

**PERANCANGAN MODEL REGRESI LOGISTIK DAN
NEURAL NETWORK UNTUK MEMPREDIKSI
RESIKO KREDIT DI KREDIT X**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat guna mencapai gelar
Sarjana dalam bidang ilmu Teknik Industri

Disusun oleh:

Nama : Arvin Hermawan

NPM : 2013610099



**PROGRAM STUDI TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG
2018**

**FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG**



Nama : Arvin Hermawan
NPM : 2013610099
Jurusan : Teknik Industri
Judul Skripsi : PERANCANGAN MODEL REGRESI LOGISTIK DAN *NEURAL NETWORK* UNTUK MEMPREDIKSI RESIKO KREDIT DI KREDIT X

TANDA PERSETUJUAN SKRIPSI

Bandung, 15 Januari 2018

Ketua Jurusan Teknik Industri

(Dr. Carles Sitompul, S.T., M.T., M.I.M.)

Pembimbing Pertama

(Sani Susanto, Ph.D.)

Pembimbing Kedua

(Titi Iswari, S.T., M.Sc., M.B.A.)



Jurusan Teknik Industri
Fakultas Teknologi Industri
Universitas Katolik Parahyangan



Pernyataan Tidak Mencontek atau Melakukan Tindakan Plagiat

Saya, yang bertanda tangan di bawah ini,

Nama : Arvin Hermawan

NPM : 2013610099

dengan ini menyatakan bahwa skripsi dengan judul :

“PERANCANGAN MODEL REGRESI LOGISTIK DAN *NEURAL NETWORK* UNTUK MEMPREDIKSI RESIKO KREDIT DI KREDIT X”

adalah hasil pekerjaan saya dan seluruh ide, pendapat atau materi dari sumber lain telah dikutip dengan cara penulisan referensi yang sesuai.

Pernyataan ini saya buat dengan sebenar-benarnya dan jika pernyataan ini tidak sesuai dengan kenyataan, maka saya bersedia menanggung sanksi yang akan dikenakan kepada saya.

Bandung, 15 Januari 2018



Arvin Hermawan
2013610099

ABSTRAK

Berdasarkan Cinthia (2014), PT. X merupakan perusahaan yang bergerak pada jasa pemberian kredit. Data-data yang didapatkan dari pengajuan kredit oleh debitur belum diolah secara optimal untuk memprediksi resiko kredit. Toleransi perusahaan untuk kredit macet adalah 3% dari seluruh debitur, tetapi debitur yang mengalami kemacetan sebesar 9,3% dari seluruh debitur. Penelitian Cinthia (2014) telah melakukan klasifikasi untuk melakukan prediksi terhadap resiko status kredit seorang calon debitur pada PT. X. Metode CART dan C 4.5 digunakan dalam penelitian tersebut. Klasifikasi tersebut dapat secara langsung melakukan klasifikasi apabila variabel prediktor memenuhi syarat-syarat tertentu. Tetapi, dalam penelitian tersebut tidak dilakukan pemisahan set data untuk pemodelan dan evaluasi sehingga tidak dapat diketahui apakah model yang dibuat sesuai digunakan dengan data lain atau tidak.

Dalam penelitian ini digunakan metode regresi logistik dan *neural network* dengan harapan akan menghasilkan performansi yang lebih baik untuk melakukan prediksi resiko kredit dengan lebih baik sehingga dapat meningkatkan kinerja sistem. Metode ini juga digunakan karena lebih relevan untuk digunakan dalam kasus melakukan prediksi resiko kredit karena dapat melihat kecenderungan data. Dalam penelitian ini juga dilakukan perbaikan pemisahan set data untuk pemodelan dan evaluasi.

Dari hasil penelitian ini dapat diketahui bahwa pemisahan set data perlu dilakukan untuk menguji pemodelan yang dilakukan terhadap set data lain. Dari hasil penelitian ini juga didapatkan tingkat akurasi pada tahap evaluasi dari regresi logistik sebesar 87% dan metode *neural network* sebesar 75%. Jadi pemodelan dengan akurasi yang lebih baik dalam melakukan prediksi resiko kredit di PT. X adalah model regresi logistik.

ABSTRACT

According to Cinthia (2014) PT. X is a creditor company. Data that obtained from submission by debtor is not optimally used for predicting credit risk. The company give tolerance about 3% of all the debtor can have bad credit, but in reality there were 9,3% debtor who had bad credit. Research by Cinthia (2014) had made classification for predicting credit risk in PT. X. CART and C 4.5 method were used in that research. Both of the method can directly classify if the terms were met. But, the research were not using data set separation for modeling and evaluating phase, so there were not known if the model can be used for another data set.

In this research logistic regression and neural network method were used with expectation can generate better performance on predicting credit risk so the model can be used for improving system performance. These methods is also more relevant to use in this case because can observe data trend. There is also improvement in this research that is separation in data set for modeling and evaluating.

From this research can be known that separation in data set is use for testing the model in another data for checking the performance. Accuracy level in evaluation phase using regression logistic model were 87% and using neural network model were 75%. So, regression logistic model were better on predicting credit risk in PT. X.

