How much of this submission has been generated by AI?

*4%

of qualifying text in this submission has been determined to be generated by AI.

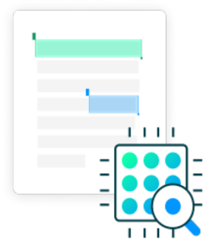
* Low scores have a higher likelihood of false positives.

Caution: Percentage may not indicate academic misconduct. Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Frequently Asked Questions

What does the percentage mean?
The percentage shown in the AI writing detection indicator and in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was generated by AI.



Our testing has found that there is a higher incidence of false positives when the percentage is less than 20. In order to reduce the likelihood of misinterpretation, the AI indicator will display an asterisk for percentages less than 20 to call attention to the fact that the score is less reliable.

However, the final decision on whether any misconduct has occurred rests with the reviewer/instructor. They should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in greater detail according to their school's policies.

How does Turnitin's indicator address false positives?
Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be AI-generated will be highlighted blue on the submission text.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

What does 'qualifying text' mean?
Sometimes false positives (incorrectly flagging human-written text as AI-generated), can include lists without a lot of structural variation, text that literally repeats itself, or text that has been paraphrased without developing new ideas. If our indicator shows a higher amount of AI writing in such text, we advise you to take that into consideration when looking at the percentage indicated.

In a longer document with a mix of authentic writing and AI generated text, it can be difficult to exactly determine where the AI writing begins and original writing ends, but our model should give you a reliable guide to start conversations with the submitting student.

ABSTRACT

Traditional methods of sports summarization heavily depend on large teams of manual editing, where enormous portions of game footage are picked over through by humans to select the most crucial moments for the compilation of highlights. This is, however, a time-consuming, resource-intensive, and uneven-coverage of events. Our work proposes a novel, multi-modal approach to sports summarization, in which all things are combined: Twitter data, audio features, and video content are finally computed to automate and enhance the entire process of sports summarization. This approach is based on the integration of diverse sources of information, which should streamline the summarization process and hence enhance efficiency in covering sports events. This methodological novelty has the potential to transform sports summarization into a scalable and efficient solution, delivering engaging and informative highlights to sports enthusiasts from all over the globe.

CHAPTER 1

INTRODUCTION

Sports Video summarization is the condensing of lengthy sports videos into shorter summaries that capture the important events of a game. In an age where people do not have the time to sit for hours together to watch their favorite game, this can help viewers efficiently enjoy the entire game in very little time.

Traditional approaches to summarizing sports videos rely heavily on manual efforts by video editors who sift hundreds of hours of video footage from multiple camera angles and positions in order to identify the significant events and then manually edit this into a highlights video. This process is not only expensive but also time consuming. It also lacks an element of audience engagement. However, this approach can be made better by leveraging the advancements in computer vision, deep learning and natural language processing. Furthermore, the audience reactions can be captured from various social media platforms to help create a better experience for the sports viewers.

The goal of this project is to create a system that can efficiently summarize sports videos utilizing the various advancements in computer vision, deep learning and natural language processing. Further, this project also aims to capture user sentiments by analyzing the tweets related to the game, thereby enabling us to effectively capture audience reactions and insights to create more engaging highlights for the viewer.

The core of this project involves a multi-modal approach that seamlessly integrates multiple modals of data - visual, audio and textual in addition to the twitter data of that game. This ensures that our system is accurate and effective in generating summaries of sports videos and also adds audience engagement and insights into our video summarization process.

Audio Analysis will be crucial for our system in order to help identify events based on valuable insights obtained from the audio track such as crowd intensity, commentator fervor, and player reactions, adding the summaries with a deeper layer of context and emotion.

Visual Analysis will play a key role in ensuring that no event goes unnoticed. The system will analyze visual elements such as scoreboards, player actions, and replays to identify pivotal moments and highlights within the videos.

Textual Analysis will be performed both on the commentary data of the game as well as on the twitter feed provided. This will ensure that there is a human element in summaries generated. The commentary will be analyzed using a large language model to classify the event as significant or not significant.

By incorporating various modalities into our project we aim to create a superior system than the traditional systems being used to generate highlights and additionally by adding audience reactions and insights from twitter we aim to generate a more enriching summary of the sports video.

CHAPTER 2

PROBLEM DEFINITION

Sports fans are unable to watch the whole match because of their busy schedule and other activities. The traditional approach of producing sports highlights is normally time consuming and the teams involved in the editing process need to put many hours in browsing through the overall in-game data and going through the diversity of possible moments and looking for those ones with the greatest potential to be included in the video summary.

The existing approaches of video summarization rely on basic algorithms that fail to capture the dynamic energy and excitement of the game. They lack the integration of content from social media platforms, which could enrich the summarization process with real time insights and reactions from the audience.

Our project addresses this challenge by proposing a multi-modal approach, leveraging the power of Twitter data, audio features, and video content to automate and enhance the summarization process.

This process enhances the accessibility and usability of sports content and it also improves user engagement and satisfaction and it provides analysis by identifying key moments through audio, video and twitter analysis to extract details from those moments.
It provides user-driven highlights based on detected events and user preference (favorite teams, players, event types). It provides concise and engaging highlights based on user choice.
Overall, it redefines the traditional method of sports content creation, providing a platform for dynamic and engaging content generation and empowers sports enthusiasts to stay connected with the latest sports events efficiently and effectively.

CHAPTER 3

LITERATURE SURVEY

The aim of the research paper [1] by Kaito Hirasawa, Keisuke Maeda,Takahiro Ogawa, and Miki Haseyama is to propose a method for the detection of important scenes in baseball videos by leveraging both Twitter data and video analysis in order to generate highlights automatically. They have proposed a method termed, Time-Lag-Aware Multimodal Variational Autoencoder (Tl-MVAE) that considers time-lags between tweets and multiple previous events in the video to improve the accuracy of scene detection. The Tl-MVAE comprises an encoder, decoder and an important scene detector. In the encoder architecture, the features extracted from tweets and videos are transformed into latent features and this is where they include the consideration of time-lags. The time-lags represent the influence of past events on tweets, assuming a Poisson distribution where events from present to past influence tweets. The decoder reconstructs the original data from the latent features, its aim is to closely resemble the original features extracted from tweets and videos. The important scene detector consists of multiple fully connected layers and takes the latent features from the encoder as input. Its objective is to classify scenes as important or normal. A predetermined threshold is used to determine if a scene is important based on the calculated probability. In order to evaluate the performance of the proposed model, they made use of specificity recorded to be 0.409.

The aim of the paper [2] is that A new approach toward summarizing football match videos through advanced technologies tries to secure the end-users' accessibility and experience in consuming video content.The paper proposes a new approach toward video summarization with the integration of deep neural networks and semantic mapping techniques. This new approach will alter the perspective of how videos can be summarized more effectively. The method may heavily depend upon the availability and quality of the training data. In case of small and biased training data, the method is likely to produce worse performance or biased summaries.

The aim of the paper [3] is to leverage large language models to aid in the generation of sports summaries. The approach mentioned in this paper uses the YOLO model in order to split the video into 20 second segments based on the bowler's position and later feeds the commentary data into a large model that is trained to identify events based on words used for each event. The approach used in this paper has an accuracy of roughly 97% which has been achieved due to the large corpus of words that has been identified for each type of event. The use of a BERT that has been fine tuned for English commentary data has also aided in achieving such a high level of accuracy. Misclassifications occur when the commentators describe the game's bigger picture, separate from the actual ball action, such as boundaries and wickets. There are also times when ball actions come to an abrupt end, which causes a discrepancy between the extracted comments and the real event sequence. The commentary's temporal misalignment separates the model's narrative depiction from the real event.

turnitin Page 11 of 43 - AI Writing Submission Submission ID trn:oid:::1:2913957965

The aim of paper [4] is that A novel approach weighted dynamic heartbeat graph for detection events from twitter stream.The text stream was systematically transformed into a series of temporal graphs. These graphs inherited temporal frequencies and co-occurrence relationships of the words appearing in the text stream. Each graph was further used to extract a heartbeat score using two features: growth factor and aggregated centrality. A rule-based classifier labeled the graphs as event candidates. Multiple event candidates were merged to extract a list of ranked topics. For the performance evaluation of the proposed approach,there are three benchmarks FA Cup, Super Tuesday, and the US Election were used.

