

KOMBINASI METODE MULTILAYER PERCEPTRON DAN TEORI FUZZY UNTUK KLASIFIKASI DATA MEDIS

Dini Adni Navastara¹, Julia Safitri², dan Diana Purwitasari³

Jurusan Teknik Informatika, Institut Teknologi Sepuluh Nopember (ITS)
Jl. Teknik Kimia, Kampus ITS Sukolilo, Surabaya-Indonesia

dini_navastara@if.its.ac.id¹

ABSTRAK

Kemajuan teknologi informasi saat ini banyak digunakan untuk membantu komputasi data dalam berbagai penelitian, salah satunya dalam bidang kesehatan (medis). Dibutuhkan peranan teknologi informasi untuk membantu komputasi dengan melakukan klasifikasi data medis berdasarkan keterangan-keterangan yang menjelaskan data tersebut. Dalam tahapan klasifikasi terkadang data masih dapat timbul beberapa ketidakpastian yang disebabkan oleh adanya informasi yang kurang tepat, ambiguitas dalam data masukan, tumpang tindih batas-batas antara kelas, dan ketidaktentuan dalam mendefinisikan fitur. Untuk mengatasi permasalahan tersebut, dilakukan implementasi metode Neuro-fuzzy yang menggunakan kombinasi *Neural Network* dan pendekatan teori *Fuzzy Set* untuk klasifikasi data medis. Neuro-fuzzy merupakan penggabungan antara sistem *Neural Network* dan sistem fuzzy. Sistem logika fuzzy memiliki kemampuan menangani data pengetahuan dalam persepsi dan penalaran seperti otak manusia tetapi tidak memiliki kemampuan untuk belajar dan beradaptasi. Sedangkan *Neural Network* memiliki kemampuan untuk belajar dan beradaptasi tetapi tidak memiliki kemampuan penalaran seperti pada sistem logika fuzzy. Salah satu algoritma yang dapat diandalkan dalam klasifikasi data dari domain *Neural Network* adalah *Multilayer Perceptron Backpropagation Network* (MLPBPN). Dari hasil uji coba didapatkan tingkat akurasi pada dataset Breast Cancer Wisconsin, Mammographic Mass, dan Pima Indians Diabetes masing-masing mencapai 97,512%, 84,666%, dan 81,613%. Selain itu, metode Neuro-Fuzzy dapat meningkatkan akurasi rata-rata sebesar 3,536% dari metode ANFIS.

Kata kunci: Neuro-fuzzy, Multilayer Perceptron Backpropagation Network, klasifikasi, Neural Network, data medis

ABSTRACT

Nowadays, information technology is currently widely used to assist data computing in various studies, one of them in the field of health (medical). It takes the role of information technology to assist computation by classifying medical data based on data descriptions. In the classification stage, sometimes the data still appear some uncertainty caused by the lack of information, ambiguity in input data, overlapping boundaries between classes, and uncertainty in defining features. To overcome those problems, we proposed a combined neural network method and fuzzy set theory for medical data classification, called Neuro-fuzzy. Fuzzy logic system is able to handle knowledge in perception and reasoning. Whereas, neural network has ability to learn and adapt. One of reliable algorithm in neural network domain for classification is Multilayer Perceptron Backpropagation Network (MLPBPN). Based on experimental results, the accuracy of proposed method for Breast Cancer Wisconsin, Mammographic Mass, and Pima Indians Diabetes datasets are 97,512%; 84,666%; and 81,613%; respectively. Whereas, the proposed method is able to increase accuracy average of 3,536% compare to ANFIS.

Keywords: Neuro-fuzzy, Multilayer Perceptron Backpropagation Network, classification, Neural Network, medical data

1. PENDAHULUAN

Data medis merupakan data yang berisi riwayat penyakit seseorang berdasarkan diagnosis, hasil pemeriksaan fisik, dan atribut-atribut lain yang menerangkan suatu penyakit. Dibutuhkan peranan teknologi informasi untuk membantu komputasi dengan melakukan klasifikasi data medis berdasarkan keterangan-keterangan yang menjelaskan data tersebut. Dalam tahapan klasifikasi data masih dapat timbul beberapa ketidakpastian yang disebabkan oleh adanya informasi yang kurang tepat, ambiguitas dalam data masukan, tumpang tindih batas-batas antara kelas, dan ketidaktentuan dalam mendefinisikan fitur (Liu, 2004). Untuk mengatasi permasalahan tersebut, maka digunakan teknik klasifikasi Neuro-fuzzy baru yang menggunakan kombinasi *Neural Network* dan pendekatan teori *Fuzzy Set* pada data medis.

