IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

InstaViral AI: A Deep Learning-Based System for Intelligent Video Analysis and Content Optimization in Instagram Reels

[1]Prof. Madhuri Kale, [2]Dnyaneshwari Dandagawhal, [3]Shravani Jadhav,

[4]Shubham Ghodele,[5]Vaishnavi Maske

[1]Professor, Department of Computer Engineering, DPCOE, Pune-412207, India

[2,3,4,5]Student, Department of Computer Engineering, DPCOE, Pune-412207, India

Abstract: Because short-form video platforms like Instagram Reels are growing so quickly, content producers frequently find it difficult to forecast the performance of their work and lack sophisticated pre-publication optimization tools. Because post-publication measurements are the main emphasis of current analytics, this leads to a crucial gap in the digital content lifecycle. The design and methods for "InstaViral AI"—an intelligent video analyzer and content optimizer for Instagram reels—are presented in this research. The suggested system analyses multimodal data inputs, such as video frames, audio, captions, and user metadata, by utilizing cutting-edge deep learning and machine learning approaches InstaViral AI will automatically recognize interesting highlights, find the best thumbnail frames, and intelligently recommend enhancements like trending hashtags, better captions, the best times to upload, and the length of the material. InstaViral AI seeks to function as a smart assistant by fusing sentiment analysis, video summarization, and predictive analytics. This will improve content discoverability and lessen the trial-and-error process that is usually involved with social media material.

Keywords- Artificial Intelligence, Video Analysis, Deep Learning, Popularity Prediction, Content Optimization, Social Media Analysis

I. INTRODUCTION

A. The Opportunity (Growth and Challenges)

With the explosive growth of short-form video platforms, particularly Instagram Reels, the number of ways you can use social media to communicate digitally is now at its highest level ever. However, the sheer amount of video content being published daily creates a major issue for both individual creators and marketers: increasing the visibility of your content so that it gains the maximum possible engagement — also referred to as going "viral." While Reels are becoming increasingly popular, many content creators have difficulty predicting how well their content will do because they don't have enough reliable, data-driven information about what works best when creating content. Because they rely heavily on trial-and-error, there is a great opportunity to develop an intelligent way to optimize content. The formatter will be responsible for producing these features, including the specific requirements noted below.

B.The Problem and Gap (Shortcomings of Existing Analytical Platforms)

The primary problem is that existing analytical platforms are unable to meet today's needs in terms of the types of analytics available. Current analytics platforms are focused on post-publishing performance metrics; they provide historical data regarding views and engagement once content has been published. These platforms offer little support during the content creation and optimisation phases; therefore, content creators lack access to intelligent tools that assist them in making improvements to their videos prior to publication. As such, there exists a large gap in the digital content life cycle where creators are struggling to compare their own content against viral trends and receive real-time, AI-based feedback to improve upon aspects of their content such as expression, audio, captions and editing [2], [3].

C.The Proposal (InstaViral AI Solution)

To bridge this crucial gap, this paper proposes **"InstaViral AI: "Intelligent Video Content Optimizer & Video Analyzer for Instagram Reel Content**". InstaViral's AI is a smart content discovery and social media assistant built with AI to optimize user engagement and content visibility. InstaViral's AI uses advanced Deep Learning/Machine Learning techniques to evaluate and assess multi-modal inputted data types (video frame visuals, audio emotion/sentiment, caption/hashtag text)**.** The objective of this paper is to detail the deep learning-based architecture and methodology for video analysis, popularity prediction, and **pre-publication content optimization**, presenting a blueprint for the system's implementation

II. Relevant Background and Work

Established studies in computer vision, natural language processing (NLP), and social media predictive a nalytics served as the foundation for the creation of InstaViral AI.
This section examines the fundamental technological underpinnings that enable the operation of the sugg ested system.

A. Analysis of Video Content

A strong framework for automatically extracting significant information from the visual and temporal c omponents of the video stream is necessary for effective content optimization for short-form video. Techniques from Computer Vision (CV) are largely responsible for this capability.

- Understanding scenes, detecting faces and objects, and measuring human-centric features like posture, gesture, and emotion are all part of CV's function in content analysis. To find interesting moments and evaluate a user's video quality versus successful material, it is ess ential to analyse these qualities.
- Predicting viewer engagement also requires examining particular human characteristics, such as bo dy language, emotional state, and facial expressions.
  To accurately and quickly measure the energy and appeal of the content creator's performance in th e reel, libraries such as MediaPipe are used to evaluate human movement and facial expressions. According to recent research, this field can reach high prediction accuracy (up to 74.7\%) by mergi ng AI and deep learning [1].

B.Predictive Social Media Analytics

Predictive analytics is the foundation of InstaViral AI's ultimate objective, which is to forecast content success and offer practical optimisation methods.

- Forecasting the reach and engagement (likes, shares, and comments) of content based on its charact eristics and context is the focus of the current research field known as **"Popularity Prediction".** To do this, it is necessary to establish a relationship between the qualities of the video, visual appeal, audio quality, and textual relevance and the patterns of audience involvement that follow [1].
- In a similar vein, trend detection is the process of spotting contemporary trends in popular social m edia content, like popular audio tracks, editing styles, or hot subjects.

This detecting feature will be incorporated into our system to make sure that recommended enhanc ements match the virality characteristics of the present platform

- A multimodal data approach is required to get high predicted accuracy [3]. To create a thorough predictor of possible virality, the model must incorporate quantitative visual features, textual features (from Sentiment Analysis and Text Summarization of captions and audio transcript [2]), and metadata (like the best times to post and the length of the video) [3]. The foundation of predictive modeling with machine learning tools like Sci-Kit is the integration of various data types.

C. The Deficit in Architecture

Although the fundamental elements covered in this section, popularity modelling algorithms and sophisticated video analysis tools, exist separately [1], [2], they do not come together to provide a cohesive pre-publication optimisation solution for content producers. The majority of the tools now in use concentrate on either post-publication performance measures or simple, low-level analysis. The absence of a unified, user-friendly system that can:

1. Instantaneously analyse user content is the architectural gap.
2. Examine it in relation to statistics on viral trends.
3. Condense the findings into practical pre-publication recommendations (such as better captions and perfect thumbnails).

By combining these various technologies into a single, integrated system, InstaViral AI bridges this gap by providing video producers with data-driven, intelligent optimisation prior to upload.

III. Methodology

A. Data Collection and Input Processing

The proposed system, InstaViral AI, begins by accepting an Instagram Reel video uploaded by the user through the Streamlit-based frontend interface. Along with the video, the user selects a content niche/category and can also provide a viral reference video URL for comparison. The uploaded video is stored temporarily on the server for processing.

The system extracts:

- Video Frames
- Audio Stream
- Metadata such as video duration and frame count

B.Multimodal Video Analysis Pipeline

The Project uses a multi-modal AI Pipeline that combines audio analysis, visual analysis, emotion detection, and video structure analysis to evaluate the viral potential of the content.