The aim of the paper [5] is to recognize and clip crucial occurrences in a cricket match by considering event-based attributes. The dataset used in this paper has been gathered from YouTube and Hotstar. The approach followed in this paper involves identifying key frames based on energy levels and then the features are extracted using CNNs(VGG16 and ResNet50). The features vectors are fed to an LSTM network to obtain word embedding and caption for each ball. The advantages of this paper is that it categorizes scenes in matches accurately, also this model covers a wide range of events including fours, sixes, wickets and player celebration. The limitations of this paper is that the ResNet model used for feature extraction requires significant processing power.

The aim of paper [6] Video summarization techniques were described, turning long videos into those with only important events.The process of video summarization involved the deep semantic features wherein the original video was cut into parts, and only relevant features were extracted from those parts.In conjunction with other techniques like DPP (Determinantal Point Process) that allow frame selection to provide diversity, LSTM approaches seem to be particularly highlighted.Limitations of activities include the requirement for large and diversified training datasets and the slow manual annotation process and the changes of video content over time.Future directions mention its potential application in sports to automate match highlights and filter video rightly.

The aim of paper [7] is to present a high-accuracy framework to automatically clip the sports video stream by using a three-level prediction algorithm based on two classical open-source structures, YOLO-v3 and OpenPose. The research paper focuses on racquet turn based events. The three-level prediction algorithm comprises low level prediction where a boolean decision for each player engaged in “playing” action or not is taken. The middle level prediction involves a boolean decision for every frame. It captures temporal aspects of player actions within the video, The high level prediction is to counter the errors that may arise in the low level prediction. The boolean decision is taken over a short video segment. In this paper, each frame in the short video segment represents whether there is a playing action or not. Confidence levels are calculated for each frame and if it crosses a certain threshold value, the segment is classified as “playing”.

The aim of paper [8] is to develop automatically detects and summarizes important events in cricket match.This model uses techniques like optical character recognition(OCR),sound detection, and replay detection to extract important events such as boundaries and wicket.This method combine both CNN and OCR for event detection and it gives higher accuracy compare to other techniques.The system splits up video into separate shots,detect keyframes using audio cues,recognizes scoreboard information using OCR, and generates highlights based on detected events.

CHAPTER 4

DATA

4.1 Overview

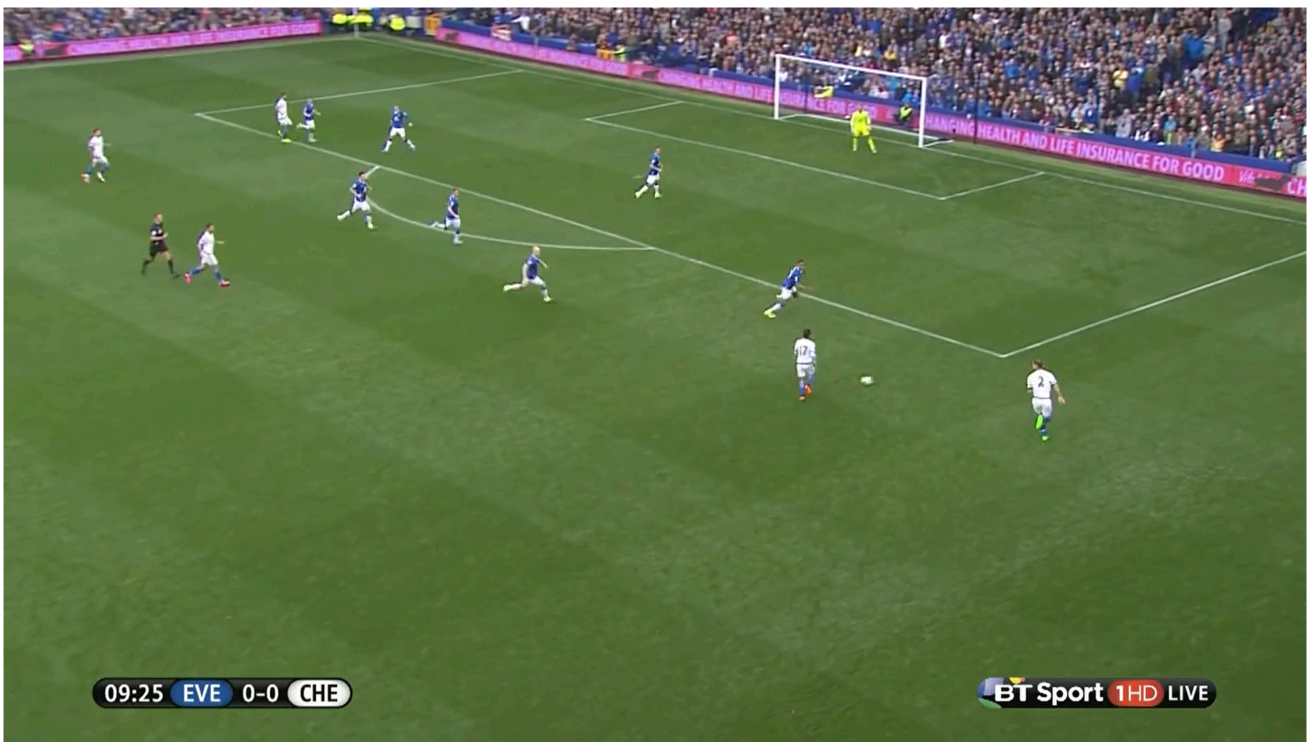
Our project has the requirement of two types of data, full length sports videos that have to be summarised and twitter data which has the features of each tweet related to the game.

For full length sports videos we will be making use of the SoccerNet dataset which is a large scale dataset for soccer video understanding. It is composed of 550 complete broadcast soccer games from the major European soccer leagues. This dataset has evolved over the years to include tasks such as action spotting, camera calibration, player re-identification and tracking. This dataset has a wide variety of soccer matches and will be beneficial for our project.

The tweets dataset consists of tweets starting from the Round of 16 till the World Cup Final that took place on 15 July, 2018 & was won by France. This dataset consists of Tweet ID which is a unique ID that is generated for every tweet. This is used to uniquely identify every tweet. Lang specifies the language of the particular tweet. The date column mentions the date along with the time that the tweet was put out. Source specifies the device used for making the tweet. Len tells us the number of characters in the tweet. Orig_Tweet has the tweet in the original form. Tweet is the original tweet after pre-processing to remove the hashtags and user mentions. Likes mentions the number of likes that the tweet has received. RTs is the number of times that the tweet has been shared. Hashtags specify the hashtags that have been mentioned in the original tweet. UserMentionNames is the name of the users that have been mentioned in the tweet. UserMentionID is the twitter ID of the user that has been mentioned in the tweet. Name is the name of the user that has made the tweet and place is the location of that user. Followers is the number of followers that the person who made the tweet has and friends is the number of friends that the tweeter has. This dataset contains the necessary fields in order to analyze the users reactions leading up to the match and also during the match.

