JCTA Journal of Computing Theories and Applications ISSN:3024-9104

Research Article

Development of a Model to Classify Skin Diseases using Stacking Ensemble Machine Learning Techniques

Oluwayemisi Jaiyeoba [1,*], Emeka Ogbuju [2], Owolabi Temitope Yomi [1], and Francisca Oladipo [3]

[1] Department of Computer Science, Federal University Lokoja, Lokoja, Nigeria; e-mail: oluwayemisi.jaiyeoba@fulokoja.edu.ng; temittope010@gmail.com
[2] Department of Computer Science, Miva Open University, Abuja, Nigeria; email: emeka.ogbuju@fulokoja.edu.ng
[3] Department of Computer Science, Thomas Adewunmi University, Oko, Nigeria; e-mail: francisca.oladipo@fulokoja.edu.ng
* Corresponding Author : Oluwayemisi Jaiyeoba

Abstract: Skin diseases are highly prevalent and transmissible. It has been one of the major health problems that most people face. The diseases are dangerous to the skin and tend to spread over time. A patient can be cured of these skin diseases if they are detected on time and treated early. However, it is difficult to identify these diseases and provide the right medications. This study's research objectives involve developing an ensemble machine learning based model for classifying Erythemato-Squamous Diseases (ESD). The ensemble techniques combine five different classifiers, Naïve Bayes, Support Vector Classifier, Decision Tree, Random Forest, and Gradient Boosting, by merging their predictions and utilizing them as input features for a meta-classifier during training. We tested and validated the ensemble model using the dataset from the University of California, Irvine (UCI) repository to assess its effectiveness. The Individual classifiers achieved different accuracies: Naïve Bayes (85.41%), Support Vector Machine (98.61%), Random Forest (97.91%), Decision Tree (95.13%), Gradient Boosting (95.83%). The stacking method yielded a higher accuracy of 99.30% and a precision of 1.00, recall of 0.96, F1 score of 0.97, and specificity of 1.00 compared to the base models. The study confirms the effectiveness of ensemble learning techniques in classifying ESD.

Keywords: Dermatology; Erythemato-Squamous Diseases; Machine Learning; Skin Diseases; Stacking.

1. Introduction

The skin is the body's biggest organ, a critical protective barrier, and a physical shield against external threats. The skin comprises three primary layers: the epidermis, the dermis, and the Hypodermis [1], [2]. The epidermis is the outermost layer of the skin, consisting of epithelial cells that protect against environmental factors such as pathogens and Ultraviolet radiation [3]. The dermis is located beneath the epidermis and contains various structures such as blood vessels, nerves, hair follicles, sweat glands, and sebaceous glands [4]. The Hypodermis is located beneath the dermis; it consists of adipose tissue (fat cells) and connective tissue. It provides insulation, energy storage, and padding for the body [5].

Skin diseases are a wide range of conditions that affect the skin's structure, function, or appearance, which may cause the individual to lose confidence or descend into depression [6]. Therefore, detecting skin diseases at an early stage is critical. Some diseases may be an inflammation of the skin condition, affecting other organs of the system if not diagnosed early, while others may come as a disease that causes an itchy rash with flaky scales caused by factors such as stress and harsh detergents. Sometimes, they appear as a chronic, itchy, and painful inflammatory autoimmune disease affecting skin, scalp, and nails; they appear purplish and flat-topped with bumps. Skin diseases encompass various conditions, each with traits and symptoms [7]. Several prevalent skin disorders include scabies, Melanoma, skin eruptions, malignant Melanoma, eczema, psoriasis, acne, warts, vitiligo, tinea corporis, squamous cell carcinoma, basal cell carcinoma, inherited conditions such as genetic skin disorders, leprosy,

Received: April, 14th 2024
Revised: April, 30th 2024
Accepted: May, 16th 2024
Published: May, 20th 2024

viral infections, seborrheic dermatitis, sickle cell anemia, lichen planus, pityriasis rosea, persistent dermatitis, Pityriasis Rubra Pilaris, herpes, seborrheic keratosis, birthmarks, among others. Only a few numbers of people can recognize a particular kind of disease without access to a field guide. The disease must be correctly recognized and classified early to choose the medication. The skin can be affected by various internal and external factors, including Genetics, Hormonal Changes, Ultra-Violent light, smoking, alcohol consumption [8], environmental factors, bacterial and viral infections, a weak immune system [9], reported fungi, heat rash, and black patches.

Erythemato-Squamous Disease (ESD) is a group of dermatological conditions characterized by skin redness (erythema) brought on by skin cell loss (squamous). These conditions, which are often known as "red-skin diseases," have many of the same clinical symptoms, such as erythema (redness), scaling, and itching [10], [11]. The diseases are Pityriasis rubra pilaris, psoriasis, lichen planus, pityriasis rosea, chronic dermatitis, and seborrheic dermatitis.

Machine learning (ML) is a branch of artificial intelligence that focuses on creating systems that can learn from problem-specific training data to automate the creation of analytical models and resolve related tasks [12]. ML relies on different algorithms to solve problems, and this algorithm learns from historical patterns to uncover insights, detect patterns, and make predictions [13], [14]. ML models are used in classification and prediction tasks, and this task often employs supervised learning techniques, where an algorithm is trained on label data to learn the relationship between input features and output labels. Examples of some of the models include Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Naïve Bayes (NB). Ensemble learning techniques provide a powerful extension to traditional machine learning methods. Ensemble techniques are procedures applied to train multiple models and combine their outputs to produce a single prediction with a higher accuracy rate [15], [16]. It comprises two types: parallel and sequential. RF and bagging are parallel techniques that train base classifiers separately and then combine their predictions. While Sequential ensembles, such as boosting methods, continuously fix prior errors in the model to improve predicted accuracy [17]. Examples of ensemble techniques used in machine learning include Bagging, AdaBoost, Gradient Boosting Machines, Extreme Gradient Boosting (XGB), Stacking Ensemble, and Voting Ensemble. This research study aims to improve the accuracy of classifying skin diseases using an ensemble method of five algorithms. The significance is in providing the right artificial intelligence diagnosis to aid the treatment of skin diseases.

1.1. Challenge and Contribution

The current method of diagnosing skin diseases involves a combination of visual inspection by healthcare professionals, patient history-taking, and diagnostic procedures such as skin biopsy or laboratory studies. While these approaches have proved useful to some extent, they are frequently time-consuming and intrusive and may not always produce accurate results. Also, some skin disorders may appear with atypical symptoms or be difficult to detect visually, creating obstacles to correct diagnosis. The Current approaches are unreliable due to time limits and human error. The use of ML techniques can be used to accelerate the diagnostic process, resulting in earlier intervention and better patient outcomes. The main contribution of this study is the utilization of an ensemble method known as stacking in classifying skin diseases. Stacking involves combining the predictive abilities of multiple base methods to improve overall accuracy. Several articles in the literature used ML techniques to diagnose ESD, such as RF, SVM, K Nearest Neighbor (KNN), DT, LR, and more. In this study, we will be stacking five (5) ML models, including NB, SVM, RF, DT, and Gradient Boosting (GB), to enhance the diagnostic accuracy of skin diseases.

The rest of the study is organized as follows: Section 2 delves into the existing literature, exploring related works. Section 3 outlines the methodology employed in our research. The findings and their analysis are elaborated upon in Section 4. While Section 5 summarizes our findings and outlines future research.

2. Related Works

ML experts have been actively studying the classification of skin diseases to achieve accurate and efficient diagnostic outcomes. Prior research has explored various methodologies to develop an artificial intelligence diagnostic system for skin cancer classification.

Study [18] Classified skin cancer from images using ML and deep learning to identify whether a tumor is malignant or benign on dermoscopic images. The ML algorithms used were LR, linear discriminant analysis, KNN classifier, DT classifier, and Gaussian Naive Bayes (GNB). Ensemble learning was adopted to maximize accuracy. The deep learning model used was the convolution neural network (CNN). The ML model was grouped into E1, E2, E3. Where E1 is a classifier that maximizes diversity and averages predictions from very different ML methods (LR, KNN, and GNB), E2 averages predictions from all the ML models used in this work (LR, KNN, Linear Discriminant Analysis (LDA), Classification and Regression Trees, and GNB), E3 gathers the predictions from the three ML methods (LR, and LDA) that exhibit the best performance. The result of the ML model shows that LR and LDA had the highest accuracy. However, the deep learning models, with an accuracy of (0.88%) performed better than the ML models, with an accuracy of (0.72%) which increased to (0.75%) with ensemble learning.

The study [19] proposed a system that uses a novel method to predict skin diseases using supervised classification techniques. The classifiers used were KNN, SVM, and RF, and the 3 different metrics were used to measure the model's performance. RF yielded the highest accuracy, with 97%. The model consisted of 9 different skin diseases with 11 features.

The study [20] used ML and multi-model ensemble techniques to classify and predict ESD diseases. They applied 6 classification techniques: LR, LDA, KNN, Classification and Regression Trees, NB, SVM, and 4 Ensemble methods. 2boosting (Ada boosting and gradient boosting) and 2 baggings (RF and Extra trees) were used to improve the algorithm's performance. The model was evaluated, and an accuracy of 99% was achieved.

