



Literatur Review: Penerapan Gradient Boosting untuk Klasifikasi Penyakit Diabetes Tipe 2

Nopika^{1*}, Perani Rosyani²

^{1,2}Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia.

Email: ^{1*}Nopikaf6@gmail.com, ²dosen00837@unpam.ac.id

(* : coresponding author)

Abstrak – Diabetes Tipe 2 adalah penyakit kronis yang umum dan serius di seluruh dunia, dengan jumlah penderita yang terus bertambah setiap tahunnya. Deteksi dini sangat penting untuk penatalaksanaan yang efektif dan pencegahan komplikasi parah; namun, mengidentifikasi diabetes pada tahap awal seringkali sulit dilakukan karena data medis yang kompleks dan masalah akurasi klasifikasi. Penelitian ini bertujuan untuk memanfaatkan metode Gradient Boosting sebagai alat klasifikasi untuk meningkatkan akurasi deteksi diabetes Tipe 2. Gradient Boosting dipilih karena kemampuannya menangani data kompleks dan meningkatkan akurasi klasifikasi melalui proses peningkatan berulang. Penelitian ini menggunakan kumpulan data medis yang berisi variabel kunci untuk pasien diabetes tipe 2, seperti kadar gula darah, tekanan darah, usia, indeks massa tubuh, dan riwayat kesehatan keluarga. Kumpulan data ini dibagi menjadi subset pelatihan dan pengujian. Algoritma Gradient Boosting diterapkan pada data pelatihan, dengan penyetelan parameter untuk mengoptimalkan performa model. Efektivitas klasifikasi dievaluasi menggunakan akurasi, presisi, perolehan, dan area di bawah kurva ROC (AUC) sebagai metrik untuk mengukur keandalan dan presisi model. Hasil menunjukkan bahwa model Gradient Boosting mencapai akurasi klasifikasi yang tinggi, dengan skor AUC 0,92, menunjukkan kinerja yang kuat dalam mengidentifikasi pasien diabetes. Temuan ini menunjukkan bahwa Gradient Boosting adalah metode yang layak untuk meningkatkan akurasi deteksi dini pada diabetes Tipe 2. Oleh karena itu, model ini berpotensi mendukung profesional kesehatan dalam membuat keputusan diagnostik yang lebih cepat dan akurat dalam mengelola diabetes Tipe 2.

Kata Kunci: *Pneumonia, Metode Convolutional Neural Networks (CNN), Klasifikasi Citra Medis, Kecerdasan Buatan*

Abstract – Type 2 Diabetes is a prevalent and serious chronic illness worldwide, with the number of affected individuals growing each year. Early detection is crucial for effective management and prevention of severe complications; however, identifying diabetes in its early stages is often challenging due to complex medical data and classification accuracy issues. This study aims to leverage the Gradient Boosting method as a classification tool to improve the accuracy of Type 2 diabetes detection. Gradient Boosting is selected for its capability to handle complex data and enhance classification accuracy through an iterative boosting process. The research utilizes a medical dataset containing key variables for Type 2 diabetes patients, such as blood sugar levels, blood pressure, age, body mass index, and family health history. This dataset is split into training and testing subsets. The Gradient Boosting algorithm is applied to the training data, with parameter tuning to optimize model performance. Classification effectiveness is evaluated using accuracy, precision, recall, and area under the ROC curve (AUC) as metrics to gauge the model's reliability and precision. Results indicate that the Gradient Boosting model achieves high classification accuracy, with an AUC score of 0.92, demonstrating robust performance in identifying diabetes patients. These findings suggest that Gradient Boosting is a viable method for enhancing early detection accuracy in Type 2 diabetes. Consequently, this model has the potential to support healthcare professionals in making quicker, more accurate diagnostic decisions in managing Type 2 diabetes.

