that has not been processed, even naive Bayes algorithm is capable of handling the data which is missing or incomplete.

Keywords: data mining, debtors, naive bayes algorithm, classification.

1. PENDAHULUAN

1.1 Latar Belakang

Untuk Salah satu pelayanan di dalam dunia perbankan adalah pemberian pinjaman kredit pada nasabah yang memenuhi syarat dalam suatu bank [1]. Pada perbankan kredit merupakan sumber utama penghasilan bank sekaligus juga sumber resiko bisnis paling besar karena dimungkinkannya kredit macet dimana resiko kredit yaitu kemungkinan penurunan hasil kredit dari peminjam yang mempunyai reputasi tidak produktif [3] Permasalahan yang sering di hadapi dalam perbankan adalah banyaknya nasabah yang menunggak dalam pembayaran.

Dalam analisa kredit pihak bank menggunakan prinsip 5c yaitu *The Five C's of Credit Analysis*. *The Five C's of Credit Analysis* merupakan alat untuk menganalisis kredit yang mencakup *character, capacity, capital, collateral* dan *condition* dari untuk menganalisa nasabah yang mengajukan kredit [4] karena dalam menganalisa data dengan jumlah yang besar sering terjadinya kesalahan dalam analisa.

Data mining telah terbukti sebagai alat yang memegang peran penting dalam dunia perbankan dan ritel, dimana data mining dapat mengidentifikasi data atau informasi yang berguna dari data yang berukuran besar. Ada banyak tehnik dalam pembangunan model yang menggunakan aplikasi dari data mining tersebut [6].

Naive bayes classifier merupakan suatu metode yang menangani masalah bersifat hipotesis dimana dapat mendisain suatu klasifikasi dalam memisahkan suatu objek. Untuk credit croring penelitian yang dilakukan oleh Leung dan ceong [3] telah mendapatkan akurasi tertinggi dari 6 model dan 22 algoritma klasifikasi pada german data set, yang menemukan bahwa naive

bayes yang mempunyai akurasi tertinggi. dan telah di buktikan dalam penelitian yang membedakan akurasi decision tree dan naive bayes dengan menggunakan 100 record dalam prediksi kelayakan kredit. Dan hasilnya yaitu naive bayes classifier yang mempunyai akurasi lebih baik (82%) daripada decision tree (76%) [6]. Dalam penelitian ini, penulis akan menggunakan metode Naive Bayes. Karena Naive Bayes Classifier merupakan metode yang diimplementasikan dan memiliki nilai akurasi yang cukup tinggi yaitu 82% dalam menilai kelayakan kredit. Hal tersebutlah yang melandasi penerapan metode tersebut untuk implementasi Sistem Pencairan Kredit [9].

1.2 Rumusan Masalah

Bagaimana Menerapkan metode *data mining* menggunakan model *Naive Bayes* untuk mengklasifikasikan data debtur Bank BRI KC Rembang. Dan membangun aplikasi yang efisien sebagai implementasi metode klasifikasi.

2. METODE

2.1 Bank

Menurut Undang-Undang No. 10 Tahun 1998 Perbankan adalah sesuatu yang menyangkut segalanya tentang bank, yaitu yang mencakup kelembagaan, kegiatan usaha, serta cara dan proses dalam melaksanakan kegiatan usaha bank. bank sendiri adalah badan usaha yang menghimpun atau menyimpan dana dari masyarakat dalam bentuk simpanan dan menyalurkannya kepada masyarakat dalam bentuk kredit ataupun bentuk lainnya, tidak lain tujuannya adalah untuk meningkatkan taraf hidup rakyat banyak [10].

2.2 Kredit

Apa itu kredit. Kredit yaitu penyediaan uang ataupun tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau

kesepakatan pinjam meminjam antara pihak peminjam dan tentunya dengan pihak bank yang mewajibkan dari pihak peminjam untuk melunasi utang yang di pinjam dari bank setelah jangka waktu tertentu dengan pemberian bunga.

