Classification of You-Tube Videos using Deep Learning and NLP

Department of Computer Science and Engineering
(Supervisor: **Dr Deepti Gupta**)

Name of Member	Roll Number	Role
Akash Kandpal	1513310027	NLP/ Deep Learning Part
Sivasish Koch	1513310214	Javascript Part for the browser
Saif Ali	1513321165	Classifcation checker (manual)
Rahul Bilra	1513310161	Scraping the data

Index

- Project Acceptance Status and Suggestions given in Previous Presentation
- Refined Problem Statement with literature Survey
- Modules Summary
- Brief Module details
- Module Progress(in Tabular/Graphical form in terms of % completion)
- Hardware /Software Requirement
- Conclusion
- References

Acceptance Status and Suggestions

- The project has been accepted with some modifications. Suggestions were :
- Reference are not proper.
- PPT not in format.
- Improve project name.

Problems in Existing System

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.

Problems in Existing System

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 can see the quality scores of the videos and like-wise decide as to which videos to watch and won't waste time.
- Building Recommendation videos will be easier after categorization of videos.

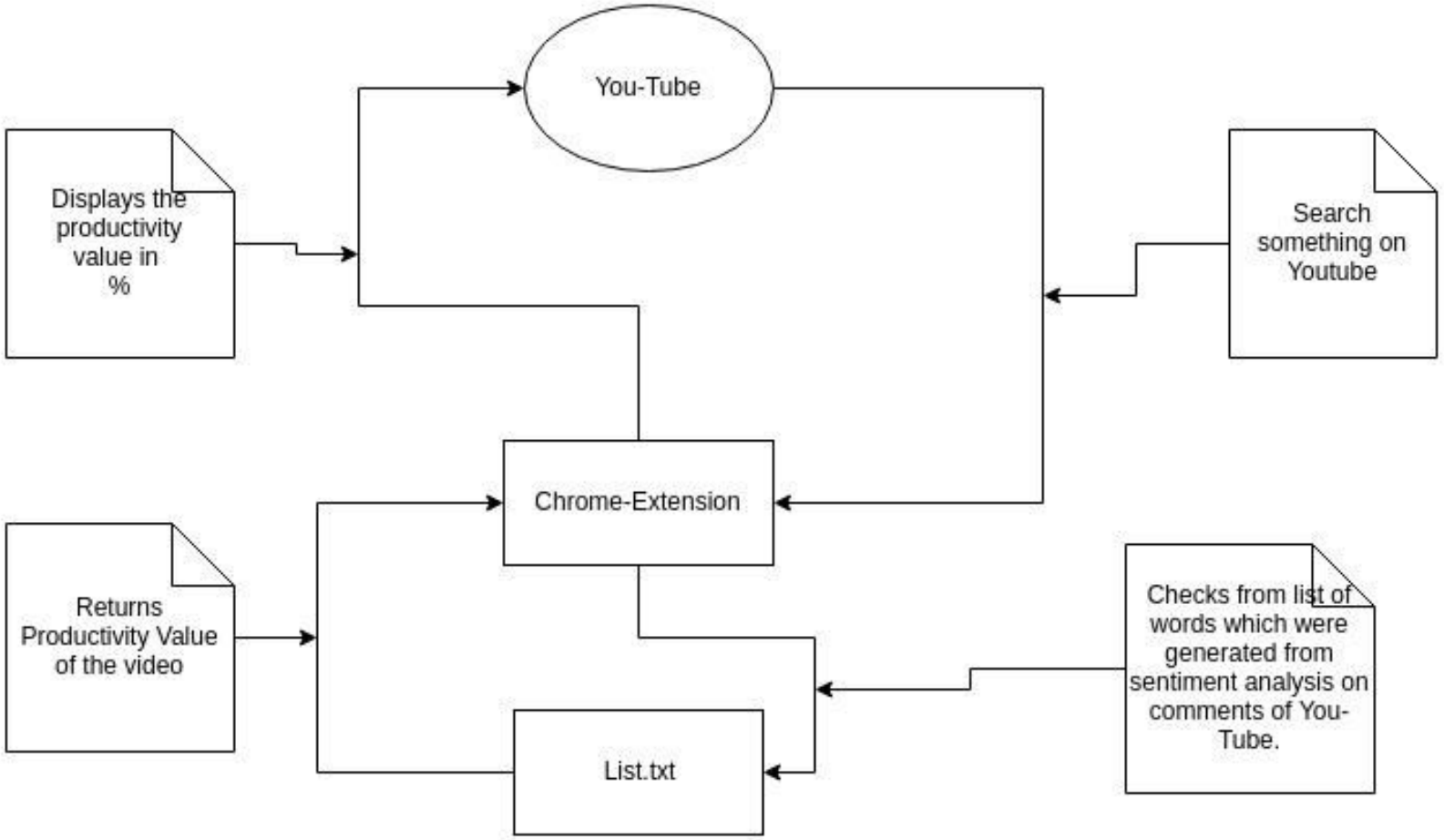
Module Summary

Module name	Progress (in %)
Clean the comments (removing stop-words) and manual cleaning (Backend)	100
ML Model for classifying these comments, using positive and negative txt files(Backend)	100
Scraper for comments from Educative Videos(Backend)	100
Script for Fetch the titles and description and match with the list of words txt file. (Backend)	100
Generate the score for education videos using separate script. (Backend)	100
AJAX script for rendering server calls(Front-End)	100
Send the score back to the browser.(Front-End)	100
Get titles and description of videos with more educative content and save their data in list of words (Backend)	100

Module Details

- Cleaning and analysis of comments is still in progress. We are collecting the most found words from the video and using them for generating the score about the educative content present in the video.
- JS scripts are written for showing the generated score back on the screen.
- Convnets are yet to be implemented after this the accuracy of videos will increase comfortably.
- Research papers like audio-visual convergence will be implemented for generating quality scores for videos.
- Deep NLP will be used for extracting better content from the comments.

Module Details



DEMO

Installation Video

Demo

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 usuage 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

- [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.

Thank You

Feedback from the Panel

- The panel wants the team to :