

## PREDIKSI RESIKO KREDIT MENGGUNAKAN NAIVE BAYES CLASSIFIER

Fajar Agung Nugroho<sup>1</sup>

Lailil Muflikhah, S.Kom, M.Sc.<sup>2</sup>, Candra Dewi, S.Kom, M.Sc.<sup>3</sup>

<sup>1)</sup>Mahasiswa Jurusan Matematika Program Studi Ilmu Komputer

<sup>2)</sup>Staf Pengajar Jurusan Matematika Program Studi Ilmu Komputer

Universitas Brawijaya Malang

Jalan Mayjen Haryono 169, Malang 65145, Indonesia

### ABSTRAK

*Dalam dunia perbankan, pemberian kredit merupakan kegiatan usaha yang mengandung resiko tertinggi yang berpengaruh terhadap kelangsungan usaha perbankan. Oleh karena itu, seorang pembuat keputusan harus melakukan evaluasi untuk pemohon kredit secara objektif, akurat dan konsisten. Pada penelitian ini akan digunakan evaluasi menggunakan metode Naive Bayes Classifier untuk memtukan status resiko kredit.*

*Pada penelitian ini, Metode Naive bayes Classsifier dibagi menjadi 3 tahap, yaitu preprocessing, training dan testing. Untuk mengetahui pengaruh jumlah data terhadap tingkat akurasi maka digunakan jumlah data training yang beragam, serta menggnakan prosentase kelas target yang sama dan berbeda. Pengujian sistem dengan menggunakan kelas target berbeda menghasilkan tingkat akurasi rata-rata prediksi resiko kredit adalah 73% dan 75% untuk pengujian menggunakan prosentase kelas target yang sama.*

Kata kunci : Perbankan, Resiko kredit, Naive Bayes Classifier

### I. PENDAHULUAN

Perbankan merupakan salah satu lembaga yang berkewajiban untuk memperlancar arus kegiatan dibidang ekonomi dan pengembangan perekonomian. Bank memiliki salah satu fungsi yaitu sebagai sumber dana bagi masyarakat yang membutuhkan dengan memberikan fasilitas pinjaman yang diperuntukan bagi nasabahnya, yang dinamakan dengan kredit. Dalam memberikan sebuah kredit, bank diwajibkan memiliki kepercayaan terhadap calon penerima pinjaman kredit bahwa dana yang akan diberikan akan dimanfaatkan sesuai dengan tujuan yang diajukan oleh calon penerima kredit dan pada akhirnya akan dikembalikan lagi kepada bank sesuai dengan perjanjian yang telah disepakati antara pihak Bank dengan peminjam

atau calon penerima kredit. Namun kredit yang diberikan kepada peminjam selalu memiliki sebuah resiko seperti contoh kredit tidak dapat dikembalikan sesuai dengan waktu yang telah disepakati, yang selanjutnya sering dikenal dengan istilah kredit macet atau kredit bermasalah. Kredit macet muncul salah satunya dapat disebabkan oleh kekurangmampuan pihak bank dalam menilai resiko calon debitur sehingga kreditur dituntut harus mampu mengevaluasi permohonan kredit secara cepat, objektif, akurat, dan konsisten. Hasil evaluasi tersebut kemudian akan digunakan dalam mengambil keputusan yang tepat untuk menerima atau menolak permohonan kredit tersebut. Prosedur dalam pengambilan keputusan

yang umumnya digunakan perbankan saat ini yaitu dengan meminta pemohon (nasabah) mengisi formulir berupa daftar pertanyaan dan melengkapi permohonan kredit dengan berkas-berkas yang diperlukan oleh perbankan, untuk kemudian dilakukan penilaian permohonan kredit tersebut [17]. Berdasarkan penilaian parameter tersebut, maka bank dapat membuat peringkat (*rating*) calon debitur untuk memutuskan kelayakan debitur menerima kredit [10].

Jumlah data pemohon kredit yang besar akan membutuhkan waktu yang cukup lama dalam melakukan analisis secara manual. Oleh karena itu diperlukan perangkat komputer dalam menganalisis data pemohon kredit. Metode yang dapat digunakan untuk melakukan analisis data salah satunya adalah *data mining*.

## II. TINJAUAN PUSTAKA

### 2.1 Kredit

Kredit merupakan penyediaan uang atau tagihan-tagihan yang dapat disamakan dengan itu berdasarkan persetujuan pinjam meminjam antara bank dengan pihak lain dalam hal mana pihak peminjam berkewajiban melunasi hutangnya setelah jangka waktu tertentu dengan jumlah bunga yang telah ditetapkan [18].

