

Penerapan Algoritma Artificial Bee Colony untuk Optimasi Bobot Awal pada Neural Network Backpropagation dalam Klasifikasi Penyakit Diabetes

Application of Artificial Bee Colony Algorithm for Initial Weight Optimization in Neural Network Backpropagation in Diabetes Disease Classification

Zainal Abidin^{1) a)}, Ricardus Anggi Pramunendar^{2) b)}

¹⁾ Sekolah Tinggi Teknik Pati

²⁾ Univ. Dian Nuswantoro

a) Jl. Raya Pati-Tayu Km. 4,5 Pati. 59119. Jawa Tengah

Email: zainal.frstd@yahoo.co.id

ABSTRACT

In medicine, patient safety and disease diagnosis are key aspects. Diabetes is a complex disease caused by insulin deficiency or resistance, and requires early detection and treatment. Machine learning is now widely used in disease diagnosis, one of which is through classification using Neural Network with Backpropagation training. However, this method often experiences slow convergence and gets stuck in local minimums. This study proposes the use of the Artificial Bee Colony (ABC) algorithm to optimize Backpropagation Neural Network (BPNN) training in diabetes classification. Based on previous studies, the ABC-BPNN approach showed high accuracy. The purpose of this study was to compare the classification accuracy of ABC-BPNN with other methods such as Tuning Neural Network and Fuzzy-MABC. As a result, ABC-BPNN showed superior accuracy.

Keywords: Diabetes, Disease Diagnosis, Machine Learning, Neural Network, Backpropagation, Artificial Bee Colony, Weight Optimization, Classification

ABSTRAK

Dalam dunia kedokteran, keselamatan pasien serta diagnosis penyakit merupakan aspek utama. Diabetes adalah penyakit kompleks yang disebabkan oleh kekurangan atau resistensi insulin, dan memerlukan deteksi serta penanganan dini. *Machine learning* kini banyak digunakan dalam diagnosis penyakit, salah satunya melalui klasifikasi menggunakan *Neural Network* dengan pelatihan *Backpropagation*. Namun, metode ini sering mengalami konvergensi lambat dan terjebak di *local minimum*. Penelitian ini mengusulkan penggunaan algoritma *Artificial Bee Colony* (ABC) untuk mengoptimalkan pelatihan *Backpropagation Neural Network* (BPNN) dalam klasifikasi diabetes. Berdasarkan studi sebelumnya, pendekatan ABC-BPNN menunjukkan akurasi tinggi. Tujuan penelitian ini adalah membandingkan akurasi klasifikasi ABC-BPNN dengan metode lain seperti *Tuning Neural Network* dan *Fuzzy-MABC*. Hasilnya, ABC-BPNN menunjukkan akurasi yang lebih unggul.

Kata kunci : Diabetes, Diagnosa Penyakit, *Machine Learning*, *Neural Network*, *Backpropagation*, *Artificial Bee Colony*, *Optimasi Bobot*, Klasifikasi

PENDAHULUAN

Seiring berkembangnya ilmu pengetahuan dan teknologi, terutama ilmu komputer yang dapat dimanfaatkan dan diterapkan pada segala macam bidang. salah satu pemanfaatan ilmu komputer yaitu dibidang kesehatan. Bidang ilmu komputer terutama sistem cerdas tumbuh pesat dalam dunia kedokteran dan kesehatan yang dapat membantu dokter untuk meningkatkan kualitas perawatan pasien [3]. Pada bidang kedokteran dan kesehatan hal yang paling penting adalah keamanan dan keselamatan. Selain itu klasifikasi dan diagnosis penyakit adalah tugas utama untuk para dokter. Diabetes merupakan salah satu penyakit kompleks yang ditandai dengan kekurangan insulin atau resistensi terhadap insulin, hormon yang sangat penting bagi metabolisme pada gula darah. Pankreas orang yang sehat memproduksi insulin untuk membantu metabolisme gula dalam darah dan menjaga tingkat glukosa pada kisaran

