# InVideo Recommendation

Aly Mohamed, Amr Sherif, Foad Osama, Youssef Roshdy

March 16, 2020

## 1 Introduction

### 1.1 Purpose

This document is mainly for the full description of the requirements for the project InVideo Recommendation. This document will explain how the cycle of the system will go on, with the assist of the overview, constraints, functional and non-functional requirements which help this document to illustrate what should the user know and how the user will use the system.

### 1.2 Scope

The project is a plugin tool, which help the users to search by a video they upload and the system will recommend videos which they desired.This system will help people to block or cut some scenes from the video they aren't interesting in it. The user will provide the system with a video as an input; the system will start processing on it ,then recommend the most similar videos related to it as the system depend on feature extraction on the video so the resulted will be more accurate to the input one rather than the other platforms which used some calculations and algorithms. The system will give the user the chance to cut some scenes from the video he upload if there are some scenes not desired to him. This software needs internet access.

### 1.3 Overview

The proposed system implements a new function for searching by a scene just like a normal search engine. It aims to find similar content from video and output as a search result. Also,a great challenge is introducing a way to block certain scenes based on the user's custom-built filter, to achieve a clean watching experience. Video platforms uses a lot of calculations and algorithms such as: watch history and search results to recommend videos to their users which in many cases doesn't returns an accurate results especially for the new users or called the cold-start users who doesn't have for example a watch history or even an email on that platform. So, we aim to solve this problems according to detection and classifications done on each video uploaded by the user and this will be done by using feature extraction on a given video so it will give more accurate results rather than the algorithms using by other platforms. This will enhance the video recommendation system and increase the accuracy of the resulted videos.

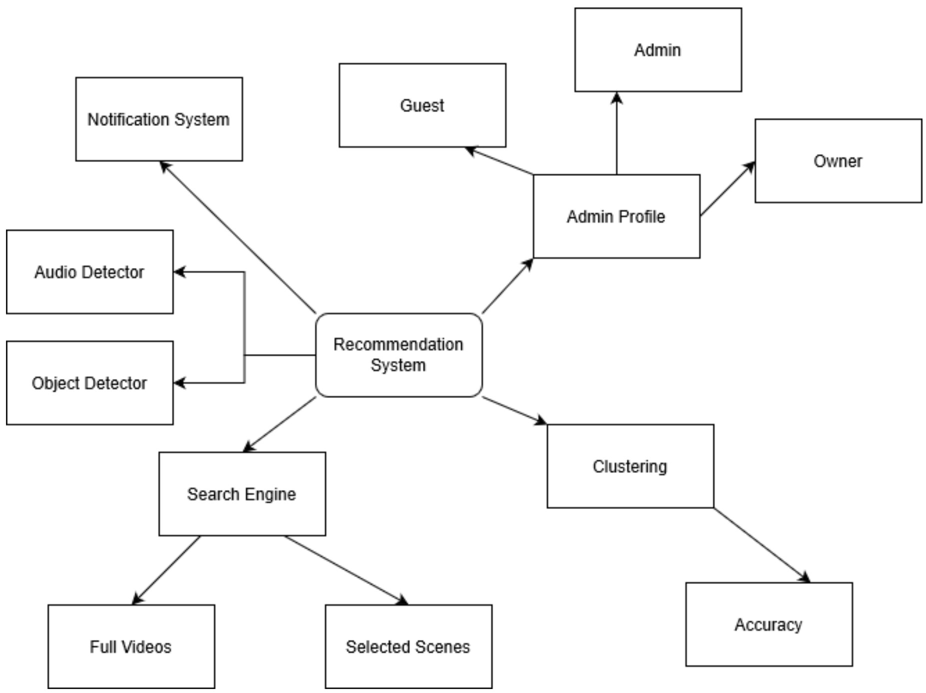


Figure 1: Context Diagram

## 1.4 Definitions and Acronyms

| Term | Definition |
|---|---|
| YoloV3/ Darknet | Used for Training and Detecting objects in videos design information. |
| Cosine Similarity | Used to compare the content of the videos. |
| Spectral Clustering | Used to split videos into separate groups with their similar videos. |
| Convolutional Neural Network | Used for Training and Detecting type of Audio Files. |
| Recurrent Neural Network | Used for Training and Detecting type of Audio Files. |

## 2 System Overview

The proposed system implements a new function for searching by a scene just like a normal search engine. It aims to find similar content from video and output as a search result. Also,a great challenge is introducing a way to block certain scenes based on the user's custom-built filter, to achieve a clean watching experience.The proposed system overview is shown in figure 1. It consists of three main phases. In the first phase, the user can start watching videos normally. The input scene will be selected from the highlights of the video or the user can select a specific interval form the current video to use as a search query. A video will be imported to be processed in the second phase either from the search or from the highlights. During the second phase, object detection can take place. After processing has occurred, a frequency table for objects has been generated.This table is used to compare the content of the video to the database videos which also has the table of data given. By similarity measurements throughout the third phase, results should appear in the form of recommended videos based on highlights from the input video, or in a form of search result, from the user's input. It is also possible to have some scenes filtered and removed from the video based on a filtering created by the user to remove a certain content.

