

Assignment -8

Task:-1

Train a Random Forest classifier and a Gaussian Naive Bayes model on the diabetes dataset. Compare their default accuracies using 5-fold cross-validation.

Task-2

Identify the top 3 most important features using Random Forest and visualize their importance.

Task-3

Use GridSearchCV to find the best hyperparameters (n_estimators, max_depth, min_samples_split) for Random Forest.

Task-4

The diabetes dataset is imbalanced. Modify the Random Forest to handle this using class_weight='balanced' and compare results.

Task-5

Combine Random Forest and Naive Bayes predictions using a voting classifier and compare its performance with individual models.