

Analisis Prediksi Diabetes Berbasis *Artificial Neural Network* Menggunakan Data *CDC Diabetes Health Indicators*

Artificial Neural Network-Based Diabetes Prediction Analysis Using CDC Diabetes Health Indicators Data

Muhammad Afandi^{1*}, Dodi Dwi Riskianto², M. Raihan Ramadhan³, Sudriyanto⁴

^{1,2,3,4} Informatika, Teknik, Universitas Nurul Jadid, Probolinggo, Indonesia

E-mail: muhammadafandi0112@gmail.com; dwidr07@gmail.com; ramaraihan559@gmail.com; sudriyanto@unuja.ac.id

Article History

Submitted : Jul 14, 2025
Revised : Nov 25, 2025
Accepted : Jan 29, 2026
Available Online : Feb 2, 2026
Published Regularly : Feb 2, 2026

Kata Kunci: *Diabetes; Artificial Neural Network; Prediksi; Indikator Kesehatan; Machine Learning*

Keywords: *Diabetic; Artificial Neural Network; Prediction; Health Indicators; Machine Learning*

Contact



Author

muhammadafandi0112@gmail.com

ABSTRAK

Diabetes melitus merupakan salah satu penyakit kronis dengan prevalensi yang terus meningkat dan memerlukan upaya deteksi dini yang efektif. Penelitian ini bertujuan untuk mengembangkan model prediksi risiko diabetes menggunakan Artificial Neural Network (ANN) berbasis indikator kesehatan non-laboratorium. Dataset yang digunakan adalah CDC Diabetes Health Indicators dengan jumlah data besar dan karakteristik kelas yang tidak seimbang. Tahapan penelitian meliputi preprocessing data yang mencakup penanganan missing value, encoding data kategorikal menggunakan one-hot encoding, normalisasi fitur numerik, serta analisis distribusi kelas target. Model ANN dilatih menggunakan arsitektur Multilayer Perceptron dengan regularisasi dropout dan L2 penalty serta optimisasi AdamW. Hasil evaluasi menunjukkan bahwa model mencapai akurasi sebesar 86,45%, presisi 85,2%, recall 82,7%, dan nilai AUC-ROC sebesar 0,89. Meskipun akurasi berada pada kisaran menengah untuk dataset berukuran besar, nilai AUC yang tinggi menunjukkan kemampuan diskriminasi model yang sangat baik. Performa ini dipengaruhi oleh keterbatasan jumlah fitur non-laboratorium yang digunakan serta distribusi kelas yang tidak seimbang. Temuan penelitian ini menunjukkan bahwa ANN berbasis indikator kesehatan sederhana berpotensi digunakan sebagai alat bantu skrining risiko diabetes pada layanan kesehatan primer. Penelitian selanjutnya disarankan untuk menerapkan teknik penyeimbangan kelas, analisis interpretabilitas model, serta validasi eksternal pada populasi Indonesia.

ABSTRACT

Diabetes mellitus is a chronic disease with increasing prevalence and requires effective early detection efforts. This study aims to develop a diabetes risk prediction model using an Artificial Neural Network (ANN) based on non-laboratory health indicators. The dataset used is the CDC Diabetes Health Indicators with a large amount of data and characteristics of classes that are not fully balanced. The research stages include data preprocessing that includes handling missing values, encoding categorical data using one-hot encoding, normalization of numerical features, and analysis of the target class distribution. The ANN model was trained using a Multilayer Perceptron architecture with dropout regularization and L2 penalty and AdamW optimization. The evaluation results show that the model achieved an accuracy of 86.45%, a precision of 85.2%, a recall of 82.7%, and an AUC-ROC value of 0.89. Although the accuracy is in the medium range for a large dataset, the high AUC value indicates excellent model discrimination ability. This performance is affected by the limited number of non-laboratory features used and the imbalanced class distribution. The findings of this study indicate that ANN based on simple health indicators has the potential to be used as a diabetes risk screening tool in primary healthcare. Further research is recommended to apply class balancing techniques, model interpretability analysis, and external validation in the Indonesian population.

