TEM Journal. Volume 12, Issue 1, pages 96-103, ISSN 2217-8309, DOI: 10.18421/TEM121-13, February 2023.

Facial Skin Type Prediction Based on Baumann Skin Type Solutions Theory Using Machine Learning

Rosayanti Efata, Widya Indriani Loka, Natasha Wijaya, Derwin Suhartono

Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, 11480, Indonesia

Abstract **– The lack of knowledge of different facial skin types is still a frequent problem in Indonesia. The purpose of this research is to build a facial skin type prediction system using machine learning to classify facial skin types based on Baumann Skin Type Solutionswhich provides information on different skin types and suitable skincare ingredients. The dataset is collected manually by distributing a questionnaire among Indonesian citizens. The prediction models are built using three machine learning methods namely SVM, XGBoost, and 1D-CNN, and compared using 5-fold stratified cross-validation. XGBoostachieved the best performance on facial skin type prediction and optimized through hyperparameter tuning using Bayesian Optimization with a result of 93.5% averaged F1-score.**

Keywords **– Skin type classification, Baumann Skin Type Solutions, XGBoost, SVM, 1D-CNN, Bayesian Optimization.**

DOI: 10.18421/TEM121-13
https://doi.org/10.18421/TEM121-13

Corresponding author: Rosayanti Efata,
Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, 11480, Indonesia.
Email: rosaefataa@gmail.com

Received: 22 September 2022.
Revised: 30 November 2022.
Accepted: 12 December 2022.
Published: 27 February 2023.

1. Introduction

A survey from a beauty clinic in Indonesia [1]showed that out of 6,460 Indonesian female respondents, 81.7% prioritized the use of skincare products. To choose and determine the skincare products that are used every day, there are things thatmust be understood first. The ingredients contained in the product, the concerns that those products are targeting, and the type of facial skin which are some of the more important things to consider in the selection of products to be used[2]. However, there are still many who do not have an understanding about diverse types of facial skin. As a result, the skincare products that are used may not be suitable for their facial skin type and can damage the main protective layer of the facial skin barrier[1], [3]. Therefore, a prediction system for facial skin types was built.

The theory of Baumann Skin Type Solutionswas chosen as the basis for the development of the facial skin type prediction system and the classification method will be selected through several process of experiments. The machine learning classification methods chosen that will be applied to the prediction system and compared are XGBoost, the development of the Gradient Tree Boosting algorithm[4], SVM [5], and deep learning CNN, specifically 1D-CNN [6]. The formulation of the problems to be discussed in this research is as follows: 1) whether it is possible to build a system that can predict facial skin type according to the theory of Baumann Skin Type Solutions by Dr. Leslie Baumann uses machine learning among Indonesian people; 2)the method that gives out the best performance in predicting facial skin type based on Baumann Skin Type Solutions theory among SVM, XGBoost, and 1D-CNN.

This research is focused on building a machine learning model that is able to predict facial skin types and initiatingthe research and collection of data related to the classification of facial skin types in among Indonesian citizens based on the theory of

Baumann Skin Type Solutions by Dr.Leslie Baumann. With the hope that it can be useful as a reference in conducting similar research, namely predicting facial skin type using questionnaire data through a machine learning approach, helping Indonesian people understand personal facial skin types so that they can carry out a facial skin care regime with appropriate products, andalso contribute to the field of facial skin type classification for aesthetic purposes and to the procurement of Indonesian facial skin type data for based on the theory of Baumann Skin Type Solutions.

2. State of the Art

A previous study on Baumann Skin Type Solutions Theory [7] on classifying facial skin types among Korean women was done using the 63 questions from the original Baumann Skin Type Solutions questionnaire. However, many questions which affect the questionnaire's length were found to have a negative impact on response rate [8]. This encouraged the idea of simplifying and decreasing the number of questions used on this research from the original of 63 questions to 24 simplified questions under the evaluation of an expert dermatologist. The research [7] also mentionsThe Fitzpatrick Scale as a popular facial skin type classification theory.However, despite the popularity, one constraint that the Fitzpatrick Scale has is the inability to predict skin reaction to laser treatments or chemical trauma. This made the Baumann Skin Type Solutions superior since it covers many factors that were not covered by the Fitzpatrick Scale.

