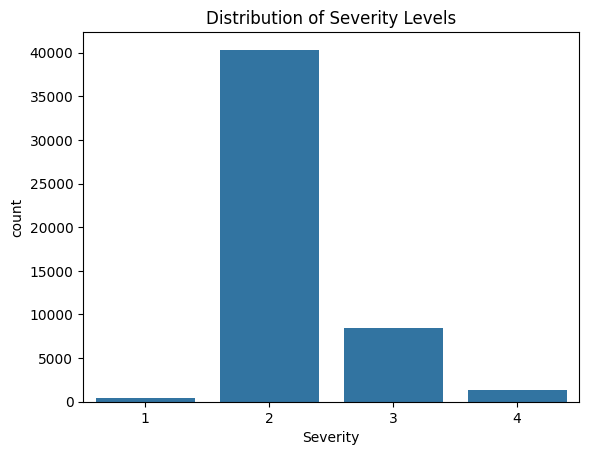
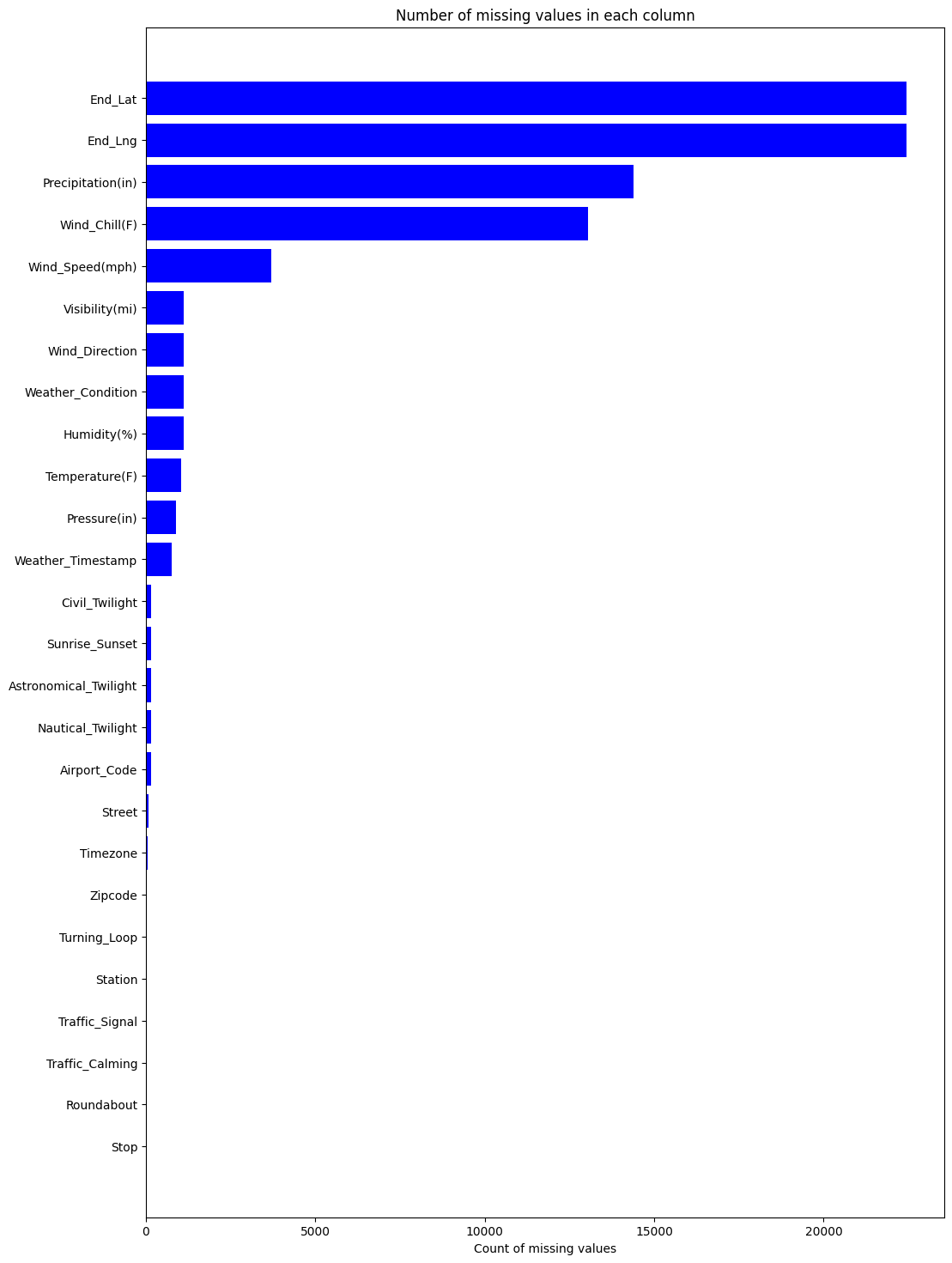
**Results**

