IETE Journal of Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tijr20

Erythemato-Squamous Diseases Prediction and Interpretation Using Explainable AI

Abhishek Singh Rathore, Siddhartha Kumar Arjaria, Manish Gupta, Gyanendra Chaubey, Amit Kumar Mishra & Vikram Rajpoot

To cite this article: Abhishek Singh Rathore, Siddhartha Kumar Arjaria, Manish Gupta, Gyanendra Chaubey, Amit Kumar Mishra & Vikram Rajpoot (2022): Erythemato-Squamous Diseases Prediction and Interpretation Using Explainable AI, IETE Journal of Research, DOI: 10.1080/03772063.2022.2114953

To link to this article: https://doi.org/10.1080/03772063.2022.2114953

Published online: 11 Sep 2022.

IETE JOURNAL OF RESEARCH
https://doi.org/10.1080/03772063.2022.2114953

Erythemato-Squamous Diseases Prediction and Interpretation Using Explainable AI

Abhishek Singh Rathore[1], Siddhartha Kumar Arjaria[2], Manish Gupta[3], Gyanendra Chaubey [2], Amit Kumar Mishra[4] and Vikram Rajpoot[5]

[1]Department of CSE, SVIIT, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore 453111, India; [2]Department of Information Technology, Rajkiya Engineering College, Banda, India; [3]Department of Electronics and Communication Engineering, GLA University, Mathura, UP, India; [4]Department of CSE, Amity University Madhya Pradesh, Gwalior, MP, India; [5]Department of Information Technology, Madhav Institute of Technology & Science, Gwalior, MP, India

ABSTRACT

Erythemato-squamous diseases (ESD) diagnosis is a significant challenge in dermatology. It is divided into six categories. Artificial intelligence models have been applied to categorize these categories. Artificial intelligent models are black boxes in nature. The objective of this study is to unbox the black-box behavior and interpret the decision-making. Random Forest and XGBoost models are applied on a standard dataset with SHAP value to get interpretability and causability of decision. The Random Forest model had a classification accuracy of 98.21%. Integration of explainability increase the transparency of result and identify the root cause of the disease in the subject. A comprehensive quantitative study will help to adopt artificial intelligence in healthcare with ethical issues like transparency, causability, and interpretability of diagnosis.

KEYWORDS

Dermatology; erythemato-squamous diseases; explainable AI; feature ranking; machine learning; SHAP value

1. INTRODUCTION

Erythemato-squamous diseases (ESDs) are one of the most prevalent types of skin diseases [1]. They are divided into six categories: psoriasis, seborrheic dermatitis, lichen planus, pityriasis rosea, chronic dermatitis, and pityriasis rubra pilaris. All these classes of skin disorders are usually related to genetic, environmental risk factors, or unknown etiology.

- Psoriasis conditions appear to be polygenic and are considered hereditary disorders. but is also dependent on certain environmental triggers. Psoriasis is T-lymphocytes-related immune disorder
- Pityriasis rosea disorder of adolescents and young adults. The cause is unknown, but it is thought to be viral or post viral.
- Lichen planus etiology is unknown but psoriasis. It is also postulated as T-cell driven immune-related disorder.
- Pityriasis Rubra is the unknown origin, two most common forms one is juvenile type (classic) is autosomal dominant another one is an adult type having no inheritance
- Chronic dermatitis is also related to a family history of asthma or any environmental or food allergies or genetically transferred as well as related to T-lymphocyte-related immune disorder
- Seborrheic dermatitis is also related to genes, cold weather, or environmental factors that affect it. Some preexisting diseases like stress, AIDS, alcoholism, and parkinson increase their chances as well as more common in the male gender with oily skin.

They all have erythema and scaling as clinical characteristics, along with minor variations. Because they have overlapping indications and symptoms, automated detection of these diseases are a difficult challenge. Another issue is when a disease exhibits the symptoms and traits of another illness in its early stages but then develops its distinct features as it advances [2]. Initially, patients are diagnosed based on twelve clinical features as shown in Table 1, afterwards, patients are tested for twenty-two histopathological features under the microscope [3], which creates another problem for automated prediction as these histopathological features have similar values for the six ESDs. Some of the common causes of these skin diseases are weak immune systems, parasites, fungus, or some other microorganisms residing on the skin, etc. Too much flushing, red or white color raised bumps, a loss of skin pigment, cracked skin, etc. are some of the common indications of skin disorder. Correct identification of these diseases and appropriate treatment selection are two steps for becoming combating the disease in the least amount of time and financial cost. Contemporary

Table 1: UCI machine learning dermatology dataset attributes

#	Attribute	Type
1	Erythema	Clinical Attribute
2	Scaling	
3	Definite Borders	
4	Itching	
5	Koebner Phenomenon	
6	Polygonal Papules	
7	Follicular Papules	
8	Oral Mucosal Involvement	
9	Knee And Elbow Involvement	
10	Scalp Involvement	
11	Family History, (0 Or 1)	
12	Melanin Incontinence	Histopathological Attributes
13	Eosinophils in the 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	Vacuolisation and Damage Of Basal Layer	
28	Spongiosis	
29	Saw-Tooth Appearance of Retes	
30	Follicular Horn Plug	
31	Perifollicular Parakeratosis	
32	Inflammatory Monoluclear Inflitrate	
33	Band-Like Infiltrate	
34	Age (Linear)	Clinical Attribute

machine learning models can perform disease diagnoses at relatively low costs. However, machine learning models are accurate, and only perform a supportive role in medical decision-making. This is due to the lack of trust and thus the lack of social acceptability of the models as these models simulate a black box behavior by hiding the details of the decision-making. Questions like "how the decision has been made" and "what is the contribution of each feature in the outcome" is considered important for developing faith in the system's outcome. There should be trust between the healthcare practitioner and the information system in taking decisions. If a healthcare practitioner is unable to understand the decisions made by machine learning models, he is unable to explain his treatment to patients [4]. The lack of explanation that certain algorithms suffer from, as well as the reality that treatment options are often less successful in ordinary clinical practice than initial evaluation, increase regulatory complexity [5]. A health practitioner must think about the origins and repercussions of medical issues, as well as the tools and models that help them make decisions [6]. Explainable Artificial Intelligence (XAI) is one way to unbox the black-box behavior of artificial intelligence models [7]. Integrating the explainability constraints using XAI into artificial intelligence models, improves the understandability of the decision process.

According to a study, being a black box is a "severe limitation" for AI in dermatology because it can't perform personalized assessments by-certified dermatologists that may be utilized to elucidate clinical facts [8]. The paper uses Shapley's value-based approach for the interpretation of the model and model outcome. Some methods of interpretation are model-specific. They are only able to explain the model for which they are designed. Another approach is Model agnostic. Model-agnostic interpretation methods can be used to interpret any machine learning model. No matter how complex the model is. Usually, the model agnostic methods perform the interpretation by analyzing feature input and output pairs. The model-agnostic approach can be applied in almost every model. SHAP [9] approach is the most used model-agnostic approach. This is a cooperative game theory-based approach to finding the impact of each player (attribute) in the game output (accuracy/precision/recall/F1-Measure). The SHAP values are good in model interpretation both globally and locally. In global interpretation, the model finds out the role of each predictor (feature) in prediction. The plot thus generated will demonstrate the positive and negative impact of each input. For a greater level of transparency, SHAP values can also provide Local interpretation. which deals with the observation on hand by calculating the SHAP values for each feature of every observation thereby showing the local interpretation of the model. The SHAP method is the post-hoc method and is applied after model training.

Though simple models like linear models as shown in Figure 1(left), are easy to interpret. The contribution of each input I_i is just multiplied with some constant weights, but the accuracy of such models is not up to the mark. Complex models yield good results but are too complex to understand. The SHAP employed a function G over model function f and current input I_i to

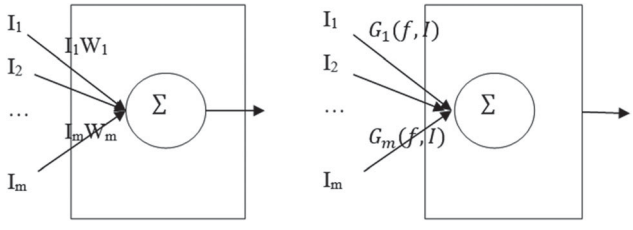


Figure 1: Simple model (left) vs SHAP value (right)

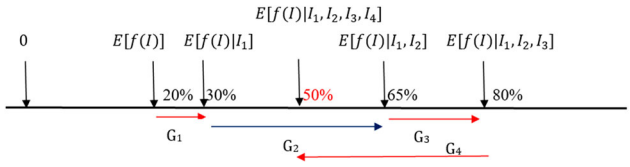


Figure 2: Conditional expectation computed in SHAP

understand the decision as shown in Figure 1(right). The function G calculates the conditional expectation of the model with all features as shown in Figure 2. The model starts computation with the null model. The SHAP values are used to find out the contribution(importance) of each feature. For it, all the possible combinations of features should be explored. The SHAP begins its journey with a model without any independent variables. This model is known as the null model. SHAP then calculates the average conditional contribution for each feature to be added in this null model in sequence.

To calculate the contribution of the feature I_1 in the decision process, SHAP computes G_1 function is shown in Equation (1):

$$G_1(f, I_1) = E[f(I)|I_1] \tag{1}$$

It signifies that we are calculating the conditional expectation of the model in the presence of I_1 feature. Similarly, G_2 is calculated as shown in

$$G_2(f, I_2) = E[f(I, I_1)|I_2] \tag{2}$$

So, the SHAP values compute the effect of each feature on the joint distribution of features. The rest of the paper is divided into five different sections.

Section 2 conducted a literature review of explainable AI in healthcare. Section 3 illustrates the methodology used in this white paper. Section 4 uses various graphs and tables to examine the results in detail. Section 5 presents the findings of Section 4 and conclusions based on the findings, exploring future aspects of the study.

2. LITERATURE REVIEW

Arora *et al.* [10] propose a bag of features to extract robust features and applied a quadratic support vector machine for the early prediction of skin diseases for binary classification. Many studies have been proposed by different authors to identify the region of interest in skin images [11–15] using deep neural networks. Problems associated with these studies include high training time and insufficient training data. Yuan *et al* [16] introduced Jaccard distance to improve the performance of image segmentation with the use of neural networks. Tyagi and Mehra [17] proposed an Intelligent Prognostics Model that combines K-means with Particle Swarm Optimization (PSO) for image segmentation to find the region of interest. Speeded-up robust features (SURF) are extracted from the region of interest (ROI) which are then filtered with PSO for classification using convolutional neural networks. Though, SURF features do speed up the training time, they result in poor performance with different illumination and are unstable with different rotation transformations.

In addition to this, authors have combined different machine learning algorithms for performance improvements [18–22]. Verma *et al.* [23] applied ensemble methods using six machine learning methods on the top fifteen features selected from the root mean square estimate. This study still needs the impact of various numbers of features on the accuracy of the classification algorithm. Verma and Pal [24] in addition to previous work, used chi-square and heat-map to filter out features for ensemble learning. In pattern analysis, features play important roles. Different approaches were reported by authors for feature selections from medical images. Some authors used handcrafted features [25–28], while others used color features [29–32], textual features [33–35], or shape features [36–38] to improve the classification accuracy.

Many studies have worked on the same data set in the field of ESD. Güvenir and Demiröz [39] proposed a data mining technique called the voting feature intervals-5 algorithm, to classify the existing data and predict the unseen data. While Menai and Altayash [40] used an Ensembled decision tree for better accuracy. Recently, several studies have been published on the same data set using machine learning techniques including SVM with optimal feature selection [41], extreme machine learning and neural network [42], Fuzzy extreme machine learning [43], decision tree [21,44], ensemble methods [3, 65], SVM [45,46], neuro-fuzzy system [47], and Bayesian Optimization [66].

