

ABSTRAK

Prayoga, I Kadek Yogi (2025), Klasifikasi Penyakit Demam Berdarah Menggunakan Algoritma Stacking Ensemble Learning. Tesis, Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha

Tesis ini telah disetujui dan diperiksa oleh Pembimbing I: Dr. I Made Gede Sunarya, S.Kom., M.Cs. dan Pembimbing II: Dr. Putu Hendra Suputra, S.Kom., M.Cs.

Kata-kata kunci: demam berdarah, *decision tree*, *naïve naves*, *KNN*, *stacking ensemble learning*

Ensemble Learning digunakan untuk meningkatkan kinerja klasifikasi dengan mengintegrasikan beberapa metode klasifikasi. Ada banyak teknik yang terkait dengan metode ini, seperti *Stacking*, *Boosting* dan *Bagging*. Dalam *Stacking Ensemble Learning*, *stacking* adalah teknik yang menggabungkan beberapa metode dasar (*base-learner*) yang dilatih untuk menghasilkan hasil akhir menjadi metode akhir (*meta-learner*). Tujuan dari penelitian ini adalah untuk mengevaluasi efektivitas klasifikasi menggunakan pengklasifikasi singular dibandingkan dengan *Stacking Ensemble Learning*. Penelitian ini menerapkan banyak prosedur, termasuk pengumpulan data, prapemrosesan data, pemodelan, dan evaluasi. Data yang akan digunakan diperoleh dari rekam medis pasien di Puskesmas I, Kecamatan Mengwi. Selanjutnya, data tersebut menjalani pemrosesan tambahan (*preprocessing*), membuang data yang memiliki nilai kosong. Data tersebut kemudian diperiksa untuk mengidentifikasi ketidakkonsistenan dan memperbaiki kesalahan, termasuk kesalahan ketik dan ketidakkonsistenan format. Teknik yang telah ditentukan sebelumnya digunakan untuk melakukan tahap pemodelan, dan hasilnya kemudian dinilai menggunakan *confusion-matrix* sebagai alat ukur. Metode *Stacking Ensemble Learning* digunakan dalam proses klasifikasi, dan selanjutnya dinilai menggunakan *confusion-matrix*. Pengujian ini menggunakan dua skenario: Klasifikasi menggunakan *single-classifier* dan *Stacking Ensemble Learning* dengan Logistic Regression. *Cross Validation* digunakan untuk langsung menerapkan algoritma ini pada dataset, dengan nilai k-fold 5 dan 10. Hasil evaluasi menunjukkan bahwa metode *Stacking Ensemble Learning* telah meningkat sebesar 8,89% dari metode *single-classifier* terendah, *Decision tree* dengan k-fold 5, dan sebesar 11,08% dari metode *single-classifier* terendah, *Decision tree* dengan k-fold 10, pada dataset 650 data.

ABSTRACT

Prayoga, I Kadek Yogi (2025), *Classification of Dengue Fever Disease Using Stacking Ensemble Learning Algorithm*. Thesis, Computer Science, Master Program, Ganesha University of Education.

This thesis has been approved and examined by Supervisor I: Dr. I Made Gede Sunarya, S.Kom., M.Cs., and Supervisor II: Dr. Putu Hendra Suputra, S.Kom., M.Cs..

Keywords: dengue fever, decision tree, naïve naves, KNN, stacking ensemble learning

Ensemble Learning is used to improve classification performance by integrating multiple classification methods. There are many techniques associated with this method, such as Stacking, Boosting and Bagging. In Stacking Ensemble Learning, stacking is a technique that combines several base methods (base-learners) that are trained to produce the final result into the final method (meta-learner). The purpose of this study is to evaluate the effectiveness of classification using a singular classifier compared to Stacking Ensemble Learning. This study applies many procedures, including data collection, data preprocessing, modeling, and evaluation. The data to be used were obtained from patient medical records at Puskesmas I, Mengwi District. Furthermore, the data undergoes additional processing (preprocessing), discarding data that has empty values. The data is then checked to identify inconsistencies and correct errors, including typos and format inconsistencies. A predetermined technique is used to perform the modeling stage, and the results are then assessed using a confusion matrix as a measuring tool. The Stacking Ensemble Learning method is used in the classification process and is then assessed using a confusion matrix. This test uses two scenarios: Classification using a single-classifier and Stacking Ensemble Learning with Logistic Regression. Cross Validation is used to directly apply this algorithm to the dataset, with k-fold values of 5 and 10. The evaluation results show that the Stacking Ensemble Learning method has increased by 8.89% from the lowest single-classifier method, Decision tree with k-fold 5, and by 11.08% from the lowest single-classifier method, Decision tree with k-fold 10, on a dataset of 650 data.