

Prediksi Customer Churn Berbasis Adaptive Neuro Fuzzy Inference System

Yayak Kartika Sari¹, Kusrini², Ferry Wahyu Wibowo³

^{1,2,3}Teknik Informatika, Fakultas Teknik, Universitas Amikom Yogyakarta
E-mail: ¹yayakkartikasari93132042@gmail.com, ²kusrini@amikom.ac.id, ³ferry.w@amikom.ac.id

Abstrak – Customer Churn adalah pelanggan yang berhenti berlangganan dan pindah pada perusahaan lain, karena berbagai faktor. Customer churn merupakan masalah yang sangat penting yang harus dihadapi oleh perusahaan karena berhentinya pelanggan akan berdampak pada retensi perusahaan. Oleh sebab itu, dibuatkan sistem prediksi customer churn untuk mengetahui tingkat pelanggan yang churn, apabila customer churn dapat diketahui terlebih dahulu, maka akan menguntungkan bagi pihak CRM untuk mengatur strategi-strategi mencegah pelanggan yang melakukan churn. Untuk menentukan prediksi customer churn menggunakan teknik data mining dengan algoritma ANFIS. Algoritma ANFIS merupakan gabungan antara jaringan syaraf tiruan dengan fuzzy inference system. Model prediksi yang dibangun dengan metode ANFIS menggunakan pembelajaran alur maju dan pembelajaran alur mundur, sehingga untuk melakukan prediksi dibutuhkan nilai parameter fuzzy baru yang diperoleh dari proses pelatihan. Setelah nilai parameter fuzzy baru didapatkan, maka akan dilakukan tahap pengujian. Pada tahap pengujian dilakukan dengan proses pembelajaran maju untuk mendapatkan nilai prediksinya, sehingga pada prosesnya nilai prediksi yang berupa angka dan status prediksi. Pelatihan dan pengujian ANFIS untuk semua produk menghasilkan perbandingan nilai error rata-rata pelatihan sebesar 8,316 %.

Kata Kunci — Data Mining, ANFIS, CRM, Customer Churn.

Abstract – Customer Churn is a customer who unsubscribes and moves to another ISP, due to various factors. Customer churn is a very important issue that should be faced by the company because the cessation of customers will impact on corporate retention. Therefore, customer churn prediction system is made to know the level of customer churn, if customer churn can be known in advance, it will be advantageous for CRM to arrange strategies to prevent customers who do churn. To determine customer churn prediction using data mining techniques with ANFIS algorithm. ANFIS algorithm is a combination of artificial neural network with fuzzy inference system. Predictive models built with the ANFIS method use advanced groove learning and retrograde learning, so to make predictions it takes the value of new fuzzy parameters obtained from the training process. After the new fuzzy parameter values are obtained, the test phase will be performed. In the testing phase is done by the process of advanced learning to get predictive value, so that in the process predictive value in the form of numbers and prediction status. ANFIS training and testing for all products resulted in comparison of the average error rate of training of 8.316%

Keywords — Data Mining, ANFIS, CRM, Customer Churn.

1. PENDAHULUAN

Perkembangan internet saat ini menjadi kebutuhan yang sangat penting bagi masyarakat. Tidak hanya sebagai sumber informasi tetapi bisa menjadi sarana pendukung kegiatan ekonomis dan hiburan. Menurut data www.internetworldstats.com, pengguna internet di Indonesia saat ini mencapai 132,7 juta jiwa dari total populasi sebesar 263 juta jiwa. Dengan perkembangan internet yang semakin luas saat ini, membuat penyedia jasa layanan internet menjadi semakin banyak dan berkembang, terutama di daerah kediri. Penyedia jasa layanan internet di daerah kediri yaitu PT. Telekomunikasi, Biznet, PT. Indosat M2, Primadona Net, CV. Global Media Data, dan lain-lain.