1. Video Processing

The video_processor.py module processes the uploaded video using OpenCV and PySceneDetect. The methodology includes:

- Extracting key frames from the video
- Detecting scene transitions
- Calculating motion intensity using optical flow
- Measuring colour vibrancy using HSV histograms
- Detecting background music using BPM analysis

2. Audio and Speech Analysis

The audio_analyzer.py module performs multilingual speech analysis using Groq Whisper large-3.

The methodology follows a two-pass approach:
1. Speech transcription in the original language (Hindi, Marathi, or English)
2. Translation into English for AI-based scoring and understanding.

The system further analyses:
1. Audio sentiment
2. Hook strength
3. Call-to-action (CTA) detection
4. Speech clarity

A hybrid rule-based and LLM -based technique is used to detect Indian-language engagement phrases such as “share karo” and “follow kara”.

3. Visual Frame Analysis

The vision_analyzer.py module sends selected key frames to the Groq Llama-4-Scout Vision model for detailed analysis.

The methodology evaluates:
1. Lighting quality
2. Visual energy
3. Frame composition
4. Text overlay presence
5. Colour vibrancy

4. Emotion Detection

The emotion_analyzer.py module performs facial emotion recognition using OpenCV Haar Cascade and EmotiEffLib ONNX models.

The methodology includes:

- Detecting human faces from frames
- Classifying emotions frame-by-frame
- Computing:
  - Smile ratio
  - Positive emotion score
  - Emotion variance

5. Feature Extraction and Virality Scoring

After all analyses are completed, the feature_extractor.py module combines outputs into a normalized 14-dimensional feature vector.

The features include:

- Audio sentiment
- Hook strength
- CTA presence
- Visual energy
- Lighting quality

- Motion intensity
- Smile ratio
- Music presence
- Emotion variance
- Scene cuts per minute

The final Viral Readiness Score is calculated using cosine similarity between the user video vector and stored viral benchmark vectors.

The cosine similarity formula is:

$$\textbf{Cosine Similarity} = \frac{A.B}{\|A\|\|B\|}$$

Where:

- A = User video feature vector
- B = Viral benchmark feature vector

The resulting similarity score is scaled between 0 to 100.

6. AI-Based Suggestion Generation

The suggestion_generator.py module uses Groq Llama-3.3-70B and Llama-3.1-8B models to generate improvement recommendations.

The methodology generates:

- Detailed improvement reports
- Optimized captions
- Hashtag recommendations
- Best posting time suggestions

The generated suggestions are personalized according to:

- Detected language
- Regional culture
- Content category
- Performance gaps identified during analysis

7. Database Storage and Result Generation

All analysis results are stored in MongoDB collections including:

- analysis_jobs
- analysis_results
- viral_benchmarks

The backend FastAPI service manages:

- Video upload
- Analysis triggering
- Result retrieval
- Benchmark management

The final output is displayed through the Streamlit dashboard using:

- Virality score gauges
- Radar charts
- Emotion timelines
- Gap analysis tables
- AI-generated reports

IV. System Design:

A. System Architecture:

The InstaViral AI system follows a modular multi-tier architecture consisting of the frontend layer, backend layer, AI analysis pipeline, database layer, file system, and external AI services. The architecture is designed to support scalable, asynchronous, and multi-modal video analysis for Instagram Reels.
The complete workflow begins when a user uploads a video through the Streamlit frontend. The request is processed by the FastAPI backend, which triggers the AI analysis pipeline in the background. The generated analysis results are stored in MongoDB and displayed to the user through an interactive dashboard.

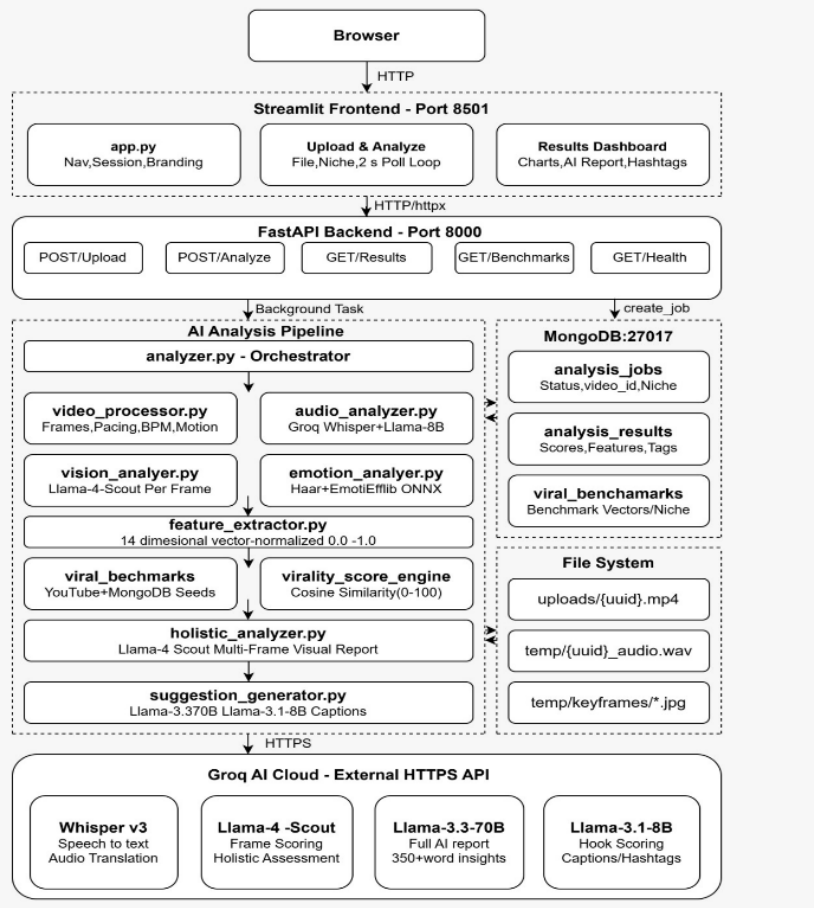

Fig. 1. System Architecture

1. Frontend Layer

The frontend of InstaViral AI is developed using Streamlit and operates on Port 8501. It provides a simple and interactive interface where users can upload Instagram Reel videos, choose content categories, and track the analysis process. The dashboard displays virality scores, charts, AI-generated reports, captions, hashtags, and other insights in a user-friendly manner. The frontend mainly includes app.py for navigation and session handling, the Upload & Analyze module for video submission, and the Results Dashboard for displaying outputs. Communication between the frontend and backend takes place through HTTP/HTTPS requests.

2. Backend Layer

The backend is implemented using FastAPI and runs on Port 8000. It serves as the communication bridge between the frontend and AI processing modules. REST APIs are used for video uploading, triggering analysis, retrieving results, accessing benchmark data, and monitoring system health. The backend also supports asynchronous background processing, enabling efficient handling of large video files without affecting frontend performance.

