

**PERBANDINGAN ANALISIS DISKRIMINAN LINEAR, REGRESI
LOGISTIK BINER DAN
RADIAL BASIS FUNCTION NEURAL NETWORK (RBFNN)
(STUDI KASUS PADA PENGKLASIFIKASIAN KETEPATAN
WAKTU PENYAMPAIAN LAPORAN KEUANGAN PERUSAHAAN
MANUFAKTUR DI BEI)**

Erna Hayati

Fakultas Ekonomi, Universitas Islam Lamongan

ABSTRAK

Terdapat beberapa metode statistika yang dapat digunakan untuk mengklasifikasikan objek ke dalam kelompok-kelompok, tiga diantaranya yaitu metode analisis diskriminan linear, regresi logistik biner dan RBFNN. Jenis data yang digunakan dalam analisis diskriminan linear adalah data non-metrik untuk variabel respon dan data metrik untuk variabel bebasnya. Sedangkan regresi logistik biner dan RBFNN bisa digunakan pada data yang berjenis non-metrik untuk variabel respon dan data metrik atau non-metrik pada variabel bebasnya. Tujuan penelitian ini adalah membandingkan ketepatan klasifikasi ketiga metode tersebut pada studi kasus pengklasifikasian ketepatan waktu penyampaian laporan keuangan perusahaan. Data yang digunakan dalam penelitian ini adalah data sekunder yang berupa laporan keuangan perusahaan manufaktur yang terdaftar di BEI. Data tersebut dibagi menjadi data training sebesar 70% dan data testing sebesar 30%. Hasil penelitian menunjukkan bahwa prosentase ketepatan klasifikasi pada data training dan data testing untuk metode analisis diskriminan linear lebih besar dibandingkan dengan metode regresi logistik biner dan RBFNN, sehingga dapat disimpulkan bahwa metode analisis diskriminan linear merupakan metode terbaik dalam mengklasifikasikan ketepatan waktu penyampaian laporan keuangan perusahaan.

Kata Kunci : Analisis Diskriminan Linear, Klasifikasi, Regresi Logistik Biner, RBFNN.

1. PENDAHULUAN

Analisis diskriminan linear, regresi logistik biner dan *radial basis fuction* (RBFNN) merupakan metode statistika yang dapat digunakan untuk mengklasifikasikan atau mengelompokkan suatu objek ke dalam kelompok-

kelompok berdasarkan variabel bebasnya. Dalam analisis diskriminan linear, jenis data yang bisa digunakan dalam analisis ini adalah non-metrik pada variabel respon dan metrik pada variabel bebasnya. Sedangkan pada metode regresi logistik dan RBFNN, jenis data

yang digunakan adalah non metrik untuk variabel respon dan data metrik atau non metrik pada variabel bebasnya.

Menurut Pohar, *et al.*(2004) analisis diskriminan linear lebih baik dalam mengklasifikasikan objek dibandingkan regresi logistik jika data berdistribusi normal, namun jika data berdistribusi non normal, maka regresi logistik lebih baik dibandingkan analisis diskriminan linear. Sedangkan menurut Marino dan Yuliani (2014) regresi logistik biner memiliki ketepatan klasifikasi yang lebih baik dibandingkan analisis diskriminan linear pada pengelompokan atau penjurusan IPA atau IPS di SMA Negeri 1 Bangorejo Banyuwangi. Penelitian yang dilakukan Arif, dkk (2014) menyimpulkan bahwa RBFNN dapat mengklasifikasikan peubah respon biner lebih baik dibandingkan regresi logistik biner.

