Arab American University

Faculty of Graduate Studies

Machine Learning Model for Prediction of Human Skin Locations and Conditions through Emissivity Measurements

Prepared By

Mamoun Mohammad Jameel Hakawati

Supervisor By

Dr. Majdi Owda

Co-Supervisor

Dr. Amani Owda

This Thesis Was Submitted in Partial Fulfillment of the Requirements for the Master's Degree in Computer Science

January / 2022

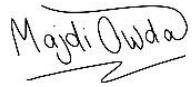
Thesis Approval

Machine Learning Model for Prediction of Human Skin Locations and Conditions through Emissivity Measurements.

By

Mamoun Mohammad Jameel Hakawati

This thesis was defended successfully on 20-Feb-2022 and approved by:

Committee members	Signature
1. Dr. Majdi Owda (Supervisor)	
2. Dr. Amani Owda (Co-Supervisor)	
3. Prof. Mohammed Awad (Internal examiner)	
4. Dr. Mohammad Jubran (External examiner)	

Declaration

I declare that the thesis titled "Machine learning model for prediction of human skin locations and conditions through emissivity measurements" is my work, and has been composed solely by myself and does not contain work from other researchers, and has not been submitted for any other degree or scientific work except the reference is made.

Name: Mamoun Hakawati

Date: 20/5/2022

Signature: 

Dedication

I dedicate this thesis to my family and friends for their unconditional love and support. To my mother and my father, for their support, which has not left me throughout my life. Also for my brothers and sisters, whose support I have not always forgotten. To my dear friends and work colleagues, for their continued support throughout my learning and work journey.

Acknowledgment

I would like to use this space to express my deep gratitude to Dr. Majdi Owda and Dr. Amani Owda for their advice, help, and valuable time; they spent reviewing and correcting my work. Dr. Majdi & Dr. Amani provided useful suggestions and advice that have had an important effect and helped in overcoming many obstacles in preparing this work in the best way possible.

Abstract

Machine Learning Model for Prediction of Human Skin Locations and Conditions through Emissivity Measurements

Prepared By: Mamoun Mohammad Jameel Hakawati

Supervisor By: Dr. Majdi Owda

Co-Supervisor: Dr. Amani Owda

In this thesis, machine-learning techniques were used to predict the human skin emissivity and to classify the human skin status based on its emissivity for wet and dry skin conditions. Predicting skin emissivity was made by selecting a specific skin location and using other locations from the same person to predict the selected location emissivity measure. Predicting measurement location emissivity will help to determine the normal emissivity value for the needed location; the implication of having this is the non-invasive diagnosis of diseased skin as the location of the skin may be infected with one of the skin diseases or thermal burns or may be covered with a hidden object, which will affect the emissivity. Regression models were used to predict the emissivity of the skin from nine locations using Linear Regression, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Multiple-Layer Perceptron Neural Network (MLPNN). Experimental results show that MLPNN gives the highest accuracy for most measurement locations, and the palm of hand achieved the highest accuracy of 91.3%. and the lowest mean absolute error of 0.0184 for the elbow location using MLPNN. For the second part of the development which is the classification of the skin conditions; many algorithms were used to predict the skin status based on the emissivity of the dry and wet skin conditions, and those are KNN, random forest, decision tree, and MLPNN.

The dataset of regression was collected from previously published studies of emissivity this dataset contains 540 records for 60 volunteers with 9 different measure locations from both genders. This dataset contains 5 features, location, ethnicity, age, mean emissivity, and gender The dataset for the classification was collected using the calibrated radiometer for a sample of 120 participants from 9 different locations in dry (normal skin) and wet skin (skin after the application of water) skin conditions from the previous studies also, This data set feature contains location, ethnicity, age, skin status, emissivity, and gender.

Results show that MLPNNs achieve accuracy of about 91.6%, recall 93.3%, precision 87.5%, and F-measure 90.3%, other detailed results will be discussed in chapter five.

Table of Contents

Contents	Page
Thesis approval	i
Declaration	ii
Dedication	iii
Acknowledgment	iv
Abstract	v
Table of Contents	vii
List of Tables	ix
List of Figures	x
List of Abbreviations	xii
Chapter One: Introduction	
1.1 Human Skin Introduction	1
1.2 Objectives	3
1.3 Contribution	4
1.4 Overview	4
Chapter Two: Literature Review	
2. Literature Review	6
2.1 Background	6
2.2 Emissivity Measures in Health Care	6
2.3 Machine Learning in Healthcare	9
2.4 Machine Learning with Skin Diseases	12
2.5 Machine Learning with Spectral Reflectance in Healthcare	14
2.6 Conclusions	18
Chapter Three: Exploratory Data Analysis	
3.1 Introduction	20
3.2 Regression Dataset Visualization	20
3.3 Classification Dataset Visualization	26
3.4 Conclusion	30
Chapter 4: The Proposed Method	
4.1 Introduction	31
4.2 Datasets	31
4.3 Data Preprocessing	33
4.3.1 Data Normalization	33
4.3.2 Data Standardization	34
4.3.3 Feature Selection	34
4.3.4 Outlier Removing	35
4.4 Building Models Phase	35
4.4.1 Classification	36

4.4.1.1 KNN	37
4.4.1.2 Decision Tree	39
4.4.1.3 Random Forest Classifier	40
4.4.1.4 Multilayer Perceptron	42
4.4.2 Regression	44
4.4.2.1 Linear Regression	46
4.4.2.2 Support Vector Machines (SVM)	46
4.5 Metrics Selection	47
4.5.1 Metrics For Classification	47
4.5.2 Metrics For Regression	49
4.6 Conclusion	49
Chapter Five: Results	
5.1 Introduction	50
5.2 Classification Results	50
5.2.1 Confusion Matrix	50
5.2.2 Accuracy	51
5.2.2.1 KNN Accuracy	51
5.2.2.2 Random Forest Accuracy	52
5.2.2.3 Accuracy Using Decision Tree	53
5.2.2.4 Accuracy Using Multilayer Perceptron Neural Network	54
5.2.3 Recall, Precision, and F-measure	56
5.3 Regression Results	58
5.3.1 Accuracy Records	58
5.3.2 Mean Absolute Error (MAE)	69
5.4 Conclusion	70
Chapter Six: Conclusion and Future Work	
Conclusion and Future Work	72
References	74
Appendix	80
الملخص	81

List of Tables

Number	Table	Page
4.1	First 5 dataset samples before data preprocessing	33
4.2	Confusion matrix description for skin status classification	48
5.1	Confusion matrix using different classification algorithms	50
5.2	Confusion matrix description for skin status classification	51
5.3	Accuracy results using MLPNN with different parameters	54
5.4	recall, precision, and F-measure for KNN, random forest, decision tree, and MLPNN algorithms	56
5.5	The measurements location numbering	59
5.6	BRF, ARF, ARO accuracy results using a linear regression algorithm	59
5.7	BRF, ARF, ARO accuracy results using KNN	60
5.8	BRF, ARF, ARO accuracy results using SVM	61
5.9	BRF, ARF, ARO accuracy results using MLPNN algorithm	62
5.10	MAE for different locations using different machine learning algorithms	69

List of Figures

Number	Figure	Page
1.1	Radiometer for emissivity measurement of the skin	3
3.1	Heat map for measurements locations	21
3.2	The 9 locations emissivity for random 13 volunteers	22
3.3	Box blot emissivity based on gender for palm, back of the hand, and fingers for regression dataset	23
3.4	Box blot emissivity based on gender for inner wrist middle, inner wrist side, and outer wrist for regression dataset	24
3.5	Box blot emissivity based on gender for volar side, dorsal surface, and elbow for regression dataset	24
3.6	Box blot emissivity based on ethnicity for palm, back of hand, and fingers	25
3.7	Box blot emissivity based on ethnicity for inner wrist middle, inner wrist side, and outer wrist	25
3.8	Box blot emissivity based on ethnicity for volar side, dorsal surface, and elbow	25
3.9	Correlation matrix between the measurements locations for classification dataset	26
3.10	Box blot emissivity based on gender for palm, back of the hand, and fingers for classification dataset	27
3.11	Box blot emissivity based on gender for inner wrist middle, inner wrist side, and outer wrist for classification dataset	27
3.12	Box blot emissivity based on gender for volar side, dorsal surface, and elbow for classification dataset	28
3.13	Box blot emissivity based on skin status for palm, back of the hand, and fingers for classification dataset	29
3.14	Box blot emissivity based on skin status for inner wrist middle, inner wrist side, and outer wrist for classification dataset	29
3.15	Box blot emissivity based on skin status for volar side, dorsal surface, and elbow for classification dataset	29
4.1	The measurement locations of the human skin emissivity	32
4.2	Basic workflow for classifying skin status	37
4.3	KNN classification method	38
4.4	Decision tree structure	40
4.5	Random forest structure	41
4.6	The Structure of MLPNNs	42
4.7	Basic workflow for skin location emissivity prediction models	45
5.1	Accuracy measures using different K numbers and different distance metrics algorithms	52

5.2	Accuracy measures using different N decision trees in the random forest	53
5.3	Accuracy measures using different min samples split for decision tree	53
5.4	Accuracy measures using 4 hidden layers and different numbers of epochs	55
5.5	Recall score using a different machine learning algorithm	57
5.6	Precision score using a different machine learning algorithm	57
5.7	F-measure score using a different machine-learning algorithm	58
5.8	BRF, ARF, ARO accuracy results using a linear regression algorithm	60
5.9	BRF, ARF, ARO accuracy results using KNN algorithm	61
5.10	BRF, ARF, ARO accuracy results using SVM algorithm	62
5.11	BRF, ARF, ARO accuracy results using MPNN algorithm	63
5.12	comparison of the different machine learning algorithms using ARO data	64
5.13	Predicted vs actual emissivity for the palm of the hand	64
5.14	Predicted vs actual emissivity for the back of the hand	65
5.15	Predicted vs actual emissivity for fingers	65
5.16	Predicted vs actual emissivity for inner wrist middle	66
5.17	Predicted vs actual emissivity for inner wrist side	66
5.18	Predicted vs actual emissivity for outer wrist	67
5.19	Predicted vs actual emissivity for volar side	67
5.20	Predicted vs actual emissivity for dorsal surface	68
5.21	Predicted vs actual emissivity for the elbow	68
5.22	MAE for different algorithms used in predicting different locations' emissivity	70

List of Abbreviations

KNN	K-Nearest Neighbors
SVM	Support Vector Machines
MLPNN	Multiple-Layer Perceptron Neural Network
PIDD	Pima Indians Diabetes Database
RGB	Red – Green – Blue
HSV	Hue – Saturation – Value
ROC	Receiver Operating Curve
ABCD	Asymmetrical Shape, Border, Color, and Diameter
MC	Monte Carlo
DRS	Diffuse Reflectance Spectroscopy
MLT	Machine Learning Techniques
WT	Wavelet Transformation
SNR	Signal-to-Noise Ratio
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
TP	True Positive
FP	False-Positive
FN	False-Negative
TN	True Negative
FNN	Feedforward Neural Network
BRF	Before Removing highly correlated Features
ARF	After removing highly correlated Features
ARO	After removing highly correlated features and Outliers

Chapter One

Introduction

1.1 Human Skin Introduction

Skin is the largest organ of the human body and consists of about 16% of the body mass and covers the internal organs [1]. Human skin plays a variety of critical functions and roles for the human body, the first function is sensory, it is working as a sense organ for skin senses such as touch, pain, cold, heat, and pressure. The second function is thermoregulation, the skin works as an absorber, heat generator, transmitter, conductor, radiator, and vaporizer. The third function is host defense since the skin prevents the entrance of foreign objects such as dust, water, and microorganisms. Fourthly, preventing excess water and food loss using the adipose tissue. Finally getting rid of excess water, acid, urea, and salt [2]. The human skin structure differs from one skin location to another and at a particular area based on its functions; the skin formation differs to meet the required roles based on the skin location and roles. In addition, there are three different sorts of human skin types namely: glabrous skin, hairy skin, and mucocutaneous [3]. The glabrous skin is usually located in the palm of hand and foot soles; it is characterized by a lack of hair follicles and a thick epidermis layer unlike the hairy skin, which contains a thin epidermis and dense hair follicles, the last sort of skin is mucocutaneous, which representing the entrances to the interior of the human body [3].

For the three types of human skin, the skin consists of three different layers:

Epidermis: the outer waterproof layer of human skin and the first defense line against foreign objects, bacteria, and harmful radiation, the dead skin cells renewal process happens in this layer by sloughing off the dead cells and replacing them with new, the epidermis contains the skin pores, which get rid of dirt and oils [3].

Dermis: The dermis is the second layer of the human skin, which consists of connective tissues; this layer houses the human body's oil glands, blood vessels, and hair follicles [3].

Hypodermis: (subcutaneous tissue). The hypodermis is the deepest layer of the human skin, this layer consists of nerve cells, fats, blood vessels, and connective tissues, one of the most important roles of this layer is storing the fats, which regulate skin tissue and cushion the body internal organs against falls and bumps [3].

Skin diseases are one of the most common diseases, and skin status may be a reliable indicator of the human body status [4]. Authors in [5] indicated a significant gradient in self-rated health by skin color; darker skin indicates poorer health. Like other organs, the skin suffers from many different diseases and these occur due to a variety of reasons such as tanning, family history, and ultraviolet radiation [6]. A person with skin disease may suffer from several problems like body changes, physical impairment, and even death in some dangerous diseases like malignant melanoma [7]. The skin emissivity differs from the healthy and non-healthy skin since the water level changes as a result of vascularization around the burns and damage [8], using a sensor with a millimeter-wave band; unlike the dielectric permittivity of the skin that can be measured using open-

ended coaxial probes [8]. Figure 1.1 [9] illustrates the emissivity radiometer that is used to measure the human hand emissivity over the frequency band 80-100 GHz.

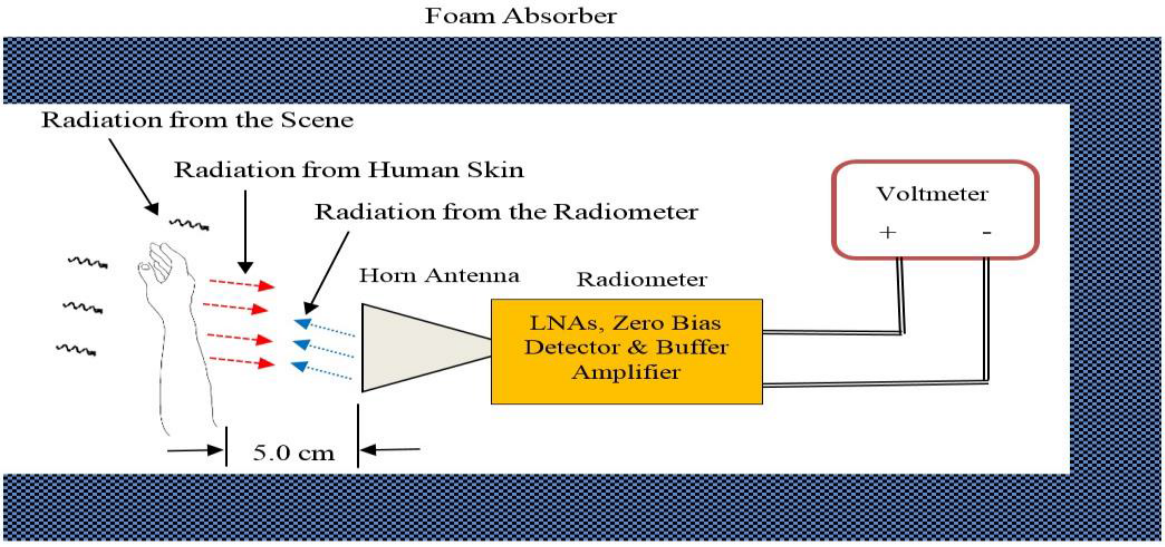


Figure 1.1: Radiometer for emissivity measurement of the skin *[9]*.

