# Abstract

Recommendation applications are becoming evermore demanding with the abundance of information is available online. In this generation of saving time, recommendation systems help develop viable suggestion for a user based on their characteristics made available through the vast collection of data that has been acquired in the past decade, and evermore in coming years. Skincare or cosmetics is an ever-expanding industry that has become more relevant recently given the amount of research made in the domain for people to increase their longevity and to get the desired skin/face that is perfect and inspiring.

Recommendation systems can be divided in to 2 main categories, these being the Content Based Filtering, and Collaborative Based Filtering. This project not only contributes in the form of preparing a dataset, but also presents a complete solution in the form of a mobile application which in rare cases work on both perspectives of a recommendation system. This application utilizes skincare data obtained from multiple sources to present with a solution that prepares models for both format and make suggestions based on the items’ characteristics and user features that relate to their previously liked items (like list) and skin specifications respectively. The models for this approach implement Tf-idf vectorizing technique along with cosine similarity taken for content-based filtering. Whereas, the other methodology follows a KNN and Decision tree approach taken for the item based and user based CBF. MAPE has been used as the evaluation criteria for the models which can be regarded as one of the basic standards for evaluating recommendation systems. Other measures discussed include hit ratio, confusion matrix and viable manual observations noted for suggesting a product. The models obtained with an accuracy of up to 80% are prepared for API usage which connect to the frontend application via a real-time database middleware.

Skincare companies would like to utilize the information they are collecting about their customers for making viable recommendations. The application in result offers a systematic solution to the skincare product user where they can rate, like and comment on the products. Thus, this application reduces the workload that a user has to do when finding and selecting the right skincare product for themselves.

# Contribution

There have been a few approaches where the author has implemented either a content-based filter or made suggestion via collaborative based filtering, but not both. (Adebo, 2020) implements content-based filtering using Tf-idf with the concept of comparing ingredients, their active chemical contents and nature of effect they had on human skins. A similar approach has been opted by (Songsri Tangsripairoj, 2018) in the form of a mobile application. This application utilizes both approaches i.e. content and collaborative based filtering and provides suggestion independently of one another. Content filtering uses the Tf-idf vector to categorize item on the basis of their category, brand and ingredients. Whilst item and user based collaborative filtering implement KNN, Decision trees and sentiment analysis respectively.

Skincare Recommendation approaches taken previously represent their work in the form of a research or just algorithm results. (Junaidi, 2021) present their solution in the form of a web application which takes into consideration only the user characteristics for user based collaborative approach. This application on the other hand presents a native mobile application in the form of a complete solution that is scalable, responsive and presents the user with multiple results and approaches.

The contribution of this project has not only been on the basis of results acquired, but also the steps taken to prepare the dataset. The collection of the right format of datasets for skincare recommendation is sparse indeed. Kaggle only presents with a cleaned and formatted data for Sephora whilst (Songsri Tangsripairoj, 2018) take into regard just the chemical ingredients and their effects on the human skin. (Junaidi, 2021) describe their working with a data set that contains only user data. Data pre-processing and engineering have played an entirely different criterion in this application. User characteristics were collected, normalized, regularized along with a similar procedure being carried out for user reviews and ratings. A list of 2000+ unique Products was accumulated alongside their reviews, ratings, URL and images. Previous approached all considered data from a single source with less complexity plus limited features to analyse making the algorithm smart, but their prediction was limited to the extent of its knowledge.

Finally, I’ve accumulated a result of multiple algorithms such as KNN, Decision trees, Vader sentiment and Tf-idf vectorization to discuss on the idea of the which is more superior based on the provided dataset and its features. These machine learning models have shed some light as to why many recommendation systems usually implement these specific algorithms such as why a natural language-based application is preferred for content based or review based recommendation system. The few literatures such as (Songsri Tangsripairoj, 2018) that exists in this domain only consider implementing a single model or algorithm.

# References

Adebo, A., 2020. *A Natural Language Processing Approach to a Skincare Recommendation Engine,* Dublin: National College of Ireland.

Junaidi, N. F. A. a. S. N. Z. M. a. A. N. a. K. H., 2021. A Personalized Web Cosmetic Recommendation System Based on Skin Type. *Applied Information Technology And Computer Science,* 2(1), pp. 225--234.

Songsri Tangsripairoj, K. K. P. P. Y. B., 2018. *SkinProf: An Android Application for Smart Cosmetic and Skincare Users.* Nakhon Pathom, IEEE, pp. 1-6.