



**PERBANDINGAN METODE REGRESI LOGISTIK BINER DAN NAIVE
BAYES DALAM KLASIFIKASI DEBITUR BERDASARKAN KUALITAS
KREDIT NASABAH**

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Statistika

JURNAL ILMIAH

Oleh

**Bonggo Bawono
B2A218007**

**PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS MUHAMMADIYAH SEMARANG
2019**

PERBANDINGAN METODE REGRESI LOGISTIK BINER DAN NAIVE BAYES DALAM KLASIFIKASI DEBITUR BERDASARKAN KUALITAS KREDIT NASABAH

Bonggo Bawono¹, Tiani Wahyu Utami², Indah Manfaati Nur³

123Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang

e-mail : bonggosaptaji@gmail.com

ABSTRAK

Permintaan kredit baru di Indonesia semakin meningkat. Peningkatan tersebut bersumber dari semua jenis penggunaan kredit, baik modal kerja, investasi, maupun konsumsi. Namun, tidak semua kredit dapat dikembalikan secara sempurna dan tepat waktu artinya akan muncul suatu risiko yang dikenal dengan risiko kredit. Risiko kredit dapat menyebabkan Kredit bermasalah, untuk menangani itu dibutuhkan metode klasifikasi yang tepat dengan mengidentifikasi dan memprediksi nasabah dengan baik sebelum memberikan pinjaman dengan cara memperhatikan data historis pinjaman. Metode klasifikasi yang sering digunakan adalah Regresi Logistik Biner dan Naive Bayes. Namun Masing-masing metode tersebut saling memiliki kelebihan dan kekurangan. Oleh karena itu dilakukan perbandingan nilai akurasi untuk mengetahui metode mana yang paling baik dalam mengklasifikasi Debitur berdasarkan kualitas kredit nasabah. Hasil penelitian ini menunjukkan bahwa, Regresi Logistik Biner memiliki nilai akurasi 99,47% sedangkan Naive Bayes 96,55% yang berarti metode Regresi Logistik Biner lebih baik dari pada Naive Bayes berdasarkan nilai akurasinya.

Kata Kunci : Kualitas Kredit, Klasifikasi, Naive Bayes, Regresi Logistik Biner

ABSTRACT

New credit demand in Indonesia is increasing. The increase is sourced from all types of credit use, whether working capital, investment, or consumption. However, not all credits can be perfectly restored and in a timely manner means that a risk is known as credit risk. Credit risk can lead to problematic credit, for which the appropriate classification method is required by identifying and predicting the customer properly before lending by taking into account the historical data of the loan. The frequently used classification methods are binary logistic regression and the Naive Bayes. However, each one of these methods has advantages and disadvantages. It is therefore a comparison of the methods to determine which method is best in classifying debtor based on the quality of customer credit. The results of this study indicate that, Binary Logistic Regression has an accuracy value of 99.47% while Naive Bayes 96.55% which means the Binary Logistic Regression method is better than Naive Bayes based on its accuracy value.

Keywords : *Quality Credit, Klasifikasi, Naive Bayes, Binary Logistic Regression*

PENDAHULUAN

Survei Perbankan Bank Indonesia pada Desember 2018 memperoleh hasil permintaan kredit baru triwulan IV-2018 meningkat 21,2% pada triwulan sebelumnya yakni 71,7%. Pertumbuhan kredit dari tahun

2016–2018 mengalami peningkatan. Peningkatan tersebut bersumber dari semua jenis penggunaan kredit, baik modal kerja, investasi, maupun konsumsi. Penggunaan kredit paling besar adalah untuk kredit investasi yakni 68,2% pada triwulan

pertama tahun 2019. Kredit investasi digunakan untuk membiayai barang-barang modal, misalnya untuk pembelian mesin-mesin, bangunan dan tanah sebagai modal yang dibiayai. Sehingga, tujuan pemberian kredit salah satunya untuk mendorong dan melancarkan perdagangan, produksi dan jasa-jasa yang kesemuanya ditujukan untuk meningkatkan taraf hidup masyarakat. Namun, pada faktanya, tidak semua kredit dapat dikembalikan secara sempurna dan tepat waktu artinya akan muncul suatu risiko yang dikenal dengan risiko kredit dimana risiko kredit dapat terjadi pada setiap bank.

