iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** I **Month of publication:** January 2025

DOI: https://doi.org/10.22214/ijraset.2025.66465

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue I Jan 2025- Available at www.ijraset.com

Automating Sentiment Analysis for Business Intelligence

S Kabilesh[1], Sanjay S Srivathsa[2], Pranav V[3], Shivaraj D[4]
Dept. of Computer Science and Engineering RNS Institute of Technology, Bangalore

Abstract: ***The growing reliance on online reviews for con- sumer decision-making has underscored the need for efficient tools to interpret and analyze user sentiments. This paper explores the application of Sentiment Analysis to iPhone reviews collected from multiple nations, leveraging advanced natural language processing techniques. The study employs Power BI for summarizing sentiment data into visually interpretable graphs and key performance indicators (KPIs), enabling faster and more informed decision-making for customers. A comparative evaluation of sentiment trends across diverse geographic regions provides valuable insights into user satisfaction and product perception. Additionally, the challenges of handling multilingual data, ensuring unbiased analysis, and addressing data privacy concerns are discussed to provide a comprehensive perspective on sentiment analysis in a global context.***

Index Terms: Sentiment Analysis, Power BI, Natural Lan- guage Processing, Consumer Insights, Multilingual Data, Data Visualization

I. INTRODUCTION

Advanced data analysis has gone up unprecedentedly with the techniques like sentiment classification, opinion mining, and predictive analytics. Traditional measures of text analysis are effective to a certain extent but fail to cope up with the changing complexity of natural language processing tasks. AI offers a paradigm shift in enabling proactive and adaptive mechanisms that can cope up with dynamic and diverse data. AI in automating sentiment analysis depends on techniques of machine learning, natural language processing, integration of PowerBI to support the enhancement of sentiment detection and inferring customer behaviors. This research discusses the assimilation of AI into frameworks so that its competencies and limitation can be accessed. This project analyze the IPhone data and develop interactive PowerBI dashboard along with AI chatbot which enhance users interaction.

II. RELATED WORK

The field of sentiment analysis has been widely studied, with numerous advancements in both the methodologies used and their applications in business intelligence. This section explores key studies related to sentiment analysis, focusing on the integration of machine learning, deep learning, and AI- powered systems to enhance business intelligence.

A. Traditional Approaches to Sentiment Analysis

The majority of the methods used by early sentiment analysis systems were lexicon- and rule-based. These systems examined texts according to the frequency and existence of predetermined lists of terms that were classified as neutral, negative, or positive. Although these techniques worked well for small datasets, they had trouble capturing subtleties like context, idioms, and sarcasm. Sentence structure and word order were not taken into consideration by these early sys- tems, which limited their comprehension of intricate linguistic aspects.

- Pang et al. (2002) developed a sentiment classifier using a bag-of-words approach.
- Turney (2002) used semantic orientation to detect positive and negative opinions in reviews.

B. Machine Learning Developments for Sentiment Analysis

NLTK and similar packages improved sentiment classifica- tion by allowing the learning of patterns from large labeled datasets. However, these methods still needed a lot of feature engineering, including the extraction of named entities and n- grams, and could not handle semantic connections within the text.

- B. Liu and L. Zhang (2012) explored techniques for review mining, focusing on product reviews and social media comments.

C. Integration of Sentiment Analysis with Business Intelli- gence

By combining sentiment analysis with business intelligence tools like Power BI organizations can display sentiment data and make

©IJRASET: All Rights are Reserved

data-driven choices. By employing visualization technologies to present sentiment data in an intelligible man- ner, businesses may enhance their marketing, product devel- opment, and customer contact strategies.

- Chaurasia et al. (2020) explored the use of Power BI for sentiment analysis in customer service feedback, showing how automated sentiment categorization and visualization optimize business strategies.

