ABSTRACT

Tea is one of the most widely consumed beverages in the world and has many health benefits. Tea quality assessment is generally carried out by experts based on aroma and taste, but this approach is subjective, requires special expertise, and tends to produce inconsistent evaluations. To overcome these limitations, the Electronic Nose (E-Nose) is used as a tool to analyze tea aroma composition more objectively. In this study, the E-Nose dataset is utilized to perform multiclass classification of Gambung green tea composition using a machine learning algorithm. The dataset used consists of 23,220 rows with 6 multiclass prediction targets A, C, D, E, and F. Several machine learning algorithms such as Random Forest, Bagging Decision Tree, Adaboost, Gradient Boosting, XGBoost, and LightGBM are optimized with hyperparameter tuning using Grid Search to improve classification accuracy, with evaluation using metrics such as accuracy, precision, recall, and F1-score. In addition, the Explainable AI (XAI) approach using the LIME method is employed to improve model transparency and interpretability, so that prediction results are easier to understand and analyze. Among all models, Random Forest performed best with 99.20% accuracy. XAI with the LIME method produced visualizations of prediction probabilities, feature contributions, and feature values for each instance of the model's prediction results. The MQ3 sensor was the most dominant feature with a major positive contribution in almost all classes. This research contributes to the development of an automated system for more objective and consistent tea classification by combining E-Nose technology, machine learning, and XAI.

Keywords: Tea, Classification, Electronic-Nose, Explainable AI, Machine Learning, Random Forest