



DLNN DAN BPNN-GA PADA PREDIKSI PENYAKIT DIABETES DI BOJONEGORO

Nita Cahyani¹, Alif Yuanita Kartini²

Corresponding author : Nita Cahyani

¹Universitas Nahdlatul Ulama Sunan Giri Bojonegoro, nitacahyani@unugiri.ac.id

²Universitas Nahdlatul Ulama Sunan Giri Bojonegoro, yuaniaalif5@gmail.com

Received : 10 April 2022, Revised : 20 November 2022, Accepted : 4 Desember 2022

Abstract

Diabetes occurs when blood sugar levels are higher than normal. Otherwise, the production of insulin is considered insufficient. It is found that in recent days, the proportion of patients with diabetes has increased more than in the world. This issue needs to be taken more seriously to ensure a reduction in the prevalence of typical diabetes. This study aims to detect diabetes early, especially in Bojonegoro, by predicting the disease, in order to reduce the number of people with diabetes. Two methods have been used in this study, namely Deep Learning Neural Network (DLNN) and weight optimization method in training Back-propagation Neural Network (BPNN) with Genetic Algorithms (GA) (BPNNGA) to predict diabetes in Bojonegoro. The results of this research present that using the DLNN and BPNNGA methods produces an accuracy rate of 94%.

Keywords: Diabetes; Deep Learning Neural Network (DLNN); Backpropagation Neural Network (BPNN); Genetic Algorithm (GA).

Abstrak

Diabetes terjadi ketika kadar gula darah lebih tinggi dari biasanya. Di sisi lain, produksi insulin dianggap tidak mencukupi. telah dicatat dalam beberapa hari terakhir bahwa pangsa pasien yang menderita diabetes telah meningkat ke tingkat yang lebih besar di seluruh dunia. Masalah ini harus ditanggapi lebih serius untuk memastikan bahwa persentase tipikal orang yang menderita diabetes berkurang. Penelitian ini bertujuan untuk mendeteksi lebih dini penyakit diabetes khususnya di Bojonegoro dengan cara memprediksi penyakit tersebut, sehingga dapat mengurangi penderita penyakit diabetes. Digunakan 2 metode dalam penelitian ini, yaitu Deep Learning Neural Network (DLNN) dan menggunakan metode optimasi bobot pada pelatihan Backpropagation Neural Network (BPNN) dengan Algoritma Genetika (GA) (BPNNGA) untuk memprediksi diabetes di Bojonegoro. Hasil penelitian ini menyajikan bahwa dengan menggunakan metode DLNN dan BPNNGA menghasilkan tingkat akurasi sebesar 94%.

Kata kunci: Diabetes; Deep Learning Neural Network (DLNN); Backpropagation Neural Network (BPNN); Genetic Algorithm (GA).

1. Pendahuluan

Diabetes Mellitus (DM) dapat merupakan hasil dari berbagai hal yang menyusun serangkaian masalah anatomi dan kimia. DM juga dapat didefinisikan sebagai gangguan fungsi insulin. Penyakit ini menyebabkan komplikasi yang secara perlahan menyerang organ tubuh manusia seperti otak (stroke),

mata (kebutaan), gagal ginjal, bahkan gagal jantung, menjadikannya salah satu penyakit paling mematikan di dunia. DM telah diklasifikasikan menjadi diabetes tipe 1, diabetes tipe 2, diabetes tipe lain, dan diabetes gestasional [1].

Di Indonesia berdasarkan World Health Organization (WHO) diprediksi penyakit

diabetes akan meningkat sebesar 8,4 juta di tahun 2000 dan meningkat menjadi 21,3 juta pada tahun 2030, dan provinsi Jawa Timur termasuk di antara 10 besar prevalensi penderita diabetes se-Indonesia atau menempati urutan ke lima dengan prevalensi 2,6% pada tahun 2018 [2]. Dalam penelitian ini khususnya di Bojonegoro berdasarkan data Rumah Sakit Umum Sosodoro Djatikusuma Bojonegoro penderita penyakit diabetes ditahun 2015 sebanyak 246 orang dan 23 orang penderita lainnya meninggal. Sehingga DM menjadi masalah kesehatan karena terus meningkatnya penyakit ini, mengakibatkan penderita individu dan kerugian ekonomi. Sehingga dalam hal ini sangat diperlukan adanya prediksi guna pendeteksian dini terhadap penyakit yang mematikan ini agar mengurangi penderita penyakit diabetes.

