| **PROJECT REQUIREMENTS SPECIFICATION**  **Indexing and Summarization of Sports**  **Videos using Multi-Modal Approach**  **UE21CS390A – Project Phase – 1**  ***Submitted by:***   | Meenal Bagare  Melvin Jojee Joseph  Naveen Reddy G  Krupashree MV | PES2UG21CS289  PES2UG21CS294  PES2UG21CS324  PES2UG21CS242 | | --- | --- |   Under the guidance of   | **Dr. Sandesh B.J**  Chairperson & Professor  PES University | | --- |   **January - May 2024**  **DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**  FACULTY OF ENGINEERING  **PES UNIVERSITY**  (Established under Karnataka Act No. 16 of 2013)  Electronic City, Hosur Road, Bengaluru – 560 100, Karnataka, India |
| --- | --- | --- | --- |

**TABLE OF CONTENTS**

| 1. Introduction | 3 |
| --- | --- |
| 1.1 Project Scope and Motivation | 3 |
| 1. Literature Survey or Existing System | 4 |
| 1. Product Perspective | 10 |
| 3.1 Product Features | 10 |
| 3.2 User Classes and Characteristics | 10 |
| 3.3 Operating Environment | 10 |
| 3.4 General Constraints, Assumptions and Dependencies | 10 |
| 3.5 Risks | 11 |
| 1. Functional Requirements | 11 |
| 1. External Interface Requirements | 12 |
| 5.1 User Interfaces | 12 |
| 5.2 Hardware Requirements | 12 |
| 5.3 Software Requirements | 12 |
| 5.4 Communication Interfaces | 13 |
| 1. Non-Functional Requirements | 14 |
| 6.1 Performance Requirements | 14 |
| 6.2 Safety Requirements | 15 |
| 6.3 Security Requirements | 15 |
| 1. Other Requirements | 16 |
| Appendix A: Definitions, Acronyms and Abbreviations | 16 |
| Appendix B: References | 17 |

# **Introduction**

This document outlines the requirements for the development of a sports video summarization , leveraging multi-modal data analysis techniques. The system aims to enrich the experience and comprehension of sports content by extracting key events from videos through the analysis of Twitter streams, Video Analysis , and Audio Analysis.

# **Project Scop**e

**Project Description:**

The project aims to develop a system for summarizing and indexing sports videos by leveraging multi-modal data analysis techniques, including analyzing peak activity in Twitter streams, Computer Vision, and Audio Analysis. By extracting key events from sports videos and generating concise summaries, the system aims to enhance the accessibility, enjoyment, and understanding of sports content for enthusiasts, analysts, and broadcasters.

**Purpose:**

The purpose of the project is to address the challenges associated with navigating and consuming lengthy sports videos by providing users with an efficient and informative summarization tool. By automating the process of identifying and highlighting key moments, the system aims to streamline the viewing experience and facilitate content discovery.

**Benefits:**

* Enhances accessibility and usability of sports content.
* Saves time for users by providing concise summaries.
* Enables efficient content navigation and discovery.
* Facilitates deeper analysis and understanding of sports events.
* Improves user engagement and satisfaction.

**Objectives and Goals:**

1. Develop algorithms for multi-modal data analysis, including analyzing peak activity in Twitter streams, Computer Vision, and Audio Analysis.
2. Implement a user-friendly interface for interacting with the summarization system.
3. Extract key events from sports videos based on multi-modal insights.
4. Generate concise and informative video summaries highlighting key moments.
5. Evaluate the system's performance using relevant metrics and user feedback.
6. Ensure scalability and adaptability to different sports and events.

**Coverage:**

* The system covers a wide range of sports and events, including but not limited to football, cricket and basketball.
* It analyzes multi-modal data sources, including peak activity in Twitter streams, audio commentary, and visual cues from sports videos.

**Limitations:**

* The system may be limited in its ability to generalize to different sports due the various rules and diverse conditions of each sport.
* The system may encounter challenges in accurately identifying key events, especially in complex or fast-paced sports scenarios.
* Performance may vary depending on the quality and availability of data sources such as Twitter streams and audio commentary.
* The system's summarization capabilities may be limited by the complexity and variability of sports events.
* It may require manual intervention or validation in certain cases to ensure the accuracy of generated summaries.