There have been many studies regarding the building of facial skin type prediction systems. However, the dataset used in most of them was in the form of image and none was found to use categorical data in building the facial skin type prediction model and so no obtainable public dataset that could be used for this research was found. Research shows that manual data collection and labelling[9]is possible to be done and thus the same method of manually collecting and labelling dataset is used in this research. The dataset is distributed into 80% training data and 20% testing data [10]. The decision to use categorical data instead of image data like other previous studies was made based on the consideration that categorical data could cover more skin features like genetics and lifestyle which cannot be provided by image data [7]. Image type dataset can only cover information skin physical conditions like using image data to classify facial skin type based on the skin texture [11].

There are many machine learning approaches that can be used for categorical data classification problems. One method was to use Support Vector Machine (SVM). SVM works by mapping the input vector into a higher dimensional space non-linearly. In the higher dimensional space, a linear decision surface or hyperplane is built to separate the data between classes [5] SVM determines this hyperplane using support vectors and margins.

The 2D data is linearly separable, because a linear straight line can be drawn to separate all blue data from all red data. However, in 3D data, it is required to find a hyperplane that can separate the data first. SVM uses a hyperplane with the largest margin to be more accurate in classifying data in the future compared to a hyperplane with a small margin. The hyperplane with the largest margin is also called the maximum marginal hyperplane (MMH). Data that falls right on the hyperplane is also called a support vector [12]. In a previous study, SVM achieved 100% training accuracy and 99.7% testing accuracy [13]**.**

Tree-based machine learning models could also be used for classification problems. Research [14] on tree-based methods comparison was done using Random Forest, Extra Tree, and XGboost with Bayesian Optimization,and using k-fold cross validation as the evaluation method. This research showed that XGBoost achieved the best performance withthe highest accuracy of 92%. The XGBoost algorithm builds a model by generating a set of weak learners through boosting. Boosting iteratively forms rules that have a certain weight or residual to be studied by weak learners, where for each incorrect prediction result, the rule will be given more weight or residual, so that in the next iteration the learner can learn the error [15].

XGBoost is well known for its scalability and speed which is capable of processing billions of data and running ten times faster than other existing methods. In addition, other innovations that complement XGBoost include handling incomplete data (sparse data awareness), a block structure that supports parallel tree creation, and continued training which means that the final model that has been created can be developed (boost) further with new data processing.

XGBoost is equipped with many hyperparameters that can be set and combined to get optimal XGBoost performance.

Table 1. XGBoost Hyperparameter Details

Hyperparameter	Default	Description
eta	0.3	Shrink the feature weight at the end of each boosting step
gamma	0	Specification of the minimum value of loss reduction required to perform a node split.
max_depth	6	The maximum level of the tree
max_child_weight	1	The minimum amount of weight a child or leaf node can have
subsample	1	Shows the fraction of observations to be a random sample for each tree.
colsample_bytree	1	Parameters for column subsampling
lambda	1	L2 weights regularization
alpha	0	L1 weights regularization
n_estimators	100	Number of gradient boosted trees created

Bayesian Optimization is an approach that uses Bayes' Theorem which directs the search to look for an objective function's maximum or minimum value [17].

In deep learning, 1-Dimensional Convolutional Neural Networkcan be used for classification problems using categorical data. In the field of image recognition and processing, Convolutional Neural Network (CNN) is an artificial neural network designed specifically to convolutionally process pixel data. CNN uses a system such as a multilayer perceptron. The neuron layer of CNN is made of an input layer, an output layer, and a hidden layer consisting of multiple convolutional layers with activation function parameters such as ReLU, pooling layer, flatten layer, fully connected layer, batch normalization layer, and dropout regularization techniques.

The convolutional layer is the first layer used for feature extraction from the input image. In this layer, mathematical calculations are carried out between the input image and the filter/kernel which moves across the entire surface of the input image and produces a feature map containing image information. The pooling layer is right after the convolutional layer and is meant to decrease the dimensions of the produced feature map. There are several types of pooling that can be used, such as max pooling, average pooling, and sum pooling. Batch Normalization is used to normalize the activation of an input volume before passing it on to the next layer. Batch Normalization standardizes the input in each batch so that the model learning process can run more stably. Dropout is a regularization technique that aims to help prevent overfitting by temporarily removing some neurons at random from the training process [18].

