

Deteksi Dini Penyakit Diabetes Menggunakan Algoritma *Neural Network* Berbasis Algoritma Genetika

Primandani Arsi^{1*)} Oman Somantri²

¹Jurusan Teknik Informatika STMIK AMIKOM Purwokerto

²Jurusan Teknik Informatika Politeknik Harapan Bersama Tegal

¹Jl. Letjen Pol Soemarto Watumas, Purwanegara – Purwokerto, Indonesia

²Jl. Mataram No.9 Pesurungan Lor, Kota Tegal, Indonesia

email: ¹ukhti.prima@amikompurwokerto.ac.id, ²oman_mantri@yahoo.com

Received: 28 Juli 2018; Revised: 1 Oktober 2018; Accepted: 20 Oktober 2018

Copyright ©2018, Politeknik Harapan Bersama, Tegal

Abstract – Diabetes or Diabetes Mellitus (DM) is a disease in which the glucose content in the blood can't be processed by the body. Based on the case of many diabetic patients, first action is needed as a solution of the problem of diabetes by predicting to detect early diabetes. This is necessary because often medical decisions are made based on experience and rational reasoning. Prediction of diabetes can be done by using some data of diabetic patients who have been stored in the database to make a pattern for the determination of diabetes with Artificial Intelligen technique so that the result of inaccurate diagnosis can be avoided. In this study the authors apply Genetic Algorithm (GA) to optimize Neural Network (NN) model by searching the best parameter value on Neural Network (NN) model. The experimental results showed a decrease in RMSE value which means an increase in predicted accuracy value, ie from 0.402 +/- 0,035 to 0.396 +/- 0,032. The optimization of the NN model prediction, the relevant policy makers can determine the prediction of diabetes more accurately.

Abstrak – Diabetes atau istilah lainnya Diabetes Melitus (DM) adalah sebuah penyakit dimana kandungan gula dalam darah tidak dapat diolah oleh tubuh. Berdasarkan kasus banyaknya penderita penyakit diabetes, maka perlu tindakan awal sebagai solusi dari permasalahan penyakit diabetes yaitu dengan melakukan prediksi guna mendeteksi dini penyakit diabetes. Hal ini perlu dilakukan karena sering kali keputusan medis yang dibuat berdasarkan pengalaman dan penalaran rasional. Prediksi penyakit diabetes dapat dilakukan dengan cara memanfaatkan beberapa data pasien penderita penyakit diabetes yang telah tersimpan dalam basis data guna dibuat suatu pola untuk penentuan penyakit diabetes dengan teknik *Artificial Intelligence* sehingga hasil diagnosa yang tidak tepat dapat dihindari. Dalam penelitian ini peneliti menerapkan Algoritma Genetika (GA) untuk mengoptimalkan model *Neural Network* (NN) dengan cara pencarian nilai parameter terbaik pada model *Neural Network* (NN). Hasil eksperimen menunjukkan terjadinya penurunan nilai RMSE yang berarti terjadi peningkatan pada nilai akurasi prediksi,

yaitu dari 0,402 +/- 0,035 menjadi 0,396 +/- 0,032. Dengan adanya optimasi prediksi model NN, maka para pengambil kebijakan terkait dapat menentukan prediksi penyakit diabetes dengan lebih akurat.

Kata Kunci – Diabetes, *Artificial Intelligen*, Algoritma Genetika (GA), *Neural Network*, RMSE.

I. PENDAHULUAN

Diabetes atau istilah lainnya Diabetes Melitus (DM) adalah sebuah penyakit dimana kandungan gula dalam darah tidak dapat diolah oleh tubuh [1]. Kondisi normal tubuh seseorang secara otomatis akan menghasilkan insulin yang berfungsi untuk memastikan kadar gula tetap berada dalam ambang batas yang normal. Ketika tubuh menghasilkan hormon insulin yang tidak cukup atau bahkan sel tidak merespon insulin yang dihasilkan, maka gula darah menjadi tinggi sehingga penyakit DM terjadi. Adapun gejala klasik yang terjadi secara umum adalah hasrat untuk buang air kecil yang terlalu sering, rasa haus yang tidak berhenti, dan tingginya ritme rasa lapar [2].

