ISSN 2277-2685
IJESR/June. 2025/ Vol-15/Issue-3s/257-263
Kanchanapally Renu Sree *et. al.*, / International Journal of Engineering & Science Research

Shoplytics

Dr P Deepthi, Kanchanapally Renu Sree, Akkala Sai Sruthi
[1]Associate Professor, Department Of Cse, Bhoj Reddy Engineering College For Women, India.
[2,3]B. Tech Students, Department Of Cse, Bhoj Reddy Engineering College For Women, India.

ABSRTACT

Shoplytics is a smart feedback analysis tool designed to automate the process of gathering and analyzing product reviews from e-commerce websites. By leveraging Python scripts and sentiment analysis models like BERT, it transforms large volumes of unstructured review data into meaningful insights.

The system scrapes user reviews, product prices, and ratings directly from product pages using Selenium, then stores them securely in MongoDB. These reviews are processed to identify sentiment—positive, neutral, or negative—and the results are presented through intuitive visualizations.

Shoplytics eliminates the need for manual data handling, improves analysis accuracy, and enhances the online shopping experience for customers and vendors alike.

Its modular design allows for flexibility, and its centralized data storage ensures organized management of large-scale feedback. By combining web scraping, sentiment classification, and report generation, Shoplytics enables e-commerce platforms to make smarter decisions, boost customer satisfaction, and tailor product recommendations based on actual consumer sentiment.

1-INTRODUCTION

Shoplytics is a comprehensive feedback intelligence tool that automatically collects and analyzes customer reviews from e-commerce platforms.

It uses machine learning models and data processing scripts to identify sentiment trends and extract valuable information, such as product ratings and pricing patterns.

This system provides a deep understanding of customer preferences, allowing businesses to enhance product offerings and shopping experiences.

Existing System

- Current systems often depend on manual methods like Excel sheets or basic scripts to analyze reviews.
- These processes are slow, error-prone, and lack scalability. There is no centralized storage, and the data is often fragmented.
- As a result, companies are unable to get timely or accurate insights.

Proposed System

The proposed system scrapes product reviews, ratings, and prices directly from websites. The data is stored in MongoDB and analyzed using pre-trained BERT models. It classifies customer reviews into sentiment categories and presents them using interactive dashboards created with Streamlit. This real-time analysis helps businesses understand trends, compare products, and improve customer engagement.

2-REQUIREMENT ANALYSIS

Functional Requirements

User Module

1. **Register**
   - Users can create an account by providing a unique username and password.
   - Upon successful registration, the account is created, and users are redirected to the login page.
2. **Login**
   - Registered users can log in using their credentials.

- The application manages sessions using Streamlit session to retain user state until logout.

3. **Input Product Details**

- Users provide the name of the product and the number of products they wish to scrape.
- The application takes these inputs to fetch product data from e-commerce websites.

4. **Scrape Reviews**

- The system uses Selenium to scrape user reviews, ratings, and prices from the target website.
- All collected data is preprocessed and stored in a MongoDB collection.

Non-Functional Requirements

- Security: Secure handling of credentials and review data using encryption techniques like MD5 hashing.
- Scalability: Designed to efficiently handle growing amounts of review data across multiple products.
- Usability: Simple and intuitive UI built using Streamlit for both technical and non-technical users.
- Maintainability: Modular code allows for quick updates and easy integration of new features.
- Performance: Optimized for fast scraping, quick sentiment processing, and responsive interface updates.

Hardware Resources

The hardware requirements are designed to ensure smooth development and testing of the platform, especially for tasks like real-time collaboration and database operations.

Processor:

- Intel i3 (min) or equivalent AMD processor.
- 8 GB (min)
- Supports parallel operations including browser automation and ML inference.

Storage:

- 256 GB (min) SSD
- Ensures faster read/write speeds for smooth operation of the development environment and databases. And access for storing reviews and user data.