Pada penelitian ini ingin dibandingkan ketepatan klasifikasi metode analisis diskriminan linear, regresi logistik biner dan RBFNN pada studi kasus pengklasifikasian ketepatan waktu pelaporan keuangan perusahaan manufaktur di BEI. Ketepatan waktu penyampaian laporan keuangan merupakan hal yang sangat penting bagi tingkat kemanfaatan dan nilai laporan

tersebut. Ketepatan waktu penyampaian laporan keuangan merupakan dasar bagi para investor untuk menentukan kebijakan di masa yang akan datang apakah akan membeli atau menjual kepemilikan sahamnya. Berdasarkan ketepatan waktu penyampaian laporan keuangan, perusahaan dikelompokkan menjadi dua yaitu tepat waktu dan tidak tepat waktu. Data inilah yang digunakan sebagai respon dalam penelitian ini. Sedangkan variabel yang mempengaruhi ketepatan penyampaian laporan keuangan yang digunakan dalam penelitian ini adalah likuiditas dan leverage keuangan.

2. TINJAUAN PUSTAKA

2.1 Analisis Diskriminan Linear

Analisis diskriminan merupakan salah satu metode dalam statistika yang dapat digunakan untuk mengevaluasi klasifikasi objek. Tujuan dari analisis ini adalah untuk mendapatkan suatu fungsi yang disebut dengan fungsi diskriminan, dimana fungsi ini dapat memisahkan objek sesuai dengan grupnya. Fungsi diskriminan ini juga dapat digunakan untuk memprediksi grup dari objek baru yang diamati (Sharma, 1996).

Jenis data yang digunakan pada analisis ini adalah data non metrik untuk

variabel dependen atau responnya dan data metrik pada variabel bebasnya. Asumsi yang harus dipenuhi dalam analisis diskriminan linear adalah data harus berdistribusi normal multivariate dan matrik varian harus sama. Bentuk umum fungsi diskriminan linear adalah sebagai berikut (Johnson dan Dean, 2002):

$$y = (\bar{x}_1 - \bar{x}_2)' S^{-1} \mathbf{x} \quad (1)$$

Dimana:

\bar{x}_1 = rata-rata kelompok pertama

\bar{x}_2 = rata-rata kelompok ke dua

S^{-1} = invers kovarian gabungan

2.2 Regresi Logistik Biner

Regresi logistik biner adalah regresi yang variabel responnya terdiri dari dua kategori sedangkan jenis data variabel bebasnya bisa berupa data metrik (interval dan ratio) dan data non metrik (nominal dan ordinal). Asumsi yang harus dipenuhi dalam regresi logistik biner adalah ketiadaan multikolinearitas antar variabel bebasnya (Hosmer dan Lemeshow, 1989). Bentuk umum regresi logistik biner adalah sebagai berikut:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)} \quad (2)$$

Dimana k adalah banyaknya variabel bebas.

2.3 Radial Basis Function Neural Network (RBFNN)

Neural Network atau jaringan syaraf tiruan merupakan model nonparametrik untuk pemrosesan informasi yang terinspirasi oleh sistem sel syaraf biologi, sama seperti otak yang memproses suatu informasi (Siang, 2005). Model *Neural Network* mempunyai beberapa desain diantaranya yaitu *Perceptron*, *Feedforward Neural Network* (FFNN), *Radial Basis Function Neural Network* (RBFNN) dan lain-lain.

RBFNN adalah salah satu model *Neural Network* dengan satu unit dalam lapisan tersembunyi, di mana fungsi aktivasinya adalah fungsi radial basis dan fungsi linier pada lapisan output. Fungsi radial basis yang sering digunakan adalah fungsi gaussian karena fungsi ini mempunyai sifat lokal, yaitu jika input dekat dengan rata-rata, maka fungsi akan menghasilkan nilai satu, sedangkan jika input jauh dari rata-rata, maka fungsi memberikan nilai nol.

2.4 Ketepatan Waktu Penyampaian Laporan Keuangan

Tujuan laporan keuangan adalah menyediakan informasi yang menyangkut posisi keuangan, kinerja, serta perubahan posisi keuangan suatu perusahaan yang bermanfaat bagi sejumlah besar pengguna dalam pengambilan keputusan ekonomi (IAI, 2007). Suatu informasi akan bermanfaat bagi para pemakai apabila tersedia tepat waktu sehingga pemakai tidak kehilangan kesempatan atau kemampuan untuk mempengaruhi keputusan yang akan diambil. Menurut Baridwan (1997) tepat waktu diartikan bahwa informasi harus disampaikan sedini mungkin untuk dapat digunakan sebagai dasar untuk membantu dalam pengambilan keputusan ekonomi dan untuk menghindari tertundanya pengambilan keputusan tersebut.