Resiko kredit dapat terjadi akibat ketidakmampuan nasabah dalam membayar kewajibannya dalam jangka waktu yang telah ditentukan dalam perjanjian pemberian kredit oleh pihak Bank kepada nasabah. Seperti pembayaran pokok pinjaman, pembayaran bunga dan lain-lain yang tidak sesuai dengan jangka waktu yang telah ditetapkan, apabila tidak dikelola dengan baik maka akan mengakibatkan kredit bermasalah (*non performing loan*) yang semakin besar (Hamonangan dkk, 2009).

Berdasarkan Peraturan Bank Indonesia No. 8/19/ 2006 Indonesia tentang Kualitas Aktiva Produktif dan Pembentukan Penyisihan Penghapusan Aktiva Produktif Bank Perkreditan Rakyat, membagi kualitas kredit dua golongan yakni kredit tidak bermasalah dan kredit bermasalah. Kredit bermasalah adalah kredit dengan kualitas kurang lancar, diragukan dan macet (Riyadi, 2006). Menurut (Siswanto Sutojo, 2002) Kredit *non performing* pada umumnya merupakan kredit yang pembayaran angsuran pokok dan atau bunganya telah lewat sembilan puluh hari lebih setelah jatuh tempo, atau kredit yang pembayarannya secara tepat waktu sangat diragukan. Kredit bermasalah perolehan laba dan pengaruh buruk bagi profitabilitas bank dan dapat menyebabkan kebangkrutan.

Dalam menghadapi masalah resiko kredit yang dialami oleh Industri perbankan saat ini salah satunya dapat diatasi dengan mengidentifikasi dan memprediksi nasabah dengan baik sebelum memberikan pinjaman dengan cara memperhatikan data historis pinjaman. Oleh karena itu klasifikasi resiko kredit dalam perbankan memiliki peran yang penting. Apabila pengklasifikasian debitur mengalami kesalahan, maka salah satu dampak yang ditimbulkan adalah kredit bermasalah. Dalam menyelesaikan permasalahan klasifikasi, penggunaan metode atau teknik bertujuan untuk mempermudah proses klasifikasi. Beberapa teknik yang digunakan dalam kasus klasifikasi yaitu *decission tree*, *clasification andassosiation rule*.

Penelitian terdahulu tentang klasifikasi kredit bank telah dilakukan oleh Twala pada tahun 2010. Penelitian tersebut menggunakan beberapa metode antara lain, ANN, Decision Tree, Naive Bayes, k-NN dan logistic discriminant. Penelitian ini menggunakan data dari kredit bank Texas, Kredit di Australia, dan Kredit di Jerman. Hasil yang didapatkan adalah metode Naive Bayes dan Decision Tree memiliki akurasi paling tinggi.

Selain itu, penelitian juga dilakukan oleh Diaprina dan Suhartono tahun 2014 melakukan perbandingan antara Regresi Logistik Biner dengan Radial Basis Function Network dengan studi kasus klasifikasi nasabah kredit dan mendapatkan hasil ketepatan klasifikasi dengan menggunakan metode Regresi Logistik Biner lebih besar dibandingkan dengan menggunakan metode Radial Basis Function Network.

Regresi logistik biner adalah salah satu metode statistika yang sering digunakan untuk mengklasifikasikan sejumlah pengamatan dengan respon biner ke dalam beberapa kelompok berdasarkan satu atau lebih variabel prediktor. Melalui metode ini

akan dihasilkan peluang dari masing-masing kategori respon yang akan dijadikan sebagai pedoman pengklasifikasian dan suatu pengamatan akan masuk ke dalam respon kategori tertentu berdasarkan nilai peluang yang terbesar (Hosmer dan Lemeshow, 2000).

Naive Bayes adalah metode yang digunakan dalam statistika untuk menghitung peluang dari suatu hipotesis, Naive Bayes menghitung peluang suatu variabel respon berdasarkan pada variabel prediktor yang dimiliki dan menentukan variabel respon yang memiliki peluang paling tinggi (Lu, Ling, & Huang, 2003). Metode Naive Bayes juga memiliki beberapa keunggulan seperti mudah serta biaya perhitungan kecil (Wu & Kumar, 2009), dapat menangani data missing (Kang, Yoo, & Han, 2012), memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar (Escalante, Morales, & Sucar, 2016). Namun Metode Naive Bayes memiliki asumsi independensi atribut (Zhang, Liu, & Alpanidis, 2017). Penelitian yang sudah dilakukan masing-masing algoritma memiliki kekurangan dan kelebihan masing-masing oleh karena itu penulis akan melakukan perbandingan terhadap algoritma Regresi Logistik Biner dan Naive Bayes, untuk klasifikasi debitur berdasarkan kualitas kredit dengan membandingkan nilai akurasi.

