
ANALISIS KLASIFIKASI CREDIT SCORING MENGGUNAKAN *WEIGHTED PROBABILISTIC NEURAL NETWORK* (WPNN)

¹Arief Rachman Hakim, ²Moch. Abdul Mukid, ³Hasbi Yasin, ⁴Sugito

^{1,2,3,4}Departemen Statistika, FSM Universitas Diponegoro

Email :Arief.rachman@live.undip.ac.id

ABSTRAK

Credit Scoring merupakan salah satu metode yang digunakan untuk menilai kelayakan dan memprediksi lebih awal adanya potensi kredit macet dari calon nasabah kredit. Dalam analisis kasifikasi Credit Scoring dapat menggunakan *Weighted Probabilistic Neural Network* (WPNN). Neural Network (NN) dikembangkan sebagai model matematika dengan prinsip kerja yang menyerupai pola pikir atau jaringan syaraf pada mahluk hidup. WPNN merupakan pengembangan daripada metode PNN dengan menambahkan faktor pembobot antara pattern layer dan summation layer. Metode ini memiliki kelebihan dalam mengatasi permasalahan yang terdapat pada *Back-Propagation* (BP) yaitu dapat mengatasi waktu pelatihan (training) yang lama, terjebak pada global minimum, serta sulitnya perancangan arsitektur jaringan. Pemilihan Klasifikasi dilakukan dengan melihat nilai *Apparent Error Rate* (APER) yang terkecil, yang dibagi kedalam nilai APER untuk Training dan Testing. Nilai APER training sebesar 0.0003 dan testing sebesar 0.0395 yang merupakan nilai APER terkecil, maka bobot yang terpilih dengan nilai bobot atau spread sebesar 0.8.

Kata kunci : Credit Scoring, Klasifikasi, WPNN, APER.

PENDAHULUAN

Dalam undang-undang Republik Indonesia nomer 10 tahun 1998, kredit diartikan sebagai suatu kegiatan 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 hutangnya setelah jangka waktu tertentu dengan pemberian bunga. Bentuk kredit juga ada dalam bentuk kredit kendaraan bermotor, praktik kredit kendaraan bermotor biasanya ditawarkan oleh banyak perusahaan-perusahaan pembiayaan atau leasing. Hal ini juga turut mengundang resiko, dimana debitur tidak mampu membayar angsuran atau dengan istilah kredit macet.

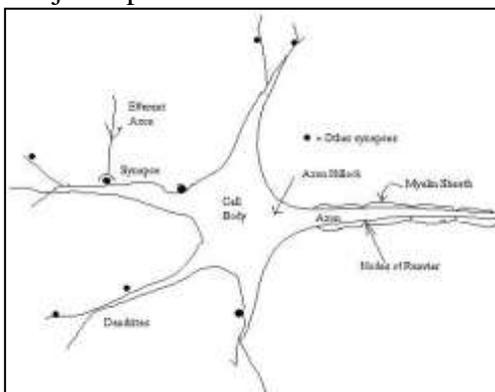
Credit Scoring merupakan salah satu metode yang digunakan untuk menilai kelayakan dari calon nasabah kredit. Penggunaan metode ini

diharapkan dapat memprediksi lebih awal adanya potensi kredit macet dari calon nasabah kredit tersebut. Metode yang dapat digunakan untuk mengklasifikasikan kondisi status kredit nasabah *Credit Scoring* dapat menggunakan *Neural Network* (NN). Salah satu model *Neural Network* yang digunakan untuk pengklasi-fikasian adalah *Weighted Probabilistic Neural Network* (WPNN). WPNN disusun dengan empat lapisan, yakni lapisan input, lapisan pola, lapisan penjumlahan dan lapisan output [1]. Pada penelitian ini dilakukan untuk menggambarkan penggunaan PNN dalam *Credit Scoring*.

Metode *credit scoring* merupakan salah satu metode statistika yang digunakan untuk memprediksi peluang calon nasabah tidak sanggup membayar kembali pinjaman, ataupun bagi nasabah kredit yang sudah terdaftar berpeluang menjadi nasabah yang pembayaran

pinjamannya menunggak [2]. Penggunaan metode *credit scoring* ini dapat memberikan informasi kepada pihak pemberi kredit atau pinjaman untuk menentukan apakah pengajuan kredit atau pinjaman dari peminjam dapat disetujui atau tidak. Data yang digunakan dalam *credit scoring* yaitu data aplikasi pinjaman nasabah, yang berisi tentang sejarah pinjaman yang bersangkutan[2].