# **Literature Survey or Existing System**

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| K Hirasawa, K.; Maeda, K.; Ogawa, T.; Haseyama, M. Detection of Important Scenes in Baseball Videos via a Time-Lag-Aware Multimodal Variational Autoencoder. *Sensors* 2021, *21*, 2045. https://doi.org/10.3390/s21062045 | Multimodal Variational Autoencoder (Tl-MVAE)" for efficient detection of crucial scenes in baseball videos. Addressing time-lags between tweets and events, using a Poisson distribution.  Encoder captures relationships , decoder for reconstruction, and significant events detector using fully connected layers. | The paper addresses the challenge of time-lags between tweets and events by introducing a Poisson distribution, enhancing the model's ability to capture temporal dependencies.  The use of MVAE allows for flexible expression of relationships between heterogeneous modalities, contributing to high-quality scene detection. | While the paper justifies the use of a Poisson distribution for modeling time-lags, the effectiveness of this assumption may vary in different scenarios or datasets.  Training and inference with such complex models may be resource-demanding, potentially limiting real-time applications or deployment on resource-constrained devices. |

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| 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 | The objective of the paper is to develop an innovative approach for summarizing football match videos using advanced technologies, with the ultimate goal of improving user accessibility and experience in consuming video content. | Innovative Approach: The paper introduces a novel approach to video summarization, leveraging deep neural networks and semantic mapping techniques. This innovative method offers a fresh perspective on how to efficiently summarize videos. | Data Dependency: The effectiveness of the approach may heavily depend on the availability and quality of training data. If the training dataset is limited or biased, it could result in suboptimal performance or biased summarizations |

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| A. Javed, K. B. Bajwa, H. Malik and A. Irtaza, "An Efficient Framework for Automatic Highlights Generation from Sports Videos," in *IEEE Signal Processing Letters*, vol. 23, no. 7, pp. 954-958, July 2016, doi: 10.1109/LSP.2016.2573042. | The objective of the paper you provided is to propose a method for automatic highlight generation from sports videos through the detection of replay segments (RSs). The paper aims to exploit two observations for replay detection | The proposed system is evaluated on a diverse dataset comprising videos from different sports categories and broadcasters, indicating its potential applicability across various sports and broadcasting styles. | OCR Accuracy: The effectiveness of SC detection using OCR depends on the accuracy of the OCR algorithm and the quality of the input images. Errors in OCR recognition could lead to mislabeling of frames, impacting the overall performance of the system. |

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| 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 | Machine learning techniques, specifically YOLO v3 and OpenPose, to recognize players' actions in sports videos and efficiently generate high-accuracy framework highlights by removing match-irrelevant frames.  Three-Level Prediction Algorithm  YOLO v3 and OpenPose  Voting Scheme | Achieves high precision, recall, and combined metrics, indicating a significant improvement over previous systems.  Demonstrates potential applications in various sports beyond the commonly addressed ones like football, basketball, and baseball despite being a framework for racquet turn-based sporting events. | The paper acknowledges that the voting scheme, particularly for the OpenPose-based system, may not always be reliable, suggesting room for improvement.  Despite the high accuracy, the paper mentions the possibility of errors due to misdetected rallies, which may require manual correction |

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| M. Sanabria, F. Precioso and T. Menguy, "Profiling Actions for Sport Video Summarization: An attention signal analysis," *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, Tampere, Finland, 2020, pp. 1-6, doi: 10.1109/MMSP48831.2020.9287062. | The paper utilizes LSTM with an attention mechanism to automatically model action profiles from soccer video events. The goal is to enhance the efficiency of human operators in summarizing soccer matches by capturing the significance of key events. Graphical action profiles offer visual insights. | The model has the ability to transfer knowledge effectively between datasets from different broadcasting companies and leagues, can effectively work with different types of data from various sources, showcasing flexibility.  Attention layer, learns to identify significant features in the input data, which is not limited to soccer-related events. | Focuses on event metadata and does not incorporate analysis of audio or visual content from soccer videos. This leads to excluding important contextual information.  The process of summarizing complex action profiles into curves may result in a loss of information. |

