Jurnal Sains Komputer & Informatika (J-SAKTI)
Volume (1) No. 1 Maret 2017
ISSN:2548-9771/EISSN: 2549-7200
http://ejurnal.tunasbangsa.ac.id/index.php/jsakti

PENERAPAN*K-NEAREST NEIGHBOR* BERBASIS *GENETIC ALGORITHM* UNTUKPENENTUAN PEMBERIAN KREDIT

Ester Arisawati
Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri
(STMIK Nusa Mandiri)
Jl. Kramat Raya No.18 Jakarta Pusat
http://www.nusamandiri.ac.id
esterarisawati@yahoo.com

Abstrak

Pembiayaan konsumen adalah kegiatan pembiayaan untuk pengadaan barang berdasarkan kebutuhan konsumen dengan pembayaran secara angsuran. Sedangkan Perusahaan Pembiayaan adalah badan usaha yang khusus didirikan untuk melakukan sewa guna usaha, anjak piutang, pembiayaan konsumen, dan atau usaha kartu kredit. Perusahaan pembiayaan akan menyetujui kredit yang diajukan konsumen setelah melakukan analisa kredit terhadap kelayakan pemberian pembiayaan konsumen, apakah disetujui dan tidak disetujui.Dalam proses analisa terhadap konsumen, terdapat beberapa yang tidak akurat, oleh karena itu konsumen tidak mampu membayar dengan tepat waktu yang mengakibatkan kredit macet. Untuk mengatasi permasalahan yang ada diperlukan suatu model yang mampu mengklasifikasikan dan memprediksi data konsumen yang bermasalah dan tidak bermasalah. Dalam penelitian ini dilakukan pengujian yaitu *k-Nearest Neighbor* dan *k-Nearest Neighbor* yang dioptimasi *Genetic Algorithm* yang diaplikasikan terhadap data konsumen yang mendapat pembiayaan kredit baik yang konsumen yang bermasalah maupun tidak. Dari hasil pengujian dengan mengukur kinerja ketiga algoritma tersebut menggunakan metode pengujian *Cross Validation*, *Confusion Matrix* dan Kurva ROC, diketahui bahwa algoritma *k-Nearest Neighbor* yang dioptimasi *Genetic Algorithm* memiliki nilai *accuracy* dan AUC paling tinggi.

Keywords*:* Kredit, *K-Nearest Neighbor*, Optimasi, *Genetic Algorithm*

Abstract

Consumer financing is financing activities for the procurement of goods based on the needs of consumers with payment in installments. While the Financing Company is a business entity specifically set up to conduct leasing, factoring, consumer finance, or business credit card. The finance company will approve the proposed consumer credit after a credit analysis of the feasibility of providing consumer financing, if approved and not disetujui.Dalam analysis process for consumers, there are some that are not accurate, therefore consumers can not afford to pay in a timely manner resulting in bad debts , To solve the problem we need a model that is able to classify and predict consumer data is problematic and not

problematic. In this research, testing ie k-Nearest Neighbor and k-Nearest Neighbor optimized genetic algorithm is applied to the data consumer that gets better the consumer credit financing is problematic or not. From the test results by measuring the performance of the three algorithms using Cross Validation testing methods, Confusion Matrix and ROC curves, it is known that the k-Nearest Neighbor algorithm optimized Genetic Algorithm has the AUC value and highest accuracy.

***Keywords**: Credit*, *K-Nearest Neighbor*, Optimasi, *Genetic Algorithm*

1. PENDAHULUAN

Penting bagi bank dan lembaga pembiayaan untuk mengevaluasi resiko kredit dilakukan dimuka bagi konsumen. Sebuah model yang baik bagi penilaian kredit akan membantu bank dan lembaga pembiayaan membuat keputusan yang tepat dalam rangka untuk menghindari potensi besarnya resiko[7] Zhang, Hifi, Chen, & Ye, 2008.

