
Implementasi Metode Naive Bayes Untuk Klasifikasi Kredit Motor

Implementation of Naive Bayes Method For Classification of Motorcycle Credit

Heryono ^{*1}, Acun Kardianawati²

^{1,2} Program Studi Sistem Informasi Universitas Dian Nuswantoro;
Jl. Imam Bonjol 207, Semarang. telp. (+6224)3517261
e-mail: ^{*1}glouse_21@yahoo.co.id, ²acunkardiana@gmail.com

Abstrak

Kredit dalam hal ini telah menjadi sumber penghasilan bagi beberapa bank ataupun perusahaan dan instansi swasta yang menyewakan jasa kreditnya kepada konsumen, meskipun begitu dapat dilihat terdapat beberapa konsumen yang memiliki kredit macet seperti tunggakan dan hal lainnya yang mengakibatkan tidak dapat melanjutkan pembayaran kredit. OTO Kredit Motor adalah salah satu perusahaan yang bergerak dibidang perkreditan kendaraan beroda dua atau motor. Dalam hal ini OTO Kredit Motor masih sangat kesulitan di dalam menentukan calon debitur yang nantinya akan layak untuk mendapatkan kredit tersebut. Oleh karena itu data mining digunakan untuk dapat memprediksi resiko terjadinya kredit macet dengan melakukan pengklasifikasian terhadap calon debitur kendaraan nantinya. Tujuan dari penelitian ini, adalah memprediksi terhadap kelayakan kredit macet menggunakan metode Naïve Bayes. Hasil dari penelitian digunakan untuk memprediksi kelayakan kredit untuk menghindari terjadinya kredit macet, dan mengevaluasi performance model Naïve Bayes.

Kata Kunci: *DataMining, Algoritma Naive Bayes, Klasifikasi, Debitur.*

Abstract

In this case the credit has become a source of income for some banks or companies also private institutions for lease their credit services to consumers, although it can be seen there are some consumers who have bad loans such as arrears and other things that can be effecting failure in resume credit payments. OTO Kredit Motor is one of the companies were engaged in the loan of motorcycles. In this case OTO Kredit Motor is still very difficult in determining the prospective borrowers who will be eligible to get the credit. Therefore, data mining is used to predict the risk of non-performing loans by classifying the prospective borrower of the vehicle. The purpose of this study, is to predict the viability of bad loans using the Naïve Bayes method. The results of the study were used to predict creditworthiness to avoid bad credit, and to evaluate the performance of the Naïve Bayes model.

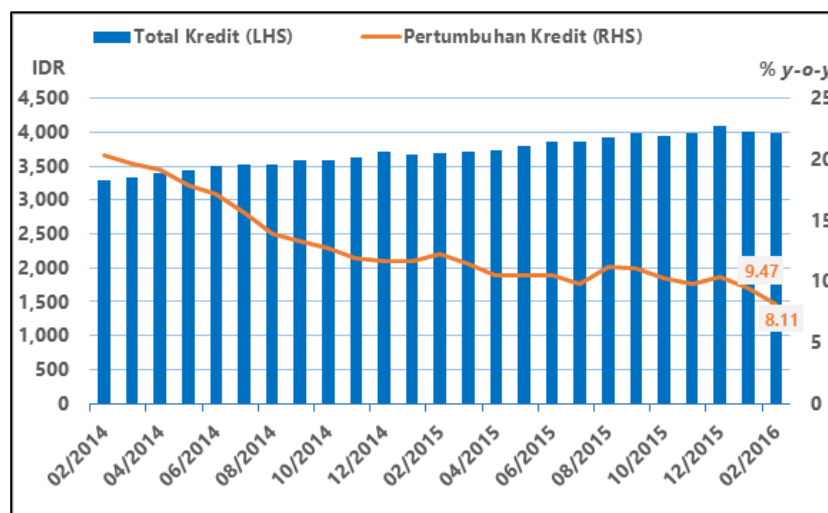
Keyword : *Data Mining, Naive Bayes Algorithm, Classification, Debtor.*

1. PENDAHULUAN

Di dalam memberikan pelayanan, perusahaan peminjaman ataupun kredit tentunya memiliki syarat yang telah ditentukan sebagaimana calon debitur yang akan menerima peminjaman kredit tersebut. Perusahaan wajib menentukan layak atau tidaknya calon debitur menerima kredit. Pada perusahaan peminjaman kredit atau perbankan khususnya, dimana pemberian peminjaman kredit kepada debitur merupakan sumber penghasilan utama yang dimiliki oleh perusahaan. Selain itu resiko untuk terjadinya kemungkinan kredit macet sangatlah riskan dan memungkinkan terjadinya penurunan pendapatan perusahaan yang meminjamkan kreditnya.

Di dalam menghadapi persaingan, perusahaan dituntut untuk dapat memprediksi kemungkinan adanya kredit macet yang terjadi atau selektif di dalam melakukan pemilihan calon debitur. Salah satunya metode data mining yang telah menjadikan perusahaan pada saat ini lebih menggunakan atau memanfaatkan sebagai salah satu media di dalam melakukan prediksi bisnis.

Dalam proses pembiayaan kredit motor yang dilakukan oleh OTO Kredit Motor masih belum cukup maksimal di dalam melakukan pemilihan dan menentukan calon debitur. Hal ini dikarenakan penganalisis perusahaan terhadap kredit tidak begitu cermat dan dinilai tidak hati-hati terhadap konsumen yang memiliki karakter yang tidak baik.