normal. Sedangkan untuk penderita diabetes pankreas tidak dapat memproduksi insulin dengan baik mengakibatkan glukosa dari aliran darah tidak dapat dihapus. Klasifikasi diabetes meliputi dua kelas utama yaitu diabetes 1 adalah gangguan autoimun dimana sel-sel beta yang memproduksi insulin dihancurkan oleh sistem kekebalan tubuh yang mengakibatkan tubuh tidak mampu memproduksi insulin. Diabetes tipe 2 yaitu jenis insulin yang diproduksi dalam jumlah yang cukup dan tidak dapat digunakan oleh tubuh untuk mengendalikan kadar gula darah. Deteksi dan pengobatan terhadap diabetes diperlukan untuk mengurangi penyakit metabolic yang berhubungan [1]. *Machine Learning* merupakan pendekatan yang semakin berkembang dan banyak digunakan untuk deteksi dan diagnosis penyakit dengan cara klasifikasi suatu pola [2]. Klasifikasi dikenal sebagai *supervised learning* yaitu untuk setiap pelatihan diberikan label yang menunjukkan kelasnya. Klasifikasi telah berhasil diterapkan pada berbagai bidang percobaan ilmiah, persetujuan kredit dan diagnosa medis [4].

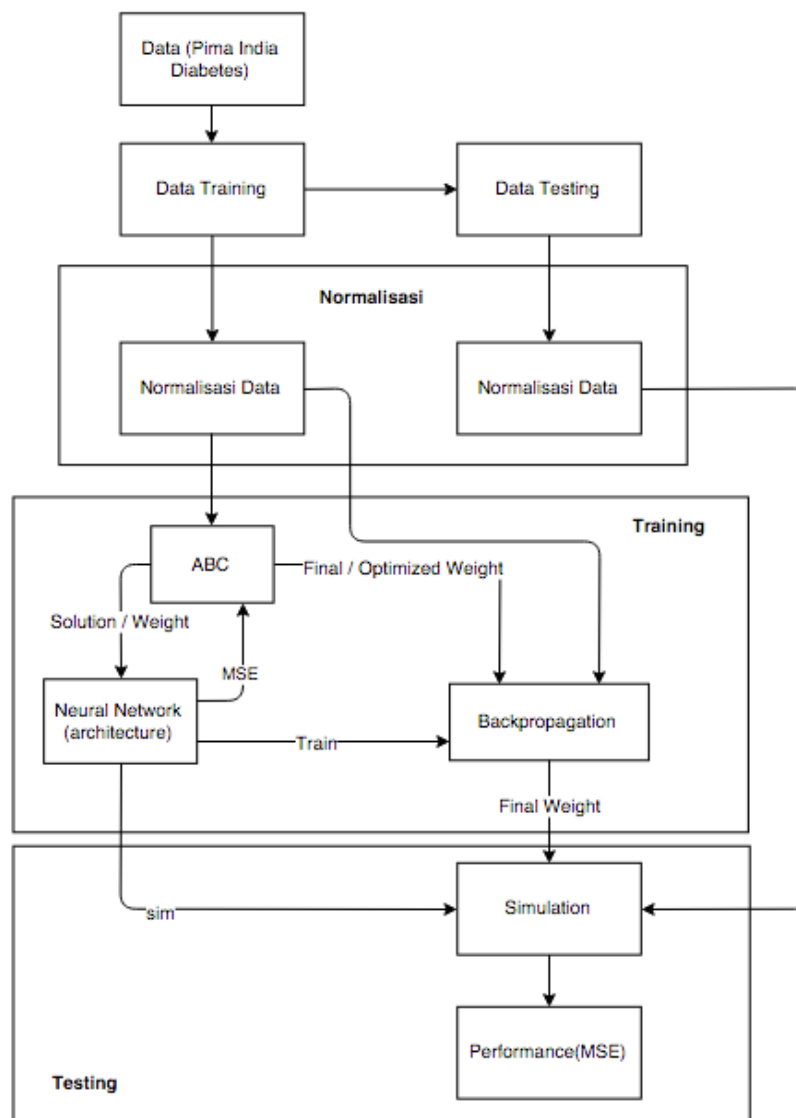
Dengan demikian masalah dalam merancang struktur *neural network backpropagation* yang mendekati optimal belum terpecahkan [5]. Untuk mengatasi masalah tersebut teknik pencarian global memiliki kemampuan untuk menghindari local minima yang digunakan untuk mengatur bobot *neural network*. Algoritma yang biasa digunakan seperti algoritma *evolutioner*, *tabu search*, *ant colony optimization* (ACO), *particle swarm optimization* (PSO) dan *artificial bee colony* (ABC) [5] [6] [7].

