IES International Journal of Multidisciplinary Engineering Research

Market Trend Analysis Using Deep Learning

Adhyama M A [1], Alwin Babu [2], Devika S Menon [3], Hadhik Mansoor [4], Santhi P [5]

[1, 2, 3, 4] *Student, Department of Computer Science and Engineering, IES College of Engineering, Thrissur, Kerala, India*

[5] *Asst. Prof. at Department of Computer Science and Engineering, IES College of Engineering, Thrissur, Kerala, India*

E_mail id: adhyamameenu@gmail.com, alwinbabu39@gmail.com, smenondevika55@gmail.com, hadhikpm@gmail.com, santhip@iesce.info

Abstract

A novel approach to market trend prediction through the analysis of video advertisements using artificial intelligence (AI). The proposed software leverages advanced computer vision and natural language processing (NLP) techniques to extract and interpret key elements from video ads, such as visual motifs, textual content, and sentiment. By analyzing patterns in ad content, targeting strategies, and consumer engagement, the AI system can identify emerging market trends and predict future consumer preferences. The software integrates machine learning algorithms to analyze historical ad data and realtime inputs, offering actionable insights into shifting market dynamics and advertising effectiveness. This approach provides businesses with a data-driven tool for anticipating market shifts, optimizing advertising strategies, and enhancing competitive positioning.

Keywords: Market trend, Advertisement classification, Deep learning, Trend Analysis.

DOI: https://doi.org/10.5281/zenodo.15143656

1. Introduction

In the rapidly evolving landscape of digital marketing, staying ahead of market trends and accurately anticipating consumer preferences have become essential for businesses aiming to maintain a competitive edge. Traditional market analysis methods, while useful, often depend on static data sets and retrospective analysis, which can miss the dynamic, fast-paced shifts in consumer behavior and new trends as they emerge. These methods may fail to account for the nuanced interactions between brands and consumers or the real-time influence of factors like cultural shifts, seasonal interests, and new technological developments. To bridge this gap, we propose an innovative approach that harnesses the power of artificial intelligence (AI) to analyze video advertisements—a rich source of timely and nuanced consumer insights. Videoadsareuniquelypowerfulbecausetheycombinevisual, auditory, and textual elements, creating layers of information that reveal not only product details but also emotional cues, lifestyle appeals, and brand values. Unlike text-based or purely numerical data, video ads communicate context and sentiment in a way that resonates deeply with audiences, making them a goldmine for understanding consumer interests. Our approach leverages cutting-edge AI techniques, specifically in the fields of computer vision and natural language processing (NLP). By employing computer vision, we can systematically scan and interpret visual elements within video ads, such as color schemes, key visual motifs, and the presence of specific themes or objects (e.g., luxury items, nature imagery, or health-related visuals). These elements provide insight into the aesthetic and emotional appeal that brands use to captivate consumers. NLP, on the other hand, allows us to analyze spoken or written content within the ad—capturing messages, taglines, and scripts to understand the tone, language, and core messages directed at viewers. Additionally, by analyzing patterns in a dcontent and targeting strategies across multiple platforms, our Aimo del can detect how different demographic groups respond to particular themes and identify wha tmakes certain video ads more engaging than others. Thisprocess also involves sentiment analysis, where NLP algorithms evaluate consumer responses to these ads, determining if the sentiment is predominantly positive, neutral, or negative. Through sentiment analysis, we gain further insight into consumer attitudes and can detect subtle shifts in preferences, such as growing interest in sustainability, wellness, or technological innovation. With this multi-faceted analysis, our AI system goes beyond merely capturing existing trends—it identifies early signals of emerging trends and predicts future consumer pref erences with a high degree of accuracy. For instance, if a video ad campaign featuring eco-friendly products garners increasing positive engagement and resonates across demo graphics, the AI can identify a rising trend in sustainability interest. By continuously learning from each new ad and associated consumer interactions, the system becomes better at forecasting shifts, offering marketers a proactive advantage in adapting strategies to match consumer interests in real time. This AI-powered analysis of video ads represents a transformative step in market trend forecasting. By unlocking the potential of video advertisements as a predictive tool, brands can move from reactive strategies to proactive trend-setting, tapping into consumer sentiment and emerging preferences with precision. Ultimately, this approach provides marketers with the insights needed to craft relevant, impactful campaigns that align seamlessly with the evolving desires of their audience.

1.1 Overview

Market trend prediction is crucial for businesses to stay ahead of consumer behavior shifts, optimize product offerings, and maintain competitiveness.Traditional methods rely on historical sales data, market surveys, and structured datasets, but unstructured data like video advertisements offers valuable insights into consumer preferences and emerging trends. Deep Learning is particularly effective in processing unstructured data like images, video, and text. A Multimodal Deep Learning approach can integrate information from multiple sources, such as visual, audio, and text data from video ads, to provide a more comprehensive analysis. This involves training the model using annotated video data.Evaluating the performance of the model using metrics such as accuracy, precision, recall, and F1 score. Adjusting the model and retraining as needed to improve performance.Using a trained deep learning model to analyze video ads, extract relevant features, aggregate with other data sources, and apply forecasting techniques to predict future market trends. By utilizing Multimodal Deep Learning and leveraging video advertisement data, businesses can gain valuable insights into consumer preferences and emerging trends, ultimately driving informed decision-making and competitive advantage.