Software Resources

The system utilizes modern and scalable tools to support end-to-end review analysis which are:

Operating System:

- Windows 11
- Ensures broad compatibility for browser drivers and Python environments

Programming Language:

- Python 3.11
- Used for scripting, scraping, sentiment analysis, and backend logic

Web Framework:

- Streamlit
- Used for building an interactive and real-time user interface

IDE:

- Visual Studio Code
- Provides efficient development tools, extensions, and version control

Database:

- MongoDB
- A NoSQL document-based database used for storing reviews and metadata

3-DESIGN

Design architecture refers to the overall structure and framework of a system, outlining how its components interact, communicate, and work together to meet functional and non-functional requirements. It serves as a blueprint for the system's development, ensuring it is scalable, maintainable, and efficient.

Software Architecture

The software architecture illustrates three primary components: User, System, and Database. Users interact with the system through a web interface to register, log in, input product names, scrape reviews, analyze feedback, and download reports. The system handles web scraping, sentiment classification using

BERT, and visualization of results. It communicates with the database, which stores review data, user credentials, and analysis outcomes. This architecture ensures seamless user interaction, secure data handling, and efficient processing for large-scale data analysis.

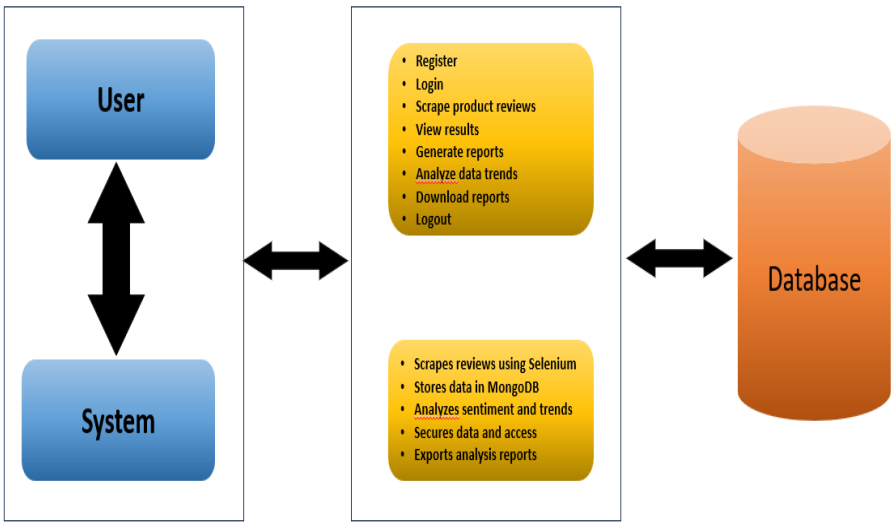


Fig. 3.1 Software Architecture

Technical Architecture

The technical architecture illustrates a system built using Streamlit, Python, and MongoDB. Streamlit serves as the frontend, enabling user interaction through a web-based interface. Python acts as the core logic layer, processing inputs, executing functionality, and managing data flow. MongoDB functions as the backend database, storing and retrieving data efficiently. The architecture ensures seamless integration between the user interface, processing logic, and data management.

