
PREDIKSI KEBANGKRUTAN PERUSAHAAN MENGGUNAKAN SUPPORT VECTOR MACHINE (SVM)

¹Lestari Handayani, ²Fitriandini
Jurusan Teknik Informatika Fakultas Sains dan Teknologi
Universitas Islam Negeri Sultan Syarif Kasim Pekanbaru - Riau
Email : ¹lestari81ok@yahoo.com, ²v3_andien@yahoo.com

ABSTRAK

Kebangkrutan perusahaan yang diakibatkan oleh kesulitan keuangan dapat dianalisa dari laporan keuangan. Laporan keuangan perusahaan masa lampau dapat memprediksi kondisi keuangan di masa yang akan datang dengan menggunakan teknik analisa laporan keuangan. Caranya dengan menghitung rasio keuangan yang diperkenalkan oleh Altman. Penelitian ini menggunakan lima variabel rasio keuangan Altman untuk memprediksi kebangkrutan perusahaan dengan menggunakan teknik data mining yaitu metode SVM. Metode yang diperkenalkan oleh Vapnik ini menemukan optimal hyperplane yang memisahkan kondisi perusahaan menjadi dua kelas yaitu bangkrut dan tidak bangkrut. Dengan menggunakan fungsi SMO WEKA, fungsi atau model yang ditemukan dari proses data latih digunakan untuk memprediksi 10 perusahaan yang belum memiliki atribut class output. Data sampel yang digunakan dalam penelitian ini sebanyak 50 perusahaan yang terdaftar di BEI dalam periode 2007-2010. Hasil pengujian 10-cross validation dengan menggunakan fungsi kernel RBF didapat parameter terbaik $C=1$ dan $\Gamma=4$, tingkat akurasi klasifikasi tertinggi mencapai 90.78%. 10 perusahaan yang diprediksi dengan metode SVM menghasilkan 9 perusahaan diklasifikasikan dengan benar dan sisanya diklasifikasikan salah.

Kata Kunci: *Hyperplane*, Prediksi Kebangkrutan, Rasio Keuangan, SVM, Weka

1. PENDAHULUAN

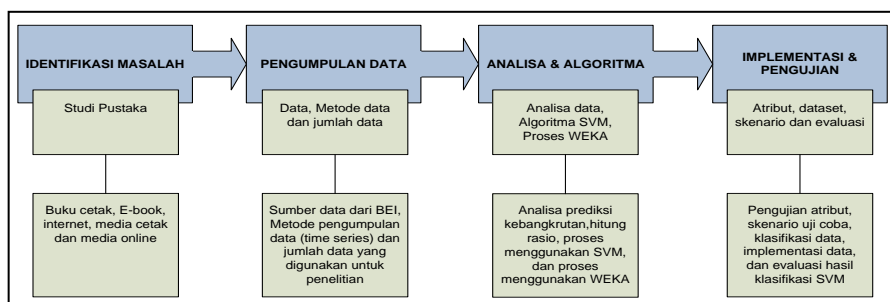
Isi mengenai *financial distress* atau kesulitan keuangan merupakan tahapan awal sebelum terjadinya kebangkrutan. Kebangkrutan yang diakibatkan oleh *financial distress* dapat dilihat dari laporan keuangan perusahaan. Munculnya berbagai model prediksi kebangkrutan merupakan antisipasi dan sistem peringatan dini terhadap *financial distress*. Model tersebut digunakan untuk mendeteksi lebih awal kondisi perusahaan, sehingga sangat memungkinkan perusahaan dan investor melakukan langkah-langkah antisipatif untuk mencegah agar krisis keuangan segera tertangani.

Menurut Santosa (2007), ANN berada dalam satu kelas dengan SVM dalam hal fungsi dan kondisi permasalahan yang diselesaikan, keduanya masuk dalam kelas *supervised learning*. Perbedaan keduanya terletak pada solusi yang dicapai, ANN menemukan solusi berupa *local optimal*, solusi dari setiap training selalu berbeda sedangkan SVM menemukan solusi *global optimal*, mencapai solusi yang sama untuk setiap running. Metode SVM sendiri mendapat perhatian bagi peneliti saat ini, studi ini didukung oleh Institute of Electrical and Electronics Engineers (IEEE) dalam International Conference on Data Mining (ICDM) identified the top 10 algorithms in data mining for presentation ICDM 2006 in Hong Kong, published in april 2009. Dalam paper tersebut disebutkan bahwa SVM terletak di urutan ketiga setelah C4.5 dan K-Means.