D. Chatbots and User Interaction in Sentiment Analysis

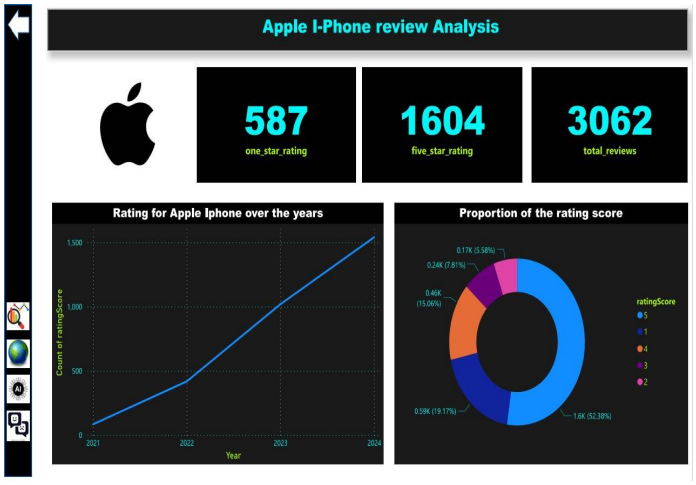
In this contemporary era, AI has revolutionized especially for improving customer experiences and making data dis- covery easier. Businesses may engage in dynamic consumer interactions with these chatbots, which offer real-time data based on sentiment analysis. By leveraging AI-driven sentiment analysis, businesses can gain deeper insights into customer emotions, preferences, and feedback. This empowers organizations to tailor their marketing strategies, improve product offerings, and provide personalized experiences. Real-time sentiment tracking en- ables proactive decision-making, helping businesses address customer concerns swiftly and enhance overall satisfaction.

E. Analyzing Opinions in Product Reviews and Online Plat- forms

Sentiment analysis has been effectively used to the study of product evaluations and social media posts, which are abundant sources of unstructured customer feedback. Large volumes of data are provided by social networking sites like Facebook and Twitter as well as review sites like Yelp and Amazon, which can be analysed to determine consumer mood. Sentiment analysis libraries have made the analysis process easier, but deep learning models like BERT and LSTM have shown great promise in classifying sentiment in these domains.

- Goh et al. (2020) developed a sentiment analysis frame- work using social media posts to gauge consumer senti- ment about brands.

In summary, the accuracy and scalability of sentiment analy- sis have been greatly improved by its progression from simple lexicon-based approaches to sophisticated machine learning and deep learning techniques. By combining sentiment analy- sis with AI-powered chatbots and business intelligence tools, companies can extract useful insights from massive amounts of textual data, revolutionizing the way that consumer input is used to inform decisions.



Analytical Page

III. METHODOLOGY

Data collection, preprocessing, sentiment classification, data aggregation, and chatbot interaction are all integrated into the multi-step process of the suggested sentiment analysis system for the iPhone sales dataset. The objective is to generate meaningful insights for business intelligence by automating the examination of customer input via surveys, reviews, and social media. This section describes the approach taken in the system's development, along with the important phases that were involved.

A. Data Collection

The methodology's initial step, data collecting, entails ob- taining client input from various sources. These sources in- clude

customer surveys, product reviews from sites like Ama- zon and Yelp, and social media sites. Web scraping methods and APIs are used to gather the data, which guarantees that it is pertinent to the iPhone sales dataset. After being gathered, the data is kept in a centralized database where it may be accessed for further processing and analysis.

- Reviews are fetched from popular e-commerce platforms like Amazon and Best Buy through their public APIs or web scraping.
- Customer survey responses are gathered using tools like online Forms and stored in a structured format.

B. Data Preprocessing

Data preprocessing is a vital process in sentiment analysis. The preprocessing pipeline has numerous stages designed to remove noise, normalize the data, and convert it into a form suitable for sentiment classification.

- Noise Removal: Special characters, URLs, and stopwords (commonly used words like ”the,” ”and,” ”is”) are re- moved from the text to avoid unnecessary noise.
- Tokenization: Text is split into smaller units (tokens), such as words or phrases, that are easier to analyze.
- Normalization: Text is converted to lowercase, and spelling variations are corrected to maintain consistency across the dataset.
- Lemmatization: Words are simplified to a common form for understanding the context(for example, ”running” becomes ”run”).

