

Klasifikasi Penyakit Diabetes Melitus Berbasis Jaringan Syaraf Tiruan Menggunakan Algoritma *Multi-Layer Perceptron*

Jose Julian Hidayat^{1*}, Daffa Eka Sujianto², Muhammad Randika Saputra³,
Erik Ahmad Ramdhani⁴, Muhamad Jihansyah⁵, Yogi Nandiya⁶

¹Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa, Bekasi, Indonesia

²Teknik Informatika, Fakultas Teknik, Universitas Airlangga, Surabaya, Indonesia

³Teknik Informatika, Fakultas Teknik, Universitas Perjuangan Tasikmalaya, Tasikmalaya, Indonesia

⁴Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Terbuka, Jakarta, Indonesia

⁵Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Tangerang Selatan, Indonesia

⁶Teknik Informatika, Fakultas Teknik Dan Ilmu Komputer, Universitas Indraprasta PGRI, Jakarta Selatan, Indonesia

Email: ^{1*}josejulianhidayat@gmail.com, ²daffaakun82@gmail.com, ³saputrandika857@gmail.com,

⁴ramdhanierik48@gmail.com, ⁵jihansyahmuhamad1@gmail.com, ⁶yoginandiya15@gmail.com

(*Email Corresponding Author: josejulianhidayat@gmail.com)

Received: May 13, 2026 | Revision: May 14, 2026 | Accepted: May 16, 2026

Abstrak

Diabetes Melitus merupakan salah satu penyakit kronis yang membutuhkan proses deteksi dini secara cepat dan akurat agar risiko komplikasi dapat diminimalkan. Penelitian ini bertujuan untuk mengklasifikasikan penyakit Diabetes Melitus menggunakan Jaringan Syaraf Tiruan dengan algoritma *Multi-Layer Perceptron*. Dataset yang digunakan terdiri dari 100.000 data dengan 9 atribut, yaitu *gender*, *age*, *hypertension*, *heart disease*, *smoking history*, *body mass index*, *HbA1c level*, *blood glucose level*, dan *diabetes* sebagai target klasifikasi. Setelah dilakukan pengecekan data, tidak ditemukan missing value, namun terdapat 3.854 data duplikat sehingga jumlah data setelah pembersihan menjadi 96.146 data. Proses penelitian meliputi *preprocessing* data, *encoding* fitur kategorikal, standarisasi fitur numerik, pembagian data *training* dan *testing*, pemodelan *Multi-Layer Perceptron*, serta evaluasi performa model. Model yang digunakan memiliki beberapa *hidden layer* dengan aktivasi *ReLU*, *batch normalization*, *dropout*, dan *output sigmoid* untuk klasifikasi *biner*. Hasil pengujian menunjukkan bahwa model memperoleh *accuracy* sebesar 0,9715548622, *precision* sebesar 0,9939810834, *recall* sebesar 0,6816037736, *F1-score* sebesar 0,8086743617, dan *ROC-AUC* sebesar 0,9749626265. Berdasarkan hasil tersebut, algoritma *Multi-Layer Perceptron* mampu memberikan performa klasifikasi yang baik, khususnya dalam membedakan pasien non-diabetes dan diabetes berdasarkan atribut kesehatan yang tersedia.

Kata Kunci: Diabetes Melitus, Jaringan Syaraf Tiruan, Multi-Layer Perceptron, Klasifikasi, Machine Learning

Abstract

Diabetes Mellitus is a chronic disease that requires fast and accurate early detection to minimize the risk of complications. This study aims to classify Diabetes Mellitus using an Artificial Neural Network with the Multi-Layer Perceptron algorithm. The dataset consists of 100,000 records with 9 attributes: gender, age, hypertension, heart disease, smoking history, body mass index, HbA1c level, blood glucose level, and diabetes as the classification target. After data checking, no missing values were found; however, 3,854 duplicate records were identified, resulting in 96,146 records after data cleaning. The research process includes data preprocessing, categorical feature encoding, numerical feature standardization, training and testing data splitting, Multi-Layer Perceptron modeling, and model performance evaluation. The proposed model consists of several hidden layers using ReLU activation, batch normalization, dropout, and a sigmoid output layer for binary classification. The testing results show that the model achieved an accuracy of 0.9715548622, precision of 0.9939810834, recall of 0.6816037736, F1-score of 0.8086743617, and ROC-AUC of 0.9749626265. These results indicate that the Multi-Layer Perceptron algorithm provides good classification performance, particularly in distinguishing non-diabetic and diabetic patients based on the available health attributes.

Keywords: Diabetes Mellitus, Artificial Neural Network, Multi-Layer Perceptron, classification, machine learning.

1. PENDAHULUAN

Diabetes Melitus merupakan salah satu penyakit kronis yang menjadi perhatian utama di bidang kesehatan global karena prevalensinya terus meningkat setiap tahun [1]. Gangguan metabolisme glukosa menyebabkan kadar gula darah tetap di atas batas normal, yang menyebabkan penyakit ini [2]. Jika tidak ditangani dengan benar, kondisi ini dapat menyebabkan berbagai komplikasi serius seperti penyakit jantung, gagal ginjal, kerusakan saraf, hingga gangguan penglihatan [3]. Untuk mengurangi risiko komplikasi, sistem deteksi dini diperlukan karena banyaknya penderita diabetes.

Pembelajaran mesin dan kecerdasan buatan telah meningkatkan proses klasifikasi penyakit, termasuk diabetes melitus. Jaringan Syaraf Tiruan adalah teknik yang banyak digunakan karena dapat mendeteksi pola kompleks dalam data medis [4]. *Perceptron Multi-Layer (MLP)* mampu melakukan pembelajaran non-linear melalui berbagai lapisan tersembunyi, yang membuatnya menjadi pilihan yang populer [5]. Kemampuan ini memungkinkan model untuk menyelidiki bagaimana usia, indeks massa tubuh, kadar *HbA1c*, glukosa darah, hipertensi, dan riwayat penyakit jantung berkorelasi dengan risiko diabetes.

Studi sebelumnya, "Prediksi Penyakit Diabetes Menggunakan Algoritma *Support Vector Machine (SVM)*" ditulis pada tahun 2023, membahas penerapan algoritma *Support Vector Machine* untuk klasifikasi penyakit diabetes mellitus [6]. Jumlah penderita diabetes yang besar dan pentingnya sistem prediksi untuk membantu proses diagnosis yang lebih cepat dan akurat adalah alasan utama dari penelitian ini. Studi ini menunjukkan bahwa sebagian besar penelitian sebelumnya menggunakan metode seperti *C4.5*, *K-Nearest Neighbor*, *Naive Bayes*, dan *SVM Linier*. Namun, karena dianggap menghasilkan akurasi yang kurang optimal, *SVM* dengan kernel *Radial Basis Function (RBF)* masih jarang digunakan. Untuk menyelesaikan masalah tersebut, penelitian ini menggunakan algoritma *SVM Radial Basis Function* berbasis pilihan maju dan melakukan evaluasi menggunakan matriks ketidakpastian, kesalahan kuadrat rata-rata, dan validasi *cross-k-fold* kelipatan 10. Hasil penelitian menunjukkan bahwa algoritma *SVM RBF* dapat digunakan secara efektif untuk memprediksi penyakit diabetes.