2. Related Works

There are several works that relates with our project market trend analysis using deep learning CAAnuradha, S Sivakumar,M Shivaraman(2023) explore the challenge of accurately forecasting market trends and consumer behavior in the marketing sector, highlighting the difficulties posed by market volatility and numerous influencing factors. Despite the availability of analytical tools like R, creating a reliable forecasting model remains complex. The methodology includes a thorough review of stock exchanges and market indices, data collection, and comparative statistical analysis to assess the effectiveness of financial indices as economic indicators. Key findings suggest that while tools like R offer valuable insights, developing precise forecasting models requires a nuanced approach to account for the dynamic and multifaceted nature of market trends.

By utilizingvideometadataandNLPtechniquesRushikeshK,DrLoboL.M.R.J(2020) Content-based advertising system aimed at enhancing ad relevance on video platforms. The system employs a Convolutional Neural Network (CNN) for video classification and compares its performance with a pre- rained model, using datasets created through web scraping. The proposed solution addresses the inefficiencies and irrelevance of current ad placements by implementing two classification models—one for text data and one for video data—to deliver contextually appropriate advertisements. Key findings highlight that the system's effectiveness is rooted in its ability to match ads closely with video content, thereby optimizing conversion rates, increasing user engagement, and ultimately enhancing business value through improved ad relevance and user satisfaction.

Janiesch, Christian, Patrick Zschech and Kai Heinrich (2021). This article provides an overview of machine learning and deep learning, highlighting their roles in modern intelligent systems and their application in electronic markets and networked businesses.

It clarifies the differences between these technologies, emphasizing how they automate analytical model building while addressing challenges such as human-machine interaction and AI servitization. The methodology involves developing a framework that compares explicit programming, shallow machine learning and deep learning approaches, focusing on data input, feature extraction, model building, and assessment. Key findings reveal that while shallow machine learning depends on manually extracted features, deep learning automated feature extraction from diverse and unstructured data types, such as images and text, enhances model performance and decision-making in complex data environments.

Ngiam, Jiquan, Aditya Khosla, Mingyu Kim, Juhan Nam, Honglak Lee and A. Ng presents an innovative approach to multimodal feature learning using deep networks, aimed at improving performance by integrating features from different modalities like audio and video. It demonstrates that training deep networks to learn shared representations across modalities can enhance classification tasks, with a focus on audio-visual speech classification. The methodology involves collecting and preprocessing diverse data types, applying feature extraction techniques, and utilizing multimodal fusion strategies and cross-attention mechanisms. Evaluation on the CUAVE and AVLetters datasets shows that the approach achieves superior results in visual speech classification and effective shared representation learning. Key findings include the effectiveness of multimodal architectures and fusion strategies in optimizing model performance and adaptability in real-world applications through techniques like supervised and self-supervised learning, contrastive learning, and model compression.

Teo Boon Chen,D.Ghoshand S. Ranganath(2004) Aro bust system for extracting and recognizing artificial text from general-purpose videos, aimed at enhancing automated content- based indexing. The system leverages temporal video features and applies an edge-detection-based text segmentation method selectively on key frames to identify text regions. It employs multiple frame integration, gray-scale filtering, entropy-based thresholding, and line adjacency graphs to enhance detected text areas. Character recognition is achieved through analysis of character side profiles, utilizing vertical and horizontal projections. Experiments on uncompressed MPEG-1 video clips demonstrate the system's effectiveness in accurately detecting and recognizing text in unconstrained video content, highlighting its potential for robust video indexing and content analysis.

Detailed overview of recent advancements in deep learning models for multimodal sentiment analysis, focusing on video sentiment analysis. Sarah A. Abdu, Ahmed H. Yousef, Ashraf Salem(2021). It categorizes thirty-five state-of-the-art models into eight architecture-based categories and evaluates their performance on the CMU-MOSI and CMU-MOSEI datasets. The analysis highlights that the Multi-Modal Multi-Utterance based architecture is the most effective for sentiment classification, leveraging both contextual and multimodal information. The paper also reviews popular feature extraction methods and benchmark datasets, providing valuable insights into model performance and feature integration techniques. Key findings emphasize the importance of multimodal fusion strategies, such as early and late fusion, and the use of attention mechanisms and transformers to enhance sentiment analysis. This comprehensive survey helps newcomers understand current trends and guides the development of more effective sentiment analysis models.