1. Pendahuluan

Diabetes melitus merupakan salah satu penyakit tidak menular yang menjadi perhatian utama kesehatan global karena prevalensinya yang terus meningkat setiap tahun, baik di negara maju maupun negara berkembang, termasuk Indonesia [1]. Peningkatan jumlah penderita diabetes berimplikasi langsung terhadap tingginya beban biaya kesehatan dan menurunnya kualitas hidup masyarakat. Oleh karena itu, upaya deteksi dini melalui prediksi risiko berbasis indikator kesehatan menjadi langkah strategis untuk mendukung pencegahan dan pengendalian penyakit ini. Namun, pendekatan konvensional dalam deteksi diabetes masih menghadapi keterbatasan, terutama dari sisi akurasi, efisiensi pemrosesan data, dan kemampuan adaptasi terhadap data berskala besar [2].

Seiring dengan perkembangan teknologi, penerapan *machine learning* dalam bidang kesehatan, khususnya untuk prediksi diabetes, telah menunjukkan kinerja yang cukup menjanjikan. Berbagai algoritma seperti *Random Forest*, *Support Vector Machine* (SVM), dan *Logistic Regression* telah banyak digunakan dalam penelitian sebelumnya untuk mengidentifikasi individu yang berisiko terkena diabetes [3]. Hasil penelitian tersebut membuktikan bahwa pendekatan berbasis *machine learning* mampu meningkatkan akurasi prediksi dibandingkan metode tradisional. Meskipun demikian, sebagian besar penelitian masih berfokus pada algoritma *machine learning* konvensional, sementara pemanfaatan pendekatan *deep learning*, khususnya *Artificial Neural Network* (ANN), belum dieksplorasi secara optimal. Namun, penelitian yang secara spesifik memanfaatkan dataset *CDC Diabetes Health Indicators* dengan pendekatan ANN masih relatif terbatas, sehingga membuka peluang pengembangan model prediksi yang lebih adaptif dan akurat.

Berdasarkan celah penelitian tersebut, penelitian ini bertujuan untuk mengembangkan model prediksi diabetes menggunakan *Artificial Neural Network* (ANN). Dataset yang digunakan berasal dari *CDC Diabetes Health Indicators* yang mencakup berbagai indikator kesehatan, seperti indeks massa tubuh (BMI), tekanan darah, aktivitas fisik, serta faktor demografis. Model ANN dirancang menggunakan arsitektur *multilayer perceptron* (MLP) yang terdiri dari lapisan input, *hidden layer*, dan *output layer*. Proses pelatihan model dilakukan dengan menggunakan

optimizer Adam dan fungsi *loss binary cross entropy* untuk meningkatkan performa klasifikasi [4].

Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan solusi berbasis kecerdasan buatan untuk mendukung pencegahan dan deteksi dini diabetes pada level kesehatan masyarakat. Selain itu, hasil penelitian ini dapat menjadi referensi bagi penelitian selanjutnya dalam mengoptimalkan penerapan metode *deep learning* pada analisis data kesehatan.

2. Metode Penelitian

2.1 Desain Penelitian

Penelitian ini menggunakan pendekatan *computational experiment* untuk mengembangkan dan mengevaluasi model prediksi risiko diabetes berbasis Artificial Neural Network (ANN) [5], [6]. Desain penelitian disusun mengikuti pedoman TRIPOD (*Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis*) guna menjamin transparansi, konsistensi, dan replikabilitas pelaporan model prediksi. Tahapan utama penelitian meliputi seleksi dataset, preprocessing data, perancangan arsitektur model, pelatihan, serta evaluasi dan validasi performa model [7].



Gambar 1. Flowchart Alur Kerja Model

2.2 Sumber Data

Dataset yang digunakan merupakan *CDC Diabetes Health Indicators* dengan total 253.680 sampel data. Dari keseluruhan fitur yang tersedia, enam fitur utama dipilih berdasarkan analisis *feature importance* menggunakan algoritma *Random Forest* sebagai dasar seleksi variabel [8]. Fitur tersebut terdiri atas dua variabel numerik, yaitu usia dan indeks massa tubuh (*Body Mass Index/BMI*), serta empat variabel kategorik, yakni jenis kelamin (*male/female*), status hipertensi (*yes/no*), aktivitas fisik (aktif/tidak aktif), dan status merokok (*current/former/never*).