Neural Network dikembangkan sebagai model matematika yang prinsip kerjanya menyerupai pola pikir manusia atau jaringan syaraf makluk hidup. NN biasa disebut juga dengan Jaringan Syaraf tiruan (JST) dimana system informasi yang memiliki karakteristik dengan jaringan syaraf pada mahluk hidup. NN memiliki asumsi bahwa sinyal yang melewati antar neuron menggunakan hubungan tertentu. Neuron sendiri merupakan proses informasi terjadi pada banyak elemen sederhana. Pada setiap penghubung antar neuron saling memiliki bobot (*weight*) yang bersesuaian, yaitu dengan mengalikan sinyal yang dikirimkan. Setiap neuron menggunakan fungsi aktivasi terhadap *input* (jumlah sinyal *input* yang terboboti) untuk menentukan sinyal *output*. Secara sederhana gambar sebuah system jaringan syaraf manusia disajikan pada Gambar 1.

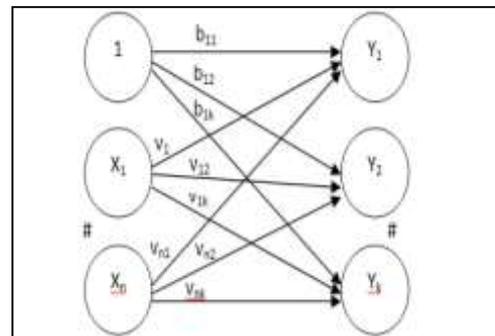


Gambar 1. Sistem Jaringan Syaraf Manusia [3].

NN memiliki dua tahap pemrosesan informasi yaitu tahap pelatihan (training) dan tahap pengujian (testing). Berdasarkan jumlah lapisan yang dimiliki *Neural Network* dapat dibedakan menjadi jaringan lapisan tunggal (*single layer*) dan jaringan multi lapis (*multi layer*).

Jaringan Lapisan Tunggal (*Single Layer*)

Sebuah jaringan disebut sebagai jaringan lapis tunggal jika jaringan tersebut tidak mempunyai lapisan tersembunyi atau hanya mempunyai satu lapisan bobot koneksi [4]. Pada jaringan ini sekumpulan *input* neuron dihubungkan langsung dengan sekumpulan *output* [5]. Arsitektur dari jaringan lapis tunggal digambarkan pada Gambar 2.

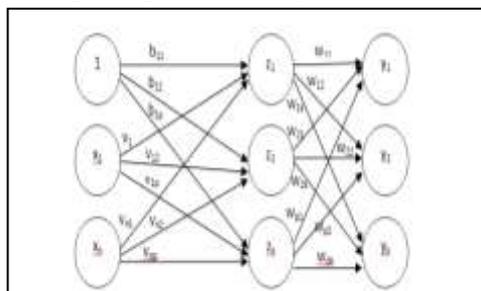


Gambar 2. Jaringan Lapis Tunggal dengan n Unit Masukan (X_1, X_2, \dots, X_n) dan k Unit Keluaran (Y_1, Y_2, \dots, Y_k).

Jaringan Multilapis (*Multi Layer*)

Jaringan multilapis merupakan perluasan dari jaringan *single layer* [5]. Jaringan multilapis terdiri dari lapisan *input* (*input layer*), lapisan tersembunyi (*hidden layer*) dan lapisan *output* (*output layer*). Lapisan tersembunyi terletak diantara lapisan *input* dan lapisan *output*. *Output* dari sebuah lapisan jaringan

tersembunyi akan menjadi *input* bagi lapisan berikutnya. Jaringan ini paling tidak mempunyai satu lapisan tersembunyi. Arsitektur jaringan multilapis digambarkan pada Gambar 3.



Gambar 3. Jaringan Multilapis dengan n Unit *Input* (x_1, x_2, \dots, x_n); p Unit *Tersembunyi* (z_1, z_2, \dots, z_p) dan k Unit *Output* (k_1, k_2, \dots, k_k).

Probabilistic Neural Network (PNN)

Probabilistic Neural Network (PNN) merupakan pengembangan algoritma NN dengan fungsi probabilistic dan juga merupakan algoritma klasifikasi. Kelebihan PNN sendiri dalam mengatasi permasalahan yang terdapat pada *Back-Propagation* (BP) yaitu dapat mengatasi waktu pelatihan (training) yang lama, terjebak pada global minimum, serta sulitnya perancangan arsitektur jaringan [6]. PNN juga kelemahan yaitu pada saat penentuan parameter *smoothing* dimana penentuan dilakukan dengan cara *trial and error* atau *user defined* [7].

