

Analisa Kelayakan Pemberian Kredit Mobil Dengan Menggunakan Neural Network Backpropagation

Amrin

Program Studi Manajemen Informatika
AMIK Bina Sarana Informatika, Jakarta

Telp : (021) 54376399

E-Mail : amrin.ain@bsi.ac.id

ABSTRAK

Problems are often encountered in the provision of credit is to determine lending decisions to someone, while other issues are not all credit payments can run well. Among the causes are errors of judgment in making credit decisions. In this study will be used back propagation neural network method to analyze the feasibility of providing car loans. From the test results to measure the performance of the method is to use testing methods Confusion Matrix and ROC curve, it is known that the method of back propagation neural network has a value of 89% accuracy and AUC value of 0.831. This shows that the model produced, including the classification is quite good because it has the AUC values between 0.8-0.9.

Keywords : neural network backpropagation, Confusion Matrix, ROC Curva

PENDAHULUAN

Kendaraan sebagai alat transportasi menjadi kebutuhan vital saat ini, terutama untuk mendukung dan memenuhi kegiatan dan mobilitas seseorang. Kebutuhan terhadap kendaraan ini khususnya kendaraan roda empat terkadang tidak dapat terpenuhi karena faktor keuangan dan mahal nya harga kendaraan tersebut. Mengingat harganya yang mahal maka akan kesulitan bagi sebagian orang untuk memilikinya, karena tidak semua masyarakat mampu membeli kendaraan secara tunai, maka mereka membelinya secara kredit.

Pemberian kredit merupakan kegiatan usaha yang mengandung resiko tinggi dan berpengaruh terhadap kesehatan dan keberlangsungan usaha perusahaan pembiayaan. Permasalahan yang sering dihadapi dalam pemberian kredit mobil adalah menentukan keputusan pemberian kredit kepada

seorang calon debitur, sedangkan permasalahan yang lain adalah tidak semua pembayaran kredit mobil oleh debitur dapat berjalan dengan baik atau sering kita sebut kredit macet. Penyebab kredit macet diantaranya adalah kesalahan penilaian dalam membuat keputusan kredit.

Pada kasus permohonan kredit oleh nasabah, pengambil keputusan harus mampu mengambil keputusan yang tepat untuk menerima atau menolak permohonan kredit tersebut.

Selama ini, proses penilaian kelayakan pembiayaan kredit nasabah bersifat konvensional, artinya perusahaan pembiayaan meminta nasabah mengisi formulir berupa daftar pertanyaan dan melengkapi permohonan kredit dengan berkas-berkas yang diperlukan, untuk kemudian dilakukan penilaian permohonan kredit tersebut. Jika salah menilai maka akan menjatuhkan kelangsungan usaha yang bersangkutan.

Untuk memecahkan masalah tersebut, pada penelitian ini akan digunakan metode klasifikasi *neural network backpropagation* untuk menilai kelayakan pemberian kredit mobil. Data yang penulis gunakan adalah data nasabah Bank Perkreditan Rakyat (BPR) Syariah Citayam Depok.

Kelebihan metode neural network diantaranya melakukan generalisasi dan ekstraksi dari pola data tertentu, mampu mengakuisisi pengetahuan walau tidak ada kepastian, menyelesaikan masalah yang tidak terstruktur dan sulit didefinisikan, dan mampu melakukan perhitungan secara paralel sehingga proses lebih singkat.

BAHAN DAN METODE

A. Data Mining

Data mining adalah rangkaian proses untuk menggali nilai tambah berupa informasi yang belum terekplorasi dari sebuah basis data, melakukan eksplorasi dengan cara-cara tertentu untuk memanipulasi data menjadi informasi yang lebih berharga dengan cara mengekstraksi dan mengenali pola penting dari basis data (Han & Kamber, 2006). Menurut Daryl Pregibons dalam (Gorunescu, 2011) "*Data mining* adalah perpaduan dari ilmu statistik, kecerdasan buatan, dan penelitian bidang *database*". Nama *data mining* berasal dari kemiripan antara pencarian informasi yang bernilai dari *database* yang besar dengan menambang sebuah gunung untuk sesuatu yang bernilai (Sumathi, 2006). Keduanya memerlukan penyaringan melalui sejumlah besar material, atau menyelidiki dengan cerdas untuk mencari keberadaan sesuatu yang disebut bernilai tadi.