The study [21] proposed a model using images to diagnose skin diseases. The study classified 4 types of skin diseases: acne, cherry angioma, Melanoma, and psoriasis using a SVM, RF, and KNN. The dataset consisted of 377 images and was split into 80% for training and 20% for testing. The approach involves employing a medium filter for resizing images and eliminating noise. These images are then converted to grayscale. Otsu's thresholding technique distinguishes between different diseases, and features are extracted using Gabor, entropy, and Sobel methods. When applying the RF algorithm, an accuracy of 84.2% is attained. For the KNN algorithm, the accuracy reached 67.1%. However, the SVM classifier outperforms the others, achieving an accuracy of 90.7%.

The study [22] proposed an Ensemble meta-strategy for the detection of ESD: Psoriasis, Seborrheic dermatitis, Pityriasis rosea, Chronic dermatitis, Lichen planus, and Pityriasis rubra pilaris. The study compared algorithms such as RF, DT, NB, KNN, and Multi-Layer Perceptron for classifying ESD based on clinical and histopathological data. The preprocessing techniques used on the dataset were data cleaning, feature selection, data transformation, data splitting, and cross-validation. The model was evaluated using accuracy, precision, recall, and F-1 score and achieved an accuracy of 97.8%.

The study [23] proposed a KNN model for accurately identifying and classifying skin lesions as normal or benign. The preprocessing techniques used in the study were image preprocessing, segmentation, feature extraction, and classification. The system achieved an accuracy of 98% in classifying skin lesions.

The study [24] proposed a SVM with different kernel functions (polynomial, radial basis function, linear, and sigmoid), LR, and GNB classifier for classifying Melanoma skin disease. The study applied preprocessing techniques such as morphological algorithms, filters, and image sharpening to enhance image quality. The study introduced an interpretable feature selection process using Recursive Feature Elimination to rank and select features based on their importance. The study evaluated the model's generalization through feature screening, model evaluation using tenfold cross-validation, and 100-fold randomization experiments. The result showed that LR and SVM linear had the highest accuracy at 74.71%, followed by GNB at 68.96%, SVM polynomial at 65.92%, SVM sigmoid at 62.93%, and SVM basis function at 61.96%.

The study [25] proposed a ML classifier to classify skin lesions. The classifiers used in the study were KNN, RF, and SVM. The dataset used was the International Skin Imaging Collaboration (ISIC), and it was preprocessed by applying a median filtering method to remove unnecessary elements such as hair, bubbles, and noise from the images before segmentation and feature extraction. The evaluation metrics used in the study included accuracy, sensitivity, and specificity for assessing the performance of various classifiers with features like the ABCD rule, Gray-Level Co-occurrence Matrix (GLCM), and shape features. The

Support Vector Machine classifier achieved an accuracy of 89.43% for the ABCD rule feature, 85.72% for the GLCM feature, and 82.31% for the shape feature in classifying melanoma skin lesions.

The study [26] proposed a ML model for skin diseases, including actinic keratosis, benign keratosis, dermatofibroma, vascular lesion, Melanoma, melanocytic nevus, basal cell carcinoma, and Squamous cell carcinoma. The study utilized two standard datasets, the ISIC 2019 challenge and the HAM10000 dataset. The preprocessing techniques used in the study included image resizing to 512 × 512, digital hair removal using Black-Hat transformation and inpainting, and noise removal with Gaussian filtering to enhance image quality for skin disease detection and classification. The evaluation metrics used in the study included accuracy, precision, recall, F1 score, categorical cross-entropy, Receiver Operating Characteristic (ROC) curves, and Area Under the ROC Curve (AUC) analysis. The study achieved a 98% accuracy for a KNN classifier, 92% accuracy for a convolutional neural network (CNN) model, 85% accuracy for a Mobile Net model with transfer learning, 95% accuracy for an optimal probability-based deep neural network, and an average accuracy of 100% for classifying specific diseases using a multiclass SVM approach.

The study [27] proposed the combination of a CNN with a SVM for the classification of Melanocytic Nevi, Benign Keratosis-like Lesions, Dermatofibroma, Vascular Lesions, Actinic Keratoses, and Intraepithelial Carcinoma, Basal Cell Carcinoma, and Melanoma. The dataset used was collected from the HAM10000 dataset consisting of 10015 images. The model was evaluated and achieved an accuracy of 85.75% with CNN+DT, 81.69% with CNN+KNN, 91.04% with CNN+SVM, and 84.37% with CNN+ Light Gradient Boosting Machine (LGBM).

The study [28] proposed a ML model for classifying the Melanoma skin disease. This classification system involves three stages: preprocessing, feature extraction using shape components, ABCD rule, and Gray-Level Co-occurrence Matrix (GLCM) features, and classification using KNN, RF, and SVM classifiers. The dataset was obtained from the PH2 database, and metrics such as accuracy, sensitivity, and specificity were assessed. The study achieved average accuracy rates of 91.169% for KNN, 87.615% for RF, and 94.817% for SVM classifiers.

The study [29] Proposed a RF model for skin cancer detection. The study utilized the ISIC- International Symposium on Biomedical Imaging (ISBI) 2016 dataset to classify benign or malignant melanoma skin diseases. Preprocessing techniques such as image resizing, noise removal with a median filter, contrast stretching based on mean and standard deviation of pixel intensities, Red Blue Green (RBG) to gray conversion, and hair removal using bottom-hat filtering were applied to the dataset. The evaluation techniques used were accuracy, sensitivity, precision, and AUR-ROC Curve. They achieved an accuracy of 93.89%.

The study [30] proposed a ML model using KNN and RF to classify skin diseases such as Acne, Melanoma, Actinic Keratosis, Cold Sore, Eczema, Psoriasis, and Rosacea. The images were obtained from an open-source benchmark dataset of 21,075 skin disease images with ten class labels. They calculated the accuracy of the classification model using performance metrics such as F1 Score, Precision, Recall, and average accuracy. An accuracy of 95.23% was achieved using KNN and 94.22% using RF.

The study [31] proposed a ML model applying three data mining techniques: SVM, DT, and RF for the classification of Pityriasis rubra, Lichen planus, Rosea pityriasis, Healthy skin, Psoriasis, Chronic dermatitis, Seborrheic dermatitis skin diseases. They applied an ensemble approach of three data mining techniques: SVM, DT, and RF. The k parts of training data were fed into these three algorithms as input to create k models of each algorithm in the ensemble model. The final ensemble models built by these algorithms are used for output, and a voting process is used to establish the majority ranking of predictions. The data was interpolated for missing values and normalized to ensure consistency and accuracy in the analysis. Evaluation techniques such as accuracy, sensitivity, and specificity were used to measure the performance of the classification models. An accuracy of 94.47% was attained for the DT model, 95.09% for the RF, and 92.63% for the SVM. When these three models were combined into an ensemble, the overall accuracy reached 96.93%.

The study [32] proposed a ML model to classify skin diseases such as Melanoma, basal cell carcinoma, vascular lesions, dermatofibroma, melanocytic nevi, benign keratoses lesions, and actinic keratoses. SVM was utilized in this study to classify these skin diseases. They

utilized the ISIC 2018 dataset and were preprocessed using Gaussian filter and contrast enhancement techniques. The model was evaluated, and an accuracy of 90.37% was achieved.

The study [33] proposed a ML technique, including LR, SVM, RF, KNN, and GB, to facilitate non-invasive melanoma skin cancer detection. They combined these models into a stacked ensemble for enhanced diagnostic accuracy. Preprocessing techniques such as Scaling data and data Augmentation were applied, and evaluation techniques such as Accuracy, F1-Score, Cohen's Kappa, Confusion Matrix, and ROC Curves were employed. The LR model achieved an accuracy of 84%, and the RF and SVM models both attained 84% accuracy. The GBM model performed slightly better, with an accuracy of 87%. The KNN model achieved an accuracy of 82%. Surpassing them all, the stacking model exhibited the highest accuracy of 88%.

The study [34] proposed a CNN and two classical ML classifiers, a SVM and KNN, for melanoma skin cancer detection. They employed and preprocessed the ISIC dataset using hair removal, lesion segmentation, and feature extraction techniques. The evaluation techniques used to assess the model's performance were accuracy, sensitivity, specificity, and AUC-ROC. They achieved an accuracy of 57.3% with KNN and 71.8% with SVM.