### 2.2 Resiko Kredit

Risiko kredit (*credit risk*) adalah risiko yang ditanggung oleh kreditur disebabkan oleh kegagalan nasabah dalam mengembalikan pinjaman karena keengganan, ketidakmampuan, kebangkrutan, atau berbagai alasan lainnya mengapa nasabah tidak dapat mengembalikan pinjamannya. Risiko kredit merupakan probabilitas kegagalan nasabah, dan probabilitas ini mewakili tingkatan risiko kredit dari nasabah. Penaksiran risiko kredit sangatlah penting bagi kreditur untuk membantu menentukan Kemungkinan risiko dan selanjutnya membuat keputusan yang benar mengenai permohonan kredit yang diajukan.

Penelitian Restu, digunakan Algoritma CART (*Classification and Regression Trees*) Dalam Menentukan Status Resiko Kredit (Restu, 2011). Dari penelitian tersebut, didapatkan tingkat akurasi sebesar 68%. [14].

Salah satu metode yang terdapat dalam *data mining* adalah Prediksi. Prediksi dalam *data mining* dapat dilakukan dengan proses klasifikasi. Pada pengklasifikasian, terdapat beberapa macam metode yang dapat digunakan, salah satunya adalah *Naïve Bayes Classifier*. *Naïve Bayes Classifier* merupakan metode pengklasifikasian statistik yang digunakan untuk memprediksi keanggotaan suatu kelas. Metode ini memiliki tingkat kecepatan dan akurasi yang lebih tinggi apabila diimplementasikan dalam *database* yang berukuran besar [11].

Terdapat parameter-parameter yang digunakan untuk melakukan prediksi resiko kredit. Beberapa parameter yang digunakan meliputi : besar tabungan, penjamin, jangka waktu kredit, aset jaminan, sejarah kredit, bunga terhadap penghasilan, tujuan kredit, lama tinggal, lama bekerja, pekerjaan, umur peminjam, kepemilikan telepon, besarnya kredit, kepemilikan rumah, status perkawinan, jumlah orang bertanggung, penjamin, jumlah kredit yg dimiliki, rencana angsuran, besar tabungan untuk jaminan[18].

### 2.3 Data

Definisi data adalah nilai yang mewakili sebuah fakta atau gambaran dari objek atau kejadian. Fakta dari transaksi dalam perusahaan di bidang perdagangan misalnya, dapat berupa transaksi penjualan yang meliputi waktu transaksi, pelaku transaksi, barang yang ditransaksikan beserta jumlah dan harganya. Dalam menyatakan data dapat berupa nilai yang berbentuk angka, deretan karakter, atau simbol [9].

## 2.4 Informasi

Data yang telah diproses dan diinterpretasikan menjadi sesuatu yang berguna disebut informasi. Informasi berguna dalam pengambilan keputusan karena memberikan tambahan pengetahuan sehingga meminimalkan ketidakpastian dari suatu masalah. Pemrosesan data menjadi informasi merupakan proses manipulasi atau perubahan data dengan tujuan meningkatkan fungsi dari data itu sendiri [9].

## 2.5 Data Mining

Pengertian *data mining* menurut beberapa referensi adalah sebagai berikut:

1. Proses menganalisis sekumpulan data hasil penelitian, dengan tujuan untuk menemukan hubungan antar data, dan untuk meringkas data sehingga data menjadi mudah dimengerti dan berguna bagi pemilik data [5].
2. Data mining merupakan sebuah proses penemuan yang efisien sebuah pola terbaik yang dapat menghasilkan sesuatu yang bernilai dari suatu koleksi data yang sangat besar [4].
3. Ekstraksi informasi atau pola yang penting atau menarik dari data yang ada di basis data yang besar sehingga menjadi informasi yang sangat berharga [2].
4. Serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual [9].

## 2.6 Klasifikasi

Klasifikasi adalah proses pencarian sekumpulan model atau fungsi yang menggambarkan dan membedakan kelas dengan tujuan agar model tersebut dapat digunakan untuk memprediksi kelas dari suatu obyek yang belum diketahui kelasnya. Klasifikasi memiliki dua proses yaitu membangun model klasifikasi dari sekumpulan kelas data yang sudah didefinisikan sebelumnya (*training data set*) dan menggunakan model tersebut untuk proses

klasifikasi tes data serta mengukur akurasi model [13].