KATA PENGANTAR

Puji dan syukur kepada Tuhan Yang Maha Esa karena dengan kasih karunia dan rahmat-Nya maka dapat terselesaikan laporan skripsi yang berjudul “Perancangan Model Regresi Logistik dan *Neural Network* untuk Memprediksi Resiko Kredit di PT. X”. Dalam penyusunan laporan ini tentunya disadari banyak kekurangan dan keterbatasan. Maka dari itu penulis akan sangat berterima kasih apabila pembaca memiliki kritik dan saran untuk laporan skripsi ini. Tak lepas juga laporan skripsi ini tidak akan tersusun apabila tidak ada pihak-pihak yang berkontribusi dan mendukung dalam penyusunannya. Penulis mengucapkan terima kasih sebesar-besarnya kepada:

1. Bapak Sani Susanto. Ph.D selaku dosen pembimbing yang telah bersedia meluangkan waktu, tenaga, dan pikiran untuk membantu memberikan masukan, pembelajaran, dan nasihat dari awal penyusunan hingga selesainya penyusunan laporan skripsi ini.
2. Ibu Titi Iswari, S.T., M.Sc., M.B.A. selaku dosen pembimbing yang telah bersedia meluangkan waktu, tenaga, dan pikiran untuk membantu memberikan masukan, pembelajaran, dan nasihat dari awal penyusunan hingga selesainya penyusunan laporan ini.
3. Bapak Dr. Carles Sitompul, S.T., M.T., M.I.M selaku koordinator skripsi yang telah mengkoordinasi skripsi agar berjalan dengan baik.
4. Bapak Fransiscus Rian Praktikto, S.T., M.T., M.S. selaku dosen penguji proposal skripsi yang telah meluangkan waktunya untuk menguji proposal sehingga menghasilkan masukan yang bermanfaat dalam penyusunan laporan skripsi ini.
5. Bapak Hanky Fransiscus, S.T., M.T. selaku dosen penguji sidang skripsi yang telah meluangkan waktunya dalam melakukan sidang skripsi sehingga didapatkan masukan untuk perbaikan penelitian.
6. Ibu Cherish Rikardo, S.Si., M.T. selaku dosen penguji proposal skripsi dan penguji sidang skripsi yang telah meluangkan waktunya untuk menguji proposal dan dalam sidang skripsi sehingga menghasilkan masukan yang bermanfaat dalam penyusunan laporan skripsi ini.

7. Bapak Alfian, S.T., M.T. selaku koordinator laboratorium analisis perancangan kerja dan ergonomi karena telah mengizinkan peminjaman ruangan untuk melakukan seminar skripsi.
8. Seluruh staf pengajar Teknik Industri Unpar karena telah memberikan ilmu-ilmu yang diharapkan akan berguna bagi penulis.
9. Seluruh staf Teknik Industri Unpar karena telah memberikan sarana prasarana bagi penulis untuk menuntut ilmu selama kuliah.
10. Keluarga penulis yang selalu memberikan dukungan moril dan materil kepada penulis selama perkuliahan.
11. Teman-teman penulis yang tidak bisa disebutkan satu per satu yang telah memberikan dukungan, dorongan, dan semangat serta masukan selama pengerjaan penyusunan laporan skripsi ini.
12. Segenap staf Tata Usaha Teknik Industri Universitas Katolik Parahyangan Bandung.

Penulis mohon maaf apabila terdapat hal-hal yang kurang berkenan bagi pembaca. Mohon juga dimaafkan apabila terdapat kesalahan, kekeliruan, atau kekurangan yang terdapat pada laporan skripsi ini. Semoga laporan ini dapat bermanfaat bagi pembaca.

Bandung, 15 Januari 2018

Arvin Hermawan
(2013610099)

DAFTAR ISI

ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR TABEL	ix
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	I-1
I.1 Latar Belakang Masalah	I-1
I.2 Identifikasi dan Perumusan Masalah.....	I-2
I.3 Pembatasan Masalah dan Asumsi Penelitian	I-5
I.4 Tujuan Penelitian	I-6
I.5 Manfaat Penelitian	I-6
I.6 Metodologi Penelitian.....	I-6
I.7 Sistematika Penulisan.....	I-9
BAB II TINJAUAN PUSTAKA	II-1
II.1 Kredit	II-1
II.1.1 Tujuan Kredit.....	II-1
II.1.2 Fungsi Kredit.....	II-2
II.1.3 Jenis Kredit	II-3
II.1.4 Prinsip Pemberian Kredit.....	II-4
II.1.3 Kredit Lancar dan Resiko Kredit.....	II-5
II.2 Pengertian <i>Data Mining</i>	II-5
II.3 CRISP-DM	II-5
II.4 Skala Pengukuran.....	II-6
II.5 Uji Korelasi.....	II-7
II.5.1 <i>Coefficient Contingency Pearson C</i>	II-8
II.5.2 <i>Somers' d_{yx}</i>	II-8
II.5.3 <i>Pearson r</i>	II-10
II.5.4 <i>Spearman rho</i>	II-10

