Vol. 06, No. 4 (2024) 1741-1750, doi: 10.24874/PES.SI.24.03.009

Proceedings on Engineering Sciences

www.pesjournal.net

MACHINE LEARNING ALGORITHMS FOR ERYTHEMATO-SQUAMOUS DISEASE CLASSIFICATION: FEATURE RANKINGS AND PERFORMANCE ANALYSIS

M. Venkata Subbarao[1]
S. Lakshmi Praveena
T. Naga Sharmila
Praveen Kumar Nalli
G. Challa Ram
D. Ramesh Varma
Srinivasarao Alluri
Tabassum Nazeer Syed

Received 19.10.2023.
Received in revised form 25.11.2023.
Accepted 31.12.2023.
UDC – 004.85

Keywords:

Dermatology, Psoriasis, Seborrheic Dermatitis, Lichen Planu, Pityriasis Rosea, Chronic Dermatitis, Pityriasis Rubra Pilari, ML Classifiers, Kruskal-Wallis Feature Ranking

A B S T R A C T

Erythemato-squamous diseases (ESDs), also known as erythrodermas, are a group of dermatological disorders characterized by both redness (erythema) and scaling (squamous) of the skin. These conditions can have various causes and implications. The implications of ESDs vary depending on the specific condition and its severity. While some may cause mild symptoms and have minimal impact on daily life, others can be chronic, recurrent, and significantly affect a person's physical and emotional well-being. Treatment options for these conditions may include topical medications, oral medications, phototherapy, and lifestyle modifications. In this paper, state of art machine learning (ML) algorithms is implemented for classification of ESD. To classify the disease a set of 11 clinical features and 23 histopathological features are considered. The performance of the ML classifiers is analyzed with individual sets of features and combination of both. Further, the performance of the ML classifiers is analyzed at different training rates to know the superior classifier for ESD classification. Furthermore, the study is extended to investigate the effectiveness of the Kruskal-Wallis algorithm in ranking the importance of features in the dataset used for disease classification. An investigation depicts that Ensemble and SVM classifiers outperformed the other ML classifiers in terms of accuracy and F1-score.

[1] Corresponding author: M. Venkata Subbarao
Email: mandava.decs@gmail.com

1. INTRODUCTION

Erythrodermas, referred to as ESDs, encompass a group of dermatological conditions characterized by both skin redness and scaling. These disorders can arise from various underlying causes and have wide-ranging implications (Banu & Toacşe, 2013). Some common ESDs and their key features are as follows.

Psoriasis: Psoriasis is a chronic autoimmune disease that causes rapid skin cell turnover, resulting in the formation of thick, scaly patches of skin. The exact cause of psoriasis is unknown, but it is believed to involve a combination of genetic and environmental factors. The disease can have significant implications on a person's quality of life, causing physical discomfort, itching, pain, and emotional distress (Singh et al., 2022).

Seborrheic dermatitis: It is a common inflammatory skin condition that mainly affects areas rich in sebaceous glands, such as the scalp, face, and chest. It is characterized by redness, greasy or flaky scales, and itching. While the exact cause is unclear, factors like yeast overgrowth, genetic predisposition, and certain neurological conditions may contribute to its development. Seborrheic dermatitis can be chronic and recurrent but is not considered a serious medical condition (Basu et al., 2015).

Eczema (atopic dermatitis): It is a chronic inflammatory skin condition that is often associated with allergies and immune system dysfunction. It leads to red, itchy, and inflamed skin, which may also develop scaling or weeping blisters. The exact cause of eczema is not fully understood, but genetic factors, environmental triggers, and abnormalities in the skin barrier function are believed to play a role. Eczema can have a significant impact on a person's quality of life, causing discomfort, disrupted sleep, and psychological distress (Badrinath et al., 2020).

Pityriasis rosea: It is a self-limiting skin rash characterized by the appearance of a larger "herald patch" followed by numerous smaller scaly patches. The exact cause of pityriasis rosea is unknown, but it is thought to be associated with viral infections, particularly human herpesvirus 6 (HHV-6) or human herpesvirus 7 (HHV-7). Pityriasis rosea typically resolves on its own within several weeks to months and does not generally have long-term implications.

Lichen planus: It is a chronic inflammatory disorder that affects the skin, mucous membranes, hair, and nails. It is characterized by small, itchy, flat-topped, polygonal bumps that can develop a lacy white pattern and may cause erosions or ulcers. The exact cause of lichen planus is unknown, but it is believed to involve an abnormal immune response. Lichen planus can be chronic and may cause discomfort or pain in severe cases.

While some diseases may cause mild symptoms and have minimal impact on daily life, others can be chronic, and recurrent, and significantly affect a person's physical and emotional well-being. Treatment options for these conditions may include topical medications, oral medications, phototherapy, and lifestyle modifications. If you suspect you have an ESD or any other medical condition, it is important to consult with a healthcare professional for an accurate diagnosis and appropriate management (Xie et al., 2012). These conditions often share similar symptoms, making their classification and diagnosis challenging. The precise symptom may differ in accordance with the underlying disease (Ravichandran et al., 2014).

Here are some general symptoms associated with these conditions:

Redness (Erythema): The affected skin usually appears red, inflamed, or irritated.

Scaling (Squamous): The skin may develop flakes, scales, or patches of thickened skin that can be white, silver, or greyish.

Itching: Itchiness is a common symptom in ESDs. The severity can range from mild to severe and may significantly affect the quality of life.