## 2.7 Sistem Rekomendasi

Sistem rekomendasi adalah teknologi yang digunakan untuk memprediksi apakah seseorang akan tertarik pada suatu barang atau dapat juga digunakan untuk mengidentifikasi barang mana yang sesuai dengan kebutuhan seseorang [6].

Sistem rekomendasi adalah program atau agen yang secara otomatis membuat daftar yang berisi sekumpulan informasi yang sesuai dengan kebutuhan seseorang. Sistem rekomendasi menerima data masukan dari pengguna dan membandingkannya dengan data yang sudah ada untuk memberikan daftar yang akan direkomendasikan. Tujuannya adalah memberikan informasi yang sesuai dengan keinginan atau kebutuhan seseorang berdasarkan data yang sudah ada [7].

## 2.8 Teori Probabilitas

Probabilitas adalah suatu ukuran tentang kemungkinan suatu peristiwa (event) akan terjadi di masa mendatang. Probabilitas dinyatakan antara 0 sampai 1 atau dalam bentuk presentase. Ada tiga hal penting dalam probabilitas yaitu percobaan (*experiment*), hasil (*outcome*), dan peristiwa (*event*). Rumus umum untuk probabilitas dinyatakan dengan persamaan

$$P(A) = \frac{n(A)}{n(S)}$$

Dimana  $n(A)$  adalah banyaknya anggota dalam peristiwa A dan  $n(S)$  adalah banyaknya anggota ruang sampel [16].

## 2.9 Teori Bayesian

Teorema Bayes atau yang disebut juga dengan aturan Bayes, adalah hasil dari pengembangan teori peluang (*probability theory*), yang berhubungan dengan peluang bersyarat (*conditional probability*).

Sebelum sampai pada teorema Bayes, akan dijelaskan dahulu mengenai peluang bersyarat. Peluang suatu kejadian A dengan syarat kejadian B dapat dinyatakan dengan notasi  $P(A|B)$ . Untuk penyelesaiannya, dari beberapa literature dinyatakan dengan persamaan

$$P(A|B) = \frac{P(B \cap A)}{P(B)}, \text{ dengan } P(B) > 0$$

Pada kondisi saling lepas (*independence*), yaitu apabila kejadian A saling lepas dengan kejadian B, maka dapat dinyatakan dalam bentuk persamaan

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ P(A|B) &= P(A) \\ P(B|A) &= P(B) \end{aligned}$$

Sedangkan pada kondisi berpotongan (*intersection*) atau tidak saling lepas, yaitu A dan B adalah dua kejadian didalam sebuah populasi, maka dapat dinyatakan dalam bentuk persamaan

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

Teorema Bayes sendiri berhubungan dengan peluang bersyarat dan dinyatakan dalam bentuk persamaan

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Teorema Bayes Pada persamaan 2.2 juga dapat digunakan untuk menyelesaikan permasalahan serupa dengan jumlah variabel lebih dari satu, dan dapat dinyatakan dalam bentuk persamaan

$$P(C|F_1, \dots, F_n) = \frac{P(C)P(F_1, \dots, F_n|C)}{P(F_1, \dots, F_n)}$$

Keterangan :

- F : Sampel data yang belum diketahui kelasnya
- C : Kelas

$P(C)$  : *Prior probability*, yaitu peluang dari kelas C.

$P(F|C)$  : *Conditional probability*, yaitu peluang atribut F pada kelas C

$P(C|F)$  : *Posterior probability*, yaitu kelas C yang dicari berdasarkan atribut F.

$P(F)$  : *Marginal probability*, yaitu peluang dari masing-masing atribut F.

## 2.10 Naïve Bayes Classifier

*Naïve Bayes Classifier* adalah metode pengklasifikasian sederhana dengan teori peluang. Meski dikatakan sederhana, metode ini baik digunakan dalam proses klasifikasi. Metode ini juga sudah terbukti efektif dalam beberapa penerapannya, termasuk dalam pengklasifikasian teks, *medical diagnosis* dan *system perform management* [15]

Metode klasifikasi ini diturunkan dari penerapan teorema Bayes dengan asumsi saling lepas (*independence*) yang kuat. *Naïve Bayes Classifier* adalah metode pengklasifikasian sederhana dari model pengklasifikasian dengan peluang, dimana diasumsikan bahwa setiap atribut contoh bersifat saling lepas satu sama lain berdasarkan atribut kelas [11].