Studi yang ditulis pada tahun 2025 yang berjudul " *Effects of Data Balancing in Diabetes Mellitus Detection: A Comparative XGBoost and Random Forest Learning Approach*" membahas bagaimana teknik *data balancing* memengaruhi seberapa baik klasifikasi diabetes mellitus menggunakan algoritma *XGBoost* dan *Random Forest*. Dataset *PIMA Indian Diabetes* yang memiliki ketidakseimbangan kelas digunakan dalam penelitian ini. Akibatnya, beberapa teknik *balancing* digunakan, termasuk *Random Under Sampling (RUS)*, *Upscaling*, *SMOTE*, *ADASyn*, *SMOTE-Tomek*, dan *SMOTEEN* [7]. Hasil penelitian menunjukkan bahwa metode *balancing* data dapat meningkatkan kinerja model klasifikasi dibandingkan dengan metode tanpa *balancing* data. Metode *SMOTE-Tomek* memiliki tingkat akurasi terbaik dengan nilai akurasi 0,8182 dan nilai *F1-score* 0,8189 pada algoritma *XGBoost*, sedangkan metode *SMOTEEN* memiliki tingkat akurasi terbaik dengan nilai akurasi 0,8199 dan nilai *F1-score* 0,8178 pada algoritma *Random Forest*. Penelitian tersebut menunjukkan bahwa peningkatan kemampuan model untuk mendeteksi diabetes secara signifikan dipengaruhi oleh penanganan ketidakseimbangan data.

Penelitian sebelumnya, dengan judul " *Evaluasi Kinerja Algoritma Ensemble Learning Pada Klasifikasi Penyakit Diabetes Berbasis Boosting Method*" membahas bagaimana metode pembelajaran kelompok berbasis penguatan diterapkan dalam klasifikasi penyakit diabetes. Studi tersebut menggunakan beberapa algoritma peningkatan, seperti *Gradient Boosting*, *Extreme Gradient Boosting (XGBoost)*, dan *CatBoost*, pada dataset 19.230 data yang memiliki dua kelas data, yaitu diabetes dan non-diabetes [8]. Untuk menilai performa masing-masing algoritma, evaluasi model dilakukan dengan metrik ketepatan, *recall*, skor F1, dan ketepatan. Hasil penelitian menunjukkan bahwa ketiga algoritma memiliki performa yang sangat baik dengan tingkat ketepatan sebesar 97%. Pada kelas diabetes, *Gradient Boosting* menghasilkan ketepatan sebesar 0,99, *recall* sebesar 0,69, dan skor F1-sebesar 0,81, sedangkan *XGBoost* menghasilkan ketepatan sebesar 0,97, *recall* sebesar 0,69, dan skor F1-sebesar 0,81. *CatBoost* menghasilkan ketepatan sebesar 0,98, *recall* sebesar 0,69, dan skor F1-sebesar 0,81. Penelitian tersebut menunjukkan bahwa metode pembelajaran kelompok berbasis peningkatan mampu melakukan klasifikasi yang baik pada penyakit diabetes. Namun, penelitian tersebut juga menemukan bahwa nilai *recall* pada kelas diabetes masih relatif rendah, sehingga kemampuan model untuk mengidentifikasi semua kasus positif belum optimal.

Penelitian sebelumnya, " *Analisis Hubungan Obesitas dan Diabetes Melitus Berdasarkan Usia dan Jenis Kelamin Menggunakan Algoritma Apriori*", yang dilakukan untuk membahas analisis hubungan antara obesitas dan diabetes melitus berdasarkan usia dan jenis kelamin menggunakan teknik *data mining association rule* [9]. Data sekunder yang digunakan dalam penelitian ini berasal dari Dinas Kesehatan Kota Makassar dari tahun 2023 hingga 2025. Data ini kemudian melalui tahap *preprocessing*, yang mencakup pembersihan, transformasi, dan diskretisasi data menggunakan teknik tertil. Algoritma *Apriori* digunakan untuk menemukan pola hubungan antar variabel dengan minimum dukungan 10% dan keyakinan 60%. Hasil penelitian menunjukkan korelasi yang signifikan antara obesitas dan diabetes melitus, terutama pada wanita dan kelompok usia lanjut. Setelah divalidasi menggunakan uji *Chi-Square*, kami menemukan aturan asosiasi dengan nilai lift ratio tertinggi sebesar 5,581, yang menunjukkan bahwa ada hubungan yang kuat antara variabel. Nilai p yang dihasilkan untuk hubungan tersebut adalah 0,05. Penelitian ini memberikan kontribusi dalam memahami faktor-faktor yang berhubungan dengan Diabetes Melitus, namun penelitian tersebut lebih berfokus pada analisis hubungan antar variabel menggunakan *association rule* dan belum melakukan klasifikasi penyakit diabetes secara otomatis menggunakan pendekatan *deep learning* seperti *Multi-Layer Perceptron* yang digunakan pada penelitian ini.

Gap penelitian lainnya terletak pada perbedaan fokus dan pendekatan yang digunakan dalam penelitian-penelitian sebelumnya terkait Diabetes Melitus. Beberapa penelitian hanya menggunakan metode pembelajaran mesin konvensional seperti *Support Vector Machine*, *XGBoost*, *Random Forest*, dan algoritma peningkatan, tetapi tidak menggunakan kemampuan pembelajaran mendalam untuk mempelajari pola non-linear yang lebih kompleks pada data kesehatan. Kemudian, algoritma *Apriori* belum melakukan klasifikasi penyakit diabetes secara otomatis, penelitian lain lebih fokus pada analisis hubungan antar variabel. Selain itu, beberapa penelitian masih memiliki keterbatasan pada tahap *preprocessing* data, seperti belum menerapkan kombinasi *encoding* fitur kategorikal, standarisasi fitur numerik, dan regularisasi model secara optimal. Sementara metrik penting lainnya seperti ketepatan, *recall*, skor F1, dan *ROC-AUC* belum dianalisis secara menyeluruh, evaluasi model pada penelitian sebelumnya biasanya berpusat pada nilai ketepatan. Oleh karena itu, penelitian yang mampu mengintegrasikan *preprocessing* data yang lebih komprehensif dengan struktur *Multi-Layer Perceptron* yang lebih stabil dan mendalam diperlukan untuk meningkatkan kinerja klasifikasi Diabetes Melitus secara lebih efisien.

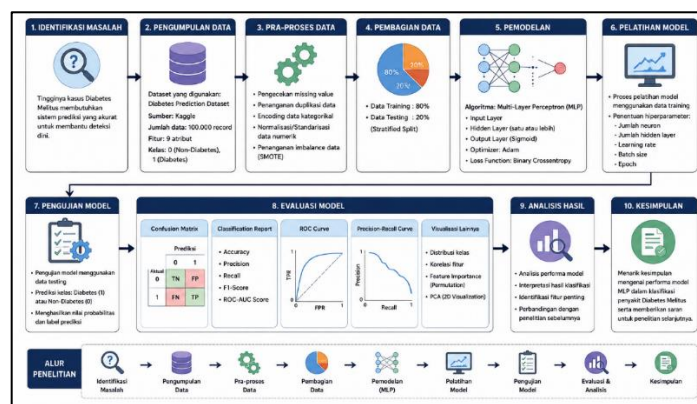
Novelty atau kebaruan pada penelitian ini terletak pada penerapan arsitektur *Multi-Layer Perceptron* yang dikombinasikan dengan *batch normalization* dan *dropout* pada dataset Diabetes Prediction Dataset berukuran besar

dengan 96.146 data setelah proses pembersihan data. Selain itu, penelitian ini tidak hanya mengevaluasi model menggunakan akurasi, tetapi juga menggunakan *precision*, *recall*, *F1-score*, *ROC-AUC*, *threshold tuning*, serta analisis kesalahan prediksi untuk memperoleh evaluasi performa yang lebih komprehensif. Penelitian ini juga melakukan analisis *feature importance* untuk mengetahui atribut kesehatan yang paling berpengaruh terhadap klasifikasi Diabetes Melitus.

Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem klasifikasi penyakit Diabetes Melitus berbasis kecerdasan buatan yang lebih akurat dan efektif serta dapat menjadi referensi untuk penelitian selanjutnya di bidang machine learning dan kesehatan digital.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk melakukan klasifikasi penyakit Diabetes Melitus menggunakan algoritma *Multi-Layer Perceptron* berbasis Jaringan Syaraf Tiruan. Proses penelitian dimulai dengan pengumpulan dataset, *preprocessing* data, pembuatan model, pelatihan model, evaluasi performa, dan analisis hasil klasifikasi. Setiap langkah penelitian dilakukan menggunakan bahasa pemrograman *Python* dengan bantuan kumpulan pembelajaran mesin dan pembelajaran mendalam seperti *TensorFlow*, *Keras*, *Scikit-learn*, *Pandas*, *NumPy*, *Matplotlib*, dan *Seaborn*.



Gambar 1. Alur dan Desain Penelitian

2.1 Pengumpulan Data

Dataset Diabetes Prediction yang dipublikasikan oleh Mohammed Mustafa adalah sumber data yang digunakan dalam penelitian ini. Data ini diperoleh dari platform Kaggle yang dapat diakses pada laman <https://www.kaggle.com/datasets/iammustafatz/diabetes-prediction-dataset>. Berdasarkan riwayat kesehatan dan kondisi tubuh pasien, kumpulan data medis dan demografis ini digunakan untuk memprediksi kemungkinan mereka menderita Diabetes Melitus. Dataset terdiri dari 100.000 data yang diklasifikasikan berdasarkan 9 atribut: *gender*, usia, hipertensi, penyakit jantung, riwayat rokok, berat badan, tingkat *HbA1c*, tingkat glukosa darah, dan diabetes. Atribut *gender* menunjukkan jenis kelamin pasien, sedangkan *age* menunjukkan usia pasien. Atribut *hypertension* dan *heart_disease* menunjukkan riwayat hipertensi dan penyakit jantung yang dimiliki pasien. Atribut riwayat merokok menunjukkan sejarah kebiasaan merokok pasien. Selain itu, nilai bmi menunjukkan indeks massa tubuh, nilai *HbA1c* menunjukkan kadar *hemoglobin* terglykasi dalam darah, dan nilai glukosa darah menunjukkan kadar glukosa darah pasien. Variabel target diabetes terdiri dari dua kelas: nilai 0 untuk orang yang tidak memiliki diabetes dan nilai 1 untuk orang yang memiliki diabetes. Karena memiliki kombinasi fitur medis dan demografis yang relevan untuk proses klasifikasi penyakit diabetes, dataset ini banyak digunakan dalam penelitian *machine learning* dan *deep learning*.

2.2 Preprocessing Data

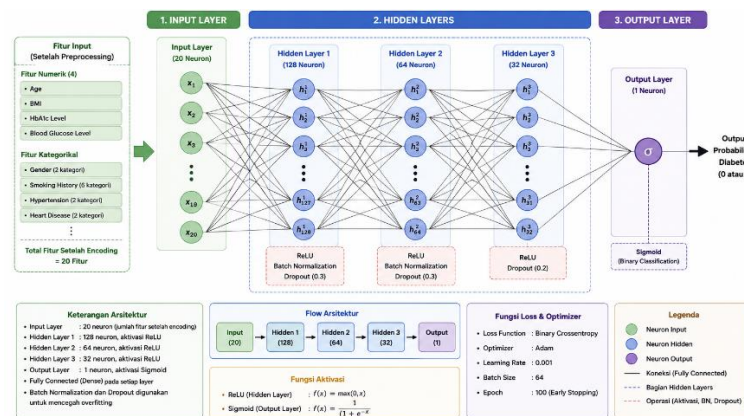
Sebelum data digunakan dalam proses pelatihan model *Multi-Layer Perceptron*, tahap *preprocessing* dilakukan untuk memaksimalkan kualitas data dan memenuhi kebutuhan algoritma *deep learning* [10]. Pada tahap ini, fitur numerik dan kategorikal dipisahkan [11]. Usia, hipertensi, penyakit jantung, bmi, tingkat glukosa darah, dan *HbA1c* adalah fitur numerik, sedangkan *gender* dan riwayat rokok adalah fitur kategorikal. Fitur kategorikal kemudian diubah menjadi data numerik menggunakan metode *One Hot Encoding*, sehingga setiap kategori dapat diproses oleh model dalam bentuk *binner* [12]. Selanjutnya, fitur numerik distandardisasi menggunakan metode *StandardScaler* untuk menyamakan rentang nilai masing-masing fitur [13]. Ini dilakukan agar proses pelatihan model menjadi lebih stabil dan tidak didominasi oleh fitur tertentu dengan skala yang lebih besar [14]. Setelah seluruh proses *preprocessing* selesai, jumlah fitur naik dari 8 fitur input menjadi 15 fitur hasil transformasi. Fitur-fitur ini kemudian digunakan sebagai data masukan untuk proses pelatihan dan pengujian model klasifikasi Diabetes Melitus menggunakan algoritma *Multi-Layer Perceptron*.

2.3 Pembagian Data Training dan Testing

Dataset yang telah diproses dibagi menjadi data *training* dan data *testing* menggunakan metode *train-test split* [15]. Sebanyak 80% data digunakan sebagai data *training* dan 20% data digunakan sebagai data *testing* [16]. Pembagian data dilakukan secara acak menggunakan parameter *random_state* untuk menjaga konsistensi hasil penelitian [17].

2.4 Perancangan dan Pelatihan Model Arsitektur Multi-Layer Perceptron

Model yang digunakan pada penelitian ini adalah *Multi-Layer Perceptron* berbasis Jaringan Syaraf Tiruan. Arsitektur model terdiri dari beberapa *hidden layer* dengan jumlah *neuron* sebanyak 256, 128, 64, dan 32 *neuron*. Fungsi aktivasi yang digunakan pada *hidden layer* adalah *ReLU*, sedangkan *output layer* menggunakan fungsi aktivasi *sigmoid* untuk klasifikasi biner [18]. Penelitian ini juga menerapkan *Batch Normalization* dan *Dropout* untuk meningkatkan stabilitas model dan mengurangi risiko *overfitting* [19].