A. *The Five C's Of Credit Analysis.*

The Five C's Of Credit Analysis berguna untuk menganalisis kredit dalam perbankan yang terdiri dari [9]:

- *Character* data yang mencakup data kepribadian dari calon nasabah kredit.
- *Capacity* merupakan kemampuan calon nasabah kredit dalam mengelola suatu usahanya.
- *Capital* kondisi atau kekayaan yang di miliki atau perusahaan yang di miliki dan di kelolanya
- *collateral* adalah jaminan yang dimiliki oleh nasabah kredit. Dan kemungkinan dapat di sita oleh bank jika sewaktu waktu nasabah tidak memenuhi kewajibannya.

Condition kondisi ekonomi dari pihak nasabah.

2.3 Klasifikasi

Data Mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu:[16] Deskripsi, Estimasi, Prediksi, Klasifikasi, Pengklusteran, dan Asosiasi. Klasifikasi merupakan bagian dari algoritma data mining, klasifikasi ini adalah algoritma yang menggunakan data dengan target (*class/label*) yang berupa nilai kategorikal/nominal. Menurut Gorunescu [17] proses klasifikasi didasarkan pada empat komponen mendasar, yaitu:

1. Kelas (*Class*)

Variabel dependen dari model, merupakan variabel kategorikal yang merepresentasikan "label" pada objek setelah klasifikasinya. Contoh kelas semacam ini adalah: adanya kelas penyakit jantung, loyalitas pelanggan, kelas bintang (galaksi), kelas gempa bumi (badai), dll.

2. Prediktor (*Predictor*)

Variabel independen dari model, direpresentasikan oleh karakteristik (atribut) dari data yang akan diklasifikasikan dan berdasarkan klasifikasi yang telah dibuat. Contoh prediktor tersebut adalah: merokok, konsumsi alkohol, tekanan darah, frekuensi pembelian, status perkawinan, karakteristik (satelit) gambar, catatan geologi yang spesifik, kecepatan dan arah angin, musim, lokasi terjadinya fenomena, dll.

3. Pelatihan dataset (*Training dataset*)

Kumpulan data yang berisi nilai-nilai dari kedua komponen sebelumnya dan digunakan untuk melatih model dalam mengenali kelas yang cocok/sesuai, berdasarkan prediktor yang tersedia. Contoh set tersebut adalah: kelompok pasien yang diuji pada serangan jantung, kelompok pelanggan supermarket (diselidiki oleh intern dengan jajak pendapat), database yang berisi gambar untuk monitoring teleskopik dan pelacakan objek astronomi, database badai, database penelitian gempa.

4. Dataset Pengujian (*Testing Dataset*)

Berisi data baru yang akan diklasifikasikan oleh (*classifier*) model yang telah dibangun di atas sehingga akurasi klasifikasi (*model performance*) dapat dievaluasi.

Berikut beberapa model (metode) klasifikasi yang paling populer:[17]

1. *Decision/classification trees*;
2. *Bayesian classifiers/Naive Bayes classifiers*;
3. *Neural networks*;
4. *Statistical analysis*;
5. *Genetic algorithms*;
6. *Rough sets*;
7. *k-nearest neighbor classifier*;
8. *Rule-based methods*;
9. *Memory based reasoning*;
10. *Support vector machines*

2.4 Naive Bayes Classifier

Bayesian Classification yaitu pengklasifikasian statistik yang bisa digunakan dalam memprediksi probabilitas keanggotaan suatu class. Naive bayes dihitung berdasarkan Teorema Bayes yang memiliki bentuk umum sebagai berikut:

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

Dalam hal ini:

X= data dengan class yang belum diketahui.

H= hipotesis data X merupakan suatu class spesifik

$P(H|X)$ = probabilitas hipotesis H berdasarkan kondisi X (posterior probability)

$P(H)$ = probabilitas hipotesis H (prior probability).

Contoh D adalah sebuah training set dari sample dan class label terkait, dan setiap tuple direpresentasikan dengan sebuah n ke D vektor atribut $X = (x_1, x_2, \dots, x_n)$. Seharusnya terdapat m class C_1, C_2, \dots, C_m . Klasifikasi bertujuan untuk menurunkan posterior maksimum, misalnya $P(C_i|X)$ maksimal. Hal ini bisa diturunkan dari teorema Bayes,