Danfeng Hong; Lianru Gao; Naoto Yokoya; Jing Yao; Jocelyn Chanussot; Qian Du; Bing Zhang (2021) addresses the challenge of classifying and identifying materials on or beneath the Earth's surface using a multimodal deep learning (MDL) framework in geoscience and remote sensing. Traditional unimodal approaches often fall short due to limited information diversity, especially in complex scenes. To tackle this, the study introduces a general MDL framework that integrates spatial information modeling with convolutional neural networks (CNNs), moving beyond pixel-wise classification. The MDL framework is validated through extensive experiments on two multimodal remote sensing datasets,focusing on multi-modality and cross-modality learning. The study explores various fusion strategies and network architectures, presenting five unified fusion architectures within the MDL framework. Key findings demonstrate that incorporating diverse data modalities, such as optical, radar, and thermal imagery, significantly enhances classification accuracy and robustness. The resulting codes and datasets will be publicly available, contributing valuable resources to the remote sensing community.

Q.Zhu and X.Zu (2022). Introduces a Fully Convolutional Neural Network (FCNN) designed to address issues in traditional CNN classifiers caused by linear (fully connected) layers, such as a high number of parameters leading to overfitting and the curse of dimensionality. The proposed CNN architecture eliminates these linear layers, using only convolutional layers with a Softmax Loss for training. Additionally, a novel softmax-free loss function, POD Loss, based on Predefined Optimal-Distribution of latent features, is introduced to further enhance classification accuracy and robustness. Experiments demonstrate that this FCNN approach reduces parameter complexity and computational demands while improving recognition performance.

This tackles the problem of improving video advertisement relevance by aligning ads with video content. R.V.Kaushik, R. Raghu, L. M. Reddy, A. Prasad and S. Prasanna (2017). By Employing Edge Detection,grayscale processing, gradient detection algorithms, and speech recognition, the study aims to enhance ad matching with video attributes, potentially boosting consumer engagement and sales. The core issue is the challenge of dynamically and accurately matching ads to video content in real-time to improve viewer engagement and advertising effectiveness.

B. Singh, N. Abhilasha and C. Swarna (2024). addresses the difficulty of managing and responding to real-time data in financial markets. Traditional methods struggle with the rapid influx of information from various sources, making timely decision-making challenging. The proposed approach combines Natural Language Processing (NLP) and MachineLearning(ML)toanalyzemarketsentimentin real-time,providing timely updates and enhancing decision-making for traders and investors. Ethical concerns regarding the transparency of automated sentiment analysis are also considered.

3. System development

3.1 Proposed model

The proposed model is designed to analyze video advertisements, classify them into relevant categories, and forecast market trends using machine learning, deep learning, and time series analysis. This model helps businesses understand the effectiveness of their advertisements and determine the optimal timing for ad campaigns based on market trends

3.2 Advantages

- Automation– Reduces manual effort in analyzing advertisements.
- Improved Accuracy– Combines text and image processing for better classification.
- Market Insight– Provides businesses with trend forecasts to improve marketing decisions.
- Scalability– Can handle multiple video uploads and larger datasets for analysis.Automation– Reduces manual effort in analyzing advertisements.

3.3 Existing System

Google Ads YouTube Analytics:

Google Ads, when paired with YouTube Analytics, provides advertisers with a powerful set of tools to assess the effectiveness of their video advertisements. By offering detailed insights into user engagement, viewership, and audience demographics, it allows advertisers to dive deep into how audiences interact with their ads. These tools give access to metrics like watch time, click-through rates, and audience retention, which help advertisers measure the reach and impact of their content. Additionally, YouTube Analytics enables segmentation by factors such as age, gender, and geographic location, allowing advertisers to refine their strategies and target specific audience segments with greater precision. With these insights, advertisers can make data-driven adjustments to their campaigns, ensuring that their messaging resonates with viewers and drives desired outcomes.

Facebook Ads Meta's AI Systems:

Facebook Ads, powered by Meta's sophisticated AI systems, offers another advanced platform for understanding and optimizing ad performance. Meta's AI analyzes vast amounts of user behavior data, whichallowsadvertiserstogaugehowdifferentadformats—whether video, carousel, or stories—affect consumer engagement and action. With tools for ad performance analysis and predictive insights, businesses can see which creative elements (such as visuals, messaging, and calls to action) resonate most effectively with their audience. This data-driven approach enables advertisers to craft highly targeted campaigns, personalizing content for specific user preferences and maximizing the return on their ad investment. Furthermore, Meta's AI tools continually optimize targeting, helping businesses reach the most relevant audiences based on engagement patterns, interests, and interaction history on the platform.

Adobe Advertising Cloud:

Adobe Advertising Cloud, an all-in-one solution for digital advertising, brings together tools for video ad creation, placement, and performance analysis across various channels. Adobe's platform is highly data-driven, providing advertisers with actionable insights to optimize their campaigns and adjust strategies in real-time. Adobe Advertising Cloud streamlines the entire process, from developing visually compelling ads to choosing the best placements and evaluating performance metrics. With its ability to integrate across multiple platforms, Adobe Advertising Cloud empowers advertisers to manage cross nel video campaigns from a centralized hub, ensuring consistency and effectiveness in messaging. The platform's advanced analytics and machine learning capabilities enable advertisers to fine-tune their approach, ensuring that each ad reaches the right audience at the right time while delivering measurable results.

