

Interpretation and Discussion

- What do the results of the confusion matrix indicate?
The confusion matrix shows how well the model classifies each category. The high number of correct predictions along the diagonal means the model is generally accurate. However, there are still a few misclassifications, suggesting that some classes overlap in features or that the model struggles with borderline cases.
- How consistent is the model's performance based on 5-Fold Cross Validation?
The 5-Fold Cross Validation results show that the model's accuracy, precision, recall, and F-1 scores remain fairly stable across all folds. The small standard deviation values indicate consistent performance, meaning the model generalizes well and is not very dependent on any specific subset of the training data.
- What insights can be derived from the learning curve?
The learning curve shows that training and validation scores converge, suggesting the model is well-fitted and not overfitting.
- How can the model be improved?
The model could be improved by performing hyperparameter tuning, adding more relevant features, or using a different classifier, such as a Decision Tree or Random Forest. Additionally, collecting more balanced data could help improve prediction accuracy, especially for minority classes.