

## PREDIKSI PINJAMAN KREDIT DENGAN SUPPORT VECTOR MACHINE DAN K-NEAREST NEIGHBORS PADA KOPERASI SERBA USAHA

Nandang Iriadi<sup>1</sup>, Henny Leidiyana<sup>2</sup>

<sup>1,2</sup>Program Studi Teknik Komputer AMIK BSI Jakarta  
Email : <sup>1</sup>nandang.ndi@bsi.ac.id, <sup>2</sup>henilei@gmail.com

### ABSTRACT

*Cooperatives as a form of organization that are important in promoting economic growth . Cooperatives be an alternative for people to get funds in an effort to improve their quality of life , day-to- day needs and develop the business . No doubt , lend funds to member cooperatives will surely emerge problems , such as members of the borrower paying the overdue installment of funds , misuse of funds for other purposes , the customer fails to develop its business so as to result in cooperative funds do not flow or it can lead to bad credit . In this research will be carried out loans prediction using data mining classification Support Vector Machine and k - Nearest Neighbors were then conducted a comparison of both methods . From the test results to measure the performance of both methods using cross validation , confusion matrix and ROC curves is known that Support Vector Machine has an accuracy value of 92.67 % followed by k -Nearest Neighbors, which has a value of 88.67 % accuracy . Thus the Support Vector Machine method is included in Verry Good Clasification because it has the accuracy of 92.67 % .*

*Keywords: comparative, Support Vector Machines, k-Nearest Neighbors, Credit Analysis*

### ABSTRAK

Koperasi sebagai salah satu bentuk organisasi yang penting dalam meningkatkan pertumbuhan ekonomi. Koperasi simpan pinjam menjadi salah satu alternatif bagi masyarakat untuk mendapatkan dana dalam upaya memperbaiki taraf kehidupan, pemenuhan kebutuhan sehari-hari dan mengembangkan usaha. Tidak dipungkiri, memberikan pinjaman dana kepada anggota koperasi pasti akan muncul permasalahan-permasalahan, seperti anggota peminjam terlambat membayarkan cicilan dana, penyalahgunaan dana untuk keperluan lain, nasabah gagal mengembangkan usahanya sehingga dapat mengakibatkan dana di koperasi tidak mengalir atau dapat mengakibatkan kredit macet. Dalam penelitian ini akan dilakukan prediksi pinjaman kredit dengan menggunakan metode klasifikasi *data mining Support Vector Machine* dan *k-Nearest Neighbor* syang kemudian dilakukan komparasi kedua metode tersebut. Dari hasil pengujian dengan mengukur kinerja kedua metode tersebut menggunakan *cross validation*, *confusion matrix* dan kurva *ROC* diketahui bahwa *Support Vector Machine* memiliki nilai akurasi 92.67% diikuti oleh *k-Nearest Neighbors* yang memiliki nilai akurasi 88,67%. Dengan demikian Metode *Support Vector Machine* tersebut termasuk dalam *Verry Good Clasification* karena memiliki nilai akurasinya sebesar 92.67%.

Kata kunci: komparasi, *Support Vector Machine*, *k-Nearest Neighbors* ,Analisa Kredit

### 1. Pendahuluan

Penelitian mengenai analisa kredit sangatlah menarik, karena disini banyak sekali permasalahan yang pelik yang belum tentu bisa diatasi dengan cara manual. Pemberian kredit kepada debitur khususnya kepada anggota koperasi dengan menggunakan metode data mining telah banyak dilakukan. Kredit yang diajukan oleh debitur memiliki resiko, karena dari sekian

banyak debitur yang mengajukan aplikasi ada kemungkinan beberapa debitur yang bermasalah dalam pembayarannya sehingga menyebabkan kredit macet, namun belum diketahui dengan menggunakan metode klasifikasi data mining yang paling akurat tingkat ketelitiannya dalam melakukan analisis kelayakan kredit tersebut. Sebagai tolak ukur bahwa debitur disetujui atau ditolak, dapat digunakan data histori debitur