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

Karena $P(X)$ bernilai konstan untuk semua kelas maka yang perlu dimaksimalkan hanya

$$P(C_i|X) = P(X|C_i)P(C_i)$$

Dengan asumsi bahwa setiap atribut bersifat independent atau tidak saling mempengaruhi antara satu atribut dengan atribut yang lain. Sehingga didapatkan persamaan:

$$P(X|C_i) = \prod_{k=1}^n P(x_k|C_i) = P(x_1|C_i) \times P(x_2|C_i) \times \dots \times P(x_n|C_i)$$

Dalam Penelitian ini melibatkan data yang bernilai kontinyu, sehingga dalam proses perhitungan *posterior probabilitynya* akan memerlukan mean (rata-rata) dan standar deviasi dari setiap atribut yang ada. Karena jika x_k bernilai kontinyu, $P(x_k|C_i)$ biasanya akan dihitung berdasarkan distribusi Gaussian (normal) dengan

μ = mean, menyatakan rata-rata dari seluruh atribut

σ = *standard deviation*, menyatakan varian dari seluruh atribut

g = Peluang

x_i = nilai Attribut Ke i

sebagai berikut:

$$g(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Sehingga $P(X_k|C_i)$ menjadi setara dengan

$$P(X|C_i) = g(x_k, \mu_{C_i}, \sigma_{C_i})$$

dan untuk mean μ dan standard deviation σ sendiri akan dihitung sebagai berikut:

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i$$

Sedangkan untuk standart deviation σ akan di hitung

$$\sigma = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2 \right]^{0.5}$$

Di Dalam dunia perbankan terdapat alat analisis yang digunakan untuk mempertimbangkan pencairan kredit yang disebut dengan *The Five C's of Credit Analysis* yang terdiri dari: [9]

3. HASIL DAN PEMBAHASAN

3.1 Analisa Data

Data yang digunakan pada penelitian ini berasal dari data debitur Bank BRI KC Rembang periode 2014 dengan jumlah data sebanyak 667 *record*. Terdiri dari 17 atribut yaitu No, Cabang, Nama Debitur, No Kredit, Alamat, Jenis Kelamin, Umur, Agama, Status, Pekerjaan, Penghasilan, Maks Kredit, Jangka Waktu Kredit, Angsuran, Suku Bunga, Outstanding, Kredit macet / lancar.

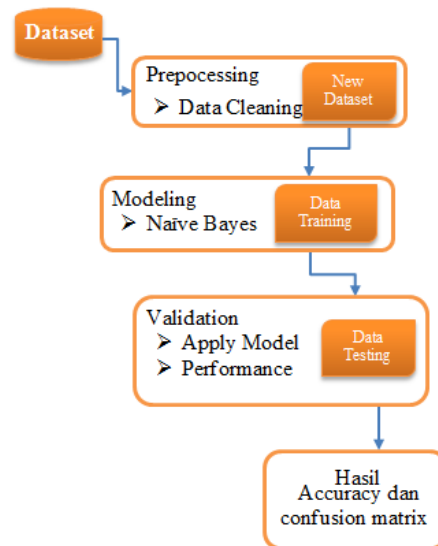
Selanjutnya pengolahan data awal Jumlah record atau data awal yang diperoleh dari pengumpulan data sebanyak 667 data/*record* dan 17 atribut, Untuk mendapatkan data yang berkualitas, beberapa teknik yang dilakukan sebagai berikut: [15]:

1. Data Cleaning

Terdiri dari Identifikasi atau menghilangkan *outliers* dan menghilangkan data *nois*. Mengisi nilai - nilai yang hilang atau data tidak

lengkap (*missing value*), Algoritma naive bayes mempunyai keunggulan tersendiri yaitu menangani data yang tidak lengkap (*missing value*), Memperbaiki data yang tidak konsisten. Dan Memecahkan redundansi yang di sebabkan oleh integrasi data.

2. *Data integration and transformation.* Langkahah selanjutnya yaitu *integration* teknik yang dipakai menganalisis data korelasi, atribut yang redundan dan duplikat data, dan *transformation.*
3. *Data reduction* Mereduksi data set, dengan mengurangi jumlah attribute dan record agar lebih sedikit namun bersifat informative, Sehingga didapatkan data sebanyak 489 data/*record* dan 9 atribut (table 3.1.)



