Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

A Hybrid Recommendation System for Pregnancy-Safe Skincare: Integrating Keyword-Based and Rule-Based Classification with Content-Based Filtering

Jihan Syahira Adnanda Putri*[1], Kholid Haryono[2]

[1,2]Informatics, Universitas Islam Indonesia, Indonesia

Email: [1]jihan.putri02@students.uii.ac.id

Received : Apr 21, 2025; Revised : Jun 21, 2025; Accepted : Jun 24, 2025; Published : Jul 9, 2025

Abstract

Pregnant women struggle to select safe skincare products, often relying on social media and blog searches, and manual ingredient checking. Choosing safe ingredients is essential, as exposure to unsafe substances may lead to teratogenic effects and endocrine disruption, which can result in fetal abnormalities such as retinoic acid embryopathy and neurodevelopmental disorders. Exposure to retinoids, for instance, has been associated with a 20–30% incidence of fetal retinoid syndrome in affected pregnancies. This study develops an integrated recommendation system using three techniques: (1) keyword-based classification with regular expressions to detect 50 unsafe ingredients across 8 categories; (2) rule-based classification using IF-THEN statements matching products with 5 pregnancy-related skin conditions; and (3) content-based filtering utilizing TF-IDF vectorization and cosine similarity for safer alternatives. The system achieved 86.25% accuracy in safety classification, with high recall (97.50%) indicating strong ability to identify safe products. However, moderate precision (79.59%) suggests some unsafe products were misclassified as safe, highlighting need for improvement in safety-critical contexts. Pilot user evaluation using ResQue framework with 10 participants yielded scores of 4.50–4.85 across 8 dimensions, achieving 4.65 overall average. This research demonstrates effective integration of multiple recommendation methods in context-sensitive applications, enabling safer product selection during pregnancy. By providing accessible, personalized, and evidence-based information, the system enables pregnant women to make informed skincare decisions and continue their routines despite limited access to healthcare services.

Keywords : *Content-based filtering, Keyword-based classification, Pregnancy-safe skincare, Recommendation system, Rule-based classification.*

1. INTRODUCTION

Human skin, the body's largest organ, protects against external threats while simultaneously regulating temperature, moisture levels, and sensory perception. Significant hormonal, immunologic, and metabolic changes affect skin appearance and function during pregnancy, leading to heightened sensitivity and conditions such as hyperpigmentation, melasma, dryness, acne, and stretch mark [1], [2].

Research indicates that nearly 70% of pregnant women develop melasma, while up to 90% experience hyperpigmentation during the first two trimesters [2], [3]. These skin changes occur because pregnancy triggers significant physiological and hormonal shifts that increase melanocyte activity, thereby raising the risk of post-inflammatory hyperpigmentation (PIH) [4]. The impact of these hormonal changes extends beyond melasma, affecting other skin conditions as well. For instance, in a comprehensive study examining 295 pregnant women with acne, researchers found varying degrees of severity: 56.6% presented with mild acne, 29.5% with moderate acne, 1.2% with severe acne, and 1.7% with very severe acne [5].

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

To address these challenges, skincare products can be selected based on their ingredients. However, these products are not without risks. Some ingredients commonly found in skincare products, such as retinol, hydroquinone, and parabens, are associated with risks ranging from endocrine disruption to teratogenic effects [6], [7], [8]. Topical tretinoin (all-trans-retinoic acid), for instance, is frequently prescribed for conditions like acne and photoaging, making it particularly relevant to women of reproductive age. Its teratogenic potential has raised concern due to its structural similarity to isotretinoin (13-cis-retinoic acid), a well-established human teratogen known to cause a distinct pattern of fetal abnormalities collectively referred to as retinoic acid embryopathy, which includes malformations of the brain, heart, ears, and thymus [9]. Moreover, in animal studies, systemic administration of all-trans-retinoic acid has demonstrated developmental toxicity equal to or greater than that of isotretinoin, leading to stage- and dose-dependent defects in the central nervous system, eyes, palate, and limbs [9]. Notably, isotretinoin itself carries a significant risk, with studies reporting a 20%–30% incidence of fetal retinoid syndrome in exposed pregnancies [10].

Chemical substances in skincare, such as microplastics, benzophenones, parabens, phthalates, and metals, may disrupt the nervous system and cross the placental barrier, affecting the embryo. Phthalates can also be detected in the urine of pregnant women, suggesting exposure during pregnancy. The effects of phthalates on fetal growth are complex and can vary, with some studies showing inconsistent results regarding their impact on small-for-gestational age (SGA) births [11], [12]. Early exposure during critical developmental phases can lead to neurodevelopmental disorders and congenital enteric neuropathies. Additionally, these neurotoxins can be secreted in breast milk, prolonging exposure to the newborn even after birth [13], [14].

The application of recommendation systems in skincare has gained considerable attention in recent years, driven by the increasing complexity of product selection and personalized beauty needs. Traditional skincare recommendation systems primarily focus on general population needs, utilizing various computational approaches to match users with suitable products. Several studies have explored different methodologies for skincare recommendations, ranging from simple content-based filtering to more sophisticated hybrid approaches. [15] developed a skincare recommendation system using deep learning techniques that analyzed user skin images to suggest appropriate products, while [16] proposed a hybrid system combining content-based and collaborative filtering for skincare analysis.

However, existing digital platforms for pregnancy-safe skincare face significant usability challenges and typically address only isolated aspects of the problem. Many current solutions require users to manually input ingredient lists and lack comprehensive support for identifying suitable product alternatives. Given the high prevalence of pregnancy-related skin conditions, an effective recommendation system should account for both product safety and therapeutic efficacy. This research gap presents an opportunity for a more integrated approach. While machine learning classification methods offer effective solutions [17], they depend on extensive labeled datasets, which are currently unavailable for pregnancy-safe ingredients and pregnancy-related skin conditions. To fully address this issue, it is important to expand the research and offer safe replacement options, since pregnant women who discover their routine skincare products contain unsafe ingredients require effective replacements. Unlike existing solutions that focus on single aspects of the problem and remain difficult for pregnant women to use efficiently, this approach uniquely combines filtering features that allow users to select their skin type and specific pregnancy-related skin conditions to receive personalized recommendations. The safety assessment includes detailed explanations of why specific ingredients are unsafe, empowering users to make informed decisions, while simultaneously providing alternative products based on ingredient similarity and other product attributes such as benefits and suitable skin types match. This comprehensive approach ensures that pregnant users not only understand the safety concerns but

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

also receive immediate access to suitable alternatives that maintain therapeutic effectiveness for their specific skin needs.

Recommendation systems provide a promising solution by generating personalized suggestions based on user preferences and product characteristics. These systems are widely used in e-commerce, content streaming, and online services, employing techniques like content-based approaches, collaborative filtering, and hybrid models [18]. The recommendation process typically involves three stages: collecting user data through explicit or implicit feedback; analyzing the data using algorithms to identify preferences; and delivering tailored recommendations [18].

According to [19], traditional recommendation systems are commonly categorized into three primary types: (1) content-based filtering, which recommends items similar to those a user has previously liked; (2) collaborative filtering, which suggests items based on the preferences of similar users; and (3) hybrid systems, which combine both approaches to leverage their respective strengths and mitigate individual limitations. Despite their widespread application across various domains, there remains a paucity of research focused on recommendation systems tailored for pregnancy-safe skincare products.

Content-based recommendation systems leverage user profiles and item characteristics to suggest relevant items based on content similarity [20]. This method is particularly effective when the system has limited data on user preferences or aims to provide personalized recommendations [21]. In skincare-related applications, feature selection can vary significantly. For instance, [22] utilized ingredients as the primary feature, whereas [23] incorporated multiple product attributes, including benefits, skin concerns, and ingredients. Their system also integrated user skin types through direct input or facial image analysis, demonstrating the potential of multi-feature recommendation systems in enhancing personalization.

According to Huang et al. [24], traditional recommendation systems often rely on single variables, which can lead to the neglect of important user preferences. Although content-based filtering may result in overspecialization, it aligns with the objective of this study: to provide product recommendations based on ingredient similarity.

To address the complex needs of pregnant women, this study introduces a novel integrated approach that combines three distinct techniques to create a comprehensive and personalized recommendation system. The system utilizes keyword-based classification to automatically flag harmful ingredients based on a curated database of unsafe components for pregnancy. Simultaneously, rule-based classification determines product suitability for common pregnancy-related skin conditions such as melasma, PIH, and acne by analyzing ingredient effectiveness based on dermatological research. To further assist users in finding safe alternatives, content-based filtering utilizing TF-IDF and cosine similarity recommends similar products free from unsafe ingredients.

This integrated system delivers an end-to-end experience where users can input their specific skin conditions, search for products by name, receive comprehensive safety and effectiveness information, and immediately access suitable alternatives when a product is flagged as unsafe. The system also features a comparison function that enables users to evaluate differences between their original selection and safer alternatives, creating a tailored decision-making experience specifically designed for pregnant users. The main contributions of this paper are mentioned below:

1. A novel integrated system that combines keyword-based safety checks, rule-based classification for treating pregnancy-related skin conditions, and content-based filtering to deliver safe and personalized skincare recommendations for pregnant women.
2. A comprehensive end-to-end solution that not only identifies unsafe products but also immediately suggests suitable alternatives based on ingredient similarity and therapeutic effectiveness for specific pregnancy-related skin conditions.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

User evaluation, conducted using pregnancy-specific user scenarios and feedback based on the ResQue Framework from Pu and Chen [25], demonstrates improved personalization, usability, and safety assurance compared to existing single-method recommendation models. Meanwhile, safety evaluation was carried out using ground truth data to calculate precision, recall, F1-score, and accuracy, providing a comprehensive assessment of the model's performance.

The selection of the methods was driven by several factors: (1) Keyword-based classification detects pregnancy-safe or unsafe ingredients through pattern matching, chosen for its minimal data requirements and interpretability when labeled datasets are unavailable. Given the critical nature of ingredient safety during pregnancy, this transparent approach enables precise, traceable decisions over potentially unreliable predictive models. (2) Rule-based classification addresses complex scenarios where ingredient safety depends on contextual factors or compound interactions by encoding domain-specific knowledge into structured rules. This method has demonstrated effectiveness in healthcare information extraction, as shown by [26] who showed their applicability in high-stakes biomedical decision-making. (3) Content-based filtering is used to generate personalized product recommendations by analyzing features such as ingredients and claimed benefits. While content-based systems typically face cold-start limitations due to the absence of user interaction data [27], this issue is addressed by initiating recommendations from products flagged as unsafe. By using these flagged products as reference points, the system can suggest safer alternatives based on ingredient similarity, enabling relevant and safety-aware recommendations even in the absence of user history.