3. AI Analysis Pipeline

The AI Analysis Pipeline is the core component of the system and is managed by analyzer.py, which acts as the orchestrator. The video_processor.py module performs frame extraction, scene detection, motion analysis, and BPM detection using OpenCV and PySceneDetect. The audio_analyzer.py module uses Groq Whisper and Llama models for speech transcription, translation, sentiment analysis, and hook detection. The vision_analyzer.py module uses the Llama-4-Scout Vision model to evaluate lighting quality, frame composition, visual energy, and colour vibrancy. The emotion_analyzer.py module detects facial emotions using Haar Cascade and EmotiEffLib ONNX models. All extracted features are combined into a normalized 14-dimensional feature vector by feature_extractor.py, and the virality_score_engine calculates the Viral Readiness Score using cosine similarity. The holistic_analyzer.py and suggestion_generator.py modules generate AI-based reports, captions, hashtags, and posting-time recommendations.

4. Database and Storage

MongoDB is used as the primary database for storing analysis jobs, viral benchmark vectors, and final analysis results. Collections such as analysis_jobs, analysis_results, and viral_benchmarks help manage data efficiently. The local file system temporarily stores uploaded videos, extracted audio files, and keyframes during processing.

5. External AI Services

The system integrates Groq AI cloud services through HTTPS APIs. Whisper v3 is used for speech recognition and translation, Llama-4-Scout for visual analysis, Llama-3.3-70B for AI report generation, and Llama-3.1-8B for caption and hashtag generation. Overall, the architecture enables InstaViral AI to perform accurate multi-modal analysis and provide personalized optimization suggestions for Instagram Reel creators.

B. UML Diagrams

1. Use Case Diagram:

The use case diagram of InstaViral AI represents the interaction between users and the AI-based system. It illustrates how the content creator uploads Reel videos for analysis and how the system generates optimization suggestions using artificial intelligence techniques.

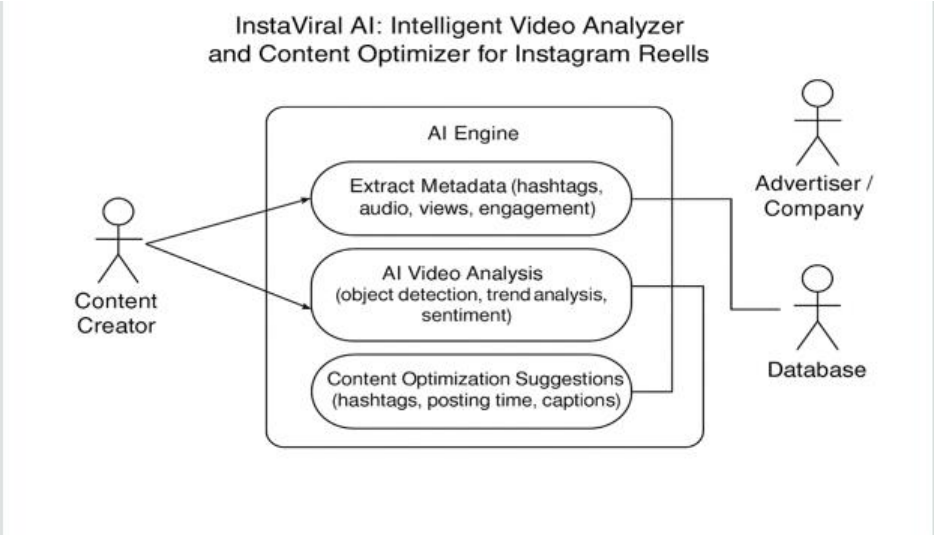

Fig. 2. Use Case Diagram of InstaViral AI

.The diagram shows that the Content Creator interacts with the AI engine by uploading Instagram Reels. The system extracts metadata, performs video analysis, and generates optimization suggestions such as captions, hashtags, and posting recommendations. Advertisers or companies can use these insights for marketing purposes, while the database stores analysis results and related information.

2. Class Diagram:

The class diagram of InstaViral AI represents the structural design of the system and shows the relationship between the main classes involved in the video analysis process. The diagram mainly includes the InstaViral AI, Video Analyzer, User, Optimized Reel, and Database classes.

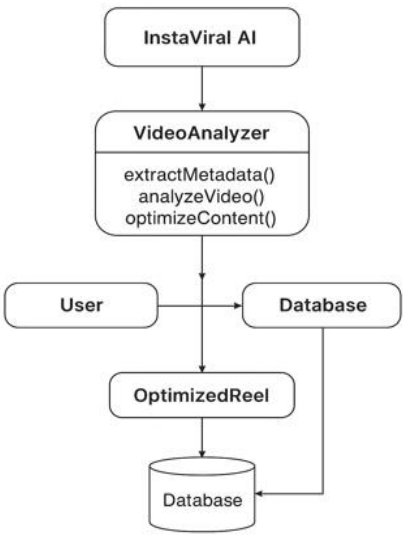

Fig. 3. Class Diagram of InstaViral AI

The Video Analyzer class acts as the core processing unit of the system and contains important functions such as extractMetadata(), analyzeVideo(), and optimizeContent(). These methods are responsible for extracting video information, performing AI-based analysis, and generating optimized content suggestions.

The User class interacts with the system by uploading Instagram Reel videos for analysis. The processed output is generated in the form of an Optimized Reel, which contains improved captions, hashtags, and engagement recommendations.

The Database class stores user data, analysis results, and optimized content information for future retrieval and processing. Overall, the class diagram illustrates the interaction between different system components and represents the internal structure of the InstaViral AI application.

3. Activity Diagram:

The activity diagram of InstaViral AI illustrates the step-by-step workflow followed by the system during Instagram Reel analysis and optimization. It represents the sequence of activities performed from video upload to result storage.

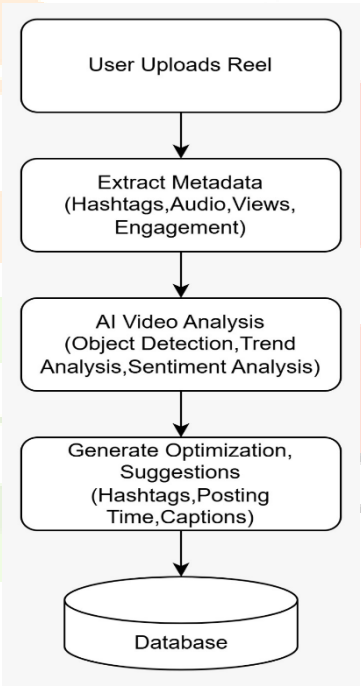

Fig. 4. Activity Diagram of InstaViral AI

The process begins when the user uploads an Instagram Reel to the system. After receiving the video, the system extracts important metadata such as hashtags, audio information, views, and engagement details.

