International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I11P202
Volume 70 Issue 11, 10-18, November 2022

Original Article

A Novel Approach to Analysis Consequence of Climate Changes on Erythemato-Squamous Diseases using Machine Learning Algorithms

Shaik Abdul Khalandar Basha[1], P. M. Durai Raj Vincent[2*], K. Srividya[3], Pusarla Samyuktha[4], Jessu Madhumathi[5]

[1,2]*School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India*
[3,4,5]*Department of CSE, Vasavi College of Engineering, Hyderabad, India.*

[2*]*Corresponding Author : pmvincent@vit.ac.in*

Received: 17 March 2022 Revised: 22 May 2022 Accepted: 02 November 2022 Published: 26 November 2022

Abstract *- Erythemato-squamous Disease is considered an arboviral disease that falls under the subset of skin disease. Nearly 15% of death caused due to skin diseases are getting registered in the WHO IIPC survey, which can be considered a major threat to public health. In our retrospective study, we investigated ESD incidence with 34 attributes and attempted to determine and diagnose the effects of climate attributes on ESD. In the paper, weather data from Chicago and turkey are taken. Among that, 22 attributes from the first data set and 34 attributes from the second dataset are considered, with this total no of 49 unique attributes taken into consideration for the analysis. In this work, the Weather-Method-Diseases framework by using Machine Learning Algorithms. Pearson's correlation coefficient factor is used to measure the strength of a relationship that exists among 49 attributes for the analysis. The finding in our research is presented, and the result analysis shows the superiority of the proposed approach; these 23attributes are highly correlated. These attributes can be considered to be the most deciding factor for predicting six different types of skin diseases in the future using this proposed framework*

Keywords *- Erythemato-squamous Diseases (ESD), W-M-D framework, Machine Learning Algorithms, Pearson's correlation coefficient.*

1. Introduction

Extreme weather and heat events are big threats to developing and developed countries regarding safety and public health[1]. Even though these countries are doing their best to examine climate change and its impacts, comparatively more attention is given to the agriculture, energy, and economic sector than the public health sector.

Therefore, the focus is moving toward improving public health during extreme heat and weather events. Precautions and predictive steps needed for making decisions with guidelines are to be implemented[2][3].

Long heat wave events and high temperatures in a single day may affect public health [4][5]. Some studies suggest a strong association between mortality and morbidity increase due to extreme weather changes [6]. Some of the other studies hint that there will be an effect on cardiovascular and respiration areas respective to age group in these extreme weather conditions [7]. Most studies have considered the health index ultimate and mortality as health indicators in the literature.

According to the literature survey, it is observed that only limited research and studies tried to examine the association between extreme heat weather events and public health(Erythemato-Squamous diseases). A group can treat a sudden climate change of people with extreme heat shock. If the body temperature raises 40'C and above, damage at the cellular level occurs that causes the failure of organs and the possibility of fatality [8] in the United States. Hyperthermia was recorded as the cause of death at 54% due to increased weather events [6]. North and south Korea used to experience the longest and hottest summer in 2012 in a decade which registered total deaths of 975 due to heat wave-related illness.[9]

The data recorded in Chicago city of the United States, from 2009 to 2013 NOAA center, shows that an increase in longer heat periods and intensity was frequently projected [10]. These happenings lead us to analyze how heat weaves or weather events affect Erythemata-Squamous Diseases (skin diseases) and illnesses correlated with extreme heat. So, a constructive implementation plan for public health precaution and predictive programs must be considered. Machine learning algorithms used to classify and analyze diseases (like EDS) will achieve hidden patterns and information.

This work mainly estimates the association between Tmin, Tmax, and Tavg and Erythemata-Squamous Diseases (skin diseases) in hot-absorbed locations in various cities. Weather data from Chicago city and dermatology data from Ankara city are considered for the result analysis.

The rest of the article deals with the theoretical background, followed by findings, research methodology, discussions, and recommendations.

2. Related Work

Climatic factors were chosen for each location and horizon, and 15 invasive ant species were examined on the size of the species. It was done with the effect of climate change to determine whether there will be an increase or decrease in taxonomic species in the coming years [11]. SDM and GCMS climatic models were applied, and Machine learning models such as SVM, SVC, ANN, and Classification Tree were deployed for analysis. The limitation in feature extraction methods has not been examined properly for better classification [12].

A machine learning algorithm for the prediction of sky diffuse and analysis of the sensitivity of metrological attributes has developed. Boosted regression tree model has been applied with logistic regression (97.78 accuracies) algorithm for the correlation of solar attributes like cloud amount, visibility, and air temperature [13]. According to the survey, Australia is mostly affected by skin cancer or skin-related diseases. According to government records and analysis, nearly 15% of deaths are caused due to skin diseases, and climatic parameters are closely correlated. This resulted in 20% country's budget being invested in skin disease prevention detection.

AVONA models are applied for self and response efficiency to gain and loss frame of skin detection and prevention [9]. A hydrological, Agronomical, and dermatological approach for weather data analysis and prediction with the help of ELM, GP, ANFIS, SVM, and ANN algorithms was developed. Soft computing techniques and some error measures like MAE, RMSE, and correlation(R) were used to test the results. ELM performed better with higher accuracy due to the considered factors for the change of environment. The dew points were measured as relative humidity to moisture level and its subset of bulb temperature[14].

The effects of hot weather on skin diseases and other infectious diseases were considered [6]. A systematic review was done on Chikungunya, Malaria, Zika virus, Nail virus, Dengue, and Skin Cancer. The study exhibits the relationship between heat respective to the above diseases mentioned. In China, extensive analysis shows the heat effects on public health. The findings were the type of diseases categorized as Mild illness, severe illness, dizziness, sweating, and repaid pulse loss.

