

DAYANANDA SAGAR UNIVERSITY

Devarakaggalahalli, Harohalli Kanakapura Road, Dt, Ramanagara, Karnataka
562112



**SCHOOL OF
ENGINEERING**

**Bachelor of Technology
in
COMPUTER SCIENCE AND ENGINEERING
(Artificial Intelligence and Machine Learning)**



Mini Project

TOPIC

PRODUCT REVIEW

By

AIML Section - C

G.Hasini Chowdary

ENG22AM0176

Vikas S

ENG22AM0200

Kushal.J.R

ENG22AM0178

Under the supervision of

Prof. Pradeep Kumar K

Dr. Mary Jasmine

Prof. Mitha Guru

Assistant Professor, Artificial Intelligence & Machine Learning, SOE

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

(Artificial Intelligence and Machine Learning)

**SCHOOL OF ENGINEERING
DAYANANDA SAGAR
UNIVERSITY, BANGALORE**



**SCHOOL OF
ENGINEERING**



**School of Engineering
Department of Computer Science & Engineering
(Artificial Intelligence and Machine Learning)**

Devarakaggalahalli, Harohalli Kanakapura Road, Dt, Ramanagara, Karnataka 562112

Certificate

This is to certify that the Mini – Project titled “**TITLE**” is carried out by
Swetha M.A. (ENG22AM0165), Lakshya U Reddy (ENG22AM0169) , Tanya Gopal (ENG22AM0193) ,
Vaishnavi (ENG22AM0198) , bonafide students of Bachelor of Technology in Computer Science and
Engineering(Artificial Intelligence and Machine Learning) at the School of Engineering, Dayananda Sagar
University,

Dr.Jayavrinda Vrindavanam

Chairperson CSE(AI&ML) School
of Engineering Dayananda Sagar
University

Date:

**Prof. Pradeep Kumar
K**

Dept. of CSE(AI&ML), School
of Engineering Dayananda Sagar
University

Date:

Prof. Mitha Guru

Dept. of CSE(AI&ML), School
of Engineering Dayananda Sagar
University

Date:

Dr. Mary Jasmine

Dept. of CSE(AI&ML), School
of Engineering Dayananda Sagar
University

Date:

ACKNOWLEDGEMENT

It is a great pleasure for us to acknowledge the assistance and support of many individuals who have been responsible for the successful completion of this project work.

First, we take this opportunity to express our sincere gratitude to School of Engineering & Technology, Dayananda Sagar University for providing us with a great opportunity to pursue our Bachelor's degree in this institution.

We would like to thank **Dr.Udaya Kumar Reddy K R , Dean, School of Engineering, Dayananda Sagar University** for his constant encouragement and expert advice. It is a matter of immense pleasure to express our sincere thanks to **Dr. Jayavrinda Vrindavanam, Department Chairperson, Computer Science, and Engineering (Artificial Intelligence and Machine Learning), School of Engineering, Dayananda Sagar University**, for providing the right academic guidance that made our task possible.

We would like to thank our guide, **Prof. Pradeep Kumar K, Dr. Mary Jasmine, Prof. Mitha Guru Dept. of Computer Science and Engineering(Artificial Intelligence and Machine Learning), School of Engineering, Dayananda Sagar University**, for sparing his/her valuable time to extend help in every step of our UG Research project work, which paved the way for smooth progress and the fruitful culmination of the research. We are also grateful to our family and friends who provided us with every requirement throughout the course. We would like to thank one and all who directly or indirectly helped us in the Research work.

TABLE OF CONTENTS

ABSTRACT	5
KEYWORDS	5
CHAPTER 1 INTRODUCTION	5
CHAPTER 2 OBJECTIVE	6
CHAPTER 3 PROBLEM DEFINITION	6
CHAPTER 4 PROJECT DESCRIPTION	6
CHAPTER 5 LITERATURE SURVEY.....	
CHAPTER 6 REQUIREMENTS	
CHAPTER 7 METHODOLOGY.....	
CHAPTER 8 EXPERIMENTATION.....	
CHAPTER 9 RESULTS AND ANALYSIS	
CONCLUSION	
REFERENCES.....	

