

Contents

Business Requirement Document (BRD)	1
Introduction.....	2
Trending News Classifier.....	2
Video Pattern Classification and Screening Module (VPCSM).....	2
Sentiment Analysis Module for Automated News Generator	2
Content Personalization.....	2
Social Media News Application.....	3
Project Overview	3
Overview:	3
Objective:	4
Trending News Classifier:	4
Video Pattern Classification and Screening Module (VPCSM):	4
Sentiment Analysis Module for Automated News Generator:	4
Content Personalization:	4
Social Media News Application:	5
Project Timeline	5
Business Requirements	5
Functional Requirements:.....	5
Trending News Classifier:	5
Video Pattern Classification and Screening Module (VPCSM):	6
Sentiment Analysis Module for Automated News Generator:	7
Content Personalization Functional Requirement:	9
Social Media News Application Functional Requirements:.....	10
Non-Functional Requirements:.....	11
User and Client Requirements:	13
User Stories:	15
Dependencies:	17
Constraints	18
Risks and Mitigation Strategies	20
Flowchart.....	27
Appendices	41
Engagement Metrics Model :-	41
Video Categorization:	42

Business Requirement Document (BRD)

Introduction

In today's fast-paced digital landscape, staying updated with real-time news has become crucial for individuals and organizations alike. This Business Requirements Document (BRD) presents a comprehensive outline of multiple projects, each aiming to address specific aspects of news discovery, video content management, sentiment analysis, and social media news curation.

Trending News Classifier

- The Trending News Classifier project seeks to develop a cutting-edge system that discovers trending news by leveraging the power of established news agencies and Twitter. By integrating influential accounts, velocity of tweets, hashtags, highest engagement of tweets, and trending topics, this system will provide users with a user-friendly interface to browse, search, and discover the most relevant and popular news.

Video Pattern Classification and Screening Module (VPCSM)

- The Video Pattern Classification and Screening Module (VPCSM) addresses the growing demand for efficient and accurate video content extraction in the digital age. Using advanced computer vision techniques and deep learning algorithms, VPCSM enables the identification of specific patterns, objects, and behaviors within video data. Additionally, it empowers sentiment analysis by transforming audio content into written text, providing insights into emotions and sentiments conveyed in videos.

Sentiment Analysis Module for Automated News Generator

- The Sentiment Analysis Module for the Automated News Generator is a critical component of an AI-driven platform that generates news articles based on trending topics and user preferences. This module utilizes Natural Language Processing (NLP) techniques and state-of-the-art sentiment analysis models to determine the emotional tone of video transcript and trending news text data. By incorporating sentiment analysis results, the Automated News Generator aims to provide news articles with appropriate emotional context to enhance user engagement and satisfaction.

Content Personalization

- The Content Personalization aimed at developing a system that provides personalized content recommendations to users based on their interests and behavior. The system will analyze user data to build individual profiles and use sophisticated algorithms to match users with relevant content. The objectives include increasing user engagement, improving conversion rates, and enhancing the overall user experience.

Key stakeholders, project scope, constraints, and a cost-benefit analysis are outlined to ensure the successful implementation of the content personalization system.

Social Media News Application

- The Social Media News Application is a versatile platform that allows users to record, upload, and share videos, photos, and descriptions of local events, effectively creating local news stories. This application employs an algorithm that promotes stories based on public engagement, and users can report inappropriate content while accessing regional, national, and global news from trusted sources through APIs. Additionally, advanced Machine Learning (ML) and Deep Learning (DL) models are integrated for content verification and description generation, ensuring reliable and relevant news content.

Each of these projects contributes to a more informed, efficient, and user-centric approach to news discovery, content management, and sentiment analysis. By harnessing the power of technology and cutting-edge algorithms, these systems strive to enhance user experiences and cater to the diverse emotional preferences of their audiences.

As these projects progress, the teams behind each endeavor will continue to refine and optimize their solutions, taking into account user feedback and emerging trends in the dynamic landscape of news consumption. The pursuit of excellence in news curation remains at the forefront of these initiatives, empowering individuals and organizations with timely, accurate, and engaging information for making informed decisions and staying connected with the world around them.

Project Overview

Overview:

- In the fast-paced digital age, the need for efficient and accurate systems to discover trending news, analyze video content, perform sentiment analysis, and curate social media news is more critical than ever. This Business Requirement Document (BRD) outlines the objectives and functionalities of four projects, namely the "Trending News Classifier," "Video Pattern Classification and Screening Module (VPCSM)," "Sentiment Analysis Module for Automated News Generator," and "Social Media News Application."

Objective:

Trending News Classifier:

- The Trending News Classifier project aims to develop a cutting-edge system capable of identifying trending news using Twitter's data. By integrating information from Journalists, News agencies, Experts, and Thought leaders, the system will extract relevant data, including hashtags, likes, and comments, for further analysis. The primary objective is to rank tweets based on engagement metrics and set thresholds to identify potential trending news.

Additionally, the system will verify the potential trending topics by comparing them with actual trending topics.

Video Pattern Classification and Screening Module (VPCSM):

- The Video Pattern Classification and Screening Module (VPCSM) is designed to analyze, classify, and screen video content efficiently. Using advanced computer vision techniques and deep learning algorithms, VPCSM's key objective is to identify specific patterns, objects, or behaviors within video data. Additionally, the module will generate data for sentiment analysis, enriching the content analysis process. Through sentiment analysis, the VPCSM will understand the emotional tone conveyed within the video content.

Sentiment Analysis Module for Automated News Generator:

- The Sentiment Analysis Module aims to enhance an Automated News Generator by providing news articles with appropriate emotional context. Using Natural Language Processing (NLP) techniques and state-of-the-art sentiment analysis models, the module will analyze video transcript and trending news text data, categorizing sentiments as positive, neutral, or negative. The objectives are to enhance news content, improve the user experience, and increase relevance by aligning the generated news articles with users' emotional preferences.

Content Personalization:

- Content Personalization system to provide personalized content recommendations based on user interests and behavior. The objectives are to increase user engagement, improve conversion rates, and enhance the user experience. The system will analyze user data, build profiles, and use advanced algorithms to match users with relevant content. Key stakeholders include Marketing, Product, Engineering, Content Team, and Legal. The project has a budget of \$100,000 and aims to generate a 20% increase in revenue through personalized recommendations. Overall, the project aims to transform the user experience and drive business growth.

Social Media News Application:

- The Social Media News Application's objective is to provide users with a platform to record, upload, and share local news stories through videos, photos, and descriptions. The application will foster user engagement through a voting system based on likes, dislikes, and neutral reactions. To ensure content appropriateness and user safety, the application will include a reporting system and an SOS feature for emergency situations. Additionally, the application will integrate Machine Learning (ML) and Deep Learning (DL) models for content verification and generation of relevant descriptions. The application will fetch and display news from trusted sources based on users' selected tags and interests.

Project Timeline

[Provide an estimated timeline for the project, including key milestones and deadlines.]

Business Requirements

Functional Requirements:

Trending News Classifier:

- ❖ Twitter API Integration:
 - The system shall register a developer account with Twitter to obtain API keys and access tokens.
 - It should use appropriate libraries or SDKs to facilitate communication with the Twitter API.
 - OAuth authentication shall be implemented for secure access to Twitter data.
- ❖ Data Retrieval:
 - Users shall specify a list of Twitter accounts (journalists, news media outlets, experts, thought leaders) to monitor for trending news.
 - The system shall fetch recent tweets from the selected accounts using Twitter API endpoints.
- ❖ Data Filtering:
 - Fetched tweets shall be filtered to exclude non-news-related content, such as personal opinions or spam.
 - Natural Language Processing (NLP) techniques may be used to categorize tweets into relevant news topics.
- ❖ Engagement Metrics Gathering:
 - The system shall extract engagement metrics (likes, comments, retweets, views) from each fetched tweet.
 - These metrics shall be recorded and associated with the corresponding tweets in the database.
- ❖ Data Storage:
 - A database shall store the retrieved tweets and their associated metrics.
 - Popular database systems like MySQL, PostgreSQL, or NoSQL databases like MongoDB can be used.
- ❖ Trending Detection Algorithm:
 - The system shall analyze engagement metrics over a specified time period to identify tweets with high engagement.
 - Tweets with significant increases in likes, comments, or retweets compared to historical averages shall be considered potential trending news.
- ❖ Trending Topic Classification:
 - The system shall categorize trending tweets into different news topics based on content and context using NLP techniques or keyword matching.
- ❖ Data Visualization:

- The system shall provide interactive visualizations (line charts, bar graphs) to present trending news data and engagement metrics.
- Users shall view trends over time and compare metrics for different news topics.
- ❖ Error Handling:
 - The system shall handle potential errors, such as API rate-limiting or connectivity problems, and display appropriate error messages.
 - Retry mechanisms may be implemented to handle temporary API failures.
- ❖ Security and Privacy:
 - API credentials shall be stored securely following best practices.
 - User data shall be stored and transmitted securely using encryption where appropriate.

Video Pattern Classification and Screening Module (VPCSM):

- ❖ Video Input Interface:
 - The VPCSM shall feature a user-friendly video input interface allowing users to upload video files or provide streaming links from various sources.
 - The system shall accept a wide range of video formats and codecs for flexibility.
- ❖ Preprocessing Module:
 - Upon receiving video input, the VPCSM shall employ a preprocessing module to standardize the data, enhancing overall quality, and addressing potential inconsistencies in frame rates, resolutions, or encoding.
- ❖ Pattern Detection and Extraction:
 - The core functionality of the VPCSM shall lie in its pattern detection and extraction capabilities.
 - The system shall use advanced computer vision techniques and deep learning algorithms to identify specific patterns, objects, or actions within the video content.
- ❖ Classification and Tagging:
 - After successful pattern extraction, the VPCSM shall perform classification and tagging of the identified patterns.
 - Each video shall be annotated with relevant tags, making it easier for users to search and retrieve specific content later.
- ❖ Screening and Filtering:
 - The system shall allow users to set custom screening criteria based on specific tags, patterns, or classifications.
 - This feature shall enable content moderation and compliance with regulatory guidelines, ensuring inappropriate or undesirable content is filtered out or flagged for review.
- ❖ Real-time Monitoring (Optional):
 - For applications requiring real-time analysis, the VPCSM can be equipped with a real-time monitoring module.

- This module can analyze live streams or videos on the fly, enabling immediate actions and responses.
- ❖ Reporting and Analytics:
 - The VPCSM shall generate comprehensive reports and analytics, providing insights into the distribution of different patterns within the video dataset.
 - This information shall be valuable for content creators, advertisers, or researchers, helping them understand audience preferences and engagement.
- ❖ Security and Privacy:
 - To maintain data security and user privacy, the VPCSM shall ensure compliance with relevant data protection laws and incorporate encryption measures for sensitive information.

Sentiment Analysis Module for Automated News Generator:

- ❖ News Article Sentiment Analysis:
 - The system shall analyze the sentiment of news articles obtained from various sources.
 - Sentiment shall be categorized as positive, neutral, or negative for each news article.
 - Sentiment scores shall be generated to quantify the emotional tone.
- ❖ Video Transcript Sentiment Analysis:
 - The system shall perform sentiment analysis on transcribed audio content from videos.
 - Sentiment shall be categorized as positive, neutral, or negative for each video transcript.
 - Sentiment scores shall be generated to quantify the emotional tone.
- ❖ Real-time Sentiment Analysis:
 - The system shall provide real-time sentiment analysis for trending news articles and video transcripts.
 - Sentiment analysis results shall be updated in real-time as new data is processed.
- ❖ Visualization of Sentiment Analysis Results:
 - The system shall display sentiment analysis results in a user-friendly interface.
 - Sentiment analysis results shall be presented using appropriate charts or graphs to visualize sentiment scores and emotional categories.
- ❖ User Personalization:
 - The system shall allow users to personalize sentiment analysis preferences based on topics or categories of interest.
 - Personalized sentiment analysis updates shall be provided to users based on their preferences.
- ❖ Verification of News Sources:
 - The system shall verify the authenticity and credibility of news articles from reliable sources for accurate sentiment analysis.
 - Unverified or potentially unreliable news articles shall be flagged or excluded from sentiment analysis results.