Many researchers are working in the field of feature selection. In addition to computer vision, researchers put their findings in feature selection using rough set methods [48–52] to find out hidden patterns. Sinha and Namdev [53] used the rough set theory to find out the best features for skin diseases. The study has not provided a detailed study of the different features. The machine learning algorithms yield accurate results, but the interpretability of the results is a more important issue and cannot be

ignored. It certainly helps to understand what is going wrong with the patients and contemporary research is also focused on tracing the etiology of the diseases. Shin [54] and Chazette & Schneider [55] highlighted the Implications of Explainable AI on trust and awareness. Lots of studies have been published dedicated to the application of XAI in the healthcare system [56–59]. Khodabandehloo *et al.* [60] developed HealthXAI using XAI for monitoring the cognitive decline in elderly people. Lauritsen *et al.* [61] proposed an XAI-based model for the prediction of acute critical illness based on early warning scores from electronic health records. Carrieri *et al.* [62] identified variations in skin microbiome composition that are connected to phenotypic differences using XAI.

3. METHODS

Fusion of machine learning models with ESDs dataset can help to better understand the function of the symptoms involved in skin health. It is critical to evaluate both the predictive strength and the interpretability of the machine learning models for the transparency of decisions [62].

In this proposed work, our objective is to use Explainable AI (XAI) to predict and identify the root cause of ESD. We propose the XAI approach to predict the disease accurately and investigate how the distribution of different symptoms drives the prediction of one of the ESD. The interpretation of the learning models is expressed in the terms of the conditional expectation of different symptoms (features) in presence of others. Two machine learning models are used for the prediction of ESD, namely – Random Forest [63], and XGBoost [64]. For the classification task, both models are hyper-tuned on the training dataset to predict ESD. These optimized models were then used to predict the disease in unseen data. The performance of both models is compared on validation accuracy, precision, and recall for the classification task. The interpretation of both models is then provided using SHAP values.

The steps of the XAI-based methodology are shown in Figure 3, starting with the data preprocessing step. For the cross-validation of the data, the whole dataset is randomly split into the train and test parts. The machine learning models (Random Forest and XGBoost) are trained using training data to learn optimized hyperplane for classifications. Both models are non-linear. To avoid overfitting of data, models are then tested on test data set using 10-fold cross-validation to get validation accuracy. Finally, SHAP values are used to interpret the predictions of the model.

Algorithm 1: Random Forest with Tuning

1. Read the dataset
2. Split the dataset in an 80:20 ratio for training and testing
3. Among 34 attributes of the dataset, select randomly $k = 6$ attributes

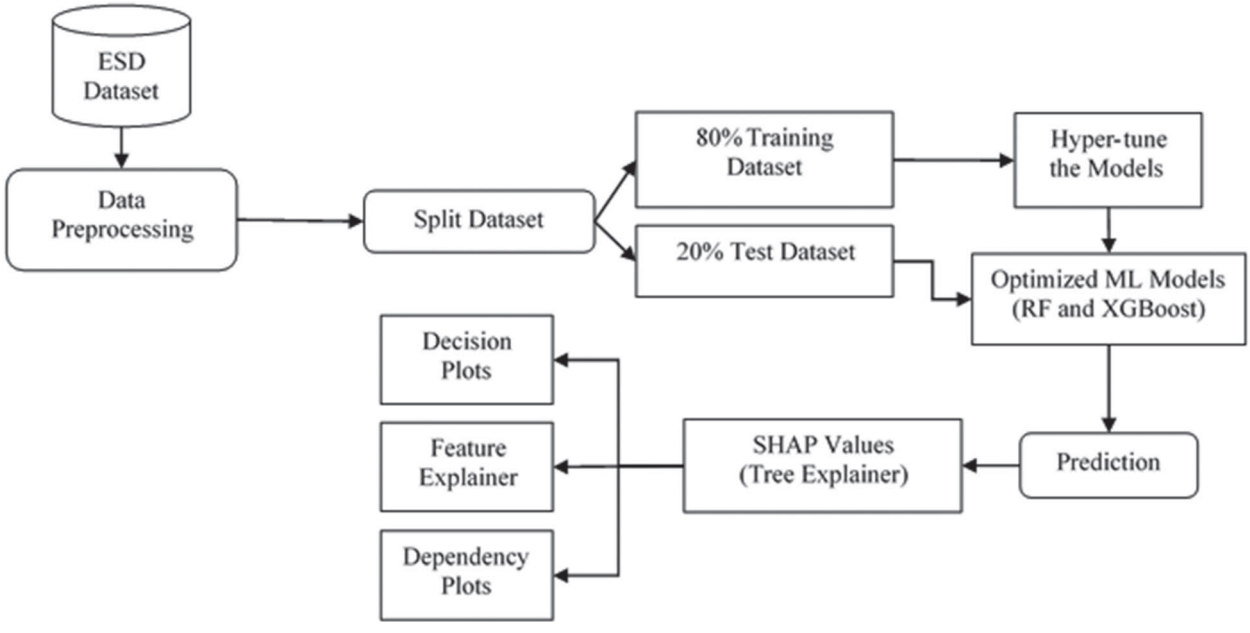


Figure 3: Proposed methodology

3.1 Generally k = Squareroot (Total Attributes)
4. Among 6 attributes, select the best split node using Gini Impurity
5. Repeat step 3 and 4 for the remaining nodes to build an individual tree
6. Repeat step 3, 4, and 5 to build a forest for n times on training data.
7. Calculate the prediction of all the trees on test data
8. Most highly voted will be the final prediction of the forest.
9. Repeat steps 3–8 for varying values of n and k to tune the parameter of the random forest.

Algorithm 2: XGBoost with Tuning

1. Read the dataset
2. Split the dataset in 70:30 ratio for training and testing
3. Train the model using parameters eta = 0.3, gamma = 0, max_depth = 6, subsample = 0.5, lambda = 1, and alpha = 0
4. Calculate the prediction of model on test dataset
5. Repeat the steps 3 and 4 for varying the values of parameters to tune the model.

Algorithm 3: Calculation of Shap Values

1. Let f is the efficient model trained from Algorithm 1 and Algorithm 2
2. I is set of the input features = $\{I_1, \ldots .., I_{34}\}$
3. Compute Shap Function $G_1(f, I_1) = E[f(I)|I_1]$
4. Compute $G_2(f, I_2) = E[f(I, I_1)|I_2]$
5. Similarly, Compute all G_i's

3.1 Dataset Description

For this study, UCI Machine Learning Dermatology data set[1] and provided by Guvenir [34] is used and multivariate in nature. Total 34 attributes are listed as shown in Table 1. ESD is categorized into 6 different classes as with number of instances is shown in Table 2.

3.2 Model Performance

After hyper-tuning, both models are fitted to the training data. The testing and cross-validation accuracy of the models are shown in Table 3. From Table 3, it is clear that both models perform well in training data, but Random Forest generalizes slightly better than XGBoost on unseen data. Since accuracy is not a global measure to compare the classification models, therefore classification report is generated for both models.

Table 2: Class-wise instance of subjects

ESD	Classes	Count of instances
Psoriasis	Class 0	112
Seborrheic Dermatitis	Class 1	61
Lichen Planus	Class 2	72
Pityriasis Rosea	Class 3	49
Chronic Dermatitis	Class 4	52
Pityriasis Rubra Pilaris	Class 5	20

Table 3: Performance comparison of random forest and XGBoost model

	Random forest	XGBoost
Testing accuracy	100%	99.50%
Leave one out validation accuracy	98.21%	97.85%

Table 4 represents the classification metric of the Random Forest model on each type of ESD. This provides a more in-depth understanding of the Random Forest classifier's behavior rather than global accuracy, which might conceal functional flaws in the classification of each ESDs. True and false positives, as well as true and false negatives, are used to establish these metrics.

- Precision is the ability of the classifier in percentage to correctness in classification.
- Recall is the ability of a classifier in percentage to correctly classify instances as positive.
- F1-Score is the harmonic mean of Precision and Recall
- Support is the number of instances of the class in the supplied dataset.

The Unweighted Average (Macro Average) and Weighted Average are also depicted in Table 4. Score 1 reflects the best score while 0 reflects the worst one. Since F1-Score are generally lower than general accuracy, weighted averages are used to compare different classifier models. While generating the classification report of the XGBoost classifier, similar classification metrics have been retrieved. Therefore, based on Table 3 accuracy comparison, Random Forest has been chosen for result interpretation.

Table 4: Classification report of random forest

	Precision	Recall	F1-Score	Support
Psoriasis	1.00	1.00	1.00	34
Seborrheic Dermatitis	0.97	1.00	0.98	29
Lichen Planus	1.00	1.00	1.00	43
Pityriasis Rosea	1.00	0.97	0.98	33
Chronic Dermatitis	1.00	1.00	1.00	25
Pityriasis Rubra Pilaris	1.00	1.00	1.00	36
Accuracy			0.99	200
Macro Average	0.99	0.99	0.99	200
Weighted Average	1.00	0.99	1.00	200

4. RESULT AND DISCUSSION

In the previous section, the dataset and methodology used in this paper have been explored. This section demonstrates the interpretation of the black-box machine learning models of the experiment through different plots and tables.

Figure 4 shows the SHAP feature summary plot. It is a measure to explain the model globally. Features with large SHAP values are more important. In chronic dermatitis, the degree of contribution of fibrosis of the papillary dermis is very high, with mild values of spongiosis, koebner phenomenon, elongation of the rete ridges, perifollicular parakeratosis, and clubbing of the rete ridges. The features are sorted down with decreasing importance. The importance of a feature can be explained by the magnitude of the feature's classification capacity. The larger its value, the more important the feature is. Figure 4 shows the SHAP feature importance for the random forest trained before for predicting the type of ESDs. Thining of the suprapapillary epidermis, clubbing of the rete ridges, spongiosis, elongation of the rete ridges, follicular horn plug is significantly observed in Psoriasis.

Koberner phenomenon develops Seborrheic Dermatitis lesions on the skins along with fibrosis of the papillary dermis, scaling, and thinning of the suprapapillary epidermis. Similarly, Lichen Planus includes the symptoms related to oral mucosal involvement, melanin incontinence, vacuolization and damage of basal layer, focal hypergranulosis, and saw-tooth appearance of rete ridges. Similarly, Pityriasis Rosea has symptoms of Koberner phenomenon, fibrosis of the papillary dermis, follicular horn plug, elongation of the rete ridges, knee and elbow involvement, and disappearance of the granular layer while follicular horn plug, perifollicular parakeratosis, follicular papules with age factor are significantly observed in Pityriasis Rubra Pilaris. Though it is important information, it fails to explore other critical things like the effect of each feature instance-wise.

