

ANALYSIS ON PUBLIC OPINION FOR YOUTUBE COMMENTS BY OPINION MINING USING CLASSIFICATION MODELS

¹Dr. D. Srinivas Reddy, ²K. Sridhar Reddy, ³Ch. Rajender

¹Associate Professor, ²Assistant Professor, ³Student ^{1,2,3}Dept. of Computer Science Engineering,
^{1,2,3}Vaageswari College of Engineering, Karimnagar, Telangana

E-Mail: srinivasreddydhava@gmail.com, sridhark529reddy@gmail.com

ABSTRACT

Opinion mining is the computational analysis of text-based expressions of people's opinions, sentiments, attitudes, and emotions. Because views are fundamental to practically all human activities and are significant determinants of our behavior, it has a wide range of applications. We seek out other people's perspectives whenever we need to make a decision. Our project can quickly process any quantity of YouTube comments and provide real-time results. The goal of this research is to develop an efficient classification model for YouTube comments that is both accurate and automatic. The current study uses Random Forest and Naive Bayes for its analysis. Here, we go through the opinion mining procedure, and the results are classified as either favorable, negative, or neutral.

Keywords: opinion mining, positive, negative, neutral, YouTube API

I. INTRODUCTION

Some of the most well-known social media websites in the world, like YouTube, give all users a voice and the ability to express their thoughts and feelings. Opinion mining can be used to extract and quantify these users' opinions and feelings. This study looks at opinion mining on YouTube comments and how effectively it can forecast how a YouTube video will perform by counting how many comments are positive, neutral, or negative. We examined four alternative prediction models that employ neutral remarks in various ways. Pre-training on YouTube comments, tweets, and a combination of tweets and comments was done on five different classifiers. The expected and actual liking proportions showed some positive association. A logistic regression classifier was the configuration with the best performance.

Opinion Mining allows you to get inside your customers' heads and find out what they like and dislike, and why, so you can create products and services that meet their needs. When you have the right tools, you can perform opinion mining automatically, on almost any form

of unstructured text, with very little human input needed. It can process thousands of pages, comments, emails, or surveys in just minutes for real-time results. The most common use of opinion mining works to categorize comments and statements on a scale of opinion polarity. This can be simply positive, negative, or neutral.

II. PROPOSED SYSTEM

In our current model we mine the YouTube comments which is useful for the content makers and YouTubers to make a review about their video. and it will be time saving process for them instead of reading all the thousands of text mentioned in the comments . We fetched the real time data base from google API key.

Using these algorithms Random Forest and Naive Bayes undergo the process of data cleaning , data preprocessing, analyzing and prediction and the results are viewed in a plot bar classifying positive ,negative and neutral comments

Our methodology in YouTube comment analysis is classifying polarity of a given data is positive , negative or neutral.

Data Collection: Existing YouTube comments data set via YouTube API

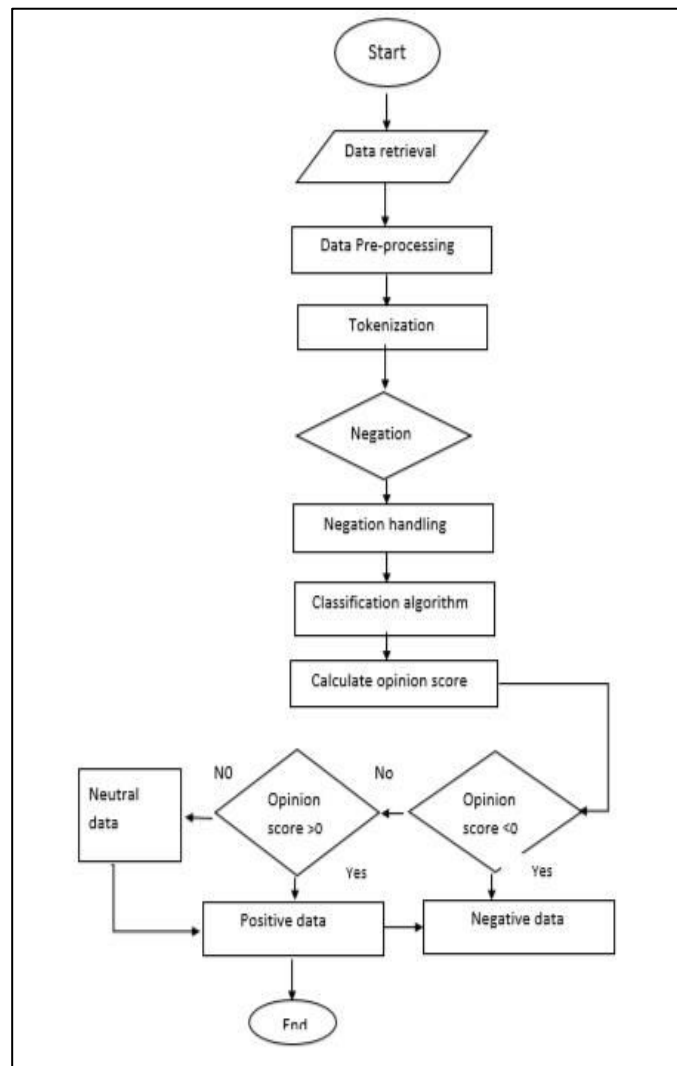
Data Pre processing : Get the “clean” data and transform it to the format we need