4.2 Dataset

A frame from a video downloaded from the SoccerNet dataset



PES UNIVERSITY

Indexing and Summarization of Sports Videos using Multi-Modal Approach

Tweets.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	ID	lang	Date	Source	len	Orig_Tweet	Tweet	Likes	RTs	Hashtags	UserMenti	UserMenti	Name	Place	Followers	Friends
2	1.0136E+18	en	02-07-2018 1.35	Twitter for Android	140	RT @Squawka:	Only two goalkeepers have saved three penalties in	0	477	WorldCup,	Squawka F	Squawka	Cayleb	Accra	861	828
3	1.0136E+18	en	02-07-2018 1.35	Twitter for Android	139	RT @FCBarcelon	scores the winning penalty to send into the quarter	0	1031	WorldCup	FC Barcelo	FCBarcelo	Febri Adity	Bogor	667	686
4	1.0136E+18	en	02-07-2018 1.35	Twitter for Android	107	RT @javierfernar	Tonight we have big game	0	488	worldcup	Javier Fern	javierferna	??		65	67
5	1.0136E+18	en	02-07-2018 1.35	Twitter Web Client	142	We get stronger	We get stronger Turn the music up now We got that	0	0	PowerByE	EXO,FIFA V	weareone	Frida Carri	Zapopan, J	17	89
6	1.0136E+18	en	02-07-2018 1.35	Twitter for Android	140	RT @Squawka:	Only two goalkeepers have saved three penalties in	0	477	WorldCup,	Squawka F	Squawka	tar		137	216
7	1.0136E+18	en	02-07-2018 1.35	Twitter for Android	140	RT	We re looking strong going into the knockout stage	0	153	BRAMEX	FIFA Worl	FIFAWorld	Ligefut		29	283
8	1.0136E+18	en	02-07-2018 1.35	Twitter for iPhone	140	RT	am happy for winning Especially since you know we	0	4	Russia	V?Deplora	ShShShShS	?a?????	Mount Oly	208	338
9	1.0136E+18	en	02-07-2018 1.35	Twitter Web Client	138	RT	When you see me When we feel the same feeling Po	0	1	PowerByE	Frida Carri	FridaCarril	STAN LEGE	Lima, Peru	7	9
10	1.0136E+18	en	02-07-2018 1.35	Twitter for Android	138	RT	Kasper Schmeichel takes the final award of the day	0	2199	Manofthe	FIFA Worl	FIFAWorld	Sky Ler		1	6
11	1.0136E+18	en	02-07-2018 1.35	Twitter Lite	139	RT	After Years Global Puma Ambassador LG Mobile Am	0	5146		???? & ???	BTSARMY	Kate	Meme City	158	245
12	1.0136E+18	en	02-07-2018 1.35	Twitter for Android	86	RT	Idek what to do anymore Spain and Portugal both o	0	1	WorldCup	Ella Laxa-F	EllaPangili	Alice	Bulacan, C	34	315
13	1.0136E+18	en	02-07-2018 1.35	Twitter for Android	135	RT @SiClancy: I k	know it been an amazing World Cup but this is bit to	0	3817	worldcup	Simon Clar	SiClancy	Goku	Barcelona	105	51
14	1.0136E+18	en	02-07-2018 1.35	Twitter for iPhone	139	RT @valpan026:	by our happy virus will be perfect choice Happy alwa	0	107	POWER,EX	Shiqi,Eunic	valpan026	ELyXiOn		12	88
15	1.0136E+18	en	02-07-2018 1.35	Twitter for iPad	166	Get the best in sp	Get the best in sports news and information by goin	0	0	sportstalk,sportstalkradio,NBA,N			JSA		1978	4288
16	1.0136E+18	en	02-07-2018 1.35	Twitter for iPad	140	RT @IanStaffs: Ju	Just worked out that if England can get past Colomb	0	7		Ian Staffor	IanStaffs	SKS Media	Mumbai, I	2147	2403
17	1.0136E+18	en	02-07-2018 1.35	Twitter for iPhone	139	RT	People getting rightly excited over Mbappe after his	0	6889		CCFCFan	CCFCFan1	??	The nether	2167	956
18	1.0136E+18	en	02-07-2018 1.35	Twitter for Android	88	RT @dw_sports:	PENALTY SHOOTOUTS IN ONE DAY	0	17	CRODEN,V	DW Sports	dw_sports	miikii	Thailand	187	188
19	1.0136E+18	en	02-07-2018 1.35	Twitter Web Client	140	RT	Edinson Cavani for Napoli Games Goals Edinson Cav	0	70		Your Villag	Onyema_D	mattorazi		85	137
20	1.0136E+18	en	02-07-2018 1.35	Twitter for Android	140	RT	Japan have never made it past the last eight so we v	0	219		FIFA Worl	FIFAWorld	Michelle??	Orange Co	109	297
21	1.0136E+18	en	02-07-2018 1.35	Twitter for iPad	120	RT @IanStaffs: W	Wonder what the odds were on Messi Ronaldo and	0	4	WorldCup	Ian Staffor	IanStaffs	SKS Media	Mumbai, I	2147	2403
22	1.0136E+18	en	02-07-2018 1.35	Twitter for iPhone	125	RT @LeJuan__Ja	Lebron news is cool and all pero let gets back to wh	0	3	WorldCup	LeJuan Jan	LeJuan__J	DD Benite	The Ohio S	171	403
23	1.0136E+18	en	02-07-2018 1.35	Twitter for iPhone	140	RT @eunicehuili:	Highly anticipating by to be played at the stadium	0	105	POWER,EX	Eunice,Rac	eunicehuili	ELyXiOn		12	88
24	1.0136E+18	en	02-07-2018 1.35	Twitter for iPhone	318	How Do You	How Do You Control the Ball Lose How Do You Pass	0	0	Spain,Russia,Comrades,Soccer,W			Fire and Br	Toledo, OH	112	372
25	1.0136E+18	en	02-07-2018 1.35	Twitter Web Client	140	RT	Kasper Schmeichel vs Croatia Saves penalty in Extra	0	15		9GAG Foot	9GAGFoot	Jay	Toronto, C	4	65

CHAPTER 3

SYSTEM REQUIREMENT SPECIFICATION

3.1 Operating Environment

- **Hardware Platform:**The software will run on a standard desktop computer, laptop, and mobile device.
- **Operating System:** Works on such operating systems as Windows, Mac OS, and Linux.
- **Software Requirements:** The user interface of the software requires modern HTML5 and CSS3-enabled web browsers.

3.2 General Constraints, Assumptions and Dependencies

- **Legal Implications**: Compliance with the data protection policies and copyright legislations, and the terms of service provided by social media platforms, such as Twitter, is required to avoid legal issues. Obtaining the right permissions and licenses for using copyrighted material such as sports broadcasts becomes a legal necessity.
- **Usage Limitations**: Success of the project basically hinges on availability of sports information, whether it is tweets or video feeds.Over-dependence on third-party APIs or data sources could be a usage limitation of rate limits or access limits to data.

3.2.1 Assumptions Made in the Project:

- **Data Availability:** This project presumes that good quality and reliable sources of data are available on time for analysis and summarization purposes.

- **Consistency in Data Format:** It assumes that a proper level of consistency in the format and structure of data sources is made available to perform the necessary processing.

3.3 Risks

- **Resource Constraints**: In terms of performance and scalability, the limited computational resources can affect system performance.
- **Data Availability**: Since the data sources are external to the system, it means issues regarding the reliability and accessibility of the data are at stake.
- **Technical Issues**: Implementing some of the most advanced techniques may suffer from technical complexities and other constraints.
- **Regulatory Compliance**: Not meeting legal regulations can expose the organization to legal risks and consequences.

3.4 Functional Requirements

- **Automated Data Analysis:** The system automatically identifies significant event in sports data by event extraction during high twitter activity and it cross verifying with commentator excitement,audiences cheers and incorporates real time scoreboard data.
- **Efficient Data Processing:** Efficiently processes large volumes of data in real-time,dynamically adjusting capabilities to handle increased twitter data during high activity.
- **Noise Handling:** Noise was well filtered out and key moments were rightly detected.

- **Highlight Selection**: It automatically selects and compiles the most interesting or important moment from the game into a highlight reel.
- **Key Moment Detection Accuracy:** The system should try to be as accurate as possible to detect key moments within the video footage.
- **Processing Speed:** The system should be more efficient than manual editing methods and generate summaries much faster than those done manually.