Dalam kurun waktu lebih kurang dari 30 tahun terakhir jumlah penderita penyakit diabetes di Negara Cina mengalami peningkatan yang signifikan, hampir 110 juta jiwa pada tahun 2017. Data ini menunjukkan bahwa Cina memiliki populasi pasien dengan penyakit diabetes terbesar di dunia. Penderita diabetes dengan jenis kelamin pria cenderung lebih tinggi dibandingkan dengan wanita. Hal ini menyebabkan perubahan pola dan gaya hidup warga sekitar yang ingin menjadi lebih sehat [3].

Data terbaru dari Federasi Diabetes Internasional (IDF) tentang diabetes di *Atlas Diabetes* mengemukakan bahwa pada tahun 2015 pasien diabetes di dunia saat ini berjumlah kurang dari 415 juta. Populasi pasien penderita penyakit diabetes akan semakin meningkat hingga 642 juta jiwa, yang berarti satu dari sepuluh orang dewasa [4].

Berdasarkan kasus banyaknya penderita penyakit diabetes, maka perlu tindakan awal sebagai solusi dari permasalahan penyakit diabetes yaitu dengan melakukan prediksi guna mendeteksi dini penyakit diabetes. Hal ini perlu dilakukan karena sering kali keputusan medis yang dibuat berdasarkan pengalaman dan penalaran rasional [5]. Oleh karena itu petugas medis dalam hal ini dokter membutuhkan sebuah

*) **Corresponding author:** (Primandani Arsi)
Email: ukhti.prima@amikompurwokerto.ac.id

model untuk memprediksi penyakit diabetes yang handal [6]. Adapun rekomendasi prediksi dari penyakit diabetes ini dapat dibuat berdasarkan data catatan sejarah medis pasien penderita diabetes [7]. Prediksi penyakit diabetes dapat dilakukan dengan cara memanfaatkan beberapa data pasien penderita penyakit diabetes yang telah tersimpan dalam basis data guna dibuat suatu pola untuk penentuan penyakit diabetes dengan teknik *Artificial Intelligence* sehingga hasil diagnosa yang tidak tepat dapat dihindari [5]. Sebuah kontribusi yang baik dalam pengambilan keputusan dan perencanaan dimasa yang akan datang akan dihasilkan dari prediksi yang akurat [8].

Berbagai penelitian sebelumnya yang berkaitan dengan prediksi sebuah penyakit telah banyak dilakukan, termasuk penggunaan algoritma *Neural Network* [9]. Algoritma *Neural Network* adalah sebuah algoritma dengan model terlatih yang memiliki kemampuan sangat baik dalam mengatasi masalah *nonlinier*, melalui pengenalan pola algoritma ini mampu melakukan pemodelan. Disamping keunggulan yang dimiliki, *neural network* mempunyai kelemahan yaitu permasalahan dalam penentuan parameter sehingga perlu dilakukan eksperimen dalam menentukan tiap parameternya [10]. Untuk efisiensi dalam pencarian parameter *neural network* tersebut diperlukan sebuah algoritma optimasi [11]. Penulis mengusulkan salah satu algoritma optimasi yaitu algoritma genetika (GA) untuk mengoptimalkan kinerja *neural network*. Algoritma genetika (GA) diterapkan pada model *Neural Network (NN)* untuk menentukan nilai parameter kontrol agar lebih optimal sehingga model yang dihasilkan menjadi handal [12]. Prinsip kerja algoritma genetika (GA) adalah dengan menghasilkan keturunan lebih baik melalui proses seleksi [11].