Gambar 2. Perancangan Arsitektur Multi-Layer Perceptron Berbasis Jaringan Syaraf Tiruan

Proses pelatihan model dilakukan menggunakan *optimizer* Adam dengan *learning rate* awal sebesar 0,001 dan fungsi *loss binary crossentropy* [20]. Model dilatih selama maksimal *100 epoch* dengan ukuran *batch* sebesar 64. Untuk meningkatkan performa pelatihan, digunakan teknik *Early Stopping* dan *Reduce Learning Rate on Plateau* agar model dapat berhenti secara otomatis ketika performa tidak mengalami peningkatan.

2.5 Evaluasi Performa Model

Evaluasi model dilakukan menggunakan beberapa metrik evaluasi yaitu *accuracy*, *precision*, *recall*, *F1-score*, dan *ROC-AUC* [21]. Selain itu, penelitian ini juga menggunakan Confusion Matrix, ROC Curve, Precision-Recall Curve, Calibration Curve, dan Threshold Tuning untuk melakukan analisis performa model secara lebih komprehensif [22].

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

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

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

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

2.6 Analisis Hasil dan Visualisasi

Tahap akhir penelitian dilakukan dengan menganalisis hasil prediksi model melalui berbagai visualisasi data seperti distribusi probabilitas prediksi, analisis kesalahan prediksi (*error analysis*), analisis *confidence* model, serta analisis feature importance untuk mengetahui fitur yang paling berpengaruh terhadap klasifikasi Diabetes Melitus. Seluruh hasil pengujian divisualisasikan dalam bentuk grafik dan tabel agar lebih mudah dipahami dan diinterpretasikan.

3. HASIL DAN PEMBAHASAN

Dataset prediksi diabetes digunakan dalam penelitian ini. Dataset ini diperoleh dari platform *Kaggle* dan digunakan untuk mengklasifikasikan penyakit diabetes melitus berdasarkan data kesehatan dan karakteristik pasien. Dataset awal terdiri dari 100.000 data, dengan klasifikasi target dari 9 atribut: gender, usia, hipertensi, penyakit jantung,

riwayat rokok, bmi, tingkat *HbA1c*, tingkat glukosa darah, dan diabetes. Variabel target terdiri dari dua kelas, dengan nilai 0 untuk non-diabetes dan nilai 1 untuk diabetes.

Tabel 1. Dataset Penelitian

Age	Gender	BMI	HbA1c level	...	Diabetes
80	Female	25.19	6.6	...	0
54	Male	27.32	6.6	...	0
28	Female	27.32	5.7	...	0
36	Male	23.45	5.0	...	0
76	Female	20.14	4.8	...	0
44	Male	19.31	6.5	...	1
67	Female	32.45	7.2	...	1
29	Male	24.89	5.4	...	0
58	Female	30.11	6.8	...	1
....
62	Male	28.76	6.9	...	1

Hasil eksplorasi data menunjukkan bahwa dataset tidak memiliki nilai yang hilang pada semua fiturnya. Oleh karena itu, semua data dapat digunakan dalam proses penelitian. Namun, setelah ditemukan 3.854 data duplikat, dilakukan pembersihan data untuk menghilangkan data tersebut. Setelah proses ini selesai, jumlah data yang digunakan dalam penelitian meningkat menjadi 96.146 data, dan proses pembersihan data dilakukan dengan tujuan meningkatkan kualitas dataset dan mengurangi kemungkinan bias dalam proses pelatihan model.

Tabel 2. Missing Value Table

Atribut	Missing Value
gender	0
age	0
hypertension	0
heart_disease	0
smoking_history	0
bmi	0
HbA1c_level	0
blood_glucose_level	0
diabetes	0

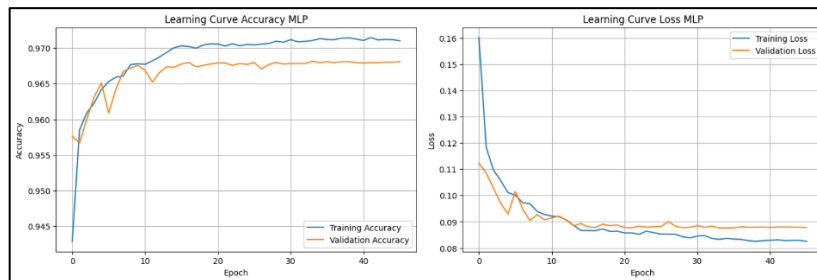
Berdasarkan hasil proses pelatihan model *Multi-Layer Perceptron* selama 46 *epoch*, performa model mengalami peningkatan yang cukup stabil pada setiap iterasi pelatihan. Nilai *accuracy training* meningkat dari 94,28% pada *epoch* pertama menjadi sekitar 97,10% pada *epoch* terakhir, sedangkan *validation accuracy* meningkat hingga mencapai 96,81%. Selain itu, nilai *loss* dan *validation loss* mengalami penurunan secara bertahap yang menunjukkan bahwa model mampu mempelajari pola data dengan baik selama proses *training* berlangsung. Penggunaan teknik *Reduce Learning Rate on Plateau* menyebabkan nilai *learning rate* menurun secara bertahap dari 0,001 menjadi 0,000031 untuk membantu model mencapai konvergensi yang lebih stabil. Hasil tersebut menunjukkan bahwa model *Multi-Layer Perceptron* memiliki kemampuan pembelajaran yang baik dan tidak mengalami *overfitting* yang signifikan karena nilai *accuracy training* dan *validation accuracy* memiliki selisih yang relatif kecil.

Tabel 3. Hasil Proses Pelatihan Model Multi-Layer Perceptron Selama 46 Epoch

Epoch	Accuracy	Loss	Val Accuracy	Val Loss	Learning Rate
0	0.942891	0.160137	0.957618	0.112298	0.001000
1	0.958509	0.118357	0.956708	0.108617	0.001000
2	0.960915	0.109613	0.960023	0.102704	0.001000
3	0.962280	0.105601	0.963144	0.097000	0.001000
4	0.964181	0.101109	0.965094	0.092946	0.001000
5	0.965286	0.100128	0.960868	0.101562	0.001000
6	0.965904	0.097233	0.964249	0.094815	0.001000
7	0.966083	0.096867	0.966719	0.090500	0.001000
8	0.967675	0.093929	0.967174	0.092707	0.001000
9	0.967789	0.092767	0.967564	0.090729	0.001000
10	0.967708	0.092167	0.966849	0.091489	0.001000
11	0.968195	0.091919	0.965224	0.092157	0.001000
12	0.968748	0.090681	0.966589	0.090571	0.001000

13	0.969398	0.088624	0.967369	0.088512	0.000500
14	0.970048	0.086691	0.967304	0.089293	0.000500
15	0.970292	0.086656	0.967759	0.088038	0.000500
16	0.970211	0.086597	0.967954	0.087810	0.000500
17	0.969967	0.087270	0.967369	0.089061	0.000500
18	0.970422	0.086300	0.967564	0.088543	0.000500
19	0.970568	0.086391	0.967759	0.088744	0.000500
20	0.970552	0.085702	0.967889	0.087774	0.000500
...
44	0.971169	0.082896	0.968019	0.087885	0.000031
45	0.971007	0.082519	0.968084	0.087813	0.000031