- ❖ Integration with Trending News and Video Modules:
 - The sentiment analysis module shall be seamlessly integrated with the Trending News and Video Pattern Classification and Screening Modules.
 - The sentiment analysis module shall communicate effectively with other modules to obtain data for analysis.
- ❖ Error Handling and Exception Management:
 - The system shall implement effective error handling and exception management to handle any unexpected issues or failures during sentiment analysis processing.
 - Error messages shall be clear and informative to assist users in understanding the nature of the error.
- ❖ Data Storage and Management:
 - The sentiment analysis module shall store user data securely and comply with data protection regulations.
 - Data storage and management processes shall be optimized for efficient retrieval and analysis of sentiment data.
- ❖ Reporting and Analytics:
 - The system shall generate comprehensive sentiment analysis reports and analytics to provide insights into emotional trends.
 - Sentiment analysis trends and patterns shall be presented to content creators, journalists, and other stakeholders for decision-making purposes.

Content Personalization Functional Requirement:

- ❖ Recommendation Algorithms:
 - Develop sophisticated algorithms to analyze user profiles and content metadata.
 - Generate personalized content recommendations based on user interests.
- ❖ Display of Personalized Recommendations:
 - Present personalized content recommendations on the homepage and user profile pages.
 - Ensure the presentation is visually appealing and user-friendly.
- ❖ User Feedback Mechanism:
 - Provide a feedback mechanism for users to rate and provide feedback on recommended content.
 - Utilize user feedback to continuously improve the quality of future recommendations.
- ❖ Compatibility with Content Management System:

- Seamlessly integrate with the existing content management system.
- Deliver personalized content recommendations across the platform.
- ❖ Data Privacy and Security:
 - Implement robust data protection measures to ensure user data privacy and security.
 - Comply with all relevant privacy regulations and standards.
- ❖ Scalability and Latency:
 - Design the system to handle increased user traffic and data efficiently.
 - Deliver real-time recommendations with minimal latency.
- ❖ Uptime and Reliability:
 - Ensure high uptime for the recommendation engine.
 - Minimize downtime and ensure continuous service availability.
- ❖ Error Handling:
 - Implement effective error handling mechanisms.
 - Provide clear error messages to users when necessary.
- ❖ Content Metadata Management:
 - Handle missing or incomplete content metadata effectively.
 - Ensure the system can work with potentially incomplete information.
- ❖ User Personalization Control:
 - Allow users to customize and control their personalization preferences.
 - Enable users to adjust the types and frequency of recommended content according to their preferences.

Social Media News Application Functional Requirements:

- ❖ User Registration and Content Upload:
 - Users shall be able to register and create accounts to access the application's features.
 - Registered users shall have the ability to upload videos and photos along with descriptions, tags, location, and time.
- ❖ Voting System:
 - The application shall include a voting system based on public engagement (likes, dislikes, and neutral reactions).
 - User-uploaded content shall be promoted based on voting results.
- ❖ Reporting System:

- The application shall include a reporting system for users to flag inappropriate content.
 - Reported content shall be reviewed by administrators for necessary actions.
- ❖ SOS Feature:
- The application shall feature an SOS feature for emergency situations.
 - Users can use this feature to seek immediate help or assistance.
- ❖ ML and DL Integration:
- The application shall integrate Machine Learning (ML) and Deep Learning (DL) models for content verification and description generation.
 - These models shall ensure content accuracy and provide relevant descriptions for user-uploaded content.
- ❖ Fetching News from Trusted Sources:
- The application shall fetch and display news from trusted sources based on user-selected tags and interests.
 - Users can access curated news content aligned with their preferences.
- ❖ User Engagement:
- Users shall be able to engage with news stories by liking, commenting, and sharing them.
 - The application shall encourage user interaction with news content.
- ❖ Vernacular Support:
- The application shall support vernacular languages to cater to a diverse user base.
 - Users can access news content in their preferred languages.
- ❖ Guest User Login:
- The application shall allow guest users to access limited features without registration.
 - Full functionality shall be available to registered users.
- ❖ News Filtering:
- The application shall enable users to filter news based on various parameters such as region, interest, time, etc.
 - Users can customize their news feed according to their preferences.

Non-Functional Requirements:

❖ Performance:

- The application shall have a fast response time for fetching and delivering breaking news updates to ensure a seamless user experience.
- It shall handle a high volume of data from news agencies and Twitter, processing and delivering updates in real-time.
- Performance testing shall be conducted to ensure optimal performance during peak usage periods.

❖ Availability:

- The application shall have a high level of availability to ensure uninterrupted availability of breaking news updates.
- Load balancing and redundancy strategies shall be implemented to handle increased user traffic and mitigate the impact of system failures.

❖ Scalability:

- The application shall be designed to accommodate increasing user traffic and a growing volume of breaking news updates.
- Horizontal and vertical scalability strategies shall be considered to ensure optimal performance as the user base expands.
- Scalability testing shall be performed to assess the application's ability to handle increasing demands.

❖ Data Privacy:

- The application shall comply with relevant data privacy regulations and guidelines, such as GDPR or CCPA.
- User consent shall be obtained for collecting, storing, and processing personal data.
- Data anonymization techniques shall be employed to protect user privacy when necessary.
- Adequate measures shall be in place to secure data storage and prevent unauthorized access or data breaches.

❖ Usability:

- The application's user interface shall be intuitive and user-friendly, requiring minimal user training.
- Users shall find it easy to navigate, access news content, and interact with the application's features.

❖ Compliance:

- The application shall adhere to industry standards and best practices for news content presentation and user engagement.
- It shall comply with relevant copyright and intellectual property regulations.

❖ Interoperability:

- The application shall be designed to integrate seamlessly with other systems and platforms, allowing for data exchange and collaboration.
- Standard data exchange formats shall be supported for easy sharing of news content with external systems if required.

❖ Maintainability:

- The application's codebase shall be well-documented to facilitate future maintenance and updates.
- Modular design and adherence to coding standards shall enhance code readability and maintainability.

❖ Portability:

- The application shall be deployable on various platforms and operating systems, such as Windows, Linux, and macOS.
- Containerization technologies like Docker shall be used for easy deployment and portability.

User and Client Requirements:

- ❖ User Searching for Breaking News:
 - Users shall be able to search for breaking news updates by entering specific keywords, topics, or categories of interest.
 - The application shall retrieve and display relevant breaking news content based on user search queries.
- ❖ Automated Breaking News Detection:
 - The application shall continuously monitor news agencies and Twitter feeds in real-time.
 - It shall use advanced algorithms to automatically detect breaking news stories based on predefined criteria, such as keywords, sources, or popularity.
- ❖ Viewing Detailed News Articles and Social Media Updates:
 - Users shall have access to detailed news articles and social media updates related to breaking news stories.
 - The application shall provide a user-friendly interface to view and engage with comprehensive information, including article content, images, embedded media, and relevant social media discussions.
- ❖ Content Personalization:
 - The application shall provide personalized news content based on users' interests and preferences.
 - Users shall receive relevant breaking news updates aligned with their selected tags and interests.
- ❖ News Source Verification:
 - Users shall be assured of news article authenticity and credibility, as the application shall fetch content from trusted sources.
 - Unverified or potentially unreliable news articles shall be clearly indicated or excluded from the content feed.
- ❖ Real-time Updates and Notifications:
 - The application shall deliver real-time breaking news updates and notifications to users.
 - Users shall receive timely alerts on critical news events and emerging trends.
- ❖ User Engagement and Interaction:
 - Users shall be encouraged to engage with breaking news stories through likes, comments, and sharing options.
 - The application shall foster an active and participatory user community.
- ❖ Data Security and Privacy:
 - Users shall have confidence in the application's data security measures.
 - Personal data, including user preferences and interactions, shall be stored and transmitted securely, with user consent and adherence to privacy regulations.

- ❖ **Trending News Filter:**
 - Users shall be able to filter trending news based on specific parameters, such as region, time, or category.
 - Customized filters shall enable users to focus on news content relevant to their preferences.
- ❖ **Feedback and Support:**
 - Users shall have the ability to provide feedback on the application's performance and content.
 - The application shall offer responsive customer support to address user queries and concerns.
- ❖ **Accessibility:**
 - The application shall be designed to be accessible to users with disabilities, following accessibility guidelines and standards.
- ❖ **Speed and Responsiveness:**
 - Users shall experience quick loading times and responsive interactions within the application.
 - The application shall be optimized to deliver a smooth and seamless user experience.

User Stories:

- As a user, I want to search for breaking news updates by entering specific keywords, topics, or categories of interest to stay informed about current events.
- As a user, I want the system to automatically detect breaking news stories from reputable sources based on predefined criteria, such as keywords or popularity.
- As a user, I want to access detailed news articles and social media updates related to breaking news stories, including content, images, and relevant discussions.
- As a user, I want the application to provide personalized news content based on my interests and selected tags.
- As a user, I want to like, comment, and share news stories to engage with the community and contribute to discussions.
- As a user, I want to filter news content based on various parameters like region, interest, or time to focus on topics of relevance.
- As a user, I want the system to verify the authenticity and credibility of news articles for accurate sentiment analysis.

- As a user, I want the application to display sentiment analysis results in an easy-to-understand interface using charts or graphs.
- As a user, I want to personalize sentiment analysis preferences to receive updates aligned with my interests.
- As a user, I want the system to handle errors effectively and display clear error messages to understand the nature of the issue.
- As a user, I want my data, including sentiment analysis results, to be stored and transmitted securely to protect sensitive information.
- As a user, I want the ability to save news articles for later reading or reference.
- As a user, I want the application to offer multiple language options to access news content in different languages.
- As a user, I want the ability to provide feedback on news articles, reporting accuracy, and overall user experience.
- As a user, I want the option to filter out or block specific news sources or topics that I find irrelevant or inappropriate.
- As a user, I want the application to offer a seamless login and registration process with secure authentication methods.
- As a user, I want the application to allow me to save my favorite news topics or tags for quick access to relevant content.
- As a user, I want the application to offer support for accessibility features, such as screen readers, for users with disabilities.
- As a user, I want the application to have a user-friendly interface that is easy to navigate and visually appealing.
- As a user, I want to see the source and credibility of each news article, including the publisher and publication date.
- As a user, I want the application to provide real-time updates on breaking news stories to stay informed about the latest developments.
- As a user, I want the system to provide a feature to share news content within the app with other users or create private groups for discussions.
- As a user, I want the system to automatically detect breaking news stories from reputable sources based on predefined criteria, such as keywords or popularity.
- As a user, I want to access detailed news articles and social media updates related to breaking news stories, including content, images, and relevant discussions.

- As a user, I want the application to provide personalized news content based on my interests and selected tags.
- As a user, I want to like, comment, and share news stories to engage with the community and contribute to discussions.
- As a user, I want to filter news content based on various parameters like region, interest, or time to focus on topics of relevance.
- As a user, I want the system to verify the authenticity and credibility of news articles for accurate sentiment analysis.
- As a user, I want the application to display sentiment analysis results in an easy-to-understand interface using charts or graphs.
- As a user, I want to personalize sentiment analysis preferences to receive updates aligned with my interests.
- As a user, I want the system to handle errors effectively and display clear error messages to understand the nature of the issue.
- As a user, I want my data, including sentiment analysis results, to be stored and transmitted securely to protect sensitive information.