yang telah disetujui oleh koperasi. Namun, perlu diperhatikan juga bahwa debitur yang telah disetujui juga tidak semuanya pembayar kredit yang baik, artinya ada beberapa debitur yang telah disetujui tapi beberapa bulan kemudian pembayarannya menunggak. Untuk memprediksi perilaku nasabah kredit ini diperlukan suatu metode atau teknik yang dapat mengolah data-data yang sudah ada di bank tersebut. Salah satu metodenya dapat menggunakan teknik *data mining*. Beberapa teknik yang digunakan untuk menganalisa untuk permasalahan kredit telah dilakukan antara lain dengan menggunakan *neural network* dan *naive bayes*.

Penelitian-penelitian Komparasi yang telah dilakukan oleh Lan Yu, Guoqing Chen, Andy Koronios, Shiwu Zhu, Xunhua Guo (2007), Henny Leidiyana (2011), Ferry Febian (2011), Wisti Dwi Septiani (2013). Penelitian ini bertujuan Untuk memprediksi perilaku nasabah kredit. Hal ini diperlukan suatu metode atau teknik yang dapat mengolah data-data yang sudah ada di koperasi tersebut. Salah satu metodenya dapat menggunakan teknik *data mining*. Metoda tersebut digunakan dalam penelitian ini untuk membandingkan (komparasi) antara metode *support vector machine* dan *k-Nearest Neighbors*, untuk mengetahui apakah kedua metode tersebut benar-benar akurat atau tidak.

## 2. Bahan dan Metode Penelitian

### 2.1. Bahan

Bahan yang digunakan dalam penelitian ini adalah data nasabah koperasi

serba usaha Ceger Jaya tahun 2011, dengan jumlah nasabah bermasalah sebanyak 400 orang. *Preprocessing* data yang dilakukan adalah mereduksi data nasabah koperasi. Data sejumlah 400 *record* direduksi menjadi 300 *record* dengan menghilangkan duplikasi data. Data tersebut selanjutnya diolah menggunakan rapid miner dengan 10 K-fold validation sehingga didapatkan data *training* dan data *testing*.

### 2.2. Metode Penelitian

Penelitian ini terdiri dari beberapa tahap seperti terlihat pada kerangka pemikiran Gambar 1.

#### a. Problem

Permasalahan (*problem*) pada penelitian ini adalah belum diketahuinya algoritma yang akurat untuk kelayakan pemberian pinjaman kredit kepada anggota koperasi, terutama untuk penentuan level resiko kredit.

#### b. Approach

Jenis penelitian yang dilakukan dalam penelitian ini adalah jenis penelitian eksperimen, yaitu menekankan pada percobaan dari teori-teori yang sudah ada. Untuk itu dibuat *approach* (model) dalam bentuk *support vector machine* dan *k-Nearest Neighbors* untuk memecahkan permasalahan kemudian dilakukan pengujian terhadap kinerja dari kedua metode tersebut.

*Support vector machine* (SVM) adalah metode pengklasifikasian yang mencari *hyperplane* terbaik untuk memisahkan data-data dengan kelas-kelas yang berbeda. Diistilahkan juga sebagai

*maximum margin classifier* karena secara simultan meminimalisasikan *classification error* dan memaksimalkan *margin* geometrinya (Sembiring, 2007). Permasalahan klasifikasi yang ada pada umumnya berusaha memisahkan sekelompok data yang berada pada kelas yang satu dengan sekelompok data yang ada di kelas yang lain. *Support vector machine* (SVM) adalah sistem pembelajaran yang pengklasifikasiannya menggunakan ruang hipotesis berupa fungsi-fungsi *linear* dalam sebuah ruang fitur (*feature space*) berdimensi tinggi, dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan mengimplementasikan *learning bias* yang berasal dari teori pembelajaran statistik. Dalam konsepnya SVM berusaha menemukan fungsi pemisah (*hyperplane*) terbaik diantara fungsi yang tidak terbatas jumlahnya. SVM adalah sebuah metode baru yang menjanjikan untuk klasifikasi data baik secara linier maupun onlinier (Han dan Kamber, 2006). Singkatnya, mesin dukungan vektor (SVM) merupakan algoritma yang bekerja sebagai berikut. Menggunakan pemetaan nonlinier untuk mengubah data pelatihan asli ke dimensi yang lebih tinggi. SVM cukup populer untuk penggunaan klasifikasi karena kelebihanannya antara lain dari segi cara kerja, SVM baik untuk klasifikasi, tidak tergantung pada jumlah fitur dan bisa mengatasi masalah dimensi. Dari segi komputasi, SVM dapat melakukan proses