Table 1. Summary of previous studies in skin disease classification

Ref	Skin Diseases	Models	Accuracy	Datasets
[18]	benign keratosis-like lesions, dermatofibroma, melanocytic nevi, vascular lesions, Actinic keratoses, intraepithelial carcinoma/Bowen's diseases, basal cell carcinoma, and Melanoma.	Ensemble Model: (LR, LDA, KNN, CART, GNB)	75%	Kaggle & International Skin Imaging Collaboration
[19]	Abscess, eczema, fungal infections, psoriasis, scabies, acne vulgaris, urticaria, alopecia areata, pruritus, and decubitus ulcer.	KNN, RF, SVM	KNN: 78%, RF: 90%, SVM: 82%.	National Data Centre in India
[20]	psoriasis, seborrheic dermatitis, lichen planus, pityriasis rosea, chronic dermatitis and pityriasis rubra pilaris	Ensemble Techniques	99%.	-
[21]	Acne, cherry angioma, Melanoma, and psoriasis.	SVM	90.7%.	Dermnet NZ & Atlas dermatologico
[22]	psoriasis, seborrheic dermatitis, lichen planus, pityriasis rosea, chronic dermatitis, and pityriasis rubra pilaris	Ensemble Meta Technique	97.8%.	UCI Repository
[23]	Normal or Benign	KNN	98%	
[24]	Melanoma	GNB, LR, SVM	GNB: 69.13% LR: 74.80% SVM: 77.27%	PH2 & ISIC 2019
[25]	Melanoma	SVM	89.43%	ISIC
[26]	melanoma, melanocytic nevus, basal cell carcinoma , actinic keratosis, benign keratosis, dermatofibroma, vascular lesion, and Squamous cell carcinoma.	KNN CNN Mobile Net	KNN: 92% CNN: 85% Mobile Net; 95%	ISIC 2019 & The HAM10000 dataset
[27]	Melanocytic Nevi, Benign Keratosis-like Lesions, Dermatofibroma, Vascular Lesions, Actinic Keratoses, Intraepithelial Carcinoma, Basal Cell Carcinoma, and Melanoma.	CNN+DT, CNN+KNN, CNN+SVM, CNN+LGBM	CNN+DT= 85.75%, CNN+KNN = 81.69%, CNN+SVM = 91.04%, CNN+LGBM = 84.37%	- The HAM10000 dataset
[28]	Melanoma	KNN, RF, and SVM	KNN= 91.169%, RF= 87.615% SVM= 94.817%	PH2 database

Ref	Skin Diseases	Models	Accuracy	Datasets
[29]	Benign or malignant Melanoma	RF	93.89%	ISIC-ISBI 2016
[30]	Acne, Melanoma, Actinic Keratosis, Cold Sore, Eczema, Psoriasis and Rosacea.	KNN and RF	KNN: 95.23%, RF: 94.22%	-
[31]	Pityriasis rubra, Lichen planus, Rosea pityriasis, Healthy skin, Psoriasis, Chronic dermatitis, Seborrheic dermatitis	Ensemble techniques (SVM, DT, and RF)	96.93%.	UCI repository
[32]	Melanoma, basal cell carcinoma, vascular lesions, dermatofibroma, melanocytic nevi, benign keratoses lesions, and actinic keratoses.	SVM	90.37%.	ISIC 2018
[33]	Melanoma	Stacked Ensemble Techniques	88%	ISIC 2018
[34]	Melanoma	SVM and KNN	KNN: 57.3%, SVM: 71.8%	ISIC

2.1. Findings

Research has contributed to advancing skin disease diagnosis using ML techniques in related studies. Figure 1 shows that the SVM is the most prevalent classifier in 8 studies. Following closely, KNN was employed in 6 studies, while both RF and ensemble techniques were utilized in 5 studies. Each model, including Linear Regression, GNB, Mobile Net, and DT, was reviewed in separate studies. The reviewed studies [20] attained the highest accuracy of 99% by applying ensemble techniques. This outstanding accuracy demonstrates the effectiveness of using multiple models to improve performance in classifying skin disorders.

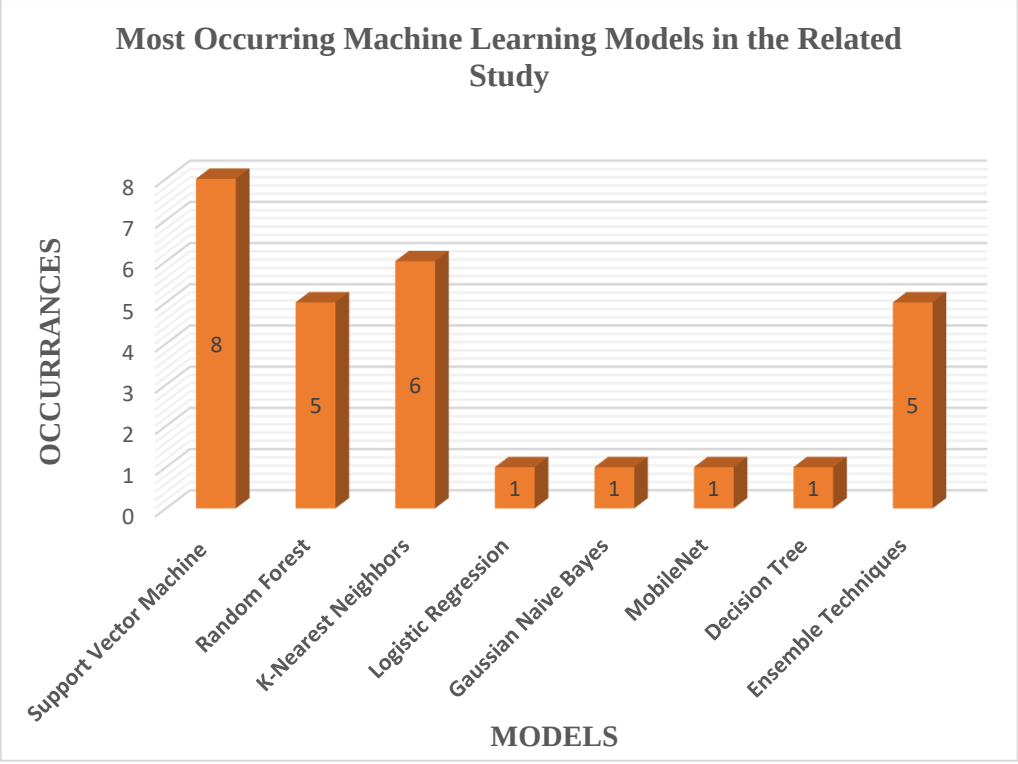

Figure 1. Most Occurring machine learning models in the related study

3. Methodology

Skin diseases are classified using many approaches and methodologies that have evolved. The study provides an overview of the methods used to classify various skin diseases.

Understanding these methodologies allows researchers to gain insight into classifying various skin diseases and explore novel strategies to improve diagnosis accuracy.

3.1. Proposed Framework

The framework, as shown in Figure 2, involves several steps. Initially, we gathered datasets containing both clinical and histopathological information. Next, we analyzed and preprocessed the dataset, addressing missing values and refining data quality through feature engineering. This preprocessing step was crucial and was performed in the third step. Following this, we divided the dataset into training (80%) and testing (20%) for model development by applying the random state method, a superparameter technique that helps split datasets optimally. We utilized five ML classifiers: NB, SVC, DT, RF, and GB. We then applied the Stacking ensemble technique to combine the predictions of these models to enhance the accuracy of skin disease classification. The model was evaluated to determine its effectiveness, employing metrics such as precision, accuracy, recall, specificity, and F1 scores. Lastly, deployment for real-world applications is planned.

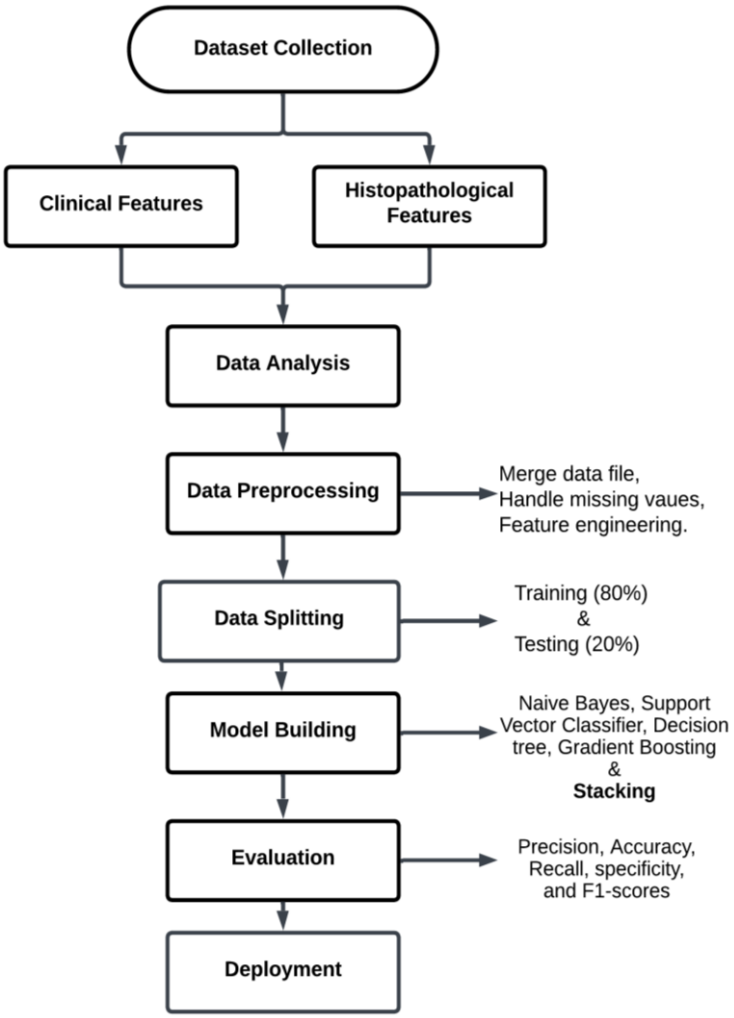

Figure 2. Proposed farmwork in the classification of skin diseases

3.2. Dataset Description

The dataset used in this study encompasses a comprehensive set of clinical and histopathological features, as shown in Table 2. It contains six classes of skin diseases, including Psoriasis, Seborrheic Dermatitis, Lichen Planus, Pityriasis Rosea, Chronic Dermatitis, and Pityriasis Rubra Pilaris. The dataset used in this study was obtained from the University of California, Irvine (UCI) repository [35] to classify ESD. The dataset comprises 366 instances stored as a CSV file, with each instance containing 34 distinct attributes. Among these attributes, 33 are linear, consisting of 12 clinical and 22 histological features and 1 attribute (Age).