Penilaian kredit sebagai teknik penilaian yang sangat instrumen penting dalam industri keuangan dan perbankan[6] Wang, Lai, & Niu, 2011. Penilaian kredit telah menjadi isu yang sangat penting karena pertumbuhan terbaru dari industri kredit, sehingga kredit departemen bank menghadapi sejumlah besar data kredit konsumen untuk proses, tetapi tidak mungkin menganalisis ini sejumlah besar data baik dalam hal ekonomi dan tenaga kerja. Dalam studi ini kami terakhir karya-karya yang telah diterapkan metode data mining dalam masalah risiko kredit evaluasi [1] Keramati, & Yousefi, 2011. KNN memiliki kelebihan antara lain yaitu ketangguhan terhadap training data yang memiliki banyak noise dan efektif apabila training data-nya besar, proses mudah direpresentasikan dibandingkan dengan metode lain [4] Nanja & Purwanto, 2015.

Genetic Algorithm kompleks dan adaptif biasanya digunakan dalam memecahkan kuatmasalah optimasi. Pada dasarnya, mereka melibatkan bekerja dengan populasi individudi mana setiap individu merupakan potensi (optimal) solusi, dan masing-masing populasi adalahhimpunan bagian dari ruang pencarian keseluruhan[3] Matic, 2010.

2. METODOLOGI PENELITIAN

Metodologi penelitian yang digunakan penulis dalam penelitian eksperimen ini dengan menggunakan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM) terdiri dari enam tahap yang merupakan sebuah proses siklis yaitu [2] Larose, 2005 :

a. *Business Understanding* (Pemahaman Bisnis) adalah:
   Pemahaman bisnis meliputi penetapan tujuan bisnis, penilaian situasi terkini, penetapan tujuan bisnis, penetapan tujuan penggalian data, dan pengembangan rencana proyek.
b. *Data Understanding* (Pemahaman Data) adalah:
   Begitu tujuan bisnis dan rencana proyek ditetapkan, pemahaman data mempertimbangkan data yang dibutuhkan. Langkah ini bisa meliputi

pengumpulan data awal, deskripsi data, eksplorasi data, dan verfikasi kualitas data. Eksplorasi data seperti peninjauan statistik rangkuman (yang meliputi tampilan visual variabel-variabel kategorik) bisa terjadi pada akhir tahap ini. Model-model seperti analisis pengelompokan (*cluster analysis*) dapat pula diterapkan dalam tahap ini, dengan tujuan mengidentifikasi pola dalam data tersebut.

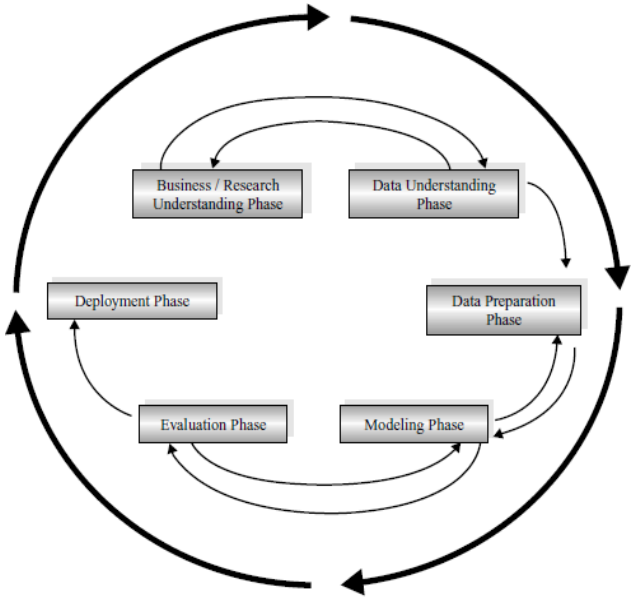
c. *Data Preparation* (Persiapan Data) adalah:
Setelah sumber data yang tersedia diidentifikasi, sumber data tersebut perlu diseleksi, dibersihkan, dibangun ke dalam wujud yang dikehendaki dan dibentuk. Pembersihan dan transformasi data dalam persiapan model data perlu dilakukan pada tahap ini. Eksplorasi data secara lebih mendalam juga dapat diterapkan dalam tahap ini, dan penggunaan model-model tambahan sekali lagi memberikan peluang untuk meihat berbagai pola berdasarkan pemahaman bisnis.

