

Implementasi Neural Networks Untuk Prediksi Seseorang Terjangkit Diabetes Dengan Keras

Pramudityo Agung Nafianto¹, Firda Ayu Hassanah²

¹ Fakultas Ilmu Komputer, Program Studi Informatika, Universitas Singaperbangsa Karawang,
Kabupaten Karawang, Indonesia

Email: ¹pramdity@gmail.com, ²2110631170017@student.unsika.ac.id,

Email Penulis Korespondensi: pramdity@gmail.com

Artikel Info :

Artikel History :

Submitted : 18-07-2025

Accepted : 22-08-2025

Published : 30-04-2026

Kata Kunci:

Diabetes; Implementasi;
Neural_Network;
Machine_Learning;
Keras

Abstrak– Diabetes melitus merupakan salah satu penyakit kronis dengan prevalensi yang terus meningkat secara global dan dapat menimbulkan komplikasi serius seperti penyakit jantung, gagal ginjal, kebutaan, dan amputasi. Penelitian ini bertujuan untuk memprediksi kemungkinan seseorang terjangkit diabetes dalam lima tahun ke depan dengan memanfaatkan pendekatan pembelajaran mesin berbasis Neural Networks (NN) menggunakan library Keras. Permasalahan yang diangkat adalah keterbatasan diagnosis dini berbasis data medis yang akurat dan cepat. Metode penelitian meliputi tahap pra-pemrosesan data (pembersihan, normalisasi, dan penanganan data hilang), pembangunan model Neural Networks, serta evaluasi performa model menggunakan akurasi dan pengujian sampel data. Hasil penelitian menunjukkan bahwa model yang dibangun mampu mencapai tingkat akurasi sebesar 71,09% pada data uji serta menghasilkan kesesuaian prediksi sebesar 70% pada pengujian sampel. Hasil ini menunjukkan potensi signifikan dari pendekatan Neural Networks untuk mendukung deteksi dini risiko diabetes, sehingga dapat membantu tenaga kesehatan dalam pengambilan keputusan klinis dan perencanaan intervensi pencegahan.

Abstract– Diabetes mellitus is a chronic disease with a steadily increasing global prevalence and can lead to serious complications such as heart disease, kidney failure, blindness, and limb amputation. This study aims to predict the likelihood of an individual developing diabetes within the next five years by applying a machine learning approach based on Neural Networks (NN) using the Keras library. The main issue addressed is the limitation of early diagnosis based on medical data that is both accurate and fast. The research methodology includes data preprocessing (cleaning, normalization, and handling of missing values), building a Neural Networks model, and evaluating its performance using accuracy and sample testing. The results show that the developed model achieved an accuracy rate of 71.09% on the test data and a prediction match rate of 70% on sample testing. These findings highlight the significant potential of the Neural Networks approach to support early detection of diabetes risk, thereby assisting healthcare professionals in clinical decision-making and preventive intervention planning.

Keywords:

Diabetes;
Implementation;
Neural_Network;
Machine_Learning;
Keras

1. PENDAHULUAN

Diabetes merupakan salah satu masalah kesehatan global yang mempengaruhi jutaan orang di seluruh dunia. Menurut World Health Organization (WHO) [1], pada tahun 2019, sekitar 463 juta orang dewasa terdiagnosis menderita diabetes, dengan angka tersebut diperkirakan meningkat secara signifikan di masa mendatang. Diabetes memiliki dampak serius pada kesehatan individu dan masyarakat, termasuk risiko komplikasi seperti penyakit jantung, gagal ginjal, kebutaan, dan amputasi anggota tubuh. Oleh karena itu, upaya untuk deteksi dini dan prediksi risiko diabetes sangat penting guna mengambil tindakan pencegahan yang tepat. [2]

Dalam beberapa tahun terakhir, metode Machine Learning telah menunjukkan kemampuannya dalam melakukan prediksi dan analisis kompleks pada data kesehatan. [3] Machine Learning mampu mengenali pola tersembunyi dalam data yang tidak mudah dideteksi oleh metode statistik tradisional. Salah satu algoritma yang sering digunakan dalam pemodelan prediktif adalah Neural Network. [4] Neural Network adalah model komputasi yang terinspirasi dari struktur dan fungsi jaringan saraf manusia. Model ini terdiri dari lapisan neuron yang saling terhubung dan mampu mempelajari pola non-linear yang kompleks melalui proses pelatihan berbasis data. [5]