2.3 Preprocessing Data

Tahap preprocessing meliputi penanganan nilai hilang menggunakan imputasi median untuk variabel numerik dan imputasi modus untuk variabel kategorik [9]. Selanjutnya, dilakukan normalisasi menggunakan *RobustScaler* yang efektif dalam menangani keberadaan *outlier*. Dataset kemudian dibagi menjadi data pelatihan, validasi, dan pengujian dengan rasio 70:15:15 menggunakan *stratified sampling* untuk menjaga proporsi distribusi kelas [10].

2.4 Arsitektur Model

Model ANN dikembangkan menggunakan arsitektur *Multilayer Perceptron (MLP)* dengan enam neuron pada lapisan input. Dua lapisan tersembunyi digunakan secara berurutan, masing-masing terdiri atas 64 dan 32 neuron dengan fungsi aktivasi ReLU. Untuk mengurangi risiko *overfitting*, diterapkan *dropout* dengan nilai 0,3 pada setiap lapisan tersembunyi serta regularisasi L2 dengan nilai λ sebesar 0,01. Lapisan output terdiri atas satu neuron dengan fungsi aktivasi sigmoid untuk menghasilkan probabilitas kelas *biner* [11].

2.5 Pelatihan Model

Pelatihan model dilakukan menggunakan *optimizer AdamW* dengan parameter $\beta_1 = 0,9$ dan $\beta_2 = 0,999$. Strategi *cyclical learning rate* diterapkan dengan rentang nilai antara 1×10^{-4} hingga 1×10^{-3} untuk mempercepat proses konvergensi [12]. Model dilatih menggunakan *batch size* sebesar 32 dengan jumlah maksimum 50 *Epoch*. Selain itu, diterapkan mekanisme *early stopping*

dengan *patience* lima *Epoch* untuk menghentikan pelatihan apabila tidak terjadi peningkatan kinerja pada data validasi.

2.6 Evaluasi dan Validasi

Evaluasi performa model dilakukan menggunakan teknik *10 fold nested cross validation* untuk memperoleh estimasi kinerja yang lebih andal dan meminimalkan bias evaluasi [13]. Selain validasi internal, model juga diuji pada data klinis retrospektif guna menilai kelayakan implementasi dalam konteks praktik nyata, dengan tetap memperhatikan prinsip etika penelitian[16].

Kinerja model dievaluasi menggunakan metrik akurasi, *presisi*, *recall*, dan AUC-ROC [14], yang masing-masing memberikan perspektif komplementer terhadap kemampuan klasifikasi model, khususnya dalam menangani data dengan distribusi kelas yang tidak seimbang [15], [17] sebagai berikut:

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

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

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

$$\text{AUC-ROC} = \int_0^1 \text{TPR}(\text{FPR})d\text{FPR} \quad (4)$$

Keterangan:

TP : *True Positive*, jumlah data yang benar-benar positif dan berhasil diprediksi positif oleh model.

TN : *True Negative*, jumlah data yang benar-benar negatif dan berhasil diprediksi negatif oleh model.

FP : *False Positive*, jumlah data yang sebenarnya negatif tetapi diprediksi positif oleh model.

FN : *False Negative*, jumlah data yang sebenarnya positif tetapi diprediksi negatif oleh model.

TPR : *True Positive Rate*, proporsi data yang sebenarnya positif dan berhasil diprediksi positif oleh model dibandingkan dengan seluruh data positif yang ada.

FPR : *False Positive Rate*, proporsi data yang sebenarnya negatif tetapi diprediksi positif oleh model dibandingkan dengan seluruh data negatif yang ada.

