

PENERAPAN *GLOBAL RIDGE-REGRESSION* PADA PERAMALAN DATA *TIME SERIES NON LINEAR* STUDI KASUS : PEMODELAN NILAI TUKAR US DOLAR TERHADAP RUPIAH

Sri Utami Zuliana¹

¹Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta
Jl. Marsda Adisucipto No. 1 Yogyakarta
Email: suzuliana@yahoo.com

Abstract

Time series modelling has two types i.e. linear and non-linear. Feed Forward Neural Networks (FFNN) has modelled linear time series well but has found difficulties to model non-linear time series. Radial Basis Function Neural Networks (RBFNN) give an alternative to model non-linear time series. This network has Radial Basis Function in the hidden layer that provides non-linear functions. The RBFNN output is a linear combination of Radial Basis Functions and output weights. An optimal output has the least square error. The weights are gotten from regression. Global-ridge regression adds a regulation parameter to give the optimal weights that produce an optimal output.

Keyword : *Radial Basis Function Neural Networks (RBFNN), non-linear, time series, global ridge-regression,*

I. PENDAHULUAN

Neural network merupakan suatu salah satu pemodelan statistik untuk data *time series*. Salah satu bentuk model *neural network* adalah jaringan *radial basis function* (RBF). Model RBF terdiri dari lapisan *input*, lapisan *hidden* dan lapisan *output*. Output pada data *time series* berbentuk dimensi satu yaitu data pada waktu ke-*t*. Sedangkan lapisan *input* dan lapisan *hidden* memiliki jumlah unit yang bervariasi dengan berbagai cara menentukan jumlah unit pada *hidden layer*, nilai varian, bobot *output*, algoritma untuk mendapatkan taksiran parameter pada model *neural network* dan kriteria kesesuaian model RBF pada pemodelan data *time series*.

¹ E-mail: suzuliana@yahoo.com

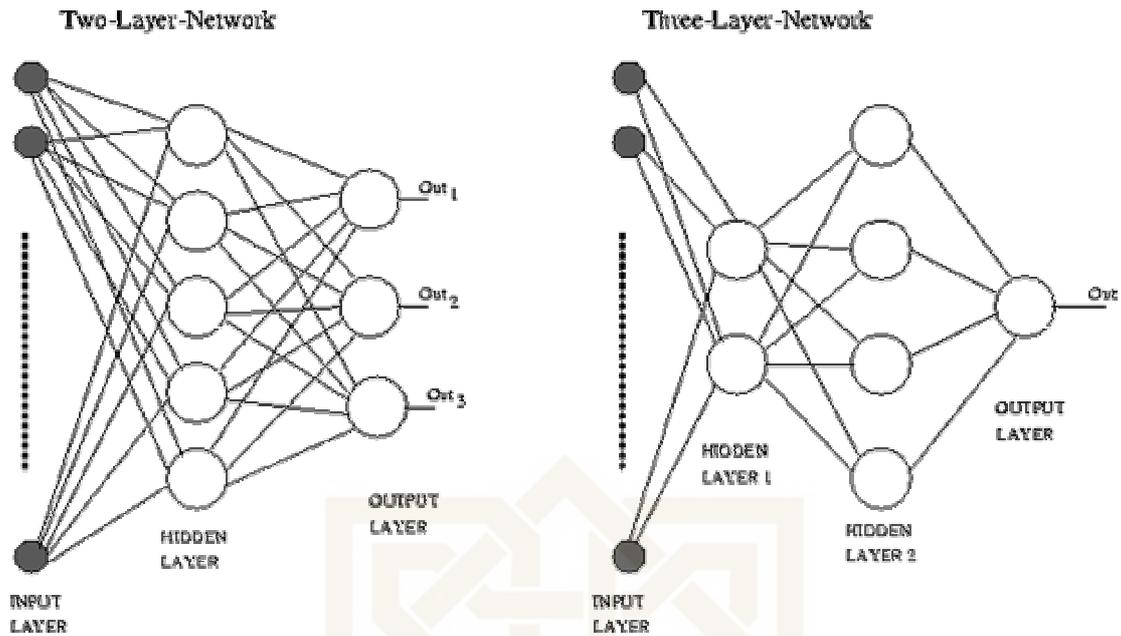
Pembatasan masalah dalam penelitian ini sebagai berikut model NN yang dibahas adalah model RBFNN dengan satu *hidden layer*, fungsi aktivasi yang digunakan adalah fungsi *Gaussian* dan residual (*error*) yang terjadi dalam RBFNN adalah $\varepsilon_i \sim \text{IIDN}(0, \sigma^2)$.

