International Journal of Trend in Scientific Research and Development (IJTSRD)
Special Issue on Innovations in Computer Science and Applications
Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

An AI-Based Automated YouTube Video Publishing and Analytics System

Abdul Amaan Shah

Department of Science and Technology,
G. H. Raisoni Skill Tech University, Nagpur, Maharashtra, India

Abstract

Digital media has allowed many people to create content by publishing their own videos online. However, the largest video-sharing service that has developed over time is YouTube which is also the largest publisher—by number of daily video uploads.

In order to be able to post videos to YouTube, there are a number of procedures associated with video publishing such as creating a title, writing a description, tagging with appropriate hashtags, creating a thumbnail image and publishing the videos at designated times.

The objective of this study is to develop an artificial intelligence-based (AI-Based) automated system called VidMind that helps simplify the overall process involved with video publishing on YouTube. When using the VidMind system, Content Creators will be able to publish videos to YouTube and have the system use AI to analyse the content of those videos for any video analytics purposes.

In addition, the information produced by VidMind will be shared with creators upon completion of all required steps to publish a video to YouTube so that the creator receives an email notification when an uploaded video has been posted.

The proposed system performs a sentiment analysis on user engagement data (comments and reactions) to analyze how an audience perceives the video currently being viewed, as well as to provide overall performance analysis of each individual video; therefore, one of the system's proposal's goals is to automate as well as to decrease the amount of manual labor needed by content creators to create and/or publish their video(s); and as a result increase the productivity level of content creators and provide improved content management practices.

In addition to providing automation of workflows, the system creates more accurate metadata (or data describing the content of the video), and ensures metadata accurately reflects what the actual context of the video is; thus increasing the overall accuracy of metadata that is produced. The proposed system contains capabilities for enhancing a creator's focus on creating content of high quality, while reducing the time they spend addressing technical issues that arise from publishing a video. The VidMind System provides the means for monitoring the video performance using engagement analytics to enable content creators to determine the level of audience engagement with their content; ultimately demonstrating how AI and automation can greatly enhance the performance, scalability, and content management practices of today's video creator.

KEYWORDS: *Artificial Intelligence; YouTube Automation; Video Publishing System; Content Creator Tools; Metadata Generation; Sentiment Analysis; Databricks; Python.*

1. Introduction

YouTube, as an example, is one of the most established, or popular, digital media platforms to have great growth over the last several years and has changed the way people create, share and consume information. YouTube is one of the largest online video-sharing sites in the world, with content creators across all subject areas creating and uploading millions of videos each day. Given the huge growth in digital content, the level of competition between content creators is much higher than before. Creating and posting a video on YouTube has far more components involved than just uploading the actual file; there are at least five additional steps that must also take place before a video is able to be seen by its intended audience, such as developing a unique and catchy title; writing an informative description of the video; choosing the appropriate hashtags (tags) ; selecting a thumbnail for the video; and scheduling when the video will go live. All of these components combined will help to make up the video's metadata. The video's metadata will ultimately allow the YouTube algorithm to recognize the video and recommend it to the most relevant audience members (i.e., users that are most likely to be interested in the video(s) post). The more creative, informative, and strong the metadata is for each video, the better the video will rank in search and the more engagement it will receive from its audience. In order to create optimal metadata, a content creator needs to have a high level of creativity, be knowledgeable about trending keywords, and have a thorough understanding of search engine optimization (SEO) principles, as many of the content creators may find it difficult to do so.

recent development of machine intelligence (MI) and the development of natural language processing (NLP) has made it possible to create automated systems that can process data and generate textual representations of that data (the content) and the content is of good quality. Therefore, more companies are using MI and NLP-based automated systems to accomplish many different types of tasks such as: generating automated content; developing recommendation systems; building chatbots; and significantly enhancing the efficacy of their digital marketing efforts. Additionally, using MI technology, companies can also create automated systems that are able to analyze video (streaming media) content and automatically generate optimal title/ description/ hashtag/ other metadata for that video. This will save the video publisher from doing this work manually and improve the quality and consistency of the metadata that is generated.

