

Aplikasi Data Mining dengan Metode *Support Vector Machine* (SVM) untuk Prediksi *Financial Distress* pada Industri Jasa *Go Public* yang Terdaftar di Bursa Efek Indonesia

Lusi Mei Cahya W^{1,*}, Albertus Daru¹, David Andrian¹

1 Fakultas Teknik, Program Studi Teknik Industri, Universitas Katolik Darma Cendika
Jl. Dr.Ir Soekarno 201 Surabaya 60117

* E-mail : loucee_mei@yahoo.com

Abstrak. *Financial Distress* adalah tahapan penurunan kondisi keuangan suatu perusahaan sebelum terjadinya kebangkrutan. Prediksi *financial distress* bermanfaat bagi perusahaan, investor dan kreditor. Penelitian ini dilakukan pada Industri Jasa *Go Public* yang terdaftar di Bursa Efek Indonesia menggunakan metode *Support Vector Machine* (SVM) dengan variabel keuangan. Seleksi variabel menggunakan *Linear Programming SVM* (LP-SVM) . Dari 16 variabel yang tersedia terpilih 10 variabel prediktor pada model. Hasil percobaan penelitian menunjukkan bahwa tingkat misklasifikasi menjadi semakin kecil ketika menggunakan variabel yang telah diseleksi baik dengan metode SVM maupun *Linear Discriminant Analysis* (LDA). Misklasifikasi terkecil dengan metode SVM diberikan pada Kernel RBF dengan parameter $\sigma=2, C=1$

Kata Kunci: *Financial Distress*, Industri Jasa, *Linear Discriminant Analysis*(LDA), *Support Vector Machine* (SVM).

1. Pendahuluan

Kondisi perekonomian negara secara makro akan mempengaruhi nilai saham. Melemahnya nilai tukar rupiah terhadap dolar Amerika Serikat (AS) berdampak negative terhadap bursa saham. Sepanjang Maret 2015 telah terlihat banyaknya dana investor asing yang cenderung keluar dari bursa saham (www.liputan6.com). Gerak nilai tukar rupiah dan sejumlah kebijakan pemerintah menjadi factor yang mendorong aksi jual investor asing. Padahal nilai pasar saham merupakan representasi dari kinerja perusahaan bagi pemegang saham. Hal ini dapat mengakibatkan kondisi perusahaan mengalami *financial distress* (gejala penurunan performa keuangan perusahaan sebelum terjadinya likuidasi atau kebangkrutan).

Financial distress juga dapat dialami oleh semua perusahaan, terutama jika negara dimana perusahaan tersebut beroperasi sedang mengalami krisis ekonomi. Gejala *financial distress* sangat penting untuk dikenali sedini mungkin sehingga tindakan-tindakan pencegahan yang mampu meminimalkan resiko menuju kebangkrutan dapat diambil. Informasi ini selanjutnya dapat digunakan oleh pihak internal (manajemen) maupun eksternal (investor, kreditor, pemerintah) perusahaan untuk membuat keputusan karena kelangsungan hidup perusahaan tentunya berimbas pada kemakmuran pihak-pihak tersebut.

Pentingnya mengenali *financial distress* ini menyebabkan telah banyak metode dan penelitian yang dilakukan selama ini. Secara umum, indikator dan objek utama dalam tiap penelitian mengenai *financial distress* adalah laporan keuangan perusahaan yang melaporkan status dan posisi perusahaan serta perubahan posisi keuangan. Dari laporan keuangan ini, dilakukan analisis seperti analisis rasio dan analisis arus kas (*liability, leverage, profitability, dll*) sehingga teridentifikasi kinerja keuangan saat itu. Selain laporan keuangan, beberapa penelitian juga menemukan bahwa struktur *corporate governance* (komite dan kepemilikan perusahaan) juga berpengaruh signifikan pada *financial distress* suatu perusahaan.