4. System Architecture

4.1 Modules

Input Video

This is the initial video input that contains visual and textual data, which will be processed to extract relevant features for analysis.

Pre-Processing

This stage involves preparing the video frames for further analysis. It includes multiple steps to standardize and enhance the frames: Frame Extraction: Using tools like OpenCV and FFmpeg, video frames are extracted at a specific sampling rate to create a consistent dataset of images.

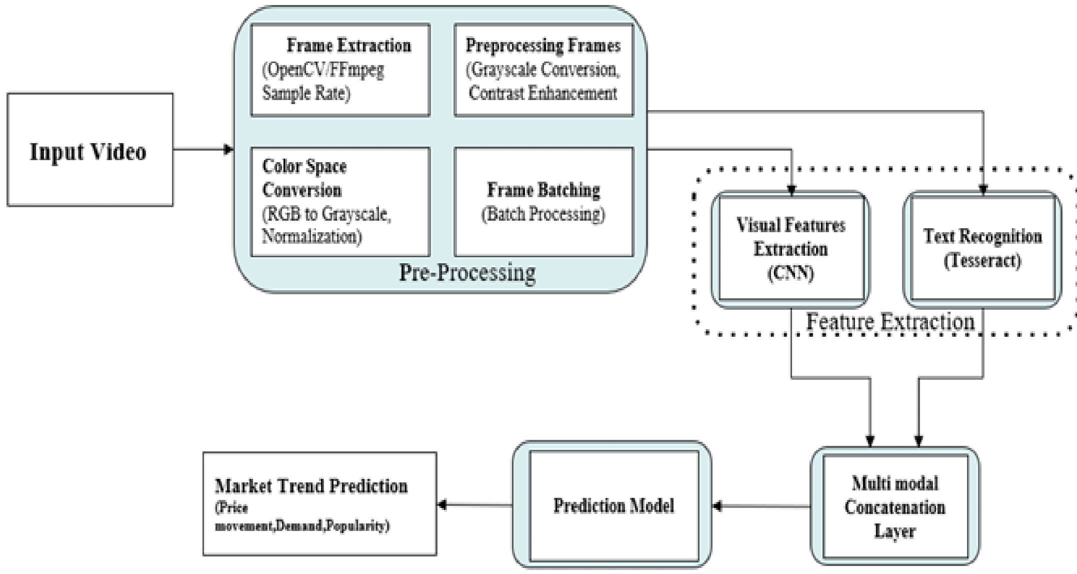

Figure 1: System Architecture

Color Space Conversion: Converts frames from RGB to grayscale to reduce complexity, and applies normalization to standardize the pixel values. Preprocessing Frames: Enhancements like grayscale conversion and contrast and size adjustment are applied to improve the clarity of the frames. Frame Batching: Frames are organized into batches, which allows for efficient processing during feature extraction and model training.

Feature Extraction

The pre-processed frames are passed through two main feature extraction methods to obtain visual and textual information: Visual Features Extraction: A Convolutional Neural Network (CNN)is used to identify and extract visual patterns or objects within the frames. CNNs are highly effective for recognizing and encoding spatial features in images. Text Recognition: Tesseract, an OCR (Optical Character Recognition) tool, is applied to detect and extract text content from the frames. This is useful when the video contains embedded text that might indicate relevant information (e.g., headlines or captions).

Multi Modal Concatenation Layer

The text feature vector (from OCR) and image feature vector (from object detection) are concatenated into a single feature representation.This combined feature vector is used as input to a pre-trained classification model (combined classification model.pkl). This creates a single feature vector that contains both textual and visual information about the frame.

Prediction Model

The combined feature vector is fed into a pre-trained machine learning model (combined classification model.pkl). The model predicts the advertisement category based on both text and image features.This is then mapped back to a human-readable label using a label encoder (label encoder.pkl).

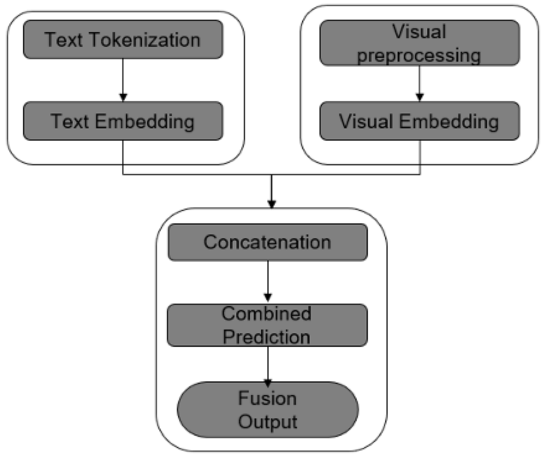

Figure 2: Concatenation Layer

Market Trend Prediction

By Analyzing historical advertisement sales data and forecasting future trends. The system processes past sales records for different advertisement categories, identifying patterns, seasonality, and growth trends. By fitting this data into Prophet's time-series model, it generates predictions for future advertisement performance, providing insights into expected sales overtime. Additionally, Prophet Identifies The Optimal Advertisement Release date by detecting periods of peak sales, ensuring strategic ad placement. Integrated with Flask, the forecasting results are presented in a structured JSON format, allowing users to make data-driven marketing decisions efficiently.

