

A Synopsis

on

**CLASSIFICATION OF YOU-TUBE VIDEOS USING DEEP LEARNING**

*in partial fulfillment of the requirement for the degree*

of

Bachelor of Technology

In

COMPUTER SCIENCE AND ENGINEERING

Submitted by

**Sivasish Koch( 1513310214 )**

**Saif Ali( 1513321165 )**

**Rahul Bilra( 1513310161)**

**Akash Kandpal( 1513310027)**

under the supervision of

**Dr. Deepti Gupta**

( Associate Professor Computer Science Dept.)



**NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY**

# GREATER NOIDA

## Index

<b>Sr.No.</b>	<b>Topics</b>	<b>Page No.</b>
1	Abstract	1
2	Existing System(Literature Survey)	2
3	Problem Statement	3
4	Objective of the project and Features	4
5	Resource Requirements	5
6	Conclusion	6
7	References	7

Supervisor Sign:

(Sign of supervisor on synopsis is mandatory while submission)

## **ABSTRACT**

You-Tube contains lots of videos which are both good and bad in terms of the quality of the content and the satisfaction of the user. We search on YouTube for a video related to our projects and/or studies but end up getting distracted by advertisements, promotional videos, music videos or any other type of video and then after an hour or two we realize that we wasted our time and regret afterward. This **reduces** our **productivity** and also makes us **mentally tired** to watch more videos regarding our subject.

We are unable to get quality content from the pool of videos. So, we came across this idea of providing quality scores to videos (currently for education domain) and further provide tags to it for categorizing the videos in different domains.

## **MODIFICATION AND IMPROVEMENT OVER THE EXISTING IMPLEMENTATION**

### **Present State:**

Software present currently are limited in scope and they don't deal directly with improving the content of the You-Tube.

Currently the systems mainly focusses on building recommendation systems for You-Tube and like-wise video sites.

Till now You-Tube itself is not considering about categorizing the videos and checking the quality so an effort is required in this area.

### **After implementation of project:**

Quality score for all videos can be seen in the You-Tube itself.

No need for an additional app and the user-experience with the You-Tube videos will be maintained.

User could save his/her time as he knows which videos are of good quality.

Building Recommendation videos will be easier after categorization of videos.

## **PROBLEM STATEMENT**

You-Tube contains lots of videos which are both good and bad in terms of the quality of the content and the satisfaction of the user. We search on YouTube for a video related to something but end up getting distracted by advertisements, promotional videos, music videos or any other type of video and then after an hour or two we realize that we wasted our time and regret afterward. This **reduces** our **productivity** and also makes us **mentally tired** to watch more videos regarding our subject.

We are unable to get quality content from the pool of videos. So, we came across this idea of providing quality scores to videos and also provide tags to it for categorizing the videos in different domains. It can act as a self-checking mechanism and prevent us from getting distracted by unwanted and promotional videos. This will **increase productivity** and keep us on track and prevent mental tiredness. It can be used in **office environment** and in **institutions like colleges** .

## **OBJECTIVE**

The following are the objectives of the A.I. :

1. Extracting comments on videos from You-Tube and using them to generate score for the video. The higher the score the better the video.
2. ML model will be able to read and understand the comments and it generates a list of words used in the positive comments and negative comments with respect to the title of the video for generating quality scores.
3. Simple approach of number of cumulative matches with the title is considered for the score. Also, we will produce list of words which belong to a particular domain.
4. Saving the time of users by providing them with the quality score for each video.
5. Till now, we are more dependent on the number of views for the quality of the video but we have considered a few factors like comments(textual analysis), views and video itself(in future) for providing score to a video.
6. Improving the quality of You-Tube as a video platform.

This project is mainly used by two types of users :

- i. Daily You-Tube users
- ii. You-Tube developers for recommendation systems on the user data.

## **TOOLS AND TECHNOLOGY USED**

### **Tool and Libraries:**

Gensim : <https://radimrehurek.com/gensim/>

Fasttext : <https://fasttext.cc/>

NLTK : <https://www.nltk.org/>

### **Technology:**

Chrome Extension

Vanilla JavaScript

Tensorflow : As a base for Keras and more optimization.

Keras : For making Deep Learning Models

Pandas : For cleaning the data

Numpy : For mathematical purposes.

Plotly : For visualising graphs.

Scikit : For Shallow learning Algorithms.

Open CV : For computer vision part.

Convnet : For Handwritten Notes detection.

Pix2pix : Extension for video analysis.

Big Huge Labs API

XML parsing

## **CONCLUSION**

Our effort has been to make it easier for improving the quality content available on You-Tube and thus enhancing the user-experience of the You-Tube users. They can easily get relevant information from the videos. This has never been approached before as previously people have tried to make recommendation system for these platforms rather than doing something about the quality of content available over there.

Also, the approach involves further advancements like usage of Deep NLP and Computer Vision for understanding the sentiments of the comments and also by analyzing the video itself frame-by-frame. So, this is a small effort from our side to improve the quality of videos and rank videos on You-Tube by providing them scores, lot of advancements will be seen in this area in the near future.



## **REFERENCES**

Research Papers involved are :

- [1]. Green BF, Wolf AK, Chomsky C, and Laughery K. Baseball: An automatic question answerer.
- [2]. Weizenbaum J. ELIZA - a computer program for the study of natural language communication between man and machine.
- [3]. Woods W. Progress in Natural Language Understanding - An Application to Lunar Geology.
- [4]. Bobrow DG, Kaplan RM, Kay M, Norman DA, Thompson H, and Winograd T. Gus, a frame-driven dialog system.
- [5]. Katz B. Annotating the World Wide Web using natural language.
- [6]. Clark P, Thompson J, and Porter B. A knowledge-based approach to question answering.
- [7]. Riloff E and Thelen M. A Rule-based Question Answering System for Reading Comprehension Tests.
- [8]. Reading Comprehension Tests as Evaluation for Computer-Based Language Understanding Systems, Vol. 6, 2000, pp. 13-19.
- [9]. Ittycheriah A, Franz M, Zhu WJ, Ratnaparkhi A and Mammone RJ. IBM's statistical question answering system.

