ABSTRACT

Tea quality, influenced significantly by aroma, is critical for consumer satisfaction. This paper introduces an innovative method by integrating Electronic Nose (E-Nose) technology with Bagging Algorithms to predict organoleptic scores in organic green teas. The study evaluates various regression models, focusing on optimizing hyperparameters within the Bagging framework. Among these models, Random Forest emerges as robust, showing exceptional predictive accuracy with low Mean Squared Error (MSE) and high R-squared values. Random Forest Regression excels at capturing intricate dataset patterns, with its ensemble approach effectively reducing overfitting and ensuring reliable predictions in complex scenarios. In contrast, the Decision Tree Regressor, despite a high R-squared value, exhibits higher MSE, indicating less precision. Support Vector Regressor (SVR) and Neural Network Regression encounter challenges, highlighting the need for models adept at nonlinear relationships. This research underscores the effectiveness of Bagging Algorithms, particularly Random Forest ensembles, in predicting organoleptic scores, offering insights for model selection in regression tasks. Beyond methodology, implications extend to the tea industry, where these models promise to transform quality control and optimize production. This study pioneers the use of Bagging Algorithms for precise tea quality assessment, aligning with efforts to elevate industry standards.

Keywords—Tea quality, Organoleptic scores, Electronic Nose (E-Nose), Machine learning, Bagging Algorithms