3. Methods

This combines two different datasets; we have collected datasets from UCI:

- http://archive.ics.uci.edu/ml/datasets/dermatology

Dermatology data sets and time series data from:

- https://www.ncdc.noaa.gov/?prior=N of hot weather.

3.1. Proposed Framework

This is the first paper that addresses how climate change affects skin diseases. Through a complete literature study, not even a single paper try to address it; This research domain is related to public health and real-time. The proposed W-A-D framework is trying to address and give a handful of research gaps for the future. The framework includes the activities of effects of climate change, the flow of the W-M-D framework starts with collecting historical data of weather datasets, which is the main centric part of the research. Suppose data availability is more and huge; the analysis process will help move forward. In that case, the next phase of module 2 is the selection of features. As every dataset is inbuilt with N number of features, it needs to be sorted out the most applicable features. In module 3, even more filtration of parameters to acquire better results module 4 is decision-making with the help of feature extraction techniques, and module 5 is in the framework to formulate the model to apply machine learning algorithm.

In the next part of the framework, we will deal with the EDS dataset's approach. Now we have to consider the dermatology data, which consists of 6 different types of skin diseases with 34 features in which one feature is linear as we have 3 levels of extraction, level 0 again sub-categories into 3 phases of extraction seen in figure 1. Selection of features in unsupervised data is difficult, consisting of 6 skin diseases without label data. to tackle such unsupervised data, first, we randomly selected a feature of $M_f$notation to formulate the graph, and next to that, randomization of each feature in 3 levels tries to form the cluster, M_f ……. Mn notation. For better prediction, data is converted into semi-supervised to label the data, which can be used for predicting value. All the 3 level and 3 phase follows the above process to apply a machine learning algorithm and combine the two phases of the framework to formulate a W-N-D framework.

3.1.1. Dermatology Data Set

This data set was posted in 1998 by 2 owners Niles and Guvernir, in the UCI repository, one of the data repositories for research. The data consist of 6 different types of skin diseases, i.e., Psoriasis, Lichen Planes, Cronic Dermatitis, Seboreic Dermatitis, Pityriasis Rosea, and Pityriasis Rubra Pilaris. There is a total of 366 instances with 34 attributes, in which 34 attributes are subdivided into 2 groups, clinical and Histopathological, where the clinical group consists of 11 attributes with extra age group attribute included and the values taken as (0,1,2,3) on the other side, Histopathological consist of 22 attributes in which is tested under microscopic experimentation and the value taken as (0,1,2,3). The total 366 rows and 34 columns with 8 missing values, a linear data type in age group column and missing value, are filled with the mean square method. Using the classification algorithm, 6 types of skin disease are clustered.

3.1.2. Weather Data Set

Time series data have been collected from Chicago's NOAA center for the weather station. This data set consists of 2945 rows and 22 columns. We have to consider columns as 22 attributes with different measurements and values listed in the weather dataset file.NOAA has collected weather data for Chicago city from 05-01-2007 to 30-09-2014 In which data is collected twice a day with min temp, max temp, avg temp, sea level, humidity, pressure, dew point, etc., out of 22 attributes in the data set we have selected the most relevant attribute which is not specific in nature, Chicago is the third most crowded city in the USA.

Furthermore, fall in the region of Illinois state. Due to vast development, the climate has been changed past few decades, and climate change is recorded. We have performed preprocessing phases on the dataset and acquired a clean and ready-to-deploy experiment dataset. At the Chicago weather station, they have collected daily twice, one in the morning and another in the evening, which is recorded as the minimum and maximum temperatures, out of which the average daily temperature is determined.

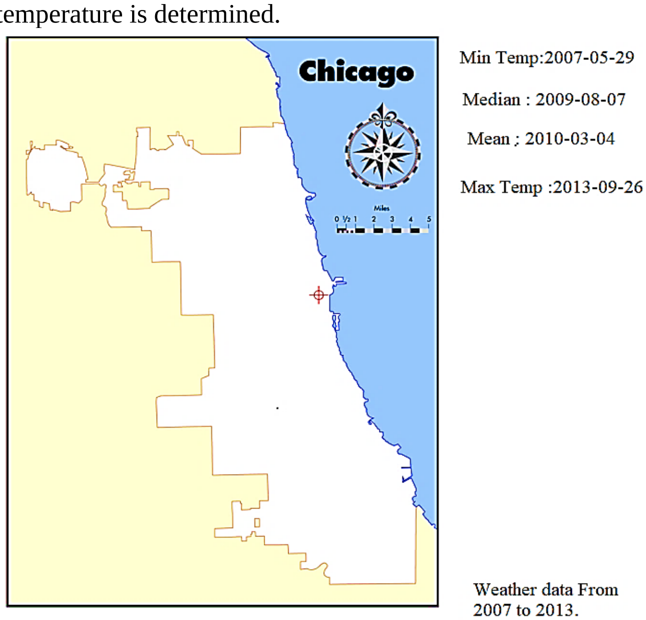


Fig. 1 Geographical map of Chicago

source: From Wikimedia Commons

3.1.3. Statistical Approach

This problem statement is the combination of two different areas: A trans-disciplinary domain effects of hot weather on types of skin diseases. That is classified into 6 different clusters already we have discussed in the above sections. The main aim of this paper is to find if hot weather really affects Erythemato-Squamous Disease; with the statistical help approach, we will explore the finding. Three phases of the statistical approach were followed

- A complete analysis of time series data
- Using classification model for dermatology data
- Merging and finding correlations among them