Penelitian ini mengusulkan penggunaan *Artificial Bee Colony* (ABC) sebagai algoritma pencarian bobot awal dari *neural network backpropagation* untuk menyelesaikan kasus klasifikasi diabetes tipe II dengan data dari pima india. Data dipilih karena merupakan data yang banyak digunakan untuk penelitian sebelumnya [8] [2]. Penggunaan *Artificial Bee Colony* sebagai algoritma pelatihan *neural network backpropagation* mempunyai performa yang baik [5]. Penggunaan algoritma yang diusulkan untuk menyelesaikan solusi dari klasifikasi penyakit diabetes. Tuning parameter yang dimaksud adalah *neural network* mempunyai parameter seperti *learning rate*, *epoch* dan jumlah *hidden layer*. Peneliti mencoba beberapa kemungkinan dari nilai parameter. Dari eksperimen yang telah dilakukan sebanyak 17 tuning menghasilkan akurasi yang bermacam-macam [9]. Tetapi penelitian tersebut memiliki kelemahan dimana tuning parameter dilakukan dengan cara manual dengan cara mencoba satu per satu setiap tuning menjadikan pada proses eksperimen kurang efisien.

METODE PENELITIAN

Penelitian ini mengusulkan ABC sebagai pencarian bobot awal yang akan digunakan untuk inialisasi bobot awal pada proses pelatihan BPNN. Proses klasifikasi yang diusulkan pada penelitian ini diperlihatkan pada gambar 1. Beberapa tahap klasifikasi antara lain:

Tahapan Penelitian



Gambar 1 : Tahapan Penelitian

1. Input parameter, pada tahap ini parameter awal yang diinisialisasi yaitu jumlah neuron input, jumlah neuron pada hidden layer, epoch, jumlah lebah, maksimal MSE, dan maksimal waktu.
2. Tahap normalisasi data training menjadi interval nol sampai satu.
3. Tahap inialisasi, merupakan proses penentuan bobot pada neural network. Proses penentuan bobot awal menggunakan Artificial Bee Colony. Bobot yang dinilai terbaik digunakan sebagai bobot awal neural network.
4. Tahap pelatihan atau training, merupakan proses pelatihan jaringan dengan metode backpropagation. pada tahap ini bobot akan disesuaikan sehingga dapat menghasilkan hasil klasifikasi yang sesuai target dengan error kecil.

5. Neural network yang sudah melalui tahap training akan diuji (tahap testing) dengan data baru untuk mengetahui akurasi dari hasil pelatihan.
6. Pengukuran akurasi training dan testing.
7. Akurasi dibandingkan dengan metode Neural Network Backpropagation setelah dan sebelum diinisialisasi bobot dengan Artificial Bee Colony.

Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data diabetes. Data digunakan untuk mengetahui pengaruh Artificial Bee Colony pada Neural Network Backpropagation untuk klasifikasi diabetes. Data yang diperoleh dari UCI Machine Learning Repository sumber asli dari *National Institute of Diabetes and Digestive and Kidney Diseases*. Data semua pasien di sini adalah perempuan berumur sekitar 21 tahun dari Pima, India.