1.2 Objectives

In this research, one of the main objectives is to classify the human skin into dry or wet, which will be a reliable indication of skin health status using KNN, random forest, decision tree, and MLPNN machine learning classification algorithms. The other objective is to predict a skin location emissivity based on the measured emissivity records for other locations using linear regression, k-nearest neighbors, support vector machines, and multiple-layer perceptron neural network machine learning algorithms. The objectives of this work can be illustrated in those points:

- Using python plotting libraries to visualize the datasets to get deep insight into the data will help us to understand the dataset and the correlation between the elements.

- Preprocessing the dataset by removing outliers, highly correlated features, and dataset standardization and normalization.
- Training a classification and regression models to distinguish between the wet skin and the dry skin based on its emissivity measure and to predict the human skin emissivity measurement.
- Comparing the results while using the different machine learning algorithms and at different stages of data preprocessing.

1.3 Contribution

To the best of the author's knowledge, we did not find in open literature studies used skin emissivity data to predict or classify the skin status, some of them used photography images, others used machine learning for medical data classification, and others to link non-medical data with some diseases. In this research, novel published datasets will be manipulated and used to learn a system to be able to predict skin emissivity and classify the skin emissivity status using different prediction models.

Applying ML techniques in this field can be used as a decision support system for the medical applications also, as a system that will save the analysis time for the authority staff while scanning the human body to search for the hidden object.

1.4 Overview

The rest of this work is coordinated as follows. Chapter 2 literature review of the related work in machine learning for healthcare applications and some techniques used to predict and classify skin diseases and machine learning with spectral reflectance in

healthcare. In chapter 3, an exploratory data analysis was made to graphically show the dataset and the element correlation to gain deeper insights from the dataset. In chapter 4, the proposed method will be illustrated for both machine learning tasks classification and regression using 4 machine learning algorithms for each task. In chapter 5, the results for classification and regression using different evaluation metrics will be presented. Chapter 6 will represent the conclusion and the future work.

Chapter Two

Literature Review

2. Literature Review

This chapter presents a summary and analysis of the relevant publications on these thesis topics about human skin emissivity, machine learning in healthcare, machine learning with skin diseases, and machine learning with spectral reflectance in healthcare.

2.1 Background

Machine learning algorithms are broadly used in the healthcare field to predict and classify healthcare data [10, 11, 12, 13, 14]. There are many different applications for machine learning in health care such as medical image analysis, diseases prediction, robotic surgery, and robotic patient support tasks [15, 16, 17, 18, 19, 20, 21, 22, 23]. This will help health care staff and the patients in their tasks. Because of this, many tools and approaches will be analyzed and discussed in this chapter with other related topics.

2.2 Emissivity Measures in Health Care

Owda et al in [24] were working to find a unique signature of human skin by understanding the emissivity and reflection of human skin and with different people categories defined by age and gender.

Human skin signature will help in distinguishing between the skin and other different objects, which may be attached to the skin within a tenth (1/10) of seconds without exposing the human body to any type of radiation. A technique was described to measure the skin emissivity over frequency band 80-100 GHz and calibrated using liquid nitrogen and ambient temperature sources applied this technique for 60 healthy participants (36 males and 24 females), the test was applied in 4 different locations of a human hand.

There is a substantial difference in the emissivity of the skin across different locations. These differences depend on gender, skin thickness, and water content, the skin emissivity changes due to the hydration level of the skin, water content, and state of health.

Owda et al in [25] used millimeter-wave emissivity using a 95GHz calibrated radiometer to test the emissivity of the human skin, in addition, a comparison was done between the half-space model –which is used for this paper tests- and the complex three-layer model with three different water contents 50%, 75% and 95% over the frequency band of 30-100 GHz and the comparison shows that the result for the two models is identical.

In addition, samples with different skin conditions were got, such as with different water content, with and without aqueous gel, wet and dry human skin. 30 healthy skin samples and skin ranges measured between 0.2 and 0.7 were used to prove that the male skin is thicker than the female skin and there is a consistent difference between them, and the outer wrist and dorsal forearm are thicker than the inner wrist and the volar forearm and there is a statistically significant difference between them.

Owda et al in [26] used the millimeter-wave band to test the human skin reflectance for 50 healthy participants over 80-100 GHz band-30 males & 20 females. A radiometer contains three radiation sources, the first one is the hot load (ambient temperature source), the second one is the cold load (liquid nitrogen) and the third one is the area of skin to be measured was used.

A digital voltmeter was used to measure the output voltage for the skin and an infrared thermometer to measure the skin temperature. The test was applied at 6 different locations of the human hand skin (palm of hand, inner wrist, volar side, back of hand, outer wrist, and dorsal side).

Many tests for each participant were done and repeated, the first test is to check the difference between the normal clean skin and the wet skin, this test shows that there is a big impact of the water application on the emissivity. Other tests for male and female shows that the reflectance for the male skin is lower than the female skin. Other tests for the hand locations skin shows that the thinner skin gives a higher reflection than the thicker skin. Owda et al in [9] were working to prove the relation between human physical activity and the electromagnetics signature and they uses the radiometer over band 80-100 GHz to measure the skin emissivity for two different locations of the human hand (the palm and the back of the hand skin). The passive millimeter-wave sensor was used because its free of artifacts, environmental and user friendly, and less complex compared with the active model, this system was made using a single channel radiometer operating over a frequency band of 80-100 GHz consisting of an amplifier, radio frequency filters, a detector, further video amplifiers, an integrator, and a data-recording device. The tests show that there are substantial differences before and after

jogging, the emissivity of the skin as the mean emissivity value increases after jogging and then retains the original value after 15 minutes of exercising when the human body is relaxed, furthermore the known difference between the palm of hand and back of the hand. Also, the mean differences in the skin emissivity values before and after the application of an aqueous gel were shown. Adding gel to the human skin will decrease emissivity.

2.3 Machine Learning in Healthcare

A model in [27 ،28] was proposed to find an approach used to early detect different types of diabetes using data mining and machine learning techniques. The proposed models depend on many features, which could have a relation to diabetes such as body mass index value, members of the family suffering from diabetes, and having an inactive lifestyle. in [27] the Pima Indians Diabetes Database (PIDD) was used as a dataset and 76.3% accuracy was achieved. The dataset in [28] was collected from 735 patients confirmed to have diabetes or prediabetes and 752 normal people, 77.87% classification accuracy was achieved using the decision tree. Velasco et al in [29] employ digital image processing to analyze the human tongue status as an indicator of the body health for monitoring the healing process and the progress or weakening of the patient's health situation since the tongue and the tongue coating colors are not affected rabidly by short-term incidents or current variations. through visual inspection, classification of tongue coat, and body color. For image acquisition, two daylight LEDs matrixes were used in the scanning to scatter light equally to the tongue surface. Afterimage acquisition, image segmentation was done by converting the images to a 2-D array to find the edge size then the images were separated from overlapping objects

by labeling the image pixels with an integer value. Where any value except the zero in the input was used as features or objects and the zero values will present the background. The non-zero pixels represent the tongue body. After isolating the background of the tongue, isolating the tongue coating from the tongue body arises to identify the tongue body and coating. The tongue images were converted from RGB (Red, Green, and Blue) to HSV (Hue, Saturation, and Value) color space to allow the program to recognize the image like the human eyes way. For coding, the python programming language and sci-kit imaging library with scipy were used to cover four tongue colors: white, pink, red, and pale; and four tongue-coating colors: thick white, thin yellow, thin white and thick yellow. The main input for the proposed system is a tongue image and the output will be a list of assessed symptoms of the patient, the proposed method produced significant results, the accuracy was 93.3% for this work. in [30] a machine-learning model that can prognosticate the likelihood of diabetes in patients was proposed using three machine learning classification algorithms, SVM, decision tree, and naive bayes.

The dataset used in [30] was generated using the WEKA tool; WEKA is software designed by the University of Waikato and includes a collection of various machine-learning methods for data clustering, visualization, classification, and regression. The dataset is called Pima Indian Diabetes Dataset (PIDD); this dataset contains medical details for 768 female instances. There are eight attributes for each instance: pregnancy number, plasma glucose concentration, diastolic blood pressure (mm Hg), skinfold thickness (mm) 2-hour serum insulin (mu U/ml), BMI (weight in kg/(height in m)2), diabetes pedigree function and age in years. Accuracy, Precision, Recall, F-Measure, and ROC accuracy measures were used to evaluate this model, naïve bise gets the

highest accuracy with 76% then the decision tree with 74% and the lowest one is SVM with 64%

In [31] a classification model was proposed to distinguish between the heart diseases based on a variety of performance metrics, in the dataset, 13 dependent metrics were used as features to predict one independent variable. The original dataset consisted of predicted values from 0 (healthy) to 4 (unhealthy), the predicted data was converted from a 0 - 4 scale to 0(healthy) or 1(unhealthy). Grid search, probability calibration, and feature selection techniques were used to achieve the optimal results, also a comparison between k-nearest neighbor, naïve bayes, AdaBoost, random forest, support vector machine, and artificial neural network and conclude that support vector machine and artificial neural network gives the best results. The accuracy was tested in four different environments, the first one by using 10-fold cross-validation in the training set, the second by using a model without parameter tuning, 3rd by using tuned parameter values, and the last environment by using calibrated parameters. For the first environment, the highest accuracy was achieved using KNN with 84% then support vectors with 82%, naive bayes with 80% and the lowest one is random forest with 67%. For the second environment, KNN achieved the highest accuracy with 82%, naive bayes with 81%, support vectors with 79%, and random forest with 70%. For the 3rd environment, the highest accuracy was achieved using support vector with 80%, KNN with 78%, and random forest with 73%. For the st environment the highest accuracy was achieved using naive bayes with 80%, support vectors with 79%, KNN 78%, and random forest 75%.

2.4 Machine Learning with Skin Diseases

Nawal in [32] proposed an image-processing model to detect three different skin diseases in addition to normal skin, they used 100 samples for the three diseases and normal skin, and those diseases are Eczema, Melanoma, and Psoriasis. The model input is colored images taken by a digital camera with different diminutions and the model will resize it to 227×227 pixels and then use a support vector machine to classify the disease's type. The model achieved 100% accuracy for the four classes.

Bannihatti et al in [33] proposed a new approach for exploring and detecting six different types of skin diseases (Pityriasis Rosea, Chronic Dermatitis, Pityriasis Rubra Pilaris, Psoriasis, Seborrheic Dermatitis, and Lichen Planus). This approach combines machine learning and computer vision, computer vision is for feature extraction from digital images, and machine learning for diseases detection. This approach was applied and tested for six different types of skin diseases with 95% accuracy.

Dmitriy et al in [34] worked to early detect melanoma disease using machine learning in the examination of dermatoscopy images, they proposed a deep convolutional neural network to distinguish between malignant and benign skin cancer with 91% accuracy.

A mobile application was proposed in [35] to detect skin cancer using live mobile phone images on time using a convolutional neural network model. The CNN model was pre-trained using 10,015 skin cancer images from [36] ; these images contained seven different types of skin lesions: melanocytic nevi, vascular lesions, dermatofibroma, benign keratosis-like lesions, basal cell carcinoma, actinic keratosis, and melanoma. A macOS with Python, TensorFlow, Keras, and sci-kit-learn were used

to implement the core algorithms. And the (MaxPool2D) after convolutional layers to prevent the overfitting by removing the unnecessary features and reducing the number of parameters, after the convolutional and max-pooling layers, the system depends on the flatting, fully connected layers, and the soft-max for the output layer. For the testing and validation, the data was split as 64% for training, 16% for validation, and 20% for testing; the proposed model achieved 75.2% accuracy.

The authors in [37] depend on the ABCD rule (Asymmetrical Shape, Border, Color, and Diameter) for the skin dermoscopy to pre-classify the skin lesions into three categories normal, abnormal and melanoma using four machine learning methods namely: Artificial neural network, support vector machines, k-nearest neighbors and decision tree. The dataset for this study was created by a group of researchers who work at the technical universities of Porto and Lisbon; this dataset is called PH2 and contains 200 images with 768x560 resolution (80 normal, 80 abnormal, and 40 melanoma). The system successfully classified 92.50%, 89.50%, 82.00%, and 90.00% of the testing data for artificial neural networks, support vector machines, k-nearest neighbors, and decision trees respectively.

Jessica et al in [38] an approach was proposed to create a system for skin diseases prediction using android software. 3,406 images in JPEG extension was used from a combination of public accessible dermatology repositories and color photo atlas of dermatology, those images consist of seven different types of skin diseases (Vitiligo, Tinea Corporis, Psoriasis, Pityriasis rosea, Chickenpox, Eczema, and Acne), all of the images were validated by a dermatologist. The dataset was split into 80% for the training purpose and 20% for testing, the model-learning rate was 0.0001, the activation

function was softmax, the optimizer was Adam, and the loss function was cross-entropy and the number of the epoch was 30. The system achieved 91.8% accuracy using oversampling, 93.6% accuracy using the imbalanced dataset and the default preprocessing of input data, 94.4% using the oversampling and data augmentation technique.

2.5 Machine Learning with Spectral Reflectance in Healthcare

Shiwei et al [39] proposed a model to analyze and improve non-invasive skin pigments detection using machine learning, the proposed model depends on a 3-layered model that includes the epidermis, dermis, and subcutis, each layer was defined with thickness, absorption coefficient, scattering coefficient, refractive index, and anisotropic factor.

The epidermis thickness varies from .027mm to .15mm, the melanin in this layer is absorbed mostly by the light propagating.

The second layer is the dermis, which consists of blood vessels, and tissue, the thickness in this layer varies from .6mm to 3mm and it gets the pigment from two types of blood hemoglobin, oxy and deoxy-hemoglobin differ in the visible light range.

The last layer is the subcutis the thickness of the subcutis is up to 5mm and it is absorbed fat, water, blood, and skin baseline. The Monte Carlo (MC) simulation was used for reconstructing skin diffuse reflectance using python 3.6 and an array for indexing was created to reduce trigonometry operations and to directly store the results GPU techniques were used to accelerate the MC 1000 times faster. 1E07 energy photons were emitted vertically for each simulation. The detection window was built as a two cm radius circle to collect the total diffuse reflectance based on the three-layered

skin model. In this study, 50,000 samples were simulated by defining 12 skin parameters with wavelengths from 450 to 700 nm. The database was generated by Nvidia GeForce GTX1060 GPU, Intel Core i7-7700HQ CPU and it's taken almost 18 hours.

The Matlab R2018 deep learning toolbox was used to build two neural networks, each network composed of 55 neurons, two hidden layers, one layer input, and one layer for output.

The dataset generated by the MC was used to train the Feedforward Neural Network(FNN), and then the FNN starts generating the dataset because the MC method consumes time and resources without noticing the difference in results. The results were tested by using root mean squared error and the model gives an acceptable result even when they tested it with measured samples.

Rajitha et al [40] proposed a model to employ the Diffuse Reflectance Spectroscopy (DRS) in real-time robotic surgery systems using supervised machine learning to distinguish between four different surfaces in the human joint tissue, these surfaces are cartilage, subchondral, meniscus, and cancellous. The use of DRS in robotic surgery systems to classify the surface type is important because the surgeons depend on the tactile feedback to determine the surface type, unlike the robot which cannot determine the surface using tactile feedback.

The data was collected from total knee replacement surgeries using an optical spectrometer called USB- 650 tide spectrometer, which contains a linear silicon charge-coupled device to detect the wavelength of dispersed light. The used light source is a

150w halogen lamp with a light ring to standardize the light spread for the samples. the data set consisted of 3043 spectra samples, with help of the DRS, spectra from 1579 cartilage, 156 cancellous bone, 1269 subchondral bone, and 39 meniscus samples were collected

Each sample was measured using a variety of wavelengths -2048 wavelengths-, each one of this wavelength considered as an attribute helps in determining the tissue sample class the wavelength reduced from 2048 to 3 classes based on each class identifiability, for the software, WEKA tool was used.

The first step for the software implementation is to create ground truth for the samples by the clinical orthopedic surgeons depending on color, shape, presentation, and the location in which it was removed from the patient.

The second step is data normalization; the normalization was done by creating a form by division by the source light spectrum followed by the application of standard normal variety, this form was used to calculate the standard deviation and average spectra for each sample.

