

PREDIKSI PENDAPATAN SEWA DENGAN DATA MINING PADA PERUSAHAAN XYZ

Meyliana; Christine Sanjaya; Agus Widodo; Marshall Martinus

Jurusan Sistem Informasi, Fakultas Ilmu Komputer, Bina Nusantara University
Jln. K.H. Syahdan No 9, Palmerah, Jakarta Barat 11480
meyliana@binus.edu

ABSTRACT

XYZ Company has a program to predict leasing income that only predict in constant condition where every tenant assumed for leasing renewal. This research is done to build accurate income prediction system that accommodate in making strategic decision towards the company. Premier data collecting is through direct interview with the company management. The analysis is through data training from the previous years to build neural network model. The analysis result shows that this model has produced error total value that is smaller than the previous error total value in years before. Therefore, it could be concluded that data mining with neural network technique that produced more accurate leasing income that could help the company making decision based on the hidden information in the database.

Keywords: *leasing income prediction, data mining, neural network, backpropagation, Levenberg Marquardt, CRISP-DM.*

ABSTRAK

Perusahaan XYZ mempunyai program untuk memprediksi pendapatan sewa yang hanya memprediksi pendapatan sewa dalam keadaan konstan di mana semua tenant dianggap pasti melakukan perpanjangan sewa. Penelitian ini dilakukan untuk membangun sistem prediksi pendapatan yang lebih akurat sehingga dapat memudahkan dalam pengambilan keputusan strategis terhadap perusahaan XYZ tersebut. Pengumpulan data primer dilakukan berdasarkan wawancara langsung dengan pihak manajemen perusahaan. Analisis dilakukan dengan melakukan training pada data beberapa tahun yang lalu untuk membangun model neural network. Hasil analisis menunjukkan bahwa model ini telah menghasilkan total nilai error yang jauh lebih kecil daripada total nilai error sebelumnya dengan menemukan pola dari data di tahun-tahun sebelumnya. Karenanya, dapat disimpulkan bahwa data mining dengan teknik neural network dapat menghasilkan prediksi pendapatan sewa yang lebih tepat sehingga dapat membantu perusahaan XYZ dalam pengambilan keputusan strategis berdasarkan informasi tersembunyi dari database yang ada.

Kata kunci: *prediksi pendapatan sewa, data mining, neural network, backpropagation, Levenberg marquardt, CRISP-DM.*

PENDAHULUAN

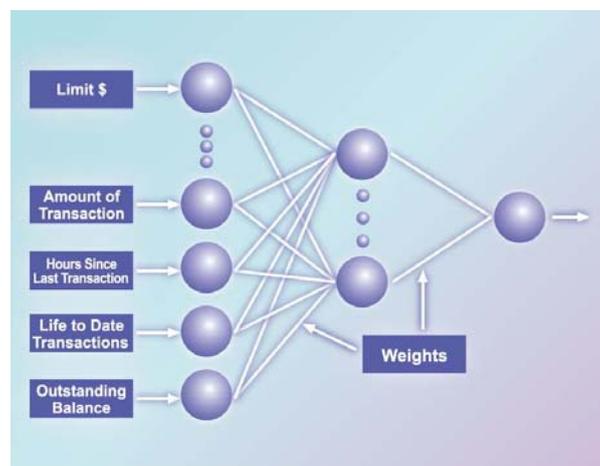
PT. XYZ adalah sebuah perusahaan yang bergerak di bidang properti, healthcare dan rumah sakit. Perusahaan ini berdiri sejak 1990 dan pada tahun 2004 telah menjadi pemimpin perusahaan properti di Indonesia dengan asset yang cukup besar. Dengan bertambahnya jumlah mall yang dimiliki, maka pengawasan dan perawatan untuk tiap mall harus semakin ditingkatkan, terutama pengawasan jumlah tenant yang menyewa tempat dan besarnya transaksi yang dihasilkan tiap mall.

PT. XYZ telah melakukan prediksi besarnya pendapatan sewa pada tahun-tahun mendatang dengan mewajibkan tiap mall untuk membuat laporan budget yang berisi laporan jumlah *tenant*, *cash flow*, *profit and loss* dan prediksi pendapatan sewa mulai dari tahun berjalan sampai 10 tahun kedepan untuk dikirimkan ke pusat. Pihak manajemen perlu menganalisa laporan tersebut untuk mengetahui kapan pencapaian BEP (Break Even Point) dari pembangunan mall, untuk mengambil keputusan dalam perluasan mall/ekspansi mall di daerah lain ataupun untuk pengambilan keputusan strategis lainnya.