Neural Network merupakan sistem komputasi cerdas untuk pemodelan data yang cukup populer untuk presisi tinggi dan kemampuan belajar yang tinggi bahkan ketika informasi yang tersedia sangat sedikit (Ghosh, 2014). Salah satu algoritma yang dapat diandalkan dalam klasifikasi data dari domain *Neural Network* adalah *Multilayer Perceptron Backpropagation Network* (MLPBPN). Teori *Fuzzy Set* sendiri sangat fleksibel dalam menangani aspek yang berbeda dari ketidakpastian atau ketidaklengkapan tentang situasi dalam kehidupan nyata. Sebuah himpunan fuzzy pada variabel numerik diwakilkan oleh fungsi keanggotaan fuzzy (Eberhart, 2007). Fungsi keanggotaan memainkan peranan yang sangat penting untuk merepresentasikan masalah dan menghasilkan keputusan yang akurat (Suyanto, 2014). Pada metode klasifikasi Neuro-fuzzy ini menggunakan fungsi keanggotaan *Generalized Bellshaped* yang dapat memberikan lebih banyak fleksibilitas untuk klasifikasi dengan memodifikasi parameter yang dimilikinya. Metode klasifikasi Neuro-fuzzy ini akan diterapkan pada tiga dataset dari *UCI Machine Learning Repository*, antara lain: Breast Cancer Wisconsin, Mammographic Mass, dan Pima Indians Diabetes. Ketiga dataset tersebut merupakan data medis yang cukup populer digunakan untuk klasifikasi data dan memiliki tidak lebih dari sepuluh atribut.

Untuk mengatasi permasalahan yang telah disebutkan sebelumnya, dilakukan implementasi metode Neuro-fuzzy untuk klasifikasi data medis yang diharapkan dapat menghasilkan komputasi yang lebih akurat dalam klasifikasi data medis dibandingkan model klasifikasi lainnya. Hasil uji coba metode Neuro-fuzzy ini akan dianalisis dengan hasil uji coba dari metode ANFIS.

2. METODOLOGI

Metode Neuro-Fuzzy menerapkan dua algoritma utama yaitu teori himpunan fuzzy dan



Gambar. 1. Diagram alir sistem Neuro-Fuzzy secara umum

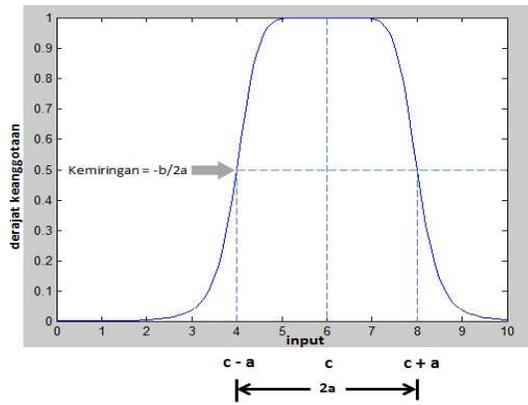
Multilayer Perceptron Backpropagation Network (MLPBPN). Sistem ini memiliki proses fuzzifikasi menggunakan fungsi keanggotaan *Generalized Bellshaped* dan memiliki proses *training* serta *testing* pada MLPBPN. Proses yang dilakukan oleh sistem Neuro-Fuzzy adalah untuk menghasilkan nilai akurasi dari klasifikasi data. Selain itu, sistem Neuro-Fuzzy juga menghasilkan RMSE dan *f-measure*. Alur sistem Neuro-fuzzy secara umum dapat dilihat pada Gambar 1.