Berdasarkan latar belakang diatas, penelitian ini bertujuan untuk mengangkat masalah tentang bagaimana SVM dapat memprediksi kebangkrutan perusahaan yang disebabkan oleh kesulitan keuangan. Metode tersebut kemudian dapat digunakan untuk mengklasifikasikan perusahaan ke dalam kategori bangkrut atau tidak bangkrut. Untuk itu, penelitian ini diberi judul "Prediksi Kebangkrutan Perusahaan Menggunakan Support Vector Machine (SVM)".

2. METODE PENELITIAN

Metodologi yang dibutuhkan untuk pengembangan penelitian ini dimulai dari identifikasi masalah, pengumpulan data, analisa klasifikasi algoritma SVM, pengujian, evaluasi hasil pengujian dan hasil pembahasan.



Gambar 1. Tahapan Metodologi Penelitian

Serangkaian kegiatan yang dilakukan pada penelitian ini yaitu studi pustaka mengenai penelitian yang sudah ada dan mengembangkan ide yang sudah ada untuk penelitian baru. Selanjutnya menyiapkan data berupa rasio keuangan yang dihitung berdasarkan laporan keuangan perusahaan. Untuk sumber data perusahaan diperoleh dari situs resmi Bursa Efek Indonesia (BEI), www.idx.co.id dan metode pengumpulan data secara *time series* periode 2007-2010. Jumlah record data yang digunakan dalam penelitian ini sebanyak 141 data dari 50 perusahaan.

Tahap analisa klasifikasi data dilakukan analisa terhadap laporan keuangan perusahaan yang sehat dan perusahaan yang telah mengalami kesulitan keuangan. Setiap laporan keuangan perusahaan pada periode 2007-2010 dihitung rasio-rasio keuangannya berdasarkan perhitungan Metode Altman. Selanjutnya data diklasifikasikan menjadi dua yaitu *data training* (pelatihan) dan *data test* (pengujian) hal ini untuk menemukan akurasi permodelan prediktif yang bisa dipakai pada proses pembelajaran pada SVM.

Teknik pembelajaran pengenalan pola oleh SVM akan mengklasifikasikan perusahaan bangkrut dan tidak bangkrut berdasarkan *data training*. Secara umum, metode ini akan mencari *hyperplane* (bidang pembatas) terbaik dengan *margin* terbesar antara kedua *class* tersebut yaitu *class 1* untuk perusahaan bangkrut dan *class -1* untuk perusahaan tidak bangkrut. *Hyperplane* ini akan menjadi fungsi keputusan $f(x)$ untuk masalah klasifikasi data yang akan diuji (*data testing*)

Implementasi dan pengujian dalam penelitian ini menggunakan aplikasi WEKA. Perancangan data set dibagi menjadi 80% *data training* dan 20% *data testing*. Nilai akurasi didapat dari hasil pengujian terhadap estimasi parameter terbaik. Pengujian dilakukan dengan menggunakan tiga *test option* yaitu *cross validation*, *use training set & use test set* dan *percentage split*. Beberapa skenario uji coba dilakukan untuk mendapatkan konstanta ongkos C dan parameter *kernel trick*. Kernel yang diujikan yaitu *RBF kernel* dan *poly kernel*. Hasil klasifikasi akan dievaluasi dengan menggunakan hasil perhitungan *confusion matrix*. Evaluasi ini membandingkan hasil akurasi klasifikasi SVM menggunakan WEKA dengan uji 10 data real kondisi perusahaan.

3. HASIL DAN ANALISIS

Analisa keuangan dilakukan dengan menggunakan rasio keuangan. Adapun atribut yang digunakan yaitu variable rasio keuangan model Altman :

1. $X1 = \text{Net Working Capital} / \text{Total Assets}$
2. $X2 = \text{Retaining Earnings} / \text{Total Assets}$
3. $X3 = \text{Earning Before Interest and Tax} / \text{Total Assets}$
4. $X4 = \text{Market Value of equity} / \text{Book Value of Debt}$
5. $X5 = \text{Sales} / \text{Total Assets}$

Berdasarkan laporan keuangan perusahaan, didapat masing-masing rasio keuangan perusahaan