d. *Modeling* (Pembuatan Model) adalah:
Metode penggalian data, seperti visualisasi (penggambaran data dan penetapan hubungan) serta analisis pengelompokan (untuk mengidentifikasi variabel mana yang berhubungan satu sama lain) bermanfaat bagi analisis awal. Alat bantu seperti induksi aturan yang digeneralisasikan dapat mengembangkan aturan-aturan asosiasi awal. Begitu pemahaman data yang lebih luas diperoleh (sering kali melalui pengenalan pola yang dipicu dengan melihat output model), model-model lebih terinci yang sesuai dengan jenis data tersebut dapat diterapkan. Pembagian data ke dalam data latihan dan data uji juga diperlukan untuk pembuatan model.

e. *Evaluation* (Evaluasi) adalah:
Hasil model sebaiknya dievaluasi dalam konteks tujuan bisnis yang ditetapkan pada tahap awal (pemahaman bisnis). Hal ini akan mengarahkan pada identifikasi kebutuhan lain (kerap kali melalui pengenalan pola), sering kali kembali ke tahap-tahap CRISP-DM sebelumnya. Perolehan pemahaman bisnis merupakan prosedur berulang dalam penggalian data, di mana hasil dari beragam visualisasi, fakta statistik, dan metode kecerdasaan buatan menunjukan hubungan-hubungan baru kepada pengguna yang memberikan pemahaman yang lebih mendalam mengenai operasi perusahaan.

f. *Deployment* (Pelaksanaan) adalah:
Pengalian data dapat digunakan baik untuk membuktikan hipotesis sebelumnya, ataupun untuk penemuan pengetahuan (pengidentifikasian hubungan yang tidak terduga dan bermanfaat), Melalui pengetahuan yang ditemukan dalam tahap awal proses CRISP-DM, model yang kuat dapat diperoleh yang mungkin kemudian dapat diterapkan pada kegiatan bisnis untuk berbagai keperluan, termasuk memprediksi atau mengidentifikasikan situasi-situasi penting. Model-model ini perlu dipantau untuk mengawasi adanya perubahan dalam operasi, karena apa yang mungkin tepat untuk saat ini mungkin tidak lagi tepat satu tahun ke depan. Jika perubahan besar benar-benar terjadi, model tersebut sebaiknya dibuat ulang. Merupakan hal yang bijaksana

untuk mencatat hasil proyek penggalian data agar bukti-bukti yang terdokumentasi tersedia untuk penelitian di masa mendatang.

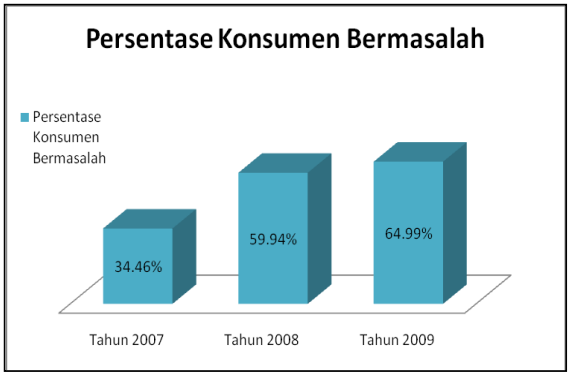


Gambar 1. Proses *Data Mining* menurut CRISP-DM

2.1. Pemahaman Bisnis

Bagian penting dari suatu penelitian penggalian data adalah dimana mengetahui untuk apa penelitian tersebut dilakukan. Berawal dari kebutuhan manajerial akan pengetahuan baru, menentukan tujuan akhir, dan membuat rencana untuk mendapatkan pengetahuan seperti itu perlu dikembangkan, berkenan dengan mereka yang bertanggung jawab untuk menggumpulkan data, menganilis data, dan membuat laporan.