Weighted Probabilistic Neural Network (WPNN)

WPNN merupakan pengembangan dari model PNN dimana dengan penambahan faktor pembobot antara lapisan pola (pattern layer) dan lapisan penjumlahan (summation layer). Algoritma yang digunakan dalam pengujian PNN, misalkan data pelatihan (training)

dinyatakan dalam bentuk matriks **M**, yang dinyatakan sebagai berikut [8]:

1. Pada tahapan ini ditetapkan: Matriks *input* **X** berukuran $q \times p$.

Vektor target **T** yang bersesuaian dengan matriks **M**, berukuran $n \times 1$,

dengan

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1j} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2j} & \dots & X_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{k1} & X_{k2} & \dots & X_{kj} & \dots & X_{kp} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{q1} & X_{q2} & \dots & X_{qj} & \dots & X_{qp} \end{bmatrix}$$

dimana $k =$ banyaknya

pengamatan data *input* (1, 2, ..., q)
 $j =$ banyaknya variabel data *input* (1, 2, ..., p)

2. Kemudian pada tahap inialisasi :

Bobot lapisan *input* : $\mathbf{V} = \mathbf{M}$

Bobot bias lapisan *input*:

$$b = \frac{\sqrt{-\ln(0.5)}}{\sigma} \text{ dengan } 0 < \sigma \leq 1$$

Bobot lapisan *output*: $\mathbf{W} = \mathbf{T}$.

dengan

$$\mathbf{V} = \begin{bmatrix} V_{11} & V_{12} & \dots & V_{1j} & \dots & V_{1p} \\ V_{21} & V_{22} & \dots & V_{2j} & \dots & V_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ V_{i1} & V_{i2} & \dots & V_{ij} & \dots & V_{ip} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ V_{n1} & V_{n2} & \dots & V_{nj} & \dots & V_{np} \end{bmatrix};$$

$$\mathbf{W} = \begin{bmatrix} W_{1h} \\ W_{2h} \\ \vdots \\ W_{ih} \\ \vdots \\ W_{nh} \end{bmatrix}$$

dimana $i =$ banyaknya baris pada bobot *input* (1, 2, ..., n),
 $h =$ kelas data yang bersesuaian dengan data *training* (1, 2, ..., r).

3. Kemudian untuk $k=1$ sampai q pada masing-masing poin, pada semua data *training* dan *testing*:

- Mencari jarak X_k terhadap $V_i, i=1, 2, \dots, n$ sebagai berikut:

$$D_{ki} = \sqrt{\sum_{j=1}^p (X_{kj} - V_{ij})^2}$$

- Mencari aktivasi $a1_{ki}$ sebagai berikut:

$$a1_{ki} = e^{-(b_i D_{ki})^2}; \text{ dengan } i = 1, 2, \dots, n$$

- Mencari $a2_{kh}$ sebagai berikut:

$$a2_{kh} = \sum_{i=1}^n (a1_{ki} * W_{ih}); \text{ dengan } h = 1, 2, \dots, r$$

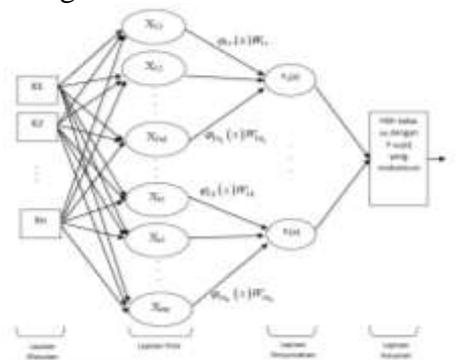
- Mencari z sedemikian hingga

$$a2_{kz} = \max(a2_{kh} | h = 1, 2, \dots, r)$$

- Menetapkan z sebagai kelas dari X_k asli
- Dengan W_{ij} merupakan matriks perkalian antara jarak mahalanobis

4. Melakukan perhitungan akurasi dari klasifikasi, salah satu metode akurasi yang sederhana adalah metode *holdout*.

Arsitektur dari WPNN secara umum dapat digambarkan pada Gambar 4 sebagai berikut :



Gambar 4. Arsitektur Jaringan WPNN

Ketepatan Klasifikasi

Klasifikasi diharapkan dapat melakukan pengelompokan atau

klasifikasi semua data set dengan benar, namun pada kenyataannya system ini tidak terhindar dari kesalahan. Salah satu bentuk kesalahannya dalam mengklasifikasikan objek baru ke dalam suatu kelas (*missclassification*).