Figure 1 presents the key fields of study that inform this research, illustrating how recommendation systems connect approaches from data mining, keyword-based, and rule-based classification. The proposed system applies these interconnected approaches using keyword-based classification to identify potentially risky ingredients, rule-based classification to determine product safety, and content-based filtering to suggest similar but safer alternatives for pregnant users.

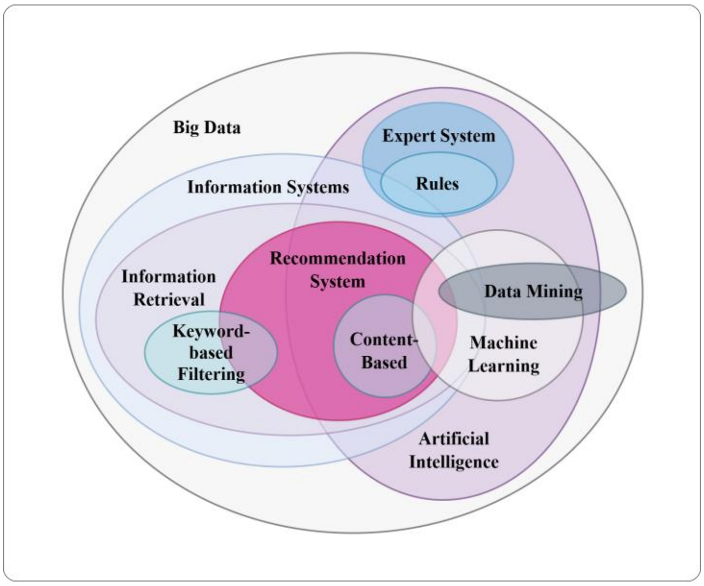

Figure 1. Field of Research

2. METHOD

The research process begins with problem identification, followed by the development of a structured recommendation system integrating content-based, rule-based, and keyword-based approaches. The system is implemented as a mobile application, and its effectiveness is evaluated using the ResQue framework.

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 3, Juni 2025, Page. 1643-1668
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

2.1. Recommendation System Development

The recommendation system is developed through several stages: the data collection and preprocessing, and the development of a content-based recommendation system. Figure 2 shows how the system gives alternatives to the users. The system architecture was adapted from the method proposed by Lee *et al*. in [28].

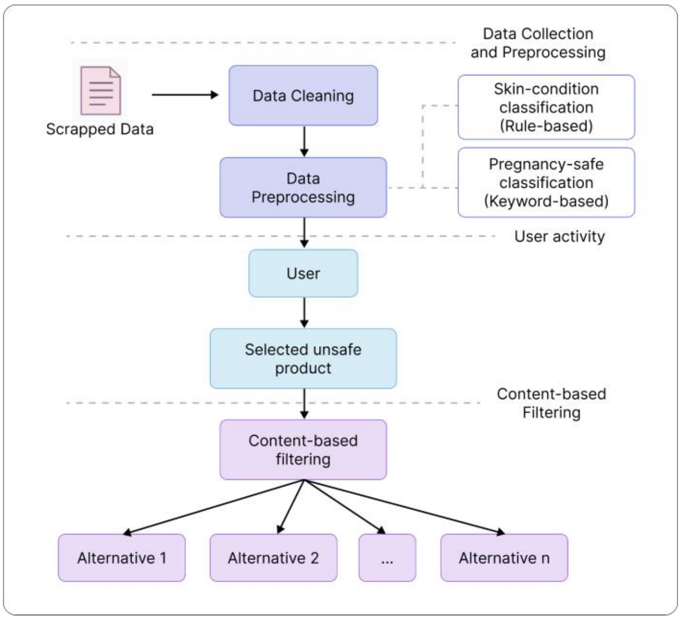

Figure 2. Framework of The Recommender System

1. Data Collection and Preprocessing

Product data was collected from skinsort.com using a web scraping approach implemented in Python 3.9, with Selenium WebDriver 4.0 handling dynamic content and BeautifulSoup4 parsing the HTML structure. The scraping process involved navigating through paginated listings and extracting structured information from individual product pages [29], including product name, brand, category, image URL, type, ingredients, good_for tags, benefits, concerns, and included/excluded features. To ensure data quality, completeness was verified after scraping. Incomplete entries were either re-scraped for critical fields, deleted if non-essential, or flagged as "unknown." Duplicate records were also identified and removed to maintain a dataset of unique entries.

After cleaning the data, a two-phase classification process was implemented during preprocessing. First, each product was labeled using a keyword-based classification approach to determine whether it is pregnancy-safe, based on the presence of specific ingredients. This method was selected for its interpretability, as the classification results are easily understood and explained [30]. The products are marked as unsafe when certain assigned ingredients are detected in their composition. Such transparency is particularly crucial in sensitive contexts like pregnancy, where users must clearly understand why a product is considered unsafe.

Following this initial safety classification, a rule-based classification approach was then applied to identify which pregnancy-related skin conditions each product might effectively treat. These conditions include melasma, hyperpigmentation, acne, PIH, and stretch marks. The rule base, which contains a set of application-specific rules, can be represented either as a list of rules or as a decision table [31]. In rule-based classification, these predefined rules are used to categorize items, typically following the form of IF-THEN statements. The rule structure used for the skin condition is shown in Equation (1).

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

$$IF\ CONTAINS\ P\ THEN\ CONDITION\ is\ Q \qquad (1)$$

The rule-based classification approach was applied to assign condition labels based on ingredient analysis. The same keyword-based classification technique was then used to detect the presence of condition-treating ingredients within each product's formulation based on the established rules.

All of these classifications and parts of the data-cleaning process were implemented primarily using regular expressions (RegEx), which are sequences of characters that define search patterns for efficient pattern matching and string manipulation in text data. RegEx is widely used in both data preprocessing and classification tasks due to its ability to identify and manipulate specific patterns in unstructured text. Varshney and Torra [31] emphasize the role of regular expressions in text classification, especially for performing operations on strings. In the context of data cleaning, RegEx enables operations such as removing unwanted characters, extracting relevant information, and standardizing data formats, thereby enhancing data quality and consistency [32].

RegEx was used as well to extract skin type information from the "good for" column by identifying patterns for dry, oily, and sensitive skin. Along with skin conditions from rule-based classification, these features were one-hot encoded to convert categorical labels into binary vectors for modeling. For example, a product for dry and sensitive skin is encoded as dry=1, oily=0, sensitive=1; and one treating acne and PIH as acne=1, PIH=1, others=0.

2. Content-based Recommendation System

The content-based filtering model in this research is implemented using TF-IDF vectorization and cosine similarity to measure product similarity based on textual features. According to Lui et al. [33], the TF-IDF method is widely used for generating sentence vectors by weighting words based on their importance, factoring in both their frequency within a specific document and their inverse frequency across the entire corpus. The primary goal of this model is to recommend alternative products that are most similar to a selected item. If a product is identified as unsafe, the system will suggest safer alternatives with closely matched content profiles.

To construct the feature vectors, several product attributes were selected based on their relevance to skincare product formulation and consumer preferences. These include ingredients, benefits, included_features, and skin concern tags such as hyperpigmentation, PIH, acne, stretch_marks, melasma, dry_skin, oily_skin, and sensitive_skin. The attributes category and type were also included to help match products with similar functional or usage classifications. These selected textual features were concatenated into a single string per product entry, forming the document corpus for vectorization. The TF-IDF vectorizer was then applied to this corpus, assigning each product a high-dimensional vector where each dimension corresponds to a weighted term. The TF-IDF value is computed using Equations (2) and (3) [34].

$$TF - IDF(t,d,D) \ = \ tf(t,d) * idf(t,D) \qquad (2)$$

Where:
tf(t, d): The term frequency, representing the count of term t in document d.
idf(t, D): The inverse document frequency, which quantifies the rarity of term t across the document corpus D. Document frequency (DF) is defined as:tf(t, d): The term frequency, representing the count of term t in document d. Document frequency (DF) is defined as:

$$TF - IDF(t,d,D) \ = \ tf(t,d) * idf(t,D) \qquad (3)$$

Where:

P-ISSN: 2723-3863
https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

t(n): The number of documents containing the term t.
D(n): The total number of documents in the corpus.

To avoid infinite values when a term does not appear in any document, the logarithm of the inverse document frequency is used in Equation (4).

$$IDF(t,D) = log((D(n))/(t(n))) \quad (4)$$

After converting products into vector representations using TF-IDF, cosine similarity is then used to measure the degree of similarity between two product vectors. This similarity is calculated using the formula in Equation (5) [35],[36].

$$similarity(A,B) = cos(\theta) = \frac{A.B}{||A||||B||} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \quad (5)$$

Where:
$\boldsymbol{A}$: Feature vector of the target product
$\boldsymbol{B}$: Feature vector of the compared product
$\boldsymbol{A}_i$, $\boldsymbol{B}_i$: The i-th elements of vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively

After computing cosine similarity scores between the selected product and all other items in the dataset, the system ranks the results from highest to lowest. Only products labeled as safe are included in the final recommendation list. Products with the highest similarity scores are considered the most comparable and are suggested as safer alternatives. To enhance interpretability, the similarity score is converted into a percentage and displayed to the user, allowing them to understand the degree of similarity between the original product and its alternatives. This process ensures that the recommended products not only align with the user's initial preferences but also exclude ingredients identified as unsafe.

3. Integration and Implementation

The recommendation process leverages outputs from the data preprocessing stage: (1) the keyword-based classification is used to flag products as pregnancy-safe or unsafe; and (2) the rule-based classification assigns each product to relevant skin conditions based on its ingredient composition.