II.5.5 <i>Eta</i>	II-10
II.6 Klasifikasi.....	II-11
II.7 Regresi Logistik	II-11
II.8 <i>Neural Network</i>	II-13
II.9 Pengukuran Performansi Model Klasifikasi	II-16
BAB III PENGUMPULAN DAN PENGOLAHAN DATA.....	III-1
III.1 Pemahaman Penelitian	III-1
III.2 Pemahaman Data	III-2
III.3 Persiapan Data	III-6
III.3.1 Uji Korelasi Nominal–Nominal	III-8
III.3.2 Uji Korelasi Ordinal–Ordinal.....	III-11
III.3.3 Uji Korelasi Interval–Interval	III-13
III.3.4 Uji Korelasi Nominal–Ordinal	III-15
III.3.5 Uji Korelasi Nominal–Interval.....	III-18
III.3.6 Uji Korelasi Ordinal–Interval	III-21
III.3.7 Hasil Uji Korelasi Semua Variabel	III-23
III.3.8 Uji Korelasi Variabel Terpilih dengan Variabel Target.....	III-23
III.3.8.1 Uji Korelasi Ordinal Terpilih – Variabel Target.....	III-24
III.3.8.2 Uji Korelasi Interval Terpilih – Variabel Target	III-25
III.4 Pemodelan.....	III-26
III.4.1 Pemodelan Regresi Logistik	III-27
III.4.2 Pemodelan <i>Neural Network</i>	III-28
III.5 Evaluasi Model.....	III-32
III.5.1 Evaluasi Model Regresi Logistik	III-32
III.5.2 Evaluasi Model <i>Neural Network</i>	III-36
III.5.3 Perbandingan Performansi Model Regresik Logistik dan <i>Neural Network</i> dengan Penelitian Sebelumnya	III-36
BAB IV ANALISIS.....	IV-1
IV.1 Analisis Variabel dan Skala Pengukuran Variabel.....	IV-1
IV.2 Analisis Persiapan Data.....	IV-3
IV.3 Analisis Pemodelan dan Evaluasi Regresi Logistik	IV-4
IV.4 Analisis Pemodelan dan Evaluasi <i>Neural Network</i>	IV-6
IV.5 Analisis Perbandingan dengan Penelitian Sebelumnya	IV-8

BAB V KESIMPULAN DAN SARAN.....	V-1
V.1 Kesimpulan.....	V-1
V.2 Saran.....	V-1
DAFTAR PUSTAKA	
LAMPIRAN	
RIWAYAT HIDUP PENULIS	

DAFTAR TABEL

Tabel I.1	Contoh Data PT. X Alamat - DP	I-3
Tabel II.1	Tabel Ilustrasi <i>Crosstab Somers' d_{yx}</i>	II-8
Tabel III.1	Tabel Suku Bunga Berdasarkan Tipe Barang dan Tenor	III-3
Tabel III.2	Tabel Skala Pengukuran Variabel	III-6
Tabel III.3	Tabel Data Nomor <i>Customer</i> untuk <i>Training Set</i>	III-7
Tabel III.4	Tabel Data Nomor <i>Customer</i> untuk <i>Test Set</i>	III-7
Tabel III.5	Tabel Data Nomor <i>Customer</i> Terdapat pada <i>Training Set dan Test Set</i>	III-8
Tabel III.6	Tabel Hasil <i>Crosstab</i> Variabel Tipe Barang – Jenis Kelamin	III-9
Tabel III.7	Tabel Hasil <i>Chi Square</i> Variabel Nominal Tipe Barang – Jenis Kelamin	III-9
Tabel III.8	Tabel Hasil Korelasi Variabel Nominal Tipe Barang – Jenis Kelamin	III-10
Tabel III.9	Rekapitulasi Korelasi Nominal–Nominal	III-10
Tabel III.10	Tabel Hasil <i>Crosstab</i> Variabel Ordinal Suku Bunga – Tenor	III-12
Tabel III.11	Tabel Hasil Korelasi Variabel Suku Bunga – Tenor	III-12
Tabel III.12	Rekapitulasi Korelasi Ordinal–Ordinal	III-13
Tabel III.13	Tabel Contoh Hasil Korelasi Interval – Interval	III-13
Tabel III.14	Rekapitulasi Korelasi Interval–Interval	III-14
Tabel III.15	Tabel Hasil <i>Crosstab</i> Variabel Tipe Barang – Tenor	III-16
Tabel III.16	Tabel Hasil <i>Chi Square</i> Variabel Tipe Barang – Tenor	III-16
Tabel III.17	Tabel Hasil Korelasi Variabel Tipe Barang – Tenor	III-17
Tabel III.18	Rekapitulasi Korelasi Nominal–Ordinal	III-17
Tabel III.19	Tabel Hasil Korelasi Variabel Tipe Barang – DP	III-19
Tabel III.20	Rekapitulasi Korelasi Nominal–Interval	III-19
Tabel III.21	Tabel Hasil Korelasi Variabel Tenor dengan DP dan Rasio	III-21
Tabel III.22	Rekapitulasi Korelasi Ordinal–Interval	III-22
Tabel III.23	Rekapitulasi Variabel Prediktor Regresi Logistik	III-23
Tabel III.24	Tabel Hasil Korelasi Variabel Status Kredit – Tenor	III-24

