

ABSTRAK

Arifiantono, Aditya Bagus. 2024. Pemodelan *Machine Learning* pada Data Pasien Diagnosis Demam *Tifoid* Menggunakan Metode *K-Nearest Neighbor*. Tugas Akhir. Program Sarjana. Program Studi Teknik Informatika. Universitas Muhammadiyah Jember.

Pembimbing:

(1) Hardian Oktavianto, S.Si., M.Kom.; (2) Guruh Wijayah, S.T., M.Kom

Diagnosis demam tifoid merupakan tantangan dalam praktek medis sehari-hari, memerlukan pendekatan yang teliti dan efektif untuk memastikan ketepatan diagnosa. Penelitian ini menggunakan pendekatan *Machine Learning* dengan metode *K-Nearest Neighbor (KNN)* dan teknik *Synthetic Minority Over-sampling Technique (SMOTE)* untuk menangani ketidakseimbangan dataset. Data yang digunakan berasal dari pasien yang terdiagnosis demam tifoid dan non-demam tifoid, yang dikumpulkan dari puskesmas di Malang, Jawa Timur. Metode *preprocessing* mencakup eliminasi atribut yang tidak relevan dan pengkodean label untuk transformasi data kategorial menjadi numerik. Hasil pemodelan menunjukkan bahwa parameter terbaik untuk *KNN* adalah *Metric = Euclidean*, *Weights = Uniform*, dan *n_Neighbors = 3*, dengan rata-rata akurasi mencapai 86%. Model ini memiliki presisi 87% untuk kelas A01 (demam tifoid), dan *recall* 100% untuk kelas A01, menunjukkan kemampuan yang baik dalam mengklasifikasikan kasus demam tifoid. Namun, model menunjukkan tantangan dalam mengidentifikasi kelas “*Unspecified*” dengan presisi dan *recall* yang rendah. Implementasi *SMOTE* berhasil meningkatkan presisi kelas A01 menjadi 95% dan *recall* kelas “*Unspecified*” menjadi 71%. Meskipun demikian, akurasi total pengujian mengalami penurunan menjadi 73% karena perubahan distribusi kelas setelah penerapan *SMOTE*. Meskipun demikian, hasil ini menunjukkan bahwa model tetap mampu mengklasifikasikan data uji secara efektif dalam konteks diagnosa demam tifoid. Penelitian ini memberikan kontribusi penting dalam meningkatkan akurasi diagnosa demam tifoid menggunakan pendekatan *Machine Learning*, yang dapat memberikan panduan berharga bagi praktisi kesehatan dalam meningkatkan efisiensi dan ketepatan diagnosa penyakit ini.

Kata kunci: demam tifoid, *Machine Learning*, *K-Nearest Neighbor (KNN)*, *SMOTE*

ABSTRACT

Arifiantono, Aditya Bagus. 2024. *Machine Learning Modeling on Typhoid Fever Diagnosis Patient Data Using the K-Nearest Neighbor Method*. Final Project. Bachelor Program. Informatics Engineering Study Program. University of Muhammadiyah Jember.

Advisors:

(1) Hardian Oktavianto, S.Si., M.Kom.; (2) Guruh Wijayah, S.T., M.Kom.

Diagnosing typhoid fever poses a significant challenge in daily medical practice, requiring meticulous and effective approaches to ensure diagnostic accuracy. This study employs Machine Learning using the K-Nearest Neighbor (KNN) method and Synthetic Minority Over-sampling Technique (SMOTE) to address dataset imbalance. Data originates from patients diagnosed with either typhoid fever or non-typhoid conditions, gathered from health centers in Malang, East Java. Preprocessing methods include eliminating irrelevant attributes and label encoding to transform categorical data into numerical form. The modeling results indicate that the optimal parameters for KNN are Metric = Euclidean, Weights = Uniform, and n_Neighbors = 3, achieving an average accuracy of 86%. The model exhibits 87% precision for class A01 (typhoid fever) and 100% recall for class A01, demonstrating robust capability in classifying typhoid cases. However, the model faces challenges in identifying the "Unspecified" class, showing low precision and recall. The implementation of SMOTE successfully enhances precision to 95% for class A01 and improves recall to 71% for the "Unspecified" class. Nevertheless, the overall testing accuracy decreases to 73% due to changes in class distribution post-SMOTE implementation. Nonetheless, these results underscore the model's effectiveness in accurately classifying test data within the context of typhoid fever diagnosis. This research makes a significant contribution to improving the accuracy of typhoid fever diagnosis using Machine Learning approaches, providing valuable guidance for healthcare practitioners to enhance efficiency and precision in disease diagnosis.

Keywords: typhoid fever, Machine Learning, K-Nearest Neighbor (KNN), SMOTE