If there are any illnesses in the family, Age features were indicated by the patient's age. A combination of clinical symptoms and histopathological findings characterizes each instance in the dataset. These were assigned a value from 0 to 3 (0 = absence of features; 1, 2 = comparative intermediate values; 3 = highest value.). Figure 3 shows the six different skin diseases along with the total occurrences of each. Table 2 outlines both the clinical and histological features of Erythemato-squamous diseases. Clinical features are those that can be easily seen on the skin's surface and are often assessed by eye inspection or patient reports. The Histopathological features provide information about microscopic changes seen under a microscope in skin tissue samples, providing insight into cellular and tissue-level anomalies related to the skin disorder. Features like "Family history" and "Age" are not physical symptoms but are necessary for understanding a person's medical background. All features in the dataset are represented as integer data types.

Table 2. Features of the Erythemato-squamous diseases

No	Features	Clinical	Histopathological
1	Erythema	☑	
2	Scaling	☑	
3	Definite borders	☑	
4	Itching	☑	
5	Koebner phenomenon	☑	
6	Polygonal papules	☑	
7	Follicular papules	☑	
8	Oral mucosal involvement	☑	
9	Scalp involvement	☑	
10	Knee and elbow involvement	☑	
11	Family history (0 or 1)	☑	
12	Melanin incontinence		☑
13	Eosinophils in infiltrate		☑
14	PNL infiltrate		☑
15	Fibrosis of the papillary dermis		☑
16	Exocytosis		☑
17	Acanthosis		☑
18	Hyperkeratosis		☑
19	Parakeratosis		☑
20	Clubbing of the rete ridges		☑
21	Elongation of the rete ridges		☑
22	Thinning of the suprapapillary epidermis		☑
23	Spongiform pustule		☑
24	Munro microabcess		☑
25	Focal hypergranulosis		☑
26	Disappearance of the granular layer		☑
27	Vacuolization and damage of basal layer		☑
28	Spongiosis		☑
29	Saw-tooth appearance of retes		☑
30	Follicular horn plug		☑
31	Perifollicular parakeratosis		☑
32	Inflammatory mononuclear infiltrate		☑
33	Band-like infiltrate		☑
34	Age	☑	

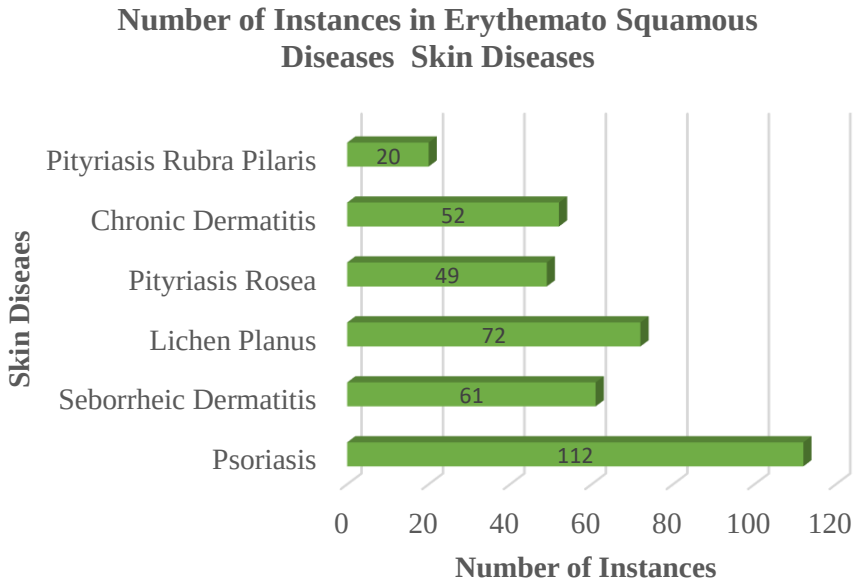

Figure 3. Number of Instances of ESD Skin Diseases

3.3. Data Preprocessing

The study began with extensive preparation of the skin disease dataset obtained from the UCI repository library. Initially, the dataset comprises of 2 separate files: the dermatology name file and the dermatology data file. The data file contained missing column names, which were provided in a separate file called the "name file." This file describes the clinical and histopathological attributes, which were then added to the corresponding columns in the data file. Also, the missing data were removed to ensure clarity and accuracy in the analysis.

3.3.1. Feature Selection

Feature selection involves identifying the most relevant attributes from a dataset to improve model performance while excluding irrelevant features [36]. By carefully selecting features, ML models can improve their generalization, making them more robust and adaptable to real-world circumstances. The preprocessing data consists of a total of 34 features. In this study, we included all 34 features, consisting of both clinical and histopathological features, in our modeling process. Their significant correlation with the target variable guided the decision to include these features.

3.3.2. Data Normalization

Normalization standardizes data attributes to a consistent scale, which is essential for ensuring meaningful comparisons across multiple features with disparate scales or units. It is commonly applied to numerical data[37]. In this study, we used 'StandardScaler' from the "scikit-learn" preprocessing module to standardize the features to ensure that all the selected features contributed equally to the analysis and to prevent any biases that could arise from differences in the scales of these features. Standardization involved adjusting the features to contain a mean value of 0 and a standard deviation of 1. This process helped to make the data more uniform and comparable for building an effective ML model.

3.4. Machine Learning Algorithms

ML algorithms are computational approaches that allow computers to learn from data and make predictions without explicitly programming each task [38]. In this study, we applied five different classification techniques, including NB, SVC, RF, DT, and GB, and then applied a stacking method to combine their predictive abilities

Naive Bayes: NB is a prediction method containing a simple probabilistic classifier based on Bayes' theorem and assuming feature independence. It calculates the likelihood of a given class label based on feature values, making it useful for text classification and other

classification tasks. [39] . Based on various symptoms or traits, Naïve Bayes is an efficient method for managing multiple features and may accurately predict the likelihood of a skin illness.

Support Vector Classifier: SVC is a supervised learning technique that maximizes the margin between classes to identify the ideal hyperplane in a high-dimensional space that divides them [40].

Random Forest: RF comprises several decision trees, each with a unique classification that affects position layout. This method evaluates sampling allocation using random sampling, which is especially suitable for minute models. In terms of classification, it selects the most common class among the trees, while for regression, it calculates the average prediction from all the trees [25].

Decision Tree: DT is a supervised learning algorithm used in ML for classification and regression problems [41]. It is composed of leaf nodes that each indicate an outcome, a dataset, inner nodes that represent the algorithm's decision, and branching structures [42]

Gradient Boosting: GB is an ensemble learning algorithm for classification and regression tasks. It produces an accurate classification by integrating the predictions of numerous weak predictive models, usually decision trees [43]

Stacking method: Stacking enhances predictive performance using numerous base models in an ensemble learning technique. The final prediction is generated by a meta-model using the base models' predictions as input features. Stacking works especially well when there is uncertainty about which individual model will perform best on unseen data or when the base models show complementary capabilities [44].

4. Performance Evaluation Measures

Evaluating the performance of a ML model is one of the most crucial aspects of creating an efficient model. It describes a model's capacity for prediction accuracy, which aids in evaluating the model's performance. These metrics aid in model comparison and hyperparameter adjustment by revealing how well the model performs on unknown data. Various metrics are used to evaluate the quality of the model and how well a model works with the data. The assessment of the models in this study was evaluated using precision, accuracy, recall, specificity, and F1 scores on the test dataset. The measurements are computed using the equations shown below.

