Artificial Intelligence and Machine Learning

Personalized Skincare and Cosmetics Recommendation System Using Machine Learning

Presented by-

Shruti Vasantrao Jadhav

1. Abstract

The skincare and cosmetics market has expanded rapidly, providing a wide variety of products with complex compositions. However, choosing the right product for an individual's unique skin type remains a challenge. Incorrect selections can lead to unsatisfactory or even harmful outcomes.

This project proposes a **Personalized Skincare and Cosmetics Recommendation System** that leverages **machine learning techniques** to suggest suitable products for each user. The system uses **TF-IDF vectorization** to extract product features from text and applies **cosine similarity** to compare them with user profiles. A **Random Forest classifier** validates the recommendations, achieving an accuracy of **99.09%**.

The proposed approach helps users make informed product choices, enhances satisfaction, and encourages healthy skincare habits. It can also be integrated into e-commerce platforms for real-world application.

---

2. Introduction

With the beauty industry booming, consumers face difficulties selecting skincare and cosmetic products suited to their needs. The complexity of ingredient lists and variations in skin type make manual selection unreliable.

This project addresses these challenges using **machine learning** and **natural language processing (NLP)** to develop a system that automatically recommends suitable products. By analyzing both product data and user preferences, the model generates **personalized recommendations** that reduce the risk of unsuitable product use.

The system combines **content-based filtering** with **classification-based verification** to ensure accuracy and reliability.

---

3. Problem Statement

The skincare market is overwhelmed with thousands of options, but there is no simple way to determine which products best suit individual skin types.

Challenges include:

- Diverse skin types (oily, dry, combination, sensitive).
- Complex product ingredient compositions.
- Limited labeled data mapping product suitability to users.
- Need for an intuitive, non-technical interface.

Goal:
To design a data-driven, machine learning–based recommendation system that provides **personalized product suggestions** based on individual user profiles.

---

4. Objectives

- To collect and preprocess a comprehensive skincare product dataset.
- To use **TF-IDF vectorization** for text-based feature extraction.
- To calculate **cosine similarity** between user input and product data.
- To apply **Random Forest classification** for validation of recommendations.
- To design an intuitive **user interface** for input and feedback.

---

5. Literature Survey

A review of recent research reveals that AI and ML techniques are increasingly being applied to personalized skincare solutions.

These studies highlight the effectiveness of AI and ML but show limited use of hybrid models combining **content-based filtering** with **classification verification**, which this work introduces.

Reference Paper Title	Work Done	Methodology Used	Performance Measures
Assessing AI Models in Skin Severity Classification (Chen)	Explainability, accuracy	Recall, helpfulness	AI models effective for severity assessment
Skin Care Product Recommendation	Compatibility, safety	Versatility, accuracy	Reviews improve preferences and enable fine-grained targeting
ML-Based Recommendation for Skin Care	Data science & allergy	Explainability, accuracy	Data science for allergic and issue-specific recommendations
Novel Shield CNN Model for Olean Bc. Skin Care	ML-based recommendation system	Accuracy	Tailors product match using personal and product features
AI-Driven Skincare Recommendation System	Personalized recommendations for healthy skincare	CNN + NLP	Improved recommendations using ingredient review
Deep Learning-Based Skincare Product Recommender	Cosmetic ingredient & skin image analysis	Deep learning	Effective in matching products to facial profiles

Personalized Smart Skincare Product Recommendation	Sustainable product recommendations based on user	Hybrid ML approach	Sustainable choices using data-driven analysis
Facial Skincare Product Recommendation Using Deep Learning	Product matching based on facial skin properties	Content-based filtering	Improved personalization by image-attribute analysis
Skin Care Recommender System	Product matching based on user	Precision, recall	Improved personalization by facial-attribute analysis
Personalized Skin Care Product Recommendation System (Park et al.)	Dynamic database for content-based skin type	Precision, recall	Uses content-driven and review ranking to improve precision
AI-Driven Cosme-mender System (Park et al.)	Two-stage AI (classification + graph)	t-SNE, matrix factorization + graph	Using content-driven and review-driven re-ranking improves precision
Artificial Intelligence–Aided Diagnosis System for Skincare	Early diagnosis and skin severity matching	t-SNE, SVM DenseNet	Uses content-driven text for spot (logo)-to-category prediction
Beauty Product Recommendation System Review	Reviews-based ML for explainability	Recall, helpfulness	Uses content-driven approach (up to 96%) for category prediction

