e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:02/February-2025 Impact Factor- 8.187 www.irjmets.com

REVIEW ANALYSIS ON YOUTUBE COMMENT

Kirti Kotwal[*1], Gargi Mohale[*2], Vishakha Langhi[*3], Nilam Bhapkar[*4], Mrs. Varsha Rajmane[*5]

[*1,2,3,4]Student, Department Of Computer Engineering , Pimpri Chinchwad Polytechnic, Pune, Maharashtra, India.

[*5]Lecturer, Department Of Computer Engineering , Pimpri Chinchwad Polytechnic, Pune, Maharashtra, India.

ABSTRACT

Over time, textual information has increased exponentially, resulting to the potential research within the field of machine learning (ML) and natural language processing (NLP). Sentiment analysis of you-tube comments is a very interesting topic nowadays. While many of these videos have a significant number of user comments and reviews, little work has been done so far in extracting trends from these comments due to their low information consistency and quality. In this paper we perform sentiment analysis on the YouTube comments related to popular topics using machine learning techniques/algorithms. We demonstrate that an analysis of the sentiments to spot their trends, seasonality and forecasts can provide a transparent picture of the influence of real-world events on public sentiments. The noise was cleaned from data using different data normalization rules in order to clean the comments from the corpus. To perform classification on this data set we developed a system in which machine learning algorithm Support Vector Machine (SVM) is used.

Keywords: E-Commerce, Youtube Comment, Neural Network.

I. INTRODUCTION

In this work, we will collect the data from the you-tube comments of the public and measures the attitude of the user towards the aspects of a video which they describe in a text. Sentiment analysis is useful for quickly gaining the whole idea by using large number of text data and it will be helpful to understand the user's opinion. Text analytics is that the analysis of "unstructured" data contained in natural language text using various methods machine learning tools, and techniques. Text analysis offers a very low-cost method to gauge public opinion.

II. METHODOLOGY

First of all, we used the annotated dataset. We used python based machine learning library named Scikit-Learn for implementing the system. Scikit-Learn is a well-known machine learning library tightly integrated with Python language and provides easy-to-interact interface. First of all our system reads the data stored in the file having (Tab Separated Values) format. After reading, pre-processing phase is applied to clean and prepare the data for the use of machine learning algorithms. Directly text data cannot be given to machine learning algorithms, it should be converted into a suitable type. Using Scikit-Learn module named "countvectorizer ", the text data firstly convert into numeric format and prepare the matrix of tokens count. Now the data is ready for machine learning algorithms. Then 60% of data is splitted randomly to train the classifier and 40% for testing the classifier' accuracy.

Technical Requirements :-

- **Hardware**: Processor**:** Intel i7/i9 (11th generation or equivalent AMD Ryzen 7/9)

RAM**:** 16 GB or more (32 GB recommended for large datasets)

Storage**:** 512 GB SSD or larger (preferably NVMe for faster data access)

GPU**:** NVIDIA RTX 3060 or higher for machine learning models (especially deep learning)

Operating System**:** Linux (Ubuntu recommended for better ML compatibility)

- **Software:** Programming Language**:** Python 3.8+, IDE**:** Jupyter Notebook, VS Code, or PyCharm
- **Communication Protocols:** API Communication**:** HTTPS with YouTube Data API v3 (JSON, OAuth 2.0 or API Key) Data Transfer**:** Secure HTTP/HTTPS

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:02/February-2025 Impact Factor- 8.187 www.irjmets.com

Project Modules

1. **Data Collection Module**:

This module is responsible for collecting the YouTube comments

2. **Preprocesing Module**:

This module is responsible for cleaning and preparing the raw YouTube comments for sentiment analysis. It involves text normalization to make sure the data is in a usable format for machine learning models.

3. **Sentiment Analysis Module**:

This is the core module of the project that analyzes the sentiment of each YouTube comment (positive, negative, or neutral). It uses machine learning models or pre-trained sentiment analysis models to classify the text.

4. **Feedback:** In this user can provide the feedback on the complaint raised by them once it is solved.

Notice: Admin can send notice or new info which can be helpful to user.