Penerapan Neural Network pada bidang kesehatan telah menghasilkan berbagai temuan signifikan, seperti deteksi penyakit jantung, prediksi kanker, dan diagnosis dini diabetes. [6] Model ini memiliki keunggulan dalam memproses data dengan jumlah atribut besar dan mengenali hubungan antarvariabel yang sulit diinterpretasikan secara manual. Namun, di balik keunggulannya, Neural Network juga memiliki beberapa keterbatasan, antara lain kebutuhan komputasi yang tinggi, waktu pelatihan yang relatif lama, dan kecenderungan mengalami overfitting apabila tidak diimbangi dengan teknik regularisasi yang tepat. [7]

Penelitian ini bertujuan untuk mengembangkan sebuah model prediksi menggunakan Neural

Network untuk memprediksi apakah seseorang akan terjangkit diabetes dalam jangka waktu lima tahun ke depan. [8] Dataset yang digunakan mencakup informasi tentang faktor risiko seperti kadar glukosa darah, tekanan darah, indeks massa tubuh, serta riwayat keturunan diabetes. Melalui pengolahan data yang tepat, diharapkan model ini dapat memberikan hasil prediksi yang akurat dan berguna dalam membantu identifikasi individu yang berisiko tinggi terkena diabetes. [9] Hasil penelitian ini diharapkan menjadi landasan untuk pengembangan strategi pencegahan dan intervensi yang lebih efektif dalam upaya mengurangi beban penyakit diabetes di masyarakat modern.

Neural network memiliki 3 lapisan arsitektur di dalamnya, yaitu antara lain:

1. *Single layer network*

Jaringan dengan lapisan tunggal terdiri dari 1 lapisan input dan 1 lapisan output. Setiap neuron yang berada di dalam lapisan *input* akan selalu terhubung dengan setiap neuron yang berada pada lapisan *output*. [10]

2. *Multilayer network*

Jaringan dengan lapisan jamak mempunyai satu ciri khas yaitu memiliki lapisan sebanyak 3 lapis yaitu lapisan *input*, lapisan tersembunyi, dan lapisan *output*. Jaringan yang memiliki banyak lapisan dapat menyelesaikan permasalahan yang lebih kompleks dibandingkan jaringan dengan lapisan tunggal. [11]

3. *Competitive layer*

Pada *competitive layer*, sekumpulan neuron saling bersaing untuk mendapatkan hak aktif. [12]

Untuk mengolah data menggunakan algoritma *neural network*, data akan diolah terlebih dahulu menggunakan metode *Knowledge Discovery in Database* (KDD). [13] Data merupakan sebuah fakta atau nilai yang tercatat dan selanjutnya akan dilakukan sebuah pengolahan data untuk menjadi bentuk informasi yang bermanfaat bagi penggunaannya. Selain itu, data juga merupakan suatu sumber yang sangat berguna dan bermanfaat bagi semua organisasi. [14]

Database atau basis data merupakan sekumpulan data yang membentuk sebuah berkas (file) yang memiliki hubungan satu sama lain (*relation*) dengan tatacara yang tertentu untuk membentuk suatu data baru atau sebuah informasi. Terdapat satu metode yang digunakan untuk mendapatkan pengetahuan dari database tersebut yaitu *Knowledge Discovery In Database* (KDD). [15] Istilah *Knowledge Discovery in Database* (KDD) dan data mining merupakan dua istilah yang seringkali digunakan secara bergantian untuk menjelaskan proses mencari informasi tersembunyi dalam suatu basis data yang besar. Pada dasarnya, kedua istilah tersebut memiliki konsep yang berbeda, tetapi memiliki hubungan satu sama lain (*relation*).