Rash: Rashes may be present, and their appearance can vary depending on the specific disease. They may be well-defined or more diffuse.

Dryness: The affected skin may become dry and rough, leading to discomfort and a tendency to crack or fissure.

Burning or stinging sensation: Some individuals may experience a burning or stinging sensation in the affected areas.

Flares and remissions: ESDs often exhibit a cyclic pattern of flares and remissions, with symptoms worsening during flare-ups and improving during periods of remission.

Location-specific symptoms: The distribution and location of symptoms may vary depending on the specific condition. For example, psoriasis commonly affects the scalp, elbows, knees, and lower back, while seborrheic dermatitis often involves the scalp, face (especially the eyebrows and nasolabial folds), and chest.

Various classification systems have been proposed to aid in the identification and management of these diseases. Here is a general overview of the classification of ESD (Polat & Güneş, 2009). The traditional classification system divides ESDs into two main categories based on the presence or absence of fungal infection. Non-infectious ESDs category includes conditions such as psoriasis, seborrheic dermatitis, lichen planus, pityriasis

rosea, and cutaneous lupus erythematosus. The infectious ESDs category includes fungal infections like tinea corporis (ringworm), tinea versicolor (pityriasis versicolor), and candidiasis. Over time, researchers have proposed modifications to the traditional classification system to incorporate new knowledge and improve diagnostic accuracy (Polat & Güneş, 2006). These revisions often consider additional factors such as histopathology, immunopathology, and genetic markers. For example, the expansion of non-infectious ESDs may include subcategories based on distinctive clinical and histological features, as well as response to treatment. This can aid in the differentiation of conditions such as psoriasis, seborrheic dermatitis, and lichen planus. Given the overlapping clinical features of ESDs, an accurate differential diagnosis is crucial. Dermatologists rely on a combination of clinical evaluation, patient history, physical examination, and, in some cases, laboratory tests to differentiate between these conditions (Abdi & Giveki, 2013). Features such as lesion morphology, distribution, and associated symptoms are carefully assessed to arrive at a proper diagnosis.

The purpose of this study is to develop and implement machine learning algorithms for accurately classifying ESDs. The primary aim is to improve the diagnostic process and enhance the accuracy of disease classification. Further, the study is extended to investigate the effectiveness of the Kruskal-Wallis algorithm in ranking the importance of features in the dataset used for disease classification. The paper could explore how this ranking technique helps in selecting the most relevant features for improved classification performance. This objective could focus on demonstrating how the reduction of irrelevant features contributes to enhanced model performance and efficiency.

The paper is structured as detailed below. Section II outlines the experimental setup, including data preprocessing, feature extraction, feature ranking with the Kruskal-Wallis algorithm, model training, and evaluation procedures. The methodology of the classifiers is presented in section III. Section IV presents the results of the experiments in a clear and interpretable manner i.e., the comparative performance of different machine learning algorithms and the impact of feature ranking on classification accuracy. Section V discusses the implications of the findings in the context of ESD classification and the broader field of medical diagnostics.

2. LITERATURE & FREAMEWORK

ML algorithms can be used to analyze clinical data, such as patient symptoms, medical history, and visual representations of skin lesions, to aid in the classification and diagnosis of ESDs.

Investigation of use of Convolutional Neural Networks (CNNs) in classifying ESDs is presented (Tekin, 2014). Using a dataset of 10,000 high-resolution images of different ESD cases, the authors trained a CNN model to detect and categorize these diseases. The methodology benefited from data augmentation and transfer learning, leveraging pretrained models for enhanced accuracy.

The model achieved a commendable 95% accuracy on the test set. However, the paper faced limitations in terms of a potentially non-diverse dataset predominantly from a single ethnic group, possibly limiting its generalizability across diverse populations. It is compared traditional ML techniques, specifically Support Vector Machines (SVM), with newer deep learning methods in classifying ESDs (Kaushik et al., 2023). The study utilized a dataset containing both clinical data and patient history for 5,000 cases. While the SVM demonstrated an accuracy of 85%, the deep learning model surpassed this with a 92% success rate. One significant advantage of this study was its incorporation of non-image data, providing a holistic approach. However, the research's limitation lay in the relatively smaller dataset, which might not capture all the nuances of ESD manifestations.

A hybrid model that combines the strengths of traditional machine learning (Random Forest) and deep learning is presented (Putatunda, 2020). The methodology utilized a diverse dataset of 15,000 entries, blending clinical images and patient medical histories. Through feature integration, the hybrid model achieved an impressive accuracy of 95%. The robustness of this research was its comprehensive dataset, which encompassed varied ethnic backgrounds and age groups. However, a limitation noted was the increased computational cost and complexity due to the hybrid nature of the model, potentially hindering its deployment in real-time or resource-limited settings.

In (Li et al., 2022), it delves into the potential of Recurrent Neural Networks (RNNs) to trace the progression of Erythemato-Squamous Diseases over time. Utilizing a time-series dataset of 7,000 patients, where images were taken at various stages of disease development, the authors established an RNN model that aimed to predict disease progression patterns. The methodology notably excelled in handling sequential data, with the model achieving 89% accuracy in predicting future disease states. Nevertheless, the paper's limitation was evident in its reliance on well-documented and time-stamped data, which may be challenging to procure in real-world scenarios.