Klasifikasi adalah teknik statistik yang biasa digunakan untuk menentukan akurasi atau prediksi dan menentukan seberapa akurat hasil prediksi dari pengamatan yang sebenarnya. Penelitian prediksi diabetes sebelumnya menggunakan pendekatan jaringan syaraf tiruan (JST) untuk mendeteksi diabetes dan menentukan jenisnya. Metode ini efisien dalam memprediksi tingkat kelangsungan hidup pasien diabetes [3]. Penelitian lain tentang diabetes dengan menggunakan metode data mining [4]–[14] serta peneliti lainnya terkait penggunaan klasifikasi [15]–[18].

Genetic Algorithm (GA) merupakan pendekatan pencarian heuristik yang dapat diterapkan pada berbagai masalah pengoptimalan. Fleksibilitas ini membuatnya menarik untuk banyak masalah pengoptimalan dalam praktiknya. Evolusi adalah dasar dari Algoritma Genetika [19]. GA mencari solusi dari titik populasi daripada dari satu titik. Pencarian dengan sekumpulan titik populasi membuat algoritme genetika cenderung tidak terjebak pada nilai optimal lokal [20], [21].

Backpropagation adalah topologi jaringan syaraf tiruan yang menyesuaikan bobot dan bias jaringan dengan menghitung gradien kesalahan. BPNN diterapkan dengan pengaturan bobot awal acak karena kerusakan simetri [22]. Namun, melatih jaringan saraf dengan bobot awal acak dapat menyebabkan

dua kelemahan utama: terjebak dalam minima lokal dan konvergen secara perlahan. Mengingat keterbatasan jaringan saraf backpropagation ini, terdapat teknik pencarian global yaitu, algoritma genetika yang dapat mengatasi kekurangan BPNN [23]–[26]. Algoritme genetika (GA) yang dikembangkan untuk meniru beberapa proses yang diamati dalam evolusi alami adalah kelas teknik algoritme pencarian global. Mereka telah ditunjukkan dalam praktiknya sangat efektif dalam optimasi fungsi, dan mencari ruang besar atau kompleks

Berdasarkan uraian tersebut, penelitian ini membandingkan metode DLNN dan metode gabungan menggunakan BPNN dan GA guna mendapatkan akurasi klasifikasi atau prediksi diabetes di Bojonegoro dengan tingkat kinerja yang lebih baik.

2. Metode

A. Sumber Data

Data dalam penelitian ini berasal dari Rumah Sakit Umum Daerah Sosodoro Djatikusuma Bojonegoro dari kuesioner yang diberikan kepada pasien diabetes atau yang tidak memiliki diabetes tetapi memiliki gejala.

B. Variabel Penelitian

Pada penelitian ini digunakan 16 variabel input yaitu Umur (X_1), Jenis Kelamin (X_2), sering mengalami buang air kecil dengan volume banyak (X_3), sering merasa haus (sering kali merasa ingin selalu minum (X_4), mengalami penurunan berat badan secara tiba-tiba (X_5), sering merasa badan terasa lemas dan tidak bertenaga (X_6), mengalami peningkatan nafsu makan yang lebih besar dari biasanya (X_7), mengalami gatal/iritasi pada kelamin (X_8), mengalami penglihatan kabur (hilangnya ketajaman penglihatan) (X_9), sering mengalami gatal dikulit (Rasa gatal yang terus-menerus) (X_{10}), sering mengalami perubahan suasana hati dan lekas marah (X_{11}), mengalami lambatnya penyembuhan luka (luka cenderung sembuh lebih lambat) (X_{12}), mengalami kelemahan sebagian anggota gerak atau badan (X_{13}), mengalami nyeri otot, kekakuan sendi, sensasi “kesemutan” pada lengan atau kaki (X_{14}), mengalami kerontokan rambut di kepala (X_{15})

dan obesitas (X_{16}) serta variabel output berupa biner yaitu tidak diabetes (0) dan diabetes (1).

C. Penelitian dan Prosedur

Langkah-langkah dan metode penelitian:

1. Pra-pemrosesan data diabetes termasuk menganalisis nilai yang hilang dan kemudian membagi data menjadi kumpulan data (70:30) untuk pelatihan dan pengujian.
2. Klasifikasi diabetes menggunakan DLNN:
 - a. Untuk membangun model DLNN dan mendapatkan parameter yang optimal dilakukan beberapa percobaan dengan menggunakan 16 neuron input dan 2 lapisan tersembunyi .
 - b. Pada penelitian ini bobot awal yang digunakan mengikuti bawaan dari software R dengan menggunakan package neuralnet.
 - c. Pada lapisan tersembunyi (*hidden layer*) dan keluaran output digunakan aktivasi fungsi sigmoid persamaan 1.

$$f(x) = \frac{1}{1+e^{-x}} \quad (1)$$

Dengan menggunakan algoritma metode backpropagation [27] (Gambar 2 a. Struktur Neural Network) dilakukan proses pembelajaran untuk proses perhitungan bias dan bobot pada data diabetes.