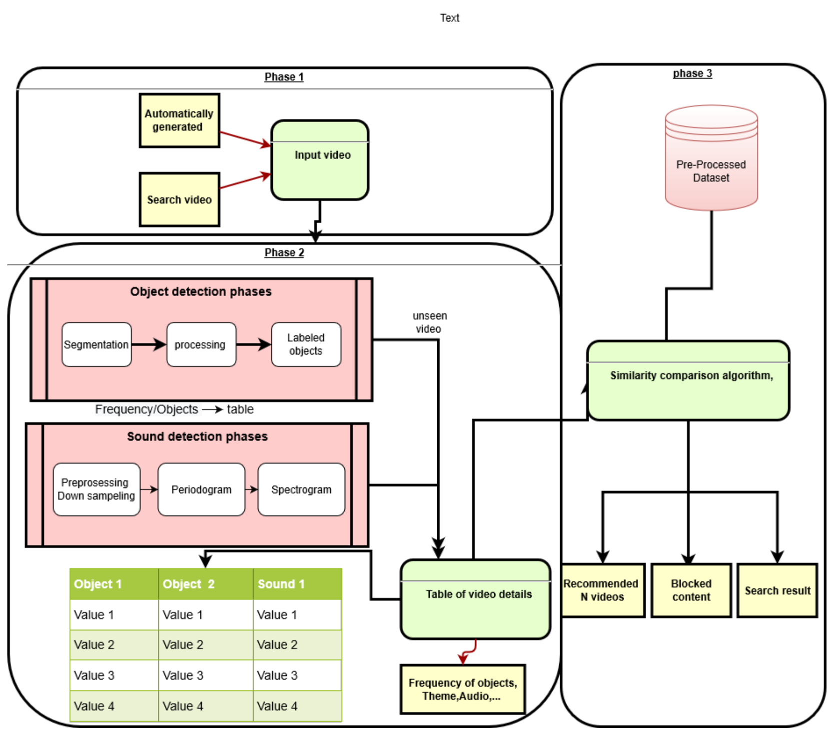


Figure 2: Proposed system overview

# 3 System Architecture

## 3.1 Architectural Design

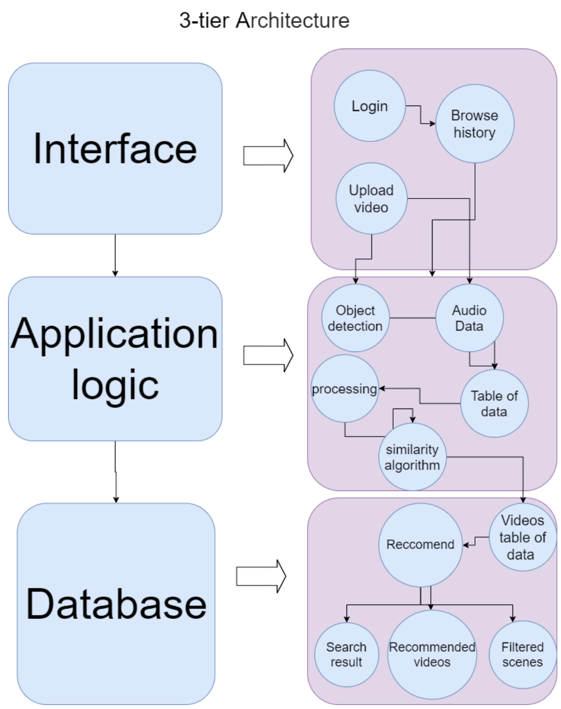


Figure 3: Architectural Design

## 3.2 Decomposition Description

Using 3-Tier architecture As shown in figure 3, Moving from top to bottom, Starting with interface, which is the plugin tool attached to any chromium based browser. The tool allows the users to upload videos or insert video's URL to get recommendation or search result, Also available in the plugin the ability to view and retrieve history. Phase two starts after the video is successfully uploaded the processing takes place in then application logic phase which contains all the algorithms required. The objects and the sounds in the video are used to create the sheet of data which will be used to compare the video relevancy with other videos. Finally The last phase takes place which will compare the sheet of data generated from the video inserted with the sheets of data stored into the database using cosine similarity. The results can be displayed to the user in form of three different outputs, recommended videos, search result or video after being filtered from the custom built filter made by the user earlier from phase 1.

