

SISTEM SEDERHANA UNTUK MEMPREDIKSI RISIKO PEMBERIAN KREDIT

D. Lusiyanti¹ dan N. Nacong²

^{1,2,3} Program Studi Matematika Jurusan Matematika FMIPA Universitas Tadulako

Jalan Soekarno-Hatta Km. 09 Tondo, Palu 94118, Indonesia.

¹Desylusiyanti@gmail.com, ²nasrianacong@gmail.co.id

ABSTRACT

Credit risk prediction is very beneficial for the bank or financing institution in making credit decisions. In the decision-making, a decision maker in a banking must be able to take the right decision to accept or reject the credit application. If the decision maker is right in making a decision, then the bank will get customers who support the health and sustainability of the banking business, and vice versa. In this study, Support Vector Machine (SVM) is implemented to predict the crediting risk. The data used is data obtained from one of the financing institutions. By using different activation functions, 80.9524% accuracy is obtained or there are 51 precisely predictable data from 63 existing data.

Keywords : Accuracy, Credit Risk, Support Vector Machine.

ABSTRAK

Prediksi risiko kredit sangat bermanfaat bagi bank atau lembaga pembiayaan dalam mengambil keputusan kredit. Dalam pengambilan keputusan, pengambil keputusan di perbankan harus dapat mengambil keputusan yang tepat untuk menerima atau menolak permohonan kredit. Jika pembuat keputusan benar dalam membuat keputusan, maka bank akan mendapatkan pelanggan yang mendukung kesehatan dan keberlanjutan bisnis perbankan, dan sebaliknya. Dalam studi ini, Dukungan Vector Machien (SVM) diimplementasikan untuk memprediksi risiko kredit. Data yang digunakan adalah data yang diperoleh dari salah satu lembaga pembiayaan. Dengan menggunakan fungsi aktivasi yang berbeda, akurasi 80.9524% diperoleh atau ada 51 data yang dapat diprediksi akurat dari 63 data yang ada.

Kata Kunci : Akurasi, Resiko Kredit, Support Vector Machine.

I. PENDAHULUAN

1.1. Latar Belakang

Pemberian kredit mengandung risiko tinggi dan sangat berpengaruh terhadap kesehatan dan keberlangsungan usaha perbankan. Hal ini tidak lepas dari fungsi dan tugas *decision maker*. Seorang *decision maker* pada suatu perbankan harus mampu mengambil keputusan yang tepat untuk menerima atau menolak permohonan kredit tersebut. Jika *decision maker* tepat dalam mengambil keputusan, maka pihak bank akan mendapatkan nasabah yang menyokong kesehatan dan keberlangsungan usaha perbankan, dan sebaliknya jika salah membuat keputusan maka akan menjatuhkan kelangsungan usaha perbankan.

Beberapa tahun terakhir, metode-metode data mining telah banyak diterapkan dalam memprediksi risiko pemberian kredit, seperti metode decision tree (Sitorus, 1995), jaringan syaraf tiruan backpropagation (Sutikno dkk, 2007), metode naive bayesian dan metode yang relatif baru yaitu radial basis function (RBF). Pada penelitian sebelumnya, Ulfa (2011) menerapkan jaringan syaraf tiruan RBF dan metode PCA dalam memprediksi risiko pemberian kredit. Implementasi RBF dipadukan dengan metode analisis komponen utama atau *Principal Component Analysis* (PCA) untuk meningkatkan preformansi.

Santosa (2010) menyatakan bahwa *Support Vector Machines* (SVM) adalah suatu teknik untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi di mana SVM berada dalam satu kelas dengan ANN dalam hal fungsi dan kondisi permasalahan yang bisa diselesaikan. Telah banyak penerapan dari teknik ini untuk menyelesaikan masalah-masalah nyata seperti masalah gene expression analysis, finansial, cuaca hingga bidang kedokteran. Terbukti SVM memberikan hasil yang lebih baik dari ANN. Oleh karena itu, sebagai pembandingan metode *Support Vector Machine* (SVM) akan diterapkan untuk memprediksi risiko pemberian kredit. Diharapkan penelitian ini dapat memberikan referensi metode yang mempunyai kinerja terbaik untuk permasalahan klasifikasi risiko kredit dan memberikan metode baru pada lembaga pembiayaan atau perbankan untuk memprediksi risiko pemberian kredit.

1.2. Tujuan Penelitian

Mendapatkan hasil klasifikasi risiko pemberian kredit dengan menggunakan Support Vector Machine

1.3. Batasan Masalah

Data yang digunakan berasal dari salah-satu Lembaga pembiayaan yang ada di Kota Palu

II. TINJAUAN PUSTAKA

2.1. Support Vector Machine

SVM adalah suatu teknik yang relatif baru untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. *SVM* pertama kali diperkenalkan oleh Vapnik, Boser, dan Guyon (1992). Prinsip dasar *SVM* adalah linier classifier, yaitu kasus klasifikasi yang secara linier dapat dipisahkan, dan selanjutnya dikembangkan agar dapat bekerja pada masalah nonlinear dengan memasukan konsep kernel trick pada ruang kerja dimensi tinggi (Nugroho,2008).