Next, the AI Video Analysis module performs detailed analysis using techniques such as object detection, trend analysis, and sentiment analysis to evaluate the viral potential of the content. Based on the analysis results, the system generates optimization suggestions including recommended hashtags, captions, and ideal posting times.

Finally, the generated analysis results and optimization data are stored in the database for future retrieval and processing. The activity diagram therefore represents the complete workflow of the InstaViral AI system in a sequential manner.

4. Component Diagram:

The component diagram of InstaViral AI illustrates the interaction between the major software modules of the system. The Streamlit-based User Interface communicates with the FastAPI Backend API, which manages requests and controls the overall workflow.

The backend interacts with the AI Processing Module containing the Video Analyzer, Audio Analyzer, Caption and Hashtag Analyzer, and ML models for Reel analysis. MongoDB is used to store analysis data and benchmark information, while the Recommendation Engine generates optimization suggestions such as captions, hashtags, and posting recommendations.

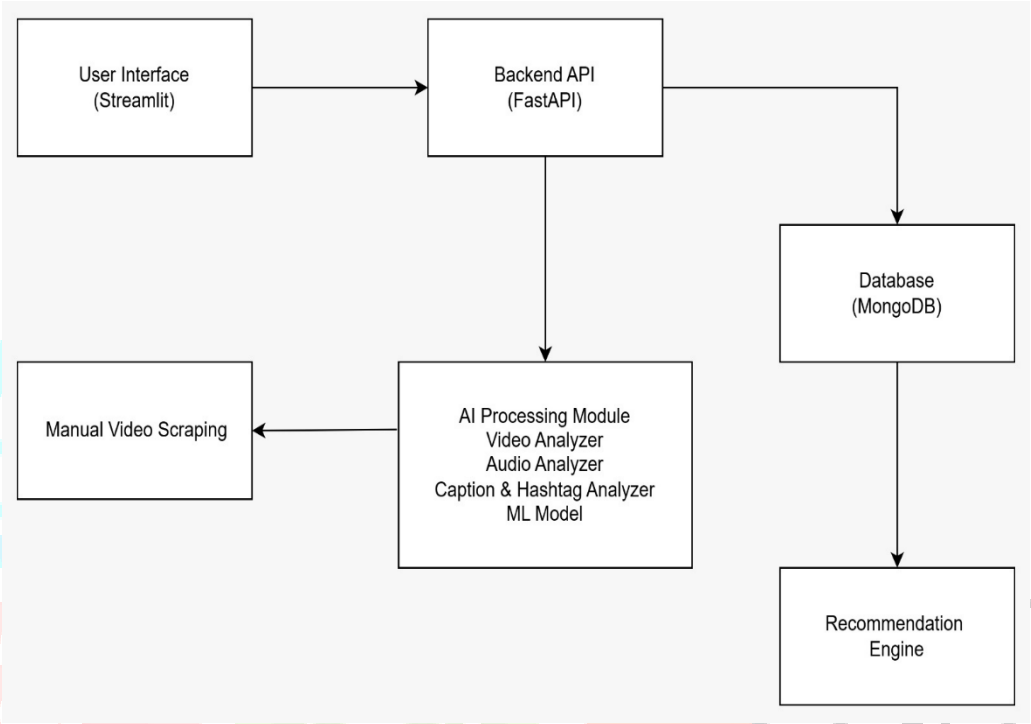

Fig. 5. Component Diagram of InstaViral AI

V. Implementation:

The implementation of InstaViral AI was carried out using a modular approach to ensure efficient processing and accurate analysis of Instagram Reel videos. The system integrates multiple technologies and AI models to perform video analysis, emotion detection, speech processing, and content optimization.

1.1 Presentation Layer — Streamlit Frontend:

The frontend is built using Streamlit and consists of two pages:

— Video upload, niche selection, viral URL input, progress polling.
Upload & Analyze Page

— Virality score gauge, radar chart, emotion timeline, gap table, AI report, caption and hashtags.
Results Dashboard Page

1.2 Application Layer — FastAPI Backend:

The backend exposes a REST API with four router modules:

Endpoint	Method	Purpose
/api/video/upload	POST	Accept and save uploaded video file
/api/analyze	POST	Trigger full AI analysis pipeline in background
/api/results/{video_id}	GET	Poll for analysis status and results
/api/benchmarks/{category}	GET	Retrieve viral benchmark data by niche
/health	GET	Backend health check

1.3 Data Layer — MongoDB:

Three collections are used:

Collection	Purpose	Key Fields
analysis_jobs	Track job status per video	video_id, status, niche, created_at
analysis_results	Store full analysis output	virality_score, feature_vector, suggestions, caption, hashtags
viral_benchmarks	Store viral reference vectors	category, feature_vector, video_url

2. Technology Stack:

Category	Technology	Version	Purpose
Frontend	Streamlit	1.56.0	Web UI framework
Frontend	Plotly	6.7.0	Interactive charts (gauge, radar, line)
Backend	FastAPI	0.115.14	REST API framework
Backend	Uvicorn	0.35.0	ASGI production server
Database	MongoDB	7.0	NoSQL document database
Database	Motor	3.7.1	Async MongoDB driver
AI — Audio	Groq Whisper large-v3	—	Speech-to-text + translation
AI — Audio	Groq Llama-3.1-8B	—	Hook/sentiment/CTA scoring
AI — Vision	Groq Llama-4-Scout	—	Frame visual analysis
AI — Report	Groq Llama-3.3-70B	—	Improvement report generation
Video	OpenCV	4.13.0.92	Frame extraction, optical flow, HSV
Video	PySceneDetect	0.6.7.1	Scene cut detection
Video	yt-dlp	2026.3.17	Viral reference video download
Audio	librosa	0.11.0	BPM and music detection
Emotion	EmotiEffLib (ONNX)	1.1.1	Facial emotion classification
ML	scikit-learn	1.8.0	Cosine similarity computation

3. The 14-Feature Virality Vector:

Every video is represented as a 14-dimensional normalised vector (values 0.0 to 1.0). The Viral Readiness Score is computed as the cosine similarity between the user's vector and the average viral benchmark vector, scaled to 0–100%

#	Feature	Source	Range
0	Audio Sentiment	Groq Llama-3.1-8B	0–1
1	Hook Strength	Groq Llama-3.1-8B	0–1
2	CTA Present	Hybrid LLM + Rule-based	0 or 1
3	Visual Energy	Groq Llama-4-Scout	0–1
4	Lighting Quality	Groq Llama-4-Scout	0–1
5	Frame Composition	Groq Llama-4-Scout	0–1
6	Colour Vibrancy	OpenCV HSV + Groq Vision	0–1
7	On-Screen Text	Groq Llama-4-Scout	0 or 1
8	Positive Emotion	EmotiEffLib ONNX	0–1
9	Emotion Variance	EmotiEffLib ONNX	0–1
10	Smile Ratio	EmotiEffLib ONNX	0–1
11	Cuts Per Minute	PySceneDetect	0–1
12	Motion Intensity	OpenCV Optical Flow	0–1
13	Music Present	librosa BPM	0 or 1

4. Multilingual Support:

One of the key differentiators of InstaViral AI is its deep support for Indian languages. The system handles Hindi, Marathi, Telugu, Tamil, Kannada, Gujarati, Punjabi, and English.