Data Mining merupakan teknologi baru yang sangat berguna untuk membantu perusahaan-perusahaan menemukan informasi yang sangat penting dari gudang data mereka. Beberapa aplikasi data mining fokus pada prediksi, mereka meramalkan apa yang akan terjadi dalam situasi baru dari data yang menggambarkan apa yang terjadi di masa lalu (Witten, Frank, & Hall, 2011).

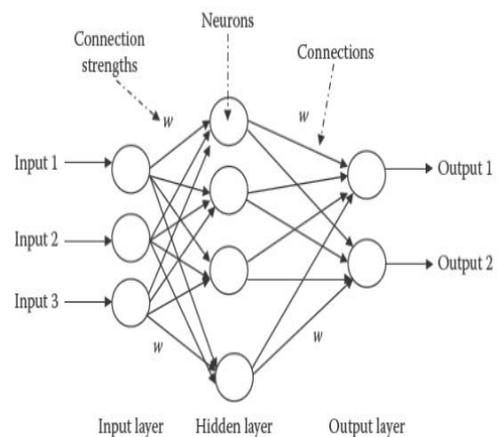
B. Neural Network

Neural network atau jaringan syaraf tiruan adalah merupakan salah satu representasi buatan dan otak manusia yang selalu mencoba mensimulasikan proses pembelajaran pada otak manusia. Istilah buatan disini digunakan karena jaringan syaraf ini diimplementasikan dengan menggunakan program komputer yang mampu menyelesaikan sejumlah proses perhitungan selama proses pembelajaran (Kusuma Dewi, 2010). *Neural network* adalah (Han, 2006) satu set unit *input/output* yang terhubung dimana tiap relasinya memiliki bobot. Hal yang perlu mendapat perhatian istimewa adalah bahwa jaringan syaraf tiruan tidak diprogram untuk menghasilkan keluaran tertentu. Semua keluaran atau kesimpulan yang ditarik oleh jaringan didasarkan pada pengalamannya selama mengikuti proses pembelajaran. Pada proses pembelajaran, kedalam jaringan syaraf tiruan dimasukkan pola-pola input (dan output) lalu jaringan akan diajari untuk memberikan jawaban yang bisa diterima (Diyah Puspitaningrum, 2006).

Neural Network dimaksudkan untuk mensimulasikan perilaku sistem biologi susunan syaraf manusia, yang terdiri dari sejumlah besar unit pemroses yang disebut

neuron, yang beroperasi secara paralel (Alpaydin, 2010). *Neuron* mempunyai relasi dengan *synapse* yang mengelilingi *neuron-neuron* lainnya. Susunan syaraf tersebut dipresentasikan dalam *neural network* berupa graf yang terdiri dari simpul (*neuron*) yang dihubungkan dengan busur, yang berkorespondensi dengan *synapse*. Sejak tahun 1950-an, *neural network* telah digunakan untuk tujuan prediksi, bukan hanya klasifikasi tapi juga untuk regresi dengan atribut target kontinu (Vecellis, 2009).

Neural network terdiri dari dua lapisan atau lebih, meskipun sebagian besar jaringan terdiri dari tiga lapisan : lapisan input, lapisan tersembunyi, dan lapisan output (Larose, 2005). Pendekatan *neural network* dimotivasi oleh jaringan saraf biologis. Secara kasar, *neural network* adalah satu set terhubung input/output unit, di mana masing-masing sambungan memiliki berat yang terkait dengannya. *Neural network* memiliki beberapa ciri yang membuat mereka populer untuk *clustering*. Pertama, *neural network* adalah arsitektur pengolahan *inheren paralel* dan terdistribusi. Kedua, *neural network* belajar dengan menyesuaikan bobot interkoneksi dengan data, Hal ini memungkinkan *neural network* untuk "menormalkan" pola dan bertindak sebagai fitur (atribut) *extractors* untuk kelompok yang berbeda. Ketiga, *neural network* memproses vektor numerik dan membutuhkan pola objek untuk diwakili oleh fitur kuantitatif saja (Gorunescu, 2011).



Gambar 1. Arsitektur Neural Network

C. Algoritma Neural Network Backpropagation

Algoritma pelatihan *backpropagation* atau ada yang menterjemahkan menjadi propagasi balik pertama kali dirumuskan oleh Paul Werbos pada tahun 1974 dan dipopulerkan oleh Rumelhart bersama McClelland untuk dipakai pada *neural network*. Metode *backpropagation* pada awalnya dirancang untuk *neural network feedforward*, tetapi pada perkembangannya, metode ini diadaptasi untuk pembelajaran pada model *neural network* lainnya (Astuti, 2009).