Penelitian ini akan dimulai dari proses pemilihan variabel rasio keuangan, pemodelan prediksi *financial distress* dengan *data mining*, serta evaluasi hasil. Pada penelitian-penelitian sebelumnya, sektor yang sering diobservasi adalah industri manufaktur. Karena itu, penelitian ini akan

menggunakan perusahaan jasa *go-public* yang terdaftar pada Bursa Efek Indonesia (BEI) sebagai objeknya, kecuali sektor perbankan .

Penelitian mengenai kebangkrutan sudah dimulai tahun 1960 dan terus berkembang hingga saat ini. Altman (1968) melakukan prediksi kebangkrutan dengan menggunakan data dari laporan keuangan dan menyatakan bahwa proses kebangkrutan merupakan proses jangka panjang sehingga dalam laporan keuangan seharusnya tercantum signal peringatan akan terjadinya kebangkrutan. Metode *Multiple Discriminant Analysis* (MDA) yang digunakan ini mengasumsikan bahwa variabel bebas harus memenuhi multivariat normal distribution dan memiliki matriks kovarian yang sama. Dari penelitian Altman muncul penelitian lain dengan pengembangan metode statistik, seperti logistik regression (Martin,1997). Adanya asumsi asumsi yang ketat dalam metode tradisional statistik seperti linieritas dan normalitas membuat aplikasinya dalam permasalahan menjadi terbatas. (Delen et al 2013).

Teknik data mining untuk membangun model prediksi financial distress mulai dilakukan , dimana kelebihan dari teknik data mining adalah tidak adanya asumsi linieritas dan normalitas seperti pada dalam metode statistik tradisional. Salah satu peran data mining adalah teknik klasifikasi yang bertujuan untuk menghasilkan model sehingga dapat menggolongkan data testing kedalam kelas se akurat mungkin. (Gong dan Huang, 2012).

Prediksi *financial distress* perusahaan merupakan permasalahan klasifikasi yang bersifat biner, dimana perusahaan yang mengalami *financial distress* dilambangkan dengan nilai 1 (kelas positif), sedangkan yang tidak mengalami *financial distress* dilambangkan dengan nilai -1 (kelas negatif), (Nisa, 2013). Kinerja model prediksi *financial distress* meliputi serangkaian proses mulai dari pemilihan sampel, pemilihan variabel, penentuan metode perhitungan, konstruksi model dan evaluasi kinerja model.

Perkembangan teknik *data mining* dan kecerdasan buatan (*artificial intelligent*) dimulai tahun 1990 an yang meliputi beragam metode, seperti *decision tree* (DT), *case-base reasoning* (CBR), *artificial neural network* (ANN), algoritma evolusioner *genetic algorithm* (GA) dan *support vector machine* (SVM). Teknik-teknik tersebut banyak diaplikasikan untuk memprediksi *financial distress* perusahaan (Chen, 2011). Pada beberapa penelitian metode SVM telah terbukti sebagai metode yang cukup handal untuk kasus klasifikasi dan regresi

Teknik pemilihan variabel rasio keuangan yang umumnya digunakan pada sebagian penelitian sebelumnya adalah menggunakan teknik statistik seperti Principal Component Analysis (PCA), Factor Analysis, t-test, Stepwise Regression dan Correlation Matrix (Tsai, 2009), Penelitian ini menggunakan Li-norm yang disebut juga Linear Programming SVM (LP-SVM)(Travalis et al,2005 dalam Santoso,2007)

2. Support Vector Machine (SVM)

SVM adalah sebuah algoritma yang diusulkan oleh Vapnik pada tahun 1995. SVM tergolong metode klasifikasi baru dan telah banyak dijadikan metode dalam sejumlah penelitian, seperti *pattern recognition*, regresi, dan estimasi. SVM menggunakan masukan atau input dari data *training* untuk menemukan *hyperplane* yang dapat mengklasifikasikan dua atau lebih tipe data untuk kemudian memproses atribut-atribut dari problem klasifikasi (Nisa, 2013). SVM dapat diterapkan pada data yang bersifat linear maupun non-linear. Untuk kasus klasifikasi, dimana datanya tidak linear dapat menggunakan metode Kernel.

