IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15599

AI-Foresight: Market Research and Trend Analyzer

Akshata Panchal[1], Vidya S[2]

Department of MCA, BIT, K.R. Road, V.V. Pura, Bangalore, India[1,2]

Abstract: In the modern digital economy, organizations generate and consume massive volumes of unstructured textual data through online news portals, blogs, forums, and social media platforms. Extracting meaningful insights from such data using traditional market research techniques is time-consuming, expensive, and often inefficient. This paper presents AI-Foresight – Market Research and Trend Analyzer, an AI-driven platform developed to automate market intelligence generation using Natural Language Processing (NLP), Machine Learning (ML), and Generative AI techniques. The proposed system performs automated data collection, text preprocessing, sentiment analysis, topic modelling, trend detection, and AI-based summarization to transform raw textual information into actionable business insights. The platform integrates tools such as BeautifulSoup, NLTK, spaCy, BERTopic, VADER, TextBlob, and GPT-based summarization models, while presenting analytical outputs through an interactive Streamlit dashboard with dynamic visualizations. Experimental evaluation demonstrates that the system efficiently processes large-scale textual datasets, accurately identifies market trends, and significantly reduces manual analytical effort. The proposed solution provides organizations, researchers, and business strategists with a scalable and intelligent framework for real-time market analysis and strategic decision-making.

Keywords: Artificial Intelligence, Market Research, Trend Analysis, Natural Language Processing, Sentiment Analysis, Topic Modelling, Generative AI, Business Intelligence.

I. INTRODUCTION

In today's rapidly evolving digital environment, organizations are continuously exposed to enormous volumes of data generated through social media platforms, blogs, online news portals, customer reviews, and public discussion forums. This data contains valuable information related to consumer behaviour, market trends, technological advancements, and business competition. However, because most of this information exists in unstructured textual form, traditional analytical systems struggle to process and interpret it effectively.

Conventional market research techniques depend heavily on manual analysis, requiring analysts to individually review thousands of articles, reports, and customer opinions. Such approaches are highly time-consuming, expensive, and vulnerable to inconsistencies and human error. In fast-moving business environments, delays in understanding emerging trends or customer sentiment can lead to missed opportunities and poor strategic decisions.

To address these challenges, this paper presents AI-Foresight: Market Research and Trend Analyzer, an intelligent AI-driven platform designed to automate market intelligence generation using advanced Artificial Intelligence and Natural Language Processing techniques. The system integrates automated data collection, text preprocessing, sentiment analysis, topic modelling, trend identification, and AI-powered summarization into a unified framework.

The proposed system transforms large volumes of unstructured textual information into meaningful and actionable insights through interactive dashboards and analytical visualizations. By automating market research workflows, AI-Foresight improves efficiency, reduces manual effort, enhances analytical accuracy, and enables organizations to make faster and more informed strategic decisions.

1.1 Project Description

AI-Foresight is a web-based AI-powered market intelligence platform developed to automate the complete lifecycle of market research and trend analysis. The system continuously gathers textual data from online sources such as news portals, blogs, and public platforms using web scraping and API-based techniques.

The collected data undergoes preprocessing operations including tokenization, stop-word removal, normalization, and lemmatization to convert raw text into structured and machine-readable formats. Machine learning and NLP techniques are then applied to identify dominant themes, emerging trends, and public sentiment. The platform further utilizes Generative AI models for automated summarization and strategic insight generation.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15599

The final analytical outputs are displayed through an interactive Streamlit dashboard featuring charts, graphs, trend indicators, and sentiment visualizations. The system significantly reduces the complexity and time associated with traditional market research while delivering real-time and reliable business intelligence.

1.2 Motivation

Modern organizations operate in highly dynamic and competitive environments where timely market intelligence plays a critical role in decision-making. Traditional research methods are inefficient for processing the massive volumes of digital content generated every day.

The motivation behind AI-Foresight is to develop a scalable and intelligent system capable of automatically identifying market trends, monitoring public sentiment, and generating meaningful insights without requiring extensive manual intervention. By leveraging AI and NLP technologies, the system enables organizations to respond quickly to market changes, improve strategic planning, and gain competitive advantages.