Tabel III.25 Tabel Hasil Korelasi Variabel Pendidik Terakhir – Status Kredit	III-24
Tabel III.26 Tabel Hasil Korelasi Variabel Tanggungan – Status Kredit	III-25
Tabel III.27 Tabel Hasil Korelasi Variabel Kondisi Rumah – Status Kredit	III-25
Tabel III.28 Tabel Hasil Korelasi DP – Status Kredit	III-25
Tabel III.29 Tabel Rekapitulasi Hasil Korelasi Interval Terpilih – Variabel Target.....	III-26
Tabel III.30 Tabel Hasil Regresi Logistik Variabel Terpilih	III-27
Tabel III.31 Tabel Hasil Model Regresi Logistik Variabel Rasio	III-27
Tabel III.32 Tabel Akurasi Model Regresi Logistik Terhadap <i>Training Set</i>	III-28
Tabel III.33 Tabel Keakurasian <i>Neural Network Training Set 1 Node</i> <i>Hidden Layer</i>	III-30
Tabel III.34 Tabel Hasil Klasifikasi <i>Neural Network Training Set 1 Node</i> <i>Hidden Layer</i>	III-30
Tabel III.35 Tabel Keakurasian <i>Neural Network Training Set 2 Node</i> <i>Hidden Layer</i>	III-30
Tabel III.36 Tabel Hasil Klasifikasi <i>Neural Network Training Set 2 Node</i> <i>Hidden Layer</i>	III-31
Tabel III.37 Tabel Keakurasian <i>Neural Network Training Set 3 Node</i> <i>Hidden Layer</i>	III-31
Tabel III.38 Tabel Hasil Klasifikasi <i>Neural Network Training Set 3 Node</i> <i>Hidden Layer</i>	III-31
Tabel III.39 Tabel Perhitungan Regresi Logistik <i>Test Set</i>	III-33
Tabel III.40 Tabel Akurasi Pemodelan <i>Neural Network 2 Node Test Set</i>	III-36

DAFTAR GAMBAR

Gambar I.1	Metodologi Penelitian.....	I-8
Gambar II.1	Ilustrasi Model <i>Neural Network</i>	II-14
Gambar III.1	Gambar Hasil <i>Neural Network 2 Node Hidden Layer</i>	III-29

DAFTAR LAMPIRAN

- Lampiran A Data Mentah
- Lampiran B Uji Korelasi Nominal - Nominal
- Lampiran C Uji Korelasi Ordinal - Ordinal
- Lampiran D Uji Korelasi Interval - Interval
- Lampiran E Uji Korelasi Nominal - Ordinal
- Lampiran F Uji Korelasi Ordinal - interval
- Lampiran G Hasil *Training Set Neural Network*
- Lampiran H Hasil *Test Set Neural Network*

BAB I

PENDAHULUAN

Pada bab ini akan dijabarkan mengenai latar belakang masalah, identifikasi masalah, pembatasan masalah dan asumsi penelitian, manfaat penelitian, metodologi penelitian, serta sistematika penulisan.

I.1 Latar Belakang Masalah

Menurut Turner (1993) teknik industri merupakan sebuah ilmu yang berdasarkan pada beberapa ilmu. Ilmu-ilmu tersebut kemudian digunakan pada teknik industri untuk merancang, memperbaiki, mengembangkan, dan mengaplikasikan sistem yang terintegrasi. Sistem terintegrasi yang dimaksud adalah apabila sistem tersebut melibatkan manusia, material, peralatan, informasi, dan energi. Ilmu-ilmu yang menjadi dasar teknik industri adalah antara lain matematika, fisika, ilmu sosial, dan manajemen.

Informasi merupakan bagian dari sistem yang menjadi bagian teknik industri, maka dari itu teknik industri juga mempelajari mengenai teknologi informasi. Teknologi informasi dalam teknik industri berguna untuk memecahkan suatu permasalahan yang terjadi dalam sistem. Selain dapat memecahkan suatu masalah, pengembangan dari suatu sistem juga dapat dilakukan dengan mengaplikasikan ilmu teknologi informasi.

Menurut Larose dan Larose (2015) *data mining* merupakan salah satu kajian teknologi informasi yang dapat digunakan untuk menyelesaikan masalah atau mengembangkan performansi dari suatu sistem. Dalam *data mining*, data yang dimiliki diolah untuk dijadikan suatu informasi yang berguna bagi pengguna data. Dalam proses *data mining*, data yang diolah merupakan data yang berjumlah cukup banyak.