### 3.2.1 Class Diagram

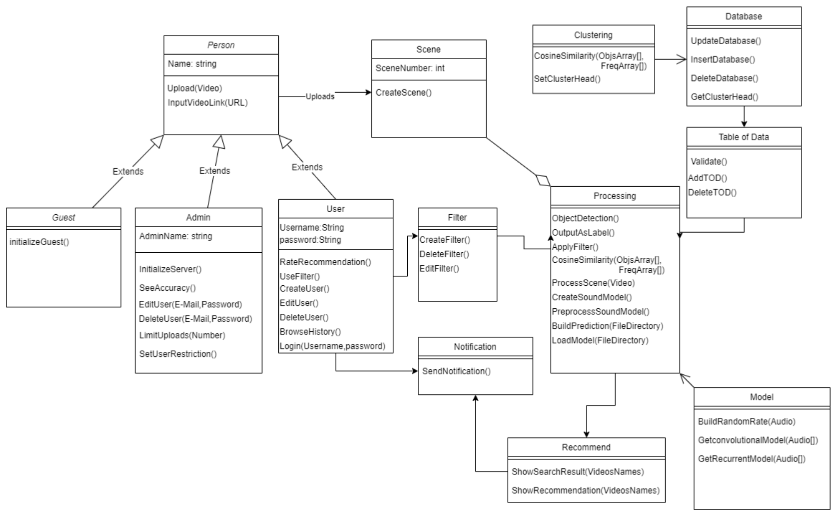


Figure 4: Class Diagram

### 3.2.2 System Activity Diagram


Figure 5: System Activity Diagram

### 3.2.3 System Sequence Diagram

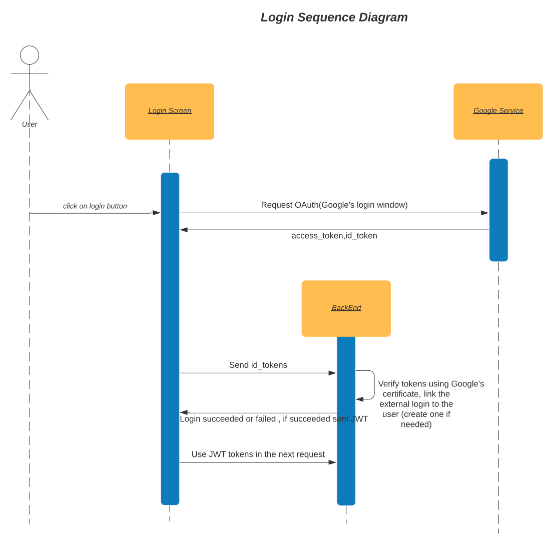


Figure 6: Login Sequence Diagram

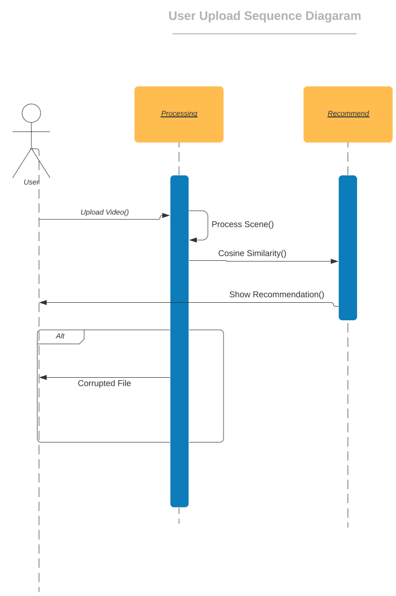


Figure 7: User Upload Sequence Diagram

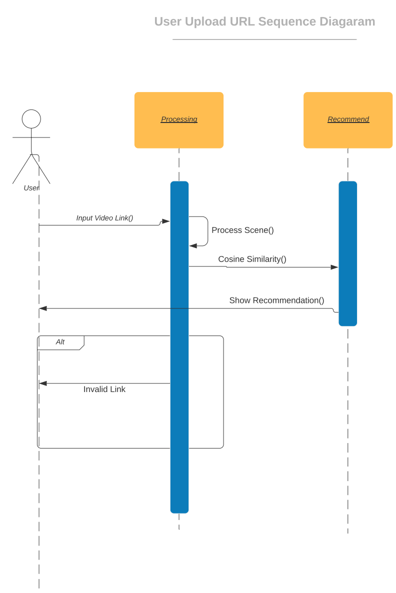


Figure 8: User URL Sequence Diagram

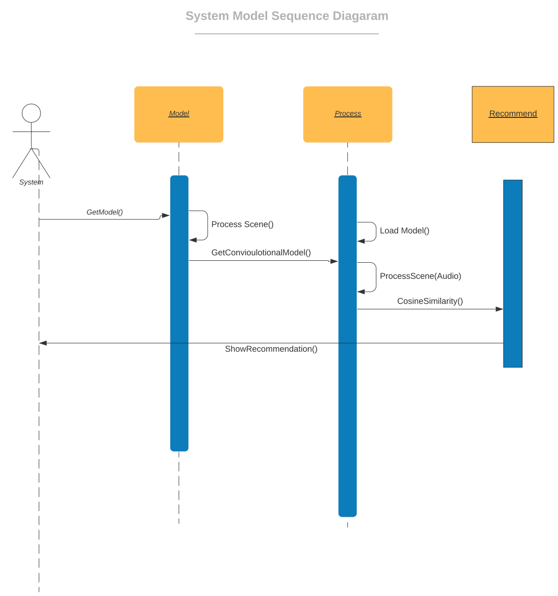


Figure 9: System Model Sequence Diagram

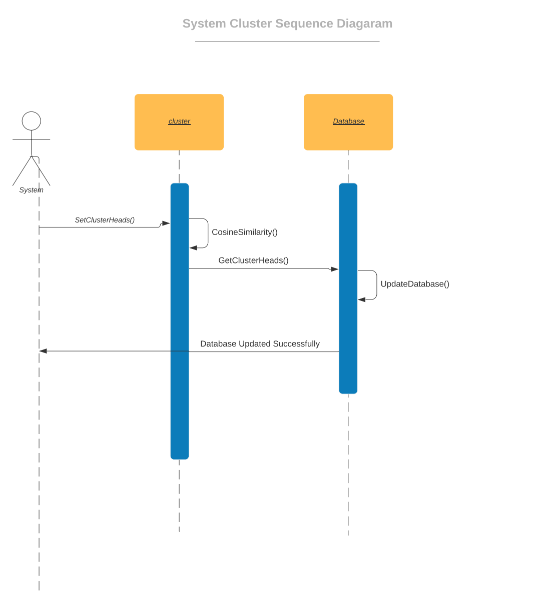


Figure 10: System Clustring Sequence Diagram

## 3.3 Design Rationale

This architecture allows our system to run seamlessly as the whole system processes are sequential.MVC architecture was not needed as its main purpose is to send data along Model,view and controller. Its better to use 3 tier architecture if the system is running in a sequential way as in this systems case. For the algorithms choice, A lot of algorithms are available to calculate similarity including Cosine similarity,Euclidean distance,Jaccard's intersection and Manhattan's distance. The Cosine similarity beats all of the above measurements where it measures the angle between the videos, rather than the distance in case of Euclidean distance. Therefore, making the similarity measurement much