II. RELATED WORK

Paper [1] discusses traditional market analytics systems such as Google Trends, which are commonly used to monitor keyword popularity and seasonal trends. While useful for trend visualization, these systems provide only surface-level insights and lack contextual understanding and intelligent recommendation capabilities.

Paper [2] explores social listening platforms such as Brand watch and Talk walker that perform sentiment analysis and social media monitoring. Although these systems efficiently collect data and identify public sentiment, they provide limited AI-driven summarization and strategic interpretation.

Paper [3] presents transformer-based language models such as GPT that demonstrate strong capabilities in contextual understanding, summarization, and text generation. However, these models are generally used as standalone tools rather than integrated components within complete market intelligence pipelines.

Paper [4] focuses on academic research involving topic modelling techniques such as Latent Dirichlet Allocation (LDA) and BERTopic for identifying hidden patterns in textual datasets. While these techniques effectively extract trends and themes, most implementations remain theoretical and are rarely integrated into real-time production systems.

Paper [5] examines business intelligence tools such as Tableau and Power BI that provide interactive dashboards and visualization capabilities. These tools depend heavily on manual preprocessing and external analytical pipelines, limiting automation and intelligent insight generation.

The proposed AI-Foresight system addresses these limitations by integrating automated data collection, NLP preprocessing, topic 766odelling, sentiment analysis, Generative AI summarization, and dashboard visualization into a unified and intelligent market analysis framework.

III. METHODOLOGY

A. System Environment

The AI-Foresight system operates within a web-based and AI-enabled environment developed using Python and modern NLP frameworks. The backend processing layer utilizes machine learning libraries and text analytics frameworks to process and analyze large-scale textual datasets.

The development environment consists of Python 3.x, Streamlit, BeautifulSoup, NLTK, spaCy, BERTopic, VADER, TextBlob, Plotly, and Hugging Face Transformers. The system is platform-independent and supports deployment on Windows, macOS, and Linux operating systems.

The visualization environment is implemented using Streamlit and Plotly to provide interactive dashboards for displaying market trends, topic clusters, sentiment scores, and AI-generated summaries.

B. System Architecture

The AI-Foresight system follows a layered architecture consisting of multiple interconnected modules:

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15599

1. Data Source Layer
This layer collects textual data from online news portals, blogs, discussion forums, and social media platforms.
2. Data Collection Layer
This layer utilizes web scraping frameworks and APIs to retrieve large volumes of structured and unstructured textual data.
3. Data Preprocessing Layer
The preprocessing module performs tokenization, stop-word removal, normalization, lemmatization, punctuation removal, and noise filtering.
4. Analytics and Intelligence Layer
This layer applies NLP and machine learning algorithms for topic modeling, trend identification, clustering, and sentiment analysis.
5. AI Summarization Layer
Generative AI models are used to create concise summaries and actionable business insights from analyzed datasets.
6. Presentation and Visualization Layer
The processed outputs are displayed through interactive dashboards using charts, graphs, trend indicators, and sentiment visualizations.

C. Workflow of the System

1. Collect textual data from online sources.
2. Preprocess and clean the collected data.
3. Convert text into vectorized numerical representations.
4. Apply topic modeling and sentiment analysis.
5. Generate AI-driven summaries and insights.
6. Visualize analytical results through dashboards.
7. Deliver actionable market intelligence to users.

D. Hardware and Software Requirements

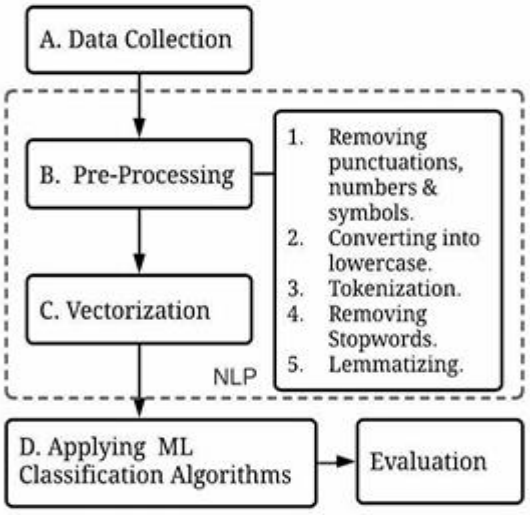
- **Hardware:** Intel Core i5/Apple M2 processor, minimum 8 GB RAM, 512 GB SSD, optional GPU support.
- **Software:** Python 3.x, BeautifulSoup, NLTK, spaCy, BERTopic, VADER, TextBlob, Streamlit, Plotly, and Hugging Face Transformers.