CNN is generally used for image data. However, Baosenguo at the Kaggle competition in 2021 developed the 1 Dimensional-CNN (1D-CNN) method for tabular data and won the best performance. The model was built with the inspiration that CNN has a reliable performance in feature extraction but is rarely used to process table data because the table data does not have a special relationship or ordering between each feature[20].

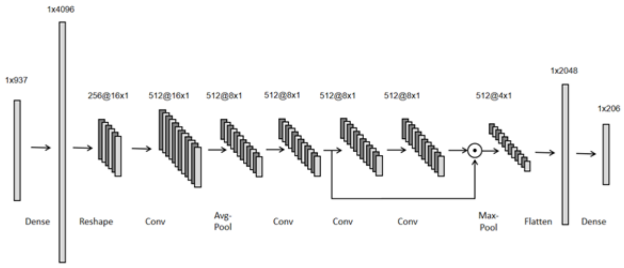

Figure 1. 1D-CNN Architecture

The key difference between 2D-CNN and 2D-CNN is that 1D-CNN, instead of using 2D matrices, uses 1D arrays for the kernel and feature maps. The raw 1D data is then processed by the CNN layer to learn the extraction of the features that will be used in the classification tasks performed by the dense layer or fully connected layer. Classification operations and feature extraction are thus combined into a singular process which can be optimized to maximize the performance of the classification. This is the key advantage of 1D-CNN which can potentially result in lower computational complexity since 1D convolution sequence is the only operation with significant cost which only consists of simple linear weighted sums of existing 1D arrays [20]. In previous research, 1D-CNN achieved a high accuracy of 97.79% on training data and 96.77% on testing data [21].

The problem of the amount of data between classes that is not balanced often occurs in a classification process. There are several strategies to overcome this problem. One strategy that can be done is oversampling. Classes with fewer data are better represented due to oversampling improving the distribution of the training data. Oversampling functions by re-sampling data fromthe minority class so that the number between classes is balanced. [12] One method that can be used for oversampling is SMOTE-N.

SMOTE-N is an extension of SMOTE which is used for oversampling by looking at the nearest neighbour with the use of Value Difference Metric.

Value Difference Metric (VDM) observes the overlap of feature values in all existing feature vectors. To generate a new feature vector in the minor class, a new feature value set can be created by taking into account the majority vote of the feature vector and its k nearest neighbours [22].

One important step in building a machine learning model is feature selection since it not only helps in reducing training speed but also lowers model complexity, increases generalization, and makes it easier to understand and increase performance metric using accuracy, precision, or recall [23]. There are various feature selection methods that can be used to select features from a categorical data type. One research [24] did a performance comparison among feature selection methods namely Recursive Feature Elimination, Random Forest, Principal Component Analysis, Genetic Algorithm, and Univariate Feature Selection. Another research on feature selection methods [23] also compared performances between RFE, Boruta, and Random Forest. The result of both researches showed that Random Forest achieved the best performances in comparison to other feature selection methods in selecting features from categorical data.

The Random Forest (RF) method uses an ensemble learning approach through bagging. Certain part of the features arerandomly chosen, and a voting procedure is then applied to classify them by taking into consideration the output of several trees which then provides the final result. This model has proven to have superior results compared to individual decision trees [25]. Feature selection method using RF is carried out by ranking of the results of measuring the importance of the feature set. The process of building a decision tree makes use of the Classification and Regression Tree (CART) algorithm. One of the two ways of measuring the importance of features is by employing the partition function by the use of Gini Index and calculating the importance of each feature based on the 'Gini Importance' score [26].

