Machine Learning Proyect

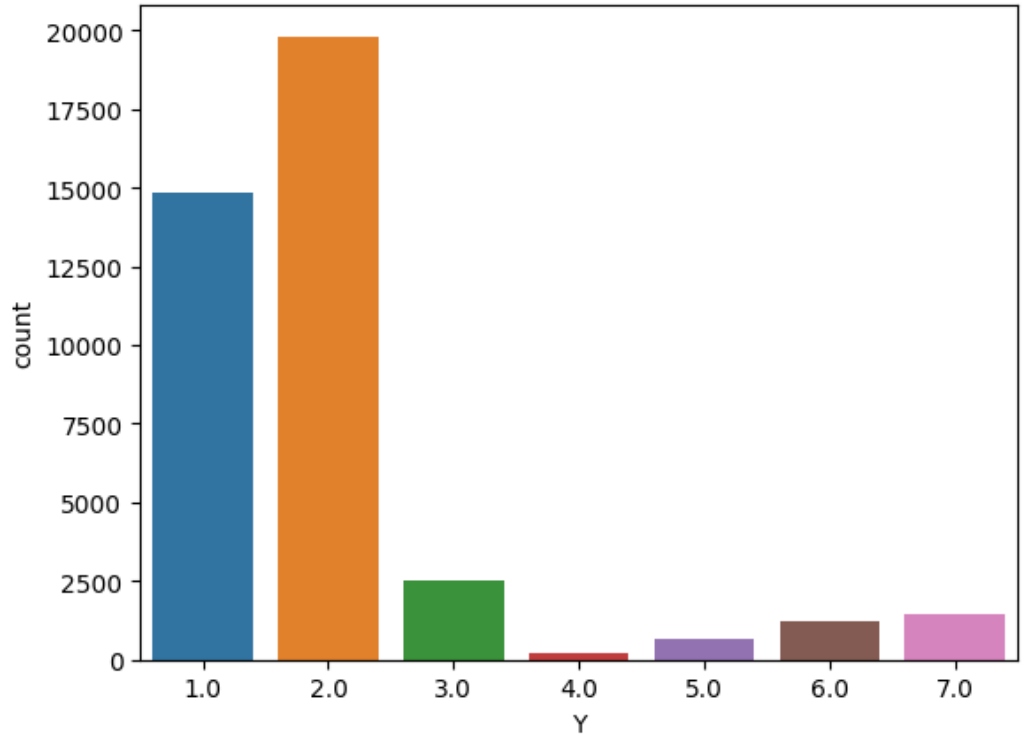
Carlos María del Pino
Lamlih Houssam

UMONS
University of Mons

Index

Exploratory Data Analysis

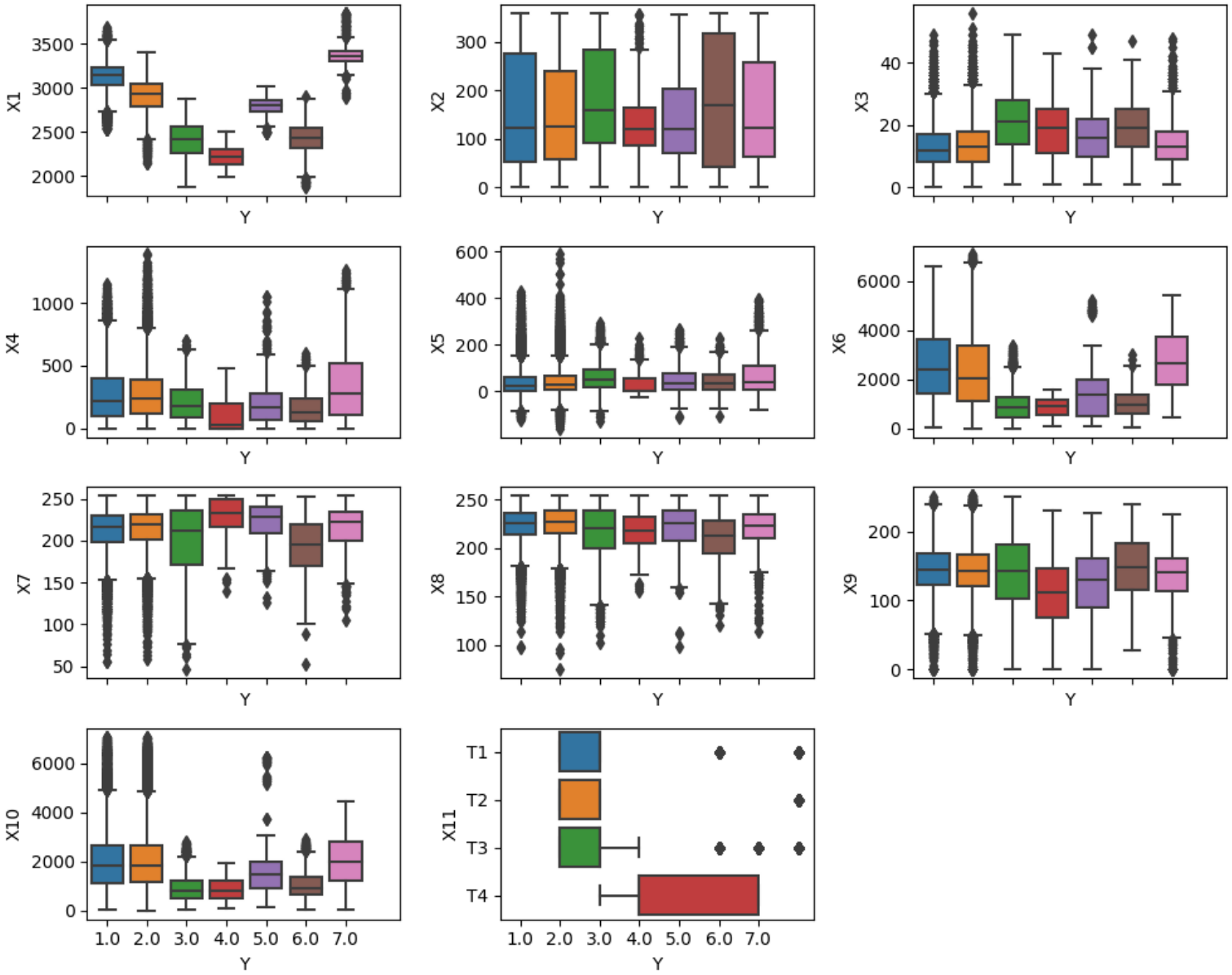
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
0	3215.0	206.0	8.0	127.0	14.0	5588.0	215.0	248.0	168.0	1057.0	T1	T29	1.0
1	3022.0	18.0	7.0	0.0	0.0	3012.0	215.0	226.0	148.0	1476.0	T1	T23	1.0
2	3008.0	18.0	19.0	382.0	92.0	4715.0	201.0	197.0	127.0	3616.0	T1	T29	1.0
3	2595.0	296.0	24.0	85.0	28.0	1252.0	144.0	224.0	216.0	780.0	T3	T33	6.0
4	3261.0	87.0	4.0	42.0	-2.0	3719.0	226.0	233.0	142.0	3784.0	T1	T38	7.0