IV. SYSTEM IMPLEMENTATION

The implementation of AI-Foresight is divided into multiple functional modules responsible for data acquisition, preprocessing, analysis, and visualization.

A. Data Collection Module

The data collection module retrieves textual information from online sources using BeautifulSoup and API-based integrations.



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

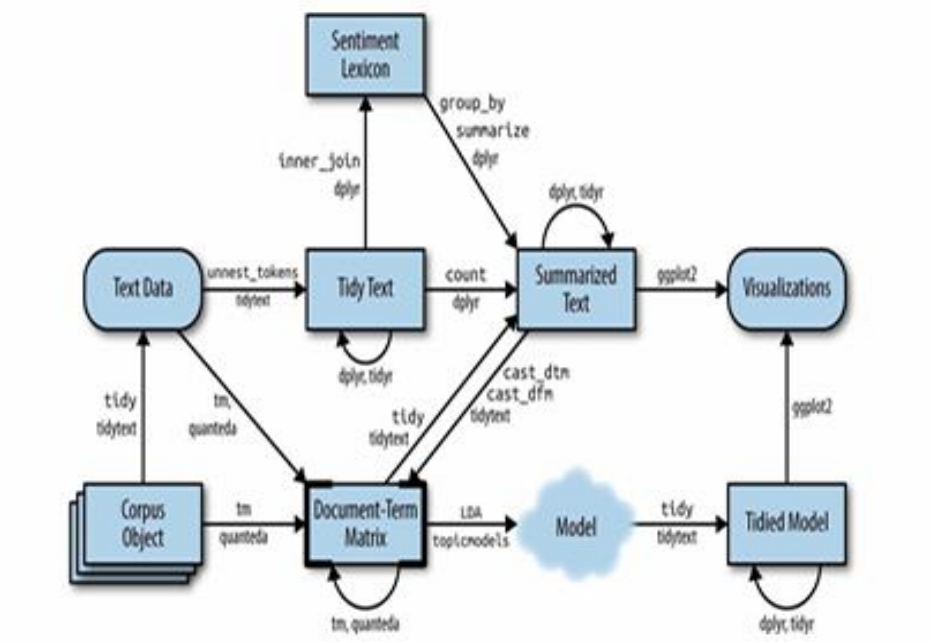
International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15599