In (Spolaôr et al., 2023), it explored the viability of using transfer learning techniques for ESD classification. Leveraging popular pre-trained networks, the authors fine-tuned these architectures using a modest dataset of 3,000 ESD images. The methodology

highlighted the practicality and cost-effectiveness of transfer learning, yielding a respectable 90% accuracy rate without necessitating extensive training data or computational power. The primary advantage was the resource efficiency of the model, making it accessible for smaller clinics or research setups. However, the study's main limitation is its dependency on pre-existing architectures, which might not be optimally designed for dermatological image nuances.

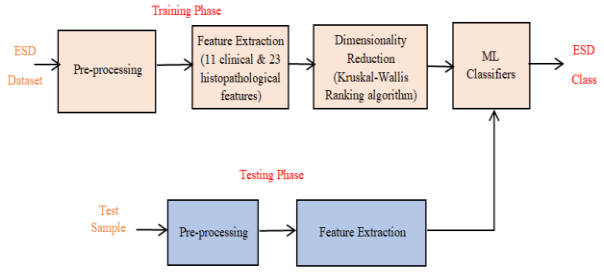

Figure 1. Proposed Framework

Figure 1 represents the framework of the proposed ESD classification. The steps involved in the framework as follows.

Data collection: Relevant clinical data and images of skin lesions are collected from patients diagnosed with different ESDs. The data may include information about symptoms, patient demographics, and any other factors that could contribute to the classification.

Preprocessing: Removal of noise, normalizing the features and handle missing values are performed in this stage. This step is crucial to ensure the quality and consistency of the data.

Feature extraction: Relevant features are extracted from the preprocessed data. In the case of ESDs, features could include characteristics of skin lesions, such as color, texture, shape, and distribution.

Training data preparation: The dataset gets split into two parts: a training set and a testing set. We use the training set to teach the machine learning model, and the testing set to see how well it works.

Model training: Many ML methods, including decision trees, SVMs, K-Nearest Neighbors (KNN), and Ensemble classifiers, can be taught using the features we found. These models understand how the features are connected to the ESD labels, learning the patterns and relationships between them.

Model evaluation: In this phase, evaluating the model using the testing dataset to assess its classification performance. Evaluation metrics like accuracy, precision, recall, and F1-score are commonly used to measure the model's effectiveness.

3. MACHINE LEARNING CLASSIFIERS

3.1 Decision Trees (DT)

Decision tree classifiers are widely used in ML for both regression and classification tasks. The DT classifier is constructed using a training dataset, which consists of labeled examples. The algorithm uses a recursive process called recursive binary splitting to partition the data based on the feature values. The goal is to find the best feature and the best split point that maximizes the information gain or Gini impurity (a measure of how well the feature splits the data based on class labels) (Danjuma & Osofisan, 2014). The splitting process continues recursively until a stopping criterion is met. This could be a maximum depth limit, a minimum number of samples required to split, or when all samples belong to the same class.

3.2 SVM Classifiers

SVMs are powerful machine learning models used for both classification and regression tasks. They operate by finding an optimal hyperplane that separates data points into different classes or predicts continuous values (Subbarao et al., 2023).

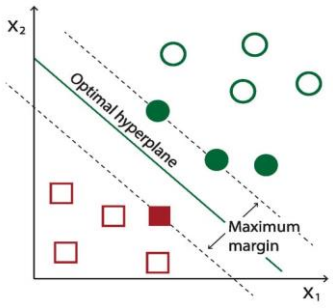

Figure 2. Hyperplane in SVM Classification

Figure 2 depicts the hyperplane representation to distinguish two classes. One of the key features of SVMs is the use of kernels. Kernels allow SVMs to operate efficiently in high-dimensional feature spaces without explicitly calculating the coordinates of the data points. The linear kernel is the simplest and most commonly used kernel. It represents a linear decision boundary in the input space. It works well when the data is linearly separable. The polynomial kernel maps the original features into a higher-dimensional space using polynomial functions. It can capture non-linear relationships between the data points. The kernel function is defined as

$$K(x,y) = (x^T y + c)^d \quad (1)$$

where, c is a constant and d is the degree of the polynomial.

The RBF kernel is popular in SVMs because of its flexibility. It defines a similarity measure between data points based on their Euclidean distance in the feature space. The kernel function is defined as

$$K(x,y) = exp^{(-\gamma * ||x-y||\hat{}2)} \quad (2)$$

where, γ is a hyperparameter that controls the influence of each training example.

The sigmoid kernel maps the features into a higher-dimensional space using a sigmoid function. It can capture non-linear relationships, but it is generally less commonly used compared to other kernels. The kernel function is defined as

$$K(x, y) = tanh(\alpha x^T y + c) \quad (3)$$

SVMs also allow the use of custom kernels, where you can define your own similarity measure. This is useful when the data has a specific structure or when domain knowledge suggests a particular kernel.

3.3 KNN Classifiers

K-Nearest Neighbors (KNN) is a non-parametric and instance-based algorithm, meaning that it doesn't make any assumptions about the underlying data distribution and instead relies on the proximity of training examples to make predictions. The choice of the K value is crucial. A small K may lead to over fitting, where the model becomes too sensitive to noise, while a large value of *k* may result in under fitting, where the model becomes too biased and fails to capture local patterns. It is typically chosen through experimentation and cross-validation. The distance metric determines how distances are calculated between data points. Euclidean distance is commonly used for continuous features, while other metrics like Manhattan distance or Hamming distance may be used for specific types of data.