Misalkan diberikan himpunan $X = \{x_1, x_2, \dots, x_i\}$, dengan $x_i \in \mathbb{R}^n, i=1, \dots, l$. Telah diketahui bahwa X berpola tertentu, yaitu apabila x_i termasuk dalam suatu kelas maka x_i diberikan label (merupakan target) $y_i = +1$, jika tidak diberi label $y_i = -1$. Dengan demikian data yang diberikan berupa pasangan $(x_1, y_1), \dots, (x_m, y_m)$.

Konsep *SVM* dapat dijelaskan secara sederhana sebagai usaha menemukan *fungsi pemisah (klasifier hyperplane)* yang optimal.. Fungsi pemisah yang kita cari adalah fungsi linier. Fungsi ini bisa didefinisikan sebagai :

$$(x) = w^T x + b$$

dimana $x, w \in \mathbb{R}^n$ adalah bobot/weight dan b adalah bias. Dalam teknik ini kita berusaha menemukan *hyperplane* terbaik yaitu *hyperplane* yang terletak ditengah-tengah antara dua set obyek dari dua kelas. Jika $w^T x + b = +1$ adalah *hyperplane* pendukung dari kelas +1 dan $w^T x + b = -1$ *hyperplane* pendukung dari kelas -1, margin antara dua kelas dapat dihitung dengan mencari jarak antara kedua *hyperplane* pendukung dari kedua kelas.

Secara spesifik, margin dihitung dengan cara berikut $(w^T x + b = +1) - (w^T x + b = -1) \Rightarrow (w(x_1 - x_2)) = 2 \Rightarrow \left(\frac{w}{|w|} (x_1 - x_2) \right) = \frac{2}{|w|}$. Dengan mengalikan b dan w dengan sebuah konstanta, akan dihasilkan nilai margin yang dikalikan dengan konstanta yang sama. Selain itu karena memaksimalkan $\frac{1}{|w|}$ sama dengan meminimumkan $|w|^2$ dan jika kedua bidang pembatas direpresentasikan dalam pertidaksamaan

$$(w^T x_i + b) - 1 \geq 0$$

maka pencarian bidang pemisah terbaik dengan nilai margin terbesar dapat dirumuskan menjadi masalah kuadrat programing $\text{Min } \frac{1}{2} |w|^2$

$$\text{dengan kendala: } y_i((w^T x_i) + b) \geq 1$$

dimana : x_i adalah data input

y_i adalah keluaran dari data x_i

w, b adalah parameter-parameter yang kita cari nilainya

Untuk menyelesaikan persamaan di atas, secara komputasi agak sulit dan perlu waktu lebih panjang. Untuk itu diperkenalkan pengali lagrangian. Inti dari Pengali lagrangian adalah mengubah persoalan titik ekstrem terkendala menjadi persoalan ekstrem bebas kendala.

Misalkan $\alpha \geq 0$ adalah pengali lagrangian maka persamaan (3) di atas berubah menjadi

$$L(w, b, \alpha) = \frac{1}{2} |w|^2 - \sum_{i=1}^l \alpha_i [y_i(x_i \cdot w + b) - 1]$$

Solusi dapat diperoleh dengan meminimalkan L terhadap variabel w, b (*primal variables*) dan memaksimalkan L terhadap variable α_i (*dual variables*). Pada saat solusi itu diperoleh (titik optimal), gradient $L = 0$. Dengan demikian, dapat diperoleh nilai α_i yang nantinya digunakan untuk menentukan w dan b dengan rumus :

$$w = \sum_{i=1}^l \alpha_i y_i x_i$$
$$b = y_i - wx_i$$

Terdapat nilai α_i untuk setiap data pelatihan. Data pelatihan yang memiliki nilai $\alpha_i > 0$ adalah *support vector*. Dengan demikian fungsi keputusan yang dihasilkan hanya dipengaruhi oleh *support vector*. Setelah solusi permasalahan kuadratik programing ditemukan (nilai α_i), maka kelas dari data pengujian x dapat ditentukan

berdasarkan nilai dari fungsi keputusan

$$f(x_d) = \sum_{i=1}^{ns} \alpha_i y_i x_i x_d + b$$

dimana : x_i = support vector

ns = jumlah *support* vektor

x_d = data yang akan diklasifikasikan.

III. METODE PENELITIAN

Penelitian ini dilakukan sesuai dengan prosedur sebagai berikut.

1. Identifikasi Permasalahan

Pada tahap pertama ini akan dilakukan identifikasi permasalahan yang akan dibahas yaitu mulai dari mencari topik, literatur-literatur yang berhubungan dengan bahan penelitian dan membuat proposal penelitian.

