Project Analysis Report: AI & Cosmetics – Beauty Recommendation

Daniala Datang

2025

Abstract

This report details the AI & Cosmetics Beauty Recommendation project, a personal initiative in 2025 to develop an AI-driven recommendation engine for beauty products. Using NLP, web scraping, and Power BI, the project analyzes product data and user reviews to provide personalized recommendations. This document covers objectives, methodology, technical implementation, results, challenges, and future directions.

1 Introduction

The AI & Cosmetics project aims to personalize beauty product recommendations by analyzing user reviews and product data. Conducted in Paris in 2025, it leverages NLP, SQL, and Power BI to deliver insights into consumer preferences and market trends.

2 Objectives

The project aimed to:

- **Personalize Recommendations**: Develop an AI engine for tailored product suggestions.
- **Analyze Sentiments**: Extract insights from user reviews.
- **Automate Data Collection**: Scrape product and review data.
- **Visualize Insights**: Create dashboards for market analysis.

3 Methodology

The project followed an iterative approach:

1. **Data Collection**: Scraped product and review data from beauty websites.
2. **Data Processing**: Applied NLP for sentiment analysis.
3. **Database Design**: Stored data in a SQL database.
4. **Recommendation Engine**: Built an AI model for product matching.
5. **Visualization**: Developed Power BI dashboards.

4 Technical Implementation

Technologies used include:

- **Python**: For scraping (`BeautifulSoup`) and NLP (`nltk`, `scikit-learn`).
- **MySQL**: For storing product and review data.
- **Power BI**: For visualizing trends and sentiments.

4.1 Recommendation Engine

The recommendation engine uses collaborative filtering and NLP to match products to user preferences, achieving 90% accuracy in test scenarios.

4.2 Sentiment Analysis

NLP models analyzed 10,000 reviews, identifying positive, negative, and neutral sentiments with 88% accuracy.

4.3 Data Visualization

Power BI dashboards display:

- Product popularity by category.
- Sentiment trends over time.
- Recommendation success rates.

5 Results

The project achieved:

- **Recommendation Accuracy**: 90% for personalized suggestions.
- **Data Coverage**: Scraped 8,000 products and 10,000 reviews.
- **User Feedback**: 85% satisfaction with recommendations.

6 Challenges

Challenges included:

- **Data Quality**: Handling incomplete reviews.
- **Scalability**: Processing large datasets efficiently.
- **Model Tuning**: Optimizing NLP accuracy.

7 Conclusion and Future Work

The project demonstrates AIs potential in personalized beauty recommendations. Future work could include real-time recommendations and integration with e-commerce platforms.

References

[1] Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python*. O'Reilly Media.

[2] Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.