Pada Gambar 1, data masukan berupa matriks yang telah melewati tahap *preprocessing* akan diubah menjadi matriks keanggotaan melalui proses fuzzifikasi. Fuzzifikasi pada sistem Neuro-Fuzzy ini menggunakan fungsi keanggotaan *Generalized Bellshaped*. Matriks keanggotaan hasil dari fuzzifikasi akan digunakan sebagai input untuk membangun pengklasifikasi MLPBPN. Pengklasifikasi MLPBPN dibangun dengan jaringan syaraf tiruan *feedforward* yang terdiri dari sejumlah neuron yang dihubungkan oleh bobot-bobot penghubung kemudian menggunakan *backpropagation* untuk *training*. Model MLPBPN ini selanjutnya digunakan untuk *testing* data uji. Kemudian dilakukan pemetaan data hasil *testing* ke dalam kelas tertentu pada tahap defuzzifikasi.

Tahap Preprocessing Dataset

Terdapat tiga dataset yang digunakan pada sistem, yaitu Breast Cancer Wisconsin (Original), Mammographic Mass, dan Pima Indians Diabetes yang masing-masing memiliki karakteristik yang berbeda. Ketiga dataset yang diambil dari laman *UCI Machine Learning Repository* tersebut merupakan data yang belum siap uji karena masih terdapat banyak *missing value*. Untuk itu dilakukan tahap pra-proses data yang mencakup proses penanganan *missing value*, penghilangan atribut, dan normalisasi.

Penanganan *missing value* pada ketiga dataset dilakukan dengan menghapus atau menghilangkan data yang mengandung *missing value*. Sedangkan penghilangan atribut dilakukan pada dataset Breast Cancer Wisconsin dalam atribut *Sample code number*. Penghilangan atribut digunakan untuk



Gambar. 2. Grafik Fungsi Keanggotaan Generalized Bellshaped dengan parameter a , b , dan c adalah 2, 4, dan 6.

membuang atribut berlebihan sehingga hanya atribut yang relevan yang diambil dari dataset yang diberikan. Kemudian untuk normalisasi dengan memetakan data dalam rentang nilai 1 sampai 10 dilakukan pada dataset Mammographic Mass dan Pima Indians Diabetes, sedangkan pada dataset Breast Cancer Wisconsin tidak dilakukan normalisasi karena rentang data pada tiap atributnya bernilai 1 sampai 10.

Tahap Fuzzifikasi

Tahap fuzzifikasi dilakukan dengan mengubah data input yang telah melalui tahap *preprocessing* menjadi matriks keanggotaan berdasarkan fungsi keanggotaan. Fungsi keanggotaan memainkan peranan yang sangat penting untuk merepresentasikan masalah dan menghasilkan keputusan yang akurat. Fungsi keanggotaan memetakan tiap elemen dari input menjadi derajat keanggotaan antara 0 dan 1. Untuk proses fuzzifikasi pada sistem Neuro-Fuzzy ini digunakan fungsi keanggotaan *Generalized Bellshaped*. Berikut merupakan rumus fungsi keanggotaan *Generalized Bellshaped* yang ditentukan oleh tiga parameter berbeda, yaitu a , b , dan c (Ghosh, 2014):

$$mf_{ij}(y_{ij}; a, b, c) = \frac{1}{1 + \left| \frac{y_{ij} - c}{a} \right|^{2b}} \quad (1)$$

dimana y_{ij} adalah data input atribut ke- i pada baris data ke- j , parameter a menunjukkan setengah lebar kurva, parameter b (bersama-sama dengan a) mengontrol kemiringan di titik potong yang berbeda, sedangkan parameter c menentukan pusat kurva. Gambar 2 merupakan grafik fungsi keanggotaan *Generalized Bellshaped* dengan parameter a , b , dan c adalah 2, 4, dan 6.

Pada proses fuzzifikasi akan dihasilkan matriks keanggotaan orde $N \times M$ yang terdiri dari derajat keanggotaan dari N atribut yang berbeda ke M baris data. Matriks keanggotaan dari proses fuzzifikasi ini akan digunakan sebagai input untuk MLPBPN.

Tahap Membangun Pengklasifikasi MLPBPN

Multilayer Perceptron adalah jaringan syaraf tiruan *feedforward* yang terdiri dari sejumlah neuron yang dihubungkan oleh bobot-bobot penghubung. Neuron-neuron tersebut disusun dalam lapisan-lapisan yang terdiri dari satu lapisan input (*input layer*), satu atau lebih lapisan tersembunyi (*hidden layer*), dan satu lapisan output (*output layer*). *Multilayer Perceptron* menggunakan teknik pembelajaran *supervised*, yaitu *backpropagation* untuk *training* (Riedmiller, 1994). *Multilayer Perceptron* yang digunakan pada sistem ini adalah *single hidden layer*.