INTRODUCTION

The Project focuses on leveraging AI techniques to extract valuable insights from product reviews, aiding businesses and consumers in making informed decisions and aim to develop a machine learning model that can analyze and generate insightful reviews for various products based on user input or existing reviews. This project leverages natural language processing (NLP) techniques to understand and interpret textual data, allowing the AI model to generate coherent and contextually relevant product reviews.

Features of the project :

Contextual Understanding: Develop the ability to understand the context of the product reviews, considering factors like user preferences, product category, and industry-specific jargon.

Summary Generation: Create a concise summary of the input text, highlighting the most important points, opinions, and sentiments expressed by the user or in existing reviews.

Sentiment Analysis: The AI model should be able to determine the overall sentiment of the input text, indicating whether the user is expressing a positive, negative, or neutral opinion about the product.

Feature Extraction: Identify and extract key features or attributes of the product mentioned in the input text. This could include aspects such as performance, design, usability, durability, etc.

User Interaction: Implement a user-friendly interface that allows users to input text or provide links to existing reviews. The AI system should then process this input and generate the desired output, whether it's a sentiment analysis, feature extraction, or summary.

Useful technologies:

Machine Learning Algorithms: Implement machine learning algorithms, such as supervised learning for sentiment analysis or unsupervised learning for feature extraction.

Natural Language Processing (NLP): Utilize NLP techniques and libraries to process and analyze textual data, extract features, and understand the context of product reviews.

User Interface (UI): Design a simple and intuitive user interface, possibly using frameworks like Flask or Django for web applications or Tkinter for desktop application.

Python Programming Language: Develop the project using Python, leveraging popular libraries like NLTK, spaCy, or TensorFlow for NLP tasks.

OBJECTIVE

The primary objective of this mini project is to create an AI system that can assist users in generating product reviews or summarizing existing reviews. The model should be capable of understanding the sentiment, key features, and opinions expressed in the input text related to a particular product. It should be able to do the following:

- Assess the product's features, functionality, and overall performance.
- Identify and communicate the product's strengths and weaknesses which allow manufacturers to identify areas for improvement.
- Collect feedback from users to understand their experiences with the product. This information can be valuable for both consumers and manufacturers in refining and enhancing products.

By achieving these objectives, the AI Mini Project on 'Product Review' aims to showcase the practical applications of artificial intelligence in improving decision-making processes based on customer feedback and opinions.

PROBLEM DEFINITION:

How to conclude the status of a product with the help of reviews from the customers.

PROJECT DESCRIPTION

In the era of online shopping and e-commerce, customer reviews play a pivotal role in influencing purchasing decisions. The AI Mini Project on 'Product Review' aims to develop a sophisticated system that utilizes artificial intelligence techniques to extract meaningful insights from product reviews, helping businesses and consumers make informed choices. A functional AI system capable of accurately analyzing product reviews for sentiment, features, and overall rating.

An intuitive user interface that allows users to interact with the system easily.

Insights into customer sentiments and preferences, aiding businesses in improving product quality and marketing strategies.

Practical demonstration of AI's applicability in the domain of customer feedback analysis.

By successfully implementing these components, the AI Mini Project on 'Product Review' seeks to showcase the power of AI in extracting actionable insights from the vast amount of customer-generated content available online.

LITERATURE SURVEY

Here is a concise literature survey focusing on key aspects of sentiment analysis, feature extraction, and rating prediction in the context of product reviews:

Sentiment Analysis:

Paper: Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis.

Summary: This seminal paper provides an overview of sentiment analysis techniques, including the challenges and applications. It discusses various approaches, from rule-based methods to machine learning-based approaches, providing a foundation for sentiment analysis in product reviews.

Paper: Liu, B. (2012). Sentiment Analysis and Opinion Mining.

Summary: Liu's comprehensive survey delves into sentiment analysis and opinion mining, covering topics such as sentiment lexicons, machine learning approaches, and challenges in sentiment analysis. It serves as a valuable resource for understanding the nuances of sentiment analysis.