2. Pengambilan dan Pengolahan Data

Pada tahap ini dilakukan penentuan data input dan target. Sebelum data diolah, akan dilakukan preprocessing data dengan merubah data kategori menjadi numerik. Agar tidak ada variable data yang nilainya mendominasi variable lainnya, maka data dinormalisasi sehingga seluruh data akan bernilai 0 dan 1. Kemudian dilakukan pembagian data menjadi training set dan testing set.

3. Proses Simulasi

Pada tahapan ini dilakukan implementasi metode *Support Vector Machien (SVM)*. Proses pengimpelentasian merupakan proses menerjemahkan algoritma dari metode *SVM* sehingga dapat diproses dalam bentuk kode program MATLAB.

Tahapan ini terdiri dari proses training dan testing. Proses training bertujuan untuk membentuk model hyperplane, langkah yang dilakukan adalah mencari support vector pada training set menggunakan quadratik programing dengan lagrange, yang outputnya berupa lagrange multiplier (alpha). Training set yang mempunyai nilai apha positif disebut dengan support vector. Selanjutnya, model hyperplane yang telah dibentuk akan diujikan pada testing set yaitu data baru yang tidak termasuk dalam proses training.

4. Analisa dan Pembahasan

Agar dapat dikenali oleh jaringan, output yang diperoleh dari proses simulasi dengan metode *SVM* harus diubah ke dalam bentuk numerik, yaitu :

- Diberi nilai "1" jika kredit diterima
- Diberi nilai "-1" jika kredit ditolak.

IV. ANALISIS DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini adalah data aplikasi kredit yang berasal dari salah satu Bank yang ada di Kota Palu. Dataset kredit ini terdiri 63 kasus aplikasi pinjaman. Variable-variable yang digunakan sebagai parameter dalam memprediksi risiko pemberian kredit terdiri dari 9 variabel yang terbagi dalam dua kelas yaitu kelas kategori dan kelas numerik.

Tabel 1 : Atribut Dataset kredit

No	Variabel	Kelas
1	Usia	Numerikal
2	Status Pernikahan dan Jenis Kelamin	Kategorikal
3	Pekerjaan	Kategorikal
4	Lama Kerja (Tahun)	Kategorikal
5	Properti atau Kekayaan	Kategorikal
6	Kepemilikan Rumah	Kategorikal
7	Jumlah Rekening Tabungan	Kategorikal
8	Jangka Waktu (bulan)	Numerikal
9	Sejarah Kredit	Kategorikal

Tahap selanjutnya yaitu melakukan preprocessing terhadap data sebelum data diolah menggunakan metode Support Vector Machine. Proses implementasi diawali dengan proses training. Pada proses training SVM menggunakan faktor rekam medis dan kelas target sebagai input.

DAFTAR PUSTAKA

- [1] Hamel, L. (2009), *Knowledge Discovery With Support Vector Machines*, A John Wiley & Sons, Inc., Hoboken, New Jersey.
- [2] J. Shawe-Taylor and Cristianini. (2004), *kernel methods for pattern analysis*, Cambridge university press, Cambridge.
- [3] Nugroho, A.S. (2008), *Pengantar Support Vector machine*, Pusat Teknologi informasi & Komunikasi BPP Teknologi, http://www.ittelkom.ac.id/stafmhd/MateriKuliah/KapitaSelekta/Kapita%20Selekta_Kedua%20-%20Support%20Vector%20Machine.pdf , diakses 23 Januari 2014.
- [4] Santosa, B. (2010), Tutorial *Support Vector Machine*, Program Studi Teknik Industri, ITS, Surabaya, (<http://oc.its.ac.id/ambilfile.php?idp=1222>), diakses 13 Januari 2014.
- [5] Sembiring, K. (2007), *Penerapan Teknik Support Vector Machine untuk Pendeteksian Intrusi pada Jaringan*, Institut Teknologi Bandung, Bandung
- [6] Sitorus. (1995), *Prediksi Resiko Kredit Dengan Metode Decision Tree*, Skripsi, Universitas Gajah Mada, Yogyakarta
- [7] Sutikno, T. & Pujianta, A. & Supanti, Y. T. (2007), "Prediksi Risiko Kredit dengan Jaringan Syaraf Tiruan Backpropagation" *Seminar Nasional Aplikasi Teknologi Informasi 2007 (SNATI 2007)* ISSN: 1907-5022, Yogyakarta.
- [8] Tomasouw, B. (2012), *Multiclass Twin Bounded Support Vector Machine Untuk Pengenaan Ucapan*", Program Magister Matematika Institut Teknologi Sepuluh Nopember, Surabaya
- [9] Ulfa, F. (2011) , *Penerapan Jaringan Syaraf Tiruan RBF Dan Metode PCA Dalam Memprediksi Risiko Pemberian Kredit*, Program Magister Matematika Institut Teknologi Sepuluh Nopember, Surabaya