II. PENELITIAN YANG TERKAIT

Model prediksi berbagai penyakit telah dilakukan oleh peneliti sebelumnya yakni, penelitian yang dilakukan oleh K. Rajesh, V. Sangeetha (2012) [6], menggunakan data Pima Indian Diabetes Database untuk membandingkan hasil dari beberapa metode diantaranya CRT, CS-RT, C4.5, K-NN, LDA, NAIVE BAYES, PLS-DA, SVM, RND TREE. Diantara 10 metode yang disusulkan algoritma RND TREE merupakan algoritma yang memiliki akurasi paling tinggi yaitu 100%. Namun algoritma tersebut memiliki kelemahan yaitu memiliki *rule data* yang banyak sehingga penulis memilih C4.5 untuk penelitiannya dimana algoritma C4.5 dinilai lebih efisien. Tingkat akurasi yang dihasilkan mencapai 91%.

A, Andriani (2013) dalam penelitiannya mengklasifikasikan data dari pasien penyakit diabetes [7]. *Confusion matrix* dan kurva ROC digunakan dalam evaluasi hasil klasifikasi tersebut. Dari hasil evaluasi didapatkan hasil *Excellent Clasification*. *Rule* hasil klasifikasi kemudian diimplementasikan pada sebuah system.

Fatmawati (2016) menerapkan teknik klasifikasi *data mining* dengan metode *Naive Bayes* dan *Decision Tree* [12]. Evaluasi dilakukan menggunakan *Confusion Matrix* dan kurva ROC. Dari eksperimen yang telah dilakukan *Naive Bayes* memiliki tingkat akurasi yang lebih baik dari pada *Decision Tree* dalam memprediksi penyakit diabetes.

Sementara itu Meng X, Huang., dkk [13] membandingkan kinerja regresi logistik, jaringan saraf tiruan (JST) dan model

pohon keputusan untuk memprediksi diabetes atau prediabetes menggunakan faktor risiko umum. Menggunakan *dataset* pasien diabetes dari Guangzhou, Cina dengan 735 pasien terdeteksi pradiabet dan 752 normal. Hasil perbandingan tiga model menunjukkan bahwa Model pohon keputusan (C5.0) memiliki akurasi klasifikasi terbaik.

Ratnaningtyas WP, dkk (2013) dalam penelitiannya menggunakan JST dalam mengklasifikasikan data penyakit TBC [9]. Gejala-gejala penyakit TBC yang digunakan sebagai input untuk mendiagnosis penyakit tersebut terdiri atas 6 variabel dengan kode target 1 untuk suspek TBC dan 0 untuk negatif TBC. Berdasarkan hasil penelitian menggunakan JST dengan metode *backpropagation* menunjukkan bahwa dengan adanya variasi jumlah *neuron*, *learning rate* dan *hidden layer* dalam arsitektur JST akan diperoleh arsitektur jaringan yang optimal, yakni 100%.

Ricardus A.P, dkk (2013), dalam penelitiannya menggunakan BP-NN berbasis *Adaboost* untuk memprediksi penyakit jantung [5]. Eksperimen menggunakan BP-NN yang dilakukan memperoleh akurasi sebesar 96,65% sedangkan hasil eksperimen yang dilakukan dengan penambahan algoritma *Adaboost* mengalami peningkatan akurasi yaitu 99,29%. Sehingga dapat disimpulkan bahwa metode *adaboost* efektif meningkatkan BP-NN.

Optimasi algoritma juga dilakukan oleh Abdul Rohman dkk, dalam penelitiannya performa algoritma c4.5 ditingkatkan dengan *Adaboost* guna prediksi penyakit jantung. Hasil penelitian menunjukkan bahwa *adaboost* memberikan pemecahan yang lebih akurat pada prediksi penyakit jantung [14].