Prediksi pendapatan sewa pada laporan tersebut hanyalah berdasarkan asumsi bahwa semua tenant pasti memperpanjang sewa. Dengan demikian didapatkan prediksi pendapatan sewa yang konstan pada tahun-tahun mendatang. Padahal dalam kenyataannya belum tentu semua tenant memperpanjang sewa. Di sinilah data mining dapat membantu PT. XYZ mengekstrak pola yang terbentuk pada tahun-tahun sebelumnya, sehingga dapat dilakukan prediksi pendapatan sewa yang lebih tepat.

Penelitian ini akan membahas bagaimana *data mining* akan menganalisa beberapa parameter yaitu faktor-faktor yang mempengaruhi pendapatan sewa untuk menghasilkan prediksi pendapatan sewa yang lebih akurat sehingga dapat mendukung manajemen PT. XYZ dalam pengambilan keputusan strategis.

Penelitian akan dilakukan dengan mengumpulkan parameter-parameter yang menentukan dalam prediksi pendapatan sewa, membangun model data mining untuk prediksi pendapatan sewa, menguji coba model data mining pada data PT. XYZ, mengevaluasi hasil uji coba model data mining pada data PT. XYZ. Teknik data mining yang digunakan adalah *Neural Network* dengan algoritma *Backpropagation Levenberg Marquardt* (Gambar 1).



Gambar 1 Arsitektur Neural Network

Neural Network adalah sistem parallel computing dari beberapa prosesor yang terhubung. *Neural Network* menawarkan model matematika yang berusaha meniru otak manusia. *Backpropagation* adalah generalisasi dari aturan pembelajaran Widrow-Hoff pada jaringan multiple-layer dan fungsi transfer nonlinear yang berbeda-beda. *Vektor input* dan *vektor target* yang berhubungan digunakan untuk melatih jaringan sampai dapat mendekati sebuah fungsi, yang menghubungkan *vektor input* dengan vektor output yang spesifik atau mengklasifikasikan *vektor input* ke dalam cara yang sesuai yang telah didefinisikan. Training *network* dengan *backpropagation* melibatkan 3 tahap, yaitu *feedforward* untuk menentukan pola training input, *backpropagation error* yang berhubungan, dan meakukan penyesuaian pada *weight*.

Algoritma training *backpropagation*, yaitu (1) *batch gradient descent*, ada 7 parameter yang berhubungan dengan algoritma ini: epochs, show, goal, time, min_grad, max_fail, learning rate (lr); (2) *gradient descent with momentum*, yang memungkinkan *network* tidak hanya memberi respon pada gradien lokal tapi juga pada tren yang terjadi pada error yang dangkal. Tanpa momentum, *network* dapat terjebak dalam minimum lokal yang dangkal. Dengan momentum, *network* dapat mencapai minimum. Algoritma ini tergantung pada 2 parameter training: *learning rate* (lr), *momentum constant* (mc).

Kedua algoritma training *backpropagation* di atas seringkali terlalu lambat untuk masalah praktis. Ada beberapa algoritma yang lebih cepat dari kedua algoritma di atas. Pada kesempatan ini hanya akan dibahas mengenai algoritma Levenberg Marquardt, karena algoritma ini merupakan metode tercepat untuk training feedforward neural network berukuran besar (sampai ratusan *weight*). Algoritma Levenberg Marquardt menggunakan pendekatan ini untuk matrix Hessian sebagai berikut:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}]^{-1} \mathbf{J}^T \mathbf{e}$$

Keterangan rumus:

x_{k+1} = bobot selanjutnya (pada epoch k+1)

x_k = bobot sekarang (pada epoch k)

J = matrix Jacobian yang mengandung turunan pertama dari network error terhadap weight dan bias.

J^T = matrix Jacobian yang di-transpose.

μ = learning rate parameter

I = matrix identitas

e = vektor dari network error.

Evaluasi model yang telah dibuat dilakukan dengan MSE (*Mean Square Error*) dan Regresi Linear.

$$mse = \frac{1}{n} \sum_{k=1}^n (t(k) - a(k))^2$$

Keterangan rumus:

n = jumlah kumpulan instance training

k = jumlah node output

$t(k)$ = target output untuk instance ke-n dan node output k

$a(k)$ = output hasil perhitungan untuk instance ke-n dan node output k

Regresi Linear

Regresi nilai R mengukur korelasi antara output dengan target. Nilai $R=1$ mempunyai arti bahwa ada hubungan yang dekat. Nilai $R = 0$ mempunyai arti bahwa hubungannya random.

Rumusan Permasalahan

PT. XYZ mewajibkan setiap mall untuk memberikan laporan pertanggungjawaban secara berkala. Dan kapan saja laporan tersebut dibutuhkan, laporan harus tersedia berdasarkan data terbaru. Dalam pembuatan laporan ini, divisi yang terlibat adalah Divisi Leasing dan Divisi FAA. Laporan ini meliputi prediksi pendapatan beberapa tahun yang akan datang. Laporan tersebut adalah hasil analisa terhadap data yang ada dengan menggunakan program yang telah berjalan. Namun program ini mempunyai keterbatasan dalam memprediksi pendapatan yaitu tidak dapat menemukan pola dari data yang telah ada. Prediksi hanya dilakukan dalam keadaan konstan di mana semua tenant dianggap pasti melakukan perpanjangan sewa.