Figure 5 visualizes all the instances of the dataset. Every instance of the dataset is represented by the point in the

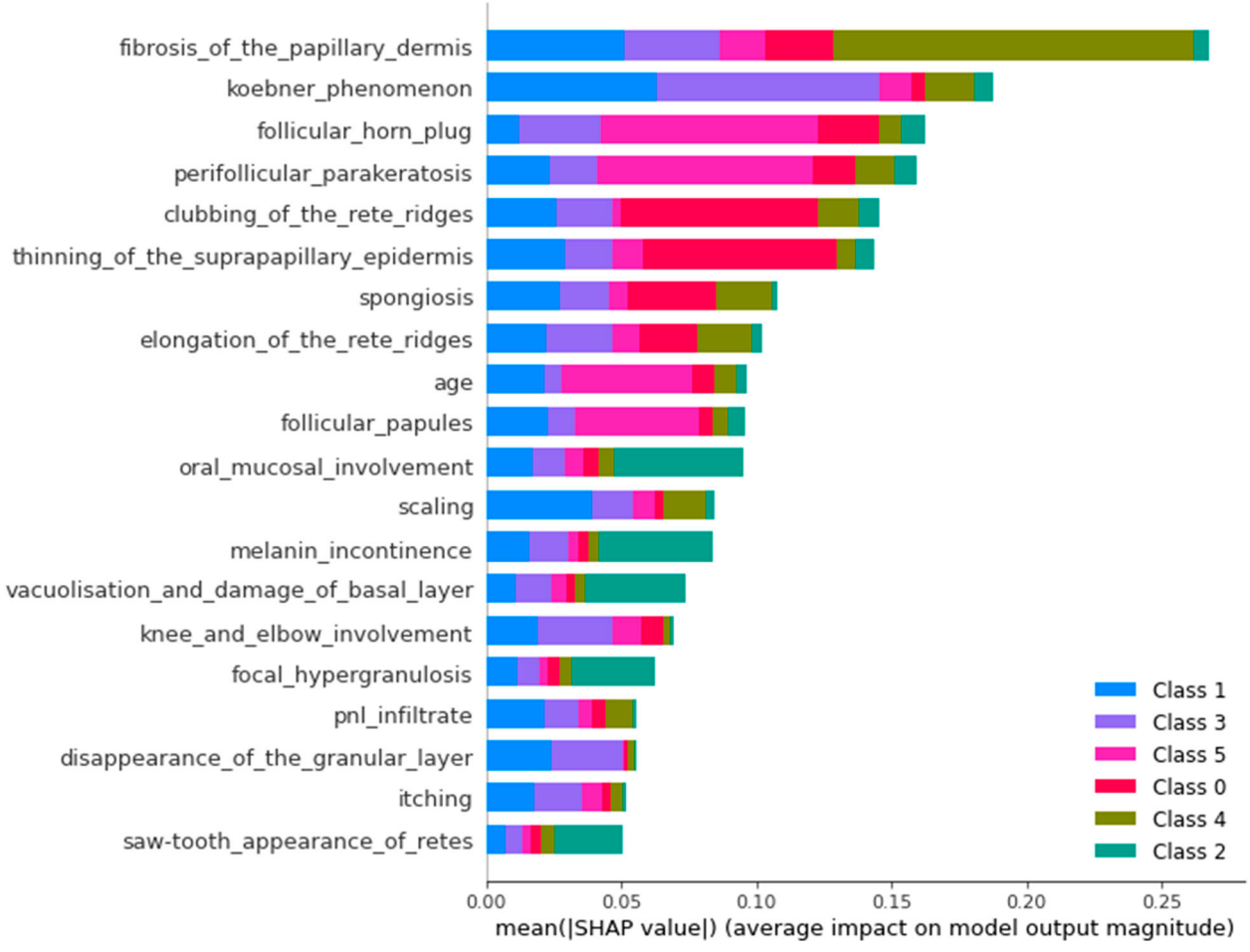


Figure 4: SHAP feature importance

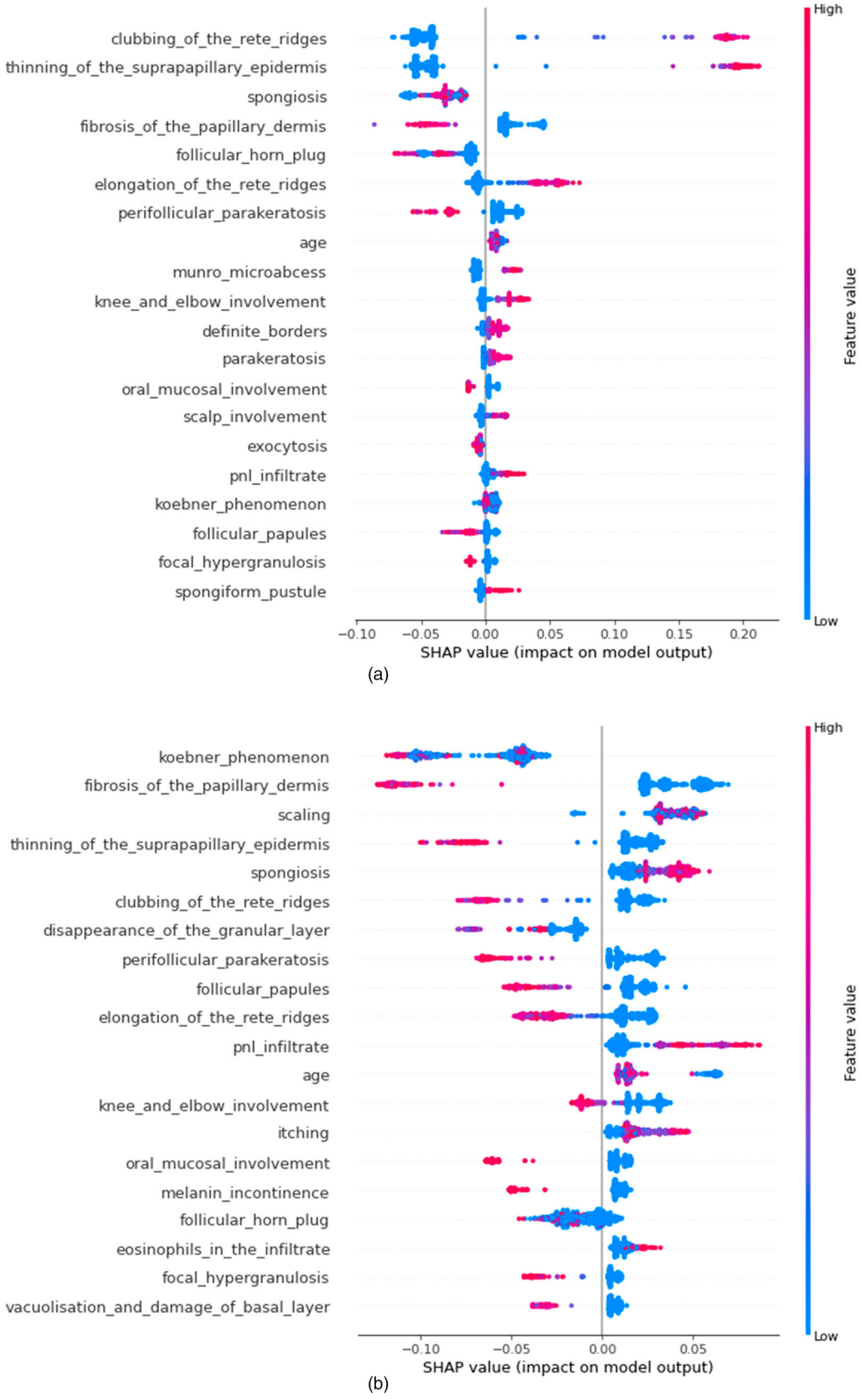
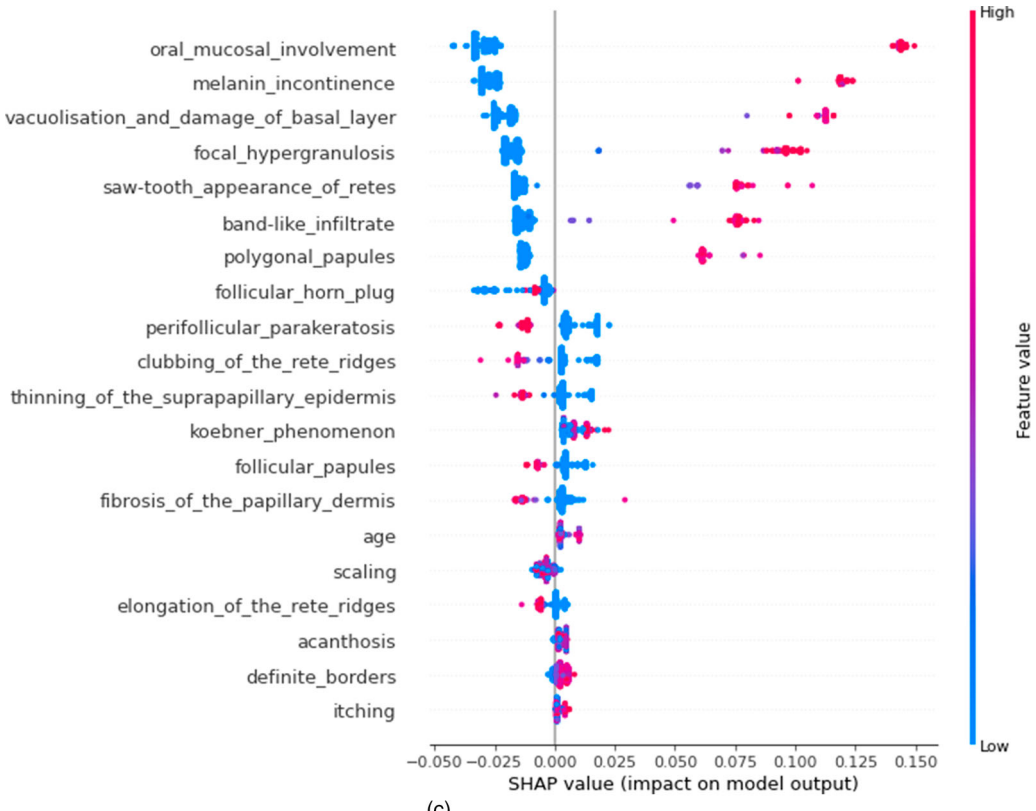
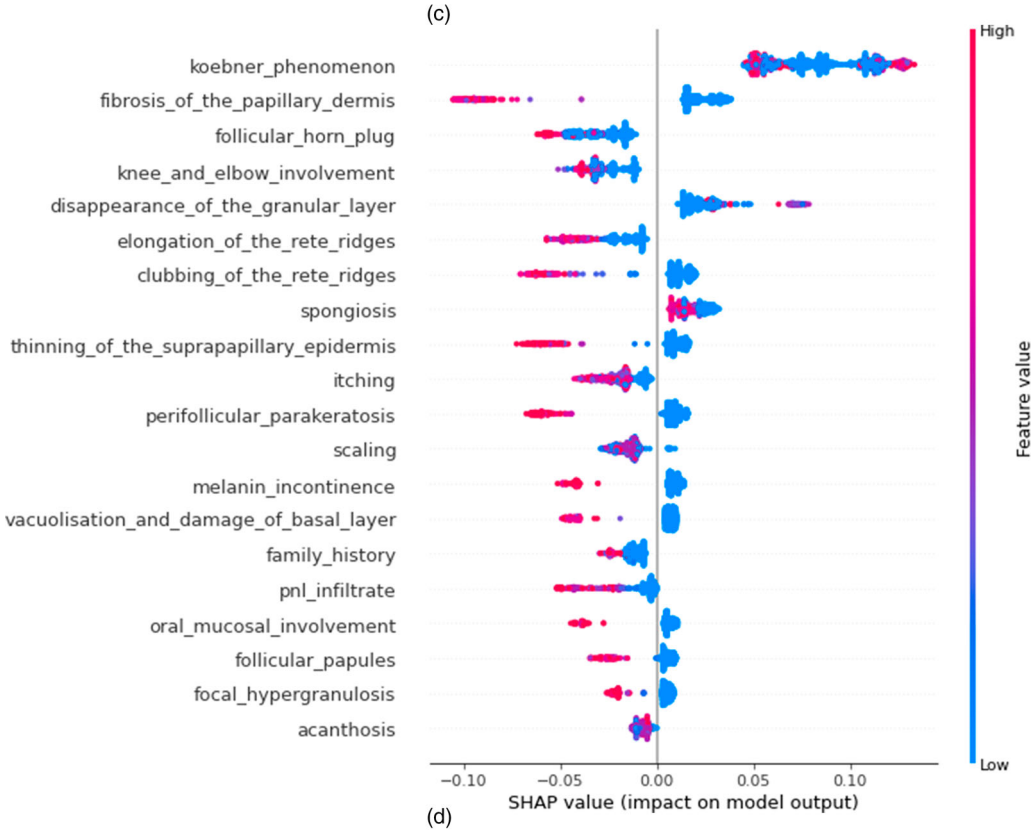


Figure 5: SHAP summary plot for ESDs, left to right and top to bottom (a) Psoriasis, (b) Seborrheic Dermatitis, (c) Lichen Planus, (d) Pityriasis Rosea, (e) Chronic Dermatitis, (f) Pityriasis Rubra Pilaris