The third step is dimensionality reduction; this step aims to reduce the number of attributes or wavelengths for each spectral sample using multiclass fisher's linear discriminant analysis and because of this, each sample become has only three attributes. The last step was for a comparison of the used classifiers to determine the best correlated with the originally known labels by using linear discriminant analysis with cross-validation technique, this includes splitting the data for training and testing 9 sets used for training and 1 set for testing. The accuracy was changed based on spectral

range, resolution, and Signal-to-Noise Ratio (SNR), so the comparison step will select the best values of spectral range, resolution, and SNR to get the best accuracy. Amongst those investigated to achieve the maximum accuracy results, the wavelength resolution must be larger than 8nm, an SNR better than 10:1, and the 800-900nm wavelength range should be used. The accuracy was over 99% using the 10-fold cross-validation.

Karadağ et al in [41] proposed a model to early detect plant diseases depending on spectral reflectance and machine learning technique, leaves status data were collected from the same environment based on disease status and mycorrhizae fungus. The dataset was collected in a climate room with controlled conditions of 26 C temperature, 60% humidity, and a constant level of lighting, a spectra radiometer kit called field spec 3 was used in this study to record the spectra reflectance from the targeted paper, the spectroradiometer was calibrated using white spectral, it should give a stable reading for the white spectral.

The data set was collected from four different groups of papers, some of these leaves were infected with fusarium disease and some others inoculated with arbuscular mycorrhizal fungi (AMF), the uninfected with fusarium disease called later F-, the papers with infection with fusarium disease called later F+, the unvaccinated with (AMF) called A- and the inoculated with (AMF) called A+. each group of the 4 contains 4 pots and from each pot 5 leaves randomly were selected to make the spectral reflectance test, this means that from all groups 80 samples were collected using 2151 different wavelength from 350 to 2500 with 1 nm spectral resolution, for the all samples they collect 172080 data point. The large number of data points, which were generated

using the spectroradiometer, have a big amount of features; this makes the classification cumbersome and adversely affects the performance and results, so the data dimensions should be reduced without losing the data information, for this, Wavelet Transformation (WT) was used to express the frequency information and the time information together. Using the WT process, the number of features was reduced from 2150 to 75, and then the features were reduced again from 75 to 4 by calculating the statistical values of wavelet coefficients. After the WT process, the model will be trained using wavelet coefficients and calculated features dataset, then for classification, k-nearest neighbor algorithm, artificial neural networks, and naive bayes and for testing the classification performance was used, the cross-validation method was used by splitting the data into training and testing sets, 75% training and 25% for testing. All the project process was performed using an i5 processor and 8 GB ram. The results show that the highest accuracy results were achieved using the KNN classification algorithm and data of wavelet coefficients, also the other classification algorithms and the data from statistics of wavelet coefficients are acceptable.

2.6 Conclusions

In this chapter, we reviewed some previous studies related to our work; we started with reviewing some works related to emissivity measures in health care and illustrated how the emissivity was used to serve the healthcare section. Then the use of machine learning in health care and algorithms was used. After that, we start reviewing previous studies about the use of machine learning specifically in skin diseases, and the methods to employ machine learning to detect or predict skin diseases. Finally, we reviewed the

use of machine learning with spectral reflectance in healthcare and how this was implemented for planet diseases prediction and non-invasive skin pigments detection.

As mentioned in the objective section, we did not find in open literature any studies that used skin emissivity data to predict or classify skin status. Our work will focus on using the emissivity records for human skin to classify the skin status and predict the normal skin emissivity.

Chapter Three

Exploratory Data Analysis

3.1 Introduction

Data visualization is the presentation of data in a pictorial or graphical format, the dataset visualization is required for a better understanding of the dataset and to uncover insights from the beginning or identify areas and patterns to dig into more. Data exploration helps to make better decisions on where to dig deeper into the data. In this chapter, we will use python libraries to visualize the dataset to exploit meaningful sights.

3.2 Regression Dataset Visualization.

To explore the relationship between the dataset features the heat map, a matrix was used for the sixty volunteers, the heat map helps in knowing the correlation ratio between the features to know if there is enough correlation to make a regression model predict the skin emissivity measurements. The heat map was generated using the python matplotlib library, which uses the hereunder equation to calculate the correlation coefficient (r) between the elements.

$$r = \frac{\sum(x_{i-}x^{-})(y_i - y^{-})}{\sqrt{\sum(x_{i-}x^{-})^2 \sum(y_i - y^{-})^2}} \quad [42] \qquad 3.1$$

While

r = correlation coefficient

x_i = values of the x-variable in a sample

x^{-} = mean of the values of the x-variable

y_i = values of the y-variable in a sample

$\bar{y}$ = mean of the values of the y-variable

Figure 3.1 shows the correlation matrix for the nine locations.

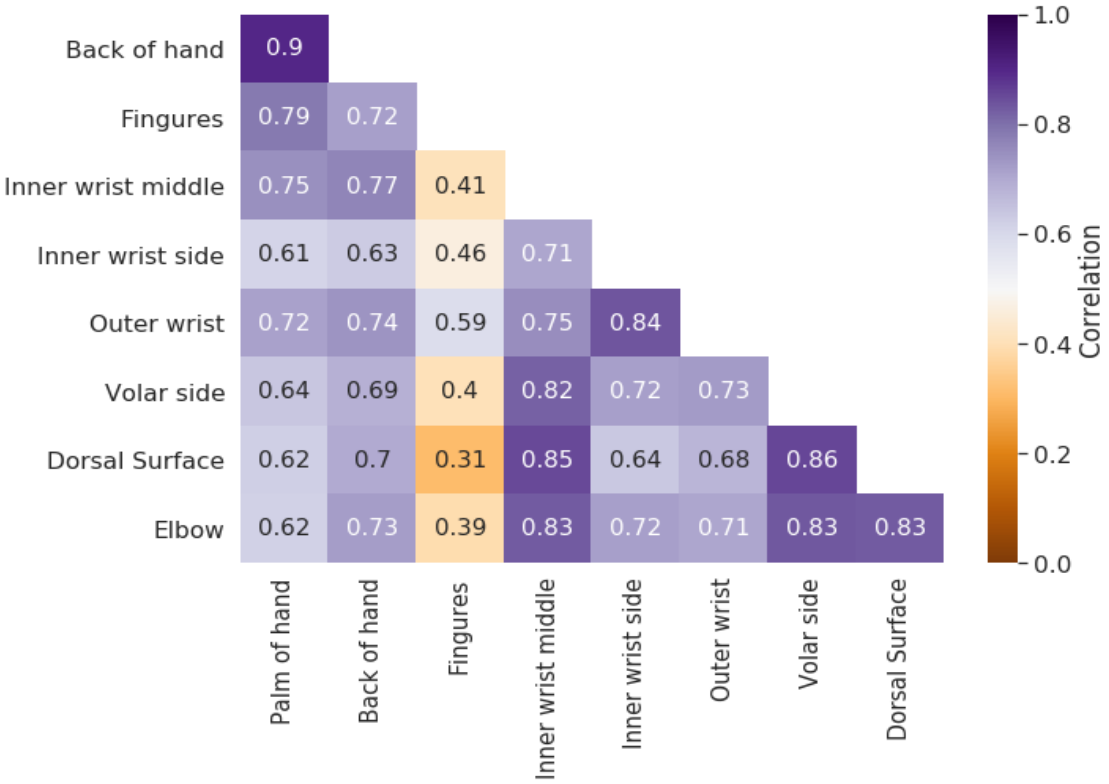


Figure 3.1: Heat map for measurements locations

Figure 3.1 shows that there is a sufficient correlation between some locations each other such as the palm and the back of the hand, this means that the location like the palm of the hand could be predicted using the back of the hand and some other highly correlated features. To go in more depth, we brought the measurements locations for random 13 volunteers and illustrate the line plot for the different locations to check the correlation between the 9 measurements locations for each other, for the same volunteer Figure 3.2 illustrates the emissivity records for the nine locations and the random thirteen volunteers.

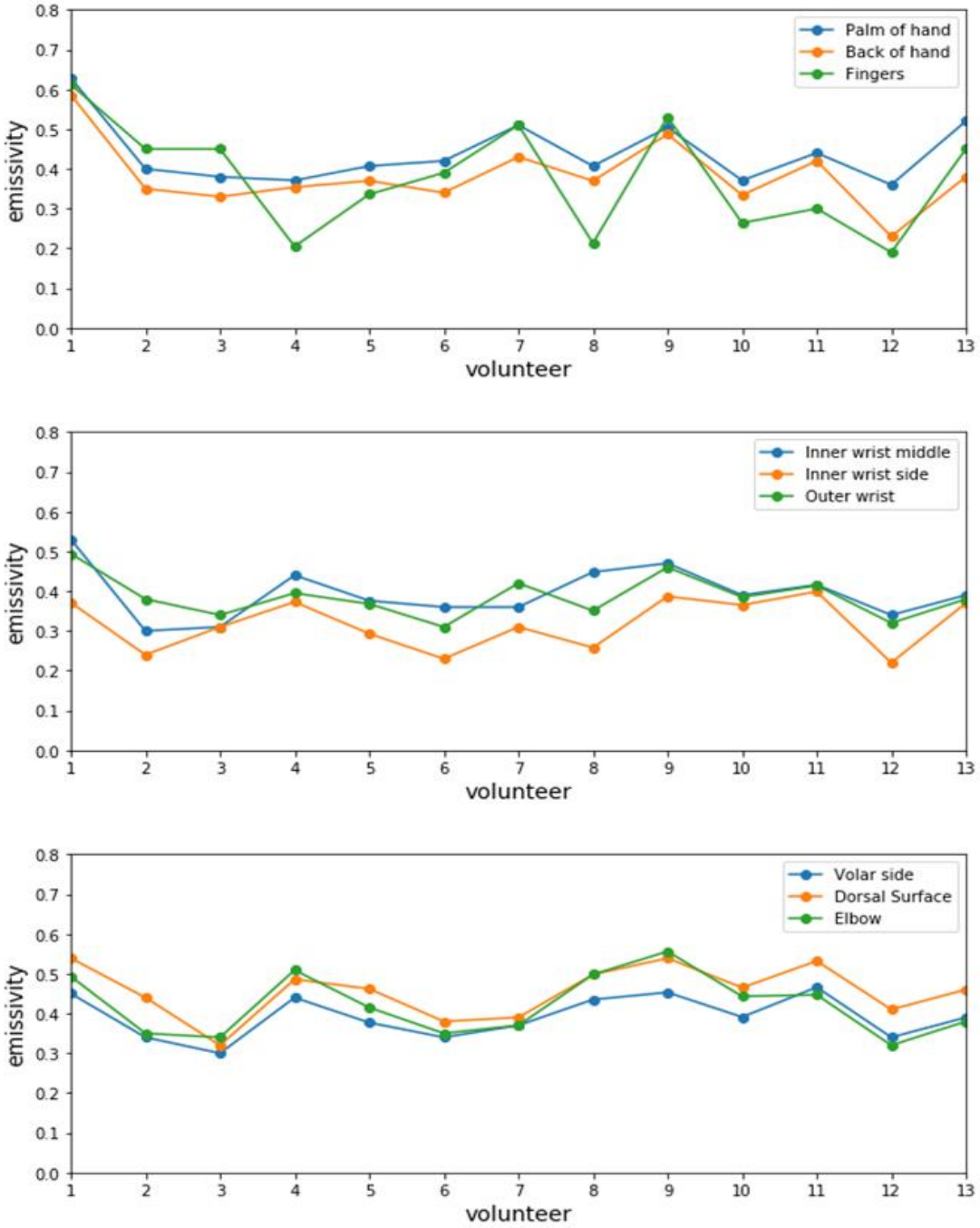


Figure 3.2: The nine locations emissivity for random 13 volunteer

Figure 3.2 shows the pattern of emissivity for each volunteer, for the volunteer number one, all emissivity records are high, this is an indicator that there is a good correlation or pattern for skin emissivity for the same volunteer – if a location is high, other locations will be high also-. Volunteer number twelve all location emissivity is low, this supports the emissivity pattern theory. Some locations such as fingers have an irregular change compared to other locations but still, there is a pattern with changing behavior.

To find the relation between the volunteer gender and the emissivity, we illustrated the box blot for all datasets at each location separately (24 females and 36 females), this will give an impression about the male and female if there is a noticed difference between both genders, Figure 3.3 – 3.5 shows the box plot based on gender for all 9 locations.

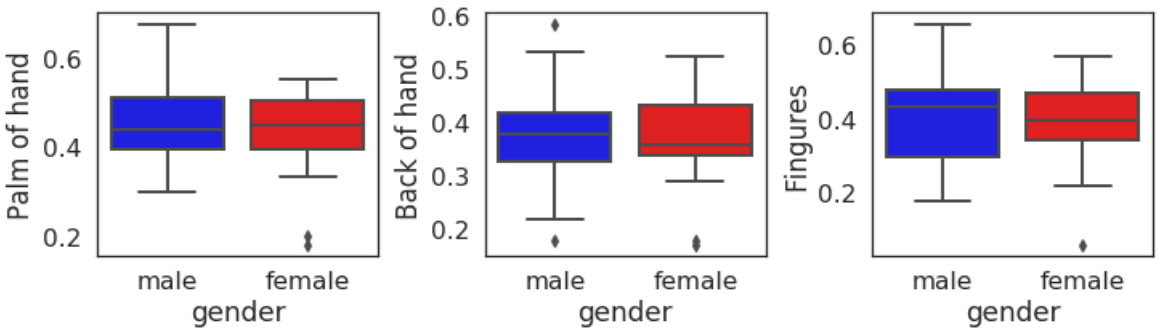


Figure 3.3: Box blot emissivity based on gender for palm, back of the hand, and fingers for regression dataset.

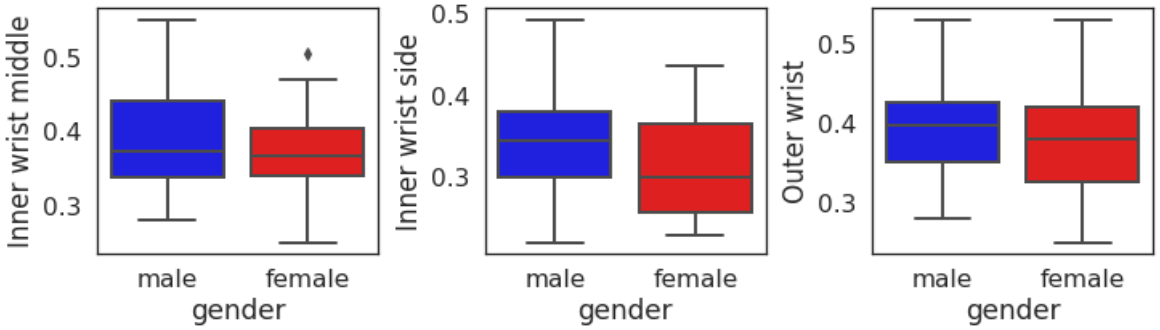


Figure 3.4: Box blot emissivity based on gender for inner wrist middle, inner wrist side, and outer wrist for regression dataset.

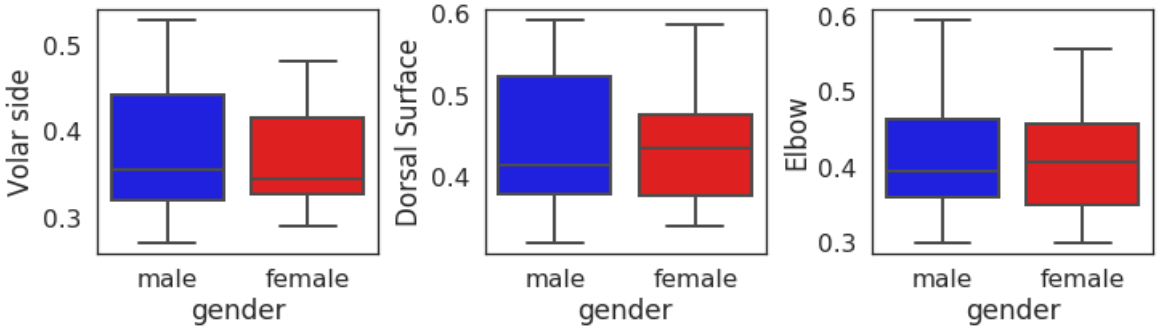


Figure 3.5: Box blot emissivity based on gender for volar side, dorsal surface, and elbow for regression dataset.