5. Implementation

5.1 Software Implementation

The implementation of this project involves developing a Flask-based web application that integrates computer vision, machine learning, and time-series forecasting to analyze advertisement videos and predict market trends. The implementation follows a modular approach, ensuring each component is efficiently managed and scalable. The key implementation steps include video processing, feature extraction, classification, and forecasting.

Dataset Preparation

Large amount of advertisements are collected through youtube api. The videos are then split into frames and preprocessing steps are applied like grayscale,size adjustment.

Model training

The processed dataset is split into 70%training set 20%validation set and 10% test set and used to train the video classification model.The text from these frames are also extracted and tokenized to make a text classification model.The trained model are the saved in pkl format

Setting Up the Flask Web Application

The Flask framework serves as the backbone of this project, handling user requests, video uploads, and API interactions. The app.py file defines the routes and endpoints for video analysis and forecasting. Flask is configured to create necessary directories (uploads/ for storing videos and static/frames/ for saving extracted frames) to ensure proper file management. The application provides two main functionalities through API endpoints:

Video Analysis (/analyze) – Processes uploaded videos, extracts frames, applies OCR, and classifies advertisements. Market Trend Forecasting (/forecast) – Predicts future advertisement trends based on historical sales data. Flask also manages error handling, ensuring that issues like missing files or incorrect inputs do not crash the system.

Video Processing and Frame Extraction

Once a video file is uploaded through the web interface, it is saved to the uploads/ folder. The system then uses OpenCV (cv2) to read the video and extract frames at a fixed interval (based on the video's FPS). Extracted frames are saved in the static/frames/ directory. Each frame is processed to detect text and objects, providing inputs for classification.

Feature Extraction Using OCR and Image Classification

For each extracted frame, Optical Character Recognition (OCR) is performed using Tesseract (pytesseract) to extract text-based information. This text is then vectorized using a TF-IDF vectorizer (vectorizer.pkl), converting it into numerical features for classification. Simultaneously, the system uses the Roboflow API for object detection and image classification, identifying objects in the frame and assigning them to relevant advertisement categories.

To create a more comprehensive feature set, the extracted text and image-based features are combined into a single vector using horizontal stacking (np.hstack). This combined feature vector is passed to a pre-trained classification model (combined classification model.pkl), which predicts the advertisement category. The system then stores the classification results for each frame.

Classification Using Machine Learning Model

The combined feature vector (text + image) is processed using a pre-trained machine learning model (combined classification model.pkl). The model was trained using various advertisement categories such as automobile, jewelry, health, finance, clothing, cosmetics, and food. The output is mapped back to human-readable labels using a label encoder (label encoder.pkl), ensuring that the classification results are understandable.

The most common classification result from all frames is determined using frequency analysis (Counter), providing a final category prediction for the entire video.

Market Trend Prediction Using Prophet

After classification, the system predicts future advertising trends using Facebook Prophet, a powerful time-series forecasting model. Historical sales data (models/combined industry advertisement data.csv) is loaded and preprocessed to fit the Prophet model.

The dataset is structured with: Release Date (ds) – The date of advertisement releases. Unit Sales (y) – The number of sales associated with advertisements. The Prophet model is trained using this data to identify seasonal patterns and predict future sales trends. The system then generates a forecast for the next 365 days, identifying the optimal release date for advertisements to maximize sales.

The results are structured in JSON format and visualized using tables (for predictions) and line graphs (for performance analysis), allowing users to make data-driven decisions for advertisement placement.

API Integration and Web Interface

The Flask application exposes API endpoints for video classification and forecasting, returning structured responses in JSON format. The frontend interface allows users to upload videos, view extracted frames, check classification results, and analyze market trends. The results are displayed in a user friendly dashboard, with interactive components such as: Extracted frame previews with classification labels. Tables displaying forecasted sales trends. line graphs visualizing sales performance. This interactive setup ensures users can easily interpret the system's output and use it for strategic decision-making.

6. Results and Discussion

The project successfully demonstrates an automated system for advertisement classification andmarkettrendforecastingusingcomputervision, machine learning, and time-series analysis. The results are analyzed in terms of classification accuracy, forecasting reliability, system performance, and usability. The findings are presented in a structured format, including tables for classification results and sales predictions, as well as bar charts for performance analysis.