Permasalahan yang akan diangkat dalam paper ini meliputi:

1. Mengkaji pemodelan statistik untuk peramalan pada model RBF.
2. Menggunakan *global ridge-regression* pada algoritma penentuan bobot *output* sebagai kandidat yang sesuai pada model RBF.

Neural network atau lebih tepatnya *artificial neural network* diterapkan pada berbagai disiplin ilmu. Aplikasi *neural network* adalah kemampuan untuk belajar (*learn*) dari data input dengan atau tanpa *teacher*. *Neural networks* terdiri *neuron* dan setiap *neuron* dihubungkan dengan *neuron* yang lain membentuk arsitektur jaringan. Komputasi terjadi dalam *neuron-neuron* yang jumlahnya banyak sekali secara parallel dan terdistribusi.

Arsitektur *jaringan* dapat dibagi menjadi tiga kelompok besar yaitu *single-layer feedforward network*, *multilayer feedforward network* dan *recurrent network*. *Single-layer feedforward network* terdiri dari lapisan input dan lapisan output. Sedangkan pada *multilayer feedforward network* terdiri dari lapisan input, satu atau lebih lapisan *hidden* dan lapisan output dengan proyeksi dari lapisan input ke lapisan *hidden* lalu ke lapisan output. *Recurrent network* terdiri dari paling sedikit satu *feedback loop*. Pada *feedback loop* terdapat *unit-delay element* yang dilambangkan dengan z^{-1} .



Gambar 1. *Multilayer feedforward network* (Schmidt, 1996)

Proses *learning* terdiri dari tiga proses yaitu jaringan distimulasi oleh lingkungan, jaringan berubah dalam parameter bebas sebagai hasil stimulasi dan jaringan merespon lingkungan dengan cara yang baru karena perubahan yang terjadi. Suatu aturan penyelesaian masalah pembelajaran yang sudah tersusun dengan baik disebut algoritma pembelajaran yang memiliki perbedaan pada cara penentuan bobot suatu unit.

Neural network mempunyai kelebihan-kelebihan diantaranya bersifat nonlinear, mampu menyimpulkan pemetaan input dan output tanpa mengetahui distribusi probabilitiknya, mampu beradaptasi dengan data yang selalu berubah dan memiliki analisis dan desain yang seragam di berbagai bidang. Sebagai pemroses informasi *neural network* memiliki sifat universal sehingga memudahkan bidang ilmu yang berbeda untuk berbagi teori dan algoritma pembelajaran.

II. PENGESTIMASIAN BOBOT OUTPUT MENGGUNAKAN *GLOBAL RIDGE-REGRESSION* PADA PEMODELAN RBFNN PADA DATA *TIME SERIES*

Supervised learning adalah masalah pengestimasian fungsi $y(x)$ yang didapatkan berdasarkan himpunan data training yang merupakan pasangan titik input-output $\{(x_i, y_i)\}_{i=1}^p$. Tujuan utama *neural network* adalah mengestimasi fungsi. Pengestimasian bobot yang tepat akan menghasilkan output fungsi dengan *error* yang minimal. Metode *weight-decay* mengestimasi bobot dengan menambahkan parameter regularisasi yang bernilai positif pada SSE (jumlah kuadrat *error*) sehingga didapatkan fungsi biaya:

$$C = \sum_{i=1}^p (\hat{y}_i - f(\mathbf{x}_i))^2 + \lambda \sum_{j=1}^m w_j^2$$

Weight-decay sama dengan *ridge regression*.

Global ridge-regression memiliki satu parameter untuk semua *basis function*, sehingga didapatkan persamaan

$$\sum_{i=1}^p f(\mathbf{x}_i) h_j(\mathbf{x}_i) + \lambda w_j = \sum_{i=1}^p \hat{y}_i h_j(\mathbf{x}_i)$$

sebanyak m persamaan sejenis dengan $1 \leq j \leq m$ yang menghasilkan m bobot.