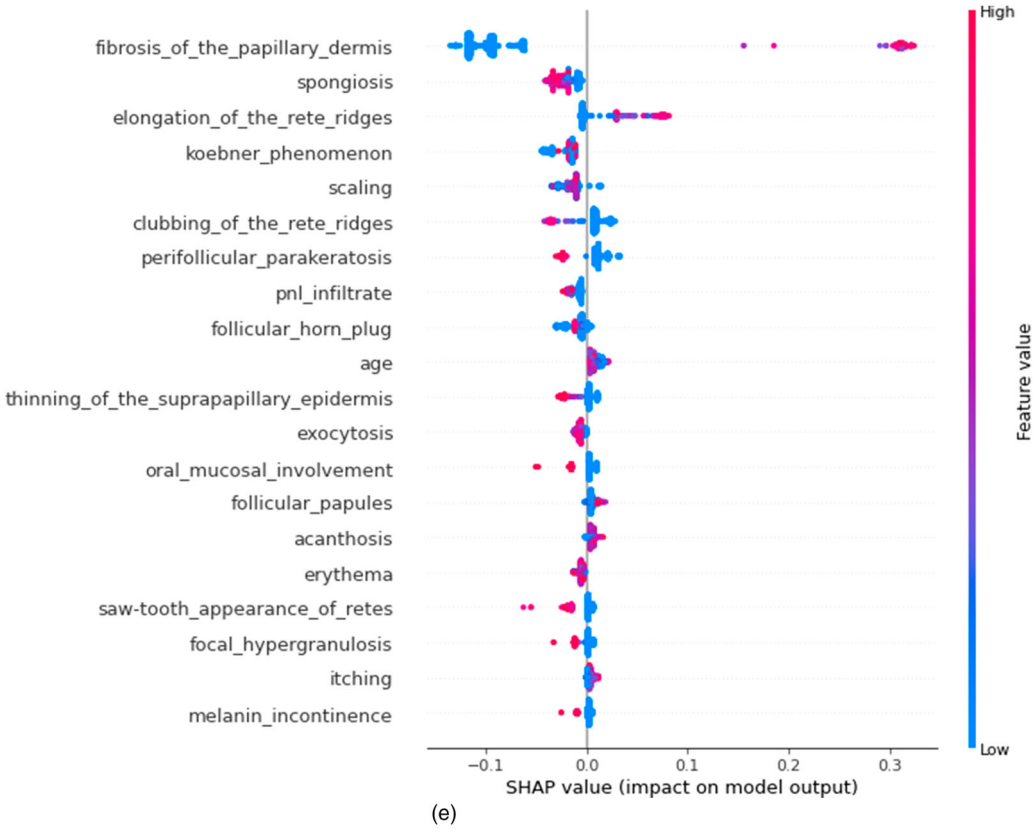


(c)

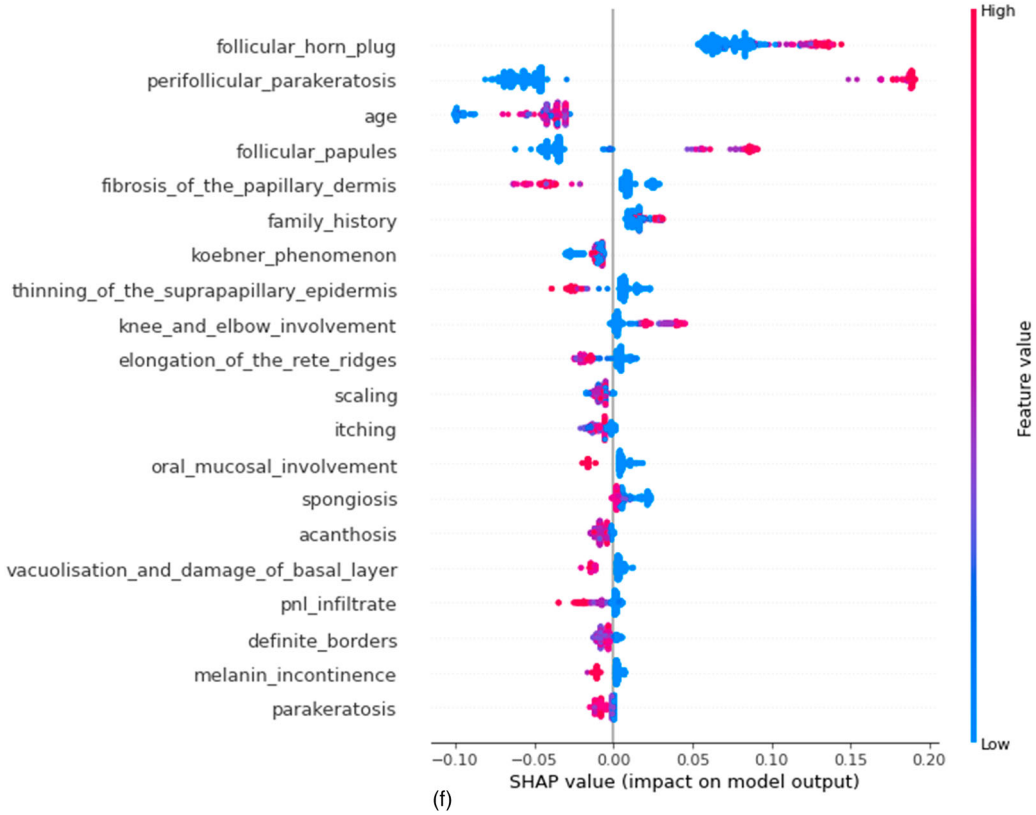


(d)

Figure 5: Continued



(e)



(f)

Figure 5: Continued

graph where features are represented on the y-axis and Shapley values are represented on the x-axis. The features are arranged in terms of their contribution to disease prediction and sorted in descending order. It is usually used for interpreting feature importance and its effect. A zero value on the x-axis shows the zero contribution of that feature value for the single data instance. The contribution increases in terms of sign and quantity when the SHAP value moves away from zero. The vertical line on zero represents no contribution whereas contributions increase as the SHAP value moves away from the zero line. Figure 5 shows the SHAP summary plot for all the six ESDs. A low number of clubbing of the rete ridges involvement, thinning of the suprapapillary epidermis, and elongation of the rete ridges lowers the risk of Psoriasis while higher values of these features increase the risk. The presence of spongiosis lowers the risk of Psoriasis. In addition to this, higher values of fibrosis of the papillary dermis, follicular horn plug, and perifollicular parakeratosis reduce the risk of Psoriasis and vice versa as shown in Figure 5(a).

Koebner phenomenon has a great impact on Seborrheic Dermatitis as shown in Figure 5(b). The presence of skin lesions on trauma lines, and the disappearance of the granular layer reduces the risk of Seborrheic Dermatitis. Higher the value of fibrosis of the papillary dermis, thinning of the suprapapillary epidermis, clubbing of the rete ridges involvement, perifollicular parakeratosis, follicular papules, and elongation of the rete ridges reduce the risk of Seborrheic Dermatitis. The presence of scaling in skin lesions, spongiosis, PNL infiltrate, and itching increases the chance of Seborrheic Dermatitis. Higher the involvement of oral mucosal, melanin incontinence, vacuolization and damage of basal layer, focal hypergranulosis, the saw-tooth pattern of rete, band-like infiltrate, and polygonal papules higher the chance of Lichen Planus as shown in Figure 5(c), while lower values reduce the risk. The presence of Koebner phenomenon and disappearance of the granular layer increases the risk of Pityriasis Rosea and fibrosis of the papillary dermis, follicular horn plug, and involvement of knee and elbow reduces the risk as shown in Figure 5(d). Higher the value of fibrosis of the papillary dermis, elongation of the rete ridges increases the probability of having Chronic Dermatitis while symptoms related spongiosis, koebner phenomenon, scaling in skin lesions suggests that there are low chances of Chronic Dermatitis in the subject as shown in Figure 5(e). If there are symptoms related to follicular horn plug, the subject has a family history, involvement of knee and elbow are seen, then there will be higher chances of Pityriasis Rubra Pilaris as shown in Figure 5(f). In addition, higher values of perifollicular parakeratosis, follicular papules increase the risk of Pityriasis Rubra Pilaris, while higher values of fibrosis of the papillary dermis reduce the risk.

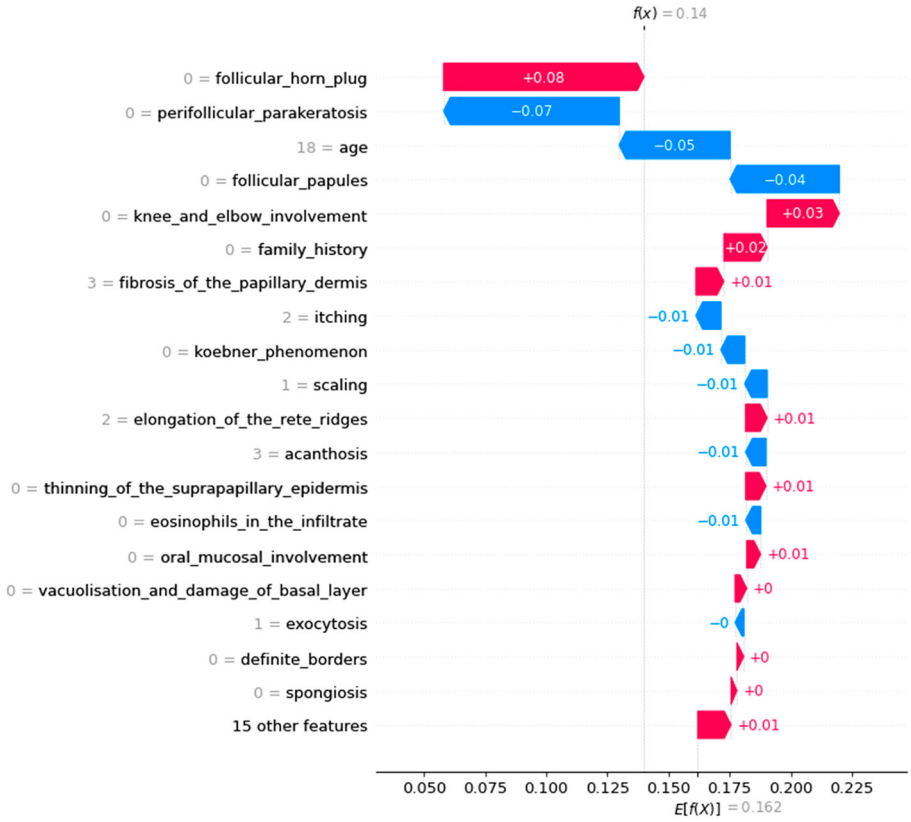
Waterfall Plot represented in Figure 6 is a measure of the explainability of the model locally. It explains the model performance instance-wise. This plot summarizes the role of each feature in predicting the class of a single instance. On the y axis, features of the instance are indicated. Each row is indicated either in red or blue color. The Red color indicates the positive contribution of the corresponding feature while the blue color indicates the negative contribution of the corresponding feature in the outcome of the given instance. The contribution of each feature is indicated positively or negatively through the value indicated in the horizontal box from the expected model output over the background dataset to the model output for this prediction. Check this, appears to be a bad statement. Higher the value more the positive or negative impact of the feature in the outcome of the given instance.

In Figure 6, f(x) is the model predict_proba value is –.098 and the $E[f(x)]$ is the base value $= 0$, Feature acanthosis of subject 56 has value 0, and then it shows the highest negative .06 value shift from the base value 0.