Menurut Prasetyo [9], matriks konfusi merupakan tabel pencatat hasil kerja klasifikasi. Tabel 1 merupakan matriks konfusi yang melakukan klasifikasi masalah biner. Setiap sel f_{ij} dalam matriks menyatakan jumlah *record* (data) dari kelas i yang hasil prediksinya masuk ke kelas j . Misalnya, sel f_{11} adalah jumlah data dalam kelas A yang secara benar dipetakan ke kelas A, dan f_{10} adalah data dalam kelas A yang dipetakan secara salah ke kelas B.

Tabel 1. Matriks Konfusi

f_{ij}	Kelas Hasil Prediksi (j)	
	Kelas = A	Kelas = B
Kelas = A	f_{11}	f_{10}
Kelas = B	f_{01}	f_{00}

Berdasarkan isi matriks konfusi, dapat diketahui jumlah data dari masing-masing kelas yang diprediksi secara benar, yaitu $(f_{11} + f_{00})$, dan data yang diklasifikasikan secara salah, yaitu $(f_{10} + f_{01})$. Maka dapat dihitung tingkat akurasi dan tingkat kesalahan prediksi atau *Apparent Error Rate* (APER):

1. Akurasi = $\frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}}$
2. *Apparent Error Rate* (APER) = $\frac{f_{10} + f_{01}}{f_{11} + f_{10} + f_{01} + f_{00}}$

METODELOGI PENELITIAN

Sumber Data dan Variabel Penelitian

Sumber data yang digunakan dalam penelitian ini adalah data kredit dari pembiayaan kendaraan bermotor dari

suatu perusahaan pada tahun 2017, dengan keterangan data sebagai berikut:

Tabel 2 Variabel data

Variabel	Keterangan
X ₁	Jumlah pinjaman
X ₂	Penghasilan
X ₃	Pengeluaran
X ₄	Lama bekerja
X ₅	Jumlah besar cicilan
Y	Status kredit yaitu lancar atau macet

Metode analisis

Pada penelitian ini untuk running data menggunakan bantuan software matlab, kemudian langkah-langkah penelitian sebagai berikut :

1. Input data status kredit nasabah sebagai variabel repon sedangkan usia debitur, lama bekerja, jumlah pendapatan bersih, jumlah pinjaman lain, jumlah nominal dalam akun tabungan, dan rasio hutang terhadap pendapatan dimasukkan sebagai variabel prediktor.
2. Melakukan Analisis deskriptif data nasabah / debitur berdasarkan status kredit.
3. Membagi data menjadi dua bagian yaitu data *training*, dan data *testing* dengan proporsi 90:10
4. Menentukan nilai σ (*smoothing parameter*) secara *trial and error*
5. Menghitung nilai APER dari data training dan testing.
6. Membuat kalsifikasi berdasarkan model WPNN
7. Membuat prediksi klasifikasi credit scoring.

HASIL DAN PEMBAHASAN

Analisis deskriptif

Data yang digunakan merupakan data kredit pembiayaan kendaraan

bermotor di suatu perusahaan. Secara deskriptif di tampilkan pada tabel 3.

Tabel 3. Statistics descriptive

Variable	status	Mean	Maximum
Jumlah Pinjaman	lancar	9450362	85000000
	macet	8416119	23973824
Penghasilan	lancar	5710065	120000000
	macet	4901628	30000000
Pengeluaran	lancar	2806512	30000000
	macet	2586667	20000000
Lama Bekerja	lancar	10,845	50
	macet	9,019	35
Cicilan	lancar	644945	7000000
	macet	580483	1860000

Tabel 3 menjelaskan variabel-variabel yang digunakan yaitu jumlah pinjaman, penghasilan pengeluaran, lama bekerja, dan cicilan. Kemudian variabel tersebut dikelompokkan berdasarkan status kredit lancar dan macet.