Pada tahap ini, data input berupa matriks keanggotaan hasil fuzzifikasi dibagi menjadi *data train* dan *data test* dengan mode pembagian *data k-fold cross-validation*. Data input dibagi menjadi data untuk *training* adalah dua per tiga dari data input dan data untuk *testing* adalah satu per tiga dari data input. Berikut ini adalah algoritma *Multilayer Perceptron Backpropagation Network* (Suyanto, 2014):

- 1) Definisikan matriks masukan dan matriks kelas
- 2) Inisialisasi jumlah node pada hidden layer
- 3) Inisialisasi nilai ambang *mean-square error* sebagai kondisi berhenti serta menentukan bobot dan bias dari *input layer* ke *hidden layer* dan dari *hidden layer* ke *output layer* dengan nilai random.
- 4) Ulangi langkah-langkah berikut selama kondisi berhenti belum terpenuhi:

Feedforward

- Tiap unit input menerima sinyal input dan meneruskan sinyal tersebut pada lapisan selanjutnya (*hidden layer*).
- Fungsi aktivasi yang digunakan pada *hidden layer* adalah fungsi aktivasi sigmoid.

$$hidden_k = \frac{1}{1 + \exp^{-hidden_Net_j}} \quad (2)$$

dimana *hidden_Net_j* adalah input jaringan yang didefinisikan dengan persamaan berikut (Ghosh, 2014):

$$hidden_Net_j = \sum_i weight_{jk} * mf_{ij}(y_{ij}) + bias_k \quad (3)$$

dimana *weight_{jk}* adalah bobot neuron j pada layer sebelumnya ke neuron k pada *hidden layer*, $mf_{ij}(y_{ij})$ adalah derajat keanggotaan dari atribut i ke kelas j hasil fuzzifikasi, dan *bias_k* adalah bias neuron k .

- Hasil keluaran *hidden layer* digunakan untuk menghitung keluaran dari *output layer* $output_l$ (Ghosh, 2014).

$$output_Net_k = \sum_j weight_{kl} * hidden_k + bias_l \quad (4)$$

$$output_l = \frac{1}{1+exp^{-output_{Net_k}}} \quad (5)$$

- Keluaran dari *output layer* dibandingkan dengan target (kelas) yang diinginkan. Selisih nilai target (kelas) dengan keluaran dari *output layer* adalah nilai error.

Backpropagation

- Hitung faktor δ output layer berdasarkan error di unit output layer $output_l$.

$$\delta_l = (1 - output_l^2) * Error \quad (6)$$

- Hitung perubahan bobot $weight_{kl}$ dengan *learning rate* η

$$\Delta weight_{kl} = \eta * \delta_l * hidden_k \quad (7)$$

- Hitung perubahan $bias_l$.

$$\Delta bias_l = \eta * \delta_l \quad (8)$$

- Nilai error yang didapat dikembalikan lagi ke lapis-lapis sebelumnya (*hidden layer*). Neuron pada lapisan tersebut akan memperbaiki nilai bobot dan biasnya.

$$\delta_k = (1 - hidden_k^2) * weight_{kl} * \delta_l \quad (9)$$

- Setelah neuron-neuron mendapatkan nilai yang sesuai dengan kontribusinya pada error, maka bobot dan bias jaringan akan diperbaiki agar error dapat diperkecil.

$$weight_{kl} = weight_{kl} + \Delta weight_{kl} \quad (10)$$

$$bias_l = bias_{kl} + \Delta bias_l \quad (11)$$

$$weight_{jk} = weight_{jk} + \Delta weight_{jk} \quad (12)$$

$$bias_k = bias_{jk} + \Delta bias_k \quad (13)$$

- Bobot dan bias yang baru (hasil perbaikan) dapat digunakan kembali untuk mengetahui apakah error yang akan dihasilkan sudah cukup kecil.

Langkah-langkah tersebut untuk satu kali siklus pelatihan (satu *epoch*). Pelatihan akan diulang beberapa kali hingga *mean-square error* yang diinginkan terpenuhi.