These safety labels and condition classifications are incorporated into the filtering logic. The cosine similarity scores are used to find the closest matches, while the safety and condition tags serve as constraints to ensure that only safe and relevant alternatives are shown. To enhance user trust and facilitate informed decision-making, the mobile interface displays both a similarity percentage and an ingredient match count for each product recommended.

2.2. System Development

The system integration is implemented in a mobile application that provides users with a seamless experience in exploring safe skincare alternatives. It was built using the Flutter framework to ensure cross-platform compatibility and deliver a responsive user interface. The application architecture consists of data storage using MySQL and an application programming interface (API) built with FastAPI, which contains the core recommendation system. Docker containers are also utilized to streamline future deployment. The data deployed using Ngrok. This implementetion is for user evaluation purposes.

The mobile application features four main screens: the home page, product detail page, alternatives page, and comparison page. Within the application, each recommended product is displayed with a cosine similarity score and an ingredient match count to help users make informed decisions.

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

Filtering options are also available, allowing users to narrow down results based on specific skin conditions.

2.3. Evaluation

1. Safety Evaluation

To evaluate the performance of the keyword-based classification system for assessing the safety of skincare products during pregnancy, a confusion matrix was used. This matrix compares the predicted classification outcomes of the system with a medically curated ground truth from a website [37]. Although validation by medical experts would enhance reliability, such review was not feasible due to ethical constraints.

The structure of the confusion matrix is presented in Table 1 and consists of four components: true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) [38]. The evaluation involved 80 skincare products, with a balanced distribution between safe and unsafe categories to ensure fair assessment. From this comparison, four key outcomes were identified as shown in table x [38].

Table 1. Confusion Matrix Structure

		Actual Values	
		Safe	Unsafe
Predicted Values	Safe	True Positive (TP)	False Positive (FP)
	Unsafe	False Negative (FN)	True Negative (TN)

In this study, accuracy, precision, and recall were used as key performance metrics to evaluate the effectiveness of the keyword-based classification system in identifying unsafe skincare products during pregnancy [39]. Accuracy reflects the overall correctness of the system's predictions, measuring the proportion of all products (both safe and unsafe) that were classified correctly as shown in Equation (6).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

While accuracy provides a general measure of performance, it can be misleading when the cost of certain errors, such as failing to detect an unsafe product is high. To address this, precision and recall are also considered. Precision measures the proportion of products predicted as unsafe that are actually unsafe, indicating the system's ability to issue reliable warnings. A high precision means the system rarely misclassifies safe products as unsafe. Recall, on the other hand, measures the proportion of truly unsafe products that are successfully detected, capturing the system's sensitivity to risk showed in Equations (7) and (8) [39].

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

$$Recall = \frac{TP}{TP+FN} \tag{8}$$

2. User Evaluation

The system was evaluated using the ResQue framework [25], a widely adopted model for assessing user satisfaction and perceived quality in recommender systems. ResQue comprises 13 constructs and 60 question items, which collectively measure key dimensions such as recommendation quality, system usability, perceived usefulness, interaction and interface design, and overall user satisfaction. The framework also assesses users' behavioral intentions, including their likelihood to reuse the system, recommend it to others, and consider purchasing recommended products.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

For this study, participants were asked to respond to 21 statements distributed across the eight ResQue dimensions, each rated using a 5-point Likert scale (see Table 3). These statements assess specific aspects such as recommendation relevance, diversity, enjoyment, ease of use, interface satisfaction, trust, and behavioral intention. The evaluation covers the following questions from provided framework taken one questions for each aspects as shown in Table 2.

Table 2. Resque Question

No	Category	Aspects	Questions
A1	Quality of Recommended Items	Accuracy	The items recommended to me matched my interests.
		Relative Accuracy	The recommendation I received better fits my interests than what I may receive from a friend.
		Familiarity	Some of the recommended items are familiar to me.
		Attractiveness	The items recommended to me are attractive.
		Enjoyability	I enjoyed the items recommended to me.
		Novelty	The recommender system helps me discover new products.
		Diversity	The items recommended to me are diverse.
		Context Compatibility	The recommendations are timely.
A2	Interaction Adequacy		The recommender provides an adequate way for me to express my preferences.
A3	Interface Adequacy		The layout of the recommender interface is attractive and adequate.
A4	Perceived Ease of Use	Ease of Initial Learning	I became familiar with the recommender system very quickly.
		Ease of Preference Elicitation	I found it easy to tell the system about my preferences.
		Ease of Preference Revision	It is easy for me to get a new set of recommendations.
		Ease of Decision Making	Finding an item to buy with the help of the recommender is easy.
A5	Perceived Usefulness		The recommended items effectively helped me find the ideal product.
A6	Control/ Transparency		The system helps me understand why the items were recommended to me.
A7	Attitudes		The recommender made me more confident about my selection/decision.
A8	Behavioral Intentions	Intention to Use the System	If a recommender such as this exists, I will use it to find products to buy.
		Continuance and Frequency	I prefer to use this type of recommender in the future
		Recommendation to Friends	I will tell my friends about this recommender.
		Purchase Intention	I would buy the items recommended, given the opportunity.

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

Table 3. Likert Scale

Category	Value
Strongly disagree	1
Disagree	2
Neither agree nor disagree	3
Agree	4
Strongly Agree	5

The evaluation involved 10 participants who met two inclusion criteria: (1) they had experienced pregnancy within the past four years or were currently pregnant, and (2) had actively used skincare products before or during pregnancy. This sample size is appropriate for a pilot quantitative study, particularly given the specialized nature of the target population and the pre-deployment status of the system. Research involving pregnant women presents unique recruitment challenges, and similar studies in digital health interventions for this population have successfully employed comparable sample sizes [40], [41]. Participants provided written informed consent, including agreement to the collection of personal data such as age, skin type, and pregnancy-related skin conditions. All data were anonymized and collected in accordance with ethical research standards.

3. RESULT

The result chapter presents the output of the implemented methodology. This chapter contains the findings from the data collection process, preprocessing techniques, safety classification algorithms, rule-based skin condition classification, and the content-based filtering approach that powers the recommendation engine.

3.1. Recommendation System Development

The development of the recommendation system involves several steps. These include data preprocessing, which covers data cleaning and safety classification using a keyword-based method, rule-based classification for classifying products to pregnancy-related skin conditions, and the implementation of a content-based filtering recommendation system.

3.1.1. Data Collection and Preprocessing

The web scraping process resulted in 26.266 skincare products from global brands, categorized into nine groups: cleansers, masks, treatments, moisturizers, sunscreen, lip care, serums, body care, and other. The data was compiled into a data frame with the following column names: category, brand, product, type, image, benefits, good_for, country, ingredients, concern, included_feature, and excluded_feature.

The raw dataset, as a result of web scraping, contained various unnecessary characters such as quotation marks, double spaces, and square brackets, which were cleaned using Pandas and RegEx. Feature extraction was followed to identify unique types, countries, and ingredients, which were then used to support filter functionality in the mobile application.

Table 4 presents a summary of missing data analysis, showing high completeness across the key attributes used in the recommendation system. Products with missing essential fields were handled accordingly: for instance, products missing ingredient data were flagged as ‘unknown’, while other missing attributes were left empty but retained for context. A total of 466 duplicate records were removed based on matching brand, product name, and ingredient lists. Additionally, text fields were standardized for consistency, and a manual spot-check was performed to verify the accuracy of the scraping process.

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

Table 4. Missing Data Analysis

Attribute	Product with Missing Values	Completeness Rate
Good_for	1037	96.1%
Benefits	1028	96.1%
Included_features	64	99.8%
Ingredients	63	99.8%
Type	0	100%

To determine product safety, a keyword-based classification approach was applied. Based on a literature review, a list of unsafe ingredients was compiled, as shown in Table 5. To prioritize precaution, any product containing one or more of these ingredients was classified as unsafe, regardless of the individual risk level posed by each component. An exception was identified in the Metals and Trace Metals category, where only titanium dioxide and zinc oxide, as listed among UV filters in the U.S. Food and Drug Administration (FDA) monograph, are classified as "Generally Recognized as Safe and Effective (GRASE)." [42], and were therefore considered safe in this system.

Table 5. Potentially Unsafe Ingredients for Pregnancy

Category	Ingredients	Source
Microplastics and nano plastics	Polyethene (PE), Polypropylene (PP), Polyvinylchloride (PVC), Polystyrene (PS), Polylactic (PLA)	[13], [43]
Parabens	Methylparaben (MtP), Butylparaben (BuP), Ethylparaben (EtP), Propyl paraben (PrP)	
Benzophenone	Benzophenone-1, Benzophenone-2, Benzophenone-3/Oxybenzone, Benzophenone-4	
Phthalates	Di-ethyl-phthalate (DEP), Dimethyl-phthalate (DMP), Din-butyl phthalate (DBP)	
Metals and Trace Metals	Lead (Ld), Aluminium (Al), Cadmium (Cd), Nickel (Ni), Arsenic (As), Mercury (Hg), Manganese (Mn), Titanium dioxide (TiO2), Chromium (Cr), Iron (Fe), Copper (Cu), Cobalt (Co)	
Vitamin A Derivatives	Retinoic Acid, Tretinoin, Retinol, Retinal, Retinyl Acetate, Retinyl Propionate, Retinyl Palmitate, Adapalene, Tazarotene, Hydroxypinacolone Retinoate	[43-45]
Hormonal Therapy	Spironolactone, Trimethoprim-Sulfamethoxazole, Dapsone, Benzoyl Peroxide	
Skin Lightening	Hydroquinone	[46]
AHA/BHA	Salicylic Acid	[13]

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

Category	Ingredients	Source
Endocrine Disruption	Octinoxate, 4-Methylbenzylidene camphor	[7]
Resorcinol Derivatives	Resorcinol, Phenylethyl Resorcinol, 4-Butylresorcinol, Hexylresorcinol	

These unsafe ingredients were detected using RegEx patterns as part of the keyword-based classification process. If a product contained any of the predefined harmful ingredients, it was flagged as ‘unsafe’. A new column labeled ‘safe’ was added to indicate the product’s safety status, while an additional column, ‘unsafe_reason’, was added to specify the exact ingredient(s) responsible for the classification.

The keyword-based classification effectively identified products containing unsafe components and categorized them accordingly. A summary of the classification results is presented in Table 6. Products labeled as ‘Unknown’ lack ingredient data, resulting in unclassified conditions.