*training* dengan cepat dan ini berguna dalam teknik *learning* ketika menghadapi masalah ketidaktegasan (Maimon, 2005). Algoritma *k-Nearest Neighbors* (k-NN) adalah metode menyediakan pendekatan sederhana untuk menghitung prediksi dalam pengamatan yang diketahui (Myatt 2007). Algoritma *k-Nearest Neighbor* (k-NN) merupakan suatu metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. k-NN merupakan salah satu metode pengklasifikasian data berdasarkan similaritas dengan label data (Larose, 2006). Untuk algoritma *K-Nearest Neighbor* banyak kasus yang dapat diselesaikan dan salah satunya adalah kasus tentang kemungkinan seorang nasabah bank akan bermasalah dalam pembayaran atau tidak (Kusrini dan Luthfi, 2009). *K-Nearest Neighbors* yaitu metode yang memperhitungkan kedekatan jarak data atau kemiripan jumlah kemunculan data antara satu data dengan data lainnya. *Support Vector Machines* yaitu metode yang mencari fungsi pemisah (*hyperplane*) terbaik untuk memisahkan data-data dengan kelas-kelas yang berbeda

### c. Development

Untuk mengembangkan aplikasi (*development*) berdasarkan model yang dibuat, digunakan *Rapid Miner*. Model yang telah dibentuk diuji tingkat akurasi dengan memasukkan data uji yang berasal dari data *training* yang

didapatkan dari metode *cross validation* dengan 10 K-fold *cross validation*.

d. Implementation

Untuk penerapan (*implementation*) yaitu penerapan hasil model yang sudah terbentuk menjadi sebuah *rule* yang digunakan untuk menganalisa anggota koperasi yang layak untuk mendapatkan pinjaman. Di sini tidak dibangun aplikasi tersendiri, melainkan aplikasi dengan cara memasukan data yang sudah dipreprocessing ke rapid miner.

e. Result

Tahap terakhir yaitu menghasilkan (*result*) algoritma klasifikasi yang paling akurat untuk prediksi penentuan pinjaman kredit pada anggota Koperasi Serba Usaha “Ceger Jaya”. Hasil akurasi yang baik pada *rule* yang terbentuk akan direkomendasikan untuk diterapkan dan implementasikan pada sebuah aplikasi yang dapat memprediksi penentuan resiko pinjaman kredit bermasalah pada anggota koperasi tersebut.

f. Measurement

Untuk pengukuran (*measurement*) digunakan metode *Comparison*, *Cross Validation*, *Confusion Matrix*, *Kurva ROC*, *T-Test*. Metode tersebut digunakan untuk membandingkan hasil dari masing-masing metode.

*Confusion Matrix* adalah alat visualisasi yang biasa digunakan pada *supervised learning* (Gorunescu, 2010). Tiap kolom pada matriks adalah contoh dalam kelas prediksi, sedangkan setiap baris mewakili kejadian di kelas yang sebenarnya. Satu

keuntungan dari *Confusion Matrix* adalah mudah untuk mengetahui jika data ada diantara dua kelas (*mislabeling*). *Confusion Matrix* berisi informasi tentang aktual (*actual*) dan prediksi (*predicted*) pada sistem klasifikasi. Kinerja sistem seperti ini biasanya dievaluasi dengan menggunakan data pada matriks.