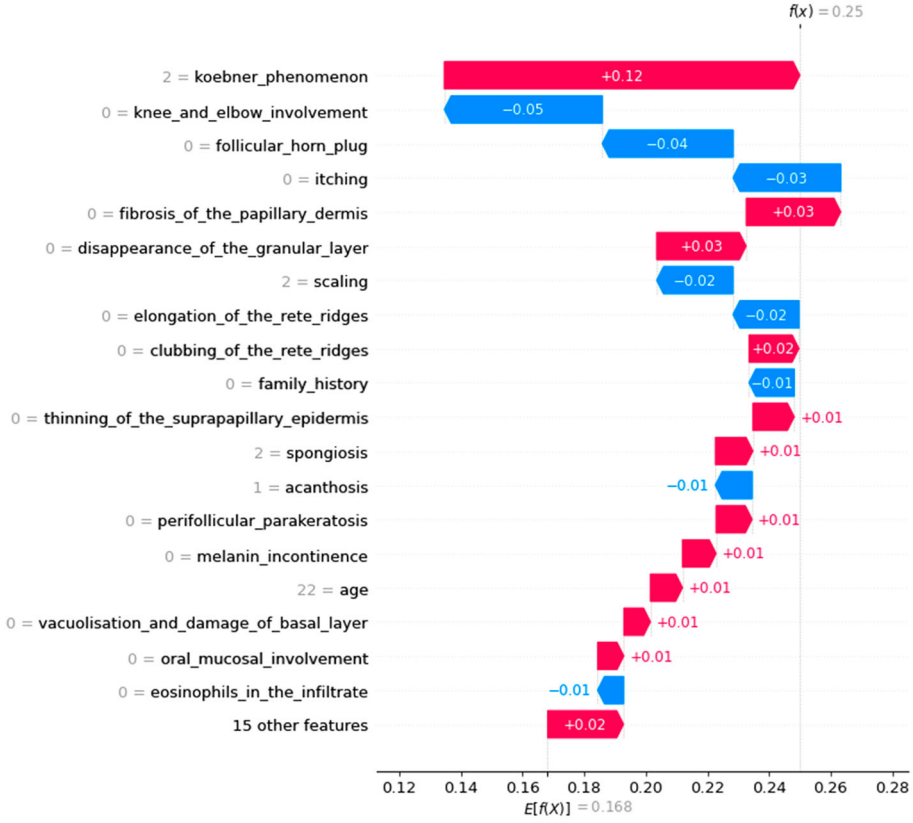
In Figure 6(f), the value given in the bottom of the figure $E[f(x)] = 0.162$ is the result from the null model. The value $f(x) = 0.14$ on the top of the model indicates $f(x) = 0.14 = E[f(x)] +$ impact of each feature.

So, the prediction can be explained for subject 20, is explained by following a combination of features:

$$
\begin{aligned}
f(x) = 0.14 = 0.162 + impact[&(follicular\ horn\ plug = 1) \\
&+ (perifollicular\ para\,\mathrm{ker}\,atosis = 2) + (age = 8) \\
&+ (follicular\_paules = 3) \\
&+ (knee\_and\_elbow\_invovement = 1) \\
&+ (family\_history = 0) \\
&+ (fibrosis\_of\_the\_papillary\_demis = 0) \\
&+ (itching = 2) + (koebner\_phenomenon = 0) \\
&+ (scaling = 1) \\
&+ (elongation\_of\_the\_rete\_ridges = 1) \\
&+ (acanthosis = 1) \\
&+ (thining\_of\_the\_\,\mathrm{sup}\,rpapillary\_epidemis = 0) \\
&+ (esoinophils\_in\_the\_\,\mathrm{inf}\,iltrate = 0) \\
&+ (oral\_mu\,\mathrm{cos}\,al\_invovement = 0) \\
&+ vacuolization\_and\_demage\_of\_
\end{aligned}
$$

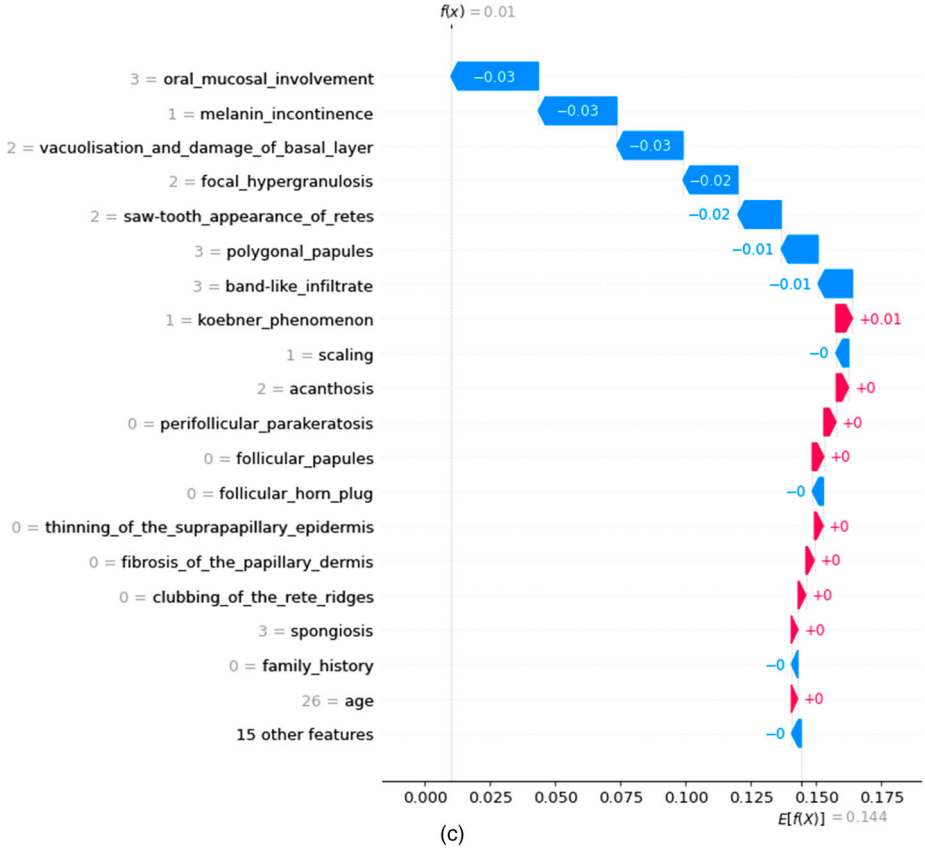


(a)

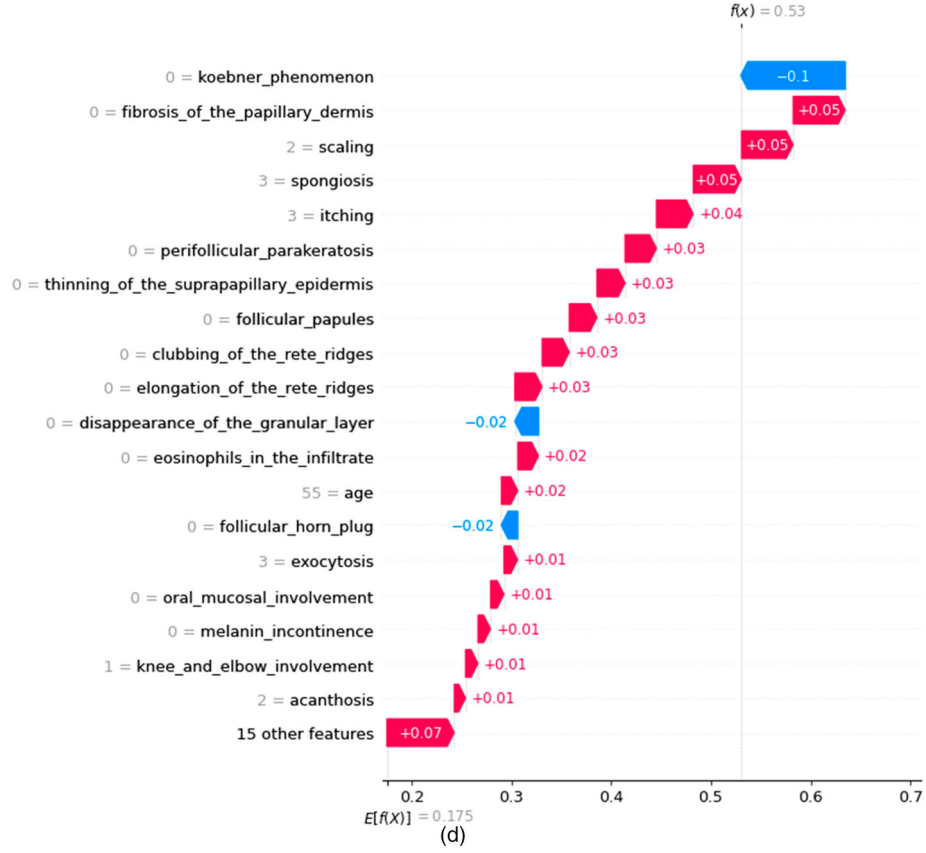


(b)

Figure 6: Waterfall plot of ESDs (a) Psoriasis, (b) Seborrheic Dermatitis, (c) Lichen Planus, (d) Pityriasis Rosea, (e) Chronic Dermatitis, (f) Pityriasis Rubra Pilaris

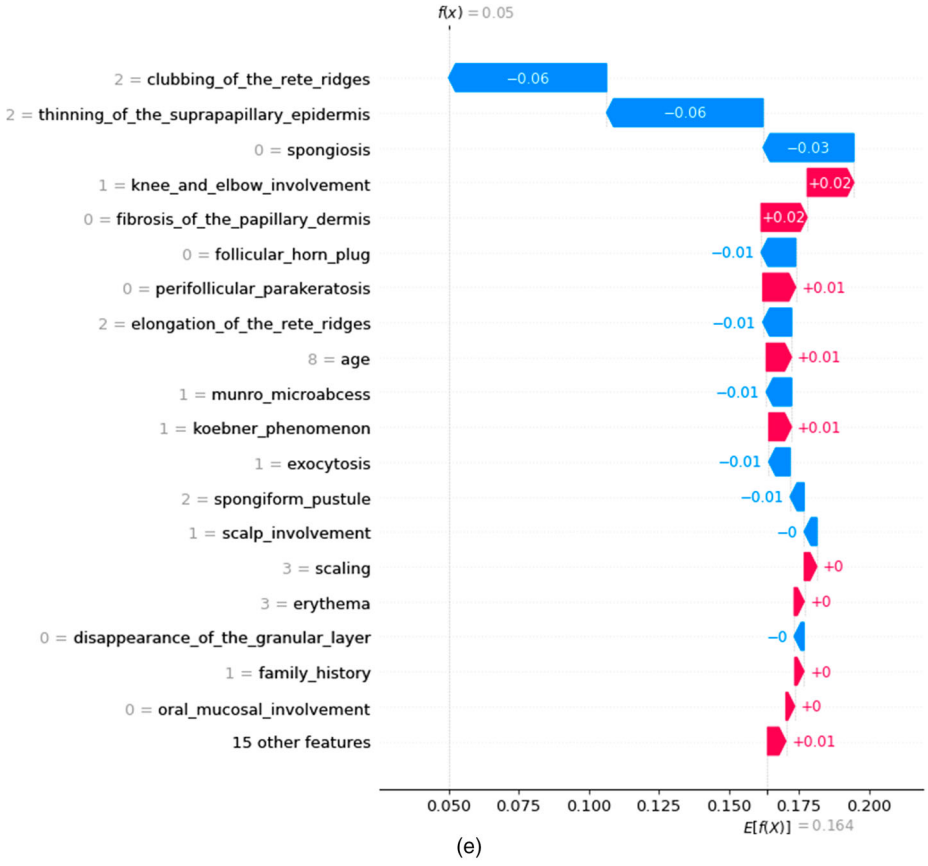


(c)