Selanjutnya, model MLPBPN ini akan digunakan untuk menguji *data test* yang telah dibagi pada mode *k-fold cross-validation*. Pengujian *data test* akan menghasilkan keluaran yang kemudian dipetakan ke dalam kelas tertentu pada tahap defuzzifikasi.

Defuzzifikasi

Tahap defuzzifikasi dilakukan dengan memetakan data hasil *testing* ke dalam kelas tertentu

dengan batas 0,5. Jika data hasil *testing* lebih dari 0,5, maka dikelompokkan ke kelas 1. Jika data hasil *testing* kurang dari 0,5, maka dikelompokkan ke kelas 0. Hal ini dilakukan karena ketiga dataset yang digunakan pada sistem hanya memiliki dua kelas.

3. HASIL DAN PEMBAHASAN

Data Uji Coba

Data yang digunakan untuk uji coba terdiri dari tiga data medis, yaitu:

1. Breast Cancer Wisconsin (Original)
Dataset Breast Cancer Wisconsin (Original) merupakan dataset untuk diagnosis kanker payudara. Jumlah instance pada dataset ini adalah 699. Dataset ini memiliki 11 atribut termasuk atribut kelas, 10 atribut yang pertama dianggap sebagai atribut masukan. Terdapat 16 instance dalam dataset ini yang mengandung nilai atribut yang hilang. Dataset Breast Cancer Wisconsin memiliki dua kelas, yaitu jinak sebanyak 458 instance dan ganas sebanyak 241 instance.
2. Mammographic Mass
Dataset Mammographic Mass merupakan patokan (benchmark) yang ditetapkan untuk deteksi kanker payudara. Dataset UCI yang diberikan adalah hasil dari penelitian yang dilakukan pada massa jaringan payudara manusia dengan menggunakan *Breast Imaging-Reporting and Data System (BI-RADS)* dikombinasikan dengan mamografi digital. Dataset ini memiliki 961 jumlah instance dan terdiri dari 6 atribut termasuk atribut kelas yang bernama *Severity*. Dataset Mammographic Mass memiliki dua kelas, yaitu jinak (normal) sebanyak 516 instance dan ganas (kanker) sebanyak 445 instance.
3. Pima Indians Diabetes.

Dataset Pima Indians Diabetes merupakan dataset untuk deteksi diabetes dan merupakan hasil dari survei penelitian yang dilakukan di *National Institute of Diabetes and Digestive and Kidney Diseases, United States* pada pasien wanita keturunan Pima Indian yang memiliki usia minimal 21 tahun. Jumlah instance pada dataset ini adalah 768. Dataset ini memiliki 9 atribut termasuk atribut kelas. Dataset Pima Indians Diabetes memiliki dua kelas, yaitu "dinyatakan negatif untuk diabetes" sebanyak 500 instance dan "dinyatakan positif untuk diabetes" sebanyak 268 instance. Pada data ini terdapat *missing value* sehingga diperlukan preprocessing untuk menangani *missing value* ini.

Sebelum dilakukan uji coba, ketiga data tersebut telah melalui tahap *preprocessing*.

Pada tahap membangun pengklasifikasi *Multilayer Perceptron Backpropagation Network*, data dibagi menjadi *data train* dan *data test* dengan teknik *pembagian k-fold cross-validation*. Proporsi pembagian data adalah *data train* dua per tiga dari jumlah masing-masing data dan *data test* sepertiga dari jumlah masing-masing data.

Skenario Pengujian

Uji coba dilakukan dengan menggunakan beberapa skenario sebagai berikut:

1. Pengujian untuk menentukan parameter fungsi keanggotaan *Generalized Bellshaped*.
Penentuan parameter *Generalized Bellshaped* ini selain mempengaruhi tingkat akurasi klasifikasi data, juga mempengaruhi pembentukan grafik fungsi keanggotaan *Generalized Bellshaped*. Fungsi keanggotaan *Generalized Bellshaped* memiliki tiga parameter berbeda, yaitu parameter a menunjukkan setengah lebar kurva, parameter b (bersama-sama dengan a) mengontrol kemiringan di titik potong yang berbeda, sedangkan parameter c menentukan pusat kurva. Pada pengujian ini, parameter a nilainya tidak lebih besar dari parameter c , karena setengah lebar kurva tidak bisa lebih dari nilai pusat kurva. Sedangkan parameter b , nilainya berada di antara parameter a dan c . Oleh karena itu, parameter pusat kurva c nilainya harus lebih besar dari parameter a dan b .
2. Pengujian untuk menentukan parameter *learning rate* awal pada MLPBPN.
Pengujian untuk menentukan *learning rate* ini dilakukan agar MLPBPN bekerja dengan stabil. *Learning rate* menunjukkan intensitas dalam proses *learning* dan konvergensi dari *training*. Jika *learning rate* terlalu kecil maka akan memperlambat proses konvergensi dan jika terlalu besar maka MLPBPN akan bekerja tidak stabil. Oleh karena itu, pengujian untuk menentukan parameter *learning rate* awal sangat diperlukan untuk membangun sistem Neuro-Fuzzy ini.
Setelah mendapat nilai parameter fungsi keanggotaan untuk masing-masing dataset, dilakukan pengujian untuk menentukan *learning rate*. Nilai *learning rate* yang akan digunakan nantinya, diambil dari tingkat akurasi yang terbaik dari masing-masing dataset.
3. Uji kinerja klasifikasi pada model Neuro-Fuzzy dan model ANFIS.
Skenario ini bertujuan untuk membandingkan hasil kinerja klasifikasi metode Neuro-Fuzzy dan ANFIS (*Adaptive Neuro-fuzzy Inference System*). ANFIS merupakan jaringan node (*processing elements*) dan link terarah (bobot) yang secara fungsional setara dengan Sistem

Inferensi Fuzzy (Ghosh, 2014). ANFIS umumnya menggunakan algoritma pembelajaran *hybrid* yang merupakan kombinasi dari *gradient descent* dan metode *least square* untuk mengadaptasi parameter dalam jaringan adaptif. Sederhananya, ANFIS adalah penggabungan mekanisme fuzzy inference system model Tagaki-Sugeno-Kang (TSK) yang digambarkan dalam arsitektur jaringan syaraf. Sedangkan Neuro-Fuzzy merupakan kombinasi dari MLPBPN dengan teori himpunan fuzzy. Karena terdapat kemiripan struktur antara kedua metode, maka ANFIS digunakan sebagai metode pembandingan.

Hasil dan Analisis Uji Coba

Pada skenario pengujian untuk menentukan parameter fungsi keanggotaan *Generalized Bellshaped* dihasilkan nilai parameter a , b , dan c dengan akurasi terbaik masing-masing untuk dataset Breast Cancer Wisconsin adalah $a = 5$, $b = 7$, dan $c = 8$ dengan akurasi sebesar 97,512%. Sedangkan untuk dataset Mammographic Mass adalah $a = 1$, $b = 2$, dan $c = 9$ dengan akurasi sebesar 84,666%. Kemudian untuk dataset Pima Indians Diabetes adalah $a = 3$, $b = 5$, dan $c = 5$ dengan akurasi sebesar 81,613%.

Pada skenario pengujian untuk menentukan parameter *learning rate* awal pada MLPBPN dihasilkan *learning rate* yang optimal untuk ketiga dataset adalah 0,1. Hasil uji tingkat akurasi berdasarkan parameter *learning rate* dapat dilihat pada Tabel 1.

Pada Tabel 1, dapat dilihat bahwa saat *learning rate* kecil tingkat akurasi rendah. Semakin besar *learning rate* semakin besar pula tingkat akurasinya hingga *learning rate* bernilai 0,1. Namun, saat *learning rate* bernilai lebih dari 0,1 tingkat akurasinya kembali rendah. Oleh karena itu, *learning rate* yang optimal untuk ketiga dataset Breast Cancer Wisconsin, Mammographic Mass, dan Pima Indians Diabetes adalah 0,1.

Hasil skenario uji kinerja klasifikasi pada model Neuro-Fuzzy dan model ANFIS dapat dilihat pada Tabel 2 yang menunjukkan hasil kinerja klasifikasi data Breast Cancer Wisconsin, Tabel 3 yang menunjukkan hasil kinerja klasifikasi data Mammographic Mass, dan Tabel 4 yang menunjukkan hasil kinerja klasifikasi data Pima Indians Diabetes.