more accurate in terms of objects included. In addition, it uses the number of common attributes divided by the total number of possible attributes, rather than Jaccard's intersection divided by the union. Therefore, the best-used similarity technique for the proposed recommendation system is the Cosine similarity.

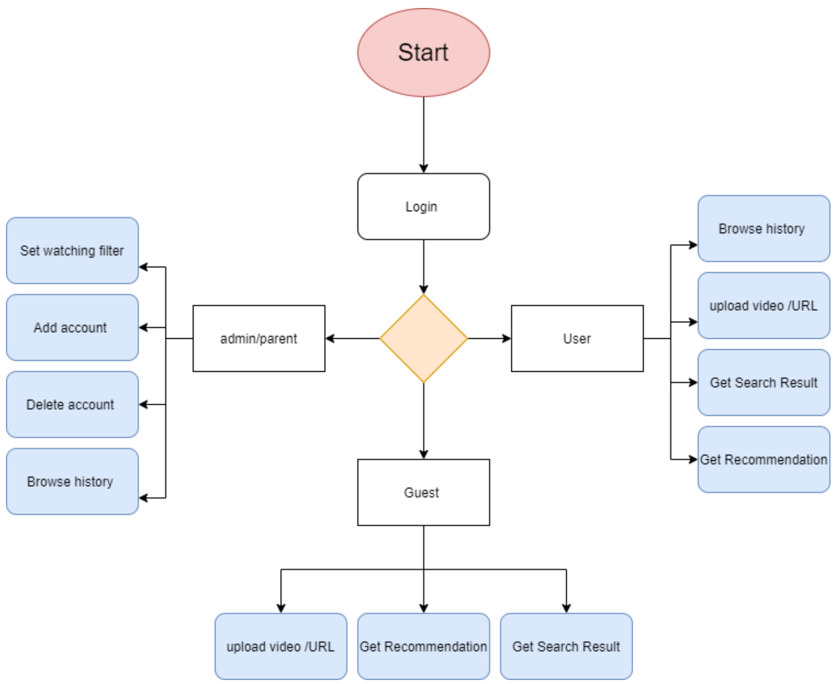


Figure 11: Process Diagram

# 4 Data Design

## 4.1 Data Description

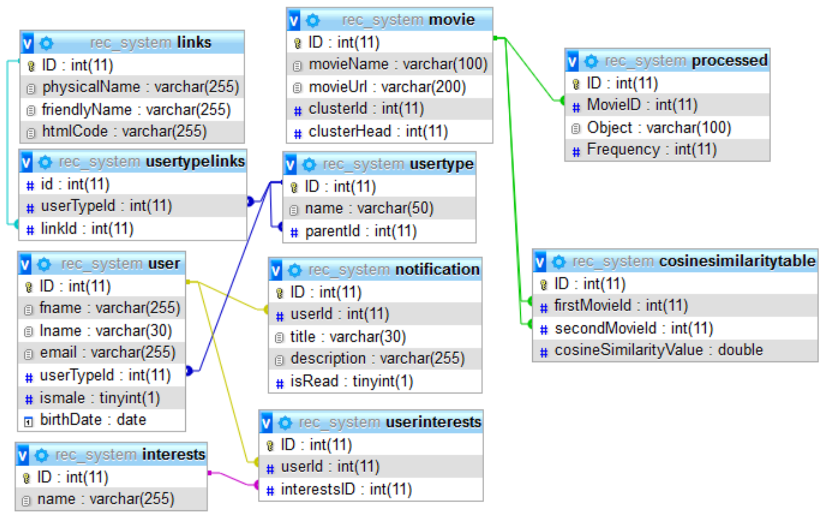


Figure 12: Database Schema

## 4.2 Data Dictionary

- user: This entity will hold the account information for every user. It will store information like: username, password and email.
- usertype: This entity will hold the type of the user.
- usertypelinks: This entity will store the allowed pages link of the plugin for each user type
- links: This entity will contain all the page links of the plugin
- userinterest: This entity will store every user interest of the videos in the plugin
- movie: This entity will hold the movie details
- interests: This entity will contain all interests users can be interested in
- cosine similarity: This entity will store the cosine similarity between videos in the database
- processed: This entity will store every object of all the videos have
- notification: This entity will hold the user notification and it's details.