Dari semua penyedia jasa layanan internet di Kediri memberikan penawaran-penawaran yang sangat menarik untuk meningkatkan pelanggan. Hal itu menjadi salah satu faktor penyebab terjadinya menurunnya pelanggan internet pada perusahaan XYZ disebabkan karena *churn*. *Customer Churn* atau turun naiknya pelanggan didefinisikan sebagai hilangnya pelanggan karena berpindah ke layanan lain.[1] Hal ini menjadi isu penting yang merupakan salah satu tantangan utama oleh perusahaan XYZ yang harus dihadapi, karena untuk memperoleh pelanggan baru memerlukan biaya 10 kali lipat lebih mahal dibandingkan biaya untuk mempertahankan pelanggan yang ada. Mahalnya untuk memperoleh pelanggan baru tentunya perusahaan lebih memilih mempertahankan pelanggan. Selain itu, terjadinya *customer churn* akan berdampak penurunan *revenue* perusahaan.[2] Maka dari itu, untuk mencegah terjadinya *customer churn* pada perusahaan XYZ maka dibutuhkanlah prediksi *customer churn*, agar perusahaan XYZ dapat mengatur strategi yang dapat mencegah *customer churn* di perusahaan XYZ.

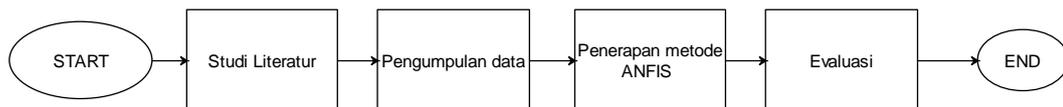
Teknik yang digunakan untuk memprediksi adalah data mining. Data mining merupakan serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual.[3] Sistem prediksi yang telah banyak diteliti antara lain analisis prediksi *churn* menggunakan metode *logistic regression* dan algoritma *decision tree*[4], prediksi *customer churn* menggunakan algoritma *fuzzy iterative dichotomiser 3*[2], IG-KNN untuk prediksi *customer churn* telekomunikasi, penggunaan *deep learning* untuk prediksi *churn* pada jaringan telekomunikasi mobile[5]. Tetapi untuk penelitian ini menggunakan metode metode ANFIS (*Adaptive Neuro Fuzzy Inference System*).

ANFIS merupakan gabungan dua sistem yaitu *neural network* dan *fuzzy logic*. Sistem saraf tiruan memberikan kemudahan dalam mengklasifikasikan suatu objek. Berdasarkan masukan sistem, *Fuzzy logic* merupakan pemakaian fungsi keanggotaan untuk menentukan seberapa besar suatu predikat memenuhi suatu fungsi yang dinyatakan kedalam “*if – then*” yang memberi kemudahan dengan tidak memerlukan analisis matematik untuk pemodelan.[6]

Pokok permasalahan dari penelitian ini adalah penggunaan metode ANFIS untuk memprediksi *customer churn* di perusahaan XYZ Kediri. Sebagai masukannya, menggunakan data *time series* selama 2 tahun dengan variabel jumlah data status pelanggan dan data keluhan. Diharapkan dengan menggunakan ANFIS, sistem prediksi untuk *customer churn* dapat dilakukan dengan baik.

2. METODE PENELITIAN

Dalam penelitian ini, terdiri dari beberapa tahapan yang diawali dengan melakukan Studi literatur, pengumpulan data, penerapan algoritma ANFIS. Alur penelitian ditunjukkan oleh gambar 1.



Gambar 1. Metode penelitian

2.1. Studi Literatur

Metode melalui studi literature yang bertujuan mendapatkan pengetahuan dari penelitian yang akan dilakukan. Studi literatur tersebut didapatkan melalui sumber paper dan buku. Adapapun hal-hal yang dapat diambil dari metode ini diantaranya adalah:[5]

- a. Studi Pendahuluan : tahap ini merupakan kegiatan untuk menemukan informasi tentang obyek permasalahan yang ada. Permasalahan-persalahan yang berkembang beberapa tahun terakhir dalam sebuah organisasi atau perusahaan khususnya mengenai *customer churn*.
- b. Studi Pustaka : tahapan ini adalah tahap untuk menemukan penelitian-penelitian yang sejenis dengan penelitian ini yang nantinya dijadikan sebagai referensi dan pendukung teori dalam menyelesaikan permasalahan yang diangkat.
- c. Perumusan Masalah : adapun pada tahapan selanjutnya setelah mendapatkan permasalahan utama dari obyek penelitian yang dilengkapi dasar teori dari studi pustaka yang mendukung maka masalah yang ada dapat dirumuskan dengan baik.