Backpropagation merupakan salah satu algoritma pembelajaran dalam jaringan syaraf tiruan. Proses pembelajaran dalam *backpropagation* dilakukan dengan penyesuaian bobot-bobot jaringan syaraf tiruan dengan arah mundur berdasarkan nilai *error* dalam proses pembelajaran (Kusrini, 2009). Penemuan algoritma *backpropagation* untuk *multilayer perceptron (MLP)*, merupakan metode yang sistematis untuk *training* sehingga bisa dilakukan dan lebih efisien. Algoritma *backpropagation* berasal dari *learning rule* Widrow

dan Hoff, disusun oleh Werbos (1974), dibuat oleh Parker (1985), Rumelhart Hinton, Williams (Rumelhart dan Williams, 1986) dan peneliti lainnya (Maimon, 2005).

Algoritma pelatihan *backpropagation* terdiri dari dua tahapan yaitu *feedforward* dan *backpropagation* dari galatnya. Langkah pembelajaran dalam algoritma *backpropagation* adalah sebagai berikut (Kusumadewi, 2010):

- Inisialisasi bobot (ambil bobot awal dengan nilai random yang cukup kecil).
- Kerjakan langkah-langkah berikut selama kondisi berhenti bernilai FALSE:
 1. Untuk tiap-tiap pasangan elemen yang akan dilakukan pembelajaran, kerjakan:

Feedforward:

- a. Tiap-tiap unit input (X_i , $i=1,2,3,\dots,n$) menerima sinyal x_i dan meneruskan sinyal tersebut ke semua unit pada lapisan yang ada di atasnya (lapisan tersembunyi).
- b. Tiap-tiap unit tersembunyi ($Z_j, j=1,2,3,\dots,p$) menjumlahkan sinyal-sinyal input terbobot:

$$z_in_j = v_{0j} + \sum x_i v_{ij}$$
 Gunakan fungsi aktivitas untuk menghitung sinyal outputnya:

$$z_j = f(z_in_j)$$
 Kirimkan sinyal tersebut ke semua unit dilapisan atasnya (unit-unit output)
- c. Tiap-tiap unit output (Y_k , $k=1,2,\dots,m$) menjumlahkan sinyal-sinyal input terbobot

$$y_in_k = w_{0k} + \sum z_j w_{jk}$$
 Gunakan fungsi aktivitas untuk menghitung sinyal output:

$$y_k = f(y_in_k)$$
 Dan kirimkan sinyal tersebut ke semua unit dilapisan atasnya (unit-unit output).

Backpropagation

- d. Tiap-tiap unit output (Y_k , $k=1,2,\dots,m$) menerima target pola yang berhubungan dengan pola input pembelajaran, hitung informasi errornya:

$$\delta_k = (t_k - y_k) f'(y - in_k)$$
 Kemudian hitung koreksi bobot (yang nanti akan digunakan untuk memperbaiki nilai w_{jk}):

$$\Delta w_{jk} = \alpha \delta_k z_j$$
 Hitung juga koreksi bias (yang nantinya akan digunakan untuk memperbaiki nilai w_{0k}):
 Kirimkan δ_k ini ke unit-unit yang ada dilapisan bawahnya.
 - e. Tiap-tiap unit tersembunyi ($Z_j, j=1,2,3,\dots,p$) menjumlahkan delta inputnya (dari unit-unit yang berada pada lapisan di atasnya):

$$\delta_in_j = \delta_k w_{ij}$$
 Kalikan nilai ini dengan turunan dari fungsi aktivasi untuk menghitung informasi error:

$$\delta_j = \delta_in_j f'(z_in_j)$$
 Kemudian hitung koreksi bobot untuk memperbaiki nilai v_{ij} :

$$\Delta v_{ij} = \alpha \delta_j x_i$$
 Hitung juga koreksi bias untuk memperbaiki nilai v_{0j} :

$$\Delta v_{0j} = \alpha \delta_j$$
 - f. Tiap-tiap unit output (Y_k , $k=1,2,\dots,m$) memperbaiki bias dan bobotnya ($j=0,1,2,\dots,p$)

$$w_{jk}(\text{baru}) = w_{jk}(\text{lama}) + \Delta w_{jk}$$
 Tiap-tiap unit tersembunyi (Z_j , $j=1,2,\dots,p$) memperbaiki bias dan bobotnya ($i=0,1,2,\dots,n$).

$$v_{ij}(\text{baru}) = v_{ij}(\text{lama}) + \Delta v_{ij}$$
2. Tes kondisi berhenti

D. Evaluasi dan Validasi Model

Untuk mengukur akurasi model maka dilakukan evaluasi dan validasi menggunakan teknik:

1. Confusion matrix

Confusion Matrix adalah alat (*tools*) visualisasi yang biasa digunakan pada supervised learning. Tiap kolom pada matriks adalah contoh kelas prediksi, sedangkan tiap baris mewakili kejadian di kelas yang sebenarnya (Gorunescu, 2011). *Confusion matrix* berisi informasi aktual (*actual*) dan prediksi (*predicted*) pada sistem klasifikasi.