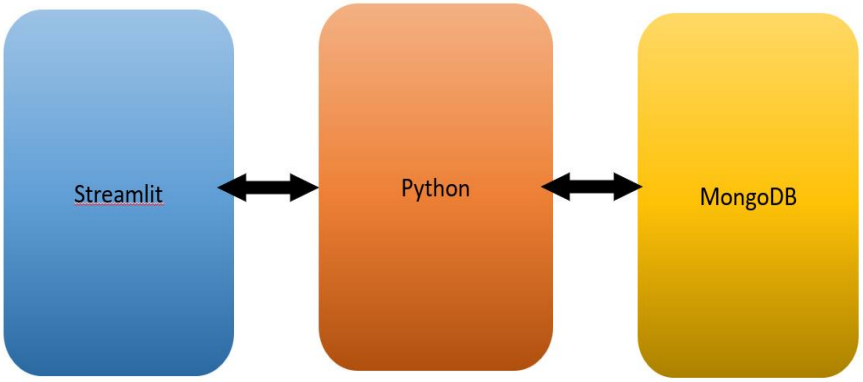


Fig. 3.2 Technical Architecture

4-IMPLEMENTATION

Frontend

- Streamlit:
  - A Python framework for creating interactive and data-driven web applications.
  - Used to create the UI for login, registration, product input, and results dashboard.
  - Features: Clean interface, real-time updates, easy integration with Python.

Backend:

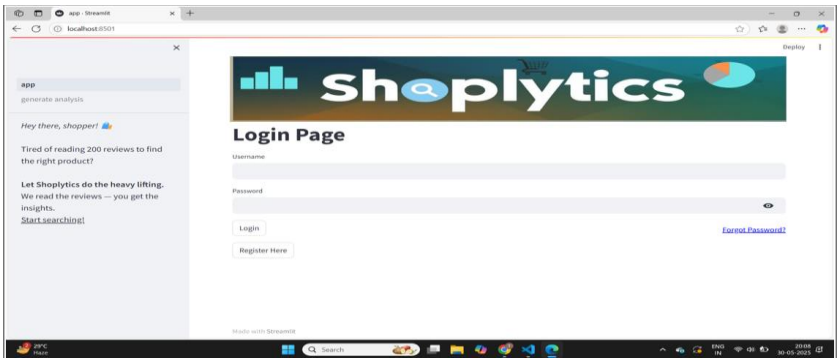
- Python:
  - Main programming language for backend operations like web scraping, data analysis.
  - Libraries used:
    - Selenium: For automated web scraping of product reviews and details from e-commerce sites.
    - BeautifulSoup (bs4): For parsing HTML content and extracting relevant review data.
    - Pandas: For data manipulation and analysis.
    - Torch (PyTorch): To run the pre-trained BERT model for sentiment analysis.
    - Transformers (Hugging Face): To utilize the BERT model for NLP-based classification.

Database:

- MongoDB:
  - NoSQL database used for storing user details, product data, and extracted reviews.
  - Suitable for handling semi-structured data like JSON objects from web scraping.
  - Features: Centralized and secure data storage, easily scalable for large datasets and collections created dynamically based on product names.