Apabila memperhatikan teorema Bayes, penyebutnya tidak tergantung pada kelas C, dan jika nilai atribut  $F_1$  sampai  $F_n$  diketahui, maka penyebutnya bernilai konstan, Sehingga persamaan tersebut dapat dinyatakan dengan persamaan

$$P(C|F_1, \dots, F_n) = \frac{1}{z} P(C)P(F_1, \dots, F_n|C) \quad (2.7)$$

Dimana z adalah factor bagi yang hanya dipengaruhi oleh  $F_1$  sampai  $F_n$ . Sedangkan untuk penyebut, dapat dinyatakan dengan persamaan  $P(C)P(F_1, \dots, F_n|C)$

$$\begin{aligned}
 &= P(C)P(F_1|C)P(F_2, \dots, F_n|C, F_1) \\
 (2.8) \\
 &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3, \dots, F_n|C, F_1, F_2) \\
 &\text{dst.}
 \end{aligned}$$

Karena adanya asumsi bahwa setiap  $F_i$  sampai  $F_n$  secara kondisional saling lepas atau setiap  $F_i$  secara kondisional saling lepas dengan  $F_j$ , dimana  $i \neq j$ , maka dapat dinyatakan dengan persamaan

$$P(F_i|C, F_j) = P(F_i|C)$$

Dalam klasifikasi dengan menggunakan teori peluang, hasil yang digunakan didapat dari hipotesis yang paling mungkin, atau dengan kata lain yang memiliki nilai peluang paling besar. Fungsi klasifikasinya dapat dinyatakan dengan persamaan

$$\begin{aligned}
 &\text{classify}(F_1, \dots, F_n) = \\
 &\text{argmax } P(C) \prod_{k=1}^n P(F_k|C)
 \end{aligned}$$

## 2.11 Evaluasi

Evaluasi ditujukan untuk mencari tingkat kesalahan (*error rate*) atau tingkat kebenaran (*accuracy rate*). Tingkat kesalahan menunjukkan berapa besar kesalahan hasil prediksi yang diberikan oleh sistem. Sedangkan tingkat kebenaran menunjukkan seberapa besar tingkat benaran hasil prediksi yang diberikan oleh sistem [1].

Dalam melakukan proses evaluasi, data yang ada harus dibagi menjadi minimal dua bagian, yang pertama digunakan sebagai data uji. Apabila tidak menggunakan data latih dan data uji secara terpisah tingkat akurasi yang diperoleh akan berlebihan. Data uji ini dapat dikatakan juga sebagai data sampel dari seluruh data [1].

Tingkat kebenaran dapat dihitung dengan membagi jumlah kebenaran dalam klasifikasi dengan jumlah data uji. Untuk menghitung tingkat kebenaran dapat menggunakan persamaan

$$\text{accuracy} = \frac{\text{jumla } h_{\text{kebenaran}}}{\text{jumla } h_{\text{data}}}$$

## III. METODOLOGI

Sistem yang dibuat pada penelitian ini merupakan sistem yang digunakan untuk memprediksi suatu resiko kredit dari calon kreditur. *Dataset* yang dipakai adalah data nasabah Bank Jerman yang diambil dari situs *UCI Machine Learning Repository*.

Langkah pertama algoritma Naïve Bayes Classifier terdapat tiga tahap proses, yang pertama adalah *preprocessing*, dalam tahap ini proses yang dilakukan adalah menghitung peluang prior. Tahap ini dihitung nilai peluang kemunculan setiap kelas target resiko kredit, yaitu kelas baik dan kelas buruk. Setelah melakukan penghitungan peluang prior, selanjutnya dilakukan penghitungan penghitungan peluang bersyarat dari setiap variabel-variabel kredit dari calon kreditur. Penghitungan ini menggunakan dua persamaan, dilihat dari jenis data apakah itu termasuk dalam data yang berjenis kategorik atau numerik, apabila data tergolong dalam jenis data kategorik maka nilai peluang dihitung dengan menggunakan persamaan teori Bayesian, sedangkan apabila data tergolong dalam jenis data numerik maka nilai peluang bersyarat dihitung menggunakan persamaan distribusi *Gaussian*. Proses penghitungan nilai peluang bersyarat dari setiap variabel-variabel kredit dari calon kreditur ini termasuk dalam proses *training*. Proses ketiga yaitu proses *testing*, dimana dalam proses *testing* ini dihasilkan sebuah keputusan atau sebuah prediksi status resiko kredit. Tahap *testing* ini dilakukan penghitungan nilai peluang akhir, yaitu perkalian nilai peluang yang didapat dari tahap *preprocessing* dan *training*, dimana semua nilai peluang prior dan bersyarat yang dihasilkan dari setiap kelas resiko kredit dikalikan. Setelah dilakukan proses perkalian, maka dihasilkan dua

