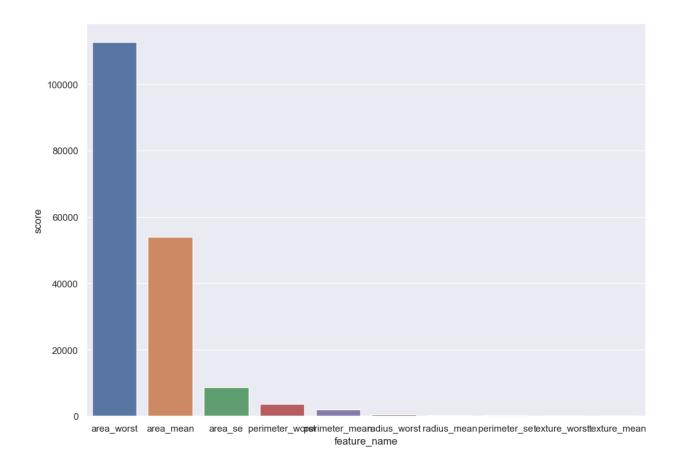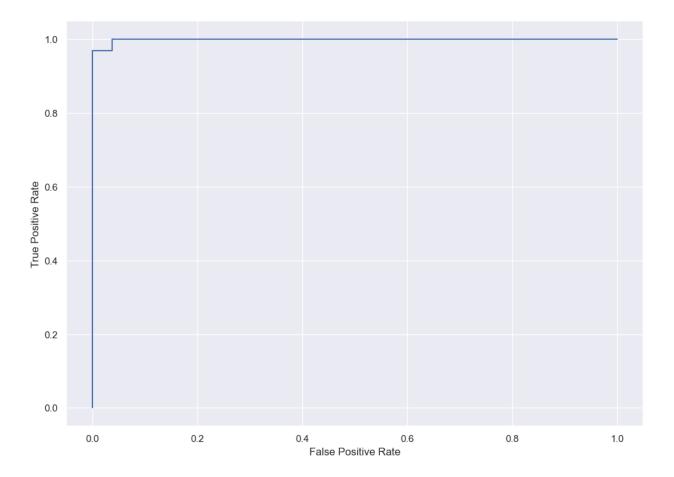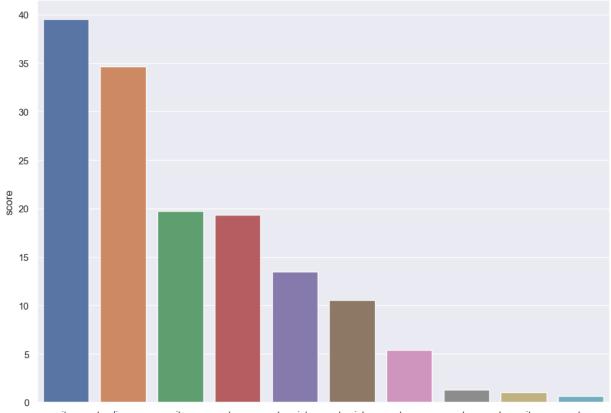Discussion:

Things I have tried:
For feature selection I tried to use "SelecteKBest" with "chi2" function, to get each feature's score. There are 30 different features in total, I separated them into 3 sets, each set has 10 features.
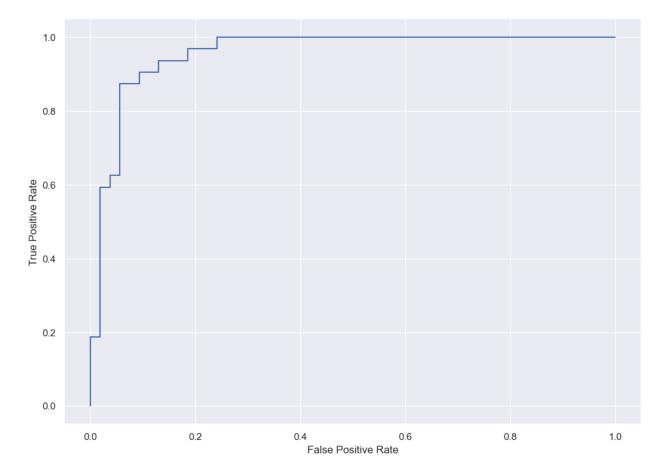
I first use the 10 features with the highest score to train my logistic regression model. The performance is way beyond my expectations. So, I kept using the second and third set as di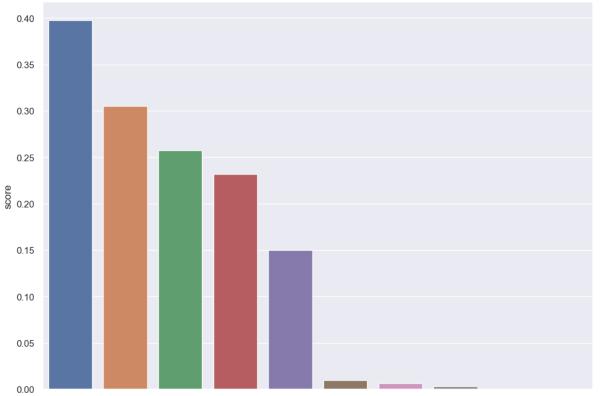fferent training dataset, thing get a bit more realistic.