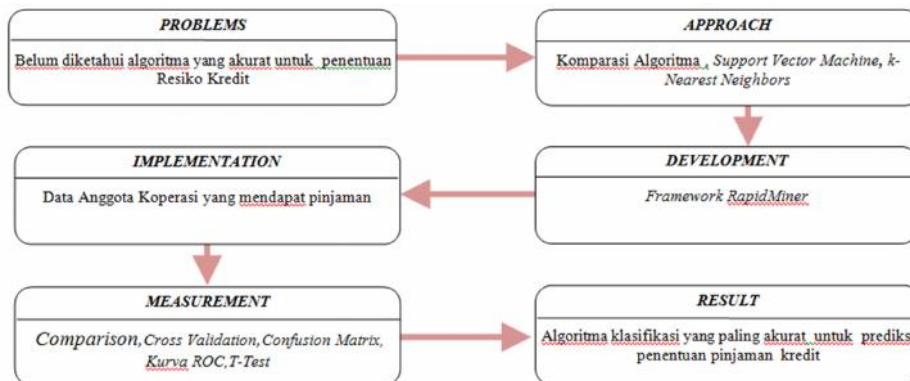
*Receiver operating characteristic* (ROC) adalah cara lain untuk menguji kinerja mengklasifikasi (Gorunescu, 2010). Kurva merupakan suatu grafik yang dapat terbentuk oleh nilai *area under curve* (AUC). Nilai Akurasi AUC dikatakan sempurna apabila nilai AUC mencapai 1.000 dan akurasinya buruk jika nilai AUC dibawah 0.500. Klasifikasi *data mining*, nilai *Area Under Curve*(AUC) dibagi menjadi 5 kelompok, nilai AUC antara 0.90-1.00 (klasifikasi sangat baik), nilai AUC antara 0.80-0.90 (klasifikasi baik), nilai AUC antara 00.70-0.80 (klasifikasi cukup), nilai AUC antara 00.60-0.70 (klasifikasi buruk), nilai AUC antara 00.50-0.60 (klasifikasi salah) (Gorunescu, 2011).

Sebuah grafik ROC adalah plot dengan tingkat positif salah (FP) pada sumbu X dan tingkat positif benar (TP) pada sumbu Y. Titik (0,1) adalah klasifikasi sempurna yang mengklasifikasikan semua kasus positif dan kasus negatif dengan benar, karena tingkat positif salah (FP) adalah 0 (tidak ada) dan tingkat positif benar (TP) adalah 1. Titik (0,0) merupakan sebuah klasifikasi yang

memprediksi setiap kasus menjadi negatif, sedangkan titik (1,1) sesuai dengan sebuah klasifikasi yang memprediksi setiap kasus menjadi positif. Titik (1,0) adalah klasifikasi yang tidak benar untuk semua klasifikasi. Dalam banyak kasus, klasifikasi memiliki parameter yang dapat disesuaikan untuk meningkatkan TP atau penurunan FP. Setiap pengaturan parameter menyediakan pasangan FP dan TP dan serangkaian pasangan tersebut dapat digunakan untuk memetakan kurva ROC. Klasifikasi non-parametrik diwakili oleh titik ROC tunggal, sesuai dengan pasangannya (FP, TP).

T-Test termasuk kedalam metode statistik induktif (inferensi). Pada dasarnya statistik inferensi (Santosa, 2010) mempelajari pengambilan keputusan tentang parameter populasi dari sampel

yang ada. Dalam sebuah kegiatan inferensi ada dua hal dasar yaitu adanya data yang berasal dari sampel dan adanya perlakuan dengan tujuan tertentu terhadap sampel tersebut. Dalam hal ini adalah melakukan pengujian apakah ada beda yang signifikan antara sebelum dan sesudah dilakukan suatu aksi. T-Test adalah salah satu metode pengujian hipotesis dengan menggunakan satu individu (objek penelitian) dikenai dua perlakuan yang berbeda. Walaupun dengan menggunakan individu yang sama tetapi sampel yang diperoleh tetap ada dua data sampel, yaitu data dengan perlakuan pertama dan data dengan perlakuan kedua. Dengan demikian *performance* dapat diketahui dengan cara membandingkan kondisi objek penelitian pertama dan kondisi objek pada penelitian kedua.