Tabel 1. Rasio Keuangan Perusahaan

| No | Nama Perusahaan | Tahun | X1 | X2 | X3 | X4 | X5 | Keterangan |
|----|-------------------------------------|-------|-------|-------|-------|-------|-------|------------|
| 1 | PT.Davomas Abadi (DAVO) | 2008 | 0.335 | 0.014 | 0.128 | 0.229 | 0.937 | Bangkrut |
| | | 2007 | 0.125 | 0.320 | 0.049 | 1.846 | 1.060 | Bangkrut |
| 2 | PT.Infoasia Teknologi Global (IATG) | 2008 | 0.063 | 0.328 | 0.032 | 1.856 | 0.944 | Bangkrut |
| | | 2007 | 0.693 | 0.612 | 0.174 | 2.817 | 0.494 | Bangkrut |
| 3 | PT.Sara Lee Body Care Indonesia | 2007 | 0.693 | 0.612 | 0.174 | 2.817 | 0.494 | Bangkrut |

| | | | | | | | | |
|---|-----------------------------|------|-------|-------|-------|-------|-------|----------------|
| | (PROD) | 2008 | 0.759 | 0.821 | 0.116 | 4.791 | 0.438 | Bangkrut |
| 4 | PT. Merck (Merk) | 2007 | 0.667 | 0.722 | 0.388 | 5.513 | 1.653 | Tidak Bangkrut |
| | | 2008 | 0.694 | 0.766 | 0.381 | 6.856 | 1.699 | Tidak Bangkrut |
| 5 | PT. Darya Varia Laboratoria | 2007 | 0.586 | 0.186 | 0.138 | 4.683 | 0.882 | Tidak Bangkrut |
| | | 2008 | 0.544 | 0.235 | 0.171 | 3.912 | 0.906 | Tidak Bangkrut |

Rasio keuangan x1, x2, x3, x4 dan x5 menjadi nilai input atribut (x_i) pada pembelajaran SVM.

3.1. Proses Prediksi Data Dengan SVM

Dalam melakukan prediksi masa depan, SVM melakukan klasifikasi data berdasarkan data historis dengan skenario yang telah diberikan. Proses pengklasifikasian ini membagi data sesuai dengan kelas bangkrut dan kelas tidak bangkrut, oleh karena itu data dibagi menjadi dua yaitu data latih dan data uji. SVM mencari *hyperplane* (garis pemisah) terbaik dari sekian banyak kemungkinan *hyperplane* atau fungsi keputusan yang memisahkan antar kelas tersebut. *Hyperplane* yang terbaik juga dilihat dari *margin* (jarak) terbesar antara titik maksimal dari masing-masing kelas dengan *hyperplane*.

Agar data dapat diklasifikasi, maka data direpresentasikan sebagai vektor berupa pasangan data (x, y). Nilai x merupakan nilai rasio keuangan perusahaan. Nilai *class* bangkrut diberi label 1 dan *class* tidak bangkrut diberi label -1, direpresentasikan sebagai berikut,

Tabel 2. Vektor Pasangan (x, y)

| Kode | (x_1, y) | (x_2, y) | (x_3, y) | (x_4, y) | (x_5, y) |
|--------|--------------|--------------|--------------|--------------|--------------|
| IATG07 | (0.125, 1) | (0.320, 1) | (0.049, 1) | (1.846, 1) | (1.060, 1) |
| MACO07 | (-0.069, 1) | (0.111, 1) | (0.212, 1) | (0.142, 1) | (0.633, 1) |
| PROD07 | (0.693, 1) | (0.612, 1) | (0.174, 1) | (2.817, 1) | (0.494, 1) |
| MERK07 | (0.667, -1) | (0.722, -1) | (0.388, -1) | (5.513, -1) | (1.653, -1) |
| DVLA07 | (0.586, -1) | (0.186, -1) | (0.138, -1) | (4.683, -1) | (0.882, -1) |

Dalam penyelesaian klasifikasi data *non linear* diatas, maka digunakan formula (1) untuk pencarian bidang pemisah terbaik dengan nilai margin terbesar, maka formula bidang pemisahanya,

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \tag{1}$$

dengan kendala $y_i(w \cdot x_i + b) \geq 1 - \xi_i, i = 1, \dots, l, \xi_i \geq 0, i = 1, \dots, l$

Dengan besaran konstanta ongkos C yang ditentukan pengguna. Pada pengujian tugas akhir ini dilakukan pengujian terhadap pencarian nilai C terbaik yang memiliki akurasi tertinggi. Konstanta ongkos C digunakan untuk meminimumkan kesalahan dalam klasifikasi data. Untuk memperoleh nilai w (bobot vektor) maka selanjutnya proses menggunakan formula *primal lagrangian* berikut

$$\min L_p(w, b, \alpha) \equiv \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i - \sum_{i=1}^n \alpha_i \{y_i (x_i^T \cdot w + b) - 1 + \xi_i\} - \sum_{i=1}^n \mu_i \xi_i \tag{2}$$