3. Hasil dan Pembahasan

3.1 Preprocessing Data

Tahapan *preprocessing* data dilakukan untuk memastikan kualitas data, mengurangi *noise*, serta menyesuaikan format data dengan kebutuhan algoritma *Artificial Neural Network* (ANN). Langkah-langkah *preprocessing* yang diterapkan adalah sebagai berikut.

1. Penanganan *Missing Value*

Dataset CDC *Diabetes Health Indicators* memiliki sejumlah nilai hilang (*missing value*) pada beberapa variabel. Untuk variabel numerik, seperti usia dan indeks massa tubuh (BMI), nilai hilang ditangani menggunakan metode imputasi median. Pendekatan ini dipilih karena median lebih robust terhadap outlier dibandingkan rata-rata. Sementara itu,

untuk variabel kategorikal, seperti jenis kelamin dan status merokok, nilai hilang diisi menggunakan modus yang paling sering muncul pada masing-masing variabel.

2. *Encoding Data Kategorikal*
Variabel kategorikal yang bersifat nominal, antara lain jenis kelamin, status merokok, dan aktivitas fisik, dikonversi ke bentuk numerik menggunakan teknik *one-hot encoding*. Teknik ini dipilih karena tidak memberikan asumsi hubungan ordinal antar kategori serta kompatibel dengan ANN yang membutuhkan input numerik.
3. *Normalisasi Data*
Variabel numerik dinormalisasi menggunakan metode min–max *scaling* ke rentang [0,1]. Normalisasi ini bertujuan untuk mempercepat proses konvergensi selama pelatihan dan mencegah dominasi fitur dengan skala nilai yang lebih besar.
4. *Analisis Keseimbangan Data*
Analisis distribusi kelas target menunjukkan bahwa dataset bersifat tidak sepenuhnya seimbang, dengan proporsi kelas non diabetes lebih besar dibandingkan kelas diabetes. Ketimpangan ini berpotensi memengaruhi performa model, khususnya pada metrik sensitivitas (*recall*) terhadap kelas minoritas.
5. *Strategi Penyeimbangan Data*
Pada penelitian ini tidak diterapkan teknik penyeimbangan data seperti *Synthetic Minority Oversampling Technique* (SMOTE) maupun penyesuaian bobot kelas (*class weight*). Model dilatih menggunakan distribusi data asli untuk mengevaluasi performa ANN dalam kondisi realistis. Namun demikian, keputusan ini berimplikasi pada keterbatasan sensitivitas model dalam mendeteksi kasus diabetes.
6. *Augmentasi Data*
Augmentasi data tidak dilakukan karena seluruh fitur yang digunakan merupakan indikator kesehatan numerik dan kategorikal yang tidak sesuai untuk teknik augmentasi sintesis seperti pada data citra atau sinyal.

3.2 Proses Pelatihan dan Validasi Model

Model ANN dilatih menggunakan arsitektur *Multilayer Perceptron* (MLP) dengan satu lapisan tersembunyi. Untuk mengurangi risiko *overfitting*, diterapkan teknik regularisasi berupa *dropout* dan *L2 penalty*. Proses optimisasi menggunakan algoritma AdamW yang dikombinasikan dengan strategi *cyclical learning rate* [3]. Pelatihan dilakukan selama 10 *epoch* dengan pembagian data ke dalam data latih dan data validasi. Hasil akurasi validasi pada setiap *epoch* disajikan pada Tabel 1.