III. ALGORITMA NEURAL NETWORK & ALGORITMA GENETIKA (GA)

A. Algoritma Neural Network (NN)

Algoritma *Neural Network (NN)* adalah sebuah algoritma yang cara kerjanya meniru fungsi organ otak manusia. Organ otak manusia ini memiliki unit pengolahan yang berjumlah jutaan, masing-masing unit ini disebut *neuron*. Unit kecil atau *neuron* ini bekerja dengan cara saling terhubung antara *neuron* satu dengan lainnya, sedangkan setiap *neuron* memperoleh *input* dari satu set *neuron* yang lainnya. Setelah mengambil *input* dari *neuron* yang lain, *input* diproses dan melewati *output* selanjutnya untuk diproses lebih lanjut [15].

Neural network (NN) memiliki beberapa algoritma yang sering digunakan salah satunya *Backpropagation*. *Backpropagation* pertama kali ditemukan oleh Paul W pada tahun 1974, setelah itu kembali dikembangkan oleh Rumelhart dan McClelland kemudian diterapkan pada algoritma *Neural Network (NN)*. Awalnya *backpropagation* dirancang untuk *Neural Network (NN) feedforward*, namun dalam perkembangannya metode tersebut diadaptasi guna proses pembelajaran algoritma *Neural Network (NN)* [16].

Algoritma *backpropagation* termasuk dalam algoritma jenis terawasi dimana cara kerja algoritma ini dengan cara meminimalisir tingkat *error* pada *output*-nya. Konsepnya adalah jika *output* dari model jaringan yang dihasilkan bernilai salah maka bobot (*weight*) dalam jaringan akan dikoreksi sehingga galat yang dihasilkan lebih kecil, dengan demikian respon jaringan yang dihasilkan akan mendekati nilai yang benar. Adapun perbaikan bobot yang dilakukan

oleh *backpropagation* adalah *hidden layer* (lapisan tersembunyi) [17].

B. Algoritma Genetika (GA)

Algoritma genetika (GA) adalah sebuah metode data mining yang efektif digunakan untuk pengenalan pola dan solusi masalah optimasi [18]. Algoritma Genetika (GA) umumnya digunakan untuk proses optimasi dan pencarian dengan prinsip kerja proses seleksi alam, merupakan bidang dari kecerdasan buatan yang berkembang yakni komputasi evolusioner. Algoritma ini memiliki konsep menyerupai proses seleksi alam dimana yang terkuat akan menjadi pemenang. Pemenang terbaik merupakan hasil yang optimal dari proses genetika yang disebut *fitness*. Penyelesaian yang didapatkan dari algoritma genetika diimplementasikan pada populasi individu yang disebut kromosom. Nilai yang diperoleh dari setiap solusi dan setiap individu bergantung pada kromosom dan dievaluasi oleh fungsi *fitness* [19].

IV. METODOLOGI PENELITIAN

A. Dataset

Dalam penelitian ini dataset yang digunakan adalah dataset publik Pima Indian Diabetes Database, diperoleh dari *UCI Machine Learning Repository*. Data yang ada pada database berjumlah 768 *record* yang terdiri dari 9 atribut. Terdiri dari 1 label atau kelas dan 8 atribut lainnya yakni Pregnant, Glucose, DBP, TSFT, INS, BMI, DPF dan Age [20]. Adapun dataset *real* tampak pada Tabel I.

B. Preprocessing Data

Preprocessing data dilakukan sebelum dataset diterapkan kedalam model yang diusulkan yakni dengan melakukan beberapa hal diantaranya;

1) Pemilihan variabel yang sesuai;

Variabel yang digunakan pada dataset ini adalah:

- Jumlah kehamilan,
- Konsentrasi glukosa plasma,
- Tekanan darah diastolic,
- Ketebalan triceps pada lipatan kulit,
- Serum insulin 2 jam,
- Berat badan,
- Silsilah diabetes (keturunan),
- Umur
- Atribut output yaitu keterangan diabetes “ya” atau “tidak” dalam bentuk biner 1 atau 0 yang artinya jika 1 adalah positif diabetes dan 0 negatif.

