ABSTRACT

Traffic congestion is a major issue in many nations, especially emerging nations, and it leads to negative consequences such air pollution, increased fuel use, and longer travel times. Solving these challenges often needs the enhancement of public transit usage, which demands superior service quality, especially accurate bus schedule predictions. Accurate travel time predictions may increase user happiness by providing reliable information about bus schedules. Prediction accuracy is influenced by environmental factors, such as the presence of dedicated bus lanes or shared lanes with private vehicles, which create dynamic conditions. Previous research [1] showed the use of machine learning techniques such as Extreme Gradient Boost, Artificial Neural Networks (ANN), and Long Short Term Memory (LSTM) models for predicting travel time on GPS-based information. Despite showing robust performance in Dublin City, their models were limited to that specific dataset and situation.

This research aims to utilize machine learning techniques XGBoost, ANN, and LSTM on a dataset customized for the bus routes in Bandung City, focusing on the Baleendah-Bandung Electronic Center route. The collection contains time data for three specific bus stops showing high passenger transfer rates. The assessment of model performance will use metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), R-squared (R^2), Mean Absolute Percentage Error (MAPE), and Median Absolute Error (MdAE) to evaluate prediction accuracy. This project aims to minimize prediction errors while developing a city-specific dataset for Bandung to improve compatibility with local conditions.

Keywords: Artificial Neural Networks (ANN), bus travel time prediction, Long Short Term Memory (LSTM), traffic congestion, XGBoost.