3.5 Software Requirements

3.5.1 TensorFlow:

- **Name and Description:** TensorFlow is a free of charge machine learning framework is open-source and developed by Google for crafting and training deep learning models.
- **Version / Release Number:** version 2.5.0.
- **Operating Systems**: Compatible with Windows and Mac OS.
- **Tools and Libraries:** TensorFlow is packed with the numerous toolkits and libraries that cover the whole depth of the deep learning tasks, including TensorFlow Keras for developing neural networks.
- **Source:** TensorFlow has GitHub source code to allow anyone to easily pick up where they are, as well as contribute and check their own issues.

3.5.2 OpenCV:

- **Name and Description:** OpenCV (The Open Source Computer Vision Library) is a suite of functions for performing real-time computer vision tasks.
- **Version / Release Number:** version 4.5.3 OpenCV
- **Operating Systems:** The software works well on all four operating systems including Windows, macOS, Linux, Android, and iOS

- **Tools and Libraries:**
- OpenCV has a suite of complex functionalities for image processing at its disposal feature identification, object detection, and video analytics.
- **Source:** OpenCV is open-source and the source code is on GitHub so developers can access it and use it for getting the documentation, source code, and community contributions.

3.5.3 MySQL Database:

- **Name and Description:** This is a free, query-driven and highly scaling RDBMS which most professionals use for structuring and storage of the information.
- **Version / Release Number:** MySQL 8.0.
- **Operating Systems:** Available for Windows OS, macOS, and Linux.
- **Tools and Libraries:** MySQL arms the user with a powerful toolbox for database management, which caters for various front end requirements such as data indexing, querying plus administration.
- **Source:** MySQL does not carry any costs and is open-source; one may download it from its official site or through the relevant repositories for different operating systems.

3.5.4 OpenAI Whisper:

- **Name and Description:**Whisper is an open-source Natural Language Processing Framework for Sentiment Analysis and Text Classification.
- **Operating systems**: It works on Windows, macOS, and Linux.
- **Usage:** It is based on various tools primarily for Sentiment Analysis, Text Classification, and Language Understanding.
- **Source:** Whisper holds open-source code on GitHub to check the documentation, source code, and to submit contributions in the community.

3.5.5 BERT:

- **Name and Description:** BERT is an acronym for Bidirectional Encoder Representations from Transformers. BERT is the pre-trained Natural Language Processing model from Google, which can be used for an understanding of human text as well as to generate human-like text.
- **Version / Release Number:**This model has been deployed with BERT v2.0.
- **Operating Systems:**It runs smoothly on Windows, macOS, Linux, Android, and iOS platforms.
- **Tools and Libraries:**It uses several state-of-the-art natural language understanding tools like Text Classification, Named Entity Recognition, and Question Answering.
- **Source:** BERT is also an open-source model, the source code is available on GitHub with access to documentation, source code, and the BERT community.

3.5.6 Python:

- **Name and Description:** Python is an easy-to-read and write language used in various applications such as web development, data analysis, and artificial intelligence, and is used for various purposes.
- **Version / Release Number:** This version is Python 3.9.2
- **Operating Systems :**It is compatible with Windows, macOS, Linux, and other Unix-like systems.
- **Tools and Libraries:**This language has a significant ecosystem of libraries and frameworks—like NumPy for numerical computing, Pandas for data manipulation, and Flask for web development.

- **Source:** Python is a free, cross-platform language. The source code is available on GitHub to access the documentation, source code, and contribute to the Python community.

3.6 Communication Interfaces

- With introduction of relating local network protocols (TCP/IP) by a server side component and a client side component, data connection will be established.
- Establishing the internal communications approach and way of communication as a must requirement will be one of the top implementations for realizing the system compatibility and interoperability.
- Buffer sizes and sampling rates will be setting obstacles related with setting data speed in bps and size of data will be minimized during analog to digital processing and digital to analog processing in this way network capacity will be improved up to the highest level.

3.7 Non-Functional Requirement

3.7.1 Performance Requirement

3.7.1.1 Processing Speed

- **Objective:** The system would need to process the sports videos and make highlights as it is needed for real-time or near real-time usage.
- **Metric**: A video having processing time over 1 minute should be excluded. The handling time should be less than a minute in a good work load.

3.7.1.2 Scalability:

- **Objective:** The system should be scalable when more users use it during different sporting events and it requires no break down in the performance.
- **Metric**: The redundancy system should be designed to achieve handling of 1000 simultaneous video processing requests with no notable discernible degradation in

performance. Scalability tests are to be done time and again for ascertaining that enough space is available with the system to handle the incoming traffic

3.7.1.3 Reliability:

- **Objective:** Another important goal is the system must be able to precisely and signify major events in a sport video.
- **Metric**: The detector needs to have a detection accuracy of over 90 percent for key games with the use of machine learning algorithms. Besides there is the service called Mean Time Between Failures (MTBF) which should be at least 500 hours.

3.7.1.4 Robustness:

- **Objective:** The system should have the ability to deal with the underlying distortions such as dissimilarity in video quality, noise, and other data patterns which may be unpredictable.
- **Metric:** The system has to keep up a constant level of performance no matter when it is the video of low resolution (e.g. SD, HD), high resolution or different commentary styles. Robustness testing to cover situations with high background noise and various audio characteristics should be included considering the broad spectrum of mobile environments that are available today.

3.8 Safety Requirements

3.8.1 Data Integrity:

- **Objective:** Identifying the accurate integrity measures that imply that data is reliable and authentic.
- **Metric:** The trust data proof protocol can expand the verification through utilization of hash functions and public key signatures as well.

- We advocate data consistency check and data integrity assessment on a time-delimited basis. It provides for data integrity in the sense that inception of the process and current time frame (of life cycle) have the same value.

3.8.2 User Safety:

- **Objective:** Come up with privacy policy as well as personal data protection mechanisms that will merge together with the product because it is a tool to be used by ordinary people.
- **Metric:** Encryption is one of the most efficient approaches to anonymize user-related data. Furthermore, the application of a dedicated storage media for data sensitive information concept is required and other implicitly used security measures to reach the desired level of security are deemed necessary.

3.8.3 Data backup and recovery:

- **Objective**: Data backup through use of data backup and recovery methods, employing data redundancies and regular service plan that entails speedups, problems solvings and elimination of data loss.
- **Metric**: For a start, you should be making it a point to do backup of data and metadata as well, and it must be done every day. Consumers need to be equipped with appropriate strategies, such as a deprivation schedule as an alternative. Hence, we are going to try it out whatever be the case after both failures and data loss which can be there. Remind the user again and again about how the process should be executed sometimes.

3.9 Security Requirements

3.9.1 Data Privacy:

- **Objective:** Protect a user's data confidentiality, in particular, any comments and tweets, and the wearables' identification information.

- **Metric:** Ascertain the application of encryptions for data encryption over the networks and for storage to resist the risk of unauthorized access. Done privacy compliance review and cross-compliance check on against regulation enforcement periodically. The system aiming to meet the privacy compliance rate of 98%, can be assessed at the end of the deployment year.

3.9.2 Access Control:

- **Objective:** The purpose of access controls is to guarantee that only authorized users can perform particular functions, for example, confidential data access is needed in the process.
- **Metric:** role -based access control (RBAC) for user roles in order to limit access to the stipulated roles. A workable grant principle is to regularly revise and update user access permissions. Keep an eye on and record who got to the private data, and, in case of the unwanted access, investigate.

3.9.3 Secure Communication:

- **Objective:** Implementation of best practices relating to security scenarios for third-party APIs, such as Twitter, customer security should be very good. API keys and tokens should be kept secret and changed periodically in order to ensure safety.
- **Metric:** There should be an expiry date for the API keys and tokens and they should be changed, updated regularly, Look to see whether communication with third-party service APIs is performed with security protocols that are capable. Monitor the usage and access to the API to trace down illegal actions and respond without delay.