2) Penentuan data input dan output;

Berdasarkan gejala dan atribut yang tersedia pada dataset untuk mendiagnosa diabetes. Maka variabel input dan output dataset sebagai berikut :

- Jumlah kehamilan menggunakan atribut numerik,
- Konsentrasi glukosa plasma menggunakan atribut numerik,
- Tekanan darah diastolik (mm Hg) menggunakan atribut numerik,
- Ketebalan triceps pada lipatan kulit bawah (mm) menggunakan atribut numerik,
- Serum insulin 2 jam (mu U/ml) menggunakan atribut numerik,

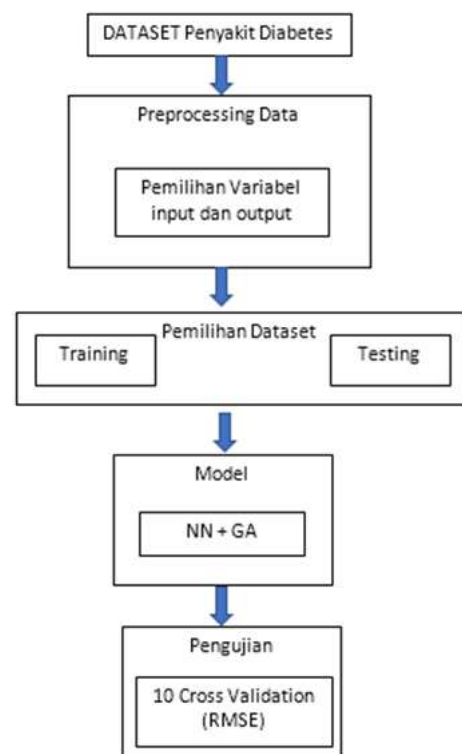
- Berat badan menggunakan (kg) atribut numerik,
- Silsilah Diabetes menggunakan atribut numerik,
- Umur (tahun),
- Output klasifikasi dengan atribut diabetes (positif/negatif).

3) Penentuan data training dan testing.

Data *training* diperoleh dari Pima Indian Diabetes Database, diperoleh dari *UCI Machine Learning Repository* yang berjumlah 768 *record*. *Dataset* kemudian dibagi menjadi dua bagian dengan perbandingan 9:1, yakni 9 bagian data *training* dan 1 bagian untuk *testing* [21].

C. Metode yang diusulkan

Algoritma Genetika (GA) diusulkan untuk dapat meningkatkan tingkat akurasi prediksi awal penyakit diabetes. Pencarian nilai parameter terbaik dilakukan oleh algoritma genetika (GA) yang nantinya akan diimplementasikan pada algoritma *Neural Network (NN)* sehingga akurasi yang dihasilkan menjadi lebih optimal. Semakin kecil nilai *Root Mean Square Error (RMSE)* yang dihasilkan maka kualitas prediksi dikategorikan baik [22]. Adapun *tools* yang digunakan untuk mendapatkan model yang diusulkan adalah *RapidMinerStudio8*. Sedangkan proses validasi model digunakan *K-Fold Validation* dengan jumlah $k=10$ [23]. Gbr. 1 menunjukkan alur metode dari model yang diusulkan dalam penelitian.



