

OUTSTANDING-3

Name – Om Patil

Batch - Machine learning (April 2021)

EDA for Gender classification for
Twitter dataset

While performing EDA and analysing the data with using `corr()` function and via some basic knowledge , dropped the unwanted columns from the gender dataset

Null values – Checked for Null values and filled them with appropriate method .

Dummy Variables – used dummies from pandas and made new columns replacing the original columns

and later trained the model via two algorithms

SVM = used Support Vector Machine because it is very efficient and gives better accuracy.