Dari Tabel 2 dapat dilihat bahwa pada data Breast Cancer Wisconsin, tingkat akurasi Neuro-Fuzzy lebih tinggi dari ANFIS. Neuro-Fuzzy memiliki akurasi sebesar 97,512%, sedangkan ANFIS memiliki akurasi sebesar 96,049%. Neuro-Fuzzy memiliki RMSE yang lebih baik dari ANFIS, yaitu 0,129. Sedangkan berdasarkan nilai *f-measure*, Neuro-Fuzzy lebih baik dibandingkan ANFIS, yaitu sebesar 98,072%.

Dari Tabel 3 dapat dilihat bahwa pada data Mammographic Mass, tingkat akurasi Neuro-Fuzzy

Tabel 1. Hasil uji tingkat akurasi berdasarkan parameter *learning rate*

Learning Rate	Akurasi pada dataset (%)		
	Breast Cancer Wisconsin	Mammographic Mass	Pima Indians Diabetes
0,01	96,629	80,977	66,837
0,05	97,072	83,683	73,449
0,08	97,362	84,296	76,285
0,1	97,512	84,666	81,613
0,5	97,218	83,435	78,307
0,8	97,072	83,191	76,265

Tabel 2. Hasil kinerja klasifikasi data Breast Cancer Wisconsin

Ukuran Kinerja	Akurasi pada dataset (%)	
	Neuro-Fuzzy	ANFIS
Akurasi	97,512	96,049
RMSE	0,129	0,160
f-measure	98,072	96,976
Waktu (detik)	11,424	2,142

Tabel 3. Hasil kinerja klasifikasi data Mammographic Mass

Ukuran Kinerja	Akurasi pada dataset (%)	
	Neuro-Fuzzy	ANFIS
Akurasi	84,666	82,948
RMSE	0,223	0,377
f-measure	84,846	83,782
Waktu (detik)	11,955	1,285

lebih tinggi dari ANFIS. Neuro-Fuzzy memiliki akurasi sebesar 84,666%, sedangkan ANFIS memiliki akurasi sebesar 82,948%. Kemudian Neuro-Fuzzy memiliki RMSE yang lebih baik dari ANFIS yaitu sebesar 0,223. Selanjutnya berdasarkan nilai *f-measure*, Neuro-Fuzzy lebih baik dibandingkan ANFIS yaitu sebesar 84,846%.

Dari Tabel 4 dapat dilihat bahwa pada data Pima Indians Diabetes, tingkat akurasi Neuro-Fuzzy lebih tinggi dari ANFIS. Neuro-Fuzzy memiliki akurasi sebesar 81,613%, sedangkan ANFIS memiliki akurasi sebesar 76,264%. Berdasarkan nilai RMSE, Neuro-Fuzzy memiliki RMSE yang lebih baik dari ANFIS yaitu sebesar 0,245, sedangkan ANFIS memiliki RMSE sebesar 0,398. Pada Tabel 4 ditunjukkan bahwa nilai *f-measure* Neuro-Fuzzy lebih besar dari ANFIS yaitu sebesar 86,926%, sedangkan nilai *f-measure* ANFIS adalah 83,112%. Untuk waktu yang dibutuhkan dalam proses klasifikasi data pada ketiga dataset, Neuro-Fuzzy membutuhkan waktu lebih lama daripada ANFIS.

Dari hasil kinerja klasifikasi pada ketiga dataset dapat dilihat bahwa nilai akurasi yang dihasilkan sistem Neuro-Fuzzy pada ketiga dataset dapat lebih baik dari metode ANFIS. Analisis tingkat

Tabel 4. Hasil kinerja klasifikasi data Pima Indians Diabetes

Ukuran Kinerja	Akurasi pada dataset (%)	
	Neuro-Fuzzy	ANFIS
Akurasi	81,613	76,264
RMSE	0,245	0,398
f-measure	86,926	83,112
Waktu (detik)	6,847	0,590

Tabel 5. Analisis tingkat akurasi metode Neuro-Fuzzy dan ANFIS

Model	Akurasi pada dataset (%)		
	Breast Cancer Wisconsin	Mammographic Mass	Pima Indians Diabetes
Neuro-Fuzzy	97,512	84,666	81,613
ANFIS	96,049	82,948	76,264
Selisih	1,463	1,718	5,349
Peningkatan Rata-Rata	1,523	2,071	7,014