The primary objective of this research is to introduce the VidMind automated YouTube video publishing system, a machine intelligence (MI) system designed to simplify the video publishing workflow for video makers. Using VidMind, a video maker can upload their finished video(s) to the VidMind website, and the system will automatically analyze

the video and produce an optimized set of metadata (title, description, hashtag, suggested thumbnails, etc.) for the video in preparation for uploading to YouTube after it is published. Once the video has been published, the video maker will get an email confirmation of the publication.

Also, the system will use sentiment analysis to analyse comments from audiences, Likes and reactions to content or

videos, in order to determine how viewers see the content. By reviewing data collected through sentiment analysis, the system will assist video creators with obtaining feedback from viewers who engaged with the video and enabling them to evaluate the success of the video that they create. The primary goal of the VidMind is to automate frequent tasks of creating videos; therefore offering an intelligent video development solution that will improve the productivity of content creators. Because of this, as the time needed to video upload and create metadata is reduced, content creators will have increased time to develop content to engage with the audience.

Moreover, not only does the system aim to simplify the video publishing process, the aim of the entire system is to also improve the overall efficiency of the process of content management for content creators. With the rate of growth of digital video content, content creators need intelligent tools to assist them in managing a large number of video files and ensuring equal quality with all their metadata. Automated system such as VidMind can help reduce repetitious manual tasks so that the content creators can work in a manner that allows for them to have more time to produce new content and engage their audience.

Digital media workflows can benefit greatly from the adoption of artificial intelligence within the marketplace and the implementation of video publishing platforms. With the ability for AI to quickly analyze large amounts of different types of media files such as audio, video, and images, identify trends and provide insight into how to make improvements in video editing and content creation that would otherwise take a lot of human labor to accomplish, AI could help to develop more intelligent and effective content management systems.

Automated publishing combined with metadata enhancement and audience satisfaction assessment on one system is a substantial improvement to the way digital media automation functions. This type of system allows for increased productivity by content producers and provides them with the ability to make more informed decisions through insight derived from data concerning how to best engage audiences through their behaviors and participating in activities.

2. Related work

The continual advancement of digital media platforms over time has led to a rising demand for intelligent systems that help content creators manage and distribute their online video content in a more efficient manner. As one of the most popular places for video sharing on the internet, dozens of millions of content creators upload videos to YouTube each day. As there is so much content on this platform, it is essential for content creators to optimize their videos to improve ranking and audience interaction. Meda experts and analysts have studied how to automate many aspects of the content management process using Artificial Intelligence and machine learning techniques. One primary aspect of this research has been to produce automated metadata. Metadata plays a fundamental role in how videos will be recommended to users and how soon after they have been uploaded, so researchers have suggested methods for generating optimally structured metadata through Natural Language Processing and keyword extraction.

Speech recognition technology has made huge strides in the area of multimedia content analysis. Using an automatic speech recognition (ASR) system to convert spoken audio into a text-based transcript enables contextual analysis of video content through the creation of summaries, captions, and other forms of metadata that enhance both accessibility and searchability of video content. In addition to transcripts, researchers are working on developing methods for the automated generation of thumbnail images that visually depict what the video is about to help attract viewers to the video content. These methods use recent advances in the fields of image generation technology and computer vision techniques to automatically create relevant thumbnails from specific video frames and to choose an appropriate thumbnail for the video. As well as providing authors with the tools to create multiple forms of metadata, some platforms also implement analytics systems that provide information on the number of views, number of likes, number of comments, and total time watched of each video that is helpful to content creators in determining how well their videos are performing and finding trends with their audience.

algorithms and sensors to facilitate the publishing process by automating the manual tasks that were once performed by people.

Despite these developments, video publishers continue to face challenges in optimizing the audience experience and maximizing engagement because they rely on time-consuming and inefficient manual methods to gather audience feedback and gather metadata.

Recently, the research community has recognized and responded to this gap in the video publishing ecosystem, as there is growing interest among publishers and developers of AI solutions to create comprehensive solutions to meet the challenges of the industry. This research will document current technologies and applications of artificial intelligence in video publishing and media management as well as highlight strengths and limitations within the existing video production software and online video platforms available today.

Despite progress in using automated video analysis, currently available products are limited to only providing either metadata or reports. They are not able to perform all functions, from video analysis and creation, through to uploading to YouTube, and then performing analysis of viewer sentiment. VidMind is a complete content creation and audience engagement solution with all automated capabilities built into one single solution.