Gambar 1 : Model untuk menemukan performance algoritma

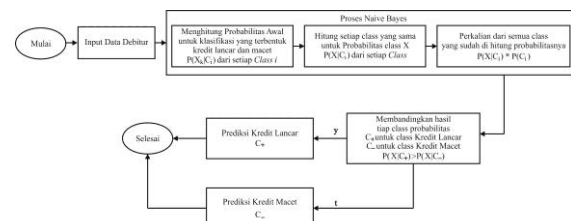
Tabel 1 : Atribut yang digunakan

No	Atribut
1	Jenis Kelamin
2	Umur
3	Status
4	Pekerjaan
5	Penghasilan(perbulan)
6	Jangka Waktu Kredit
7	Angsuran
8	Suku Bunga
9	Kredit macet / lancar

A. Model dan Metode yang di usulkan

1. *Modeling*

Metode yang akan digunakan dalam penelitian ini adalah Algoritma Naive Bayes. Dalam pemodelan ini, algoritma naive bayes akan di cari Performance Vector (accuracy) dan ConfusionMatrix.



Gambar 2 : Model algoritma naive bayes

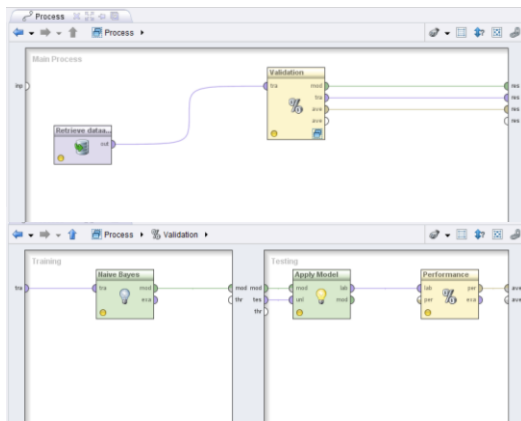
Mulai Identifikasi sampel dari *data set* baca data. selanjutnya $P(X_i|C_i)$ menghitung jumlah class dari klasifikasi yang sudah terbentuk yaitu class Kredit Lancar dan Kredit Macet untuk setiap class. Kemudian $P(X|C_i)$ menghitung jumlah kasus yang sama dari kelas yang sama X, dalam kasus *data set* pada penelitian ini terdiri dari 2 class yaitu debitur dengan kelas kredit Lancar yang dinyatakan dengan simbol (+) dan debitur dengan kelas kredit Macet yang dinyatakan dengan simbol (-). Kemudian hitung $P(X|C_i), i = +, -$ untuk setiap kelas atau atribut. Setelah itu dibandingkan, jika $P(X|C_+) > P(X|C_-)$ maka kesimpulannya adalah C_+ atau pada penelitian ini berarti kredit Lancar. Jika $P(X|C_+) < P(X|C_-)$ maka kesimpulannya C_- atau kredit Macet.

3.2 Hasil dan Pembahasan

Data akhir berjumlah 489 data/record, dan Atribut yang akan digunakan adalah 2 jenis type atribut yaitu attribute kategorikal ini adalah salahsatu tipe attribute yang domainnya merupakan sebuah himpunan atau memiliki batas, nilai modelnya yaitu jenis kelamin, status, dan pekerjaan. Dan yang kedua adalah Atribut Numerik adalah attribute yang domainnya bilangan riil atau integer, nilai modelnya yaitu umur, penghasilan(perbulan), jangka waktu kredit, angsuran, suku bunga. Dan kredit macet/lancer yang merupakan label target.

A. Eksperimen dan Pengujian Metode

Eksperimen penelitian dilakukan menggunakan *Rapid Miner*. Dalam *Rapid Miner* terdapat beberapa operator pengolahan data sebelum dilakukan proses *mining* menggunakan Naive Bayes yaitu *Retrieve Data*, *Validation*, *Naive Bayes Classifier*, *apply Model* dan *Performance*. Kemudian dilakukan pengujian model. Naive Bayes. Tampilan proses utama dapat dilihat seperti pada gambar berikut :



Gambar 3 : Desain model Naive Bayes

B. Percobaan dan Hasil

Percobaan pertama yang akan dilakukan menggunakan yang belum melalui pengolahan data. Jumlah data sebanyak 667 record Terdiri dari 17 atribut yaitu No, Cabang, Nama Debitur, No Kredit, Alamat, Jenis