Data yang digunakan memiliki 8 atribut dan 1 sebagai label kelas. Sampel data diperlihatkan pada 14table 3.1 dan keterangan atribut sebagai berikut :

1. A1 : Jumlah / Berapa kali Hamil
2. A2 : Kadar Glukosa Plasma pada waktu 2 jam (test toleransi glukosa)
3. A3 : Diastolic tekanan darah (mm Hg)
4. A4 : Ketebalan lipatan kulit trisep (mm)
5. A5 : 2-jam serum insulin (mu U/ml)
6. A6 : Body Mass Index (weight in kg/(height in m)^2)
7. A7 : Fungsi silsilah diabetes
8. A8 : Umur (years)
9. Kelas (0 or 1)

Tabel 1. Sampel Data

No	A1	A1	A3	A4	A5	A6	A7	A8	Kelas
1	6	148	72	35	0	33.6	0.627	50	1
2	1	85	66	29	0	26.6	0.351	31	0
3	8	183	64	0	0	23.3	0.672	32	1
4	1	89	66	23	94	28.1	0.167	21	0
5	0	137	40	35	168	43.1	2.288	33	1
6	5	116	74	0	0	25.6	0.201	30	0
7	3	78	50	32	88	31	0.248	26	1
8	10	115	0	0	0	35.3	0.134	29	0
9	2	197	70	45	543	30.5	0.158	53	1
10	8	125	96	0	0	0	0.232	54	1
11	4	110	92	0	0	37.6	0.191	30	0
12	10	168	74	0	0	38	0.537	34	1
13	10	139	80	0	0	27.1	1.441	57	0

Eksperimen

Eksperimen dilakukan dengan mengumpulkan data-data yang dijadikan penelitian, kemudian membuat program dengan menggunakan *Matlab r2015a*, yang nantinya akan digunakan untuk implementasi dari metode Artificial Bee Colony pada Neural Network Backpropagation. Setelah itu data diproses dengan program yang telah dibuat untuk mengetahui hasil klasifikasi, dari perhitungan akurasi akan diketahui kinerja dari metode Neural Network Backpropagation dan Artificial Bee Colony yang diterapkan pada Neural Network Backpropagation.

HASIL DAN PEMBAHASAN

Pengolahan Data Set

Data yang diperoleh diolah menjadi dataset penelitian. Dataset digunakan untuk proses training dan testing. Struktur dari dataset terdiri dari input dan target. Input terdiri dari 8 atribut dan terdapat 1 output.

Tabel 2. Dataset Penelitian

No	A1	A1	A3	A4	A5	A6	A7	A8	Kelas
1	6	148	72	35	0	33.6	0.627	50	1
2	1	85	66	29	0	26.6	0.351	31	0
3	8	183	64	0	0	23.3	0.672	32	1
4	1	89	66	23	94	28.1	0.167	21	0
5	0	137	40	35	168	43.1	2.288	33	1
6	5	116	74	0	0	25.6	0.201	30	0
7	3	78	50	32	88	31	0.248	26	1
8	10	115	0	0	0	35.3	0.134	29	0
9	2	197	70	45	543	30.5	0.158	53	1
10	8	125	96	0	0	0	0.232	54	1
11	4	110	92	0	0	37.6	0.191	30	0
12	10	168	74	0	0	38	0.537	34	1
13	10	139	80	0	0	27.1	1.441	57	0

Normalisasi Data

Setelah data diolah menjadi dataset tahap selanjutnya adalah normalisasi data. Normalisasi data bertujuan untuk mengoptimalkan pada proses komputasi. Normalisasi menghasilkan data menjadi range lebih kecil. Pada penelitian ini range data hasil normalisasi yaitu [-1 1]. normalisasi menggunakan formula (16)

$$y = \frac{(y_{max}-y_{min})*(x-x_{min})}{(x_{max}-x_{min})} + y_{min} \quad (16)$$

contoh :

$$x_{max} = 543$$

$$x_{min} = 0$$

$$y_{max} = 1, y_{min} = -1$$

$$x = 6$$

$$\begin{aligned} y &= \frac{(y_{max} - y_{min}) * (x - x_{min})}{(x_{max} - x_{min})} + y_{min} \\ &= \frac{(1 - (-1)) * (6 - 0)}{(543 - 0)} + (-1) \\ &= \frac{(2) * (6)}{543} + (-1) \\ &= \frac{12}{543} + (-1) \\ &= -0,98 \end{aligned}$$