Dependencies:

- The successful integration with the Twitter API is dependent on obtaining necessary API keys and access tokens from Twitter.
- The availability and reliability of real-time breaking news updates depend on the performance of the Twitter API and the system's ability to handle high volumes of data.
- The accuracy of sentiment analysis results relies on the effectiveness of NLP techniques and the periodic retraining of sentiment analysis models.
- The seamless integration between the sentiment analysis module and other relevant modules (Trending News and Video Pattern Classification) is crucial for cohesive data analysis.
- The system's scalability is dependent on the use of appropriate horizontal and vertical scaling strategies to accommodate increasing user traffic and data volumes.

- Compliance with data privacy regulations (e.g., GDPR or CCPA) requires obtaining user consent for data collection and implementing data anonymization techniques.

Constraints

❖ Budget

- The project's budget must be carefully considered and adhered to throughout the development process. Any expenditures related to hardware, software, personnel, and other resources must align with the allocated budget. Proper financial planning and control are essential to ensure the project's successful completion without exceeding the available funds.

❖ Technology Constraints

- While implementing the video analysis system, specific technology requirements and limitations need to be taken into account. The system must be designed to utilize efficient algorithms and hardware capable of real-time processing to handle the computational complexity associated with video analysis. Additionally, compatibility with different devices and operating systems should be considered to ensure broader accessibility.

❖ Legal and Regulatory Compliance

- Compliance with legal and regulatory requirements is of utmost importance. Data privacy and security must be given high priority to protect user information during video analysis. The project must adhere to data privacy regulations such as GDPR (General Data Protection Regulation) or any other relevant regional laws to safeguard user data and maintain confidentiality. Furthermore, other legal and ethical considerations must be taken into account during the development and deployment of the system.

❖ Speech Extraction Challenges

- During speech extraction from videos, various challenges must be addressed to ensure accurate results. Background noise or music present in videos can interfere with the speech extraction process, necessitating the implementation of noise reduction techniques. Additionally, diverse accents and dialects spoken in the videos may pose challenges for speech recognition systems, affecting the accuracy of extracted speech. Moreover, when multiple speakers talk simultaneously, the system must be capable of separating their speech to provide precise results.

❖ Speech to Text Conversion Challenges

- The speech to text conversion process faces specific constraints that must be dealt with effectively. Variations in pronunciation due to diverse accents and dialects can impact the accuracy of transcription, necessitating the use of language models trained to handle such variations. Furthermore, background noise in the environment where speech recognition takes place must be minimized to ensure high-quality output. Additionally, the system must

handle homophones and ambiguities effectively to avoid misinterpretations during transcription.

❖ Project Timeline

- The project is bound by a predefined timeline that must be strictly adhered to. Timely completion of the project is crucial for the successful integration of the sentiment analysis module. Any delays in the development process may impact the overall project delivery and subsequent implementation.

❖ Data Privacy Regulations

- To maintain user trust and uphold legal obligations, the system must comply with data privacy regulations such as GDPR or other relevant laws. Adequate measures should be in place to protect user data and ensure confidentiality throughout data processing, storage, and transmission.

❖ Available Resources

- The project's progress and outcome are directly affected by the availability of resources. This includes manpower, technology, and infrastructure. Proper resource allocation and management are necessary to efficiently utilize the available resources and achieve project goals.

❖ Scalability Limitations

- The system's scalability must be considered, particularly in terms of hardware and server capabilities. Efforts should be made to ensure that the system can handle a potentially large volume of data and users without compromising its performance.

❖ Quality of Data Sources

- The accuracy and reliability of sentiment analysis results heavily rely on the quality and credibility of the data sources used for trending news and video content. It is crucial to source data from reputable and relevant sources to obtain meaningful insights.

❖ Language Support

- While the sentiment analysis module is designed to analyze sentiments in the primary language(s) supported by the natural language processing algorithms, there may be limitations in analyzing sentiment for languages other than the primary ones. Consideration should be given to expand language support based on user requirements and demand.

❖ Network Connectivity

- The system's performance may be influenced by the quality of network connectivity, particularly when accessing external data sources or APIs for news collection. Steps should be taken to handle potential network issues gracefully and ensure a seamless user experience.

❖ Mobile Device Capabilities

- As the application relies on mobile devices for recording and uploading content, the system's functionalities should be optimized to work efficiently with varying device capabilities.

- ❖ **User Engagement for Promotion Algorithm**
 - The application's promotion algorithm is dependent on user engagement. Strategies to encourage user participation and interactions should be devised to enhance the effectiveness of the promotion algorithm.
- ❖ **User Input and Cooperation for Reporting System and SOS Feature**
 - The successful functioning of the application's reporting system and SOS feature depends on user input and cooperation. User-friendly interfaces and incentives may be incorporated to encourage users to actively participate in reporting and utilizing the SOS feature.
- ❖ **Training Data for ML and DL Models**
 - The performance of the application's ML (Machine Learning) and DL (Deep Learning) models is dependent on the quality and quantity of the training data. Efforts should be made to curate diverse and relevant data to improve the models' accuracy and generalizability.
- ❖ **Availability and Reliability of APIs for News Fetching**
 - The application's news fetching feature relies on external APIs to gather information. Contingency plans should be in place to handle scenarios where APIs may become unavailable or unreliable to ensure uninterrupted news retrieval functionality.

Risks and Mitigation Strategies

- ❖ **Data Inaccuracy or Misinformation**
 - There is a risk of inaccurate or misleading information being disseminated as breaking news, leading to a loss of credibility and trust in the system.
 - **Mitigation Strategy:**
 - To address this risk, we will implement verification mechanisms and content moderation processes to ensure the accuracy and credibility of breaking news sources. By using reputable news agencies and conducting fact-checking procedures, we can minimize the chances of false or misleading information being published. Regular reviews of the verification process will be conducted to improve its effectiveness.
- ❖ **Technical Failures or System Downtime**
 - Technical failures or system downtime can impact the availability and reliability of the platform, leading to user dissatisfaction and potential revenue loss.
 - **Mitigation Strategy:**
 - To mitigate this risk, we will implement redundancy, monitoring, and maintenance procedures. Redundancy measures will help minimize single points

of failure, ensuring that critical components have backups in case of failure. Regular monitoring will allow us to proactively detect and address technical issues before they escalate. Scheduled maintenance activities will be conducted during off-peak hours to minimize user disruption.

❖ Risk Mitigation: Data Privacy and Security Concerns

- There is a risk of data privacy and security breaches when processing sensitive video content, potentially exposing user information, personal details, or confidential data.
- Mitigation Strategy:
 - To mitigate data privacy and security concerns, we will implement robust encryption measures to protect user data during transmission and storage. Adherence to relevant data protection laws and industry best practices will ensure that user information is handled with utmost care. Regular security audits and vulnerability assessments will be performed to identify and address potential weaknesses in the system's security. Access controls and user authentication mechanisms will restrict data access to authorized personnel only, reducing the risk of unauthorized data breaches.

❖ Risk Mitigation: Computational Complexity

- The complexity of video pattern analysis and classification may lead to performance issues or delays in real-time processing.
- Mitigation Strategy:
 - To address computational complexity, we will utilize optimized algorithms and hardware accelerators to improve the efficiency and speed of video processing tasks. Employing parallel processing techniques will help distribute computational load across multiple units, ensuring faster analysis. Regular monitoring of system performance and conducting performance optimizations will maintain smooth operations and prevent performance bottlenecks.

❖ Risk Mitigation: Accuracy of Speech Extraction

- Challenges in speech extraction from videos, such as background noise, accents, and overlapping speakers, may lead to inaccuracies in the extracted audio content.
- Mitigation Strategy:
 - To ensure accurate speech extraction, we will use state-of-the-art voice activity detection (VAD) algorithms to identify and isolate speech segments. Implementing speaker diarization techniques will help separate speech from multiple individuals in the video. Audio source localization will be utilized to isolate the primary speech source in videos with multiple audio sources. Regular updates and fine-tuning of speech recognition models using diverse datasets will improve accuracy.

❖ Risk Mitigation: Challenges in Sentiment Analysis

- Performing sentiment analysis on video content can be challenging due to the complexity of emotions expressed through visual and auditory cues.

- Mitigation Strategy:
 - To address challenges in sentiment analysis, we will employ multimodal sentiment analysis techniques that combine information from speech, facial expressions, and visual scene analysis. Using pre-trained models and transfer learning, we will leverage existing knowledge in sentiment analysis and fine-tune them with video-specific data for improved accuracy. Implementing natural language processing techniques will help identify sentiment-bearing words and phrases in the transcribed text from the video.

- ❖ Risk Mitigation: Overfitting and Generalization
 - There is a risk of overfitting to the training data, leading to poor generalization on unseen data.
 - Mitigation Strategy:
 - To mitigate overfitting, we will use techniques like cross-validation and data augmentation during model training to improve model generalization. Regular monitoring of model performance on validation and test datasets will help detect signs of overfitting and prompt retraining if necessary. Ensuring that the training dataset is diverse and representative of real-world scenarios will enhance model robustness.

- ❖ Risk Mitigation: Lack of Training Data
 - Acquiring a comprehensive and diverse dataset for training the models can be challenging, particularly for specialized domains.
 - Mitigation Strategy:
 - To address the lack of training data, we will collaborate with content providers, industry partners, or relevant organizations to obtain labeled datasets for training. Data augmentation techniques will be considered to artificially increase the size and diversity of the training dataset. Leveraging transfer learning and pre-trained models will reduce the need for extensive training data, as we fine-tune them with domain-specific data.

- ❖ Risk Mitigation: Ethical Considerations
 - The analysis and classification of video content may raise ethical concerns related to privacy, bias, and potential misuse of the technology.
 - Mitigation Strategy:
 - To address ethical considerations, we will establish clear guidelines and ethical standards for the use of the VPCSM, ensuring compliance with ethical and legal regulations. Regular reviews and updates of the system's ethical guidelines will be based on emerging best practices and societal norms. Thorough audits of the system's decision-making processes will be conducted to identify and address any biases in the algorithms.

- ❖ Risk Mitigation: Integration and Compatibility Issues

- Integrating the VPCSM with existing systems and platforms may present challenges related to compatibility and data exchange.
- Mitigation Strategy:
 - To mitigate integration and compatibility issues, we will develop well-documented APIs and data formats to facilitate seamless integration with other modules or external applications. Thorough compatibility testing with various video formats and codecs will ensure smooth data interchange. Providing comprehensive documentation and technical support will assist users in the integration process.
- ❖ Risk Mitigation: Legal and Copyright Issues
 - Analyzing video content may involve copyrighted materials, leading to potential legal and copyright infringement issues.
 - Mitigation Strategy:
 - To address legal and copyright issues, we will implement content filtering mechanisms to identify and flag copyrighted materials, preventing unauthorized usage. Collaboration with content creators or copyright holders will be sought to obtain proper permissions and licensing agreements for using copyrighted materials. Users and content creators will be educated about copyright regulations and the importance of adhering to intellectual property rights.
- ❖ Risk: Inappropriate Content Being Posted and Shared
 - Mitigation: Implement content moderation mechanisms to detect and remove inappropriate content promptly. Enforce community guidelines to discourage users from posting harmful content.
- ❖ Risk: User Data Being Compromised
 - Mitigation: Implement stringent data protection measures, including encryption and secure user authentication, to safeguard user data. Conduct regular security audits to identify and fix vulnerabilities.
- ❖ Risk: SOS Feature Being Misused
 - Mitigation: Implement user verification and validation processes for accessing the SOS feature. Provide clear guidelines on its appropriate usage to prevent misuse.
- ❖ Risk: ML and DL Models Making Errors in Content Verification and Description Generation
 - Mitigation: Continuously update and fine-tune ML and DL models with new data to improve accuracy. Implement human oversight to verify critical content.
- ❖ Risk: APIs Breakdown or Providing Inaccurate/Unreliable News
 - Mitigation: Implement robust API monitoring and error handling mechanisms. Collaborate with API providers to ensure data reliability and availability.
- ❖ Risk: Inaccurate Sentiment Analysis Results
 - Mitigation: To address this risk, we will implement advanced natural language processing (NLP) techniques and utilize reliable sentiment analysis models. Conducting thorough