Biasanya nilai w besar maka persamaan *primal lagrangian* diubah menjadi dual lagrangian (3)

$$L_D(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^n \alpha_i \alpha_j y_i y_j k(x_i^T x_j) \tag{3}$$

Setelah diturunkan *lagrangian multipliers*, maka nilai w dapat diketahui

$$w = \sum_{i=1}^n \alpha_i y_i x_i \tag{4}$$

Pada persamaan *non linear separable data problem* (3), bentuk primal dari formula pencarian *hyperplane* terbaik sulit dipecahkan maka digunakan bentuk dual yang merubah nilai w dalam bentuk α . Perubahan nilai w ke dalam bentuk α merupakan masalah *Quadratic Programming* (QP). Untuk menyelesaikan permasalahan tersebut dalam penelitian ini digunakan algoritma dekomposisi yang dikenalkan oleh Platt (1998) yaitu SMO.

Permasalahan ini diselesaikan dengan menggunakan *kernel trick*. Dalam proses transformasi data ke dalam dimensi ruang fitur (*feature space*), data tersebut dapat dipisahkan sesuai dengan kelasnya. Fungsi *kernel trick* tersebut digunakan untuk transformasi data $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$. Titik data pertama pada Tabel 2 ada vektor 0.125, 0.320, 0.049, 1.846, 1.060 dikalikan dengan 0.125, 0.320, 0.049, 1.846, 1.060 maka akan ketemu nilai tertentu. Perkalian vector tersebut merupakan perkalian vektor ukuran (1x5)(5x1)=1x1.

Pada Tabel 2 ada 5 titik data dalam 5 dimensi. Dengan menggunakan *kernel polynomial* pada eksponen= 2 akan diperoleh matriks kernel 5x5. Berikut rumus fungsi kernel polykernel dan hasil perhitungan,

$$K(x,y) = (< x, y > +1)^e \tag{5}$$

Berikut contoh perhitungan K,

$$x_1x_1' = [0.125 \quad 0.320 \quad 0.049 \quad 1.846 \quad 1.060][0.125 \quad 0.320 \quad 0.049 \quad 1.846 \quad 1.060]' = 4.65;$$

$$(x_1x_1' + 1)^2 = 31.94$$

$$x_1x_2' = [0.125 \quad 0.320 \quad 0.049 \quad 1.846 \quad 1.060][-0.069 \quad 0.111 \quad 0.212 \quad 0.142 \quad 0.633]' = 0.97;$$

$$(x_1x_2' + 1)^2 = 3.88$$

Dengan melakukan perhitungan yang sama untuk seluruh data x_i maka didapat nilai matriks K sebagai berikut:

$$K = \begin{bmatrix} 31.94 & 3.88 & 49.20 & 175.89 & 114.90 \\ 3.88 & 2.19 & 3.13 & 8.67 & 4.98 \\ 49.20 & 3.13 & 101.29 & 335.56 & 230.17 \\ 175.89 & 8.67 & 335.56 & 0.00128 & 32.55 \\ 114.90 & 4.98 & 230.17 & 32.55 & 581.07 \end{bmatrix}$$

Matriks k sebagai pengganti *dot product* $x_i^T x_j$ pada persamaan (16). Nilai a_i pada persamaan (16) tersebut menunjukkan *support vector* fungsi keputusan dari pelatihan. *Support vector* didapat dari hasil perhitungan menggunakan alat bantu WEKA.

3.2. Proses Penyelesaikan Dengan WEKA

Dalam penyelesaian SVM *non linier* tersebut, penelitian ini dilakukan melalui simulasi komputer menggunakan SVM *non linear*, untuk itu dipilih alat bantu *datamining open source* yaitu WEKA 3.6. Aplikasi WEKA yang digunakan adalah *Explorer* dan modul yang digunakan yaitu *preprocessing* dan *classify*.

Data dalam kasus prediksi kebangkrutan perusahaan ini merupakan *non linearly separable*, untuk proses klasifikasinya data dipetakan ke *feature space* agar dapat dipisahkan secara linier. Oleh karena itu perlu kernel agar data dapat dipetakan berdasarkan parameter-parameter yang digunakan. Selanjutnya dilakukan perubahan pada konstanta C untuk mendapatkan keakurasian tertinggi. Fungsi yang digunakan untuk menemukan hyperplane terbaik adalah SMO.