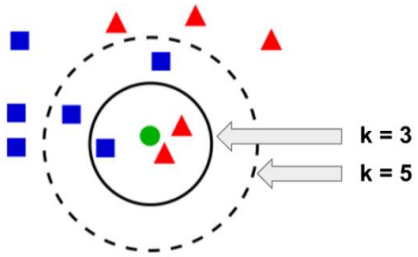

Figure 3. KNN Classification

Figure 3 represents the classification with different values of *k*. Further, it's often recommended to scale the features to ensure that they contribute equally to the distance calculation. This is important when features have different scales or units. Common scaling techniques include normalization or standardization (subtracting mean and dividing by standard deviation). KNN classifiers are complex, especially for large datasets. Techniques like KD-trees or ball trees can be used to optimize the nearest neighbor search process and improve efficiency.

3.4 NN Classifiers

Classification with Neural Network (NN) classifiers involves training a NN model to classify input data into different classes. NNs are composed of interconnected nodes (neurons) organized in layers, and they learn to extract relevant features from the input data through a training process. Figure 4 depicts basic architecture of NN classifier.

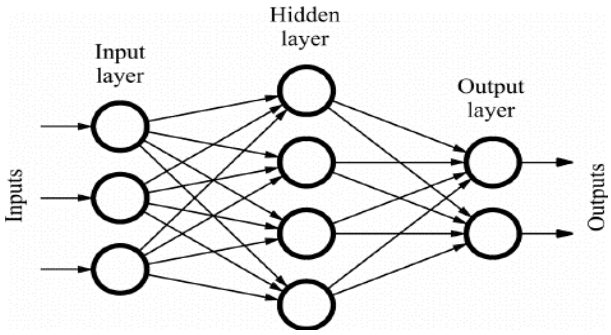

Figure 4. NN Classifier

3.5 Ensemble Classifiers

Classification with ensemble classifiers involves combining multiple individual classifiers to make more accurate predictions than any single classifier alone (Subbarao et al., 2023). Ensemble classifiers are known for their ability to improve predictive performance, increase robustness, and reduce over fitting.

3.5 Kruskal-Wallis Feature Ranking Algorithm

Kruskal-Wallis is a non-parametric statistical test used in machine learning for feature ranking when dealing with non-normally distributed data or categorical variables. It extends the one-way ANOVA test and evaluates whether there are significant differences in the distributions of a continuous target variable across different groups or levels of a categorical feature. In contrast to ANOVA, Kruskal-Wallis doesn't assume data normality, making it suitable for a wider range of datasets. It operates by ranking the data values and comparing the average ranks between groups. Kruskal-Wallis is useful when the target variable violates assumptions of parametric tests or when the data exhibits non-linear relationships. By assessing feature significance based on target variable distribution differences among groups, Kruskal-Wallis helps in selecting relevant features for classification or regression tasks, improving model interpretability and generalization.

4. RESULTS & DISCUSSIONS

To identify ESDs using ML classifiers a dataset of size 358 samples that includes 11 clinical features (e.g., age, gender, symptoms) and 23 histopathological features (e.g., cell counts, tissue characteristics) for patients with ESDs. The simulated dataset information is shown in Table 1.

The simulations are carried with two different training rates 80% and 70%. In each case 10% data has taken for the validation. The test set for evaluating the models are 20% and 30% respectively. The training set is used to train the models, the validation set is used for hyper parameter tuning, and the test set is reserved for final evaluation. MATLAB software is used for the simulation.

Table 1. Dataset Information

S.No	Skin Diseases Class	Quantity
1	Psoriasis	111
2	Pityriasis rubra pilaris	20
3	Chronic dermatitis	48
4	Lichen planus	71
5	Pityriasis rosea	48
6	Seborrheic dermatitis	60
	Total	358

Table 2 and 3 depicts the statistical information of the clinical and histopathological features for the input dataset respectively.

Table 2. Statistical Information of the clinical features

S.No	Clinical Features	Max.	Min.	Mean	Median
1	Erythema	3	0	2.078212	2
2	Scaling	3	0	1.807263	2
3	Definite borders	3	0	1.569832	2
4	Scalp involvement	3	0	0.530726	0
5	Oral Mucosal involvement	3	0	0.379888	0
6	Koebner phenomenon	3	0	0.636872	0
7	Knee and elbow involveme	3	0	0.622905	0
8	Follicular papules	3	0	0.170391	0
9	Polygonal papules	3	0	0.449721	0
10	itching	3	0	1.354749	1
11	Family history	1	0	0.122905	0

Table 3. Histopathological Features

S.No	Clinical Features	Max.	Min.	Mean	Median
1	Melanin incontinence	0.410615	0	3	0
2	Vacuolisation damage basal layer	0.460894	0	3	0
3	Eosinophils infiltrate	0.142458	0	2	0
4	PNL infiltrate	0.547486	0	3	0
5	Follicular horn plug	0.106145	0	3	0
6	Inflammatory mononuclear infiltrate	1.877095	2	3	0
7	acanthosis	1.955307	2	3	0
8	hyperkeratosis	0.511173	0	3	0
9	parakeratosis	1.287709	1	3	0
10	Clubbing rete ridges	0.662011	0	3	0
11	Elongation rete ridges	0.98324	0	3	0
12	Thinning suprapapillary epidermis	0.642458	0	3	0
13	Spongiform pustule	0.298883	0	3	0
14	Focal hypergranulosis	0.399441	0	3	0
15	Disappearance granular layer	0.47486	0	3	0
16	Fibrosis papillary dermis	0.315642	0	3	0
17	Spongiosis	0.949721	0	3	0
18	Saw tooth appearance retes	0.458101	0	3	0
19	exocytosis	1.368715	2	3	0
20	Perifollicular parakeratosis	0.117318	0	3	0
21	Munro microabcess	0.368715	0	3	0
22	Band like infiltrate	0.558659	0	3	0
23	Age	36.29609	35	75	0