Variabel jumlah pinjaman dengan status kredit lancar memiliki rata rata sebesar Rp. 9.450.362,- dan pinjaman dengan status kredit macet memiliki rata rata Rp. 8.416.119,-. Variabel penghasilan dengan status kredit lancar memiliki rata rata sebesar Rp. 5.710.065,- dan dengan status kredit macet memiliki rata rata penghasilan Rp. 4.901.628,-. Kemudian pada variabel pengeluaran merupakan besarnya pengeluaran bulanan yang dikeluarkan nasabah, pada variabel ini kredit dengan status lancar memiliki rata-rata pengeluaran sebesar Rp. 2.806.512,- dan status macet memiliki rata-rata pengeluaran sebesar Rp. 2.586.667,-. Variabel lama bekerja dengan status kredit lancar memiliki rata rata sebesar 10,8 bulan dan status kredit macet memiliki rata rata lama bekerja 9 bulan. Variabel yang kelima yaitu variabel cicilan menerangkan besarnya jumlah angsuran yang harus

dibayarkan oleh nasabah kepada perusahaan secara rutin, dengan status kredit lancar memiliki rata-rata cicilan sebesar Rp. 644.945,- dan dengan status kredit macet memiliki rata-rata cicilan sebesar Rp 580,483,-.

Apabila secara keseluruhan persentase banyaknya kredit lancar sebanyak 95% dan kredit dengan status macet 5 % .

Klasifikasi kredit dengan WPNN

Pada penelitian ini pertama data dibagi menjadi dua yaitu training dan testing, dengan pembagian 90 % untuk data training dan 10% untuk data testing, serta kelas yang dilatih adalah 2. Kemudian klasifikasi dilakukan dengan menentukan parameter-parameter bobot yang optimal, yaitu menggunakan nilai *Apparent Error Rate (APER)* terkecil.

Tabel 4. Hasil akurasi dan nilai APER.

spread	Akurasi		APER	
	training	testing	training	testing
0.2	0.9994	0.9579	0.0006	0.0421
0.4	0.9994	0.9526	0.0006	0.0474
0.6	0.9994	0.9553	0.0006	0.0447
0.8	0.9997	0.9605	0.0003	0.0395
1	0.9991	0.9579	0.0009	0.0421

Tabel 4 menjelaskan besarnya nilai akurasi atau ketepatan klasifikasi serta nilai APER, baik dari data training maupun data testing. Apabila melihat nilai APER training dan testing terkecil maka bobot yang terpilih dengan nilai bobot atau spread sebesar 0.8. Berikut gambar 5 merupakan output untuk bobot atau spread sebesar 0.8 :

```

=====
Analisis klasifikasi dengan Model WPNN
=====
Tanggal Analisis   :02-Jan-2018 11:56:44
Sigma              = 8.000000e-001
Persentase Data Latih = 0.90
Banyaknya Kelas   = 2
-----
Klasifikasi dengan Data Latih
-----
APER              = 0.0003
Akurasi           = 0.9997
Matriks Konfusi  =
3269              0
1                 162
-----
Klasifikasi dengan Data Uji
-----
APER              = 0.0395
Akurasi           = 0.9605
Matriks Konfusi  =
363              0
15              2
    
```

Gambar 5 output WPNN untuk bobot 0.8

Untuk klasifikasi data latih di peroleh nilai matriks konfusi yang artinya data diasumsikan awalnya kredit lancar kemudian hasil klasifikasinya dengan menggunakan WPNN diperoleh bahwa kredit lancar juga yaitu sebanyak 3269, kemudian data diasumsikan awalnya lancar kemudian hasil klasifikasinya dengan menggunakan WPNN diperoleh bahwa kredit macet yaitu sebanyak 0. Data yang diasumsikan keditnya lancar namun hasilnya klasifikasinya dengan WPNN ternyata kredit macet ada sebanyak 1, lalu data yang di asumsikan macet dan ternyata hasil klasifikasi macet ada sebanyak 162.

Klasifikasi data uji atau data testing dapat dilihat pada matriks konfusi gambar 5 yang artinya data diasumsikan awalnya lancar kemudian hasil klasifikasinya di peroleh lancar juga yaitu sebanyak 363, kemudian data diasumsikan awalnya lancar kemudian hasil klasifikasinya di peroleh ternyata macet juga yaitu sebanyak 0. Data yang diasumsikan keditnya lancar namun hasilnya klasifikasinya ternyata kredit macet ada sebanyak 15, lalu data yang di asumsikan macet dan ternyata hasil klasifikasi macet ada sebanyak 2.