Feedforward:

- Pada *hidden layer* akan menerima sinyal dari setiap masukan X_i , ($X_i, i = 1, 2, \dots, n$).
- ($Z_j, j = 1, 2, \dots, p$) menjumlahkan sinyal-sinyal masukan (*input*) terbobot di setiap *hidden layer*,

$$z_in_j = b1_j + \sum_{i=1}^n x_i v_{ij}, \quad (2)$$
menerapkan fungsi aktivasi untuk menghitung sinyal *output*,

$$z_j = f(z_in_j), \quad (3)$$

Semua lapisan keluaran (*output*) menjadi masukan (*input*) yang telah menerima sinyal.

- $Y_k (k = 1, 2, \dots, m)$ menjumlahkan sinyal-sinyal masukan yang terboboti di setiap keluaran,

$$y_in_k = b2_k + \sum_{i=1}^p z_i w_{jk}, \quad (4)$$

menerapkan fungsi aktivasi untuk menghitung sinyal *output*,

$$y_k = f(y_in_k), \quad (5)$$

langkah 2 (b) dikerjakan sejumlah lapisan tersembunyi (dalam penelitian ini menggunakan 2 lapisan tersembunyi)

Backpropagation:

- $Y_k (k = 1, 2, \dots, m)$

Menghitung informasi *error* setelah setiap keluaran mendapatkan sasaran bentuk yang terdapat hubungan dengan bentuk masukan pelatihan,

$$\delta 2_k = (t_k - y_k) f'(y_in_k), \quad (6)$$
mengupdate *error* dengan cara menghitung informasi pembetulan bobot.

$$\Delta w_{jk} = \alpha \delta_k z_j, \quad (7)$$

Selanjutnya memperbaharui bias

$$b2_k, \quad \Delta b2_k = \alpha \delta_k, \quad (8)$$

Langkah 2 (d) dikerjakan sejumlah 2 lapisan tersembunyi.

- Menjumlahkan delta $Z_j (j = 1, 2, \dots, p)$ masukan pada setiap unit tersembunyi

$$\delta_in_j = \sum_{k=1}^m \delta 2_k w_{jk}, \quad (9)$$

kalikan nilai ini dengan turunan dari fungsi aktivasinya untuk menghitung *error*,

$$\delta 1_j = \delta_in_j f'(z_in_j) \quad (10)$$

Memperbaharui bobot v_{ij} dengan menghitung pembetulan bobot,

$$\Delta v_{ij} = \alpha \delta 1_j x_j \quad (11)$$

Memperbaharui bias $b1_j$ dengan menghitung pembetulan bias,

$$\Delta b1_j = \alpha \delta 1_j \quad (12)$$

- Memperbaharui bias dan bobot ($j = 1, 2, \dots, p$) pada setiap unit keluaran ($Y_k, k = 1, 2, 3, \dots, m$),

$$w_{jk}(new) = w_{jk}(old) + \Delta w_{jk} \quad (13)$$

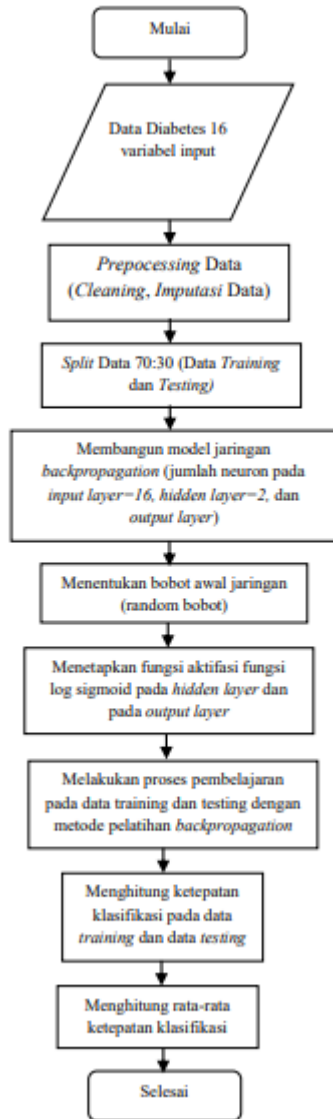
$$b2_k(new) = b2_k(old) + \Delta b2_k \quad (14)$$

memperbaharui bias dan bobot ($i = 0, 1, 2, \dots, n$) pada setiap unit tersembunyi ($Z_j, j = 1, 2, 3, \dots, p$),