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| Rahul. S. Bhat, Jayanth. 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 | Recognizing and clipping crucial occurrences in a cricket match by considering event-based attributes. Dataset Selection-Gathering data from various sources like YouTube and Hotstar. Key frames are identified using Energy levels. Features are extracted from the frames using CNNs(VGG16 and ResNet50). Features vectors are fed to a LSTM network to obtain the word embeddings and a caption for each ball is obtained. | The method categorizes scenes in cricket matches accurately, improving the performance of the summarization. By leveraging Deep Learning models such as FCNN and LSTM, the approach can effectively extract features and generate captions. The model covers a wide range of events in cricket matches, including wickets, fours, sixes, player celebrations, thereby enhancing viewer experience. | The ResNet model used for feature extraction requires significant processing power. The use of Deep Learning models like 3D-CNN and LSTM may introduce complexity training of the system. The approach may be tailored specifically for cricket videos, limiting its applicability to other sports. |

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| 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 | Multi-modal architecture for generating cricket highlights efficiently by identifying key events through commentary text and visual data analysis. This involves splitting the video into non-replay deliveries, transcribing commentary using Automated Speech Recognition (ASR), preprocessing the text, and utilizing a Large Language Model (LLM) to predict event occurrences. Video analysis focuses on cues like replays, bowler positions, commentary. | This approach identifies key events through multi-modal analysis of commentary text and visual data. By utilizing cues like replays and bowler positions, the system can accurately extract pivotal moments, enhancing the viewer experience. The use of Automated Speech Recognition (ASR) for transcribing commentary streamlines the process, while the LLMs improves event prediction accuracy. | Potential inaccuracies arise in event identification due to off-topic conversations being classified as relevant events, introducing noise. While the architecture excels in bowler and replay detection, it may face challenges in certain settings, leading to inferior performance. The system's scalability to encompass other sports needs further exploration. |

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| Sushant Gautam, Cise Midoglu, Saeed Shafiee Sabet, Dinesh Baniya Kshatri, and Pål Halvorsen. 2022. Soccer Game Summarization using Audio Commentary, Metadata, and Captions. In Proceedings of the 1st Workshop on User-centric Narrative Summarization of Long Videos (NarSUM '22). Association for Computing Machinery, New York, NY, USA, 13–22. https://doi.org/10.1145/3552463.3557019 | This approach creates soccer game summarization by creating an automated pipeline that utilizes audio commentary, metadata, and captions to generate text summaries. The methodology involves extending existing datasets with ground truth summaries, designing the summarization pipeline, and conducting a comparative analysis of alternative methods. By integrating NLP tools and exploring multimodal inputs, the study seeks to enhance summarization. | The automated pipeline streamlines the summarization process, reducing manual effort and enhancing scalability. Incorporating Natural Language Processing tools enables the analysis of complex data sources, improving the quality of the output. The method's comparative analysis helps identify effective summarization approaches. | Limited availability of comprehensive public datasets containing diverse information, the complexity of integrating multimodal data sources, and the need for robust multilingual support. Automation may still require manual intervention, impacting scalability. Adapting the approach to other sports or domains pose challenges |

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| 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 | Automatically detection and summarizing important event in cricket match .goal to provide high accuracy by using computer vision technologies and machine learning,  It uses method like video shot detection,sound detection,score board recognition(OCR),CNN,  highlight generation. These methods are used to create highlights. | Efficiency:  The use of video shot detection reduces processing time by breaking down the full video into smaller fragments, focusing on frames with significant changes.  - Incorporating techniques such as sound detection and optical character recognition (OCR) ensures a comprehensive identification of key events, including boundaries, sixes, and wickets.  - The algorithm generates highlights objectively based on detected runs and wickets. | - The accuracy of the system is directly dependent on the quality of the input scoreboard image for OCR. Poor quality or unclear scoreboards may lead to errors in event detection.  -Processing Time challenges for Long Matches |