Gambar 1. Pertumbuhan Kredit tiap dua bulan.

Hal ini dapat dilihat dari pertumbuhan kredit yang menurun salah satu dari dampak pendapatan operasional yang ikut menurun dikarenakan kredit macet selain itu data ketika melakukan survey terhadap lapangan tidak sesuai dengan data yang diberikan calon debitur, sehingga data menjadi tidak valid dan tidak dapat dijadikan tolak ukur. Di dalam melakukan analisis diperlukan beberapa aspek yang nantinya dapat dijadikan sebuah pertimbangan bagi penganalisis di dalam menentukan layak atau tidaknya diberikan. Ketika aspek aspek yang ditentukan telah dipenuhi dan data yang di cocokan dengan data di lapangan valid maka nantinya data tersebut dapat diolah dan dijadikan tolak ukur di dalam menentukan konsumen atau calon debitur kedepan.

Tujuan yang ingin dicapai adalah dapat melakukan implementasi menggunakan Metode Naïve Bayes dalam klasifikasi kelayakan calon kredit motor pada OTO Kredit Motor. Untuk memberikan gambaran dengan lebih terfokus, data yang digunakan meliputi: data debitur atau konsumen, data kriteria atau atribut, proses pembuatan laporan mengenai informasi data debitur.

Metode yang digunakan adalah metode Naïve Bayes yang mempunyai tahapan-tahapan antara lain: proses pembersihan data (Data Cleaning), integrasi data (Data Integration), seleksi data (Data Selection), transformasi data (Data Transformation), proses mining, evaluasi pola (Patterin Evaluation), presentasi pengetahuan (Knowledge Presentation).

Fungsi kredit bagi masyarakat untuk [1]:

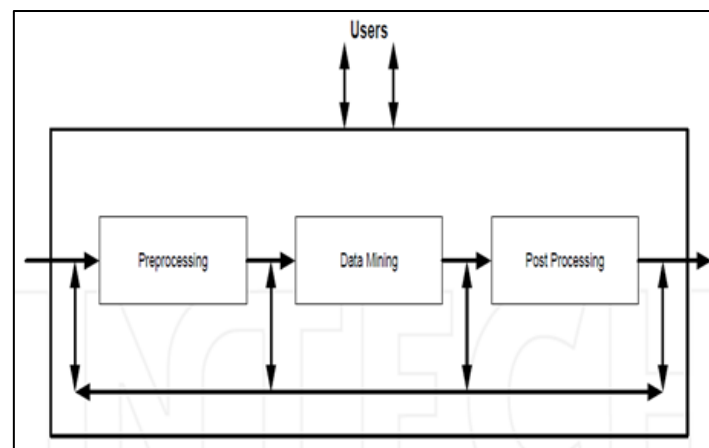
1. Sebagai motivator dan dinamisator dalam peningkatan kegiatan perdagangan dan perekonomian.
2. Menambah lapangan kerja bagi masyarakat.
3. Memperlancar arus barang dan arus uang.
4. Meningkatkan hubungan internasional.

5. Memaksimalkan produktivitas dana yang ada.
6. Meningkatkan daya guna barang.
7. Meningkatkan semangat berusaha bagi masyarakat.
8. Memperbesar modal kerja perusahaan.
9. Meningkatkan IPC (income per capita) masyarakat.
10. Mengubah pola berpikir atau bertindak masyarakat untuk lebih ekonomis.

Pada dasarnya terdapat dua fungsi yang saling berkaitan dengan kredit, yaitu *profitability* dan *safety*". *Profitability* yaitu, tujuan untuk memperoleh hasil dari kredit berupa keuntungan dari bunga yang harus dibayar nasabah. Sedangkan *safety* merupakan keamanan dari prestasi atau fasilitas yang diberikan harus benar-benar terjamin sehingga tujuan *profitability* dapat tercapai tanpa hambatan yang berarti. [2].

Klasifikasi merupakan bagian dari algoritma data mining, klasifikasi ini adalah algoritma yang menggunakan data dengan target (class/label) yang berupa nilai kategorikal/nominal. Proses klasifikasi didasarkan pada empat komponen mendasar, yaitu Kelas (Class), Prediktor (Predictor), Pelatihan dataset (Training dataset), Dataset Pengujian (Testing Dataset).[3]

Knowledge Discovery in Database (KDD) merupakan sebuah proses dengan beberapa tingkatan, tidak sepele, interaktif dan berulang untuk identifikasi pola yang dipahami, sah, baru dan secara potensial berguna mulai dari sekumpulan data yang sangat besar [4]. KDD dikarakteristikan sebagai sebuah proses yang terdiri dari beberapa tahap operasional : Preprocessing, Data Mining dan Post Processing



Gambar 2. Tahap *Knowledge Discovery in Database* (KDD)

Data Mining merupakan proses pengekstraksian informasi dari sekumpulan data yang sangat besar melalui penggunaan algoritma dan teknik penarikan dalam bidang statistik, pembelajaran mesin dan sistem manajemen basis data [5].

Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. Definisi lain mengatakan Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya [3]. Naive Bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan [6].