Penyelesaian tunggal untuk persamaan di atas dapat diselesaikan dalam notasi vektor sebagai berikut:

$$\mathbf{H}^T \mathbf{f} + \lambda \hat{\mathbf{w}} = \mathbf{H}^T \hat{\mathbf{y}}$$

Dalam persamaan ini \mathbf{H} adalah matriks desain dan \mathbf{f} adalah perkalian matriks desain dan vektor bobot. Dapat dibentuk persamaan normal:

$$\hat{\mathbf{w}} = (\mathbf{H}^T \mathbf{H} + \lambda \mathbf{I}_m)^{-1} \mathbf{H}^T \hat{\mathbf{y}}$$

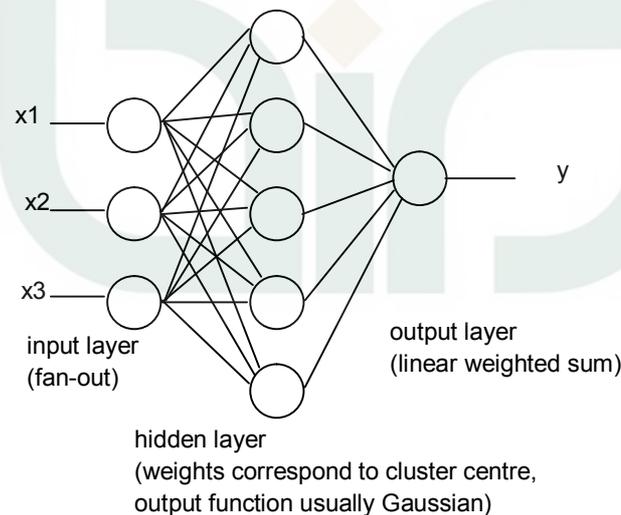
yang menggunakan metode *global ridge-regression*.

Pemodelan *neural networks* pada data *time series* memiliki ciri khusus yaitu memiliki satu unit *output* yaitu data pada saat t . Peramalan data *time series* bertujuan

untuk memberikan dugaan tentang output di masa yang akan datang dengan menggunakan data-data sebelumnya. Peramalan *time series* sering digunakan untuk menaksir data beberapa tahap waktu di masa mendatang.

Model *Neural Networks* sering digunakan untuk mempelajari data *time series* sedemikian hingga mampu melakukan identifikasi dan pendekatan model dari data *time series*. Untuk memahami data *time series*, setiap unit dalam *Neural Networks* akan ditraining dengan suatu metode pelatihan. Data *time series* dibagi menjadi dua bagian yaitu data *training* dan data *validasi*. Pada data *training* dilakukan algoritma *training* dan selanjutnya dilakukan pengecekan *error* dan dipilih bobot terbaik dari keseluruhan data. Data *validasi* tidak digunakan selama pelatihan. Dengan demikian pada data *validasi*, nilai bobot tidak akan berubah.

Salah satu kelebihan dari *Neural Networks* adalah mampu menyelesaikan model *time series* nonlinear dengan baik. *Radial basis function* adalah salah satu pilihan fungsi aktivasi unit *hidden* pada *Neural Networks*. *Radial basis function neural networks* (RBFNN) terdiri dari tiga lapisan yaitu satu lapisan *input*, satu lapisan *hidden* dan satu lapisan *output*. Fungsi Gaussian adalah fungsi aktivasi yang akan digunakan dalam penelitian ini. Fungsi ini berbentuk $\phi(\|\underline{\mathbf{x}} - \underline{\mathbf{x}}_k\|) = \exp(-\|\underline{\mathbf{x}} - \underline{\mathbf{x}}_k\|^2 / \beta^2)$. Fungsi Gaussian ini biasa juga disebut fungsi distribusi normal.



Gambar 2. Radial Basis Function Neural Networks (Wells dan Browne, -)

Tujuan pemodelan adalah mendekati $\{y(t)\}$ dengan model berdimensi hingga yaitu memberikan output $\{\hat{y}(t)\}$ yang akan menghasilkan barisan ralat peramalan output $\{e(t)\}$. Output sistem $\{y(t)\}$ dapat disajikan sebagai

$$y(t) = \hat{y}(t) + e(t)$$

Pembelajaran RBFNN terdiri dari penentuan parameter yang tidak diketahui yaitu menentukan jumlah unit hidden, pusat dan nilai varian tiap-tiap unit hidden serta bobot lapisan output. Untuk beberapa algoritma, langkah-langkah ini dilakukan secara bersamaan.