Hal ini tentunya sangat tidak tepat karena belum tentu semua tenant pasti melakukan perpanjangan sewa. Memang sebagian besar tenant melakukan perpanjangan sewa. Namun bukan berarti pendapatan akan konstan selama beberapa tahun. Dengan demikian, dibutuhkan sistem prediksi pendapatan yang lebih akurat dalam memprediksi pendapatan sewa sehingga dapat memudahkan dalam pengambilan keputusan strategis terhadap mall tersebut. Contohnya pada suatu kasus, manajemen PT. XYZ hendak menganalisis suatu mall. Tentu mereka memerlukan laporan pertanggungjawaban mall tersebut. Mereka juga perlu melihat prediksi pendapatan sewa pada tahun-tahun mendatang sehingga mereka dapat membuat perencanaan pengembangan mall dan perluasan mall. Bila diperlukan dalam mencapai target pengembangan mall, apakah harga *base rent* perlu dinaikkan/diturunkan.

METODE

Berdasarkan arsitektur data mining CRISP-DM, maka dilakukan tahap-tahap data mining sebagai berikut: *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation* dan *deployment*. Keenam tahap ini dijabarkan lebih detil pada subbab di bawah.

Business Understanding

Keterbatasan program yang telah berjalan dalam memprediksi pendapatan dapat diatasi dengan data mining. Data mining akan dilakukan dengan software MATLAB dari data hasil analisa. Dengan kemampuan menganalisis pola data yang tersembunyi, diharapkan data mining dapat menjawab permasalahan PT. XYZ di atas sehingga dapat dihasilkan prediksi pendapatan sewa yang lebih tepat. Dengan demikian akan mempermudah PT. XYZ dalam mengambil berbagai keputusan strategis seperti menurunkan/menaikkan harga kios, memprediksi waktu tercapainya BEP, memprediksi waktu yang tepat dalam memperluas mall atau membangun mall baru. Berikut adalah tabel Gantt Chart yang menggambarkan tahapan bagaimana data mining dapat menjawab permasalahan PT. XYZ.

Tabel 1 Gantt Chart Data Mining

No.	Tugas	Waktu (hari)	Minggu 1					Minggu 2					Minggu 3					Minggu 4				
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	Bussiness Understanding	4																				
	Analisis Kebutuhan	2	■	■																		
	Definisi Kebutuhan	2		■		■																
2	Data Understanding	2																				
	Entity Relationship Diagram						■	■														
3	Data Preparation	2																				
	Data Selection (Query)								■	■												
4	Modeling	10																				
	Training & Prediksi Neural Network									■	■	■	■	■	■	■	■	■	■	■	■	
5	Evaluation	1																				
	Perbandingan MSE																			■		
6	Deployment	1																				
	Final Report & Presentation																				■	
Total Waktu yang dibutuhkan		20																				

Data Understanding

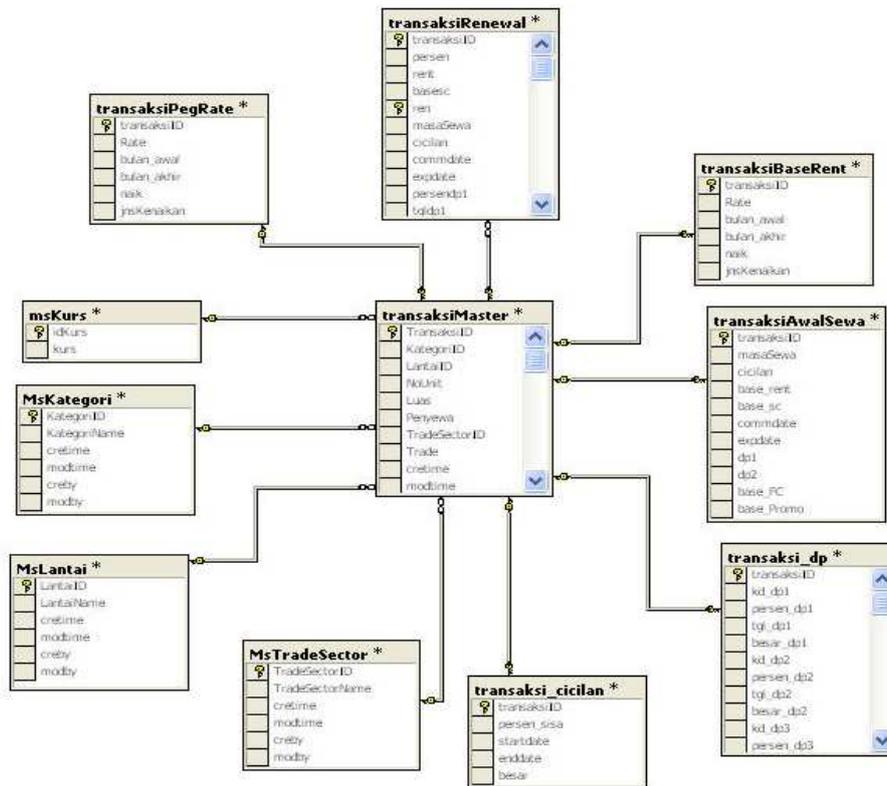
Data yang digunakan dalam penelitian ini terdiri dari data primer dan data sekunder. Data primer adalah data yang diperoleh dari responden melalui wawancara dengan berpedoman pada daftar pertanyaan yang telah dipersiapkan sebelumnya. Pengumpulan data primer dilakukan berdasarkan wawancara langsung dengan pihak manajemen dari PT XYZ. Sedangkan data sekunder adalah data yang diperoleh dari PT XYZ serta dari pustaka yang relevan dengan penelitian. Berikut ini (Gambar 2) adalah relasi antar data yang kami peroleh dari PT. XYZ. Penjelasan atribut dan tipe atribut dari tiap tabel dapat dilihat pada lampiran. Dari data-data yang didapatkan di atas, dipelajari relasinya dan selanjutnya akan dilakukan pengolahan untuk mempersiapkan data yang akan digunakan dalam data mining.