# 5 Component Design

## 5.1 Machine Learning

### 5.1.1 Spectral Clustering

It clusters all videos into number of clusters to simplify the time of comparisons.

$$d_i = \sum_{j=1|(i,j)\epsilon E}^{n} w_{ij}$$

Figure 13: Spectral Clustering Equation

### 5.1.2 Fast Fourier Transform

It converts audio signals from its original domain to a representation in the frequency domain.

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j\left(\frac{2\pi}{N}\right)nk} (k = 0,1,\ldots,N-1)$$

Figure 14: Fast Fourier Transform Equation

## 5.2 Neural Network

### 5.2.1 ReLU

ReLU stands for Rectified Linear Unit, it is used in Convolutional neural network. We use it to eliminate the negative values in the neural network.

$$f(x) = x^{+} = \max(0, x)$$

Figure 15: Relu Equation

# 6 Human Interface Design

## 6.1 Overview of User Interface

The User interface is going to be flexible and easy. First the user will sign in with google, then the user will choose between inserting video url or upload a video from his pc, after this the system will show the user the recommended videos.

## 6.2 Screen Images

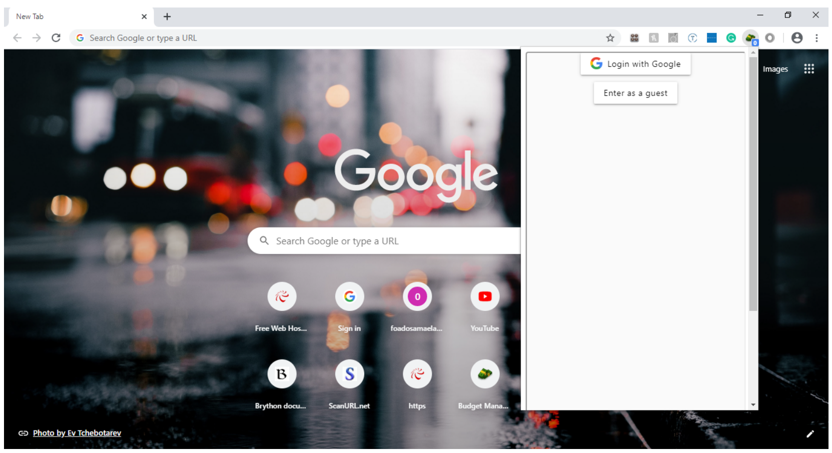


Figure 16: Sign In screen

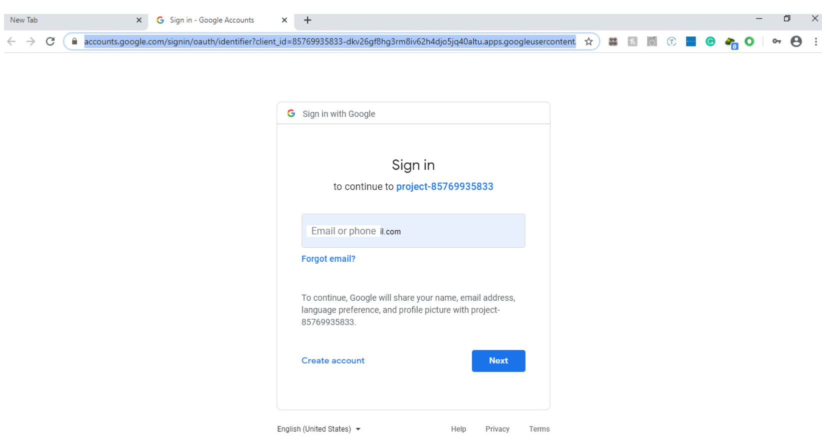


Figure 17: Sign In

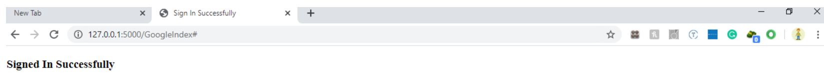


Figure 18: Sign Successfully message

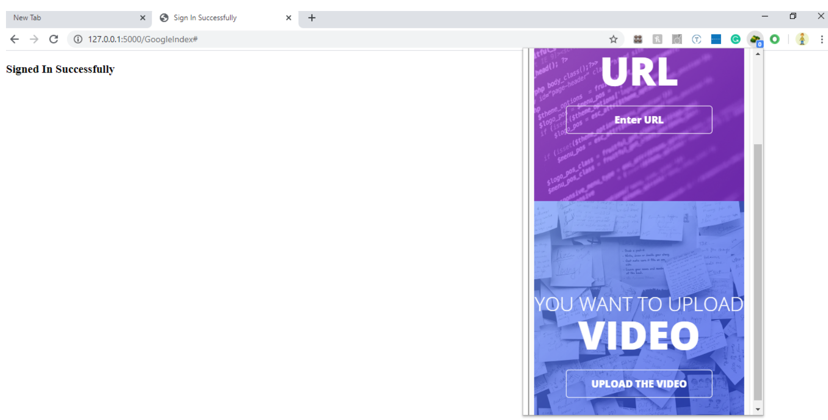


Figure 19: Video link or upload

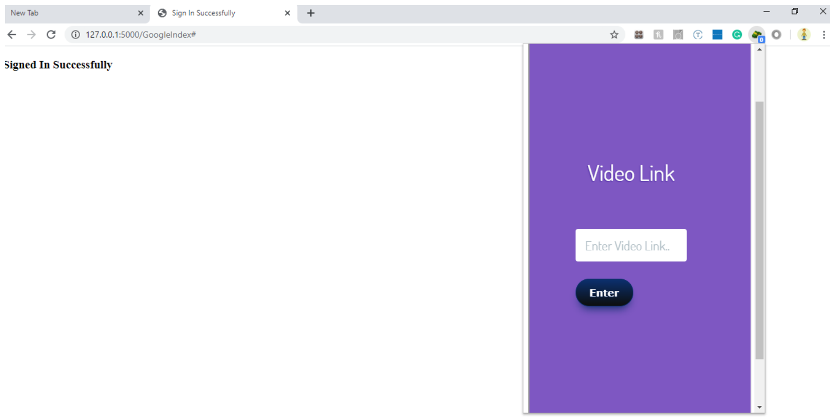


Figure 20: Insert Link

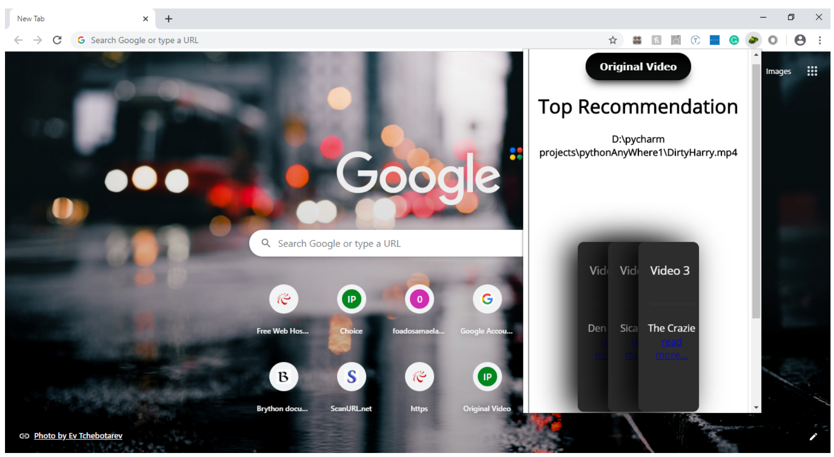


Figure 21: Recommended Videos

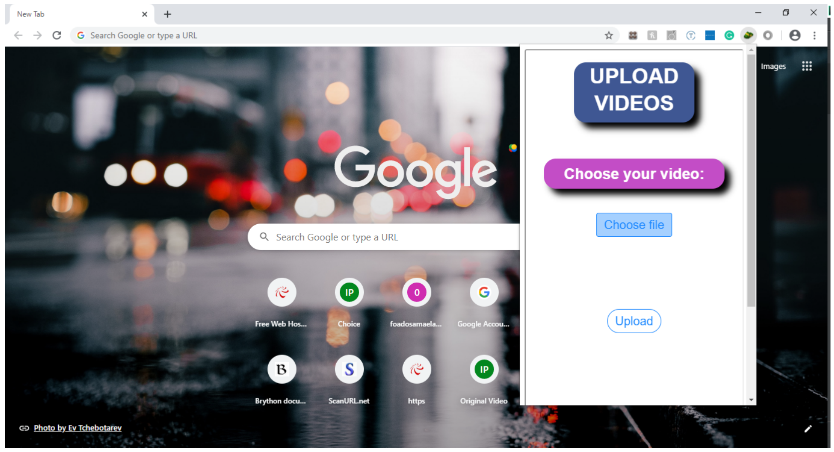


Figure 22: upload video

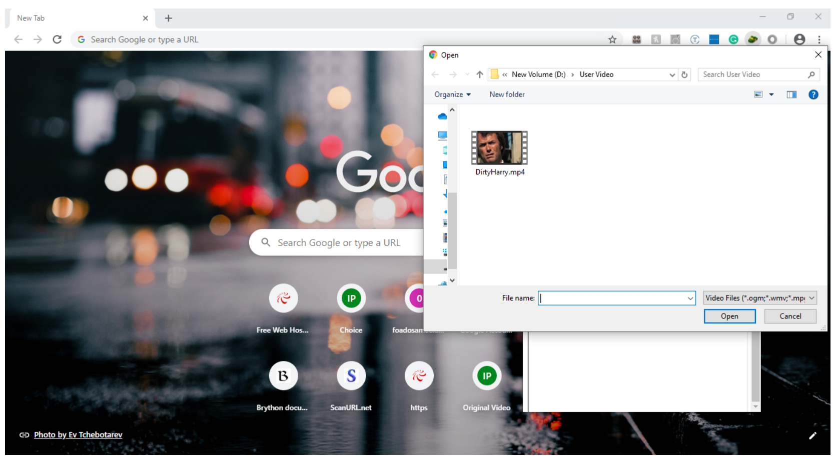


Figure 23: Choose video file

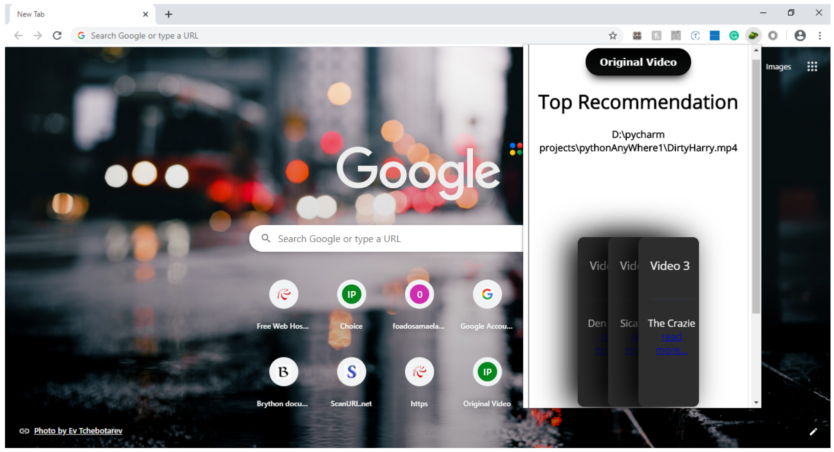


Figure 24: Recommended videos