Dalam penelitian ini digunakan metode *k-means clustering* sebagai prosedur pemilihan pusat. Sedangkan untuk pemilihan jumlah *basis function* (atau unit hidden) untuk jaringan RBF digunakan *constructive learning*. Metode *Global ridge-regression* digunakan untuk penentuan bobot lapisan. Pemrograman menggunakan perangkat lunak MATLAB yang diperkenalkan oleh Mark J. L. Orr. (1997). Program-program utama yang dijalankan dalam pemrograman ini adalah **rbfDesign** dan **globalRidge**. Estimasi bobot dicari melalui program **rbfDesign** sedangkan parameter regularisasi didapatkan dengan menggunakan program **globalRidge**. Untuk mengetahui ada atau tidaknya relevansi antara banyaknya unit *hidden* dengan output yang dihasilkan dihitung kriteria MAPE, MAD dan MSE. Program **rbfDesign** digunakan untuk membangkitkan matriks desain untuk RBFNN dengan input, pusat, nilai varian, tipe fungsi dan unit bias pilihan yang tetap. Penggunaan program ini dilakukan pada pelatihan dan peramalan. Programnya adalah:

$$\mathbf{H} = \mathbf{rbfDesign} (\mathbf{X}, \mathbf{C}, \mathbf{R}, \mathbf{options})$$

dengan \mathbf{X} adalah matriks $n \times p$ [$\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_p$] yang merupakan p vektor input data training dengan n unit input, \mathbf{C} adalah matriks $n \times m$ [$\mathbf{c}_1 \ \mathbf{c}_2 \ \dots \ \mathbf{c}_m$] dengan kolomnya adalah titik pusat unit *hidden* ke - m (berdimensi n) dan \mathbf{R} adalah himpunan nilai varian [$\mathbf{r}_1 \ \mathbf{r}_2 \ \dots \ \mathbf{r}_m$] adalah nilai varian unit *hidden* berdimensi m sedangkan **options** digunakan untuk

menentukan tipe RBF yaitu untuk fungsi Cauchy dituliskan '**c**', fungsi multiquadric dituliskan '**m**' dan fungsi *inverse multiquadric* dituliskan '**i**' sedangkan untuk fungsi Gaussian dituliskan '**g**'. Bila **options** tidak dituliskan maka defaultnya adalah fungsi Gaussian dan bila dituliskan '**b**' pada *options string* maka unit *bias* akan ditambahkan pada jaringan yaitu matriks **H** akan mendapatkan satu kolom tambahan. Matriks desain H ini dapat digunakan untuk mencari vektor bobot optimal dengan menggunakan persamaan normal.

$$\mathbf{w} = \text{inv}(\mathbf{H}' * \mathbf{H}) * \mathbf{H}' * \mathbf{y};$$

Untuk mendapatkan output yang mendekati data yang sesungguhnya digunakan program *ridge regression (Tikhonov regularization)* sehingga didapatkan parameter regularisasi yang dapat meningkatkan pencocokan.

Ridge regression atau *weight-decay* memiliki dua bentuk yaitu *global ridge* dengan parameter tunggal dan *local ridge* dengan m parameter untuk m *basis function*. Pencocokan data yang optimal dilakukan dengan mencegah *over-fitting* yaitu dengan menambahkan parameter regularisasi positif λ pada SSE sehingga dihasilkan fungsi biaya sebagai berikut:

$$C = \sum_{i=1}^p (f(\mathbf{x}_i) - y_i)^2 + \lambda \sum_{j=1}^m w_j^2$$

Global ridge-regression menghasilkan vektor bobot optimal dengan persamaan normal sebagai berikut:

$$\hat{\mathbf{w}} = (\mathbf{H}^T \mathbf{H} + \lambda \mathbf{I}_m)^{-1} \mathbf{H}^T \mathbf{y}$$

dengan \mathbf{I}_m adalah matriks identitas dimensi m . Pemanggilan program ini ditulis sbb:

$$[\mathbf{l}, \mathbf{e}, \mathbf{L}, \mathbf{E}] = \text{globalRidge}(\mathbf{H}, \mathbf{Y}, \mathbf{l}, \text{options}, \mathbf{U})$$