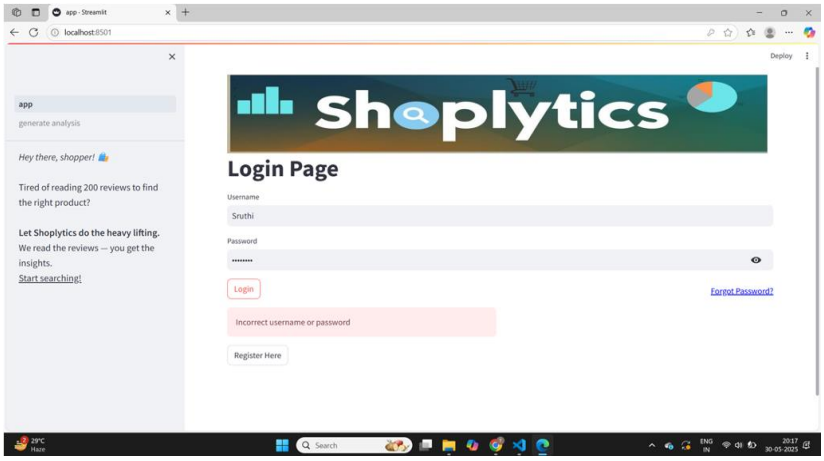
5-SCREENSHOTS

Login Page



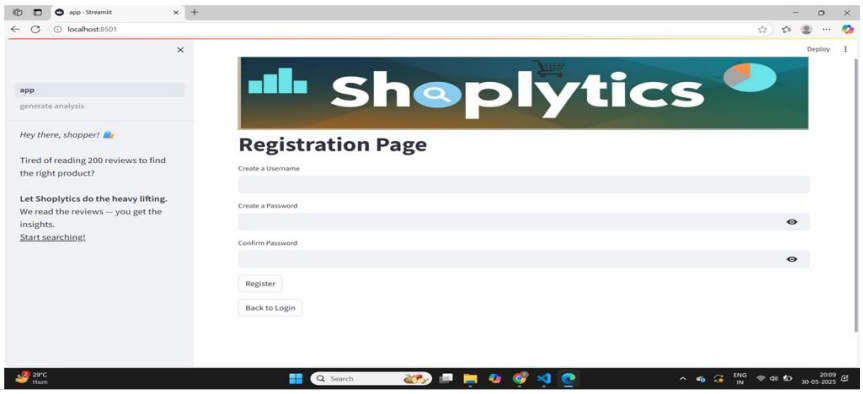
Screenshot 1 Login Page

Login Page with Validation Error



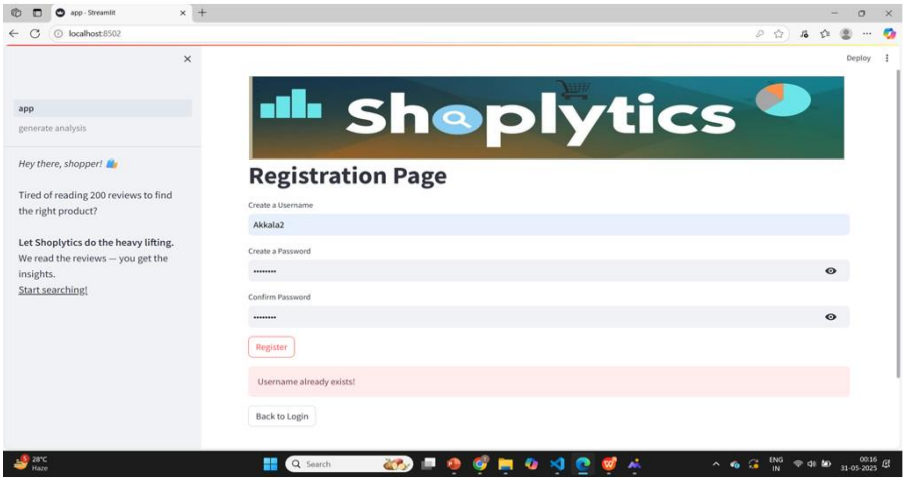
Screenshot 2 Login Page with Validation Error