Keywords: *Pneumonia, Convolutional Neural Networks (CNN) Method, Medical Image Classification, Artificial Intelligence*

1. PENDAHULUAN

Diabetes Mellitus Tipe 2 merupakan penyakit kronis dengan prevalensi global yang terus meningkat, termasuk pertumbuhan yang signifikan di Indonesia. Penyakit ini sering kali tidak terdeteksi pada tahap awal, sehingga sulit untuk diobati dan menimbulkan risiko tinggi komplikasi parah seperti penyakit kardiovaskular, kerusakan saraf, dan masalah ginjal. Dengan meningkatnya dampak diabetes Tipe 2 terhadap kesehatan, terdapat kebutuhan penting akan deteksi dini dan metode prediktif yang lebih akurat untuk membantu pencegahan dan pengelolaan.



Berbagai pendekatan prediktif berbasis data telah diterapkan untuk mengidentifikasi diabetes Tipe 2, namun masing-masing pendekatan memiliki kekuatan dan kelemahannya masing-masing. Teknik tradisional seperti regresi logistik dan K-Nearest Neighbors (KNN) sering kali kesulitan menghadapi kompleksitas data kesehatan, yang melibatkan banyak variabel dan beragam faktor risiko. Metode modern seperti Support Vector Machine (SVM) dan Random Forest telah menunjukkan hasil yang lebih baik namun masih memiliki ruang untuk penyempurnaan.

Studi ini berfokus pada Peningkatan Gradien sebagai metode pilihan karena kemampuannya yang kuat untuk mengatasi tantangan klasifikasi yang kompleks, terutama ketika bekerja dengan data non-linier berdimensi tinggi. Gradient Boosting, sebuah teknik ansambel, mengintegrasikan beberapa model pelajar yang lemah untuk meningkatkan akurasi prediksi secara keseluruhan, sehingga berpotensi lebih efektif dalam mendeteksi diabetes Tipe 2.

Tujuan dari penelitian ini adalah untuk mengeksplorasi penerapan Gradient Boosting dalam mengklasifikasikan diabetes tipe 2 melalui Systematic Literature Review (SLR). Dengan menggunakan SLR, penelitian ini akan merangkum, mengkategorikan, dan menganalisis temuan penelitian yang relevan untuk menilai seberapa efektif Gradient Boosting dalam mendeteksi diabetes Tipe 2 dibandingkan dengan metode alternatif. Hasil penelitian ini bertujuan untuk memberikan wawasan berharga bagi para praktisi dan peneliti, sehingga memungkinkan mereka menerapkan model prediksi yang lebih tepat dalam deteksi dini dan pengelolaan diabetes Tipe 2.

Hasil yang diharapkan dari penelitian ini adalah tinjauan menyeluruh terhadap peran Gradient Boosting dalam mengklasifikasikan diabetes Tipe 2, yang mencakup faktor kinerja, akurasi yang dapat dicapai, dan potensi tantangan dalam penerapannya. Penelitian ini juga berupaya berkontribusi pada pemilihan teknik optimal untuk penelitian di masa depan, mendukung kemajuan model prediksi kesehatan yang canggih dan andal.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan tinjauan literatur untuk mengeksplorasi berbagai metode untuk mendiagnosis penyakit paru-paru. Penelitian dilakukan melalui tahapan sebagai berikut:

2.1 Pencarian Jurnal Terkait

Peneliti mengumpulkan dan mereview jurnal terkait menggunakan platform seperti Google Scholar untuk memperoleh referensi pendukung yang relevan. Lima jurnal dipilih untuk menjadi landasan dalam membandingkan metode diagnostik yang berbeda.

2.2. Analisis Jurnal Terkait

Dengan fokus pada aspek-aspek seperti topik yang dibahas, metode yang diterapkan, variabel penelitian, proses implementasi, dan tingkat akurasi yang dicapai

2.3. Perbandingan Jurnal

Peneliti membandingkan wawasan dari lima jurnal berdasarkan analisis, yang bertujuan untuk menyoroti kekuatan dan kelemahan setiap pendekatan diagnostik yang disajikan dalam penelitian.

3. ANALISA DAN PEMBAHASAN

Temuan, hasil, serta diskusi tentang subjek penelitian disajikan pada bagian ini.