Accuracy: The accuracy of the model's predictions is determined by calculating the percentage of correctly identified occurrences out of all instances[45].

$$\text{accuracy} = \frac{\text{True Negatives } + \text{TruePositives}}{\text{True Negative} + \text{True Positive} + \text{False Negative} + \text{False Positive}} \quad (1)$$

Precision: Precision is the number of accurately identified positive samples divided by the total number of positive samples[46].

$$\text{precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (2)$$

Recall: Recall measures how successfully a model can identify all relevant instances of a specific class in a dataset. It's calculated by taking the number of correctly predicted positive instances and dividing it by the total number of actual positive instances[47].

$$\text{recall} = \frac{\text{True Positives}}{\text{False Positives} + \text{False Negatives}} \quad (3)$$

Specificity: Specificity shows how well the model can detect true negatives by calculating the percentage of accurately identified negative cases out of all negative cases[48].

$$\text{specificity } = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}} \quad (4)$$

F1-scores: F1.- Score is a measure that combines precision and recall into a single number. It strikes a balance between these two measurements, accounting for both false positives and false negatives[49].

$$F1 - scores = \frac{precision\ x\ recall}{precision + recall} \quad (5)$$

5. Results and Discussion

This section presents the findings from our study on the classification of skin diseases using stacking ML approaches and compares them with existing techniques. Figure 4 illustrates the accuracy performance of five distinct ML models across both training and testing datasets. In the modeling process for skin disease classification using ensemble stacking techniques, we fine-tuned the parameters of the individual classifiers to achieve optimal performance. The DT model's criterion parameter is set to 'gini,' while the 'random state' is set to 0. SVC uses a 'sigmoid' kernel and a 'random state' of 42. GB model used 100 decision trees as base estimators, using the 'n_estimators' parameter. RF model was set to 100 decision trees and 'entropy' as the splitting criterion. Although NB did not undergo hyperparameter tuning, we ensured a 'random state' of 42 for consistency in its performance evaluation.

GB was selected as the meta-learner because of its ability to handle a variety of datasets and capture complex correlations between target variables and features. It shows resilience to overfitting, especially when using weak base learners in ensemble learning. Also, because of its capacity for reducing errors, it is a good fit for improving model accuracy. Certain parameters were selected to balance complexity and maximize the model's performance, such as n_estimators=10, learning_rate=0.1, criterion='friedman_mse', random state=0, and subsample=0.6 were carefully selected. This parameter helps to improve the performance of the stacking ensemble technique for predictive modeling.

NB achieved an accuracy of 90.33% on the training set and 85.41% on the testing set. SVC emerges as the best base model, achieving an accuracy of 100% on the training set and 98.61% on the testing set. RF and GB also achieved high accuracy, with RF achieving 98.31% accuracy on training and 97.91% on testing and GB attaining 97.12% and 95.83%, respectively. Despite slightly lower accuracies than SVC, DT achieved a training accuracy of 96.00% and a testing accuracy of 95.13%. However, a significant improvement in accuracy is observed through the stacking method, which combines the results from all five models, achieving a remarkable increase to 99.85% in training accuracy and 99.30% in testing accuracy. This significant improvement shows the effectiveness of ensemble learning techniques in enhancing performance by utilizing the strengths of different models.

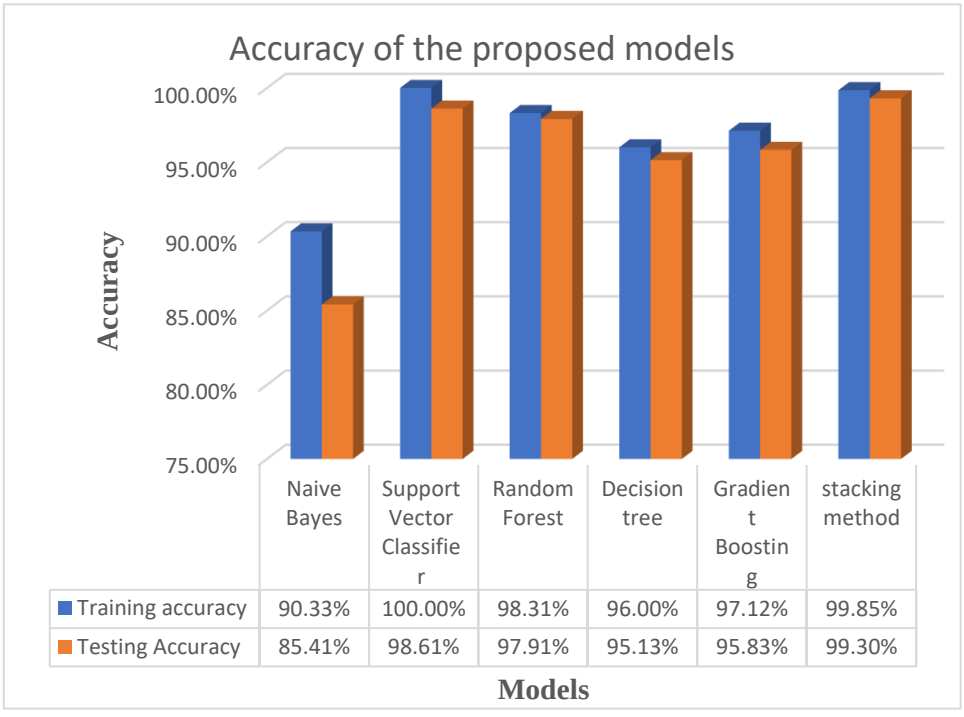

Figure 4. Comparison of different Individual Classifiers and Stacking model classification performance.

The confusion matrix of the models used in our study was to gain insights into how well they perform in classifying skin diseases. Each confusion matrix provides a comprehensive overview of the model's ability to classify instances across different classes, showing strengths and weaknesses. The value in each row represents the corresponding actual labels, and the values in each row represent the corresponding predicted labels. The diagonal cells represent the instances where the predicted class matches the actual class. The off-diagonal entries indicate misclassification. Where class 0 represents Psoriasis, class1 represents Seborrheic Dermatitis, class2 represents Lichen Planus, class3 represents Pityriasis Rosea, class4 represents Chronic Dermatitis, class5 represents Pityriasis Rubra Pilaris

Figure 5(a) presents the confusion matrix for the NB classifier's performance in classifying skin diseases. The confusion matrix wrongly categorized 17 instances from Class 1 as Class 3 and 1 instance as Class 4. It also misclassified 1 instance from Class 2 as Class 3. Figure 5(b) presents the confusion matrix for SVC, where 4 instances from Class 3 were misclassified as Class 1, and 1 instance from Class 1 was misclassified as Class 3.

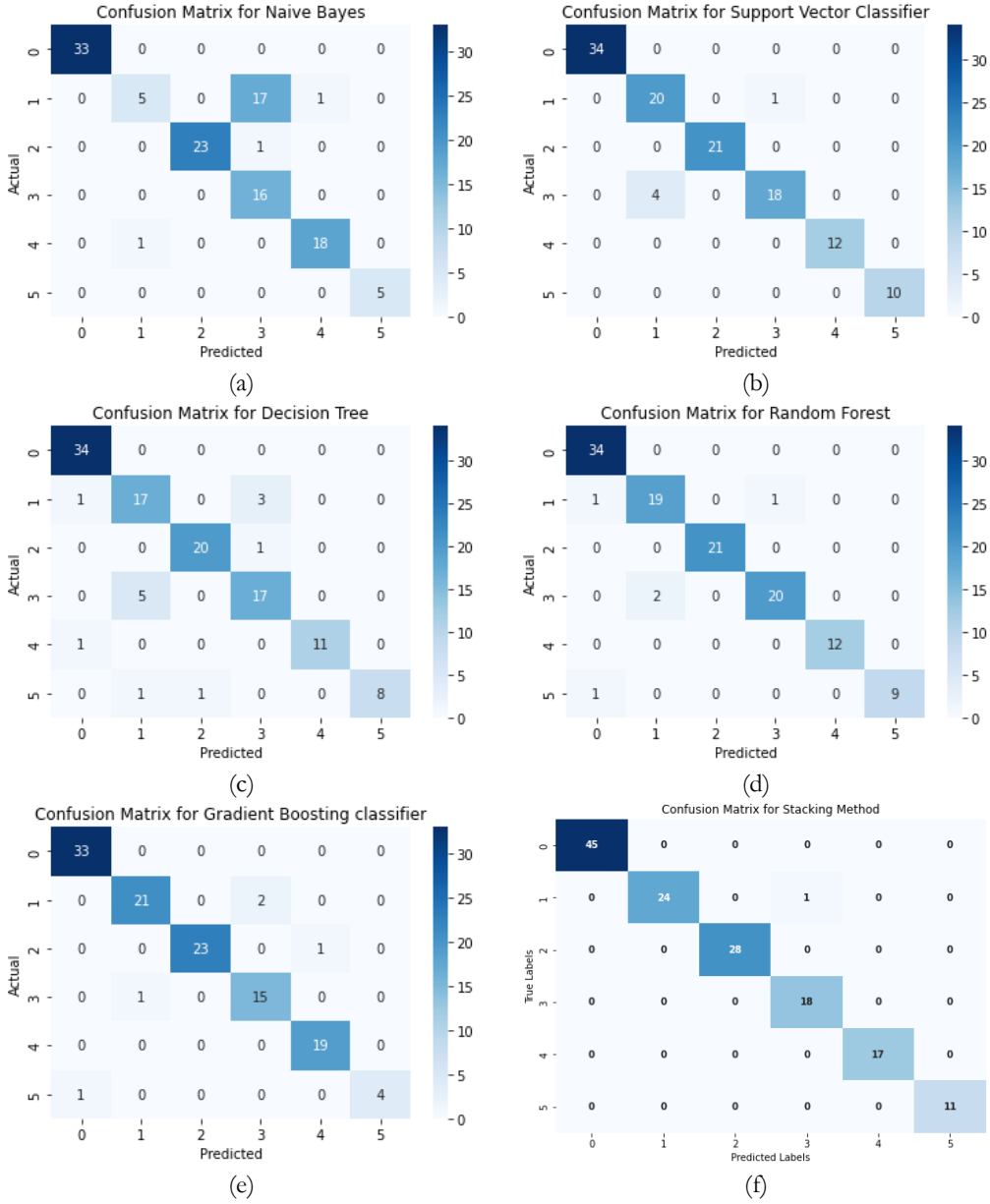

Figure 5. Confusion matrix results (**a**)Naïve Bayes; (**b**)SVM; (**c**)Decision tree; (**d**)Random Forest; (**e**)Gradient Boosting; (**f**)Stacking.