2. METODE PENELITIAN

2.1 Metode Pengumpulan Data

Untuk mendapatkan data-data yang diperlukan, menggunakan cara: Pertama dengan Studi pustaka, yaitu suatu cara untuk memperoleh data melalui bahan-bahan yang telah tertulis dalam jurnal, skripsi, buku, serta catatan-catatan baik tercetak maupun elektronik yang berkaitan dengan permasalahan yang diteliti sebagai landasan teorinya. Kedua dengan Survei yang dilakukan pada OTO Kredit Motor, dengan cara melihat dan datang langsung ke OTO Kredit Motor untuk mendapatkan data-data yang diperlukan. Diantaranya prosedur perkreditan yang berjalan, penyusunan laporan kredit yang terjadi. Ketiga dengan wawancara yang dilakukan dengan mengajukan beberapa pertanyaan yang berkaitan dengan pengumpulan data terhadap narasumber dari pihak OTOKredit Motor.

2.2 Data Mining Menggunakan Metode Naïve Bayes

Metode yang digunakan untuk data mining adalah metode Naive Bayes. Di dalam penerapan data mining menggunakan metode Naive Bayes mempunyai tahapan-tahapan sebagai berikut:

1. Pembersihan data (Data Cleaning) : pembersihan data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan.
2. Integrasi data (Data Integration) : integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru.
3. Seleksi data (Data Selection) : data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database.
4. Transformasi data (Data Transformation) : data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam Data Mining.
5. Proses Mining : merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data. Beberapa metode yang dapat digunakan berdasarkan pengelompokan Data Mining.
6. Evaluasi pola (Pattern Evaluation) : untuk mengidentifikasi pola-pola menarik ke dalam knowledge based yang ditemukan.
7. Presentasi pengetahuan (Knowledge Presentation).

3. HASIL DAN PEMBAHASAN

3.1 Pemahaman Bisnis (*Business Understanding*)

Pemahaman bisnis (*business understanding*), tahap pertama dalam proses CRISP-DM yang juga dapat disebut sebagai tahap pemahaman bisnis (penelitian).

1. Menentukan Tujuan Bisnis

Di dalam melakukan penentuan bisnis terdapat beberapa data yang perlu diolah seperti data kreditur atau data konsumen dari bulan januari hingga maret 2017 tujuannya adalah untuk menentukan kreditur yang baik untuk perusahaan serta meningkatkan pendapatan perusahaan di dalam melakukan selektif pemilihan kreditur.

2. Melakukan Penilaian Situasi

Pada tahapan ini dengan melakukan pengklasifikasian terhadap kelayakan calon debitur, hal ini sangat berkaitan dengan proses bisnis yang terjadi pada Oto Kredit Motor.

3. Menentukan Tujuan Data Mining

Tujuan pada penelitian yang dibuat penulis adalah menemukan pengetahuan baru terhadap pola klasifikasi di dalam penentuan kelayakan penerimaan calon kredit motor pada Oto Kredit Motor.

3.2 Pemahaman Data (Data Understanding)

Data subset yang digunakan berupa dokumen excel sejumlah 115 record dengan data yang terbagi pada bulan Januari, Februari dan Maret.

1. Pengumpulan data awal

Dataset debitur Oto Kredit Motor dimana data yang diperoleh diambil dari bulan Januari, Februari dan Maret berdasarkan format dokumen excel atau .xlsx. Selain itu kegiatan yang dilakukan berupa wawancara, observasi serta melakukan studi pustaka.

2. Mendeskripsikan Data

Dataset dari debitur memiliki beberapa atribut diantaranya Jenis Kelamin, Usia, Status, Pekerjaan, Penghasilan, Masa pembayaran.

3. Evaluasi kualitas data

Pada dataset penelitian yang dilakukan tidak ditemukan data yang kosong ataupun null.

4. Pemilihan Atribut

Atribut yang digunakan adalah Jenis Kelamin, Usia, Status, Pekerjaan, Penghasilan, Masa pembayaran.

3.3 Persiapan Data

Persiapan data mencakup semua kegiatan untuk membangun dataset calon debitur yang akan diterapkan ke dalam alat pemodelan, dari data mentah awal berupa dataset calon debitur dan selanjutnya akan melakukan proses data mining.

1. Seleksi Data

Atribut yang digunakan adalah Jenis Kelamin dari calon debitur, Usia debitur, Status pernikahan, Pekerjaan yang dimiliki saat ini, Penghasilan per bulan dan Masa pembayaran atau rentang waktu pembayaran.

2. Pengolahan data mentah (Preprocessing Data)

Pada tahap ini merupakan tahap untuk memastikan data calon debitur yang dipilih telah layak untuk dilakukan proses pengolahan

3. Transformasi data

Data yang berjenis numerikal seperti masa pembayaran harus dilakukan proses inialisasi data terlebih dahulu ke dalam bentuk nominal. Untuk melakukan inialisasi masa pembayaran dapat dilakukan dengan:

- a. Calon debitur dengan masa pinjam atau pembayaran 15 bulan dan 17 bulan dapat dikategorikan < 21 bulan.
- b. Calon debitur dengan masa pinjam atau pembayaran 21 bulan dan 27 bulan dapat dikategorikan 21-27 bulan.

- c. Calon debitur dengan masa pinjam atau pembayaran 33 bulan dapat dikategorikan > 27 bulan.

3.4 Pemodelan

Pemodelan adalah fase yang secara langsung melibatkan teknik data mining yaitu dengan melakukan pemilihan teknik data mining dan menentukan algoritma yang akan digunakan.

1. Pemilihan teknik pemodelan

Naive Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu.