Lets us first discuss dermatology data; this is unsupervised data with 366 instances with 34 attributes. For the unsupervised data type, we have applied 5 unsupervised algorithms: hear Random Forest, Gaussian Naive Base, Multinomial Naive Base, Bernoulli Naive Base, and Support Vector Clustering. Study analysis of 5 different algorithms has been applied for better classification of the dermatology data set, as we have already discussed that this data set had 6 different types of Erythemato Squamous Diseases. All the algorithms have performed the best error rate and clustered the diseases with respective instances. Every algorithm aims to cluster the diseases. We have divided the data into train and test phases 0.8 and 0.2 for better classification. We have taken a ratio of 8:2. It can be changed on user requirements and names with train labels and test labels to convert unsupervised to supervised data for prediction. The type of disease is by means of coding. On the other hand, we have calculated the Accuracy, Confusion matrix, and Classification report. We have calculated the feature weights and plotted decision boundaries for every classifier; this can be seen in the results and discussion. Now let us examine the other side of the problem statement: time series data. As the data sets consist of 22 attributes of weather data, Herein weather data have been divided into training and testing phase with 0.7 and 0.3 as ratios; after that few preprocessing steps have been applied for cleaning of data in which the data set has some missing Values. Now select the attributes present in the data sets, i.e. (PreciTotal, StnPressure, SeaLeavel, ResultSpeed, ResultDir, AvgSpeed, Cool, Sunrise, Sinsetamd Tavg) to fix some NA's in PrecipTotal and Stnpressure a mean value (0,29.28)is assigned to merge the training and testing phase. Next, scatter the result of weather data intended to calculate the approach applied separately for each data set. A clear analysis of data sets has been explained, combining the two data sets into A single. Find the correlation among them with the help of the Pearson correlation coefficient; this can be seen in the discussion section

3.1.4. Derivation of Multivariate Case Correlation

Let us consider $M_0, M_1, M_2 \ldots M_f$. Are Independent normal derivation and dependent normal derivation.

$M_d = M_0 + \beta_1 M_1 + \ldots\ldots\ldots \beta = \beta^1 M.$

Applying Regression function M_d on M^1 is clearly $\beta^1 M$.

$$Z(M_d) = 1 + \beta^1\beta = 1 + \tau^2 \ldots\ldots \text{(Say 1)}$$

$cor(M_d, M) = \beta$

$cor(M_d, \beta^1 M) = \beta^1\beta = \tau^2$

Now Multiples correlation derivation M_d with Y is

$$Z = {}^{Z}\!/_{(1-Z^2)^{\frac{1}{2}}} \ldots\ldots \text{(Say 2)}$$

At bivariate case

$$\tau = {}^{Z}\!/_{(1-Z^2)^{\frac{1}{2}}} \quad (\tau, P > 0) \ldots\ldots\ldots\ldots \text{(Say 3)}$$

Twice bivariate case

$$\text{Ʈ}_{dd} = \sum M_d^2 - X_n^2/(1-Z^2) \ldots\ldots\ldots \text{(Say 4)}$$

Simple Regression M_d on $M_1, M_2, \ldots\ldots\ldots M_F$ (β unknown)

Sum of Squares Ʈ_{dd} into: regression SS $M_d^l A M_d$ and Residual SS $M_d^l (I-A) M_d$

(Considering Equations 1,2,3,4)

Appling Sum of Squares Regression

$(1-Z^2)^{1/2n} \sum^{\alpha}_{i=0} Z^{2i}/i! \; r(1/2+i)/r(1/2^n) \quad e^{1/2x} \; 1/2^{p+i-1}/2^{1/2p+I} r(1/2^{p+1}) = \sum b_i g_p + 2(x)$

Squared multiple correlation coefficient

$(1-Z^2)^{1/2n} \sum^{\alpha}_{i=0} Z^{2i}/i! \; r(1/2+i)/r(1/2^n) \quad \mu^{1/2p+i-1}(1-\mu)^{1/2(n-p)}-1/B(1/2^{p+i}, ½^{(n-1)})$

$= \sum^{\alpha}_{i=0} b_i \, \mu^{1/2p+i-1}(1-\mu)^{1/2(n-p)}-1/B(1/2^{p+i}, ½^{(n-1)})$

3.1.5. Analysis of Methods

The mean, min, and max temperatures of Chicago have been recorded. We are going to go through the study, i.e., min, temp on 2007-05-09 in ORD terminal Airport Chicago with address 750 in the first quarter in the year 2007-08-24 have recorded in the place of south point Chicago with 514 address, a median of the city in the year 2009-08-07 in the location of south stony with 314 address, mean of the temp recorded in the year 2010-03-04 at North oak plane avenue Chicago at the end of the 3rd quarter in the year 2001-09-16 at West Street and last max temp of the city has recorded in the year 2013-09-16 at North cannon drive Chicago.[13]a complete temperature analysis is related to hot weather, which is the summer period of Chicago from May to September

```
      Date                                                          Address
Min.   :2007-05-29   ORD Terminal 5, O'Hare International Airport, Chicago, IL 60666, USA: 750
1st Qu.:2007-08-24   South Doty Avenue, Chicago, IL, USA                                 : 542
Median :2009-08-07   South Stony Island Avenue, Chicago, IL, USA                         : 314
Mean   :2010-03-04   4100 North Oak Park Avenue, Chicago, IL 60634, USA                  : 185
3rd Qu.:2011-09-16   4200 West 127th Street, Alsip, IL 60803, USA                        : 183
Max.   :2013-09-26   2200 North Cannon Drive, Chicago, IL 60614, USA                     : 163
                     (Other)                                                             :8369
```

Fig. 2 Chicago Min, 1st quart, Median, Mean, 3rd quart, Max Temperature with Location.