$$v_{ij}(new) = v_{ij}(old) + \Delta v_{ij} \quad (15)$$

$$b_{1j}(new) = b_{1j}(old) + \Delta b_{1j} \quad (16)$$

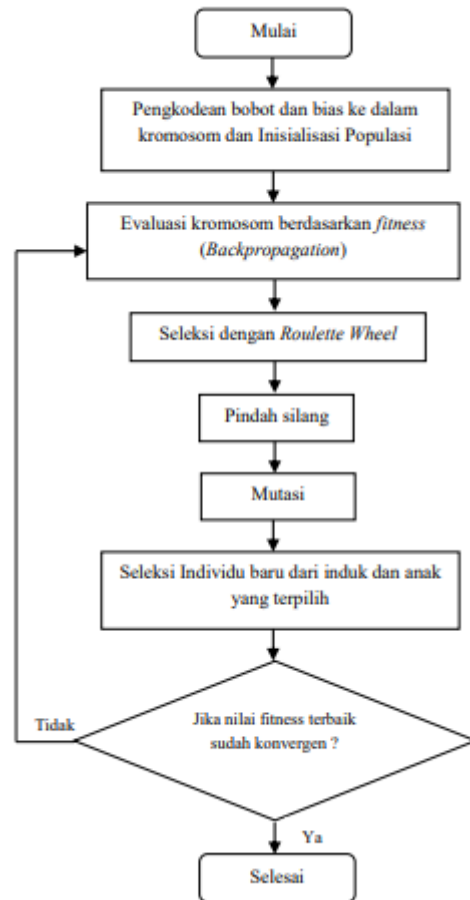
- Uji keadaan berhenti.
- d. Berdasarkan parameter dan arsitektur yang didapatkan dari pelatihan data training juga dilakukan pengujian dengan data testing.
- e. Melakukan perhitungan evaluasi performansi klasifikasi.



Gambar 1 Flow Chart DLNN

3. Algoritma dasar GA [28]–[30].
 - I. Optimasi bobot GA pada BPNN
 - a. Mulai.
 - b. Menginisialisasi populasi secara acak.
 - c. Menentukan fungsi fitness dari masalah.
 - d. Menentukan fitness dari populasi
 - e. **Ketika** konvergen atau optimal tidak tercapai, lakukan! :

- f. Memilihan orang tua dari populasi.
 - g. Pemindahan Silang (crossover) untuk generasi populasi baru.
 - h. Melakukan mutasi pada populasi baru
 - i. Menghitung fitness populasi baru
 - j. Selesai
 - k. Jika optimal tercapai, akan menampilkan hasil akhir
 - l. Berhenti
- II. Menerapkan hasil bobot yang telah di optimasi oleh GA ke algoritma BPNN.



Gambar 2 Diagram Alir Optimasi Bobot Backpropagation dengan GA (BPNNNGN)

4. Memilih metode yang terbaik dari hasil evaluasi masing-masing metode dengan evaluasi performansi pada persamaan berikut. [31], [32] :

$$AccuracyRate = \frac{TP+TN}{TP+TN+FN+FP} \quad (1)$$

$$ErrorRate = \frac{FN+FP}{TP+TN+FN+FP} \quad (2)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

$$AUC = \frac{1+Sensitivity-(1-Specificity)}{2} \quad (5)$$

3. Pembahasan

Bagian pembahasan diuraikan proporsi data diabetes dan hasil kinerja klasifikasi pada metode DLNN dan BPNNGA (optimasi bobot)

Berdasarkan Tabel 1 diketahui jumlah penderita penyakit diabetes sebanyak 35 pasien, sedangkan jumlah pasien tidak diabetes sebanyak 72 pasien. Tabel 1 disajikan proporsi data diabetes.

Tabel 1. Tabel Proporsi Status Prediksi Diabetes

Negatif Diabetes	Positif Diabetes
67,29%	32,71%

Proporsi Negatif Diabetes: $\left(\frac{72}{35+72} \times 100 = 67.29\%\right)$ dan Proporsi Positif Diabetes: $\left(\frac{35}{35+72} \times 100 = 32.71\%\right)$. Tabel 1 diketahui bahwa proporsi positif diabetes sebesar 32,71% sedangkan proporsi negatif diabetes sebesar 67,29%.

Pada metode DLNN digunakan beberapa percobaan untuk mendapatkan parameter yang terbaik dalam menghasilkan perpoma klasifikasi yang tinggi. Dengan menggunakan software R dengan learning rate 0,01 dan threshold 0,01 didapatkan hasil struktur NN yang optimal 16-4-24-1. 16 merupakan jumlah data input yang terdapat pada penelitian ini, 4 adalah jumlah neuron pada lapisan tersembunyi pertama, 24 adalah jumlah neuron pada lapisan tersembunyi kedua dan 1 output. Tabel 2 menyajikan hasil klasifikasi DLNN.

