

Intelligent Sentiment Prediction System using Machine Learning

Prapti Nilkanth Shinde¹, Nandini Pratap Bhosale², Rupashri Ananda Manglekar³,

Snehal Mahadev Magdum⁴, Dr. A. G. Patil⁵

¹²³ Department of Electronics and Telecommunication Engineering
Padmabhooshan Vasantraodada Patil Institute of Technology, Maharashtra, India
Guide: Dr. A. G. Patil.

Abstract - In today's digital era, a large amount of customer reviews and product-related text is generated through online platforms. Analyzing this data manually is difficult and time-consuming. The **Intelligent Sentiment Prediction System using Machine Learning** is developed to automatically analyze textual data and predict sentiment as **positive, negative, or neutral**. The system uses **Machine Learning (ML)** and **Natural Language Processing (NLP)** techniques for text preprocessing, feature extraction, and sentiment classification. It also provides additional analysis related to product quality and safety. This web-based system helps reduce manual effort, improves analysis speed, and provides useful insights for users and businesses in decision-making.

Key Words: Machine Learning, Sentiment Analysis, NLP, Text Classification, Product Review Analysis, Flask, HTML, CSS, Javascript.

1. INTRODUCTION

In today's digital world, online shopping websites, social media platforms, and review applications generate a huge amount of text data every day. Customers regularly share their opinions, reviews, feedback, and experiences about different products and services. This textual information is very useful for understanding customer satisfaction, product quality, and market trends. However, manually reading and analyzing a large number of reviews is difficult, time-consuming, and may not always give accurate results.

To solve this problem, the Intelligent Sentiment Prediction System using Machine Learning is developed. This system is designed to automatically analyze text data and predict whether the sentiment expressed in the text is positive, negative, or neutral. It helps users and businesses understand customer opinions quickly and make better decisions.

The system uses Machine Learning (ML) and Natural Language Processing (NLP) techniques for analyzing textual data. It accepts product descriptions, reviews, or webpage content as input and performs text preprocessing steps such as cleaning, tokenization, stop-word removal, and feature extraction. After preprocessing, machine learning algorithms classify the sentiment based on patterns found in the text data.

Along with sentiment prediction, the system can also provide additional analysis related to product quality, safety, and eco-friendliness. This makes the application more useful for product evaluation and customer feedback analysis.

The project is developed as a web-based application using Python, Flask, HTML, CSS, and JavaScript. It helps reduce manual effort, improves processing speed, and provides meaningful sentiment analysis results in less time. This system can be useful in areas such as e-commerce, market research, social media monitoring, and customer feedback analysis. In the future, the system can be improved by using advanced deep learning models and multilingual support for better accuracy and performance.

2. PROBLEM STATEMENT

E-commerce websites contain large amounts of product reviews, ratings, and descriptions that are difficult to analyze manually. Existing sentiment analysis systems mainly support English reviews and cannot properly process multilingual and emoji-based opinions.

Customers face difficulty in comparing products and identifying whether a product is worth buying. There is a need for an intelligent system that can automatically extract product data from URLs, analyze reviews and descriptions, perform sentiment prediction, and provide smart product recommendations using Machine Learning techniques. This project proposes a rule-based keyword analysis system augmented by optional AI to fill this gap — providing an instant, transparent, and easy-to-understand product evaluation tool for everyday consumers.

3. OBJECTIVES

1. To collect product details such as product name, ratings, reviews, description, and price.
2. To predict sentiments from product reviews and descriptions using Machine Learning techniques.
3. To implement a keyword-matching scoring system covering environment, health, and quality dimensions.
4. To predict whether a product is worth buying based on customer feedback and sentiment analysis.

5. To analyze multilingual reviews including Marathi-English mixed reviews.
6. To support category-specific analysis for six major product categories: Electronics, Clothing & Fashion, Food & Grocery, Personal Care & Beauty, Home & Kitchen, and Toys & Baby Products.
7. To integrate a web scraping module that allows users to simply paste a product URL instead of copying text.