The data recorded in the station have been showing Result 1 and the temperature with Tmin =41C, Tmax=57C and Tavg=50, which is normal in the city of Chicago and predicting the value of temperature in the city Tmin=76C, Tmax=96C and Tavg=84Cin the upcoming year of the city another attribute of temp Dew Point(Min, Mean, Median, Max)Heat(Min, mean, median, Max)snowfall and Wet Bulb Length =10566&10566 Total Perfection (Min, median, mean, max) pressure of station with values of (Min, Mean, Median, Max), Sealevel might be one of the ones the changing factor in future of the problem statement. As we are dealing with and working on Erythemato-Squamous diseases, effects can be seen on the sea level. Which will be discussed in a further section Sea Level (Min, Median, Mean, Max), Wind speed receded in summer with wave speed in random direction with the hot image (Min, Median, Mean, Max), the direction of Air (Min.Median.max. mean) and hot air and temperature avgspeed of (Min, Median, Mean, Max) Chicago city commonly tends with mortality rate more in number according to NOAA center USA.

We have examined the weather data analysis and its factors, and the next move to the dermatology dataset is exploring the hidden patterns of data with complete analysis and finding the correlation. As dermatology data is purely unsupervised, learning for such kind of data classification algorithm is suitable for grouping in clusters 5 different types of unsupervised machine learning algorithms have to consider, i.e. (Random Forest, Gaussian Naive Base, Multinomial Naive Base, Bernoulli Naive Base and Support Vector Clustering) for better classification all the algorithm are selected from classification and clustering the 6 different types of Erythemato-Squamous diseases. We have already discussed data sets in the previous section. As the data is unsupervised, it is converted into labeled data as supervised learning since time series data in supervised data provides better analysis data which can be helpful for prediction on the same type of data sets.

3.1.6. Five Classification Algorithms

- Random Forest
- Gaussian Naive Base
- Multinomial Naive Base
- Bernoulli Naive Base
- Support Vector Clustering

Each algorithm generates 3 reports according to the coding we have done and discussed below

Random Forest

Accuracy for classifiers, as a result, shows that the accuracy of Random Forest is “1.0” and generates confusion matrix in an array format with base int64 and calculate (precision, recall, F1score, supports/n/n) in the below result show.

Gaussian Naive Base

Gaussian Naive Base Algorithm is the most commonly used algorithm for signals and dimensionally reduction. The experiment accuracy rate is "1.0" and generates a confusion matrix in an array with base int64 (precision, recall, F1 score, supports/n/n); results can be seen in table 5**.**

Multinomial Naive Base

Multinomial Naive Base Algorithm work on multiple attributes and more than 3 to 6 clustering in heavy data sets. With the help of this algorithm, we have achieved "0.9729727" with a confusion matrix with array base int64 (precision, recall, F1score, supports/n/n).

Bernoulli NaiveBase

Bernoulli Naive Base Algorithm is good at classification problems with an accuracy rate of "0.9455459" with a confusion matrix with array base int64 (precision, recall, F1 score, and supports/n/n).

Support Vector Clustering

Support Vector Clustering algorithm is the best algorithm for our problem statement with accuracy "1.0" and confusion matrix with array base int64 (precision, recall, F1score, supports/n/n)

The total weight of 34 attributes have been calculated, and the weights of the attribute are measurements of correlation among them depending on the weight generated in dermatology data; the most confident attribute is selected to find the correlation concerning weather data of Chicago

```
('Weights of the selected features : \n',
array([0.00783455, 0.01368266, 0.00857602, 0.02855975,
0.04130304,
       0.03696483, 0.00974073, 0.0048016 , 0.00749536,
0.01520533,
       0.00290068, 0.0310845 , 0.01537896, 0.03650129,
0.07188722,
       0.01603034, 0.0034784 , 0.00338888, 0.00855527,
0.0746063 ,
       0.06261587, 0.07827947, 0.00321859, 0.00624083,
0.02000498,
       0.01165534, 0.04464991, 0.02301745, 0.04305388,
0.01631028,
       0.01991595, 0.00310465, 0.06029351, 0.16966357]))
```

Fig. 3 Weights of 34 Attributes

4. Results and Discussion

4.1. Temp Analysis

A complete analysis of the weather data of Chicago city of USA, is shown in tables 1 and 2, a normal calculation of time series data to find Mean, Median, Tmin, Tavg, and Tmax. We have considered 9 parameters of weather attribute out of 22 attributes .author has to consider the dew point for his experiment, and analysis with the dew point shows positive and high correlation factors [23].

Through analysis, relative humidity is considered one of the attributes that affect NTS diseases[24]. Climate change is not limited to one area, even if it affects pregnancy. Women at pregnancy time with the increase of temp might be very risky with an increase of preterm birth and stillbirth with a sample of 17,171 pregnant women considered for analysis outcome shows the very dangerous result, i.e., increased absorption.[25]

Various analyses on different allergic datasets of weather locations have been considered for the experiment. Throughout the analysis, authors found a positive finding that Air Pollution, sea level, and wind speed are relative attributes that cause allegoric diseases. Skin diseases are the root causes of skin cancer pigmentation and many risk factors. The findings of the study found that 3% increase in skin diseases around the world.[26]

In Australia [27], nearly 15% of deaths rate recorded in the last 2 decades due to vast changes in climate and adversely affecting public health.out of the Extensive study we have made through the search on reputed journals, authors have not worked or examined hot weather and its e/effects on Erythemato-Squamous diseases. We have gone through 350 plus journals, but our study outcomes project interest in the research gap. We have identified some supporting frameworks to approach such kind of problem statement. We discuss this in the previous section. Spectral analysis of image data for skin spot dictions with the help of Fourier analysis and filter techniques like K-law with nor linear has option 95.4% confidence level.[28]

Further deeper analysis should be carried out on this. However, before the start of this problem statement, we have calculated the correlation coefficient factor as we have considered 49 attributes, out of which 34 attributes are Erythemato-Squamous diseases, and [29][27][30] authors have expensively analyzed about 34 attributes and applying data mining techniques for extraction for our finding we have considered this attributes and other we have considered 9 attributes of weather data of 22 attributes after combing different tow data set total 49 attributes and its correlation coefficients and R values we can see in below result.