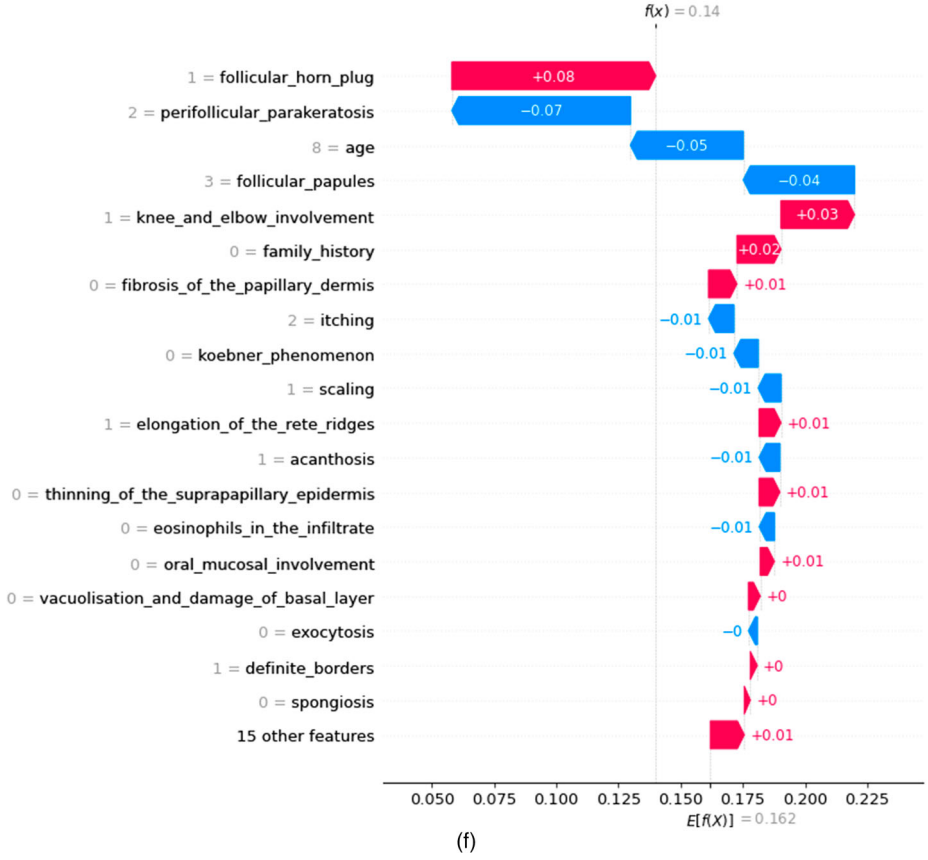


(d)

Figure 6: Continued



(e)



(f)

Figure 6: Continued

$$
\begin{aligned}
&basal\_layer = 0 \\
&+ (exocytosis = 0) + (definite\_borders = 1) \\
&+ (spongiosis = 0) + (15\ other\ features)] \\
f(x) = 0.14 = 0.162 + [&(+0.08) + (-0.07) + (-0.05) \\
&+ (-0.04) + (+0.03) + (0.02) + (0.01) \\
&+ (-0.01) + (-0.01) + (-0.01) \\
&+ (+0.01) + (-0.01) + (0.01) + (-0.01) \\
&+ (0.01) + (0) + (-0) + (0) + (0) \\
&+ (0.01)] \\
f(x) = 0.14 = 0.162 + &[-.022]
\end{aligned}
$$

Next, we focus on the dependence of individual features on ESDs to check whether feature values are constant, or whether their value varies with other features. In Figure 5, the value and impact of features are discussed. To extract

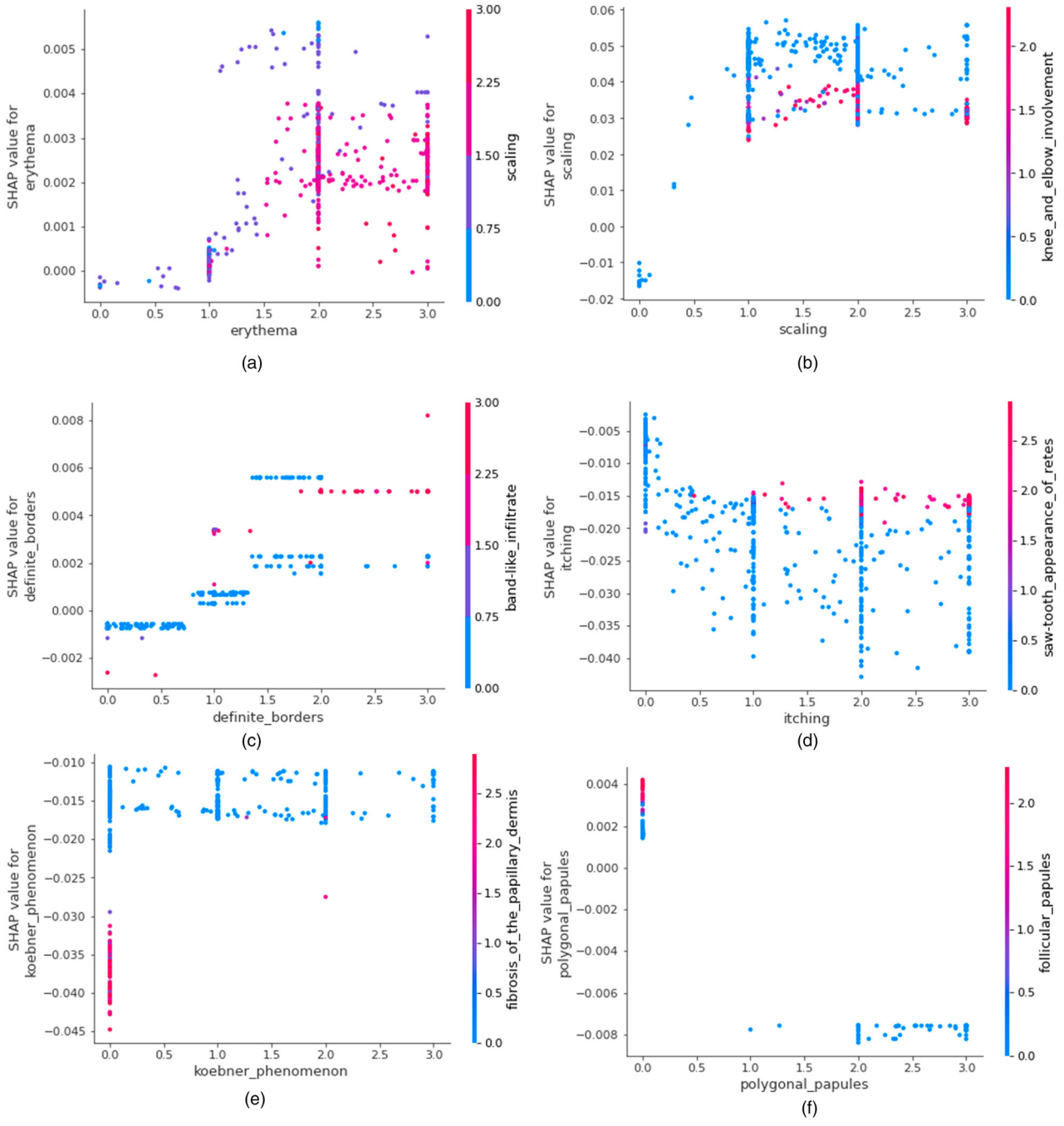


Figure 7: Dependence plot ESDs with feature Interaction, (a) Psoriasis, (b) Seborrheic Dermatitis, (c) Lichen Planus, (d) Pityriasis Rosea, (e) Chronic Dermatitis, (f) Pityriasis Rubra Pilaris

the exact relationship SHAP dependence plots are plotted as shown in Figure 7. It provides the comparative analysis of a feature with another feature and shows if these two features have an interaction effect or not. This analysis proves to be very useful for the selecting most prominent features for ESD diagnosis. The x-axis of the plot shows the feature, and the y-axis represents the predicted SHAP value of that feature. Figure 7(a) suggests that the higher the value of erythema, the higher the chances of Psoriasis. Here, the spread of instances on the same horizontal locations suggests that while predicting Psoriasis, erythema interacts with other features "scaling". So, this figure suggests that the chances of having Psoriasis are high when erythema values are high with low symptoms of scaling. Figure 7(b) says that mild to high scaling in skin indicates higher chances of Seborrheic Dermatitis, even in moderate involvement of knee and elbow joints. Figure 7(c) suggests if there is a high chance of lichen planus lesions with well definite borders and band-like infiltrate at dermal-epidermal junctions. There is a very low chance of Pityriasis Rosea if it is mild itchy with a "sawtooth" pattern of the retes (refer to Figure 7(d)). Figure 7(e) suggests that the presence of a linear exposure or irritation in the papillary dermis's fibrosis points to a reduced outcome of having Chronic Dermatitis. It has further been observed that higher values of polygonal papules reduce the chance of Pityriasis Rubra Pilaris as shown in Figure 7(f), but in the absence of polygonal papules, if follicular papules are present the chances of Pityriasis Rubra Pilaris increases. To understand the effect of all the features collectively, the SHAP values are summed up to understand the decision or the effect of features on the decision as shown in Figure 8.

Each feature value in Figure 8, represents a force that increases/decreases the model's prediction. "Baseline" represents the average of all the predictions made by the model. Red arrows push the model's prediction while the blue arrow lowers down the prediction, where the size of the arrow represents the magnitude of force. Two subjects randomly chosen are considered for interpretation of prediction as depicted in Figure 8. The baseline probability for Psoriasis is 0.1638, while the model prediction for the first subject is 0.05 and less than the baseline value. The risk of Psoriasis increased with mild symptoms in knee and elbow involvement and the absence of fibrosis of the papillary dermis. This risk is offset by intermediate values of clubbing of the rete ridges involvement, thinning of the suprapapillary epidermis, elongation of the rete ridges, and absence of spongiosis and follicular horn plug. The second subject has a higher risk of Seborrheic Dermatitis with a predicted value of 0.53 than the baseline value of 0.1753. High-value spongiosis, itching, and scaling with the absence of fibrosis of the papillary dermis increased the risk of Seborrheic Dermatitis in subject 2. In this work, a total of 100 trees are trained to predict one of the ESDs and interpret the prediction globally as well as local.

From the recent studies, authors are focused on the classification accuracy of ensembled [65] and Bayesian optimization-based models [66] and achieved 99.45 and 99.07 testing accuracy respectively on the same data set. Arjaria *et al.* [67] used ensembled methods on same dataset and achieved 94% classification accuracy.

Verma *et al.* [24] used Support Vector Machines to classify ESD with a classification accuracy of 97.2% and identified important features using Univariate features, Feature Importance, and Correlation & Heat Map. Listing of important features:

- Univariate feature important table
  - Band-like infiltrate
  - Perifollicular parakeratosis
  - Fibrosis of the papillary dermis
  - Vacuolization and damage of basal layer
  - polygonal papules

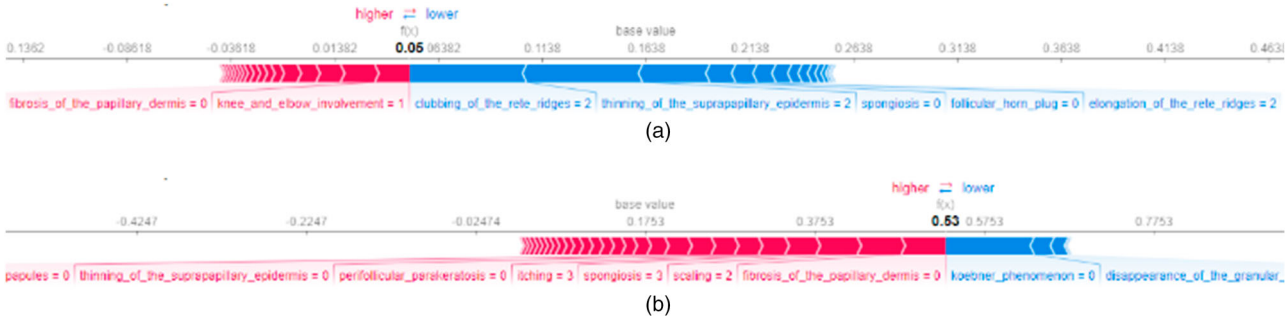


Figure 8: Force plot for understanding the prediction for randomly chosen two subjects

Table 5: Summary of feature involvement in each ESD

Feature	Psoriasis	Seborrheic Dermatitis	Lichen Planus	Pityriasis Rosea	Chronic Dermatitis	Pityriasis Rubra Pilaris
Scaling	L	H + I		L + R	M + R	L + R
Erythema	D					
Definite borders						L
Polygonal papules			M + II			
Koebner phenomenon		H + R	L	H + I	M + R	M + R
Itching	L	M + I		M + R		L + R
Follicular papules	L	M + DI	L	L + DI		H + R
Scalp involvement						
Knee and elbow	L	M + DI		M + R		M + I
Oral mucosal	L	M + DI	H + II	L		
Family history				L + R		M + I
Age	L	M + I		L		H + R
Melanin incontinence	L	M + DI	H + II	M + DI		
Fibrosis of the papillary dermis	M + DI	H+ DI	L	H + DI	H + I	M + DI
PNL infiltrate	L	M + I		L + R	L	
Eosinophils in the infiltrate						
Exocytosis						
Parakeratosis						L
Hyperkeratosis						
Acanthosis				L + R		L + R
Clubbing of the rete ridges involvement	H + II	M + DI	L	M + DI	M + DI	
Spongiform pustule						
Thinning of the suprapapillary epidermis	H + II	M + DI	L	M + DI		M + DI
Elongation of the rete ridges	M + II	M + DI		M + R	M + II	L
Munromicroabscess						
Vacuolization and damage of basal layer	L	M	H + II	M + DI		
Disappearance of the granular layer	L	M + R		M + I		
Focal hypergranulosis	L	M	M + II	L + DI		
Spongiosis	M + R	M + I		L	M + R	L
Perifollicular parakeratosis	M + DI	M + DI	L	M + DI	L + DI	H + II
Follicular horn plug	M + R	L + R	L	M + R	L + R	H + I
Saw-tooth appearance of rete ridges	L	L	M + II			
Inflammatory mononuclear infiltrate						
Band-like infiltrate			M + II			

Notes: Where table can be read by using the following nomenclature: II: Chances Increase with Increase in Value; DI: Chances decrease with an Increase in Value; I: No matter value is high or low, the chances of increase; R: No matter value is high or low, chances reduce; L: L affects the prediction; H: High effect on prediction; M: Medium/Intermediate effect on prediction; D: Affecting prediction when interacting with another feature.