2.2 Pengumpulan Data

Data masukan yang akan digunakan pada sistem prediksi adalah data *time series* bulanan dari data status pelanggan dan keluhan dan selanjutnya akan dilakukan perhitungan prediksi menggunakan metode ANFIS untuk perusahaan XYZ di Kediri. Data yang digunakan untuk prediksi merupakan data dari bulan januari 2015 sampai bulan desember 2016.

Sistem analisa yang digunakan untuk memprediksi *customer churn* pada perusahaan XYZ di Kediri menggunakan metode ANFIS. Untuk mencari nilai *standart devisiasi* dan *mean*, maka data di cluster menggunakan algoritma Fuzzy C Means. Kemudian dilakukan perhitungan untuk mendapatkan nilai setiap *cluster*. Tabel 1. merupakan data hasil clustering *Fuzzy C Means* yang akan diproses menggunakan perhitungan ANFIS.

Tabel 1. Clustering data Pelanggan

Bulan	Derajat keanggotaan	
	Data Status Pelanggan	Keluhan
Januari 2015	0,044	0,956
Februari 2015	0,108	0,892
Maret 2015	0,13	0,87
April 2015	0,169	0,831
Mei 2015	0,529	0,471
Juni 2015	0,567	0,433
Juli 2015	0,598	0,402
Agustus 2015	0,72	0,28
September 2015	0,73	0,27
Oktober 2015	0,742	0,258
November 2015	0,813	0,187
Desember 2015	0,822	0,178
Januari 2016	0,844	0,156
Februari 2016	0,849	0,151
Maret 2016	0,897	0,103
April 2016	0,914	0,087
Mei 2016	0,963	0,037
Juni 2016	0,967	0,033
Juli 2016	0,967	0,33
Agustus 2016	0,967	0,033
September 2016	0,967	0,033
Oktober 2016	0,969	0,031
November 2016	0,969	0,031
Desember 2016	0,044	0,956

Data hasil *clustering* menggunakan *Fuzzy C Means*, selanjutnya dihitung nilai mean (a) dan standart devisiasi (c). Setelah dilakukan perhitungan, didapatkan *mean* dan *standart devisiasi* pada tabel 2.

c=	0,3716	0,3384	a=	0,1657	0,1175
	0,4550	0,8183		0,1303	0,1083

2.3 Penerapan Metode ANFIS

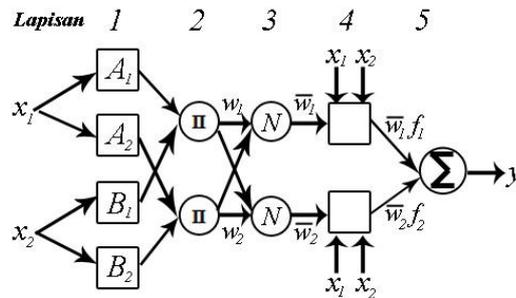
ANFIS adalah arsitektur yang secara fungsional sama dengan *fuzzy rule base* model Sugeno. Arsitektur ANFIS juga sama seperti jaringan syaraf tiruan dengan fungsi radial dengan sedikit batasan tertentu. Bisa dikatakan bahwa ANFIS adalah suatu metode yang mana dalam melakukan penyetulan aturan digunakan algoritma pembelajaran terhadap sekumpulan data. Pada ANFIS juga memungkinkan aturan-aturan untuk beradaptasi. [6]

Agar jaringan dengan fungsi basis radial ekuivalen dengan *fuzzy* berbasis aturan model sugeno orde 1 ini, diperlukan :

- 1) Aturan-aturan harus memiliki metode agegrasi yang sama untuk menghasilkan semua outputnya.

- 2) Jumlah fungsi aktivasi harus sama dengan jumlah aturan fuzzy (IF - THEN).
- 3) Jika ada beberapa input pada basis aturannya, maka tiap-tiap fungsi aktivasi harus sama dengan fungsi keanggotaan tiap – tiap input-nya.
- 4) Fungsi aktivasi dan aturan-aturan fuzzy harus memiliki fungsi yang sama untuk neuron-neuron dan aturan-aturan yang ada di sisi output-nya.