2. Kurva ROC (*Receiver Operating Characteristic*)

Kurva ROC menunjukkan akurasi dan membandingkan klasifikasi secara visual. ROC mengekspresikan *confusion matrix*. ROC adalah grafik dua dimensi dengan *false positives* sebagai garis horisontal dan *true positives* sebagai garis vertikal (Vecellis, 2009). *The area under curve* (AUC) dihitung untuk mengukur perbedaan performansi metode yang digunakan. AUC dihitung menggunakan rumus: (Liao, 2007)

$$\theta^r = \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m \psi(x_i^r, x_j^r)$$

Dimana

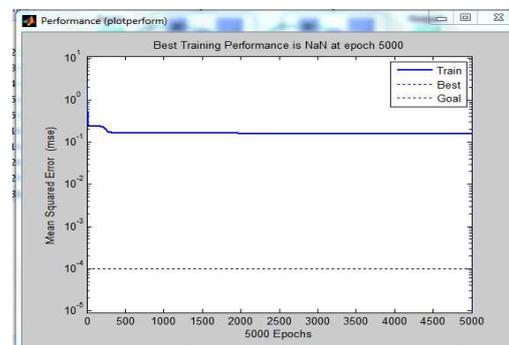
$$\psi(X, Y) = \begin{cases} 1 & Y < X \\ \frac{1}{2} & Y = X \\ 0 & Y > X \end{cases}$$

Performance keakuratan AUC dapat diklasifikasikan menjadi lima kelompok yaitu (Gorunescu, 2011):
 0.90 – 1.00 = *Exellent Clasification*
 0.80 – 0.90 = *Good Clasification*
 0.70 – 0.80 = *Fair Clasification*
 0.60 – 0.70 = *Poor Clasification*
 0.50 – 0.60 = *Failure*

HASIL DAN PEMBAHASAN

Pada penelitian ini data yang digunakan sebanyak 500 data kredit baik yang bermasalah maupun yang tidak bermasalah. Dari data tersebut, 80% data (400 data) digunakan sebagai data training, dan 20% data (100 data) sebagai data testing. Variabel input pada penelitian ini terdiri dari tujuh variabel, yaitu: 1. Status Perkawinan, 2. Jumlah Tanggungan, 3. Status Tempat Tinggal, 4. Pekerjaan, 5. Status Pekerjaan, 6. Penghasilan, 7. Uang Muka, Sedangkan variabel output adalah variabel Keputusan. Perangkat lunak yang digunakan untuk menganalisa adalah *matlab versi 7.7* dan *SPSS versi 17*.

Proses pelatihan jaringan syaraf tiruan menggunakan 80% dari total data, pada proses ini akan dilakukan pelatihan dengan arsitektur neural network yang berbeda-beda. Setiap arsitektur neural network akan menghasilkan bobot pelatihan yang terakhir yang akan digunakan sebagai bobot awal saat melakukan pengujian. Hasil pelatihan oleh *matlab versi 7.7* untuk arsitektur jaringan 7-2-1 diperlihatkan oleh gambar di bawah ini:



Gambar 2. Mean Square Error dengan arsitektur Jaringan 7-2-1

Mencari network terbaik dilakukan dengan cara mengubah jumlah neuron/node pada lapisan tersembunyi secara trial and error, maka di dapatkan konfigurasi terbaik. Berikut tabel Mean Square Error(MSE) dari beberapa variasi jaringan yang dilatih menggunakan matlab versi 7.7 pada data training dengan maximal epoch(iterasi) = 5000, learning rate (α) = 1, momentum = 0.8, dan target error = 0.0001:

Tabel 1. Hasil Pelatihan Arsitektur Jaringan

Arsitektur Jaringan	MSE
7-2-1	0,158
7-3-1	0,158
7-4-1	0,154
7-5-1	0,154
7-10-1	0,158
7-15-1	0,157
7-20-1	0,243
7-25-1	0,164
7-30-1	0,213

Berdasarkan tabel di atas, maka didapatkan konfigurasi jaringan terbaik sebagai berikut:

1. Jumlah neuron lapisan input = 7
2. Jumlah neuron lapisan tersembunyi = 4
3. Jumlah neuron lapisan output = 1
Dengan nilai MSE sebesar 0,154.