Data Preparation

Dari data yang telah didapatkan dalam tahap *Data Understanding*, selanjutnya dilakukan query untuk mendapatkan variabel-variabel input yang berpengaruh terhadap pendapatan sewa sebagai variabel outputnya. *Query* ini dapat dilihat pada lampiran. Dari variabel-variabel ini, akan dibuat sebuah model dengan teknik *data mining* Neural Network.

Variabel Input

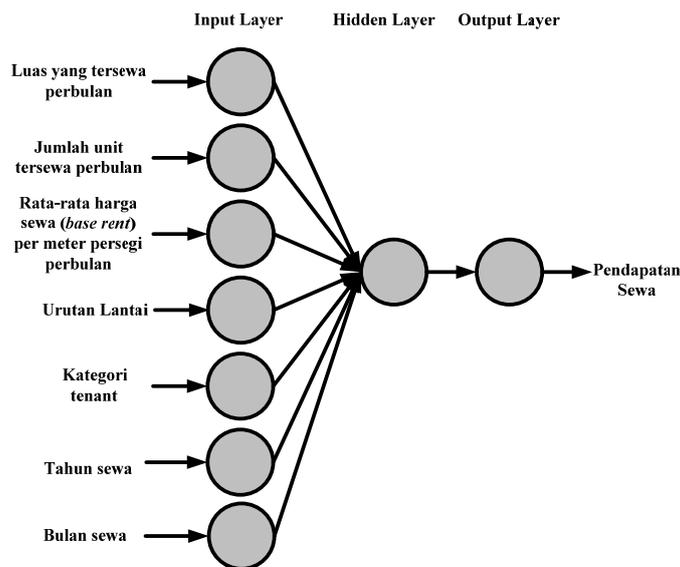
Variabel yang akan diamati untuk memprediksi pendapatan sewa adalah luas yang tersewa perbulan, jumlah unit tersewa perbulan, rata-rata harga sewa (*base rent*) per meter persegi perbulan, urutan lantai, kategori *tenant*, tahun sewa, bulan sewa. Pada kesempatan ini tidak diperhitungkan faktor eksternal (kurs) karena sudah tercakup di dalam variabel input rata-rata harga sewa (*base rent*) per meter persegi perbulan.



Gambar 2 Entity Relationship Diagram dari Data PT. XYZ

Variabel Output

Dari variabel-variabel input di atas, akan diprediksi variabel outputnya yaitu berupa pendapatan sewa.



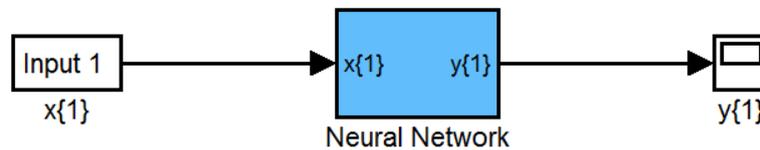
Gambar 3 Variabel Input dan Variabel Output

Modeling

Pada tahap ini, akan dibuat model data mining dengan teknik Neural Network. Sebenarnya data mining memiliki berbagai teknik. Adapun yang paling sesuai untuk permasalahan ini adalah Neural Network dan Regresi Linear. Namun banyak penelitian yang terdahulu telah membuktikan bahwa Neural Network menghasilkan prediksi yang lebih tepat dibandingkan dengan Regresi Linear (Mohd, 2009).