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| A. S. Parihar, R. Mittal, P. Jain and Himanshu, "Survey and Comparison of Video Summarization Techniques," *2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP)*, Chennai, India, 2021, pp. 268-272, doi: 10.1109/ICCCSP52374.2021.9465347 | The objective of the paper is to analyze and compare various architectures for video summarization.  The paper explores and compares various techniques and methodologies for video summarization, including supervised and unsupervised approaches, utilizing architectures such as encoder-decoder networks, attention-based models, and reinforcement learning frameworks. | -Efficiently condenses large video content into concise summaries.  -For viewers it provides the most relevant and engaging content. This helps maintain user interest and engagement, leading to a more satisfying viewing experience. | -Different algorithms or human annotators may produce varying summaries, introducing subjectivity and bias into the summarization process.  -Resource Intensive: Some summarization techniques require significant computational resources. |

| Paper Details | Objective of paper, Techniques/Methods | Advantages | Limitations |
| --- | --- | --- | --- |
| 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 | Weighted dynamic heartbeat graph for detecting events in a dynamic text stream like social media data like twitter, Weighted Dynamic Heartbeat Graphs (WDHG) capture evolving word usage over time. Key features, Growth Factor (GF) and Aggregated Centrality (AC), enabling the detection of significant events by classifying and ranking strong snapshots. | WDHG efficiently detects and highlights emerging topics in dynamic text streams.  suitable for processing large volumes of Twitter data in real-time event detection scenarios. | Dependency on Keyword Frequency: The method heavily relies on keyword frequency, which may lead to biased results, especially in noisy or diverse text streams.  The approach may face challenges in handling bursty events where topics rapidly gain popularity and decline. |

# **Product Perspective**

# **Product Features**

* **Scoreboard Analysis:**

Extracts and interprets scoreboard information to provide insights into the match's progress.

* **Video Summarization:**

Utilizes video processing techniques to generate concise summaries of key moments and highlights.

* **Multimodal Analysis:**

Integrates information from video, text, and scoreboard sources for comprehensive analysis.

# **User Classes and Characteristics**

* **Casual Viewers**: These users may have a general interest in cricket but may not be deeply familiar with the game's intricacies. They may prefer concise summaries that highlight key moments such as wickets, boundaries, and significant milestones.
* **Enthusiastic Fans**: These users are passionate about cricket and may have a good understanding of the game. They may seek more detailed summaries that include insights into player performances, strategic decisions, and match dynamics
* **Coaches and Players**: Coaches and players use video summarization for performance analysis and tactical planning. They require comprehensive summaries that provide detailed insights into player performances, strategies employed, and areas for improvement.
* **Media Professionals**: Journalists, broadcasters, and content creators often use cricket video summarization to produce news articles, match reports, and highlight reels. They need access to curated summaries with high-quality visuals and relevant statistics for their content production.

# **Operating Environment**

* **Hardware Platform**: Compatible with standard desktop computers, laptops, and mobile devices.
* **Operating System**: Supports multiple operating systems, including Windows, macOS, and Linux.
* **Software Components**: Requires web browsers with modern HTML5 and CSS3 support for the user interface.

# **General Constraints, Assumptions and Dependencies**

* **Legal Implications:**

Compliance with data protection regulations, copyright laws, and terms of service of social media platforms (such as Twitter) is crucial to avoid legal issues.

Obtaining necessary permissions and licenses for using copyrighted content (such as sports broadcasts) is essential to ensure legal compliance.

* **Usage Limitations:**

The project's success may depend on the availability and access to real-time sports data, including Twitter feeds and live video streams.

Dependence on third-party APIs or data sources may introduce usage limitations, such as rate limits or data access restrictions.

* **Assumptions Made in the Project:**

Availability of Data: Assumes the availability of sufficient and reliable data sources in real-time for analysis and summarization.

Consistency in Data Format: Assumes a level of consistency in the format and structure of data sources for effective processing.