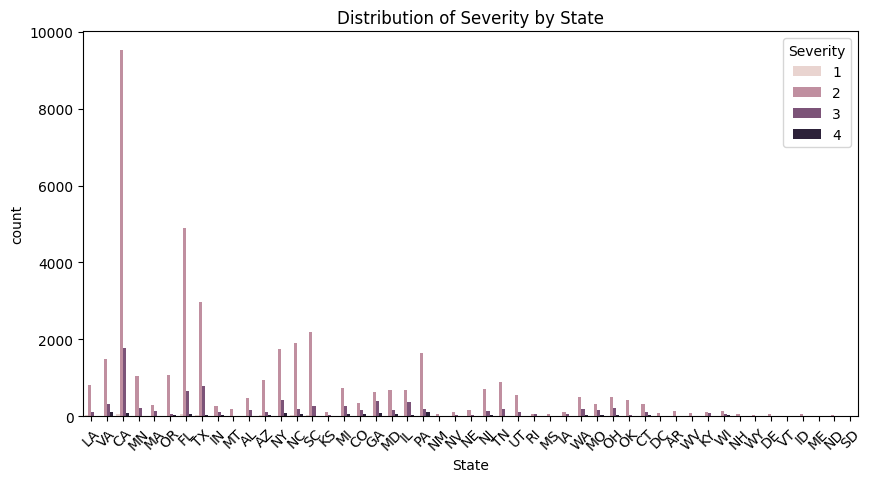
Distribution of Target Variable:

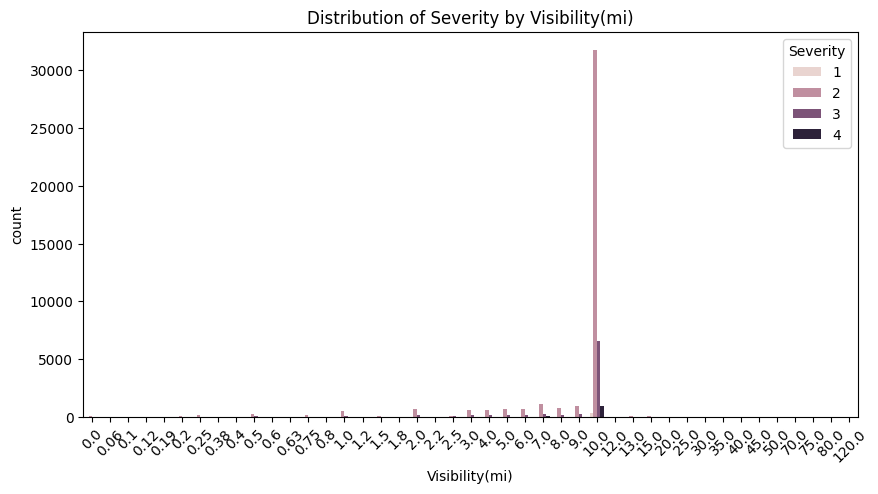


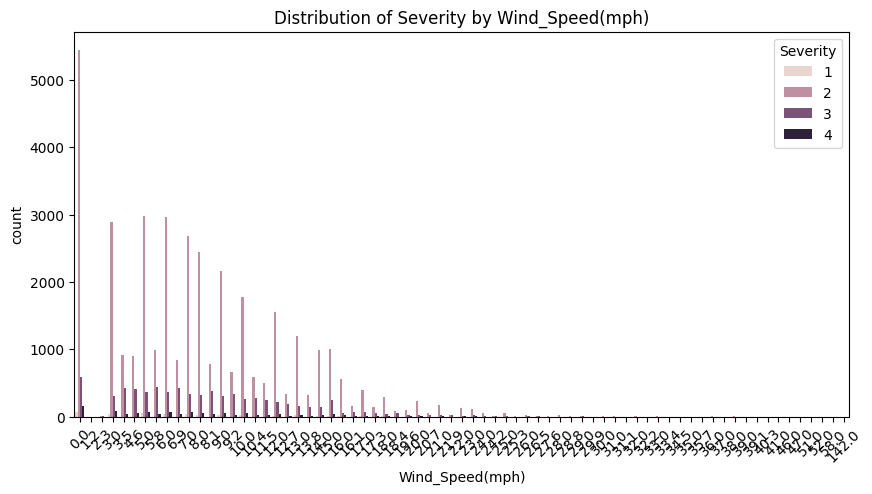
Cleaning of missing values from the dataset:

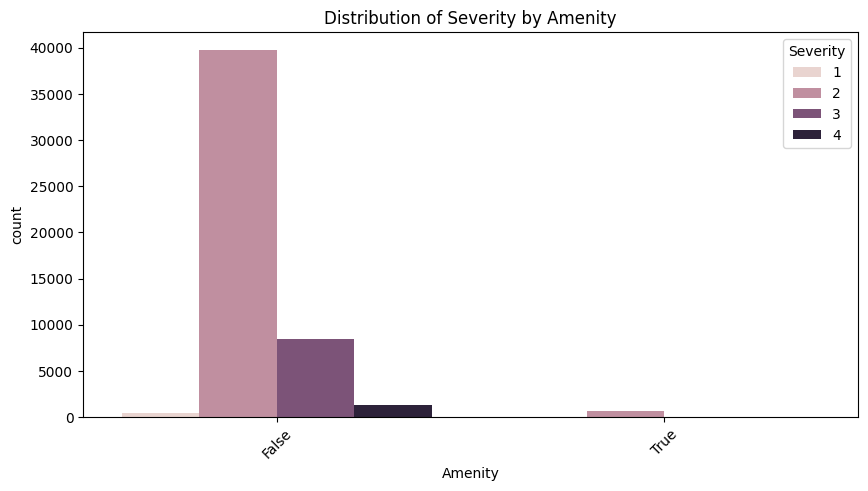


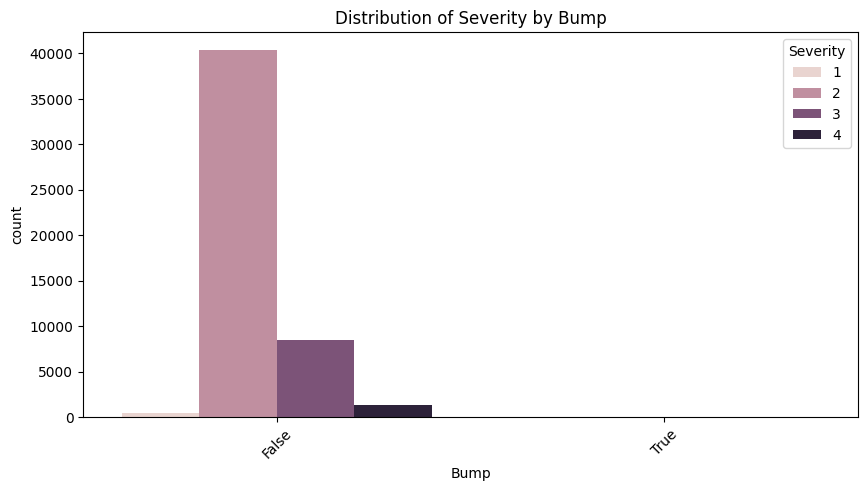
**Research Question 1**: Primary factors contributing to the “Severity” of traffic collisions in the United States?

