

Task 16: Hyperparameter Tuning using GridSearchCV

Report

Objective:

The objective of this task is to improve machine learning model performance by identifying optimal hyperparameters using GridSearchCV with cross-validation.

Tools & libraries:

- Python
- scikit-learn
- Pandas
- Jupiter Notebook

Objectives:

The task began by loading the Breast Cancer dataset and preparing it for model training. The dataset was split into training and testing sets to evaluate model performance on unseen data. A baseline Random Forest model was first trained using default hyperparameters to establish reference performance. A parameter grid was then defined containing different values for key hyperparameters such as number of estimators, maximum depth, and minimum samples split. GridSearchCV with cross-validation was applied to systematically test all parameter combinations and identify the optimal configuration. The best model obtained from GridSearchCV was used for prediction on the test set. Finally, the performance of the tuned model was compared with the baseline model to evaluate improvement.

Conclusion:

In this Task Hyperparameter tuning using GridSearchCV improved the model's predictive performance by identifying the optimal parameter combination. The tuned Random Forest model achieved better accuracy than the default configuration. Cross-validation ensured reliable evaluation and reduced overfitting. This task demonstrates the importance of model optimization using scikit-learn in practical machine learning workflows.