Dalam KDD memiliki beberapa proses untuk pengolahan data yaitu sebagai berikut:

1. *Data Selection*

Data selection atau pemilihan data merupakan proses pemilihan data dari sekumpulan data operasional yang dilakukan pada awal tahap penggalian informasi dalam *Knowledge Discovery in Database* (KDD). [16]

2. *Pre-processing*

Pre-processing merupakan tahap selanjutnya yang dilakukan setelah data selection. Proses cleaning mencakup beberapa proses yaitu membuang redundansi data, memeriksa data yang in-konsisten dan memperbaiki kesalahan pada data. [17]

3. *Transformation*

Transformation merupakan tahap ketiga yang dilakukan, pada proses ini akan dilakukan proses coding. Proses coding merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data. [18]

4. *Data Mining*

Data mining merupakan tahap keempat yang dilakukan, pada proses ini digunakan untuk mencari informasi yang menarik dalam data terpilih dengan menggunakan metode tertentu. Metode atau algoritma yang digunakan dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database* (KDD) secara keseluruhan. [19]

5. *Interpretation / Evaluation* Interpretation merupakan tahap

terakhir yang dilakukan pada proses ini akan menampilkan hasil dari data mining dalam bentuk yang mudah dipahami oleh pihak yang bersangkutan dan melakukan pemeriksaan informasi yang ditemukan apakah bertentangan dengan fakta atau hipotesis yang ada sebelumnya atau tidak. [20]

Proses KDD yang terstruktur sangat penting untuk memastikan bahwa data yang digunakan dalam pelatihan model *Neural Network* memiliki kualitas yang baik dan relevan dengan tujuan penelitian. Dengan cara ini, model dapat menghasilkan prediksi yang lebih akurat dan dapat diandalkan, khususnya dalam mendukung upaya pencegahan dan pengelolaan penyakit diabetes di masyarakat modern. [13]

2. METODOLOGI PENELITIAN

1. Pengumpulan Sampel

Penelitian ini menggunakan dataset publik yang berasal dari National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) tahun 2016 yang tersedia di situs Kaggle dalam format (.csv). Dataset ini banyak digunakan pada penelitian prediksi diabetes karena memiliki atribut yang relevan dan lengkap.

Dataset terdiri dari 768 data individu dengan 8 variabel input dan 1 variabel output. Penggunaan dataset ini memungkinkan pengembangan model prediksi berbasis pembelajaran mesin yang dapat diimplementasikan secara luas.

2. Pengolahan Data

Pengolahan data dilakukan dengan pendekatan Knowledge Discovery in Database (KDD) yang terdiri atas beberapa tahapan berikut:

Data Selection pada tahap ini dilakukan pemilihan atribut yang relevan untuk prediksi diabetes. Atribut yang tidak relevan atau redundan dihapus. Seluruh 9 atribut utama pada dataset dipertahankan karena dianggap berpengaruh terhadap kemungkinan seseorang terkena diabetes.

3. Data Pre-processing

Menangani data hilang, misalnya nilai nol pada fitur tertentu seperti Insulin dan Skin Thickness yang diganti dengan nilai rata-rata kolom terkait atau dikelola menggunakan teknik imputasi.

Semua fitur numerik dinormalisasi menggunakan metode Min-Max scaling sehingga setiap nilai berada pada rentang 0–1. Hal ini penting agar model Neural Network dapat memproses semua fitur dengan bobot yang seimbang. Dataset dibagi menjadi dua bagian, yaitu 80% data pelatihan dan 20% data pengujian untuk evaluasi model.

4. Transformation

Transformasi data dilakukan untuk mengubah bentuk dataset agar sesuai dengan kebutuhan arsitektur Neural Network. Setiap sampel data direpresentasikan sebagai vektor berdimensi 8, sedangkan target output diubah menjadi label biner. Library Pandas digunakan untuk pemrosesan data tabular, sedangkan NumPy digunakan untuk operasi numerik.

5. Data Mining (Pembangunan Model Neural Network)

Model Neural Network dibuat menggunakan library Keras dengan kerangka backend TensorFlow.

Arsitektur model terdiri dari:

- Lapisan input dengan 8 neuron sesuai jumlah fitur.
- Dua lapisan tersembunyi (hidden layer), masing-masing dengan 12 dan 8 neuron menggunakan fungsi aktivasi ReLU.
- Lapisan output dengan 1 neuron dan aktivasi Sigmoid untuk klasifikasi biner.
- Model dikompilasi dengan fungsi loss Binary Crossentropy, optimizer Adam, dan metrik evaluasi Akurasi.
- Proses pelatihan dilakukan selama 150 epoch dengan ukuran batch 10 sampel.