Table 6. Safety Classification Overview

Condition	Total
Safe	19.958
Unsafe	6.245
Unknown	63

The top five most common reasons for a product being marked as unsafe are shown in Figure 3. Salicylic Acid is the most frequently found unsafe ingredient in the dataset, followed by Methylparaben, Copper, Retinol, and Retinyl Palmitate.

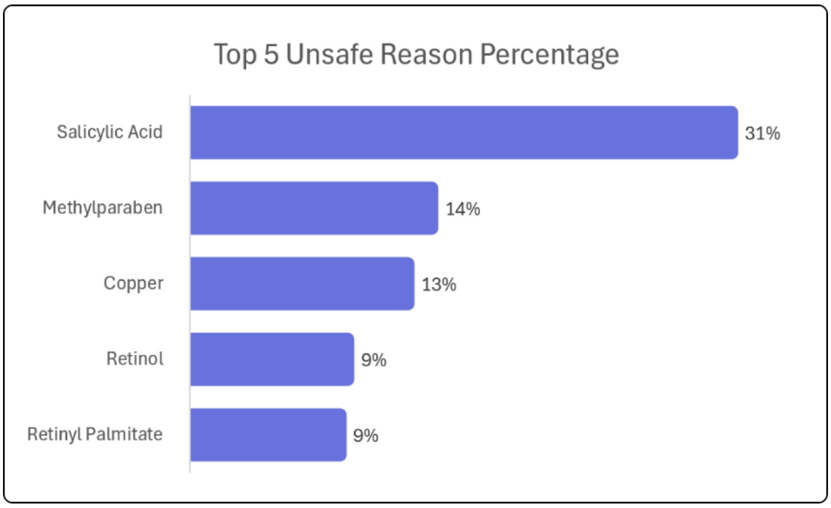

Figure 3. Top Five Unsafe Reason Percentage

To associate skincare products with pregnancy-related skin conditions, this system implemented a rule-based classification system derived from a review of dermatology and cosmetic science literature [47-51]. Each rule maps one or more ingredients to a specific skin condition such as acne, melasma, hyperpigmentation, PIH, or stretch marks based on ingredients that have been reported to show clinical or anecdotal efficacy in treating those conditions.

Table 7 presents the rule base for classifying ingredients used to treat these skin conditions during pregnancy, excluding percentage restrictions due to the limited information provided on product labels [47-51].

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

Table 7. Rule-based Skin Condition

Rule	Content
Rule 1	IF CONTAINS Zinc Coceth Sulfate OR Portulaca OR Paris Polyphylla OR Bayberry OR Camellia Dianshan OR Pomegranate OR Nigella Sativa OR Resveratrol OR Hop OR Matricaria Recutita OR Licorice OR Avocado OR Green Tea OR Fruit Acid OR Alfa Hydroxy Acids OR Salicylic Acid OR Niacinamide OR Retinol OR Zinc THEN CONDITION is Acne
Rule 2	IF CONTAINS Niacinamide OR (Niacinamide AND Desonide) OR (Ellagic Acid AND Salicylic Acid) OR Turmeric THEN CONDITION is Hyperpigmentation
Rule 3	IF CONTAINS Azelaic Acid OR (Azelaic Acid AND Glycolic Acid) OR Mulberry OR Licorice OR (Licorice AND Bellis Perennis AND Emblica) OR (Kojic Acid AND Ascorbic Acid AND Hydroquinone) OR (Kojic Acid AND Hydroquinone) OR (Ellagic Acid AND Arbutin) OR Ellagic Acid OR Arbutin OR Green Tea OR Soy OR (Ascorbic Acid AND Trichloroacetic Acid) OR (Ascorbic Acid AND Vitamin C) OR Ascorbic Acid THEN CONDITION is Melasma
Rule 4	IF CONTAINS Azelaic Acid OR Niacinamide THEN CONDITION is PIH
Rule 5	IF CONTAINS Tocopherol OR Almond Oil THEN CONDITION is Stretch Marks

For example, Rule 1 labels a product as targeting acne if it contains any of the following ingredients: Niacinamide, Zinc, Salicylic Acid, or Green Tea, compounds widely cited for their anti-acne properties. Rule 3 identifies products for melasma if they include combinations such as Azelaic Acid and Glycolic Acid, or Kojic Acid and Vitamin C, which are commonly used as depigmenting agents in clinical treatments.

These rules were implemented using RegEx patterns that scanned product ingredient lists. If an ingredient matched a rule, the product was flagged with the corresponding skin condition in a new 'condition' column. The rule-based classification successfully assigned skin condition tags based on predefined ingredient-condition mappings. From the database, it was found that 14.495 products matched the classification rules outlined in Table 8. The overview of the distribution percentages is illustrated in the pie chart in Figure 4.

Table 8. Total Classification for Each Skin Condition

Condition	Total
Stretch Marks	5.945
Acne, Hyperpigmentation, PIH	5.016
Acne	2.569
Melasma	852
Melasma, PIH	103
Hyperpigmentation	10

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

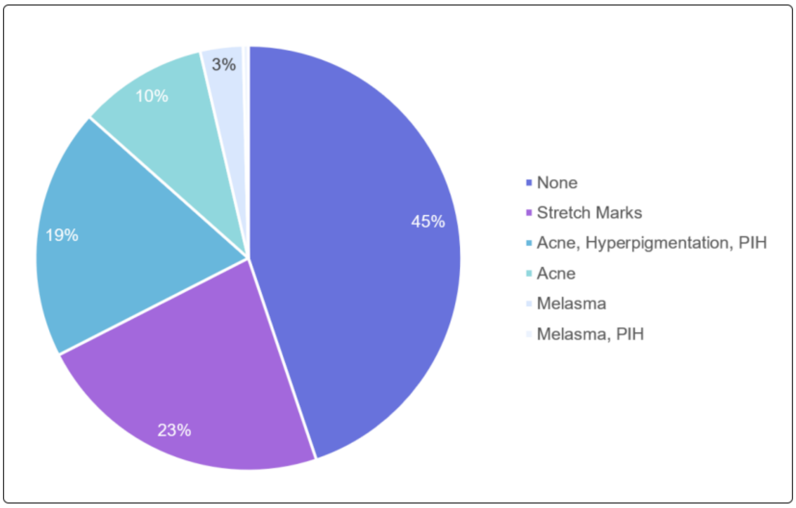

Figure 4. Distribution of Skin Conditions in Percentage

Once the new columns for safety and skin conditions have been added, one-hot encoding is applied to transform categorical data into a binary format. The columns transformed to one-hot encoding are the 'condition' column and the skin type information inside the 'good_for' column.

3.1.2. Content-Based Recommendation System

In developing this recommendation system, a content-based method was used and contains two main steps: vectorization of product features and similarity calculation. The system transforms textual product descriptions and features into numerical vectors that can be compared mathematically, enabling the identification of similar products based on multiple characteristics.

The TF-IDF Vectorizer plays a crucial role in the recommendation system by converting textual descriptions of skincare products into a numerical format. In this system, TF-IDF is used to vectorize each feature in the recommendation system.

To illustrate this process, consider Table 9, which outlines the TF-IDF analysis for the product Advanced Snail 96 Mucin Power Essence. The chosen sample ingredients for this product include: Snail Secretion Filtrate; Betaine; Butylene Glycol; 1,2-Hexanediol; Sodium Polyacrylate; and Phenoxyethanol. Next, the product weights are calculated. For example, three products are considered:

1. Product 1: Snail Mucin 95 + Peptide Essence
2. Product 2: Snail Repair Intensive Ampoule
3. Product 3: Snail Mucin 5000 Ampoule

Table 6 presents the TF-IDF calculations for these ingredients using Equations (2), (3), and (4).

Table 9. TF-IDF Vectorization

Ingredients	Product 1	Product 2	Product 3	Df	Idf	Tf.idf		
						Doc 1	Doc 2	Doc 3
Snail Secretion Filtrate	1	1	1	3	$log(3/3) = 0$	0	0	0
Betaine	1	1	0	2	$log(3/2) = 0.1761$	0.1761	0.1761	0
Butylene Glycol	1	1	1	3	$log(3/3) = 0$	0	0	0

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 3, Juni 2025, Page. 1643-1668
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

Ingredients	Product 1	Product 2	Product 3	Df	Idf	Tf.idf		
						Doc 1	Doc 2	Doc 3
1,2-Hexanediol	1	1	0	2	$log(3/2) = 0.1761$	0.1761	0.1761	0
Sodium Polyacrylate	1	1	1	3	$log(3/3) = 0$	0	0	0
Phenoxyethanol	1	1	0	2	$log(3/2) = 0.1761$	0.1761	0.1761	0
Sodium Polyacrylate	1	1	1	3	$log(3/3) = 0$	0	0	0

Once the TF-IDF matrix has been established, Cosine Similarity is used to evaluate the degree of similarity between products based on their features and benefits. This is accomplished by calculating the cosine of the angle between two vectors, with each vector representing a product's TF-IDF values as illustrated in Equation (5). A cosine value closer to 1 indicates that the two products are highly similar regarding their key attributes. This similarity metric enables the recommendation system to identify and suggest products that align closely with a user's preferences. The resulting similarity scores are stored in a matrix, allowing the system to retrieve and recommend products with high cosine similarity scores. This ensures that the recommendations are relevant and tailored to meet the user's specific skincare needs, including considerations for safety during pregnancy.