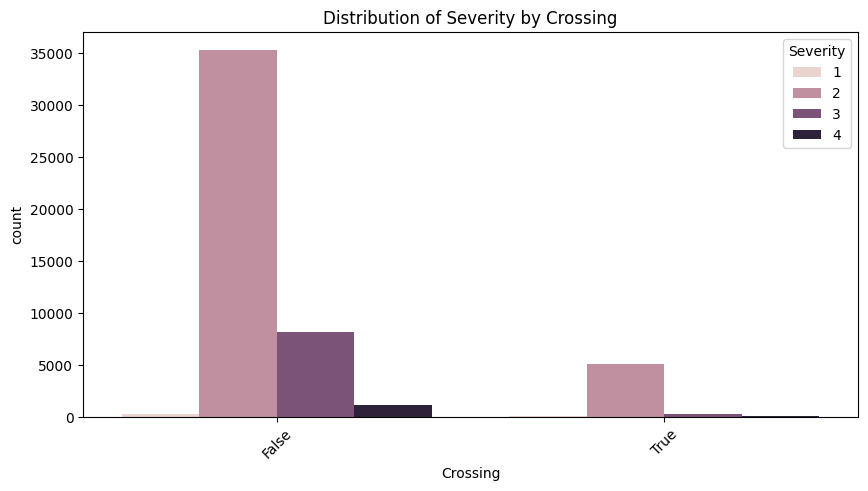


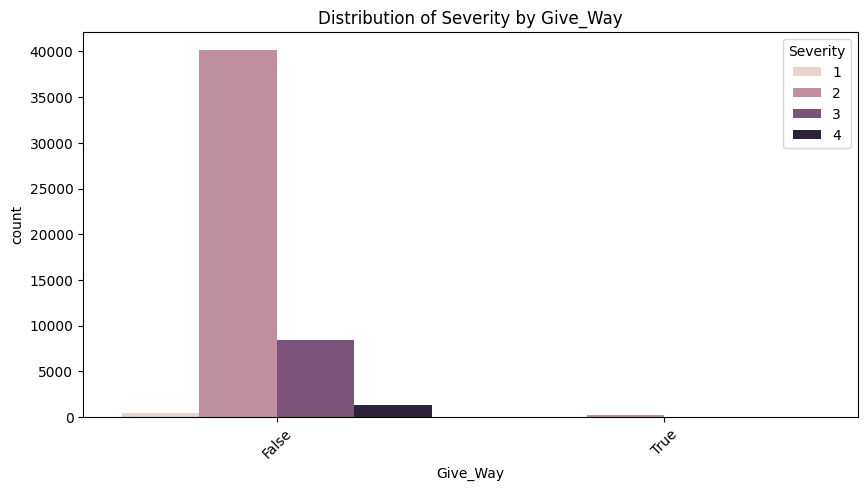


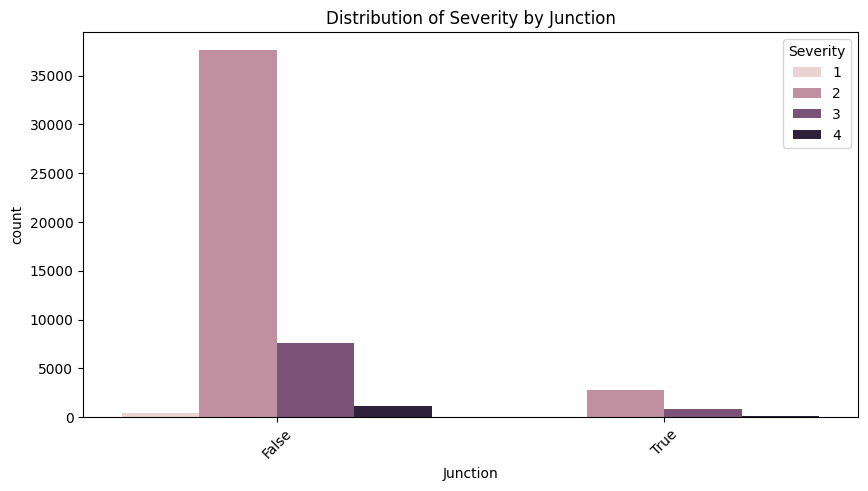


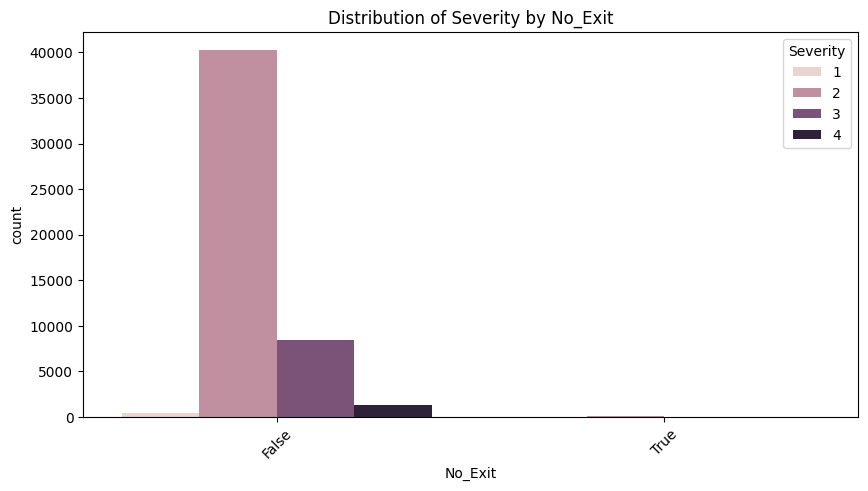


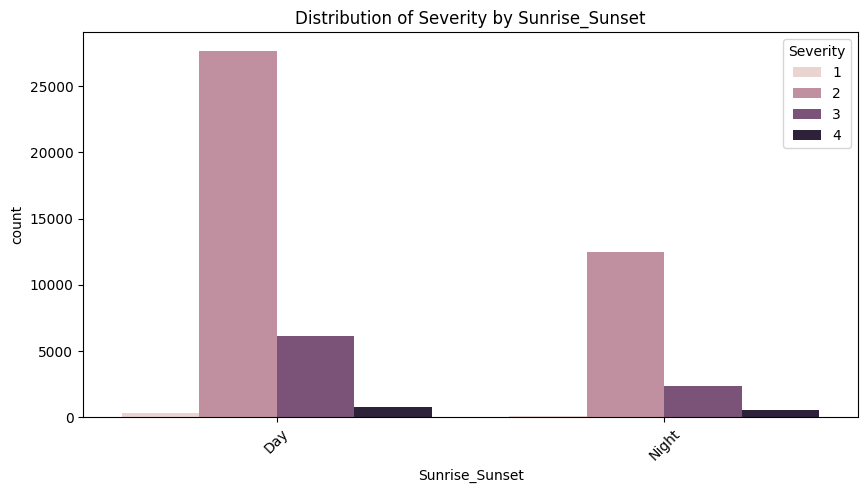






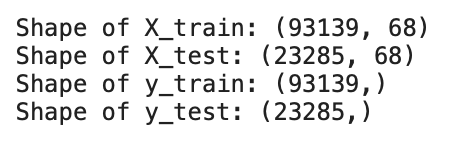