# **Risks**

* **Resource Constraints:** Limited computational resources may impact system performance and scalability.
* **Data Availability:** Reliance on external data sources poses risks related to data reliability and accessibility.
* **Technical Challenges:** Implementing advanced techniques may encounter technical complexities and limitations.
* **Regulatory Compliance:** Non-compliance with legal regulations may result in legal risks and consequences.

# **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: Effectively deals with filtering out noise and maintains accuracy in identifying key moments.
* Highlight Generation: Automatically selects and compile the most important or exciting moment from the game into a highlight reel.
* Accuracy in Key Moment Detection: The system should strive for high accuracy in identifying key moments within the game footage
* Faster Processing Speed than Manual Methods: The system should generate summaries significantly faster than manual editing, ideally near real-time for live games.

# External Interface Requirements

# User Interfaces

* The system shall provide a user-friendly graphical user interface (GUI) accessible via web or desktop application.
* Required screen formats shall adhere to GUI standards for styles, ensuring consistency and ease of use.
* Screen layouts shall be intuitive, with standard functions such as navigation menus, search bars, and help documentation available.
* Error messages shall be clear, concise, and descriptive, guiding users in resolving issues effectively.

# Hardware Requirements

* The system shall be compatible with standard computing hardware, including desktops, laptops, and servers.
* It should support various input devices such as keyboards, mouse, touchscreens, and microphones for user interaction.
* Hardware requirements shall be minimal to ensure accessibility and scalability across different computing environments

# Software Requirements

The software requirements are as follows:

1. **TensorFlow:**

* **Name and Description:** TensorFlow is an open-source machine learning framework developed by Google for building and training deep learning models.
* **Version / Release Number:** TensorFlow 2.5.0
* **Databases:** TensorFlow does not directly interact with databases, but it can be used alongside databases for data preprocessing and model training.
* **Operating Systems:** Compatible with Windows, macOS, Linux
* **Tools and Libraries:** TensorFlow includes a wide range of tools and libraries for deep learning tasks, including TensorFlow Keras for building neural networks.
* **Source:** TensorFlow is available on GitHub for source code access, contributions, and issue tracking.

1. **OpenCV:**

* **Name and Description:** OpenCV (Open Source Computer Vision Library) is a library of programming functions for real-time computer vision tasks.
* **Version / Release Number:** OpenCV 4.5.3
* **Databases:** OpenCV does not directly interact with databases but can be used for image and video processing tasks.
* **Operating Systems:** Compatible with Windows, macOS, Linux, Android, iOS
* **Tools and Libraries:** OpenCV provides a comprehensive suite of tools and algorithms for image processing, including feature detection, object recognition, and video analysis.
* **Source:** OpenCV is open-source and available on GitHub for access to source code, documentation, and community contributions.

1. **MySQL Database:**

* **Name and Description:** MySQL is an open-source relational database management system (RDBMS) widely used for indexing and storing structured data.
* **Version / Release Number:** MySQL 8.0
* **Operating Systems:** Compatible with Windows, macOS, Linux
* **Tools and Libraries:** MySQL provides a robust set of tools for managing databases, including data indexing, querying, and administration.
* **Source:** MySQL is open-source and available for download from the official website or package repositories for various operating systems.

# **Communication Interfaces**

* The system shall support communication via local area network protocols (e.g., TCP/IP) for data exchange between client and server components.
* Standard communication standards and protocols shall be adhered to, ensuring compatibility and interoperability with other systems.
* Line speed and buffer size requirements shall be optimized for efficient data transmission and processing, minimizing latency and maximizing throughput.

# **Non-Functional Requirements**

# **Performance Requirement**

**6.1.1 Processing Speed:**

* Objective: The system should process sports videos and generate highlights in real-time or near-real-time.
* Metric: The processing time for a given video should not exceed 1 minute. The average processing time should be within 30 seconds.

**6.1.2 Scalability:**

* Objective: The system should be scalable to accommodate an increasing number of users and diverse sports events without compromising performance.
* Metric: The system should support 1000 simultaneous video processing requests without significant degradation in performance. Scalability tests should be conducted periodically to ensure continued performance under increased loads.