For example, using the ingredients illustrated in TF-IDF vectorization, the following query vectors are defined for the main product and Products 1, 2, and 3:

1. Main (A): [0.1761,0.1761,0.1761,0,0.1761,0.1761]
2. Product 1 (B): [0,0.1761,0,0.1761,0,0.1761]
3. Product 2 (C): [0,0.1761,0,0.1761,0,0.1761]
4. Product 3 (D): [0,0,0,0,0,0]

The similarity is calculated using the formula in Equation (5) for the main product and each of the Products:

$$similarity(A.B) = \frac{(0.1761\times0)+(0.1761\times0.1761)+(0.1761\times0)+(0\times0.1761)+(0.1761\times0)+(0.1761\times0.1761)}{\sqrt{(5x(0.1761^2)}\ x\ \sqrt{(3x(0.1761^2)}}$$

$$similarity(A.B) = 0.062/(0.3939\ x\ 0.305)\ \approx\ 0.5167$$

$$similarity(A.C) = \frac{(0.1761\times0)+(0.1761\times0.1761)+(0.1761\times0)+(0\times0.1761)+(0.1761\times0)+(0.1761\times0.1761)}{\sqrt{(5x(0.1761^2)}\ x\ \sqrt{(3x(0.1761^2)}}$$

$$similarity(A.C) = \frac{0.062}{0.3939\ x\ 0.305}\ \approx\ 0.5167$$

$$similarity(A.D) = \frac{0}{0.3939\ x\ 0} = 0$$

According to the results, Products 1 and 2 have a similarity level with the core product (Advanced Snail 96 Mucin Power Essence) that is similar (around 0.5167). As such, these two products will be pushed higher on the recommendations when users search for other products similar to the main product.

In the system, two types of similarity metrics are used to evaluate product recommendations: ingredient match count and feature similarity score. The ingredient match count is determined by counting the number of identical ingredients shared between the original product and its recommended alternatives. Meanwhile, the feature similarity score is retrieved from the cosine similarity results generated by the TF-IDF vectorization of ingredient lists. This score reflects the overall similarity

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

between products based on their textual ingredient representation and is used to rank the most relevant alternatives.

3.2. System Integration

The mobile application was developed to integrate all components of the system. Figure 5 illustrates the flow of the application in showing the recommendation.

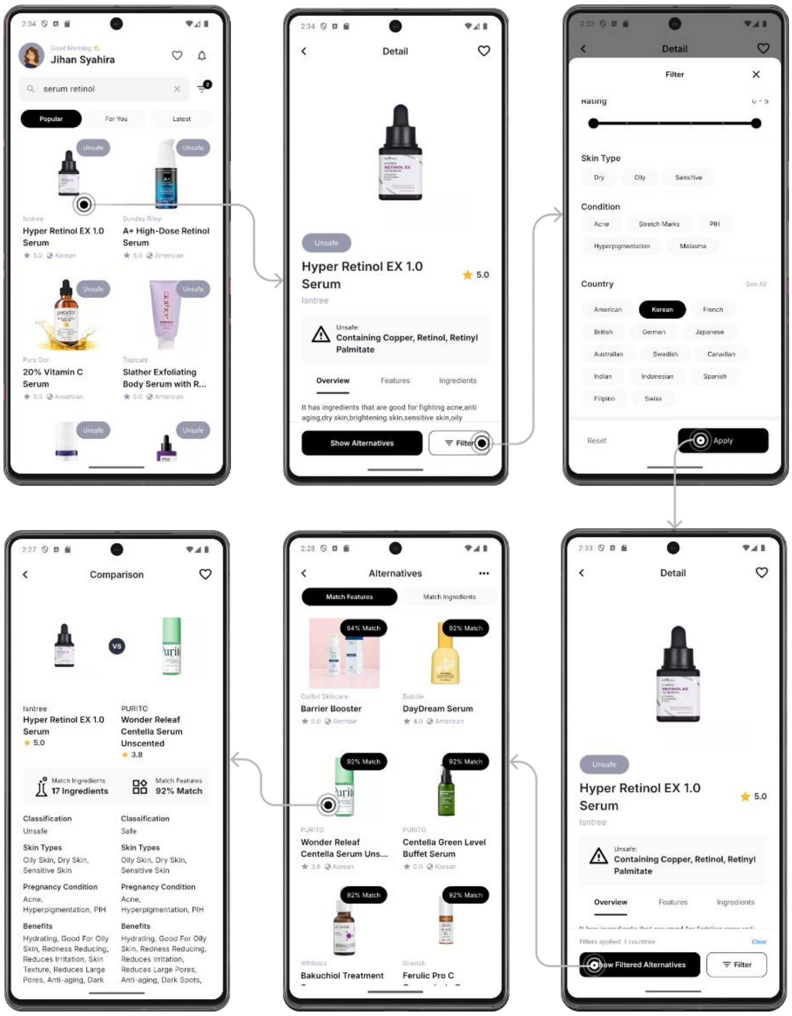

Figure 5. Recommendation System Integration in a Mobile Application

The developed system integrates the techniques used to provide a personalized recommendation system for pregnancy-safe skincare. As illustrated in Figure 6, the user journey begins on the home page. A search bar is also available, enabling users to locate specific products efficiently. Upon selecting a product, users are directed to the product detail page. This page presents comprehensive information, including the product's country of origin, compatibility with different skin types, key features (benefits, concerns, and skin conditions that can be solved), and ingredient composition. The page displays the product's safety status, determined through rule-based classification techniques. If a product is labeled as unsafe, information on specific harmful ingredients is provided.

To facilitate safer choices, the interface includes a "Show Alternatives" button. By selecting this option, users are redirected to the alternative products page, where a list of safer and similar products is generated using a content-based recommendation system. This system evaluates the similarity of product features and ingredients to the original unsafe item. Additionally, users can apply filters to tailor

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

the recommendations to their individual needs, which also includes the pregnancy-specific skin conditions identified during the data preprocessing phase. To assist users in making informed decisions, the system also offers a comparison page, allowing direct comparison between the unsafe product and selected alternatives. The comparison includes matched ingredients and a similarity percentage, thereby supporting a transparent evaluation of product compatibility and safety.

3.3. Evaluation

1. Safety Evaluation

The proposed keyword-based classification system demonstrated strong performance in identifying safe and unsafe skincare products for pregnancy. The evaluation dataset comprised 80 skincare products, which were classified by the system and compared against manually validated ground truth labels. Out of these products, the system correctly classified 39 safe products and 30 unsafe products, resulting in an overall accuracy of 86.25%. The detailed performance metrics are presented in the confusion matrix shown in Table 10.

Table 10. Confusion Matrix Result

		Actual Values	
		Safe	Unsafe
Predicted Values	Safe	TP = 39	FP = 10
	Unsafe	FN = 1	TN = 30

The system achieved a precision of 79.59% and a recall of 97.50%. While the high recall value indicates the system successfully identified most truly safe products, the precision reveals a concerning limitation: of all products classified as safe, only 79.59% were actually safe according to the ground truth. This means that 10 out of 49 products labeled as safe were actually unsafe.

A comprehensive analysis of the 11 misclassified products revealed several important patterns. The single false negative occurred when a safe product was incorrectly classified as unsafe due to the system's intention to minimize potential risks by adopting a highly cautious ingredient list. The ten false positive cases, where unsafe products were predicted as safe, revealed two primary issues. First, several ingredients present in the unsafe products according to the ground truth were not included in the system's unsafe ingredient list. These included glycolic acid, gluconolactone, salicylic acid derivatives, licorice root extract, willow bark extract, dihydroxyacetone, melasyl, and certain chemical sunscreen compounds without specific classifications. Second, the system flagged several ingredients as potentially unsafe that were considered safe in the ground truth dataset, such as methylparaben, ethylparaben, copper, manganese, Phenylethyl Resorcinol, Hexylresorcinol, and Benzophenone-3. Additionally, two misclassifications were attributed to incomplete ingredient information provided by the original data source

2. User Evaluation

The pilot user evaluation was conducted using the ResQue framework. Given the initial evaluation's limited sample size, descriptive statistics were applied to analyze the questionnaire responses. Data was collected after participants completed a guided task scenario involving unsafe product detection, alternative exploration, and product comparison. Figure 6 presents the demographics of the respondents who participated in this pilot study, including their skin types, age distribution, and various skin conditions experienced during pregnancy.

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

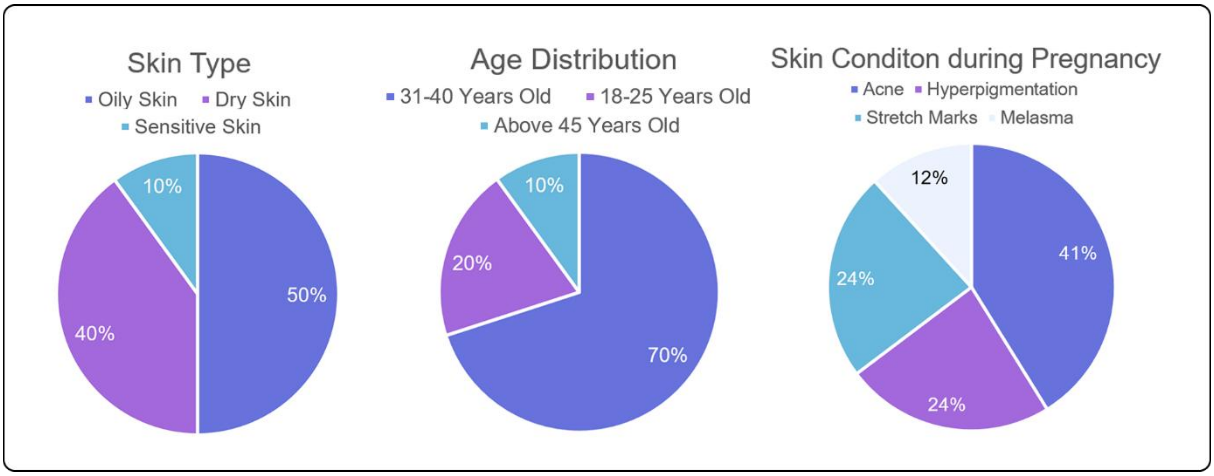

Figure 6. Demography Report (n=10)

Table 11 presents the results of the ResQue questions categorically. Overall, participants perceived the skincare recommendation system as an effective tool for pregnant women, with a total average score of 4.65 out of 5 across all categories. This positive trend was maintained across different evaluation aspects.

Table 11. Questionnaire Results

	A1	A2	A3	A4	A5	A6	A7	A8
Total Question(s)	7	1	1	4	1	1	1	4
Average	4.59	4.50	4.60	4.85	4.80	4.50	4.60	4.78
Std. deviation	0.59	0.67	0.49	0.36	0.40	0.67	0.49	0.42

The highest-rated categories based on overall averages were Perceived Ease of Use (A4) with an average score of 4.85, followed by Perceived Usefulness (A5) at 4.80, and Behavioral Intentions (A8) at 4.78. Category A4 assessed accuracy, familiarity, attractiveness, enjoyability, novelty, diversity, and context compatibility, while A5 measured how effectively the system helped users find ideal products, and A8 reflected users' intentions to continue using the system and recommend it to others.