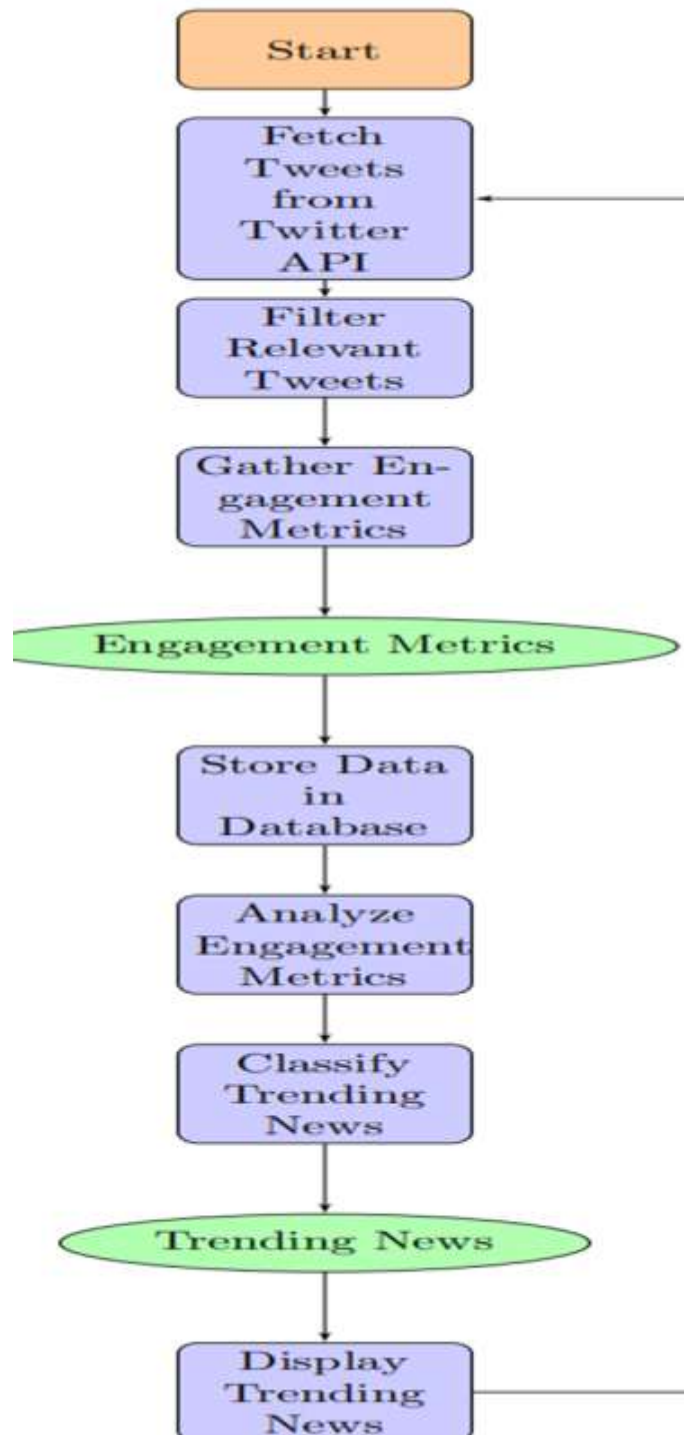
testing and validation of the sentiment analysis algorithms using a diverse dataset will ensure accurate results. Regular monitoring and updates of the sentiment analysis models will improve their effectiveness over time.

- ❖ Risk: Insufficient Data Sources for Certain Topics
 - Mitigation: To mitigate this risk, we will perform extensive research and collaborate with various reputable sources to ensure a sufficient number of reliable websites for each topic. In cases where data sources are limited, we will consider alternative sources or expand the topic scope to ensure comprehensive coverage.
- ❖ Risk: Technical Challenges in Data Collection and Processing
 - Mitigation: To address technical challenges during data collection and processing, we will allocate sufficient resources and employ experienced data analysts. Conducting regular code reviews and testing will help identify and resolve any issues early in the development process.
- ❖ Risk: Performance Bottlenecks During Peak Usage
 - Mitigation: To mitigate performance bottlenecks during peak usage, we will conduct thorough performance testing to identify potential issues. Implementing load balancing techniques will help distribute incoming requests evenly across servers, ensuring a smooth user experience during high traffic periods.
- ❖ Risk: Limited Language Support for Sentiment Analysis
 - Mitigation: To enhance language support for sentiment analysis, we will consider using multi-language sentiment analysis models or collaborating with NLP experts. Continuous efforts will be made to expand language capabilities and accommodate diverse user demographics.
- ❖ Risk: Dependency on External APIs and Data Sources
 - Mitigation: To reduce the impact of external API or data source failures, we will implement error handling and fallback mechanisms. Exploring options for local caching of data will also minimize dependency on external services.
- ❖ Risk: Incompatibility with New Data Formats or News Sources
 - Mitigation: Designing the system to be adaptable and easily extensible will help address incompatibility issues with new data formats or news sources. Regular updates and maintenance of the system will ensure support for emerging trends and technologies.
- ❖ Risk: Delayed Project Timelines
 - Mitigation: To avoid delays in the project timelines, we will set clear milestones and allocate sufficient resources for each stage of development. Regular monitoring of progress and prompt resolution of any roadblocks will be critical in ensuring timely completion.
- ❖ Risk: User Bias in Data Input
 - Mitigation: To mitigate user bias in data input, we will implement user education and validation mechanisms. Encouraging users to provide unbiased and relevant data for sentiment analysis will be a focus, and measures will be in place to filter out irrelevant or misleading information.
- ❖ Risk: Inadequate User Adoption

- Mitigation: Implement a comprehensive user onboarding and training program to ensure users understand the system's benefits and capabilities. Gather feedback from users and continually improve the user experience to enhance adoption.
- ❖ Risk: Data Privacy Compliance
 - Mitigation: Regularly review and update data privacy policies to align with changing regulations and best practices. Conduct periodic audits to ensure compliance with data privacy requirements.
- ❖ Risk: Resource Constraints
 - Mitigation: Plan resource allocation carefully, considering the project's scope and timeline. Optimize resource utilization and prioritize tasks based on their impact on project success.
- ❖ Risk: Vendor or Partner Dependence
 - Mitigation: Diversify vendor or partner collaborations to reduce dependence on a single entity. Maintain clear contractual agreements and establish contingency plans in case of unforeseen issues with vendors or partners.
- ❖ Risk: Market Competition
 - Mitigation: Continuously monitor market trends and competitor activities. Regularly update and enhance the platform's features and offerings to stay ahead of the competition.
- ❖ Risk: Budget Overruns
 - Mitigation: Develop a detailed budget plan and closely monitor project expenses. Implement cost-control measures and seek approval for any deviations from the budget.
- ❖ Risk: Scope Creep
 - Mitigation: Define and document the project's scope clearly. Establish a change control process to evaluate and approve any scope changes, ensuring they align with project goals.
- ❖ Risk: Changing Technology Landscape
 - Mitigation: Stay abreast of technological advancements and trends. Conduct regular technology assessments and updates to ensure the system remains relevant and competitive.