- Dark blue => strongly related
- Light blue => modernly related
- Red => not related

The values are attributes we have considered from 1 to 49 in terms of Y. The suffix represents the attribute name, and in brackets no, of times it has been related to other attributes' total counts.

Table 1. Total summary of weather data

Heat	Cool	Result Dir	Avg Speed	Total chicago[dates and Tmin, Tmax, Tavg]
Min. : 0.000	Min. : 0.000	Min. : 1.00	Min. : 2.100	Min. :2007-05-29
Median: 0.00	Median: 9.000	Median: 19.00	Median: 7.100	Median:2009-08-07
Mean: 0.9582	Mean: 8.489	Mean: 17.52	Mean: 7.435	Mean:2010-03-04
Max. :15.00	Max. :19.000	Max. :36.00	Max. :15.000	Max. :2013-09-26

Table 2. Total summary of weather data

Dew Point	Precip Total	Stn Pressure	Sea Level	Result Speed
Min :39.00	Min :0.0000	Min. :28.89	Min :29.60	Min: 0.100
Median: 60.00	Median:0.0000	Median: 29.26	Median: 29.98	Median: 5.500
Mean: 59.76	Mean:0.1806	Mean: 29.25	Mean: 29.97	Mean: 5.977
Max. :73.00	Max. :3.9700	Max. :29.59	Max. :30.33	Max. :14.600

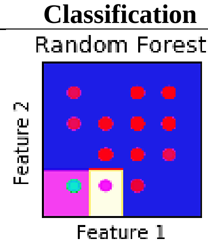
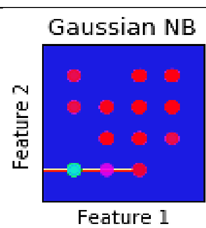
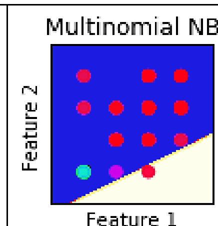
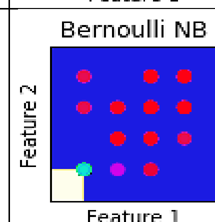
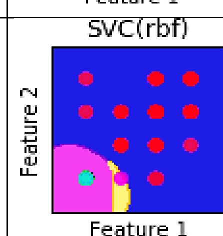
Table 3. A complete analysis among 5 different machine-learning algorithms

Classifier	Precision	Recall	F1-score support	Support/n/n
Random Forest	1.00	1.00	1.00	26\n,16\n,..74/n
Gaussian NB	0.90	0.87	0.93	26\n,16\n,..74/n
Multinomial NB	0.90	0.93	0.97	26\n,16\n,..74/n
Bernoulli NB	0.78	0.93	0.85	26\n,16\n,..74/n
SVC(rbf)	1.00	1.00	1.00	26\n,16\n,..74/n

Table 4 . Evaluation Metric Based on Confusion Matrix.

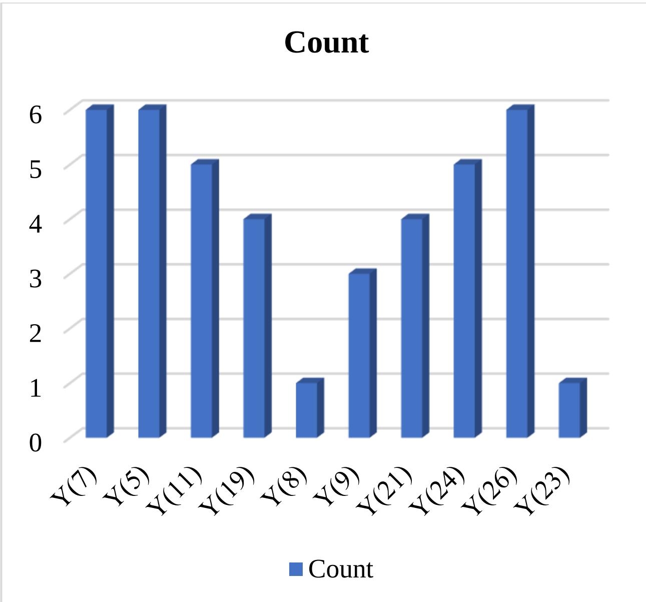
Parameters	Formula
Accuracy	$A = \frac{TP + TN}{Total}$
Recall	$R = \frac{TP}{Actual\ True}$
Precision	$P = \frac{TP}{Predicted\ True}$
F Score (FS)	$FS = 2 \times \frac{R \times P}{R + P}$

Table 5. Classification with 5 different algorithms

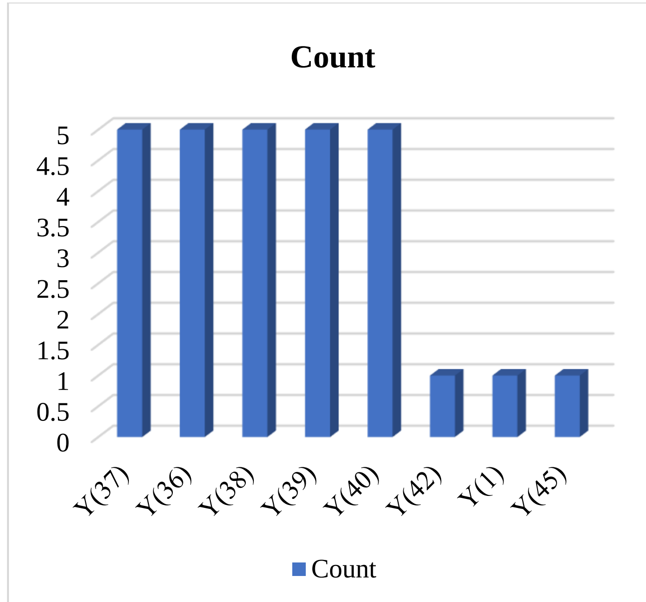
Classifier	Accuracy	Classification
Random Forest	1.0	
Gaussian NB	0.972972972972973	
Multinomial NB	0.9864864864864865	
Bernoulli NB	0.918918918918919	
SVC(rbf)	1.0	

4.2. Dermatology Analysis

Our research finds that 14 attributes of dermatology data and 7 attributes of weather data are mostly correlated with P value=0.8655142, and these 23attributes are highly correlated. Using the proposed framework, these attributes can be considered the most deciding factor for predicting six different types of skin diseases in the future.