Tabel 6. Analisis waktu pada metode Neuro-Fuzzy dan ANFIS

Model	Waktu pada dataset (detik)		
	Breast Cancer Wisconsin	Mammographic Mass	Pima Indians Diabetes
Neuro-Fuzzy	11,424	11,955	6,847
ANFIS	2,142	1,285	0,590
Selisih	9,282	10,670	6,257
Rata-Rata		8,736	

akurasi metode Neuro-Fuzzy dan ANFIS dapat dilihat pada Tabel 5, dimana nilai akurasi yang dihasilkan sistem Neuro-Fuzzy pada ketiga dataset dapat meningkat rata-rata 3,536% dari metode ANFIS. Hal ini disebabkan pada sistem Neuro-Fuzzy parameter fungsi keanggotaan *Generalized Bellshaped* untuk tiap dataset, *learning rate* awal, dan bobot serta bias telah ditentukan.

Kemudian untuk nilai RMSE yang dihasilkan sistem Neuro-Fuzzy pada ketiga dataset juga lebih baik dari metode ANFIS. Hal ini disebabkan node *pada hidden layer* ditentukan di awal dan *learning rate* awal juga telah ditentukan untuk ketiga dataset. Selain itu, pada tahap awal membangun MLPBPN juga telah ditentukan batas nilai *mean square error* dan jumlah *epoch* yang digunakan untuk pelatihan.

Sementara itu, untuk waktu yang dibutuhkan dalam proses klasifikasi data, Neuro-Fuzzy membutuhkan waktu lebih lama dari ANFIS. Analisis waktu pada metode Neuro-Fuzzy dan ANFIS dapat dilihat pada Tabel 6. Waktu yang dibutuhkan metode Neuro-Fuzzy lebih lama rata-rata 8,736 detik dari ANFIS. Hal ini disebabkan ANFIS menggunakan *Recursive Least Square Estimator* (RLSE) untuk

memperbaiki parameter konsekuen dan *gradient descent* untuk memperbaiki parameter premis.

4. KESIMPULAN

Dari hasil selama proses perancangan, implementasi, dan pengujian dapat diambil kesimpulan bahwa tingkat akurasi dan nilai RMSE yang dihasilkan metode Neuro-Fuzzy pada ketiga dataset dapat lebih baik dari metode ANFIS. Tingkat akurasi terbaik yang dihasilkan metode Neuro-Fuzzy pada dataset Breast Cancer Wisconsin sebesar 97,512%, sedangkan tingkat akurasi terbaik pada dataset Mammographic Mass sebesar 84,666%, dan tingkat akurasi terbaik pada dataset Pima Indians Diabetes sebesar 81,613%. Metode Neuro-Fuzzy dapat meningkatkan akurasi rata-rata sebesar 3,536% dari metode ANFIS, namun metode Neuro-Fuzzy membutuhkan waktu lebih lama rata-rata 8,736 detik dari ANFIS dalam proses klasifikasi data.

Saran yang dapat diberikan untuk pengembangan dan perbaikan sistem di masa yang akan datang adalah mengembangkan cara penentuan parameter fungsi keanggotaan Generalized Bellshaped menjadi otomatis agar tidak melakukan trial and error. Selain itu untuk pengembangan sistem selanjutnya, juga dapat digunakan data medis atau data lain selain data medis untuk dilakukan klasifikasi dengan metode Neuro-Fuzzy.

DAFTAR PUSTAKA

- Eberhart, R., & Shi, Y. 2007. Computational Intelligence: Concepts to Implementations. USA: Elsevier.
- Ghosh, S., Biswas, S., Sarkar, D., Sarkar, Partha P., Sept. 2014. "A Novel Neuro-fuzzy Classification Technique for Data Mining". Egyptian Informatics Journal, 129-147.
- Liu B. 2004. Uncertainty Theory: An Introduction to Its Axiomatic Foundations. Berlin: Springer Verlag.
- Riedmiller, M. 1994. Advanced Supervised Learning in Multi-layer Perceptrons — From Backpropagation to Adaptive Learning Algorithms. Computer Standards & Interfaces, 265-278.
- Suyanto. 2014. Artificial Intelligence: Searching Reasoning Planning and Learning. Bandung: Informatika.