Tabel 2. Normalisasi Data

A1	A2	A3	A4	A5	A6	A7	A8
-0.29	0.49	0.18	0.11	-1.00	0.00	-0.53	-0.03
-0.88	-0.14	0.08	-0.08	-1.00	-0.21	-0.77	-0.67
-0.06	0.85	0.05	-1.00	-1.00	-0.31	-0.49	-0.63
-0.88	-0.10	0.08	-0.27	-0.78	-0.16	-0.92	-1.00
-1.00	0.38	-0.34	0.11	-0.60	0.28	0.89	-0.60
-0.41	0.17	0.21	-1.00	-1.00	-0.24	-0.89	-0.70
-0.65	-0.21	-0.18	0.02	-0.79	-0.08	-0.85	-0.83
0.18	0.16	-1.00	-1.00	-1.00	0.05	-0.95	-0.73
-0.76	0.99	0.15	0.43	0.28	-0.09	-0.93	0.07
-0.06	0.26	0.57	-1.00	-1.00	-1.00	-0.87	0.10
-0.53	0.11	0.51	-1.00	-1.00	0.12	-0.90	-0.70
0.18	0.70	0.21	-1.00	-1.00	0.13	-0.61	-0.57
0.18	0.40	0.31	-1.00	-1.00	-0.19	0.16	0.20

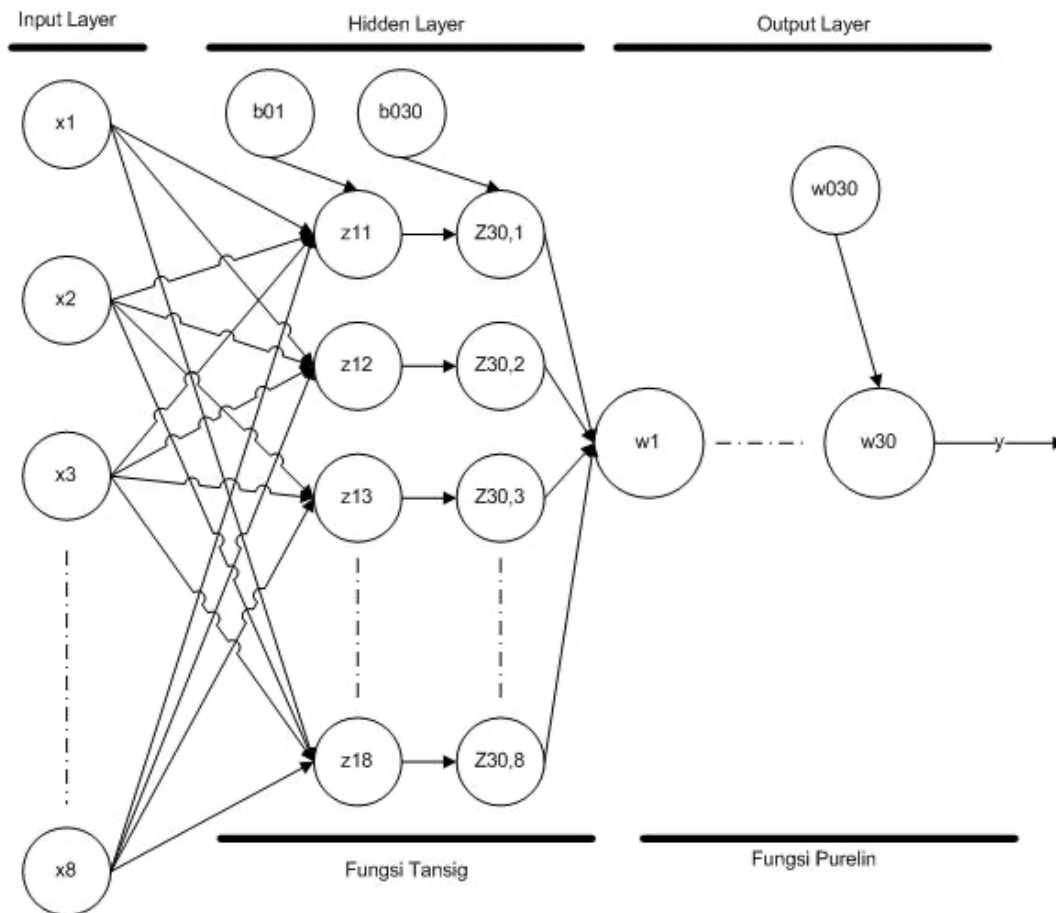
Pembentukan dan Pelatihan Neural Network