Dengan mengumpukan data konsumen kredit yang didapat dari PT AEON Credit Service Indonesia cabang Tangerang diketahui bahwa jumlah konsumen bermasalah tiap tahunnya meningkat. Jumlah konsumen bermasalah tahun 2007 adalah 419 dari 1216 konsumen, tahun 2008 adalah 585 dari 976 konsumen dan tahun 2009 adalah 310 dari 477 konsumen. Dari data tahun 2007 sampai tahun 2009 didapat tingkat tingginya persentasi kredit macet yang menjadi permasalahan.



Sumber: PT AEON Credit Service Indonesia (2009)

Gambar 2. Grafik Peningkatan Presentase Konsumen Bermasalah

2.2. Pemahaman Data

Penggalian data berorientasi pada tugas, tugas bisnis yang berbeda membutuhkan kelompok data yang berbeda pula. Hal yang pertama dilakukan dalam proses penggalian data adalah memilih data yang berkaitan dari banyak *database* yang tersedia untuk menggambarkan pemahaman bisnis yang diberikan dengan tepat. Ada tiga hal yang penting untuk dipertimbangkan dalam pemilihan data. Hal yang pertama adalah menetapkan deskripsi masalah dengan ringkas dan jelas. Hal yang kedua adalah mengidentifikasi data yang relevan untuk mendeskripsikan masalah. Dan yang terakhir hal yang ketiga adalah variabel yang terpilih untuk data yang relevan sebaiknya independen satu sama lain.

Berdasarkan *database* data yang didapat adalah jenis data kuanlitatif (*quantitative data*) menggunakan nilai numerik. Data tersebut berupa data diskret (bilangan bulat) atau kontinu (bilangan rill). Yang berisi dari sejumlah data-data kredit konsumen yang telah diketahui statusnya baik dan buruk. Dalam menentukan kelayakan konsumen penerima kredit, tiga belas atribut predictor dan satu atribut kelas. Dibawah ini adalah atribut-atribut yang menjadi parameter terlihat pada :

Tabel 1. Atribut, Nilai dan Keterangan

No	Atribut	Nilai	Keterangan
1	*Age* (umur)	20 tahun	umur konsumen
2	*Sex* (jenis kelamin)	1	laki-laki
		2	perempuan
3	*Marry Status*	1	belum menikah
		2	menikah
4	*Education Level*	1	SLTP
	(tingkat pendidikan)	2	SLTA
		3	Diploma 1/2/3
		4	S1
		5	S2
		6	S3
		7	Master
		8	yang lainnya
5	*Live Year* (lama tinggal)	10 tahun	lama tinggal per tahun
6	*Now House Owner Relate Type*	1	milik sendiri
	(status dari kepemilikan	2	milik keluarga
	rumah tinggal)	3	milik saudara
		4	sewa atau kontrak
		5	kredit
		6	milik kantor atau perusahaan
		7	hipotek bank
7	*Live Person* (jumlah tanggungan)	3 orang	jumlah tanggungan
8	*Work Year* (lama bekerja)	3 tahun	lamanya bekerja
9	*Salary* (gaji)	Rp 3.000.000	besarnya gaji per bulan
10	*Othering* (pendapatan lainnya)	Rp 1.000.000	besarnya pendapatan lainnya per bulan
11	*Business Type*	1	perusahaan swasta
	(Jenis Usaha/ Pekerjaan)	2	pemerintah
		3	perusahaan negara
		4	kaum *professional*
		5	lainnya
12	*Employments Type*	1	pekerja tetap
	(Jenis Pekerjaan)	2	pekerja kontrak
13	*Status*	Good	konsumen yang tidak bermasalah
		Bad	konsumen yang bermasalah

Sumber: PT AEON Credit Service Indonesia (2009)

2.3. Persiapan Data

Merapikan data terpilih untuk mendapatkan kualitas yang lebih baik adalah tujuan dari prapengolahan data. Dimana ada beberapa data yang terpilih mungkin mempunyai format-format yang berbeda karena mereka dipilih dari sumber data yang berbeda-beda. Dapat dikatakan secara umum, pembersihan data berarti menyaring, menggabungkan, dan mengisi kembali nilai-nilai yang hilang (imputasi – *imputation)*.