- Important Attributes using feature importance
  - Clubbing of the rete ridges involvement
  - fibrosis of the papillary dermis
  - polygonal papules
  - band-like infiltrate
  - thinning of the suprapapillary epidermis
- Correlation and heat map
  - Thinning of the suprapapillary epidermis
  - Clubbing of the rete ridges involvement
  - PNL infiltrate
  - Scalp involvement
  - Fibrosis of the papillary dermis

In this study, we achieved 97.8% classification accuracy with Random Forest. The major focus of the study is explainability. The impact of features is explained that can be used by health practitioners to describe what is going wrong with patients. Both predictive strengths and the transparency of the models are evaluated critically to get actionable insights. The summary of the overall study is depicted in Table 5.

5. CONCLUSION

Artificial intelligent models are widely used in the domain of healthcare. Predictions of AI models are sufficiently accurate; still, many healthcare practitioners avoid the use of such models because of the black-box nature of the machine learning tools involved. The proposed work is a small step to explain the results of machine learning models in the domain of healthcare. Erythemato-Squamous Diseases symptoms are classified using XGBoost and Random Forest models and then the effect of each feature is explained for the prediction of the disease locally as well as globally in this paper. ESDs have overlapping features *i.e.* different types of ESDs have similar features, hence the major contribution of this work lies in explaining the presence and impact of each feature to understand the etiology underlying the development of the disease. The unboxing of black-box behavior is summarized in Table 5. This work highlighted the interpretability of decisions that promotes the trust of medical practitioners and patients in artificial intelligence. Additionally, in the future, we can use dermatology images

and integrate deep neural networks with XAI. The fusion of AI and healthcare can improve the quality, cost, and reach of health care while providing support to healthcare practitioners.

NOTE

1. Available at https://archive.ics.uci.edu/ml/datasets/dermatology with Source Ilter and Guvenir [34].

ACKNOWLEDGEMENTS

Dataset is downloaded from UCI Machine Learning Repository by D Dua, C Graff (2019), Irvine, CA: University of California, School of Information and Computer Science, available at http://archive.ics.uci.edu/ml.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author(s).

ORCID

Gyanendra Chaubey http://orcid.org/0000-0003-0583-2618

REFERENCES

1. A. M. Elsayad, M. Al-Dhaifallah, and A. M. Nassef, "Analysis and diagnosis of erythemato-squamous diseases using CHAID Decision Trees," in *15th International Multi-Conference on Systems, Signals & Devices*, 2018, pp. 252–62. DOI:10.1109/SSD.2018.8570553.

2. N. Badrinath, G. Gopinath, K. S. Ravichandran, J. Premaladha, and R. Krishankumar, "Classification and prediction of erythemato-squamous diseases through tensor-based learning," *Proc. Natl. Acad. Sci. India A Phys. Sci.*, Vol. 90, pp. 327–35, 2020. DOI:10.1007/s40010-018-0563-x.

3. L. Nanni, "Letters: An ensemble of classifiers for the diagnosis of erythemato-squamous diseases," *Neurocomputing*, Vol. 69, no. 7–9, pp. 842–5, 2006. DOI:10.1016/j.neucom.2005.09.007.

4. E. Vayena, A. Blasimme, and I. G. Cohen, "Machine learning in medicine: Addressing ethical challenges," *PLoS Med.*, Vol. 15, no. 11, 2018. DOI:10.1371/journal.pmed.1002689.g001.

5. S. Reddy, S. Allan, S. Coghlan, and P. Cooper, "A governance model for the application of AI in health care," *J. Am. Med. Inform. Assoc.*, Vol. 27, no. 3, pp. 491–7, 2020. DOI:10.1093/jamia/ocz192.

6. S. Sunarti, F. Fadzlul Rahman, M. Naufal, M. Risky, K. Febriyanto, and R. Masnina, "Artificial intelligence in healthcare: Opportunities and risk for future," *Gac. Sanit.*, Vol. 35, pp. 67–70, 2021. DOI:10.1016/j.gaceta.2020.12.019.

7. G. Yang, Q. Ye, and J. Xia, "Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond," *Inf. Fusion.*, Vol. 77, pp. 29–52, 2022. DOI:10.1016/j.inffus.2021.07.016.

8. A. Gomolin, E. Netchiporouk, R. Gniadecki, and V. I. Litvinov, "Artificial intelligence applications in dermatology: Where do we stand?," *Front. Med.*, Vol. 7, pp. 100, 2020. DOI:10.3389/fmed.2020.00100.

9. S. M. Lundberg, and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, Vol. 30, I. Guyon, *et al.*, Eds. Curran Associates, Inc., 2017.

10. G. Arora, A. K. Dubey, Z. A. Jaffery, and A. Rocha, "Bag of feature and support vector machine based early diagnosis of skin cancer," *Neural Comput. Appl.*, 1–8, 2020. DOI:10.1007/s00521-020-05212-y.

11. P. Kharazmi, M. I. AlJasser, H. Lui, Z. J. Wang, and T. K. Lee, "Automated detection and segmentation of vascular structures of skin lesions seen in dermoscopy, with an application to basal cell carcinoma classification," *IEEE. J. Biomed. Health. Inform.*, Vol. 21, no. 6, pp. 1675–84, 2016. DOI:10.1109/JBHI.2016.2637342.

12. B. Cheng, R. J. Stanley, W. V. Stoecker, and K. Hinton, "Automatic telangiectasia analysis in dermoscopy images using adaptive critic design," *Skin Res. Technol.*, Vol. 18, no. 4, pp. 389–96, 2012. DOI:10.1111/j.1600-0846.2011.00584.x.

13. L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE Trans. Med. Imaging*, Vol. 36, no. 4, pp. 994–1004, 2016. DOI:10.1109/TMI.2016.2642839.

14. C. Sagar, and L. M. Saini, "Color channel based segmentation of skin lesion from clinical images for the detection of melanoma," in *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems* (ICPEICES), July 2016, pp. 1–5. DOI:10.1109/ICPEICES.2016.7853624.

15. M. A. Taufiq, N. Hameed, A. Anjum, and F. Hameed, "m-Skin doctor: A mobile enabled system for early melanoma skin cancer detection using support vector machine," in *Ehealth 360°*, Cham: Springer, 2017, pp. 468–475.

16. Y. Yuan, M. Chao, and Y.-C. Lo, "Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance," *IEEE Trans. Med. Imaging*, Vol. 36, no. 9, pp. 1876–86, 2017. DOI:10.1109/TMI.2017.2695227.

17. A. Tyagi, and R. Mehra, "An optimized CNN based intelligent prognostics model for disease prediction and classification from dermoscopy images," *Multimed. Tools. Appl.*, Vol. 79, no. 35, pp. 26817–35, 2020. DOI:10.1007/s11042-020-09074-3.

18. N. C. F. Codella, Q.-B. Nguyen, S. Pankanti, D. A. Gutman, B. Helba, and A. Halpern, "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM J. Res. Dev.*, Vol. 61, no. 4/5, pp. 1–5, 2017.

19. F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng, and A. Bovik, "Melanoma classification on dermoscopy images using a neural network ensemble model," *IEEE Trans. Med. Imaging*, Vol. 36, no. 3, pp. 849–58, 2016. DOI:10.1109/TMI.2016.2633551.

20. T. Y. Tan, L. Zhang, C. P. Lim, B. Fielding, Y. Yu, and E. Anderson, "Evolving ensemble models for image segmentation using enhanced particle swarm optimization," *IEEE. Access.*, Vol. 7, pp. 34004–19, 2019. DOI:10.1109/ACCESS.2019.2903015.

21. K. Polat, and S. Günecs, "A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems," *Expert. Syst. Appl.*, Vol. 36, no. 2, pp. 1587–92, 2009. DOI:10.1016/j.eswa.2007.11.051.

22. C.-L. Chang, and C.-H. Chen, "Applying decision tree and neural network to increase quality of dermatologic diagnosis," *Expert. Syst. Appl.*, Vol. 36, no. 2, pp. 4035–41, 2009. DOI:10.1016/j.eswa.2008.03.007.

23. A. K. Verma, S. Pal, and S. Kumar, "Prediction of skin disease using ensemble data mining techniques and feature selection method – a comparative study," *Appl. Biochem. Biotechnol.*, Vol. 190, no. 2, pp. 341–59, 2020. DOI:10.1007/s12010-019-03093-z.

24. A. K. Verma, and S. Pal, "Prediction of skin disease with three different feature selection techniques using stacking ensemble method," *Appl. Biochem. Biotechnol.*, Vol. 190, no. 2, pp. 341–59, 2020. DOI:10.1007/s12010-019-03222-8.

25. J. Amin, M. Sharif, M. Yasmin, T. Saba, M. A. Anjum, and S. L. Fernandes, "A new approach for brain tumor segmentation and classification based on score level fusion using transfer learning," *J. Med. Syst.*, Vol. 43, no. 11, pp. 1–16, 2019. DOI:10.1007/s10916-019-1453-8.

26. T. Saba, A. Sameh, F. Khan, S. A. Shad, and M. Sharif, "Lung nodule detection based on ensemble of hand crafted and deep features," *J. Med. Syst.*, Vol. 43, no. 12, pp. 1–12, 2019. DOI:10.1007/s10916-019-1455-6.

27. T. Saba, A. Sameh Mohamed, M. El-Affendi, J. Amin, and M. Sharif, "Brain tumor detection using fusion of hand crafted and deep learning features," *Cogn. Syst. Res.*, Vol. 59, pp. 221–30, 2020. DOI:10.1016/j.cogsys.2019.09.007.

28. M. A. Khan, M. Sharif, T. Akram, M. Raza, T. Saba, and A. Rehman, "Hand-crafted and deep convolutional neural network features fusion and selection strategy: An application to intelligent human action recognition," *Appl. Soft. Comput.*, Vol. 87, pp. 105986, 2020. DOI:10.1016/j.asoc.2019.105986.

29. R. B. Oliveira, A. S. Pereira, and J. M. R. S. Tavares, "Computational diagnosis of skin lesions from dermoscopic images using combined features," *Neural Comput. Appl.*, Vol. 31, no. 10, pp. 6091–111, 2019. DOI:10.1007/s00521-018-3439-8.

30. F. Afza, M. A. Khan, M. Sharif, and A. Rehman, "Microscopic skin laceration segmentation and classification: A framework of statistical normal distribution and optimal feature selection," *Microsc. Res. Tech.*, Vol. 82, no. 9, pp. 1471–88, 2019. DOI:10.1002/jemt.23301.

31. S. A. Khan, M. Nazir, M. A. Khan, T. Saba, K. Javed, and A. Rehman, "Lungs nodule detection framework from computed tomography images using support vector machine," *Microsc. Res. Tech.*, Vol. 82, no. 8, pp. 1256–66, 2019. DOI:10.1002/jemt.23275.