Preprocessing also involves filtering out irrelevant or dupli- cate data points to ensure that only high-quality data is used in the analysis.

C. Sentiment Classification

The sentiment classification model, which gives each textual item a sentiment label (positive, negative, or neutral), is the central component of the sentiment analysis system. The system utilizes a pre-trained model, like BERT, that has been optimized for sentiment analysis using the dataset of iPhone sales.

D. Data Aggregation and Visualization

Following sentiment classification, the outcomes are com- bined to produce insightful information. In this step, the sentiment data is arranged according to a number of crite- ria, including time periods, product categories (such iPhone models), or geographical areas.

- Aggregation: Sentiment results are aggregated by fac- tors such as time (weekly, monthly) or specific product features (battery life, camera quality). This allows for the identification of trends and patterns in customer sentiment over time.
- Visualization: Data is visualized using business intel- ligence tools like Power BI. Sentiment trends, such as the overall sentiment score for a specific product or feature, are presented in easily interpretable charts, graphs, and dashboards. These visualizations help stake- holders quickly understand customer sentiment and make informed decisions.

Stakeholders can examine the sentiment data in real time by using interactive visualizations made using Power BI dash- boards. It is possible to create these dashboards to highlight particular features of the dataset, such as sentiment trends for particular product models or customer worries on a particular feature.

E. Chatbot Integration

The interactive chatbot, which enables users to query the sentiment analysis results in real-time, is a crucial component of the system. In order to respond to inquiries regarding the attitude surrounding iPhone sales, the chatbot is constructed using a refined GPT-3.5 model that has been trained on domain-specific data.

- Chatbot Development: The GPT-3.5 model is fine-tuned with a dataset of customer queries and sentiment-related information, enabling the chatbot to provide context- aware and accurate responses.
- Real-Time Queries: Users can interact with the chatbot to ask specific questions such as ”What is the overall sentiment trend for the latest iPhone model?” or ”How do customers feel about battery life?” The chatbot retrieves insights from the database and provides answers based on the aggregated sentiment data.
- Feedback Loop: The chatbot also allows users to provide feedback on sentiment predictions, which can be used to improve the model over time. This feedback loop ensures that the system continues to learn and adapt to changing customer sentiment.

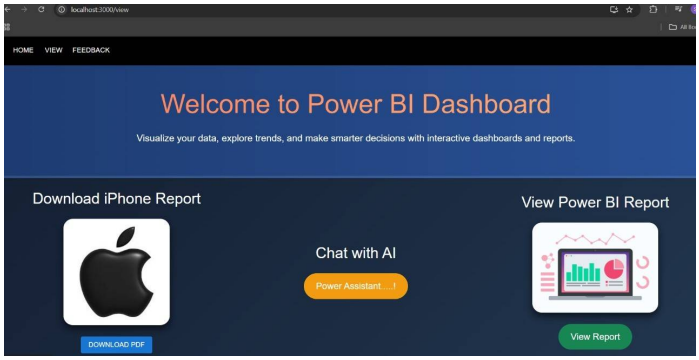
F. Automation and Scalability

To guarantee scalability, the entire procedure is automated, from data gathering to chatbot interaction. Data collection and processing, sentiment analysis, result aggregation, and visualization are all done using automated pipelines. Because of this

automation, the system can manage massive data quantities and grow to meet the expanding demands of the company.

- Automation: The use of automation tools ensures that the sentiment analysis process is fast and efficient, requiring minimal manual intervention.
- Scalability: The system is designed to scale with the growth of the iPhone sales dataset, making it adaptable to different product categories and markets.

The methodology aimed to provide a comprehensive evalu- ation of AI's efficacy in addressing cybersecurity challenges.



View Page

IV. EXPERIMENTAL SETUP

The hardware and software elements of the sentiment anal- ysis system, as well as the datasets and evaluation criteria that are used to gauge the system's effectiveness, make up the experimental setup. This section outlines the tools, technolo- gies, and resources needed for implementation as well as the environment in which the system was created and tested.