Arsitektur jaringan terbaik yang didapat saat pelatihan di atas akan digunakan sebagai arsitektur jaringan untuk mencari nilai keputusan kredit pada data testing. Bobot awal pada data testing adalah bobot terakhir saat pelatihan dari arsitektur jaringan terbaik.

untuk mengukur ketepatan dan keakuratan model dilakukan pengujian dengan *confusion matrix* dan kurva ROC sebagai berikut:

1. *Confusion Matrix*

Berikut nilai dari confusion matrix model neural network backpropagation:

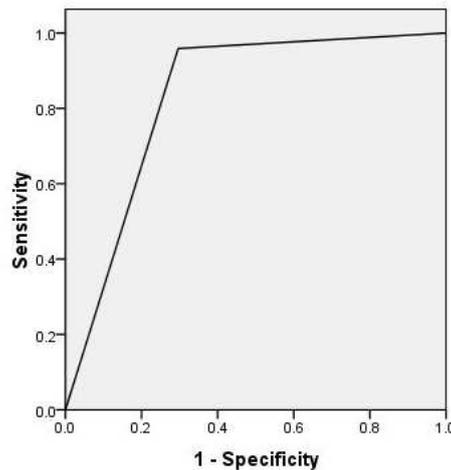
Tabel 2. *Confusion Matrix*

Accuracy : 89%			
	True Y	True N	Class Precision
Pred. Y	70	8	89,74%
Pred. N	3	19	86,36%
Class	95,89%	70,37%	
Recall			

Dari tabel di atas terlihat bahwa tingkat akurasi model sebesar 89%.

2. Kurva ROC

Berikut adalah grafik kurva ROC dengan menggunakan SPSS 17.



Gambar 3. Kurva ROC

Adapun nilai AUC (Area Under the Curva) sebesar 0,831. Seperti pada tabel berikut:

Tabel 3. *Area Under the Curva (AUC)*

Test Result Variable(s): y				
Area	Std. Error ^a	Asymptotic Sig. ^b	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.831	.055	.000	.723	.940

Model yang dihasilkan termasuk klasifikasi cukup baik karena memiliki nilai AUC antara 0.8-0.9, yaitu sebesar 0.831.

KESIMPULAN

Kesimpulan yang dapat diambil berdasarkan penelitian ini adalah bahwa performa model *neural network backpropagation* untuk pemberian kredit mobil yang dibentuk dari data training dan divalidasi pada data testing memberikan tingkat akurasi kebenaran sebesar 89% dengan nilai area under the curva (AUC) sebesar 0,831. Hal ini menunjukkan bahwa model yang dihasilkan termasuk katagori klasifikasi cukup baik karena memiliki nilai AUC antara 0.8-0.9.

DAFTAR PUSTAKA

- [1] Alpaydin, Ethem. (2010). *Introduction to Machine Learning*. London: The MIT Press.
- [2] Gorunescu, Florin (2011). *Data Mining: Concepts, Models, and Techniques*. Verlag Berlin Heidelberg: Springer
- [3] Han, J., & Kamber, M. (2006). *Data Mining Concept and Tehniques*. San Fransisco: Morgan Kauffman.
- [4] Kusrini, & Luthfi, E. T. (2009). *Algoritma Data Mining*. Yogyakarta: Andi Publishing.
- [5] Kusumadewi, Sri (2010). *Pengantar Jaringan Syaraf Tiruan*. Yogyakarta. Teknik Informatika FT UII.
- [6] Larose, D. T. (2005). *Discovering Knowledge in Data*. New Jersey: John Willey & Sons, Inc.
- [7] Maimon, Oded & Rokach, Lior. (2005). *Data Mining and Knowledge Discovey Handbook*. New York: Springer
- [8] Puspitaningrum, Diah (2006). *Pengantar Jaringan Syaraf Tiruan*. Yogyakarta: Andi Offset.
- [9] Sumathi, & S., Sivanandam, S.N. (2006). *Introduction to Data Mining and its Applications*. Berlin Heidelberg New York: Springer
- [10] Vercellis, Carlo (2009). *Business Intelligent: Data Mining and Optimization for Decision Making*. Southern Gate, Chichester, West Sussex: John Willey & Sons, Ltd.
- [11] Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining: Practical Machine Learning and Tools*. Burlington: Morgan Kaufmann Publisher.