CHAPTER 6

SYSTEM DESIGN

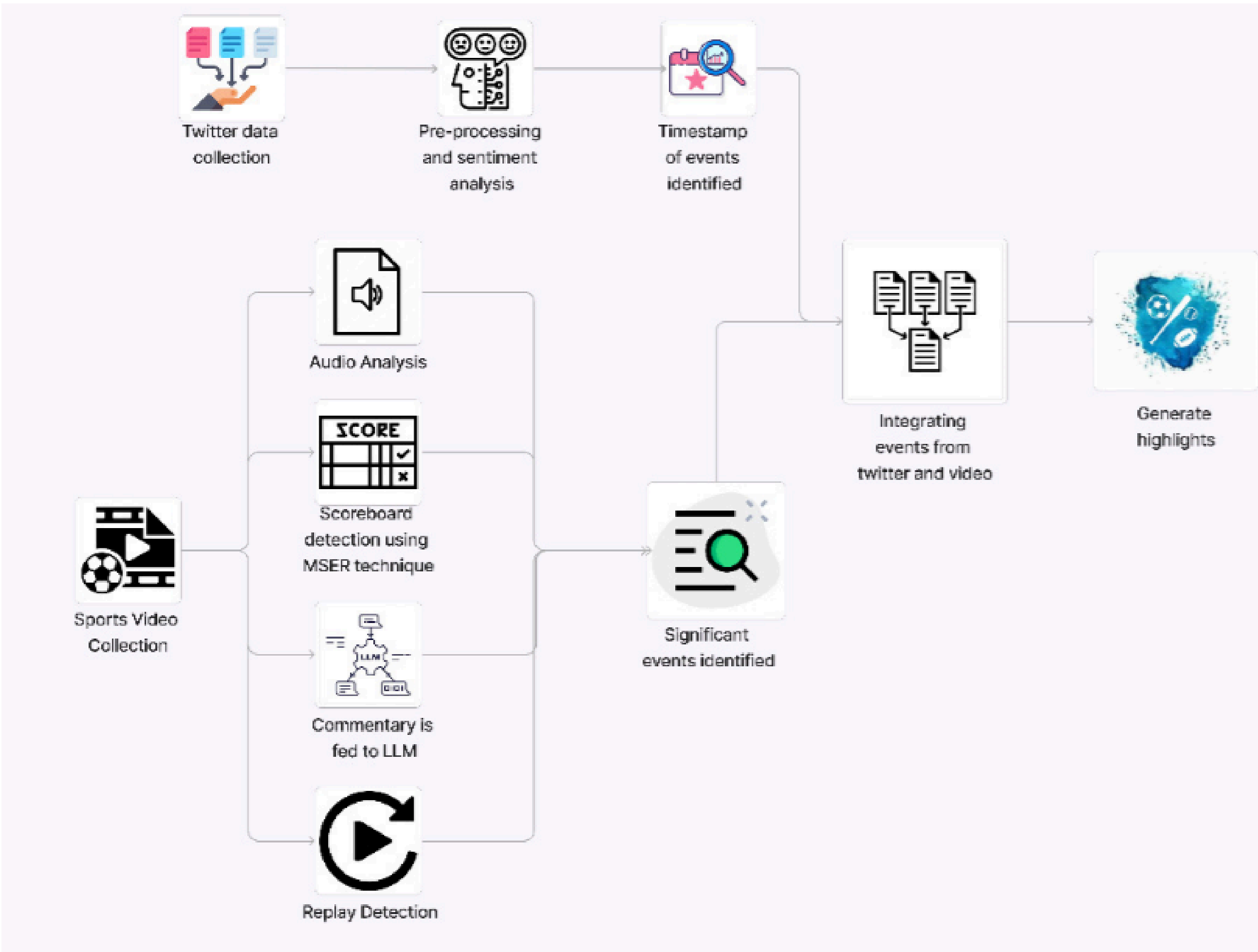


Fig 6.1 High Level Architecture design diagram

The figure 6.1, shows the high level architecture diagram for our project Indexing and Summarization of Sports Videos using Multi-Modal Approach

Twitter Data Collection: In this step we gather data related to sports events from twitter with the help of hashtags and keywords.

Preprocessing and sentiment analysis: After the twitter data is collected, we standardize it by preprocessing, and also removing unnecessary data. Then we also perform sentiment analysis to understand the tweeter's sentiment towards the event.

Timestamp of events identified: This component captures timestamps associated with significant events identified in the twitter data. It helps us index the events identified in twitter data with the video footage . This is where we will be incorporating the time lag model to ensure accurate event indexing.

Sports Video Collection: This component , collecting video footage of sports events.

Audio Analysis: This component helps us identify significant events in the video footage by analyzing the crowd noise by analyzing the changes in the audio intensity.

Scoreboard detection : This component focuses on identifying and analyzing the scoreboard within the video frames using MSER as it needs less computational power compared to OCR. Once identified, the content of the scoreboard is extracted and analyzed for score updates.

Commentary fed to LLM: In this component, commentary data that is fetched from video footage is fed into a large language model to identify significant events within the sports video.

Replay Detection: replays within the video footage, are detected by identifying scene changes, indicating the replay of specific significant events.

Significant events identified: After the video analysis done to identify significant events from various models we aggregate all of the significant events identified.

Integrating events from twitter and video : The significant events identified from both video and twitter data are combined together to facilitate the final process.

Generation of highlights: Based on the identified significant events, this component generates highlights clips from the video footage.

Conceptual or Logical View (UML Component Diagram):

- **Logical User Groups:**

- System Administrators
- Analysts
- Commentators/Broadcasters

- **Application Components:**
  - Twitter and Video Data Collection Component
  - Event Detection Component: Includes sentiment analysis, WDHG event detection, scene detection, scoreboard analysis, and audio analysis.
  - Event Linking and Analysis Component :Matches events detected in the video footage with the ones in twitter data, generates highlights, and integrates with LLM for commentary.
  - Highlights UI
- **Data Components:**
  - Social Media Data: Tweets collected for analysis.
  - Video and Audio Data: Game footage and audio streams processed for event detection.
  - Event Data: Structured data representing detected events and associated metadata.
- Interfacing Systems:
  - External APIs
  - Large Language Model
- **Process - Runtime View :**
  - Sequence of Interactions: User interacts with the User Interface Component.
  - Data flows through Data Collection, Event Detection, and Event Linking Components.
  - Results are presented back to the User Interface for visualization.
- **Physical - Distributed System View :**

- Server Node hosting application components.
- Client Devices accessing the User Interface Component.
- External Services (Twitter, LLM Service) integrated via network interfaces.

- **Security - System Security Features:**
  - Twitter users data encryption
  - Access Control: Assign roles and a set of permissions to various user groups.
  - Secure communication with external APIs using authentication and authorization mechanisms.
  - Data Privacy: Compliance with data privacy regulations and best practices.

