

My choice of Dataset:

People use emails all day, not only for works, but also for communicating with friends. If a lot of spam email sent to your email account, you may have the chance to lose your focus on your works and may result in very bad situation. That is why a spam email filter should be developed and why I chose to do this project.

My choice of Model:

There are 6 models in total I have used in this project:

- SVM-Linear
- SVM-polynomial
- SVM-sigmoid
- Multinomial Naive-Bayes model
- Gaussian Naive-Bayes model
- Neural Network

After performance comparison, neural Network model I built have the best performance among all.