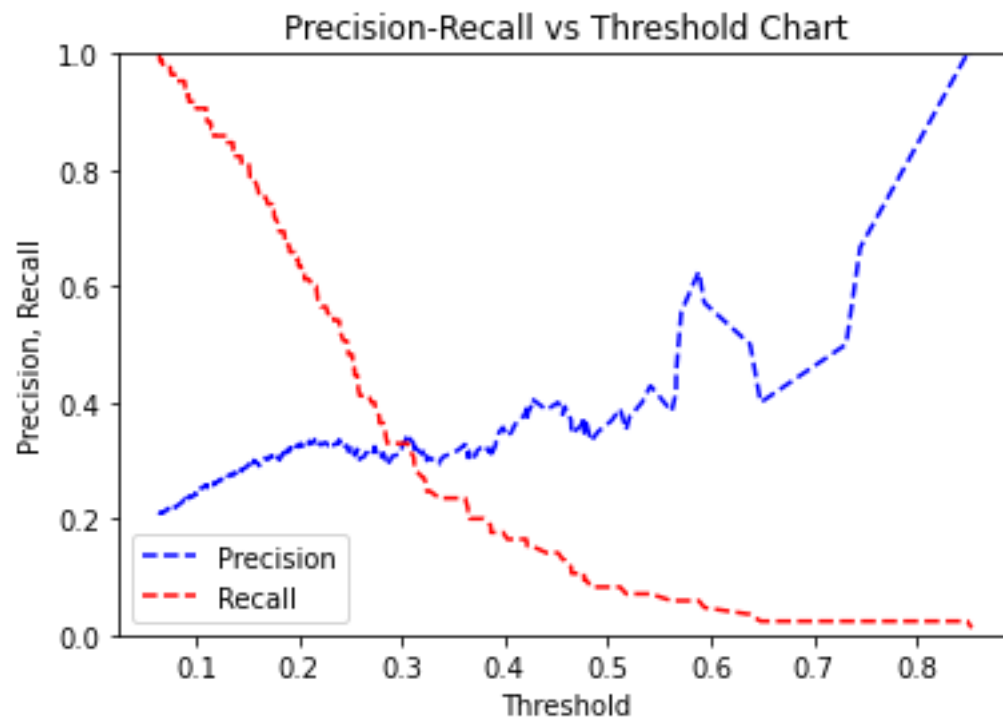


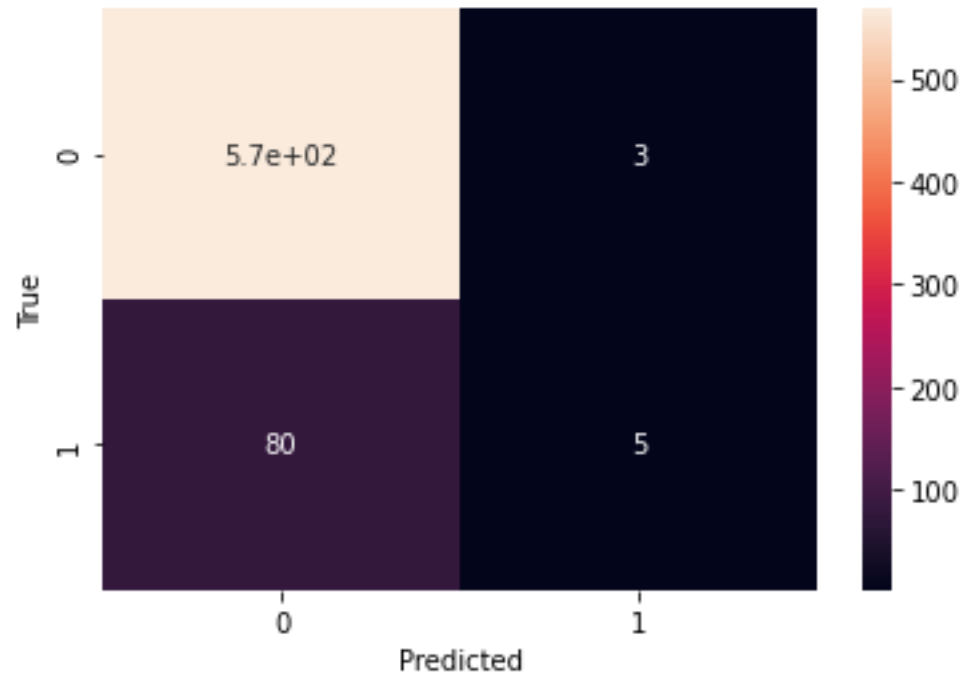
The goal of the project is to classify the water usage (potable or not) based on chemical and physical properties. To be more specific the goal is to classify if the water is potable without misclassifying not safe to potable this is why we used precision as our metric for the project.

We started building the a logistic regression model and studied the behavior of recall and precision with changing the threshold to decide the threshold for the project .



From the plot we noticed there as a spike at 0.59 so we set that as the threshold and started building the model.

After testing the model with the test data we got the following confusing matrix



TP(safe water predicted potable) = 5

FP(unsafe water predicted potable) = 3

Precision = 62.5%

Not the result we are aiming for that why we will try to build other models to get better performance.