Tabel 2. Kinerja Hasil Klasifikasi DLNN

Evaluasi Klasifikasi	Data Testing (%)	Data Training (%)
AccuracyRate	0,83	0,94
Sensitivity	0,75	0,83
Specificity	0,88	1,00
AUC	0,81	0,99

Nilai AccuracyRate data testing pada table 2 sebesar 0,94 artinya prosentase classifier benar melakukan prediksi sebesar 94 %.

Sensitivity 0,75 artinya persentase data diabetes yang diprediksi sebagai diabetes sebesar 75 %. Specificity sebesar 0,88 artinya prosentase data tidak diabetes diprediksi sebagai tidak diabetes sebesar 88 %.

Tabel 3 Menyajikan hasil evaluasi kinerja klasifikasi BPNN sebelum optimasi bobot dan sesudah optimasi bobot (BPNNGA) pada data diabetes. Hasil evaluasi didapatkan setelah pembentukan model dengan mendapatkan parameter NN yang paling optimal.

Tabel 3. Kinerja Hasil BPNN Sebelum dan Sesudah Optimasi Bobot dengan GA

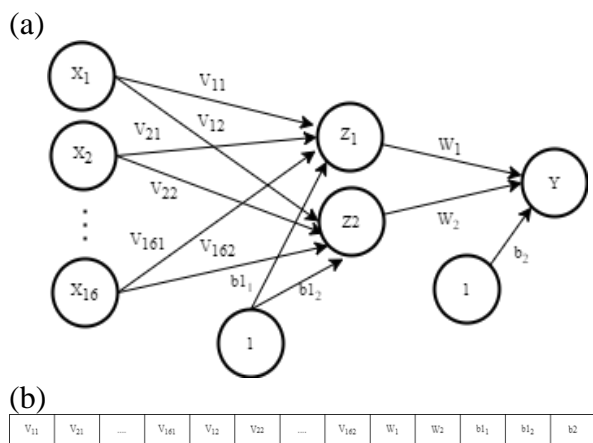
Metode	Kinerja Klasifikasi			
	Akurasi	AUC	sensitivity	specificity
BPNN Training (Optimasi GA)	0,94	0,99	0,83	1
BPNN Testing (Optimasi GA)	0,81	0,72	0,50	0,96
BPNN Training	0,93	0,97	0,83	0,98
BPNN Testing	0,80	0,75	0,67	0,87

Berdasarkan Tabel 3 diketahui BPNN dengan optimasi GA untuk data training diketahui nilai Akurasi, AUC, Sensitivity dan Specificity secara berurutan sebesar 0,94, 0,99, 0,826, dan 1 sedangkan BPNN untuk data training secara berurutan sebesar 0,93, 0,51, 0,95, dan 0,92. Hasil kinerja data testing untuk metode BPNN dengan optimasi GA secara berurutan untuk nilai akurasi, AUC, Sensitivity dan Specificity yaitu 0,81, 0,53, 0,86, dan 0,79. Akurasi sebesar 0,81 artinya 81 persen classifier benar melakukan prediksi. Nilai Sensitivity sebesar 0,86 artinya 86 persen data positif yakni penderita diabetes diprediksi sebagai penderita diabetes. Nilai Specificity sebesar 0,79 artinya sebesar 79 persen data negatif yaitu tidak diabetes diprediksi sebagai tidak diabetes. Nilai AUC memberikan informasi kinerja classifier. Sedangkan untuk metode BPPN untuk data testing secara berurutan untuk nilai akurasi, AUC, Sensitivity dan Specificity sebesar 0,81, 0,44, 0,73, dan 0,84.

Berikut tahapan-tahapan untuk mendapatkan parameter bobot pelatihan BPNN yang paling optimal dalam meningkatkan nilai akurasi yang lebih baik, yaitu dilakukan optimasi bobot pada pelatihan BPNN dengan GA.

Tahap 1. Langkah pertama dalam masalah GA adalah inialisasi populasi [30] yaitu banyaknya kromosom populasi yang akan digunakan sebanyak 50, batas iterasi sebanyak 100, dalam pindah silang digunakan sebesar 0,8 peluang, dan dalam mutasi digunakan sebanyak 0,1 peluang.

Tahap 2. Mengeset tahap awal nilai dalam kromosom (inialisasi kromosom) dengan memakai bilangan real (*real valued*) [30] yang digunakan dalam proses pengkodean dengan mengoptimasi parameter bobot ANN. Proses bias dan bobot pada ANN digambarkan kedalam bentuk kromosom GA contoh pembentukannya pada Gambar 1. Dengan nilai parameter bervariasi, dari sejumlah populasi di generet dengan membentuk 50 kromosom.



Gambar 2. (a) Bentuk *Neural Network* dengan 16 variabel masukan (*input*) dan (b) Penyusunan Kromosom *Genetic Algorithm*.