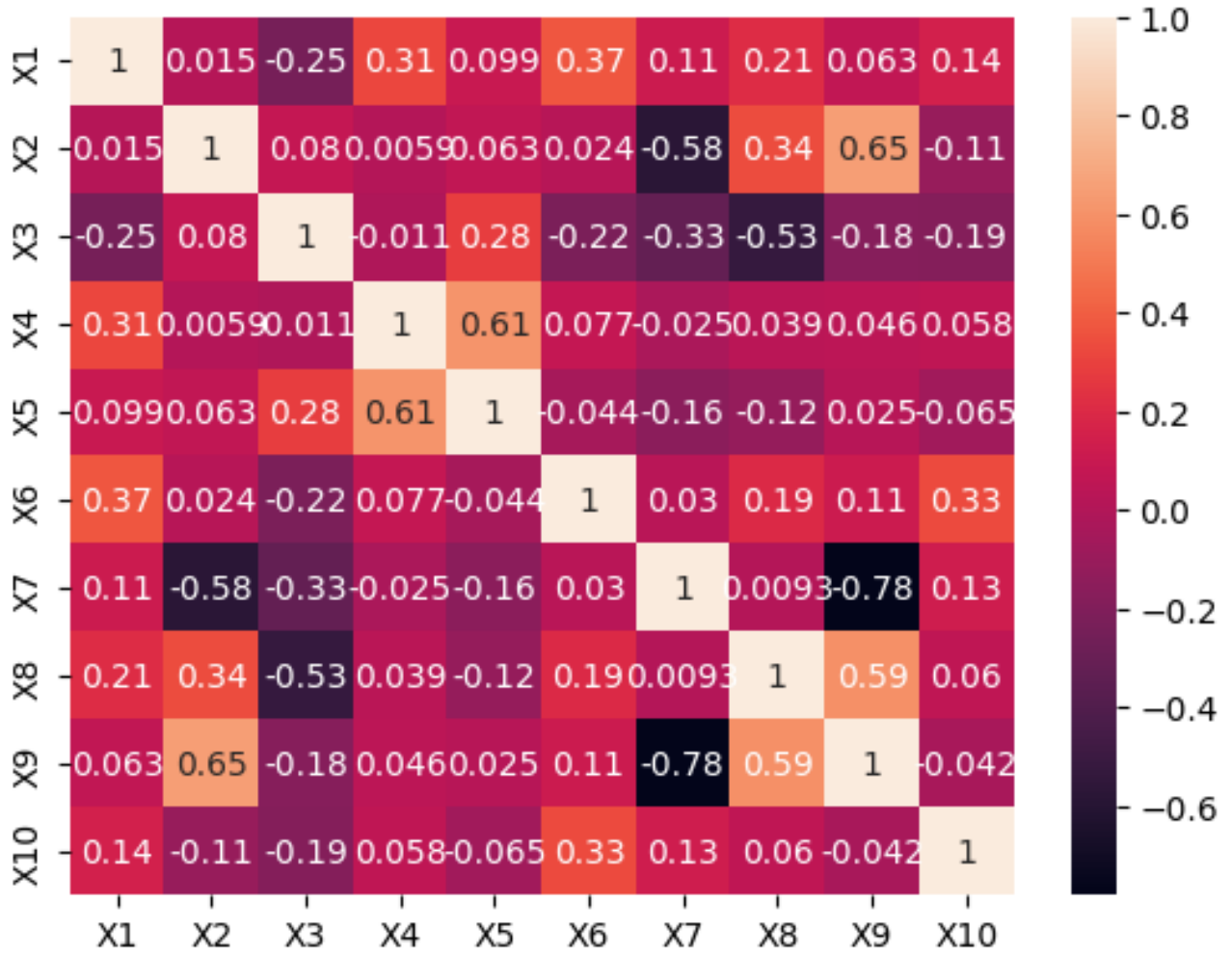
Exploratory Data Analysis

- Initial understanding of the dataset
- X1, X3, X6 or X10 are going to be suitable for using for predicting the value of Y
- X2 or X9 are not adequate to predict the label