- Groq Whisper large-v3 auto-detects language without any manual selection.
- Two-pass strategy: native transcript shown on dashboard; English translation used for LLM scoring.
- CTA detection covers 30+ Hindi and Marathi phrases (e.g. 'follow kara', 'share karo', 'dekho').
- AI suggestions adapt tone and examples to the detected language and regional culture.
- Captions and hashtags generated in the creator's language (Marathlish / Hinglish / English).

Language-specific hashtag packs included (e.g. #marathireels, #hindireels, #maharashtra).

5. System Output:

Output	Description
Viral Readiness Score	0–100% score computed via cosine similarity against viral benchmarks
Score Label	Strong / Good / Fair / Needs Work with specific next-step action
Feature Radar Chart	14-feature spider diagram — your video vs viral target
Emotion Timeline	Line chart showing emotion positivity across the video duration
Feature Gap Table	Per-feature breakdown with Critical / Weak / OK / Strong status
AI Improvement Report	350+ word structured report from Groq Llama-3.3-70B (5 sections)
Holistic Assessment	Frame-specific visual observations from Groq Llama-4-Scout Vision
Suggested Caption	Ready-to-post caption in creator's language (Hinglish/Marathlish)
Hashtag Pack	15 optimised hashtags including Indian niche and regional tags

Best Posting Time	Recommended day and time for Indian Instagram audience
Full Transcript	Original language + English translation side by side
Scene Analysis	Per-frame description of what AI sees in each key frame

VI. Results and Analysis:

1. Homepage of Instaviral AI:

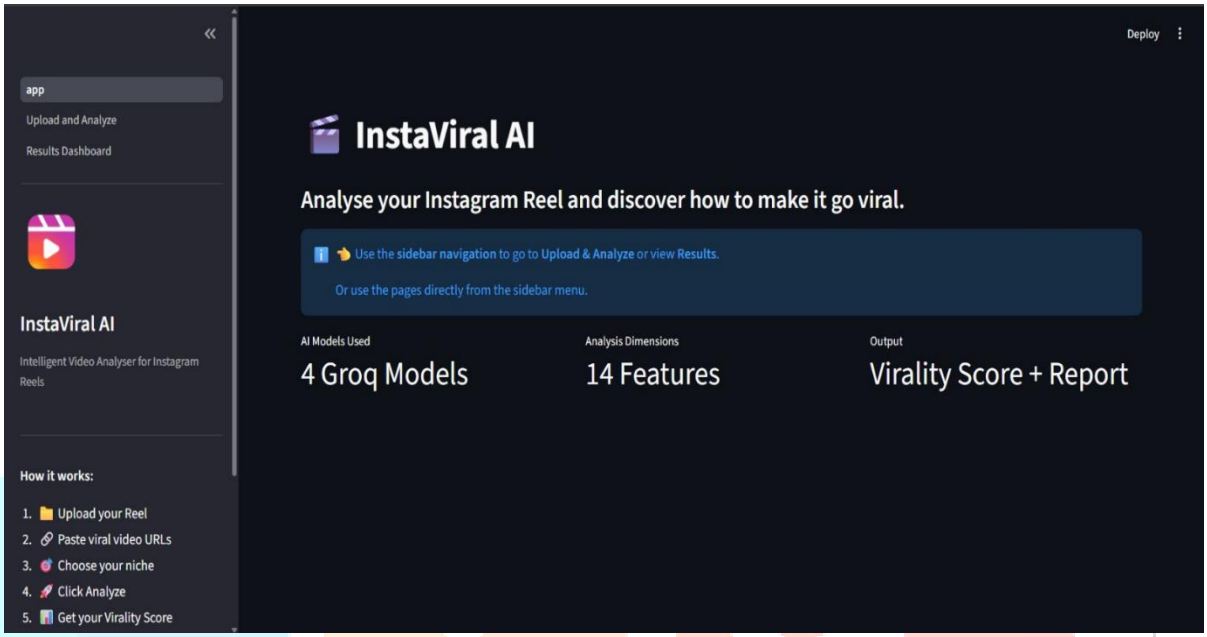

Fig. 6. Screenshot Of Homepage Dashboard

The above screen represents the homepage dashboard of the InstaViral AI system. It provides an overview of the platform, including the AI models used, analysis features, and generated outputs. The dashboard guides users through the Reel analysis process and helps creators understand how the system predicts and improves content virality.

2. Upload and Analyze Module:

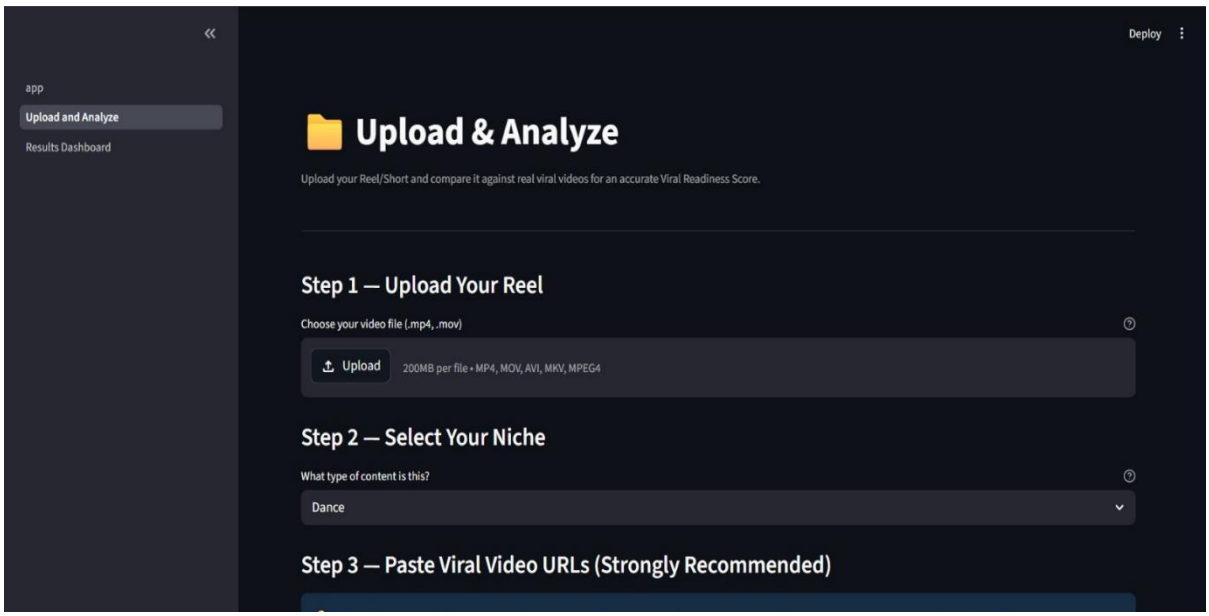

Fig. 7. Screenshot Of Upload and Analyze

The above screen represents the main upload interface of the InstaViral AI system. Users can upload their Reel or short video, select the content niche, and provide viral reference video URLs for comparison. This module helps the system analyze video performance and generate an accurate Viral Readiness Score.