Tahap 3. Fungsi *fitness function* yang digunakan dalam penelitian ini adalah nilai dari akurasi. Fungsi *Fitness* juga disebut sebagai fungsi evaluasi mengevaluasi seberapa dekat dengan solusi optimal atau solusi keinginan yang telah dicapai oleh GA [30]. Menurut aturan seleksi alam, kromosom yang dianggap paling cocok

dikirim ke tingkat berikutnya dan sisanya dibunuh. Kromosom yang diwariskan ke generasi berikutnya bertindak sebagai kromosom induk yang selanjutnya akan digunakan untuk reproduksi kromosom anak baru dengan dua metode yaitu *crossover* dan *mutasi*.

Tahap 4. Menyeleksi menggunakan *roulette wheel*. Pemilihan dalam metode ini didasarkan pada proporsionalitas kebugaran dari suatu kromosom individu. Semakin bugar kromosom individu, semakin tinggi peluang kromosom terpilih sebagai induk (orang tua).

Tahap 5. *Crossover* atau proses pindah silang. Diberikan bilangan uniform 0,1. Jika nilai uniform lebih kecil dari peluang pindah silang 0,8 maka kromosom terseleksi atau dipilih menjadi calon orang tua. Proses *crossover* atau pindah silang dapat terjadi jika isi dari parameter kromosom awal berpindah dengan kromosom kedua sehingga menghasilkan kromosom baru.

Tahap 6. Tahap mutasi, dengan menggunakan mutasi, GA bisa mendapatkan keragaman populasi yang lebih baik untuk mendapatkan set solusi yang lebih baik mutasi yang digunakan dalam penelitian ini adalah uniform. Operator mutasi ini memasukkan batas atas dan bawah dari pengguna dan mengganti nilai gen yang dipilih dengan nilai acak uniform yang terletak di antara kedua batas tersebut. Ini juga terbatas pada populasi gen tipe integer dan float.

Tahap 7. Tahapan akhir melakukan iterasi sampai ke-100 dan dengan nilai *fitness* tertinggi menghasilkan populasi baru dari 50 kromosom. Populasi terbaru didapatkan dengan mengurutkan dari semua nilai *fitness* dari semua kromosom yang terbentuk; diawali dari 50 kromosom, kromosom pindah silang (*crossover*), kromosom mutasi.

Berdasarkan analisis hasil optimasi bobot BPNN dapat dikatakan bahwa dengan menggunakan GA diperoleh hasil kinerja klasifikasi yang lebih baik atau nilai kinerja klasifikasi mengalami peningkatan, hal ini dapat dibuktikan dari hasil nilai AUC sebesar 0,97 setelah menggunakan GA pada optimasi

bobot BPNN nilai AUC menjadi 0,99, begitupun hasil kinerja klasifikasi yang lain, nilai specificity dari 0,98 menjadi 1. Dalam penelitian ini metode optimasi GA mampu meningkatkan tingkat kinerja klasifikasi pada BPNN. Jika dibandingkan hasil kinerja klasifikasi antara dua metode DLNN dan BPNN pada tabel 2 dan table 3 dapat disimpulkan untuk data training menghasilkan nilai kinerja klasifikasi yang sama dengan Akurasi 0,94, AUC 0,99, sensitivity 0,83 dan specificity 1.

4. Penutup

Hasil penelitian ini memaparkan tentang metode Deep learning Neural Network (DLNN) dan BPNN dengan Algoritma Genetika (BPNNGA) pada prediksi penyakit diabetes di Bojonegoro pada periode agustus sampai dengan september 2021. Hasil penelitian menyimpulkan bahwa pelatihan data training untuk metode DLLN dan BPNNGA menghasilkan akurasi klasifikasi yang sama yaitu sebesar 94% atau sebanyak 19 pasien kelas diabetes diprediksi diabetes dan 48 pasien kelas tidak diabetes diprediksi tidak diabetes serta jumlah pasien tidak diabetes yang diprediksi sebagai diabetes adalah 4. Artinya ada 4 kesalahan prediksi pada data training, untuk data tesing metode DLLN lebih unggul dibandingkan metode BPNNGA yaitu akurasi sebesar 83% atau sebanyak 21 pasien kelas diabetes diprediksi sebagai diabetes dan sebanyak 9 pasien kelas tidak diabetes diprediksi tidak diabetes dan jumlah kelas pasien tidak diabetes diprediksi diabetes sebanyak 3 pasien serta jumlah kelas diabetes diprediksi tidak diabetes sebanyak 3 pasien. Sehingga kedua metode sudah cukup baik, studi ini masih mempunyai perkembangan untuk penelitian tingkat lanjut dengan menambahkan variabel dan jumlah data yang telah digunakan.