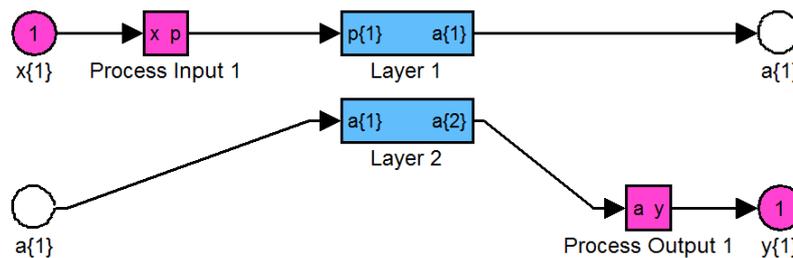
Adapun algoritma yang dipakai adalah algoritma Levenberg Marquardt pada metode training Backpropagation karena algoritma ini merupakan metode tercepat untuk training feedforward neural network berukuran besar (sampai ratusan weight) dan merupakan algoritma tercepat untuk penyelesaian masalah yang menggunakan pendekatan fungsi (Demuth, 2009).

Data dari tahap *Data Preparation* akan digunakan 70% untuk training dan sisanya untuk validasi dan *testing*. Dari hasil training akan didapatkan model neural network yang dapat digunakan untuk mencari output dari input yang baru.



Gambar 4 *Neural Network Model*

Gambar 4 memperlihatkan model utama *Neural Network* secara umum. *Neural Network* akan melakukan proses training dengan algoritma Levenberg-Marquardt sampai dicapai kondisi tertentu yaitu ketika *epoch* (perulangan) maksimum dicapai, waktu pada *time* telah dicapai, *performance* sudah diminimalisasi sesuai *goal*, *performance* gradient di bawah *min_grad*, mu melebihi *mu_max*, *performance* validasi bertambah lebih dari *max_fail* sejak terakhir kali berkurang.

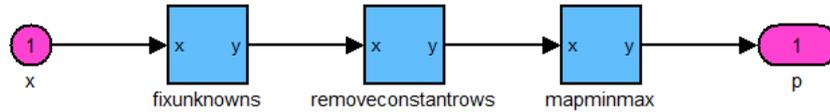


Gambar 5 *Neural Network Model Detail*

Sebelum dilakukan proses training, variabel input akan dinormalisasi terlebih dahulu (*preprocessing*). Setelah proses training selesai variabel output pun akan dikembalikan ke nilai sebenarnya (*postprocessing*). Sedangkan proses training akan terbagi menjadi 2 tahap. *preprocessing*, proses training dan *postprocessing* akan dibahas pada subbab berikut ini.

Preprocessing

Sebelum data digunakan dalam proses training, data perlu dikonversi terlebih dahulu. Pada software MATLAB secara *default*, data pada *network* dengan *backpropagation* training akan dilakukan 3 *preprocessing function*, yaitu: *fixunknowns*, *removeconstantrows* dan *mapminmax*.



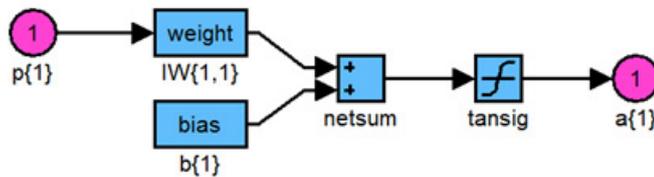
Gambar 6 Preprocessing Function

Proses Training

Pada proses training ini, variabel input akan mengalami 2 tahap, yaitu:

Tahap 1

Proses mengubah nilai input pada layer hidden menjadi nilai input bagi layer output melalui fungsi transfer Hyperbolic Tangent Sigmoid (*tansig*). Di sini kami memilih fungsi transfer *tansig* karena fungsi *tansig* ini menghasilkan prediksi yang lebih akurat daripada *logsig* dan fungsi transfer lainnya (Zhang, 2009).



Gambar 7 Proses Training Neural Network Tahap 1

$p\{1\}$ = variabel input (luas / jumlahUnit / rataHarga / lantai / kategori / tahun / bulan)

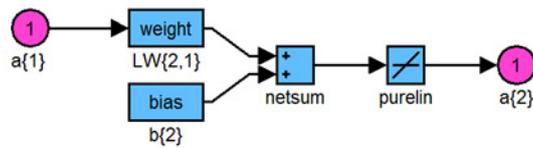
$$\text{netsum} = \text{luas}_1 * \text{weightLuas} + \text{jumlahUnit}_1 * \text{weightJumlahUnit} + \text{rataHarga}_1 * \text{weightRataHarga} + \text{lantai}_1 * \text{weightLantai} + \text{kategori}_1 * \text{weightKategori} + \text{tahun}_1 * \text{weightTahun} + \text{bulan}_1 * \text{weightBulan} + \text{bias}$$

$$a\{1\} = \text{tansig}(\text{netsum}) = \frac{2}{1 + e^{-2 * \text{netsum}}} - 1$$

Proses training Tahap 1 ini akan diulang sebanyak jumlah input. Lalu nilai hasil dari $a\{1\}$ dilanjutkan ke sejumlah layer hidden, untuk selanjutnya diproses lagi pada proses training tahap 2.