**6.1.3 Reliability:**

* Objective: The system should reliably detect and highlight significant events in sports videos.
* Metric: The system should have a detection accuracy of at least 90% for major sports events. Additionally, the Mean Time Between Failures (MTBF) should be at least 500 hours.

**6.1.4 Robustness:**

* Objective: The system should gracefully handle variations in video quality, noise, and unexpected data patterns.
* Metric: The system should maintain a consistent performance level, even with videos of varying resolutions (e.g., SD, HD), frame rates, and commentary styles. Robustness testing should include scenarios with high background noise and diverse audio qualities.

# **Safety Requirements**

**6.2.1 Data Integrity:**

* Objective: Ensuring the integrity of the data is crucial to providing accurate and reliable sports highlights.
* Metric: Implement checksums and digital signatures to verify data integrity. Establish regular data integrity checks and audits. The system should maintain a data integrity rate of 99.99% over a specified period.

**6.2.2 User Safety:**

* Objective: Implement measures to protect user privacy by anonymizing and securing any personal information collected during the extraction of tweets corresponding to significant events.
* Metric: Anonymize user-related data using encryption techniques, and ensure that personally identifiable information is securely stored. Conduct regular privacy audits to verify compliance with privacy policies.

**6.2.3 Backup and Recovery:**

* Objective: Implementing robust backup and recovery mechanisms to prevent data loss and ensure system stability.
* Metric: Regularly backup system data, including event metadata, comments, and generated highlights. Establish a backup retention policy, and ensure the ability to restore the system to a fully operational state within [specific time, e.g., 4 hours] in case of a failure or data loss. Test backup restoration processes periodically.

# **Security Requirements**

**6.3.1 Data Privacy:**

* Objective: Ensure the privacy of user data, especially comments, tweets, and any personally identifiable information.
* Metric: Implement encryption mechanisms for data transmission and storage to protect against unauthorized access. Regularly conduct privacy impact assessments and compliance checks with relevant data protection regulations. The system should maintain a privacy compliance rate of 98% or above.

**6.3.2 Access Control:**

* Objective: Implement access controls to ensure that only authorized users can perform certain actions, such as accessing sensitive data.
* Metric: Implement role-based access control (RBAC) to restrict access based on user roles. Regularly review and update user access permissions. Monitor and log user access to sensitive data, and investigate any unauthorized access incidents.

**6.3.3 Secure Communication:**

* Objective: Follow best practices for integrating with third-party APIs, especially Twitter, to ensure the security of communication and data exchange. Protect API keys and tokens, and regularly rotate them for added security.
* Metric: Regularly update and secure API keys and tokens. Ensure that communication with third-party APIs is encrypted using secure protocols (e.g., HTTPS). Monitor API usage and access logs to detect and respond to any suspicious activity promptly.

