MINI PROJECT SYNOPSIS

On

"Comment Spam Classifier"

BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING

Submitted by-

MADHUR AGRAWAL

5th SEMESTER
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INTRODUCTION

Due to the dwelling problem of increasing comments over the internet, it has become a major concern for all the people on YouTube. These spam comments may fraud, phish, and even worse hack into a user's system and make it unrecoverable.

To recover from this problem, I have developed a Comment Classifier System which is capable of classifying the comments as Spam and not spam.

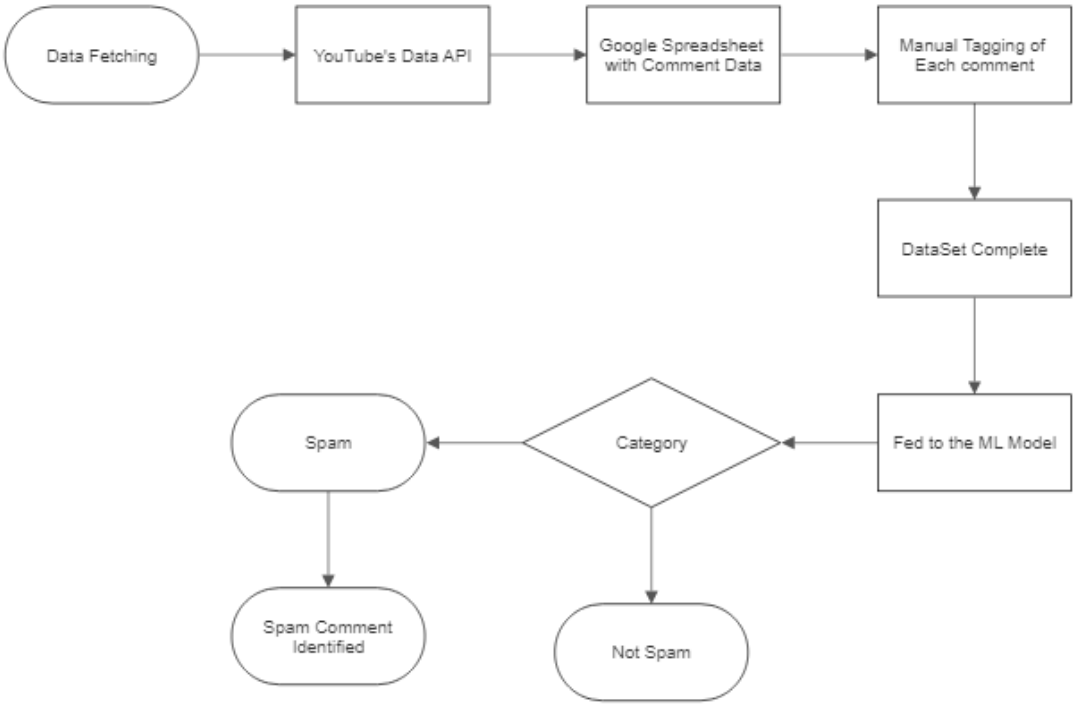
The project is used by the Logistic regression model of Machine Learning Techniques.

This project is also developed with the help of Data Analytics which has helped in the process of mining a data stream as well as purifying and construction of data set.

The various Libraries used are:

- Pandas
- Numpy
- Tensorflow
- Scikit Learn
- Matplotlib

OVERVIEW



The above flowchart shows the pathway to how the system works.

1. **Data Mining:** Data is mined by using a Google script to fetch the comments from the YouTube using the Data API.

2. **Data Set Preparation:** Data set is prepared by manually tagging each and every comment.

3. **Model Training:** The Machine Learning Model is trained with the help of ML algorithms.

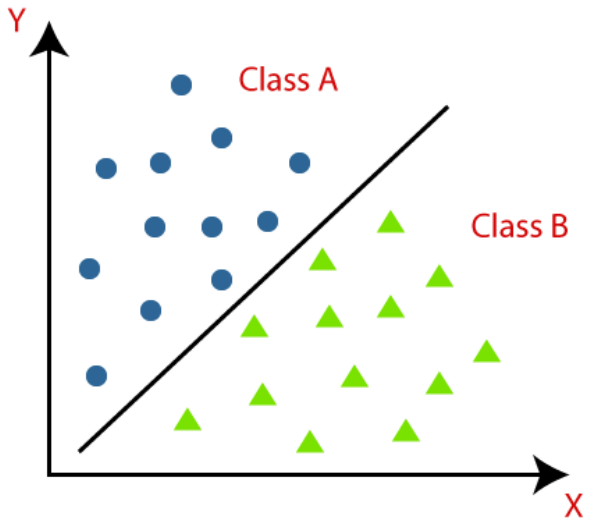
4. **Testing and Results:** The model testing and results are given by

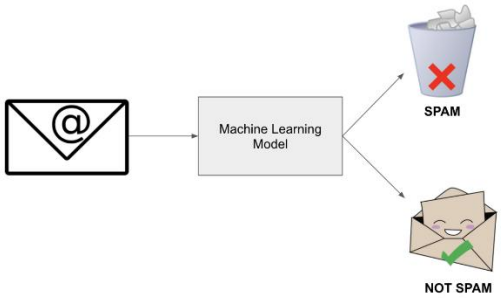
the model by giving new input to the machine learning model.

About Classification Supervised Machine Learning

- Classification is a type of Supervised Learning algorithm which is used when output variable is a category.

- Categories are represented by Binary Values i.e., **0 and 1**





Summary and Conclusion

- The model trained produces an accuracy ~91%
- As the dataset grows, it can learn further and be more precise
- Further script can also be applied using YouTube's API to apply the model to the YouTube accounts so as to make it work automatically.

Thank You!

References

- My Teachers who guided and helped me through the project.
- Google Images for the images (All images belong to their respective owners)
- Google Colab
- Libraries
- YouTube