Berdasarkan grafik *Learning Curve Accuracy MLP* pada gambar 3, terlihat bahwa nilai keakuratan pelatihan dan validasi meningkat seiring waktu pelatihan. Pada awal pelatihan, ketepatan model masih berada di sekitar 94%. Setelah beberapa periode, ia kemudian meningkat dan stabil di sekitar 97%. Selain itu, kurva ketepatan pelatihan dan validasi memiliki pola yang hampir sama, yang menunjukkan bahwa model dapat digeneralisasi dengan baik dan tidak mengalami *overfitting* yang signifikan. Selain itu, ada fluktuasi kecil dalam akurasi validasi, yang menunjukkan bahwa kinerja model cukup stabil selama pelatihan. Sementara itu, grafik *Learning Curve Loss MLP* menunjukkan bahwa nilai kehilangan pengujian dan kehilangan pengajaran mengalami penurunan secara konsisten selama proses pengajaran. Pada awal pelatihan, kehilangan pelatihan berada pada nilai sekitar 0,16 dan kemudian menurun hingga mendekati 0,08 pada interval terakhir. Kehilangan validasi juga menunjukkan pola penurunan yang konsisten dan berada di sekitar 0,08 hingga 0,09. Tidak ada *underfitting* atau *overfitting* pada model ini, karena jarak antara kehilangan pembelajaran dan kehilangan validasi sangat kecil. Hasil menunjukkan kemampuan model *Multi-Layer Perceptron* untuk mempelajari pola data dengan baik dan menghasilkan kinerja klasifikasi yang stabil baik untuk instruksi maupun validasi data.



Gambar 3. Hasil Pelatihan Model

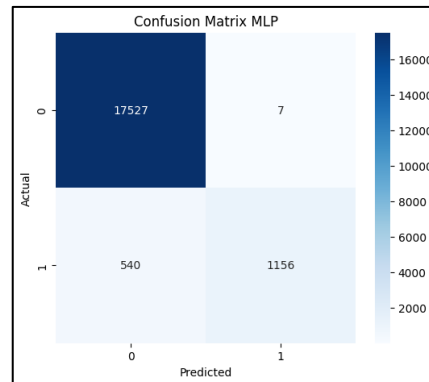
Hasil evaluasi model *Multi-Layer Perceptron* menunjukkan nilai ketepatan sebesar 97,15%. Ini menunjukkan bahwa model mampu mengklasifikasikan penyakit Diabetes Melitus dengan sangat tepat. Dengan nilai *precision* sebesar 99,39%, sebagian besar data yang diprediksi sebagai diabetes adalah pasien diabetes. Selain itu, berdasarkan nilai *recall* sebesar 68,16%, model dapat menemukan sekitar 68% kasus diabetes positif dari semua data pasien diabetes yang tersedia. Dalam proses klasifikasi, keseimbangan yang cukup baik antara ketepatan dan *recall* ditunjukkan oleh skor F1-nya sebesar 80,86%. Sebaliknya, nilai *ROC-AUC* sebesar 97,49% menunjukkan bahwa model memiliki kemampuan yang luar biasa untuk membedakan antara orang yang menderita diabetes dan orang yang tidak menderita diabetes. Berdasarkan laporan klasifikasi, kinerja model pada kelas non-diabetes (kelas 0) menunjukkan nilai *precision* sebesar 97,01%, *recall* sebesar 99,96%, dan skor F1-sebesar 98,46%. Hasil menunjukkan bahwa model sangat baik dalam mengenali pasien non-diabetes. Pada kelas diabetes (kelas 1), kinerja model menunjukkan nilai *precision* sebesar 99,39%, *recall* sebesar 68,16%, dan skor F1-sebesar 80,86%. Nilai *precision* yang sangat tinggi menunjukkan bahwa prediksi pasien diabetes Secara keseluruhan, hasil evaluasi menunjukkan bahwa algoritma *Multi-Layer Perceptron* dapat mendeteksi penyakit Diabetes Melitus dengan hasil klasifikasi yang baik dan stabil berdasarkan data kesehatan pasien.

Tabel 4. *Multi-Layer Perceptron Classification Report*

Class	Precision	Recall	F1-Score	Support
0	0.9701112526	0.9996007756	0.9846352631	17534
1	0.9939810834	0.6816037736	0.8086743617	1696
Accuracy	-	-	0.9715548622	19230
Macro Avg	0.9820461680	0.8406022746	0.8966548124	19230
Weighted Avg	0.9722164649	0.9715548622	0.9691162985	19230

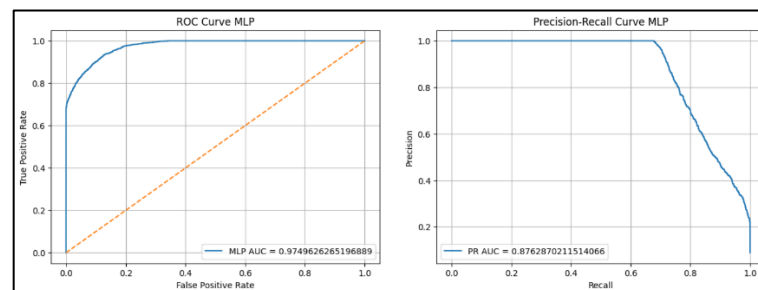
Berdasarkan hasil *Confusion Matrix* pada gambar 4, model *Multi-Layer Perceptron* berhasil mengklasifikasikan 17.527 data non-diabetes dengan benar dan hanya melakukan kesalahan prediksi sebanyak 7 data non-diabetes yang diprediksi sebagai diabetes. Pada kelas diabetes, model mendeteksi 1.156 data diabetes dengan benar, tetapi masih ada

540 data diabetes yang diprediksi sebagai non-diabetes. Hasil ini menunjukkan kemampuan model yang sangat baik untuk mengklasifikasikan data non-diabetes. Secara keseluruhan, *confusion matrix* menunjukkan bahwa model mampu melakukan klasifikasi yang baik dengan data kesehatan yang digunakan untuk membedakan pasien diabetes dan non-diabetes.



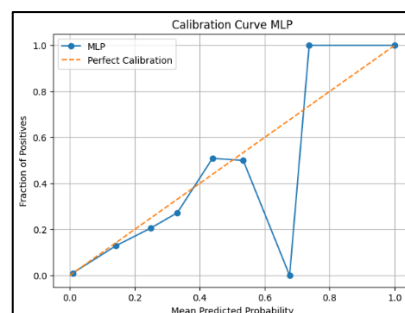
Gambar 4. *Confusion Matrix*

Ditunjukkan oleh grafik *ROC Curve MLP*, kurva *ROC* berada jauh di atas garis diagonal, menunjukkan bahwa model *Perceptron Multilayer* memiliki kemampuan yang luar biasa untuk membedakan antara kelas diabetes dan non-diabetes. Seperti yang ditunjukkan oleh nilai *ROC-AUC* sebesar 0,9749, model menunjukkan tingkat diskriminasi yang sangat tinggi dalam proses klasifikasi. Semakin mendekati nilai 1, tingkat kemampuan model untuk membedakan dua kelas semakin baik. Hasil menunjukkan bahwa model dapat membuat prediksi yang akurat dan stabil dari kumpulan data tentang diabetes melitus. Di sisi lain, grafik *Curvature Precision-Recall MLP* menunjukkan hubungan antara nilai *precision* dan *recall* pada berbagai ambang klasifikasi. Kurva menunjukkan bahwa model mampu mempertahankan nilai *precision* tinggi pada sebagian besar nilai *recall*. Meskipun kumpulan data tidak seimbang, model mendeteksi kelas positif diabetes dengan baik, seperti yang ditunjukkan oleh nilai *PR-AUC* sebesar 0,8762. Hasil menunjukkan bahwa model *Multi-Layer Perceptron* dapat menemukan pasien diabetes dengan tingkat prediksi yang tinggi.



