Analysis of different classification models for Email Spam Filtering

Manoj Battula
mbattul1@ufl.edu

***Abstract*— Emails have been a major form of communication over the last decade for business and personal uses. But people are bombarded with spam emails frequently these days. This project focuses on developing a spam email filter by training supervised machine learning models such as Naïve Bayes, K-Nearest Neighbors, Support Vector Machine, decision tree etc., on the dataset of known spam keywords and classifying the emails as spam/non-spam. Comparing the efficiency and accuracy of different spam filtering models and analyzing their advantages and drawbacks is the key aspect of the project.**

Index Terms— Email, Spam Filtering, K-Nearest Neighbors, Support Vector Machine.

I. PROJECT PLAN

The timeline of the project is divided in to the following phases.

A. Literature search

The first step of the project is to find the initial literature available on the email spam filtering.

Email spam filtering can be achieved using many different techniques such as hashing, Analyzing text information embedded into images, Classification based on machine learning models. There is a lot of research been done on the classification approach for detecting spam emails.

It would take around one week to analyze all the research papers in this field and create a base to start implementing the different models.

B. Implementation/ programming

After analyzing the different models and knowing the efficient libraries to be used to develop the models, programming will be started. Right now, I prefer to develop the project in python. ScikitLearn library is preferred for the machine learning models that are used in the implementation. Data set of spam keywords is collected from online sources and their frequency of occurrence as spam is also to be collected.

The dataset has to be modified based on the model that is being trained. For example, in Naïve Bayes, probability of a word being spam has to be calculated for the dataset of words.

This phase involves a lot of programming and data collection. So, it would take around two weeks to complete this phase of project.

C. Comparative study

Spam mails are detected using different classification model approaches in the previous phase. Now, a comparative analysis of the models is to be performed. Accuracy of the models is calculated by using the machine learning metrics such as precision and recall. Also, the efficiency of the models is studied by comparing the running times of their respective algorithms.

The comparison is performed by depicting the results in graphs. Right now, I prefer to use Python's data visualization library, matplotlib. This phase takes around 2 weeks.

D. Final Report and Presentation

In the final phase, a project report is generated describing all the phases in more detail specifying the implementation details. Also, graphs are depicted in the report showing comparative analysis of algorithms developed using different classification models. Also, a you tube video presentation is created which explains the spam filtering project, the implementation details, advantages and drawbacks of different methods.

REFERENCES

[1] A study on email spam filtering techniques
https://pdfs.semanticscholar.org/d63e/ff5dd2588a7a6289b08e73620238e03880bd.pdf

[2] An evaluation of Naïve Bayesian Anti-Spam Filtering Techniques
http://csiweb.ucd.ie/UserFiles/publications/UCD-CSI-2007-4.pdf

[3] K-nearest neighbors classifiers
http://digital.cs.usu.edu/~erbacher/publications/Bayes-Vikas2.pdf

[4] Spam filtering using K-NN
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.215.4173&rep=rep1&type=pdf

[5] SVM based spam filter with Active and Online Learning
http://trec.nist.gov/pubs/trec15/papers/hit.spam.final.final.pdf