Problem optimasi menggunakan SVM untuk kasus klasifikasi dengan dua kelas dimana data tidak dapat dikelompokkan secara benar dapat dirumuskan sebagai berikut

$$\min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l t_i \quad (1)$$

Subject to

$$y_i(\omega x_i + b) \geq 1, i = 1, \dots, l$$

Dimana

x_i = data input

y_i = output dari data x

w_i, b = parameter yang akan dicari nilainya

C = parameter yang akan dicari oleh user

Menggunakan Lagrange multiplier pada (Pers 1) dapat diformulasikan sebagai dual Lagrange sebagai berikut

$$\max \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l y_i y_j \alpha_i \alpha_j x_i^T x_j \quad (2)$$

Subject to

$$\sum_{i=1}^l \alpha_i y_i = 0$$

$$0 \leq \alpha_i, i = 1, \dots, l$$

Decision function yang akan dihasilkan memenuhi rumusan

$$g(x) = \text{sign}(\sum_{i=1}^l y_i \alpha_i^* k(x_i, x)) + b^* \quad (3)$$

Untuk mengatasi permasalahan yang bersifat tidak linier, dapat digunakan metode kernel. Dengan metode kernel suatu data x diinput space di mapping ke feature space F yang lebih tinggi. Suatu kernel map mengubah problem yang tidak linier menjadi linier dalam space baru.

Fungsi kernel yang biasanya dipakai dalam literatur SVM (Haykin,1999) dalam Santosa (2007)

1. Linier : $x^T x$

2. Polynomial : $(x^T x + 1)^p$

3. Radial basis function (RBF) :

$$\exp\left(-\frac{1}{2\sigma^2} \|x - x_l\|^2\right)$$

4. Tangent hyperbolic (sigmoid) :

$$\tanh((\beta x^T x + \beta_i))$$

Pemilihan jenis fungsi kernel yang akan digunakan untuk substitusi dot product di feature space akan sangat bergantung pada data.

3. Metode

3.1 Data Penelitian

Dalam penelitian ini akan digunakan data laporan keuangan dari perusahaan *go public* yang ada di Indonesia dengan alasan untuk kemudahan akses informasi. Data laporan keuangan tersebut diperoleh melalui *website* Bursa Efek Indonesia (www.idx.co.id). Perusahaan yang akan diamati adalah perusahaan jasa yang terdiri dari 3 sektor yaitu Sektor properti dan *real estate*, sektor infrastruktur, utilitas, dan transportasi, sektor perdagangan, jasa dan informasi sebanyak 60 perusahaan

Perusahaan yang menjadi obyek amatan dalam penelitian ini adalah perusahaan yang memiliki laporan keuangan lengkap yang terdiri atas laporan rugi laba, neraca, laporan perubahan ekuitas, laporan arus kas dan catatan keuangan dari auditor independence. Laporan keuangan yang akan digunakan dalam laporan ini adalah laporan keuangan tahunan yang telah diaudit pada periode 2011 sampai 2013. Alasan digunakan laporan tahunan yang telah diaudit terdapat opini kewajaran sebuah laporan keuangan dari auditor.

3.2 Identifikasi Variabel

Pemilihan variabel yang akan digunakan dalam penelitian ini mengacu pada berbagai hasil peneliti pendahulu. Variabel variabel yang akan digunakan pada penelitian ini meliputi variabel keuangan dan variabel non keuangan, yang bersifat kuantitatif. Variabel keuangan yang bersifat kuantitatif meliputi beberapa rasio keuangan, antara lain rasio likuiditas, rasio solvabilitas, rasio profitabilitas, rasio pemanfaatan asset dan rasio investor. Variabel yang terpilih beserta sumbernya ditunjukkan pada tabel