3. Viral Reference URL Input

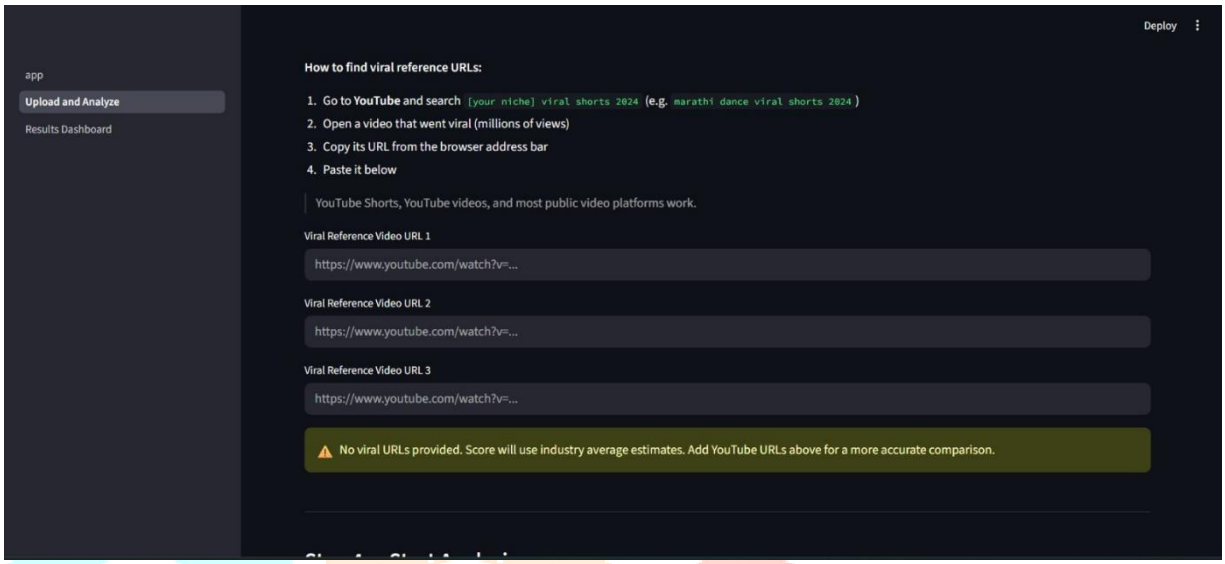

Fig. 8. Screenshot of Viral Reference URL Input

The above screen shows the viral reference video section where users can paste URLs of trending videos. These reference videos help the AI engine compare patterns, engagement trends, and content style to provide better optimization suggestions and analysis results.