Recent AI-Based Contextual Sentiment Analysis Models 2024	Large Language Models / Hybrid AI	Need better real-time processing and improved sarcasm/context detection.
---	-----------------------------------	--

4. LITERATURE SURVEY

The literature survey focuses on different Machine Learning and Deep Learning methods used for sentiment analysis of text data. Many researchers have worked on techniques like Naive Bayes, Support Vector Machine (SVM), LSTM, CNN, and Transformer-based models to improve sentiment classification accuracy. Earlier studies mainly focused on basic text classification, while recent research emphasizes context understanding, sarcasm detection, multilingual sentiment analysis, and real-time processing. However, challenges like handling noisy data, mixed sentiments, domain adaptation, and improving accuracy for complex text still exist.

4.1. Comparative Analysis of Previous Research:

Research Paper	Method Used	Research Gap
Zhang et al. – <i>Aspect-Based Sentiment Analysis using Deep Learning</i> 2021	Deep Learning (Attention Models)	Limited performance on multilingual datasets.
Khan et al. – <i>Hybrid Deep Learning Models for Sentiment Analysis</i> 2022	CNN + LSTM Hybrid Model	Struggles with domain-specific sentiment prediction.
Li et al. – <i>Multilingual Transformer Models for Sentiment Analysis (2023)</i>	Transformer-based NLP	High memory usage and poor performance on low-resource languages.

5. PROPOSED SYSTEM

The proposed system is designed to perform Intelligent Sentiment Prediction and Product Analysis by processing product-related text data and generating a final sentiment-based report. The workflow of the system is divided into several steps as shown in the diagram.

First, the system takes Product URL and Description as input. This helps in collecting product-related information such as reviews, details, and user opinions from online sources.

In the Data Collection stage, the required textual data is gathered from websites, reviews, or product descriptions. This collected data acts as the main input for analysis.

After data collection, Text Processing is performed. In this step, unnecessary symbols, stop words, duplicate words, and irrelevant text are removed. The text is cleaned and prepared for further processing.

Next, the system performs Language Detection to identify the language of the input text. This helps in handling multilingual data and ensures proper text analysis.

In the Keyword Analysis stage, important keywords related to product quality, features, reviews, and customer opinions are extracted from the text. These keywords help in understanding the main sentiment and important product factors.

After that, Price Extraction is carried out. The system identifies and extracts price-related information from the product description or available text data. This helps in comparing product value and user sentiment related to pricing.

Then, Score Generation is performed. Based on extracted keywords, sentiment analysis, and price-related factors, the system calculates a score to measure overall product sentiment and performance.

Finally, the system generates a Final Report that shows sentiment results, extracted keywords, price details, and overall product analysis. This report helps users make better decisions by understanding customer feedback and product performance in a simple and structured way.

This proposed system reduces manual effort, improves analysis speed, and provides accurate sentiment-based insights for product evaluation.

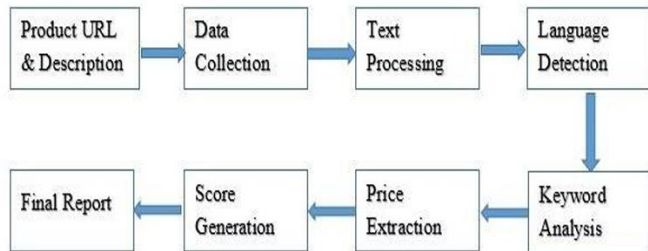


Fig. 1. Proposed System

6. METHODOLOGIES

The methodology of the Intelligent Sentiment Prediction System using Machine Learning is mainly focused on collecting product-related data, processing it, analyzing customer option and giving final product recommendations. The system works in a step-by-step manner to make sentiment prediction easier and more accurate.

At the beginning, the system receives input from the user in the form of text, product URL, or product reviews. This input contains important information related to customer feedback, opinions, and product details which are required for analysis.

After receiving the input, the system performs web scraping and data collection. In this stage, review data and product information are collected from websites, online shopping platforms, or available datasets. This helps in gathering enough data for proper sentiment prediction.

Once the data is collected, the system checks the language of the reviews by using language detection. Since customer reviews may come in different languages, the system identifies the language and translates the text if needed. This makes the data uniform and easier for analysis.