Dengan penyaringan data, data yang terpilih dicari pencilan (*outlier*-nilai yang jauh berbeda dibanding nilai-nilai lainnya dalam data) dan redundansinya. *Outlier* berbeda jauh dari sebagian besar data, atau data yang jelas-jelas berada di luar kisaran kelompok data terpilih.

Dengan penghalusan data, nilai-nilai yang hilang dari data terpilih ditemukan dan nilai-nilai yang baru atau masuk akal kemudian ditambahkan. Sebuah nilai yang hilang sering kali menyebabkan tidak adanya solusi ketika algoritma penggalian data diterapkan untuk menemukan pola-pola pengetahuan.

2.4. *Modeling*

Modeling adalah suatu tahapan dimana peranti lunak penggalian data digunakan untuk memproduksi hasil untuk berbagai situasi. Analisis pengelompokan dan eksplorasi *visual* data biasanya diterapkan lebih dulu. Bergantung pada jenis data, berbagai model baru kemudian diterapkan, dengan tujuan memungkinkan pengguna untuk bekerja dengan data guna memperoleh pemahaman. Teknik yang digunakan dalam penggalian data ini adalah klasifikasi.

Klasifikasi (*Classification*), metode-metode ditujukan untuk pembelajaran fungsi-fungsi berbeda yang memetakan masing-masing data terpilih ke dalam salah satu dari kelompok kelas yang telah ditetapkan sebelumnya. Penelitian ini menggunakan model klasifikasi yaitu metode *k-Nearest Neighbor* (k-NN) yang adalah merupakan suatu metode yang paling sering digunakan untuk klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut.

2.5. Evaluasi

Tahap evaluasi data sangatlah penting, dengan mengasimilasikan pengetahuan dari data yang telah digali. Untuk mengevaluasi pola dengan menggunakan *software* RapidMiner. Evaluasi dan validas menggunakan metode *cross validation, confusion matrix* dan kurvaROC. Evaluasi yang baik mengarahkan pada keputusan-keputusan bisnis yang produktif, sementara analisis evaluasi yang buruk mungkin melewatkan informasi yang bermanfaat.