3. Methodology

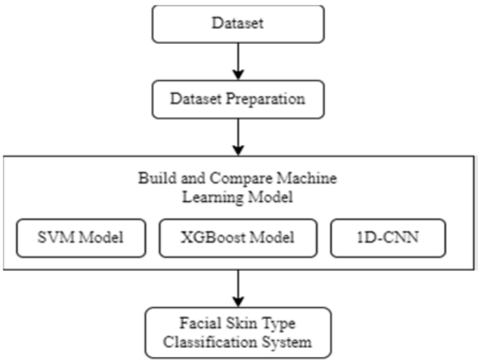

Figure 2. General Experiment Flow

Figure 2 visualizes the outline of the experimental flow of this study. Experiments were conducted on several machine learning models to determine the most optimal model as a facial skin type prediction system. The models are: SVM, XGBoost, and 1D-CNN. Model selection is done by comparing the evaluation values from cross-validation on each different model. The cross-validation used in this study implements stratified sampling or also called stratified k-fold cross-validation. This technique is applied so that the proportion between classes in each fold remains in accordance with the original proportion of the dataset.

3.1. Dataset

Dataset was collected by distributing questionnaires containing a list ofquestions referring to the book by Dr. Leslie Baumann, namely Baumann Skin Type Solutions, which is the base theory of skin types in this study.The questions will later become the features of the dataset.The responses that have been collectedare then labeled manually based on Baumann Skin Type Solutions.It has been evaluated by a dermatologist. This dataset consists of 300 instancesof categorical data with the following details: 1) Oiliness category consists of 167 Oilyand 133 Dry instances, 2) Sensitivity category consists of 94 Sensitiveand 206 Resistant instances, 3) Pigmentation consists of 224 Pigmentedand 76 Non-Pigmented instances, 4)Tightness consists of87 Wrinkled and 213 Tight instances.

3.2. Dataset Preparation

Dataset that has been collected through questionnaire will go through a series of preprocessing stages.

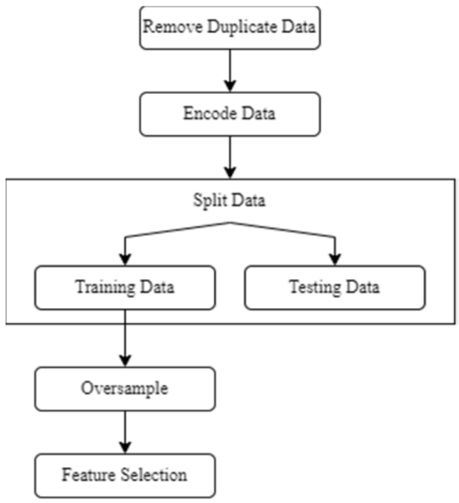

Figure 3. Dataset Preparation Flow

First, duplicate data is removed to obtain unique data to be processed. It helps the model to generalize better even on new data. Next, the data is encodedusing Ordinal Encodertoget the form that can be processed by machine learning models. Encoded data is then split into 2 parts with 80%being training data and 20%being testing data. Due to the imbalance class problem, oversampling is done on training datausing Synthetic Minority Over-sampling Technique for Nominal (SMOTE-N) method with k-neighbors = 5.

The last step of dataset preparation is feature selection. The feature selection process is carried out to reduce computational costs and improve model performance.There are two stages of feature selection: 1) manual feature selection and 2) tree-based intrinsic feature selection. Manual feature selection is done by eliminating features that have missing values of around 70% of the total data. On the other hand, tree-based feature selection is done using Random Forest with the calculation of Gini Importance score.

3.3. Support Vector Machine (SVM)

In this study, SVM is implemented as a part of the evaluation to compare the performance of traditional machine learning algorithms against other classifier model.

Kernel hyperparameter tuning is done in the construction of SVM model. Based on the experimental results, polynomial kernel is the most optimal kernel in this dataset.

3.4. eXtreme Gradient Boosting (XGBoost)

The XGBoost method is also implemented as a part of the evaluationto compare the performance of a tree boosting method in classifying to compare with other models.In this study, Hyperparameter Tuning using Bayesian Optimization is carried out to obtain the most optimal XGBoost model.

