WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Real-Time Trend Analysis and Sentiment Classification Using Machine Learning Across Social Media Paltform

Amullya Patil , Anupriya S Katnoor, Priyanka Karoor, Dr Prakash Pattan

Dept. Of computer science and engineering PDA collage of engineering kalaburagi, india

Abstract—

In an era of rapidly evolving digital content, real-time understanding of emerging trends across online platforms is crucial for media analysts, marketers, and decision-makers. This paper presents the design and implementation of a real-time, storage-free Trend Analyzer system that extracts, processes, and visualizes trending topics from multiple sources including Twitter, Google Trends, Reddit, and online news outlets. The system leverages API integration for continuous data streaming and applies advanced machine learning techniques such as BERT, RoBERTa, and GPT- based models for sentiment analysis and topic classification. Unlike conventional systems that rely on stored datasets, this solution processes data on-the-fly, enhancing privacy and reducing infrastructure costs. Visualization is achieved through a responsive, browser-based dashboard built using HTML, CSS, and JavaScript, providing dynamic graphs, sentiment scores, and trend timelines. Performance testing demonstrates low latency, high classification accuracy, and user-friendly visual outputs. The architecture is modular, scalable, and suitable for applications ranging from crisis monitoring and market research to influencer analysis and news trend forecasting. This work demonstrates the feasibility and impact of an AI-powered, real-time trend intelligence system that requires no persistent data storage.

Keywords— Real-time trend analysis, sentiment classification, machine learning, BERT, GPT, RoBERTa, social media analytics, API integration, data visualization, no-storage architecture, trend detection, cross-platform monitoring.

I. Introduction

In today's digital age, social media platforms have emerged as powerful arenas where individuals, communities, and organizations actively engage in discussions, share opinions, and influence public discourse on a global scale. Platforms like Twitter, Reddit, Google Trends, and news outlets generate a continuous, vast stream of data representing the collective voice of millions of users. This ever-evolving digital ecosystem captures real-time reactions to current events, viral content, product launches, political developments, and social movements, making it a goldmine of information for businesses, researchers, and policymakers alike. However, the challenge lies in extracting meaningful insights from this vast ocean of unstructured and noisy data, which requires advanced analytical techniques capable of processing information rapidly and accurately. Real-time trend analysis coupled with sentiment classification offers an innovative solution by harnessing the power of machine learning and natural language processing (NLP). These technologies analyze large-scale textual data to identify emerging topics, measure public sentiment, and categorize trends across various domains such as entertainment, technology, fashion, and politics. State-of-the-art models like BERT, RoBERTa, and GPT have significantly advanced the field by enabling a deeper understanding of language context, semantics, and subtle emotional cues. By integrating data from multiple social media APIs, the system provides a holistic and dynamic picture of trending conversations, allowing stakeholders to monitor and react to shifts in public opinion swiftly and effectively. One of the key strengths of this approach is its real-time processing capability. Unlike traditional methods that rely on batch processing or stored datasets, the system continuously ingests live data streams, preprocesses and cleans the input, applies sophisticated machine learning models for sentiment and trend classification, and updates a visual dashboard instantaneously. This architecture not only ensures timely insights but also enhances privacy and security by avoiding the storage of raw user data, thus minimizing risks related to data breaches and regulatory compliance. The preprocessing steps— such as tokenization, stop-word removal, and lemmatization improve the quality of input data, which is crucial for maintaining high accuracy in sentiment detection and trend identification.

II. Related works

Trend detection and analysis using machine learning have become vital tools for understanding evolving user interests across various digital platforms. Chen and Zhang (2023) investigated cross-platform trend detection by combining Twitter and YouTube data streams, utilizing machine learning techniques to enhance the tracking of emerging topics. Their study emphasized the importance of integrating multiple data sources to provide a more comprehensive picture of trending content, which directly supports the multi-platform approach used in our system. Zhang, Wu, and Liu (2022) focused on trend detection within music streaming platforms. By applying machine learning models to analyze user listening patterns, they successfully identified shifts in musical preferences and emerging genres. This domain-specific study demonstrated how machine learning can adapt to niche areas,