Exploratory Data Analysis

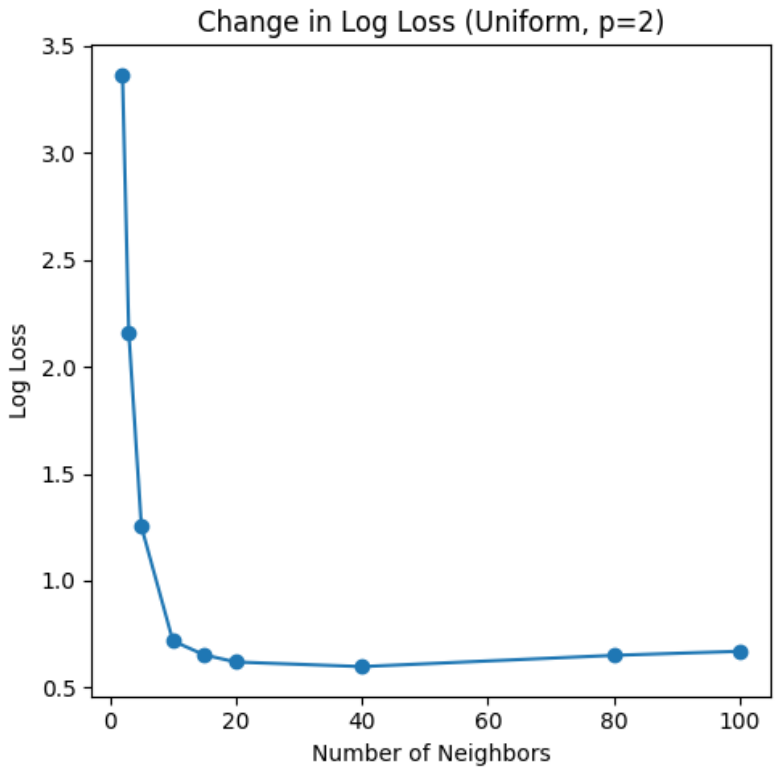


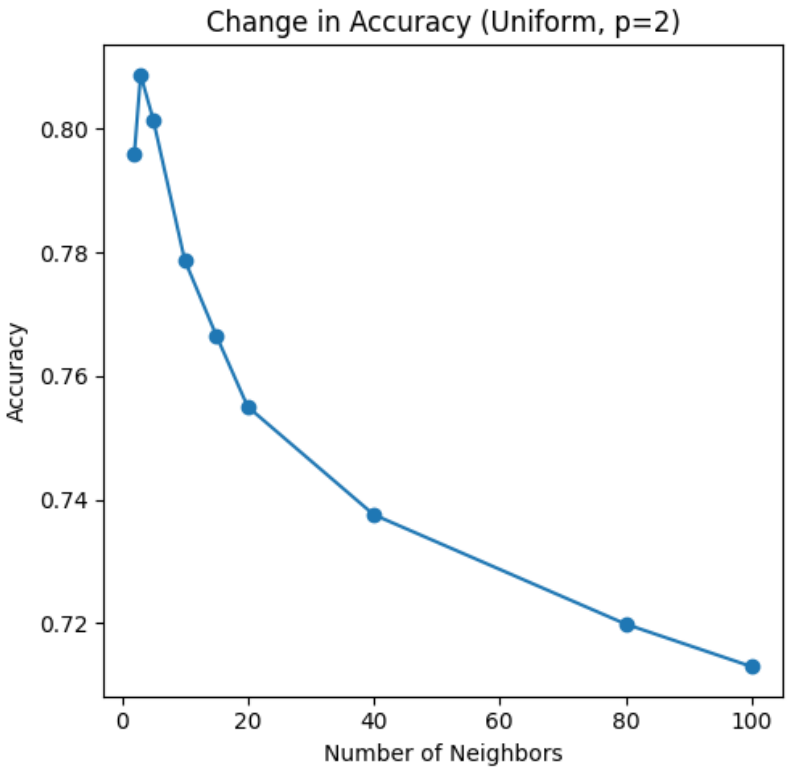