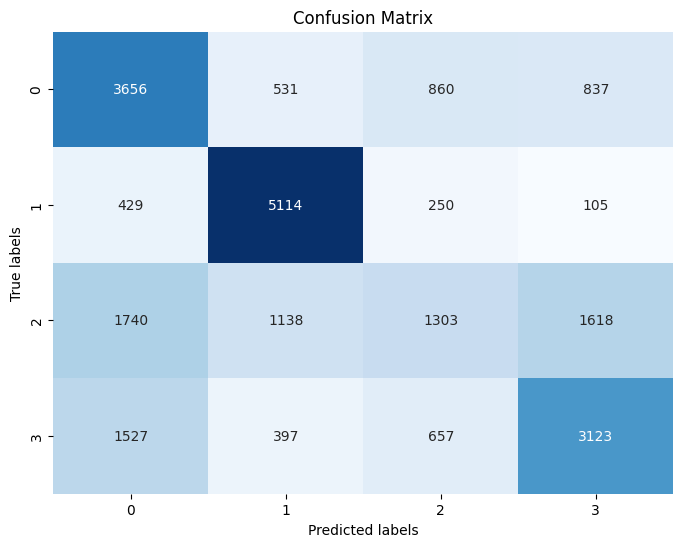
**Research Question 2 and 3:** Can machine learning models effectively classify collision severity based on historical data? How does the model's performance vary across different variants of ML algorithms?