Data mining memiliki fungsi klasifikasi. Klasifikasi dapat digunakan untuk mengklasifikasikan variabel target yang kategorikal dengan mempertimbangkan beberapa variabel prediktor. Klasifikasi dalam *data mining* dapat membantu perusahaan untuk mengetahui resiko dari suatu kejadian berdasarkan beberapa variabel jika data terlalu banyak atau variabelnya juga terlalu banyak. Jika

klasifikasi tidak dilakukan maka perusahaan akan sulit mengetahui resiko dari suatu kejadian dan akan lebih beresiko untuk mengalami kerugian. Dalam klasifikasi terdapat beberapa algoritma yang dapat digunakan untuk melakukannya dan dapat disesuaikan dengan kasus yang terjadi.

Salah satu cara untuk mengembangkan suatu sistem adalah memperbaiki sistem tersebut atau melakukan pembaharuan dengan metode yang lebih baik. Penelitian ini dilakukan untuk memperbaiki proses. Pengolahan menggunakan metode yang lebih relevan. Dari hasil algoritma metode-metode yang berbeda akan didapatkan hasil dengan nilai performansi berbeda. Dari hasil performansi itu akan dapat dilihat metode klasifikasi apa yang cocok diaplikasikan untuk suatu kasus.

Klasifikasi dalam *data mining* berguna juga untuk mengetahui peluang atau kecenderungan dalam suatu perusahaan pemberi jasa kredit apabila seorang calon debitur akan mengalami pembayaran yang terhambat atau lancar berdasarkan pengolahan data historis. Klasifikasi digunakan agar mengetahui berdasarkan beberapa variabel dapat ditentukan kemungkinan yang akan terjadi. Objek dari penelitian ini merupakan penelitian yang dilakukan oleh Cinthia (2014). Dalam penelitian tersebut, klasifikasi dilakukan untuk membantu perusahaan dalam melakukan prediksi terhadap calon debitur apabila akan mengalami pembayaran yang lancar atau macet.

PT. X ini tentunya memiliki banyak nasabah sehingga data yang dimiliki cukup banyak. Oleh karena itu penelitian tersebut menggunakan metode klasifikasi dalam masalah tersebut untuk mengetahui resiko kredit dari seorang calon debitur. Terdapat metode yang berbeda – beda dalam metode klasifikasi sehingga metode yang telah digunakan akan mungkin dapat ditingkatkan dan diperbaiki apabila terdapat metode dengan performansi yang lebih baik.

I.2 Identifikasi dan Perumusan Masalah

Objek penelitian adalah penelitian yang dilakukan sebelumnya oleh Cinthia (2014). Dalam penelitian tersebut diteliti resiko status kredit pada PT. X. PT. X merupakan perusahaan yang menawarkan jasa kredit yang terdapat di Bandar Jaya, Lampung Tengah. Sebelum melakukan pengajuan kredit, terdapat beberapa data yang harus diisi terlebih dahulu. Sebelum perusahaan melakukan pemberian keputusan untuk memberikan kredit atau tidak tentunya perusahaan

harus terlebih dahulu melakukan pertimbangan. Data yang dimiliki oleh debitur tentunya dapat membantu perusahaan untuk pengambilan keputusan tersebut sehingga dapat meminimalisir penerimaan calon debitur yang cenderung akan mengalami kemacetan dalam pembayaran.

Data debitur yang dimiliki oleh PT. X tidak sedikit dan bervariasi maka tentunya akan mengalami kesulitan untuk melakukan penentuan peluang apakah seorang debitur akan mengalami kemacetan dalam pembayaran atau tidak. Data yang dimiliki oleh PT. X cukup banyak maka *data mining* dapat digunakan untuk mengolah data tersebut menjadi informasi yang berguna. Fungsi klasifikasi dalam *data mining* dapat digunakan untuk menggolongkan berdasarkan data yang ada, apakah suatu data tersebut tergolong dalam suatu hasil tertentu atau tidak.

Klasifikasi pada PT. X memiliki variabel target berupa apakah kredit dari debitur lancar atau macet. Ketika akan mengajukan kredit tentunya seorang calon debitur mengisi banyak data untuk mendaftarkannya sebagai debitur. Ada kemungkinan data tersebut dapat membantu perusahaan untuk memprediksi seseorang apakah akan mengalami kredit macet atau tidak. Hal tersebut tentunya dapat membantu perusahaan untuk mengurangi resiko kerugian oleh perusahaan. Tabel I.1 merupakan contoh data yang didapatkan dari debitur dari data alamat hingga pembayaran DP yang dilakukan. Seluruh data mentah dapat dilihat pada Lampiran A.

