

## ABSTRAK

**Mahendra , I Gusti Agung Putu** (2024), Peningkatan Kinerja Metode Naïve Bayes Menggunakan Adaboost Untuk Klasifikasi Pada Dataset Penyakit Diabetes Melitus Tipe II. Tesis, Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha

Tesis ini sudah disetujui dan diperiksa oleh Pembimbing I: Dr. I Gede Aris Gunadi, S.Si. M.Kom. dan Pembimbing II : Dr. I Made Agus Wirawan, S.Kom., M.Cs.

*Kata-kata kunci:* Diabetes, Naïve Bayes, Adaboost, Cross Validation, Confusion Matrix

Pemanfaatan teknologi, terutama di ilmu kesehatan, semakin disesuaikan dengan penggunaan model *machine learning*. Penerapan pembelajaran mesin dapat menyederhanakan proses prediksi dan penanganan penyakit. Terdapat beberapa metode yang digunakan dalam klasifikasi penyakit diabetes melitus adalah *Random Forest*, *Naïve Bayes*, *Id3*, *C4.5* dan lain-lain. Optimasi *Naïve Bayes* dengan *Adaboost* diperlukan karena meskipun *Naïve Bayes* memiliki keunggulan minim parameter, namun rentan terhadap akurasi yang terpengaruh oleh jumlah fitur yang terlalu banyak. *Adaboost* digunakan untuk mengatasi sensitivitas terhadap jumlah fitur yang berlebihan, meningkatkan performa model, dan mengoptimalkan kemampuannya menghadapi dataset yang kompleks. Penelitian ini bertujuan untuk menganalisis hasil klasifikasi metode *Naive Bayes* dengan bantuan metode *Adaboost*. Penelitian ini memiliki tahapan yang terdiri dari pengumpulan data, data *preprocessing*, modeling, dan tahap evaluasi. Data yang digunakan berasal dari rekam medis pasien Puskesmas I, II, III Kecamatan Mengwi. Data tersebut telah diolah (*preprocessing*) dengan langkah-langkah seperti menghapus nilai kosong, memeriksa konsistensi data, dan memperbaiki kesalahan penulisan dan format yang tidak seragam. Proses klasifikasi menggunakan metode *Naïve Bayes* dan *Naïve Bayes* dengan *Adaboost*, yang selanjutnya dievaluasi menggunakan *confusion matrix*. Dalam pengujian, dua skenario digunakan: *Naïve Bayes* dan *Naïve Bayes* berbasis *Adaboost*. Algoritma-algoritma tersebut diimplementasikan langsung pada dataset dengan menggunakan *cross validation* dengan nilai k-fold 5 dan k-fold 10. Hasil evaluasi menunjukkan bahwa metode *Naïve Bayes* mengalami peningkatan akurasi sebesar 5.92% pada k-fold 5 dan 5.93% pada k-fold 10 pada dataset dengan 890 data, metode *Naïve Bayes* mengalami peningkatan dengan penambahan metode *Adaboost* dalam klasifikasi penyakit diabetes terbukti memberikan nilai akurasi yang lebih tinggi dibandingkan dengan *Naïve Bayes* tanpa *Adaboost*.

## ABSTRACT

**Mahendra, I Gusti Agung Putu** (2024), *Performance Enhancement of Naïve Bayes Method Using Adaboost for Classification in Type II Diabetes Mellitus Disease Dataset*. Thesis, Computer Science, Graduate Program, Ganesha University of Education.

*This thesis has been approved and examined by Supervisor I: Dr. I Gede Aris Gunadi, S.Si., M.Kom., and Supervisor II: Dr. I Made Agus Wirawan, S.Kom., M.Cs.*

*Keywords: Diabetes, Naïve Bayes, Adaboost, Cross Validation, Confusion Matrix.*

*The use of technology, especially in health sciences, is increasingly being adapted to the use of Machine Learning models. The application of machine learning can simplify the process of predicting and treating disease. There are several methods used in classifying diabetes mellitus, namely Random Forest, Naïve Bayes, Id3, C4.5 and others. Naïve Bayes optimization with Adaboost is necessary because although Naïve Bayes has the advantage of minimal parameters, it is susceptible to accuracy being affected by too many features. Adaboost is used to overcome sensitivity to an excessive number of features, improve model performance, and optimize its ability to deal with complex datasets. This research aims to analyze the classification results of the Naive Bayes method with the help of the Adaboost method. This research has stages consisting of data collection, data preprocessing, modeling, and evaluation stages. The data used comes from medical records of patients at Puskesmas I, II, III Mengwi District. The data has been processed (preprocessing) with steps such as deleting empty values, checking data consistency, and correcting writing errors and non-uniform formatting. The classification process uses the Naïve Bayes method and Naïve Bayes with Adaboost, which is then evaluated using a confusion matrix. In testing, two scenarios were used: Naïve Bayes and Adaboost-based Naïve Bayes. These algorithms are implemented directly on the dataset using cross validation with k-fold values of 5 and k-fold 10. The evaluation results show that the Naïve Bayes method has increased accuracy by 5.92% at k-fold 5 and 5.93% at k-fold 10 on a dataset with 890 data, the Naïve Bayes method has improved with the addition of the Adaboost method in the classification of diabetes which is proven to provide higher accuracy values compared to Naïve Bayes without Adaboost.*