## 6.3 Screen Objects and Actions

- Figure 16(Sign In screen): Selecting button sign in with google.
- Figure 17(Sign In): Go to sign in form of google.
- Figure 18(Sign Successfully message): Show the user a message to notify him that he is signed in
- Figure 19(Video link or upload): The user chooses whether to insert a video link or upload a video
- Figure 20(Insert Link): The user insert the link that he wanted to get for recommended videos.
- Figure 21(Recommended Videos): The system shows recommended videos.
- Figure 22(upload video): The user chooses to upload a video.
- Figure 23(Choose Video file): The user chooses a video file from his PC.
- Figure 23(Recommended videos): The system show the user the recommended videos for the video the user uploaded.

# 7 Requirements Matrix

| Code | Name | Type | Description | Test Strategy |
|---|---|---|---|---|
| F1 | Show Recommendation | Required | It allows the user to see the recommended videos based on the relevance calculated from the similarity. | Must give N number of recommendations to the user. |
| F2 | Select scenes | Required | It allows the user to trim the video and select the specific frames he wants to search with. | User can select scenes from the video he chose. |
| F3 | Search with video | Required | It allows the user to search with a video that he selected to get similar videos. | Results are returned to the user after he searched using his video. |
| F4 | Show history | Required | It allows the user to select videos from history and search with it to get similar videos. | User was successfully shown the history of videos he received before. |
| F5 | Upload video | Required | It lets the user to upload his own video. | Upload was successful and an error wasn't returned. |