Gbr. 1 Metode yang diusulkan

TABEL I
 DATA REAL

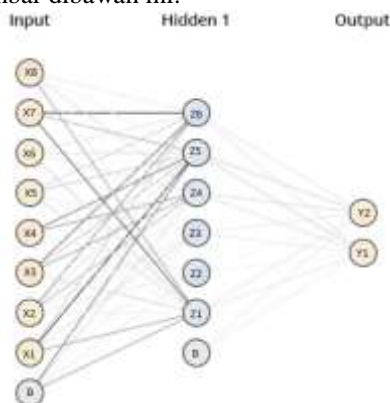
| No | Pregnant | Glucose | DBP | TSFT | INS | BMI | DPF | Age | CLASS |
|-----|----------|---------|-----|------|-----|------|-------|-----|-----------------|
| 1 | 6 | 148 | 72 | 35 | 0 | 33,6 | 0,627 | 50 | tested_positive |
| 2 | 1 | 85 | 66 | 29 | 0 | 26,6 | 0,351 | 31 | tested_negative |
| 3 | 8 | 183 | 64 | 0 | 0 | 23,3 | 0,672 | 32 | tested_positive |
| 4 | 1 | 89 | 66 | 23 | 94 | 28,1 | 0,167 | 21 | tested_negative |
| 5 | 0 | 137 | 40 | 35 | 168 | 43,1 | 2,288 | 33 | tested_positive |
| 6 | 5 | 116 | 74 | 0 | 0 | 25,6 | 0,201 | 30 | tested_negative |
| 7 | 3 | 78 | 50 | 32 | 88 | 31 | 0,248 | 26 | tested_positive |
| 8 | 10 | 115 | 0 | 0 | 0 | 35,3 | 0,134 | 29 | tested_negative |
| 9 | 2 | 197 | 70 | 45 | 543 | 30,5 | 0,158 | 53 | tested_positive |
| 10 | 8 | 125 | 96 | 0 | 0 | 0 | 0,232 | 54 | tested_positive |
| 11 | 4 | 110 | 92 | 0 | 0 | 37,6 | 0,191 | 30 | tested_negative |
| 12 | 10 | 168 | 74 | 0 | 0 | 38 | 0,537 | 34 | tested_positive |
| 13 | 10 | 139 | 80 | 0 | 0 | 27,1 | 1,441 | 57 | tested_negative |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 762 | 2 | 88 | 58 | 26 | 16 | 28,4 | 0,766 | 22 | tested_negative |
| 762 | 9 | 170 | 74 | 31 | 0 | 44 | 0,403 | 43 | tested_positive |
| 763 | 9 | 89 | 62 | 0 | 0 | 22,5 | 0,142 | 33 | tested_negative |
| 764 | 10 | 101 | 76 | 48 | 180 | 32,9 | 0,171 | 63 | tested_negative |
| 765 | 2 | 122 | 70 | 27 | 0 | 36,8 | 0,34 | 27 | tested_negative |
| 766 | 5 | 121 | 72 | 23 | 112 | 26,2 | 0,245 | 30 | tested_negative |
| 767 | 1 | 126 | 60 | 0 | 0 | 30,1 | 0,349 | 47 | tested_positive |
| 768 | 1 | 93 | 70 | 31 | 0 | 30,4 | 0,315 | 23 | tested_negative |

V. HASIL DAN PEMBAHASAN

Eksperimen ini menggunakan komputer dengan spesifikasi processor CPU Intel Core i3 2.10 GHz, RAM 4 GB, serta sistem operasi Windows 7 Ultimate 64-bit. Setelah eksperimen dilakukan, didapatkan hasil yang berkaitan dengan model yang telah diusulkan. Ekperimen dilakukan melalui dua tahapan, yaitu (1) Penerapan Algoritma Neural Network (NN); yaitu dengan melakukan ekperimen parameter Neural Network (NN) terbaik. (2) penerapan Neural Network (NN) dan Algoritma Genetika (GA); yaitu dengan pencarian parameter terbaik Neural Network (NN) terbaik berbasiskan GA.

A. Hasil Eksperimen Neural Network (NN) Terbaik

Eksperimen yang telah dilakukan dengan menerapkan algoritma Neural Network (NN) terhadap dataset yang telah ditentukan sebagai data training diperoleh arsitektur model seperti gambar dibawah ini:



Gbr. 2 Arsitektur model Neural Network (NN) terbaik.

Gambar 2 menunjukkan bahwa data input yang digunakan sebanyak delapan neuron input sedangkan output dihasilkan dua neuron serta dihasilkan juga enam neuron hidden layer. Pada ekperimen ini ditetapkan nilai parameter yang terbaik adalah Training Cycles = 100, Learning rate = 0.1, dan Momentum = 0.1 sehingga RMSE = 0.402 +/- 0.035.