A. Hardware Specifications

The experiments were conducted on a system with the following hardware specifications, designed to handle large- scale data processing and machine learning model training efficiently:

- CPU: Intel Core i7 (8th Gen) with 8 cores and 16 threads, capable of handling multi-threaded operations and processing large datasets.
- RAM: 16 GB DDR4, ensuring smooth data preprocess- ing, model training, and inference tasks.
- GPU: NVIDIA GeForce GTX 1660 Ti with 6 GB GDDR6, used to accelerate deep learning tasks, such as fine-tuning the sentiment analysis model and running inference on large datasets.
- Storage: 512 GB SSD, providing fast data read/write speeds, essential for handling large amounts of textual data and model checkpoints.

The sentiment analysis system may be effectively experi- mented with and refined thanks to the hardware configuration that is tuned for both the data preparation pipeline and the deep learning model training process.

B. Software and Tools

The software environment is made to facilitate business in- telligence, data preprocessing, and machine learning processes. The sentiment analysis system was developed and assessed using the following tools and libraries:

Programming Language

Python 3.8 was used for devel- oping the sentiment analysis model, data preprocessing, and chatbot functionality. Python's extensive libraries and frameworks make it suitable for handling NLP and machine learning tasks.

NLP Libraries

- Transformers: The Hugging Face Transformers li- brary was used to load and fine-tune pre-trained models like BERT and GPT-3.5 for sentiment clas- sification.

- spaCy: Used for tokenization, lemmatization, and text preprocessing tasks.
- NLTK: The Natural Language Toolkit (NLTK) was used for additional text processing, such as stopword removal and part-of-speech tagging.

Machine Learning Frameworks

- TensorFlow: TensorFlow was used to build and train deep learning models for sentiment classification, providing an efficient and scalable framework for model deployment.
- PyTorch: PyTorch was used as an alternative deep learning framework for training the models, known for its flexibility and ease of use in research settings.

Business Intelligence Tool

– Power BI: Power BI was used to create interactive visualizations and dashboards, allowing users to ex- plore sentiment trends and insights in real-time.

Database Management

– MySQL: MySQL was used as the primary database for storing raw data, sentiment analysis results, and aggregated insights, ensuring efficient querying and data management.

Chatbot Development Framework:

– Rasa: Rasa was used to develop the interactive chatbot, providing an open-source framework for building conversational AI.

These tools and frameworks were integrated to create a comprehensive sentiment analysis system, ensuring efficient data processing, model training, and user interaction.

C. Dataset

Customer input about iPhone sales, collected from surveys, product reviews, and social media, makes up the main dataset utilized for sentiment analysis. Prior to being entered into the system for analysis, the dataset underwent pre-processing and cleaning. Three primary categories comprise the dataset:

- Social Media Posts: Data collected from Twitter, Face- book, and other social media platforms using their respec- tive APIs. This data consists of user posts, comments, and hashtags related to iPhones.
- Product Reviews: Customer reviews from e-commerce platforms like Amazon, Best Buy, and Walmart. These reviews contain detailed feedback on various iPhone models, focusing on aspects such as performance, battery life, and camera quality.
- Surveys: Structured survey responses gathered from cus- tomers, providing insights into their satisfaction with iPhones and the factors influencing their purchase deci- sions.

Based on the feedback's content, the dataset is cleaned and tagged for sentiment classification (positive, negative, and neutral). To assess model performance during training, the training set is further separated into a training subset and a validation subset.

D. Evaluation Metrics

Standard classification measures, which offer information on the precision, dependability, and efficacy of the system in predicting sentiment labels, are used to assess the sentiment analysis system's performance. The evaluation metrics listed below are employed:

- Accuracy: Measures the percentage of correctly pre- dicted sentiment labels across all text samples in the dataset.
- F1-Score: The harmonic mean of precision and recall, providing a balanced metric for evaluating model perfor- mance, especially when dealing with imbalanced classes.
- Confusion Matrix: A matrix shows the true positive and negative value out of the total reviews collected from the users.

These metrics are calculated during both the training and testing phases to assess the system's ability to classify senti- ment accurately and effectively.