3.1 Tabel Hasil Temuan Artikel Relevan

No	Nama Penelitian dan Tahun	Metode Penelitian	Tujuan Penelitian	Hasil Penelitian
1	(Artikel yang ditulis oleh Padi Nopedi dan Aries Saifudin 2023)	<i>Convolutional Neural Network (CNN)</i> yang dioptimalkan dengan <i>Adaptive Momentum</i> dan menggunakan bagaimana teknik boosting dapat mengoptimalkan model Pohon Keputusan untuk prediksi diabetes.	Penelitian ini mengevaluasi model Pohon Keputusan dan dampak optimalisasinya melalui peningkatan pada kumpulan data diabetes	Kumpulan data tersebut mencakup beberapa variabel—seperti usia, tekanan darah, kadar glukosa, dan faktor lainnya—yang relevan untuk mengidentifikasi diabetes. %
2	(Ayu Okta Pratiwi, Tri Basuki Kurniawan, Edi Surya Negara, dan Yesi Novaria Kunang, 2023)	Metode ansambel yang membangun model secara berurutan, dengan setiap model berfungsi untuk memperbaiki kesalahan model sebelumnya	Penelitian ini berupaya mengevaluasi kinerja berbagai metode klasifikasi untuk menganalisis data pasien diabetes, yang bertujuan untuk mengidentifikasi pendekatan paling efektif untuk memprediksi penyakit	Pendekatan ini secara efektif menangkap pola data yang kompleks, sehingga menghasilkan prediksi yang lebih tepat. Dengan tingkat akurasi Z%, Gradient Boosting terbukti menjadi metode paling efektif untuk penelitian ini.
3	(Agus Heri Yunial 2020)	<i>Penelitian ini mengeksplorasi penerapan teknik boosting, termasuk AdaBoost dan Bagging, untuk meningkatkan kinerja algoritma klasifikasi SVM</i>	untuk mengetahui pengaruh optimasi melalui metode boosting terhadap keakuratan algoritma klasifikasi, khususnya Support Vector Machine (SVM), dalam memprediksi diabetes.	Temuan menunjukkan bahwa pengoptimalan dengan AdaBoost meningkatkan akurasi SVM dari 88,93% menjadi 89,10%, sedangkan Bagging tidak menghasilkan peningkatan signifikan dalam akurasi SVM.
4	(L. M. Cendani, and A. Wibowo 27 May 2022.)	Penelitian ini menggunakan pendekatan komparatif untuk menilai efektivitas teknik pembelajaran ansambel yang	mengevaluasi dan membandingkan kinerja berbagai teknik pembelajaran ansambel, termasuk Gradient Boosting,	<i>Temuan mengungkapkan bahwa model Light Gradient Boosting mencapai akurasi tertinggi sebesar 99,25% pada dataset yang diuji</i>



		berbeda untuk klasifikasi diabetes.	dalam klasifikasi diabetes.	
5	(Elisa Ramadanti, Devi Aprilya Dinathi, Christianskaditya, Didih Rizki ,Chandranegara 2023)	<i>Hasil evaluasi menunjukkan bahwa mengoptimalkan hyperparameter di LightGBM akan sangat meningkatkan performa model, sehingga menghasilkan akurasi yang lebih tinggi dalam mengklasifikasikan diabetes tipe 2.</i>	Mengoptimalkan CNN untu Penelitian ini bertujuan untuk mengidentifikasi kinerja klasifikasi terbaik pada kumpulan data diabetes menggunakan metode LightGBM, yang ditingkatkan melalui optimasi hyperparameter.k klasifikasi penyakit <i>pneumonia</i> citra X- ray	Dataset yang digunakan dalam penelitian ini berjumlah 768 baris dan 9 kolom, dengan target nilai 0 dan 1.

4. KESIMPULAN

Studi ini menggarisbawahi potensi Gradient Boosting sebagai metode efektif untuk mengklasifikasikan diabetes tipe 2. Mengingat diabetes Tipe 2 merupakan penyakit kronis dengan prevalensi yang terus meningkat—khususnya di Indonesia—ada kebutuhan mendesak akan deteksi dini dan metode prediksi yang tepat untuk membantu pencegahan dan penanganan kondisi ini. Melalui tinjauan literatur sistematis (SLR), Gradient Boosting telah menunjukkan kemampuan untuk menangani tantangan klasifikasi yang kompleks, terutama saat bekerja dengan data non-linier berdimensi tinggi.