nilai peluang akhir dari setiap kelas resiko kredit, dan untuk menentukan hasil presiksi resiko kredit yang dilakukan oleh sistem, maka dipilih nilai yang terbesar dari hasil penghitungan nilai peluang akhir.

#### IV. HASIL PENGUJIAN

Dalam pengujian program algoritma Naïve Bayes Classifier ini digunakan 20 *record data* yang diambil dari data nasabah penghajuan kredit bank Jerman, yang diambil dari <http://mllearn.ics.uci.edu/databases/statlog/german/>, yang masing-masing datanya sudah memiliki kelas resiko kredit (baik dan buruk).

Parameter pengujian algoritma Naïve Bayes Classifier ini adalah tingkat akurasi terhadap jumlah *record data training* yang digunakan.

Pada pengujian ini dilakukan variasi jumlah *record data training* yang terdiri dari 200, 300, 500, 600, 700 *record data training*. Pengujian dilakukan dengan 2 skenario, yaitu menggunakan *data training* dengan komposisi kelas target yang sama dan berbeda. Hasil uji coba algoritma Naïve Bayes Classifier ditunjukkan oleh tabel 4.1

**Tabel 4.1** Perhitungan akurasi hasil uji coba dengan algoritma Naïve Bayes Classifier dengan kelas target berbeda.

Record	Kelas Aktual	Kelas Hasil Uji Coba				
		200	300	500	600	700
1	2	2	2	2	2	2
2	1	2	2	2	2	2
3	2	1	2	2	2	2
4	2	2	2	2	2	2
5	1	2	2	2	2	2
6	2	2	2	2	2	2
7	2	2	2	2	2	2
8	2	2	2	2	2	2
9	1	2	2	2	2	2
10	2	2	2	2	2	2
11	2	1	2	2	2	2
12	2	2	2	2	2	2
13	1	2	2	2	2	2
14	2	2	2	2	2	2

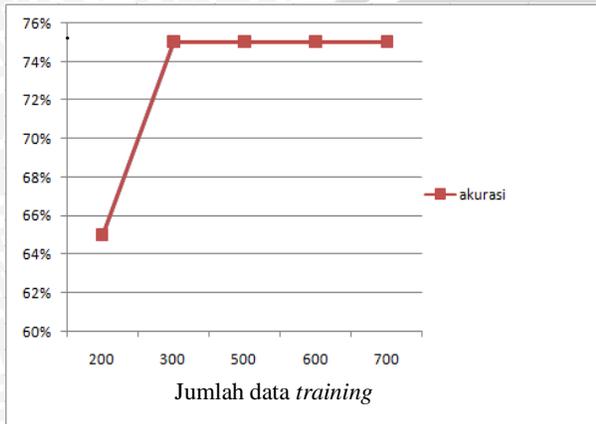
15	2	2	2	2	2	2
16	1	2	2	2	2	2
17	2	2	2	2	2	2
18	2	2	2	2	2	2
19	2	2	2	2	2	2
20	2	2	2	2	2	2
Jml.Kelas Benar		13	15	15	15	15
Akurasi (%)		65%	75%	75%	75%	75%