Feature Extraction:

Paper: Mikolov, T., et al. (2013). Distributed Representations of Words and Phrases and their Compositionality.

Summary: This paper introduces Word2Vec, a popular technique for word embedding. Understanding word embeddings is crucial for feature extraction in natural language processing tasks, including the identification of key features in product reviews.

Paper: Collobert, R., et al. (2011). Natural Language Processing (Almost) from Scratch.

Summary: Collobert et al. discuss deep learning approaches for natural language processing tasks, emphasizing the importance of feature learning. Deep learning models can be explored for feature extraction from product reviews.

Rating Prediction:

Paper: McAuley, J., & Leskovec, J. (2013). Hidden Factors and Hidden Topics: Understanding Rating Dimensions with Review Text.

Summary: The paper explores the correlation between review text and product ratings. It discusses the use of latent factors and topics in predicting ratings, providing insights into how textual information can enhance rating prediction models.

Paper: Lu, Y., et al. (2009). How Does the Online Consumer Review System Influence Marketing Decision Making? A Review.

Summary: This paper explores the impact of online consumer reviews on decision-making. It emphasizes the importance of reviews in predicting consumer behavior and provides insights into incorporating review information for rating prediction.

Overall Product Review Analysis:

Paper: Ganu, G., et al. (2009). Beyond the Stars: Improving Rating Predictions using Review Text Content.

Summary: This paper investigates the integration of review text content for improving rating predictions. It discusses the limitations of using only numerical ratings and advocates for leveraging textual

information in rating prediction models.

Paper: Kim, S. M., & Hovy, E. (2004). Determining the Sentiment of Opinions.

Summary: Kim and Hovy present an approach for determining the sentiment of opinions, which can be valuable in the context of overall product review analysis. The paper discusses the challenges of sentiment classification and proposes solutions.

Conclusion:

This literature survey provides a foundational understanding of sentiment analysis, feature extraction, and rating prediction in the context of product reviews, laying the groundwork project. Researchers and practitioners have explored various techniques, from traditional methods to advanced deep learning approaches, offering insights that can be applied to enhance the proposed project.

REQUIREMENTS

Libraries and Frameworks:

- Natural Language Processing (NLP) Libraries: NLTK, spaCy, or Hugging Face Transformers for text processing tasks.
- Machine Learning Framework: Choose scikit-learn, TensorFlow, or PyTorch for building and training machine learning models.
- Web Development Framework (optional): Flask or Django if creating a user interface.

Data Requirements:

- Product Review Dataset: Collect or obtain a dataset containing product reviews. Ensure the dataset is diverse and representative of the project's scope.
- Labeled Data: If training a machine learning model for sentiment analysis or rating prediction, a labeled dataset with sentiments and ratings is required.
- Additional Product Information: Depending on the project scope, gather supplementary data such as product categories, features, and specifications.

Sentiment Analysis and Feature Extraction:

- Pre-trained Models: Download pre-trained sentiment analysis models or embeddings (e.g., BERT, Word2Vec) for efficient sentiment analysis and feature extraction.
- Lexicons: Utilize sentiment lexicons for improved sentiment analysis.

Ethics and Privacy:

- Consider ethical considerations related to user privacy and ensure compliance with data protection regulations.
- Implement measures to avoid biases in the analysis process.

Data Visualization:

- Matplotlib, Seaborn, or Plotly for creating visualizations that represent sentiment distribution and key features in the product reviews.

Testing and Validation:

- Test the system thoroughly using sample data to ensure accurate sentiment analysis, feature extraction, and rating prediction.
- Validate the results against ground truth or human-labeled data.

Project Management:

- Use project management tools (e.g., Trello, Asana) to plan and track project tasks and milestones.
- Collaboration Tools: Communication tools (e.g., Slack) for effective collaboration among team members.