When examining individual question responses as shown in Figure 7, several items received consistently high ratings, with multiple participants giving the maximum score of five. Q13: "It is easy for me to get a new set of recommendations." received the maximum scores, with all participants rating it as five. The following questions were also highly rated, with more than 8 out of 10 participants giving a maximum score of 5:

- Q1: "The items recommended to me matched my interests."
- Q6: "The recommendation system helps me discover new products."
- Q10: "The layout of the recommendation system interface is attractive and adequate."
- Q12: "I found it easy to tell the system about my preferences."
- Q14: "Finding an item to buy with the help of the recommender is easy."
- Q18: "If a recommender such as this exists, I will use it to find products to buy."
- Q21: "I would buy the items recommended, given the opportunity."

These highly-rated questions span across different evaluation categories, including discovery capability, interface design, ease of interaction, perceived usefulness, and purchase intention.

Conversely, Interaction Adequacy (A2) and Control/Transparency (A6) received the lowest average scores (both 4.50). Category A2 contains Q9: "The recommender provides an adequate way for me to express my preferences," while A6 contains Q16: "The system helps me understand why the items were

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

recommended to me." Several statements received at least one neutral ratings (score 3), indicating areas for improvement:

- Q3: "Some of the recommended items are familiar to me."
- Q7: "The items recommended to me are diverse."
- Q8: "The recommendations are timely."
- Q9: "The recommender provides an adequate way for me to express my preferences."
- Q16: "The system helps me understand why the items were recommended to me."

The overall distribution of scores across all questions showed that 70% of responses were rated as 5 (strongly agree), 27% as 4 (agree), and only 3% as 3 (neutral), with no responses below 3. The strengths is in perceived usefulness and behavioral intentions, and weakness in interaction adequacy and control/transparency.

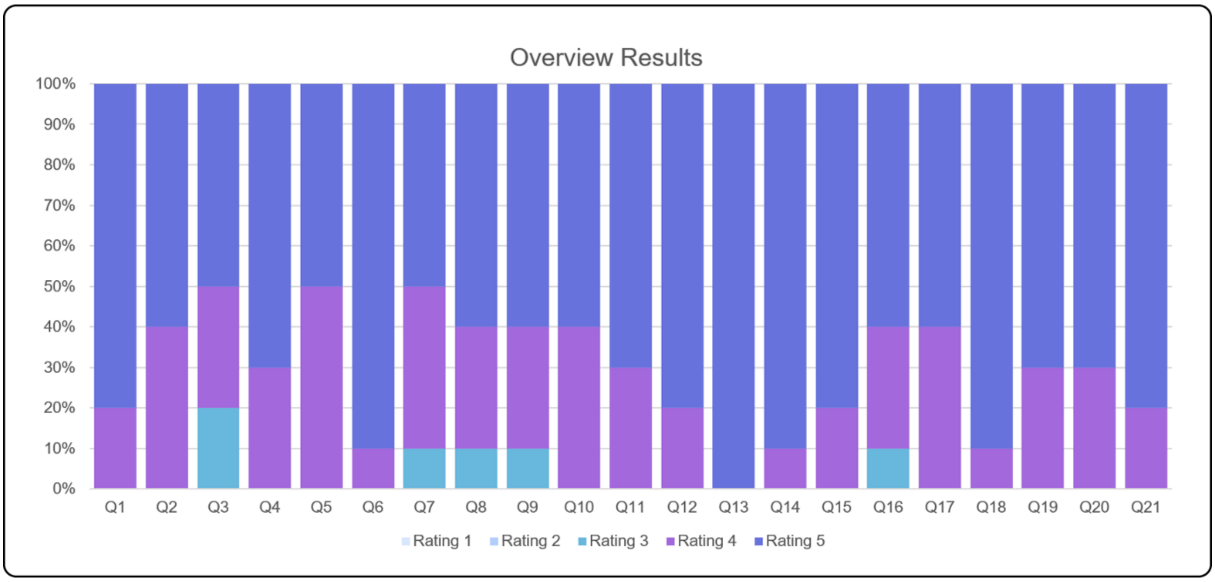

Figure 7. Overview Results for Each Question

3. DISCUSSION

This chapter analyzes the findings obtained from the development and evaluation of the system. It discusses the user evaluation results, compares the system with existing research in the field, and addresses the limitations of the current implementation while proposing directions for future work. The discussion provides insights into how the system's performance aligns with user needs and expectations, particularly regarding recommendation quality, interface usability, and safety considerations specific to pregnant women.

3.1. Evaluation Findings

1. Safety Evaluation

The safety evaluation achieved 86.25% accuracy with high recall (97.50%) but concerning precision (79.59%), meaning 10 out of 49 products labeled as safe in this system were actually unsafe. This represents a critical safety failure where pregnant users could unknowingly use harmful products.

Manual analysis revealed discrepancies between the ground truth and system classifications. Some ingredients marked unsafe in ground truth were not detected by the system due to differing safety standards. For example, glycolic acid was flagged as unsafe in ground truth but research shows alpha hydroxy acids are FDA Pregnancy Category B and safe at concentrations up to 10% with pH > 3.5 [43], with minimal skin penetration reducing systemic absorption. Conversely, some undetected ingredients

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

may warrant examination due to limited pregnancy safety documentation, indicating gaps in available safety data.

Data source limitations also contributed to errors, where some websites provided incomplete ingredient listings with critical components missing, leading to misclassification. The findings highlight the need for comprehensive ingredient databases, more complete data sources, and standardized safety criteria to address the complexity of ingredient safety classification in pregnancy applications.

2. User Evaluation

Beyond technical performance metrics, user evaluation revealed highly positive results for the skincare recommendation system, with an overall average score of 4.65 out of 5 across all ResQue categories. The exceptionally strong performance across most categories demonstrates that the system successfully addresses the critical needs of pregnant women seeking safe skincare products.

The perfect score for Q13 ("It is easy for me to get a new set of recommendations"), where all 10 participants rated it as 5, represents a significant achievement in user experience design. This unanimous positive response indicates that the system's core functionality for recommendation retrieval is highly intuitive and accessible, a crucial capability for pregnant women who may be navigating skincare concerns for the first time during pregnancy.

The outstanding performance in Perceived Ease of Use (A4) with an average of 4.85 (SD=0.36) validates the effectiveness of the content-based filtering approach in supporting user interaction across all ease-of-use dimensions. The low standard deviation indicates consistent positive experiences across participants, suggesting robust system design. The high scores for preference expression (Q12) and recommendation access (Q13) demonstrate that the system successfully addresses the technical barriers that often prevent pregnant women from finding suitable skincare products.

The high scores for Q21 ("I would buy the items recommended, given the opportunity") and Q18 ("If a recommender such as this exists, I will use it to find products to buy") indicate strong user trust in the system's safety assessments and practical value. This trust establishment is particularly significant given that skincare safety during pregnancy typically requires extensive research and verification from healthcare providers or reliable sources. The strong performance in Behavioral Intentions (A8: 4.78, SD=0.42) suggests that participants not only found the system useful but would actively recommend it to others, a critical factor for adoption in the pregnancy community where peer recommendations carry significant weight.

Despite the overall positive results, several areas present opportunities for improvement. The lowest-scoring category was Control/Transparency (A6: 4.50), with 10% of participants providing neutral responses to Q16 ("The system helps me understand why the items were recommended to me"). This finding highlights an improvement opportunity in the system's ability to provide clear rationales for its recommendations, more than just giving the percentage of similarity score based on features, total match ingredients and comparison with the chosen recommended product. For pregnant women, understanding the safety reasoning behind product suggestions is essential too for building trust and enabling informed decision-making, even tho the unsafe reason already provided. Similarly, the neutral responses (10%) for Q9 ("The recommender provides an adequate way for me to express my preferences") indicate that while current preference expression works, more sophisticated mechanisms could better capture the nuanced needs of pregnant women, such as trimester-specific concerns or varying sensitivity levels.

The neutral responses for Q3 (familiarity), Q7 (diversity), and Q8 (timeliness) suggest that the recommendation algorithm could be refined to better balance introducing new products with maintaining familiar options, and to improve contextual relevance based on pregnancy stage and seasonal factors. However, the lower performance in Interaction Adequacy (A2: 4.50, SD=0.67)

 Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

suggests that while basic interactions work well, more sophisticated preference communication mechanisms could be beneficial.

The overwhelmingly positive results, with 97% of responses rated as 4 or 5 (agree or strongly agree), strongly validate the effectiveness of developing specialized recommendation systems for specific user groups with unique safety requirements. The findings suggest that content-based filtering approaches, when properly implemented with domain-specific knowledge, can effectively serve specialized populations better than general-purpose recommendation systems. The increase in confidence when making decisions or selections after using the recommendation feature (A7: 4.60), combined with strong behavioral intentions, suggests that pregnant women are willing to adopt and recommend specialized tools that cater to their specific needs.

3.2. Comparison with Existing Research

Academic research addressing recommendation systems for safe skincare during pregnancy remains limited. Only one study by [52] proposed a prototype-based classification method to assess product safety, but their approach has significant limitations. Their unsafe ingredients list fails to comprehensively cover potentially harmful ingredients identified through extensive literature review, creating safety assessment gaps. Additionally, their system lacks personalization capabilities, missing integration of individual skin conditions and pregnancy-related skin changes crucial for effective recommendations.

Existing research focused on single-brand analysis using only the System Usability Scale (SUS) for evaluation. This study takes a different approach by implementing ResQue, a comprehensive framework specifically designed for recommendation system testing that examines both system effectiveness and interface aspects. By using ResQue, this approach helps understand the potential of this novel system more comprehensively.

Commercial solutions exist but carry significant limitations. The website gravidabeauty.com helps pregnant women verify ingredient safety but requires users to manually locate and copy-paste ingredient information from other sources, which is a time-consuming and error-prone process. Similarly, 15minutesbeauty offers static lists categorizing products as safe or unsafe for pregnancy, but users must manually search lengthy lists, receive no personalization, and receive no guidance toward suitable alternatives when discovering unsafe products.