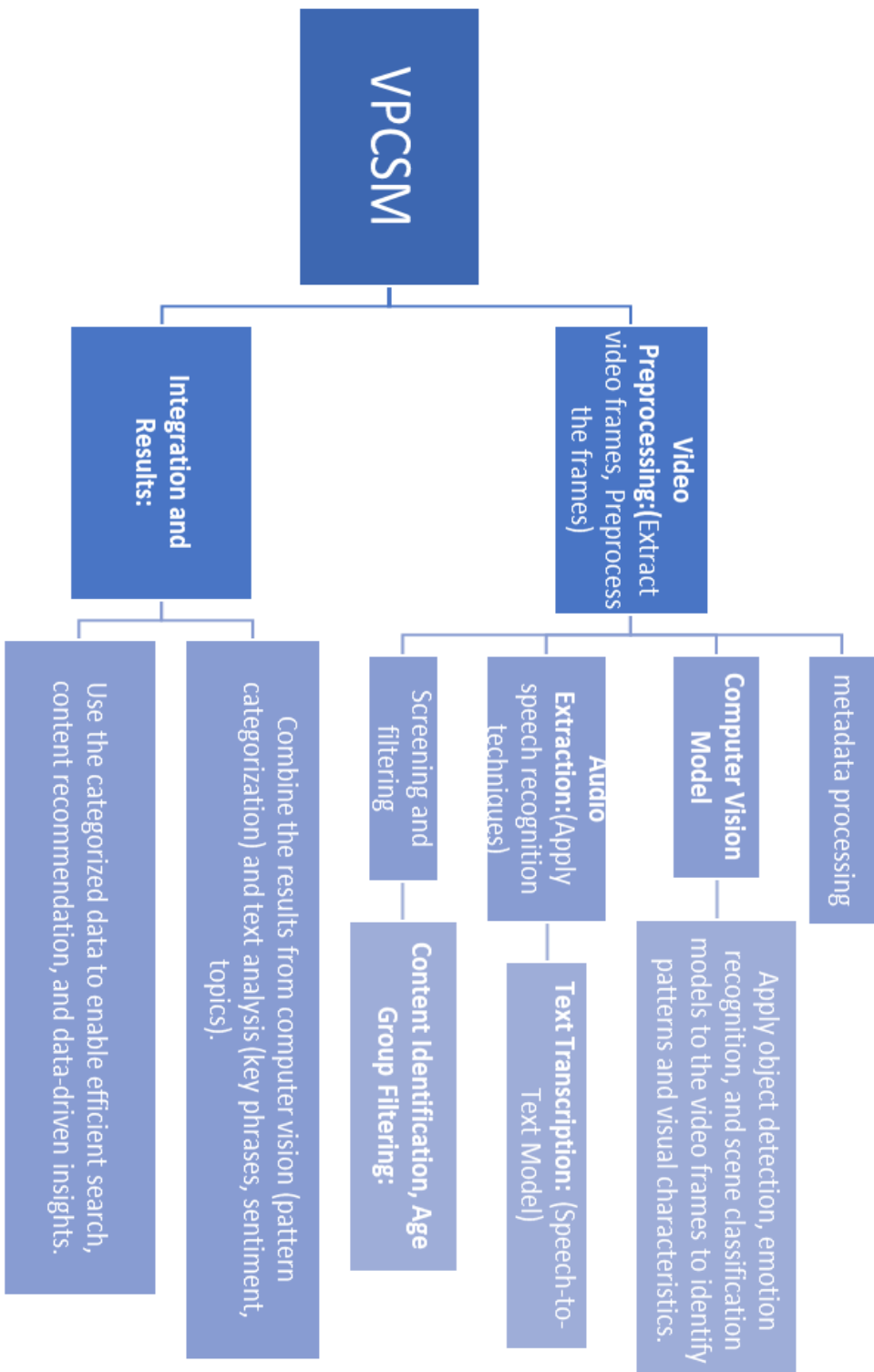
Flowchart

1. Trending News Classifier Work Flow

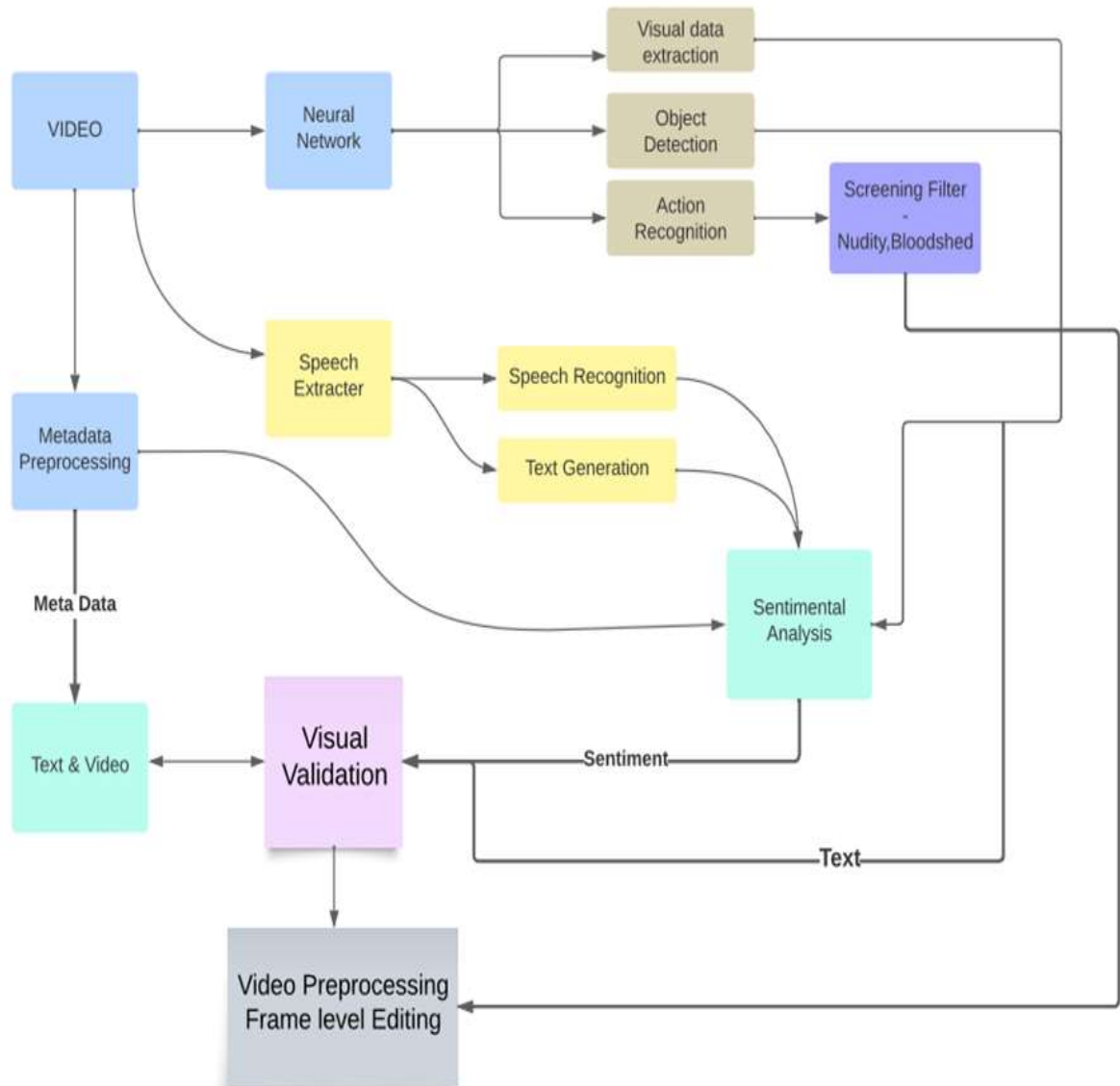


2. Video Pattern Classification and Screening Module (VPCSM)