Tahap 2

Proses mengubah nilai input pada layer output menjadi nilai output melalui fungsi transfer *Pure Linear*. Fungsi *Pure Linear* ini banyak digunakan pada output layer dan menghasilkan nilai output sesuai dengan yang diharapkan.



Gambar 8 Proses Training *Neural Network* Tahap 2

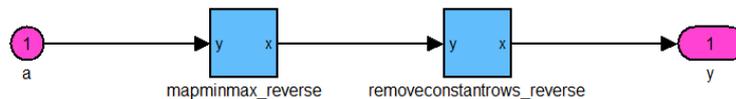
Dari n layer hidden, akan dilakukan perhitungan sbb:

$$\text{netsum} = a\{1\} * \text{weight}_1 + a\{2\} * \text{weight}_2 + \dots + a\{n\} * \text{weight}_n$$

$$a\{2\} = \text{purelin}(\text{netsum}) = \text{netsum}$$

Dari proses training tahap 2 akan dihasilkan nilai output yang selanjutnya akan dikonversi balik ke nilai sebenarnya pada proses *postprocessing*.

Postprocessing



Gambar 9 Postprocessing Function

Setelah data digunakan dalam proses training, data perlu dikonversi balik ke dalam nilai sebenarnya. Pada software MATLAB secara default, data pada network dengan backpropagation training akan dilakukan 2 *postprocessing function*, yaitu: *removeconstantrows_reverse* dan *mapminmax_reverse*.

Evaluation

Pada tahap training, variabel output yang telah disiapkan akan dibandingkan dengan hasil output dari model neural network. Perbandingan tersebut akan dihitung dengan menggunakan rumus MSE (*Mean Square Error*). Semakin kecil nilai MSE-nya, maka semakin baik model *Neural Network* hasil training tersebut.

Deployment

Untuk mendapatkan kinerja model yang lebih baik, tahap Modeling dapat diulang kembali dengan melakukan training lagi dengan menambah jumlah neuron pada hidden layer. Selain itu, kinerja model juga dapat ditingkatkan dengan mengulang dari tahap *Data Preparation* jika memang ada data yang akan ditambahkan.

HASIL DAN PEMBAHASAN

Spesifikasi Perangkat Uji Coba

Uji coba dilakukan pada laptop yang mempunyai spesifikasi, yaitu: (1) spesifikasi *hardware*, berupa Intel(R) Core(TM)2 Duo CPU T6600 @ 2.20 GHz, 2 Gb Memory (RAM),

dan 300 Gb Harddisk; (2) spesifikasi *software*, berupa 32-bit Operating System, Windows Vista™ Ultimate, Matlab Versi 7.7.0.471 (R2008b), dan Microsoft Office Excel 2007.

Data Input

Uji coba dilakukan dengan data input yang terdiri dari 7 variabel yaitu: tahun sewa, bulan sewa, urutan lantai, kategori tenant, jumlah unit tersewa perbulan, luas yang tersewa perbulan dan rata-rata harga sewa (*base rent*) per meter persegi perbulan. Data input ini terdiri dari 1200 data yang merupakan data selama 5 tahun (tahun 2003 sampai 2007). Data input tersebut akan dibagi secara random menjadi: data untuk training (70%): 840 data, data untuk validation (15%): 180 data, data untuk testing (15%): 180 data.

Data input disiapkan dalam format Excel (7 baris x 1200 kolom), lalu data tersebut dikonversi terlebih dahulu menjadi format TXT Tab Delimited sehingga dapat di-load langsung oleh MATLAB menjadi matrix sebesar 7 baris x 1200 kolom.

```
inputs 7x1200 =
| tahun1           ..... tahun1200           |
| bulan1          ..... bulan1200            |
| lantai1         ..... lantai1200           |
| kategori1        ..... kategori1200          |
| jumlah_unit1     ..... jumlah_unit1200        |
| luas_tersewa1    ..... luas_tersewa1200       |
| rerata_harga_sewa1 ..... rerata_harga_sewa1200     |
```

Data Output

Untuk training diperlukan data output sebenarnya dari data input di atas sebagai validasi apakah model yang dibangun sudah tepat. Data outputnya adalah variabel pendapatan. Data output disiapkan dalam bentuk Excel (1 baris x 1200 kolom), lalu data tersebut dikonversi terlebih dahulu menjadi format TXT Tab Delimited sehingga dapat di-load langsung oleh MATLAB menjadi matrix sebesar 1 baris x 1200 kolom.

```
targets 1x1200 =
| pendapatan1           ..... pendapatan1200           |
```

Hasil Training

Setelah data input dan data output disiapkan, maka training pun dilakukan dengan mencari pendekatan jumlah hidden neuron yang tepat. Training dilakukan dengan nilai $\mu=1$ agar konvergensi tidak terlalu cepat pada algoritma training Levenberg Marquardt (Demuth, 2009). Tabel 2 adalah ringkasan dari hasil training data (detil hasil training data dapat dilihat pada Lampiran).