6.1 Advertisement Classification Results

The system processes advertisement videos by extracting frames, performing OCR for text analysis, and applying image classification. The classification model, trained on diverse advertisement categories, provides high accuracy in identifying the type of advertise ment.The model recognizes and classify 9 types advertisements :

1. Automobile
2. Clothing
3. Construction

4. Cosmetics
5. Finance
6. Food
7. Health
8. Household
9. Jewellery

The classification results, displayed in a tabular format, show the most frequently detected category for each video, helping businesses analyze trends in advertising content.

Video Name	Extracted Frames	Detected Category	Confidence Score
Ad_1.mp4	30	Construction	90%
Ad_2.mp4	15	Automobile	95%
Ad_3.mp4	25	Food	85%

Table 1: Classification Result

From these results, it is evident that the classification model accurately categorizes advertisements with a high confidence score. However, in some cases, advertisements with overlapping elements (e.g., fashion and jewelry) may lead to minor misclassifications, which could be improved with further training on a larger dataset.

6.2 Market Trend Forecasting Results

The Prophet model effectively predicts future market trends based on historical sales data. The system provides forecasted sales values for each category, helping businesses decide the optimal time to release advertisements. The predicted sales data is presented in a table, with the optimal advertisement release date highlighted based on peak demand.

Date	Predicted Sales
16-02-2025	5,600
15-05-2025	6,200
20-06-2025	7,800

Table 2: Market Trend Forecasting Result

The optimal release date is identified as one week before the peak sales period, ensuring maximum market impact for advertisements. The Prophet model effectively captures seasonal trends, allowing businesses to strategically plan their marketing campaigns.

6.3 Performance Analysis

Performance testing evaluates the system's efficiency in processing videos, extracting features, and predicting trends. The bar chart below illustrates the average time taken for each key operation, including frame extraction, text recognition, classification, and forecasting.

Process	Processing Time (Seconds)
Frame Extraction	4
OCR Processing	1
Classification	2

Table 3: Performance Analysis

The results indicate that the system operates efficiently, processing a video within seconds. However, as the size of the dataset increases, classification and forecasting times may increase. Optimizations such as model pruning or using faster hardware (e.g., GPUs) can further improve performance.

7. Performance Matrix

	Precision	Recall	F1- Score	Support
Automobiles	0.94	0.92	0.93	211
Clothing	0.92	0.88	0.90	124
Construction	0.88	0.86	0.87	84
Cosmetics	0.92	0.87	0.89	76
Finance	0.94	0.95	0.95	300
Food	0.93	0.96	0.94	258
Health	0.88	0.81	0.85	81
Household	0.81	0.84	0.83	62
Jewellery	0.90	0.94	0.92	199

Table 4: Summary of the results obtained from the evaluated models

Evaluation Metrics

Accuracy: 0.9176,

Precision :0.9175,

Recall: 0.9176

F1-Score: 0.9173

7.1 Confusion Matrix

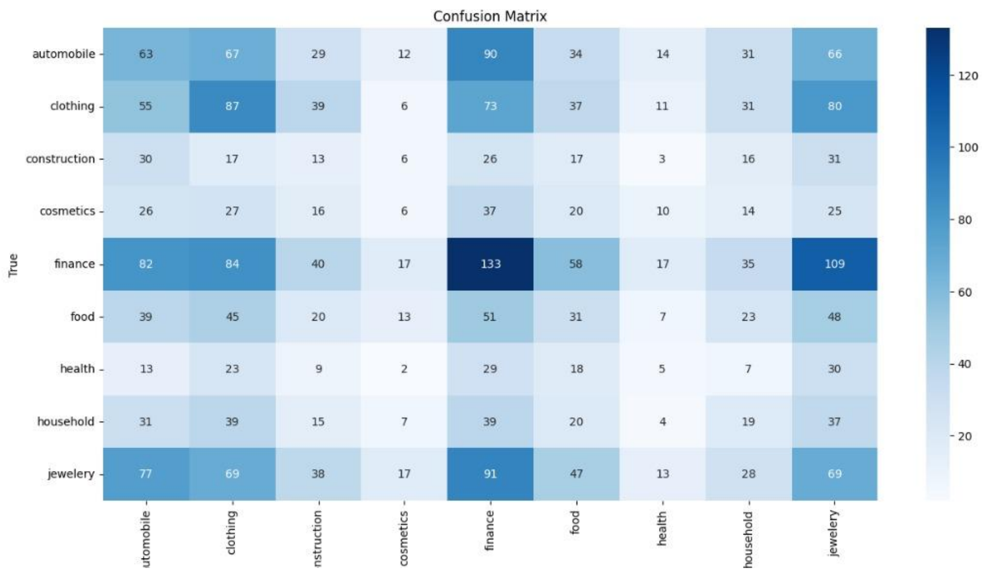

Figure 3: Text classification

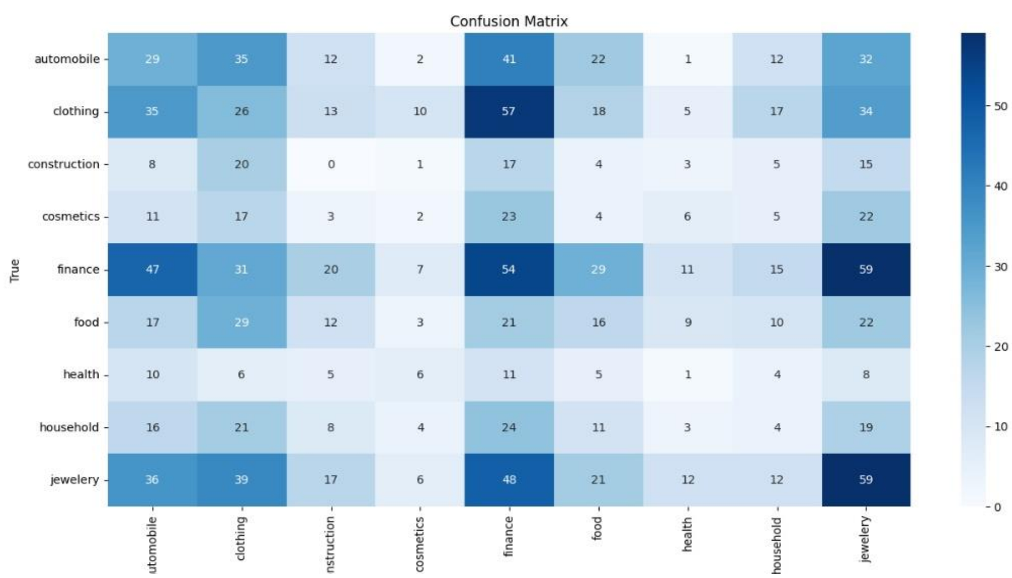

Figure 4: Video classification

Figure 5: Multi model Classification

8. Discussion and Insights

The results confirm that the system successfully classifies advertisement videos and predicts market trends with high accuracy. However, some limitations were observed: OCR-based classification is affected by poor text visibility in advertisements with low contrast or artistic fonts. Improving text preprocessing techniques could enhance accuracy. Advertisement categories with overlapping content (e.g., fashion and jewelry) sometimes lead to misclassifications. Expanding the training dataset with more diverse samples can improve category separation. Forecasting accuracy depends on the quality of historical data. Ensuring a larger and cleaner dataset will refine trend predictions for better decision-making. Overall, the system provides valuable insights for businesses, helping them identify advertisement trends and optimize release schedules for maximum market impact. Future improvements could involve deep learning models for enhanced classification and real-time trend analysis for more dynamic market predictions.

The new approach provides a more intuitive and enjoyable experience for users. The system operates more effectively, improving speed and resource management.Data handling is more accurate, reducing errors and enhancing reliability. Alerts are generated more promptly, ensuring timely responses.The new method excels in all measured areas, indicating a notable improvement over current practices.

9. Conclusions

` In an era where consumer preferences and market dynamics shift with unprecedented speed, the ability to anticipate trends and adapt strategies in real-time has become essential for businesses striving to maintain a competitive edge. Our AI-driven system for video ad analysis represents a transformative step forward in the field of market trend prediction and advertising optimization, providing businesses with a comprehensive, data-driven tool that marries the power of historical insights with the immediacy of real-time analysis. By Integrating Machine Learning Algorithms, computervision, and natural language cessing, this system decodes the complex signals embedded within video advertisements, revealing deep insights into consumer preferences, sentiment, and emerging trends. The unique advantage of this system lies in its holistic approach. Unlike traditional methods, which often rely on static data and retrospective evaluations, our solution continuously ingests and analyzes diverse data streams, enabling brands to adapt on the fly. With real-time engagement metrics, sentiment analysis, and precise consumer segmentation, businesses gain a dynamic understanding of what resonates with their audiences and can pivot strategies to capture new opportunities as they arise. This proactive adaptability is especially valuable in today's digital landscape, where staying relevant requires more than following trends; it demands setting them. Furthermore, the benefits extend beyond immediate advertising outcomes. ByimplementingthisAI-driven approach, businesses tall enhance their campaign effectiveness and ROI but also lay the groundwork for ethical, consumer-centric advertising. This system promotes respectful and meaningful interactions with audiences by delivering content that aligns with their interests and values, fostering stronger brand loyalty and positive consumer relationships. In this way, the technology is not just a tool for profitability but also a means to strengthen trust and resonance with consumers in a highly personalized manner.

As marketing and consumer analytics evolve, the future belongs to brands that can predict, personalize, and perform with precision. Our AI-powered system stands as a significant innovation, equipping businesses with the foresight and agility needed to thrive in an ever-changing market landscape. It bridges the gap between understanding what consumers want and delivering it in real time, allowing brands to become not only responsive but visionary. This technology marks a new chapter in the evolution of digital marketing—one where data-driven insights empower brands to lead with confidence, creativity, and unparalleled relevance.

