

PEMANFAATAN STATUS KREDIT NASABAH UNTUK MENGEVALUASI PEMBIAYAAN KPR PADA BANK MUAMALAT INDONESIA MENGGUNAKAN *DATA MINING*

Tri Wahyudi^{1*}, Richardus Eko Indrajit², Muh. Fauzi³

¹Pasca Sarjana Ilmu Komputer, STMIK Nusa Mandiri, Jakarta, Indonesia

²ABFI Institute Perbanas, Jakarta, Indonesia

³STMIK Bumigora Mataram, Mataram – Indonesia

¹ Jl. Kramat Raya No.16AB, RT.5/RW.7, Kwitang, Senen, Kota Jakarta Pusat

² Jl. HR Rasuna Said, Karet Kuningan, Kecamatan Setiabudi, Kota Jakarta Selatan

³Jl. Ismail Marzuki No.22, Cilinaya, Cakranegara, Kota Mataram, Nusa Tenggara Barat

*E-mail : triwahyudi100390@gmail.com

ABSTRAK

Rumah adalah salah satu kebutuhan *primer* yang wajib dimiliki semua orang, tetapi tidak semua orang mampu untuk membeli rumah secara tunai. Bagi masyarakat yang tidak mampu untuk membeli rumah dengan cara cash bisa melalui bank sebagai salah satu penyalur penyedia pembiayaan pembelian rumah yang pembayaran bisa angsuran setiap bulan, dari penyaluran kredit pembiayaan rumah ada beberapa nasabah yang melakukan pembayaran angsuran tidak tepat waktu sebagai menjadi status kredit menjadi macet atau tidak lancar. Teknik klasifikasi *data mining* dapat membantu untuk mengevaluasi faktor apa saja yang mempengaruhi kredit macet nasabah.

Hasil dari penelitian yang dilakukan ini dengan menggunakan decision tree menghasilkan nilai akurasi sebesar 90%.

Kata Kunci : Data Mining, Status Kredit, Rumah, Pembiayaan

ABSTRACT

Home is one of the primary needs that everyone must have, but not everyone can afford to buy a house in cash. For people who can't afford to buy a house by cash can be through the bank as one of the providers of home financing providers that payments can be installments every month, from the distribution of home financing loans there are some customers who make installments not timely payment into credit status becomes stuck or not smoothly. Data mining classification techniques can help to evaluate what factors affect the customer's bad credit. The result of this research by using decision tree yields accuracy value of 90%.

Keywords : Data Mining, Credit Status, Home, Financing

PENDAHULUAN

Manusia hidup tidak terlepas dari kebutuhan yang harus dipenuhi untuk keberlangsungan hidup, kebutuhan manusia terdiri dari kebutuhan primer, sekunder dan tersier. Kebutuhan primer adalah kebutuhan pokok yang harus dipenuhi oleh manusia seperti makan, minum, pakaian dan rumah, dari kebutuhan pokok seperti makan, minum, pakaian dan rumah hanya rumah yang tidak semua orang mampu untuk membeli secara tunai, karena keterbatasan kemampuan tiap individu untuk membeli secara tunai maka banyak yang memanfaatkan perbankan untuk membeli rumah secara kredit atau biasa yang disebut dengan KPR.

Banyak perbankan yang menyalurkan pembiayaan kredit rumah secara angsuran dimana nasabah tiap bulannya membayar cicilan, setiap bulan nasabah mempunyai kewajiban membayar cicilan atau angsuran, namun ada beberapa nasabah yang membayar tidak tepat waktu, hal itu tentu berdampak pada kredit macet perbankan yaitu ketidakmampuan nasabah dalam membayar cicilan atau angsuran yang dapat menyebabkan status kredit nasabah berubah menjadi macet,

Dari status kredit bisa dievaluasi dengan menggunakan data mining, dimana data mining mengacu pada kelas yang luas dari kegiatan yang memiliki kesamaan pencarian lebih cara

yang berbeda untuk memproses atau data paket statistik atau ekonometrik dengan tujuan membuat presentasi akhir memenuhi kriteria desain tertentu .