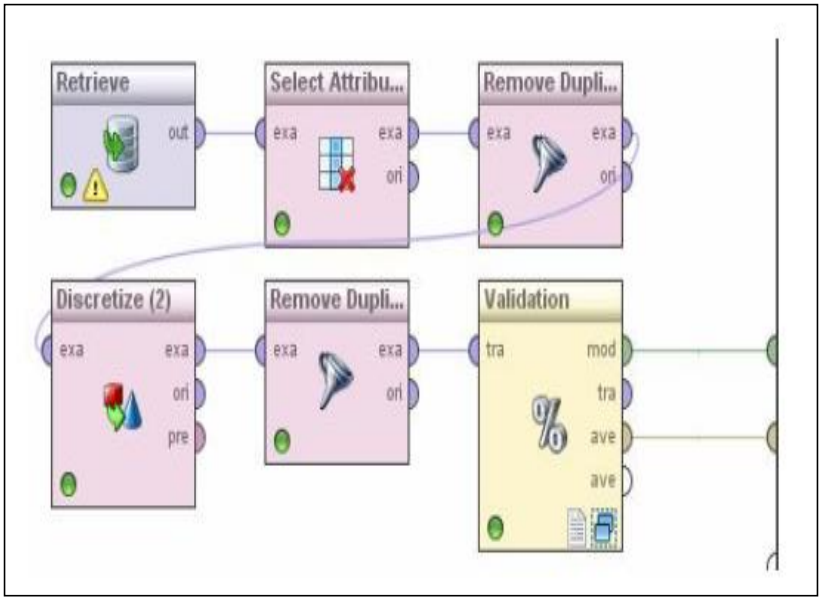
2.6. Pelaksanaan

Setelah melewati tahap *modeling* dan *evaluation* dan tahap-tahap sebelumnya, selanjutnya pada tahap ini ditetapkan model yang dianggap paling akurat untuk diterapkan dalam penentuan kelayakan pemberian kredit terhadap konsumen.

3. HASIL DAN PEMBAHASAN

Tujuan dari penelitian ini adalah mengetahui tingkat keakurasian dari algoritma klasifikasi *data mining* pada data konsumen dalam bentuk kredit yaitu semua data yang telah disetujui oleh pihak perusahaan pembiayaan. Untuk menentukan tingkat keakurasian maka hasil dari analisis algoritma *k-Nearest Neighbor* dan yang sudah dioptimasi dengan *Genetic Algorithm* akan dibandingkan atau dikomparasikan.

Sebelum melakukan komparasi, masing-masing algoritma akan dilakukan pengujian kinerjanya. Cara standar untuk memprediksi tingkat kesalahan pada sampel menggunakan *10-Fold Cross Validation*. Data tersebut dibagi secara acak menjadi 10 bagian dimana kelas diwakili disekitar sama proporsi seperti *data set* lengkap. *10-Fold Cross Validation* telah menjadi metode standar dan cukup untuk mendapatkan perkiraan kesalahan yang dapat diandalkan. Dengan desain modelnya seperti dibawah ini:



Gambar 3. Desain Model Validasi
Sumber: Witten, 2011

3.1. Pengujian Model *k-Nearest Neighbor*

Nilai *accuracy*, *precision*, dan *recall* dari *data training* dapat dihitung dengan menggunakan RapidMiner. Hasil pengujian dengan menggunakan model *k-Nearest Neighbor* didapatkan hasil *accuracy* = 54.32%, *precission* = 36.39%, *recall* = 41.92% seperti pada Tabel 2. dibawah ini:

1. *Confusion Matrix*

Perhitungan berdasarkan data *training* pada Tabel 2., diketahui dari 477 data, 189 diklasifikasikan *bad* sesuai dengan prediksi yang dilakukan dengan metode *k-Nearest Neighbor*, lalu 97 data diprediksi *bad* tetapi ternyata *good*, 70 data *class good* diprediksi sesuai, dan 121 data diprediksi *good* ternyata *bad*.

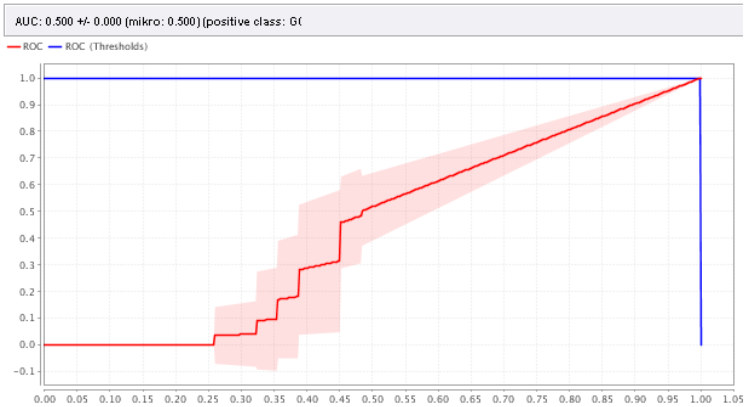
Tabel 2. Model *Confusion Matrix* untuk Metode *K-Nearest Neighbor*

