

*PREDICTIVE ANALYSIS LAB PROJECT*

*REPORT*

*GROUP 2*

|  |  |
| --- | --- |
| NAME | SAPID |
| 1.SHUBH DULWANI | 500107409 |
| 2.VARTUL CHANANA | 500107363 |

E commerce Product Recommendation System Using GoogleCollab

1. Introduction

The project integrates sentiment analysis with collaborative and content-based filtering to enhance the accuracy of recommendations. By leveraging customer reviews, the system refines suggestions, helping users make informed decisions.

2. Objectives

Develop a personalized product recommendation system.

Combine collaborative and content-based filtering.

Incorporate sentiment analysis to evaluate customer reviews.

3. Data Collection

Sources: E-commerce platforms, review websites.

Methods:

APIs: Extract product and user data.

Web Scraping: Collect customer reviews and product metadata.

Databases: Store data in CSV format for easy preprocessing.

Privacy Compliance: User data anonymized; datasets securely stored.

4. Data Preprocessing

Cleaning:

Handled missing values.

Removed duplicates.

Ensured consistent product and review data.

Transformation:

Scaled numerical data and encoded categorical features.

Segmentation:

Divided data into users, products, and interactions (e.g., purchases, clicks, reviews).

5. Exploratory Data Analysis (EDA)

Statistical Insights:

User demographics, product categories, and purchase frequency.

Visualizations:

Heatmaps and bar charts for user behavior.

Pie charts for product popularity.

Correlation matrices for interaction patterns.

6. Recommendation Algorithm

Collaborative Filtering:

User-based: Suggests items based on similar users.

Item-based: Recommends items similar to those already interacted with.

Content-Based Filtering:

Considers product features (e.g., category, price) and customer reviews.

Hybrid Model:

Combines collaborative and content-based filtering for higher accuracy.

Sentiment Analysis:

Analyzes reviews to classify sentiments (positive, neutral, negative), refining recommendations.

7. Model Development

Validation: Cross-validation to ensure robustness.

Metrics: RMSE, precision, recall, and F1-score for evaluation.

8. Model Evaluation

Performance evaluated using:

Precision and Recall: Relevance of recommendations.

F1 Score: Balance of precision and recall.

AUC-ROC Curve: Classification accuracy.

RMSE and MAE: Regression accuracy.

9. Deployment

Integration: Implemented using a Flask API for real-time recommendations.

Scalability: Deployed on AWS with load balancing for handling high traffic.

10. Monitoring and Maintenance

Feedback Collection: Evaluate relevance of recommendations from users.

Performance Monitoring: Track response time and engagement metrics.

11. Iterative Improvements

Regular updates with new data and feedback.

Experimentation with algorithms and features for better performance.

12. Tools and Technologies

Languages: Python, R.

Libraries: Scikit-learn, TensorFlow, Pandas, NumPy.

Databases: MySQL.

Conclusion

The personalized recommendation system effectively integrates sentiment analysis with collaborative and content-based filtering. It provides more accurate and relevant product suggestions. Future iterations will focus on incorporating additional data sources and refining algorithms to adapt to evolving user preferences.