Kelamin, Umur, Agama, Status, Pekerjaan, Penghasilan, Maks Kredit, Jangka Waktu Kredit, Angsuran, Suku Bunga , Outstanding, dan keterangan Kredit macet / lancar.

accuracy: 79.84%			
	true Kredit lancar	true Kredit macet	class precision
pred. Kredit lancar	194	10	95.10%
pred. Kredit macet	41	8	16.33%
class recall	82.55%		44.44%

Gambar 4 : Tabel confusion matrix hasil pengujian akurasi data awal

Setelah itu pada data awal dilakukan *preprocessing* yaitu pemilihan data dengan dan dilakukan pemilihan atribut yang digunakan pada penelitian. Kemudian di dapatkan Jumlah data sebanyak 489 data/record dan terdiri dari 9 atribut Jenis Kelamin Umur, Status, Pekerjaan, Penghasilan(perbulan), Jangka Waktu Kredit, Angsuran, Suku Bunga, dan keterangan Kredit macet / lancar

accuracy: 88.61%			
	true Kredit lancar	true Kredit macet	class precision
pred. Kredit lancar	378	22	94.50%
pred. Kredit macet	28	11	28.21%
class recall	93.10%		33.33%

Gambar 5 : Tabel confusion matrix hasil pengujian akurasi data yang melalui tahap *preprocessing* data

Untuk hasil *accuracy* percobaan pertama 79,84% dapat dikategorikan dalam *Fair classification*, dan percobaan kedua dengan hasil akurasi 88,61% sudah termasuk dalam *good classification*.

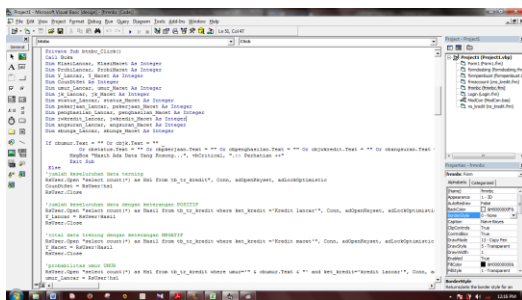
accuracy: 90.28%			
	true Kredit lancar	true Kredit macet	class precision
pred. Kredit lancar	346	23	93.77%
pred. Kredit macet	15	7	31.82%
class recall	95.84%		23.33%

Gambar 6 : Tabel confusion matrix hasil pengujian *accuracy* data yang sudah dikonversi/diklasifikasikan.

Acuuracy pada percobaan dengan data yang sudah di konversi ternyata diperoleh akurasi tertinggi yaitu 90,28% dan merupakan rentang akurasi *Excellent classification*.

C. Implementasi Algoritma Naive Bayes menggunakan Microsoft Visual Basic 6.0

Untuk proses prediksi debitur *kredit lancar* dan *kredit macet*, maka peneliti menggunakan Microsoft Visual Basic 6.0 untuk mengimplementasi penelitian yang telah dilakukan peneliti dengan algoritma *naive bayes*. Prediksi tersebut disesuaikan dengan karakteristik yang dihasilkan dari analisis sebelumnya.



Gambar 7 : Tampilan editor source code VB 6.0



Gambar 8 : Tampilan pada GUI VB 6.0

Pada GUI (*Graphical User Interface*) Gambar 4.9. diatas, aplikasi dijalankan dengan memasukkan inputan yang ada. Pada form ini menampilkan hasil output dari penginputan data calon debitur yang telah di proses dengan algoritma *Naive Bayes*. Proses klasifikasi dipengaruhi oleh atribut atribut terpilih yang mendukung untuk menentukan kelas nasabah lancar dan macet.

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Pada tahap penelitian membuktikan bahwa algoritma Naive Bayes dapat diterapkan untuk menilai kelayakan kredit pada BANK BRI KC Rembang dan untuk pengolahan data awal merupakan tahapan yang sangat mempengaruhi hasil akurasi yang baik sehingga akurasi akhir yang dihasilkan termasuk kategori *Excellent*. Tahap pengolahan data sangat mempengaruhi nilai hasil akurasi , dan Evaluasi menggunakan *confusion matrix* dengan menggunakan algoritma *Naive Bayes* menghasilkan akurasi data awal sebesar 79,84%, sedangkan data yang sudah melalui tahap pengolahan data *preprocessing* yaitu sebesar 88,61%, dan terahir data yang sudah melalui pengolahan data dan konversi data memiliki *accuracy* paling tinggi yaitu 90,28%, jadi penilaian kelayakan kredit menggunakan algoritma naive bayes pada data debitur BANK BRI akan lebih tinggi jika di lakukan pengolahan data dan konversi data.