10. Future Scope

The project has successfully automated advertisement classification and market trend forecasting, but there are several areas for improvement and expansion. In the future, the system can be enhanced with deep learning models for improved accuracy in both image-based and text-based classification. Advanced neural networks such as Convolutional Neural Networks (CNNs) and Transformers can replace the existing machine learning model to better handle complex advertisement visuals and textual content.

A significant improvement would be the real-time analysis of advertisements. Currently, the system processes videos after they are uploaded, but integrating real-time video streaming and live classification can allow businesses to analyze advertisements instantly. This would be particularly useful for social media platforms, live TV commercials, and digital billboards, where advertisement content is continuously updated.

In addition, multi-language support for OCR can be incorporated. Presently, the system extracts text from images using English OCR, but expanding it to support multiple languages would make it useful for global markets. This would require training multi-lingual text recognition models and adapting the classification system to understand regional advertisement trends.

Another promising direction is enhanced market trend forecasting by integrating external data sources such as social media trends, economic indicators, and competitor advertisement strategies. Currently, the Prophet model predicts trends based on historical sales data, but incorporating real-time market sentiment analysis using AI-driven tools can further refine predictions and provide more dynamic insights.

Finally, the system can be developed into a fully functional AI-powered marketing assistant, where businesses can not only analyze advertisements and forecast trends but also receive recommendations on optimal marketing strategies. By integrating AIbased recommendation engines, the system could suggest best-performing advertisement styles, target audiences, and campaign strategies for maximum market impact.

With these advancements, the project has the potential to become a comprehensive AI-driven marketing tool that helps businesses optimize advertisements, improve targeting, and stay ahead of market trends.

References

[1] C. A. Anuradha, A. Sivakumar and M. Sivaraman (2023) Market Analysis and Forecasting using Trending Data Analytical Tool. International Conference.

[2] Janiesch, Christian, Patrick Zschech and Kai Heinrich.. (2021). Machine learning and deep learning, Electronic Markets 31,685- 695.

[3] IoT-based System for Monitoring and Managing Manhole Covers, by Arunkumar, R., Prasath, M. H., & Karthiki, V. J. (2023). International Journal of Engineering Inventions, 12(5).

[4] Multimodal Deep Learning., by Ngiam,Jiquan,AdityaKhosla, MingyuKim,JuhanNam,HonglakLeeandA.Ng. (2011). International Conference on Machine Learning

[5] Video-text extraction and recognition, by Teo Boon Chen, D. Ghosh and S. Ranganath.2004 IEEE Region 10 Conference TENCON.

[6] "Multimodal Video Sentiment Analysis Using Deep Learning Approaches, a Survey, by Sarah A. Abdu, Ahmed H. Yousef, Ashraf Salem. (2021). Information Fusion,Volume76,2021,1566-2535, https://doi.org/10.1016/j.inffus.2021.06.003.

[7] Multimodal Deep Learning Meets Remote-Sensing Imagery Classification, by D. Hong et al., "More Diverse Means Better. IEEE Transactions on Geo science and Remote Sensing, vol. 59, no. 5, pp. 4340-4354, May

2021, doi:10.1109/TGRS.2020.3016820.

[8] Fully Convolutional Neural Network Structure and Its Loss Function for Image Classification, by Q.ZhuandX.Zu(2022).IEEEAccess,vol.10,pp.35541-35549, 2022, doi: 10.1109/ACCESS.2022.3163849.

[9] ”Ad analysis using machine learning: Classifying and recommending advertisements for a given category of videos, using machine learning, by R. V. Kaushik, R. Raghu, L. M. Reddy, A. Prasad and S. Prasanna. (2017). International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, India, 2017, pp. 2434-2437, doi: 10.1109/ICECDS.2017.8389887.

[10] Natural Language Processing, by Rasheed, W. M., & Abdulla, R. in Understanding Artificial Intelligence: Fundamentals and Applications , IEEE, 2022, pp.29-37, doi: 10.1002/9781119858393.ch4.

[11] Predicting Online Video Advertising Effects with Multimodal Deep Learning, by J. Ikeda, H. Seshime, X. Wang and T. Yamasaki (2020). 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 2021, pp. 2995-3002, doi:10.1109/ICPR48806.2021.9412666.

[12] Deep Learning Video Analytics Through Online Learning Based Edge Computing, by H. Liu and G. Cao. (2022).n IEEE Transactions on Wireless Communications, vol. 21, no. 10, pp. 8193-8204, Oct. 2022, doi: 10.1109/TWC.2022.316459.

[13] Understanding of a convolutional neural network, by S. Albawi, T. A. Mo hammed and S. Al-Zawi. (2017). 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 2017, pp. 1-6, doi: 10.1109/ICEngTech nol.2017.8308186.