**DATABRICKS NOTEBOOKS**

The first notebook below has the code for training of the models, calculating different accuracy scores as well as displaying of the confusion matrices

* + <https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/3634097844023146/150675221134314/8712538996775317/latest.html>

The next notebook below shows how Data Preprocessing was achieved in PySpark

* + <https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/1222114862673243/2771296308904893/3779371744576095/latest.html>

The next notebook below shows how Misclassification data was retrieved by using PySpark

* + <https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/436154438772259/3379983875769477/40505063256021/latest.html>