Analysis: Train a classifier to classify the comments as: positive, negative and neutral

Prediction: The results are viewed in a plot bar

III.IMPLEMENTATION

In this project it is going to provide information about the opinions of the people posted in YouTube as their comments in the form of positive ,negative and neutral format.we give the url of the video it will fetch the comments clean, analyze and then opinions are fetched in the form of a plot bar



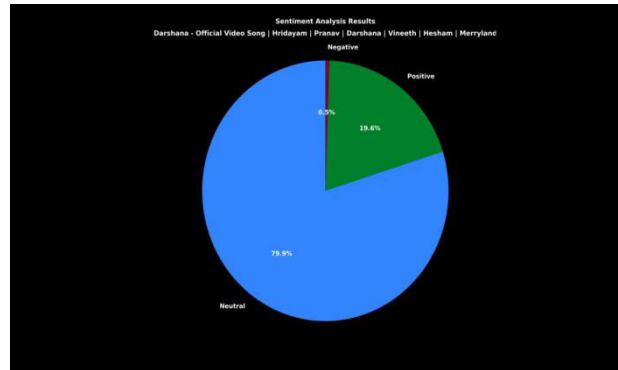
Flow chart

IV.RESULTS

✧ GETTING THE COMMENTS WITHOUT REPLIES

\$ “python yt_comments_analyzer.py -u https://www.youtube.com/watch?v=W3B73VDCDfM ”

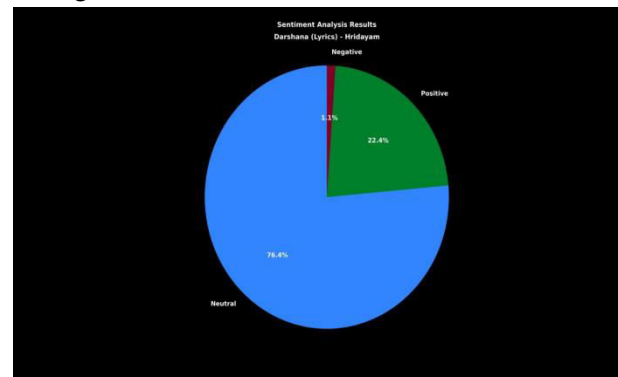
Get comments for the given video .



✧ GETTING THE COMMENTS WITH REPLIES

\$ “python yt_comments_analyzer.py -u https://www.youtube.com/watch?v=W3B73VDCDfM -ir -o result_chart.png”

Get comments with replies for the given video



The comments will be fetched using YouTube API key. Fetched data will be preprocessed and classified . The results can be displayed in plot bar for comments without replies and with replies.

CONCLUSION

Using a straightforward prediction based on the number of comments classified as positive, neutral, and negative, our study looked into the prediction of like proportions of YouTube videos. The like ratios of YouTube videos and the expected like ratios based on the sentiment of their comments categorized by the classifiers employed in this investigation showed some correlation.

FUTURE SCOPE

This project involves almost all the feature for the opinion mining for YouTube comments. In future we can also add a new feature which can process emoji's from the data. Certain emojis couldn't be properly encoded in our chosen file format (UTF- 8) so those emoji characters had to be deleted.

REFERENCES

1. Apoorv Agarwal, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca Passonneau. 2011. Sentiment Analysis of Twitter Data. In Proceedings of the Workshop on Languages in Social Media (Portland, Vol.49, No.2(VII) July–December 2019

1. Oregon) (LSM '11). Association for Computational Linguistics, USA, 30–38.
2. Rishabh Ahuja, Arun Solanki, and Anand Nayyar. 2019. Movie Recommender System Using K-Means Clustering AND K-Nearest Neighbor. In 2019 9th International Conference on Cloud Computing, Data Science Engineering (Confluence). 263–268.
3. Khin Zezawar Aung and Nyein Nyein Myo. 2017. Sentiment analysis of students' comment using lexicon based approach. In 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS).
4. Hanif Bhuiyan, Rajon Bardhan, and Dr. MD Rashedul Islam. 2017. Retrieving YouTube video by sentiment analysis on user comment. 474–478. Zhalaing Cheung, Khanh Linh Phan, Ashesh
5. Mahidadia, and Achim Hoffmann. 2005. Feature Extraction for Learning to Classify Questions. In AI 2004: Advances in Artificial Intelligence, Geoffrey I. Webb and Xinghuo Yu (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1069–1075.
6. Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. 2011. Natural Language Processing (Almost) from Scratch. J. Mach. Learn. Res. 12, null (Nov. 2011), 2493–2537.
7. Mohit Iyyer, Peter Enns, Jordan Boyd-Graber, and Philip Resnik. 2014. Political Ideology Detection Using Recursive Neural Networks. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, Baltimore, Maryland, 1113– 1122.
8. Ammar Ismael Kadhim. 2019. Survey on supervised machine learning techniques for automatic text classification. Artificial Intelligence Review 52, 1 (2019), 273– 292.
9. Anthony Khoo, Yuval Marom, and David Albrecht. 2006. Experiments with Sentence Classification. In Proceedings of the Australasian Language Technology Workshop