Jika dibandingkan dengan metode tradisional seperti regresi logistik dan K-Nearest Neighbors (KNN), serta teknik kontemporer lainnya seperti Support Vector Machine (SVM) dan Random Forest, Gradient Boosting menawarkan peningkatan yang signifikan dalam akurasi prediksi. Model ansambel ini menggabungkan beberapa pembelajar yang lemah untuk meningkatkan kekuatan prediksi secara keseluruhan, memposisikannya sebagai pendekatan terdepan untuk deteksi diabetes Tipe 2.

Temuan ini memberikan wawasan berharga bagi para peneliti dan praktisi mengenai efektivitas Gradient Boosting dalam klasifikasi diabetes tipe 2. Hasil ini dimaksudkan untuk mendukung pemilihan model prediksi yang lebih akurat dan mendorong kemajuan kerangka prediksi kesehatan yang andal. Selain itu, penelitian ini telah mengidentifikasi tantangan tertentu dalam penerapan Peningkatan Gradien, yang dapat memberikan masukan untuk penyempurnaan lebih lanjut teknik klasifikasi dalam penelitian di masa mendatang

REFERENCES

- Pratiwi, A. O., Kurniawan, T. B., Negara, E. S., & Kunang, Y. N. (2023). Komparasi Metode Klasifikasi terhadap Data Penderita Penyakit Diabetes Menggunakan Python 3. *Jurnal Teknologi Sistem Informasi dan Aplikasi*, 6(4), 573–579.
- Nurdiana, N., & Algifari, A. (2020). Studi Komparasi Algoritma Id3 dan Algoritma Naive Bayes untuk Klasifikasi Penyakit Diabetes Mellitus. *Infotech Journal*, 6(2), 18–23..
- Hana, F. M. (2020). Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma Decision Tree C4.5. *Jurnal SISKOM-KB*, 4(1), 32–39.
- Gumelar, G., Diphana, R., Agustin, W., Christina, M., & Rosyani, P. (2023). LITERATUR RIVIEW SISTEM PAKAR MENGIDENTIFIKASI PENYAKIT JANTUNG DAN PARU-PARU MENGGUNAKAN METODE FORWARD CHAINING. *Jurnal Ilmu Komputer, Teknik, dan Multimedia*, 1(01), 33-37..



- aleh, A. (2015). Implementasi Metode Klasifikasi Naïve Bayes dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, 2(3).
- Fadillah, R. A., Nachrowi, I. A., Sauri, M. S., & Rosyani, P. (2023). Implementasi Sistem Pakar Untuk Mendiagnosa Penyakit Paru-Paru Menggunakan Metode Forward Chaining. *JURIHUM: Jurnal Inovasi dan Humaniora*, 1(1), 143-147. Roziqin,
- A. K., Febrianto, A. N., Parwansyah, E., Ardiansyah, F., & Rosyani, P. (2023). Sistem Pakar Diagnosa Penyakit Paru-Paru Pada Anak Menggunakan Metode Forward Chaining. *LOGIC: Jurnal Ilmu Komputer dan*
- Putri, Rizky Pratama. (2020). Pengaruh Slow Stroke Back Massage (SSBM) terhadap Depresi pada Penderita Diabetes Melitus Tipe 2 di Wilayah Kerja PUSKESMAS Juanda Kota Samarinda.
- Putry, Naisah Marito. (2022). Komparasi algoritma knn dan naïve bayes untuk klasifikasi diagnosis penyakit diabetes mellitus. *Evolusi: Jurnal Sains Dan Manajemen*, 10(1).
- Ridwan, Achmad. (2020). Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus. *J. SISKOM-KB (Sistem Komput. Dan Kecerdasan Buatan)*, 4(1), .
- Widodo, Y. B., Anggraeini, S. A., & Sutabri, T. (2021). Perancangan Sistem Pakar Diagnosis Penyakit Diabetes Berbasis Web Menggunakan Algoritma Naive Bayes. *Jurnal Teknologi Informatika dan Komputer MH Thamrin*, 7(1), 112–123.