| F6 | Insert video link | Required | Lets the user insert the wanted video URL. | Video is successfully uploaded and processed. |
| F7 | Create filter | Required | This function let the user creates a filter for mature content. | User was able to create a suitable filter for his needs. |
| F8 | Show Accuracy | Required | This function lets the admin show the accuracy of the system such as Clustering, Training sets. | Accuracy is returned to the admin without any problems. |
| F9 | Sign up | Not Required | It lets the user sign up to start the system functions. | Account was created successfully and saved. |
| F10 | Login | Not Required | It lets the user login with his user-name and password to start the system functions. | User can successfully login to the system without any problems. |

# 8 APPENDICES

## 8.1 Definitions, Acronyms, Abbreviations

- YOLO: You only look once Library.
- Flask: Python web framework.
- Darknet: Training module.
- CNN: Convolutional Neural Network.

- RNN: Recurrent Neural Network.

# 9 References

## References


[1] Xiaoyuan Su and Taghi M. Khoshgoftaar, “A Survey of Collaborative Filtering Techniques,” Advances in Artificial Intelligence, vol. 2009, Article ID 421425, 19 pages, 2009. https://doi.org/10.1155/2009/421425.

[2] Shi, Y., Larson, M., Hanjalic, A. (2014). Collaborative Filtering beyond the User-Item Matrix. ACM Computing Surveys, 47(1), 1–45. doi:10.1145/2556270.

[3] Tariq, S., Saleem, M., Shahbaz, M. (2019). User Similarity Determination in Social Networks. Technologies, 7(2), 36.

[4] Wang, W., Zhang, G., Lu, J. (2015). Collaborative Filtering with Entropy-Driven User Similarity in Recommender Systems. International Journal of Intelligent Systems 30(8), pp. 854-870.

[5] Pazzani, M. and Billsus, D. 2007 Content-Based Recommendation Systems. The Adaptive Web. (May 2007), 325-341.