2. Penghitungan Data Mining

a. Baca Data Training

Adapun data latih yang digunakan dapat dilihat pada gambar berikut :

No	CIF	Nama	Jenis Kelamin	Usia	Status	Pendidikan	Gaji	Nilai_jaminan	Tempo	Jumlah Pinjam	Tujuan_pinjam	Angsuran	Margin
1	102003920	ABDURRAHMAN NIAM	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	PNS UMUM	2,200,000.00	33	10,000,000.00	Barang Konsumtif	303030.30	75757.58
2	102003921	DEVI HARYANTO	Perempuan	30 - 40 Tahun	Kawin	D3	KARYAWAN SWASTA	2,200,000.00	33	8,316,217.00	Barang Konsumtif	252006.58	63001.64
3	102003922	DIANI TRI RAVITA PUTRANTI	Perempuan	20 - 29 Tahun	Belum Kawin	SARJANA	WIRASWASTA	2,750,000.00	27	15,000,000.00	Barang Konsumtif	555555.56	138888.89
4	102003923	KHAIUR RAZIQIN	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	3,250,000.00	21	20,000,000.00	Barang Konsumtif	952380.95	238095.24
5	102003924	HARYONO	Laki-Laki	30 - 40 Tahun	Kawin	SMA	KARYAWAN SWASTA	2,500,000.00	27	7,000,000.00	Barang Konsumtif	259259.26	64814.81
6	102003925	WIDIANI FITRI	Perempuan	30 - 40 Tahun	Belum Kawin	SARJANA	WIRASWASTA	2,200,000.00	27	12,000,000.00	Barang Konsumtif	444444.44	111111.11
7	102003926	FEBIAN MARTIN	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	3,750,000.00	15	13,000,000.00	Barang Konsumtif	866666.67	216666.67
8	102003927	SUPANGI PARDIONO	Laki-Laki	30 - 40 Tahun	Kawin	D1	KARYAWAN SWASTA	3,250,000.00	21	16,000,000.00	Barang Konsumtif	761904.76	190476.19
9	102003928	ARDAN NUROKHMAT	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	2,500,000.00	27	5,000,000.00	Barang Konsumtif	185185.19	46296.30
10	102003929	QORYN NOURMALA SARI	Perempuan	20 - 29 Tahun	Belum Kawin	MAGISTER	KARYAWAN SWASTA	3,250,000.00	21	10,500,000.00	Barang Konsumtif	500000.00	125000.00
11	102003930	DEVINA EFRILLA WARMAN	Perempuan	30 - 40 Tahun	Kawin	SARJANA	WIRASWASTA	2,200,000.00	33	11,000,000.00	Barang Konsumtif	333333.33	83333.33
12	102003931	DESI RISMAYA	Perempuan	30 - 40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	2,200,000.00	15	7,000,000.00	Barang Konsumtif	466666.67	116666.67
13	102003932	IMAM SAFARUDIN	Laki-Laki	>40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	2,750,000.00	21	15,000,000.00	Barang Konsumtif	714285.71	178571.43
14	102003933	RENTI HARDIANA	Perempuan	30 - 40 Tahun	Kawin	SARJANA	WIRASWASTA	3,250,000.00	21	10,000,000.00	Barang Konsumtif	476190.48	119047.62
15	102003934	TINA KARLINA	Perempuan	30 - 40 Tahun	Kawin	D3	KARYAWAN SWASTA	2,500,000.00	27	6,000,000.00	Barang Konsumtif	222222.22	55555.56
16	102003935	DIAN ROMANSYAH	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	PNS UMUM	2,200,000.00	23	10,000,000.00	Barang Konsumtif	434782.61	108695.65
17	102003936	ABDUL SYAHRONI	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	PNS UMUM	3,750,000.00	29	12,000,000.00	Barang Konsumtif	413793.10	103448.28
18	102003937	WIEN AYU LIANI	Perempuan	>40 Tahun	Kawin	SMA	KARYAWAN SWASTA	3,250,000.00	17	7,000,000.00	Barang Konsumtif	411764.71	102941.18
19	102003938	NISA DEWI	Perempuan	30 - 40 Tahun	Kawin	SMP	KARYAWAN SWASTA	2,500,000.00	23	8,500,000.00	Barang Konsumtif	369565.22	92391.30
20	102003939	INDRA AFRIANA	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	WIRASWASTA	3,250,000.00	29	13,000,000.00	Barang Konsumtif	448275.86	112068.97
21	102003940	AGUNG SUPRIYONO	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	2,200,000.00	23	12,000,000.00	Barang Konsumtif	521739.13	130434.78
22	102003941	WAWAN BASUKI	Laki-Laki	>40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	2,200,000.00	35	12,000,000.00	Barang Konsumtif	342857.14	85714.29
23	102003942	HENDAR ARIAWAN	Laki-Laki	30 - 40 Tahun	Kawin	SMA	KARYAWAN SWASTA	2,750,000.00	29	8,000,000.00	Barang Konsumtif	275862.07	68965.52
24	102003943	AHMAD SUBANDI	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	WIRASWASTA	3,250,000.00	21	8,560,000.00	Barang Konsumtif	407619.05	101904.76
25	102003944	YUJIAR ADI	Laki-Laki	20 - 29 Tahun	Belum Kawin	SARJANA	KARYAWAN SWASTA	2,500,000.00	27	17,000,000.00	Barang Konsumtif	629629.63	157407.41
26	102003945	YULIUS KRISTI	Laki-Laki	30 - 40 Tahun	Kawin	MAGISTER	PNS UMUM	3,250,000.00	21	15,500,000.00	Barang Konsumtif	738095.24	184523.81
27	102003946	SULISTYAWAN SETIAWAN	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	2,200,000.00	15	9,800,000.00	Barang Konsumtif	653333.33	163333.33
28	102003947	ZULHAM FATURRAHMAN	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	WIRASWASTA	2,200,000.00	27	15,000,000.00	Barang Konsumtif	555555.56	138888.89
29	102003948	YUSIANA DWI	Perempuan	30 - 40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	3,750,000.00	21	12,000,000.00	Barang Konsumtif	571428.57	142857.14
30	102003949	AGUS PRASTOMO	Laki-Laki	30 - 40 Tahun	Kawin	D3	WIRASWASTA	3,250,000.00	27	8,000,000.00	Barang Konsumtif	296296.30	74074.67
31	102003950	FERRY LEVEUNO	Laki-Laki	20 - 29 Tahun	Belum Kawin	SARJANA	KARYAWAN SWASTA	3,051,500.00	21	10,000,000.00	Barang Konsumtif	476190.48	119047.62
32	102003951	RETNO PAMUNGKAS	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	2,200,000.00	33	15,000,000.00	Barang Konsumtif	454545.45	113636.36
33	102003952	ABU SHAFAN	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	WIRASWASTA	2,200,000.00	21	14,500,000.00	Barang Konsumtif	690476.19	172619.05
34	102003953	RIYANTO SOFYAN	Laki-Laki	20 - 29 Tahun	Belum Kawin	SARJANA	KARYAWAN SWASTA	3,750,000.00	21	12,000,000.00	Barang Konsumtif	571428.57	142857.14
35	102003954	KWOK JESUSANTO	Laki-Laki	30 - 40 Tahun	Kawin	SMP	KARYAWAN SWASTA	3,250,000.00	21	6,500,000.00	Barang Konsumtif	309523.81	77380.95
36	102003955	HANAFI YUSUF	Laki-Laki	30 - 40 Tahun	Kawin	SARJANA	KARYAWAN SWASTA	3,750,000.00	17	8,000,000.00	Barang Konsumtif	470588.24	117647.06