Tabel I.1 Contoh Data PT. X Alamat – DP

Customer ke-	Alamat	Dealer/Toko	Nama Barang	Tipe Barang	Suku Bunga	Harga Barang	DP
1	BANDARJAYA TIMUR	STANDAR FURNITURE	KORSI ROTAN	4	0,0255	5500000	
2	LK VI Bandar Jaya	Meubel Jati Makmur	Rak TV	4	0,0255	2300000	
3	DUSUN ANOMAN	C-SHOP	LAPTOP T/L 74513 TOSHIBA	3	0,0280	5400000	
...
332	SUMBER AGUNG MATARAM	PUTRA BARU	KTM ROMAWI JATI	4	0,0255	10000000	1000000
333	PONCOWATI	PUTRA BARU	KTM SOKO MATAHARI	4	0,0255	4300000	

(Cinthia, 2014)

Penelitian oleh Cinthia (2014) dilakukan karena perusahaan memiliki dari jumlah semua debitur hanya boleh 3% saja yang mengalami kredit macet, tetapi dalam kenyataannya terdapat 9,3% debitur yang memiliki kredit macet. Hal tersebut tentunya cukup signifikan dari batas toleransi yang diperbolehkan dan

juga menjadi landasan penelitian tersebut. Penelitian sebelumnya dilakukan dengan menggunakan metode klasifikasi untuk dapat mengetahui data yang ada dapat memperkirakan apakah seseorang cenderung akan mengalami kemacetan pembayaran kredit atau tidak. Namun dalam penelitian tersebut proses pemodelan dan evaluasi yang dilakukan menggunakan set data yang sama.

Menurut Wendler dan Grottrup (2016) untuk dilakukan pemodelan klasifikasi, sebuah set data perlu dilakukan evaluasi terhadap set data yang berbeda untuk mengecek bagaimana model tersebut dapat memprediksi dari variabel data yang bukan hasil dari pemodelannya. Apabila data tersebut dievaluasi dengan menggunakan data yang sama maka akan cenderung terjadi *overfitting* atau kecenderungan sebuah model mengikuti pola data dari hasil pemodelan. Ketika pemodelan tersebut terjadi kasus seperti itu maka dikhawatirkan akan mempengaruhi performansi proses prediksi apabila terdapat data baru untuk dilakukan klasifikasi.

Dalam penelitian ini dalam langkah pemodelan akan dilakukan perbaikan dengan memisahkan set data yang ada menjadi set data untuk pemodelan dan satu set data untuk mengevaluasi data yang sudah dimodelkan. Klasifikasi pada penelitian ini diharapkan dapat membantu perusahaan untuk menentukan apakah akan memberikan kredit atau tidak kepada debitur.

Penelitian yang dilakukan oleh Cinthia (2014) sebelumnya dilakukan dengan menggunakan metode *decision tree* CART dan C 4.5. Hasil keakurasian dari metode tersebut cukup tinggi sebesar 92,5% untuk CART dan 90,7% untuk metode C 4.5. Seperti yang sebelumnya dipaparkan, akan tetapi keakurasian tersebut belum dievaluasi dengan set data yang baru atau berbeda dengan yang digunakan dalam pemodelan.

Menurut Wendler dan Grottrup (2016) dalam suatu permasalahan dapat dilakukan klasifikasi dengan menggunakan beberapa metode. Metode tersebut belum tentu dipastikan metode mana yang memiliki performansi yang lebih baik. Hal tersebut harus diujikan dan dievaluasi terlebih dahulu untuk menentukan manakah metode yang cocok digunakan dalam situasi tersebut.

Dalam penelitian ini akan membandingkan hasil pemodelan klasifikasi dengan menggunakan perbaikan pemisahan data set. Pemodelan klasifikasi juga akan menggunakan metode yang berbeda untuk dijadikan pembanding dari model

yang didapatkan sebelumnya. Metode yang digunakan adalah regresi logistik dan *neural network*.

Metode regresi logistik dan *neural network* juga digunakan dalam penelitian ini karena cenderung lebih relevan dengan keadaan yang sebenarnya. Dalam penggunaan *decision tree*, klasifikasi tersebut langsung menentukan apakah seorang calon debitur akan macet atau tidak. Dalam kenyataannya seorang calon debitur hanya memiliki kecenderungan saja apakah akan mengalami kredit macet atau tidak.

Metode regresi logistik digunakan karena menurut Wendler dan Grottrup (2016) regresi logistik memiliki kelebihan untuk menghasilkan model yang akan memiliki keakurasian yang cukup tinggi. Dalam penentuan status kredit tentunya perusahaan akan mengharapkan prediksi yang akurat sehingga metode regresi logistik cocok digunakan untuk pemodelan klasifikasi.

Metode *neural network* menurut Wendler dan Grottrup (2016) dapat mengatasi hubungan kompleks dari variabel dikarenakan sistem pembelajaran dari *network* yang terus dilakukan pembaharuan dengan lebih tepat. Data yang dimiliki cukup banyak dan cenderung akan mengalami hubungan yang kompleks maka metode ini juga akan digunakan untuk pemodelan klasifikasi.

Berdasarkan identifikasi masalah yang telah dijelaskan, maka dibuatlah beberapa rumusan masalah seperti berikut:

1. Bagaimana perbedaan performansi akurasi apabila pemodelan dan evaluasi dilakukan dengan menggunakan set data yang berbeda?
2. Bagaimana perbandingan model regresi logistik dan *neural network* dalam memprediksi resiko kredit pada PT. X?

