

INTISARI

IMPLEMENTASI *LIGHT GRADIENT BOOSTING MACHINE* (LIGHTGBM) PADA ANALISIS KLASIFIKASI DATA TIDAK SEIMBANG

Oleh

Dina Lestari

20/462309/PA/20281

Pada analisis klasifikasi, permasalahan yang kerap muncul ialah kondisi data tidak seimbang atau *imbalance data*. Data tidak seimbang terjadi ketika suatu kelas data memiliki jumlah data yang berbeda dibandingkan dengan kelas lainnya. Keadaan ini berdampak pada kinerja algoritma *machine learning* karena model akan terfokus pada jumlah kelas yang lebih banyak atau kelas mayoritas dibandingkan dengan kelas minoritas yang dapat menyebabkan pengambilan keputusan yang bias. Penanganan dengan pendekatan algoritma yang digunakan untuk menyelesaikan permasalahan tersebut adalah *ensemble learning* dan modifikasinya. *Light Gradient Boosting Machine* (LightGBM) adalah algoritma dalam *ensemble learning* yang memiliki kerangka yang cepat dan efisiensi dalam komputasi. Modifikasi yang dilakukan ialah dengan menggunakan LightGBM Weighted dan LightGBM Focal Loss. Pada tugas akhir ini dilakukan analisis pada data HTRU2 dan data *Default of Credit Card* dengan menggunakan LightGBM dan modifikasinya, kemudian dibandingkan dengan model dasarnya yakni *decision tree* dan model *ensemble learning* yang lain yakni *Gradient Boosting*. Dari analisis yang dilakukan diperoleh kesimpulan bahwa LightGBM Weighted menghasilkan performa yang paling baik.

ABSTRACT

IMPLEMENTATION OF LIGHT GRADIENT BOOSTING MACHINE (LIGHTGBM) FOR IMBALANCE DATA CLASSIFICATION ANALYSIS

By

Dina Lestari

20/462309/PA/20281

In classification analysis, a common issue is the presence of imbalanced data. Imbalanced data occurs when one class of data has a different number of samples compared to other classes. This condition affects the performance of machine learning algorithms because the model tends to focus more on the majority class rather than the minority class, which can lead to biased decision-making. The algorithmic approach used to address this issue includes ensemble learning and its modifications. Light Gradient Boosting Machine (LightGBM) is an ensemble learning algorithm known for its speed and computational efficiency. Modifications include using Weighted LightGBM and LightGBM with Focal Loss. In this case study, an analysis was conducted on HTRU2 and Default of Credit Card data using LightGBM and its modifications, and then compared with the basic model, which is the decision tree, and another ensemble learning model, Gradient Boosting. From the analysis, it was concluded that Weighted LightGBM produced the best performance.