dengan **H** adalah matriks desain $p \times m$, **Y** adalah vektor output dengan p baris dan n kolom (n adalah dimensi ruang output), **l** adalah dugaan nilai parameter regularisasi positif λ , **options** adalah pilihan metode yang digunakan untuk meramalkan *error* pada data uji dan **U** adalah maktriks penalty. Optimisasi dikerjakan menggunakan satu

dugaan awal dan satu pilihan metode untuk meminimalkan *error* pada himpunan data validasi di waktu mendatang. Metode untuk meminimalkan *error* digunakan untuk mendapatkan nilai optimal dari λ . Pilihan metode yang digunakan dalam program ini adalah UEV, FPE, GCV dan BIC dengan pemilihannya dituliskan pada **options**. Jika pada **options** tidak dituliskan pilihannya maka defaultnya adalah GCV.

III. IMPLEMENTASI PADA DATA FINANSIAL

Studi Kasus : Data nilai tukar US Dolar terhadap rupiah dari tanggal 19 Mei 2011 sampai dengan 31 Mei 2012

3.1. Deskripsi Data

Untuk implementasi akan diterapkan pada data finansial yang berbentuk *time series* yaitu data nilai tukar US Dolar terhadap dalam rupiah dari tanggal 19 Mei 2011 sampai dengan tanggal 31 Mei 2012 yang merupakan data harian sebanyak 257 data. Data ini didapatkan dari <http://www.bankofcanada.ca/>.

Pergerakan nilai tukar US Dolar terhadap dalam rupiah dari tanggal 19 Mei 2011 sampai dengan tanggal 31 Mei 2012 bersifat fluktuatif, disamping itu kurva yang ditunjukkan cenderung non-linear.

3.2. Model Radial Basis Function Neural Networks

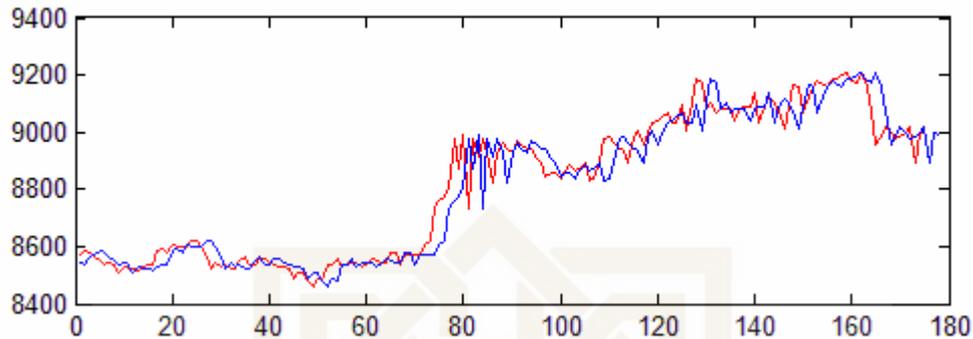
Dalam penelitian ini banyaknya input ditentukan sebanyak tiga input sedangkan unit *hiddennya* ditentukan sebanyak lima unit *hidden* yang menggunakan fungsi aktivasi Gaussian dan nilai varian pada RBF didefinisikan dengan standard deviasi, jarak rata-rata ke pusat cluster dan jarak maksimum ke pusat cluster.

Untuk mengimplementasikan pemodelan RBF, data dibagi dua kelompok data yang tidak overlap yaitu data training dan data testing. Data training terdiri dari 178 data sedangkan 89 data sebagai data testing.

3.3. Pembahasan

Setelah program dijalankan didapatkan estimasi bobot yang dicari melalui program **rbfDesign** yaitu $w_1=-9.6 \times 10^{-3}$; $w_2=-9.6 \times 10^{-3}$; $w_3=-9.6 \times 10^{-3}$; $w_4=-9.6 \times 10^{-3}$;

$w_5=1.0727$ dan $w_b=1.0 \times 10^4$. Sedangkan untuk parameter regularisasi dihasilkan 0.0048. Penghitungan kriteria didapatkan MAPE 0.0619, MAD 669.9648 dan MSE 5.5855. Plot perbandingan antara data estimasi dan data asli dapat dilihat sebagai berikut :



Gambar 4. Plot Data nilai tukar US Dolar terhadap rupiah dari tanggal 19 Mei 2011 sampai dengan 31 Mei 2012 untuk hasil pemodelan RBFNN dengan 5 unit *hidden* dan distance RBF didefinisikan sebagai jarak maksimum dari pusat cluster. Garis merah adalah data estimasi dibandingkan dengan garis biru sebagai data asli.