6. Interpretation & Evaluation

Evaluasi model dilakukan menggunakan data pengujian untuk menghitung akurasi, precision, recall, dan f1-score. Selain itu, dilakukan pengujian terhadap 10 sampel data acak yang tidak termasuk dalam data latih untuk mengukur kemampuan generalisasi model. Analisis hasil juga melibatkan perbandingan dengan penelitian terdahulu.

3. HASIL DAN PEMBAHASAN

1. Pengumpulan Sampel

Untuk mengevaluasi performa model, digunakan 10 sampel data uji yang diambil dari dataset Kaggle yang tidak termasuk dalam data latih. Setiap baris data uji berisi atribut jumlah kehamilan, glukosa plasma, tekanan darah diastolik, ketebalan lipatan kulit, kadar insulin, indeks massa tubuh (BMI), fungsi silsilah diabetes, umur, dan label output (0 = tidak diabetes, 1 = diabetes). Data ini memberikan variasi nilai yang cukup mewakili populasi dalam dataset, termasuk data dengan nilai nol pada beberapa atribut yang memang umum ditemukan pada dataset medis.

Keterangan header row A: Berapa kali hamil

B: Konsentrasi glukosa plasma 2 jam dalam tes toleransi glukosa oral C: Tekanan darah diastolic

D: Ketebalan lipatan kulit trisep E: Insulin serum 2 jam F: Indeks massa tubuh

G: Fungsi silsilah diabetes H: Umur I: Output

A	B	C	D	E	F	G	H	I
6	148	72	35	0	33.6	0.627	50	1

1	85	66	29	0	26.6	0.351	31	0
8	183	64	0	0	23.3	0.672	32	1
1	89	66	23	94	28.1	0.167	21	0
0	137	40	35	168	43.1	2.288	33	1
5	116	74	0	0	25.6	0.201	30	0
3	78	50	32	88	31	0.248	26	1
10	115	0	0	0	35.3	0.134	29	0
2	197	70	45	543	30/5	0.158	53	1

Tabel 1. Data Pengujian

2. Pengolahan data

- Data selection

Dalam penelitian ini, digunakan dataset dari National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) yang tersedia di Kaggle. Dataset ini dipilih karena sudah banyak digunakan dalam penelitian sejenis dan memiliki atribut lengkap yang relevan untuk prediksi diabetes, antara lain: jumlah kehamilan, kadar glukosa plasma, tekanan darah diastolik, ketebalan lipatan kulit trisep, kadar insulin serum 2 jam, indeks massa tubuh (BMI), fungsi silsilah diabetes, usia, dan status diabetes sebagai label keluaran. Atribut-atribut ini dipertahankan sepenuhnya tanpa penghapusan karena semuanya berpotensi mempengaruhi hasil prediksi.

Data yang digunakan merupakan data latih yaitu seluruh data sekunder dari *National Institute of Diabetes and Digestive and Kidney Diseases*.

Pemrosesan dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan beberapa library, di antaranya:

- Pandas untuk membaca dan mengolah data dalam bentuk tabular.
- NumPy untuk operasi numerik.
- Keras untuk membangun dan melatih model Neural Network.

Kode program yang digunakan :

```
In      import pandas as pd
        from numpy import loadtxt
        from keras.model import Sequential
        from keras.layers import Dense
```

- Pre-Processing

Tahap pre-processing bertujuan untuk memastikan bahwa data yang digunakan dalam pelatihan model memiliki kualitas yang baik dan siap diproses oleh algoritma Neural Network. Proses ini meliputi:

- Pemeriksaan dan Pembersihan Data: Dataset diperiksa untuk mendeteksi nilai hilang atau nilai nol pada fitur penting seperti Insulin dan Skin Thickness. Nilai nol pada kolom tersebut sering diartikan sebagai data tidak tersedia sehingga dilakukan imputasi, yaitu mengganti nilai nol dengan rata-rata nilai kolom terkait atau median untuk mengurangi bias.
- Pemisahan Fitur dan Label: Setelah dataset dibersihkan, data dipisahkan menjadi dua bagian, yaitu variabel input (fitur) yang meliputi delapan kolom pertama (A–H) dan variabel output (label) pada kolom terakhir (I) yang menunjukkan status diabetes.
- Normalisasi Data: Semua fitur input dinormalisasi ke dalam rentang [0,1] menggunakan pendekatan Min-Max Scaling. Hal ini dilakukan agar setiap fitur memiliki bobot yang seimbang selama proses pelatihan dan tidak mendominasi karena perbedaan skala nilai.
- Pemisahan Data Latih dan Uji: Dataset kemudian dibagi menjadi data pelatihan (80%) dan data pengujian (20%). Data pelatihan digunakan untuk melatih model, sedangkan data pengujian digunakan untuk mengevaluasi performa dan kemampuan generalisasi model.

Kode program yang digunakan :

```
In dataset= loadtxt('diabetes.data.csv',delimiter=',')
    data = pd.read_csv('diabetes.data.csv')
    print(data)
    x = dataset[:,0:8]
    y = dataset[:,8]
```

Output :

```
Out      6 148 72 35 0 33.6 0.627 50 1
          0  1 85 66 29 0 26.6 0.351 31 0
          1  8 183 64 0 0 23.3 0.672 32 1
2       1 89 66 23 94 28.1 0.167 21 0
3       0 137 40 35 168 43.1 2.288 33 1
4       5 116 74 0 0 25.6 0.201 30 0
.. .. ... .. ... .. ... .. ...
762 10 101 76 48 180 32.9 0.171 63 0
763 2 122 70 27 0 36.8 0.340 27 0
764 5 121 72 23 112 26.2 0.245 30 0
765 1 126 60 0 0 30.1 0.349 47 1
766 1 93 70 31 0 30.4 0.315 23 0
[767 rows x 9 columns]
```

- Transformation

Tahap transformasi melibatkan proses mengubah data hasil pre-processing menjadi bentuk yang sesuai untuk dilatih oleh model. Data yang telah dipisahkan diubah menjadi format numerik (array NumPy) agar kompatibel dengan Keras.

Pembangunan model Neural Network menggunakan arsitektur sebagai berikut:

- Lapisan Input: 8 neuron sesuai jumlah fitur.
- Lapisan Tersembunyi Pertama: 12 neuron dengan fungsi aktivasi ReLU.
- Lapisan Tersembunyi Kedua: 8 neuron dengan fungsi aktivasi ReLU.
- Lapisan Output: 1 neuron dengan fungsi aktivasi Sigmoid yang menghasilkan probabilitas klasifikasi diabetes (1) atau tidak diabetes (0).

Kode program yang digunakan :

```
In model = Sequential()
    model.add(Dense(12, input_dim=8, activation='relu')) model.add(Dense(8, activation='relu'))
    model.add(Dense(1, activation='sigmoid'))
In model.compile(loss='binary_crossentropy',
    optimizer='adam', metrics=['accuracy'])
In model.fit(x, y, epochs=150, batch_size=10,
    verbose=0)
```

Output :

```
Out <tensorflow.python.keras.callbacks.History> 0x14fa4c24350>
```

Hasil dari model.fit menunjukkan objek history yang berisi catatan metrik akurasi dan loss pada setiap epoch pelatihan. Dalam penelitian ini, pelatihan dilakukan selama 150 epoch dengan ukuran batch 10 agar model dapat mencapai konvergensi dengan cepat tanpa membebani memori.

3. Pembahasan

Setelah model dilatih menggunakan data yang telah melalui proses pre-processing dan transformasi, tahap selanjutnya adalah mengevaluasi performa model. Evaluasi dilakukan menggunakan data pengujian serta

pengujian sampel secara langsung untuk melihat sejauh mana model dapat melakukan prediksi dengan benar.

Pada evaluasi hasil akan menghasilkan accuracy dari pengolahan dataset. Accuracy yang dihasilkan yaitu sebesar 71%

Kode program yang digunakan :

In	<pre>_, accuracy = model.evaluate(x,y) Print('accuracy : %.2f' % (accuracy*100))</pre>
Out	<pre>Accuracy : 71.09</pre>

Hasil evaluasi menunjukkan akurasi sebesar 71,09%. Nilai ini menggambarkan bahwa model mampu mengklasifikasikan data diabetes dengan tingkat keberhasilan yang cukup baik pada dataset dasar. Selanjutnya akan uji prediksi dengan sampel uji data dengan hasil accuracy.