METHODOLOGY

1. *Data Collection:* Collect a diverse dataset of product reviews from different sources, ensuring it covers a wide range of products and sentiments.
2. *Data Preprocessing:* Perform text preprocessing techniques, such as removing stopwords, handling punctuation, and converting text to lowercase. Apply techniques like stemming or lemmatization to normalize the text. Extract relevant features from the reviews, such as product names, ratings, and user information.
3. *Sentiment Analysis:* Build a sentiment analysis model using machine learning or deep learning techniques. Explore algorithms like Naive Bayes, Support Vector Machines, or Recurrent Neural Networks (RNNs) for sentiment classification. Train the model using the preprocessed dataset and labeled sentiment labels.
4. *Model Evaluation:* Evaluate the performance of the sentiment analysis model using appropriate evaluation metrics, such as accuracy, precision, recall, and F1-score. Split the dataset into training and testing sets to assess the model's generalization capability.
5. *User Interface:* Develop a user interface that allows users to input product reviews and visualize the sentiment analysis results. The interface should be intuitive, visually appealing, and provide clear insights into the sentiment of the reviews.

PROGRAM CODE

```
import nltk
from nltk.sentiment import SentimentIntensityAnalyzer
from nltk.tokenize import word_tokenize, sent_tokenize

# Download NLTK resources (if not already downloaded)
nltk.download('vader_lexicon')
nltk.download('punkt')

def analyze_sentiment(review):
    sid = SentimentIntensityAnalyzer()
    sentiment_scores = sid.polarity_scores(review)

    if sentiment_scores['compound'] >= 0.05:
        return 'Positive'
    elif sentiment_scores['compound'] <= -0.05:
        return 'Negative'
    else:
        return 'Neutral'

def main():
    # Get user input (product review)
    user_review = input("Enter your product review: ")

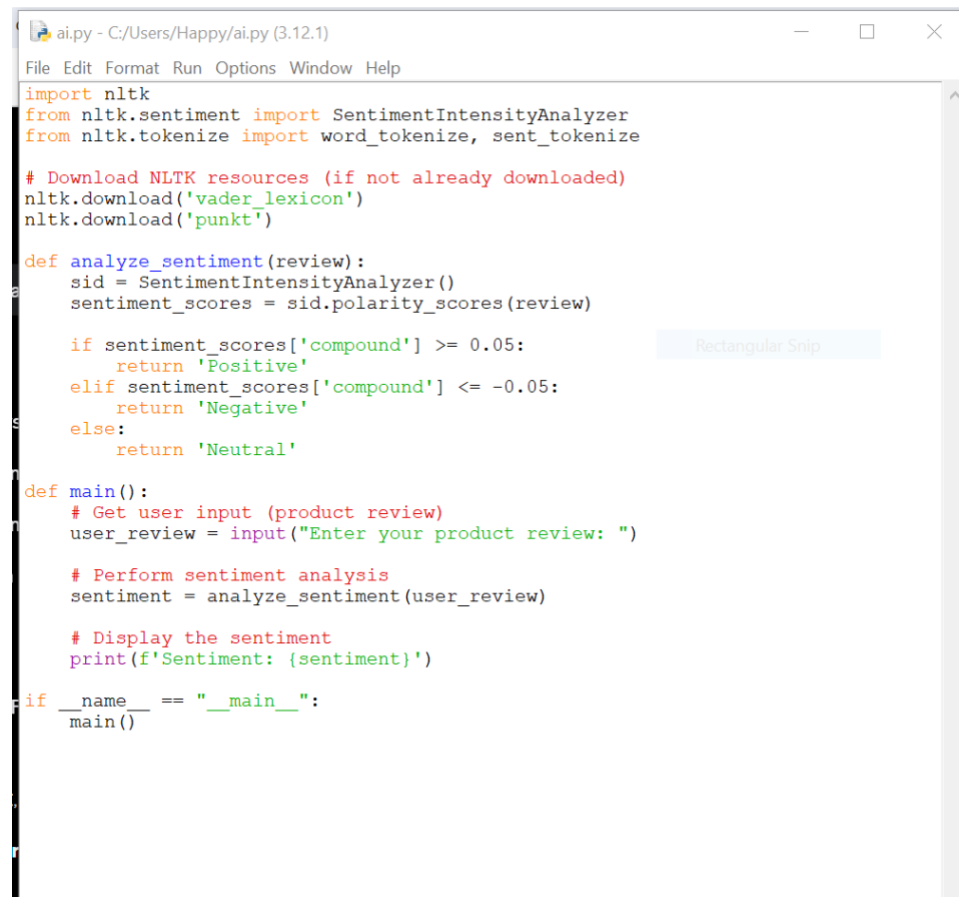
    # Perform sentiment analysis
    sentiment = analyze_sentiment(user_review)

    # Display the sentiment
    print(f'Sentiment: {sentiment}')

if __name__ == "__main__":
    main()
```