Berikut merupakan arsitektur jaringan ANFIS :



Gambar 2. Arsitektur Jaringan ANFIS [6]

Arsitektur jaringan ANFIS (Gambar 2) terdiri dari lapisan berikut :

- 1) Tiap-tiap neuron i pada lapisan pertama adaptif terhadap parameter suatu fungsi aktivasi. Output dari tiap neuron berupa derajat keanggotaan yang diberikan oleh fungsi keanggotaan input, yaitu: $\alpha_{A1}(X_1)$, $\alpha_{B1}(X_2)$, $\alpha_{A2}(X_1)$, atau $\alpha_{B2}(X_2)$.
- 2) Tiap-tiap neuron pada lapisan ke dua berupa neuron tetap yang outputnya adalah hasil dari masukkan. Biasanya digunakan operator AND. Tiap-tiap node merepresentasikan α predikat dari aturan ke-i.
- 3) Tiap-tiap neuron pada lapisan ke tiga berupa node tetap yang merupakan hasil perhitungan rasio dari α predikat (w), dari aturan ke-1 terhdap jumlah dari keseluruhan α predikat.

$$W_i = \frac{w_i}{w_1+w_2}, \text{ dengan } i= 1,2 \dots\dots\dots (1)$$

Hasil ini dikenal dengan nama *normalised firing strength*.

- 4) Tiap-tiap neuron pada lapisan ke empat merupakan node adaptif terhadap suatu output.

$$\bar{W}_i Y_i = W_i(C_{i1}X_1 + C_{i2} X_2 + C_{i0}); \text{ dengan } i=1,2 \dots\dots\dots (2)$$

Dengan W_i adalah *normalised firing strength* pada lapisan ke tiga dan $\{C_{i1}, C_{i2}, C_{i3}\}$ adalah parameter-parameter pada neuron tersebut. Parameter-parameter pada lapisan tersebut disebut dengan nama *consequent parameter*.

- 5) Tiap-tiap neuron pada lapisan ke lima adalah node tetap yang merupakan jumlahan dari semua masukkan.

Algoritma *hybrid* akan mengatur parameter-parameter C_{ij} secara maju (*forward*) dan akan mengatur parameter-parameter $\{a_i, b_i, c_i\}$ secara mundur (*backward*). Pada langkah maju, input jaringan akan merambat maju sampai pada lapisan ke empat, dimana parameter-parameter C_{ij} akan diidentifikasi dengan metode *least squares estimator* (LSE). Sedangkan langkah mundur, error sinyal akan merambat mundur dan parameter-parameter (a_i, b_i, c_i) akan diperbaiki dengan menggunakan metode *gradient-descent*(GD). [7]

2.3.1 Langkah Maju Dengan Metode LSE

Dari arsitektur ANFIS, bila parameter tetap maka keluaran keseluruhan dapat dinyatakan dengan kombinasi linear dari parameter konsekuensi. [7]

$$f = f_2 = \frac{W_1}{W_1+W_2} f_1 + \frac{W_2}{W_1+W_2} \dots\dots\dots (3)$$

2.3.2 Langkah Balik Dengan Metode GD

Terdapat beberapa metode penurunan gradient dalam masalah optimasi, antara lain : [7]

- 1) Metode penurunan tercuram (*Steepest Descent*, SD)
- 2) Metode Newton yang terdiri atas Newton klasik, Newton termodifikasi, dan Quasi Newton.
- 3) Metode *Conjugate Gradient*.

3. HASIL DAN PEMBAHASAN

Setelah melakukan pembagian (*clustering*) data menggunakan metode *Fuzzy C Means* dan mendapatkan nilai *Standar deviasi* dan *Mean*, selanjutnya kita akan mencari nilai inferensi dengan menggunakan metode *Adapotive Neuro Fuzzy Inference System* (ANFIS). Pada perhitungan metode ANFIS ini akan dilakukan perhitungan pada setiap lapisannya. Sistem inferensi *fuzzy* yang diterapkan adalah inferensi *fuzzy* model Sugeno.

3.2 Lapisan Pertama (Proses Fuzifikasi)

Pada lapisan pertama terjadi proses *fuzifikasi*. Proses ini adalah untuk memetakan inputan data kedalam himpunan *fuzzy*. Dalam proses ini akan dilakukan perhitungan fungsi keanggotaan *fuzzy* untuk mentransformasi masukan himpunan klasik ke derajat tertentu. Fungsi keanggotaan yang digunakan adalah IV-15 jenis *Generalized-Bell* dimana pada fungsi keanggotaan ini terdapat dua parameter yaitu *mean* dan *standar deviasi*. Data *output* lapisan pertama ditunjukkan dalam tabel 2.