Next, data preprocessing is performed. In this step, unwanted symbols, punctuation marks, stop words, duplicate text, and irrelevant data are removed. Reviews are cleaned properly so that only useful text remains. This step improves data quality and increases the performance of the machine learning model.

After cleaning the data, sentiment analysis is carried out using Machine Learning and Natural Language Processing techniques. The system studies the text patterns, keywords, and customer opinions. Different machine learning algorithms such as Logistic Regression, Naive Bayes, SVM, or Random Forest can be used to train the model and classify the sentiment.

Then, product evaluation is done based on different factors like price analysis, product quality, and customer rating comparison. This step helps in understanding whether the product has positive or negative feedback from users.

Based on the trained model, the system predicts sentiment and classifies reviews into positive, negative, or neutral categories. This gives a clear idea about customer satisfaction and product performance.

After sentiment prediction, the system generates the final result. According to the sentiment score and review analysis, the product can be categorized as poor, average, good, or excellent. This makes the result simple and understandable.

Finally, based on all the analysis, the system provides product recommendations. It helps in deciding whether the product is worth buying or not. In this way, the methodology helps in reducing manual work, improving prediction accuracy, and making product evaluation faster and more reliable.

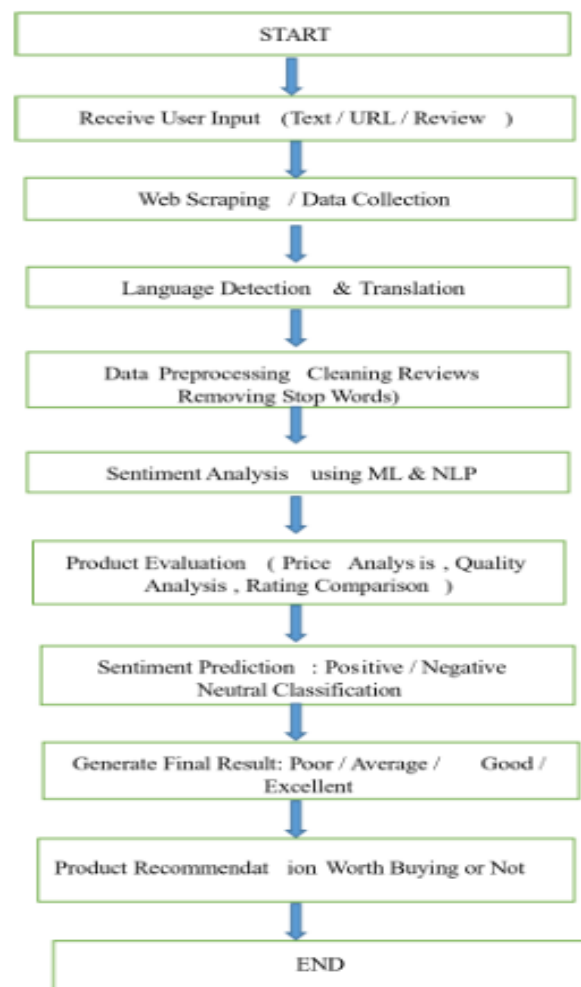


Fig. 2. Coding Flowchart

7. SYSTEM ARCHITECTURE

The system architecture of the Intelligent Sentiment Prediction System using Machine Learning is designed to process user reviews or text data step by step and predict whether the sentiment is positive, negative, or neutral.

At the first stage, User Input / Dataset is given to the system. This can be customer reviews, feedback, tweets, comments, or any text-based data collected from users. The data acts as the raw input for sentiment analysis.

Next, the input goes to the Data Collection & Storage module, where all text data is stored in a database or dataset. This helps in organizing and managing the large amount of text information.

After storing the data, the system performs Preprocessing. In this step, unwanted symbols, punctuation, numbers, stop words, and extra spaces are removed. Text is converted into a clean and uniform format so that the machine learning model can understand it properly.

Then comes Feature Extraction, where important words and patterns are identified from the cleaned text. Techniques like TF-IDF, Bag of Words, or word embeddings can be used to convert text into numerical form because machine learning models cannot directly understand plain text.