4. Analysis Results Dashboard:

Fig. 9. Screenshot of Output

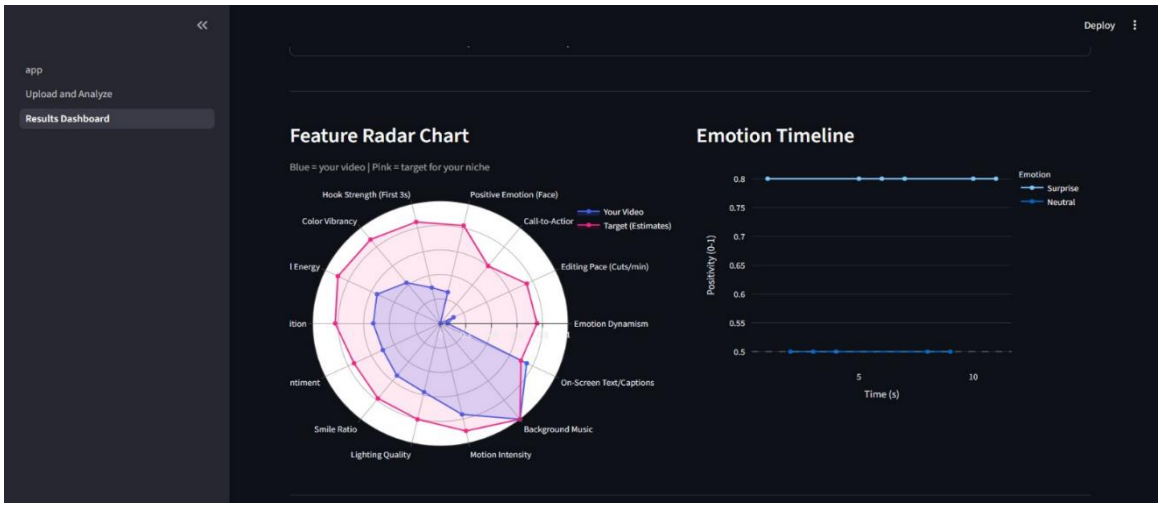

Fig. 10. Screenshot of Output

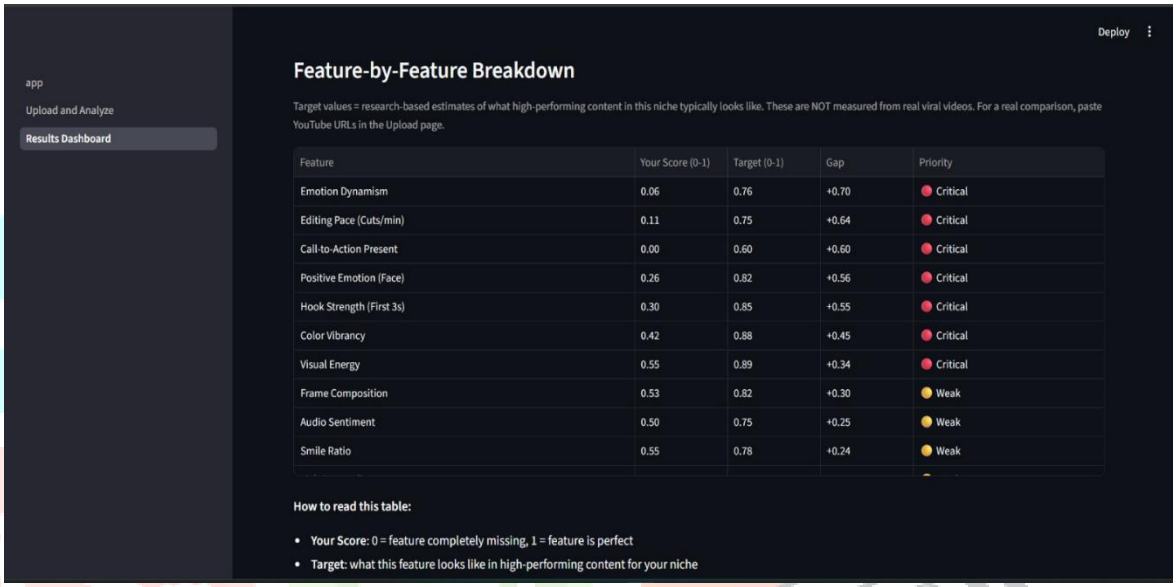

Fig. 11. Screenshot of Output

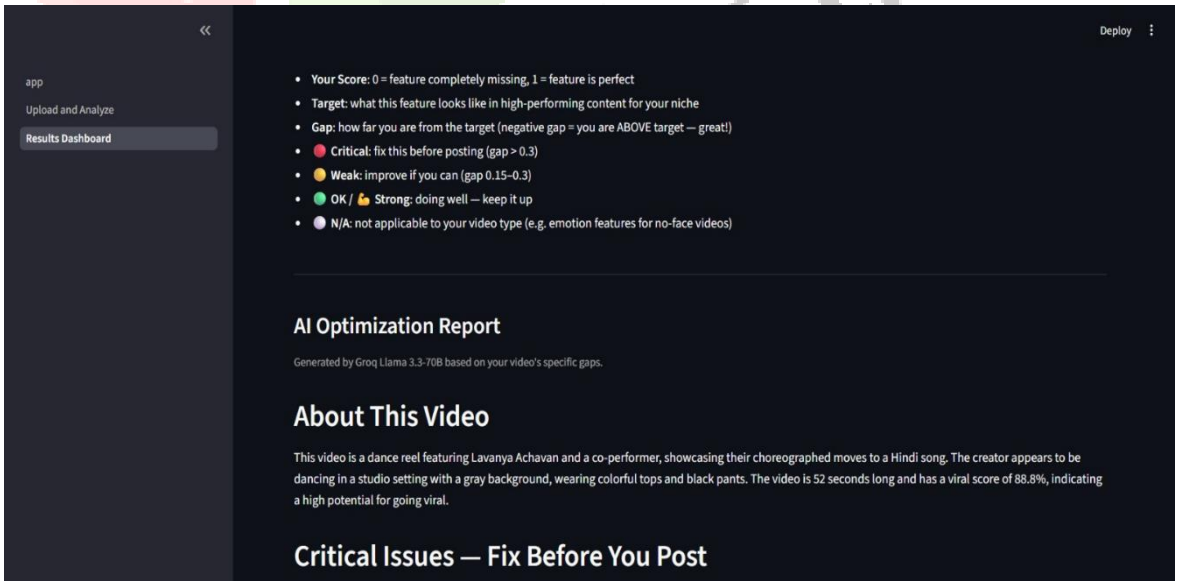

Fig. 12. Screenshot of Output

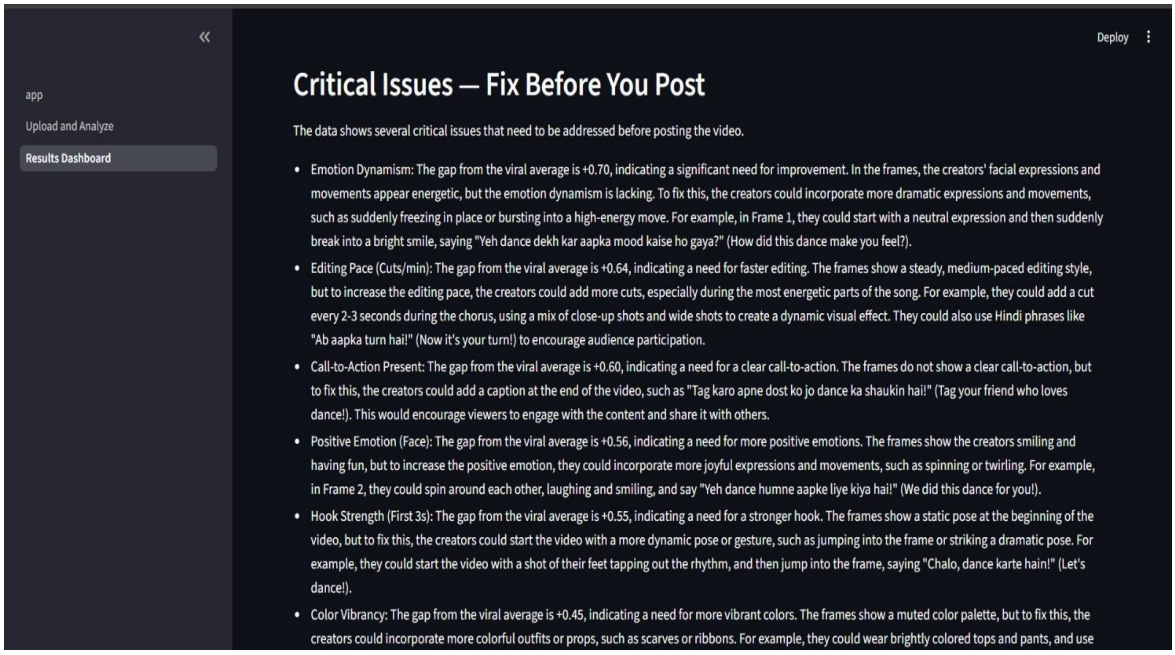

Fig. 13. Screenshot of Output

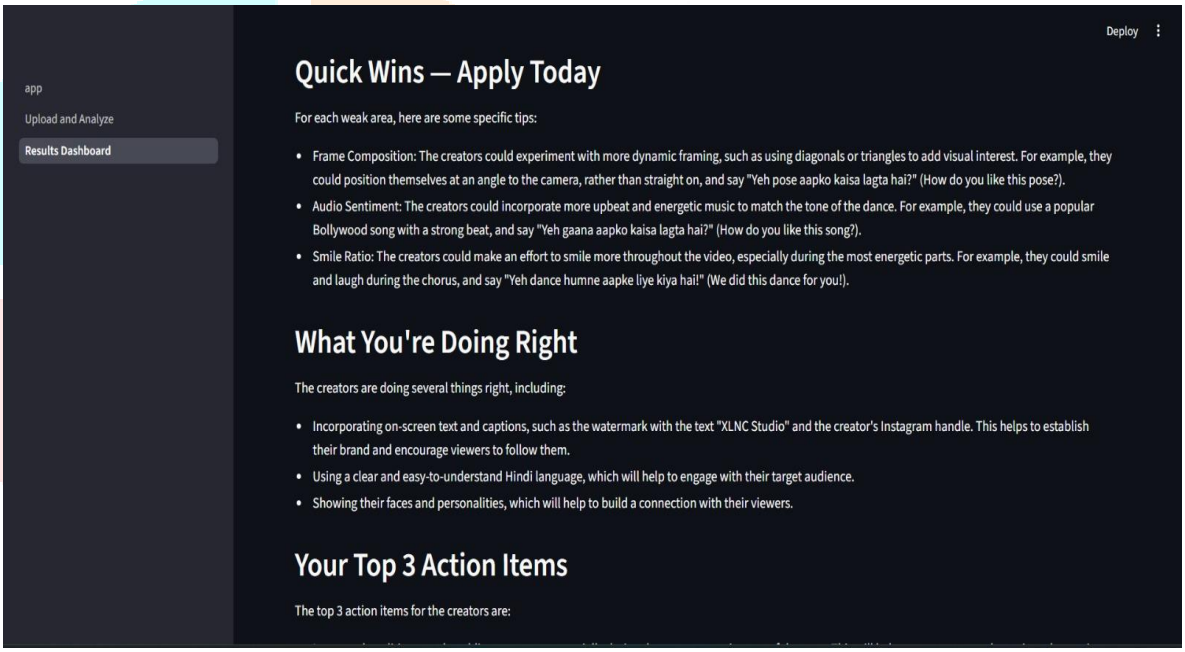

Fig. 14. Screenshot of Output

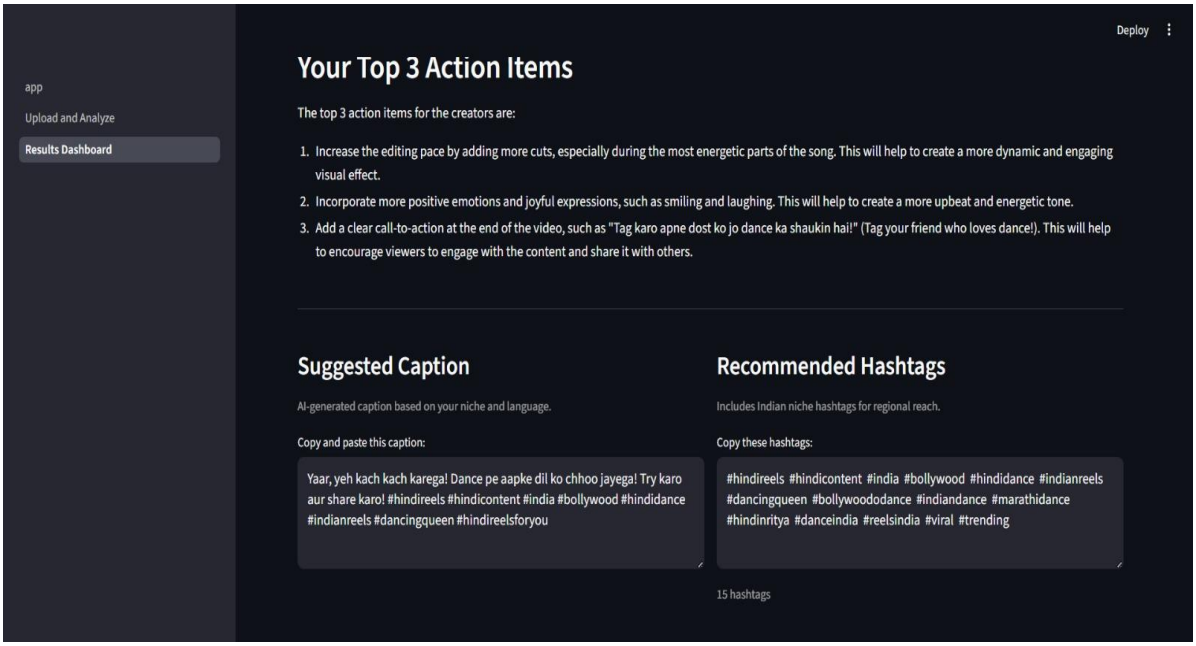

Fig. 15. Screenshot of Output

The above screenshots represent the analysis results generated by the InstaViral AI system after processing the uploaded Instagram Reel. The dashboard provides a complete performance evaluation of the video using AI-based analysis techniques.

The system generates a Viral Readiness Score along with graphical insights such as feature radar charts, emotion timelines, and feature-by-feature performance breakdowns. It also identifies critical issues, suggests optimization improvements, and provides actionable recommendations to improve engagement and virality.

Additionally, the AI engine generates optimized captions, recommended hashtags, posting suggestions, and personalized content improvement tips based on the selected niche and video analysis results. These outputs help content creators understand the strengths and weaknesses of their content and improve their chances of reaching a larger audience.

VII. Future Scope:

- Analyze video as it records using a mobile app interface. Real-time analysis
- Extend support to YouTube Shorts, Moj, and Josh platforms. More platforms
- Compare two versions of the same reel to pick the stronger one. A/B testing module
- Crawl trending audio and hooks weekly and integrate into suggestions. Trend detection
- Historical tracking of virality score improvement over time. Creator dashboard
- Deploy on AWS/GCP for public access with user authentication. Cloud deployment
- Apply niche-specific feature weights (e.g. music matters more for Dance). Weighted scoring