Tabel 2.Data Output Lapisan Pertama

Derajat keanggotaan				
Bulan	μ_{a1}	μ_{a2}	μ_{b1}	μ_{b2}
Januari 2015	5,4363	0,7322	5,6331	8,9807
Februari 2015	0,2616	0,3087	0,5257	0,0796
Maret 2015	0,2616	0,3087	0,9998	0,0488
April 2015	0,2616	0,3087	0,0661	0,8888
Mei 2015	0,3586	0,1543	0,9998	0,0488
Juni 2015	0,3586	0,1543	0,9949	0,0469
Juli 2015	0,3873	0,5208	0,0345	0,3687
Agustus 2015	0,4032	0,1729	0,2195	0,1495
September 2015	0,4631	0,6529	0,622	0,0722
Oktober 2015	0,4919	0,7015	0,5722	0,0757
November 2015	0,5779	0,2511	0,2412	0,0242
Desember 2015	0,6132	0,2686	0,2536	0,1315
Januari 2016	0,6132	0,2686	9673	0,0529
Februari 2016	0,8171	0,9985	0,9761	0,0451
Maret 2016	0,8544	0,9869	0,5402	0,0328
April 2016	0,8786	0,4468	0,5402	0,0328
Mei 2016	0,8895	0,9645	0,5402	0,0328
Juni 2016	0,8895	0,9645	0,5402	0,0328

Juli 2016	0,8895	0,9645	0,5402	0,0328
Agustus 2016	0,8895	0,9645	0,5257	0,0796
September 2016	0,8895	0,9645	0,0466	0,8144
Oktober 2016	0,8895	0,9645	0,5402	0,0328
November 2016	0,9645	0,5621	0,1245	0,2928
Desember 2016	0,9649	0,5621	0,069	0,8341

3.3 Lapisan Kedua dan Lapisan Ketiga

Output dari lapisan kedua adalah hasil perkalian dari semua sinyal yang masuk. Masing-masing keluaran simpul menyatakan derajat pengaktifan dari aturan fuzzy. $w_1 = \mu A_1 (\mu B_1)$. dan, $w_2 = \mu A_2 (\mu B_2)$.

Pada lapisan ke tiga Setiap node pada lapisan ini adalah node nonadaptif yang menampilkan fungsi derajat pengaktifan ternormalisasi (*normalized firing strength*) yaitu membagi w_i dengan jumlah total \bar{W}_1 dan \bar{W}_2 . Data *output* lapisan pertama ditunjukkan dalam tabel 3.

Tabel 3. Output Lapisan Kedua Dan Output Lapisan Ketiga

Bulan	Output lapisan ke dua		Output lapisan ke tiga	
	W1	W2	\bar{W}_1	\bar{W}_2
Januari 2015	0,0134	0,192	0,0651	0,9349
Februari 2015	0,0173	0,0142	0,0593	0,9407
Maret 2015	0,0415	0,7855	0,0501	0,9499
April 2015	0,0445	0,6045	0,0685	0,9315
Mei 2015	0,0666	0,4688	0,1243	0,8757
Juni 2015	0,0885	0,0258	0,774	0,226
Juli 2015	0,1201	0,1646	0,4219	0,5781
Agustus 2015	0,1375	0,0246	0,8484	0,1516
September 2015	0,1394	0,0061	0,9582	0,0418
Oktober 2015	0,1555	0,0353	0,8149	0,1851
November 2015	0,2615	0,0151	0,9455	0,0545
Desember 2015	0,2815	0,0531	0,8413	0,1587
Januari 2016	0,288	0,0471	0,8594	0,1406
Februari 2016	0,3568	0,0072	0,9801	0,0199
Maret 2016	0,3585	0,0075	0,9794	0,0206
April 2016	0,4615	0,0324	0,9345	0,0655
Mei 2016	0,4676	0,0768	0,859	0,141
Juni 2016	0,4746	0,0147	0,97	0,03
Juli 2016	0,4805	0,0316	0,9382	0,0618
Agustus 2016	0,4805	0,0316	0,9382	0,0618
September 2016	0,4805	0,0316	0,9382	0,0618
Oktober 2016	0,4805	0,0316	0,9382	0,0618
November 2016	0,5931	0,7855	0,9766	0,0234
Desember 2016	0,7976	0,045	0,9466	0,0534

3.4 Lapisan Keempat (Proses Defuzifikasi)

Pada lapisan ini dilakukan perhitungan mengubah hasil fuzzy kebentuk menjadi himpunan klasik (*crisp*). Pada lapisan ini dilakukan perhitungan LSE untuk mendapatkan nilai parameter koefisien.