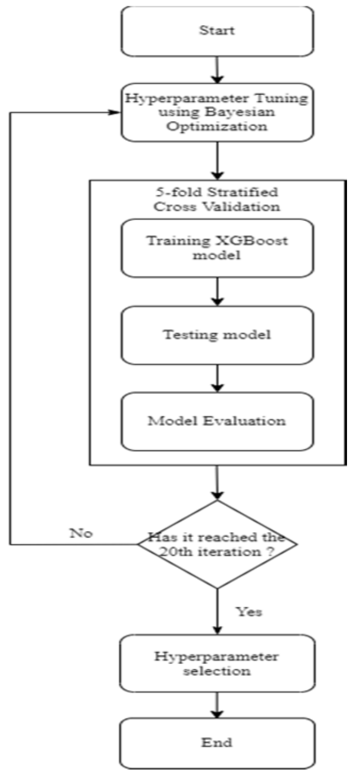

Figure 4. XGBoost Hyperparameter Tuning Scheme

Figure 4 shows the implementation of the Hyperparameter Tuning process. There will be 20 iterations to find the most optimal hyperparameter. In each iteration, Bayesian Optimization tries to optimize the F1-score value of the model through the resulting hyperparameters.

To find out whether the hyperparameters of the model are the most optimal, a comparison between the average F1-score values of the 5-fold stratified cross-validation generated in each iteration is carried out in Figure 4. After completing the iteration, the hyperparameter that makes the model result the highest F1-score value will be selected as the hyperparameter for the skin type prediction system.

3.5. Convolutional Neural Network (CNN)

CNN is used as part of the evaluation to compare the performance of deep learning algorithms with other classifier models. The CNN used in this study is a 1D-CNN, where the model only has 1 dimension in the convolutional layer.

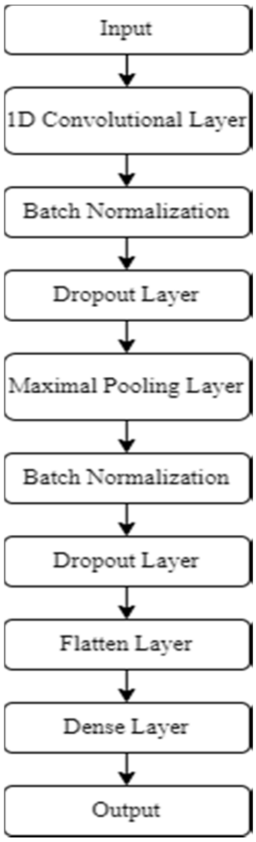

Figure 5. 1D-CNN Architecture

The architecture obtained from experimental results consists of 1 convolutional layer, 1 pooling layer, and a fully-connected layer. Input feature extraction uses the convolutional layer. This layer has a size of 2 with 32 kernels. In the pooling layer, Maximal Pooling with a pool size of 2, is used by taking the largest value from each pool to decrease the feature map dimension. In addition, to carry out the classification process, the fully-connected layer consists of 1 dense layer. The architecture uses the Batch Normalization to normalize the activation of an input volume before proceeding to the next layer in each batch. Dropout is also implemented to remove some neurons randomly from the training process to avoid overfitting. The 1D-CNN model training was carried out for 100 epochs with a batch of 64.

Each layer in 1D-CNN can use different activation functions. The ReLU Activation Function is used to get the Maximal Pool on the Feature Map generated from the Convolutional Layer Filter. In addition, the ReLU Activation Function is also applied in the Fully Connected Layer to convert the input with the added weight of the node into output for other nodes. The activation function used to get the final output in this model architecture is the Sigmoid Function. Adam Optimizer is also used in this architecture. To calculate the loss for each epoch, binary cross entropy is used.

4. Results and Discussion

Several experimental scenarios were carried out to determine the combination of dataset preparation methods and machine learning algorithms to be used for the prediction system.

4.1. ModelComparison Result

Scenarios are evaluated using 5-fold stratified cross validation. Table 2. presents the optimum results of the evaluation on each model using training data of each category after the experiment was carried out.

Table 2.Comparison of the Results of the Evaluation of the Classification Method for Each Category

Category	**Method**		
	SVM	**XGBoost**	**CNN**
Oiliness	91.70%	92.11%	90.71%
Sensitivity	95.78%	96.88%	93.14%
Pigmentation	97.46%	98.67%	93.67%
Tightness	83.44%	93.15%	85.27%

XGBoost outperforms the other 2 methods with F1-score of 92.11% for Oiliness category using oversampling and feature selection with 8 selected features, 96.88% for Sensitivity category using oversampling and feature selection with 13 selected features, 98.67% for Pigmentation category using oversampling and feature selection with a total of 7 features selected, and 93.15% for Tightness category using oversampling and feature selection with 13 selected features.