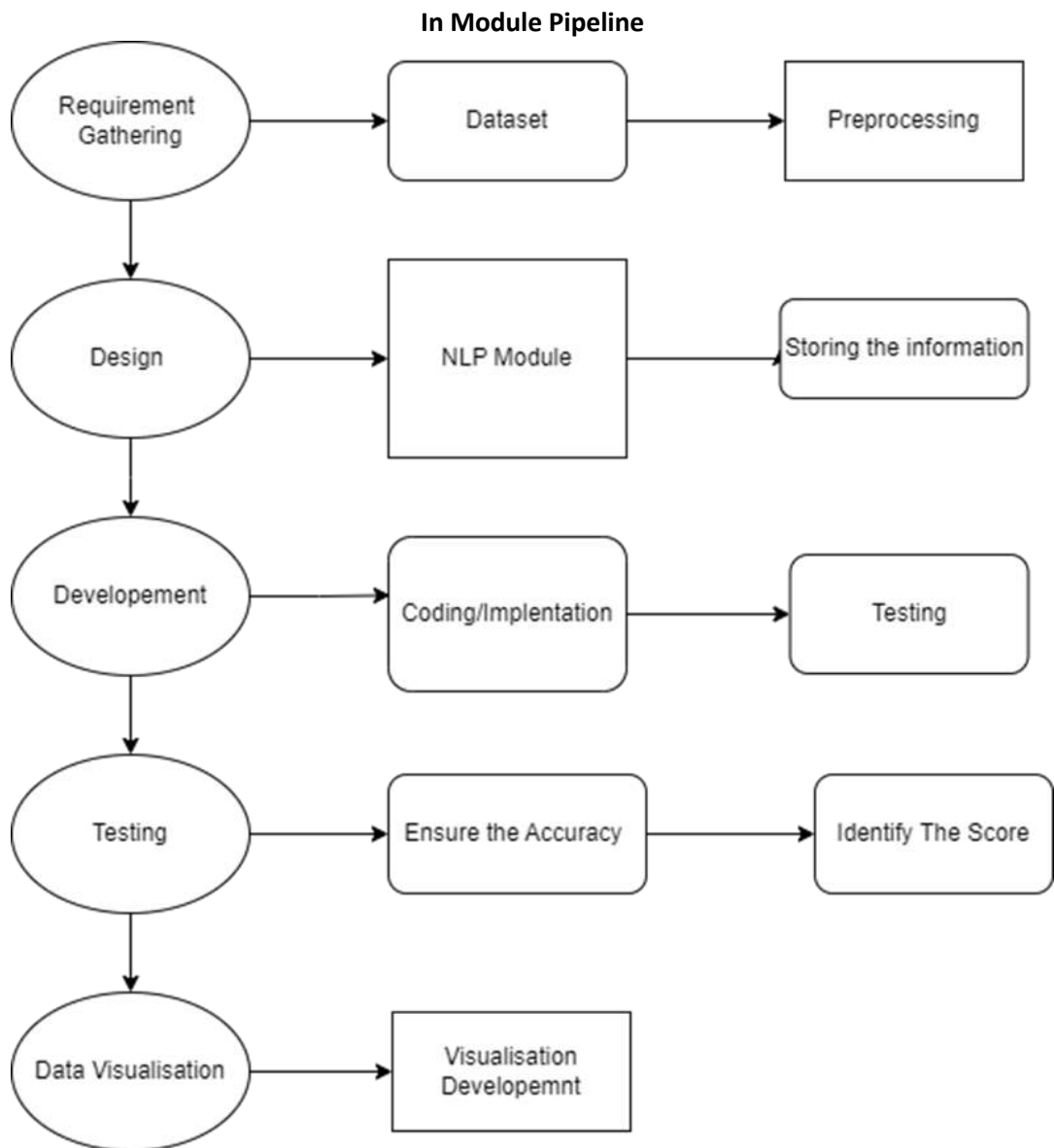
VPCSM Overview



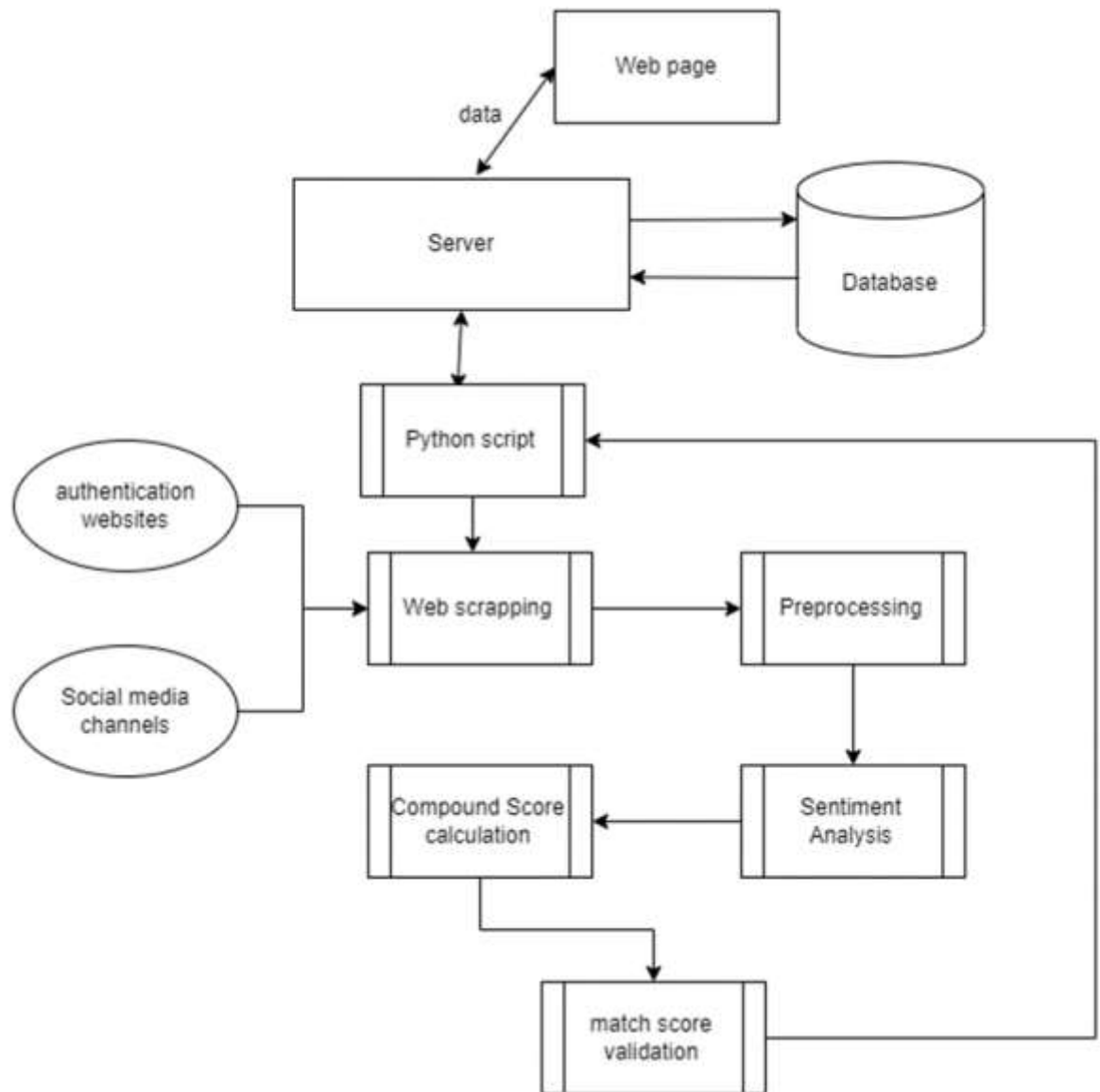
workflow diagram



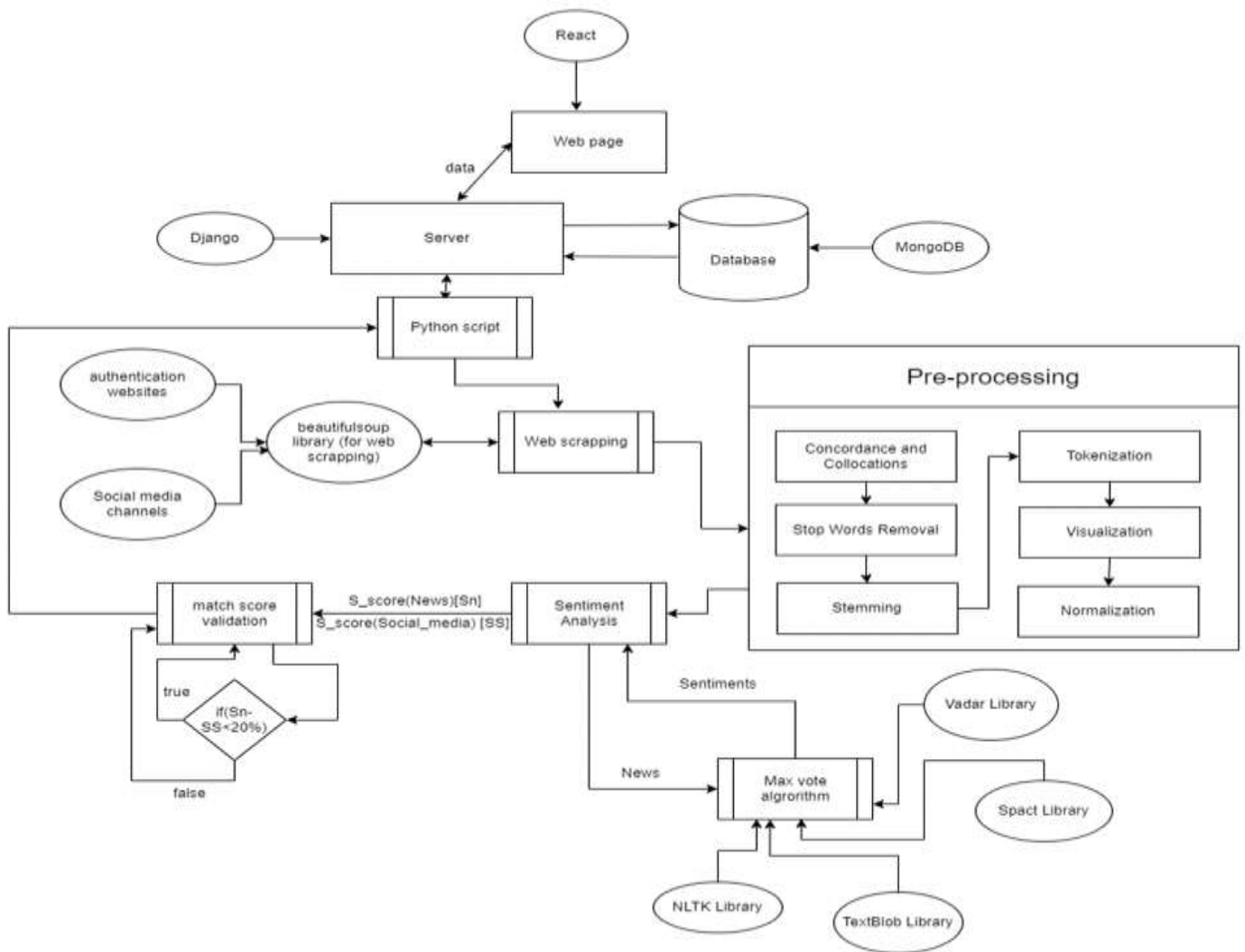
3. Sentiment Analysis Module for Automated News Generator