E. Experiment Workflow

The workflow for the experimental setup includes the fol- lowing steps:

- Data Collection: Gather customer feedback data from social media, product reviews, and surveys.

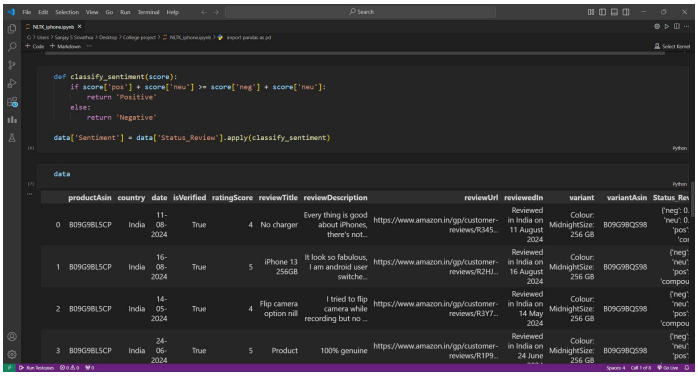
- Data Preprocessing: Clean and preprocess the data by removing noise, tokenizing text, and normalizing it for sentiment classification.
- Model Training and Fine-Tuning: Use pre-trained mod- els (e.g., BERT, GPT-3.5) and fine-tune them on the iPhone-specific dataset to classify sentiment.
- Data Aggregation and Visualization: Aggregate senti- ment data and generate visualizations using Power BI.
- Chatbot Integration: Implement a chatbot using fine- tuned GPT-3.5 to provide real-time sentiment insights.
- Evaluation: Evaluate the performance of the system using the predefined metrics and fine-tune the model if necessary.

V. RESULTS AND ANALYSIS

This section offers a thorough evaluation of the sentiment analysis system's performance along with the results it pro- duced. Information gleaned from the interactive chatbot and visualizations, along with potential future improvements and places for improvement, the system's efficacy in assessing customer sentiment for iPhone sales is examined.

A. Calculating Polar Scores

The SentimentIntensityAnalyzer library is used to assess the client reviews. The library provides a sentiment classification approach that assigns a numerical value to a review's positive, negative, and neutral aspects. We classify the statement as either positive or negative based on a threshold value for implication based on the score mentioned above. Power BI is used to import the dataset for additional sentiment analysis and visualisation.



Calculating Status of Sentiment

B. Sentiment Trend Analysis and Visualizations

To find trends over time and across several product cate- gories, sentiment analysis findings were combined and shown using Power BI. The following observations were made:

- Positive Sentiment Trend: There was a noticeable spike in positive sentiment during the release of new iPhone models, especially in the months following the announce- ment. This indicates that customers generally have a favorable reaction to new product launches.
- Negative Sentiment Trend: Negative sentiment in- creased following reports of issues with battery life and software glitches. This indicates that customers were dissatisfied with certain features and performance issues, especially in the first few months of product usage.
- Neutral Sentiment Trend: Neutral sentiment was con- sistently present, particularly in reviews where customers mentioned both pros and cons without a strong leaning toward positive or negative feedback.

The visualizations enabled stakeholders to quickly identify sentiment patterns and make informed decisions regarding product improvements and marketing strategies.

C. Chatbot Interaction Results

The effectiveness of the interactive chatbot, which was driven by a refined GPT-3.5 model, in responding to user inquiries about sentiment patterns was assessed. The chatbot was able to provide precise answers to questions like:

- ”What is the overall sentiment trend for iPhones?” The chatbot provided an accurate summary of sentiment over the past year, highlighting spikes in positive senti- ment following new releases.
- ”How do customers feel about the iPhone battery life?” The chatbot successfully identified negative sen- timent associated with battery performance issues and provided specific details from customer reviews.
- ”What are the main reasons for customer dissatis- faction?” The chatbot retrieved relevant feedback about issues such as overheating, slow performance, and battery drain.