Exploratory Data Analysis



Methodology

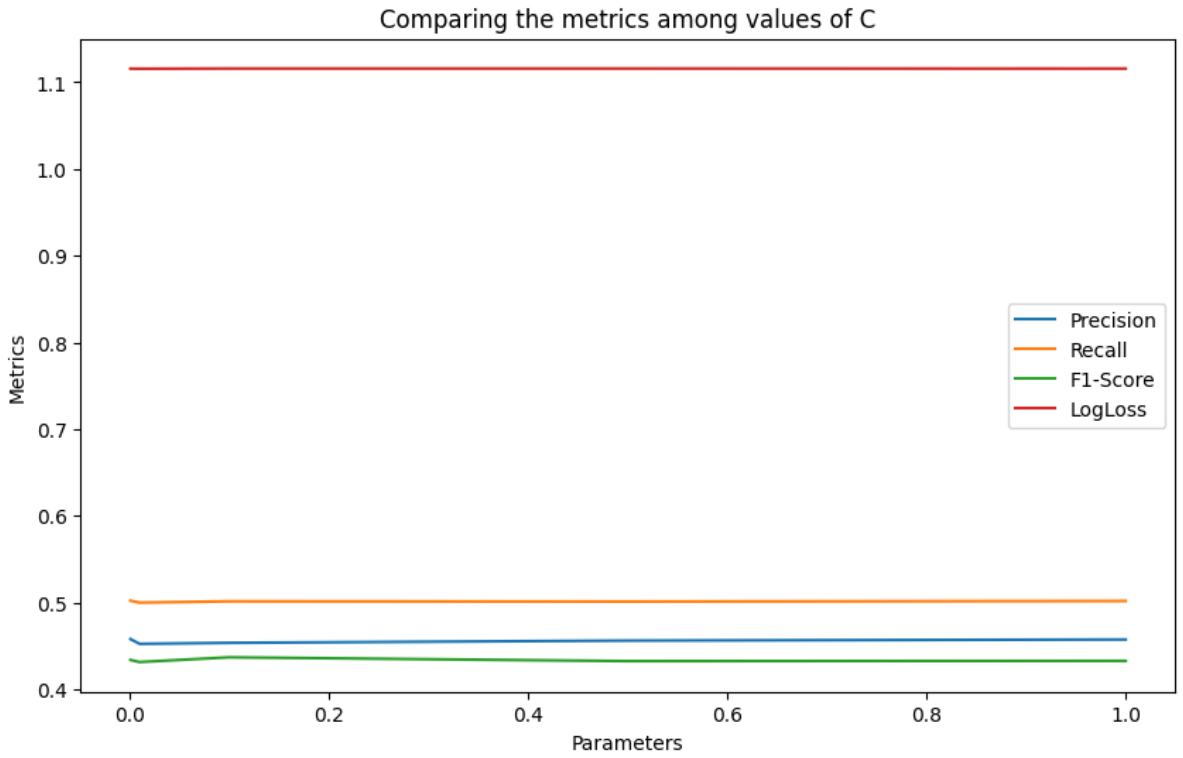
KNN





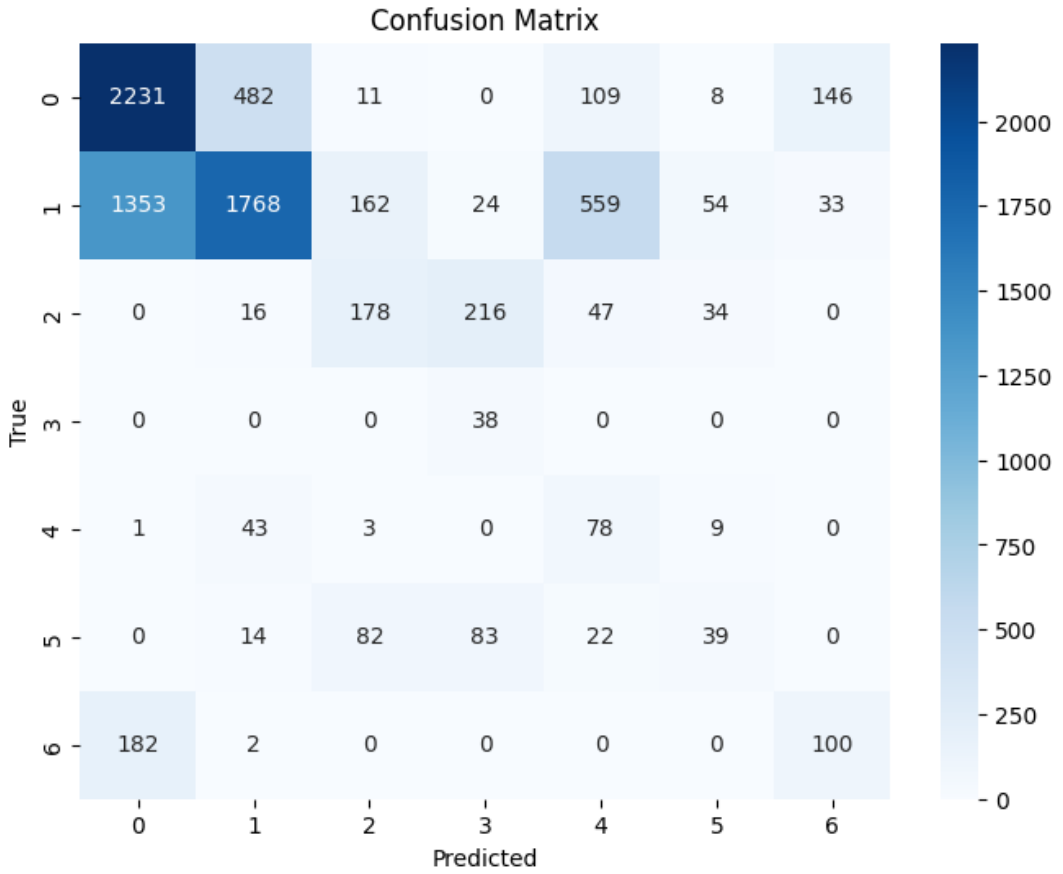