High Level Diagram

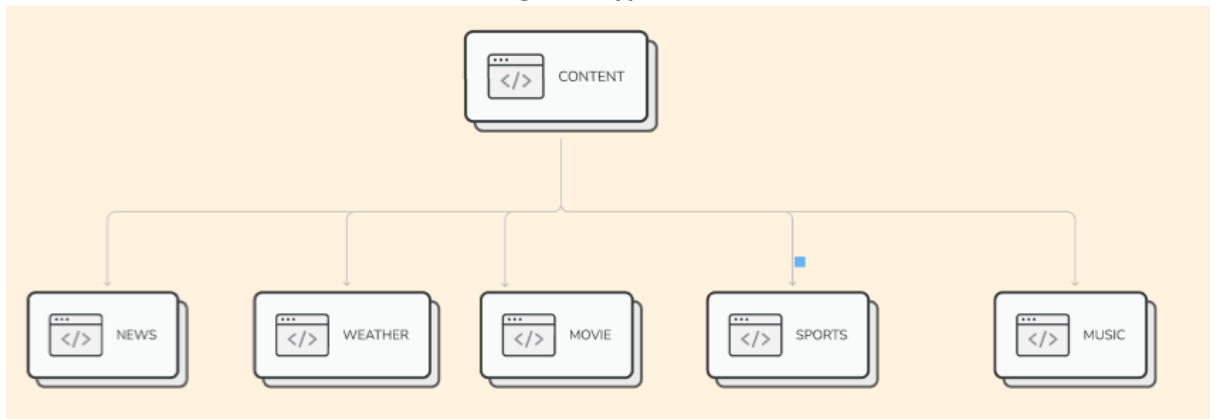


Low Level Diagram

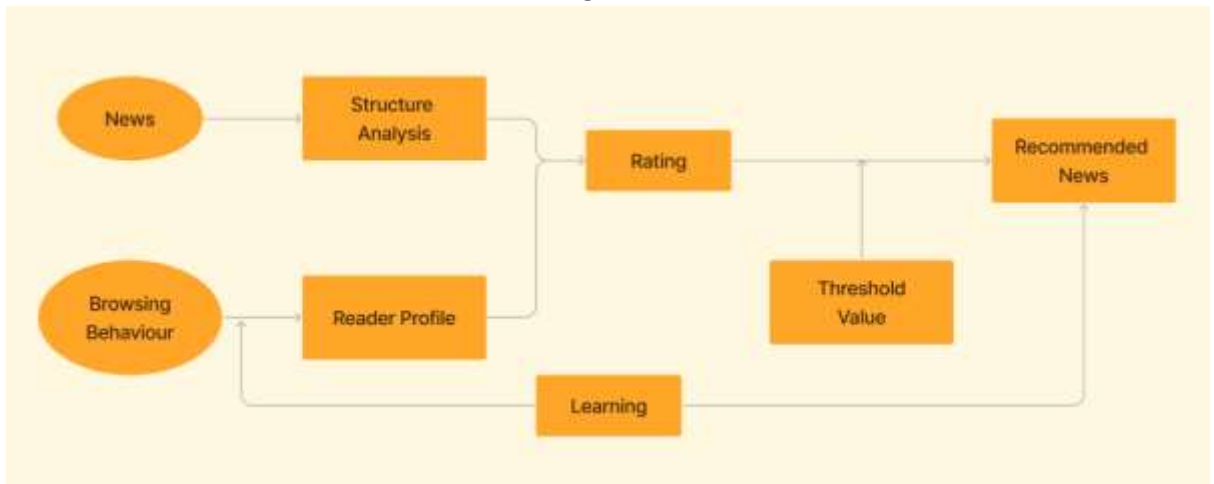


4. Content Personalization

Categories Type

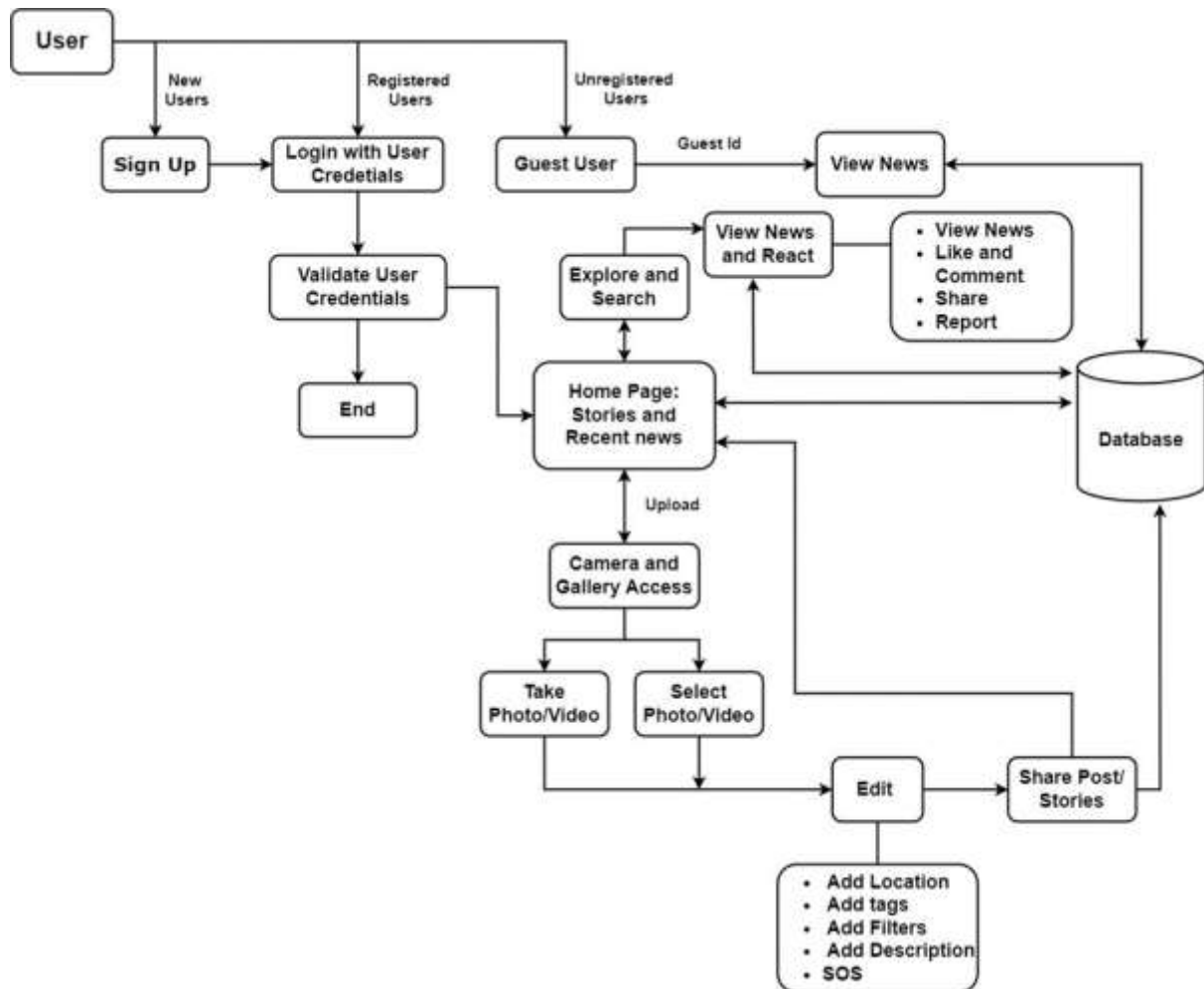


Flow Diagram

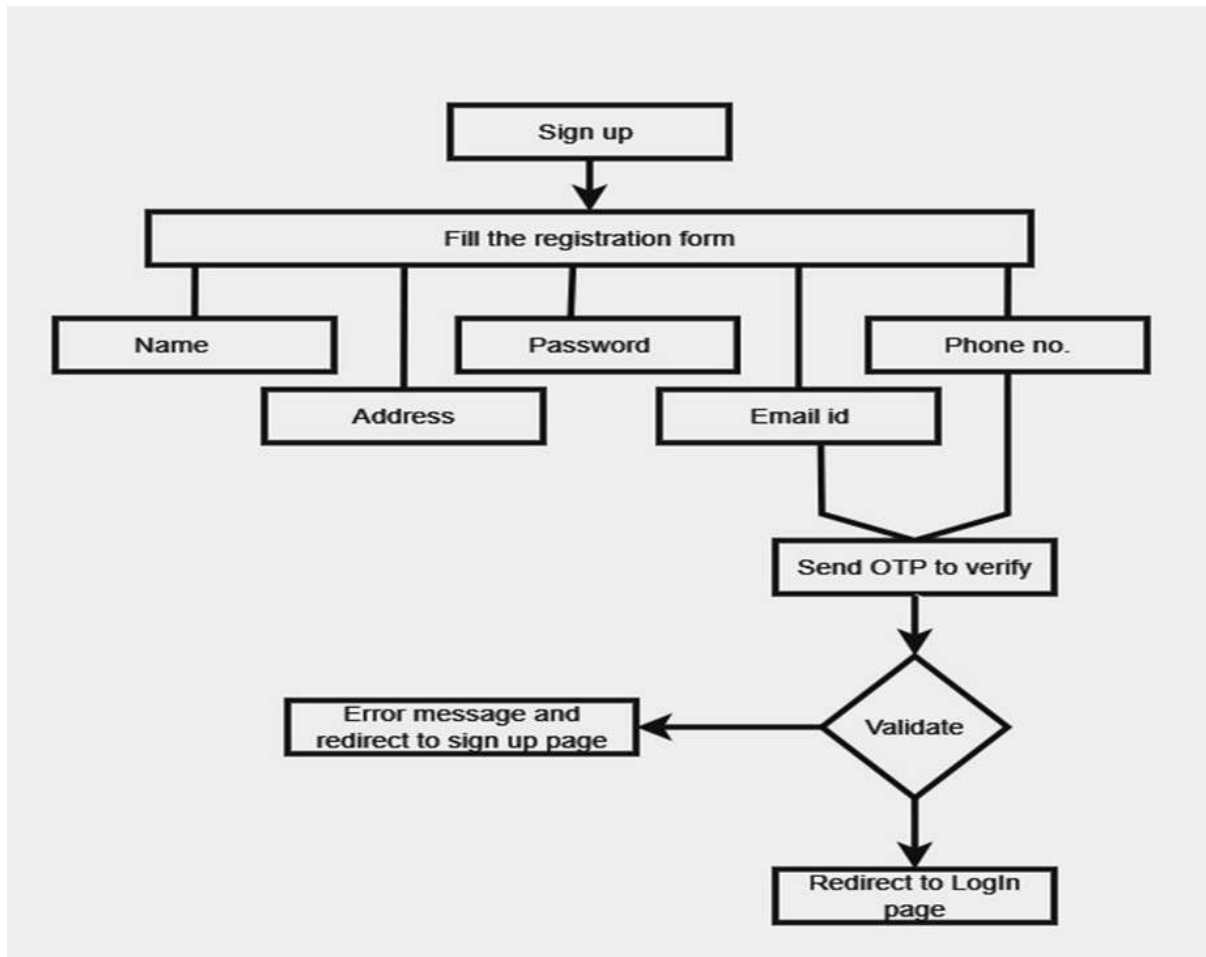
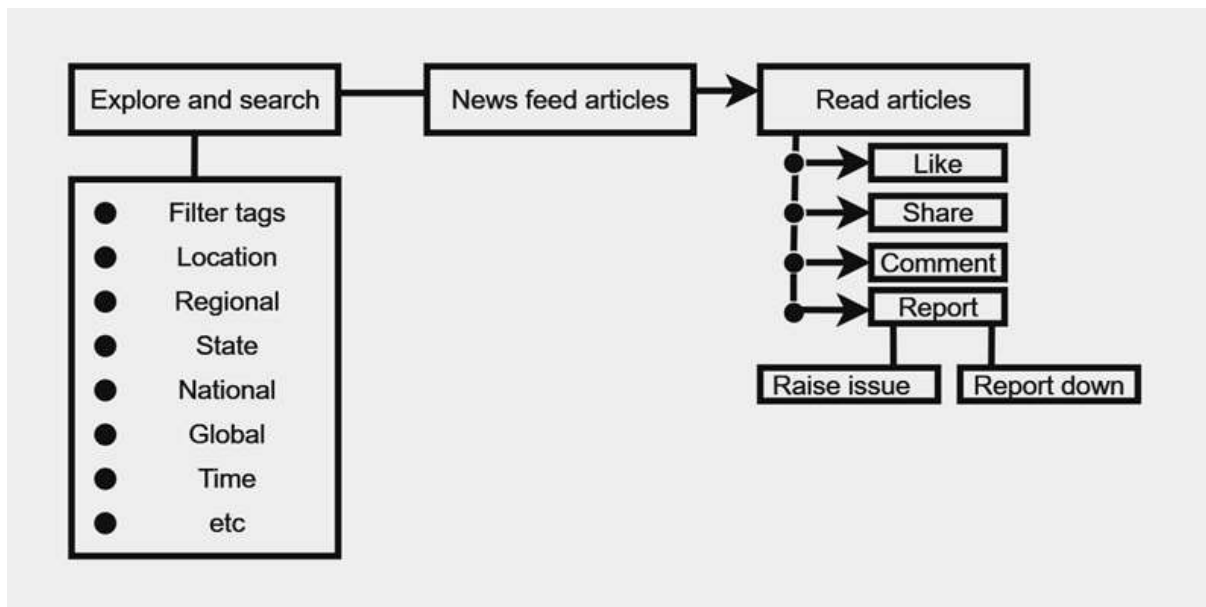


