

A Recommender System for Youtube Tags using Figurative Assessment - Abhishek Jha

About the speaker

My name is Abhishek Jha I am currently pursuing Btech In Computer Science from Atharva College of Engineering , Malad. I currently do projects and also freelanced for a web development project. I like to try out new and exciting technologies that may make my work simpler and interesting. I started using python in my projects as it removed the unwanted complexities such as syntax, long package import statements.

I am currently in a team named CodeStrike, a chapter of Codechef where we encourage our students in college to try out such languages to grasp the basics of programming. I am also currently doing image processing using OpenCV for my final year project in Augmented Reality domain.

Abstract

YouTube is one of the leading platforms for visual content creators that allows user to add, view, rate and share the available content. It is accessible around the globe, which makes it a great space for people who are willing to put their content out in front of the world. The content on YouTube is segregated into various domains like Entertainment, News, etc. Content segregation helps the video recommendation algorithm in providing proper suggestions to the users. Content creators are the individuals that upload original content containing news, live incident recordings, entertainment, etc. regularly. According to the statistics provided by YouTube, a total of 400 hours of content is uploaded every minute making it the second most accessed website. Making content is not the only challenge that the content creators face, the other challenge is to reach out to the target audience. The content creator has to add tags to the video before publishing which helps the SEO categorize the video. Social websites such as YouTube allow users to add common flexible tags to user-created resources, hence generating specifiers for content without any extra effort. The currently used ways to provide content specific tags are not optimal and can be improved.

In this project we have studied trending videos of YouTube, and extracted the tags, respective to the categories, that will help the video do better on the SEO, and make them available to all content creators. Video classification on the basis of various filters requires the videos to be tagged using various metadata labels. Appropriate metadata labels increase the targeted audience reach. Usually, it is observed that the user interaction i.e likes, dislikes and comments increase with the presence of right content. We have analyzed the top 16000 trending videos from the Great Britain and the United States, to extract what tags do the top content creators add, for their respective videos. Our algorithm works on the scipy stack, that uses pandas, numpy, and scipy, and the web-app is implemented using django. We have used all the open source libraries and languages to achieve the above system. Using these libraries helped us in creating reusable and a flexible system. Content uploaded with the suggested tags result in more user interaction and an increase in specific audience reach.