Table 4 presents the performance of different ML algorithms with different feature sets. It presents the performance of the ML classifiers with 11 clinical features and 23 histo-pathological features and combined 34 features. The performance is also measured at 80% training rate and 70% training rate.

Table 4. Performance of Different ML classifiers with Different feature sets

Category	**Classifier**	**Performance with Clinical Features (11)**	**Performance with Histo-pathological Features (23)**	**Performance with all features (34)**	**Performance with Clinical Features (11)**	**Performance wit Histo-pathological Features (23)**	**Performance with all features (34)**
		% of Training - 80			%Training rate -70		
Decision Trees	Fine	81.7	93	95.9	80.6	89.7	94.5
	Medium	83.1	93	95.9	80.6	89.7	94.5
	Coarse	70.4	77.5	78.1	70.8	79.4	78
SVM	Linear	88.7	95.8	98.6	86.1	95.3	95.4
	Cubic	87.3	95.8	95.9	81.9	94.4	91.7
	Quadratic	87.3	94.4	97.3	81.9	95.3	95.4
	Coarse Gaussian	87.3	87.3	93.2	81.9	86	89.9
	Medium Gaussian	90.1	93	94.5	87.5	96.3	94.5
	Fine Gaussian	64.8	38	30.1	55.6	43	34.9
Ensemble	Subspace KNN	87.3	85.9	97.3	880.6	92.5	94.5
	Bagged Trees	88.7	94.4	97.3	86.1	93.5	97.2
	Boosted Trees	88.7	95.8	30.1	83.3	91.6	30.3
	RusBoosted Trees	90.1	95.8	100	81.9	95.3	100
	Subspace Discriminant	88.7	97.2	98.6	83.3	93.5	97.2
KNN	Fine	88.7	88.8	93.2	77.8	88.8	94.5
	Medium	87.3	95.8	94.5	81.9	92.5	94.5
	Coarse	80.3	85.9	83.6	62.5	76.6	77.1
	Cosine	87.3	94.4	94.5	83.3	93.5	93.6
	Cubic	87.3	97.2	94.5	80.6	90.7	93.6
	Weighted	88.7	97.2	94.5	84.7	89.7	93.6
Neural Networks	Narrow	84.5	95.8	94.5	86.1	90.7	93.6
	Medium	87.3	95.3	94.5	83.3	95.3	93.6
	Wide	85.9	97.2	93.2	84.7	94.4	93.6
	Bi layered	88.7	95.8	93.2	79.2	93.5	92.7
	Tri layered	85.9	94.4	93.2	86.1	91.6	90.8

Table 5. Performance metrics of ML classifiers with all 34 Features

Category	Classifier	Accuracy (%)	Precision	Recall	F1 Score	Accuracy (%)	Precision	Recall	F1 Score
		% of Training - 80				70			
Decision Trees	Fine	95.9	0.96	0.96	0.96	94.5	0.92	0.92	0.92
	Medium	95.9	0.96	0.96	0.96	94.5	0.92	0.92	0.92
	Coarse	78.1	0.63	0.56	0.60	78	0.64	0.56	0.59
SVM	Linear	98.6	0.98	0.99	0.98	95.4	0.95	0.96	0.95
	Cubic	95.9	0.95	0.96	0.96	91.7	0.86	0.93	0.89
	Quadratic	97.3	0.97	0.97	0.97	95.4	0.95	0.96	0.95
	Coarse Gaussian	93.2	0.86	0.95	0.90	89.9	0.80	0.93	0.86
	Medium Gaussian	94.5	0.91	0.96	0.93	94.5	0.92	0.95	0.93
	Fine Gaussian	30.1	0.16	0.05	0.07	34.9	0.21	0.38	0.27
Ensemble	Subspace KNN	97.3	0.97	0.95	0.95	94.5	0.947	0.92	0.933
	Bagged Trees	97.3	0.96	0.98	0.96	97.2	0.975	0.97	0.972
	Boosted Trees	30.1	0.05	0.05	0.05	30.3	0.16	0.05	0.076
	RusBoosted Trees	100	1	1	1	100	1	1	1
	Subspace Discriminant	98.6	0.98	0.99	0.98	97.2	0.97	0.96	0.964
KNN	Fine	93.2	0.945	0.94	0.942	94.5	0.937	0.954	0.945
	Medium	94.5	0.93	0.968	0.95	94.5	0.937	0.953	0.944
	Coarse	83.6	0.70	0.69	0.695	77.1	0.619	0.613	0.615
	Cosine	94.5	0.93	0.968	0.948	93.6	0.926	0.943	0.934
	Cubic	94.5	0.93	0.968	0.948	93.6	0.926	0.943	0.934
	Weighted	94.5	0.93	0.968	0.948	93.6	0.926	0.943	0.934
Neural Networks	Narrow	94.5	0.935	0.948	0.941	92.7	0.958	0.963	0.960
	Medium	94.5	0.94	0.945	0.942	92.7	0.915	0.933	0.924
	Wide	93.2	0.922	0.930	0.926	92.7	0.915	0.933	0.924
	Bi layered	93.2	0.922	0.930	0.926	92.7	0.915	0.933	0.924
	Tri layered	93.2	0.927	0.932	0.923	89.9	0.89	0.90	0.895