Gambar 3. Data training pada bulan Januari

b. Kriteria dan Probabilitas

Probabilitas Kriteria Jenis Kelamin berdasarkan dengan data debitur pada gambar3 didapat sejumlah data (data training) yaitu dengan total 36 data, dari data tersebut terdapat 1 data yang berjenis kelamin Laki – laki dengan data klasifikasi tidak lancar, 6 data yang berjenis kelamin Laki – laki dengan data klasifikasi kurang lancar dan 18 data yang berjenis kelamin Laki – laki dengan data klasifikasi lancar. Untuk berjenis kelamin Perempuan terdapat 0 data dengan data klasifikasi tidak lancar, 1 data yang berjenis

kelamin Perempuan dengan data klasifikasi kurang lancar dan 10 data yang berjenis kelamin Perempuan dengan data klasifikasi lancar.

Tabel 1. Kriteria Probabilitas Jenis Kelamin pada bulan Januari

Jenis Kelamin	Jumlah data klasifikasi			Probabilitas		
	Tidak lancar	Kurang lancar	Lancar	Tidak lancar	Kurang lancar	Lancar
Laki-laki	1	6	18	1	0.85	0.64
Perempuan	0	1	10	0	0.15	0.36
Jumlah	1	7	28	0.03	0.19	0.78

Probabilitas Kriteria Usi berdasarkan gambar 3 dimana yang berusia 20 – 29 Tahun terdapat 1 data dengan data klasifikasi tidak lancar, 0 data yang berusia 20 – 29 Tahun dengan data klasifikasi kurang lancar dan 4 data yang berusia 20 – 29 Tahun dengan data klasifikasi lancar. Untuk yang berusia 30 -40 Tahun terdapat 0 data dengan data klasifikasi tidak lancar, 7 data yang berusia 30 - 40 Tahun dengan data klasifikasi kurang lancar dan 21 data yang berusia 30 - 40 Tahun dengan data klasifikasi lancar. Untuk yang berusia >40 Tahun terdapat 0 data dengan data klasifikasi tidak lancar, 0 data yang berusia >40 Tahun dengan data klasifikasi kurang lancar dan 3 data yang berusia >40 Tahun dengan data klasifikasi lancar.

Tabel 2. Kriteria Probabilitas Usia pada bulan Januari

Usia	Jumlah data klasifikasi			Probabilitas		
	Tidak lancar	Kurang lancar	Lancar	Tidak lancar	Kurang lancar	Lancar
20-29 tahun	1	0	4	1	0	0.14
30-40 tahun	0	7	21	0	1	0.75
>40 tahun	0	0	3	0	0	0.1
Jumlah	1	7	28	0.03	0.19	0.78

Probabilitas Kriteria Status berdasarkan gambar 3 dimana yang berstatus kawin dengan 0 data dengan data klasifikasi tidak lancar, 7 data yang berstatus Kawin dengan data klasifikasi kurang lancar dan 23 data yang berstatus Kawin dengan data klasifikasi lancar. Untuk yang berstatus Belum Kawin terdapat 1 data dengan data klasifikasi tidak lancar, 0 data yang berstatus Belum Kawin dengan data klasifikasi kurang lancar dan 5 data yang berstatus Belum Kawin dengan data klasifikasi lancar.