In Figure 5(c), a DT was employed, showing misclassifications of 1 instance from Class 1 as Class 0 and 3 instances from Class 1 as Class 3. While Class 2 misclassified 1 instance as Class 3, class 3 misclassified 5 instances as Class 1, class 4 misclassified 1 instance as Class 0, and Class 5 misclassified 1 instance as Class 1 and Class 2 each. The confusion matrix for the RF classifier, as shown in Figure 5(d), displayed misclassifications, including 1 instance each from Class 1 and Class 5 labeled as Class 0 and 2 instances from Class 3 labeled as Class 1. Also, 1 instance from Class 5 was misclassified as Class 0. GB was also employed, and the confusion matrix is shown in Figure 5(e), which misclassified 2 instances from Class 1 as Class 3, 1 instance from Class 2 as Class 4, 1 instance from Class 3 as Class 1, and 1 instance from Class 5 misclassified as class 0. However, when these models were combined and stacked, the performance improved. The confusion matrix shown in Figure 5(f) only misclassifies one instance in Class 1 as Class 3.

The classification performance of different ML algorithms was assessed to diagnose the different skin conditions. We compared the accuracy, precision, specificity, recall, and F1-score of five classifiers, including stacking the 5 models: NB, SVC, RF, DT, and GB. The stacking approach outperformed the other classifiers, followed by the SVC. Figure 6 illustrates the performance of skin disease classification using accuracy, precision, Fi-score, recall, and specificity. The Stacking Method achieved an accuracy of 99.30%, precision of 1.00, recall of 0.96, specificity of 1.00, and F1- score of 0.97. The Support Vector Machine achieved an accuracy of 98.61%, a precision of 100%, a recall of 83%, a specificity of 0.898, and an F1-score of 0.95. while the Random Forest achieved an accuracy of 97.91%, a precision of 0.94, a recall of 0.90, a specificity of 0.877, and an F1-score of 0.93. Naive Bayes achieved an accuracy of 85.41%, with 1.00 precision, 0.60 recall, a specificity of 0.905, and F1- score of 0.61. Gradient Boosting achieved an accuracy of 95.83%, with a precision of 0.97, recall of 0.95%, specificity of 0.869, and F1- score of 0.91. Decision Tree model demonstrated an accuracy of 95.13%, precision of 0.94%, recall of 0.74%, specificity of 0.874 and F1- score of 0.79. The stacking method's precision values were consistently high in all classes, showing its effectiveness in differentiating between various skin disorders. Furthermore, the SVC showed impressive recall and precision. The NB classifier showed lower precision and recall values than other algorithms. Overall, the findings demonstrate the efficacy of advanced ML approaches like stacking techniques in accurately diagnosing ESD skin diseases.

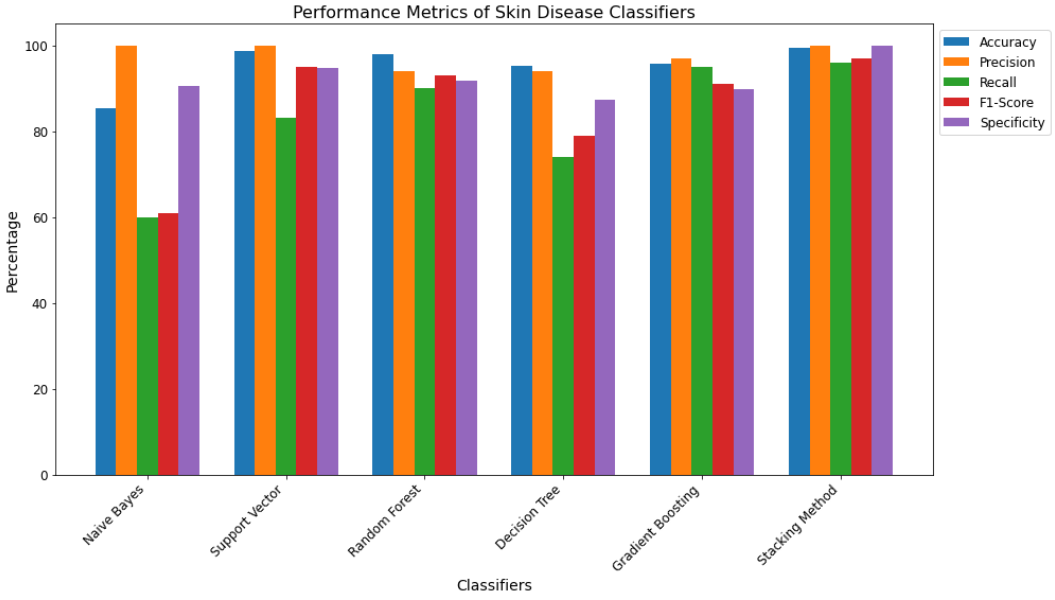

Figure 6. Performance Comparison of Individual Classifiers and Stacked Ensemble Technique in Accuracy, Precision, Recall, and F1-score.

We have compared our results with some well-established methods in ML, which were evaluated using the same ESD dataset. Table 3 shows the various techniques used in classifying skin diseases. Among the models discussed [50] employed an ensembled approach, combining SVM, KNN, DT, NB, and MLP, yielding an accuracy of 92.9% with a total of 34 features; 22 were selected in each subset and eventually narrowed down to 9 features for the

base classifier. [41] improved accuracy to 99% by employing LR, SVM, and KNN on the ESD dataset; they utilized 15 features out of 34 features. [51] achieved 99.07% accuracy by combining a hybrid multiclass SVM with Bayesian Optimization on the ESD dataset containing 34 features. [22] achieved 98.21% and 97.85% accuracy using RF and XGB on the ESD dataset with 6 attributes randomly selected out of 34 attributes for RF and all 34 attributes for the XGB model. In our model, we applied the Staking Ensemble technique by combining all five algorithms, NB, SVC, RF, DT, and GB, to obtain an accuracy of 99.30% on the ESD dataset, using all 34 features. Our methodology outperformed the four state-of-the-art methods.

Table 3. Comparison with State-of-the-art performance

Methods	Accuracy	Precision	Specificity	Recall	F1
A hybrid of multiclass Support Vector Machine with Bayesian Optimization [51]	99.07%	-	-	-	-
Bagging, Boosting, and Stacking (SVM, KNN, DT, NB, and MLP) [50]	92.9%	0.86	0.97	0.85	0.84
LR, SVM, and KNN [41]	99%				
RF	98.21%	0.99	-	0.99	0.99
XGB[52]	97.85%				
Proposed - Stacking Methods (NB, SVC, RF, DT, GB)	99.30%	1.00	1.00	0.96	0.97

5.1. Deployment

The model was deployed on a web application – ESD App, shown in Figure 7. The front end was created using HTML, which accepts user input value and sends it to the back end, which was programmed with Python. In this application, users can easily input various symptom levels using a simple scale from 0 to 3, indicating the severity of each symptom except for family history and age, which are entered differently. A score of 0 means no symptoms, while 1 or 2 suggests moderate symptoms, and 3 indicates severe symptoms. Family history is indicated with a binary choice of 0 or 1, and the patient's age is entered into a designated field. Once all features are imputed, users will click "Classify Erythemato Squamous diseases" to begin the classification process. The result will display the disease type determined by the inputted features, giving users useful information about the categorization outcome.

6. Conclusion and Future Work

Skin disease has been the most common health problem many countries face due to the long exposure to the sun. Skin disease must be diagnosed and treated early to avoid severe consequences. Any wrong diagnoses can also harm the patient's health. This study uses clinical and histopathological attributes to classify ESD, such as Lichen planus, Pityriasis rosea, Pityriasis rubra pilaris, psoriasis, psoriasis, seborrheic dermatitis, and chronic dermatitis. We employed a stacking ensemble technique to classify skin diseases using five base models: NB, SVC, DT, RF, and GB. The study was performed in several phases; the initial phase involved the collection of the dataset from UCI, followed by the preprocessing of the data, and next was the extraction of features and normalization. The classification was done, and the effectiveness of the techniques was evaluated using accuracy, precision, specificity, recall, and f1 score. Based on the findings, the base models achieved different accuracies in classifying the 6 different skin diseases, and SVC achieved the highest with 98.61%. However, the stacking ensemble techniques significantly improved the overall classification performance compared to individual base models with an accuracy of 99.30%. By employing the predictions provided by individual models as input for a meta-classifier, the meta-classifier can learn from these predictions, improving its capacity to produce more accurate results.