B. Hasil Eksperimen Neural Network (NN) dan GA

Proses pengujian guna mendapatkan nilai parameter terbaik dari Neural Network (NN) dilakukan dengan setting parameter algoritma genetika yakni Population size = 5, Max generations = 50, Mutation type = Gaussian mutation, Crossover prob = 0.9 dan Selection type = Roulette wheel

TABEL II
 HASIL EKSPERIMEN NN DENGAN GA

| Parameter | Algoritma NN + Algoritma Genetika |
|---------------|-----------------------------------|
| Learning rate | 0,1985425571526106 |
| Momentum | 0,39084178927133306 |
| RMSE | 0,396 +/- 0,032 |

Berdasarkan pengamatan terhadap ekperimen yang telah dilakukan, dengan adanya optimasi pada nilai parameter Neural Network (NN) dengan algoritma genetika (GA) mengalami peningkatan akurasi menjadi 0,396 +/- 0,032. Dengan nilai parameter terbaik untuk Learning rate sebesar 0,1985425571526106 dan Momentum terbaik sebesar 0,39084178927133306.

C. Evaluasi

Dalam penelitian ini Algoritma Genetika (GA) diterapkan kedalam model *Neural Network (NN)* dengan tujuan untuk mengoptimalkan tingkat akurasi prediksi deteksi awal penyakit diabetes. Adapun teknik optimasi dilakukan dengan cara pencarian nilai parameter terbaik model *Neural Network (NN)*. Hasil eksperimen menunjukkan terjadinya penurunan nilai RMSE yang berarti terjadi peningkatan pada nilai akurasi prediksi, yaitu dari 0,402 +/- 0,035 menjadi 0,396 +/- 0,032

TABEL III
 PERBANDINGAN NN DAN NN+GA

| Model | RMSE |
|-------|-----------------|
| NN | 0,402 +/- 0,035 |
| NN+GA | 0,396 +/- 0,032 |

VI. KESIMPULAN

Eksperimen yang telah dilakukan dengan mengusulkan Algoritma Genetika (GA) untuk diterapkan kedalam model yang digunakan yakni *Neural Network (NN)* dalam memprediksi deteksi awal penyakit diabetes, dapat disimpulkan bahwa Algoritma Genetika (GA) dapat mengoptimalkan model *Neural Network (NN)* dengan cara pencarian nilai parameter terbaik pada model *Neural Network (NN)*. Sehingga terjadi peningkatan akurasi prediksi. Dengan adanya optimasi prediksi model tersebut, maka para pengambil kebijakan terkait dapat menentukan prediksi penyakit diabetes dengan lebih tepat. Eksperimen dalam penelitian ini belum cukup untuk dapat menentukan ketepatan akurasi prediksi. Pada penelitian selanjutnya diharapkan untuk dapat menerapkan algoritma optimasi lain guna mengetahui perbedaan tingkat akurasi yang terjadi sehingga nilai RMSE menjadi lebih optimal.