Pada tahap ini dijelaskan pembentukan arsitektur dan pelatihan neural network backpropagation. Arsitektur jaringan yang digunakan yaitu 1 input layer, 30 hidden layer 1 output layer.

Langkah 1

Pada tahap ini jaringan dibentuk untuk digunakan sebagai fungsi yang mengevaluasi bobot atau sumber makanan yang dibuat dari ABC. perintah yang digunakan untuk membentuk jaringan backpropagation dengan matlab adalah newff() dengan berbagai parameter tambahan seperti fungsi

aktivasi, fungsi pelatihan jaringan dan lain lain. Pada penelitian ini pembentukan neural network adalah sebagai berikut. Arsitektur jaringan dapat dilihat pada gambar 3.



Gambar 3. Pembentukan NN

keterangan :

net : jaringan backpropagation yang terdiri dari n layer

minmax(pn) : matriks berordo $R \times 2$ yang berisi nilai minimal dan maksimal R buah elemen input

[30 1] : merupakan jumlah unit hidden layer. Berdasarkan arsitektur yang dibuat mempunyai 30 hidden layer input dan 1 layer output.

Tansig : fungsi aktivasi yang dipakai untuk fungsi tangen sigmoid. Fungsi aktivasi tangen sigmoid tepat digunakan untuk kasus yang inputnya bervariasi sehingga tepat digunakan untuk menyelesaikan kasus pada penelitian ini. Selain itu berdasarkan eksperimen yang dilakukan fungsi aktivasi tangen sigmoid dipilih karena mempunyai performa yang lebih baik dibanding fungsi aktivasi logsig.

Y : output dari hasil simulasi

ObjVal : nilai error atau MSE dari hasil simulasi

Tahapan 2

Pada tahap ini pencarian bobot awal dengan ABC dimulai dengan inialisasi parameter-parameter dari ABC seperti jumlah colony, dan maksimal cycle.

```
NP=50;  
FoodNumber=NP/2;  
maxCycle=300;  
  
objfun='nndiabetesFcn';  
D=301;
```

keterangan :

NP : Jumlah koloni lebah

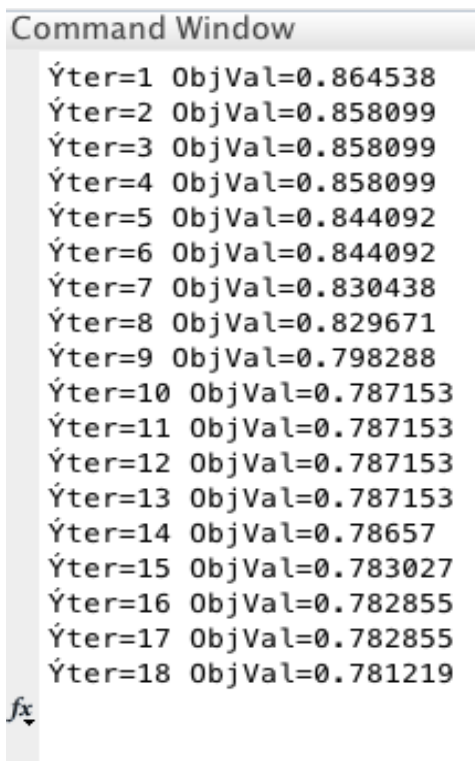
FoodNumber : Jumlah Makanan (Solusi awal yang dibuat)

maxCycle : jumlah iterasi pencarian

objfun : fungsi yang dicari nilai minimalnya (arsitektur neural network pada 4.2.2.1)

D : Dimensi masalah (jumlah parameter neural network yang dicari seperti bobot dan bias)

Pada gambar 4 ditunjukkan proses pencarian bobot dengan ABC. Proses yang ditampilkan berupa urutan iterasi dan nilai ObjVal. ObjVal merupakan nilai MSE paling kecil pada setiap urutan iterasi. Semakin kecil nilai ObjVal maka semakin baik bobot yang diperoleh. Proses pencarian selesai sampai jumlah iterasi sama dengan maxCycle.