Tabel 1. Variabel Keuangan dan Rasio

Kriteria	No	Sub kriteria	Deskripsi
Rasio Likuiditas	1	Working Capital –to Total Asset	(Current asset – Current Liabilities)/ Total Asset
	2	Current Ratio	Current Asset / Current Liabilities
	3	Quick Ratio	(Current asset – Inventory) / Current Liabilities)
	4	Account Receivable Turnover	Sales/Account Receivable
	5	Inventory Turnover	Cost of goodsold/ Inventory
Rasio Solvabilitas	6	Debt to Equity	Total Liabilities/Total Equity
	7	Leverage Ratio	Total Liabilities/Total Asset
Rasio Profitabilitas	8	Gross Profit Margin	Gross Profit/Sales
	9	Net Profit Margin	Net Income/Sales
	10	Operating Profit Margin	Operating Income/Sales
	11	ROE	Net Income/Total Equity
	12	ROA	Net Income / Total Asset
Rasio Pemanfaatan Akses	13	Asset Turnover Rate	Sales/Total Asset
	14	Working Capital Turnover Rate	Sales/Working Capital
	15	Fixed Asset Turnover Rate	Sales/Fixed Asset
Rasio Investor	16	EPS	(Laba bersih - deviden saham) /jumlah lembar saham biasa yang beredar

3.3 Transformasi Data

Transformasi data diperlukan untuk membuat rentang data yang sebelumnya tidak sama untuk setiap variabel menjadi seragam sesuai dengan rentang yang telah ditentukan. Teknik yang digunakan dalam melakukan transformasi data dalam penelitian ini adalah scaling. Dalam penelitian ini akan digunakan skala(-1,1), dimana batas bawah (BB) adalah -1 dan batas atas (BA) adalah 1.

Jika nilai maksimum tiap kolom adalah Xmax dan nilai minimalnya adalah Xmin maka untuk mengubah data dalam skala baru adalah (Santoso,2007).

$$\hat{x} = \frac{X - X_{min}}{X_{max} - X_{min}} \times (BA - BB) + BB$$

3.4 Seleksi Variabel

Rasio-rasio finansial yang terdapat dalam tabel 3.1 akan dijadikan variabel input dalam memprediksi kondisi kesulitan keuangan perusahaan. Variabel ini akan digunakan sebagai input variabel dalam proses seleksi variabel. Metode seleksi variabel yang akan digunakan adalah LP-SVM.

3.5 Pembagian Data Training dan Data Testing

Pembagian data training dan data testing dilakukan dengan proporsi yang sama, dimana data tahun 2011 dan tahun 2012 digunakan untuk data training sedang data testing menggunakan data tahun 2013. Jumlah data yang diambil sebanyak 60 perusahaan.

3.6 Pengujian model Menggunakan SVM

Untuk pengujian model dengan SVM diambil dengan beberapa jenis kernel yaitu : Kernel linear, Kernel Polynomial dan Kernel RBF. Pengujian dilakukan pada dua jenis data yaitu ; Data awal sebelum seleksi variabel dan data yang sudah dilakukan seleksi variabel. Hasil dari klasifikasi selanjutnya dibandingkan persentase misklasifikasi dari tiap jenis Kernel beserta parameternya. Adapun parameter yang digunakan adalah P=0,5,1,,2 dan konstanta C=1,10.

4. Hasil dan Pembahasan

Hasil dari seleksi variabel menggunakan LP-SVM adalah sebagai berikut .Dari 16 variabel terdapat 10 variabel yang selalu muncul pada setiap tahun amatan yaitu : *Current Ratio, Account Receivable Turnover, Inventory Turnover Rate, Debt to Equity, Leverage ratio, Gross Profit Margin, Return on Equity, Return on Assets, Fixed Asset turnover rate, Earning per share (EPS)*.

Hasil uji model menggunakan SVM dapat dilihat pada tabel berikut :

Tabel 2 Hasil penentuan model dengan menggunakan variabel yang belum diseleksi

Kernel	Misklasifikasi (%)	Kernel	Misklasifikasi (%)
Linear,C=1	29	Linear,C=10	34
Polynomial derajat 2, C=1	30	Polynomial derajat 2, C=10	30
Polynomial derajat 3, C=1	45	Polynomial derajat 3, C=10	45
RBF, $\sigma=0,5,C=1$	55	RBF, $\sigma=0,5,C=10$	67
RBF, $\sigma=1,C=1$	42	RBF, $\sigma=1,C=10$	65
RBF, $\sigma=2,C=1$	33	RBF, $\sigma=2,C=10$	37