Methodology

Logistic Regression



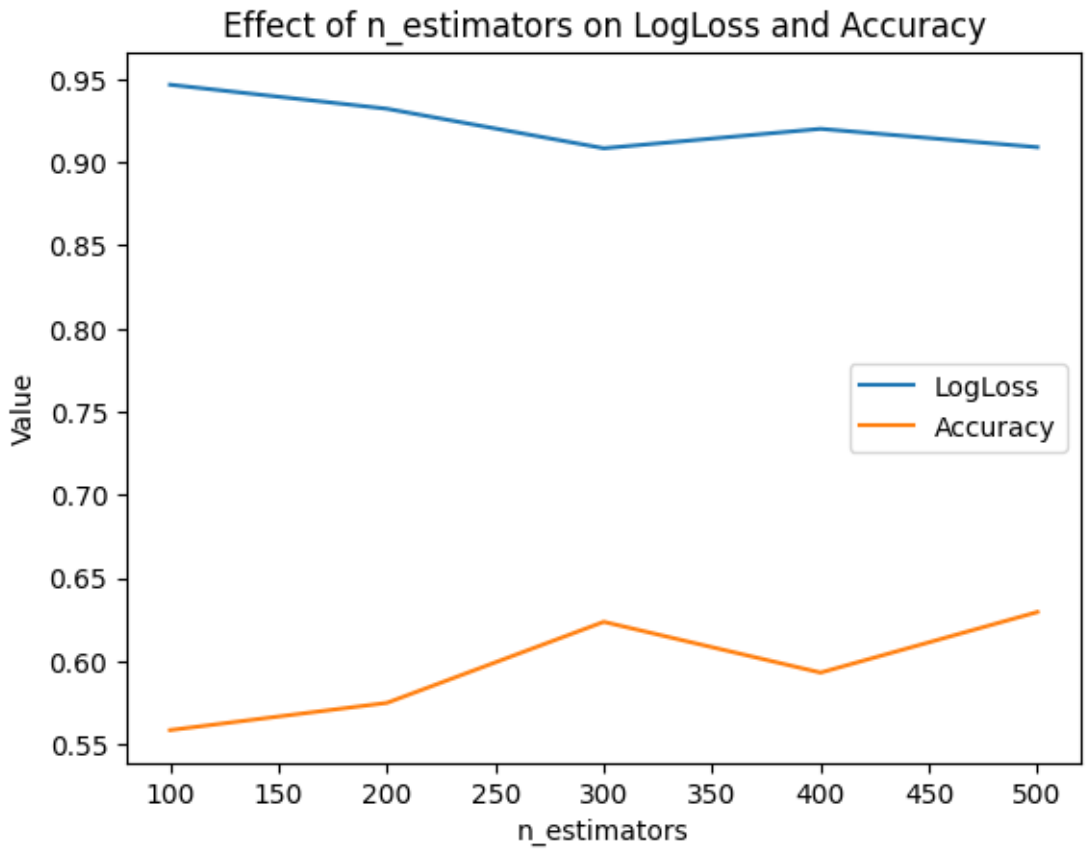
Methodology

Naive Bayes