Tabel 2 Hasil Training dengan $\mu=1$

Jumlah Hidden Neuron	Best Performance Epoch	MSE (Mean Squared Error)	Regression
20	505	44.457,5	1
50	1000	159.654,5	1

75	26	7.421.107,9	0,99996
100	85	33.859,8	1
124	572	1.397,6	1
125	191	1,9	1
126	13	1.972,5	0,99988
150	13	14.345,2	1

Dari Tabel 2 dapat disimpulkan bahwa MSE (*Mean Squared Error*) akan semakin kecil ketika jumlah hidden neuron bertambah sampai nilai tertentu. Pada uji coba diatas MSE terkecil dicapai pada jumlah hidden neuron = 125 yaitu yang terjadi pada epoch (iterasi) ke-361 dan Regresi = 1. Model inilah yang akan dipakai untuk prediksi data di tahun berikutnya.

Hasil Prediksi

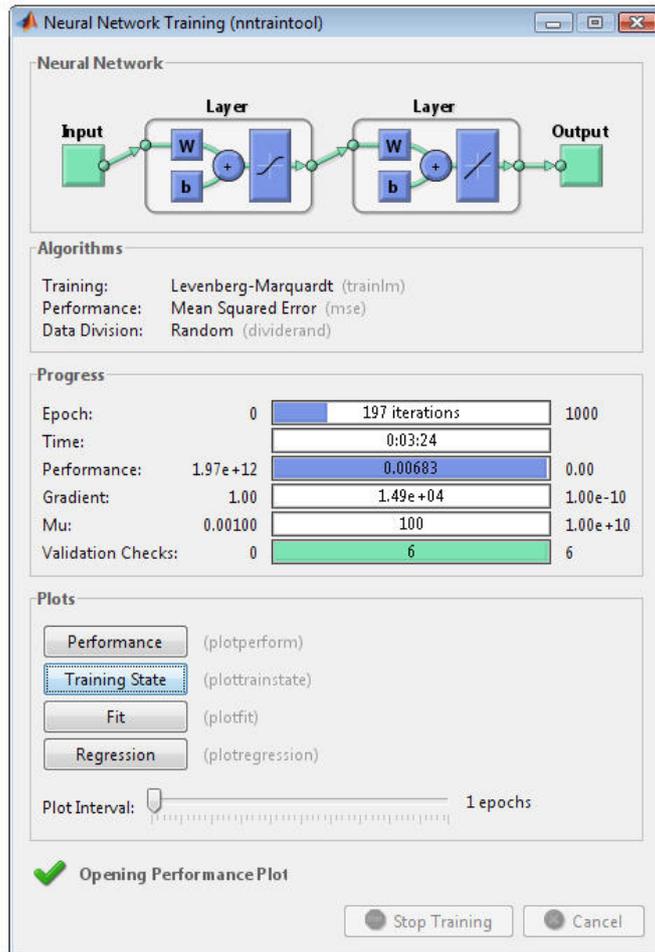
Untuk prediksi, telah disediakan data input untuk satu tahun berikutnya (tahun 2008). Data input untuk prediksi ini terdiri dari 240 data. Data input prediksi terdiri dari 7 variabel dan mempunyai format yang sama dengan format Data Input sebelumnya. Data input disiapkan dalam format Excel (7 baris x 240 kolom), lalu data tersebut dikonversi terlebih dahulu menjadi format TXT Tab Delimited sehingga dapat di-load langsung oleh MATLAB menjadi matrix sebesar 7 baris x 240 kolom.

```
inputs27x240 =
| tahun1           ..... tahun240
| bulan1          ..... bulan240
| lantai1         ..... lantai240
| kategori1       ..... kategori240
| jumlah_unit1    ..... jumlah_unit240
| luas_tersewa1   ..... luas_tersewa240
| rerata_harga_sewa1 ..... rerata_harga_sewa240
|
```