Ada beberapa faktor yang mempengaruhi ketepatan waktu penyampaian laporan keuangan suatu perusahaan. Dalam penelitian ini ada dua faktor yang diajukan yaitu likuiditas dan leverage. Likuiditas adalah ketersediaan sumber daya atau kemampuan suatu perusahaan dalam memenuhi kewajiban jangka pendeknya yang telah jatuh tempo, dengan cara melihat aset lancar

perusahaan relatif terhadap hutang lancarnya. Rasio likuiditas yang biasanya digunakan untuk mengukur tingkat likuiditas suatu perusahaan yaitu rasio lancar dan rasio *Quick* (Hanafi dan Halim, 2005). Sedangkan rasio leverage merupakan rasio pengungkit yang menggunakan uang pinjaman (*debt*) untuk memperoleh keuntungan (Ang, 1997). Rasio leverage ini diprosikan dengan *debt to equity ratio* (DER).

3. METODOLOGI PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder yang berupa laporan keuangan 20 perusahaan manufaktur yang terdaftar di BEI. Data laporan keuangan yang digunakan adalah data tahun 2011 sampai 2014. Data ini diperoleh dari situs resmi Bursa Efek Indonesia (BEI) yaitu www.idx.co.id.

Variabel respon (Y) dalam penelitian ini adalah ketepatan waktu pelaporan keuangan perusahaan. Variabel ini terdiri dari dua kategori yaitu kategori 1 untuk perusahaan yang tepat waktu dan kategori 0 untuk perusahaan yang tidak tepat waktu. Sedangkan untuk variabel bebasnya ada dua yaitu likuiditas dan leverage

keuangan. Kedua variabel bebas ini jenis datanya adalah data metrik.

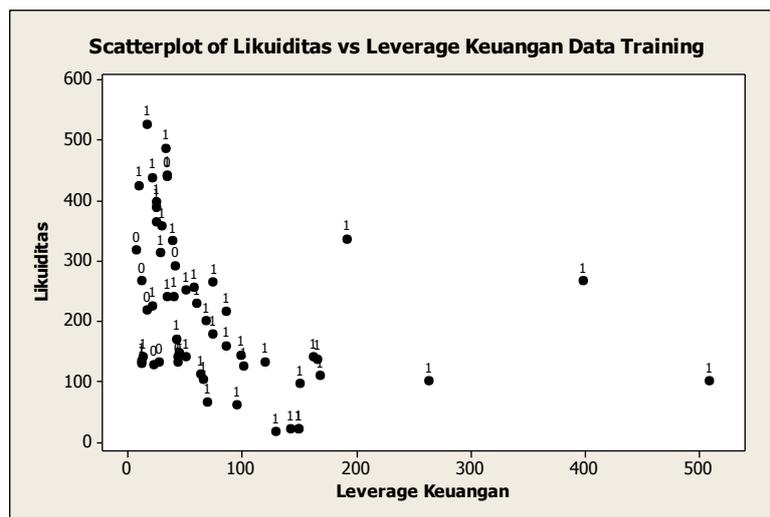
Data laporan keuangan tersebut, untuk selanjutnya dibagi menjadi dua yaitu sebanyak 70% data sebagai data *training* dan 30% data sebagai data *testing*. Langkah pertama yang dilakukan dalam penelitian ini adalah membuat model fungsi diskriminan linear, regresi logistik dan RBFNN pada data *training*, kemudian model ini digunakan untuk

mengklasifikasikan data *testing*. Pengolahan data dalam penelitian ini menggunakan bantuan *software* SPSS 20 dan R 2.1.4.