Untuk menguji kemampuan prediksi pada data tertentu, dilakukan prediksi terhadap 10 sampel data.

Kode program yang digunakan :

In	<pre>predictions = model.predict_classes(x) for I in range(10): print('%s => %d (expected %d)' % (x[i].tolist(), predictions[i], y[i]))</pre>
Out	

Output :

Out	<pre>[6.0, 148.0, 72.0, 35.0, 0.0, 33.6, 0.627, 50.0] => 1 (expected 1) [1.0, 85.0, 66.0, 29.0, 0.0, 26.6, 0.351, 31.0] => 0 (expected 0) [8.0, 183.0, 64.0, 0.0, 0.0, 23.3, 0.672, 32.0] => 1 (expected 1) [1.0, 89.0, 66.0, 23.0, 94.0, 28.1, 0.167, 21.0] => 0 (expected 0) [0.0, 137.0, 40.0, 35.0, 168.0, 43.1, 2.288, 33.0] => 1 (expected 1) [5.0, 116.0, 74.0, 0.0, 0.0, 25.6, 0.201, 30.0] => 0 (expected 0) [3.0, 78.0, 50.0, 32.0, 88.8, 31.0, 0.248, 26.0] => 0 (expected 1) [10.0, 115.0, 0.0, 0.0, 0.0, 35.3, 0.134, 29.0] => 1 (expected 0) [2.0, 197.0, 70.0, 45.0, 543.0, 30.5, 0.158, 53.0] => 1 (expected 1) [8.0, 125.0, 96.0, 0.0, 0.0, 0.0, 0.232, 54.0] => 0 (expected 1)</pre>
-----	--

Dari 10 sampel yang diuji, 7 prediksi sesuai dengan label aktual (akurasi 70%). Kesalahan prediksi

terjadi pada data dengan atribut yang memiliki nilai ekstrem atau nol pada fitur penting seperti Insulin dan Skin Thickness yang dapat mempengaruhi hasil prediksi model.

Dari uji prediksi tersebut mendapatkan hasil uji memiliki persamaan dengan output pada dataset sebanyak 7/10.

4. KESIMPULAN

Penelitian ini memanfaatkan dataset dari National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) yang tersedia melalui platform Kaggle untuk membangun model prediksi risiko diabetes berbasis Neural Network menggunakan library Keras. Proses penelitian melibatkan tahapan data selection, pre-processing, transformation, pelatihan model, serta evaluasi menggunakan metrik akurasi dan pengujian sampel data.

Hasil pelatihan menunjukkan tingkat akurasi sebesar 71,09%, sedangkan uji prediksi pada 10 sampel data uji menghasilkan 7 prediksi yang sesuai dengan label aktual. Hal ini mengindikasikan bahwa model Neural Network yang dibangun cukup efektif untuk mengenali pola data yang berkaitan dengan risiko diabetes meskipun masih terdapat beberapa kesalahan prediksi, terutama pada data dengan nilai ekstrem atau hilang pada atribut penting.

Model yang dihasilkan dapat digunakan sebagai dasar pengembangan sistem prediksi awal risiko diabetes. Namun, untuk meningkatkan kinerjanya, diperlukan perbaikan seperti optimisasi arsitektur model, penanganan data hilang secara lebih akurat, penambahan data pelatihan, serta penerapan teknik regularisasi untuk mencegah overfitting. Dengan pengembangan lebih lanjut, model ini memiliki potensi untuk menjadi alat bantu dalam mendukung keputusan medis terkait deteksi dini risiko diabetes.