3. Research Methodology

The research methodology for the VidMind System focuses on developing a comprehensive automatic framework that allows for simple video publishing on YouTube. The VidMind Systems creates fully automated ways to help with the management of responsibilities for content providers through video processing, artificial intelligence technology, and automated publishing capabilities.

The first step in the research methodology is to upload and prepare the video. The author or content provider uploads an edited video into the VidMind System. Once the video has been uploaded, the video will be processed using video processing libraries to extract key attributes from the video (e.g., audio tracks, frames). The audio and frames extracted from the video provide additional contextual information that can be used to evaluate both the video visually and the sound of the video.

Once the audio has been extracted from the video, a text document or transcript is created from the audio. The audio conversion that is created allows for the audio information extracted from the audio to be converted to a written document for the purpose of evaluating the context of the video.

Once the transcript is created, Natural Language Processing (NLP) will be applied to the transcript.

After generating the required Metadata, the system is setup to operate through the YouTube Data API remotely. Therefore when a video is fully created and ready to launch you will create its title, description, tags and thumbnail (the image that represents the video). After the video has been published YouTube will send the system engagement metrics (such as views, likes and comments) in order for the system to store them in a structured manner so that they may be used for analysis in the future to assess video performance. The final step of the process is to conduct sentiment analysis of user comments so that machine learning models can use those comments to evaluate viewer impressions of the video. This analysis provides insight into viewer perceptions of the video as well as allows publishers to adjust their content based upon viewer behaviours.

The proposed system employs a modular architecture to allow for optimal processing rate and scalability with each component performing specific functions, such as processing video pre-productions, generating a transcript, creating metadata, and automating the publishing process. This modular configuration provides improved flexibility in the overall system's functions, and additional modifications/upgrades can be made on a stand-alone basis of each component without negatively impacting the overall operation of the system as a whole.

This also includes several app features that manage the data that has been processed, as well as the engagement metrics that have been acquired from YouTube, using various data storage mechanisms, all of which will eventually be used to perform more analyses on your uploaded videos' performance over time. Additionally, as part of the methodology, there is an emphasis on automating the processes and integrating the technology, so that each step in the entire process of the system will occur in a sequence that will ensure the most efficient way of completing the processes of analyzing a video, producing the metadata about it, and then distributing it through to publication. Therefore, through integration, the complete system has been made more reliable and effective due to the amount of minimal manual intervention that will be required within the system.

4. System Architecture

The proposed VidMind system automates the publishing and metadata generation of videos for YouTube creators with its architectural design. The VidMind system consists of many components including, but not limited to, video processing tools, Artificial Intelligence models, Data Storage and Analytics modules. The VidMind system functions as a unified automated workflow by integrating these components together.

To begin the video upload to VidMind, the media creator produces and/or edits their video content. After the creator uploads their video content to VidMind, the system gets triggered to establish a processing pipeline in which the content will be analysed and the video will generate the required metadata for publishing.

The first phase of the VidMind System's Architecture includes Video Processing and Content Extraction. Within VidMind, the Video Processing Library consists of the tools used to extract Audio tracks and Keyframes from the uploaded Video. The Keyframe and Audio Track extraction allows the system to analyse both the Video's Audio and Video components. In the second phase, we process the extracted Audio using Speech Recognition Models which allow us to transform the word that is spoken in the video into text transcripts. By doing so, we understand the subject matter and context of the video. The transcripts will then be provided to the Natural Language Processing Models for further processing.

NLP models are capable of analyzing transcripts to determine key features such as keywords, phrases and relevant context. The results of these analyses allow the system to create metadata elements like titles, descriptions, and hashtags automatically for use with the video on YouTube. These metadata elements enhance the video's ability to be found on YouTube once it has been posted. Additionally, the architecture includes an automated module for generating thumbnails by selecting or creating attractive images from a video frame. Thumbnail images play an important role in getting potential viewers to see the video and improve the possibility that they will click on it. After the metadata and thumbnail have been generated, the system automatically uploads the video to YouTube with the appropriate metadata using the YouTube Data API before publication.