IV. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Berdasarkan analisis data nilai tukar US Dolar terhadap rupiah dengan menggunakan software MATLAB dapat disimpulkan sebagai berikut :

1. Berdasarkan plot data time series terlihat bahwa pergerakan nilai tukar US Dolar terhadap dalam rupiah dari tanggal 19 Mei 2011 sampai dengan tanggal 31 Mei 2012 bersifat fluktuatif dan cenderung non-linear.
2. Berdasarkan estimasi bobot dengan **globalRidge** dihasilkan parameter regularisasi 2.7850×10^{-4} .
3. Berdasarkan analisis data dengan **program rbfDesign** dihasilkan estimasi bobot sbb : $w_1=-9.6 \times 10^{-3}$; $w_2=-9.6 \times 10^{-3}$; $w_3=-9.6 \times 10^{-3}$; $w_4=-9.6 \times 10^{-3}$; $w_5=1.0727$ dan $w_b=1.0 \times 10^4$. Sedangkan untuk parameter regularisasi dihasilkan 0.0048.

4.2. Saran

Berdasarkan penelitian yang telah dilakukan, penulis memiliki beberapa saran sbb :

1. Perlu diadakan penelitian lebih lanjut tentang pemodelan terbaik menggunakan *Radial Basis Functions Neural Network (RBFNN)* dengan metode pengestimasian bobot yang lainnya.
2. Perlu diadakan penelitian lebih lanjut tentang pemodelan terbaik menggunakan *Radial Basis Functions Neural Network (RBFNN)* dengan dua atau lebih *hidden layer*
3. Perlu diadakan penelitian lebih lanjut tentang pemodelan terbaik menggunakan *Radial Basis Functions Neural Network (RBFNN)* dengan output lebih dari satu.

DAFTAR PUSTAKA

- D., 2002. *Radial Basis Function*, Handbook of Neural Network Signal Processing, Windale Technologies
- Bishop, C.M., 1995. *Neural Network for Pattern recognition*, Oxford Clarendon Press
- Broomhead, D.S. and Lowe, D., 1998. *Multivariable Functionnal Interppopulation and Adaptive Anders, U. and Korn, O., 1977. Model selection in NN, neural Networks*, 12, 309-323
- Back, Andrew *Network. Complex Systems*, 2.321-355
- Fine, T.L., 1999. *Feedforward Neural Network Methodology Springer*, New York.
- Haykin, S., 1999. *Neural Networks : A Comprehensive Foundation*, second edition Prentice-Hall, Oxford.
- Orr, Mark J. L., 1997. *MATLAB Routines for Subset Selection and Ridge Regression in Linear Neural Networks*; Center for Cognitive Science Edinburgh University Scotland, UK
- Ripley, B.D., 1996. *Pattern Recognition and Neural Networks* . Cambridge University Press, Cambridge.
- Schmidt, Albrecht, 1996, "A Modular Neural Network Architecture with Additional Generalization Abilities for High Dimensional Input Vectors", thesis,

Manchester Metropolitan University. Yang diunduh di <http://www.teco.uni-karlsruhe.de/~albrecht/neuro/html/img4.gif>, 28feb08 7:51pm

Sundararajan, N., Lu, Ying Wei., Saratchandran, P., 1999. *Radial Basis Function Neural Networks with Sequential Learning* : MRAN and its applications; World Scientific; Singapore

V.David Sanchez A., 2002. *Searching for a solution to the automatic RBF network design problem*, Advanced Computational Intelligent Systems, Pasadena, CA 91116-6130, USA; Neurocomputing 42, 147-170

Wei, William W.S., 1994. *Time Series Analysis : Univariate and Multivariate Methods*, Addison-Wesley Publishing Company, USA.

Wells, Ian dan Browne, Tony, Lecture Note “*Radial Basis Function (RBF) Networks*”. Yang diunduh di www.computing.surrey.ac.uk/courses/csm10/NeuralNetworks/RBFNetworks.ppt, 28feb08 7:55pm