Tabel 3. Kriteria Probabilitas Status pada bulan Januari

Status	Jumlah data klasifikasi			Probabilitas		
	Tidak lancar	Kurang lancar	Lancar	Tidak lancar	Kurang lancar	Lancar

Belum kawin	0	7	23	0	1	0.82
Kawin	1	0	5	1	0	0.18
Jumlah	1	7	28	0.28	0.19	0.78

Probabilitas Kriteria Pekerjaan berdasarkan gambar 3 dimana yang memiliki pekerjaan PNS 0 data dengan data klasifikasi tidak lancar, 1 data yang memiliki pekerjaan PNS dengan data klasifikasi kurang lancar dan 3 data yang memiliki pekerjaan PNS dengan data klasifikasi lancar. Untuk kriteria pekerjaan karyawan swasta terdapat 1 data dengan data klasifikasi tidak lancar, 5 data yang memiliki pekerjaan karyawan swasta dengan data klasifikasi kurang lancar dan 17 data yang memiliki pekerjaan karyawan swasta dengan data klasifikasi lancar. Untuk Kriteria Pekerjaan Wiraswasta yang memiliki pekerjaan wiraswasta terdapat 0 data dengan data klasifikasi tidak lancar, 1 data yang memiliki pekerjaan wiraswasta dengan data klasifikasi kurang lancar dan 8 data yang memiliki pekerjaan wiraswasta dengan data klasifikasi lancar.

Tabel 4. Kriteria Probabilitas Pekerjaan pada bulan Januari

Pekerjaan	Jumlah data klasifikasi			Probabilitas		
	Tidak lancar	Kurang lancar	Lancar	Tidak lancar	Kurang lancar	Lancar
PNS	0	1	3	0.00	0.14	0.10
Karyawan Swasta	1	5	17	1	0.71	0.60
Wiraswasta	0	1	8	0.00	0.14	0.28
Jumlah	1	7	28	0.3	0.19	0.78

Probabilitas Kriteria Penghasilan berdasarkan gambar 3 dimana yang memiliki penghasilan <10 terdapat 1 data dengan data klasifikasi tidak lancar, 7 data yang memiliki penghasilan <10 juta dengan data klasifikasi kurang lancar dan 23 data yang memiliki penghasilan <10 juta dengan data klasifikasi lancar. Untuk yang memiliki penghasilan 10 juta terdapat 0 data yang memiliki penghasilan 10 juta dengan data klasifikasi tidak lancar, 0 data yang memiliki penghasilan 10 juta dengan data klasifikasi kurang lancar dan 3 data yang memiliki penghasilan 10 juta dengan data klasifikasi lancar. Untuk kriteria penghasilan >10 juta terdapat 0 data yang memiliki penghasilan >10 juta dengan data klasifikasi tidak lancar, 0 data yang memiliki penghasilan >10 juta dengan data klasifikasi kurang lancar dan 2 data yang memiliki penghasilan 10 juta dengan data klasifikasi lancar.

Tabel 5. Kriteria Probabilitas Penghasilan pada bulan Januari

Penghasilan	Jumlah data klasifikasi			Probabilitas		
	Tidak lancar	Kurang lancar	Lancar	Tidak lancar	Kurang lancar	Lancar
< 10 juta	1	7	23	1	1	0.82
10 juta	0	0	3	0.00	0.00	0.10
> 10 juta	0	0	2	0.00	0.00	0.07
Jumlah	1	7	28	0.02	0.19	0.78

Probabilitas Kriteria Masa Angsuran berdasarkan gambar 3 dimana yang memiliki masa angsuran <21 bulan terdapat 0 data dengan data klasifikasi tidak lancar, 1 data yang memiliki masa angsuran <21 bulan dengan data klasifikasi kurang lancar dan 4 data yang memiliki masa angsuran <21 bulan dengan data klasifikasi lancar. Untuk yang memiliki masa angsuran 21-27 terdapat 1 data yang memiliki masa angsuran 21-27 bulan dengan data klasifikasi tidak lancar, 4 data yang memiliki masa angsuran 21-27 bulan dengan data klasifikasi kurang lancar dan 21 data yang memiliki masa angsuran 21-27 bulan dengan data klasifikasi lancar. Kemudian untuk kriteria masa angsuran >27 bulan terdapat 0 data dengan data klasifikasi tidak lancar, 2 data yang memiliki masa angsuran >27 bulan dengan data klasifikasi kurang lancar dan 3 data yang memiliki masa angsuran >27 bulan dengan data klasifikasi lancar.