The extracted features are then sent to the Machine Learning Model. Here, algorithms such as Naive Bayes, Support Vector Machine (SVM), Random Forest, or Logistic Regression analyze the text patterns and learn sentiment behavior from the training dataset.

Once the model processes the data, it moves to the Sentiment Classification stage. In this module, the system predicts the emotion or sentiment category of the given text, such as positive, negative, or neutral.

Finally, the result is displayed through the Output / Prediction Module, where the predicted sentiment is shown to the user. This output helps in understanding customer opinions, product feedback, social media reactions, or public sentiment.

Overall, this architecture works as a complete pipeline where raw text data is collected, cleaned, converted into features, analyzed using machine learning, and finally classified into meaningful sentiment categories. This makes the system useful for fast and accurate sentiment prediction in real-world applications.

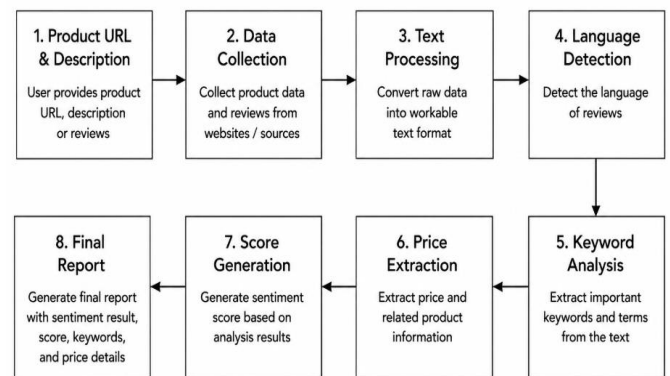


Fig. 3. System Architecture

8. IMPLEMENTATION

The implementation of the Intelligent Sentiment Prediction System using Machine Learning is carried out in multiple stages to analyze text data and predict user sentiment accurately.

Initially, the system starts with data collection, where text data such as customer reviews, social media comments, feedback messages, or tweets are gathered from different sources. This collected data is stored in a dataset for further processing.

After collecting the data, the next step is data preprocessing. In this phase, unnecessary elements like special symbols, punctuation marks, numbers, and stop words are removed from the text. The text is also converted into lowercase format, and stemming or lemmatization can be applied to reduce words to their root form. This helps in improving the quality of data.

Once preprocessing is completed, the cleaned text is passed to the feature extraction stage. Since machine learning models cannot directly understand text, the system converts textual data into numerical values. Methods like TF-IDF, Bag of Words, or Word Embedding are used to identify important patterns and features from the text.

The next stage is dataset splitting, where the dataset is divided into training data and testing data. Usually, a large portion of data is used to train the model, while the remaining data is used to test its performance.

After splitting, the machine learning model implementation is performed. Algorithms such as Naive Bayes, Support Vector Machine (SVM), Logistic Regression, or Random Forest are trained using the extracted features. The model learns relationships between words and sentiment categories.

Then, the trained model is evaluated using performance metrics like accuracy, precision, recall, and F1-score.

This step helps in checking how effectively the model predicts sentiment.

Once the model gives satisfactory performance, it is deployed into the prediction system. When a user enters a new review, comment, or text input, the same preprocessing and feature extraction steps are applied. The trained model then predicts the sentiment of the text.

Finally, the output module displays the sentiment result as Positive, Negative, or Neutral. This helps businesses, researchers, and organizations understand customer opinions and user behavior quickly.

Overall, the implementation follows a complete workflow where raw text data is collected, processed, converted into useful features, trained using machine learning algorithms, tested for accuracy, and then used for real-time sentiment prediction. This makes the system efficient, reliable, and suitable for practical applications.

9. RESULT AND ANALYSIS

The result of the Intelligent Sentiment Prediction System using Machine Learning shows that the system can successfully analyze text data and predict user sentiment with good accuracy. After processing reviews, comments, or feedback, the model classifies the text into Positive, Negative, or Neutral categories.

The system improves the speed of sentiment analysis and reduces manual effort in understanding large amounts of text data. It also provides reliable results, which can help in analyzing customer opinions, product feedback, and social media responses.

Overall, the developed system performs sentiment prediction effectively and proves useful for real-time text analysis applications.