TINJAUAN PUSTAKA

1. Regresi Logistik Biner

Regresi logistik biner merupakan suatu metode analisis regresi yang menggambarkan pola hubungan antara satu variabel respon (dependen) dengan satu atau lebih variabel prediktor (independen). Variabel respon terdiri dari dua kategori yaitu “sukses” dan “gagal” yang dinotasikan dengan $y = 1$ (sukses) dan $y = 0$ (gagal). Sehingga variabel akan

mengikuti distribusi Bernoulli untuk setiap observasi tunggal. Fungsi probabilitas untuk setiap observasi adalah

$$f(y) = \pi^y(1 - \pi)^{1-y} \text{ dengan } y = 0,1$$

Analisis regresi logistik merupakan suatu teknik untuk menganalisis data yang peubah responnya memiliki dua atau lebih kategori dengan satu atau lebih peubah penjelas yang berskala kategorik atau kontinu. Hosmer dan Lemeshow (2000) menjelaskan bahwa model regresi logistik dibentuk dengan nilai $P(Y = 1|x)$ sebagai $\pi(x)$, yang dinotasikan sebagai berikut:

$$\pi(x) = \frac{\exp(g(x))}{1 + \exp(g(x))}$$

Suatu fungsi dari $\pi(x)$ dicari dengan menggunakan transformasi logit, yaitu $g(x)$ yang dapat dinyatakan sebagai berikut:

$$g(x) = \ln \frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \dots + \beta_p x_p$$

2. Naive Bayes

Naive Bayes merupakan sebuah pengklasifikasi probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua variabel prediktor independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kategori variabel respon (Patil dan Shereker, 2013). 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 di masa sebelumnya (Bustami, 2013).

Penghitungan *Naive Bayes* dapat dijelaskan dengan C_i adalah hipotesis

data X yang merupakan suatu variabel respon $P(C_i|X)$ adalah peluang hipotesis variabel respon berdasarkan kategori X (*posteriori probability*). $P(C_i)$ adalah peluang hipotesis variabel respon (*prior probability*). $P(X|C_i)$ adalah peluang data X berdasarkan kategori pada hipotesis variabel respon. $P(X)$ adalah jumlah probabilitas data X yang nilainya adalah 1. Sehingga didapatkan rumus penghitungan *Naive Bayes* dituliskan pada persamaan.

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

Klasifikasi adalah proses pencarian sekumpulan model atau fungsi yang menggambarkan dan membedakan kategori variabel respon data dengan tujuan agar model tersebut dapat digunakan untuk memprediksi kategori variabel respon dari suatu objek yang belum diketahuikategori variabel responnya. Selain itu, klasifikasi *Naive Bayes* terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam basis data dengan jumlah yang besar (Han dan Kamber, 2006). Salah satu penerapan theorem Bayes dalam klasifikasi adalah *Naive Bayes*.

$$P(X|C_i) = \prod_{k=1}^n P(X_k | C_i)$$

Untuk menentukan nilai variabel respon pada variabel prediktor tertentu digunakan persamaan:

$$arg\max_{c \in C} = \frac{P(X|C_i) P(C_i)}{P(X)}$$

3. Klasifikasi

Klasifikasi merupakan proses untuk menemukan sebuah model atau fungsi untuk menjelaskan dan membedakan kategori data atau konsep yang bertujuan untuk memprediksi atau memperkirakan kategori dari suatu

objek dimana kategori variabel responnya belum diketahui (Han et al., 2012). Metode klasifikasi terdiri dari *Naive Bayes Classifier*, *Decision tree*, *k-Nearest Neighbors* (k-NN), *Neural Network*, Jaringan Syarat Tiruan (JST), Analisis statistik yaitu Regresi Logistik, Algoritma Genetik, *Rough Sets*, *Metode Rule Based* dan *Support Vector Machines* (SVM) (Gorunescu, 2011).