Figure 3.3 – 3.5 shows the box plot for the 9 locations based on gender, for most locations the female emissivity range is smaller than the male range, also by calculating the arithmetic mean for all emissivity records we find the arithmetic mean for the male samples are higher than female for all emissivity locations. for the palm of the hand the male emissivity is higher by 5.46% than female, back of the hand higher by 2.68%, fingers by 2.1%, inner wrist middle by 3.89%, inner wrist side by 8.7%, outer wrist 4.74%, volar side 4.17%, dorsal surface 2.34% and for elbow 2.56%.

To find if ethnicity affects the skin emissivity anyhow, we illustrated the box blot for the 9 locations based on ethnicity, Figure 3.6 – 3.8 shows the box blot emissivity measures for 9 locations for all datasets based on ethnicity.

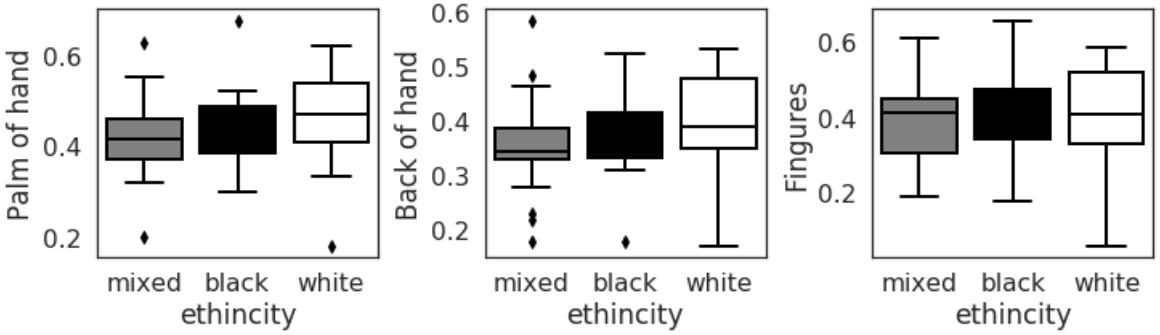


Figure 3.6: Box blot emissivity based on ethnicity for palm, back of hand, and fingers.

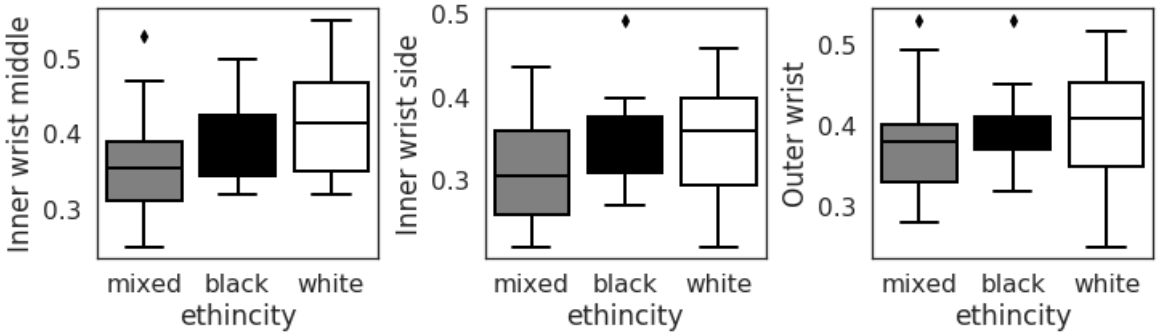


Figure 3.7: Box blot emissivity based on ethnicity for inner wrist middle, inner wrist side, and outer wrist.

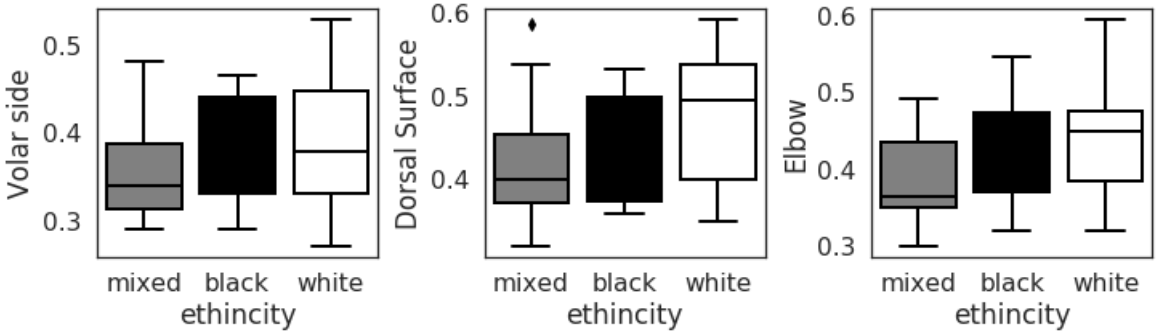


Figure 3.8: Box blot emissivity based on ethnicity for volar side, dorsal surface, and elbow

Figures 3.6 – 3.8 show a comparison for the emissivity based on the volunteer ethnicity: mixed, black, and white for each location separately using box blot for the 9 locations.

By calculating the arithmetic mean for the emissivity samples, the mixed ethnicity for the palm is lower by 5.7% than black and lower by 9.79% than white, for the back of the hand mixed lower than black by 5% and lower than white by 12.7%. for outer wrist lower than black by 3.1% and lower than white by 2.3%, for inner wrist middle lower than black by 8.6% and lower than white by 16.3%, for inner wrist side lower than black by 14.5% and lower than white by 12.7%. for outer wrist lower than black by 6.15% and lower than white by 6%, for volar side lower than black by 9.7% and lower than white by 11%, for dorsal surface lower than black by 6% and lower than white by 14.5%, and finally for elbow lower than black by 11.2% and lower than white by 14.7%.

3.3 Classification Dataset Visualization.

The classification dataset is labeled based on the skin status to dry and wet which will be the classification process target; firstly, we will illustrate the correlation between the locations to prove the correlation, Figure 3.9 shows the correlation matrix between the measurements locations for the classification dataset.

Figure 3.9: Correlation matrix between the measurements locations for the classification dataset.

As shown in Figure 3.9, some locations have a high correlation between each other such as the palm of the hand and back of the hand; some other locations did not have a high correlation with each other, such as fingers and dorsal surface. The correlation matrix for the classification dataset seems like the correlation matrix for the regression dataset.

In addition, we illustrated the box plot for the classification data set based on gender to prove that the male emissivity is higher than the female as in the regression dataset; Figure 3.10 – 3.12 shows the box blot based on gender for the classification dataset.

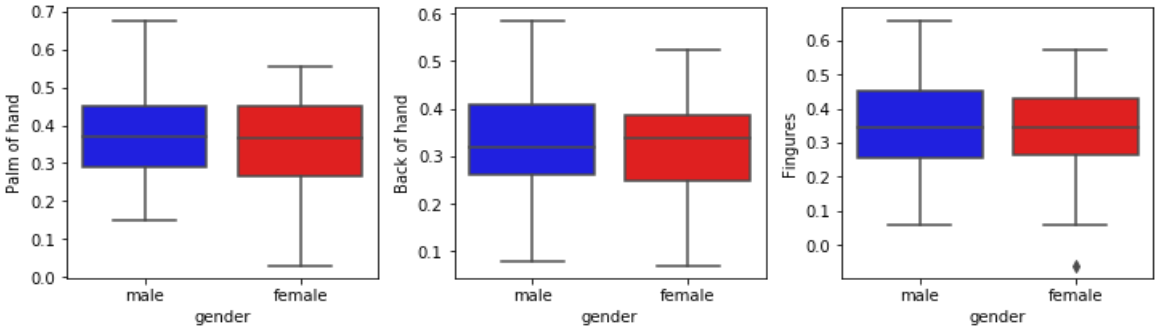


Figure 3.10: Box blot emissivity based on gender for palm, back of the hand, and fingers for classification dataset.

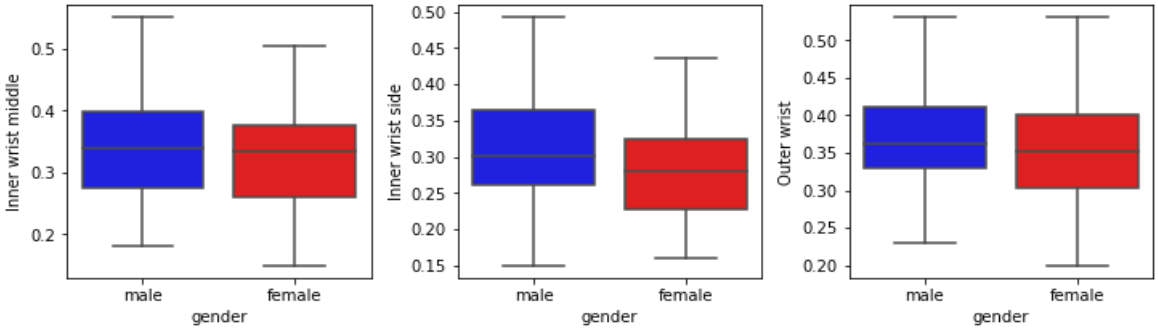


Figure 3.11: Box blot emissivity based on gender for inner wrist middle, inner wrist side, and outer wrist for classification dataset.

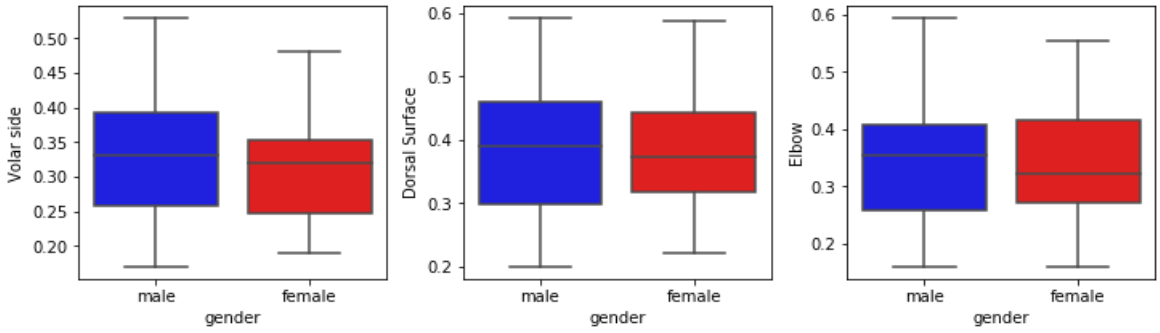


Figure 3.12: Box blot emissivity based on gender for volar side, dorsal surface, and elbow for classification dataset.

Figures 3.10 – 3.12 show the boxplot for the classification dataset based on gender, from the illustration, it's clear that the male emissivity is higher than the female in general, after calculating the arithmetic mean for the male and female separately, we find the arithmetic mean for the male's samples are higher than female for all emissivity locations. For the palm of the hand the male emissivity is higher by 2.4% than female, back of the hand higher by 1%, fingers by 0.9%, inner wrist middle by 1.5 %, inner wrist side by 3%, outer wrist by 1.9%, volar side 1.59%, dorsal surface 1.05 % and for elbow 1.6%.

To classify the dataset to wet and dry we should prove that there is a difference in emissivity between both conditions, to check that, we illustrated the box plot for the classification data based on the skin condition, Figure 3.13 – 3.15 shows the box plot for classification data based on status.

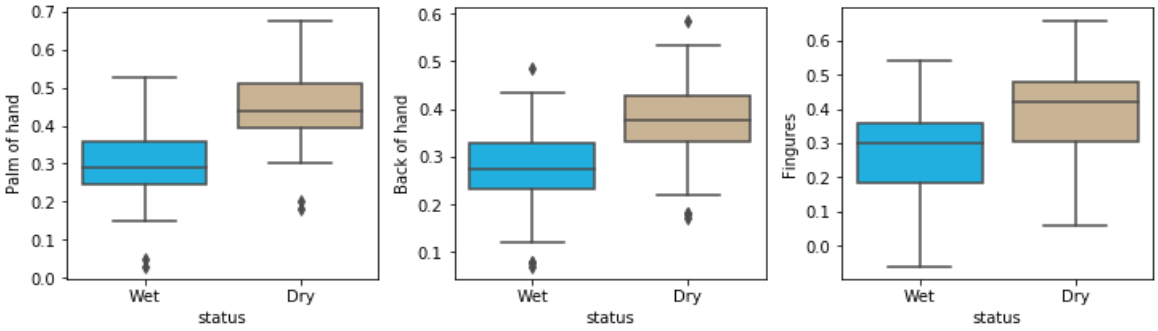


Figure 3.13: Box blot emissivity based on skin status for palm, back of the hand, and fingers for classification dataset.

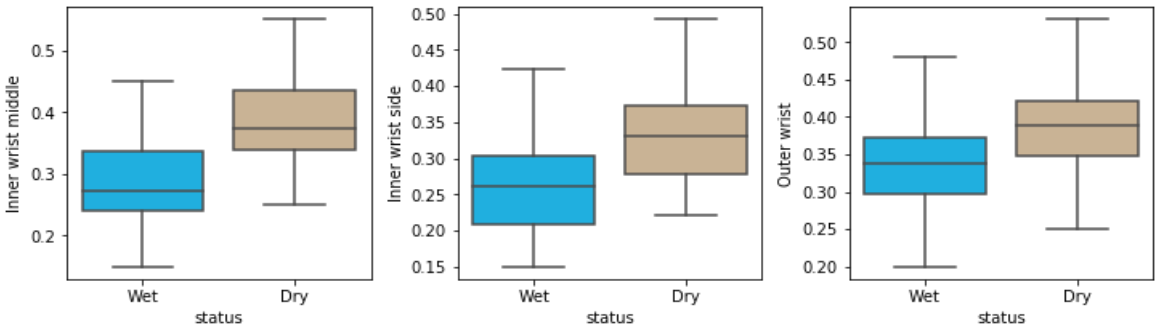


Figure 3.14: Box blot emissivity based on skin status for inner wrist middle, inner wrist side, and outer wrist for classification dataset.

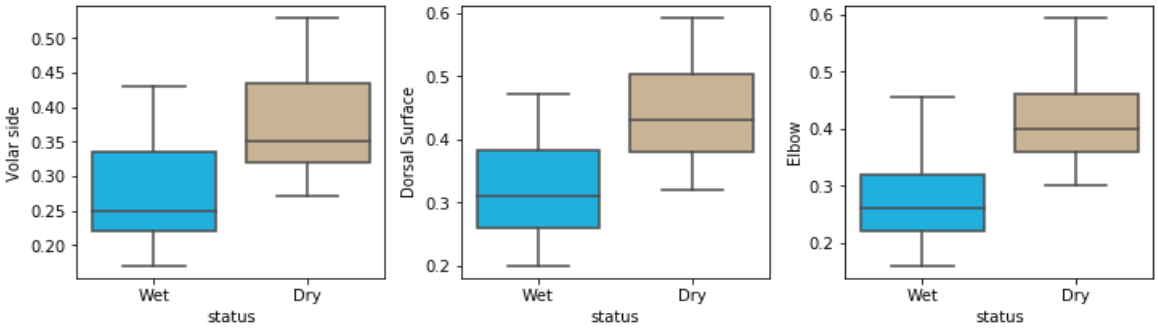


Figure 3.15: Box blot emissivity based on skin status for volar side, dorsal surface, and elbow for classification dataset.

Figures 3.13 – 3.15 show the box plot for the classification dataset based on the skin status, the illustration shows that the emissivity for the dry skin is higher than the wet skin for all measurement locations, to check the difference ratio we calculated the arithmetic mean for the wet skin and the dry skin separately.

The arithmetic means for the dry samples are higher than wet for all emissivity locations. for the palm of hand the dry emissivity is higher by 15% than wet, back of hand higher by 10%, fingers by 12%, inner wrist middle 10 %, inner wrist side 7%, outer wrist 5%, volar side 10%, dorsal surface 12 % and for elbow 14%.

3.4 Conclusion

In this chapter, we made an exploratory data analysis; this helps us to focus on the data patterns and make a decision about how to use machine learning to extract the knowledge from the data. After visualizing the data, we noticed a clear pattern for the skin emissivity for all locations to the same person in the regression dataset, which will give the ability to predict some measurement locations based on other locations. In addition, we noticed a clear difference between wet skin and dry skin, which will give the ability to predict the skin status based on its emissivity.