DAFTAR PUSTAKA

- [1] A. Zamora-Kapoor, A. Fyfe-Johnson, A. Omidpanah, D. Buchwald, and K. Sinclair, "Risk factors for pre-diabetes and diabetes in adolescence and their variability by race and ethnicity," *Prev. Med. (Baltim.)*, vol. 115, no. August, pp. 47–52, 2018.
- [2] A. A. Aljumah, M. G. Ahamad, and M. K. Siddiqui, "Application of data mining: Diabetes health care in young and old patients," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 25, no. 2, pp. 127–136, 2012.
- [3] H. Wu, S. Yang, Z. Huang, J. He, and X. Wang, "Type 2 diabetes mellitus prediction model based on data mining," *Informatics Med. Unlocked*, vol. 10, no. August 2017, pp. 100–107, 2018.
- [4] "No Title." International Diabetes Federation (IDF), 2015.
- [5] R. Premunendar, I. Dewi, and H. Asari, "Penentuan Prediksi Awal Penyakit Jantung Menggunakan Algoritma Back Propagation Neural Network dengan Metode Adaboost," *Semantik*, vol. 2013, no. November, pp. 298–304, 2013.
- [6] K. Rajesh and V. Sangeetha, "Application of Data Mining Methods and Techniques for Diabetes Diagnosis," (*sem qualis*) *Int. J. Eng. Res. Innov. Technol.*, vol. 2, no. 3, pp. 224–229, 2012.
- [7] A. Andriani, "Sistem prediksi penyakit diabetes berbasis decision tree," *J. Bianglala Inform.*, vol. I, no. 1, pp. 1–10, 2013.
- [8] S. Kang, "Personalized prediction of drug efficacy for diabetes treatment via patient-level sequential modeling with neural networks," *Artif. Intell. Med.*, vol. 85, pp. 1–6, 2018.
- [9] R. W. Purnamasari, E. Sugiharti, and Dwijanto, "Unnes Journal of Public Health," *UNNES J. Math.*, vol. 2, no. 2, pp. 0–6, 2013.
- [10] O. Somantri and C. Supriyanto, "Algoritme Genetika untuk Peningkatan Prediksi Kebutuhan Permintaan Energi Listrik," *Jnteti*, vol. 5, no. 2, pp. 108–114, 2016.
- [11] A. M. Zamani and B. Amaliah, "Implementasi Algoritma Genetika pada Struktur Backpropagation Neural Network untuk Klasifikasi Kanker Payudara," *J. Tek. POMITS*, vol. 1, no. 1, pp. 1–6, 2012.
- [12] F. Fatmawati, "PERBANDINGAN ALGORITMA KLASIFIKASI DATA MINING MODEL C4.5 DAN NAIVE BAYES UNTUK PREDIKSI PENYAKIT DIABETES," *J. Techno Nusa Mandiri*, vol. 13, no. 1, pp. 50–59, Mar. 2016.
- [13] X.-H. Meng, Y.-X. Huang, D.-P. Rao, Q. Zhang, and Q. Liu, "Comparison of three data mining models for predicting diabetes or prediabetes by risk factors," *Kaohsiung J. Med. Sci.*, vol. 29, no. 2, pp. 93–99, Feb. 2013.
- [14] A. Rohman, V. Suhartono, and C. Supriyanto, "Penerapan algoritma c4.5 berbasis," vol. 13, pp. 13–19, 2017.
- [15] A. Shukla, R. Tiwari, and R. Kala, *Real Life Applications of Soft Computing*. 2012.
- [16] E.D Astuti, *Pengantar Jaringan Saraf Tiruan*. Wonosobo: Star Publishing, 2009.
- [17] M. Badrul, "Optimasi Algoritma Neural Network Dengan Algoritma Genetika Dan Particle Swarm Optimization Untuk Memprediksi Hasil Pemilukada," *J. Pilar Nusa Mandiri*, vol. 13, no. 1, pp. 1–11, 2017.
- [18] M. Badrul, "Optimasi Neural Network Dengan Algoritma Genetika Untuk Prediksi Hasil Pemilukada," vol. 3421, pp. 229–242, 2016.
- [19] T. Weise, "Global Optimization Algorithms - Theory and Application," 2007.
- [20] D. Sisodia and D. S. Sisodia, "Prediction of Diabetes using Classification Algorithms," in *Procedia Computer Science*, 2018, vol. 132, no. Iccids, pp. 1578–1585.
- [21] H. Jiawei and M. Kamber, *Data mining: concepts and techniques*, vol. 5. 2001.
- [22] Susanti, S. Martha, and E. Sulistianingsih, "K Nearest Neighbor dalam Imputasi Missing Data," *Bimaster*, vol. 7, no. 1, pp. 1–18, 2018.
- [23] G. Jiang and W. Wang, "Error estimation based on variance analysis of k-fold cross-validation," *Pattern Recognit.*, vol. 69, pp. 94–106, 2017.