MASTER CLASS DIAGRAM

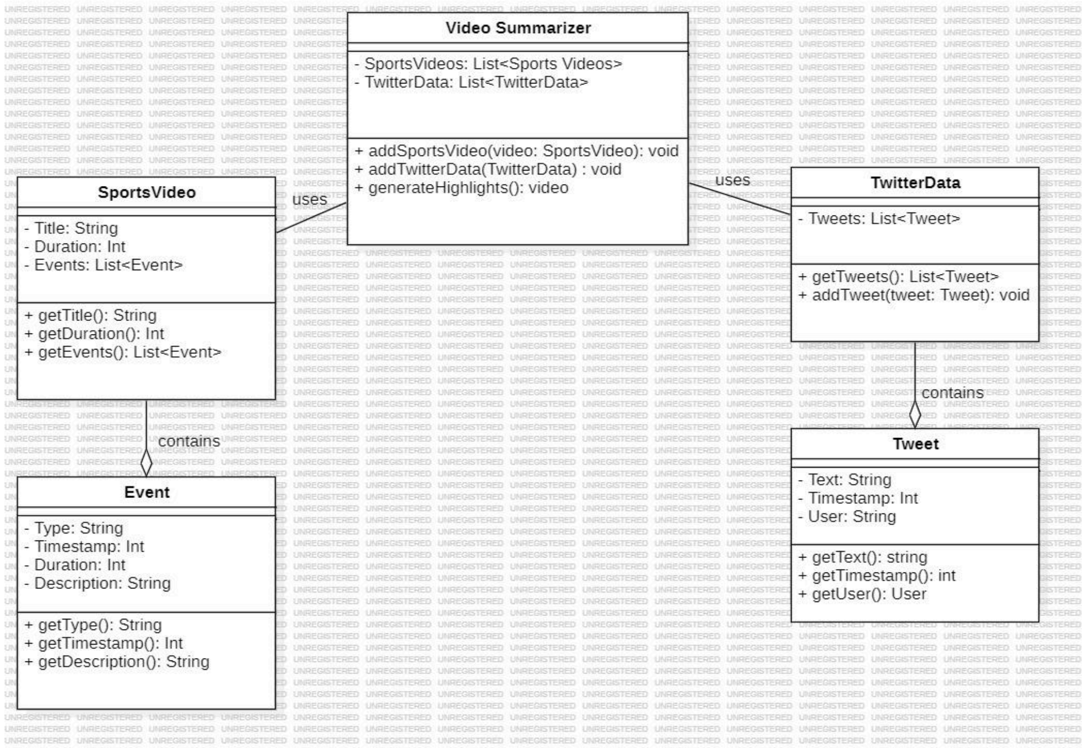


Fig 6.2 Master class diagram

According Fig 6.2 there are 5 classes:

1. **The Video Summarizer:**
   - This is the main class of our system and it uses the sports videos and twitter data
   - Both of these components are used by this class in order to generate summaries.
2. **Sports Video:**
   - This class contains a list of sports videos that is provided to the system in order to generate summaries.
3. **Twitter Data:**

- This class contains the tweets associated with each video and it is provided to the video summarizer to generate summaries that incorporate the user's reactions and insights.

4. **Events:**
   - This class contains various events that will be used by the system to flag as an important event or not

5. **Tweets:**
   - This class contains the individual tweets that comprise the twitter data.

USE CASE DIAGRAM

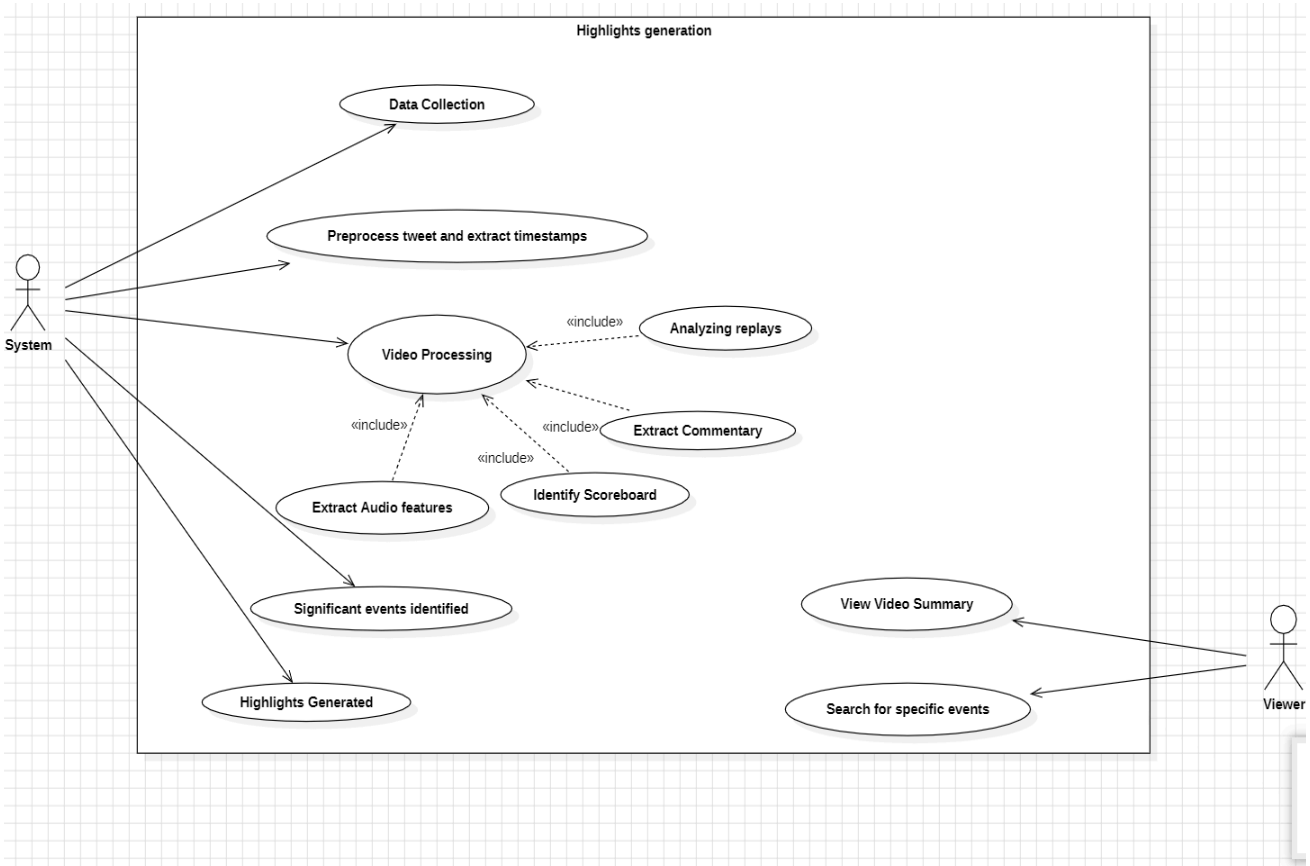


Fig 6.3 Use case diagram

In the figure there are a few actors and use cases as below:

Actors:

- **System**: It interacts with the twitter and video data, performs analysis and identifies significant events which will be aggregated to provide the video summarization.
- **Viewer**: The viewer is the user, analyst or broadcaster/commentators who will make use of the system to view highlights.

Use cases:

- **Data Collection :** This use case represents the process of gathering data of video footage and twitter from their respective sources.
- **Video Processing :** This use case signifies the analysis of sports video. There are sub-processes associated with this use case.
- **Extract Commentary :** The system extracts audio commentary from the video to feed it to the LLM to identify important moments based on the commentator's emphasis.
- **Identify Scoreboard :** The system identifies and extracts information from the scoreboard regarding the score updates.
- **Extract Audio Features :** This use case which is the sub process analyzes crowd noise to identify exciting moments from the game.
- **Analyzing replays:** This use case refers to the systems ability to analyze replays within video footage.
- **View Video Summary :** This use case of the Viewer allows them to view the generated highlights video.
- **Search for specific events:** The viewer can search for specific events in the footage as we extract the timestamp associated with each significant event.

