**SpamFilter Documentation**

To make SpamFilter using Python we can use Multiple ML and Deep Learning Framework. In this app I’ve Used **SVM (Support Vector Machine)** Algorithm which is 98% Accuracy for the given Data.

Data I’ve used from the given link <http://www.dt.fee.unicamp.br/~tiago/smsspamcollection/>. I downloaded it as a .txt format from a website.

**Step 1:**

**Data Preprocessing**

As it was txt file I used **Data Manipulation (Slicing)** to separate the Label and the Features which are Type and Message respectively. By using Pandas dataframe I approached to divide the Data into 2 columns [ Message , Type ]. Message contains the string and Type is the category of Message whether it is Spam or yam.

**Step 2:**

**Data Splitting and Model Selection**

Divide the Data for training and testing. As we already have our own data to use we need to train with ML or DL models to get output. In this app I’ve used **Scikit learn and Pandas** Libraries.

**train\_test\_split**  is a method which is responsible for dividing your data into training and testing sets with the testing ratio you define. I gave a 20% **Testing ratio** which means I will have 2 random sets of data 80% for training and 20% for Testing. By using SVM Model I tain the Dataset and fit the Test set into that Model

**Step 3:**

**Model Accuracy**

Model Score

After Training your Dataset with model you can test the remaining 20% of Testing data to Test using  **model.score(X\_test,y\_test))**

This model has 98% Accuracy using SVM algo

**Step 4 :**

**Final Output**

To check the input message Accuracy Score if it is more than 0.7 Print it as SPAM or else Print NOT SPAM