I.3 Pembatasan Masalah dan Asumsi Penelitian

Dalam penelitian ini didapati beberapa batasan dan asumsi. Batasan dan asumsi ini dibuat agar penelitian tidak keluar dari konteks penelitian dan dapat menjawab tujuan diadakannya penelitian. Batasan dalam penelitian ini adalah:

1. Data yang digunakan adalah data historis PT. X tahun 2012-2013 sesuai dengan penelitian Cinthia (2014).
2. Penelitian dilakukan dengan metode *data mining*.
3. Tahapan pemodelan klasifikasi dilakukan dengan menggunakan bantuan *software* SPSS.

4. Penelitian hanya dilakukan hingga tahap evaluasi dari pemodelan. Terdapat pula asumsi yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan tidak berubah.
2. Kondisi dan persyaratan dari PT. X tidak berubah.

I.4 Tujuan Penelitian

Berdasarkan identifikasi masalah tersebut maka dapat dijadikan beberapa tujuan dari penelitian. Adapaun tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut:

1. Mengetahui perbedaan performansi akurasi apabila pemodelan dan evaluasi dengan menggunakan set data yang berbeda.
2. Mengetahui perbandingan model regresi logistik dan *neural network* dalam memprediksi resiko kredit pada PT. X.

I.5 Manfaat Penelitian

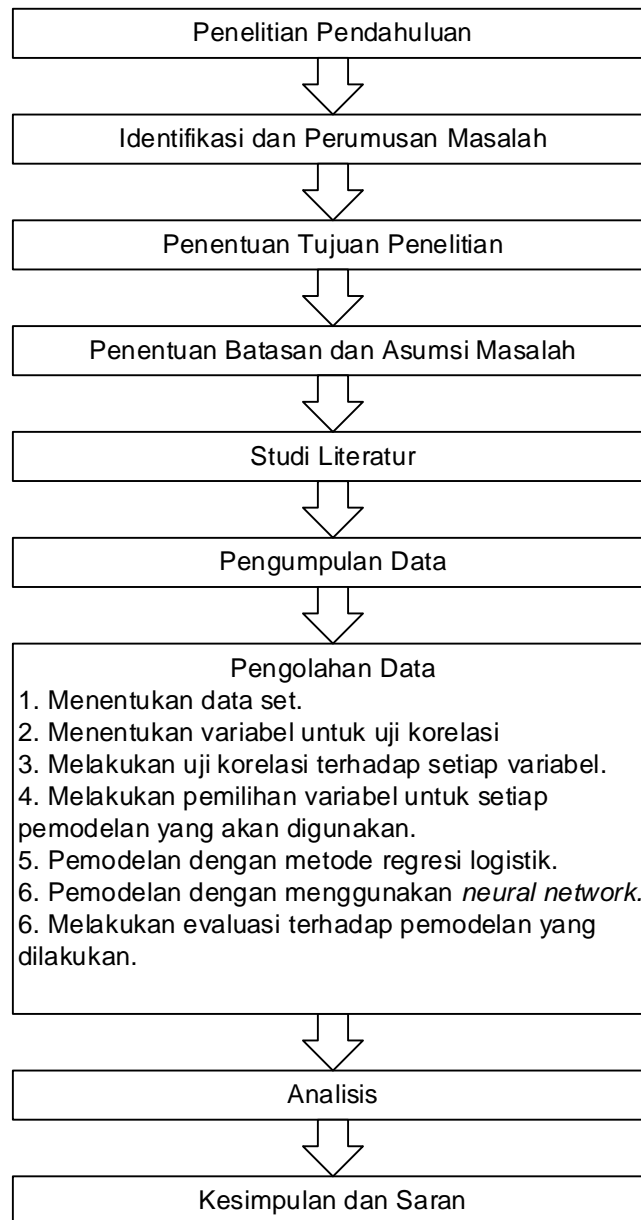
Penelitian ini diharapkan memberikan manfaat baik bagi peneliti dan pembaca. Berikut merupakan hal-hal yang diharapkan dapat menjadi manfaat dari hasil penelitian ini.

1. Peneliti dapat melakukan implementasi tahapan – tahapan dan ilmu-ilmu yang telah diperoleh dalam bidang *data mining*.
2. Peneliti lebih dapat memahami metode yang akan digunakan dalam *data mining* beserta dengan algoritma yang digunakan dalam metode tersebut.
3. Pembaca dapat menambah ilmu mengenai *data mining* dan mengetahui proses yang harus dilaksanakan dalam *data mining*.
4. Penelitian dapat dijadikan referensi atau bahan pembelajaran apabila ingin melakukan penelitian.

I.6 Metodologi Penelitian

Pada bab ini akan dibahas mengenai metodologi penelitian yang akan dilakukan. Gambar I.1 merupakan tahapan metodologi penelitian yang akan dilakukan. Berikut merupakan penjelasan dari tahapan metodologi penelitian yang akan dilakukan.