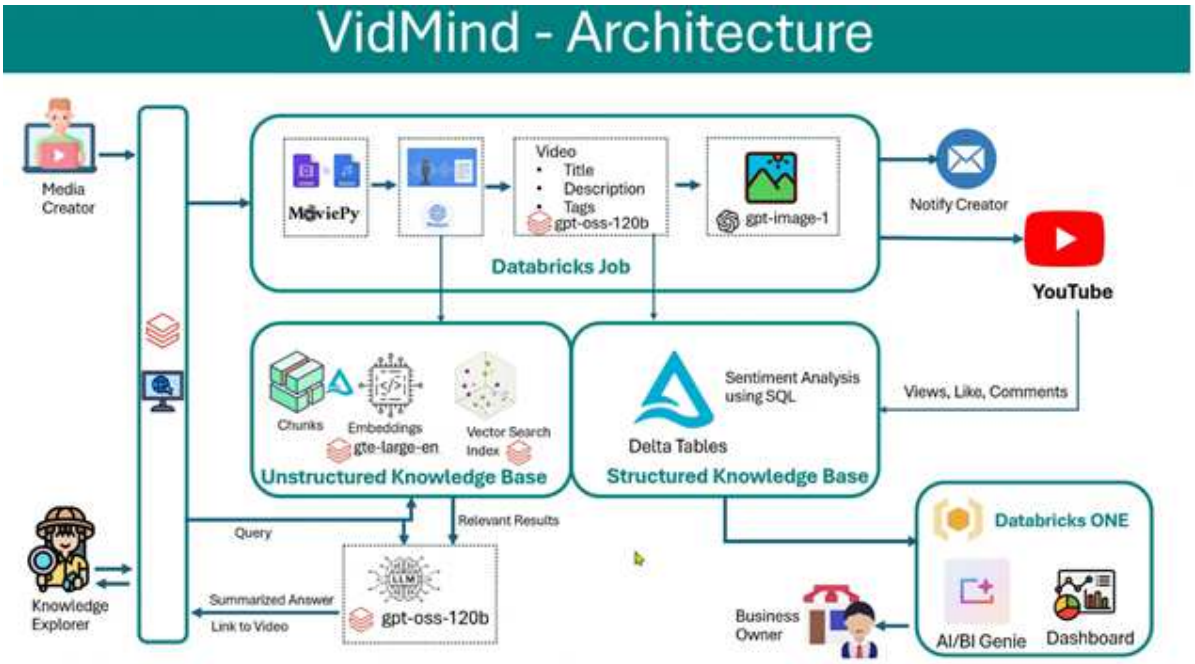

Fig 1: VidMind Architecture

The video will be uploaded, an email will be sent to the creator notifying them of the upload's success. Along with being published, an architecture for analytics (data) used to collect engagement metrics (likes, views, and comments); this data will be transferred into a structured data storage area so that it can be analyzed, as well as performing a sentiment analysis on user comments in order to analyze audience responses.

The results of the sentiment analysis will assist video content creators in assessing the effectiveness of their video content, as well as developing an accurate portrayal of how their viewers feel in relation to their video content.

The results of the analytics process will be made available to creators and administrators through dashboards to provide a comprehensive view of video performance, as well as engagement metrics related to video content.

5. Implementation Details

The VidMind system utilizes multiple technologies and software applications created together to create an automated video publishing system, by capturing video uploads and producing relevant data to support the automated publishing of videos to YouTube as well as analyzing how well the audience engages with those videos.

Python was selected as the primary programming language for implementing the applications, as it provides an abundance of libraries for the processing of digital video and audio, Natural Language Processing, and artificial intelligence. Also, it has the ability to integrate with many artificial intelligence models and to integrate with external application programming interfaces for building intelligent automated systems.

The first step in implementing the system is to process the uploaded video, which requires the utilization of video processing library functions to receive content from the uploaded video; for example, to extract the audio track from the uploaded video, to extract video frames from the uploaded video, etc. These audio and video components will contain all the data needed for additional analysis.

Once the audio has been obtained from the uploaded video, this audio content will then be processed using speech recognition to convert the sounds (the voice) into a textual version (or transcript) of what was said. This transcript of what was said allows the system to comprehend the meaning of the spoken words within the video and to determine the most relevant "keywords" associated with the video. These transcripts then go through a Natural Language Processing process to detect important words and phrases, and the context of the information provided in the transcript, as it relates to the subject matter of the video.