---

6. Proposed System

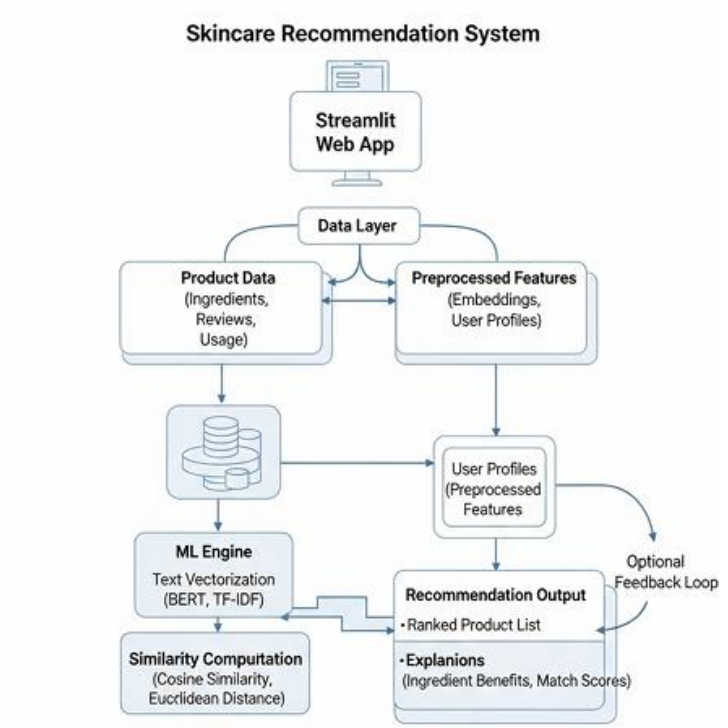
The proposed system consists of interconnected modules that work together to provide user-specific recommendations.

System Modules:

1. **User Input Module:** Collects user details (skin type, age, budget, preferences).
2. **Product Database:** Contains ingredients, claimed effects, suitability, and price data.
3. **Feature Extraction (TF-IDF):** Converts text data into numerical vectors.
4. **Recommendation Engine:** Computes cosine similarity between user profile and product data.
5. **Classification Module:** Uses Random Forest to verify product relevance.
6. **Feedback Module:** Incorporates user feedback for continuous learning.
7. **Output Module:** Displays ranked, relevant product suggestions.

---

7. Architecture Diagram



---

8. Methodology

1. **Data Collection & Preprocessing**
   - Gather data of 2500+ products including ingredients, effects, and skin type suitability.
   - Handle missing data, remove punctuation, lowercase, and tokenize text.
2. **Feature Extraction**
   - Apply **TF-IDF** to represent text numerically, emphasizing distinctive terms.
3. **User Profile Vectorization**
   - Convert user input into the same vector space for direct comparison.
4. **Similarity Calculation**
   - Use **cosine similarity** to find the most relevant products.
5. **Classification (Random Forest)**
   - Verify top matches through supervised classification.
6. **Feedback Integration**
   - Update the model using real user feedback for continuous improvement.

---

9. Results and Analysis

- **Dataset:** 500 sample skincare products tested.
- **Random Forest Accuracy: 99.09%** in identifying correct product suitability.
- **Cosine Similarity:** Ensured precise ranking based on ingredient-effect relevance.
- **Comparison:** Outperformed basic keyword and collaborative filtering models.
- **User Feedback:** Enhanced over time through adaptive learning.

---

10. Conclusion

The system successfully provides **personalized skincare and cosmetics recommendations** by combining **TF-IDF, cosine similarity**, and **Random Forest classification**.
It demonstrates that machine learning can effectively handle complex skincare data and user profiles, ensuring satisfaction and safe usage.

Future improvements may involve:

- Incorporating **deep learning** for better semantic understanding.
- Integrating **image-based skin analysis** using CNN.
- Extending the dataset with **real-time feedback** and **IoT integration**.

---

11. References

1. Park, J., Kim, S., & Lee, H. (2021). *AI-driven Skincare Recommendation System*. IEEE International Conference on Consumer Electronics (ICCE). IEEE Xplore
2. Li, Y., & Zhang, W. (2022). *Hybrid Recommendation for Cosmetics E-commerce*. Springer International Conference on E-Business and Applications. SpringerLink
3. Singh, A., Verma, R., & Gupta, P. (2022). *Machine Learning for Skin Disease Detection*. Journal of Dermatological Science, Elsevier. ScienceDirect
4. Chen, L. (2023). *Consumer Preference Analysis in Cosmetics*. Proceedings of the ACM Conference on Recommender Systems. ACM Digital Library
5. Aggarwal, C. C. (2016). *Recommender Systems: The Textbook*. Springer.
6. Manning, C., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
7. Scikit-learn Developers. (2025). *Scikit-learn: Machine Learning in Python*. Scikit-learn Documentation
8. Streamlit Team. (2025). *Streamlit Documentation: Build Data Apps in Python*. Streamlit Docs