Gambar 1. Tahapan Penelitian

### 3. Pembahasan dan Hasil

Gagasan bahwa pikiran manusia mengorganisir pengetahuan dengan menggunakan proses alami klasifikasi tersebar luas. Tapi ketika membahas kata klasifikasi, kita berbicara tentang taksonomi.

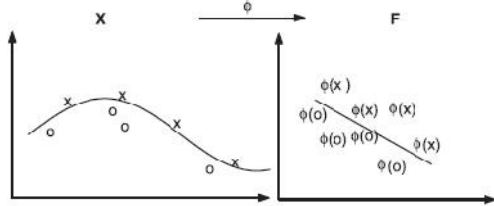
*Taksonomi* (Tassein = mengklasifikasikan dan nomos = ilmu pengetahuan, hukum) muncul pertama sebagai ilmu mengelompokkan organisme hidup (*alfa taksonomi*), tetapi kemudian dikembangkan sebagai ilmu klasifikasi pada umumnya,

termasuk di sini prinsip-prinsip klasifikasi (skema taksonomi) juga. Dengan demikian, klasifikasi (taksonomi) adalah proses menempatkan objek tertentu (konsep) dalam satu set kategori, berdasarkan masing-masing objek (konsep) property (Gorunescu, 2011). Untuk melakukan evaluasi pada komparasi algoritma klasifikasi *data mining* dengan mengambil kasus Komparasi, maka penulis menggunakan metode pengukuran *Confusion Matrix*, *ROC Curve* dan *T-Test*.

**3.1 Pembuatan Model**

Pemodelan dengan metode SVM dilakukan dengan cara berusaha untuk menemukan fungsi pemisah (pengklasifikasi / hyperplane) yang terbaik diantara fungsi yang tidak terbatas jumlahnya untuk memisahkan dua obyek, diilustrasikan pada gambar 2. Dalam penelitian ini adalah pemisahan menjdai dua kelas yaitu kelas lancar dan kelas bermasalah. Karena data tidak dapat dipisah secara linier maka digunakan metode kernel yang dihitung menggunakan persamaan 1.

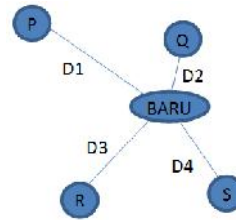
$$K(x, x') = \Phi(x)^T \Phi(x') \dots\dots\dots (1)$$



Gambar 2. SVM

Pemodelan menggunakan k-NN menggunakan prinsip kerja algoritma dengan mencari kelompok k objek dalam data training yang paling dekat atau mirip dengan objek pada data baru atau data

testing, hal tersebut diilustrasikan pada Gambar 3.



Gambar 3. K-NN

Untuk menentukan kedekatan itu dilakukan pengukuran jarak kedekatan antara data baru (data testing) dengan data lama (data training). Pada penelitian ini digunakan metode Euclidean untuk pengukuran jarak yang dicari menggunakan rumus 2, dimana a = a1,a2, ..., an, dan b = b1, b2, ..., bn mewakili n nilai atribut dari dua record.

$$\sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \dots\dots(2)$$

Penelitian ini menggunakan model yang sudah tersedia pada aplikasi rapid miner. Data yang sudah siap dimasukkan untuk diproses oleh rapid miner berdasarkan algoritma yang dipilih. Dalam penelitian ini digunakan algoritma SVM dan k-NN. Nilai akurasi model metode *support vector machine* sebesar 92.67%, dan metode *k-nearest neighbors* sebesar 88.67%.

**3.2 Pengujian Model**

Model dengan pemilihan algoritma yang telah dibentuk diuji tingkat akurasinya dengan memasukan data uji yang berasal dari data *training*, digunakan metode *cross validation* untuk menguji tingkat akurasi, yaitu menggunakan *10 fold cross validation*. Nilai akurasi model untuk metode *support*

*vector machine* sebesar 92.67%, dan metode *k-nearest neighbors* sebesar 88.67%.