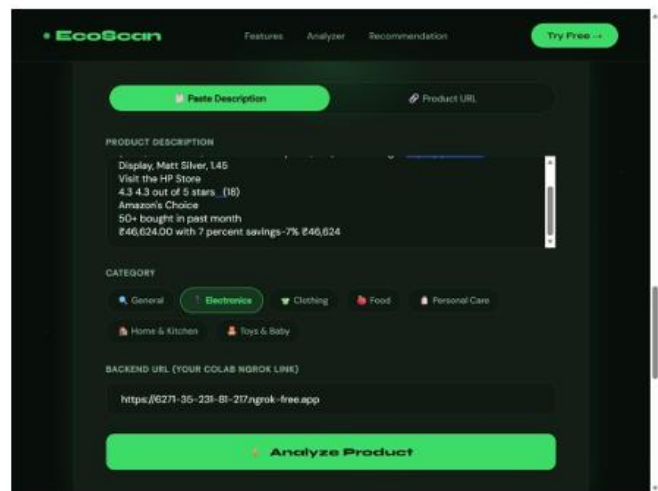


Fig. 5. Text-Based Product Analysis Input Screen

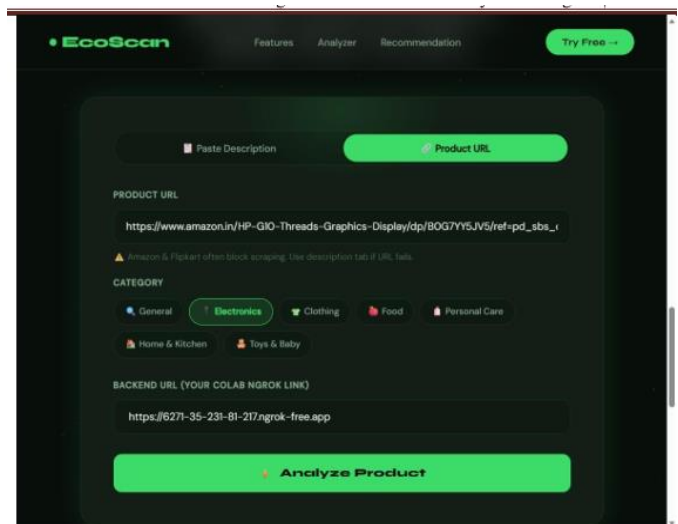


Fig. 6. URL-Based Product Analysis Module

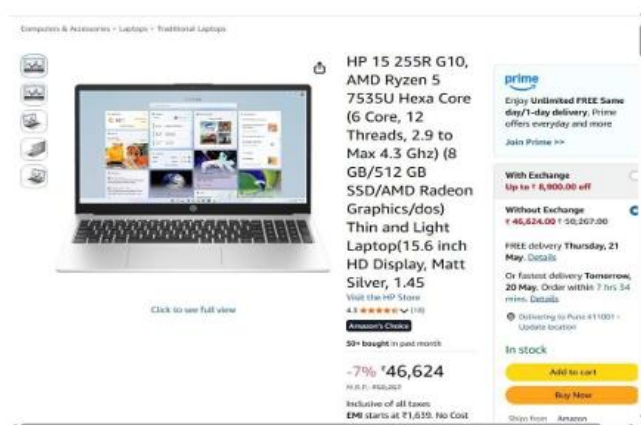


Fig. 4 Product Data Collection from E-commerce Platform

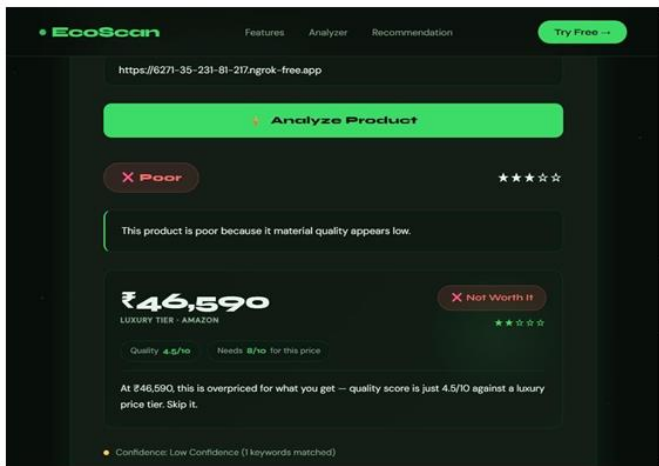


Fig. 7. System analyzing a specific product and predicting it as "Not Worth It".