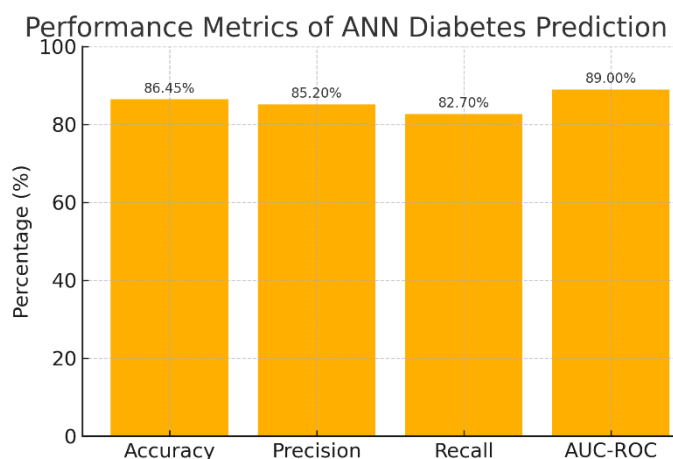
Tabel 1. *Epoch* dan Akurasi Validasi

<i>Epoch</i>	Akurasi Validasi (%)
1	85,94
2	85,80
3	85,91
4	85,87
5	85,94
6	85,95
7	86,07
8	85,83
9	86,05
10	86,04

Hasil menunjukkan bahwa akurasi validasi relatif stabil pada kisaran 85,80% hingga 86,07%. Stabilitas ini mengindikasikan bahwa model mampu mempertahankan kemampuan generalisasi yang baik tanpa mengalami *overfitting* atau degradasi performa signifikan selama pelatihan.

3.3 Hasil Kinerja Model

Evaluasi kinerja model dilakukan menggunakan beberapa metrik utama, yaitu akurasi, presisi, *recall*, dan AUC-ROC. Hasil pengujian menunjukkan bahwa model mencapai akurasi sebesar 86,45%, yang mencerminkan proporsi prediksi benar secara keseluruhan. Presisi sebesar 85,2% menunjukkan bahwa sebagian besar prediksi positif yang dihasilkan model memang merupakan kasus diabetes yang benar, sehingga kesalahan positif palsu dapat diminimalkan. *Recall* sebesar 82,7% menunjukkan tingkat sensitivitas model dalam mendeteksi individu yang benar-benar berisiko diabetes. Nilai *recall* yang sedikit lebih rendah dibandingkan presisi mengindikasikan adanya pengaruh ketidakseimbangan kelas, dimana model cenderung lebih optimal pada kelas mayoritas sebagaimana ditunjukkan Gambar 2.



Gambar 2. Hasil evaluasi model *Artificial Neural Network* pada dataset *CDC Diabetes Health Indicators*

Nilai AUC-ROC sebesar 0,89 menandakan kemampuan diskriminasi model yang sangat baik dalam membedakan kelas diabetes dan non diabetes. Nilai ini menunjukkan bahwa meskipun akurasi berada pada kisaran 86%, model memiliki kapasitas pemisahan kelas yang kuat. Akurasi yang tidak terlalu tinggi pada dataset berukuran besar dapat disebabkan oleh beberapa faktor, antara lain ketidakseimbangan distribusi kelas, keterbatasan jumlah fitur yang digunakan, serta kompleksitas ANN yang relatif sederhana. Penggunaan enam indikator non-laboratorium memang meningkatkan kepraktisan, namun membatasi informasi prediktif yang dapat dipelajari model.

3.4 Kontribusi Fitur dan Relevansi Klinis

Model dikembangkan menggunakan enam indikator kesehatan non laboratorium, yaitu usia, indeks massa tubuh (BMI), jenis kelamin, aktivitas fisik, hipertensi, dan status merokok. Hasil pelatihan menunjukkan bahwa usia dan BMI memiliki kontribusi paling dominan terhadap prediksi risiko diabetes. Temuan ini konsisten dengan literatur yang menyebutkan bahwa peningkatan usia dan obesitas merupakan faktor risiko utama diabetes. Variabel perilaku, seperti aktivitas fisik dan status merokok, meskipun kontribusinya lebih kecil, tetap berperan dalam meningkatkan sensitivitas model terhadap variasi risiko individu. Pemilihan fitur-fitur tersebut juga memiliki relevansi klinis yang tinggi karena dapat diperoleh dengan mudah melalui anamnesis dan pemeriksaan rutin, sehingga mendukung penerapan model di layanan kesehatan primer.