Tabel 1 adalah *confusion matrix* dari perhitungan berdasarkan data *training* dengan data sebanyak 300 *record*, diketahui dari 300 data, 0 diklasifikasikan bermasalah sesuai dengan prediksi yang dilakukan dengan metode *support vector machine*, lalu 1 data diprediksi bermasalah tetapi ternyata lancar, 278 data *class* lancar diprediksi

sesuai, dan 21 data diprediksi lancar ternyata bermasalah.

Dengan metode *k-Nearest Neighbors*, menghasilkan kondisi seperti pada Tabel 2. Diketahui dari 300 data, 10 diklasifikasikan bermasalah sesuai dengan prediksi yang dilakukan dengan metode *k-Nearest Neighbors*, lalu 23 data diprediksi bermasalah tetapi ternyata lancar, 256 data *class* lancar diprediksi sesuai, dan 11 data diprediksi lancar ternyata bermasalah.

Tabel 1. Model *Confusion Matrix* untuk Metode *Support Vector Machine*

Accuracy : 92.67% +/- 1.33% (mikro : 92.67%)

	True bermasalah	true lancar	class precision
Pred.Bermasalah	0	1	0.00%
Pred.Lancar	21	278	92.98%
Classrecall	0.00%	99.64%	

Tabel 2. Model *Confusion Matrix* untuk Metode *k-Nearest Neighbors*

Accuracy : 88.67% +/- 5.42% (mikro : 88.67%)

	True bermasalah	true lancar	class precision
Pred.Bermasalah	10	23	30.30%
Pred.Lancar	11	256	95.88%
Classrecall	47.62%	91.76%	

Hasil perhitungan divisualisasikan dengan kurva ROC yang di *generate* dari *rapid miner*. Gambar 4 merupakan kurva ROC untuk algoritma *Support Vector Machine*, mengekspresikan *confusion matrix* dari Tabel 1. Gambar 5 merupakan kurva ROC untuk algoritma *k-Nearest Neighbors*, mengekspresikan *confusion matrix* dari Tabel 2. Garis horizontal adalah *false positives* dan garis vertikal *true positives*. *True positives* adalah jumlah record positif yang diklasifikasikan sebagai positif, *false positives* adalah jumlah record negatif yang diklasifikasikan sebagai positif.

### 3.3 Hasil Komparasi

Model yang dihasilkan menggunakan

metode *Support Vector Machine* dan *k-Nearest Neighbors* diuji menggunakan metode *Cross Validation*. Perbandingan nilai *accuracy*, *precision*, *sensitivity*, dan *recall* disajikan pada Tabel 4. Metode *Support Vector Machine* memiliki nilai *accuracy*, *precision*, *sensitivity* dan *recall* yang paling tinggi dan predikat sangat baik, diikuti dengan metode *k-Nearest Neighbors* yang paling rendah untuk kesemua penilaian. Perbandingan nilai AUC antara *Support Vector Machine*, dan *k-Nearest Neighbors* disajikan pada Tabel 3. Nilai AUC dari metode *k-Nearest Neighbors* memperoleh hasil yang baik dibandingkan dengan hasil dari metode *Support Vector Machine*.

Dari hasil komparasi kedua metode maka didapat hasil komparasi antara *support vector machine* dan *k-Nearest Neighbors* berupa grafik *ROC curve* yang dapat dilihat pada gambar 6. Gambar tersebut merupakan kurva perbandingan kedua metode. Kurva atas merupakan hasil penggambaran dengan metode k-NN, dan yang bawah adalah metode SVM. Terlihat bahwa area di bawah kurva dengan metode k-NN lebih luas dibandingkan metode SVM. Ini berarti

tingkat optimistic k-NN lebih tinggi dari SVM.