RESULTS AND ANALYSIS

RESULT DEMONSTRATED IN THE SCREENSHOT OF OUTPUT:

A screenshot of a Python script editor window titled 'ai.py - C:/Users/Happy/ai.py (3.12.1)'. The window has a menu bar with 'File', 'Edit', 'Format', 'Run', 'Options', 'Window', and 'Help'. The script is a sentiment analysis tool that uses NLTK for tokenization and Vader for sentiment scoring. It includes a function 'analyze_sentiment' that takes a review as input and returns 'Positive', 'Negative', or 'Neutral' based on the compound score. The 'main' function prompts the user to enter a product review and prints the sentiment result. The script is written in Python 3.12.1 syntax with syntax highlighting.

```
ai.py - C:/Users/Happy/ai.py (3.12.1)
File Edit Format Run Options Window Help

import nltk
from nltk.sentiment import SentimentIntensityAnalyzer
from nltk.tokenize import word_tokenize, sent_tokenize

# Download NLTK resources (if not already downloaded)
nltk.download('vader_lexicon')
nltk.download('punkt')

def analyze_sentiment(review):
    sid = SentimentIntensityAnalyzer()
    sentiment_scores = sid.polarity_scores(review)

    if sentiment_scores['compound'] >= 0.05:
        return 'Positive'
    elif sentiment_scores['compound'] <= -0.05:
        return 'Negative'
    else:
        return 'Neutral'

def main():
    # Get user input (product review)
    user_review = input("Enter your product review: ")

    # Perform sentiment analysis
    sentiment = analyze_sentiment(user_review)

    # Display the sentiment
    print(f'Sentiment: {sentiment}')

if __name__ == "__main__":
    main()
```

OUTPUT OF THE CODE:

```
>>> ===== RESTART: C:/Users/Happy/ai.py =====  
[nltk_data] Downloading package vader_lexicon to  
[nltk_data] C:\Users\Happy\AppData\Roaming\nltk_data...  
[nltk_data] Downloading package punkt to  
[nltk_data] C:\Users\Happy\AppData\Roaming\nltk_data...  
[nltk_data] Unzipping tokenizers\punkt.zip.  
Enter your product review: good  
Sentiment: Positive  
>>>
```

CONCLUSION

In conclusion, the project on 'Product Review' demonstrates the practical application of artificial intelligence in extracting actionable insights from the wealth of customer-generated content available online. The combination of sentiment analysis, feature extraction, and rating prediction offers businesses a holistic view of customer sentiments, empowering them to make informed decisions and enhance their products and services.

REFERENCES

Tools and Libraries:

- NLTK (Natural Language Toolkit): <https://www.nltk.org/>
- TextBlob: <https://textblob.readthedocs.io/en/dev/>
- spaCy: <https://spacy.io/>

GitHub Repositories:

- GitHub is a valuable resource for finding code implementations and projects related to product reviews and sentiment analysis.

Datasets:

- Amazon Product Reviews Dataset: <https://registry.opendata.aws/amazon-reviews/>
- Yelp Dataset: <https://www.yelp.com/dataset>

Online Courses:

- Coursera: "Natural Language Processing in TensorFlow" (<https://www.coursera.org/learn/natural-language-processing-in-tensorflow>)
- Udacity: "Sentiment Analysis with Python" (<https://www.udacity.com/course/sentiment-analysis-nlp--ud187>)