VIII. Conclusion:

InstaViral AI successfully addresses a real and growing need among Indian content creators by combining multiple AI technologies into a single, accessible web application. The system's multilingual support, transparent scoring methodology, and culturally-aware suggestions make it uniquely suited to the Indian Instagram ecosystem. The use of Groq's high-speed inference for both audio and vision analysis keeps the analysis time under 5 minutes on a standard laptop — making it practical for everyday creator use.
The project demonstrates a successful integration of FastAPI, Streamlit, MongoDB, OpenCV, PySceneDetect, EmotiEffLib, librosa, and multiple large language models into a coherent, production-quality architecture. It serves as a strong foundation for a commercial-grade content intelligence platform for Indian social media creators.

IX. References:

1. Z. He and D. Li, “The short video popularity prediction using AI and deep learning,” J. Name Stand. Abbrev., 2024.
2. L. Rivadeneira and I. Loor, “Evidential reasoning approach for predicting popularity of Instagram posts,” J. Name Stand. Abbrev., 2024.
3. A. Bielski and T. Trzcinski, “Understanding multimodal popularity prediction of social media videos with self-attention,” J. Name Stand. Abbrev., 2018.
5. Groq AI Documentation — https://console.groq.com/docs
4. Meta Llama 4 Scout Model Card — https://huggingface.co/meta-llama
5. OpenAI Whisper Paper — Radford et al., 2022 — Robust Speech Recognition via Large-Scale Weak Supervision
6. PySceneDetect Documentation — https://www.scenedetect.com
7. EmotiEffLib — Efficient Facial Emotion Recognition — https://github.com/sb-ai-lab/EmotiEffLib
8. FastAPI Documentation — https://fastapi.tiangolo.com
9. Streamlit Documentation — https://docs.streamlit.io
10. MongoDB Motor Documentation — https://motor.readthedocs.io
11. librosa — Audio and Music Signal Analysis — https://librosa.org
12. Scikit-learn Cosine Similarity — https://scikit-learn.org