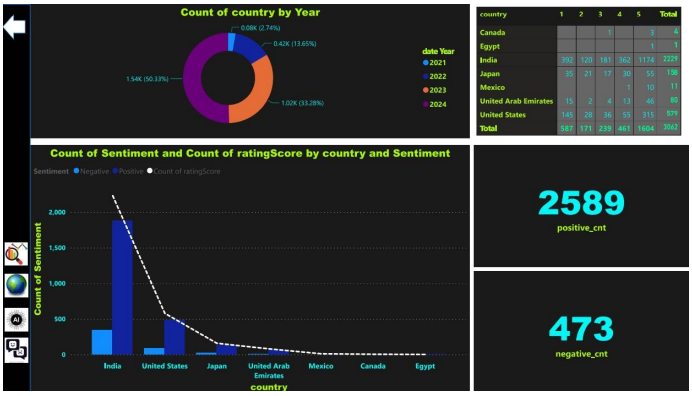
User feedback indicated that the chatbot was intuitive and effective in providing insights, allowing non-technical users to explore sentiment data easily.

D. Error Analysis

Despite its great accuracy, the sentiment analysis system had certain shortcomings, especially when it came to complicated or confusing words. The following situations had the highest frequency of these errors:

- Sarcasm and Irony: The model sometimes misclassified sarcastic or ironic comments as neutral or positive, as it struggled to detect the underlying sentiment in such contexts.
- Contextual Ambiguity: Sentences where sentiment was context-dependent (e.g., ”I love the iPhone, but it drains quickly”) were sometimes misclassified, as the model had difficulty distinguishing between positive and negative sentiments in mixed statements.

We also want to express our gratitude to the companies and people who made the iPhone sales dataset and customer feedback data available to us so that we could conduct insightful research and derive useful conclusions.
Lastly, we would like to thank our friends and family for their constant encouragement and support during this research adventure. Without the assistance and efforts of all the people and resources listed above, this endeavor would not have been feasible.



PowerBI Detailed Dashboard

VI. CONCLUSION

This research presents an advanced sentiment analysis sys- tem designed to extract actionable insights from customer feedback related to iPhone sales. By leveraging cutting-edge natural language processing (NLP) techniques, machine learn- ing models, and an interactive chatbot, the system effectively automates the process of sentiment classification, making it easier for businesses to analyze large volumes of unstructured text data from sources such as social media, reviews, and surveys.

The algorithm performed exceptionally well in identifying positive comments and was highly accurate in categorizing consumer attitudes as either positive, negative, or neutral. By empowering stakeholders to make data-driven decisions, the integration of Power BI for sentiment trend visualization sig- nificantly increased the system’s usefulness. A GPT-3.5-based chatbot was also included, enabling dynamic interaction with the sentiment data and yielding perceptive and contextually aware insights.

For companies looking to better understand consumer senti- ment and enhance their goods and services, this project shows the enormous promise of AI-powered sentiment analysis tools. This study’s methodology is flexible and scalable, making it applicable to different datasets and domains. To further improve the system’s capabilities and accuracy.

In summary, the suggested sentiment analysis system is a potent instrument for improving customer satisfaction and business intelligence, giving businesses the capacity to promptly and precisely gauge customer sentiment, pinpoint important problems, and effectively address market demands.