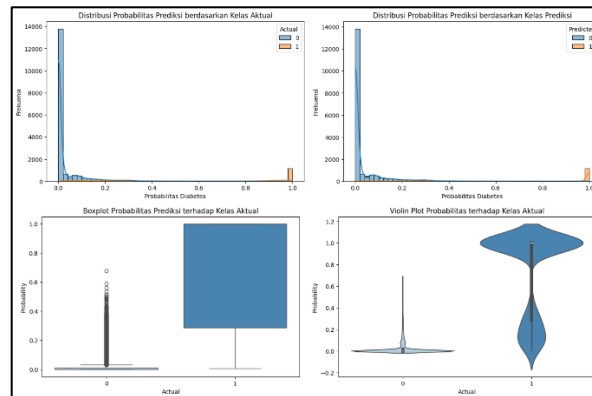
Gambar 5. *ROC Curve dan Curvature Precision-Recall MLP*

Berdasarkan grafik *Curvature Calibration MLP* pada gambar 6, sebagian besar titik prediksi model berada di dekat garis *Calibration Perfect*. Semakin dekat kurva model dengan garis diagonal, semakin baik kalibrasi model dalam memperkirakan probabilitas suatu kelas. Ini menunjukkan bahwa probabilitas prediksi yang dihasilkan oleh model *Multi-Layer Perceptron* cukup sesuai dengan kondisi data aktual. Secara keseluruhan, model menunjukkan kemampuan kalibrasi yang cukup baik untuk memprediksi kemungkinan pasien menderita Diabetes Melitus, meskipun ada beberapa titik yang berbeda dalam probabilitas. Hasil ini menunjukkan bahwa model tidak hanya memiliki kinerja klasifikasi yang baik, tetapi juga dapat menghasilkan probabilitas prediksi yang stabil dan akurat.



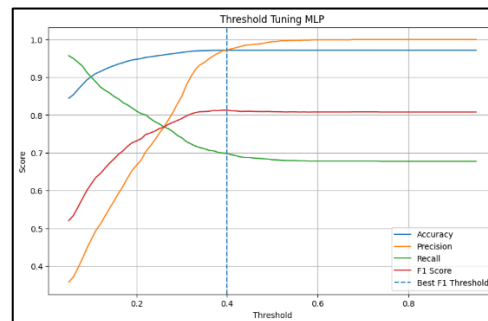
Gambar 6. *Calibration Curve MLP*

Berdasarkan visualisasi distribusi probabilitas prediksi pada gambar 7, terlihat bahwa model *Multi-Layer Perceptron* mampu memisahkan kelas diabetes dan non-diabetes dengan cukup baik. Sebagian besar data non-diabetes memiliki probabilitas prediksi yang sangat rendah dan mendekati nilai 0, sedangkan data diabetes cenderung memiliki probabilitas tinggi mendekati nilai 1. Hasil histogram menunjukkan bahwa mayoritas prediksi model sesuai dengan kelas aktual sehingga menunjukkan performa klasifikasi yang stabil. Selain itu, *boxplot* dan *violin plot* memperlihatkan perbedaan distribusi probabilitas yang cukup jelas antara kedua kelas, di mana kelas non-diabetes terkonsentrasi pada probabilitas rendah sedangkan kelas diabetes memiliki distribusi probabilitas yang lebih tinggi. Hasil tersebut menunjukkan bahwa model mampu memberikan tingkat keyakinan prediksi yang baik dalam membedakan pasien diabetes dan non-diabetes berdasarkan data kesehatan yang digunakan.



Gambar 7. Distribusi Probabilitas

Berdasarkan grafik *Threshold Tuning MLP* gambar 8, terlihat bahwa perubahan nilai *threshold* memberikan pengaruh terhadap performa model pada setiap metrik evaluasi. Semakin tinggi nilai *threshold*, nilai *precision* cenderung meningkat, sedangkan nilai *recall* mengalami penurunan. Hal tersebut menunjukkan bahwa model menjadi lebih selektif dalam memprediksi kelas diabetes sehingga jumlah prediksi positif yang salah semakin sedikit, namun beberapa kasus diabetes menjadi tidak terdeteksi.



Gambar 8. *Threshold Tuning MLP*

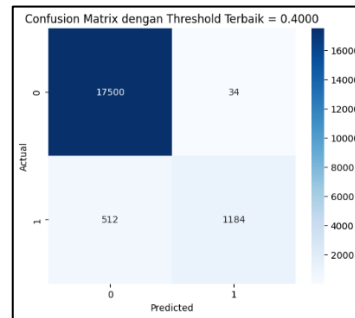
Pada gambar 8 juga menggambarkan nilai *F1-score* mengalami peningkatan hingga mencapai titik optimal pada *threshold* sekitar 0,40 yang ditunjukkan oleh garis putus-putus biru. Pada *threshold* tersebut diperoleh keseimbangan terbaik antara *precision* dan *recall* sehingga menghasilkan performa klasifikasi yang lebih optimal. Hasil ini menunjukkan bahwa pemilihan *threshold* yang tepat sangat berpengaruh terhadap kemampuan model dalam mendeteksi penyakit Diabetes Melitus secara lebih efektif dan seimbang.

Tabel 5. *Classification Report* pada titik optimal di threshold sekitar 0,40

Class	Precision	Recall	F1-Score	Support
0	0.9715745059	0.9980609102	0.9846396219	17534
1	0.9720853859	0.6981132075	0.8126286891	1696
Accuracy	-	-	0.9716068643	19230
Macro Avg	0.9718299459	0.8480870589	0.8986341555	19230
Weighted Avg	0.9716195632	0.9716068643	0.9694690269	19230

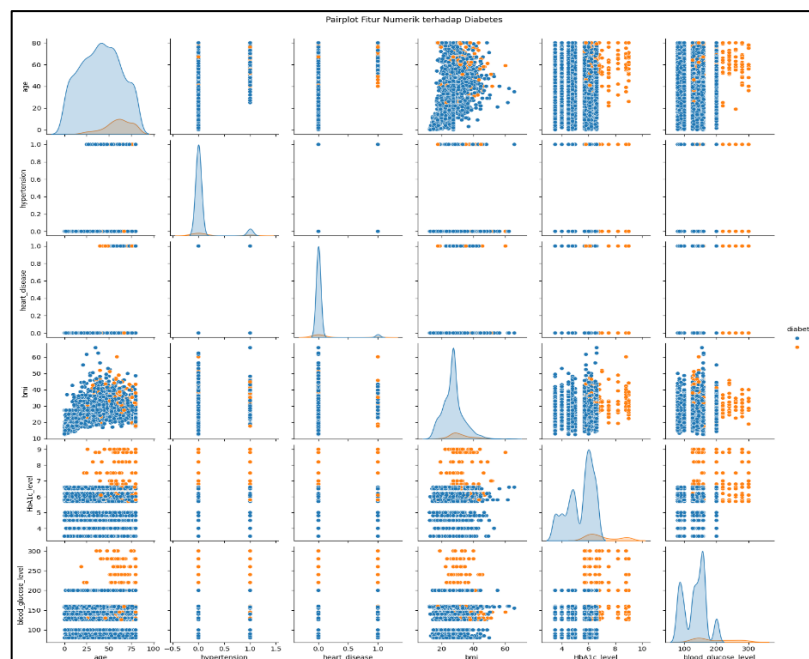
Berdasarkan hasil Laporan Klasifikasi pada *threshold* terbaik sebesar 0,40, model *Multi-Layer Perceptron* memperoleh ketepatan sebesar 97,16%, recall sebesar 99,80%, dan skor F1 sebesar 98,46% pada kelas non-diabetes (kelas 0). Sementara itu, pada kelas diabetes (kelas 1), model memperoleh ketepatan sebesar 97,15%, dan recall sebesar 99,80%,