4. HASIL PENELITIAN DAN PEMBAHASAN

4.1 Diskripsi Data

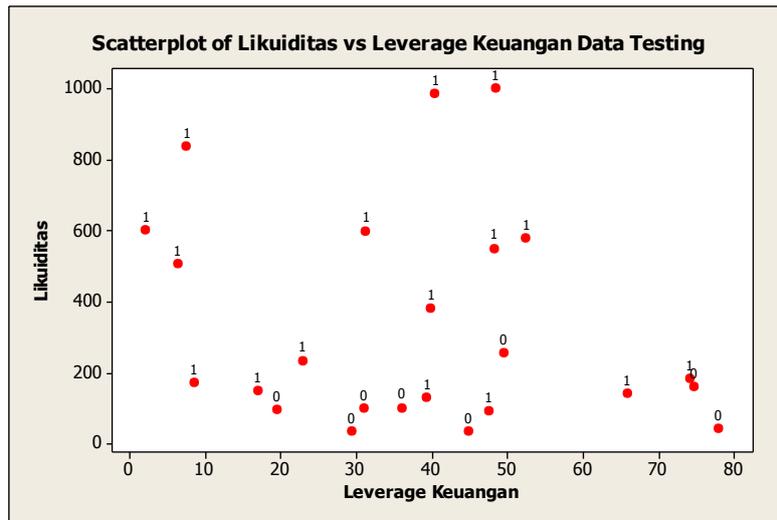
Berikut ini adalah gambar *scatterplot* data *training*:



Gambar 1. Scatterplot Data Training

Dari Gambar 1 di atas, dapat kita ketahui bahwa pada data *training* terjadi *overlapping*, dimana kategori 1 yaitu tepat waktu berbaur dengan kategori 0 yaitu tidak tepat waktu. Kategori 1 dan 2

hampir tidak memiliki perbedaan yang nyata. Sedangkan berikut ini merupakan gambar *scatterplot* dari data *testing*:



Gambar 2. Scatterplot Data Testing

Berdasarkan gambar di atas dapat diketahui bahwa terjadi *overlapping* juga pada data *testing* seperti halnya pada data *training*.

4.2 Pembentukan Fungsi Diskriminan Linear

Dengan menggunakan *software* R 2.1.4 , diperoleh fungsi diskriminan linear dari data *training* adalah sebagai berikut:

$$y = 0.0001201104 \text{ Likuiditas} + 0.0112124354 \text{ Leverage Keuangan}$$

Fungsi tersebut digunakan untuk mengklasifikasikan data *testing*. Berikut ini adalah hasil ketepatan klasifikasi data *training* dan *testing* yang dihasilkan dari fungsi diskriminan linear tersebut:

Tabel 1. Ketepatan Klasifikasi Data *Training* dan *Testing* pada Analisis Diskriminan Linear

Ketepatan Klasifikasi	
<i>Training</i>	<i>Testing</i>
85,71%	66,67%

Dari tabel di atas dapat diketahui bahwa ketepatan klasifikasi data *training* 85,71%, sedangkan pada data *testing* sebesar 66,67%.

4.3 Pembentukan Regresi Logistik Biner

Sebelum membentuk model regresi logistik biner, terlebih dahulu dilakukan uji secara serentak. Uji ini digunakan untuk melihat apakah terdapat peubah yang berpengaruh terhadap model. Pengujian secara serentak dengan menggunakan uji rasio *likelihood* (G). Dari hasil output SPSS diketahui bahwa

nilai $-2 \text{ Log Likelihood Block Number} = 0$ (44,356) mengalami penurunan pada nilai $-2 \text{ Log Likelihood Block Number} = 1$ (34,377). Dari penurunan nilai ini dapat kita simpulkan bahwa sedikitnya ada satu peubah bebas yang berpengaruh terhadap peubah respon. Maka model regresi logistik pada data *training* yang dapat dibentuk adalah sebagai berikut:

$$\text{Ln} \left(\frac{\pi}{1-\pi} \right) = -1.513507 + 0.003898 \text{ Likuiditas} + 0.056806 \text{ Leverage Keuangan}$$

Selanjutnya dilakukan uji kelayakan model (*goodness of fit*). Pengujian kelayakan model atau *goodness of fit* digunakan untuk mengetahui apakah model yang dihasilkan sudah sesuai atau belum. Hasil uji kelayakan model menunjukkan model regresi logistik biner yang digunakan sesuai atau telah cukup mampu menjelaskan data karena nilai *p-value Chi-square* lebih dari 0,05 ($0,649 > 0,05$).