5. Social Media News Application

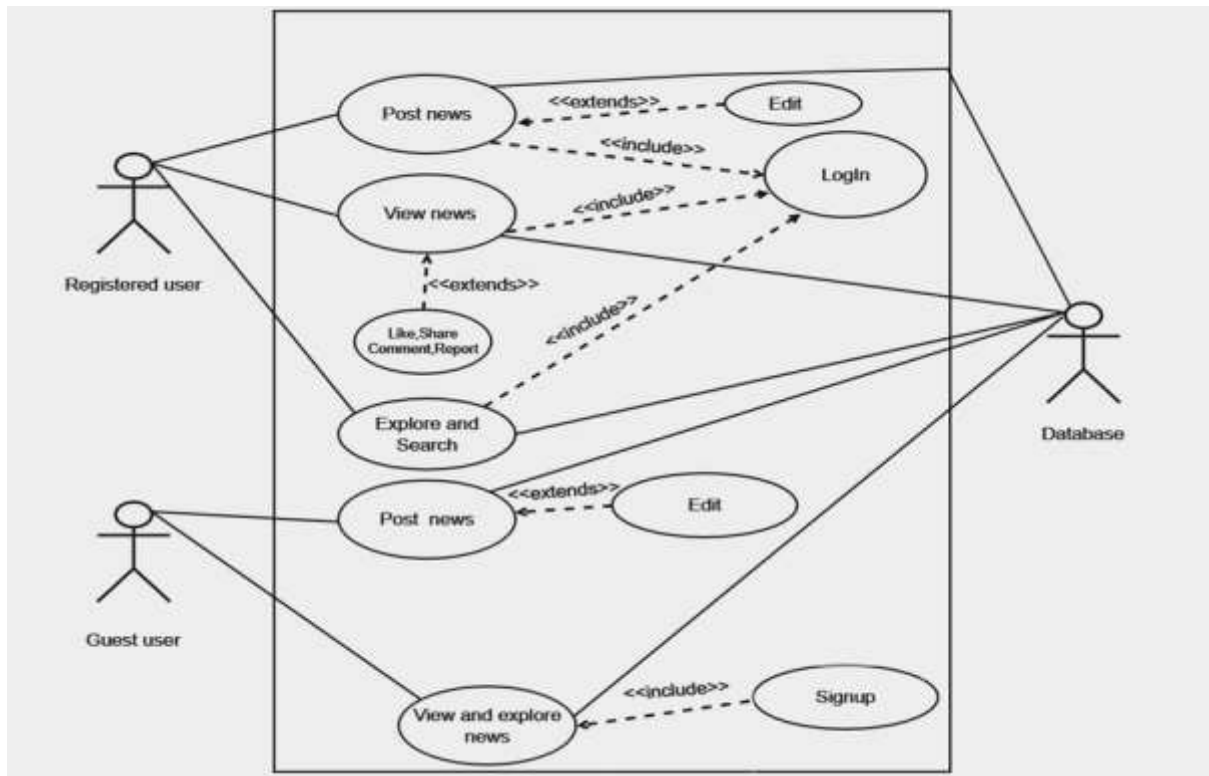
High Level Diagram



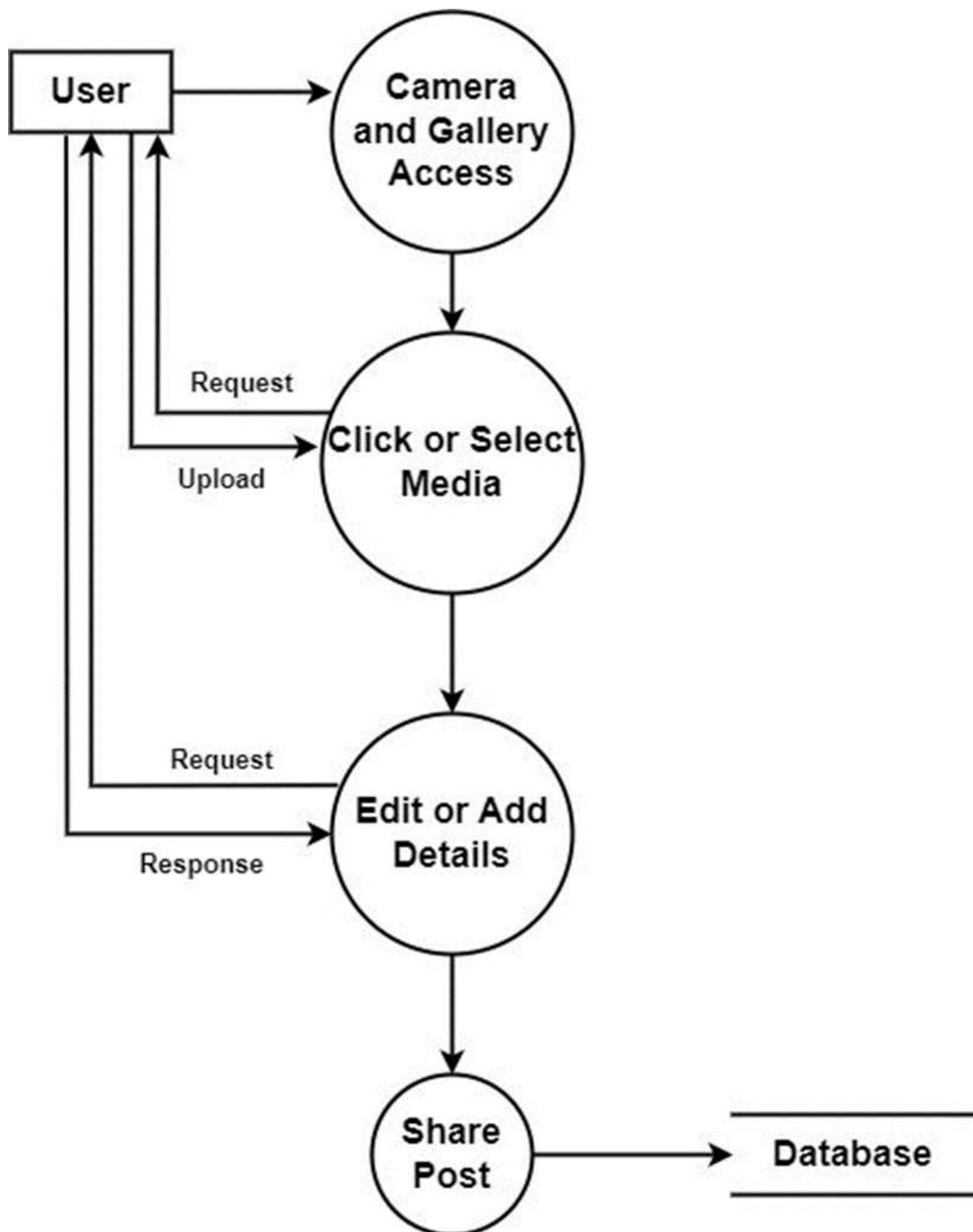
Low Level Diagram

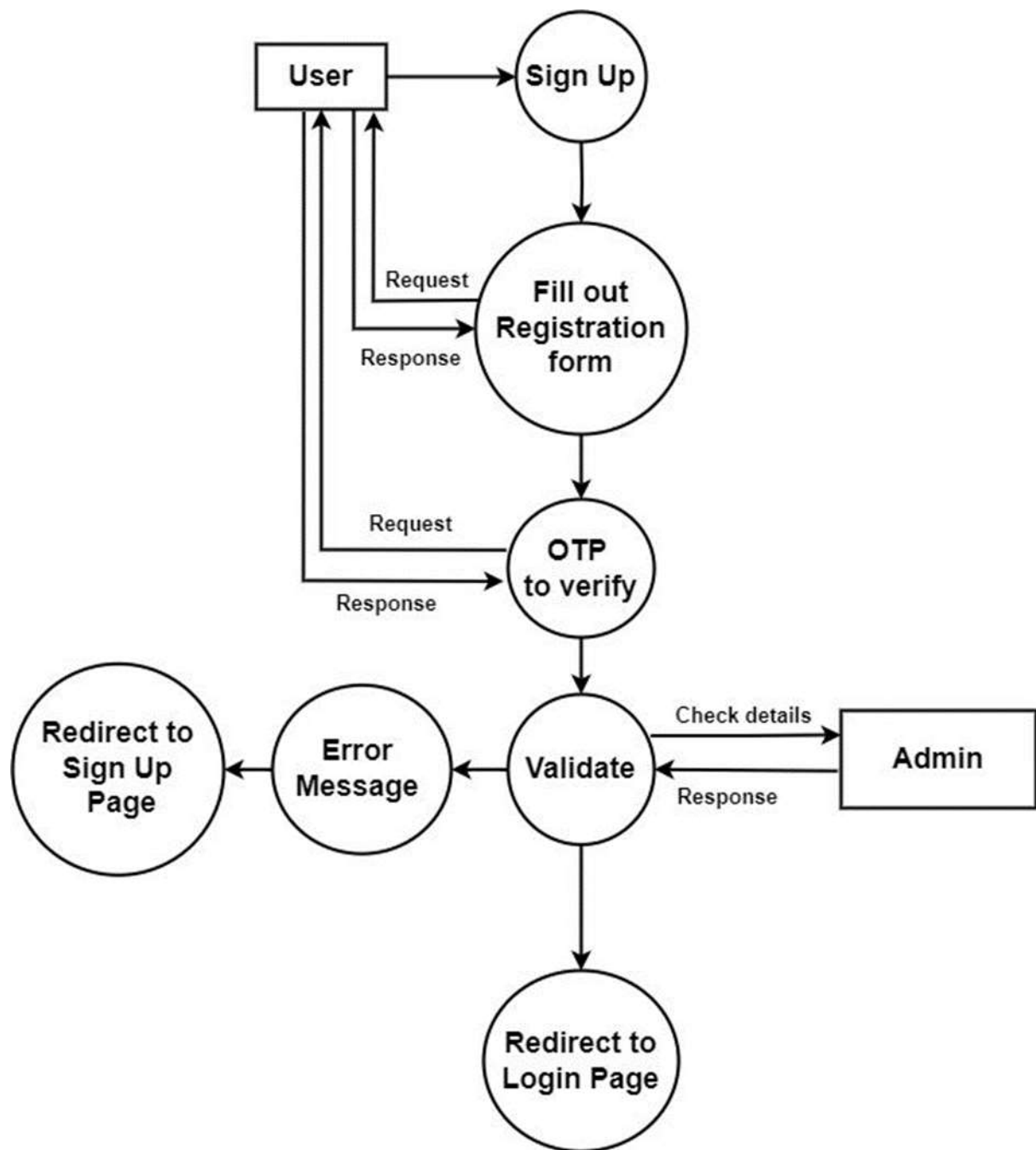


Use Case Diagram



Data Flow Diagram





Link of Mega BRD :

Appendices

Engagement Metrics Model :-

An engagement metrics model is a framework used to measure and analyze the level of audience engagement and interaction with content on social media platforms. These is the model for finding the news that is trending on Twitter or can become breaking news in future.

Points that can be affective :-

1. Total Engagements :-

- a. Tweets or news topics with high total engagement are likely to be more popular and trending. The higher the number of likes, retweets, and replies, the more attention the content is receiving from the Twitter community.
- b. Total engagement provides insights into the level of interest and relevance of a news topic to the audience. Higher engagement suggests that the news resonates with users and holds their attention.
- c. significant surge in total engagement in a short time might indicate that the content is going viral. Viral content tends to spread rapidly through social media platforms and can be a strong indicator of trending news.

2. Like Tweet Ratio :-

- a. Tweets with a high like-to-retweet ratio indicate that the content is resonating strongly with the audience, as more people are choosing to retweet the news rather than simply liking it. Such tweets may be indicative of trending news that has widespread interest and is being shared across different user networks.
- b. Early in the trending process, before news becomes mainstream, tweets with high like-to-retweet ratios might indicate emerging trends. These tweets might be from influential users or thought leaders who first shared the news, and as it gains momentum, more users begin to retweet it.
- c. By examining the engagement patterns, we can get insights into the audience's sentiment towards specific news topics. For example, high retweets but low likes might suggest controversial or divisive news, while high likes and retweets indicate widespread approval.

3. Engagement Rates :-

- a. High engagement rates indicate that the content is resonating well with the audience. When analyzing trending news, tweets or topics with high engagement rates are likely to be more popular and relevant
- b. Low engagement rates may indicate that the content needs improvement or adjustments in targeting the right audience. Analyzing engagement rates helps content creators refine their strategies to create more engaging content.

4. Reply Rates :-

- a. A high reply rate can also indicate that the news topic is generating discussions, debates, and even disagreements among users. Controversial news tends to have a higher reply rate as people share their opinions and thoughts.
- b. Monitoring the reply rate over time allows us to track how news topics are evolving and gaining attention. A sudden increase in the reply rate may indicate that a news topic is becoming more relevant and trending.
- c. A high reply rate may indicate that the news topic is thought-provoking and stimulating discussions among users. Such content often has the potential to go viral and become trending news.

Video Categorization:

❖ OBJECTIVE :

- Video categorization, also known as video content categorization or video classification, is the process of automatically assigning predefined categories or labels to videos based on their content.
- It is an essential task in VPCSM to categorize the video provided by news reporter or user
- News videos can be categorized based on various factors, including their content, format, and intended audience.
- NEWS Videos can be categorized on the following bases.
 - **Subject:** News videos can be categorized based on the main topic or subject they cover. For example, categories could include politics, economy, international relations, sports, entertainment, technology, health, environment, and more.
 - **Geographical Location:** News videos can be categorized by the geographical location they focus on, such as local news, national news, or international news.
 - **Language:** For multilingual news platforms, videos can be categorized by the language in which they are presented, such as English news, Spanish news, or Mandarin news.

- **Emotion:** Videos can be categorized based on their emotional tone, such as serious news, light-hearted news, inspiring stories, or tragic events.

The categorization of news videos can vary depending on the specific needs and goals of the news organization,

Some common categories are:

1. **Politics:** News related to government policies, elections, political events, and activities of politicians and political parties.
2. **Economy:** News about financial markets, business trends, economic indicators, corporate earnings, and economic policies.
3. **International Relations:** News covering global events, diplomatic relations, conflicts, treaties, and international agreements.
4. **Sports:** News related to sporting events, scores, player transfers, team updates, and sports-related controversies.
5. **Entertainment:** News about movies, music, celebrity updates, TV shows, events in the entertainment industry, and award ceremonies.
6. **Technology:** News on advancements in technology, gadget launches, software updates, cybersecurity, and innovations.
7. **Health:** News related to public health, medical research, disease outbreaks, vaccinations, and healthcare policies.
8. **Science:** News covering scientific discoveries, space exploration, research findings, and breakthroughs in various scientific fields.
9. **Education:** News related to education policies, school and university updates, educational trends, and learning innovations.
10. **Crime:** News about criminal activities, law enforcement operations, legal proceedings, and crime prevention efforts.
11. **Weather:** News and updates on weather conditions, forecasts, and weather-related events.
12. **Health and Wellness:** News about public health, medical breakthroughs, healthcare policies, fitness trends, nutrition, and mental health.
13. **Environment and Sustainability:** News about climate change, environmental conservation efforts, natural disasters, renewable energy, and sustainable practices.
14. **Travel and Tourism:** News related to travel destinations, tourism trends, travel advisories, and hospitality industry updates.

15. **Arts and Culture:** News on art exhibitions, cultural events, literary festivals, theater, and other cultural activities.
16. **Human Rights and Social Issues:** News stories highlighting social justice, human rights violations, activism, and social impact initiatives.
17. **Religion and Spirituality:** News about religious events, interfaith dialogues, religious leaders, and spiritual practices.
18. **Law and Justice:** News on legal developments, court proceedings, landmark judgments, and law enforcement activities.
19. **Fashion and Beauty:** News about fashion trends, designer collections, beauty products, and fashion shows.
20. **Business:** News about business ,startup companies, entrepreneurial success stories, and innovative products/services.

❖ ALGORITHMS USED FOR VIDEO CATEGORIZATION

- **Hidden Markov Models (HMM):** HMMs are commonly used for modeling temporal patterns in video data. They can capture sequential information and have been used in tasks like action recognition and gesture recognition within videos.
- **Long Short-Term Memory (LSTM):** LSTM is a type of Recurrent Neural Network (RNN) that is well-suited for sequential data like videos. LSTMs can learn long-range dependencies in video frames, making them useful for tasks requiring temporal understanding, such as action recognition or video captioning.
- **Convolutional Neural Networks (CNN):** CNNs are powerful deep learning models primarily designed for image recognition tasks. In the context of video categorization, one common approach is to extract visual features from individual frames using a pre-trained CNN (e.g., VGG, ResNet, or Inception), and then aggregate the features across frames using techniques like temporal pooling or 3D convolutions to model the video content.
- **Two-Stream Networks:** Two-stream networks combine information from both spatial (frames) and temporal (optical flow) domains to capture spatial and motion information within videos. They often consist of two separate CNNs, one processing frames and the other processing optical flow data.
- **3D Convolutional Neural Networks (3D CNN):** 3D CNNs extend traditional 2D CNNs to process spatiotemporal data directly. They take both spatial and temporal dimensions into account simultaneously, making them suitable for video categorization tasks where motion information is crucial.

❖ Potential Use Cases

- **Personalized News Recommendations:** Utilize sentiment analysis results to provide personalized news recommendations to users based on their emotional preferences. For example, if a user tends to prefer positive news, the system can recommend news articles with a positive sentiment.
- **Emotional Context for News Headlines:** Add emotional context to news headlines by incorporating sentiment analysis insights. This will help create headlines that accurately convey the emotional tone of the news content.
- **Filtering Negative or Controversial Content:** Implement a feature to filter out news articles with negative or controversial sentiments. This ensures that the generated news content is balanced and appropriate for the target range of news content. For example, the system can offer categories like "Positive News," "Neutral Updates," and "Critical Analysis."
- **Trending Topics Analysis**et audience.
- **Real-time Sentiment Analysis for Breaking News:** Perform real-time sentiment analysis on trending news articles, especially during breaking news events, to quickly gauge public reactions and emotions related to the news.
- **Sentiment-based Content Categorization:** Categorize news articles based on their sentiment to provide users with a. Analyze sentiment across different trending topics to identify popular and emotionally charged subjects. This information can help media organizations prioritize their news coverage.
- **Emotional Tone in Video Content:** Apply sentiment analysis to video transcripts to understand the emotional tone of video content. The module can categorize videos based on their sentiment, enabling users to watch videos with emotional context.