yang menunjukkan bahwa model memiliki kemampuan yang luar biasa untuk mengenali pasien diabetes yang divalidasi dengan *confusion matrix* pada gambar 9.



Gambar 9. *Confusion Matrix Threshold Terbaik*

Berdasarkan visualisasi *pairplot* fitur numerik terhadap Diabetes Melitus pada gambar 10, terlihat bahwa beberapa atribut memiliki hubungan yang cukup kuat terhadap kelas diabetes. Fitur *HbA1c_level* dan *blood_glucose_level* menunjukkan perbedaan distribusi yang cukup jelas antara pasien diabetes dan non-diabetes, di mana pasien diabetes cenderung memiliki nilai yang lebih tinggi dibandingkan pasien non-diabetes. Selain itu, atribut *age* dan *bmi* juga menunjukkan bahwa pasien dengan usia lebih tua dan nilai indeks massa tubuh yang lebih tinggi memiliki kecenderungan lebih besar mengalami diabetes. Pada atribut *hypertension* dan *heart_disease*, terlihat bahwa pasien dengan riwayat hipertensi dan penyakit jantung lebih banyak ditemukan pada kelas diabetes. Secara keseluruhan, visualisasi *pairplot* menunjukkan bahwa kombinasi beberapa fitur kesehatan memiliki pengaruh terhadap klasifikasi Diabetes Melitus dan mampu membantu model *Multi-Layer Perceptron* dalam mempelajari pola data secara lebih efektif.



Gambar 10. *Pairplot Fitur Numerik Terhadap Diabetes*

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, algoritma *Multi-Layer Perceptron* berbasis Jaringan Syaraf Tiruan mampu digunakan untuk melakukan klasifikasi penyakit Diabetes Melitus dengan performa yang baik. Dataset yang digunakan terdiri dari 100.000 data dengan 9 atribut, kemudian setelah proses pembersihan data diperoleh 96.146 data yang digunakan dalam penelitian. Tahap *preprocessing* dilakukan menggunakan *One Hot Encoding* untuk fitur kategorikal dan *StandardScaler* untuk fitur numerik sehingga menghasilkan 15 fitur yang digunakan pada proses pelatihan model. Model *Multi-Layer Perceptron* yang dikembangkan menggunakan beberapa *hidden layer*, fungsi aktivasi *ReLU*, *batch normalization*, *dropout*, dan *optimizer Adam*. Berdasarkan hasil evaluasi, model memperoleh *accuracy* sebesar 97,15%, *precision* sebesar 99,39%, *recall* sebesar 68,16%, *F1-score* sebesar 80,86%, dan *ROC-AUC* sebesar 97,49%. Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam membedakan pasien diabetes dan non-diabetes berdasarkan atribut kesehatan yang tersedia. Berdasarkan hasil *confusion matrix* dan berbagai visualisasi

evaluasi seperti *ROC Curve*, *Precision-Recall Curve*, *Calibration Curve*, serta distribusi probabilitas prediksi, model menunjukkan performa klasifikasi yang stabil dan mampu mempelajari pola data dengan baik. Selain itu, hasil *threshold* tuning menunjukkan bahwa penyesuaian *threshold* dapat meningkatkan keseimbangan antara *precision* dan *recall* sehingga performa model menjadi lebih optimal dalam mendeteksi kasus diabetes. Secara keseluruhan, penelitian ini membuktikan bahwa algoritma *Multi-Layer Perceptron* dapat digunakan secara efektif untuk klasifikasi penyakit Diabetes Melitus dan berpotensi menjadi pendukung sistem deteksi dini penyakit diabetes berbasis kecerdasan buatan di bidang kesehatan.

5. SARAN

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat digunakan untuk pengembangan penelitian selanjutnya. Penelitian berikutnya dapat menggunakan teknik penanganan data tidak seimbang seperti *SMOTE*, *ADASYN*, atau metode *balancing* lainnya untuk meningkatkan kemampuan model dalam mendeteksi kelas diabetes sehingga nilai *recall* dapat menjadi lebih baik. Selain itu, penelitian selanjutnya dapat membandingkan performa *Multi-Layer Perceptron* dengan algoritma *deep learning* maupun *machine learning* lainnya seperti *Convolutional Neural Network*, *Long Short-Term Memory*, *XGBoost*, *Random Forest*, atau *ensemble learning* untuk memperoleh model klasifikasi yang lebih optimal. Penggunaan metode optimasi *hyperparameter* juga dapat dilakukan untuk meningkatkan performa model secara lebih maksimal. Penelitian berikutnya juga dapat menggunakan dataset yang lebih besar dan lebih beragam agar model memiliki kemampuan generalisasi yang lebih baik. Selain itu, penambahan fitur kesehatan lain seperti riwayat keluarga, pola aktivitas fisik, pola makan, dan tekanan darah dapat membantu meningkatkan kualitas prediksi model dalam mendeteksi penyakit Diabetes Melitus secara lebih akurat.