Table 6. Feature Rankings from Kruskal-Wallis Algorithm

Rank	Feature	Score
1	Vacuolization damage basal layer	132.1187
2	Saw tooth appearance rates	132.0557
3	Melanin incontinence	131.205
4	Polygonal papules	131.1427
5	Fibrosis papillary dermis	130.8129
6	Perifollicular parakeratosis	130.621
7	Band like infiltrate	129.9214
8	Focal hypergranulosis	129.0225
9	Clubbing rete ridges	127.7499
10	Thinning suprapapillary epidermis	126.1007
11	Oral mucosal involvement	125.2916
12	Elongation rete ridges	115.1145
13	Follicular horn plug	107.1768
14	Knee and elbow involvement	98.7087
15	Follicular papules	87.6082
16	exocytosis	83.6033
17	Scalp involvement	83.0768
18	spongiosis	79.3603
19	Munro microabcess	77.6436
20	PNL infiltrate	64.9176
21	Koebner phenomenon	51.2436
22	Spongiform pustule	50.9313
23	Definite borders	50.0011
24	Disappearance granular layer	45.9134
25	itching	40.7342
26	parakeratosis	40.3839
27	scaling	34.1644
28	Family history	25.2672
29	Eosinophils infiltrate	19.1216
30	age	17.9604
31	acanthosis	16.946
32	erythema	16.8651
33	hyperkeratosis	16.5467
34	Inflammatory mononuclear infiltrate	16.5164

Table 6 represents the rankings of different features obtained by Kruskal-Wallis algorithm.

Table 7 represents the performance metrics of ML classifiers with best 30 features Similarly, Table 8 and 9 represents the performance of ML classifiers with best 25 and 20 features respectively.

Table 7. Performance Metrics with best 30 features

Category	Classifier	Accuracy (%)	Precision	Recall	F1 Score
Decision Trees	Fine	93.2	0.92	0.93	0.92
	Medium	93.2	0.92	0.93	0.92
	Coarse	78.1	0.63	0.56	0.60
SVM	Linear	93.2	0.91	0.93	0.92
	Cubic	94.5	0.93	0.93	0.93
	Quadratic	93.2	0.91	0.93	0.92
	Coarse Gaussian	93.2	0.91	0.93	0.92
	Medium Gaussian	93.2	0.91	0.93	0.92
	Fine Gaussian	74	0.61	0.69	0.65
Ensemble	Subspace KNN	93.2	0.918	0.922	0.919
	Bagged Trees	93.2	0.92	0.93	0.92
	Boosted Trees	93.2	0.92	0.93	0.92
	RUS Boosted Trees	95.9	0.956	0.953	0.954
	Subspace Discriminant	94.5	0.937	0.937	0.93
KNN	Fine	95.9	0.955	0.952	0.953
	Medium	93.2	0.924	0.92	0.921
	Coarse	79.5	0.63	0.59	0.609
	Cosine	93.2	0.924	0.92	0.923
	Cubic	94.5	0.937	0.937	0.937
	Weighted	94.5	0.937	0.937	0.937
Neural Networks	Narrow	94.5	0.88	0.91	0.89
	Medium	94.5	0.95	0.95	0.95
	Wide	94.5	0.95	0.95	0.95
	Bi layered	91.8	0.91	0.89	0.90
	Tri layered	94.5	0.95	0.943	0.94

Table 8. Performance Metrics with best 25 features

Category	Classifier	Accuracy (%)	Precision	Recall	F1 Score
Decision Trees	Fine	95.9	0.96	0.96	0.96
	Medium	95.9	0.96	0.96	0.96
	Coarse	78.1	0.63	0.56	0.60
SVM	Linear	100	1	1	1
	Cubic	97.3	0.97	0.97	0.97
	Quadratic	98.6	0.98	0.98	0.97
	Coarse Gaussian	94.5	0.90	0.95	0.93
	Medium Gaussia	97.3	0.94	0.98	0.96
	Fine Gaussian	57.5	0.44	0.73	0.55
Ensemble	Subspace KNN	97.3	0.968	0.968	0.968
	Bagged Trees	95.9	0.964	0.962	0.963
	Boosted Trees	95.9	0.965	0.962	0.963
	RUS Boosted Trees	98.6	0.987	0.983	0.985
	Subspace Discriminant	97.3	0.968	0.968	0.968
KNN	Fine	97.3	0.968	0.968	0.968
	Medium	97.3	0.968	0.968	0.968
	Coarse	82.2	0.66	0.58	0.617
	Cosine	97.3	0.968	0.968	0.968
	Cubic	97.3	0.968	0.968	0.968
	Weighted	97.3	0.968	0.968	0.968
Neural Networks	Narrow	95.9	0.95	0.96	0.954
	Medium	95.9	0.95	0.96	0.956
	Wide	95.9	0.95	0.96	0.956
	Bi layered	94.5	0.94	0.945	0.942
	Tri layered	95.9	0.95	0.96	0.947