4.2. XGBoostModel Evaluation

Deriving from the outcomes of the scenario experiments above, the XGBoost model has the highest 5-fold stratified cross validation results in each selected scenario in each category of facial skin type. Therefore, XGBoost will be used as a predictive system model.

In building the prediction system model, training is carried out with 80% of the data from the dataset. The data was also previously used in hyperparameter tuning for the XGBoost model.

		Actual	
		Oily	*Dry*
Prediction	*Oily*	20	2
	Dry	4	21

Figure 6. Confusion Matrix Category Oiliness

Prediction \ Actual	Sensitive	Resistant
Sensitive	18	6
Resistant	1	17

Figure 7. Confusion Matrix Category Sensitivity

Prediction \ Actual	Pigmented	Non-Pigmented
Pigmented	7	1
Non-Pigmented	0	7

Figure 8. Confusion Matrix Category Pigmentation

Prediction \ Actual	Wrinkle	Tight
Wrinkle	16	0
Tight	2	42

Figure 9. Confusion Matrix Category Tightness

Evaluation of the XGBoost model that has been built is also carried out on the testing dataset. The data is taken from 20% of the total data, where the part of the data is new data that has not been used in the training or validation process. Figure 6 to 9 show the confusion matrix in each category.

Table 3. Selected Model Evaluation Matrix

Category	Precision	Recall	F1-score	Accuracy
Oiliness	0.91	0.84	0.87	0.87
Tightness	1	0.95	0.98	0.97
Sensitivity	0.74	0.94	0.83	0.83
Pigmentation	1	0.88	0.93	0.93

Table 3 shows the evaluation matrix of the model testing on the facial skin type prediction system. The model for the Oiliness category produces an F1-score and an accuracy of 0.87. In the Tightness category, the model produces an F1-score of 0.98 and an accuracy of 0.97. In addition, the model in the Sensitivity category produces an F1-score and an accuracy of 0.83. The Pigmentation category model produces an F1-score and an accuracy of 0.93. The outcomes of the model evaluation in each category show that the model has a satisfactory performance for classifying even on new data.

5. Conclusion

The lack of knowledge about different facial skin types is still a frequent problem in Indonesia. The purpose of this research is to build a facial skin type prediction system using machine learning to classify facial skin types based on Baumann Skin Type Solutions that can help to solve the skin type classifying problem. Based on the trials that have been conducted, it is concluded that building a facial skin type prediction system based on Baumann Skin Type Solutions using machine learning approach is possible to be done. In this research, three different machine learning models, namely SVM, XGBoost with Bayesian Optimization, and 1D-CNN were built and compared to find out which model performs best on the collected BSTS dataset. XGBoost gave out the best performance based on the evaluation result of the testing dataset with an F1-score of 0.87 for the Oiliness category, 0.83 for the Sensitivity category, 0.93 in the Pigmentation category, and 0.98 for the Tightness category.

6. Further Works

The data used in this study are obtained from Indonesian respondents and thus the models were only able to learn Indonesian facial skin type characteristics. Collecting data from other various geographical locations can result in a more diverse dataset which can help the model to learn better about different facial skin types globally. This can also increase the size of the dataset which can help the learning process of machine learning models. The bigger the dataset size, the better the model can learn from the knowledge source that is provided.

References

[1]. PT.Zulu Alpha Papa. (2021). ZAP Beauty Index Research. Diambilkembalidari ZAP Clinic. Retrieved from: https://zapclinic.com/zapbeautyindex [accessed: 15 August 2022].

[2]. Putri, A. S. (2020). 5 Hal yang Harus Diperhatikan Ketika Memilih Produk Perawatan Wajah. Diambilkembalidari Fimela. Retrieved from: https://www.fimela.com/beauty/read/4161870/5-hal-yang-harus-diperhatikan-ketika-memilih-produk-perawatan-wajah [accessed: 21 August 2022].

[3]. Harper's Bazaar . (2017, April 24). What Happens When You Use The Wrong Skincare For Your Age? Retrieved from Harpers Bazaar Arabia. Retrieved from: https://www.harpersbazaararabia.com/beauty/expertise/what-happens-when-you-use-the-wrong-skincare-for-your-age [accessed: 27 August 2022].