Dalam pengujian pertama dilakukan *test option cross validation* dengan *10-folds*. Penentuan uji coba terhadap pencarian parameter C, γ (kernel RBF) dan eksponen (kernel Polynomial) dilakukan pada rentang nilai C= 0.01, 0.1, 1, 2, 3, 4 dan 5. Selanjutnya dilakukan pencarian parameter pada range nilai yang lebih kecil.

Hasil pencarian parameter pada range yang lebih kecil, hasilnya menunjukkan jika dilakukan pencarian parameter dengan range yang lebih kecil antara range parameter C yaitu 1 - 2 dan G = 3 - 4, nilai akurasi yang diperoleh yaitu 90.78%. Setelah dilakukan pencarian parameter C pada range yang lebih kecil antar 1-2 dan G=3-4, nilai keakurasian yang didapat tetap sama yaitu 90.78%. Untuk selanjutnya pengujian *test option use training set* dan *percentage split* dilakukan dengan menggunakan kernel RBF, *test option 10-cross validation*, konstanta C = 1 dan G = 4.

Setelah melakukan pengujian *10-fold cross validation* perbandingan kernel Polynomial dan kernel RBF menghasilkan nilai akurasi yang lebih tinggi dalam pelatihan dan pengujian data. Akurasi nilai data latih kernel RBF mencapai 90.78 % dan polynomial kernel 89.36%.

4. KESIMPULAN

Berdasarkan dengan data real kondisi perusahaan, kedua fungsi RBF dan Polynomial kernel tersebut menghasilkan jumlah data yang benar dan salah yang sama dalam prediksi kondisi 10 perusahaan, namun jika dievaluasi menggunakan *confusion matrix*, fungsi kernel RBF lebih tinggi tingkat akurasi nya yaitu mencapai 91.50% dibandingkan dengan poly kernel 90.44%.

Kernel RBF lebih cocok digunakan dalam kasus prediksi keuangan perusahaan dengan akurasi yang didapat pada *test option cross validation* mencapai 90,87%. Hal ini dipengaruhi oleh data yang digunakan dalam proses pembelajaran, konstanta C dan parameter kernel.

Hasil pengujian 10 data perusahaan yang baru untuk diprediksi, SVM mengklasikasikan secara benar 9 perusahaan dan sisanya *misclassification*.

REFERENSI

[1] Han, Jiawei, Micheline Kamber and Jian Pei. *Ebook Data Mining : Concepts and Techniques, third Edition. [Online]* Available, http://books.google.co.id/books?id=6hkR_ixby08C&pg=PA12&hl=id&source=gbs_toc_r&cad=4#v=onepage&q&f=false, 2001.

- [2] Hsu, chih-wei, chih-jen li. *A Comparison of Methods for Multi-Class Support Vector Machines. IEEE Transactions on Neural Network*, 13(2):415-425.2002. [Online] Available <http://www.csie.ntu.edu.tw/~cjlin/papers/multisvm.pdf>. Oktober 2011.
- [3] Santosa, Budi. "Data Mining Teknik Pemanfaatan Data Untuk Keperluan Bisnis ", Yogyakarta: Graha Ilmu. 2007.
- [4] Sembiring, Krisantus, (2007). "Penerapan Teknik Support Vector Machine untuk Pendeteksian Instrusi Jaringan". Tugas Akhir, Teknik Informatika, ITB. [Online] Available <http://digilib.itb.ac.id/gdl.php?mod=browse&op=read&id=jbptitbpp-gdl-krisantuss-28957>, diakses 23 september 2011.
- [5] Shin, Kyung-Shin, Taik Soo Lee, & Hyun-Jung Kim, "An application of support vector machine in bankruptcy prediction model". *Expert System with Application* 28 (2005) 127-135. 2005.
- [6] Rifqi, Muhammad, (2009). Analisis Perbandingan Model Prediksi Financial Distress Altman, Ohlson, Zmijewski dan Springate dalam Penerapannya di Indonesia. Skripsi. Universitas Indonesia. [Online] Available <http://lontar.ui.ac.id/opac/themes/libri2/detail.jsp?id=126451&lokasi=lokal> diakses 4 oktober 2011.
- [7] _____, Laporan Keuangan Perusahaan Bursa Efek Indonesia (BEI), http://202.155.2.90/corporate_actions/new_info_jsx/jenis_informasi/01_laporan_keuangan/02_Soft_Copy_Laporan_Keuangan/. 3 Agustus 2011.
- [8] _____, Daftar nama perusahaan bangkrut listed BEI. <http://www.idx.co.id/MainMenu/Education/WhatisEquities/tabid/88/lang/id-ID/language/id-ID/Default.aspx>. 3 Agustus 2011.