# **Appendix A: Definitions, Acronyms and Abbreviations**

1. **Multi-modal Data Analysis**: Analysis of data from multiple sources or modalities, such as Twitter streams, video analysis, and audio analysis, to extract insights or information.
2. **Computer Vision**: A field of artificial intelligence that focuses on enabling computers to interpret and understand visual information from the real world, often involving image or video processing.
3. **Audio Analysis**: The process of analyzing audio signals to extract meaningful information or features, such as speech recognition, sound classification, or sentiment analysis.
4. **Twitter Streams:** Continuous, real-time flow of tweets or messages posted on the social media platform Twitter.
5. **Project Scope:** The boundaries and extent of the project, defining what will and will not be included in the development effort.
6. **Project Description:** A summary of the project's goals, objectives, and intended outcomes.
7. **Purpose:** The reason or objective behind undertaking the project.
8. **Benefits:** The positive outcomes or advantages expected from the successful implementation of the project.
9. **Objectives and Goals:** Specific targets or aims that the project intends to achieve.
10. **Coverage:** The range or extent to which the system will operate, including supported sports and data sources.
11. **Limitations:** Restrictions or constraints that may affect the system's functionality or performance.
12. **Literature Survey or Existing System:** A review of existing research or systems relevant to the project.
13. **Paper Details:** Information about research papers, including their objectives, techniques/methods used, advantages, and limitations.
14. **Encoder:** A component in a machine learning model responsible for transforming input data into a latent representation.
15. **Decoder:** A component in a machine learning model responsible for reconstructing output data from a latent representation.
16. **Variational Autoencoder (VAE):** A type of generative model in machine learning that learns a low-dimensional representation of input data.
17. **Poisson Distribution:** A probability distribution that expresses the likelihood of a given number of events occurring in a fixed interval of time or space.
18. **User Interface:** The point of interaction between users and a computer system, including visual elements such as screens, buttons, and menus.
19. **Machine Learning**: A subset of artificial intelligence that focuses on algorithms and models that enable computers to learn from data and make predictions or decisions.
20. **Deep Learning:** A subfield of machine learning that utilizes neural networks with multiple layers to learn representations of data.
21. **Convolutional Neural Network (CNN):** A type of deep neural network commonly used for image analysis and recognition tasks.
22. **Recurrent Neural Network (RNN):** A type of neural network designed to handle sequential data by maintaining a memory of previous inputs.
23. **Long Short-Term Memory (LSTM):** A type of recurrent neural network architecture capable of learning long-term dependencies in sequential data.
24. **Graphical User Interface (GUI):** A type of user interface that allows users to interact with electronic devices using graphical elements such as windows, icons, and buttons.
25. **OCR (Optical Character Recognition):** The process of converting images of text into machine-encoded text.
26. **NLP (Natural Language Processing):** A field of artificial intelligence focused on the interaction between computers and human language, including tasks such as text analysis, translation, and sentiment analysis.
27. **API (Application Programming Interface):** A set of rules and protocols that allows different software applications to communicate with each other.
28. **TensorFlow:** An open-source machine learning framework developed by Google for building and training deep learning models.
29. **OpenCV (Open Source Computer Vision Library):** A library of programming functions for real-time computer vision tasks.
30. **MySQL:** An open-source relational database management system widely used for indexing and storing structured data.

# **Appendix B: References**

<https://ieeexplore.ieee.org/abstract/document/10393344>

“Automatic Highlight Generation of Soccer Videos”

<https://link.springer.com/article/10.1007/s10462-023-10444-0>

“Video summarization using deep learning techniques: a detailed analysis and investigation”

<https://ieeexplore.ieee.org/document/9287062>

“Profiling Actions for Sport Video Summarization: An attention signal analysis”

<https://dl.acm.org/doi/abs/10.1145/3552463.3557019>

“Soccer Game Summarization using Audio Commentary, Metadata, and Captions”

<https://arxiv.org/ftp/arxiv/papers/2101/2101.08434>

“VIDEO SUMMARIZATION: STUDY OF VARIOUS TECHNIQUES”

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7479531>

“An Efficient Framework for Automatic Highlights Generation from Sports Videos”

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5711541>

“Multi-Modal Summarization of Key Events and Top Players in Sports Tournament Videos”

<https://ieeexplore.ieee.org/abstract/document/7321723>

“Soccer Video Summarization using Video Content Analysis and Social Media Streams”

<https://ieeexplore.ieee.org/document/8711625>

“[A multimodal Approach for Automatic Cricket Video Summarization](https://ieeexplore.ieee.org/document/8711625)”

<http://www.jeffreynichols.com/papers/summary-iui2012.pdf>

“[Summarizing Sporting Events Using Twitter](http://www.jeffreynichols.com/papers/summary-iui2012.pdf)”

<https://ieeexplore.ieee.org/document/7321723>

“[Soccer Video Summarization using Video Content Analysis and Social Media Stream](https://ieeexplore.ieee.org/document/7321723)”

<https://link.springer.com/article/10.1007/s11063-020-10200-3>

“[Content-Aware Summarization of Broadcast Sports Videos: An Audio-Visual Feature Extraction Approach](https://link.springer.com/article/10.1007/s11063-020-10200-3)”