[4]. Chen, T., &Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acmsigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

[5]. Cortes, C., &Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

[6]. Li, Z., Yang, W., Peng, S., & Liu, F. (2020). A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *arXiv e-prints*, arXiv-2004.

[7]. Ahn, S. K., Jun, M., Bak, H., Park, B. D., Hong, S. P., Lee, S. H., ... & Park, K. (2017). Baumann skin type in the Korean female population. *Annals of Dermatology*, *29*(5), 586-596.

[8]. Kato, T., & Miura, T. (2021). The impact of questionnaire length on the accuracy rate of online surveys. *Journal of Marketing Analytics*, *9*(2), 83-98.

[9]. Chang, C. C., Hsing, S. T., Chuang, Y. C., Wu, C. T., Fang, T. J., Chen, K. F., & Choi, B. (2018, May). Robust skin type classification using convolutional neural networks. In *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)* (pp. 2011-2014). IEEE.

[10]. Chin, C. L., Chin, M. C., Tsai, T. Y., & Chen, W. E. (2018, September). Facial skin image classification system using convolutional neural networks deep learning algorithm. In *2018 9th International Conference on Awareness Science and Technology (iCAST)* (pp. 51-55). IEEE.

[11]. Nusantara, T. F. B., Atmaja, R. D., & Azizah, A. (2018). Klasifikasi Jenis Kulit Wajah Pria Berdasarkan Tekstur Menggunakan Metode Gray Level Co-occurrance Matrix (glcm) Dan Support Vector Machine (svm). *eProceedings of Engineering*, *5*(2).

[12]. Han, J., Kamber, M., & Mining, D. (2012). Concepts and techniques. *Morgan Kaufmann*, *340*, 94104-3205.

[13]. Krysovatyy, A., Lipyanina-Goncharenko, H., Sachenko, S., &Desyatnyuk, O. (2021). Economic Crime Detection Using Support Vector Machine Classification. In *MoMLeT+ DS* (pp. 830-840).

[14]. Budholiya, K., Shrivastava, S. K., & Sharma, V. (2020). An optimized XGBoost based diagnostic system for effective prediction of heart disease. *Journal of King Saud University-Computer and Information Sciences*, 1-96.

[15]. Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

[16]. Liu, J., Zhang, X., Li, Y., Wang, J., & Kim, H. J. (2019). Deep learning-based reasoning with multi-ontology for IoT applications. *IEEE Access*, *7*, 124688-124701.

[17]. Brownlee, J. (2019). *Probability for machine learning: Discover how to harness uncertainty with Python*. Machine Learning Mastery.

[18]. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., ... & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern recognition*, *77*, 354-377.

[19]. Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, *81*, 84-90.

[20]. Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., & Inman, D. J. (2021). 1D convolutional neural networks and applications: A survey. *Mechanical systems and signal processing*, *151*, 107398.

[21]. Hussain, S., Nanda, S. K., Barigidad, S., Akhtar, S., Suaib, M., & Ray, N. K. (2021, December). Novel Deep Learning Architecture for Predicting Heart Disease using CNN. In *2021 19th OITS International Conference on Information Technology (OCIT)* (pp. 353-357). IEEE.

[22]. Chawla, N. V., Bowyer, K. W., Hall, L. O., &Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357.

[23]. Chen, R. C., Dewi, C., Huang, S. W., &Caraka, R. E. (2020). Selecting critical features for data classification based on machine learning methods. *Journal of Big Data*, *7*(1), 1-26.

[24]. Ahuja, R., & Sharma, S. (2021). Exploiting Machine Learning and Feature Selection Algorithms to Predict Instructor Performance in Higher Education. *Journal of Information Science and Engineering*, *37*(5), 993-1009.

[25]. Reddy, V. S., Prasad, V. K., Wang, J., & Reddy, K. T. (2021). *Soft Computing and Signal Processing: Proceedings of 3rd ICSCSP 2020.* Volume 1. Springer.

[26]. Yao, R., Li, J., Hui, M., Bai, L., & Wu, Q. (2020). Feature selection based on random forest for partial discharges characteristic set. *IEEE Access*, *8*, 159151-159161.