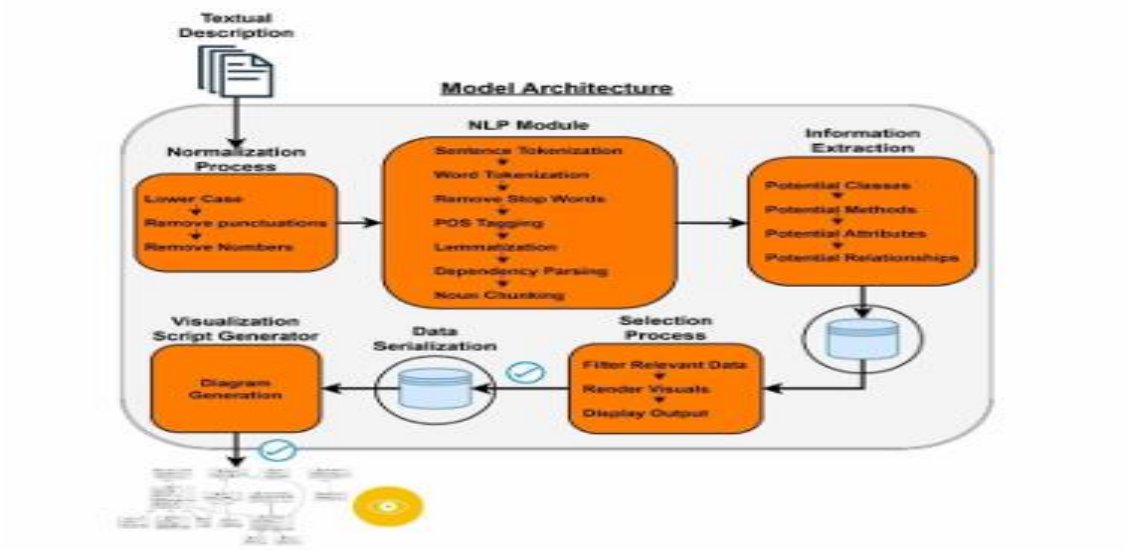
B. NLP Preprocessing Module

The preprocessing module removes unwanted characters, stopwords, punctuation, and performs text normalization and lemmatization.



C. Machine Learning Module

The machine learning module performs classification and analytical processing using logistic regression and NLP-based vectorization.



D. Dashboard Visualization

- Displays trending keywords, sentiment analysis charts, topic clusters, and data source distribution.
- Provides statistical visualizations and AI-generated summaries for better market insight analysis.

V. RESULTS AND DISCUSSION

The AI-Foresight system successfully demonstrated the capability to automate market research and trend analysis using AI and NLP technologies. The platform efficiently collected and processed large-scale textual datasets obtained from multiple online sources.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15599

The preprocessing module effectively removed noise and converted unstructured text into machine-readable formats suitable for analysis. Topic modeling algorithms accurately identified emerging market trends and thematic patterns, while sentiment analysis models successfully classified public opinion into positive, negative, and neutral categories.
The integration of Generative AI enabled the automatic generation of concise summaries and strategic insights from complex analytical outputs. Interactive dashboards developed using Streamlit and Plotly provided clear visualization of trends, sentiment distributions, and statistical analytics.

Experimental testing confirmed that the system reduced manual analytical effort, improved processing speed, and delivered accurate and meaningful market intelligence suitable for business decision-making.

The implementation results demonstrate that AI-Foresight provides a scalable and intelligent framework for automated market analysis, strategic planning, and business intelligence generation.

Fig 1 Fetching Top Keywords

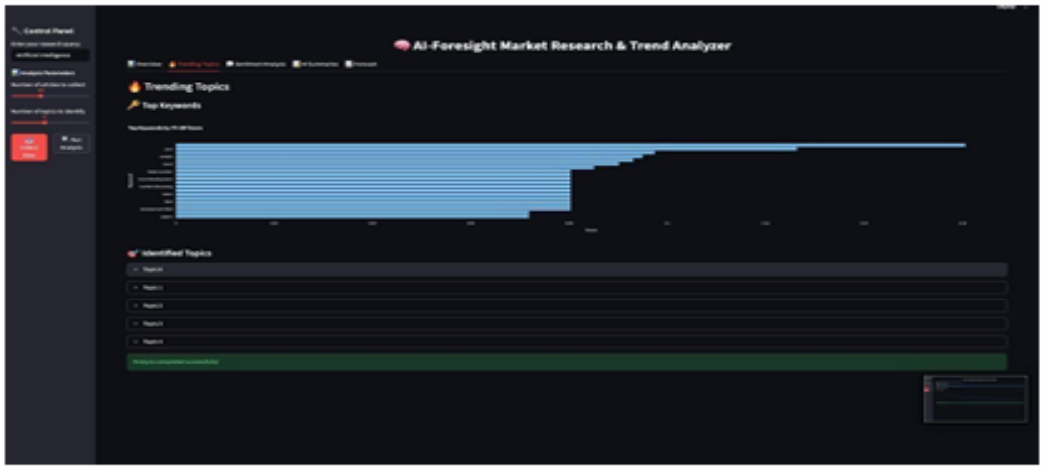
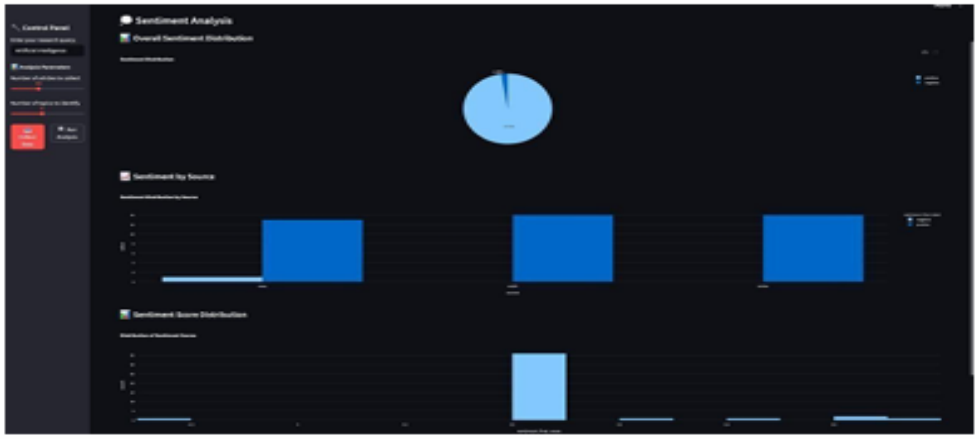