Deployment Diagram

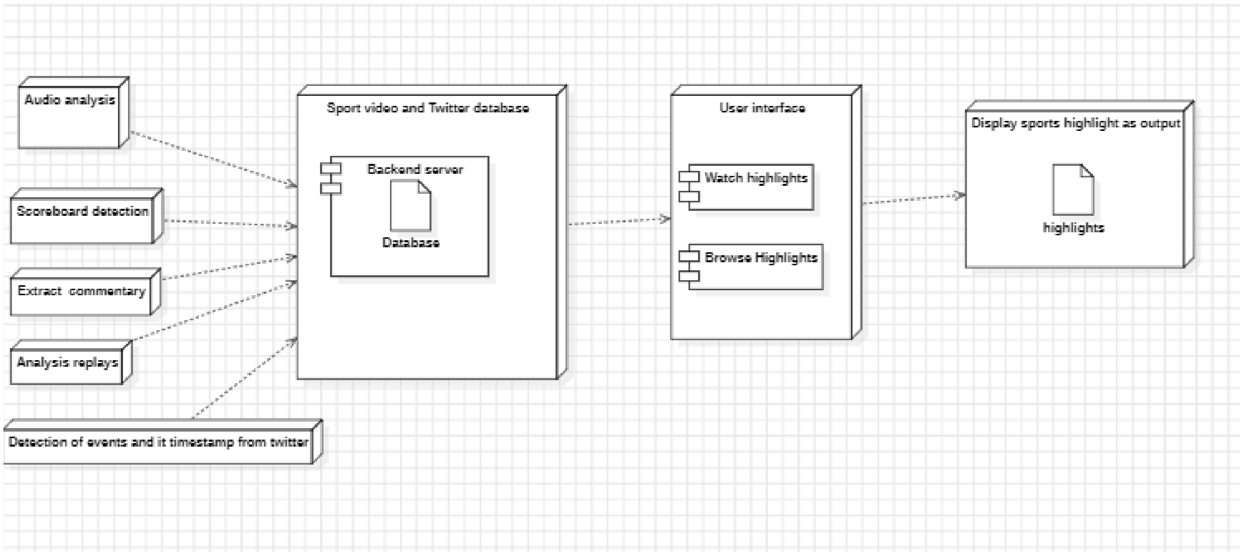


Fig 6.4 Deployment Diagram

- Backend server- this is main program that runs the entire system, it houses the analysis software that analyzes the commentary and identifies the events
- Database: This stores the sports videos and Twitter data.
- Audio analysis: This module extracts the commentary from the sports videos.
- Scoreboard detection: This module finds the scoreboard in the video footage.
- Extracting tweets and its timestamp from twitter: Extract the tweets and its timestamp and process it to identify significant events.
- User interface: This is the main user interface that the user can use to interact with the system.

Swimlane Diagram

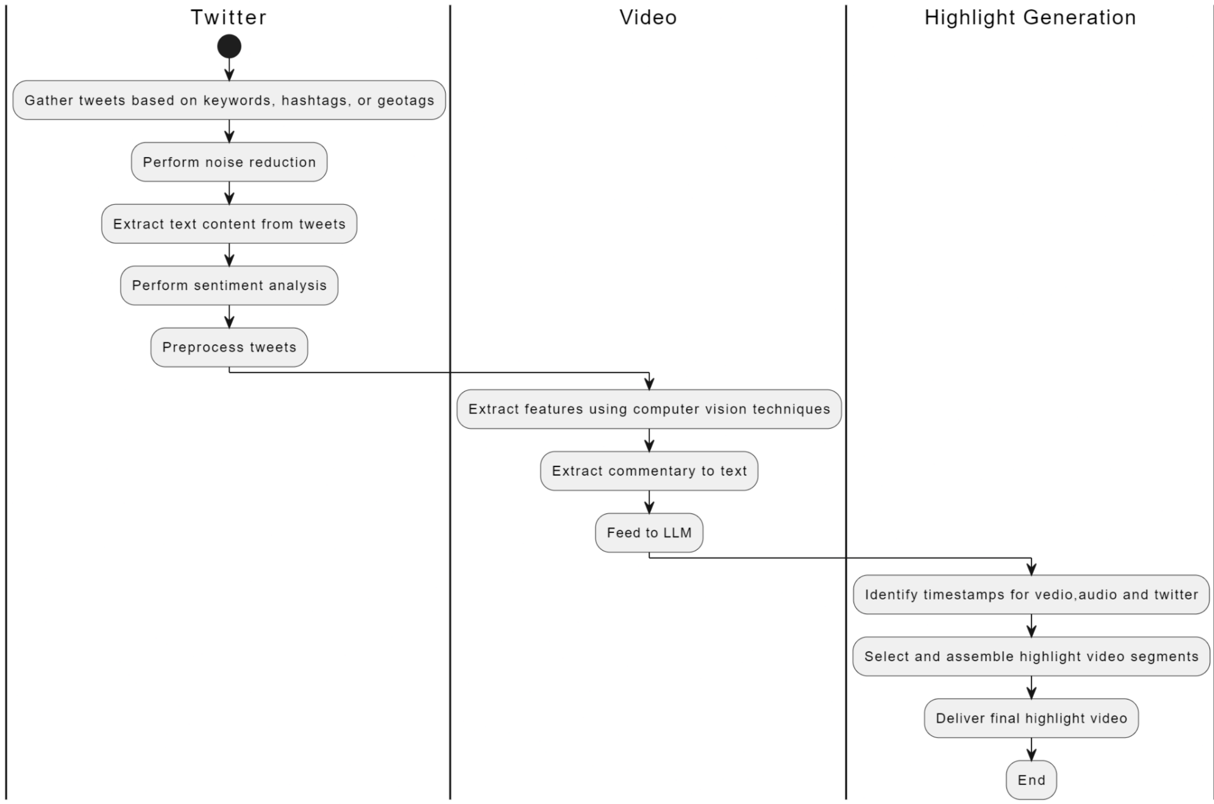


Fig 6.5 Swimlane diagram

CHAPTER 7

IMPLEMENTATION/ PSEUDOCODE

Pseudocode for script to generate commentary from video file using OpenAI's whisper model

```
\\import necessary libraries
import whisper
import os
import wave
import moviepy.editor as mp

video_file = "Brazil_vs_Belgium.mp4"

# Convert video to audio
audio_file = "Brazil_vs_Belgium.wav"
video_clip = mp.VideoFileClip(video_file)
video_clip.audio.write_audiofile(audio_file)

# Split audio into  chunks
def split_audio(input_audio, output_dir, chunk_duration=30):
    # Open the input audio file
    with wave.open(input_audio, 'rb') as wf:
```

```
# Get the audio file properties
framerate = wf.getframerate()
nframes = wf.getnframes()
duration = nframes / float(framerate)

# Define time interval for splitting audio (in seconds)
interval = chunk_duration

# Initialize variables for timestamp tracking
current_time = 8
chunk_number = 1

# Split audio into chunks and save each chunk as a separate file
while current_time < duration:
  # Read audio chunk
  start_frame = int(current_time * framerate)
  end_frame = min(int((current_time + interval) * framerate), nframes)
  wf.setpos(start_frame)
  audio_chunk = wf.readframes(end_frame - start_frame)

  # Save audio chunk as a separate WAV file
  output_file = os.path.join(output_dir, f"chunk_{chunk_number}.wav")
  with wave.open(output_file, 'wb') as wf_out:
          wf_out.setnchannels(wf.getnchannels())
          wf_out.setsampwidth(wf.getsampwidth())
```

```
                wf_out.setframerate(wf.getframerate())
                wf_out.writeframes(audio_chunk)

        # Update current time and chunk number
        current_time += interval
        chunk_number += 1

output_dir = "audio_chunks"

# Split the audio into 60-second chunks
split_audio(audio_file, output_dir, chunk_duration=60)

# Load the base model
model = whisper.load_model("base")

# Transcribe each audio chunk and write the results to a CSV file
with open("commentary.csv", "w") as file:
    file.write("Minute,Commentary\n")
    for i in range(1, 111):
    result = model.transcribe(f"{output_dir}/chunk_{i}.wav")
    file.write(f"{i},{result['text'].replace(',', '')}\n")
```

PES UNIVERSITY

Indexing and Summarization of Sports Videos using Multi-Modal Approach

Explanation of the pseudocode

This is the pseudocode to generate the commentary data when a video is given as input. This script will be used to get the commentary data that has to be fed to the large language model.