Table 9. Performance Metrics with best 20 features

Category	Classifier	Accuracy (%)	Precision	Recall	F1 Score
Decision Trees	Fine	95.9	0.96	0.96	0.96
	Medium	95.9	0.96	0.96	0.96
	Coarse	78.1	0.63	0.56	0.60
SVM	Linear	97.3	0.97	0.98	0.97
	Cubic	93.2	0.89	0.93	0.91
	Quadratic	95.9	0.95	0.96	0.96
	Coarse Gaussian	94.5	0.91	0.95	0.93
	Medium Gaussian	93.2	0.89	0.94	0.91
	Fine Gaussian	32.9	0.19	0.38	0.26
Ensemble	Subspace KNN	98.6	0.987	0.983	0.984
	Bagged Trees	97.3	0.977	0.976	0.976
	Boosted Trees	95.9	0.965	0.962	0.963
	RUS Boosted Trees	98.6	0.987	0.983	0.984
	Subspace Discriminant	97.3	0.968	0.968	0.968
KNN	Fine	93.2	0.919	0.946	0.932
	Medium	94.5	0.932	0.968	0.949
	Coarse	79.5	0.648	0.67	0.658
	Cosine	94.5	0.932	0.968	0.949
	Cubic	94.5	0.932	0.968	0.968
	Weighted	94.5	0.932	0.968	0.968
Neural Networks	Narrow	95.9	0.95	0.96	0.954
	Medium	95.9	0.95	0.96	0.956
	Wide	95.9	0.95	0.96	0.956
	Bi layered	94.5	0.94	0.945	0.942
	Tri layered	95.9	0.95	0.96	0.947

Within Tables 7, 8, and 9, SVMs employing linear and cubic kernel functions, as well as Ensemble classifiers, notably Rusboosted trees, consistently outperformed other machine learning models, even when utilizing a reduced set of features. This highlights the robustness and efficiency of these techniques in various scenarios, showcasing their potential as top-performing options in a range of predictive tasks.

From the investigations, it is observed that by considering all the features for classification Linear SVM, Rusboosted Trees and subspace discriminant ensemble classifiers outperformed all other ML classifiers. After applying feature selection through kruskal-wallis algorithm, the performance of the most of the classifiers is increased because of reduction of redundant features. It is also observed that, Linear SVM achieved 100% accuracy with best 25 features and the maximum accuracy of all other classifiers is observed with 25 best features. From the investigations it is observed that feature selection algorithm also play a key role in identification of the ESD with high accuracy. The analysis underscores the significant potential of ML classifiers in enhancing the ESDs. Nonetheless, their effective deployment is contingent upon several pivotal factors. First and foremost, prioritizing data quality and reliability is imperative, given that ML models heavily rely on input data. Secondly, the prudent selection of suitable ML models customized for the specific ESD detection task is pivotal for achieving optimal performance. Finally, the interpretability of these models is essential for comprehending their decision-making processes and fostering trust in their outcomes.

5. CONCLUSION

This paper presents detailed investigations of state-of-the-art ML models for the classification of ESDs. Leveraging a rich dataset comprising 11 clinical features and 23 histo-pathological features, the study investigates the performance of machine learning classifiers, both individually and in combination of feature sets. Through meticulous experimentation, the classifiers' effectiveness is examined under varying training rates of 80% and 70%. This extensive analysis revealed valuable insights into selecting the most suitable classifier for ESD classification. Furthermore, this study ventured into assessing the effectiveness of the Kruskal-Wallis algorithm in ranking the importance of features within the dataset—a critical step in understanding the disease classification process. Significantly, the findings highlight that Ensemble and SVM classifiers outperformed other ML models in terms of F1 score and % of accuracy. In the study, Linear SVM, Rusboosted trees, and subspace discriminant ensemble classifiers excelled in ESD classification when considering all features. Feature selection via Kruskal-Wallis improved classifier performance. Linear SVM achieved 100% accuracy with the best 25 features. These results underscore the potential of these classifiers in aiding the accurate identification of ESD, thus offering promising prospects for improved diagnostic and clinical decision-making in the field of dermatology.

References:

Abdi, M. J., & Giveki, D. (2013). Automatic detection of erythemato-squamous diseases using PSO–SVM based on association rules. *Engineering Applications of Artificial Intelligence*, *26*(1), 603–608. doi: 10.1016/j.engappai.2012.01.017.

Badrinath, N., Gopinath, G., Ravichandran, K. S., Premaladha, J., & Krishankumar, R. (2018). Classification and prediction of Erythemato-Squamous diseases through Tensor-Based Learning. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, *90*(2), 327–335. doi: 10.1007/s40010-018-0563-x.

Banu, S. M., & Toacşe, G. (2013). A mobile/desktop medical application for automatic differential diagnosis of psoriasis lesions. *2013 Second International Conference on E-Learning and E-Technologies in Education (ICEEE)*, Lodz, Poland, 186-191. doi: 10.1109/icelete.2013.6644371.

Basu, A., Roy, S. S., & Abraham, A. (2015). A Novel Diagnostic Approach Based on Support Vector Machine with Linear Kernel for Classifying the Erythemato-Squamous Disease. *2015 International Conference on Computing Communication Control and Automation*, Pune, India. doi: 10.1109/iccubea.2015.72.