Ucapan Terimakasih

Kami ucapkan terimakasih kepada civitas akademika Universitas Nahdlatul Ulama Sunan Giri, dan bantuan dana penelitian Ristekdikti Nomor: 29/LPPM/Unugiri-Bjn/KP.5/VII/2021.

Referensi

- [1] E. Decroli, *Diabetes Melitus Tipe 2*. Padang: Pusat Penerbitan Bagian Ilmu Penyakit Dalam, 2019.
- [2] Kementerian Kesehatan RI., “Infodatin tetap produktif, cegah, dan atasi Diabetes Melitus 2020,” *Pusat Data dan Informasi Kementerian Kesehatan RI*. pp. 1–10, 2020.
- [3] N. Pradhan, G. Rani, V. Dhaka, and R. Poonia, “Diabetes prediction using artificial neural network,” 2020, pp. 327–339. doi: 10.1016/B978-0-12-819061-6.00014-8.
- [4] G. Alfian *et al.*, “Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features,” *Biocybernetics and Biomedical Engineering*, vol. 40, no. 4, pp. 1586–1599, 2020, doi: <https://doi.org/10.1016/j.bbe.2020.10.004>.
- [5] M. Yousef, M. Ferwana, S. Sakr, and R. Shammari, “Early prediction of diabetes by applying data mining techniques,” *Computer Methods and Programs in Biomedicine*, vol. 171, p. 3, Apr. 2019, doi: 10.1016/j.cmpb.2018.12.008.
- [6] M. Raja and D. Pandian, “PSO-FCM based data mining model to predict diabetic disease,” *Computer Methods and Programs in Biomedicine*, vol. 196, p. 105659, Jul. 2020, doi: 10.1016/j.cmpb.2020.105659.
- [7] Md. M. Islam, Md. J. Rahman, D. Chandra Roy, and Md. Maniruzzaman, “Automated detection and classification of diabetes disease based on Bangladesh demographic and health survey data, 2011 using machine learning approach,” *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, no. 3, pp. 217–219, 2020, doi: <https://doi.org/10.1016/j.dsx.2020.03.004>.
- [8] A. Oleiwi, L. Shi, Y. Tao, and L. Wei, “A Comparative Analysis and Risk Prediction of Diabetes at Early Stage using Machine Learning Approach,” *International Journal of Future*

Generation Communication and Networking, pp. 4151–4163, Aug. 2020.

- [9] H. Yang *et al.*, “Risk Prediction of Diabetes: Big data mining with fusion of multifarious physical examination indicators,” *Information Fusion*, vol. 75, pp. 140–149, 2021, doi: <https://doi.org/10.1016/j.inffus.2021.02.015>.
- [10] F. Liang *et al.*, “Insulin-resistance and depression cohort data mining to identify nutraceutical related DNA methylation biomarker for type 2 diabetes,” *Genes and Diseases*, vol. 8, no. 5, pp. 669–676, 2021, doi: [10.1016/j.gendis.2020.01.013](https://doi.org/10.1016/j.gendis.2020.01.013).
- [11] S. Perveen, M. Shahbaz, A. Guergachi, and K. Keshavjee, “Performance Analysis of Data Mining Classification Techniques to Predict Diabetes,” *Procedia Computer Science*, vol. 82, no. March, pp. 115–121, 2016, doi: [10.1016/j.procs.2016.04.016](https://doi.org/10.1016/j.procs.2016.04.016).
- [12] H. Thakkar, V. Shah, H. Yagnik, and M. Shah, “Comparative anatomization of data mining and fuzzy logic techniques used in diabetes prognosis,” *Clinical eHealth*, vol. 4, no. 2021, pp. 12–23, 2021, doi: [10.1016/j.ceh.2020.11.001](https://doi.org/10.1016/j.ceh.2020.11.001).
- [13] C. Fiarni, E. M. Sipayung, and S. Maemunah, “Analysis and prediction of diabetes complication disease using data mining algorithm,” *Procedia Computer Science*, vol. 161, pp. 449–457, 2019, doi: [10.1016/j.procs.2019.11.144](https://doi.org/10.1016/j.procs.2019.11.144).
- [14] B. Davazdahemami and D. Delen, “The confounding role of common diabetes medications in developing acute renal failure: A data mining approach with emphasis on drug-drug interactions,” *Expert Systems with Applications*, vol. 123, pp. 168–177, 2019, doi: <https://doi.org/10.1016/j.eswa.2019.01.006>.
- [15] N. Cahyani, K. Fithriasari, I. Irhamah, and N. Iriawan, “On the Comparison of Deep Learning Neural Network and Binary Logistic Regression for Classifying the Acceptance Status of Bidikmisi Scholarship Applicants in East Java,” *Matematika*, vol. 34, no. 3, pp. 83–90, 2018, doi: [10.11113/matematika.v34.n3.1141](https://doi.org/10.11113/matematika.v34.n3.1141).
- [16] N. Cahyani, S. S. Pangastuti, K. Fithriasari, I. Irhamah, and N. Iriawan, “Classification of Bidikmisi Scholarship Acceptance using Neural Network Based on Hybrid Method of Genetic Algorithm,” *Indonesian Journal of Statistics and Its Applications*, vol. 5, no. 2 SE-Articles, pp. 396–404, Jun. 2021, doi: [10.29244/ijasa.v5i2p396-404](https://doi.org/10.29244/ijasa.v5i2p396-404).
- [17] N. Iriawan *et al.*, “On The Comparison: Random Forest, SMOTEBagging, and Bernoulli Mixture to Classify Bidikmisi Dataset in East Java,” *2018 International Conference on Computer Engineering, Network and Intelligent Multimedia, CENIM 2018 - Proceeding*, pp. 137–141, 2018, doi: [10.1109/CENIM.2018.8711035](https://doi.org/10.1109/CENIM.2018.8711035).
- [18] W. Wibowo, S. Abdul-Rahman, and N. Cahyani, “Multilevel Logistic Regression and Neural Network-Genetic Algorithm for Modeling Internet Access BT - Soft Computing in Data Science,” 2019, pp. 169–180.
- [19] O. Kramer, *Studies in Computational Intelligence 679 Genetic Algorithm Essentials*. 2017.
- [20] Y. Ge, “Genetic Algorithms and Engineering Design (Book Review),” *International Journal of Human-Computer Interaction*, vol. 9, no. 4, pp. 457–458, 1997, doi: [10.1207/s15327590ijhc0904_9](https://doi.org/10.1207/s15327590ijhc0904_9).
- [21] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed. USA: Addison-Wesley Longman Publishing Co., Inc., 1989.
- [22] Y. T. Chang, J. Lin, J. S. Shieh, and M. F. Abbod, “Optimization the initial weights of artificial neural networks via genetic algorithm applied to hip bone fracture prediction,” *Advances in Fuzzy Systems*, vol. 2012, 2012, doi: [10.1155/2012/951247](https://doi.org/10.1155/2012/951247).
- [23] A. J. Al-Shareef and M. F. Abbod, “Neural Networks Initial Weights