Ada tiga tahap dalam proses pembuatan model klasifikasi (Vercellis, 2009):

1. Tahap Pembelajaran
2. Tahap Pengujian
3. Tahap Prediksi
4. **Klasifikasi**

Beberapa orang menganggap bahwa istilah kredit dalam konteks perbankan memiliki arti yang sama dengan hutang piutang pada umumnya, namun apabila dilihat lebih lanjut keduanya memiliki pengertian yang jauh berbeda (Wijaya, 2011). Dalam bahasa Latin, kredit disebut “*creder*” yang artinya percaya, yaitu kepercayaan dari kreditor bahwa debiturnya akan mengembalikan pinjaman beserta bunganya sesuai dengan perjanjian kedua belah pihak (Kasmir, 2005). Sedangkan menurut H. V. Rivai & Veithzal, 2006 kredit adalah penyerahan barang, jasa, atau uang dari satu pihak (kreditor/atau pemberi pinjaman) atas dasar kepercayaan kepada pihak lain (debitur atau pengutang/*borrower*) dengan janji membayar dari penerima kredit kepada pemberi kredit pada tanggal yang telah disepakati kedua belah pihak. Selain pengertian kredit di atas, berikut adalah beberapa pengertian kredit menurut:

Pengertian kredit menurut Undang-undang Perbankan No. 10 Tahun 1998 adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu berdasarkan persetujuan atau kesepakatan pinjam-meminjam antara bank dan pihak lain yang mewajibkan pihak peminjam untuk melunasi utangnya setelah jangka waktu tertentu dengan jumlah bunga imbalan atau pembagian hasil keuntungan (Indonesia, 1998).

Ketentuan Pasal 1 angka 3 Peraturan Bank Indonesia 5 Oktober 2006, Nomor 8/19/PBI/2006 tentang Kualitas Aktiva Produktif dan Pembentukan Penyisihan Penghapusan Aktiva Produktif Bank Perkreditan Rakyat, kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dengan pihak peminjam yang mewajibkan pihak peminjam untuk melunasi utangnya setelah jangka waktu tertentu dengan pemberian bunga.

Setelah membahas pengertian kredit maka akan dilanjutkan dengan unsur-unsur kredit. Unsur-unsur yang terkandung dalam pemberian suatu fasilitas kredit adalah sebagai berikut (Suyatno, dkk, 2003):

Kepercayaan yaitu keyakinan dari si pemberi kredit bahwa prestasi yang diberikannya, baik dalam bentuk uang, barang, atau jasa, akan benar-benar diterimanya kembali dalam jangka waktu tertentu di masa yang akan datang.

5. Kualitas Kredit

Jumlah kredit yang disalurkan sangat berpengaruh terhadap hidup

matinya lembaga keuangan. Banyaknya jumlah kredit yang disalurkan juga harus memperhatikan kolektibilitas kredit tersebut. Berdasarkan Peraturan Bank Indonesia No. 8/19/ 2006 Indonesia tentang Kualitas Aktiva Produktif dan Pembentukan Penyisihan Penghapusan Aktiva Produktif Bank Perkreditan Rakyat, membagi tingkat kolektibilitas kredit menjadi 4 (empat) golongan yaitu sebagai berikut:

1. Lancar

Kredit Lancar yaitu kualitas kredit apabila tidak terdapat tunggakan pembayaran pokok dan/atau bunga.

2. Kurang Lancar

Kredit Kurang Lancar adalah *sub-standard* yaitu kualitas kredit yang tingkat pengembaliannya mencerminkan keadaan yang kurang baik karena terdapat tunggakan pembayaran dalam jangka waktu tertentu yang ditetapkan.

3. Diragukan

Kredit Diragukan adalah *doubtfull loan* yaitu kredit yang digolongkan diragukan karena kredit yang bersangkutan tidak memenuhi kriteria lancar dan kredit kurang lancar serta akan mempengaruhi kualitas kolektibilitas kredit, tetapi berdasarkan penilaian, dapat disimpulkan bahwa kredit masih dapat diselamatkan dan agunan (nya) senilai sekurang-kurangnya 75% dari utang peminjam, termasuk bunga tersebut atau kredit tidak dapat diselamatkan, tetapi agunannya

masih bernilai sekurang-kurangnya 100% dari utang peminjam.