offering precise trend predictions relevant to the music industry. Similarly, Patel and Desai (2022) compared trend detection efficacy across social media platforms such as Twitter and YouTube. Their findings highlighted the distinct characteristics and challenges posed by each platform's data, underscoring the need for customized machine learning models tailored to platform-specific content formats and user behavior. In the fashion industry, Liu, Zhou, and Zhang (2021) leveraged machine learning algorithms to predict upcoming fashion trends. Their work showed the potential for AI-driven analysis to forecast consumer preferences based on historical and real-time data, thereby aiding brands in product development and marketing strategies. Another notable contribution came from Zhang, Sun, and Li (2021), who applied machine learning techniques to music streaming data for trend detection, further reinforcing the value of specialized algorithms in understanding cultural consumption patterns.Ahmad, Al-Khatib, and Alwan (2021) explored trend detection and analysis exclusively on Twitter, employing various machine learning approaches to identify real-time events and sentiment shifts. Their research underlined the platform's fast-paced nature and the complexities involved in handling noisy, high-volume data. Earlier, Garofalo and Paredes (2020) examined video virality prediction on YouTube through machine learning models, providing insights into factors that drive content popularity and user engagement on video-sharing platforms. Together, these works provide a solid foundation in the application of machine learning for trend detection across different domains and platforms. Our research builds on this foundation by offering a unified, real-time, multi-source trend analysis system that incorporates advanced natural language processing techniques for sentiment classification and emphasizes privacy and usability in its design.

III. Summary and Limitations of Current Methods

The existing method for real-time trend analysis and sentiment classification leverages a modular architecture that combines real-time data ingestion, natural language processing (NLP), machine learning-based classification, and live visualization. By collecting data from multiple platforms such as Twitter, Google Trends, and Reddit through APIs, the system performs on-the-fly analysis to identify trending topics and classify public sentiment as positive, negative, or neutral. It uses powerful transformer- based models like BERT and RoBERTa for sentiment scoring and applies machine learning classifiers like Random Forest for trend domain categorization. The results are then visualized through a dashboard in real-time, allowing users to interpret and act on the insights instantly. A key advantage of this system is its storage-free design, which prioritizes user privacy by avoiding data persistence. The preprocessing pipeline, which includes tokenization, stop-word removal, and lemmatization, enhances model input quality, while live rendering through WebSockets or Fetch APIs ensures that insights are updated dynamically and visually intuitive. The dashboard's features, such as keyword bars, word clouds, and sentiment polarity graphs, are designed to make complex data accessible to non-technical users. Despite these strengths, several limitations are evident in the current approach. One of the primary challenges is the dependence on third-party APIs, which are subject to rate limits, latency issues, and potential access restrictions. This affects the consistency and reliability of data flow. Additionally, the sentiment classification models, although advanced, struggle with detecting subtle linguistic cues such as sarcasm, slang, or cultural references, leading to occasional inaccuracies. Another significant limitation lies in the static trend classification model, which uses predefined rule-based or keyword-mapped logic. This approach lacks the flexibility to dynamically learn and categorize new or evolving topics, especially those that cross multiple domains (e.g., political memes or viral activism campaigns). Furthermore, the absence of data storage restricts retrospective analysis and long-term trend monitoring, which can be crucial for historical comparisons or research studies. Lastly, under conditions of high data velocity or volume, the system may experience increased latency or reduced responsiveness due to processing bottlenecks. These constraints underscore the need for more scalable infrastructure and adaptive learning mechanisms in future iterations.