3.5 Implikasi Praktis

Hasil penelitian ini menunjukkan bahwa model ANN berbasis indikator kesehatan sederhana berpotensi digunakan sebagai alat bantu skrining risiko diabetes. Penggunaan fitur non-laboratorium memungkinkan penerapan model secara cepat dan efisien, terutama di fasilitas

kesehatan dengan keterbatasan sumber daya. Integrasi model ke dalam sistem rekam medis elektronik dapat mendukung sistem pendukung keputusan klinis dengan memberikan peringatan dini bagi individu berisiko tinggi, sehingga intervensi preventif dapat dilakukan lebih awal [18].

3.6 Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan. Pertama, model belum dilengkapi dengan analisis interpretabilitas kuantitatif sehingga kontribusi masing-masing fitur belum dapat dijelaskan secara mendalam. Kedua, ketidakseimbangan kelas belum ditangani melalui teknik penyeimbangan data, yang berpotensi memengaruhi sensitivitas model terhadap kelas minoritas. Ketiga, dataset yang digunakan berasal dari populasi luar negeri, sehingga generalisasi hasil pada populasi Indonesia masih memerlukan validasi eksternal.

3.7 Arah Penelitian Selanjutnya

Penelitian selanjutnya disarankan untuk menerapkan metode interpretabilitas seperti SHAP atau LIME guna meningkatkan transparansi model. Selain itu, validasi eksternal menggunakan data populasi Indonesia serta penerapan strategi penyeimbangan kelas, seperti SMOTE atau class *weight*, diharapkan dapat meningkatkan akurasi dan sensitivitas model. Pengembangan lebih lanjut juga dapat mencakup penambahan variabel klinis yang relevan dan integrasi model ke dalam sistem pendukung keputusan klinis berbasis rekam medis elektronik untuk mendukung skrining diabetes secara luas.

4. Kesimpulan

Penelitian ini berhasil mengembangkan model prediksi risiko diabetes berbasis *Artificial Neural Network* menggunakan enam indikator kesehatan non-laboratorium, yaitu usia, indeks massa tubuh, jenis kelamin, aktivitas fisik, hipertensi, dan status merokok. Hasil evaluasi menunjukkan bahwa model memiliki performa yang cukup baik dengan akurasi 86,45% dan kemampuan diskriminasi yang tinggi, ditunjukkan oleh nilai AUC-ROC sebesar 0,89. Meskipun demikian, capaian akurasi yang belum optimal pada dataset berukuran besar mengindikasikan adanya pengaruh ketidakseimbangan kelas dan keterbatasan kompleksitas model serta fitur yang digunakan. Model cenderung lebih optimal dalam memprediksi kelas mayoritas, sehingga sensitivitas terhadap kelas minoritas masih dapat ditingkatkan. Namun, penggunaan indikator non-laboratorium memberikan keunggulan praktis karena memungkinkan penerapan model secara cepat dan efisien di layanan kesehatan primer. Secara keseluruhan, penelitian ini menunjukkan bahwa pendekatan ANN dengan fitur kesehatan sederhana layak dipertimbangkan sebagai alat bantu skrining awal risiko diabetes. Untuk meningkatkan akurasi, sensitivitas, dan keandalan model, penelitian selanjutnya perlu mengintegrasikan teknik penyeimbangan kelas, metode interpretabilitas model, serta validasi menggunakan data populasi lokal. Integrasi ke dalam sistem pendukung keputusan klinis berbasis rekam medis elektronik juga menjadi arah pengembangan yang potensial.

Daftar Pustaka

- [1] S. I. Oktora and D. B. Butar, "Determinants of Diabetes Mellitus Prevalence in Indonesia," *J. Kesehat. Masy.*, vol. 18, no. 2, pp. 266–273, 2022.
- [2] A. R. Kandula, V. V. Samhitha, V. Harshitha, R. N. Sindhuri, and S. A. Nagavalli, "Machine Learning based Screening for Diabetes Risk Prediction," in *2024 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS)*, IEEE, 2024, pp. 1188–1193.
- [3] N. Nagarjuna and H. N. Lakshmi, "Predictive Modeling of Diabetes Mellitus Utilizing