Chapter Four

The Proposed Method

4.1 Introduction

This chapter illustrates the proposed method, which aims to predict and classify human skin emissivity using machine-learning techniques on a dataset from previous studies of human skin emissivity. It begins by forming the dataset and describing the preprocessing steps. Then the chapter illustrates the deployed models: the classification models, which will classify the skin status for healthy (normal or dry skin) and unhealthy (wet skin) based on emissivity record, and the regression model to predict the normal skin emissivity record for a specific skin location based on the emissivity records from other locations for the same volunteer. Finally, it illustrates the metrics used to measure models' performance. The classification and regression were implemented using Python 3 with Pandas, Seaborn, Sklearn, Matplotlib, and NumPy libraries; we used a platform of intel core i7 processor 2.6 GHz with 12 GB RAM running Ubuntu 19 desktop for applying this approach.

4.2 Datasets

The dataset of regression was collected from previously published studies of emissivity [24, 25, 26, 9]. This dataset contains 540 records for 60 volunteers with 9 different measure locations from both genders. The dataset was classified for 9 different locations and those are the Palm, the Back of the hand, the Fingers, the Inner wrist middle, the Inner wrist side, the Outer wrist, the Volar side, the Dorsal Surface, and the Elbow. Those locations are illustrated in Figure 4.1. The dataset contains 5 features, location

which is the measuring location illustrated in Figure 4.1, ethnicity, which represents the volunteer ethnicity (black, white, or mixed), Age, which represents the volunteer age, mean emissivity which is the measured emissivity, and the volunteer gender. The dataset for the classification was collected using the calibrated radiometer for a sample of 120 participants from 9 different locations in dry (normal skin) and wet skin (skin after the application of water) skin conditions. The data set features contain location, ethnicity, age, skin status, emissivity, and gender.

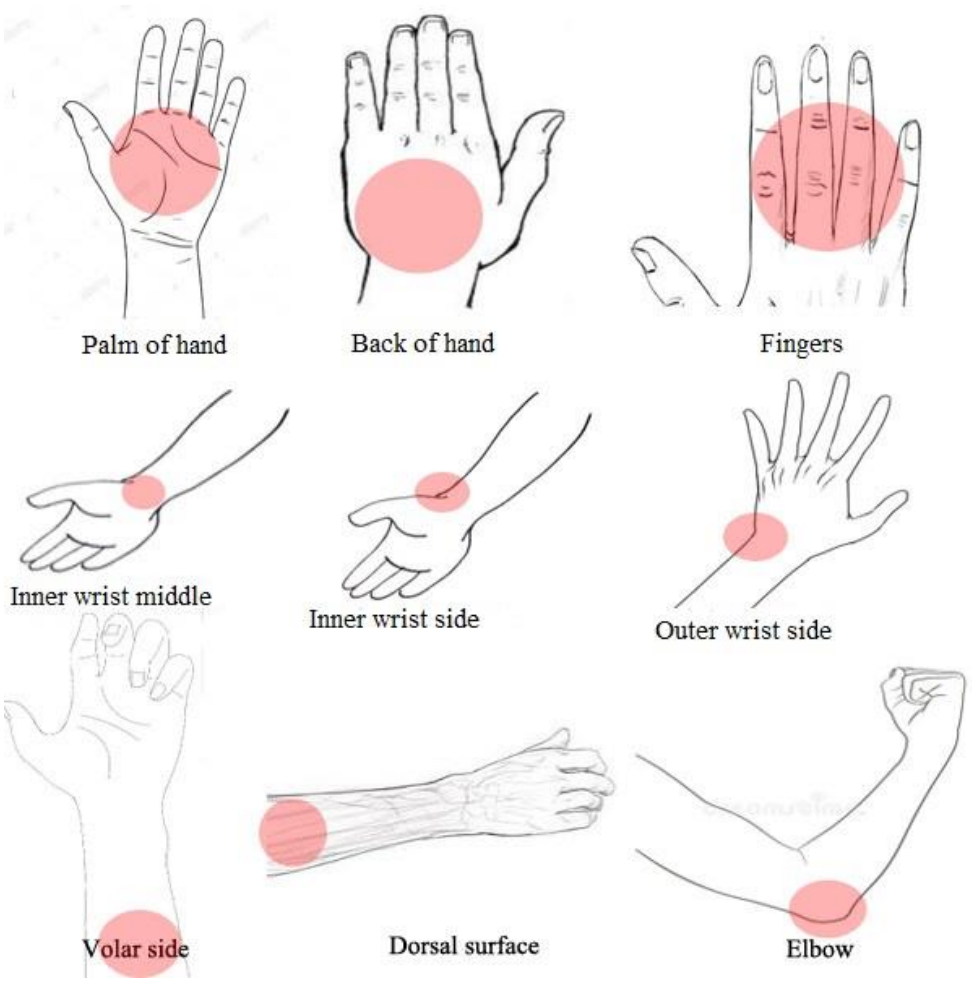


Figure 4.1: The measurement locations of the human skin emissivity

The dataset is classified and stored for each volunteer using a CSV file with 540 rows and 5 columns as shown in Table 4.1, which represents the first five records of the data.

Table 4.1: First 5 dataset samples before data preprocessing

Location	Ethnicity	Age	Mean Emissivity	Gender
Palm of hand	black	20	0.51	male
Palm of hand	mixed	20	0.44	male
Palm of hand	white	21	0.576	male
Palm of hand	mixed	22	0.36	male

4.3 Data Preprocessing

Data preprocessing is an important step to train the machine-learning model using the optimal shape of the data, without preprocessed data, the machine learning model may not work effectively as required and this results in a miss and bad results.

Different preprocessing sub-steps may be used depending on the nature of the dataset [43], data normalization, removing highly correlated features, removing the features with little correlation with the targeted feature, outliers removing were used in this thesis, also, a test to check the missing values was done in the original dataset. The following section will describe these steps in detail.

4.3.1 Data Normalization

Data normalization is a preprocessing technique used to prevent some features dominates all other features; data normalization aims to have features with the same scale to be of the same importance, there are many types of data normalization methods such as standardization and max-min normalization [43]. In this work, the range of all features normalized to be between [0-1], the max-min normalization method was used

to perform a linear transformation on the data, max-min normalization method was calculated using equation 4.1 [44] .

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad 4.1$$

While x' is the normalized value, x is the original feature value, x_{min} is the minimum value of the feature and x_{max} is the maximum value of the feature.

4.3.2 Data Standardization

The used dataset contains categorical features such as gender and ethnicity, and some used algorithms such as the SVM cannot work directly with categorical data and require all input variables and all output variables to be numeric. We converted the categorical data into numeric discrete data by using the integer-encoding technique [45].

4.3.3 Feature Selection

The highly correlated features will provide the same information, and this will multiply the value of the same information and let this information dominate the model, so we have to keep just one of the highly correlated features. In addition, the features with small relation to the target will not improve the model, so we have to remove the features with little correlation to the target to simplify the model and to make the algorithm learning faster.

To isolate the highly correlated feature and the features with little correlation with the target; we used the filter method feature selection which depends on the correlation coefficient to apply a threshold to remove the features with a correlation higher than

90% and a threshold to remove the features with a correlation less than 25% with the target.

4.3.4 Outlier Removing

The outliers are any data point with an abnormal distance from other data points, data outliers removing can improve the results accuracy [46], so we determined the outliers using Z-score [47] to label and identify the outliers, and we removed the data point which is bigger than 97% of the dataset; also, we removed the outlier's values, which are less than 97% of remaining dataset.

4.4 Building Models Phase

In this research, two types of prediction models were built, first one is classification to predict if the skin is healthy or not based on the emissivity measures and the second is the regression model that is used to predict the emissivity of normal skin measures for specific location based on the emissivity measures of other locations for the same person. We used KNN, random forest classifier, decision tree, and multiple-layer perceptron neural network algorithms for classification purposes and linear regression, KNN, SVM, and multiple-layer perceptron neural network for regression.

4.4.1 Classification

The classification is a machine-learning task to assign classes for a group of data, four different machine-learning algorithms were tested to build the classification model for binary classifying the status of the human skin for healthy (normal skin) or unhealthy skin (wet skin). The classification task starts with preparing the data set from the previous study of skin emissivity by forming, labeling, and specifying the features for the dataset and storing it with a CSV file to be readable by the python libraries. The second step is preprocessing the dataset to normalize and standardize the data, feature selection, and outlier removal; we tested the accuracy before and after removing the outlier and highly correlated feature. The preprocessed dataset was split for 70% training and 30% for testing, the training data set was used as data input for one of the used classification algorithms which are KNN, decision tree, random forest classifier, and multiple-layer perceptron neural network, those algorithms parameters will be discussed in details at the next sections. The testing data will be used to check the trained model's accuracy using many different metrics after finishing the learning phase; finally, we will find the best machine-learning algorithm to solve the classification issue for this type of data. Figure 4.2 shows the basic workflow for classification the skin status.

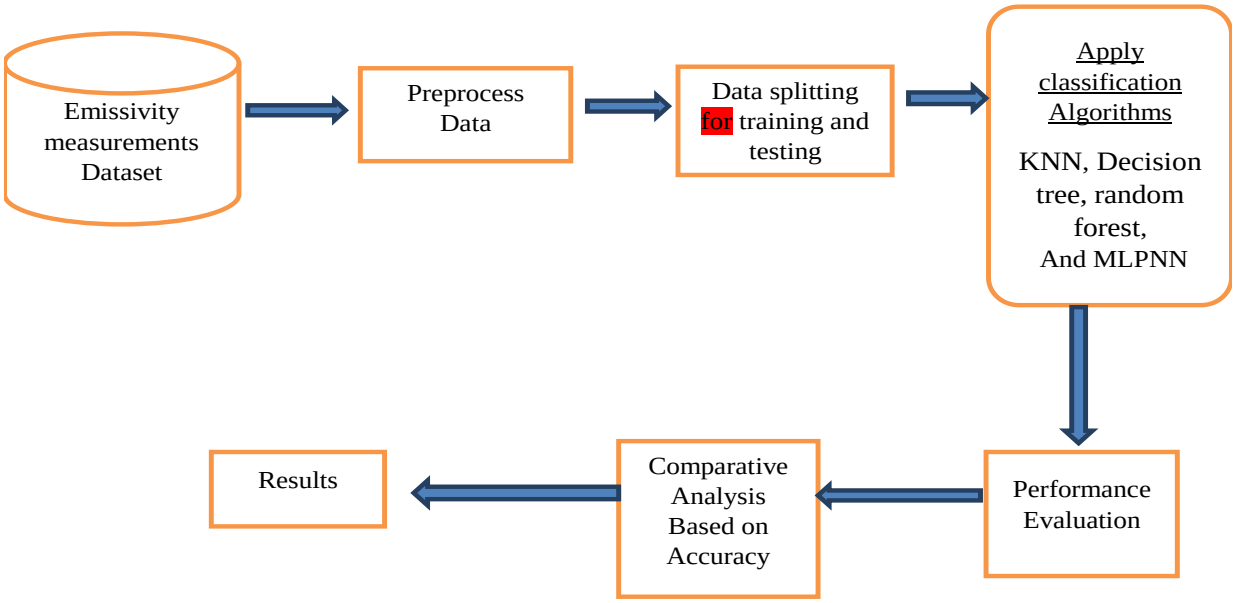


Figure 4.2: Basic workflow for classifying skin status

4.4.1.1 KNN

KNN is a supervised machine-learning algorithm developed in the 1950s by Joseph Hodges and Evelyn Fix, expanded later to handle classification and regression [48]. KNN assumes that similar data points exist in close proximity. The main concept of KNN depends on finding the distance between the data point and other k neighbors using one of the distance calculation equations. When a new data point comes for labeling and the number of k is 5, it checks the nearest 5 neighbors to find the nearest 3 neighbors with the same label, the dataset will be classified as those 3 neighbors. Figure 4.3 shows the process for the KNN algorithm.

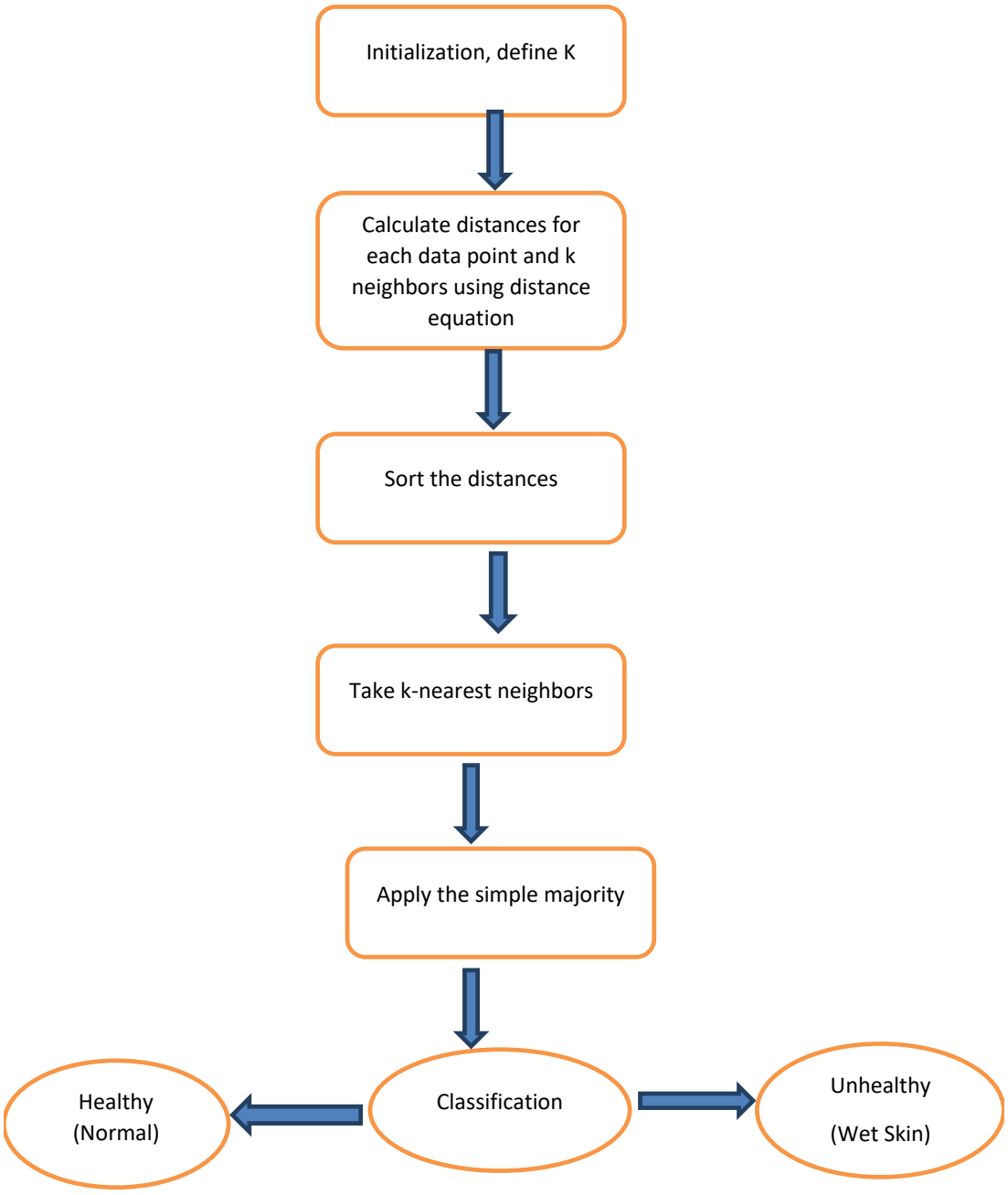


Figure 4.3: KNN classification method

We used the Manhattan distance to calculate the distance between each data point and the k neighbors; equation 4.2 shows the Manhattan distance [49].

$$D = \sum_{i=1}^{n} |x_{i-y_i}| \qquad 4.2$$

While D is Manhattan distance, n is the number of dimensions, x_i , y_i are the kth attributes respectively.

The output of the KNN classification algorithm is a class membership, healthy or unhealthy, with k =5, if the data point is classified as “healthy”, that means that 3 of the 5 neighbors at least are labeled as healthy (normal skin).