Registration page



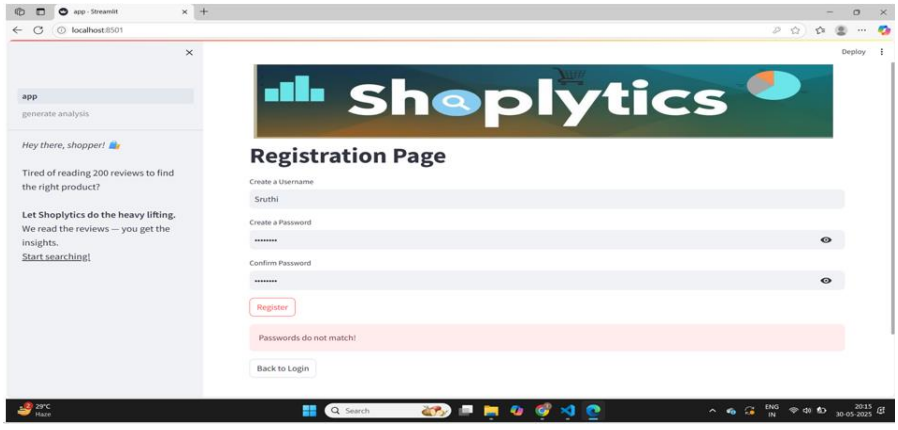
Screenshot 3 Registration page

Registration Page with Existing Username



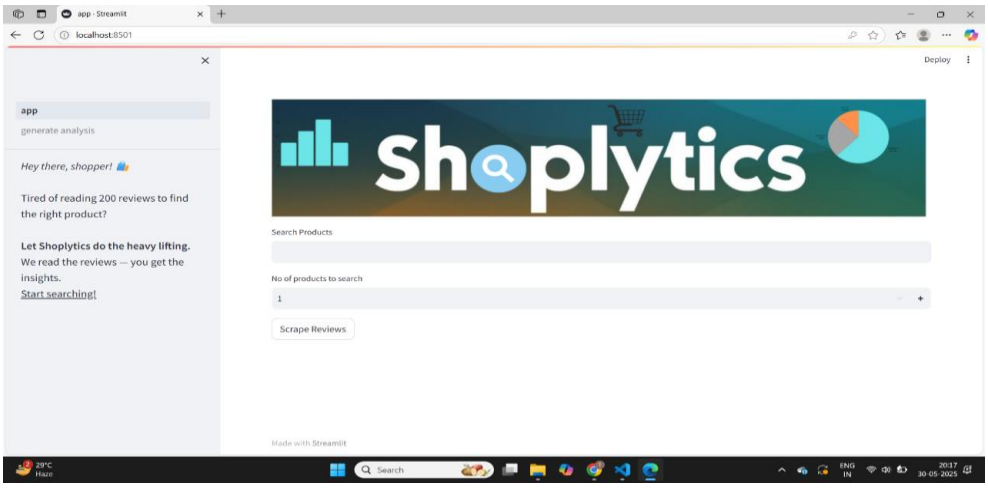
Screenshot4 Registration Page with Existing Username

Registration Page with Validation Error



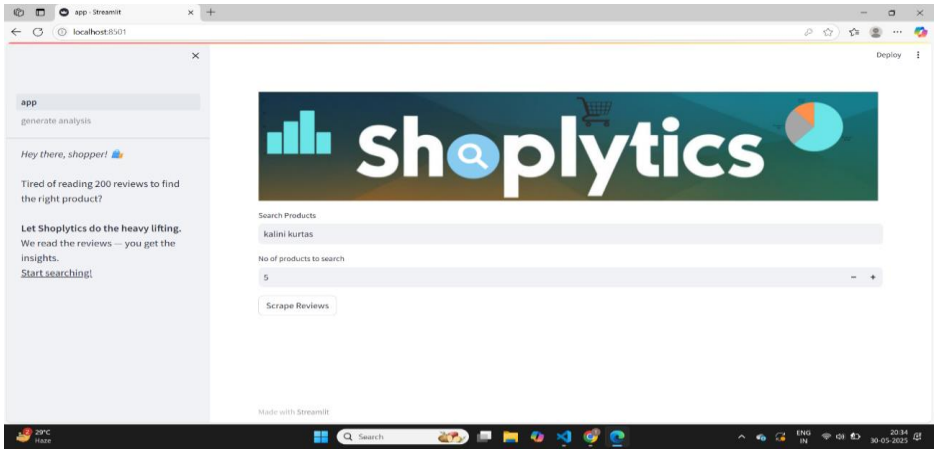
Screenshot 5 Registration Page with Validation Error