©IJRASET: All Rights are Reserved

VII. ACKNOWLEDGMENT

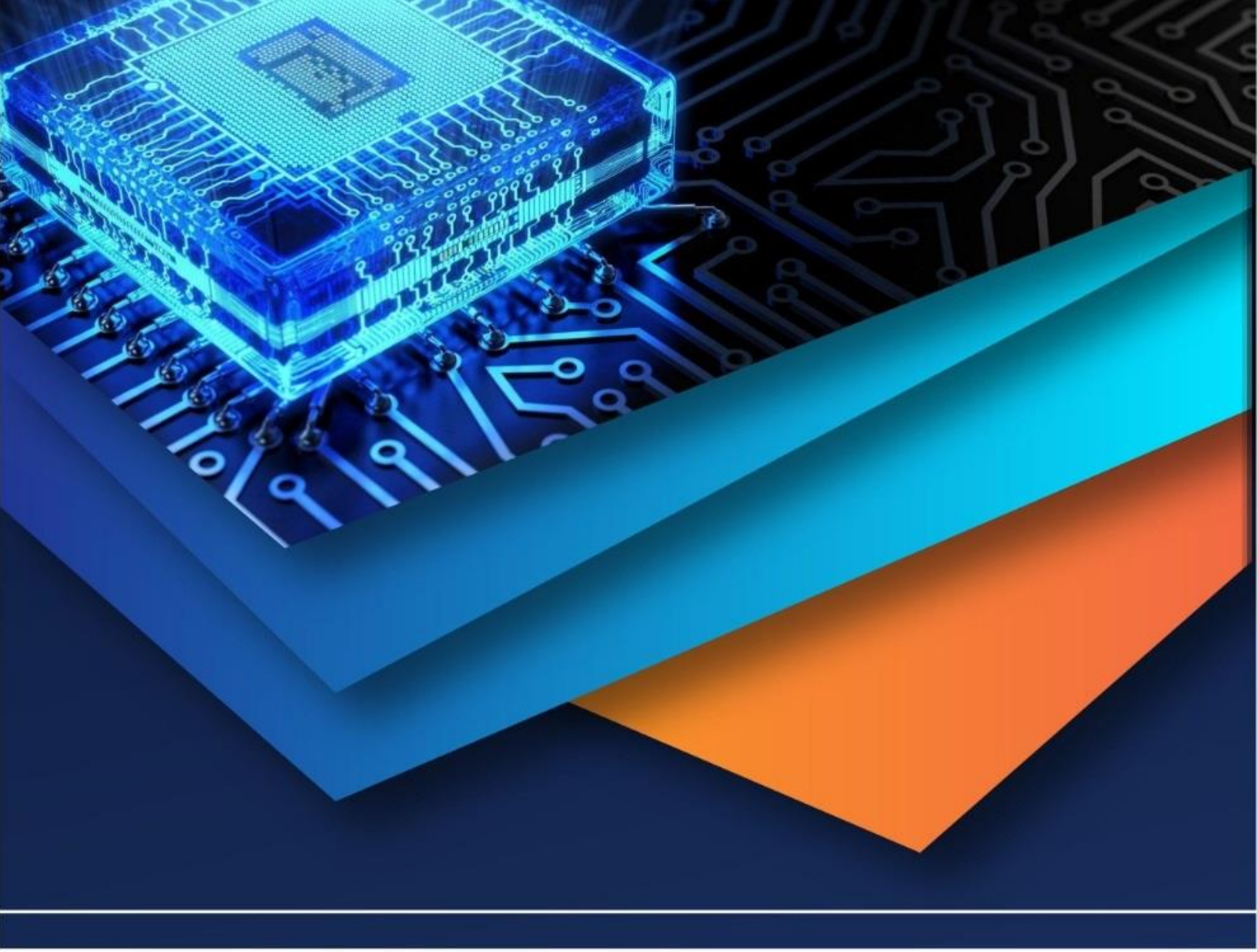
The members of the research team and our guide played a crucial role in this effort through their cooperation and ded- ication. Their contributions to preparation, model evaluation, and data collection were invaluable.

We express our gratitude to the creators and maintainers of the technologies and tools we employed, such as Power BI, OpenAI's GPT-3.5 model. Our research would not have been possible without these resources.

REFERENCES

[1] Zhang, H., Wang, L., & Liu, M. (2018). A Survey on Sentiment Analysis Techniques and Applications. Expert Systems with Applications, 39(10), 8537-8549.
[2] Liu, B. (2015). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge University Press.
[3] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language Models are Few-Shot Learners. In Proceedings of NeurIPS 2020, 33, 1877-1901.
[4] Smith, R., & Johnson, T. (2022). Visualizing Sentiment Trends with Power BI. Proceedings of the International Conference on Data Science, 45-60.
[5] Prabowo, R., & Thelwall, M. (2009). Sentiment Analysis: A Combined Approach. Journal of Informetrics, 3(2), 143-157.
[6] Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.
[7] Microsoft. (2019). Power BI: A Powerful Business Intelligence Tool. Microsoft Power BI Documentation.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)