4.4.1.2 Decision Tree

A decision tree is a machine-learning technique represented as an upside-down structured tree where the features are represented with an internal node, decision rules are represented with branches, and the outcome is represented with leaves. The decision tree contains two different nodes, which are decision and leaf, the decision nodes have multiple branches and are used to decide while the leaf node didn’t contain further branches and represents the output of decisions. The decision tree simply asks a question and depending on this question's answer, which may be (yes/no) it split the tree into other subtrees. Figure 4.4 represents the general structure for the decision tree algorithm [50].

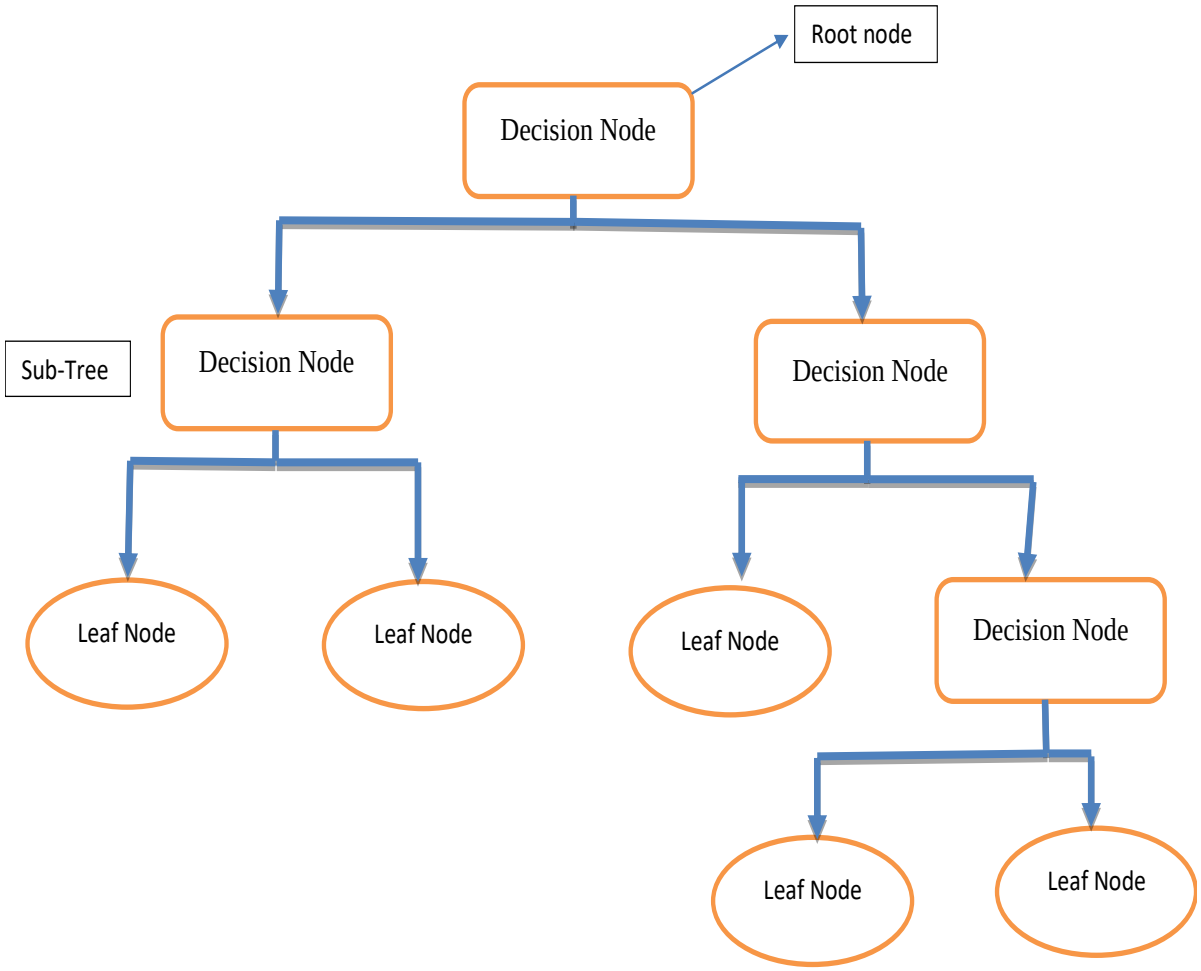


Figure 4.4: Decision tree structure

The decision tree algorithm is easy to understand because while making a decision it mimics human thinking and it shows the tree-like structuring.

4.4.1.3 Random Forest Classifier

Random forest is one of the common machine learning classifiers, it creates a collection of trees trained separately in parallel, each tree will be trained on all subsets of the training dataset using different subsets of dataset features, the final decision of the model will be the majority voting of all trees. The random forest algorithm is a collection of decision trees, this type of classification is called ensembling learning

since a lot of classifiers are stacked together to improve classification performance. Figure 4.5 shows the main structure of the random forest [51].

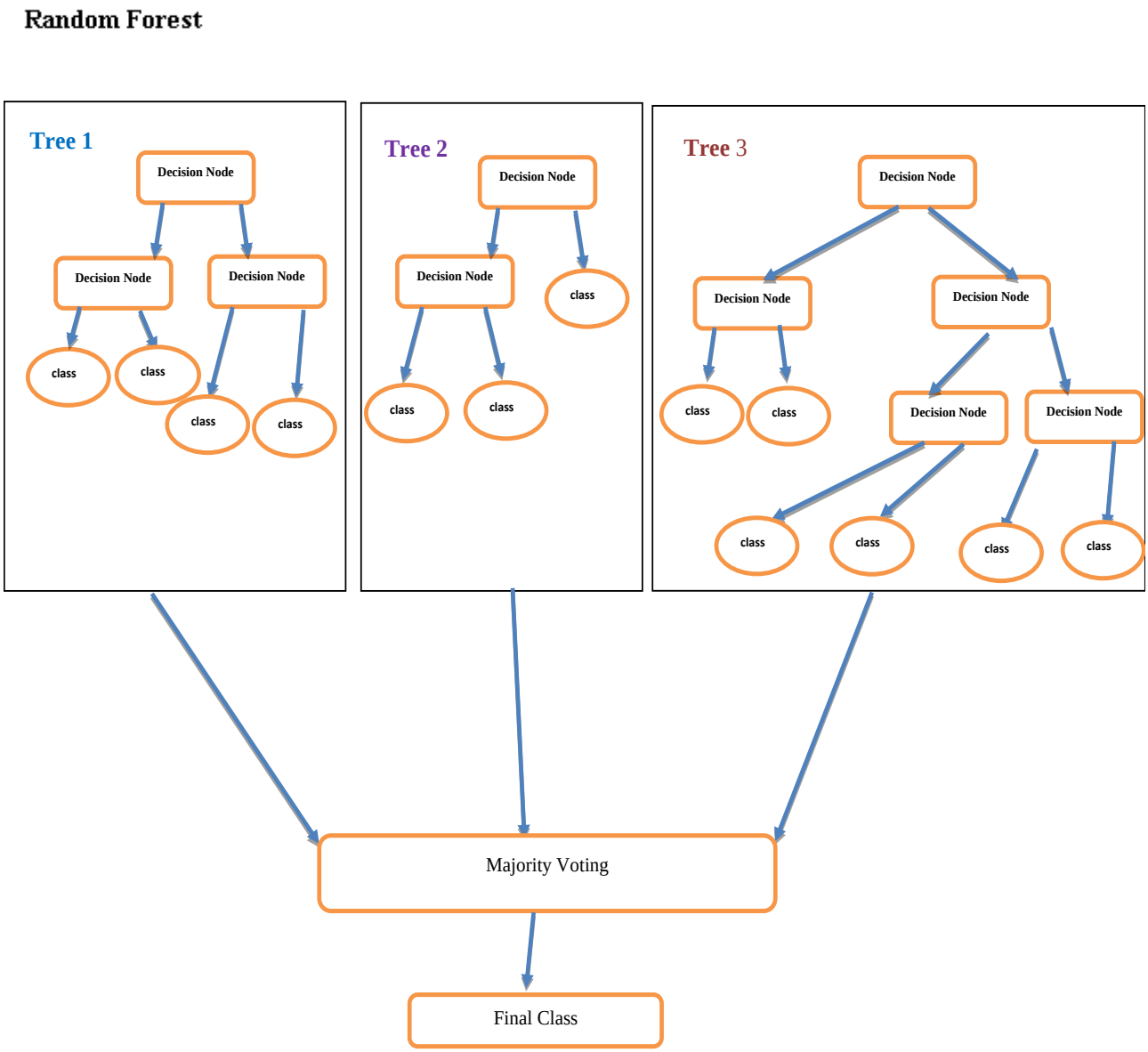


Figure 4.5: Random forest structure

Figure 4.4 shows the structure of the decision tree, the forest consists of a much-separated decision tree that works separately in parallel, and gives different results, the trees together present a forest and the result will be the majority voting of all trees.

4.4.1.4 Multilayer Perceptron

The neural network concept depends on simulating the human brain's biological neurons [52] the neural network can learn from the information to be able to predict some values based on the information, which was trained on. One and the most used neural network types are called multilayer perceptron neural network (MLPNN) [53], it works based on a method called feed-forward which consists of at least three layers; the input layer which accepts the variables, the hidden layer which performs the learning process, and the output layer which shows prediction the results. Each layer consists of many neurons and the number of neurons is not necessary to be the same for all layers, each neuron is connected with all neurons in the next layer, so it is called a fully connected algorithm, each neuron–neuron connection has a weight value, which should differ during the learning iterations. The structure of MLPNNs is depicted in Figure 4.6.

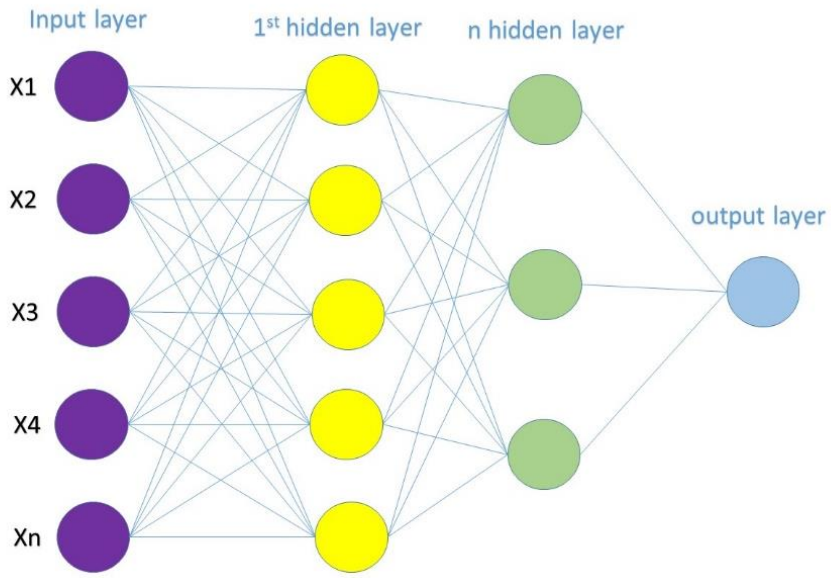


Figure 4.6: The structure of MLPNNs

MLPNN consists of two main phases; the first phase is the forward, which predicts the output and calculates the error then the error will be sent to the second phase, which is

backward propagation. The error will be used to adjust the weights of the connection to improve the output results in the next iteration, this process will continue until finishing the iteration or until achieving the required results.

The MLPNN training process starts while providing the network with the input data and the initial weights, then the algorithm will start adjusting the weights during the forward and backward propagation to minimize the error ratio between the actual and the desired values.

The MLPNN aimed to find the optimal mapping between the inputs and the corresponding output, the final output of the MLPNN was calculated using equation 4.3 [53].

$$Y_{ij} = \sum_{i=1}^{P} w_{ij}.x_i \qquad 4.3$$

Where w_{ij}: is the connection weight between the ith node in the input layer and the jth node in the hidden layer, and x_i: is the ith input.

To stop the training process, there is a certain threshold θ is set depending on the error of the MLPNN which represents the difference between the desired and actual output.

The error is calculated using equation 4.4 [54] :

$$MSE = \frac{1}{\mathrm{n}}\sum_{i=1}^{n}(y_d - y_i)^2 \qquad 4.4$$

Where n sample data points, y_d is actual value and y_i is predictive value.

The output of the MLPNNs is the weighted sums of the inputs which are calculated using the following equation: 4.5 [55]:

$$Y_{ij} = w_{ij}.x_i \qquad 4.5$$

Where w_{ij}: represents the connection weight between the node number I in the input layer and the node number j in the hidden layer, and x_i: is the ith input.

4.4.2 Regression

Classification task used to predict a discrete outcomes class label, in our case to binary, classify the human skin for healthy or unhealthy. Since the second task in our work is to predict the value of human skin emissivity, a regression model is used to predict the human skin emissivity using four different algorithms. The regression predicts continuous values for outcomes by utilizing the relationship between two or more quantitative variables, there are two types of variables one is the dependent variable (response variable) and the other one is independent (predictor variable). In the regression process, the regression algorithm investigates the relations between the elements so that the dependent variable value could be predicted by the independent variable.

Some machine learning could be used for both prediction tasks, classification, and regression such as KNN and multi-layer perceptron, we used linear regression, KNN, SVM, and MLPNN for the regression model to predict the value of skin emissivity measure on a specified skin location based on measures from other locations, Figure 4.7 shows the proposed method for regression.

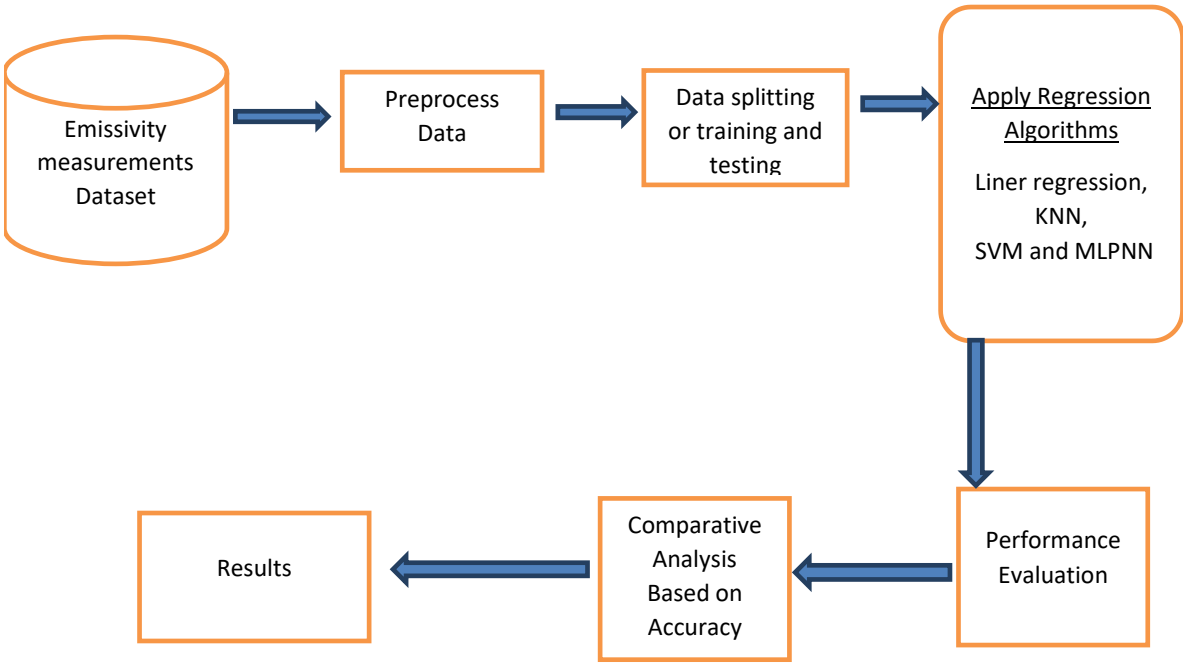


Figure 4.7: Basic workflow for skin location emissivity prediction models.