IV. Proposed system

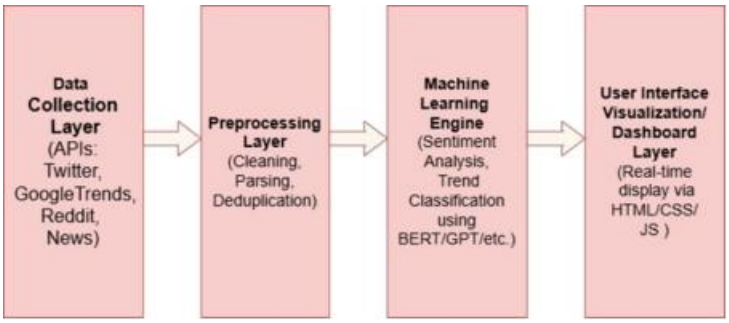


Fig. 2. Architecture of Proposed system

The proposed system aims to provide a real-time, storage-free solution for detecting and analyzing trending topics and public sentiment across multiple social media and web platforms. Unlike traditional analytics tools that rely on stored data and batch processing, this system operates on live data streams from sources such as Twitter, Google Trends, Reddit, and news websites using publicly available APIs. It emphasizes immediacy, privacy, and intelligent interpretation of data. At its core, the system features a data ingestion pipeline that captures high-velocity text content, followed by a robust preprocessing module. This module performs tokenization, stop- word removal, lemmatization, and normalization, preparing the text for accurate analysis. The cleaned data is then passed to machine learning models for sentiment classification and trend categorization. Transformer-based models like BERT or RoBERTa are employed to classify sentiment with a high degree of linguistic understanding, while domain-specific classifiers like Random Forests help categorize trending topics into sectors such as politics, entertainment, or technology. One of the key innovations of the system is its no-database architecture, which enhances user privacy and minimizes storage costs while ensuring real-time feedback. The output is displayed on an interactive dashboard that features sentiment polarity graphs, word clouds, trending keyword bars, and filtering options for topic or platform. This

allows users to derive immediate, actionable insights from fast-changing social data. The system is also designed to be modular and scalable. Its architecture supports the integration of new data sources and model upgrades without affecting the core functionality. The intuitive interface caters to both technical and non-technical users, making it suitable for use in journalism, digital marketing, public relations, political analysis, and disaster response. Overall, the proposed system bridges the gap between raw data and meaningful insight in a fast, secure, and user-centric manner.

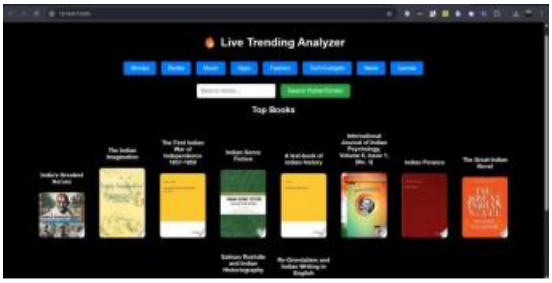
V. System Specifications

The system is designed to run on hardware equipped with an Intel i7 processor or an equivalent having 8 or more cores, along with a minimum of 16 GB RAM (32 GB recommended for optimal performance). Storage requirements include a solid-state drive (SSD) with at least 256 GB of free space to accommodate logs and model files. For machine learning inference, an NVIDIA GTX 1660 or higher GPU is recommended. The backend is developed using Python 3.8 or higher, utilizing either Flask or FastAPI for building RESTful APIs. The frontend comprises HTML, CSS, and JavaScript, with React.js used to build an interactive dashboard. Data visualization is powered by libraries such as Chart.js, D3.js, WordCloud, and Plotly. The machine learning and NLP components rely on Hugging Face Transformers (including models like BERT and RoBERTa), Scikit-learn for traditional algorithms, and NLTK and SpaCy for text preprocessing tasks. The system pulls data from multiple APIs and sources, including Tweepy for Twitter, Pytrends for Google Trends, PRAW for Reddit, and various news APIs using either the Requests or AIOHTTP libraries. Initially deployed on localhost for development purposes, the application is intended for cloud hosting on platforms like AWS or GCP. For security, no databases are used, and all API keys are securely stored in .env files