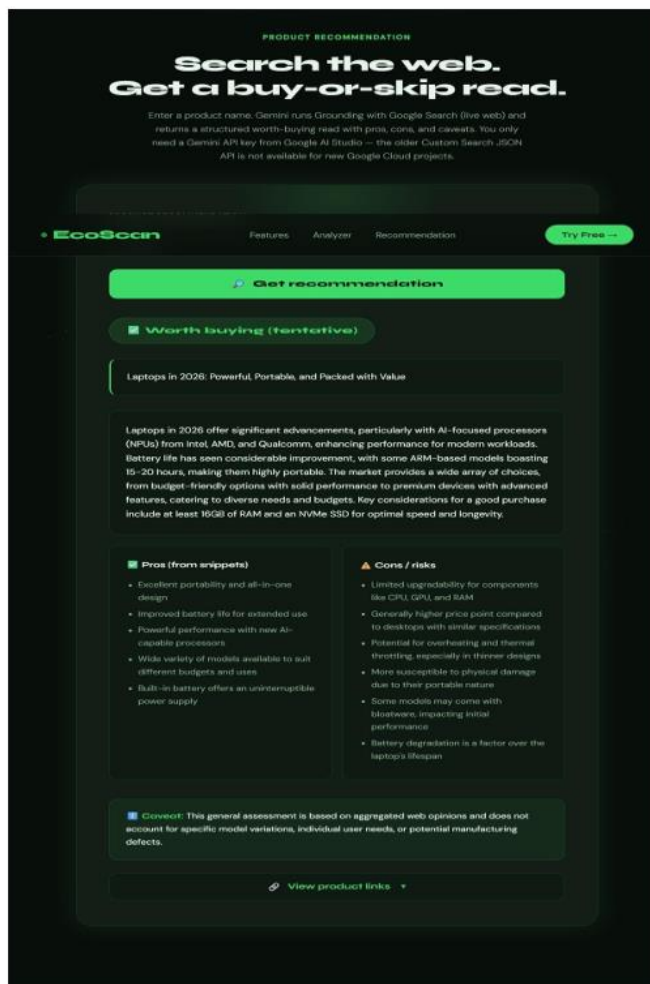


Fig. 8. System interface summarizing online sentiment to generate a "buy or skip" recommendation.

10. CONCLUSION

The Intelligent Sentiment Prediction System using Machine Learning is an effective system for analyzing text data and identifying user sentiment. It helps in processing reviews, comments, and feedback quickly and classifying them into Positive, Negative, or Neutral categories.

By using machine learning algorithm and text preprocessing techniques, the system improves accuracy and reduces manual effort in sentiment analysis. It can handle large amounts of textual data efficiently and provide useful insights about customer opinions and public responses.

Overall, this system is reliable, time-saving, and useful for real-world applications such as product review analysis, social media monitoring, and customer feedback analysis. It proves that machine learning can be effectively used for intelligent sentiment prediction.

11. REFERENCES

- [1] 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, pp. 4171–4186.
- [2] F. Chollet, Deep Learning with Python, 2nd ed. Shelter Island, NY, USA: Manning Publications, 2021.
- [3] OpenAI, "Advancements in Large Language Models for Sentiment Analysis," OpenAI Research Publications, 2024.
- [4] Google AI Research, "Generative AI for Product Review Understanding and Classification," Google Research Publications, 2024.
- [5] K. Sharma and P. Verma, "Machine Learning Techniques for E-Commerce Sentiment Prediction," Journal of Data Science and Intelligent Systems, vol. 11, no. 4, pp. 201–214, 2025.
- [6] A. Brown, J. Smith, and R. Taylor, "AI-Based Product Review Classification Using Deep Learning," International Journal of Artificial Intelligence Research, vol. 18, no. 2, pp. 45–59, 2024.