4. Macet

Kredit Macet adalah *bad debt* yaitu kredit yang tidak memenuhi kriteria lancar, kredit kurang lancar dan kredit diragukan serta akan berpengaruh terhadap kualitas kolektibilitas kredit, dan/atau, memenuhi kriteria diragukan, tetapi dalam jangka waktu 21 bulan sejak digolongkan diragukan belum ada pelunasan atau usaha penyelamatan kredit atau, penyelesaiannya telah diserahkan kepada Pengadilan Negeri atau Badan Urusan Piutang Negara (BUPN) atau telah diajukan penggantian ganti rugi kepada perusahaan asuransi kredit. Lalu, membagi kualitas kredit menjadi dua golongan yakni kredit tidak bermasalah dan kredit bermasalah. Kredit bermasalah adalah kredit dengan kualitas kurang lancar, diragukan dan macet (Riyadi, 2006).

6. Non Performing Loan

Non Performing Loan (NPL) sering disebut kredit bermasalah, dapat diartikan sebagai rasio yang menunjukkan kemungkinan terjadinya risiko tidak tertagihnya piutang terhadap sejumlah pinjaman yang telah diberikan (V. Rivai, Veithzal, & Idroes, 2007). Kredit bermasalah adalah kredit dengan kualitas kurang lancar, diragukan dan macet (Riyadi, 2006). Menurut Dendawijaya, kemacetan fasilitas kredit disebabkan oleh 2 faktor yaitu (Dendawijaya, 2005):

1. Dari pihak perbankan

Dalam hal ini pihak analisis kredit kurang teliti baik dalam mengecek kebenaran dan keaslian dokumen maupun salah dalam menghitung rasio-rasio yang ada. Akibatnya, apa yang seharusnya terjadi, tidak diprediksi sebelumnya.

2. Dari pihak Debitur

Kemacetan kredit yang disebabkan debitur diakibatkan oleh adanya unsur kesengajaan ataupun tidak sengaja. Kriteria penilaian tingkat kesehatan rasio NPL dapat dilihat pada tabel dibawah ini (B. Indonesia, 2011):

Tabel 1. Tingkat Kesehatan Rasio NPL

Rasio	Predikat
$NPL \leq 5\%$	Sehat
$NPL \geq 5\%$	Tidak Sehat

METODE PENELITIAN

1. Sumber Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari sebuah Bank. Data yang memiliki 8 Variabel prediktor yang terdiri dari Plafond, Saldo September 2017, Tunggakan Pokok, Tunggakan Bunga, Suku Bunga, Jaminan, Nilai Jaminan, Jangka Waktu (bulan) dan satu variabel respon yakni Kualitas Kredit. Jumlah keseluruhan dataset yang digunakan ada 3618 data.

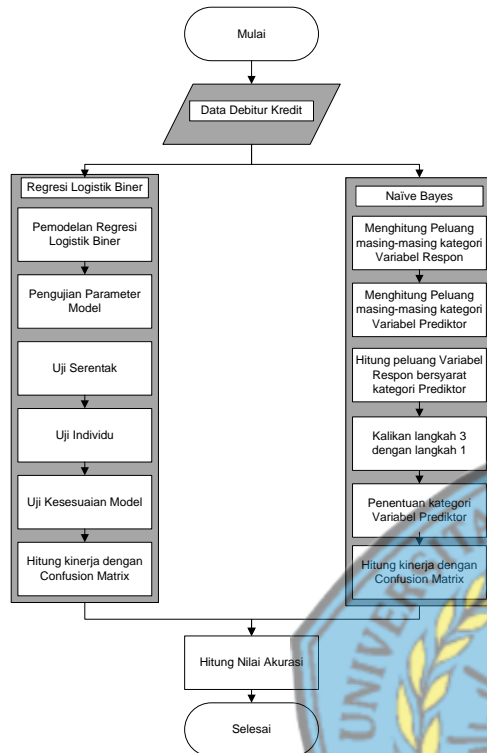
2. Variabel Penelitian

Variabel penelitian terdiri dari variabel prediktor sebanyak 8 variabel dan variabel respon sebanyak 1 variabel dengan jenis data kategorik. Variabel respon antara lain Plafond, Saldo September 2017, Tunggakan Pokok, Tunggakan Bunga, Suku Bunga,

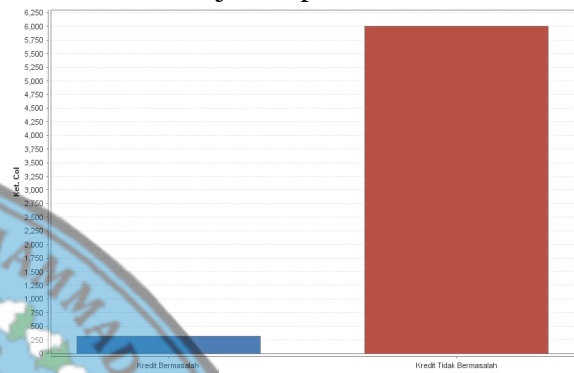
Jaminan, Nilai Jaminan, Jangka Waktu (bulan variabel prediktor yakni kualitas kredit.

3. Langkah-langkah penelitian

Adapun langkah-langkah penelitian ini adalah:



Berdasarkan analisis deskriptif yang dilakukan menggambarkan bahwa debitur di sebuah bank tersebut memiliki jumlah total kredit atau pinjaman yang diberikan oleh pihak Bank rata-rata 23.250.815 juta rupiah dengan rata-rata waktu pinjaman 28 bulan. Debitur juga memberikan jaminan rata-rata 24.866.215 juta rupiah. Selain itu, suku bunga yang diberikan bank kepada debitur pun rata-rata 13%. Sedangkan terdapat beberapa debitur yang memiliki tunggakan pokok dan tunggakan bunga mencapai 107.096.582 juta rupiah dan 9.155.689 juta rupiah.



Gambar 1 Kualitas Kredit Nasabah Sebuah Bank

HASIL DAN PEMBAHASAN

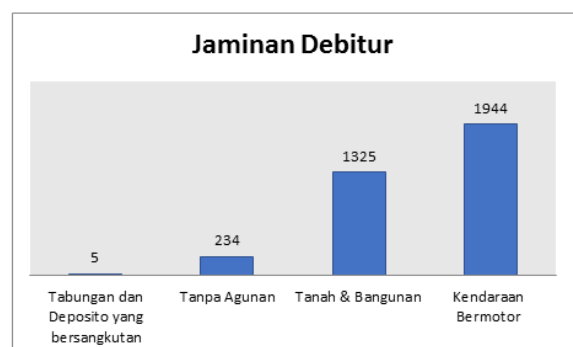
1. Preprocessing data

Analisis yang dilakukan adalah setelah mendapatkan data maka melakukan proses pembersihan data untuk melihat apakah ada missing value atau tidak. Lalu hasil yang didapat tidak ada missing value sehingga lanjut ke analisis berikutnya yakni analisis deskriptif.

2. Analisis Deskriptif

Variabel	Min	Max	Rata-rata	Standar deviasi
(X ₁)	1.000.000	600.000.000	23.250.815	3.4514198
(X ₂)	500	525.000.000	1.6754887	2.8948620
(X ₃)	0	107.096.582	930.571,58	4021121.32
(X ₄)	0	9.155.689	439.715,25	
(X ₅)	6	36	13	-
(X ₆)	5	192	28	-
(X ₇)	0	89.2740.000	24.866.215	-

Analisis deskriptif variabel respon dapat dijelaskan berdasarkan gambar kualitas kredit di atas. Variabel kualitas kredit memiliki dua kategori yakni kredit bermasalah dan kredit tidak bermasalah. Gambar menunjukkan bahwa sebanyak 94,22% nasabah merupakan debitur dengan kredit tidak bermasalah dan sisanya sebesar 5,78% merupakan debitur dengan kredit bermasalah..



Gambar 2 Kondisi Jaminan yang digunakan Debitur

Berdasarkan kondisi jaminan yang digunakan oleh debitur dapat dijelaskan bahwa sebanyak 54,01% debitur menggunakan kendaraan bermotor sebagai agunan dan 38% menggunakan tanah & bangunan serta sebanyak 7% debitur tidak menggunakan agunan dan hanya 0,9% debitur menggunakan tabungan dan deposito.

Model yang digunakan untuk klasifikasi dengan menggunakan Regresi Logistik Biner adalah

$$\pi(x) = \frac{\exp(-0,405 + 0,000x_1 + 0,000x_2 - 0,000x_3 - 0,000x_4 + 5,178x_8)}{1 + \exp(-0,405 + 0,000x_1 + 0,000x_2 - 0,000x_3 - 0,000x_4 + 5,178x_8)}$$

Setelah mendapatkan model regresi logistik di atas maka dilakukan klasifikasi dengan menggunakan data testing. Berikut hasil dari klasifikasi dengan menggunakan regresi logistik biner.

Tabel 2 Confusion Matrix Klasifikasi Kualitas Kredit dengan Regresi Logistik Biner

	Prediksi Kredit tidak	Prediksi Kredit	Jumlah
	Bermasalah	Bermasalah	
Kenyataan Kredit tidak Bermasalah	3399	9	3408
Kenyataan Kredit Bermasalah	10	200	210
Jumlah	3409	209	3618

Berdasarkan tabel confusion matrix didapatkan bahwa ketepatan prediksi dengan menggunakan regresi logistik biner berdasarkan dengan nilai akurasi sebesar 99,47%. Serta, berdasarkan nilai sensitivitas dapat digambarkan bahwa dengan menggunakan regresi logistik biner mampu memprediksi tepat kelas kredit tidak bermasalah dengan kredit yang sebenarnya merupakan kredit tidak bermasalah sebesar 99,71%. Sebaliknya dengan prediksi yang harusnya kredit bermasalah sebesar 95,73%.

Tabel 2 Confusion Matrix Klasifikasi Kualitas Kredit dengan Naive Bayes

	Prediksi Kredit tidak	Prediksi Kredit	Jumlah
	Bermasalah	Bermasalah	
Kenyataan Kredit tidak Bermasalah	3336	52	3388
Kenyataan Kredit Bermasalah	73	157	230
Jumlah	3409	209	3618

Berdasarkan tabel *confusion matrix* didapatkan bahwa ketepatan prediksi dengan menggunakan naive bayes berdasarkan dengan nilai akurasi sebesar 96,55%. Serta, berdasarkan nilai sensitivitas dapat digambarkan bahwa dengan menggunakan naive bayes mampu memprediksi tepat kelas tidak kredit bermasalah dengan kredit yang sebenarnya merupakan kredit tidak bermasalah sebesar 98,46%. Sebaliknya dengan prediksi yang harusnya kredit bermasalah sebesar 75,11%.

Tabel 3 Perbandingan Akurasi

Evaluasi	Regresi Logistik Biner	Naive Bayes
Akurasi	99,47 %	96,55 %

Berdasarkan perbandingan nilai akurasi diatas dapat diketahui yang memiliki nilai akurasi paling tinggi adalah dengan menggunakan regresi logistik biner. Regresi logistic biner dapat memprediksi tepat untuk kredit bermasalah dan kredit tidak bermasalah sebesar 99,47%.

KESIMPULAN

1. Gambaran umum debitur di sebuah bank adalah sudah 94,22% debitur merupakan pemilik kredit dengan kategori kredit tidak bermasalah yang artinya bank tersebut memiliki debitur yang sehat.
2. Model yang didapatkan dengan menggunakan regresi logistik biner adalah

$$\pi(x) = \frac{\exp(-0,405 + 0,000x_1 + 0,000x_2 - 0,000x_3 - 0,000x_4 + 5,178x_8)}{1 + \exp(-0,405 + 0,000x_1 + 0,000x_2 - 0,000x_3 - 0,000x_4 + 5,178x_8)}$$

Model yang didapat dengan menggunakan naïve bayes adalah

$$P(X|C_i) = P(X_1|C_i) * P(X_2|C_i) * P(X_3|C_i) * P(X_4|C_i) * P(X_5|C_i) * P(X_6|C_i) * P(X_7|C_i) * P(X_8|C_i) * P(X_9|C_i) * P(X_{10}|C_i)$$

Dengan $P(C_1) = 0,94$ dan $P(C_2) = 0,06$

3. Hasil Perbandingan metode Regresi logistik biner dengan naïve bayes menunjukkan bahwa regresi logistik biner memiliki akurasi lebih tinggi dibandingkan dengan naïve bayes yakni sebesar 99,47%

SARAN

Sebagai pengembangan metode diperlukan beberapa teknik evaluasi lagi agar dapat melihat benar-benar terhadap perbedaan antara metode yang dibandingkan. Serta melakukan komparasi dengan metode klasifikasi lainnya agar dapat digunakan untuk klasifikasi debitur secara akurat.

DAFTAR PUSTAKA

- Bustami 2013, Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi, *TECHSI : Jurnal Penelitian Teknik Informatika*, Vol. 3, No.2, Hal. 127-146
- D. W. Hosmer, and S. Lemeshow. 2000. *Applied Logistic Regression*. New York: Inc, Jhon Willey and Sons.
- Dendawijaya, 2005. *Manajemen Perbankan*, Edisi Kedua, Cetakan Kedua, Ghalia Indonesia, Bogor Jakarta
- Diaprina, S. R., & Suhartono. 2014. Analisis Klasifikasi Kredit Menggunakan Regresi Logistik Biner Dan Radial Basis Function Network di Bank X Cabang Kediri. *Jurnal Sains Dan Seni POMITS*, 3(2), 1–6.
- Gorunescu, F. 2011. *Data Mining: Concepts, Models and Techniques*. *Data mining - Concepts, Models and Technique*. Springer. <https://doi.org/10.1007/978-3-642-19721-5>
- Hamonangan, dkk. 2009. Pengaruh Capital Adequacy Ratio, Debt to Equity Ratio, Non Performing Loan, Operating Ratio dan Loan to Deposit Ratio terhadap Return On Equity (ROE) Perusahaan Perbankan yang Terdaftar di Bursa Efek Indonesia, *Jurnal Akuntansi* 13, Universitas Sumatera Utara, Medan.
- Han, J., Kamber, M., & Pei, J. 2012. *Data Mining: Concepts and Techniques*. San Francisco, CA, itd: Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-381479-1.00001-0>
- Indonesia, G. B. (2006). *Peraturan Bank Indonesia Nomor: 8/19/PBI/2006 tentang Kualitas Aktiva Produktif dan Pembentukan Penyisihan Penghapusan Aktiva Produktif Bank Perkreditan Rakyat*. Jakarta.
- Indonesia, R. Undang-Undang Republik Indonesia Nomor 10 Tahun 1998 Tentang Perubahan Atas Undang-Undang Nomor 7 Tahun 1992 Tentang Perbankan, Pub. L. No. 10, Sekretariat Negara (1998). Indonesia.
- Kang, H., Yoo, S. J., & Han, D. 2012. Expert Systems with Applications Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems With Applications*, 39(5), 6000–6010. <https://doi.org/10.1016/j.eswa.2011.11.107>
- Kasmir. 2005. *Bank dan Lembaga Keuangan Lainnya*. Jakarta: RajaGrafindo Persada.
- Kasmir. 2011. *Analisis Laporan Keuangan*. Jakarta: Rajawali Pers.
- Lu, J., Ling, C. X., & Huang, J. 2003. Comparing Naive Bayes , Decision Trees , and SVM with AUC and Accuracy. *The Third IEEE International Conference on DataMining*, 11–14.
- Lu, D., & Weng, Q. 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5), 823–870. <https://doi.org/10.1080/01431160600746456>
- Pattekari, S. A., Parveen, A., 2012, Prediction System for Heart Disease

- Using Naive Bayes, *International Journal of Advanced Computer and Mathematical Sciences*, ISSN 2230-9624, Vol. 3, No 3, Hal 290-294.
- Patil, T. R., Sherekar, M. S., 2013, Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification, *International Journal of Computer Science and Applications*, Vol. 6, No. 2, Hal 256-261.
- Ridwan, M., Suyono, H., Sarosa, M., 2013, Penerapan Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier, *Jurnal EECCIS*, Vol 1, No. 7, Hal. 59-64.
- Rivai, H. V., & Veithzal, A. P. 2006. *Credit management handbook: teori, konsep, prosedur, dan aplikasi panduan praktis mahasiswa, bankir, dan nasabah*. Jakarta: Raja Grafindo Persada.
- Rivai, V., Veithzal, A. P., & Idroes, F. N. 2007. *Bank and financial institution management*. Jakarta: Raja Grafindo Persada.
- Sutojo, Siswanto. 2002. *Menangani Kredit Bermasalah: konsep, teknik, dan kasus*. Jakarta: Gramedia.
- Suyatno, T., Sukada, M., Chalik, H. A., & Ananda, C. T. Y. 2003. *Dasar-dasar perkreditan* (Edisi Keem). Jakarta: PT. Gramedia Pustaka Utama.
- Twala, V. J. 2010. Multiple classifier application to credit risk assessment, *51*(8), 3968–3982. <https://doi.org/10.1016/j.cstda.2006.05.025>
- Wijaya, A. 2011. *Penyelesaian Kredit Bermasalah Pada PT. Bank Perkreditan Rakyat XYZ di Depok*. Universitas Inonesia.
- Wu, X., & Kumar, V. 2009. *The Top Ten Algorithm in Data Mining*. Taylor & Francis Group, LLC.
- Zhang, C., Liu, C., Zhang, X., & Alpanidis, G. 2017. An Up-to-Date Comparison of State-of-the-Art

Classification Algorithms. *Expert Systems With Applications*. <https://doi.org/10.1016/j.eswa.2017.04.003>