3.5 Lapisan Kelima

Pada lapisan ini hanya ada satu node tetap yang fungsinya untuk menjumlahkan semua masukan yang berasal dari lapisan keempat.

Tabel 4. *Output* Lapisan Kelima

Bulan	Data Masukan	Hasil Prediksi
Januari 2015	360	301
Februari 2015	480	350
Maret 2015	490	253
April 2015	500	338
Mei 2015	600	349
Juni 2015	660	499
Juli 2015	660	499
Agustus 2015	660	360
September 2015	660	499
Oktober 2015	670	512
November 2015	730	412
Desember 2015	750	379
Januari 2016	770	549
Februari 2016	780	405
Maret 2016	880	566
April 2016	980	675
Mei 2016	980	679
Juni 2016	990	534
Juli 2016	990	752
Agustus 2016	1100	791
September 2016	1110	574
Oktober 2016	1270	657
November 2016	1300	691
Desember 2016	1430	887

3.6 Evaluasi

Setelah mendapatkan hasil % *error* prediksi *customer churn* dengan menggunakan faktor data status pelanggan dan data keluhan, maka dilakukan perhitungan *error* dengan pengurangan data sebenarnya dengan hasil peramalan.

Tabel 5. Perbandingan nilai *error*

Bulan	<i>Error</i> %
Juli 2016	9,5 %
Agustus 2016	13,7 %
September 2016	8,2 %
Oktober 2016	6,3 %
November 2016	7,8 %
Desember 2016	4,4 %

4. SIMPULAN

Berdasarkan penelitian yang telah dilakukan, dengan menggunakan metode ANFIS dapat membantu perusahaan XYZ untuk mengetahui prediksi pelanggan yang melakukan *churn*. Dengan mengetahui prediksi pelanggan yang melakukan *churn*, maka akan memberikan keuntungan untuk manajemen, khususnya divisi *Customer Relationship Management* (CRM), dapat menciptakan promosi yang bertujuan untuk meningkatkan loyalitas pelanggan dan meningkatkan strategi penjualan untuk mendapatkan pelanggan baru dan melakukan retensi pelanggan. Pelatihan dan pengujian ANFIS untuk semua produk menghasilkan perbandingan nilai *error* rata-rata pelatihan sebesar 8,316 %.

5. SARAN

Saran yang dapat diberikan bagi pengembangan penelitian di masa mendatang terkait prediksi *customer churn* dengan menambah parameter agar menghasilkan nilai prediksi yang lebih baik lagi.

DAFTAR PUSTAKA

- [1] Fikrieabdillah. 2016. *Penggunaan Deep Learning untuk Prediksi Churn pada Jaringan Telekomunikasi Mobile*. e-Proceeding of Engineering : Vol.3, No.2.
- [2] Herawati, Meyrina, Imam Mukhlash dan Inu L. Wibowo. 2016. *Prediksi Customer Churn Menggunakan Algoritma Fuzzy Iterative Dichotomiser 3*. J. Math. and I t s Appl : Vol . 13, No. 1.
- [3] Kusrini dan Emha Taufiq Lutfi. 2009. *Algoritma Data Mining*. Yogyakarta : Andi Offset
- [4] Olivia, Cici, Indwiarti dan Yulian Sibaroni. 2015. *Analisis Prediksi Churn Menggunakan Metode Logistic Regression Dan Algoritma Decision Tree*. E-Proceeding Of Engineering : Vol.2, No.2.
- [5] Arifin, Muhammad. 2015. *Ig-Knn Untuk Prediksi Customer Churn Telekomunikasi*. Jurnal Simetris, Vol 6 No.1.
- [6] Jang, S.; Sun T., dan Mizutani E. (1997). *Neuro-Fuzzy and Soft Computing A Computational Approach to Learning and Machine intelligence*, Prentice Hall, Inc. ISBN 0132610663.
- [7] Kusumadewi, Sri dan Sri Hartati. 2010. *Neuro-Fuzzy Integrasi Sistem Fuzzy & Jaringan Syaraf*. Yogyakarta : Graha Ilmu.