Model regresi logistik biner yang terbentuk digunakan untuk mengklasifikasikan data *testing*. Hasil ketepatan klasifikasi data *training* dan *testing* yang dihasilkan dari regresi logistik biner tersebut dapat dilihat pada tabel berikut ini:

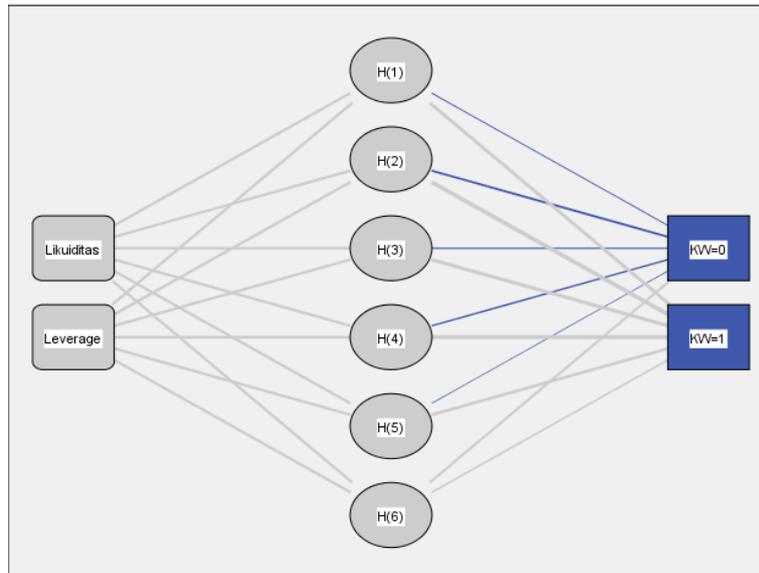
Tabel 2. Ketepatan Klasifikasi Data *Training* dan *Testing* pada Regresi Logistik Biner

Ketepatan Klasifikasi	
<i>Training</i>	<i>Testing</i>
80,36%	66,67%

Dari tabel di atas dapat diketahui bahwa ketepatan klasifikasi pada regresi logistik biner untuk data *training* 80,36%, sedangkan pada data *testing* sebesar 66,67%.

4.4 Pembentukan Model *Radial Basis Function Neural Network* (RBFNN)

Berikut ini adalah gambar arsitektur dari model RBFNN data *training*:



Gambar 3. Arsitektur Model RBFNN Data *Training*

Model RBFNN terbaik untuk data *training* adalah pada neuron ke 6, sehingga model arsitektur jaringan yang terbentuk 2-6-1. Model RBFNN yang telah terbentuk selanjutnya digunakan untuk mengklasifikasikan data *testing*.

Tabel 3. Ketepatan Klasifikasi Data *Training* dan *Testing* pada RBFNN

Ketepatan Klasifikasi	
<i>Training</i>	<i>Testing</i>
85,71%	62,5%

Berdasarkan Tabel 3 di atas dapat diketahui bahwa ketepatan klasifikasi pada model RBFNN untuk data *training* sebesar 85,71% dan pada data *testing* sebesar 62,5%.

4.5 Pemilihan Metode Terbaik

Berdasarkan Gambar 1 dan Gambar 2 dapat diketahui bahwa data *training* dan data *testing* mengalami *overlapping*, dimana kategori tepat waktu (1) dan kategori tidak tepat waktu (0) berbaur dan hampir tidak dapat dibedakan kategori yang satu dengan kategori yang lain. Oleh karena itu dibutuhkan suatu metode yang terbaik untuk mengklasifikasikan dengan tepat kedua kategori tersebut. Dalam penelitian ini terdapat tiga metode yang diusulkan dalam mengklasifikasikan data ketepatan waktu penyampaian laporan keuangan yaitu analisis diskriminan linear, regresi logistik biner dan *Radial Basis Function Neural Network* (RBFNN).