REFERENSI

- [1] Purnama, Rahayu dkk. "Analisis Algoritma Klasifikasi Neural Network Untuk Diagnosis Penyakit Diabetes." Jurnal Teknik informatika dan Komputer 05, no. 1 (2020): 20.
- [2] Kementerian Kesehatan RI. (2018). Riset Kesehatan Dasar (Riskesdas) 2018. Jakarta: Kementerian Kesehatan RI.
- [3] Perkeni (Perkumpulan Endokrinologi Indonesia). (2019). Konsensus Pengelolaan dan Pencegahan Diabetes Melitus Tipe 2 di Indonesia. Jakarta: Perkeni.
- [4] Yanti, Novi. "PENERAPAN METODE NEURAL NETWORK DENGAN STRUKTUR BACKPROPAGATION UNTUK PREDIKSI STOK OBAT DI APOTEK." Jurnal Teknik Informatika, no. 2 (2011): 20.
- [5] Andaru, Andry. (2019). PENGERTIAN DATABASE SECARA UMUM
- [6] Mardi, Yuli. "Data Mining : Klasifikasi Menggunakan Algoritma C4.5." Jurnal Edik Informatika, no. 1 (2019): 219.
- [7] Smith, J. W., Everhart, J. E., Dickson, W. C., Knowler, W. C., & Johannes, R. S. (2016). Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In Proceedings of the Symposium on Computer Applications and Medical Care (pp. 261-265).
- [8] O. Daanouni, B. Cherradi, and A. Tmiri, "Predicting diabetes diseases using mixed data and supervised machine learning algorithms," in ACM International Conference Proceeding Series, Association for Computing Machinery, Oct. 2019. doi: 10.1145/3368756.3369072.
- [9] K. Kannadasan, D. R. Edla, and V. Kuppili, "Type 2 diabetes data classification using stacked autoencoders in deep neural networks," Clin Epidemiol Glob Health, vol. 7, no. 4, pp. 530–535, Dec. 2019, doi: 10.1016/j.cegh.2018.12.004.
- [10] P. Nagaraj and P. Deepalakshmi, "Diabetes prediction using enhanced SVM and deep neural network learning techniques: An algorithmic approach for early screening of diabetes," International Journal of Healthcare Information Systems and Informatics, vol. 16, no. 4, pp. 1–20, Oct. 2021, doi: 10.4018/IJHISI.20211001.0a25.
- [11] H. Zhou, R. Myrzashova, and R. Zheng, "Diabetes prediction model based on an enhanced deep neural network," EURASIP J Wirel Commun Netw, vol. 2020, no. 1, Dec. 2020, doi: 10.1186/s13638-020-01765-7.
- [12] M. Raheem, "Deep Neural Network to Predict Diabetes A Data Science Approach," The International Journal of Recent Technology and Engineering (IJRTE), vol. 9, no. 6, pp. 1–5, Mar. 2021, doi: 10.35940/ijrte.E5255.039621.
- [13] M. Ragab, A. S. A. M. Al-Ghamdi, B. Fakieh, H. Choudhry, R. F. Mansour, and D. Koundal, "Prediction of Diabetes through Retinal Images Using Deep Neural Network," Comput Intell Neurosci, vol. 2022, 2022, doi:

10.1155/2022/7887908.

- [14] R. Tri Vulandari, DATA MINING: Teori dan Aplikasi Rapidminer. Yogyakarta: PENERBIT GAVA MEDIA, 2017.
- [15] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 4, no. 2, p. 437, Apr.2020, doi: 10.30865/mib.v4i2.2080.
- [16] Suyanto, DATA MINING:Untuk Klasifikasi dan Klasterisasi Data. Bandung: Informatika Bandung, 2017.
- [17] W. Budiharto, Machine Learning & Computational Intelligence. Yogyakarta: C. V ANDI OFFSET, 2016.
- [18] B. Gunawan Sudarsono, M. Ignatius Leo, A. Santoso, and F. Hendrawan, "ANALYSIS DATA MINING NETFLIX DATA USING THE RAPID MINER APPLICATION," Journal of Business and Audit Information Systems, vol. 4, pp. 13–21, 2021.
- [19] K. Yun, A. Huyen, and T. Lu, "DEEP NEURAL NETWORKS FOR PATTERN RECOGNITION."
- [20] N. R. Dzakiyullah, A. Pramuntadi, and A. K. Fauziyyah, "Semi-Supervised Classification on Credit Card Fraud Detection using AutoEncoders," Journal of Applied Data Sciences, vol. 2, no. 1, pp. 1–07, 2021