Tabel 3 Komparasi Nilai AUC

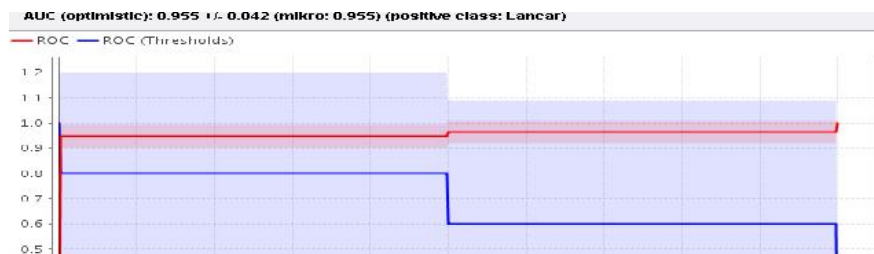
	SVM	KNN
AUC	0.641	0.955

Tabel 4 Komparasi Nilai Accuracy, Precision, dan Recall

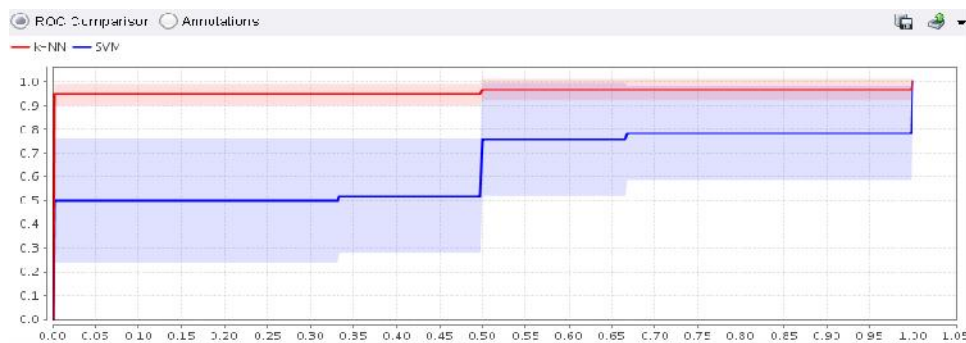
	SVM	KNN
Accuracy	92.67%	88.67%
Precision	92.98%	95.99%
Recall	99.64%	91.75%



Gambar 4. Kurva ROC pada *Support Vector Machine*



Gambar 5. Kurva ROC pada *k-Nearest Neighbors*



Gambar 6. Grafik Komparasi *ROC Curve* Metode *Support Vector Machine* dan *k-Nearest Neighbors*

Melihat dari kedua hasil pengujian dari tabel 3 dan tabel 4 (*accuracy* dan AUC) serta gambar 4 dan 5 ternyata masih belum bisa ditentukan algoritma apa yang terbaik. Menurut pengujian *accuracy*, algoritma

terbaik adalah *Support Vector Machine* sedangkan menurut pengujian *ROC Curve* (AUC), algoritma yang terbaik adalah *k-Nearest Neighbors*. Untuk penentuan lebih lanjut akan digunakan pengujian dengan



memanfaatkan uji statistik yaitu dengan menggunakan uji T-Test (Santosa, 2010).

Tabel 5 Hasil uji T-Test pada *Support Vector Machine* dan *k-Nearest Neighbors* T-Test Significance

	0.927+/-0.013	0.890+/-0.033
0.927+/-0.013		0.005
.890+/-0.033		

Tabel 5 memperlihatkan hasil uji T-Test pada *Support Vector Machine* dan *k-Nearest Neighbors*. Berdasarkan hasil tersebut dapat dianalisis bahwa algoritma *Support Vector Machine* memiliki perbedaan nilai yang sangat dominan atau signifikan karena memiliki nilai probabilitas < 0,05 yaitu 0,005 terhadap algoritma *k-Nearest Neighbors*. Sehingga dapat disimpulkan bahwa algoritma yang paling akurat untuk memprediksi pinjaman kredita dalam *Support Vector Machine* karena memiliki nilai uji *accuracy* tertinggi yaitu 92.67% dan nilai uji statistik T-Test paling dominan terhadap algoritma lainnya walaupun memiliki nilai uji AUC tidak paling tinggi yaitu 0,927 Diikuti dengan algoritma *k-Nearest Neighbors* yang memiliki perbedaan nilai yang sangat dominan atau signifikan terhadap algoritma *Support Vector Machine*.