Berikut ini adalah tabel hasil perbandingan ketepatan klasifikasi tiga metode tersebut:

Tabel 4. Perbandingan Ketepatan Klasifikasi Data *Training* dan *Testing*

Metode	Ketepatan Klasifikasi	
	<i>Training</i>	<i>Testing</i>
Analisis Diskriminan Linear	85,71%	66,67%
Regresi Logistik Biner	80,36%	66,67%
RBFFNN	85,71%	62,5%

Berdasarkan Tabel 4 dapat kita ketahui bahwa prosentase ketepatan klasifikasi untuk metode analisis diskriminan linear nilainya paling besar untuk semua data (*training* dan *testing*). Maka dapat disimpulkan bahwa metode analisis diskriminan merupakan metode terbaik dalam mengklasifikasikan perusahaan menurut ketepatan waktu penyampaian laporan keuangan perusahaan.

5. PENUTUP

5.1 Simpulan

Berdasarkan hasil analisis dan pembahasan dapat disimpulkan bahwa analisis diskriminan linear merupakan metode terbaik dibandingkan dengan metode regresi logistik biner dan

RBFFNN dalam mengklasifikasikan perusahaan berdasarkan ketepatan waktu penyampaian laporan keuangan. Hal ini bisa dilihat dari prosentase ketepatan klasifikasi yang besar pada metode analisis diskriminan untuk semua data, baik data *training* maupun data *testing*.

5.2 Saran

Penelitian ini hanya menggunakan dua variabel yang mempengaruhi ketepatan waktu penyampaian laporan keuangan. Untuk penelitian selanjutnya dapat menambahkan variabel yang lainnya.

DAFTAR PUSTAKA

- Ang, Robert. 1997. *Buku Pintar Pasar Modal Indonesia*. Mediasoft Indonesia.
- Arif, N.F, Samingun, H. dan Henny, P. 2014. Perbandingan Model Regresi Logistik dan Model *Radial Basis Function Neural Network* untuk Pengklasifikasian Peubah Respon Biner. *Jurnal Mahasiswa Statistik*. Vol. 2, No.1: 29-32.
- Baridwan, Zaki. 1997. *Intermediate Accounting*. Edisi Tujuh. Cetakan Pertama. Yogyakarta: BPFE.

- Hanafi, Mamduh M. dan Halim, A. 2005. *Analisis Laporan Keuangan*. Edisi Kedua. Yogyakarta : UPP AMP YKPN.
- Hosmer, D. W. dan Lemeshow, S..1989. *Applied Logistic Regression*. New York : John Willey and Sons, Inc.
- Ikatan Akuntan Indonesia. 2007. *Standar Akuntansi Keuangan*. Jakarta: Penerbit Salemba Empat.
- Johnson, R.A dan Dean, W.W. 2002. *Applied Multivariate Statistical Analysis*, 5th edition. New Jersey: Prentice Hall.
- Marino, I.M.T. dan Yuliani, S.D. 2014. Perbandingan Analisis Diskriminan Linear, Diskriminan *Robust* dan Regresi Logistik Biner (Studi Kasus Pada Jurusan IPA/IPS Siswa Tingkat SMA Negeri 1 Bangorejo Banyuwangi). *Prosiding Seminar Nasional Matematika*. Universitas Jember, hal. 192-200.
- Pohar, M, Mateja, B. dan Sandra, T. 2004. Comparison of Logistic Regression and Linear Discriminant Analysis : A Simulation Study, *Metodološki zvezki*, Vol. 1, No. 1 : 143-161.
- Sharma, S. 1996. *Applied Multivariate Techniques*. New York: John Wiley & Sons.
- Siang, J.J. 2005. *Jaringan Syaraf Tiruan Dan Pemrogramannya Menggunakan MATLAB*. Yogyakarta : Andi Offset.