Accuracy:54.32% (mikro:54.30%)			**+/-6.39%**
	True Ya	True Tidak	Class precission
pred. Ya	189	97	66.08%
pred. Tidak	121	70	36.65%
class recall	60.97%	41.92%	

Sumber: Hasil Penelitian (2009)

2. Kurva ROC

Hasil perhitungan divisualisasikan dengan kurva ROC. Perbandingan kedua metode komparasi bisa dilihat pada gambar 3. yang merupakan kurva ROC untuk algoritma *k-Nearest Neighbor*. Kurva ROC pada gambar 3. mengekspresikan *confusion matrix* dari Tabel4. Garis horizontal adalah *false positives* dan garis vertikal *true positives*.



Gambar 4. Kurva ROC dengan Metode *k-Nearest Neighbor*
Sumber: Hasil Penelitian (2009)

3.2. Pengujian Model *k-Nearest Neighbor* yang dioptimasi *Genetic Algorithm*

Nilai *accuracy*, *precision*, dan *recall* dari *data training* dapat dihitung dengan menggunakan RapidMiner. Hasil pengujian dengan menggunakan model *k-Nearest Neighbor* yang dioptimasi *Genetic Algorithm* didapatkan hasil *accuracy* = 68.56%, *precission* = 86.96%, *recall* = 11.98% seperti pada Tabel 3. dibawah ini:

1. *Confusion Matrix*

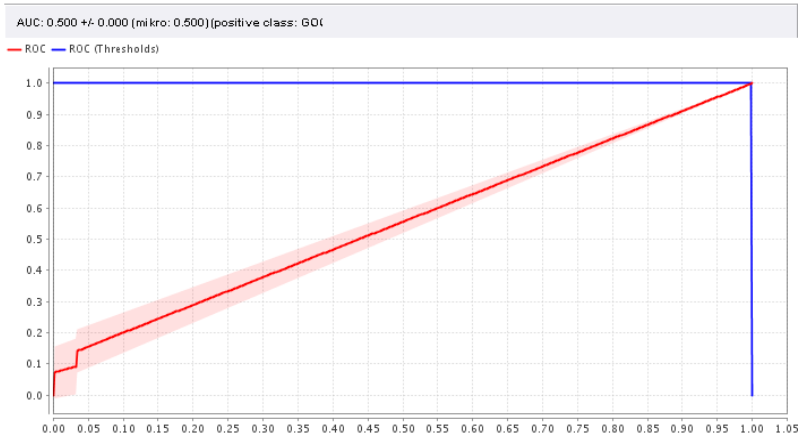
Perhitungan berdasarkan data *training* pada Tabel 3., diketahui dari 307 data, 147 diklasifikasikan *bad* sesuai dengan prediksi yang dilakukan dengan metode *k-Nearest Neighbor*, lalu 147 data diprediksi *bad* tetapi ternyata *good*, 20 data *class good* diprediksi sesuai, dan 3 data diprediksi *good* ternyata *bad*.

Tabel 3. Model *Confusion Matrix* untuk Metode *k-Nearest Neighbor* yang dioptimasi *Genetic Algorithm*

Accuracy:68.56% (mikro:68.55%)			**+/-2.55%**
	True Ya	True Tidak	Class precission
pred. Ya	307	147	67.62%
pred. Tidak	3	20	86.96%
class recall	99.03%	11.98%	

2. Kurva ROC

Hasil perhitungan divisualisasikan dengan kurva ROC. Perbandingan kedua metode komparasi bisa dilihat pada gambar 4. yang merupakan kurva ROC untuk algoritma *k-Nearest* Neighbor yang dioptimasi Genetic *Algorithm*. Kurva ROC pada gambar 4. mengekspresikan *confusion matrix* dari Tabel 4. Garis horizontal adalah *false positives* dan garis vertikal *true positives*.