1. **Penelitian Pendahuluan**
Di dalam tahapan ini penelitian pendahuluan dilakukan dengan meneliti terlebih dahulu objek penelitian yang akan diteliti. Kemudian dilakukan identifikasi masalah terhadap permasalahan yang terjadi pada objek penelitian atau dalam kasus ini adalah penelitian oleh Cinthia (2014).
2. **Identifikasi dan Perumusan Masalah**
Setelah melakukan penelitian pendahuluan maka dilakukan identifikasi masalah terhadap objek penelitian tersebut. Setelah masalah ditentukan kemudian dilakukan perumusan masalah terhadap objek penelitian, yaitu penelitian Cinthia (2014).
3. **Penentuan Tujuan Penelitian**
Dalam tahapan ini akan dilakukan tujuan dari penelitian. Tujuan penelitian ini didapatkan dari identifikasi dan perumusan masalah yang sebelumnya sudah dilakukan identifikasi.
4. **Penentuan Batasan dan Asumsi Masalah**
Dalam tahapan ini dilakukan batasan dan asumsi dengan masalah yang terkait. Batasan dan asumsi ini dilakukan dengan tujuan memfokuskan penelitian terhadap tujuan permasalahan.
5. **Studi Literatur**
Setelah tujuan penelitian ditentukan maka selanjutnya akan dilakukan pembelajaran terhadap literatur terkait yang berhubungan dengan permasalahan. Studi literatur juga berguna untuk memperkuat penelitian karena telah teruji berdasarkan landasan teori oleh ahli.
6. **Pengumpulan Data**
Pada tahap ini data dikumpulkan sesuai yang dibutuhkan dengan tujuan penelitian dan yang diperlukan untuk melakukan untuk pengolahan data. Data yang dikumpulkan sesuai dengan penelitian Cinthia (2014).



Gambar I.1 Metodologi Penelitian

7. Pengolahan Data

Tahapan dalam pengolahan data yang pertama merupakan pemisahan set data untuk pemodelan dan evaluasi. Kemudian dilakukan persiapan data untuk pemodelan agar dapat dilakukan pemodelan dengan baik. Persiapan tersebut merupakan penentuan variabel, melakukan uji korelasi, dan penentuan variabel. Sehingga dapat dilakukan pemodelan berdasarkan data yang telah disiapkan. Kemudian akan dilakukan

evaluasi terhadap model yang sudah dibuat dan akan dilakukan perbandingan apabila set data dilakukan pemisahan serta perbandingan dari kedua model yang sudah dibuat.

8. Analisis

Setelah data diolah akan dilakukan analisis terhadap penelitian yang telah dilakukan. Analisis akan berisi mengenai data yang akan digunakan, persiapan data, pemodean data, dan hasil evaluasi dari proses-proses tersebut.

9. Kesimpulan dan Saran

Setelah penelitian dilakukan akan ditutup dengan kesimpulan dan saran. Kesimpulan akan berisi mengenai jawaban dari tujuan penelitian dan saran akan berisi mengenai saran untuk penelitian lanjutan.

I.7 Sistematika Penulisan

Dalam penelitian ini sistematikan penulisan dibuat dengan tujuan untuk memudahkan penyampaian dari penelitian dengan membagi penelitian menjadi beberapa bagian. Berikut merupakan sistematika penulisan dari penelitian ini.

BAB I PENDAHULUAN

Dalam bab ini akan berisi mengenai latar belakang masalah dari penelitian, identifikasi masalah dari penelitian dan perumusan masalah, batasan dan asumsi permasalahan, tujuan penelitian, manfaat penelitian, metodologi penelitian, dan juga sistematikan penulisan.

BAB II TINJAUAN PUSTAKA

Bab II ini akan berisi mengenai teori yang digunakan dalam penelitian ini. Teori yang akan dipaparkan adalah *data mining*, CRISP-DM, skala pengukuran, uji korelasi, klasifikasi, regresi logistik, *neural network*, dan pengujian performansi model klasifikasi.

BAB III PENGUMPULAN DAN PENGOLAHAN DATA

Bab ini akan berisi pengumpulan data yang didapatkan dari Cinthia (2014). Kemudian dilanjutkan dengan pemahaman data dan pemilihan variabel yang mungkin untuk dijadikan variabel prediktor dari model klasifikasi. Setelah itu

dilakukan pembagian data set secara acak. Pengujian korelasi dilakukan untuk melihat variabel yang saling tidak berkorelasi tinggi untuk dilakukan pemilihan variabel yang akan digunakan untuk proses pemodelan. Lalu dilakukan pengujian korelasi terhadap variabel target. Pemodelan akan dilakukan dari kedua metode, yaitu dengan regresi logistik dan *neural network*. Setelah itu dilakukan evaluasi terhadap pemodelan tersebut.

BAB IV ANALISIS

Bab ini akan berisi analisis dari penelitian yang telah dilakukan. Analisis tersebut akan mengenai data yang akan digunakan, persiapan data yang dilakukan, pemodelan dengan menggunakan regresi logistik, pemodelan dengan menggunakan *neural network*, dan evaluasi dari hasil pemodelan tersebut.

BAB V KESIMPULAN DAN SARAN

Bab ini akan berisi mengenai kesimpulan dari penelitian yang telah dilakukan. Kesimpulan ini merupakan jawaban dari tujuan penelitian yang telah dilakukan. Terdapat juga saran dari penelitian yang telah dilakukan.