

Task 15

Machine Learning

Upload the .py or .ipynb extension file to GitHub public repo "100DaysofBytewise" and share the link in the submission form by July 26, 2024.

Dataset : Adult Income Dataset

1. Applying Cross-Validation to Random Forest Classifier

- **Exercise:** Implement a random forest classifier and evaluate the model using k-fold cross-validation. Analyze the cross-validation scores to assess model stability.

2. Investigating Overfitting and Underfitting in Gradient Boosting Machines

- **Exercise:** Train a gradient boosting classifier with varying numbers of estimators and learning rates. Evaluate the model for overfitting and underfitting by comparing training and validation performance.

3. Evaluating Precision, Recall, and F1-Score for Random Forests

- **Exercise:** Implement a random forest classifier and calculate precision, recall, and F1-score. Discuss the trade-offs between these metrics and their importance for classification tasks.

4. ROC Curve and AUC for Gradient Boosting Classifier

- **Exercise:** Implement a gradient boosting classifier and plot the ROC curve. Compute the AUC and interpret how well the model distinguishes between classes.

5. Model Performance Comparison with Different Metrics

- **Exercise:** Compare the performance of different classifiers (e.g., SVM, random forest, gradient boosting) using cross-validation. Evaluate and compare the models based on accuracy, precision, recall, F1-score, and ROC-AUC.