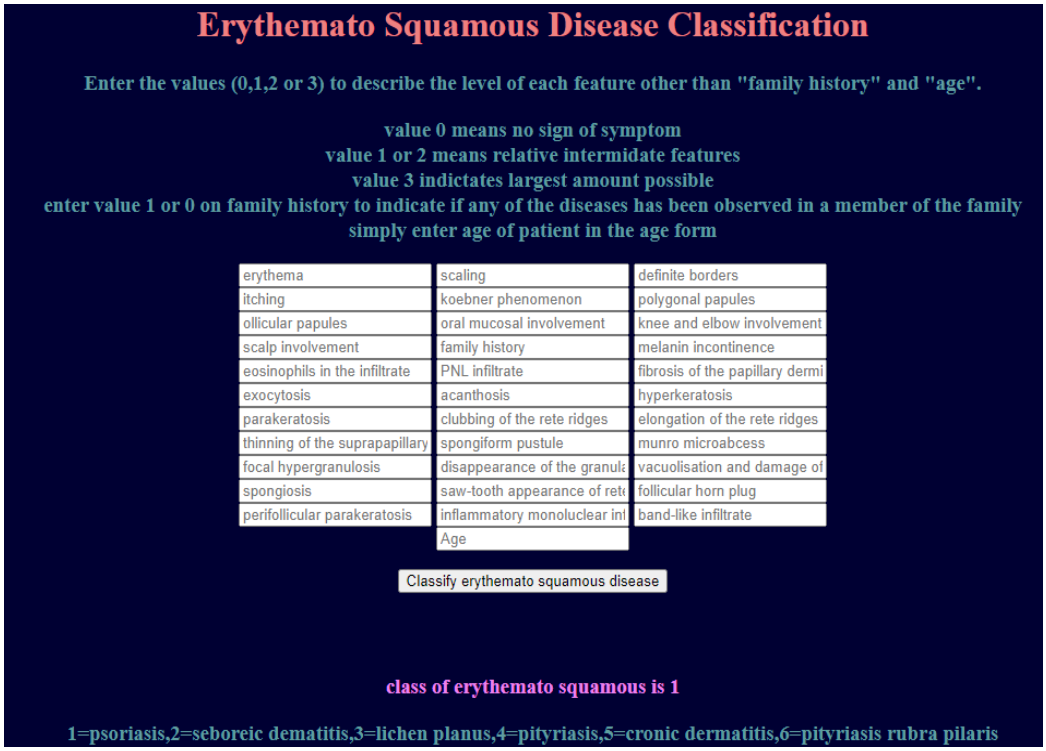

Figure 7. Web application for the classification of Erythemato Squamous Disease

Future research should focus on increasing the dataset size to improve the model's ability to generalize well. While ML can identify fundamental features and patterns from data, deep learning can more effectively capture complex correlations in dermatological images. Deep learning models, such as CNN, consist of numerous layers of interconnected nodes that learn hierarchical data representations independently. This hierarchical feature learning allows deep learning models to detect fine features and nuances in skin lesions, resulting in more accurate and robust disease classification. Future deep-learning studies can improve skin disease classification, diagnosis, and patient care.

Author Contributions: Conceptualization: Oluwayemisi Jaiyeoba and Emeka Ogbuju; methodology, Oluwayemisi Jaiyeoba, Emeka Ogbuju, and Francisca Oladipo.; software: Oluwayemisi, Jaiyeoba, and, Owolabi Temitope Yomi.; validation: Oluwayemisi Jaiyeoba, Owolabi Temitope Yomi, and Emeka Ogbuju,.; formal analysis: Oluwayemisi Jaiyeoba, Emeka Ogbuju, and Francisca Oladipo.; investigation: All,; resources: Oluwayemisi Jaiyeoba Emeka Ogbuju; writing—original draft preparation: All.; writing—review and editing: All.; visualization: Oluwayemisi Jaiyeoba and, Owolabi Temitope Yomi.; supervision: Emeka Ogbuju and Francisca Oladipo.

Funding: This research received no external funding.

Acknowledgments: We sincerely thank Jahlics Technology for providing a pleasant environment in which to conduct this literature study.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] D. J. Tobin, "Biochemistry of human skin—our brain on the outside," *Chem. Soc. Rev.*, vol. 35, no. 1, pp. 52–67, 2006, doi: 10.1039/B505793K.

[2] O. Arda, N. Göksügür, and Y. Tüzün, "Basic histological structure and functions of facial skin," *Clin. Dermatol.*, vol. 32, no. 1, pp. 3–13, Jan. 2014, doi: 10.1016/j.clindermatol.2013.05.021.

[3] A. Lefèvre-Utile, C. Braun, M. Haftek, and F. Aubin, "Five functional aspects of the epidermal barrier," *Int. J. Mol. Sci.*, vol. 22, no. 21, pp. 1–14, 2021, doi: 10.3390/ijms222111676.

[4] D. R. Bergfelt, "Anatomy and Physiology of the Mare," in *Equine Breeding Management and Artificial Insemination*, Elsevier, 2009, pp. 113–131. doi: 10.1016/B978-1-4160-5234-0.00011-8.

[5] R. Wong, S. Geyer, W. Weninger, J. Guimberteau, and J. K. Wong, "The dynamic anatomy and patterning of skin," *Exp. Dermatol.*, vol. 25, no. 2, pp. 92–98, Feb. 2016, doi: 10.1111/exd.12832.
[6] M. Maranduca *et al.*, "Skin - a vast organ with immunological function (Review)," *Exp. Ther. Med.*, pp. 18–23, Mar. 2020, doi: 10.3892/etm.2020.8619.
[7] M. S. Pitchiah and T. Rajamanickam, "Efficient Feature Based Melanoma Skin Image Classification Using Machine Learning Approaches," *Trait. du Signal*, vol. 39, no. 5, pp. 1663–1671, Nov. 2022, doi: 10.18280/ts.390524.
[8] N. Hameed, A. M. Shabut, M. K. Ghosh, and M. A. Hossain, "Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques," *Expert Syst. Appl.*, vol. 141, p. 112961, Mar. 2020, doi: 10.1016/j.eswa.2019.112961.
[9] Z. Li, Q. Ji, X. Yang, Y. Zhou, and S. Zhi, "An Identification Method of Feature Interpretation for Melanoma Using Machine Learning," *Appl. Sci.*, vol. 13, no. 18, p. 10076, Sep. 2023, doi: 10.3390/app131810076.
[10] B. Lahijanian, F. V. Farahani, and M. H. F. Zarandi, "A new multiple classifier system for diagnosis of erythemato-squamous diseases based on rough set feature selection," in *2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Jul. 2016, pp. 2309–2316. doi: 10.1109/FUZZ-IEEE.2016.7737981.
[11] N. Badrinath, G. Gopinath, K. S. Ravichandran, and R. G. Soundhar, "Estimation of automatic detection of erythemato-squamous diseases through AdaBoost and its hybrid classifiers," *Artif. Intell. Rev.*, vol. 45, no. 4, pp. 471–488, Apr. 2016, doi: 10.1007/s10462-015-9436-8.
[12] S. Badillo *et al.*, "An Introduction to Machine Learning," *Clin. Pharmacol. Ther.*, vol. 107, no. 4, pp. 871–885, Apr. 2020, doi: 10.1002/cpt.1796.
[13] A. N. A. Thohari, L. Triyono, I. Hestiningsih, B. Suyanto, and A. Yobioktobera, "Performance Evaluation of Pre-Trained Convolutional Neural Network Model for Skin Disease Classification," *JUITA J. Inform.*, vol. 10, no. 1, p. 9, May 2022, doi: 10.30595/juita.v10i1.12041.
[14] B. Mahesh, "Machine Learning Algorithms - A Review," *Int. J. Sci. Res.*, vol. 9, no. 1, pp. 381–386, 2020, doi: 10.21275/ART20203995.
[15] G. Brown, "Ensemblelearning," pp. 1–24, 2010, [Online]. Available: papers://5e3e5e59-48a2-47c1-b6b1-a778137d3ec1/Paper/p1908.
[16] O. Sagi and L. Rokach, "Ensemble learning: A survey," *WIREs Data Min. Knowl. Discov.*, vol. 8, no. 4, pp. 1–18, Jul. 2018, doi: 10.1002/widm.1249.
[17] I. D. Mienye and Y. Sun, "A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects," *IEEE Access*, vol. 10, no. August, pp. 99129–99149, 2022, doi: 10.1109/ACCESS.2022.3207287.
[18] S. Bechelli and J. Delhommelle, "Machine Learning and Deep Learning Algorithms for Skin Cancer Classification from Dermoscopic Images," *Bioengineering*, vol. 9, no. 3, p. 97, Feb. 2022, doi: 10.3390/bioengineering9030097.
[19] K. Meena, N. N. K. Veni, B. S. Deepapriya, P. A. H. Vardhini, B. J. D. Kalyani, and L. Sharmila, "A novel method for prediction of skin disease through supervised classification techniques," *Soft Comput.*, vol. 26, no. 19, pp. 10527–10533, Oct. 2022, doi: 10.1007/s00500-022-07435-8.
[20] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 3, p. 160, May 2021, doi: 10.1007/s42979-021-00592-x.
[21] A. Lefèvre-Utile, C. Braun, M. Haftek, and F. Aubin, "Five Functional Aspects of the Epidermal Barrier," *Int. J. Mol. Sci.*, vol. 22, no. 21, p. 11676, Oct. 2021, doi: 10.3390/ijms222111676.
[22] A. Ahmed, H. Ahmad, M. Khurshid, and K. Abid, "Classification of Skin Disease using Machine Learning," *VFAST Trans. Softw. Eng.*, vol. 11, no. 1, pp. 109–122, Mar. 2023, doi: 10.21015/vtse.v11i1.1204.
[23] Z. Li, Q. Ji, X. Yang, Y. Zhou, and S. Zhi, "An Identification Method of Feature Interpretation for Melanoma Using Machine Learning," *Appl. Sci.*, vol. 13, no. 18, p. 10076, Sep. 2023, doi: 10.3390/app131810076.
[24] Z. Li, Q. Ji, X. Yang, Y. Zhou, and S. Zhi, "An Identification Method of Feature Interpretation for Melanoma Using Machine Learning," *Appl. Sci.*, vol. 13, no. 18, 2023, doi: 10.3390/app131810076.
[25] A. Murugan, S. A. H. Nair, and K. P. S. Kumar, "Detection of Skin Cancer Using SVM, Random Forest and kNN Classifiers," *J. Med. Syst.*, vol. 43, no. 8, p. 269, Aug. 2019, doi: 10.1007/s10916-019-1400-8.
[26] M. Ahammed, M. Al Mamun, and M. S. Uddin, "A machine learning approach for skin disease detection and classification using image segmentation," *Healthc. Anal.*, vol. 2, no. October, p. 100122, Nov. 2022, doi: 10.1016/j.health.2022.100122.
[27] M. Kalaiyarivu and N. J. Nalini, "Skin Disease Classification Using Combined Machine Learning And Deep Learning Models," *J. Pharm. Negat. Results*, vol. 13, no. 9, pp. 3248–3256, 2022, doi: 10.47750/pnr.2022.13.S09.406.
[28] M. S. Pitchiah and T. Rajamanickam, "Efficient Feature Based Melanoma Skin Image Classification Using Machine Learning Approaches," *Trait. du Signal*, vol. 39, no. 5, pp. 1663–1671, 2022, doi: 10.18280/ts.390524.
[29] A. Javaid, M. Sadiq, and F. Akram, "Skin Cancer Classification Using Image Processing and Machine Learning," in *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST)*, Jan. 2021, pp. 439–444. doi: 10.1109/IBCAST51254.2021.9393198.
[30] O. K. Pal, "Skin Disease Classification: A Comparative Analysis of K-Nearest Neighbors (KNN) and Random Forest Algorithm," in *2021 International Conference on Electronics, Communications and Information Technology (ICECIT)*, Sep. 2021, pp. 1–5. doi: 10.1109/ICECIT54077.2021.9641120.
[31] B. Gaye, D. Zhang, and A. Wulamu, "Improvement of Support Vector Machine Algorithm in Big Data Background," *Math. Probl. Eng.*, vol. 2021, pp. 1–9, Jun. 2021, doi: 10.1155/2021/5594899.
[32] S. R. Guha and S. M. Rafizul Haque, "Performance Comparison of Machine Learning-Based Classification of Skin Diseases from Skin Lesion Images," in *Lecture Notes in Electrical Engineering*, vol. 637, Springer Singapore, 2020, pp. 15–25. doi: 10.1007/978-981-15-2612-1_2.