4.2 Saran

Berdasarkan proses penelitian yang telah dilakukan, agar penelitian ini terus berkembang, berikut saran yang diusulkan peneliti:

1. Pada penelitian ini dapat dikembangkan dengan metode klasifikasi data mining lainnya agar bias dilakukan perbandingan.
2. Untuk mengembangkan penelitian ini juga bisa dilakukan penambahan *record* data atau atribut lebih banyak untuk mendapatkan nilai akurasi yang lebih baik.
3. diharapkan dapat digunakan oleh masyarakat yang membutuhkan dan pihak bank untuk mengetahui dan meningkatkan penilaian kredit.

5. DAFTAR PUSTAKA

- [1] Kurniasih, Christin H, *Analisis 5C Sebagai Penentu Kelayakan Pemberian Kredit Pada Nasabah PT BPR Ambarawa Hartasarana*. STIE AMA Salatiga, 2010.
- [2] K. Leung, F. Cheong and C. Cheong, "Consumer Credit Scoring using an Artificial Immune System Algorithm," in *IEEE Congress on Evolutionary Computation (CEC 2007)*, Melbourne, 2007.
- [3] cahyani Damayanti, "Manfaat Laporan Keuangan Debitur Perusahaan Dagang Sebagai Bahan Pertimbangan Dalam Efektivitas Pengambilan Keputusan Pemberian Kredit Investasi," Bandung, 2007.
- [4] thamrin Abdullah dan Francis Tantri, 2012 *Bank dan Lembaga Keuangan, Ed.1-1*, Jakarta, Penerbit: Rajawali Pers.
- [5] H. Zhang and S. Sheng, "Learning Weighted Naive Bayes with Accurate Ranking," in *Proceedings of the Fourth IEEE International Conference on Data Mining (ICDM'04)*, Canada, 2004.
- [6] E. D. Madyatmaja and M. Aryuni, "COMPARATIVE STUDY OF DATA MINING MODEL FOR CREDIT APPLICATION SCORING IN BANK," *Journal of Theoretical and Applied Information Technology*, vol. 59, no. 2, pp. 269-274, 2014.
- [7] H. C. Koh, W. C. Tan and C. P. Goh, "Credit Scoring Using Data Mining Techniques," Singapore, 2004.
- [8] J. Zurada and K. N. Kunene, "Comparisons of the Performance of Computational Intelligence Methods for Loan Granting Decisions," in *Proceedings of the 44th Hawaii International Conference on System Sciences*, Hawaii, 2011.
- [9] I Wayan Budi Sentana, Erma Sulistyorini, and Ni Ketu Warastini, "Implementasi The Five C's Of Credit Analysis Dan Naive Bayes," *Konferensi Nasional Sistem dan Informatika*, 2011.
- [10] Manalu, Rumondang, *Hubungan Aspek Kelayakan Kredit Dengan Keputusan Pemberian Kredit di Bank Rakyat Indonesia Unit Nanggulan Salatiga*, STIE AMA Salatiga, 2009.
- [11] M. Berry, *Data Mining Techniques*, John Wiley & Sons, 2004.
- [12] Sumarni Adi, "Klasifikasi Data Nap (Nota Analisis Pembiayaan) Untuk," Yogyakarta, 2014.
- [13] Bustami, "Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi," *TECHSI*, vol. III, pp. 11-14, Oktober 2014.
- [14] J. Lin and J. Yu, "Weighted Naive Bayes Classification Algorithm Based on Particle Swarm Optimization," *IEEE*, pp. 444-447, 2011
- [15] F. Gorunescu, *Data Mining Concept Model Technique*, Craiova, Romania: Springer, 2011.
- [16] D. T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining*, Hoboken: John Wiley & Sons Inc., 2005.
- [17] F. Gorunescu, *Data Mining Concept Model Technique*, Craiova, Romania: Springer, 2011.