- Machine Learning Techniques,” *CVR J. Sci. Technol.*, vol. 26, no. 1, pp. 112–117, 2024.
- [4] T. Elansari, M. Ouanan, and H. Bourray, “Modeling of Multilayer Perceptron Neural Network Hyperparameter Optimization and Training,” 2023.
- [5] D. Kurniawan, T. T. Wulansari, and N. A. D. Febrianti, “Model Prediction Using Artificial Neural Network (ANN) to Strengthen Diagnostic Analysis of Diabetes Melitus,” *ComTech Comput. Math. Eng. Appl.*, vol. 15, no. 2, pp. 83–91, 2024.
- [6] S. M. Lawan and S. J. Kawu, “Artificial Neural Network (ANN) Prediction of Guide Vane for Effective Utilization of Pump as Turbine (PAT),” in *2024 IEEE International Conference on Power and Energy (PECon)*, IEEE, 2024, pp. 34–39.
- [7] P. Dhiman *et al.*, “Methodological conduct of prognostic prediction models developed using machine learning in oncology: a systematic review,” *BMC Med. Res. Methodol.*, vol. 22, no. 1, p. 101, 2022.
- [8] S. Demir and E. K. Sahin, “Random forest importance-based feature ranking and subset selection for slope stability assessment using the ranger implementation,” *Avrupa Bilim ve Teknol. Derg.*, no. 48, pp. 23–28, 2023.
- [9] S. Prakash, S. Singh, and A. Mankar, “Bridging data gaps: A comparative study of different imputation methods for numeric datasets,” in *2024 International Conference on Data Science and Network Security (ICDSNS)*, IEEE, 2024, pp. 1–7.
- [10] K. Song, R. Solozabal, H. Li, M. Takáč, L. Ren, and F. Karray, “Robustly train normalizing flows via kl divergence regularization,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, pp. 15047–15055.
- [11] T. Filimonova, H. Samoylenko, A. Selivanova, Y. Yurchenko, and A. Parashchak, “CONSTRUCTION OF A NEURAL NETWORK FOR HANDWRITTEN DIGITS RECOGNITION BASED ON TENSORFLOW LIBRARY APPLYING AN ERROR BACKPROPAGATION ALGORITHM,” *Eastern-European J. Enterp. Technol.*, vol. 126, no. 2, 2023.
- [12] Z. Lin, S. Zhang, Y. Zhou, H. Wang, and S. Wang, “Learning rate burst for superior SGDM and AdamW integration,” *J. Intell. Fuzzy Syst.*, p. JIFS-239157, 2024.
- [13] J. Qiu, “An analysis of model evaluation with cross-validation: techniques, applications, and recent advances,” *Adv. Econ. Manag. Polit. Sci.*, vol. 99, pp. 69–72, 2024.
- [14] M. Owusu-Adjei, J. Ben Hayfron-Acquah, T. Frimpong, and G. Abdul-Salaam, “Imbalanced class distribution and performance evaluation metrics: A systematic review of prediction accuracy for determining model performance in healthcare systems,” *PLOS Digit. Heal.*, vol. 2, no. 11, p. e0000290, 2023.
- [15] G. S. Sahoo, S. Dass, and F. Prakash, “The Synergy of Advanced Language Models, Sequential Skeleton Features, and GAN-Enhanced Data Augmentation for Unprecedented Precision,” in *2023 International Conference on System, Computation, Automation and Networking (ICSCAN)*, IEEE, 2023, pp. 1–6.
- [16] T. F. Monaghan *et al.*, “Foundational statistical principles in medical research: sensitivity, specificity, positive predictive value, and negative predictive value,” *Medicina (B. Aires)*, vol. 57, no. 5, p. 503, 2021.
- [17] R. H. Stern, “Interpretation of the area under the ROC curve for risk prediction models. arXiv,” *arXiv Prepr. arXiv2102.11053*, 2021.
- [18] P. H. Artanti, “Penerapan Neural Network dengan optimasi Ant Colony Optimization dan Backpropagation untuk membangun model prediksi diabetes tahap awal,” 2023, *Universitas Islam Negeri Maulana Malik Ibrahim*.