**Tabel 4.2** Perhitungan akurasi hasil uji coba dengan algoritma Naïve Bayes Classifier dengan kelas target sama.

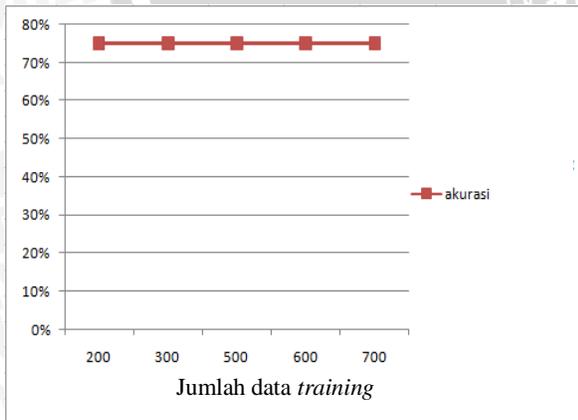
Record	Kelas Aktual	Kelas Hasil Uji Coba				
		200	300	500	600	700
1	2	2	2	2	2	2
2	1	2	2	2	2	2
3	2	2	2	2	2	2
4	2	2	2	2	2	2
5	1	2	2	2	2	2
6	2	2	2	2	2	2
7	2	2	2	2	2	2
8	2	2	2	2	2	2
9	1	2	2	2	2	2
10	2	2	2	2	2	2
11	2	2	2	2	2	2
12	2	2	2	2	2	2
13	1	2	2	2	2	2
14	2	2	2	2	2	2
15	2	2	2	2	2	2
16	1	2	2	2	2	2
17	2	2	2	2	2	2
18	2	2	2	2	2	2
19	2	2	2	2	2	2
20	2	2	2	2	2	2
Jml.Kelas Benar		15	15	15	15	15
Akurasi (%)		75%	75%	75%	75%	75%

Dari data pada Tabel 4.1, terdapat angka 1 dan 2 yang memiliki arti bahwa angka tersebut merupakan kelas yang dihasilkan pada proses pengujian. Angka 1 menunjukkan kelas resiko kredit baik, sedangkan angka 2 menunjukkan kelas resiko kredit buruk. Jumlah kelas benar pada keterangan Tabel 4.1 adalah merupakan

jumlah kelas data pengujian yang bernilai sama dengan kelas data asli. Prosentase akurasi kebenaran pada Tabel 4.1 dihitung dengan menggunakan persamaan perhitungan akurasi. Berdasarkan hasil ujia coba pada Tabel 4.1 dan 4.2 maka pengaruh penggunaan jumlah *record data training* dan penggunaan prosentase kelas target yang sama dan berbeda terhadap tingkat akurasi kebenaran ditunjukkan oleh gambar 4.1 dan 4.2



**Gambar 4.1** Grafik Akurasi Naïve Bayes Classifeir dengan prosentase kelas target berbeda



**Gambar 4.2** Grafik Akurasi Naïve Bayes Classifeir dengan prosentase kelas target sama

Dari Tabel 4.1 dapat dilihat bahwa kenaikan jumlah record data *training* berbanding lurus dengan tingkat akurasi. Pada skenario pengujian 1 dengan jumlah record data *training*

200 mendapatkan nilai akurasi sebesar 65%, untuk jumlah record data *training* 300 mendapatkan nilai akurasi sebesar 75%, jumlah record data *training* 500 mendapatkan nilai akurasi sebesar 75%, jumlah record data *training* 600 mendapatkan nilai akurasi sebesar 75%, jumlah record data *training* 700 mendapatkan nilai akurasi sebesar 75%. Sedangkan pada skenario pengujian kedua yaitu menggunakan jumlah kelas target yang sama, didapat hasil nilai akurasi yang stabil, yaitu dengan jumlah *record data training* 200 sebesar 75%, jumlah *record data training* 300 sebesar 75%, jumlah *record data training* 500 sebesar 66%, jumlah *record data training* 600 sebesar 75%, jumlah *record data training* 700 sebesar 75%. Sehingga dari hasil dua skenario pengujian algoritma Naïve Bayes Classifier pada penentuan status resiko kredit ini didapatkan tingkat akurasi terbesar adalah 75%.

## V. KESIMPULAN

Kesimpulan yang diperoleh dari penelitian ini adalah sebagai berikut :

1. Telah diimplementasikan metode *Naïve Bayes Classifier* untuk prediksi resiko kredit berdasarkan variabel pada data pengujian kredit nasabah bank Jerman dengan melalui beberapa tahap, diantaranya tahap *preprocessing* meliputi proses penentuan nilai peluang kemunculan kelas target , *training* meliputi penentuan nilai peluang bersyarat dari setiap variabel persyaratan kredit nasabah, sedangkan tahap *testing* menentukan nilai peluang total dari masing-masing kelas target, serta tahap penentuan hasil prediksi beserta tingkat akurasi maksimum yang didapatkan untuk membantu menentukan persetujuan pengajuan kredit dari nasabah.
2. Pada percobaan yang telah dilakukan dengan menggunakan 2 skenario pengujian prediksi status resiko kredit, didapat nilai akurasi tertinggi sebesar 75% pada data *training* sebanyak yang menggunakan prosentase kelas target yang

seimbang dan data *testing* sebanyak 20 *record*. Hal ini menunjukkan bahwa sistem sudah berjalan dengan cukup baik.