Home Page



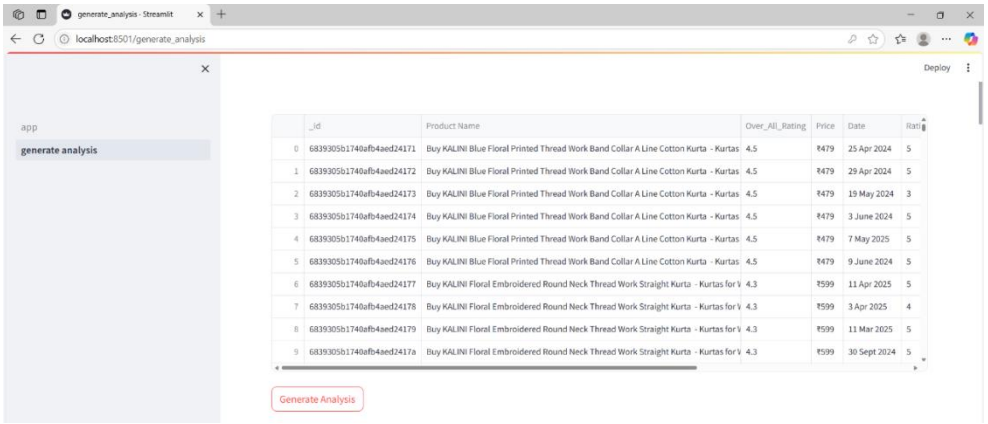
Screenshot 6 Home Page

Product Input Page

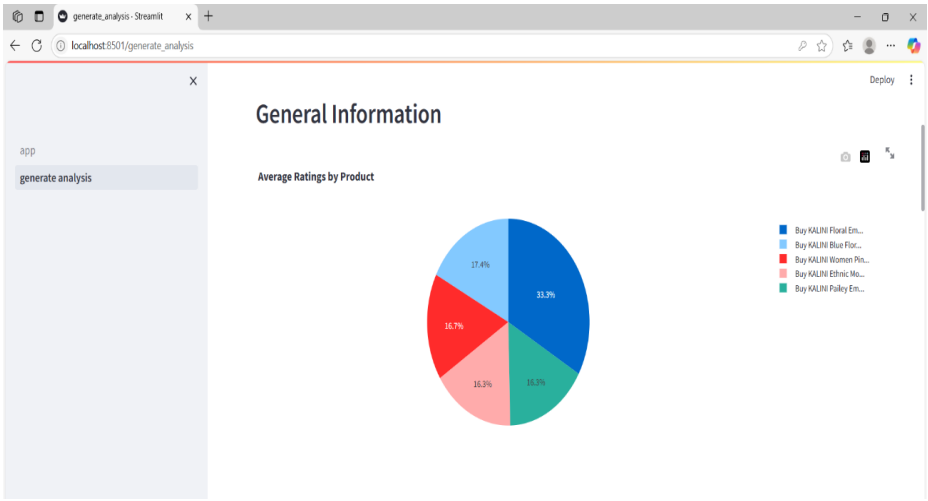


Screenshot 7 Product Input Page

Product Detail Page



Screenshot 8 Product Detail Page



Screenshot 9 Average Rating by Product

Screenshot 10 Average Price Comparision Between Products

6-CONCLUSION

"Shoplytics" enhances e-commerce intelligence by automating customer review analysis. It provides actionable feedback using sentiment analysis and ensures efficient data handling through MongoDB. This results in better product decisions and an improved shopping experience.

REFERENCES

- Sentiment Analysis on Product Reviews - R. S. Gupta, A. Sharma, and K. Mehta. 2024. *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 5, pp. 1207-1216.
- Sentiment Analysis on E-commerce Reviews and atings Using ML & NLP Models to Understand Consumer Behavior - P. Kumar, R. Gupta, and N. Reddy. 2024. *IEEE Transactions on Artificial Intelligence*, vol. 41, no. 3, pp. 897-909.
- Sentiment Analysis to Review Products Based on Machine Learning - A. Verma, T. S. Bansal, and V.R. Rao. 2024. *IEEE Access*, vol. 12, pp. 5421-5435.

Sentiment Analysis on Large Scale Amazon Product Reviews - M. S. Joshi, S. R. Jain, and P. N. Sharma. 2024. *IEEE Transactions on Big Data*, vol. 22, no. 4, pp. 1430-1442.