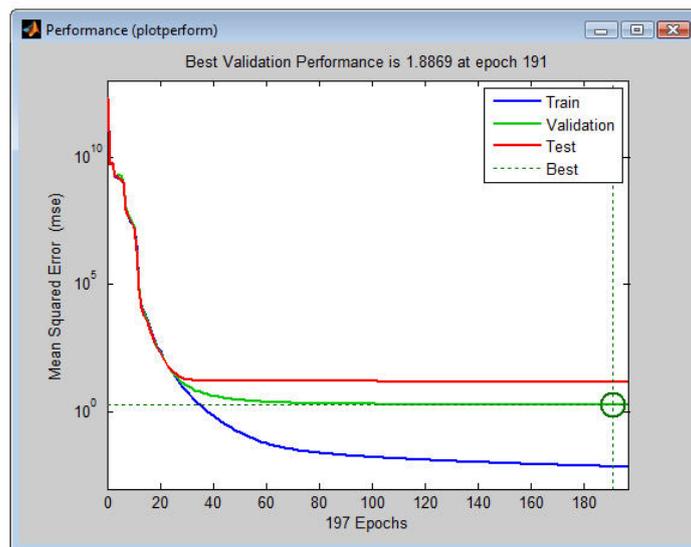
Hasil prediksi dari input tahun 2008 dapat dilihat pada gambar di atas. Dapat dilihat bahwa MSE yang dicapai telah mendekati nol yaitu 1,8869 pada *epoch* (perulangan) ke-191. Nilai MSE ini didapatkan dari hasil prediksi dibandingkan dengan output sebenarnya untuk mengetahui ketepatan hasil prediksi Gambar 10.

Perbedaan antara hasil prediksi dengan output sebenarnya dihitung dalam kolom Error Prediksi. Perbedaan antara hasil prediksi terdahulu (yang merupakan nilai konstan pendapatan sewa terakhir sebelumnya) dengan output sebenarnya dihitung dalam kolom Error Prediksi Tedahulu. Perbandingan tersebut dapat dilihat pada Lampiran. Dari perbandingan tersebut didapatkan nilai total errornya, yaitu total prediksi eror sebesar 4.900,90, dan total prediksi eror terdahulu sebesar 2.852.446,54.

Dapat dilihat bahwa perbandingannya antara error prediksi dengan error prediksi terdahulu adalah mendekati 17:10000. Atau dengan kata lain, dapat disimpulkan bahwa hasil prediksi pendapatan sewa dengan *data mining* 582 kali lebih tepat dibandingkan dengan prediksi yang terdahulu Gambar 11.



Gambar 10 Hasil Prediksi dengan Hidden Neuron=125, $\mu=1$



Gambar 11 Plot Performance Hasil Prediksi dengan Hidden Neuron=125, $\mu=1$

PENUTUP

Prediksi pendapatan sewa dengan data mining menghasilkan prediksi yang lebih tepat daripada prediksi pendapatan sewa yang terdahulu di mana nilainya konstan berdasarkan pendapatan bulan terakhir di tahun sebelumnya. Teknik data mining yang digunakan adalah *Neural Network* dengan algoritma *backpropagation* Levenberg Marquardt karena algoritma ini merupakan algoritma yang terbaik dalam pendekatan fungsi.

Untuk mendapatkan model yang tepat, dilakukan training berkali-kali dengan variasi nilai μ dan hidden neuron. Model Neural Network mencapai nilai MSE minimum yaitu ketika $\mu=1$ dan jumlah hidden neuron = 125. Kemudian model tersebut dipakai dalam melakukan prediksi pendapatan sewa dalam tahun selanjutnya. Dari uji coba didapatkan bahwa model ini telah menghasilkan total nilai error yang jauh lebih kecil daripada total nilai error sebelumnya dengan menemukan pola dari data di tahun-tahun sebelumnya. Prediksi pendapatan sewa dengan data mining ini 582 kali lebih tepat dibandingkan dengan prediksi yang terdahulu. Dengan demikian, manfaat dari penelitian ini telah tercapai sehingga dapat membantu PT XYZ dalam pengambilan keputusan strategis berdasarkan informasi tersembunyi dari database yang ada.

Prediksi pendapatan sewa yang telah dibuat dapat dikembangkan dengan menggunakan teknik training Bayesian Regularization karena teknik ini dapat menghasilkan MSE yang lebih kecil dibandingkan dengan *early stopping* pada beberapa kasus terutama pada data yang berjumlah sedikit atau bila data mengandung sedikit variasi (Demuth, 2009). Teknik Bayesian Regularization ini akan lebih lama prosesnya daripada *early stopping*. Selain itu, prediksi pendapatan sewa ini juga dapat dikembangkan cakupannya meliputi seluruh anak perusahaan.

DAFTAR PUSTAKA

- Demuth, H., et al. (2009). *Neural Network Toolbox™ 6 User's Guide*. Natick: MathWorks.
- Mohd, Z., & Azlan, et al. *Application of Regression and ANN Techniques for Modeling of the Surface Roughness in End Milling Machining Process*, IEEE Computer Society, DOI 10.1109/AMS.2009.76.
- Zhang, H., et al. (2009). Multi-objective simultaneous prediction of waterborne coating properties. *Journal of Mathematical Chemistry* 46, 1050-1059.