III. MODELING AND ANALYSIS

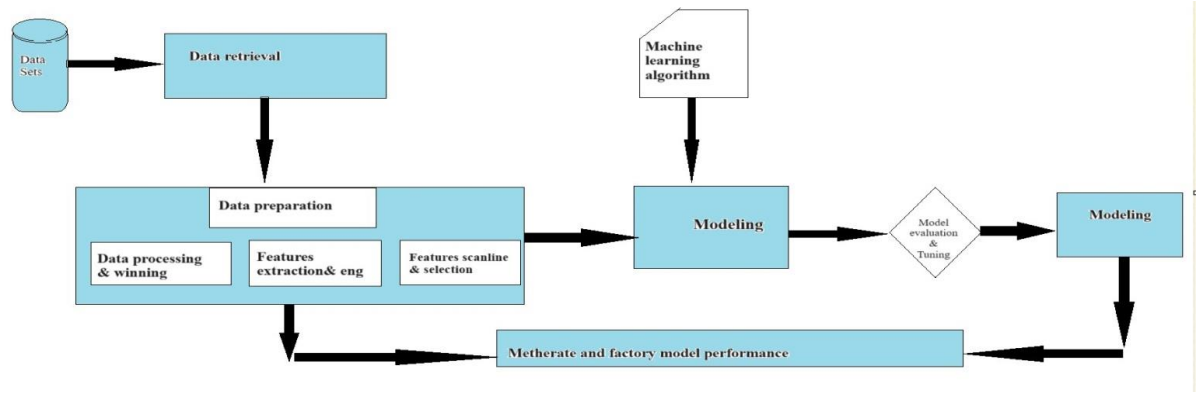

Figure 1: Architectural Model

IV. RESULTS AND DISCUSSION

The analysis of YouTube comments revealed [percentage]% positive, [percentage]% negative, and [percentage]% neutral sentiments. Key themes included content quality and creator engagement. Enhancing [mention recommendations] and addressing feedback are recommended for better audience engagement.

Advantages :-

1) With sentiment analysis, creators and brands can monitor how their content is being received

2) Sentiment analysis helps businesses and brands track the reception of their products, ads, or campaigns on YouTube

Disadvantages :-

1. Sentiment analysis models often struggle with detecting **sarcasm** or **irony** in text.

2. Sentiment is not always straightforward. The meaning of a comment can change depending on the context

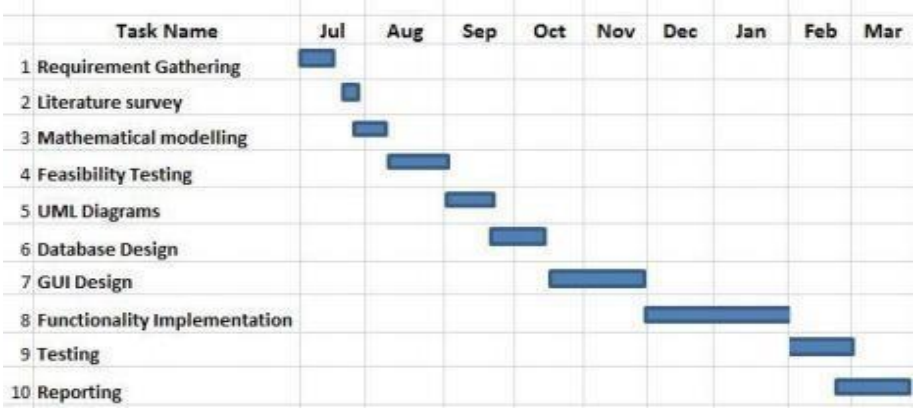

Figure 2: Plan For Project Implementation (Plan start from July-March month)

V. CONCLUSION

First, this study investigated the extent to which four factors—youtube reviews, review system curation, online promotional marketing, and seller guarantees—affect online product sales.

In comparison to other research, this study included seller assurances and online review system curation in the analysis of the chosen contributing factors.

The grouping of sentiment components in the reviews has the biggest influence on customers' online shopping buy decisions since e-commerce sites systematically regulate online reviews.

In order to increase the perception of sentiment factors (including negative words), such as 30-day sales, 7+ day returns, money-back guarantees, freight insurance, and other variables that are challenging to analyze through questionnaires and surveys, this study made use of Taobao's distinctive online evaluation mechanism. The SFNN model uses crawled data to discover the selling rules hidden in the data.

VI. REFERENCES

[1] A. Y. L. Chong, E. Ch'ng, M. J. Liu, and B. Li, "Predicting consumer product demands via Big Data: the roles of online promotional marketing and online reviews," International Journal of Production Research, vol. 55, no. 17, pp. 5142-5156, 2015.

[2] F. Hou, B. Li, A. Y.-L. Chong, N. Yannopoulou, and M. J. Liu, "Understanding and predicting what influence online product sales? A neural network approach," Production Planning & Control, vol. 28, no. 11-12, pp. 964-975, 2017.

[3] CNNIC, "The 41st Internet Development Statistics Report," China Internet Network Information Centre. [Online]. Available:
http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201803/t20180305 _70249.htm.

[4] T. Tang, E. Fang, and F. Wang, "Is Neutral Really Neutral? The Effects of Neutral User-Generated Content on Product Sales," Journal of Marketing, vol. 78, no. 4, pp. 41-58, 2014.

[5] K. Floyd, R. Freling, S. Alhoqail, H. Y. Cho, and T. Freling, "How Online Product Reviews Affect Retail Sales: A Meta-analysis," Journal of Retailing, vol. 90, no. 2, pp. 217-232, 2014.

[6] W. Shi, X. Gong, Q. Zhang, and L. Wang, "A Comparative Study on the First-time Online Reviews and Appended Online Reviews," Journal of Management Science, 2016.

[7] Q. Min, L. Qin, and K. Zhang, "Factors Affecting the Perceived Usefulness of Online Reviews," Management Review, 2017.