METHODOLOGY

Detection of spam comments in youtube videos

Sparsh Ahuja	2017258
Siddhant lohia	2017254
Saurabh kumar	2017232

METHODOLOGY

We below describe our adopted methodology specific to our project **Detection of spam comments in youtube videos**

Where our purpose is to identify spam comments which have been defined as those which have a promotional intent or those who deem to be contextually irrelevant for a given video.

The prospects of monetisation through advertising on popular social media channels over the years has attracted an increasingly larger number of users. This has in turn led to the growth of malicious users who have begun to develop automated bots, capable of large-scale orchestrated deployment of spam messages across multiple channels simultaneously. The presence of these comments significantly hurts the reputation of a channel and also the experience of normal users.

In this work, we try to put up a method to detect such comments by applying conventional machine learning algorithms efficiently and procuring the dataset required for it using YOUTUBE APIs

OVERVIEW DIAGRAM

This diagram gives a brief overview of the steps involved in the process we have decided to use in this project
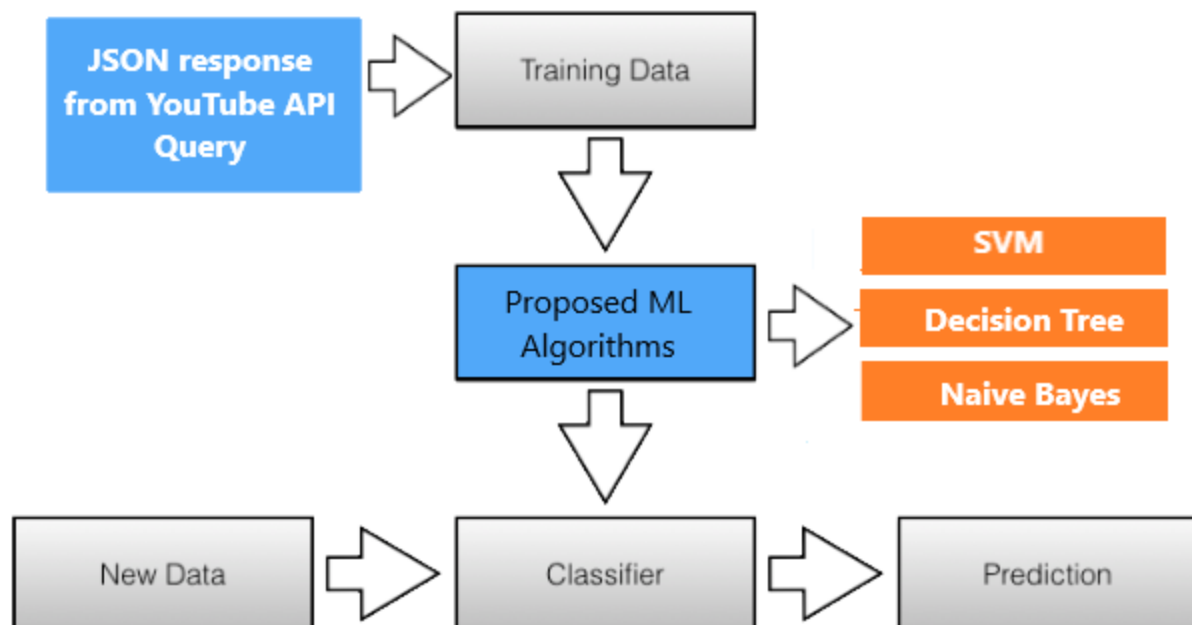
DETAILS

PROCURING DATASET

A comment dataset from music videos of some popular singers was used, Using the standard youtube APIs ,if successful it returns a JSON response body with the following structure which is used to populate the training as well as testing dataSet

```
Response

If successful, this method returns a response body with the following structure:

{
  "kind": "youtube#commentListResponse",
  "etag": etag,
  "nextPageToken": string,
  "pageInfo": {
    "totalResults": integer,
    "resultsPerPage": integer
  },
  "items": [
    comment Resource
  ]
}
```

PRE-PREPROCESSING

For the Pre-processing phase, the raw dataset will be executed and the data cleaning methods which are **tokenization, stopwords removal** and **stemming** are performed. The clean dataset will be used for the next process of feature selection and extraction

- **Tokenization**
  It is the process of dividing text into a set of meaningful pieces. These pieces are called **tokens**. For example, we can divide a chunk of text into words, or we can divide it into sentences. Depending on the task at hand, we can define our own conditions to divide the input text into meaningful tokens
- **Stopwords removal**
  Stop words are useful syntactically and grammatically, but don't tell anything about the document content and they are topic-neutral: stop words have the same likelihood of occurring in both relevant and non-relevant documents - so not very useful for Information retrieval.
- **Stemming**
  Stemming is the process of reducing a word to its word stem that affixes to suffixes and prefixes or to the roots of words known as a lemma.

FEATURE EXTRACTION AND CLASSIFICATION PHASE

Results in Feature Extraction and Feature Selection As features identified from the literature review, various features may be extracted from YouTube classification purposes. Besides, the data already consists of two (2) classes where the classes are **spam** and **ham** . Thus, easy to choose features that certainly label as spam. YouTube comments may contain hyperlinks, text, uppercase and lowercase characters. However, those uppercase characters do not exist after the preprocessing phase. After preprocessing, in our project we decide to use keywords as feature selection. The aim for feature extraction is to explore the advantages of new features in order to gain high accuracy

1. **Support Vector Machines**
   SVM is successfully suitable in differentiating positive and negative problems such as spam. SVM is a supervised learning model that analyzes data used for classification and regression.
   A Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyperplane. In other words, given labeled training data (*supervised learning*), the algorithm outputs an optimal hyperplane which categorizes new examples.

2. **Naive Bayes**
   NB is a well-known classifier relying on Bayes theorem. **Naive Bayes** classifier assumes that the features do not correlate with each other. Therefore, a probability score is calculated with multiplication of some conditional probabilities.

3. **Decision tree**
   Decision Tree classifier aims to reach a classification decision with the help of a decision tree structure it constructed. Nodes in the decision tree structure generally represent feature values and leaves in the decision tree structure represent specific class labels.

CONCLUSION

Finally the we will analyse the results of our proposed solution to the problem of **Detection of spam comments in youtube videos**
And we will compare all the three algorithms listed and their efficiencies and publish our result in the final project report which we will submit.