After the creation of metadata for a video, metadata will then be uploaded to YouTube via the YouTube Data API utilizing an API Call. During this API call, a title, description, tags (if applicable), and thumbnail will also have been previously generated. Once the video has been uploaded to YouTube successfully, an email will be sent to the user confirming that their video has been successfully uploaded. Additionally, information about how well viewers are engaged with the video (e.g., total views, total likes, total dislikes, total comments and total performance metrics) is stored within analytic databases.

In addition to creating metadata and automating the publishing of videos, the system will utilize sentiment analysis on the comments made by viewers. This will be accomplished through machine learning, as all viewer comments will be analyzed through a set of models based on user comment text and sentiment (positive, negative or neutral). Sentiment analysis results will be a helpful reference for all users to see how well their viewers feel about their videos, which provides additional value to content creator's in reviewing and assessing the quality of their content. Therefore, implementing the VidMind system is an example of howthe utilization of Artificial Intelligence (AI) and Automation can streamline the overall processes associated with digital content Management and Publishing.

6. Results and Discussion

The VidMind system's application of AI for automation of a video publishing workflow exemplifies its use case as a

workflow tool for content producers. A primary benefit of the VidMind system is the decreased workload associated with producing the necessary metadata to publish a video on YouTube by automating the production of the metadata, for example, producing the title(s), description(s) and hashtag(s) needed to properly identify what was produced.

By utilizing the automated generation of metadata in this manner, content creators are able to reduce time spent on that part of the video production process and rely on AI technology to provide metadata that is consistent with what is available at the time of publishing the video.

Furthermore, when a content creator uploads a completed video into the VidMind platform, the automation of uploading the completed video and entering the required metadata about the uploaded video into the VidMind platform eliminates any need for manual upload of a completed video by the content creator and eliminates any manual entry of metadata for that completed video prior to the video being published on YouTube. This makes it very quick and easy for content producers to use the VidMind system. Finally, the thumbnail generation module within the system provides a content producer with a visual representation of the video for thumbnail creation.

To summarize, the findings show that VidMind automates the process of publishing videos as well as providing insights about the creator's work. This development will also allow content leaders/partners to use AI tools to increase their frequency of work, reduce the amount of manual effort they expend, and allow them to manage their content more effectively than they currently do.

7. Conclusion and Future work

As a result of the increase in the number of video-sharing sites that enable users and businesses to share their content with others and develop a digital audience, this has created an opportunity to develop intelligent systems that make it easier for content creators to upload their content. As video uploads rise on sites like YouTube, intelligent systems that automate the publishing process of video content need to be developed.

The VidMind platform will be an AI-powered automated system for assisting content creators to better manage their video publishing processes. VidMind is a complete solution that provides an all-in-one platform by automating the generation of video metadata, uploading videos to YouTube, and analysing audience interaction by analysing emotion.

By automating these tasks the VidMind platform will reduce the amount of time and effort needed for the content creator to do manual work. As a result of this, the content creator will have more time to focus on producing high-quality videos. Because the use of AI technology will allow VidMind to help content creators generate metadata for their videos that will allow them to find their videos more easily on the video-sharing sites they upload to. The content creator will also have access to valuable data through the results of the sentiment analysis performed on their videos regarding how their target audience is responding to their content, allowing them to optimise their content for their audience.

Over the coming years, there is potential for the system to be improved by adding new features that facilitate the automated editing of videos, allow for real-time detection of trending content, provide recommendations based on preferences of individual users, or analyze audiences' behaviour patterns through the use of predictive analytics. These advancements can lead to the creation of a fully integrated video CMS, which will rely on sophisticated software technologies.

Additionally, future enhancements could be realised by looking at ways to optimise existing automated systems and through sourcing new methods for generating metadata, such as retraining machine learning models, or conducting sentiment analysis, and by sourcing new datasets that could either create new or support the functionality of existing solutions..

References

[1] Brown, T., et al. (2020). *Language Models are Few-Shot Learners.*

[2] Pang, B., & Lee, L. (2008). *Opinion Mining and Sentiment Analysis.* Foundations and Trends in Information Retrieval.

[3] Apache Software Foundation. *Apache Spark Documentation.*

[4] Databricks. *Machine Learning and Data Engineering Guide.*

[5] YouTube Developers. *YouTube Data API Documentation.*

[6] Jurafsky, D., & Martin, J. (2019). *Speech and Language Processin*