The integrated approach addresses these limitations through several key advantages. It provides comprehensive safety coverage through an extensively researched ingredient database surpassing previous work. The system emphasizes personalization by integrating individual skin conditions and user needs into its recommendation algorithm. Unlike fragmented solutions that leave users stranded after identifying unsafe products, this system provides transparent explanations about problematic ingredients and seamlessly offers tailored alternative recommendations with optional filtering capabilities. The cross-brand approach ensures comprehensive options across manufacturers and price points, while detailed product comparisons with similarity percentages enable informed decision-making.

However, the approach presents disadvantages. While the system provides alternatives to unsafe products, individual skin variability remains a fundamental challenge. Although the algorithm seeks the most similar alternatives to prevent breakouts, skin compatibility varies significantly between individuals, meaning recommended products may still require user trial and adaptation. The recommendation methodology relies on traditional approaches rather than machine learning due to the absence of verified databases for training in pregnancy skincare safety and skin condition related to pregnancy. Additional challenges include increased system complexity compared to simple checkers, critical dependence on database accuracy requiring continuous updates, potential user adoption barriers

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

among those preferring familiar manual methods, significant computational resource requirements, and privacy considerations for handling sensitive pregnancy-related information.

3.3. Limitations and Future Work

This study has several important limitations that should be acknowledged. The biggest challenge is the reliance on web-scraped data for both product information and safety evaluations. This approach creates potential accuracy concerns since product information online may be outdated, inconsistent across different websites, or lack verification from authoritative sources. Product formulations change regularly, items get discontinued, and sometimes the information available online simply isn't accurate, which raises questions about whether users receive the most current and correct information about recommended skincare products.

The evaluation was conducted with a relatively small group of ten participants. While this was sufficient for initial testing and met the specific participant criteria, it limits how well the findings represent the broader population of pregnant women. The participants were fairly similar to each other, which means the results might not capture the full range of preferences, skin types, and needs that exist among pregnant women from different backgrounds and circumstances. The content-based filtering algorithm showed some limitations in providing diverse recommendations, as users gave neutral responses about the variety in suggestions they received. The system also scored lowest in Control/Transparency, with an average of 4.50, indicating that users wanted more understanding and control over how recommendations were being made.

These limitations point to several exciting directions for future research. One of the most important improvements would be incorporating user preferences more effectively. Future versions could include visual interfaces where users can select preferences by looking at images, or guided assessment that ask specific questions about pregnancy-related skin issues like melasma, stretch marks, or increased sensitivity. Personalizing recommendations based on pregnancy trimester and lactation status presents another significant opportunity for system improvement. To enable this functionality, a user login feature would be required, allowing pregnant and lactating users to indicate their current stage before accessing the app. This distinction is important because certain ingredients may be considered safe during specific trimesters but not others, or may pose different risks during breastfeeding. If this feature is implemented, the safety evaluation component would also need to account for ingredient risks specific to each stage of pregnancy, such as early, middle, and late trimesters, as well as during the lactation period, since safety considerations can vary significantly throughout these phases.

Integration with computer vision technology offers particularly innovative possibilities for the future. Advanced object detection models could analyze skin images captured through smartphone cameras to identify skin conditions, texture changes, pigmentation issues, and other visual indicators of skin health. This approach could automatically detect pregnancy-related skin changes such as melasma, acne flare-ups, or dryness patterns, allowing the system to adjust recommendations based on actual visual assessment of skin conditions rather than what users report about their skin. This could make recommendations much more accurate and responsive to how skin changes throughout pregnancy, since some pregnant women don't really understand the skin they have in some cases.

Future research should also include longer-term studies following users over extended periods, testing with larger and more diverse groups of participants, and getting feedback from healthcare professionals to validate safety and skin condition rules. Working with dermatologists and obstetricians could significantly improve the system's credibility and ensure the safety evaluations are as reliable as possible. For the system, the enhancement should focuses on harmonizing safety databases, improving data extraction processes to develop a more comprehensive and reliable pregnancy safety assessment system.

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

4. CONCLUSION

This paper presents a personalized recommendation system for pregnancy-safe skincare products. By integrating three distinct techniques: keyword-based classification for safety identification, rule-based classification for pregnancy-related skin condition matching, and content-based filtering for alternative product recommendations. The system provides comprehensive guidance for pregnant women seeking safe skincare options.

The comprehensive evaluation approach employed in this research demonstrates the system's effectiveness from both technical and user perspectives. Safety evaluation revealed strong performance with 86.25% accuracy in safety classification, achieving 79.59% precision and 97.50% recall. This performance indicates reliable identification of safe products while highlighting areas for improvement in precision. The analysis revealed discrepancies between safety databases and incomplete source data as contributing factors to classification errors, demonstrating the complexity of automated pregnancy safety assessment and the importance of comprehensive ingredient databases.

The pilot user evaluation using the ResQue framework demonstrated the system's effectiveness, with an overall score of 4.65 out of 5. Particularly high scores were observed in Perceived Ease of Use, Perceived Usefulness, and Behavioral Intention, indicating that users found the system both useful and user-friendly. The system received maximum scores for ease of preference revision and achieved high ratings across several key performance areas. Users found the system easy to use for obtaining recommendations, reported that recommended items matched their interests, and indicated the system helped them discover new products.

The interface layout was rated as attractive and adequate, and users found it straightforward to express their preferences. Participants also reported that finding items to purchase became easy with the system's assistance. Importantly, users expressed strong intent to use the system if available and indicated willingness to purchase recommended items when needed. These results confirm that the system successfully meets its primary objectives of providing relevant, accessible, and user-friendly skincare recommendations for pregnant women.

The integrated approach addresses the critical gap in skincare safety information for pregnant women by combining safety assessment with personalized recommendations tailored to specific skin conditions prevalent during pregnancy. The mobile application implementation enhances accessibility and provides a user-friendly interface for navigating product information, analyzing ingredient safety, and comparing alternatives.

This research contributes to the field of recommendation systems by demonstrating the effective integration of classification and filtering methods for sensitive use cases and offers a practical solution to a significant healthcare challenge faced by pregnant women.

CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

ACKNOWLEDGEMENT

Thanks to all the respondents for their valuable participation in this research.

REFERENCES

[1] K. H. Tyler, *Cutaneous Disorders of Pregnancy*. Cham, Switzerland: Springer International Publishing, 2020, doi: 10.1007/978-3-030-49285-4.
[2] S. Musaei, "The Effect of Pregnancy on the Skin," *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, vol. 2, no. 1, pp. 17-23, 2023.
[3] Bolanca et al., "Chloasma – the mask of pregnancy," *Coll. Antropol.*, vol. 32, pp. 139–141, 2008.

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

[4] L. Zhao, X. Fu, and H. Cheng, "Prevention of Melasma during Pregnancy: Risk factors and Photoprotection-Focused Strategies," *Clinical Cosmetic and Investigational Dermatology*, vol. 17, pp. 2301-2310, 2024, doi: 10.2147/ccid.s488663.

[5] Ö. Kutlu et al., "Acne in pregnancy: A prospective multicenter, cross-sectional study of 295 patients in Turkey," *International Journal of Dermatology*, vol. 59, no. 9, pp. 1098–1105, Jun. 2020, doi: 10.1111/ijd.14999.

[6] C. Marie, R. Garlantézec, R. Béranger, and A. Ficheux, "Use of Cosmetic Products in Pregnant and Breastfeeding Women and Young Children: Guidelines for Interventions during the Perinatal Period from the French National College of Midwives," *Journal of Midwifery & Women's Health*, vol. 67, no. S1, 2022, doi: 10.1111/jmwh.13428.

[7] H. W. Lim et al., "Photoprotection in pregnancy: addressing safety concerns and optimizing skin health," *Frontiers in Medicine*, vol. 12, Mar. 2025, doi: 10.3389/fmed.2025.1563369.

[8] A. L. Williams, N. D. Pace, and J. M. DeSesso, "Teratogen update: Topical use and third-generation retinoids," *Birth Defects Research*, vol. 112, no. 15, pp. 1105–1114, Jul. 2020, doi: 10.1002/bdr2.1745.

[9] K. D. Loureiro et al., "Minor malformations characteristic of the retinoic acid embryopathy and other birth outcomes in children of women exposed to topical tretinoin during early pregnancy," *American Journal of Medical Genetics Part A*, vol. 136A, no. 2, pp. 117–121, Jun. 2005, doi: 10.1002/ajmg.a.30744.

[10] P. Brzezinski et al., "A Multicenter Cohort Study Evaluating the Teratogenic Effects of Isotretinoin on Neonates," *Children*, vol. 9, no. 11, p. 1612, 2022, doi: 10.3390/children9111612.

[11] P. A. Bommarito et al., "Prenatal exposure to consumer product chemical mixtures and size for gestational age at delivery," *Environmental Health*, vol. 20, no. 1, Jun. 2021, doi: 10.1186/s12940-021-00724-z.

[12] H. Li et al., "Maternal cosmetics use during pregnancy and risks of adverse outcomes: a prospective cohort study," *Scientific Reports*, vol. 9, no. 1, May 2019, doi: 10.1038/s41598-019-44546-z.

[13] K. Jones, L. M. Wessel, K. Schäfer, and M. Á. Tapia-Laliena, "Use of cosmetics in pregnancy and neurotoxicity: Can it increase the risk of congenital enteric neuropathies?," *Biomolecules*, vol. 14, no. 8, p. 984, 2024, doi: 10.3390/biom14080984.

[14] L. Almeida-Toledano et al., "Effect of prenatal phthalate exposure on fetal development and maternal/neonatal health consequences: A systematic review," *The Science of the Total Environment*, vol. 950, p. 175080, Jul. 2024, doi: 10.1016/j.scitotenv.2024.175080.

[15] J. Lee, H. Yoon, S. Kim, C. Lee, J. Lee, and S. Yoo, "Deep learning-based skin care product recommendation: A focus on cosmetic ingredient analysis and facial skin conditions," *Journal of Cosmetic Dermatology*, vol. 23, no. 6, pp. 2066–2077, Feb. 2024, doi: 10.1111/jocd.16218.