32. S. Iqbal, M. U. Ghani Khan, T. Saba, Z. Mehmood, N. Javaid, and A. Rehman, "Deep learning model integrating features and novel classifiers fusion for brain tumor segmentation," *Microsc. Res. Tech.*, Vol. 82, no. 8, pp. 1302–15, 2019. DOI:10.1002/jemt.23281.

33. S. Chatterjee, D. Dey, and S. Munshi, "Optimal selection of features using wavelet fractal descriptors and automatic correlation bias reduction for classifying skin lesions," *Biomed. Signal. Process. Control.*, Vol. 40, pp. 252–62, 2018. DOI:10.1016/j.bspc.2017.09.028.

34. J. Jaworek-Korjakowska, and P. Kłeczek, "Automatic classification of specific melanocytic lesions using artificial intelligence," *BioMed Res. Int.*, 2016. DOI:10.1155/2016/8934242.

35. S. Woo, and C. Lee, "Incremental feature extraction based on decision boundaries," *Pattern Recognit.*, Vol. 77, pp. 65–74, 2018. DOI:10.1016/j.patcog.2017.12.010.

36. M. S. Haji, M. H. Alkawaz, A. Rehman, and T. Saba, "Content-based image retrieval: A deep look at features prospectus," *Int. J. Comput. Vis. Robot.*, Vol. 9, no. 1, pp. 14–38, 2019. DOI:10.1504/IJCVR.2019.098004.

37. Z. Mehmood, F. Abbas, T. Mahmood, M. A. Javid, A. Rehman, and T. Nawaz, "Content-based image retrieval based on visual words fusion versus features fusion of local and global features," *Arab. J. Sci. Eng.*, Vol. 43, no. 12, pp. 7265–84, 2018. DOI:10.1007/s13369-018-3062-0.

38. M. Nasir, M. Attique Khan, M. Sharif, I. U. Lali, T. Saba, and T. Iqbal, "An improved strategy for skin lesion detection and classification using uniform segmentation and feature selection-based approach," *Microsc. Res. Tech.*, Vol. 81, no. 6, pp. 528–43, 2018. DOI:10.1002/jemt.23009.

39. H. A. Güvenir, G. Demiröz, and N. Ilter, "Learning differential diagnosis of erythemato-squamous diseases

using voting feature intervals," *Artif. Intell. Med.*, Vol. 13, no. 3, pp. 147–65, 1998. DOI:10.1016/s0933-3657(98)00028-1.

40. K. Maghooli, M. Langarizadeh, L. Shahmoradi, M. Habibi-Koolaee, M. Jebraeily, and H. Bouraghi, "Differential diagnosis of erythmato-squamous diseases using classification and regression tree," *Acta Inform. Med.*, Vol. 24, no. 5, pp. 338, 2016. DOI:10.5455/aim.2016.24.338-342.

41. J. Xie, W. Xie, C. Wang, and X. Gao, "A novel hybrid feature selection method based on IFSFFS and SVM for the diagnosis of erythemato-squamous diseases," in *Proceedings of the First Workshop on Applications of Pattern Analysis*, 2010, pp. 142–51.

42. S. O. Olatunji, and H. Arif, "Identification of erythemato-squamous skin diseases using extreme learning machine and artificial neural network," *ICTACT J. Softw. Comput.*, Vol. 4, no. 1, pp. 627–32, 2013. DOI:10.21917/ijsc.2013.0090.

43. K. S. Ravichandran, B. Narayanamurthy, G. Ganapathy, S. Ravalli, and J. Sindhura, "An efficient approach to an automatic detection of erythemato-squamous diseases," *Neural Comput. Appl.*, Vol. 25, no. 1, pp. 105–14, 2014. DOI:10.1007/s00521-013-1452-5.

44. K. Danjuma, and A. O. Osofisan, "Evaluation of predictive data mining algorithms in erythemato-squamous disease diagnosis." 2015.

45. M. J. Abdi, and D. Giveki, "Automatic detection of erythemato-squamous diseases using PSO–SVM based on association rules," *Comput. Biol. Med.*, Vol. 35, no. 5, pp. 421–33, 2005. DOI:10.1016/j.engappai.2012.01.017.

46. E. D. Übeyli, "Multiclass support vector machines for diagnosis of erythemato-squamous diseases," *Expert. Syst. Appl.*, Vol. 35, no. 4, pp. 1733–40, 2008. DOI:10.1016/j.eswa.2007.08.067.

47. E. D. Übeylı, and İ Güler, "Automatic detection of erythemato-squamous diseases using adaptive neuro-fuzzy inference systems," *Comput. Biol. Med.*, Vol. 35, no. 5, pp. 421–33, 2005. DOI:10.1016/j.compbiomed.2004.03.003.

48. L. Polkowski, *Rough Sets*. Heidelberg: Physica-Verlag, 2002. DOI:10.1007/978-3-7908-1776-8.

49. G. Peters, P. Lingras, D. Ślęzak, and Y. Yao. *Rough Sets: Selected Methods and Applications in Management and Engineering*. London: Springer London, 2012. DOI:10.1007/978-1-4471-2760-4.

50. T. Y. Lin, and N. Cercone. *Rough Sets and Data Mining*. Boston, MA: Springer US, 1997. DOI:10.1007/978-1-4613-1461-5.

51. J. Liu, Y. Lin, Y. Li, W. Weng, and S. Wu, "Online multi-label streaming feature selection based on neighborhood rough set," *Pattern Recognit.*, Vol. 84, pp. 273–87, 2018. DOI:10.1016/j.patcog.2018.07.021.

52. D. Moitra, and R. K. Mandal, "Automated grading of non-small cell lung cancer by fuzzy rough nearest neighbour method," *Netw. Model. Anal. Health Inform. Bioinform.*, Vol. 8, no. 1, pp. 1–9, 2019. DOI:10.1007/s13721-019-0204-6.

53. A. K. Sinha, and N. Namdev, "Feature selection and pattern recognition for different types of skin disease in human body using the rough set method," *Netw. Model. Anal. Health Inform. Bioinform.*, Vol. 9, no. 1, pp. 1–11, 2020. DOI:10.1007/s13721-020-00232-z.

54. D. Shin, "The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI," *Int. J. Hum. Comput. Stud.*, Vol. 146, pp. 102551, 2021. DOI:10.1016/j.ijhcs.2020.102551.

55. L. Chazette, and K. Schneider, "Explainability as a non-functional requirement: Challenges and recommendations," *Requir. Eng.*, Vol. 25, no. 4, pp. 493–514, 2020. DOI:10.1007/s00766-020-00333-1.

56. A. Holzinger, C. Biemann, C. S. Pattichis, and D. B. Kell, "What do we need to build explainable {AI} systems for the medical domain?" arXiv preprint, 2017. arXiv:1712.09923.

57. S. Tonekaboni, S. Joshi, M. D. McCradden, and A. Goldenberg, "What clinicians want: Contextualizing explainable machine learning for clinical end use," in *Machine Learning for Healthcare Conference*, F. Doshi-Velez, *et al.*, Eds., October 2019, pp. 359–80.

58. A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, "Causability and explainability of artificial intelligence in medicine," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, Vol. 9, no. 4, pp. 1312, 2019. DOI:10.1002/widm.1312.

59. M. S. Hossain, G. Muhammad, and N. Guizani, "Explainable AI and mass surveillance system-based healthcare framework to combat COVID-I9 like pandemics," *IEEE Netw*, Vol. 34, no. 4, pp. 126–32, 2020.

60. E. Khodabandehloo, D. Riboni, and A. Alimohammadi, "HealthXAI: Collaborative and explainable AI for supporting early diagnosis of cognitive decline," *Future Gener. Comput. Syst.*, Vol. 116, pp. 168–89, 2021. DOI:10.1016/j.future.2020.10.030.

61. S. M. Lauritsen, M. Kristensen, M. V. Olsen, M. S. Larsen, K. M. Lauritsen, and M. J. Jørgensen, "Explainable artificial intelligence model to predict acute critical illness from electronic health records," *Nat. Commun.*, Vol. 11, no. 1, pp. 1–11, 2020. DOI:10.1038/s41467-020-17431-x.

62. A. P. Carrieri, N. Haiminen, S. Maudsley-Barton, L.-J. Gardiner, B. Murphy, and A. E. Mayes, "Explainable AI reveals changes in skin microbiome composition linked to phenotypic differences," *Sci. Rep.*, Vol. 11, no. 1, pp. 1–18, 2021. DOI:10.1038/s41598-021-83922-6.

63. T. K. Ho, "Random decision forests," in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Vol. 1, 1995, pp. 278–82. DOI:10.1109/ICDAR.1995.598994.

64. T. Chen, and C. Guestrin, "XGBoost," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–94. DOI:10.1145/2939672.2939785.

65. S. Shastri, P. Kour, and S. Kumar, "GBoost: A novel grading-AdaBoost ensemble approach for automatic identification of erythemato-squamous disease," *Int. J. Inf. Technol.*, Vol. 13, no. 3, pp. 959–71, 2021. DOI:10.1007/s41870-020-00589-4.

66. 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, pp. 103223, 2022. DOI:10.1016/j.bspc.2021.103223.

67. S. K. Arjaria, V. Raj, S. Kumar, P. Shrivastava, M. Kumar, and J. S. Cherian, "Prediction of skin diseases using machine learning," in *Ethical Implications of Reshaping Healthcare with Emerging Technologies*, T. H. Musiolik, and A. Dingli, Eds. IGI Global, 2022, pp. 154–78. DOI:10.4018/978-1-7998-7888-9.ch008.

AUTHORS

Abhishek Singh Rathore received his BE degree in computer science and engineering from Rajiv Gandhi Technical University, Bhopal in 2005, and ME degree in computer engineering from Devi Ahilya University, Indore in 2008. He received PhD in computer science and engineering from Maulana Azad National Institute of Technology, Bhopal in 2015. His areas of interest are machine learning, probabilistic models, and quantum computing.

Email: abhishekatujjain@gmail.com

Siddhartha Kumar Arjaria received his B Tech degree in computer science and engineering from Institute of Technology, GGU, Bilaspur in 2001 and M Tech degree in computer technology from Pt. Ravishankar University Raipur in 2006. He received PhD in computer science and engineering, from National Institute of Technology, Bhopal, India in 2020. His areas of interest are data mining, soft computing and machine learning.

Email: arjarias@gmail.com

Manish Gupta received his BE degree in electronics engineering from Jiwaji University, Gwalior, India in 2000 and M Tech degree in digital communication from UPTU, Lucknow India in 2006. He received PhD in electronics engineering from RTU, Kota, India in 2016. His area of interest is signal and image processing.

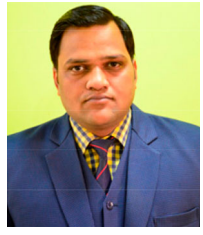
Corresponding author. Email: manish.gupta@gla.ac.in

Gyanendra Chaubey received his BTech degree in information technology from Rajkiya Engineering College, Banda, India in 2020. His areas of interest are machine learning, deep learning, data mining, and artificial intelligence.

Email: gyanendrachaubey68@gmail.com

Amit Kumar Mishra received his BTech degree in information technology from Mahatma Gandhi Chitrakoot Gramodaya Vishwavidyalaya, Chitrakoot, Satna, India in 2006 and MTech degree in computer science and engineering from VTU, Belgaum, Karnataka, India in 2009. He received PhD in computer science and engineering from Banasthali Vidyapith, Tonk, Rajasthan, India in 2021. His areas of interest are social network analysis, soft computing and wireless sensor network.

Email: akmishra1@gwa.amity.edu



Vikram Rajpoot received his BE degree in information technology from Madhav Institute of Technology & Science, Gwalior, India in 2010 and MTech degree in information technology from Samrat Ashok Technological Institute, Vidisha, India in 2012. He received PhD in information technology from MGCGV Chitrakoot Satna, India in 2019. His areas of interest are frequent pattern mining, machine learning and artificial intelligence.

Email: vikramraj@mitsgwalior.in