VI. Methodology

The proposed system is designed to perform real-time trend detection and sentiment classification across major online platforms including Twitter, Reddit, Google Trends, and news portals. It focuses on providing accurate, fast, and privacy- conscious insights without the need to store user or content data. The architecture of the system integrates powerful backend services with an interactive frontend dashboard, all built with scalability and extensibility in mind. Data acquisition begins with real-time API integrations to fetch trending topics, keywords, and user-generated content. Tools such as Tweepy for Twitter, Pytrends for Google Trends, PRAW for Reddit, and HTTP-based connectors for news APIs allow the system to capture a broad spectrum of current digital conversations and events. Once the data is collected, it undergoes rigorous preprocessing to improve the quality and relevance of the input for machine learning models. This includes the removal of stop-words, tokenization, lemmatization, and deduplication of redundant entries. These steps are essential to convert noisy social media text into structured input suitable for natural language processing. The preprocessed data is then analyzed by sentiment classification models based on transformer architectures such as BERT and RoBERTa, which are fine-tuned to identify sentiment polarity—positive, negative, or neutral. These models help the system interpret the emotion and opinion embedded in real-time textual data, including subtle tones like sarcasm or ambiguity. Simultaneously, a rule-based trend categorization module maps keywords and content patterns to predefined domains such as politics, entertainment, sports, and technology, allowing users to explore trends by interest. The results are displayed on a user-centric dashboard developed using React.js, supported by visualization libraries like Plotly and Chart.js. The dashboard enables users to view keyword clouds, sentiment line graphs, and platform comparisons through intuitive and interactive elements. Filters for category, date, and sentiment type make the system highly navigable and useful for both technical and non-technical users. The real-time responsiveness of the system is made possible by optimizing API calls and minimizing latency across each processing step. To ensure user privacy and data security, the system avoids storing any content, working purely in memory with encrypted API keys and no permanent databases. It is lightweight enough to run on standard hardware configurations, with optional GPU support for faster model inference. In essence, this system brings together advanced machine learning, real-time data engineering, and thoughtful UI/UX design to create a tool that helps users understand digital trends as they unfold. Its modular architecture allows for easy integration of additional data sources or future improvements in model accuracy and processing speed. The rule-based trend classification enhances transparency and interpretability, while the use of cutting-edge NLP models ensures depth in sentiment detection. This blend of technical rigor and user-friendly interaction positions the system as a robust solution for media monitoring, public opinion tracking, and event response in an increasingly dynamic digital world.

VII. Result and Discussion



The developed web application, Live Trending Analyzer, successfully demonstrates real-time visualization and classification of trending topics across multiple domains including news, books, movies, music, apps, fashion, games, and technology. The interface is intuitive, category-driven, and dynamically displays data fetched via APIs and scraping techniques. Results are visually represented in card-based layouts which enhance content readability and user interaction. The News section, as shown in the first result screen, presents up-to-date headlines using a clean, light-themed interface. Trending news items are fetched and rendered as clickable cards containing thumbnails and headlines. Each card is responsive and designed to maintain readability across devices. The headlines are automatically truncated to maintain layout consistency while ensuring that users get a brief yet clear idea of the story content. The Books section, depicted in the second result screen, highlights trending and top-rated books. This section uses a dark theme to differentiate the content contextually and support better visual focus. Book covers along with their titles are displayed in a horizontally scrollable format or grid, enabling efficient browsing. The dark UI enhances the color contrast of book covers, making the display more appealing.

A unique feature implemented in the movies section (integrated with the books view) is the search and sentiment analysis functionality using RottenTomatoes. Users can input a movie name, and the application fetches reviews and performs real-time sentiment analysis. This interaction bridges trending detection with sentiment classification—aligning with the core objective of this research. The sentiment result helps users determine the overall public opinion around a movie, supported by machine learning-based classification of review polarity (positive, negative, or neutral).

Overall, the application has proven effective in providing an integrated environment to track, visualize, and analyze trends across domains. Its modular design ensures easy scalability to other categories or data sources. User interface feedback also suggests high usability and engagement, with distinct visual styling (light/dark themes), responsive cards, and actionable search inputs that enhance the analytical experience for users.