[16] K. S. Y. Putri, N. I. M. A. D. Suarjaya, and N. W. O. Vihikan, "Sistem Rekomendasi Skincare Menggunakan Metode Content Based Filtering dan Collaborative Filtering," *Decode Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 3, pp. 764–774, Aug. 2024, doi: 10.51454/decode.v4i3.601.

[17] K. M. Veena, K. M. Shenoy, and K. B. A. Shenoy, "Performance comparison of machine learning classification algorithms," in *Communications in computer and information science*, 2018, pp. 489–497. doi: 10.1007/978-981-13-1813-9_49.

[18] N. S. Suryawanshi, "Predicting consumer behavior in e-commerce using recommendation systems," *Int. J. Innov. Sci. Res. Technol.*, vol. 4, no. 9, pp. 806-813, 2024, doi: 10.38124/ijisrt/IJISRT19SEP1550.

[19] Y. Afoudi, M. Lazaar, and M. A. Achhab, "Hybrid recommendation system combined content-based filtering and collaborative prediction using artificial neural network," *Simulation Modelling Practice and Theory*, vol. 113, p. 102375, Jul. 2021, doi: 10.1016/j.simpat.2021.102375.

[20] C. Dm, S. Kk, S. Gv, and S. Goravanakolla, "Content based book recommendation system," *International Research Journal Of Modernization In Engineering Technology And Science*, Aug. 2023, doi: 10.56726/irjmets44243.

[21] D. Chopra, P. Arora, and Jagan Institute of Management Studies, Rohini, Sector 5, Delhi, "The Manager's Reading List: A Personalized book recommendation system for management growth," *Journal of Harbin Engineering University*, vol. 44, no. 6, Jun. 2023.
[22] T. Z. L. Kyaw, S. Uttama, and P. Panwong, "Leveraging Ingredient Profiles in Content-Based Skincare Product Recommendation," in *Proc. 8th International Conference on Information Technology (InCIT)*, 2024, pp. 319-324, doi: 10.1109/incit63192.2024.10810620.
[23] S. Tanveer, S. Khatoon, H. U. Begum, and U. Zainab, "Korean Skin Care Recommendation System," *Department of CSE*, vol. 12, no. 2, 2024.
[24] Z. Huang, B. Jin, H. Zhao, Q. Liu, D. Lian, B. Tengfei, and E. Chen, "Personal or General? A Hybrid Strategy with Multi-factors for News Recommendation," *ACM Transactions on Office Information Systems*, vol. 41, no. 2, pp. 1-29, 2022, doi: 10.1145/3555373.
[25] P. Pu, L. Chen, and R. Hu, "A user-centric evaluation framework for recommender systems," in *Proceedings of the Fifth ACM Conference on Recommender Systems*, pp. 157–164, Oct. 2011, doi: 10.1145/2043932.2043962.
[26] G. Karystianis, K. Thayer, M. Wolfe, and G. Tsafnat, "Evaluation of a rule-based method for epidemiological document classification towards the automation of systematic reviews," *Journal of Biomedical Informatics*, vol. 70, pp. 27–34, Apr. 2017, doi: 10.1016/j.jbi.2017.04.004.
[27] Y. Deldjoo, M. Elahi, M. Quadrana, and P. Cremonesi, "Using visual features based on MPEG-7 and deep learning for movie recommendation," *International Journal of Multimedia Information Retrieval*, vol. 7, no. 4, pp. 207–219, Jun. 2018, doi: 10.1007/s13735-018-0155-1.
[28] G. Lee, X. Jiang, and N. Parde, "A Content-based Skincare Product Recommendation System," *2023 22nd International Conference on Machine Learning and Applications (ICMLA)*, pp. 2039–2043, Dec. 2023, doi: 10.1109/icmla58977.2023.00308.
[29] M. Khder, "Web scraping or web crawling: state of art, techniques, approaches and application," *International Journal of Advances in Soft Computing and Its Applications*, vol. 13, no. 3, pp. 145–168, Nov. 2021, doi: 10.15849/ijasca.211128.11.
[30] Cristina Arhiliuc and Raf Guns, "Content-based classification of research articles: comparing keyword extraction, BERT, and random forest classifiers", in *Proceedings of ISSI 2023 – the 19th International Conference of the International Society for Scientometrics and Informetrics*, Jul. 2023, vol. 1, pp. 43–63. doi: 10.5281/zenodo.8305874.
[31] A. K. Varshney and V. Torra, "Literature review of the recent trends and applications in various fuzzy Rule-Based systems," *International Journal of Fuzzy Systems*, vol. 25, no. 6, pp. 2163–2186, May 2023, doi: 10.1007/s40815-023-01534-w.
[32] J. Friedl, Mastering regular expressions. "O'Reilly Media, Inc.," 2006.
[33] H. Liu, X. Chen, and X. Liu, "A study of the application of weight distributing method combining sentiment dictionary and TF-IDF for text sentiment analysis," *IEEE Access*, vol. 10, pp. 32280–32289, Jan. 2022, doi: 10.1109/ACCESS.2022.3160172.
[34] H. Kim, J. Baek, and K. Chung, "Optimization of Associative Knowledge Graph using TF-IDF based Ranking Score," *Applied Sciences*, vol. 10, no. 13, p. 4590, 2020, doi: 10.3390/app10134590.
[35] H. J. Armadhani, P. Permana, and A. T. Wibowo, "Movie Recommendation System Based on Synopsis Using Content-Based Filtering with TF-IDF and Cosine Similarity," *Intl. Journal on ICT*, vol. 9, no. 2, pp. 1-14, 2023, doi: 10.21108/ijoict.v9i2.747.
[36] G. Yunanda, D. Nurjanah, and S. Meliana, "Recommendation System from Microsoft News Data using TF-IDF and Cosine Similarity Methods," *Building of Informatics Technology and Science (BITS)*, vol. 4, no. 1, 2022, doi: 10.47065/bits.v4i1.1670.
[37] Christine, "15 minute beauty fanatic," *15 Minute Beauty Fanatic*. https://www.15minutebeauty.com/ (accessed Jun. 18, 2025).
[38] S. Sathyanarayanan, "Confusion Matrix-Based Performance Evaluation Metrics," *African Journal of Biomedical Research*, pp. 4023–4031, Nov. 2024, doi: 10.53555/ajbr.v27i4s.4345.
[39] B. Gaye and A. Wulamu, "Sentiment analysis of text classification algorithms using Confusion Matrix," in *Communications in computer and information science*, 2019, pp. 231–241. doi: 10.1007/978-981-15-1922-2_16.

Vol. 6, No. 3, Juni 2025, Page. 1643-1668
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4646

[40] D. R. Kim et al., "An Open Label Pilot Study of Transcranial Magnetic Stimulation for Pregnant Women with Major Depressive Disorder," *Journal of Women S Health*, vol. 20, no. 2, pp. 255–261, Feb. 2011, doi: 10.1089/jwh.2010.2353.

[41] C. Knight-Agarwal, D. L. Davis, L. Williams, R. Davey, R. Cox, and A. Clarke, "Development and pilot testing of the Eating4Two mobile phone app to monitor gestational weight gain," *JMIR Mhealth and Uhealth*, vol. 3, no. 2, p. e44, Jun. 2015, doi: 10.2196/mhealth.4071.

[42] S. Q. Wang and H. W. Lim, "Highlights and implications of the 2019 proposed rule on sunscreens by the US Food and Drug Administration," *Journal of the American Academy of Dermatology*, vol. 81, no. 2, pp. 650–651, Apr. 2019, doi: 10.1016/j.jaad.2019.04.007.

[43] I. B. Putra, N. K. Jusuf, and N. K. Dewi, "Skin changes and safety profile of topical products during pregnancy," *J. Clin. Aesthet. Dermatol.*, vol. 15, no. 2, pp. 49–57, 2022.

[44] V. M. Patel, R. A. Schwartz, and W. C. Lambert, "Safety of topical dermatologic medications in pregnancy, " *J. Drugs Dermatol.*, vol. 15, no. 7, pp. 830–834, 2016.

[45] S. Ly, K. Kamal, P. Manjaly, J. S. Barbieri, and A. Mostaghimi, "Treatment of acne vulgaris during pregnancy and lactation: A Narrative review," *Dermatology and Therapy*, vol. 13, no. 1, pp. 115-130, 2022, doi: 10.1007/s13555-022-00854-3.

[46] L. L. Bio and J. J. Cies, "Lack of pregnancy warnings on over-the-counter dermatologic products containing potentially harmful hydroquinone," *Journal of Perinatology*, vol. 37, no. 7, pp. 778-781, 2017, doi: 10.1038/jp.2017.35.

[47] J. Zhao, Y. Wang, L. Jiang, and Y. Mu, "The application of skin care product in acne treatment," *Dermatologic Therapy*, vol. 33, no. 6, 2020, doi: 10.1111/dth.14287.

[48] A. K. Mohiuddin, "Skin lightening & management of hyperpigmentation," *Pharmaceutical Sciences & Analytical Research Journal*, vol. 2, no. 2, 2019, doi: 10.28933/ajodrr-2019-04-2001.

[49] M. Rendon, M. Berneburg, I. Arellano, and M. Picardo, "Treatment of melasma," *Journal of the American Academy of Dermatology*, vol. 54, no. 5, pp. S272-S281, 2006, doi: 10.1016/j.jaad.2005.12.039.

[50] K. Korgavkar and F. Wang, "Stretch marks during pregnancy: a review of topical prevention," *British Journal of Dermatology*, vol. 172, no. 3, pp. 606-615, 2014, doi: 10.1111/bjd.13426.

[51] J. Á. G. Hernández, D. M. González, M. P. Castillo, and T. F. Falcón, "Use of a specific anti-stretch mark cream for preventing or reducing the severity of striae gravidarum. Randomized, double-blind, controlled trial," *International Journal of Cosmetic Science*, vol. 35, no. 3, pp. 233-237, 2012, doi: 10.1111/ics.12029.

[52] N. S. Wahyuni and N. F. P. Putra, "Rancangan Aplikasi Rekomendasi Produk Skincare untuk Ibu Hamil Menggunakan Metode Prototype," *Repeater*, vol. 2, no. 4, pp. 136-148, 2024, doi: 10.62951/repeater.v2i4.209.