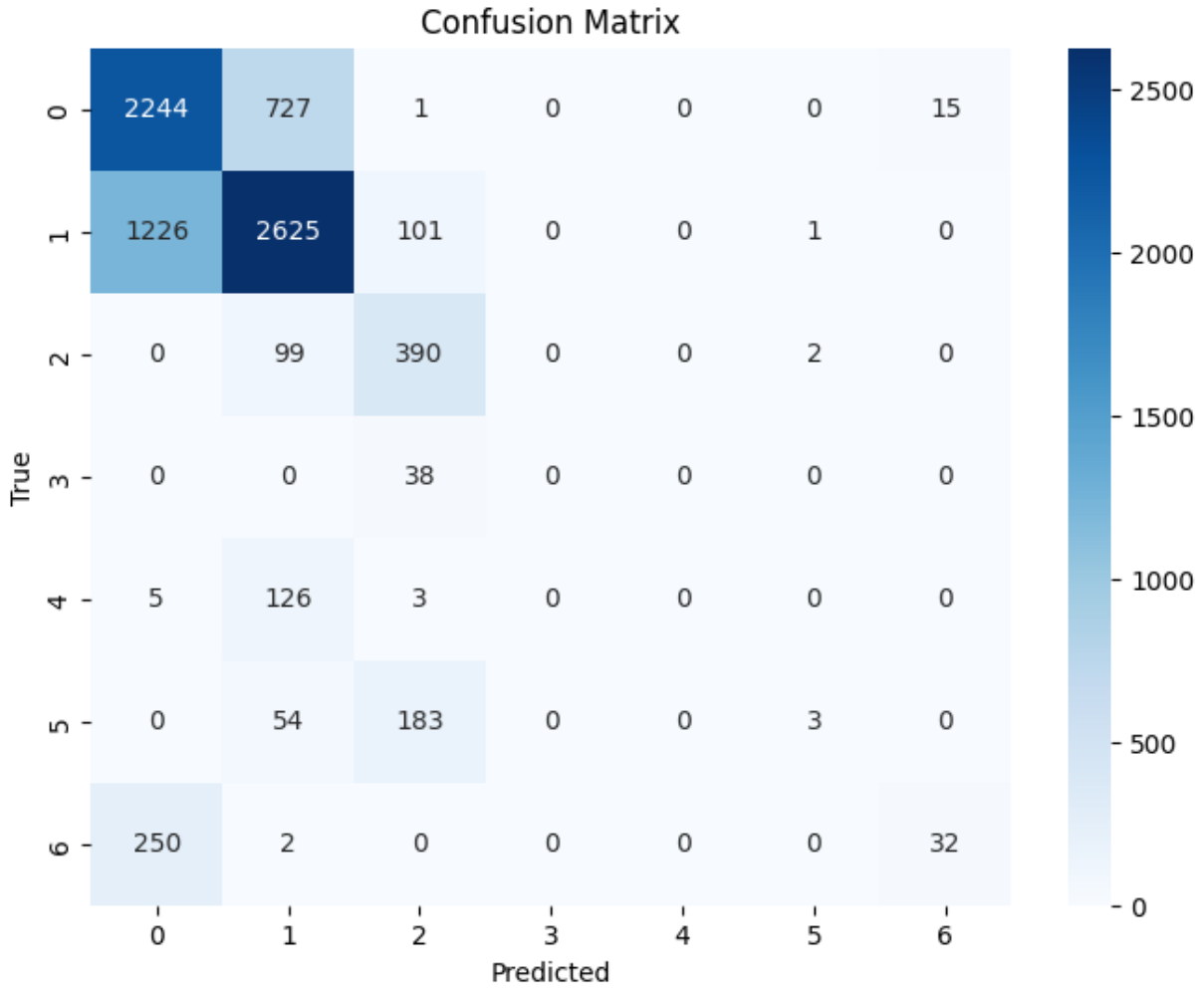
Methodology

Random Forest



Methodology

Random Forest



Conclusions