Fig 2 Sentiment Analysis



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15599

Fig 3 Data Source Distribution.

Fig 4 Statistical Analysis using Pie Chart.

VI. CONCLUSION

AI-Foresight: Market Research and Trend Analyzer demonstrates the practical application of Artificial Intelligence, Natural Language Processing, and Machine Learning in transforming traditional market research methodologies into intelligent automated systems.

The proposed platform successfully integrates automated data collection, NLP preprocessing, sentiment analysis, topic modelling, AI-driven summarization, and dashboard visualization within a unified framework. By automating the complete workflow of market intelligence generation, the system significantly reduces manual effort, improves analytical efficiency, and enhances the accuracy of decision-making.

The project highlights the ability of AI-driven systems to process and interpret massive volumes of unstructured textual data while generating actionable business insights in real time. The experimental outcomes confirm that AI-Foresight is reliable, scalable, and suitable for real-world applications in business intelligence, strategic planning, and market trend analysis.

VII. FUTURE WORK

Future enhancements of AI-Foresight can significantly improve the predictive capabilities, scalability, and analytical intelligence of the platform.

Advanced deep learning architectures such as Long Short-Term Memory (LSTM) networks and Transformer-based models can be integrated to improve temporal trend forecasting and predictive analytics.

The platform can also support multilingual NLP processing to analyze content from multiple languages and global markets. Real-time alert systems can be implemented to notify users about sudden market shifts, emerging opportunities, and sentiment changes.

Cloud-based deployment and distributed processing frameworks can further improve scalability, reliability, and performance for enterprise-level applications.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.15599

Future versions of the system may also integrate advanced business intelligence tools, enhanced visualization techniques, and enterprise-grade security mechanisms to support large-scale organizational deployments.

REFERENCES

[1]. S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*, O'Reilly Media, 2009.
[2]. C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text," in *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 2014.
[3]. J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of NAACL-HLT*, 2019.
[4]. T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," in *Proceedings of ICLR Workshops*, 2013.
[5]. M. Grootendorst, "BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure," 2022. [Online]. Available: https://maartengr.github.io/BERTopic/
[6]. Hugging Face Inc., "Transformers Documentation," Hugging Face, 2024. [Online]. Available: https://huggingface.co/docs/transformers
[7]. Streamlit Team, "Streamlit Documentation," Streamlit, 2024. [Online]. Available: https://streamlit.io
[8]. Plotly Technologies Inc., "Plotly Python Graphing Library Documentation," Plotly, 2024. [Online]. Available: https://plotly.com/python/
[9]. BeautifulSoup Documentation, "Beautiful Soup 4 Documentation," 2024. [Online]. Available: https://www.crummy.com/software/BeautifulSoup/bs4/doc/
[10]. spaCy Team, "spaCy Industrial-Strength Natural Language Processing," 2024. [Online]. Available: https://spacy.io
[11]. NLTK Project, "Natural Language Toolkit Documentation," 2024. [Online]. Available: https://www.nltk.org
[12]. Scikit-learn Developers, "Scikit-learn: Machine Learning in Python," 2024. [Online]. Available: https://scikit-learn.org
[13]. T. Brown et al., "Language Models are Few-Shot Learners," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, 2020.
[14]. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
[15]. Python Software Foundation, "Python Documentation," 2024. [Online]. Available: https://docs.python.org/3/