Gambar 4. Pencarian bobot dengan ABC

Tahapan 4

Setelah diperoleh jaringan dan bobot akhir, jaringan siap untuk diuji untuk mengetahui apakah training berhasil menghasilkan jaringan yang bagus atau tidak. Pengujian menggunakan perintah sim(). Hasil klasifikasi berdasarkan data input dapat dilihat pada table 4.6. Data yang ditampilkan berupa atribut, kelas data asli dan data hasil simulasi dengan neural network dengan bobot yang diinisialisasi dengan ABC.

Operasi Hidden Layer :

$$\begin{aligned} Z_{in_1} &= v_{01} + v_{11} * x_1 + v_{21} * x_2 + v_{31} * x_3 + \dots + v_{81} * x_8 \\ &= 2,14 + -0,17 * -0,29 + -0,49 * -0,36 + -0,13 * 0,18 \dots + -1,36 * -0,03 \\ &= -10,30 \end{aligned}$$

.....

diteruskan sampai Z_{in_8}

Fungsi aktivasi pada hidden layer (menggunakan fungsi aktivasi tansig)

$$Z_1 = \frac{2}{(1 + \exp(2 * 10,3049)) - 1} = 1,000334$$

.....

diteruskan sampai Z_8

Operasi pada Output Layer :

$$\begin{aligned} y_{in} &= w_0 + w_1 * z_1 + w_2 * z_2 \dots w_8 * z_8 \\ &= 8,73 + 0,32 * 1,000334 + -2,94 * 1,000340 \\ &= 0,837 \end{aligned}$$

Fungsi aktivasi pada layer output (menggunakan fungsi aktivasi purelin)

$$\begin{aligned} y &= x \\ &= 0,837 \end{aligned}$$

dan diteruskan untuk data index selanjutnya sehingga menghasilkan

$$\text{error} = 1 - 0,837 = 0,17$$

Tabel 4. Hasil Klasifikasi

A1	A2	A3	A4	A5	A6	A7	A8	kelas (asli)	output
6	148	72	35	0	33.6	0.627	50	1	0.837023254
1	85	66	29	0	26.6	0.351	31	0	0.154214796
8	183	64	0	0	23.3	0.672	32	1	0.817274864
1	89	66	23	94	28.1	0.167	21	0	0.006509259
0	137	40	35	168	43.1	2.288	33	1	0.999319721
5	116	74	0	0	25.6	0.201	30	0	0.048547965
3	78	50	32	88	31	0.248	26	1	0.703933576
10	115	0	0	0	35.3	0.134	29	0	0.015535233
2	197	70	45	543	30.5	0.158	53	1	1.002410958
8	125	96	0	0	0	0.232	54	1	1.049601594
4	110	92	0	0	37.6	0.191	30	0	0.023335649
10	168	74	0	0	38	0.537	34	1	1.003660018
10	139	80	0	0	27.1	1.441	57	0	-0.000806816
1	189	60	23	846	30.1	0.398	59	1	0.97665175

Pada hasil pelatihan dan pengujian didapatkan hasil performa yang dapat dilihat pada table 4.7. Hasil tersebut menunjukkan bahwa berdasarkan parameter-parameter yang dipilih menghasilkan error (MSE) dari klasifikasi sebesar 0,0611 atau 6,1 persen

Tabel 5. Hasil Pengujian ABC-BPNN

Parameter	Nilai
Learning Rate	0.15
Epoch	5000
Minimum error	0.00001
MSE	0.0611

Hasil

Dari hasil eksperimen yang dilakukan berikut ini adalah hasil perbandingan berupa akurasi BPNN sebelum dan sesudah diterapkan pencarian bobot awal dengan ABC. Dari table 4.6 dapat dilihat ABC-BPNN mempunyai hasil akurasi yang paling besar yaitu 93,9 % jika dibandingkan dengan metode (Tuning BPNN, Fuzzy M-ABC).