VIII. Applications

The system helps journalists and media outlets track breaking news and public responses across various social media platforms. It provides instant insights into which stories are gaining traction and the general sentiment of the audience, making it easier for media professionals to prioritize coverage and respond quickly to emerging developments. Companies use this technology to monitor brand mentions and customer opinions in real time. By analyzing public sentiment, businesses can engage proactively with their customers, protect their reputation, and address issues or capitalize on positive feedback promptly, ultimately enhancing customer satisfaction and loyalty. Public agencies and government bodies utilize the system to understand how citizens react to new policies, legislation, or public announcements. By analyzing trending topics and sentiments, they can tailor their communication strategies to better meet public needs and improve the effectiveness of their programs. In times of crisis such as natural disasters or political unrest, the system identifies the most pressing concerns and public emotions. This enables authorities and relief organizations to respond more swiftly, allocate resources effectively, and communicate important information to affected communities in real time. Researchers benefit from access to continuously updated, preprocessed social data that reflects public discourse on various topics. This real-time data aids in studies related to linguistics, social behavior, psychology, and political science, saving time on data cleaning and improving the reliability of their findings. Financial analysts and investors monitor social sentiment surrounding companies, products, and economic events. By detecting shifts in public opinion early, they can anticipate market trends and make more informed investment decisions ahead of traditional financial indicators. Event organizers and marketers use the system to gather immediate feedback on public reaction to campaigns, product launches, or large-scale events. This feedback helps assess the success of initiatives and guides improvements for future projects, ensuring better audience engagement and impact.

IX. Conclusion

The real-time trend analysis and sentiment classification system presented here offers a powerful tool for monitoring public discourse across multiple social media platforms. By leveraging advanced machine learning models and efficient data processing techniques, the system provides timely and accurate insights into emerging trends and public opinions. Its ability to integrate data from diverse sources, perform sentiment analysis, and visualize results in an intuitive dashboard makes it valuable for a wide range of applications—from media monitoring and brand management to crisis response and market research. While current challenges such as model limitations and dependency on external APIs exist, the system's privacy-focused, real-time architecture ensures it remains both effective and scalable. Overall, this approach represents a significant advancement in automated social media analysis, empowering users with actionable intelligence to better understand and respond to the dynamic flow of information in today's digital world.

REFERENCES

- Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. Science, 333(6051), 1878-1881.
- Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 492-499.
- Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I.
- M. (2010). Predicting elections with Twitter: What 140 characters reveal about political sentiment. Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 178-185.
- Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: Real-time event detection by social sensors. Proceedings of the 19th International Conference on World Wide Web, 851-860.
- O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. Proceedings of the Fourth International AAAI Conference on Weblogs and
- Social Media, 122-129.
- Hu, Y., Manikonda, L., & Kambhampati, S. (2014). What we Instagram: A first analysis of Instagram photo content and user types. Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media.
- Zubiaga, A., Liakata, M., & Procter, R. (2016). Learning reporting dynamics during breaking news for rumour detection in social media. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics.
- Garofalo, M. F., & Paredes, A. R. L. T. Z. (2020). Predicting Video Virality in YouTube: A Machine Learning Approach. Journal of Computer-Mediated Communication.
- Liu, X., Zhou, Z., & Zhang, L. (2021). Fashion Trend Prediction Using Machine Learning Algorithms. Journal of Fashion Technology & Textile Engineering.
- Zhang, L. P., Sun, M. X., & Li, H. Y. (2021). Trend Detection in Music Streaming Platforms Using Machine Learning. Data Mining and Knowledge Discovery.
- Ahmad, H. H. Z., Al-Khatib, N., & Alwan, M. A. S. (2021). Trend Detection and Analysis in Twitter Data Using Machine Learning Approaches. Journal of Computational Science.
- Patel, R. J., & Desai, M. S. (2022). Trend Detection in Social Media: Twitter vs. YouTube. Computers in Human Behavior.
- Zhang, F. Z., Wu, L. F., & Liu, Y. (2022). Trend Detection in Music Streaming Platforms Using Machine Learning. Journal of Music Technology & Education.
- Chen, L., & Zhang, M. (2023). Cross-Platform Trend Detection: Using Machine Learning to Track Trends in Twitter and YouTube. Information Systems Frontiers.
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1-8.