Tabel 6. Hasil Perbandingan Seluruh Pengujian

	<i>Support Vector Machine</i>	<i>k-Nearest Neighbors</i>
<i>Accuracy</i>	92.67%	88.67%
AUC	0.927	0.890
T-Test	dominan	Tdk dominan

Melihat hasil perbandingan dari seluruh pengujian yang telah dilakukan seperti terangkum pada Tabel 6, menyatakan bahwa algoritma *Support Vector Machine* sangat akurat untuk menangani masalah prediksi. Berdasarkan AUC dapat disimpulkan bahwa Algoritma *Support Vector Machine*

termasuk dalam *Excellent Clasification* dan *k-Nearest Neighbors* termasuk dalam *Good Clasification*.

#### 4. Kesimpulan

Dari penelitian berjudul Prediksi Pinjaman Kredit dengan Support Vector Machines dan K-Nearest Neighbors pada Koperasi Serba Usaha, dapat ditarik kesimpulan bahwa metode *Support Vector Machine* dapat digunakan untuk menentukan resiko peminjaman kredit pada koperasi dan memprediksi pelanggan potensial lebih handal dibandingkan dengan metode k-NN dan cara konvensional dengan akurasi 92.67%.

#### Daftar Pustaka

- Febrian, Ferry. 2011. Analisis Komparasi Algoritma Klasifikasi Data Mining Pada Akseptasi Data Fakultatif Reasuransi Jiwa. [Tesis]. Jakarta : Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer Eresha.
- Gorunescu, Florin. 2011. Data Mining: Concepts, Models, and Techniques. Verlag Berlin Heidelberg : Springer
- Han, J., & Kamber, M. 2006. Data Mining Concept and Tehniques. San Fransisco: Morgan Kauffman.
- Leidiyana, Heny. 2011. Komparasi Algoritma Klasifikasi *Data Mining* Dalam Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor. [Tesis]. Jakarta : STMIK Nusa Mandiri.
- Larose, D. T. 2006. Data Mining Methods And Models. Canada: John Wiley & Sons, Inc.
- Kusrini, & Luthfi, E. T. 2009. Algoritma Data Mining. Yogyakarta: Andi Publishing

- Maimon, Oded&Rokach, Lior. 2005. Data Mining and Knowledge Discovey Handbook. New York: Springer
- Myatt, Glenn J. 2007. Making Sense of Data: A Practical Guide to Exploratory Data Analysis and Data Mining. New Jersey: John Wiley & Sons, Inc.
- Santosa. 2010. Statistik Parametrik - Konsep dan Aplikasi dengan SPSS.Jakarta: PT. Elex Media Komputindo.
- Septiani, Dwi Wisti. 2013. Analisa Dan Komparasi Metode Klasifikasi Data Mining Algoritma C4.5, *Naïve Bayes*, Dan *Neural Network* Untuk Prediksi Penyakit Hepatitis. [Tesis]. Jakarta : STMIK Nusa Mandiri.
- Sembiring, K. 2007. Penerapan Teknik Support Vector Machine untuk Pendeteksian Intrusi pada Jaringan. Bandung : ITB.
- Sumathi, & S., Sivanandam, S.N. 2006. Introduction to Data Mining and its Applications. Berlin Heidelberg New York: Springer.
- Vercellis, Carlo 2009. Business Intelligent: Data Mining and Optimization for Decision Making. Southern Gate, Chichester, West Sussex: John Willey & Sons, Ltd.
- Witten, I. H., Frank, E., & Hall, M. A. 2005. Data Mining: Practical Machine Learning and Tools. Burlington: Morgan Kaufmann Publisher.
- Wu, Xindong & Kumar, Vipin. 2009. The Top Ten Algorithms in Data Mining. Boca Raton: CRC Press