Prediksi Klasifikasi Analisis Kredit Scoring

Resiko kredit dari seorang nasabah dapat diprediksikan, bahwa kredit tersebut akan berpotensi kredit lancar atau kredit macet. Hal ini tentunya juga harus didukung dengan data yang ada seperti variabel X_1, X_2, X_3, X_4, X_5 serta jaringan hasil pelatihan dari WPNN sendiri. Misalkan seseorang dengan jumlah pinjaman (X_1) sebesar Rp. 20.000.000,- , penghasilan (X_2) sebesar Rp. 4.000.000,- perbulan, pengeluaran (X_3) sebesar Rp. 3.000.000,- perbulan, Lama bekerja (X_4) sebanyak 12 bulan, dan dengan cicilan (X_5) sebesar Rp.

1.500.000,- perbulan, maka dari hasil prediksi analisis credit scoring diperoleh bahwa status “**kredit dalam perhatian khusus**” hal ini mengindikasikan bahwa seseorang nasabah tersebut berdasarkan kondisinya saat ini, masih berpeluang untuk mengalami kredit macet kedepannya. Prediksi ini juga bisa dilakukan menggunakan contoh yang berbeda berdasarkan keadaan masing masing variabel. Berikut Gambar 6 hasil output dari hasil analisis menggunakan program Gui MATLAB :

Gambar 6. Prediksi analisis kredit scoring nasabah

KESIMPULAN

Berdasarkan analisis deskriptif didapatkan hasil bahwa secara keseluruhan persentase banyaknya kredit lancar sebanyak 95% dan kredit dengan status macet 5 %. Melihat nilai APER training sebesar 0.0003 dan testing sebesar 0.0395 merupakan nilai APER terkecil, maka bobot yang terpilih dengan nilai bobot atau spread sebesar 0.8.

Klasifikasi data *training* di peroleh bahwa data diasumsikan awalnya kredit lancar kemudian hasil klasifikasinya kredit lancar, ada sebanyak 3269, data yang diasumsikan awalnya lancar kemudian hasil klasifikasinya diperoleh bahwa kredit macet yaitu sebanyak 0. Data yang diasumsikan keditnya lancar namun hasilnya klasifikasinya kredit macet, ada sebanyak 1, lalu data yang di asumsikan macet dan ternyata hasil klasifikasi macet ada sebanyak 162. Klasifikasi data *testing* yaitu data diasumsikan awalnya lancar kemudian hasil klasifikasinya diperoleh lancar, ada sebanyak 363, kemudian data yang diasumsikan awalnya lancar kemudian hasil klasifikasinya diperoleh ternyata macet, ada sebanyak 0. Data yang diasumsikan keditnya lancar namun hasilnya klasifikasinya ternyata kredit macet ada sebanyak 15, lalu data yang di asumsikan macet dan ternyata hasil klasifikasi macet ada sebanyak 2.

Prediksi klasifikasi credit scoring yang dilakukan dengan simulasi seorang nasabah dengan X_1 sebesar Rp. 20.000.000,- , X_2 sebesar Rp. 4.000.000,- perbulan, X_3 sebesar Rp. 3.000.000,- X_4 sebanyak 12 bulan, dan X_5 sebesar Rp. 1.500.000,- perbulan, diperoleh bahwa status “**kredit dalam perhatian khusus**” yang artinya ada indikasi bahwa seseorang nasabah tersebut berdasarkan kondisinya saat ini, masih berpeluang untuk mengalami kredit macet kedepannya.

DAFTAR PUSTAKA

- [1] Mishra, M., Jena, A. R., Das, Raja. 2013, A Probabilistic Neural Network Approach For Classification of Vehicle. *International Journal of Application or Innovation in Engineering & Management* Vol. 2, No. 7 : 367-371.

- [2] Mester, L. J. 1997. *What's The Point of Credit Scoring? Business Review (September)* 3-16.
- [3] Warsito, B. 2009. *Kapita Selekt Statistika Neural Network*. Semarang: BP Undip.
- [4] Fausset, L. 1994. *Fundamental of Neural Network: Architecture, Algorithm, and Application*. New Jersey: Prentice-Hall.
- [5] Siang, J. J. 2005. *Jaringan Syaraf Tiruan*. Yogyakarta: Penerbit Andi.
- [6] Specht, D.F. 1990. *Probabilistic Neural Networks*. Neural Networks Vol. 3: 109-118.
- [7] Yang, H. 2012. An Improved Probabilistic Neural Network with GIA Optimization. *International Conference on Intelligent Computation Technology and Automation*: 76-79.
- [8] Yasin, H. 2017. Klasifikasi Data Berat Bayi Lahir Menggunakan *Weighted Probabilistic Neural Network (WPNN)*. Media Statistika vol 10 (1): 61-70.
- [9] Prasetyo, E. 2012. *Data Mining: Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi Publisher.