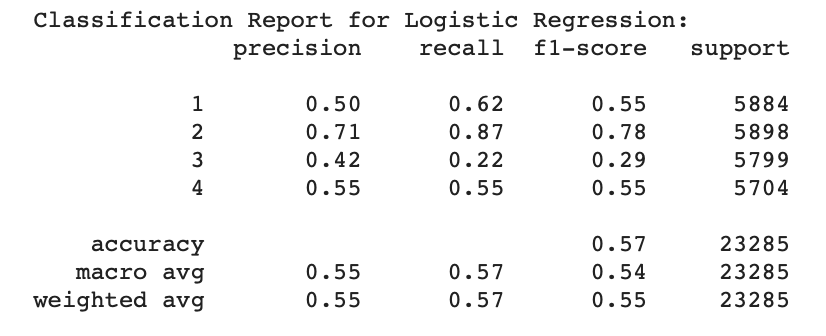
After SMOTE oversampling and 80:20 train test split:



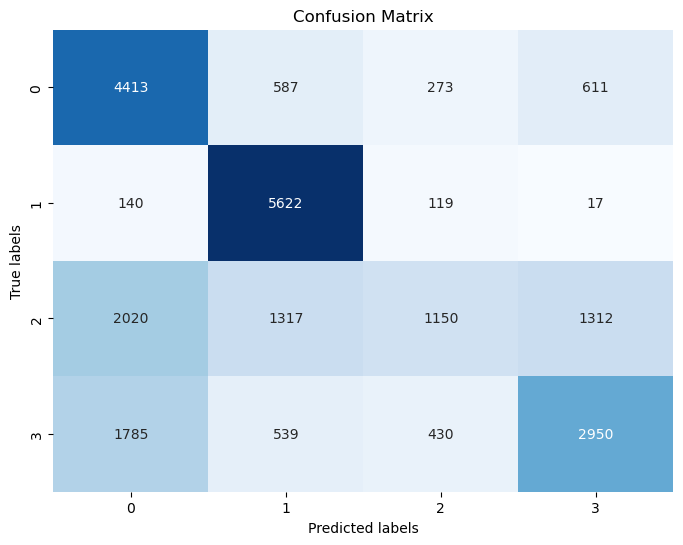
Model Evaluation after PCA:

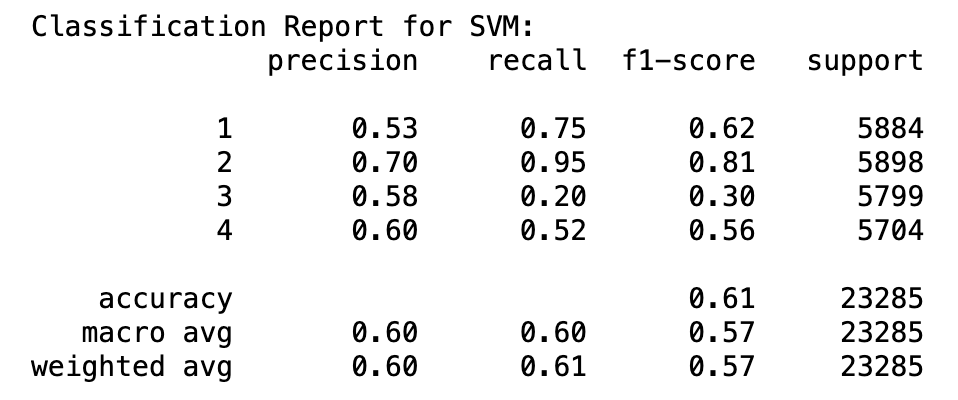
Logistic Regression:



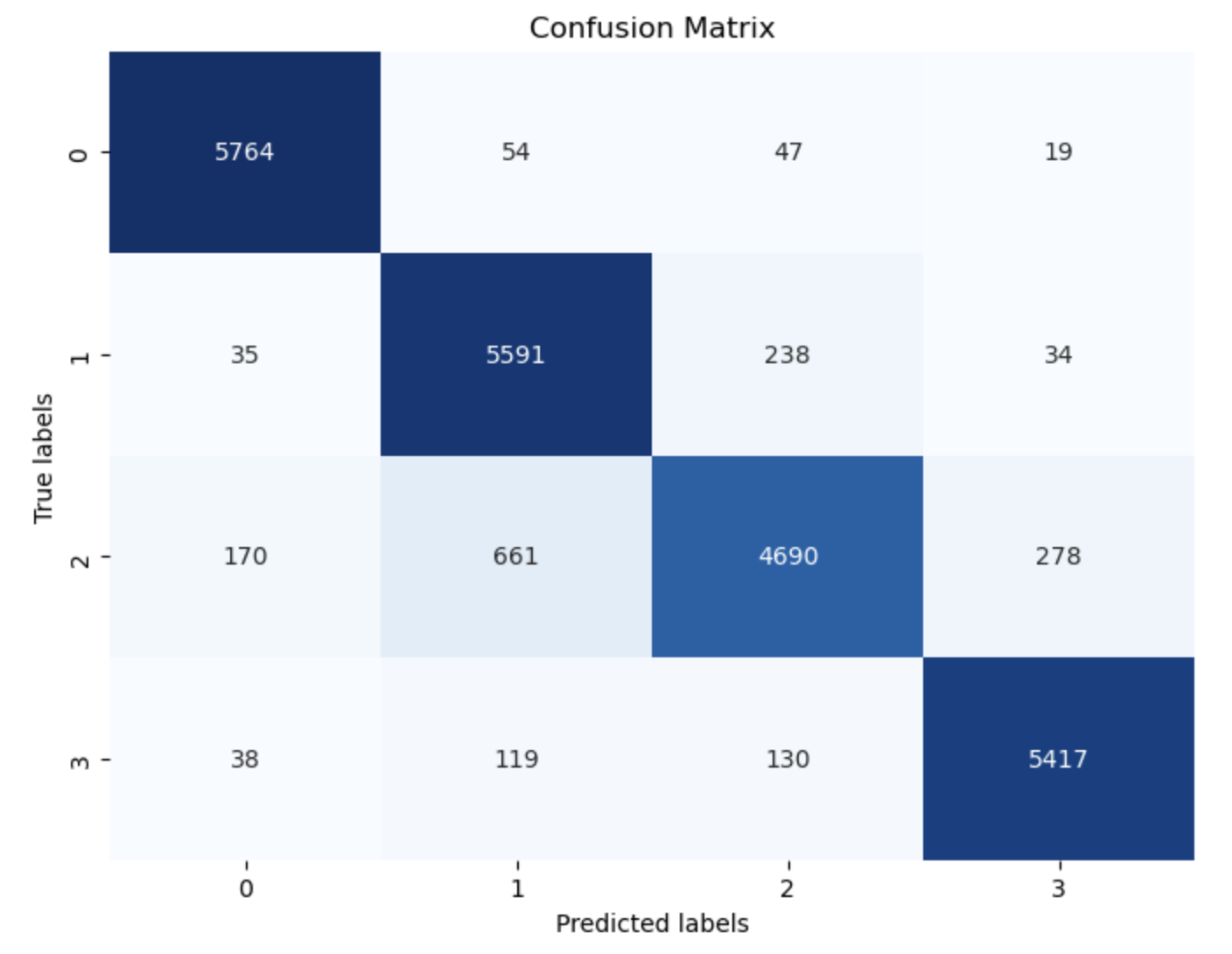


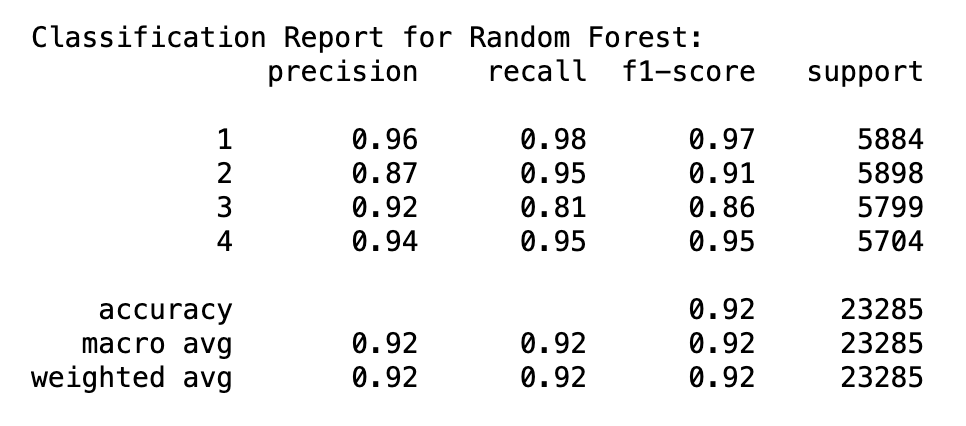
SVM:





Random Forests:





Adaboost:

