Volume XI Issue 3 Year 2026 | Page 817-828 | ISSN: 2527-9866
Received: 21-06-2026 | *Revised: 20-07-2026* | *Accepted: 02-08-2026*

Implementation Of The Naïve Bayes Method For Skincare Product Recommendations According To Skin Type At Dermakila Clinic

Afilda Maharani Asmaraloka[1], Supriyono[2], Zainur Romadhon[3]
[1,2,3] Information System, Muria Kudus University, Kudus, Indonesia
e-mail: afilda088@gmail.com[1], supriyono.si@umk.ac.id[2], zainur.romadhon@umk.ac.id[3]
*Correspondence: afilda088@gmail.com

Abstract: *Advances in information technology have encouraged the implementation of recommendation systems in various fields, including beauty and skincare. Selecting skincare products that are not suitable for an individual's skin type and condition may reduce treatment effectiveness and potentially lead to skin problems. Dermakila Clinic offers a wide range of skincare products with diverse characteristics, creating a need for a system that can assist users in selecting products that best suit their skin needs. This study aims to implement the Naïve Bayes method in developing a skincare product recommendation system based on user characteristics, including age range, gender, skin type, and skin concerns. The research applies a Data Mining approach using the Naïve Bayes classification algorithm. The dataset consists of 960 skincare product records that have undergone preprocessing and data transformation stages. The system was developed as a web-based application to provide users with fast and accurate product recommendations. The experimental results demonstrate that the Naïve Bayes method achieved an accuracy of 88%, with a precision of 89%, a recall of 88%, and an F1-score of 88%. These findings indicate that the Naïve Bayes method is effective for implementing a skincare product recommendation system at Dermakila Clinic.*

Keywords: *Data Mining, Naïve Bayes, Recommendation System, Skincare Products, Skin Type.*

1. Introduction

The rapid development of information technology has driven the application of intelligent systems across various fields, including the beauty and skincare sector. One widely used implementation is the recommendation system, which functions to assist users in determining choices based on their characteristics and needs. In skincare, selecting products that do not match one's skin type and condition can reduce treatment effectiveness and potentially cause various skin problems, such as acne, irritation, dehydration, and excessive oil production. Therefore, a system is needed that can provide accurate and objective skincare product recommendations to help users obtain products suitable for their skin condition.

Dermakila Clinic is a beauty clinic that offers a variety of skincare products with diverse characteristics and benefits. The large number of available product options often makes it difficult for users to determine the products that best suit their skin needs. In addition, the product selection process still relies on direct consultation with experts, which requires considerable time and resources. This condition indicates the need for a recommendation system that can help users obtain product recommendations quickly, accurately, and efficiently. Furthermore, the implementation of web-based information systems has been shown to improve service efficiency, data management, and user accessibility across various application domains [1]

Various approaches have been developed to build recommendation systems, including collaborative filtering, which generates recommendations by measuring the similarity between users or items. This approach has been widely applied in recommendation systems and has demonstrated promising performance in generating personalized recommendations [2]. In the development of recommendation systems, one widely used method is Naïve Bayes. According to [3], the Naïve Bayes method is a classification algorithm that applies Bayes' Theorem to calculate the probability of a class based on the attributes possessed by the data. This method is widely used because of its simple and efficient computation process and its ability to produce good accuracy for various classification problems. Recent studies have also demonstrated that the Naïve Bayes algorithm remains a competitive classification method, achieving effective performance in comparison with other machine learning algorithms such as Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) in text classification tasks [4]. This method is also widely applied in the field of data mining owing to its simple and efficient computation process and its ability to yield good accuracy across various classification problems. These capabilities make Naïve Bayes a relevant method for application in skincare product recommendation systems based on user characteristics.

Several previous studies have shown that the Naïve Bayes method performs well in classification and recommendation system problems. Applied the Naïve Bayes algorithm to a skincare product recommendation system based on skin type and obtained good classification results [5]. However, this study differs from previous research in several important aspects. The previous study developed a skincare recommendation system using data collected through observations, interviews, and customer surveys, combined with ingredient recommendations based on skin type. In contrast, the proposed study utilizes a dataset of 960 skincare recommendation records obtained from Dermakila Clinic. Furthermore, this study employs a more comprehensive set of predictive attributes, including age range, gender, skin type, and eleven skin complaint indicators, namely acne, blackheads, enlarged pores, excess oil, dull skin, dark spots, redness, irritation, dry skin, dehydration, and small bumps. The recommendation results are classified into twelve skincare product classes and implemented in a web-based recommendation system. The performance of the proposed model is evaluated using a confusion matrix with accuracy, precision, recall, and F1-score metrics. Therefore, this study provides a more comprehensive and personalized skincare product recommendation model by considering multiple user characteristics and skin conditions.

In addition, [6] used the Naïve Bayes method to identify facial skin types for skincare product selection and demonstrated that the algorithm can effectively support the decision-making process. In another study, [7] applied the Naïve Bayes method to classify sentiment in skincare product reviews and obtained effective results in grouping review data. Similar results were also shown by [8], who applied Naïve Bayes to sentiment analysis of beauty products based on user reviews. Beyond the skincare field, [9] demonstrated that the Naïve Bayes method is capable of generating product recommendations based on user characteristics with a good level of accuracy. These findings indicate that the Naïve Bayes algorithm has strong capability in processing attribute-based data to produce accurate classifications and recommendations .

Although various studies have applied the Naïve Bayes method to recommendation systems and skincare data analysis, most of these studies have focused on sentiment analysis, skin type identification, or product recommendation for different research objects. Research specifically addressing skincare product recommendations based on user characteristics—including age range, gender, skin type, and skin complaints—at Dermakila Clinic remains limited. Recent studies have also demonstrated that web-based recommendation systems continue to evolve by incorporating various intelligent recommendation techniques to generate more personalized recommendations based on user characteristics. This indicates that recommendation systems remain an active research area with broad applications across different domains. Therefore, this study was conducted to fill this research gap through the development of a web-based skincare product recommendation system using the Naïve Bayes algorithm.

This study aims to apply the Naïve Bayes method in building a skincare product recommendation system for Dermakila Clinic and to evaluate the performance of the resulting model. The results of this study are expected to help users select skincare products that suit their skin characteristics and to support the improvement of consultation service effectiveness at Dermakila Clinic.

2.Methods

The research method used in this study consists of several interrelated stages to produce a skincare product recommendation system based on the Naïve Bayes method. The research stages begin with the collection of a dataset from Dermakila Clinic containing information on user characteristics and skincare product recommendations. The obtained data then undergoes a preprocessing stage to ensure data quality and consistency before being used in the classification process [10].

After the preprocessing stage is completed, the dataset is divided into training data and testing data using a data splitting method. The training data is used to build the Naïve Bayes classification model, while the testing data is used to measure the model's ability to provide appropriate skincare product recommendations. Subsequently, the classification results are evaluated using a confusion matrix to determine the model's performance level based on the values of accuracy, precision, recall, and F1-Score. The overall sequence of research stages is shown in Figure 1.

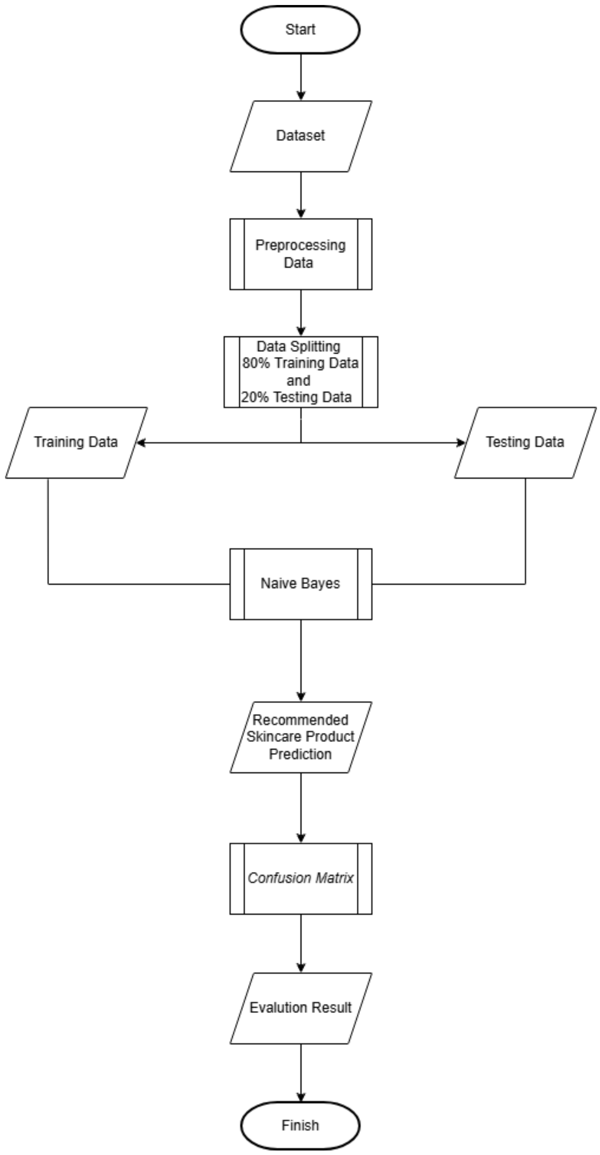

Figure 1. Research Methodology Flow

Dataset

The data preparation stage begins with the collection of a dataset obtained from Dermakila Clinic. The dataset used consists of 960 data entries and contains information on user characteristics as well as the recommended skincare products. User characteristics consist of attributes for age range, gender, and skin type, along with ten skin condition attributes represented in binary form (0 and 1), namely acne, blackheads, enlarged pores, excess oil, dull skin, dark spots, redness, irritation, dry skin, dehydration, and small bumps (fungal acne). Meanwhile, the skincare product attribute is used as the class (target) in the classification process [11]. The description of the attributes used in this study is shown in Table 1.

Table 1. Description Attribute Dataset Dermakila Clinic

Attribute	Data Type	Description
Age range	Categorical	13–15, 16–20, 21–25, etc.
Gender	Categorical	Male, Female
Skin type	Categorical	Normal, Oily, Dry, Sensitive, Combination
Skin complaints	Binary	1 = Yes, 0 = No
Product	Label	Acne Series, Brightening Series, etc.

The collected dataset was subsequently prepared for the preprocessing stage before being used to build the Naïve Bayes classification model [12]. Due to space limitations in this article, Table 2 only displays a portion of the data as a representative sample of the total 960 data used.

Data Preprocessing

The data preprocessing stage was carried out to improve data quality before it was used in the Naïve Bayes model training process [13]. This process includes data cleaning, data transformation, label encoding, and dataset splitting. Data cleaning was performed by checking data completeness and removing duplicate data or data that did not meet the research requirements. The next stage was the transformation of categorical data using the label encoding technique. The attributes of age range, gender, skin type, and skincare product were converted into numerical form so that they could be processed by the Naïve Bayes algorithm. After all data had been successfully transformed, the dataset was divided into training data and testing data using the holdout method with an 80:20 ratio. A total of 80% of the data was used for the model training process, while 20% of the data was used for the testing and performance evaluation process. The preprocessing results produced a dataset ready for use in the classification and skincare product recommendation process.

Data Splitting

In the data splitting stage, the dataset was divided into two parts, namely training data and testing data, with an 80:20 ratio. This ratio was chosen because it is a commonly used approach in machine learning modeling and is able to provide a balance between the model training and testing processes. The selection of the composition of training data and testing data is a factor that can affect the performance of the classification model, so the determination of the data splitting ratio needs to be carried out appropriately [14]. A total of 80% of the data was used as training data so that the Naïve Bayes model could learn the pattern of relationships between user characteristics and the recommended skincare products. Meanwhile, the remaining 20% of the data was used as testing data to evaluate the model's performance in providing skincare product recommendations. Of the total 960 data used, 768 data were allocated as training data and 192 data as testing data. The testing data was used to provide an objective assessment of the model's ability to generate skincare product recommendations based on user characteristics that had not been previously learned.

Naïve Bayes

At this stage, the model was built using training data that had undergone the preprocessing process. The Naïve Bayes algorithm was used to perform classification and generate skincare product recommendations based on user characteristics, including age range, gender, skin type, and skin complaints. The Naïve Bayes method is a classification algorithm based on Bayes' Theorem, which assumes that each attribute is independent of the other attributes.

This algorithm works by calculating the probability of a class occurring based on the attributes possessed by the input data. The class with the highest probability value is selected as the skincare product recommendation result. The Naïve Bayes method was chosen because it has a simple computation process, is capable of handling categorical data well, and has fairly effective performance in various classification problems [15]. Probability calculations are performed based on the training data available in the system, and Bayesian prediction is based on Bayes' theorem with the following general formula [16]. The formula for the Naïve Bayes method used in this study is:

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \tag{1}$$

Confusion Matrix

Model evaluation was carried out using a confusion matrix to measure the performance of the Naïve Bayes algorithm in generating skincare product recommendations. The measurement was carried out using the metrics of accuracy, precision, recall, and F1-Score. Accuracy was used to measure the overall accuracy of the recommendations, while precision and recall were used to evaluate the model's ability to recommend skincare products that match user characteristics. In addition, F1-Score was used to measure the balance between precision and recall so as to provide a more comprehensive picture of the model's performance. The evaluation formulas used to calculate accuracy, precision, recall, and F1-Score are shown in the following equations.

a. Accuracy
   Accuracy is used to measure the model's performance based on the proportion of data correctly classified relative to the total data tested. In multiclass classification cases, the accuracy calculation is expressed through equation (2) [17].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

b. Precision
   Precision measures the accuracy of the positive predictions generated by the classification model. The precision value can be calculated using equation (3) [18]

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

c. Recall
   Recall measures the model's success rate in identifying data belonging to the positive class. The recall value can be calculated using equation (4) [19]

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

d. F1-Score
   F1-Score is an evaluation metric that combines precision and recall through their harmonic mean to assess model performance in a balanced manner. The F1-Score value can be calculated using equation (5).

$$F1\text{-}Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \tag{5}$$

3. Results and Discussion

Dataset

The dataset used in this study consists of 960 skincare recommendation records obtained from Dermakila Clinic. The data were collected from historical customer consultation records containing user characteristics, including age range, gender, skin type, skin concerns, and the corresponding skincare product recommendations, which served as the target classes in the Naïve Bayes classification model. Before being used for model development, all records were anonymized to protect customer privacy and underwent data selection and preprocessing to ensure completeness and consistency. The recommendation labels were derived from skincare products recommended during routine consultation services provided by experienced beauticians and therapists at Dermakila Clinic. These consultation services are conducted following the clinic's standard skincare procedures under the supervision of a dermatologist. The skincare products used as the target classes are presented in Table 3.

Table 2. Target Skincare Product Classes

Product Skincare
Acne Series
Serum Exfoliating
Hydroprotect Barrier Serum
Probiotic Skin Defence Serum
Barrier Moist Cream
Brightening Night Cream
Serum Vitamin C
RetiWhite Night Cream
Derma Acne Facial Wash
Derma Bright Facial Wash
Derma Liquid Cleanser
Cover Up Sunscreen

To prevent the classification model from being biased toward a particular product class, a balanced dataset was constructed by selecting 80 records for each of the twelve skincare product classes. This sampling strategy ensured equal class representation during model training and evaluation, thereby improving the fairness and reliability of the classification results. In this study, the recommendation problem was formulated as a single-label classification task. Each consultation record was associated with one primary skincare product recommended during the consultation process at Dermakila Clinic. Therefore, each data instance contained only one target class, and the Categorical Naïve Bayes model was trained to predict a single skincare product based on the user's characteristics and skin concerns. The recommendation provided by the system is intended to represent the primary product recommendation rather than a complete skincare routine or a combination of multiple products.

Preprocessing Data

The preprocessing stage was carried out to prepare the dataset before it was used in the training and testing of the Naïve Bayes model. The research dataset consisted of 960 data entries containing the attributes of age range, gender, skin type, various skin complaints such as acne, blackheads, enlarged pores, excess oil, redness, irritation, dark spots, dull skin, dry skin, dehydration, and small bumps (fungal acne), as well as skincare products as the target class. At this stage, data quality checks were performed to ensure data completeness and the appropriateness of the format of each attribute. The categorical attributes, namely age range, gender, skin type, and skincare product, were transformed into numerical form using the Label Encoding technique so that they could be processed by the Categorical Naïve Bayes algorithm. The encoded values were used solely as categorical identifiers required by the classifier and were not interpreted as continuous numerical or ordinal values. Furthermore, duplicate records were identified and removed during the preprocessing stage before the dataset was divided into training and testing subsets. This procedure was performed to prevent duplicate-data leakage, ensuring that identical records did not appear simultaneously in both subsets. The preprocessing results showed that all data had been successfully transformed and were ready to be used in the data splitting stage and the model training process.

Data Splitting

In the data splitting stage, the preprocessed dataset was divided into two subsets, namely training data and testing data, using the train_test_split function with an 80:20 ratio. A total of 768 records (80%) were allocated to the training set, while the remaining 192 records (20%) were assigned to the testing set. The parameter random_state = 42 was applied to ensure the reproducibility of the experimental results, while stratified sampling (stratify = y) was employed to preserve the class distribution in both subsets. Since a balanced dataset was used, each of the twelve skincare product classes consisted of 64 training records and 16 testing records after the splitting process. The training dataset was used to build the Categorical Naïve Bayes classification model, whereas the testing dataset was used to evaluate the model's performance on previously unseen data [20].

Naïve Bayes Classification

In the classification stage, the testing data was processed using the Naïve Bayes model that had been built from the training data. The model performed predictions on each data entry based on the attributes used in the study, namely age range, gender, skin type, and the various skin conditions experienced by the user. Based on the probability calculation process, the system determines the class with the highest posterior probability value as the prediction result [21]. This study employed the Categorical Naïve Bayes (CategoricalNB) algorithm implemented in the Scikit-learn library because all predictor variables were categorical after the preprocessing stage. The classifier used the default Laplace smoothing parameter (alpha = 1.0) to prevent zero probability estimates for unseen categorical feature values, thereby improving the robustness of the classification process.

The classification results show that the Categorical Naïve Bayes model was able to identify the relationships between user characteristics and appropriate skincare products. Each testing instance was classified into one of the twelve skincare product classes based on the highest posterior probability. The prediction results were subsequently used as the basis for evaluating the model's performance using the metrics of accuracy, precision, recall, and F1-Score [22]. The classification performance was evaluated using the classification_report function from the Scikit-learn library. Precision, recall, and F1-score were calculated for each skincare product class, together with the macro average and weighted average values to provide a comprehensive evaluation of the proposed model.

Table 3. Classification Performance for Each Skincare Product Class

Skincare Product	Precision	Recall	F1-score	Support
Acne Series	0.93	0.81	0.87	16
Serum Exfoliating	1.00	1.00	1.00	16
Hydroprotect Barrier Serum	0.57	0.50	0.53	16
Probiotic Skin Defence Serum	1.00	0.94	0.97	16
Barrier Moist Cream	1.00	0.94	0.97	16
Brightening Night Cream	0.93	0.88	0.90	16
Serum Vitamin C	1.00	0.94	0.97	16
RetiWhite Night Cream	1.00	0.94	0.97	16

Derma Acne Facial Wash	0.89	1.00	0.94	16
Derma Bright Facial Wash	1.00	0.94	0.97	16
Derma Liquid Cleanser	0.84	1.00	0.91	16
Cover Up Sunscreen	0.52	0.69	0.59	16
Accuracy			0.88	192
Macro avg	0.89	0.88	0.88	192
Weighted avg	0.89	0.88	0.88	192

Table 3 presents the classification performance for each skincare product class. Most skincare product classes achieved high precision, recall, and F1-score values, indicating that the proposed Categorical Naïve Bayes model was able to classify skincare recommendations accurately. Several product classes, such as Hydroprotect Barrier Serum and Cover Up Sunscreen, obtained relatively lower F1-score values of 0.53 and 0.59, respectively. This indicates that these classes share similar user characteristics and skin concern attributes with other product classes, making them more difficult to distinguish. Overall, the proposed model achieved a macro-average precision of 0.89, recall of 0.88, and F1-score of 0.88. Since the dataset was balanced, the weighted-average values were identical to the macro-average values, indicating consistent classification performance across all product classes.

Confusion Matrix

Model performance evaluation was carried out using a confusion matrix to determine the ability of the Naïve Bayes algorithm to classify data into the appropriate skincare product classes. The confusion matrix was used to compare the prediction results generated by the model with the actual classes in the testing data. Through this method, it is possible to determine the number of data correctly classified as well as the data that experienced classification errors for each skincare product class.

Based on the test results on 192 testing data, most of the data were correctly classified. This is indicated by the dominance of values along the main diagonal of the matrix, which represents the number of predictions that match the actual class. The confusion matrix was interpreted by examining the values of true positive, true negative, false positive, and false negative so that the performance of the classification model could be determined [23], [24]. Although there were still several classification errors in some skincare product classes, the number of correct predictions was more dominant compared to incorrect predictions. The confusion matrix results indicate that the Naïve Bayes model has a good ability to recognize the pattern of relationships between user characteristics and the recommended skincare products.

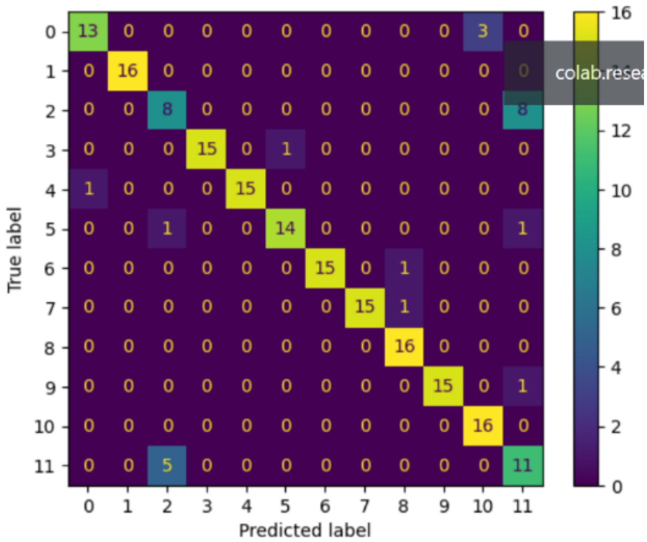

Figure 2. Confusion Matrix of Naïve Bayes Classification Results

Discussion

The proposed recommendation system was designed to predict a single primary skincare product based on the consultation records available in the dataset. Therefore, the recommendation generated by the system represents the primary product recommended for a user's skin condition rather than a complete skincare routine. The current model does not explicitly consider additional clinical factors such as ingredient compatibility, allergy history, contraindications, or combinations of complementary skincare products because these variables were not available in the dataset used in this study. Future research may incorporate these clinical factors and adopt a multi-label recommendation approach to provide more comprehensive and personalized skincare recommendations.

4. Conclusions

The results of this study demonstrate that the proposed Categorical Naïve Bayes model was able to classify skincare product recommendations based on user characteristics, achieving an accuracy of 88%, with a precision of 89%, recall of 88%, and F1-score of 88%. These results indicate that the proposed approach shows promising performance for supporting skincare product recommendations within the scope of the dataset used in this study.

However, this study has several limitations. First, the dataset was collected from consultation records at a single skincare clinic, which may limit the generalizability of the proposed model to other clinical settings. Second, the recommendation labels were derived from routine consultation records conducted under dermatologist supervision and were not independently validated by dermatology experts specifically for this study. Third, the model evaluation was performed using the holdout validation method with an 80:20 data split. Finally, the proposed system focuses on predicting the primary skincare product recommendation and does not evaluate clinical safety aspects such as ingredient compatibility, allergy history, contraindications, or multi-product skincare combinations. Future research should incorporate larger multi-clinic datasets, additional clinical variables, and more robust validation methods to improve the reliability and applicability of the recommendation system.

References

[1] R. Setiawan, Z. Romadhon, and A. R. Hakim, "Implementation of the Prototype Method in Developing a Questionnaire Information System for Integrated Survey Services," *SISTEMASI*, vol. 14, no. 2, p. 835, Mar. 2025, doi: 10.32520/stmsi.v14i2.5069.

[2] Z. Romadhon, E. Sediyono, and C. E. Widodo, "Various Implementation of Collaborative Filtering-Based Approach on Recommendation Systems using Similarity," *Kinet. Game Technol. Inf. Syst. Comput. Netw. Comput. Electron. Control*, pp. 179–186, Jul. 2020, doi: 10.22219/kinetik.v5i3.1062.

[3] K. Rahmat S, I. Kanedi, and J. Fredricka, "Application of the Naive Bayes Method in the Classification of Employee Status at PT. Prosperous New Light," *J. Media Comput. Sci.*, vol. 2, no. 1, Jan. 2023, doi: 10.37676/jmcs.v2i1.3395.

[4] N. Aditiya and P. Setiaji, "Analisis Sentimen Kepuasan Masyarakat terhadap Aplikasi 'INFO BMKG' menggunakan Naive Bayes, SVM, dan KNN," vol. 14.

[5] F. Ratnawati, A. Tedyyana, and J. Custer, "LightGBM for Liver Disease Detection with Hybrid Hyperparameter Optimization", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 10, no. 4, pp. 1025 - 1033, Aug. 2026..

[6] H. Hasnidar and P. T. Prasetyaningrum, "Sistem Pakar Pengidentifikasian Jenis Kulis Wajah Dalam Pemilihan Msglow Series Menggunakan Naïve Bayes," *J. Sains Dan Teknol. JSIT*, vol. 2, no. 2, pp. 88–103, Aug. 2023, doi: 10.47233/jsit.v2i3.204.

[7] K. Nurfebia and S. Sriani, "Sentiment Analysis of Skincare Products Using the Naive Bayes Method," *J. Inf. Syst. Inform.*, vol. 6, no. 3, pp. 1663–1676, Sep. 2024, doi: 10.51519/journalisi.v6i3.817.

[8] Pramesti Melinea Berlianti and Erwin Yudi Hidayat, "Implementasi Naïve Bayes Classifier untuk Sentimen Produk Kecantikan Berdasarkan Ulasan Female Daily," *Indones. J. Comput. Sci.*, vol. 13, no. 6, Dec. 2024, doi: 10.33022/ijcs.v13i6.4499.

[9] C. Therence and J. Tjen, "Analysis of Cosmetic Product Opinions on E-Commerce Based on Naïve Bayes Classifier," in *Proceedings of the International Conference on Digital Business Innovation and Technology (ICONBIT)*, 2024.

[10] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, Sep. 2021, doi: 10.30645/j-sakti.v5i2.369.

[11] P. Simanjuntak, H. Pangaribuan, and M. T. Syastra, "Data Mining Rekomendasi Pemakaian Skincare," *MEANS Media Inf. Anal. Dan Sist.*, pp. 80–83, Jul. 2021, doi: 10.54367/means.v6i1.1224.

[12] I. Amillina and A. Qoiriah, "Penerapan Algoritma Naïve Bayes dalam Klasifikasi Tingkat Kepuasan Siswa terhadap Pembelajaran Daring," *J. Ilm. Teknol. Inf. Dan Robot.*, vol. 3, no. 2, Dec. 2021, doi: 10.33005/jifti.v3i2.59.

[13] E. Alshdaifat, D. Alshdaifat, A. Alsarhan, F. Hussein, and S. M. F. S. El-Salhi, "The Effect of Preprocessing Techniques, Applied to Numeric Features, on Classification Algorithms' Performance," *Data*, vol. 6, no. 2, p. 11, Jan. 2021, doi: 10.3390/data6020011.

[14] R. Oktafiani, A. Hermawan, and D. Avianto, "Pengaruh Komposisi Split data Terhadap Performa Klasifikasi Penyakit Kanker Payudara Menggunakan Algoritma Machine Learning," *J. Sains Dan Inform.*, pp. 19–28, Jun. 2023, doi: 10.34128/jsi.v9i1.622.

[15] Muhammad Hidayatullah, Hanif Rahmat, and Retno Mayapada, “Sistem Pakar Diagnosis Penyakit Kulit Pada Manusia Dengan Metode Naïve Bayes Menggunakan Shiny,” *J. Tek. Mesin Elektro Dan Ilmu Komput.*, vol. 3, no. 3, pp. 47–59, Nov. 2023, doi: 10.55606/teknik.v3i3.2444.

[16] Y. Yulindawati, S. Lailiyah, A. Yusnita, and A. Hafifah, “REKOMENDASI PEMILIHAN JUDUL TUGAS AKHIR MENGGUNAKAN METODE NAÏVE BAYES,” *J. Inf. Syst. Manag. JOISM*, vol. 5, no. 2, pp. 171–175, Jan. 2024, doi: 10.24076/joism.2024v5i2.1383.

[17] G. Naidu, T. Zuva, and E. M. Sibanda, “A Review of Evaluation Metrics in Machine Learning Algorithms,” in *Artificial Intelligence Application in Networks and Systems*, R. Silhavy and P. Silhavy, Eds., Cham: Springer International Publishing, 2023, pp. 15–25.

[18] O. Peretz, M. Koren, and O. Koren, “Naive Bayes classifier – An ensemble procedure for recall and precision enrichment,” *Eng. Appl. Artif. Intell.*, vol. 136, p. 108972, Oct. 2024, doi: 10.1016/j.engappai.2024.108972.

[19] C. Sitaula and T. B. Shahi, “Monkeypox Virus Detection Using Pre-trained Deep Learning-based Approaches,” *J. Med. Syst.*, vol. 46, no. 11, p. 78, Oct. 2022, doi: 10.1007/s10916-022-01868-2.

[20] M. A. Prasetyo, A. K. Zyen, and R. H. Kusumodestoni, “OPTIMASI ALGORITMA NAIVE BAYES BERBASIS KERNEL UNTUK KLASIFIKASI PENYAKIT HATI,” *J. Inform. Teknol. Dan Sains Jinteks*, vol. 6, no. 3, pp. 748–756, Sep. 2024, doi: 10.51401/jinteks.v6i3.4783.

[21] A. Afandi, H. Bedi Agtriadi, L. Luqman, and M. Susanti, “Advanced Credit Scoring with Naive Bayes Algorithm: Improving Accuracy and Reliability in Financial Risk Assessment,” *J. E-Komtek Elektro-Komput.-Tek.*, vol. 8, no. 2, pp. 399–409, Dec. 2024, doi: 10.37339/e-komtek.v8i2.2160.

[22] H. La Ariandi, A. Setyanto, and S. Sudarmawan, “Classification Of Malaria Types Using Naïve Bayes Classification,” *J. Indones. Sos. Teknol.*, vol. 5, no. 5, pp. 2311–2327, May 2024, doi: 10.59141/jist.v5i5.1088.

[23] K. Kenny, D. Arisandi, and T. Sutrisno, “Evaluasi Penilaian Kinerja Karyawan dengan Metode Naïve Bayes,” *Comput. J. Comput. Sci. Inf. Syst.*, vol. 8, no. 1, pp. 110–118, Apr. 2024, doi: 10.24912/computatio.v8i1.20287.

[24] I. G. A. P. Mahendra, I. M. A. Wirawan, and I. G. A. Gunadi, “Enhancement performance of the Naïve Bayes method using AdaBoost for classification of diabetes mellitus dataset type II,” International Journal of Advances in Applied Sciences, vol. 13, no. 3, pp. 733–742, Sep. 2024, doi: 10.11591/ijaas.v13.i3.pp733-742.