Graph 1. Most frequent number of attributes repeated in EDS data
$Y_0, Y_1, \ldots .. Y_{34}$(Erythemato-Squamous diseases attributes)



Graph 2. Most frequent number of attributes repeated in weather data
$Y_{35}, Y_{36} \ldots \ldots Y_{49}$(weather data attributes)

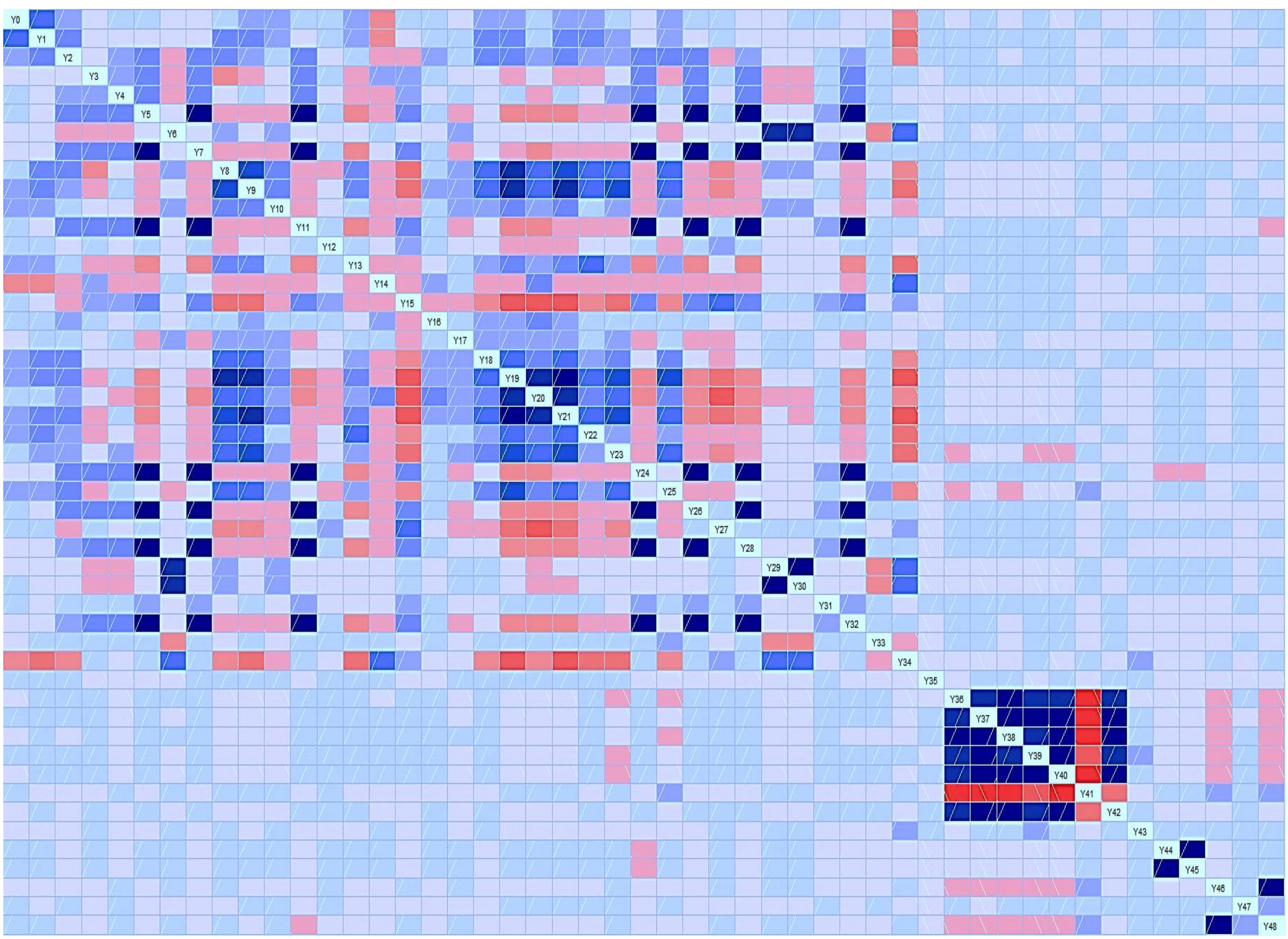


Fig. 4 A Correlation Matrix of dermatology and weather data sets

Table 6. All abbreviations on paper

Abbreviations	Explanation
ESD	Erythemato-squamous Diseases
W-M-D	Weather Method Diseases
SVM	Support vector machine
AVONA	Analysis of variance
ELM	Elaboration like hood model
GP	Genome profiling
ANFIS	The adaptive neuro-fuzzy interface system
MAE	Mean absolute error
RMSE	Root mean square error
DES	Data encripton standred
SDM	Sustainable development management
DLNN	Dynamic learning neural network

5. Conclusion

This paper analyzes the effects of hot weather on Erythemato-Squamous diseases (skin diseases) at the variant location.

We found a correlation among 49 attributes of different data sets and systematic analysis of data sets in a detailed manner and applied machine learning algorithm on dermatology data set. In our analysis, 14 attributes of dermatology and 7 attributes of weather data set are mostly correlated.

In the future, the research gap for predicting 6 types of Erythemato-Squamous diseases (skin diseases) due to climate change needs to be addressed.