[6] F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh, Recommendation systems: Principles, methods and evaluation, Egyptian Informatics Journal, Volume 16, Issue 3, 2015, Pages 261-273,

[7] Imani, M., Ghassemian, H. (2014). Feature extraction using partitioning of feature space for hyperspectral images classification. 2014 Iranian Conference on Intelligent Systems (ICIS).

[8] Li, Y., Wang, H., Liu, H., Chen, B. (2017). A study on content-based video recommendation. 2017 IEEE International Conference on Image Processing (ICIP).

[9] Yoshida, T., Irie, G., Arai, H., Taniguchi, Y. (2013). Towards semantic and affective content-based video recommendation. 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW).

[10] Jain, S., Pawar, T., Shah, H., Morye, O., Patil, B. (2019). Video Recommendation System Based on Human Interest. 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT).

[11] Kumar, Y., Sharma, A., Khaund, A., Kumar, A., Kumaraguru, P., Shah, R. R., Zimmermann, R. (2018). IceBreaker: Solving Cold Start Problem for Video Recommendation Engines

[12] Bhabad, D. T., Therese, S., Gedam, M. (2017). Multimedia based Information Retrieval Approach based on ASR and OCR and Video Recommendation System

[13] Zongxian Li1,2, Sheng Li1, Lantian Xue1,2, Yonghong Tian1,2† 1 National Engineering Laboratory for Video Technology, School of EECS, Peking University, Beijing, China 2 Pengcheng Laboratory, Shenzhen, China

[14] Seko, S., Motegi, M., Yagi, T., Muto, S. (2011). Video content recommendation for group based on viewing history and viewer preference. 2011 IEEE International Conference on Consumer Electronics (ICCE).

[15] Baviskar, P., Gunjal, P., Sirohiya, R., M anwar, S. (2017). A Survey on "User Search Recommendation System for Videos".International Journal of Innovative Research in Science, Engineering and Technology.

[16] Zheng, L., Min, F., Zhang, H., Chen, W. (2016) Fast Recommendations with the M-Distance, IEEE Conference 2016.

[17] N.A., L. (2009). Hidden Markov Model for Content-Based Video Retrieval. 2009 Third Asia International Conference on Modelling Simulation

[18] Hiwatari, Y., Fushikida, K., Waki, H. (n.d.). An index structure for content-based retrieval from a video database. Proceedings Third International Conference on Computational Intelligence and Multimedia Applications. ICCIMA'99

[19] William I. Grosky, Multimedia Information Systems, IEEE Multimedia, Spring 1994.

[20] Rune Hjelsvold, Roger Midtstraum and Olav Sandatå, Searching and Browsing a Shared Video Database, Multimedia database systems, Kluwer Academic Publishers 298-317, 1996.

[21] M. Flicker. Query by image and video content: the qbic system. IEEE computer, 23-32, 1995.

[22] Eitetsu Oomoto and Katsumi Tanaka, OVID: Design and Implementation of a Video-Object Database System, A Guided Tour of Multimedia Systems and Applications, IEEE Computer Society Press, 1995.

[23] Mallik, A., Chaudhury, S., Jain, A., Matela, M., Poornachander, P. (2007). Content Based Re-ranking Scheme for Video Queries on the Web. 2007 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology - Workshops.

[24] H. Ghosh, S. Chaudhury, K. Kashyap, and B. Maiti. Ontology specification and integration for multimedia applications. In Ontologies, volume 14 of Integrated Series in Information Systems, pages 265–296. Springer, 2007.

[25] Covington, P., Adams, J., Sargin, E. (2016). Deep Neural Networks for YouTube Recommendations. Proceedings of the 10th ACM Conference on Rec- ommender Systems - RecSys '16.

[26] Kosamkar, P., Potey, M. (2012). Feature based retrieval for animation video. Proceedings of the International Conference on Advances in Computing, Communications and Informatics - ICACCI '12.

[27] P. Geetha, Vasumathi Narayanan, "A Survey of Content-Based Video Retrieval," IEEE 2008.

[28] Parmar, M., Angelides, M. C. (2015). MAC-REALM: A Video Content Feature Extraction and Modelling Framework. The Computer Journal, 58(9), 2135–2171.

[29] Moens, M.F., Poulisse, G.J. and Vrt, M.M. (2012) State of the art on semantic retrieval of AV content beyond text resources.

[30] Yarmohammadi, H., Rahmati, M., Khadivi, S. (2013). Content based video retrieval using information theory. 2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP).

[31] Deldjoo, Y., Elahi, M., Cremonesi, P., Garzotto, F., Piazzolla, P., Quadrana, M. (2016). Content-Based Video Recommendation System Based on Stylistic Visual Features. Journal on Data Semantics, 5(2), 99–113.

[32] Bai, X., Wang, M., Lee, I., Yang, Z., Kong, X., Xia, F. (2019). Scientific Paper Recommendation: A Survey. IEEE Access, 1–1.

[33] Chaudhary, Pankaj and Deshmukh,A.(2015) "A Survey of Content Aware Video based Social Recommendation System." 2015.