

In the data preprocessing part, I noticed that some features were not numeric, and I first converted them to numeric. I then extracted some recurring features from the data while examining the data.

I have observed that the training time of SVM during the training phase is also more than others. At the same time, I observed that it did not give a good result when training models in a normal way without applying a CV during the training phase.

During the testing phase, it was observed that the Logistic regression model gave a better accuracy.

The test times were almost the same. When I retrained the top 5 features with this model and used them for the test, I observed that the training time was reduced by half, but as a result of the test, it decreased by 1 point.

As a result of this experiment, I observed that both the SVM model lasted too long during the training phase and the classification score gave low results compared to other algorithms.