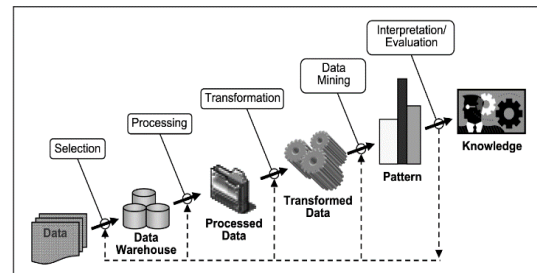
Kondisi kemacetan kredit berdampak dari penurunan ekonomi dan penurunan keuangan rumah tangga, isu penggunaan status kredit pribadi sebagai kriteria dalam proses perekrutan telah menarik peningkatan pengawasan. Bank muamalat indonesia sebagai salah satu perbankan yang memiliki program kredit kpr tentunya tidak menginginkan pembiayaan-nya yang telah disalurkan menjadi macet atau gagal bayar nasabah, yang akan berdampak pada kondisi keuangan menjadi tidak sehat. oleh karena itu pemanfaatan status kredit macet nasabah yang diolah dengan data mining akan memperkecil resiko untuk terjadi kredit macet di kemudian hari pada nasabah baru yang akan mengajukan pembiayaan kpr.

Kredit

Kredit merupakan suatu fasilitas keuangan yang memungkinkan seseorang atau badan usaha untuk meminjam uang untuk membeli produk dan membayarnya kembali dalam jangka waktu tertentu. Berdasarkan UU No. 10 tahun 1998 menyebutkan bahwa kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam untuk melunasi utangnya setelah jangka waktu tertentu dengan pemberian bunga. Kredit dapat diklasifikasi menjadi 2 kategori yaitu kredit lancar dan kredit macet, agar kredit yang diajukan tidak termasuk kategori macet maka setiap pengajuan kredit melalui proses analisa kelayakan kredit.

Data Mining

Data mining adalah proses menganalisa data yang sudah ada dalam *database* dan pola tersembunyi di dalamnya. *Data mining* menggunakan teknik statistik dan kecerdasan buatan [23]. *Data mining* mengacu pada kelas aktivitas yang umum dan memiliki persamaan dalam pencarian melalui berbagai cara untuk memproses secara statistik. *Data mining* merupakan bagian dari proses *Knowledge Discovery Database (KDD)* dimana *Knowledge Discovery Database (KDD)* memiliki tahapan proses pada gambar dibawah



Gambar 1. Proses *Knowledge Discovery Database (KDD)*

Tahapan Proses *Knowledge Discovery Database (KDD)* pada gambar 1 sebagai berikut:

a. Data Selection

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

b. Processing

Processing dimana proses pembersihan data mencakup membuang duplikasi data, memeriksa data inkonsisten, dan memperbaiki kesalahan data.

c. Transformation

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam *Knowledge Discovery Database (KDD)* merupakan proses kreatif dan sangat bergantung pada jenis atau pola informasi yang dicari dalam basis data.

d. Proses mining

Proses *mining* adalah proses mencari pola atau informasi dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery Database (KDD)* secara keseluruhan.

e. Interpretation (evaluasi)

Tahapan ini merupakan bagian dari proses *Knowledge Discovery Database (KDD)* yang disebut *Interpretation*. Pola ini melihat apakah ada sesuatu yang baru dan menarik dan dilakukan *iterasi* jika diperlukan. Tahap ini akan diketahui apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

Jika semua proses *Knowledge Discovery Database (KDD)* telah selesai akan menjadi sebuah *knowledge* yang akan digunakan oleh

user atau pengguna untuk mengambil sebuah keputusan.

Nasabah

Nasabah menurut kamus besar bahasa Indonesia (KBBI) adalah orang yang berhubungan dengan atau menjadi pelanggan bank (dalam hal keuangan). Nasabah mempunyai peranan yang penting dalam perbankan karena nasabah yang mempunyai modal atau uang yang disimpan pada suatu perbankan. Menurut undang-undang No. 10 tahun 1998 tentang perubahan atas undang-undang No. 7 tahun 1992 tentang perbankan diatur perihal nasabah yang terdiri dari dua pengertian yaitu :

1. Nasabah penyimpan adalah nasabah yang menempatkan dananya di bank dalam bentuk simpanan berdasarkan perjanjian bank dengan nasabah yang bersangkutan.
2. Nasabah debitur adalah nasabah yang memperoleh fasilitas kredit atau pembiayaan berdasarkan prinsip syariah atau dipersamakan dengan itu berdasarkan perjanjian bank dengan nasabah yang bersangkutan.

Decision Tree

Sebuah pohon keputusan (DT) *classifier* memiliki struktur seperti pohon, yang berisi dua jenis node - node *internal* dan node daun. *Node internal* sesuai dengan tes untuk sampel pada *atribut individu* atau beberapa dan simpul daun memberikan label kelas untuk sampel berdasarkan *distribusi kelas* catatan itu. Sebuah pohon keputusan mengklasifikasikan sampel secara *top-down*, dimulai dari simpul akar dan. Pembangunan pohon keputusan didasarkan pada *internal node* membelah *rekursif*. Pemilihan perpecahan *atribut* pada *node internal* sangat penting selama proses konstruksi dan menentukan untuk sebagian besar struktur akhir dari pohon keputusan. Banyak upaya harus dilakukan pada aspek ini dan seperangkat kriteria dibagi, seperti indeks (*Breiman, Friedman, Olshen et al. 1984*), gain informasi (*Quinlan 1993*) dan uji Chi-Square (*Biggs & Ville 1991; Kass 1980*), yang tersedia. Teori entropi diadopsi untuk memilih pemecahan atribut yang tepat dengan algoritma C4.5 terkenal (*Quinlan 1993*). Biarkan N menjadi Bab 2. Teknik Klasifikasi dalam Mengontrol Risiko Kredit 123 ukuran dataset D dan N_j jumlah sampel di kelas j . Asumsi bahwa ada label kelas K , teori entropi menyatakan bahwa rata-rata jumlah informasi yang

diperlukan untuk mengklasifikasikan sampel adalah sebagai berikut:

$$Info(D) = -\sum_{j=1}^k \frac{N_j}{N} \log_2 \left(\frac{N_j}{N} \right) \quad (1)$$

Ketika dataset D dibagi menjadi beberapa himpunan bagian D_1, D_2, \dots, D_n menurut hasil dari atribut X , gain informasi didefinisikan sebagai:

$$Gain(X, D) = Info(D) - \sum_{i=1}^n \frac{N_i}{N} Info(D_i) \quad (2)$$

di mana N_i adalah jumlah sampel dalam bagian D_i . ID3, pendahulu C4.5, nikmat semua *atribut* dengan *gain* terbesar, yang bias terhadap *atribut-atribut* yang memiliki banyak hasil. C4.5 dengan demikian berlaku *Gain_ratio*, bukannya *Gain*, sebagai kriteria

$$Gain_ratio(X, D) = Gain(X, D) / \left(-\sum_{i=1}^n \frac{N_i}{N} \log_2 \left(\frac{N_i}{N} \right) \right) \quad (3)$$

C4.5 rakus node dengan lahap sampai nilai sepele dari *Gain_ratio* adalah tercapai. Sebuah prosedur *prune* kemudian dilakukan untuk menghindari menghasilkan pohon kompleks yang *overfits* data. Untuk simpul daun dengan N sampel, E dari yang kesalahan klasifikasi, C4.5 memperkirakan klasifikasi tingkat kesalahan p di bawah asumsi *distribusi binomial*:

$$CF = \sum_{i=0}^E C_N^i P^i (1-p)^{N-i} \quad (4)$$

di mana CF adalah faktor kepercayaan dan biasanya ditetapkan pada 0,25. C4.5 kemudian mensubstitusi *subtree* dengan simpul daun atau cabang jika p menurun sehingga meningkatkan kemampuan prediksi untuk sampel baru. Meskipun sampel N ini yang berangkat pada simpul daun tidak diambil secara independen, yang melanggar asumsi distribusi binomial, jenis pemangkasan Strategi mencapai kinerja yang sangat baik seperti yang ditunjukkan secara empiris.

METODE

Metode data mining yang digunakan dengan menggunakan teknik *Classification* yang bertujuan untuk memprediksi kelas dari objek label kelasnya. Algoritma yang digunakan adalah *decision tree* lalu dilakukan pengujian dengan menggunakan *confusion matrix* dan *ROC curve* untuk mengetahui nilai akurasi dan *AUC (Area Under Curve)* yang didapatkan.

Dataset yang digunakan menggunakan 100 data, dengan memiliki 7 atribut atau variabel yang terdiri dari 6 atribut sebagai *predictor* dan 1 atribut tujuan *Atribut predictor*

Berdasarkan data yang telah ada maka dapat dihitung nilai *accuracy*, *sensitivity*, *specificity*, *ppv* dan *npv*

Tabel 1. Atribut dan field

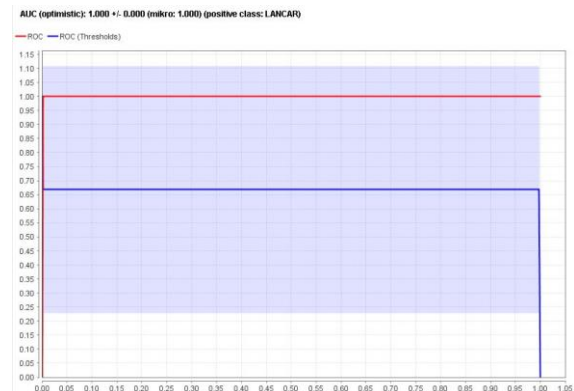
No	Atribut	Nilai Atribut
1	Jenis Kelamin	Laki-laki
		Perempuan
2	Usia	20 - 35 tahun
		36 - 50 tahun
		> 50 tahun
3	Jumlah Pinjaman	250 - 500 juta
		500 - 750 juta
		750 - 1 miliar
		> 1 miliar
4	Tenor	5 tahun
		10 tahun
		15 tahun
		> 15 tahun
5	Jumlah Angsuran Per bulan	3 - 5 juta
		5 -7 juta
		7 - 9 juta
		> 9 juta
6	Tunggakan	3 - 5 juta
		5 - 7 juta
		7 - 9 juta
		> 9 juta
7	Status Kredit	Lancar
		Macet

Tabel 2 hasil akurasi (%)

	Hasil
<i>accuracy</i>	0,900
<i>sensitivity</i>	0,8983
<i>specificity</i>	0,9032
<i>ppv</i>	0,9464
<i>npv</i>	0,8235

Untuk pengukuran AUC (*Area Under Curve*) berbentuk grafik yang lebih jelas bisa dilihat pada gambar dibawah

Gambar 3 nilai akurasi AUC (*Area Under Curve*)



Dari gambar 3 nilai akurasi dari AUC (*Area Under Curve*) dengan nilai sebesar 1,000 yang berarti memiliki nilai klasifikasi yang sangat baik

Hasil pengujian diatas yang menggunakan *confusion matriks* dan RUC untuk mendapatkan nilai akurasi terbukti bahwa hasil *algoritma decision tree* sangat baik untuk digunakan dalam klasifikasi data mining dimana untuk nilai *accuracy* menghasilkan nilai 0,900 dan pada nilai AUC (*Area Under Curve*) menghasilkan nilai 1,000

HASIL DAN PEMBAHASAN

Hasil pengujian yang dilakukan untuk mengukur tingkat akurasi dan AUC (*Area Under Curve*) dari hasil pengujian didapatkan nilai akurasi dari pengujian. Hasil dibawah merupakan hasil nilai akurasi yang telah dilakukan

accuracy: 90.00% +/- 3.33% (mikro: 90.00%)

	true MACET	true LANCAR	class precision
pred. MACET	53	3	94.64%
pred. LANCAR	6	28	82.35%
class recall	89.83%	90.32%	

Gambar 2. Hasil *Confusion Matriks*

Jumlah true positif (TP) 53, true negatif (TN) 28, false positif (FP) 3, false negatif (FN) 6.

KESIMPULAN DAN SARAN

Dari hasil yang telah diuji untuk status kredit nasabah dengan menggunakan *algoritma decision tree* menunjukkan memiliki tingkat akurasi yang tinggi, semua itu bisa dibuktikan nilai akurasi menghasilkan nilai 0,900 dan AUC menghasilkan nilai 1,000. Dengan nilai AUC yang semakin tinggi dengan menggunakan *algoritma decision tree* maka *algoritma decision*

tree mendapatkan nilai “*excellent classification*”.

References

- B.Ozzisikyilma.2009.Analysis, Characterization and Design of Data Mining Applications and Applications to Computer Architecture, *s*, no. 191
- D. Fieldhouse, I. Livshits, & J. Macgee. 2015.Aggregate Fluctuations , Consumer Credit and Preliminary and Incomplete. vol. 2, pp. 1–26
- D. Lee & W. Van der Klaauw. 2010. Federal Reserve Bank of New York Staff Reports An Introduction to the FRBNY Consumer Credit Panel,” *FRB New York Staff Rep. No. 479*.
- E. Damberg. 2014. Data Mining for Description and Prediction of Antibiotic Treated Healthcare-Associated Infections.
- K. D. Hoover & S. J. Perez. 2000. Three attitudes towards data mining, vol. 7, no. 2, pp. 195–210,
- M. T. Bingamawa & H. A. Santoso. 2016. Implementation of Naïve Bayes Algorithm To Determine Customer Credit Status in Pt . Multindo pp. 1–9, *Calif*., vol. 36, no. 11.
- S. Kelly-louw *et al.* 2009. Measures in South African Consumer Credit Legislation Aimed At the Prevention of Reckless Lending and Over-,” no. 2008, pp. 1–21
- T. Menzies & Y. Hu. 2003. Data Mining for Very Busy People,” *Computer (Long Beach Calif)*. vol. 36, no. 11