Danjuma, K., & Osofisan, A. O. (2014). Evaluation of Predictive Data Mining Algorithms in Erythemato-Squamous Disease Diagnosis. *International Journal of Computer Science Issues*, *11*(6), 85-94.

Kaushik, B., Vijayvargiya, A., Uppal, J., & Gupta, A. (2023). Comparative Analysis of Machine Learning Approaches for Classifying Erythemato-Squamous Skin Diseases. In: Kadry, S., Prasath, R. (eds) *Mining Intelligence and Knowledge Exploration*. MIKE 2023. *Lecture Notes in Computer Science*, *13924*. Springer, Cham. doi: 10.1007/978-3-031-44084-7_7

Li, Z., Koban, K. C., Schenck, T. L., Giunta, R. E., Li, Q., & Sun, Y. (2022). Artificial intelligence in dermatology image analysis: current developments and future trends. *Journal of clinical medicine*, *11*(22), 6826. doi: 10.3390/jcm11226826

Polat, K., & Güneş, S. (2006). The effect to diagnostic accuracy of decision tree classifier of fuzzy and k-NN based weighted pre-processing methods to diagnosis of erythemato-squamous diseases. *Digital Signal Processing*, *16*(6), 922–930. doi: 10.1016/j.dsp.2006.04.007.

Polat, K., & Güneş, S. (2009). A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems. *Expert Systems with Applications*, *36*(2), 1587–1592. doi:10.1016/j.eswa.2007.11.051.

Putatunda, S. K. (2020). A Hybrid Deep Learning Approach for Diagnosis of the Erythemato-Squamous Disease. *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, Bangalore, India. https://doi.org/10.1109/conecct50063.2020.9198447.

Ravichandran, K. S., Narayanamurthy, B., Ganapathy, G., Ravalli, S., & Sindhura, J. (2014). An efficient approach to an automatic detection of erythemato-squamous diseases. *Neural Computing and Applications*, *25*(1), 105–114. doi: 10.1007/s00521-013-1452-5.

Singh, S. K., Sinha, A., & Yadav, S. S. (2022). Performance Analysis of Machine learning Algorithms for Erythemato-Squamous Diseases Classification. *2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*. doi:10.1109/icdcece53908.2022.9793000.

Spolaôr, N., Lee, H., Mendes, A. I., Nogueira, C., Parmezan, A. R. S., Takaki, W. S. R., Coy, C. S. R., Wu, F., & Fonseca-Pinto, R. (2023). Fine-tuning pre-trained neural networks for medical image classification in small clinical datasets. *Multimedia Tools and Applications*. doi: 10.1007/s11042-023-16529-w.

Subbarao, M. V., Ram, G. C., & Varma, D. (2023). Performance Analysis of Pistachio Species Classification using Support Vector Machine and Ensemble Classifiers. *2023 International Conference on Recent Trends in Electronics and Communication (ICRTEC)*, Mysore, India. doi: 10.1109/icrtec56977.2023.10111889.

Subbarao, M. V., Rani, U., Mythreya, V. H., Kumar, G. P., Priyakanth, R., & Abudhagir, U. S. (2023). Investigations on Feature Ranking Algorithms for Date Fruit Genotype Detection. *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India. doi: 10.1109/icces57224.2023.10192766.

Tekin, A. (2014). Classification of Erythemato-Squamous diseases using association rules and fuzzy C-Means clustering. *Arabian Journal for Science and Engineering*, *39*(6), 4699–4705. doi: 10.1007/s13369-014-1168-6.

Xie, J., Lei, J., Xie, W., Gao, X., Shi, Y., & Liu, X. (2012). Novel hybrid feature selection algorithms for diagnosing erythemato-squamous diseases. In: He J, Liu X, Krupinski EA, Xu G (eds) HIS, 7231. Springer, Berlin, 173–185. doi: 10.1007/978-3-642-29361-0_21.

M. Venkata Subbarao
Department of ECE,
Shri Vishnu Engineering College for Women, Bhimavaram, India
mandava.decs@gmail.com
ORCID 0000-0001-5840-2190

S. Lakshmi Praveena
Department of ECE,
Shri Vishnu Engineering College for Women, Bhimavaram, India
sreepada.praveena@gmail.com
ORCID 0009-0005-1621-5430

T. Naga Sharmila
Department of ECE,
Shri Vishnu Engineering College for Women, Bhimavaram, India
sharmilatanniru24@gmail.com
ORCID 0009-0003-2894-2227

Praveen Kumar Nalli
Department of AI,
Shri Vishnu Engineering College for Women, Bhimavaram, India
praveenkumar.nalli@svecw.edu.in
ORCID 0000-0002-7443-7841

G. Challa Ram
Department of ECE,
Shri Vishnu Engineering College for Women, Bhimavaram, India
challaram.grandhi@gmail.com
ORCID 0000-0002-3790-1318

D. Ramesh Varma
Department of ECE,
Shri Vishnu Engineering College for Women, Bhimavaram, India
varmaramesh422@gmail.com
ORCID 0000-0002-9517-4967

Srinivasarao Alluri
Department of ECE,
Koneru Lakshmaiah Education Foundation (KLEF), Green Fields, Vaddeswaram (V), Guntur District, Andhra Pradesh, India
asrao.81@gmail.com
ORCID 0000-0002-3790-1318

Tabassum Nazeer Syed
The Bradley Department of Electrical and Computer Engineering,
Virginia Polytechnic Institute and State University, USA
tabassumnazeer8@vt.edu