References

[1] O. Babalola, A. Razzaque, and D. Bishai, "Temperature Extremes and Infant Mortality in Bangladesh : Hotter Months, Lower Mortality," *PLOS ONE*, vol. 14, no. 5, pp. 1–9, 2018. Crossref, https://doi.org/10.1371/journal.pone.0189252

[2] A. A. Jacobsen, A. Galvan, C. C. Lachapelle, C. B. Wohl, R. S. Kirsner, and J. Strasswimmer, "Defining the Need for Skin Cancer Prevention Education in Uninsured, Minority, and Immigrant Communities," *JAMA Dermatol*, vol. 152, no. 12, pp. 1342–1347, 2016. Crossref, https://doi.org/10.1001/jamadermatol.2016.3156

[3] Balmain BN, Jay O, Sabapathy S, Royston D, Stewart GM, Jayasinghe R, and Morris NR, "Altered Thermoregulatory Responses in Heart Failure Patients Exercising in the Heat," *Physiological Reports,* vol. 4, no. 21, 2016. Crossref, https://doi.org/10.14814/phy2.13022.

[4] Noah Scovronick, Francesco Sera, Fiorella Acquaotta, Diego Garzena, Simona Fratianni, Caradee Y. Wright, and Antonio Gasparrini, "The Association Between Ambient Temperature and Mortality in South Africa: A Time-Series Analysis," *Environmental Research*, vol. 161, pp. 229–235, 2018. Crossref, https://doi.org/10.1016/j.envres.2017.11.001

[5] Eric Lavigne, Antonio Gasparrini, Xiang Wang, Hong Chen, Abderrahmane Yagouti, Manon D Fleury and Sabit Cakmak, "Extreme Ambient Temperatures and Cardiorespiratory Emergency Room Visits: Assessing Risk by Comorbid Health Conditions in a Time Series Study," *Environmental Health: A Global Access Science Source*, vol. 13, no. 1, pp. 1–8, 2014. Crossref, https://doi.org/10.1186/1476-069X-13-5

[6] Li Bai, Gangqiang Ding, Shaohua Gu, Peng Bi, Buda Su, Dahe Qin, Guozhang Xu, and Qiyong Liu, "The Effects of Summer Temperature and Heat Waves on Heat-Related Illness in a Coastal City of China, 2011 – 2013," *Environmental Research*, vol. 132, pp. 212–219, 2014. Crossref, https://doi.org/10.1016/j.envres.2014.04.002

[7] C. Bertelsmeier, G. M. Luque, and F. Courchamp, "The Impact of Climate Change Changes Over Time," *Biological Conservation*, vol. 167, pp. 107–115, 2013. Crossref, https://doi.org/10.1016/j.biocon.2013.07.038

[8] Z. C. Félix Garza, J. Liebmann, M. Born, P. A. J. Hilbers, and N. A. W. van Riel, "A Dynamic Model for Prediction of Psoriasis Management by Blue Light Irradiation," *Frontiers in Physiology*, vol. 8, no. 1, pp. 1–17, 2017. Crossref, https://doi.org/10.3389/fphys.2017.00028

[9] K. Mohammadi, S. Shamshirband, S. Motamedi, D. Petković, R. Hashim, and M. Gocic, "Extreme Learning Machine Based Prediction of Daily Dew Point Temperature," *Computers and Electronics in Agriculture*, vol. 117, pp. 214–225, 2015. Crossref, https://doi.org/10.1016/j.compag.2015.08.008Get

[10] S. Kundu, D. Khare, and A. Mondal, "Future Changes In Rainfall, Temperature and Reference Evapotranspiration in the Central India by Least Square Support Vector Machine," *Geoscience Frontiers*, vol. 8, no. 3, pp. 583–596, 2017. Crossref, https://doi.org/10.1016/j.gsf.2016.06.002

[11] J. M. Lobo, "The Use of Occurrence Data to Predict the Effects of Climate Change on Insects," *Current Opinion in Insect Science*, vol. 17, pp. 62–68, 2016. Crossref, https://doi.org/10.1016/j.cois.2016.07.003

[12] S. Lou, D. H. W. Li, J. C. Lam, and W. W. H. Chan, "Prediction of Diffuse Solar Irradiance using Machine Learning and Multivariable Regression," *Applied Energy*, vol. 181, pp. 367–374, 2016. Crossref, https://doi.org/10.1016/j.apenergy.2016.08.093

[13] Terence Dwyer, James M Stankovich, Leigh Blizzard, Liesel M FitzGerald, Joanne L Dickinson, Anne Reilly, Jan Williamson, Rosie Ashbolt, Marianne Berwick, and Michèle M Sale, "Does the Addition of Information on Genotype Improve Prediction of the Risk of Melanoma and Nonmelanoma Skin Cancer Beyond that Obtained from Skin Phenotype?," *American Journal of Epidemiology*, vol. 159, no. 9, pp. 826–833, 2004. Crossref, https://doi.org/10.1093/aje/kwh120

[14] B. H. Kaffenberger, D. Shetlar, S. A. Norton, and M. Rosenbach, "The Effect of Climate Change on Skin Disease in North America," *Journal of the American Academy of Dermatology*, vol. 76, no. 1, pp. 140–147, 2017. Crossref, https://doi.org/10.1016/j.jaad.2016.08.014

[15] H. A. Güvenir and N. Emeksiz, "Expert System for the Differential Diagnosis of Erythemato-Squamous Diseases," *Expert Systems with Applications*, vol. 18, no. 1, pp. 43–49, 2000. Crossref, https://doi.org/10.1016/S0957-4174(99)00049-4