REFERENCES

- [1] N. K. Y. Yuniari, "Hubungan Motivasi Diri Dengan Kepatuhan Diet Pada Penderita Diabetes Melitus Tipe 2," *J. Penelit. Keperawatan Kontemporer*, vol. 6, no. 2 SE-Articles, pp. 158–169, Apr. 2026, doi: 10.59894/jpkk.v6i2.1293.
- [2] A. Setianto, L. Maria, and A. Firdaus, "Analisis Faktor Yang Mempengaruhi Kestabilan Gula Darah Penderita Diabetes Mellitus Pada Usia Dewasa Dan Lansia," *J. Ilm. Kesehat. Media Husada*, vol. 12, no. 2 SE-Articles, Nov. 2023, doi: 10.33475/jikmh.v12i2.334.
- [3] E. Erdaliza, M. Mitra, N. Rany, Y. Harnani, and A. Rienarti Abidin, "Faktor risiko yang berhubungan dengan komplikasi Diabetes Mellitus Tipe 2," *J. Kesehat. komunitas (Journal community Heal.*, vol. 10, no. 3, pp. 534–545, Dec. 2024, doi: 10.25311/keskom.Vol10.Iss3.2039.
- [4] R. M. Simanjourang, Amran Sitohang, Abdi Sembiring, and Sartika Simanjourang, "Implementasi Jaringan Syaraf Tiruan Dalam Klasifikasi Citra Medis," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 24, no. 1, pp. 64–72, Feb. 2025, doi: 10.53513/jis.v24i1.10690.
- [5] D. Arifuddin, K. Kusnari, and K. Kusnari, "Perbandingan Performansi Algoritma Multiple Linear Regression dan Multi Layer Perceptron Neural Network dalam Memprediksi Penjualan Obat: Comparison of the Performance of Multiple Linear Regression Algorithms and Multi Layer Perceptron Neural Networks in ," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 2 SE-, pp. 722–737, Apr. 2025, doi: 10.57152/malcom.v5i2.1952.
- [6] H. S. W. Hovi, A. Id Hadiana, and F. Rakhmat Umbara, "Prediksi Penyakit Diabetes Menggunakan Algoritma Support Vector Machine (SVM)," *Informatics Digit. Expert*, vol. 4, no. 1 SE-Articles, pp. 40–45, Jan. 2023, doi: 10.36423/index.v4i1.895.
- [7] F. O. Aghware *et al.*, "Effects of Data Balancing in Diabetes Mellitus Detection: A Comparative XGBoost and Random Forest Learning Approach," *NIPES J. Sci. Technol. Res.*, vol. 7, no. 1, pp. 1–11, Mar. 2025, doi: 10.37933/nipes/7.1.2025.1.
- [8] J. J. Hidayat, M. R. Saputra, A. R. Sigand, A. L. N. Fadilah, M. D. I. Amin, and A. R. Ramadhan, "Evaluasi Kinerja Algoritma Ensemble Learning Pada Klasifikasi Penyakit Diabetes Berbasis Boosting Method," *J. Surya Inform.*, vol. 16, no. 1 SE-Articles, pp. 71–80, May 2026, doi: 10.48144/suryainformatika.v16i1.2424.
- [9] E. Y. Kotte, F. I. Rachman, M. Faisal, and T. Wahyuni, "Analisis Hubungan Obesitas dan Diabetes Melitus Berdasarkan Usia dan Jenis Kelamin Menggunakan Algoritma Apriori," *Arus J. Sains dan Teknol.*, vol. 4, no. 1 SE-Artikel, pp. 71–80, Apr. 2026, doi: 10.57250/ajst.v4i1.2556.
- [10] A. Ristyawan, A. Nugroho, and T. K. Amarya, "Optimasi Preprocessing Model Random Forest untuk Prediksi Stroke," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 12, no. 1, Mar. 2025, doi: 10.35957/jatisi.v12i1.9587.
- [11] A. Candra, M. Erkamim, M. Muharrom, and E. Prayitno, "Klasifikasi Stunting Pada Balita Berdasarkan Status Gizi Menggunakan Pendekatan Support Vector Machine (SVM)," *J. Ilm. Fiso*, vol. 16, no. 2, pp. 171–181, 2024.
- [12] C. Herdian, A. Kamila, and I. G. Agung Musa Budidarma, "Studi Kasus Feature Engineering Untuk Data Teks: Perbandingan Label Encoding dan One-Hot Encoding Pada Metode Linear Regresi," *Technol. J. Ilm.*, vol. 15, no. 1, p. 93, Jan. 2024, doi: 10.31602/tji.v15i1.13457.
- [13] K. S. Arlandy, A. Faqih, and A. R. Rinaldi, "Mengoptimalkan Kinerja Naïve Bayes Pada Ancaman Modern Dengan Menggunakan PCA Pada Data Intrusion Detection System (IDS)," *J. Ilm. Ilk. - Ilmu Komput. Inform.*, vol. 8, no. 1, pp. 25–37, Jan. 2025, doi: 10.47324/ilkominfo.v8i1.303.

- [14] C. A. Fatikasari, A. M. Rahmania, L. Laili, R. A. Mega Saputri, and M. Arifin, "Analisis Segmentasi dan Prediksi Pola Pembelian IC Label Gamis Menggunakan Hybrid K-Means Random Forest," *J. Komput. Teknol. Inf. Sist. Komput.*, vol. 5, no. 1 SE-Articles, pp. 355–364, May 2026, doi: 10.62712/juktisi.v5i1.1007.
- [15] A. Ichwani, R. I. Kesuma, A. Setiawan, I. E. Wicaksono, and R. Hanifah, "Preventing Data Leakage in Classification via Integrated Machine Learning Pipelines: Preprocessing, Feature Transformation, and Hyperparameter Tuning," *J. Tek. Inform.*, vol. 7, no. 1, pp. 391–410, Feb. 2026, doi: 10.52436/1.jutif.2026.7.1.5490.
- [16] A. Z. Kamalia, Choiriyatun Nisa Latansa, and Zaenur Rozikin, "Klasifikasi Kondisi Pasar Harga Emas ANTAM Indonesia Menggunakan Algoritma Decision Tree," *J. Komput. Teknol. Inf. Sist. Inf.*, vol. 4, no. 3, pp. 2087–2098, Jan. 2026, doi: 10.62712/juktisi.v4i3.800.
- [17] S. Ngarifatul Khofiyah, A. Hamdi, and K. Nur Isnaini, "Analisis Pola Konsumsi Air Menggunakan Algoritma Random Forest Classifier Pada Distribusi Air Bersih Desa Rempoah Baturraden," *J. Algoritm.*, vol. 22, no. 2, Nov. 2025, doi: 10.33364/algoritm/v.22-2.3001.
- [18] M. Wahyuni, "Klasifikasi Penyakit Daun Tomat dengan Perbandingan Fungsi Aktivasi Multi Layer Perceptron," *J. Minfo Polgan*, vol. 13, no. 2 SE-, pp. 1988–1998, Dec. 2024, doi: 10.33395/jmp.v13i2.14351.
- [19] A. Diriyah, Khalida Zia Qinthara, Ajeng Windi Setianingsih, Yemima Perangin-Angin, and I Gde Eka Dirgayussa, "Peningkatan Akurasi Klasifikasi Digit melalui Modifikasi CNN dengan Batch Normalization, Dropout, dan Data Augmentation," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.)*, vol. 10, no. 3 SE-Articles, pp. 320–328, doi: 10.30998/hf6zvd70.
- [20] Y. Prasetyo, M. Faris Al Fatih, M. J. Iqbal An Zida, M. Isnaini, and A. Mustika Rizki, "Analisis Perbandingan Optimizer Sgd Dan Adam Pada Model Cnn Untuk Klasifikasi Jamur Edible Dan Poisonous," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 1, pp. 1348–1354, Jan. 2025, doi: 10.36040/jati.v9i1.12659.
- [21] A. B. Indrawan, D. Maulana, and M. Z. Abdurrohman, "Analisis Sentimen Terhadap Program Makan Bergizi Gratis Menggunakan Metode Logistic Regression," *Progresif J. Ilm. Komputer; Vol 22, No 2 April, 2026*, [Online]. Available: <https://ojs.stmik-banjarbaru.ac.id/index.php/progresif/article/view/3610>
- [22] Z. Rozikin and J. J. Hidayat, "Perbandingan Metode Oversampling SMOTE dan ADASYN pada Klasifikasi Diabetes Menggunakan Algoritma CatBoost," *J. Manaj. Inform. Teknol.*, vol. 6, no. 1, pp. 151–164, 2026, doi: 10.51903/mifortekh.v6i1.1157.