

Cosmetic Recommendation Approach Using Machine Learning

Authors : Shakila N. Senarath, Dinesh Asanka, Janaka Wijayanayake

Abstract : The necessity of cosmetic products is arising to fulfill consumer needs of personality appearance and hygiene. A cosmetic product consists of various chemical ingredients which may help to keep the skin healthy or may lead to damages. Every chemical ingredient in a cosmetic product does not perform on every human. The most appropriate way to select a healthy cosmetic product is to identify the texture of the body first and select the most suitable product with safe ingredients. Therefore, the selection process of cosmetic products is complicated. Consumer surveys have shown most of the time, the selection process of cosmetic products is done in an improper way by consumers. From this study, a content-based system is suggested that recommends cosmetic products for the human factors. To such an extent, the skin type, gender and price range will be considered as human factors. The proposed system will be implemented by using Machine Learning. Consumer skin type, gender and price range will be taken as inputs to the system. The skin type of consumer will be derived by using the Baumann Skin Type Questionnaire, which is a value-based approach that includes several numbers of questions to derive the user's skin type to one of the 16 skin types according to the Bauman Skin Type indicator (BSTI). Two datasets are collected for further research proceedings. The user data set was collected using a questionnaire given to the public. Those are the user dataset and the cosmetic dataset. Product details are included in the cosmetic dataset, which belongs to 5 different kinds of product categories (Moisturizer, Cleanser, Sun protector, Face Mask, Eye Cream). An alternate approach of TF-IDF (Term Frequency - Inverse Document Frequency) is applied to vectorize cosmetic ingredients in the generic cosmetic products dataset and user-preferred dataset. Using the IF-IPF vectors, each user-preferred products dataset and generic cosmetic products dataset can be represented as sparse vectors. The similarity between each user-preferred product and generic cosmetic product will be calculated using the cosine similarity method. For the recommendation process, a similarity matrix can be used. Higher the similarity, higher the match for consumer. Sorting a user column from similarity matrix in a descending order, the recommended products can be retrieved in ascending order. Even though results return a list of similar products, and since the user information has been gathered, such as gender and the price ranges for product purchasing, further optimization can be done by considering and giving weights for those parameters once after a set of recommended products for a user has been retrieved.

Keywords : content-based filtering, cosmetics, machine learning, recommendation system

Conference Title : ICICT 2022 : International Conference on Information and Computer Technology

Conference Location : Kuala Lumpur, Malaysia

Conference Dates : February 15-16, 2022