[33] I. A. Alfi, M. M. Rahman, M. Shorfuzzaman, and A. Nazir, "A Non-Invasive Interpretable Diagnosis of Melanoma Skin Cancer Using Deep Learning and Ensemble Stacking of Machine Learning Models," *Diagnostics*, vol. 12, no. 3, p. 726, Mar. 2022, doi: 10.3390/diagnostics12030726.

[34] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 3, p. 160, May 2021, doi: 10.1007/s42979-021-00592-x.

[35] H. Altay Gü Venir, G. Güls¸en Demiröz, and I. : Lter, "Learning differential diagnosis of erythemato-squamous diseases using voting feature intervals," *Artif. Intell. Med.*, vol. 13, pp. 147–165, 1998.

[36] I. D. Mienye and Y. Sun, "A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects," *IEEE Access*, vol. 10, no. 3, pp. 99129–99149, 2022, doi: 10.1109/ACCESS.2022.3207287.

[37] S. García, J. Luengo, and F. Herrera, "Instance Selection," in *Intelligent Systems Reference Library*, vol. 72, 2015, pp. 195–243. doi: 10.1007/978-3-319-10247-4_8.

[38] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 3, pp. 1–21, 2021, doi: 10.1007/s42979-021-00592-x.

[39] A. R. Ismail, A. Zainul Fanani, G. F. Shidik, and Muljono, "Implementation Of Naive Bayes Algorithm With Particle Swarm Optimization In Classification Of Dress Recommendation," in *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*, Sep. 2020, pp. 174–178. doi: 10.1109/iSemantic50169.2020.9234293.

[40] B. Gaye, D. Zhang, and A. Wulamu, "Improvement of Support Vector Machine Algorithm in Big Data Background," *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/5594899.

[41] J. Singh Kushwah, A. Kumar, S. Patel, R. Soni, A. Gawande, and S. Gupta, "Comparative study of regressor and classifier with decision tree using modern tools," *Mater. Today Proc.*, vol. 56, no. xxxx, pp. 3571–3576, 2022, doi: 10.1016/j.matpr.2021.11.635.

[42] M. Bansal, A. Goyal, and A. Choudhary, "A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning," *Decis. Anal. J.*, vol. 3, no. May, p. 100071, Jun. 2022, doi: 10.1016/j.dajour.2022.100071.

[43] B. Petracchi, E. Torti, E. Marenzi, and F. Leporati, "Acceleration of Hyperspectral Skin Cancer Image Classification through Parallel Machine-Learning Methods," *Sensors*, vol. 24, no. 5, p. 1399, Feb. 2024, doi: 10.3390/s24051399.

[44] A. Q. Md, S. Kulkarni, C. J. Joshua, T. Vaichole, S. Mohan, and C. Iwendi, "Enhanced Preprocessing Approach Using Ensemble Machine Learning Algorithms for Detecting Liver Disease," *Biomedicines*, vol. 11, no. 2, p. 581, Feb. 2023, doi: 10.3390/biomedicines11020581.

[45] Y. N. Fu'adah, N. C. Pratiwi, M. A. Pramudito, and N. Ibrahim, "Convolutional Neural Network (CNN) for Automatic Skin Cancer Classification System," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 982, no. 1, p. 012005, Dec. 2020, doi: 10.1088/1757-899X/982/1/012005.

[46] K. Shehzad *et al.*, "A Deep-Ensemble-Learning-Based Approach for Skin Cancer Diagnosis," *Electronics*, vol. 12, no. 6, p. 1342, Mar. 2023, doi: 10.3390/electronics12061342.

[47] M. Q. Khan *et al.*, "Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer," *IEEE Access*, vol. 7, pp. 90132–90144, 2019, doi: 10.1109/ACCESS.2019.2926837.

[48] H. Wang, Q. Qi, W. Sun, X. Li, and C. Yao, "Classification of clinical skin lesions with double-branch networks," *Front. Med.*, vol. 10, no. 1, Jun. 2023, doi: 10.3389/fmed.2023.1114362.

[49] A. N. A. Thohari, L. Triyono, I. Hestiningsih, B. Suyanto, and A. Yobioktobera, "Performance Evaluation of Pre-Trained Convolutional Neural Network Model for Skin Disease Classification," *JUITA J. Inform.*, vol. 10, no. 1, p. 9, 2022, doi: 10.30595/juita.v10i1.12041.

[50] O. Obe, O. Owolafe, E. Charles Igodan, and A. Favour-Bethy Thompson, "Erythemato-Squamous Disease Prediction using Ensemble Multi-Feature Selection Approach," *Int. J. Comput. Sci. Inf. Secur.*, vol. 20, no. June, pp. 95–106, 2022.

[51] A. M. Elsayad, A. M. Nassef, and M. Al-Dhaifallah, "Bayesian optimization of multiclass SVM for efficient diagnosis of erythemato-squamous diseases," *Biomed. Signal Process. Control*, vol. 71, p. 103223, Jan. 2022, doi: 10.1016/j.bspc.2021.103223.

[52] A. S. Rathore, S. K. Arjaria, M. Gupta, G. Chaubey, A. K. Mishra, and V. Rajpoot, "Erythemato-Squamous Diseases Prediction and Interpretation Using Explainable AI," *IETE J. Res.*, pp. 1–20, Sep. 2022, doi: 10.1080/03772063.2022.2114953.