- Optimisation,” in *2010 12th International Conference on Computer Modelling and Simulation*, 2010, pp. 57–61. doi: 10.1109/UKSIM.2010.19.
- [24] D. Venkatesan, K. Kannan, and R. Saravanan, “A genetic algorithm-based artificial neural network model for the optimization of machining processes,” *Neural Computing and Applications*, vol. 18, no. 2, pp. 135–140, 2009, doi: 10.1007/s00521-007-0166-y.
- [25] A. Karegowda, A. S. Manjunath, and J. M.A, “Application of Genetic Algorithm Optimized Neural Network Connection Weights for Medical Diagnosis of PIMA Indians Diabetes,” *International Journal on Soft Computing (IJSC)*, vol. 2, Apr. 2011, doi: 10.5121/ijsc.2011.2202.
- [26] O. Mohammed, M. Nordin, S. Sulaiman, and W. Fatimah, “Study of Genetic Algorithm to Fully-automate the Design and Training of Artificial Neural Network,” vol. 9, Jan. 2009.
- [27] S. Kusumadewi, *Membangun Jaringan Syaraf Tiruan (menggunakan MATLAB Dan Excel Link)*. Yogyakarta: Graha Ilmu, 2004.
- [28] V. Trevino and F. Falciani, “GALGO: an R package for multivariate variable selection using genetic algorithms,” *Bioinformatics (Oxford, England)*, vol. 22, no. 9, pp. 1154–1156, 2006, doi: 10.1093/bioinformatics/btl074.
- [29] Z. Zukhri, *Algoritma Genetika: Metode Komputasi Evolusioner untuk Menyelesaikan Masalah Optimasi*. 2014.
- [30] S. D. Imanuel and U. K. Chakraborty, “Genetic Algorithm: An Approach on Optimization,” *Proceedings of the 4th International Conference on Communication and Electronics Systems, ICCES 2019*, no. July, pp. 701–708, 2019, doi: 10.1109/ICCES45898.2019.9002372.
- [31] D. Chico, N. Tötsch, and G. Jurman, “The matthews correlation coefficient (Mcc) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation,” *BioData Mining*, vol. 14, pp. 1–22, 2021, doi: 10.1186/s13040-021-00244-z.
- [32] N. P. Tigga and S. Garg, “Prediction of Type 2 Diabetes using Machine Learning Classification Methods,” *Procedia Computer Science*, vol. 167, no. 2019, pp. 706–716, 2020, doi: 10.1016/j.procs.2020.03.336.