The regression process starts with collecting and forming the data from previous studies of emissivity, this data will organize and labeled to be used by the python libraries. After collecting the dataset, this data needs to be normalized, standardized, and clean of outliers and highly correlated features. The dataset after preprocessing phase is ready to be split for training and testing data and then entered for one of the four regression algorithms, the performance of the machine learning algorithm will be measured using model accuracy by subtracting the mean absolute percentage error of 100%. and error metrics, and finally, the results of the different regression algorithms will be compared.

As mentioned above, two prediction algorithms are used for classification and regression, which are KNN and multilayer perceptron neural network, the linear regression and SVM will be described in section (4.4.2.2).

4.4.2.1 Linear Regression

Linear regression is the simplest prediction algorithm, the concept of linear regression depends on the model of a relationship between the dependent and independent variables, by fitting a linear equation to the needed data, the linear regression could be single or multilinear regression based on the number of independent variables. In this work, we will use multi-linear regression since the number of independent variables is more than one, the multi-linear regression algorithm is formulated as in equation 4.6 [56] .

$$\hat{Y} = a_0 + \sum_{j=1}^{m} a_j X_j, \qquad 4.6$$

Where $\hat{Y}$ is output for the regression model, X_j is the model inputs variables m is the number of partial coefficients and a_0, a_1 …..a_m are the partial regression coefficients.

4.4.2.2 Support Vector Machines (SVM)

The SVM [57] is a well-known machine learning technique developed at the beginning to handle classification issues, after that its principle was extended to involve regression [58]. There are many types of SVM kernels such as linear, polynomial, and Gaussian.

To categorize the data points, the SVM mapped each data record to a high-dimensional feature space for the data, a separator between the categories is found even for not separated linearly datasets. Then the dataset will be reshaped in such a way that the separator could be created as a hyperplane. Any new record will be analyzed to find where it should belong using the characteristics of new data.

In this study, we used the linear kernel to find a function that gives the optimal results for the regression challenge as equation 4.7 [59].

$$f(x) = \omega . \emptyset(x) + b \qquad 4.7$$

While ω is the vector weight, $\emptyset(x)$ is the high-dimensional feature spaces that are nonlinearly mapped from the inputs and b represents the scalar threshold.

4.5 Metrics Selection

There are several metrics associated with data classification and regression that statistically measure its performance [60]. This thesis will focus on the following metrics for classification: confusion matrix, true positive (TP), false positive (FP), false negative (FN), true negative (TN), accuracy, sensitivity (recall), specificity, and precision. For regression, the following were used: accuracy (score coefficient), mean absolute error, mean squared error and root mean squared error.

4.5.1 Metrics For Classification

TP: the samples number was correctly classified as dry (healthy).

FP: the samples number was incorrectly classified as dry (healthy).

TN: the samples number was correctly classified as wet (unhealthy).

FN: the samples number was incorrectly classified as wet (unhealthy).

A Confusion Matrix: is a table that is used to show the classification model results. The table contains two-diminution, the rows represent the predicted class values, and the columns represent the real class values. The confusion matrix is used to compute most

of the performance measures. Table 4.2 describes the confusion matrix for classification the of skin status.

Table 4.2: Confusion matrix description for skin status classification

		Predicted Classes	
		Dry	Wet
Actual Classes	Dry	TP	FP
	Wet	FN	TN

Accuracy: The main metric that is used to measure the performance in class pattern recognition and classification, which represented by the following formula:

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN} + \mathrm{TN}} \qquad 4.8$$

Sensitivity or Recall: The percentage of records that are classified correctly as healthy to all records that are classified as healthy. It is the percentage of records that are predicted to a certain class correctly to all records predicted in that class. It is calculated using equation 4.9.

$$Recall = \frac{T\mathrm{P}}{T\mathrm{P} + F\mathrm{N}} \qquad 4.9$$

F-measure: it measures the balance between precision and sensitivity (recall). The F-measure is calculated using equation 4.10:

$$\text{F-measure} = \frac{2 * recall * precision}{recall + precision} \qquad 4.10$$

Precision: The percentage of records correctly predicted as healthy to all records predicted in healthy class and it can be described as shown in equation 4.11:

$$Precision = \frac{TP}{TP + \mathrm{FP}} \qquad 4.11$$

4.5.2 Metrics For Regression

Accuracy: the regression model accuracy is calculated by subtracting the mean absolute percentage error MAPE of 100%, equation 4.13 [61] shows the MAPE

$$MAPE = \frac{\sum_{t=1}^{n} y_{i-} \hat{y_i}}{n} \qquad 4.13$$

Mean Absolute Error (MAE) is the mean of the absolute value of the errors, equation 4.14 shows how to calculate MAE.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left(|y_{i-} \hat{y_i}| \right) \qquad 4.14$$

4.6 Conclusion

This chapter summarizes the methodology of prediction and classification of the human skin emissivity; the first step of the methodology is data preparation and exploration, the second step is data preprocessing for machine learning; the third step is to use of different machine learning algorithms for predicting the human skin emissivity and then to classify dry skin (normal or healthy skin) and wet skin (unhealthy skin), and finally use different performance metrics to compare between different machine learning algorithms performance and select the best approach.

Chapter Five

Results

5.1 Introduction

After deploying the models for the classification and regression, we used the mentioned metrics to evaluate the models' efficiency with different classification and regression algorithms. In this chapter, we will illustrate the results for classification using accuracy, confusion matrix, precision, recall, and F-score. For the regression, we calculate the accuracy, MAE, for each regression algorithm and each skin location prediction.

5.2 Classification Results

5.2.1 Confusion Matrix

The confusion matrix shows a summary of the testing data of the correctly and incorrectly predicted samples, Table 5.1 shows the optimal confusion matrix while using the different classification algorithms.

Table 5.1: Confusion matrix using different classification algorithms

		Predicted Classes							
		KNN		**Random Forest**		**Decision Tree**		**MLPNN**	
		Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet
Actual Classes	Dry	43%	10%	46%	6%	43%	13%	46%	6%
	Wet	13%	33%	20%	26%	6%	36%	3%	43%

Table 5.1 shows the confusion matrix for the classification algorithms results, the cells with red colors represent the percentage of the samples that were predicted wet while the desired is dry or predicted dry while it is wet. The green cells give the percentage of the samples which was predicted correctly. For each classification algorithm, the top-left cell represents the sample percentage correctly categorized as dry (**TP**), the top right represents the percentage of the samples incorrectly categorized as dry (**FP**), the bottom left represents the samples percentage correctly categorized as wet (**TN**) and the bottom right the percentage of the samples incorrectly categorized as wet (**FN**).

5.2.2 Accuracy

5.2.2.1 KNN Accuracy

The accuracy for each classification algorithm with different parameters was tested to find the optimal algorithm parameters Figure 5.2 shows the accuracy for the KNN algorithm while using different values of K neighbors and with different distance metrics.

Table 5.2: Confusion matrix description for skin status classification

KNN algorithm results										
K / Metrics	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Euclidean	56%	60%	63%	63%	70%	73%	66%	70%	60%	56%
Manhattan	66%	66%	63%	66%	73%	80%	83%	80%	73%	66%
Minkowski	66%	63%	66%	66%	73%	83%	80%	80%	76%	73%
Hamming	50%	56%	66%	63%	70%	66%	63%	63%	56%	60%

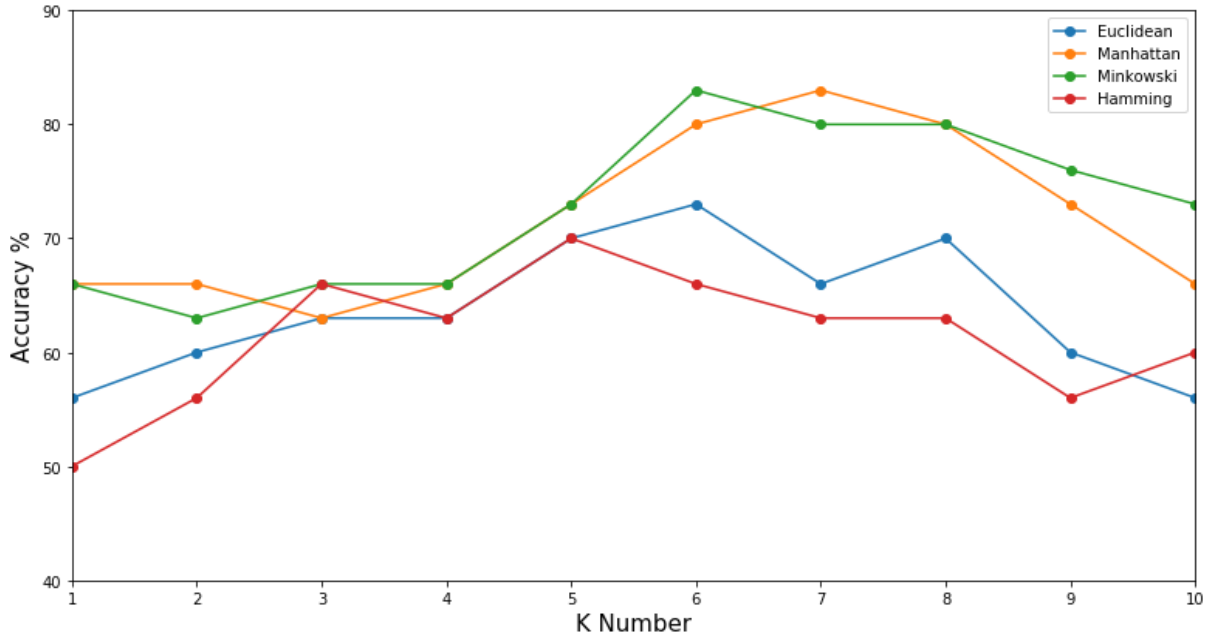


Figure 5.1: Accuracy measures using different K numbers and different distance metrics algorithms

Table 5.2 and Figure 5.1 show the results for the KNN algorithm while using four different distance metrics algorithms which are Euclidean, Manhattan, Minkowski, and Hamming, the results also show the accuracy using a different number of K neighbors. The best accuracy results were gained using the Minkowski algorithm and while using 6 K, the accuracy starts to go down while using a larger k number of 6 for all distance metrics.

5.2.2.2 Random Forest Accuracy

The accuracy for the random forest was tested using different values of N estimators, which is the number of trees to be used for prediction voting. The accuracy starts to be stable while using 12 N estimators, using more than 15 N estimators didn't improve the accuracy while it starts slowing the code, Figure 5.2 shows the accuracy while using 1,2,3,4,5,6,7,8,9,10,11,12,13 N numbers.

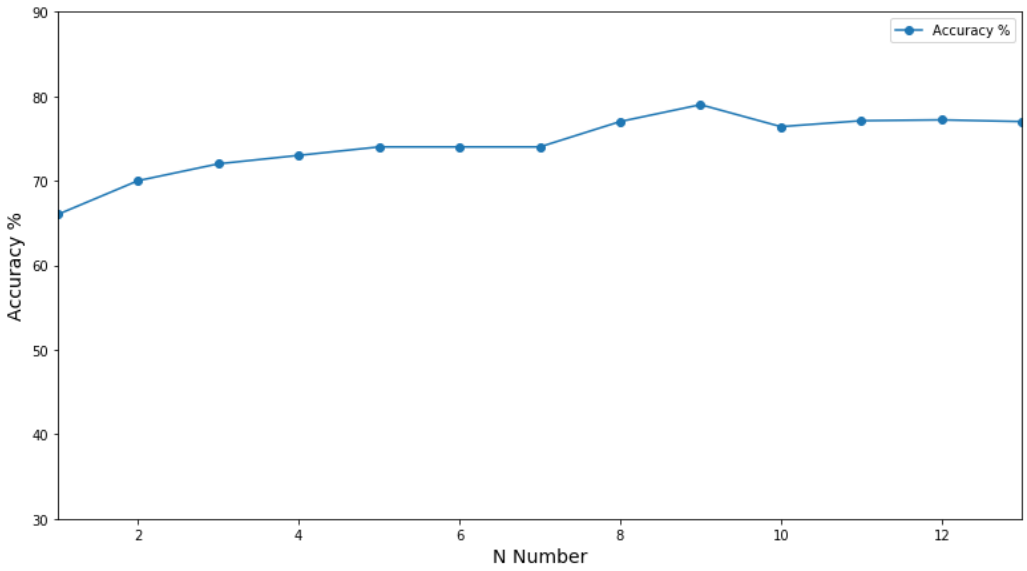


Figure 5.2: Accuracy measures using different N decision trees in the random forest.

5.2.2.3 Accuracy Using Decision Tree

Decision tree accuracy was tested using different values of min samples split which refers to the minimum number of samples per node, Figure 5.3 shows the accuracy for the different min split values.

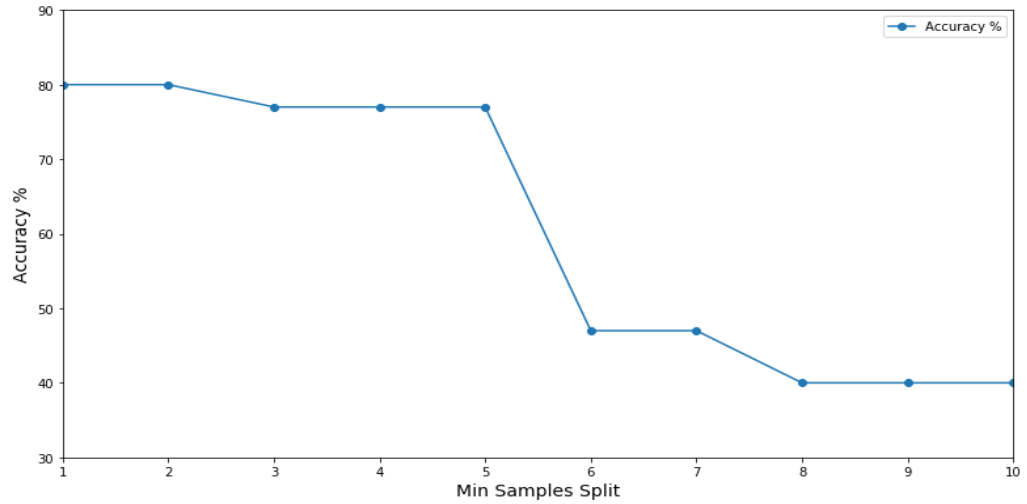


Figure 5.3: Accuracy measures using different min samples split numbers for decision tree.

The accuracy in our case starts being down while increasing the value of the mini-split number, the best accuracy was gained using 10 samples split number and starts suddenly decreasing while using more than 50 min samples split.

5.2.2.4 Accuracy Using Multilayer Perceptron Neural Network

The accuracy while using a multilayer perceptron neural network affects many different parameters such as the number of hidden layers, number of neurons, number of epochs, and activation algorithm. Table 5.3 and Figure 5.4 show the accuracy measures while using a different number of the hidden layer, different scale of epochs, and a different number of neurons, H represents the number of hidden layers and N represents the number of neurons for each layer.

Table 5.3: Accuracy results using MLPNN with different parameters.

	H / N	1	2	3	4	5	6
10 epoch	**10**	44.4%	69.7%	69.5%	50.5%	50.5%	50.2%
	100	61.2%	72.6%	69.5%	72.5%	72.5%	71.4%
	200	71.2%	72.5%	72.2%	72.2%	72.2%	72.3%
	300	61.5%	69.2%	63.3%	63.2%	62.2%	63.5%
	400	61.5%	61.1%	63.5%	63.2%	61.6%	61.4%
	500	61.6%	50.4%	61.4%	63.2%	63.4%	61.5%
	H / N	**1**	**2**	**3**	**4**	**5**	**6**
100 epoch	**10**	72.6%	72.2%	69.3%	72.3%	80.5%	80.05%
	100	72.6%	80.3%	82.6%	82.4%	82.5%	83.3%
	200	72.6%	82.5%	88.9%	91.6%	88%	77.2%
	300	71.7%	82%	88.6%	91.6%	88.3%	80.7%
	400	72.6%	82.4%	88.6%	90.5%	89.5%	79.2%
	500	72.6%	86%	88.4%	90.5%	87%	91.6%

	H / N	1	2	3	4	5	6
500 epoch	**10**	72.6%	74.2%	72.6%	77.4%	77.2%	77%
	100	80.2%	80.4%	82.7%	84%	84.6%	84%
	200	80.6%	82.7%	88.9%	90.5%	88.6%	88.6%
	300	82.4%	82.4%	88.9%	88.9%	89.8%	87.8%
	400	80.2%	82.7%	88.8%	88.9%	87.8%	89.8%
	500	82.4%	89.8%	88.9%	88.6%	88.8%	88.8%

Table 5.3 shows the accuracy results for different deep learning parameters, the best accuracy results were obtained using 4 hidden layers and using 200-300 neurons trained for 100 epochs, Figure 5.4 shows the accuracy results while using 4 hidden layers and while using different epochs and different numbers of neurons.