Tabel 6. Perbandingan Hasil

Metode	MSE	Akurasi	Studi
Neural Network Based On Tuning Parameter	16	84 %	[9]

Fuzzy, Modified-ABC	17.32	82.68 %	[1]
ABC-BPNN	6.11	93.9 %	Penelitian ini

KESIMPULAN DAN SARAN

Kesimpulan

Dari hasil penelitian dan pengujian klasifikasi diabetes dengan Artificial Bee Colony – Backpropagation Neural Network, pada akhir laporan penulis dapat memberikan kesimpulan sebagai berikut:

1. Didapatkan model klasifikasi penyakit diabetes dengan menggunakan algoritma ABC-BPNN dengan akurasi 93,9 % lebih baik dari penelitian sebelumnya.
2. Dari hasil eksperimen telah diketahui pengaruh penerapan ABC terhadap BPNN. Jika dibandingkan dengan metode (Tuning BPNN, Fuzzy M-ABC) tingkat error pada ABC-BPNN mempunyai nilai yang lebih kecil sebesar 6,11. Dari hasil tersebut peningkatan akurasi dari metode sebelumnya diatas 5 persen.

DAFTAR PUSTAKA

- [1] Fayssal Beloufa , "Design of fuzzy classifier for diabetes disease using Modified Artificial Bee Colony algorithm," *computer methods and programs in biomedicine* 112, 2013.
- [2] Mohapatra P and Chakravarty S, "An improved cuckoo search based extreme learning machine for medical data classification," *Swarm and Evolutionary Computation*, 2015.
- [3] Srinivasa Rao B, "An Approach for Heart Disease Detection by Enhancing Training Phase of Neural Network Using Hybrid Algorithm," *IEEE International Advance Computing Conference (IACC)* , 2014.
- [4] Waheed Ghanem, "Improving accuracy of applications fingerprinting on local networks using NMAP-AMAP-ETTERCAP as a hybrid framework ," *Control System, Computing and Engineering (ICCSCE)* , 2013.
- [5] Celal Ozturk and Dervis Karaboga, "Hybrid Artificial Bee Colony Algorithm for Neural Network Training," 2011.
- [6] Blum C and Socha K, "Training feed-forward neural networks with ant colony optimization: An application to pattern classification," vol. 233-238, 2005.
- [7] Celal Oxturk and Karabota Dervis, "Neural networks training by artificial bee colony algorithm on pattern classification," vol. 19(3), pp. 279–292, 2009.

- [8] Mehdi Khashei, Saeede Eftekhari, and Jamshid Parvizian, "Diagnosing Diabetes Type II Using a Soft Intelligent Binary Classification Model," in *Review of Bioinformatics and Biometrics (RBB)*, vol. Vol 1, 2012.
- [9] Wuttichai Luangruangrong, Annupan Rodtook, and Sanon Chimmanee, "Study of Type 2 Diabetes Risk Factors Using Neural Network For Thai People and Tuning Neural Network Parameters," in *IEEE International Conference on Systems, Man, and Cybernetics*, Seoul, 2012.
- [10] T. Sutojo, Vincent Suhartono, and Edy Mulyanto, *Kecerdasan Buatan*. Semarang: Andi, 2011.
- [11] Diabetes care 28, "Diagnosis and classification of diabetes mellitus," 2005.
- [12] Dervis Karaboga and Celal Ozturk, "Hybrid Artificial Bee Colony Algorithm for Neural Network Training," *Evolutionary Computation*, 2011.
- [13] jok jeng Siang, *Jaringan syaraf tiruan dengan menggunakan matlab*. Yogyakarta: Penerbit andi, 2005.