Gambar 5. Kurva ROC dengan Metode *k-Nearest Neighbor* yang dioptimasi *Genetic Algorithm*

Sumber: Hasil Penelitian (2009)

3.3. Analisis Hasil

Dari hasil analisis model yang dihasilkan dengan metode algoritma *k-Nearest Neighbor dan k-Nearest Neighbor* yang dioptimasi *Genetic Algorithm* diuji menggunakan metode *Cross Validation* maka dapat dirangkumkan :

Tabel 4. Komparasi Nilai A*ccuracy, Precision* dan *Recall*

	Accuracy	**AUC**
K-Nearest Neighbor	54.32%	0.500
K-Nearest Neighbor Berbasis Algoritma Genetika	68.56%	0.500

Sumber: Hasil Penelitian (2009)

Tabel 4. Membandingan *accuracy* , precision dan *recall* dari tiap metode. Terlihat bahwa nilai *accuracy K-Nearest Neighbor* yang dioptimasi *Genetic Algorithm* paling tinggi hasil penggujiannya.

4. SIMPULAN

Berikut ini kesimpulan yang penulis ambil setelah melakukan penelitian :

1. Pengujian model dengan menggunakan *k-Nearest Neighbor* dengan menggunakan data kredit Tahun 2009. Model yang dihasilkan diuji untukmendapatkan nilai *accuracy, precision* dan AUC dari setiap algoritma sehingga didapat pengujian dengan menggunakan *k-Nearest Neighbor* didapat nilai *accuracy* adalah 54.32 % dengan nilai *precision 36.39* % dan nilai AUC adalah 0.500
2. Pengujian dengan mengunakan *k-Nearest Neighbor* yang dioptimasi *Genetic Algorithm* didapatkan nilai accuracy 68.56% dengan nilai *precision* 86.96% dan nilai AUC adalah 0.500

Maka dapat disimpulan pengujian pengujian data kredit Tahun 2009 menggunakan *K-Nearest Neighbor* yang dioptimasi *Genetic Algorithm* lebih baik dari pada *K-Nearest Neighbor* sendiri.

Pada bagian ini, penulis memberikan saran-saran berdasarkan permasalahan serta kesimpulan yang penulis dapat selama penelitian, yaitu :

1. Penelitian ini diharapkan bisa digunakan pada perusahaan pembiayaan untuk lebih meningkatkan akurasi analisa kelayakan kredit bagi konsumen yang hendak mengajukan kredit.
2. Penelitian ini dapat dikembangkan dengan metode optimasi lainnya seperti *Particle Swarm Optimization* (PSO), *Ant Colony Optimization* (ACO), dan lainnya.

3. Penelitian ini dapat dikembangkan dengan metode klasifikasi data mining lainnya seperti *Neural Network*, *Naive Bayes* dan lainnya untuk melakukan perbandingan.

DAFTAR PUSTAKA

[1]. Keramati, A., & Yousefi, N., (2011).*A Proposed Classification of Data Mining Techniques in Credit Scoring*. Malaysia
[2]. Larose, D. T. (2005).*Discovering Knowledge in Data*. New Jersey: John Wille& Sons, Inc.
[3]. Matic, Dragan. 2010. Genetic Algorithm For Composing Music. Yugoslav Journal Of Operations Research,157-177.
[4]. Nanja, M. Purwanto. (2015). Metode K-Nearest Neighbor Berbasis Forward Selection Untuk PrediksI. Jurnal Pseudocode,53-64.
[5]. Witten, I. H., Frank, E., & Hall, M. A. (2011). Data Mining Practical Machine Learning Tools And Techniques. Burlington, Usa: Morgan Kaufmann Publishers.
[6]. Wang, Q., Lai, K. K., & Niu, D., (2011). *Green Credit Scoring System and its Risk Assessement Model withSupport Vector Machine*. China. IEEE.
[7]. Zhang. D., Leung, S. C. H., Ye, Zhime., (2008) *A Decision TreeScoring Model Based on Genetic Algoritm and K-means Algorithm*. IEEE.