We first import the required libraries and get the audio from the video file using the moviepy library. Then we split the audio into 60 seconds to enable quicker processing by the Whisper model, this is done by the wave library in python. We then save the chunks locally on our system.

We then load the base Whisper model and feed the audio file one by one and store the output in a CSV file along with the minute of the commentary.

Output of the Script on Brazil vs Germany 2018 world cup match

Minute	Commentary
1	and Liverpool. Looks like Belgium won the toss. He's an asa made captain by Roberto Martinez. Danieli Osato of Italy leading the video assistant referee team there. All kitted out as usual. Belgium by the way will stop with a team which finished
2	the game. Since winning the 2002 final against Germany Brazil have lost all four games in the knockout phase when they face European opposition which is a bit of a surprising one really but there it is. Again quickly underway and Marcelo back fr
3	it. Cheechy very impressive at his press conference yesterday. Spoke a lot of common sense once things to be cool. I think four years ago in that traumatic 7-1 defeat in the semi-final against Germany Brazil rather drowned in their own emotion.
4	that's not a good start for Fan and Junior only him and Marcelo survivors of that seven one game and that's not a good start for Fan and Junior and what is his first start of this world cup he's been coming off the bench till now fact that unity play
5	the pitch is good as wide as they can. And Belgium took the view looking at the Mexico going that they did create things that name are to get this one into woods. Gabrielle Jesus but too close to up Krot wild that time. You can put the flick on and
6	the first time in a knockout game where a team had been two nil down in the second half and gone on to win in 90 minutes. Bit luckless Japan they played wonderfully. And then the first time in a knockout game where a team had been two nil do
7	two earlier points in the proceedings. Jesus is on his stop this as a real battle between the two numbers. Today. And as I name. I'm for lining and he picks up. De Bruyne. I have caught my spits. I was marnier with the head of down towards. Luke
8	and he's getting a word of warning here from the Serbian referee. Putting himself about a bit bit free with the elbow there and the referee's seen that he's three yards away. Already he knows that the referees on his case and watching. So I think
9	the left hand side because Marcelo as you will have noticed likes to play almost like a left winger himself when he can. So I've got him and name on Coutinho at times posing a threat down that side. They might have their work cut out. Those Bel
10	the way this is some start. Here's the Chadly effort swinging wide. The other end Flicton off Chargo Silver and against the Pose. He can't believe it. Look at that face. Tells the story. Somebody up there I think liked Belgium then. Of course Chargo
11	the game. Fagna. It's interesting. Lukaku said it. His press conference yesterday. Brazil have got three out of four very experienced defenders. We do respect them. So you're kind of thinking well will they target the other one? That's Fagna. We'll
12	the ball. But already some alarms are the back for Belgium. His william about to take this corner kick towards the near post and again Brazil mess up the chance. Palinho. It felt to him and he miskecked it. Now Namer. But that Belgian defense lo
13	the ball. Well it's breathless stuff early on here. And what I think most people would agree on neutrals anyway on paper. This is the most attractive looking of the quarter finals jesus going for the heart of the Belgian defense. Away by a bettonge
14	the player. But a lively display as a substitute including a winning goal. There's one in a start here. This company is leadership qualities vital in this Belgian squad. Of course he's not the captain anymore. That's a lovely ball to Brian to for Linde. I c
15	the league. They've got nine six footers in this team. Brazil have only got three. Flick down the knee post and they have the lead. That's all company. And Brazil can see first. Extraordinary. The near post had a dead end. And it's just taken a little f
16	and it's a gift really yet another own goal. It's breaking records for own goals this World Cup. He never altogether. Just has a chance to win. Chargo Silva. They will always believe that they are going to score. And of course they nearly always do.
17	still it won't go in. Well there have been some hair raising escapes back there. Gabrielle Jesus this time who couldn't just stick out of foot. Leading a bit of a charmed life Neymar looking to dance through them all. Marcelo charged down by Marij
18	the goal. Don't think he got any kind of touch. And anyway the decisive one was off for N'Ginho in the end. The problem for Belgium is Brazil have a long time to do something about this. The one thing he won't do is panic. Chichu who a few year
19	the other 31 nations at this well-cut that preparation was that thorough. They actually lent out the work to various clubs around Brazil. So they're all in on this. All the clubs in Brazil as well. It's not just the national team but there's a bit of trouble
20	the ball. He's not going to get anything there. And Strenker was top scorer in the qualifiers for Brazil in the end. They skated it. One of the South American group that tough South American group by ten points. Was that a little handoff by compan
21	regular. Let them move the whole idea what you've been doing to those of these teams on Harp and get plenty of half of course we've got quality. . Fin. Genie. And flan and Genie up. the ball. He's a big one for a hundred and forty one million po
22	the ball. I've got over 60 caps today. Lot of experience so they shouldn't freeze. Lukaku. It's continue again. Pulling all the strings from that midfield position this time it's easy for Yann Vatongan. Over 100 caps for their Tongan in this experience l
23	the game. Good crossing and again this time Miranda stops it getting to Luke Haku. You'd be amazed if we've seen the last goal in this game. You can remind me about that later if they turn out to be famous last words. I don't know. I've got an e
24	the ball so far. I'm not sure if it's going to be good. My cello. I'm going to play it wide now. William is in rather less of the ball so far. Another useful one but here's De Bruyne out and he's got a lot of space just like Mexico did in the ball. Belgian m
25	the last time. Did Brian had to take it. And he calls problems from yet another corner. Not this time. It's a rare misshead from the master craftsman from Manchester City. The master craftsman from Manchester City. Debrainer has the most assi

commentary

REFERENCES/BIBLIOGRAPHY

[1] Hirasawa K, Maeda K, Ogawa T, Haseyama M. "Detection of Important Scenes in Baseball Video Videos via a Time-Lag-Aware Multimodal Variational Autoencoder." Sensors (Basel). 2021 Mar 14;21(6):2045. doi: 10.3390/s21062045. PMID: 33799412; PMCID: PMC7999231.

[2] Raj, R., Bhatnagar, V., Singh, A. K., Mane, S., & Walde, N. (2021, January 21). Video Summarization: Study of various techniques. arXiv.org. https://arxiv.org/abs/2101.08434

[3] H. Sattar, M. S. Umar, E. Ijaz and M. U. Arshad, "Multi-Modal Architecture for Cricket Highlights Generation: Using Computer Vision and Large Language Model," 2023 17th International Conference on Open Source Systems and Technologies (ICOSST),2023, pp. 1-6, doi: 10.1109/ICOSST60641.2023.10414235

[4] Z. Saeed, R. Ayaz Abbasi, M. I. Razzak and G. Xu, "Event Detection in Twitter Stream Using Weighted Dynamic Heartbeat Graph Approach [Application Notes]," in IEEE Computational Intelligence Magazine, vol. 14, no. 3, pp. 29-38, Aug. 2019, doi: 10.1109/MCI.2019.2919395

[5] R. S. Bhat, J. O, P. P. P, P. Kumar Vedurumudi and D. K. N, "Cricket Video Summarization Using Deep Learning," 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), Lonavla, India, 2023, pp. 1-6, doi: 10.1109/I2CT57861.2023.10126359

[6]Video Summarization Study Of Various Techniques Proceedings of IRAJ International Conference,26th May, 2019

[7]C. Yan, X. Li and G. Li, "A New Action Recognition Framework for Video Highlights Summarization in Sporting Events," 2021 16th International Conference on Computer Science & Education (ICCSE), Lancaster, United Kingdom, 2021, pp. 653-666, doi: 10.1109/ICCSE51940.2021.9569708.

[8]A. Bhalla, A. Ahuja, P. Pant and A. Mittal, "A Multimodal Approach for Automatic Cricket Video Summarization," 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2019, pp. 146-150, doi: 10.1109/SPIN.2019.8711625