## VI. DAFTAR PUSTAKA

- [1] Gay, L.R. 1976. *Educational Research : Competencies for Analysis and Application*. Ohio: Bell & Howell Company.
- [2] Giri, Yudho, S., 2003. *Data Mining – Menggali Informasi yang Terpendam*. <http://ikc.cbn.net.id/populer/yudho/yudho-datamining.zip>., diakses tanggal 12 September 2011
- [3] Giudichi, P.2003. *Applied Data Mining Statistical Method for Business and Industry*. John Wiley & Sons. Chicester. Hal 130-132.
- [4] Hand, David, Heikki Mannila dan Padharic Smyth.2001. *Principles of Data Mining*. Massachusetts:MIT Press Cambridge.
- [5] Han, Jiawei dan Michele Kamber, 2000, *Data Mining : Concepts and Techniques*. Morgan Kaufmann Publishers.
- [6] Han, Eui-Hong. Dan G.Karypis. 2005. *Feature Based Recommendation System*. Internet : <http://glaros.dtc.umn.edu/gkhome/fetch/papers/fbrsCIKM05.pdf>..Tanggal akses 12 September 2011
- [7] Haruechaiyasak, C., Meu-Ling Shyu, dan Shu-Ching Chen. 2004. *A Data Mining Framework for Building AWeb-Page Recommender System*. Internet: <http://www.eng.miami.edu/~shyu/paper/2004/iri04-rec.pdf>.. Tanggal akses 12 September 2011
- [8] Kantardzic, Mehmed. 2003. *Data Mining : Concepts, Models, Methods, and Algoritms*. US: A JOHN WILLEY & SONS, Inc.
- [9] Khusnawi. 2007. *Pengantar Solusi Data Mining*. STMIK AMIKOM. Yogyakarta. <http://p3m.amikom.ac.id/p3m/56%20-%20PENGANTAR%20SOLUSI%20DATA%20MINING.pdf>., tanggal akses: 12 September 2011
- [10] Kristijadi, Emanuel. 2003. *Sistem Informasi Credit Scoring Berbasis Database Untuk Retail Banking*. STIE Perbanas. Surabaya. <http://ekristijadi.files.wordpress.com/2007/10/sniktiviireff0737-emanuel-kristijadi.pdf>., tanggal 12 September 2011
- [11] McCallum, Anrew dan K. Nigam.2002. *A Comparison of Event Models for Naïve Bayes Text Classification*.Internet : <http://www.kamalnigam.com/papers/multinomial-aaaiws98.pdf>., tanggal akses 12 September 2011
- [12] Panda,Mrutunjaya. 2007. *Network Intrusion Detection Using Naïve Bayes*. Behampur,India : Departement Of Computer Science
- [13] Rachli, Muhammad. 2007. *Email Filtering Menggunakan Naïve Bayeian*. Tugas Akhir Departemen Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Bandung. Bandung.
- [14] Restu. 2010. *Prediksi Resiko Kredit Menggunakan Algoritma CART dan CMAR*. Skripsi Program Studi Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Brawijaya. Malang.
- [15] Rish, Iriana. 2001. *IBM Research Report : An empiricial study of the naïve bayes classifier*. Internet: <http://www.research.ibm.com/people/r/rish/papers/RC22230.pdf>.. Tanggal akses: 12 September 2011

[16] Suharyadi.2009.*Statistika Untuk Ekonomi dan Keuangan Modern*.Jakarta:Salemba Empat

[17] Sutikno, T., Pujianta, A., Tri Y.S. 2007. *Prediksi Resiko Kredit Dengan Jaringan Syaraf Tiruan Backpropagation* Universitas Ahmad Dahlan. Yogyakarta. <http://journal.uii.ac.id/index.php/Snati/article/viewFile/1679/1461>., diakses tanggal: 12 September 2011

[18] Suyatno.2003. *Dasar-dasar Kredit*.Internet : <http://www.scribd.com/doc/13857422/Dasar-Dasar-Kredit>. Diakses tanggal 12 September 2011

[19] Tan, Pang-Ning, Michael Steinbach dan Vipin Kumar. 2004. *Introducing to data Mining*. New York