Tabel 6. Kriteria Probabilitas Masa Angsuran pada bulan Januari

Masa Angsuran	Jumlah data klasifikasi			Probabilitas		
	Tidak lancar	Kurang lancar	Lancar	Tidak lancar	Kurang lancar	Lancar
< 21	0	1	4	0	0.14	0.14
21-27	1	4	21	1	0.57	0.75
>27	0	2	3	0	0.28	0.10
Jumlah	1	7	28	0.02	0.19	0.78

c. Menghitung Probabilitas Posterior

Probabilitas Prior untuk menentukan class terhadap temuan kasus baru, dengan cara terlebih dahulu menghitung Probabilitas Posteriornya, hal tersebut dilakukan apabila ditemukan kasus baru dalam pengolahan data. Berikut tabel probabilitas posterior untuk menghitung kasus baru yang ditemukan:

Tabel 7. Penghitungan Probabilitas Posterior pada bulan Januari

Data X		P(X Ci)		
Atribut	Nilai (Value)	Tidak lancar	Kurang lancar	Lancar
Jenis Kelamin	Laki-laki	0.10	0.85	0.64
Usia	30 - 40 Tahun	0.00	1	0.75
Status	Kawin	0.00	1	0.82
Pekerjaan	Karyawan Swasta	1	0.71	0.60
Penghasilan	<10 Juta	1	1	0.82
Masa Angsuran	<21	0.00	0.14	0.14

Pada Bulan Januari

$$\begin{aligned}
 P(X| \text{Kredit} = \text{Lancar}) &= P(\text{Jenis Kelamin} = \text{Laki-laki} | \text{Kredit} = \text{Lancar}) * P(\text{Usia} = 30 - 40 \\
 \text{Tahun} | \text{Kredit} = \text{Lancar}) * P(\text{Status} = \text{Kawin} | \text{Kredit} = \text{Lancar}) * P(\text{Pekerjaan} = \\
 \text{Karyawan Swasta} | \text{Kredit} = \text{Lancar}) * P(\text{Penghasilan} < 10 \text{ Juta} | \text{Kredit} = \text{Lancar}) * P(\text{Masa} \\
 \text{Angsuran} = < 21 | \text{Kredit} = \text{Lancar}) \\
 &= 0.64 * 0.75 * 0.82 * 0.60 * 0.82 * 0.14 \\
 &= \mathbf{0.02}
 \end{aligned}$$

$$\begin{aligned}
 P(X| \text{Kredit} = \text{Kurang Lancar}) &= P(\text{Jenis Kelamin} = \text{Laki-laki} | \text{Kredit} = \text{Kurang Lancar}) * \\
 P(\text{Usia} = 30 - 40 \text{ Tahun} | \text{Kredit} = \text{Kurang Lancar}) * P(\text{Status} = \text{Kawin} | \text{Kredit} = \text{Kurang} \\
 \text{Lancar}) * P(\text{Pekerjaan} = \text{Karyawan Swasta} | \text{Kredit} = \text{Kurang Lancar}) * P(\text{Penghasilan} \\
 < 10 \text{ Juta} | \text{Kredit} = \text{Kurang Lancar}) * P(\text{Masa Angsuran} = < 21 | \text{Kredit} = \text{Kurang Lancar})
 \end{aligned}$$

$$= 0.85 * 1 * 1 * 0.71 * 1 * 0.14$$

$$= \mathbf{0.08}$$

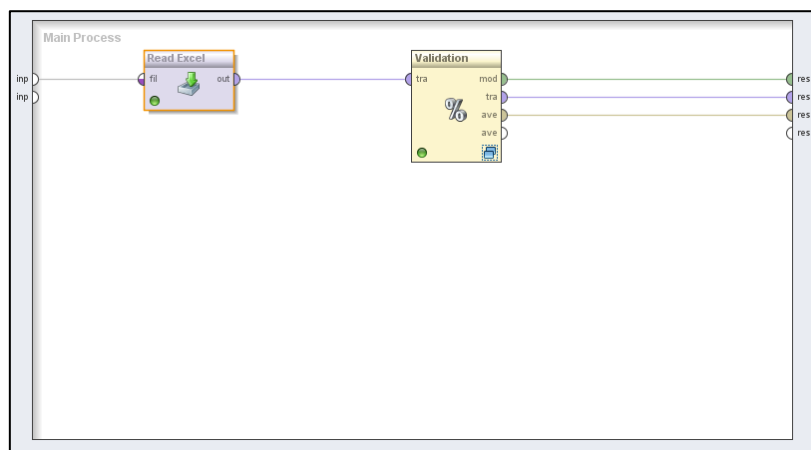
$$P(X | \text{Kredit} = \text{Macet}) = P(\text{Jenis Kelamin} = \text{Laki-laki} | \text{Kredit} = \text{Macet}) * P(\text{Usia} = 30 - 40 \text{ Tahun} | \text{Kredit} = \text{Macet}) * P(\text{Status} = \text{Kawin} | \text{Kredit} = \text{Macet}) * P(\text{Pekerjaan} = \text{Karyawan Swasta} | \text{Kredit} = \text{Macet}) * P(\text{Penghasilan} < 10 \text{ Juta} | \text{Kredit} = \text{Macet}) * P(\text{Masa Angsuran} < 21 | \text{Kredit} = \text{Macet})$$

$$= 0.10 * 0.00 * 0.00 * 1 * 1 * 0.00$$

$$= \mathbf{0}$$

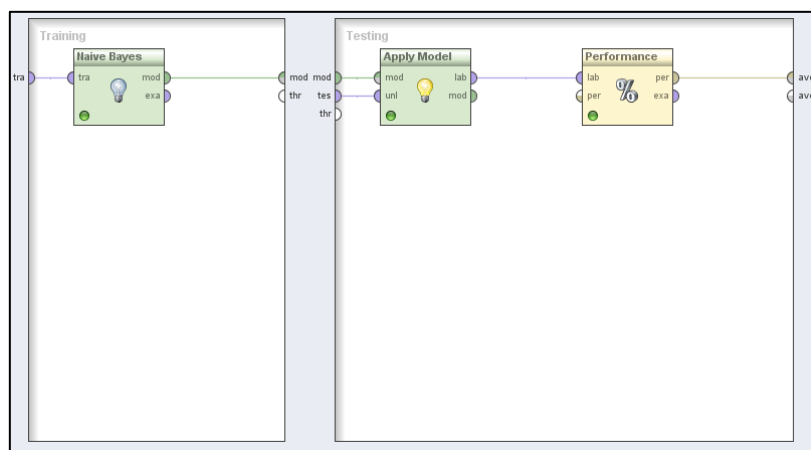
d. Implementasi dengan RapidMiner

Berikut adalah pengolahan data dengan menggunakan naïve bayes pada RapidMiner:



Gambar 4. Pemodelan naïve bayes dengan rapidminer

Pemodelan adalah tahapan (langkah) dalam membuat model dari suatu sistem nyata (realitas). Rapidminer adalah sebuah lingkungan machine learning data mining, text mining dan predictive analytics.



Gambar 5. Proses Training and Testing

Tujuan utama penelitian ini adalah untuk mengetahui nilai akurasi dari algoritma naïve bayes yang digunakan untuk mengklasifikasi kelayakan kredit. Di dalam kolom training terdapat algoritma klasifikasi yang diterapkan yaitu Naïve bayes. Sedangkan di dalam kolom testing terdapat Apply Model untuk menjalankan model naïve bayes dan Performance untuk mengukur performa dari model Naïve bayes tersebut.

e. Hasil pengujian dan percobaan

Hasil dari pengujian model yang telah dilakukan adalah untuk mengukur tingkat akurasi pada gambar 4 merupakan pengujian tools rapidminer dengan jumlah data 28 record. Berikut gambar yang didapat:

accuracy: 65.00% +/- 21.67% (mikro: 64.29%)				
	true Kurang Lancar	true Lancar	true Tidak Lancar	class precision
pred. Kurang Lancar	0	1	0	0.00%
pred. Lancar	5	18	1	75.00%
pred. Tidak Lancar	1	2	0	0.00%
class recall	0.00%	85.71%	0.00%	

Gambar 6. Performance Vector

Jumlah True Positive (TP) adalah 18 record diklasifikasikan sebagai kredit Lancar ,False Negative (FN) sebanyak 5 record diklasifikasikan sebagai kredit Kurang Lancar tetapi kredit Kurang Lancar dan False Negative (FN) sebanyak 1 record diklasifikasikan sebagai kredit Lancar tetapi kredit Tidak Lancar. Berdasarkan tabel 4.5 tersebut menunjukkan bahwa, tingkat akurasi dengan menggunakan algoritma Naïve Bayes Classifier adalah sebesar 65,00%

4. KESIMPULAN

Berdasarkan hasil pembahasan dari uraian yang telah disampaikan diatas, maka dapat ditarik kesimpulan sebagai berikut:

1. Hasil penelitian untuk perhitungan terhadap probabilitas pada bulan Januari diketahui bahwa nilai $P(X| \text{Kurang Lancar})$ lebih besar daripada nilai $P(X| \text{Lancar})$ dan $P(X| \text{Macet})$, sehingga dapat digolongkan Kurang Lancar kemudian pada bulan Februari dan Maret diketahui bahwa nilai $P(X| \text{Lancar})$ lebih besar daripada nilai $P(X| \text{Kurang Lancar})$ dan $P(X| \text{Macet})$ sehingga dapat diambil kesimpulan bahwa dalam kasus Kredit tersebut akan masuk kedalam klasifikasi tingkat kredit Lancar.
2. Hasil pengujian dari penelitian untuk nilai akurasi algoritma klasifikasi Naïve Bayes Classifier adalah 65.00%.
3. Berdasarkan hasil penelitian dari bulan januari, februari hingga maret terjadi perubahan tingkat kelancaran kredit hal ini ditunjukkan dari besar nilai lancar yang meningkat.

5. SARAN

Penelitian nantinya dapat dikembangkan lebih lanjut, dengan mendapatkan data yang bisa dijadikan sebagai tolak ukur serta lebih valid sehingga dapat meningkatkan tingkat akurasi nantinya serta mengembangkan metode lainnya di dalam klasifikasi kelayakan kredit dengan naïve bayes.

DAFTAR PUSTAKA

- [1] Malayu S.P. Hasibuan, “Dasar-Dasar Perbankan”. PT Bumi Aksara : Jakarta, 2009.
- [2] Rivai, Veithzal dan Andria Permata Veithzal, 2006. Credit Manajemen Handbook, Edisi Pertama, Jakarta.

- [3] Bustami, "Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi," TECHSI, vol. III, pp. 11-14, Oktober 2014.
- [4] Silvia Rissino, Germano Lambert-Torres, (2009). Rough Set Theory-Fundamental Concepts, Principals, Data extraction, and Applications, Data Mining and Knowledge Discovery in Real Life Application, February 2009 I-Tech, Vienna, Atria.
- [5] Shyara Taruna R, Saroj Hiranwal, (2013). Enhanced Naive Bayes Algorithm for Intrusion Detection in Data Mining, International Journal of Computer Science and information Technologies, Vol. 4, 2013.
- [6] Shadab Adam Pattekari, Asma Parveen, (2012), Prediction System for Heart Disease Using Naive Bayes, International Journal of Advanced Computer and Mathematical Sciences, ISSN 2230-9624, Vol. 3, Issue 3, 2012