[16] A. M. Hashem, T. Abujamel, R. Alhabbab, M. Almazroui, and E. I. Azhar, "Dengue Infection in Patients with Febrile Illness and its Relationship to Climate Factors: A Case Study in the City of Jeddah, Saudi Arabia, for the Period 2010–2014," *Acta Tropica*, vol. 181, pp. 105–111, 2018. Crossref, https://doi.org/10.1016/j.actatropica.2018.02.014

[17] Qi Zhao, Yi Zhao, Shanshan Li, Yajuan Zhang, Qingan Wang, Huiling Zhang, Hui Qiao, Wuping Li, Rachel Huxley, Gail Williams, Yuhong Zhang, and Yuming Guo, "Impact of ambient Temperature on Clinical Visits for Cardio-Respiratory Diseases in Rural Villages in Northwest China," *Science of the Total Environment*, vol. 612, pp. 379–385, 2018. Crossref, https://doi.org/10.1016/j.scitotenv.2017.08.244

[18] Gopalakrishnan.S, Dr.Ebenezer Abishek.B, Dr. A. Vijayalakshmi, Dr. V. Rajendran, "Analysis and Diagnosis Using Deep -Learning Algorithm on Erythemato-Squamous Disease," *International Journal of Engineering Trends and Technology*, vol. 69, no. 3, pp. 52-57, 2021. Crossref, https://doi.org/10.14445/22315381/IJETT-V69I3P210

[19] M. Grasso, M. Manera, A. Chiabai, and A. Markandya, "The Health Effects of Climate Change: A Survey of Recent Quantitative Research," *International Journal of Environmental Research and Public Health*, vol. 9, no. 5, pp. 1523–1547, 2012. Crossref, https://doi.org/10.3390/ijerph9051523

[20] S. C. Harrison and W. F. Bergfeld, "Ultraviolet light and Skin Cancer in Athletes," *Sports Health*, vol. 1, no. 4, pp. 335–340, 2009. Crossref, https://doi.org/10.1177/1941738109338923

[21] E. D. Übeyli and I. Güler, "Automatic Detection of Erythemato-Squamous Diseases using Adaptive Neuro-Fuzzy Inference Systems," *Computers in Biology and Medicine*, vol. 35, no. 5, pp. 421–433, 2005. Crossref, https://doi.org/10.1016/j.compbiomed.2004.03.003

[22] Sang Gyu Lee, Sung Kyeom Kim, Hee Ju Lee, Hee Su Lee, and Jin Hyoung Lee, "Impact of Moderate and Extreme Climate Change Scenarios on Growth, Morphological Features, Photosynthesis, and Fruit Production of Hot Pepper," *Ecology and Evolution*, vol. 8, no. 1, pp. 197–206, 2018. Crossref, https://doi.org/10.1002/ece3.3647

[23] Brecht Devleesschauwer, Massimiliano Marvasi , Mihai C Giurcanu, George J Hochmuth, Niko Speybroeck, Arie H Havelaar, and Max Teplitski, "High Relative Humidity Pre-Harvest Reduces Post-Harvest Proliferation of Salmonella in Tomatoes," *Food Microbiology*, vol. 66, pp. 55–63, 2017. Crossref, https://doi.org/10.1016/j.fm.2017.04.003

[24] L. Cecchi, G. D'Amato, and I. Annesi-Maesano, "External Exposome and Allergic Respiratory and Skin Diseases," *The Journal of Allergy and Clinical Immunology*, vol. 141, no. 3, pp. 846–857, 2018. Crossref, https://doi.org/10.1016/j.jaci.2018.01.016

[25] S. He, T. Kosatsky, A. Smargiassi, and M. Bilodeau-bertrand, "Heat and Pregnancy-Related Emergencies : Risk of Placental Abruption During Hot Weather," *Environment International*, vol. 111, pp. 295-300, 2018. Crossref, https://doi.org/10.1016/j.envint.2017.11.004

[26] R. Gordon, "Skin Cancer: An Overview of Epidemiology and Risk Factors," *Seminars in Oncology Nursing*, vol. 29, no. 3, pp. 160–169, 2013. Crossref, https://doi.org/10.1016/j.soncn.2013.06.002

[27] A. Sitek, I. Rosset, E. Zadzińska, A. Kasielska-Trojan, A. Neskoromna-Jędrzejczak, and B. Antoszewski, "Skin Color Parameters and Fitzpatrick Phototypes in Estimating the Risk of Skin Cancer: A Case-Control Study in the Polish Population," *Journal of the American Academy of Dermatology*, vol. 74, no. 4, pp. 716–723, 2016. Crossref, https://doi.org/10.1016/j.jaad.2015.10.022

[28] E. Guerra-rosas and J. Álvarez-borrego, "Methodology for Diagnosing of Skin Cancer on Images of Dermatologic Spots by Spectral Analysis," *Biomedical Optics Express*, vol. 6, no. 10, pp. 3876–3891, 2015. Crossref, https://doi.org/10.1364/BOE.6.003876

[29] E. Tuba, I. Ribic, R. Capor-Hrosik, and M. Tuba, "Support Vector Machine Optimized by Elephant Herding Algorithm for Erythemato-Squamous Diseases Detection," *Procedia Computer Science*, vol. 122, pp. 916–923, 2017. Crossref, https://doi.org/10.1016/j.procs.2017.11.455

[30] J. K.Williams, D.A.Ahijevych, C.J.Kessinger, T.R.Saxen, M.Steiner, and S.Dettling, "A Machine Learning Approach to Finding Weather Regimes and Skillful Predictor Combinations for Short-Term Storm Forecasting," *National Center for Atmospheric Research*, pp. 1–6, 2008.

[31] J. Rocklöv, K. Ebi, and B. Forsberg, "Mortality Related to Temperature and Persistent Extreme Temperatures: A Study of Cause-Specific and Age-Stratified Mortality," *Occupational and Environmental Medicine*, vol. 68, no. 7, pp. 531–536, 2011. Crossref, https://doi.org/10.1136/oem.2010.058818