Tabel 3 Hasil penentuan model dengan menggunakan variabel yang telah diseleksi

Kernel	Misklasifikasi (%)	Kernel	Misklasifikasi (%)
Linear,C=1	29	Linear,C=10	29
Polynomial derajat 2, C=1	30	Polynomial derajat 2, C=10	30
Polynomial derajat 3, C=1	35	Polynomial derajat 3, C=10	35
RBF, $\sigma=0,5,C=1$	38	RBF, $\sigma=0,5,C=10$	57
RBF, $\sigma=1,C=1$	32	RBF, $\sigma=1,C=10$	55
RBF, $\sigma=2,C=1$	28	RBF, $\sigma=2,C=10$	32

Bila dibandingkan antara hasil percobaan pada tabel 3 dan 4 dapat disimpulkan bahwa tingkat misklasifikasi semakin kecil dengan menggunakan variabel yang telah diseleksi menggunakan LP-SVM. Tingkat misklasifikasi terkecil diberikan oleh Kernel RBF dengan parameter , $\sigma=2,C=1$ sebesar 28%.

Pengujian model juga dilakukan dengan menggunakan LDA. Hasil misklasifikasi dari model adalah sebagai berikut.

Tabel 4. Hasil Uji Model Dengan Menggunakan LDA

	Misklasifikasi (%)
LDA , sebelum seleksi variabel	38
LDA, setelah seleksi variabel	30

5. Kesimpulan dan Saran

Kesimpulan dari penelitian ini adalah sebagai berikut :

1. Dari proses seleksi variabel menggunakan LP-SVM didapatkan 10 variabel sebagai variabel prediktor.
2. Proses seleksi variabel dengan menggunakan LP-SVM dapat memperkecil tingkat misklasifikasi pada model prediksi financial distress
3. Pada percobaan prediksi financial distress dengan menggunakan SVM dengan seleksi variabel, misklasifikasi terkecil diberikan oleh Kernel RBF dengan parameter , $\sigma=2,C=1$
4. Pada percobaan prediksi financial distress dengan menggunakan LDA ,misklasifikasi terkecil terjadi pada variabel yang telah diseleksi.

Saran untuk penelitian selanjutnya adalah untuk variabel yang digunakan dapat ditambahkan dengan variabel kualitatif lainnya dengan metode seleksi variabel yang berbeda.

6. Daftar Referensi

- [1] Altman, E. I., "Financial Ratios Discriminant Analysis and The Prediction of Corporate Bankruptcy", *Journal of Finance*, Vol. 23, hal. 1-22,1968.
- [2] Martin, D., "Early Warning of Bank Failure: A Logit Regression Approach", *Journal of Banking and Finance*, Vol. 1, hal. 249-276, 1997.
- [3] Delen, D.,Kuzey, C. dan Uyar, A. "Measuring Firm Performance Using Financial Ratio: A Decision Approach", *Expert Systems with Applications*, Vol. 40, hal. 3970-3983,2013.
- [4] Gong, R. Dan Huang, S. H., "A Kolmogorov-Smirnov Statistic Based Segmentation Approach to Learning from Imbalanced Datasets: With Application in Property Refinance Prediction", *Expert Systems with Applications*, Vol. 39, hal. 6192-6200,2012.
- [5] Nisa, U. Z., "Model Prediksi Finansial Distress Pada Perusahaan Manufaktur Go Public di Indonesia", Thesis Program Magister Bidang Optimasi Sistem Industri Jurusan Teknik Industri Institut Teknologi Sepuluh Nopember,2013.
- [6] Chen, M-Y., "Bankruptcy Prediction in Firms with Statistical and Intelligent Techniques and a Comparison of Evolutionary Approaches", *Computers and Mathematics with Applications*, Vol. 62, hal. 4514-4524,2011
- [7] Santosa, B.. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*, Graha Ilmu, Yogyakarta,2007.
- [8] Tsai,C.F, "Feature Selection in Bancrupy Prediction", *Knowledge Based System*,Vol 22,hal 120-127,2009.