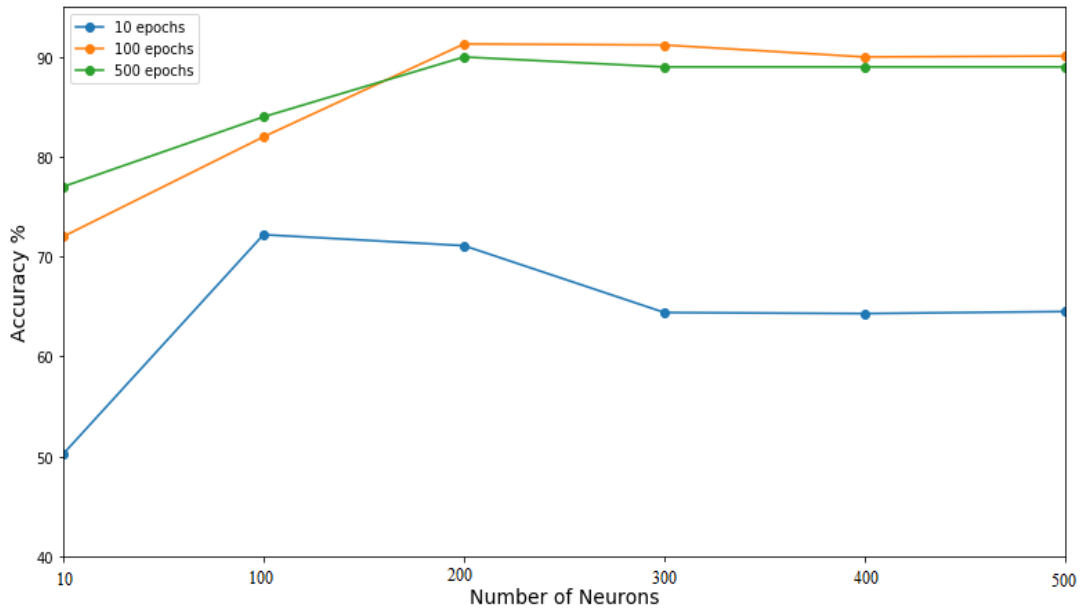


Figure 5.4: Accuracy measures using 4 hidden layers and different numbers of epochs.

For 10 epochs number, the accuracy is not high since the epochs number is not enough to complete the learning state, and the model while using 10 epochs will be under a fitting state. While using 100 epochs the accuracy keeps increasing until reaches 200

neurons with 91.6% accuracy, while using 500 epochs the accuracy is not better than 100 epochs, since the model starts being in an over fitting state because of that, using 100 epochs, 4 hidden layers and 200 neurons for each layer is enough.

5.2.3 Recall, Precision, and F-measure

The recall, precision, and F-measure were calculated with optimal results for all classification algorithms, using the mentioned equations: 4.9, 4.10, and 4.11, Table 5.4 and Figure 5.5 show the recall, precision, and F-measure for KNN, random forest, decision tree, and MLPNN algorithms.

Table 5.4: Recall, precision, and F-measure for KNN, random forest, decision tree, and MLPNN algorithms.

	Recall	Precision	F-measure
KNN	76%	81%	78.4%
Random-forest	70%	87.5%	77.7%
decision tree	86.6%	76.4%	81.1%
MLPNN	93.3%	87.5%	90.3%

Figures 5.5, 5.6, and 5.7 show the recall, precision, and F-measure sequentially result using Minkowski distance with 7 K neighbors for KNN algorithm, 15 trees to be used for prediction voting for random forest algorithm, 10 samples split number for the decision tree, and 100 epochs, 4 hidden layers and 200 neurons for each layer while using MLPNN.

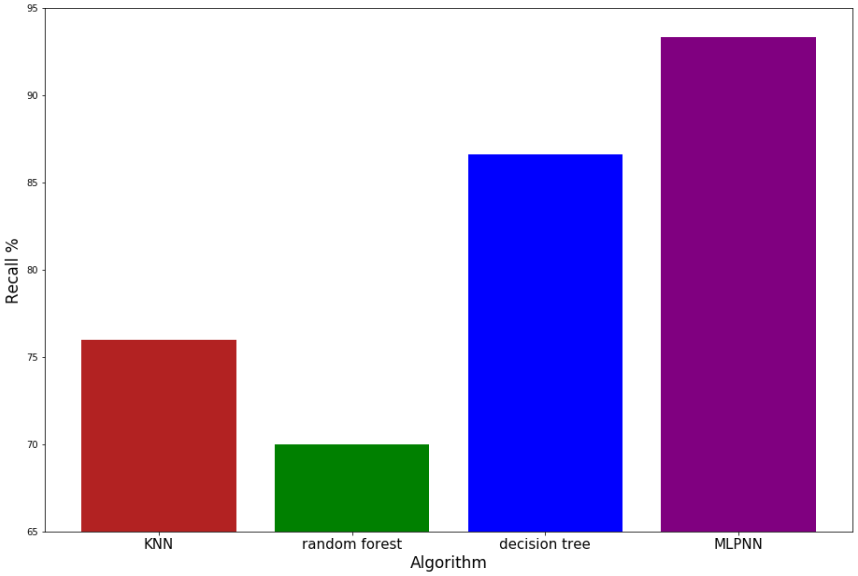


Figure 5.5: Recall score using a different machine learning algorithm.

The highest recall score was achieved using MLPNN with 93.3%, then decision tree with 86.6%, KNN with 76%, and the lowest recall was achieved using random forest with 70%.

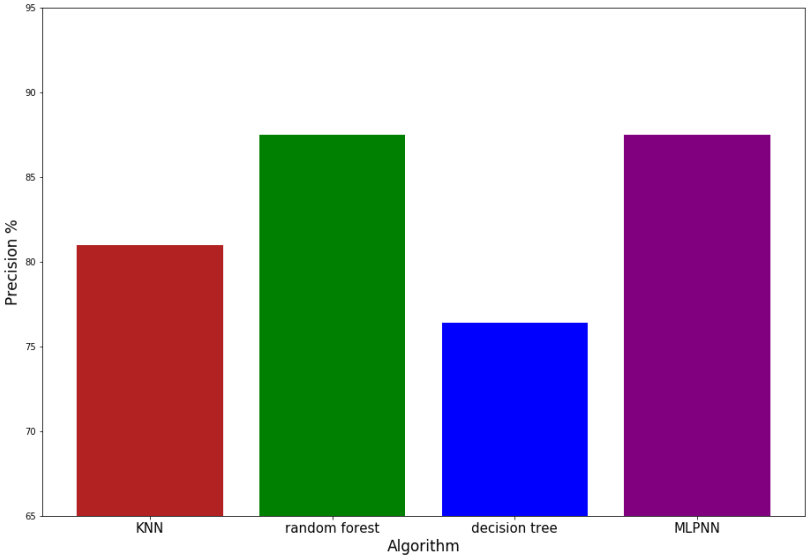


Figure 5.6: Precision score using a different machine learning algorithm.

Both MLPNN and random forest achieved the same Precision score with 87.5%, then KNN with 81%, and the lowest Precision score was achieved using the decision tree with 76.4%.

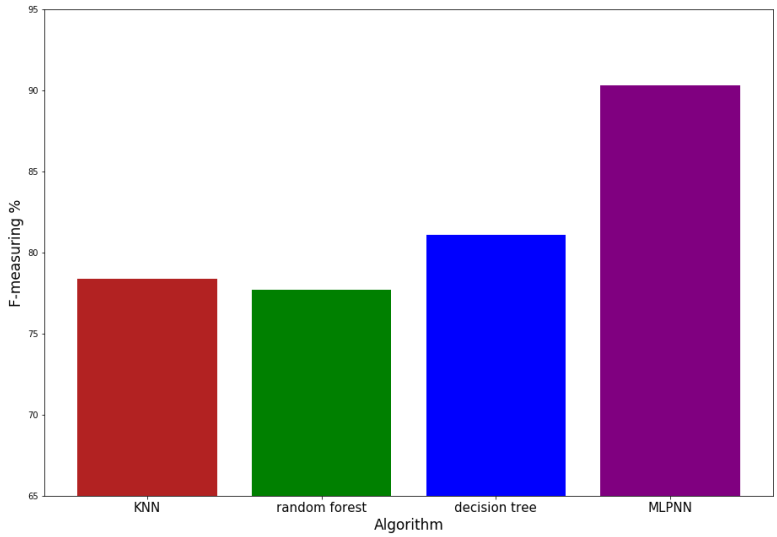


Figure 5.7: F-measure score using a different machine-learning algorithm.

MLPNN achieved the highest F-measure score with 90.3%, then the decision tree with 81.1%, and KNN with 78.4%, and the lowest F-measure score was achieved using a random forest classifier with 77.7%.

Figures 5.5, 5.6, and 5.7 show that the MLPNN achieved the highest score for recall with a 7.44% difference from the random forest and the highest score for F-measure with a difference of 10.7% than decision tree and equals random forest with precision.

5.3 Regression Results

5.3.1 Accuracy Records

We removed the highly correlated features for each measurement location separately since the highly correlated features when predicting the palm of hand measures differ

from the highly correlated features when predicting the back of hand measures for example.

We made the regression task using the four different algorithms and using the data after or before removing highly correlated data or removing outliers. Table 5.6 and Figure 5.8 show the accuracy results using the linear regression since BRF shows the accuracy before removing highly correlated features, ARF shows the accuracy after removing the highly correlated features and ARO shows the accuracy after removing outliers. Table 5.5 shows each location and its number in the Figures.

Table 5.5: The measurements location numbering.

Location	Palm of hand	Back of hand	Fingers	Inner wrist middle	Inner wrist side	Outer wrist	Volar side	Dorsal Surface	Elbow
number	1	2	3	4	5	6	7	8	9

Table 5.6: BRF, ARF, ARO accuracy results using a linear regression algorithm.

	BRF	ARF	ARO
Palm of hand	85.4%	87.1%	88.2%
Back of hand	86.4%	82.2%	83.2%
Fingers	45.3%	76.6%	76.7%
Inner wrist middle	67.3%	84%	84.6%
Inner wrist side	61.7%	61.7%	71.3%
Outer wrist	69%	70.6%	70.8%
Volar side	49.8%	83.1%	82.3%
Dorsal Surface	69.1%	76.4%	74.6%
Elbow	88.8%	82.1%	80.2%

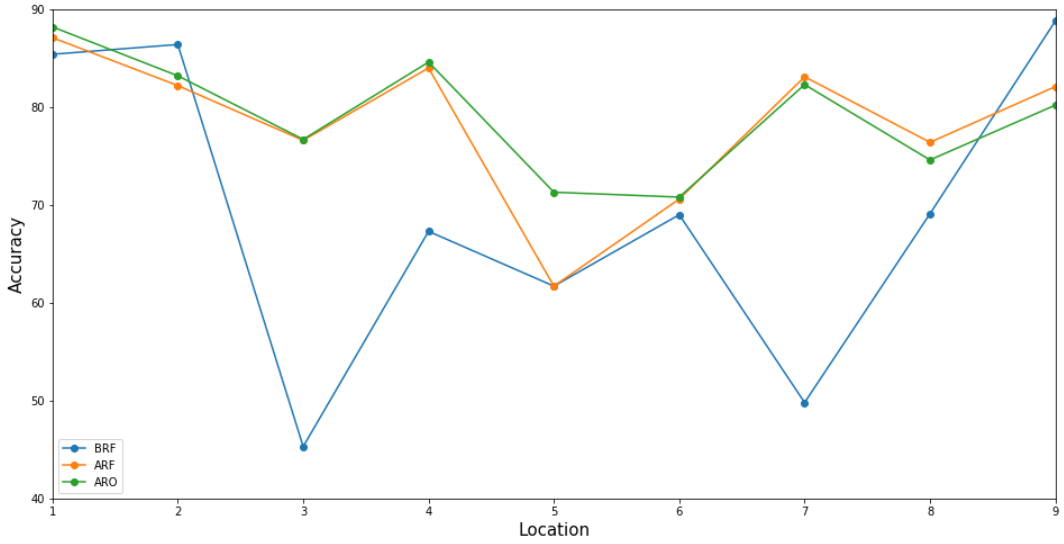


Figure 5.8: BRF, ARF, ARO accuracy results using a linear regression algorithm.

Also using the KNN algorithm we did the regression after and before removing the outlier and highly correlated features, Table 5.7 and Figure 5.9 show the accuracy results for the KNN algorithm.

Table 5.7: BRF, ARF, ARO accuracy results using KNN.

	BRF	ARF	ARO
Palm of hand	67.9%	81.1%	86.3%
Back of hand	67.5%	88.4%	81.9%
Fingers	43.6%	56.2%	55.3%
Inner wrist middle	51.5%	85.2%	81%
Inner wrist side	56.8%	56.8%	52.4%
Outer wrist	46.1%	60.2%	62.7%
Volar side	60.5%	72.7%	75.7%
Dorsal Surface	67.3%	74.8%	85.3%
Elbow	81.8%	82.7%	71.2%

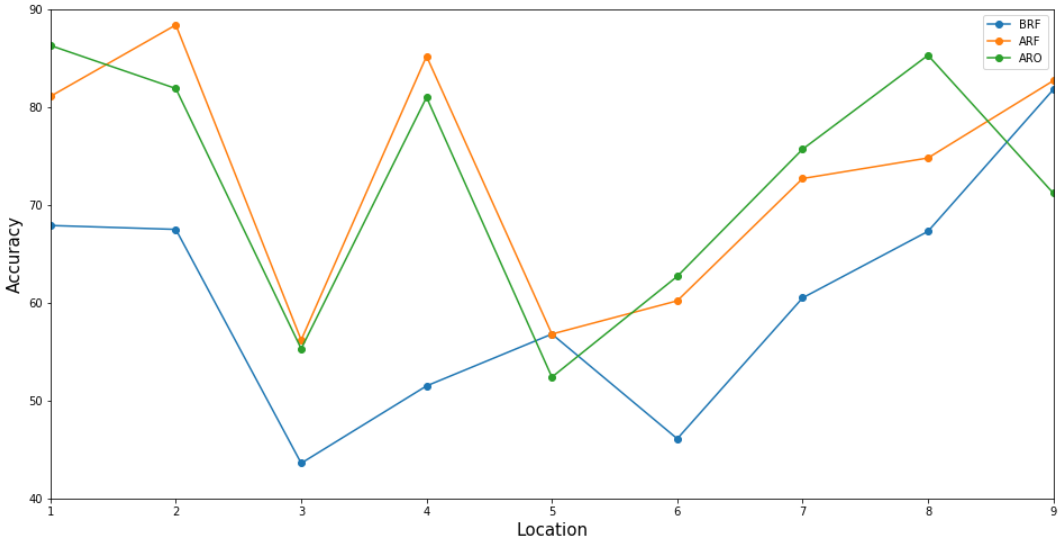


Figure 5.9: BRF, ARF, ARO accuracy results using the KNN algorithm.

The SVM also did the regression before and after removing the highly correlated and outlier features, Table 5.8 and Figure 5.10 show the results.

Table 5.8: BRF, ARF, ARO accuracy results using SVM

	BRF	ARF	ARO
Palm of hand	55.1%	70.4%	54%
Back of hand	61.1%	70.4%	77%
Fingers	39%	44.1%	52.9%
Inner wrist middle	16.3%	23%	45%
Inner wrist side	12.4%	45.9%	45.9%
Outer wrist	20.2%	32%	32.9%
Volar side	12.6%	35.6%	26.4%
Dorsal Surface	10%	20.4%	6.9%
Elbow	18.3%	60.2%	60.9%

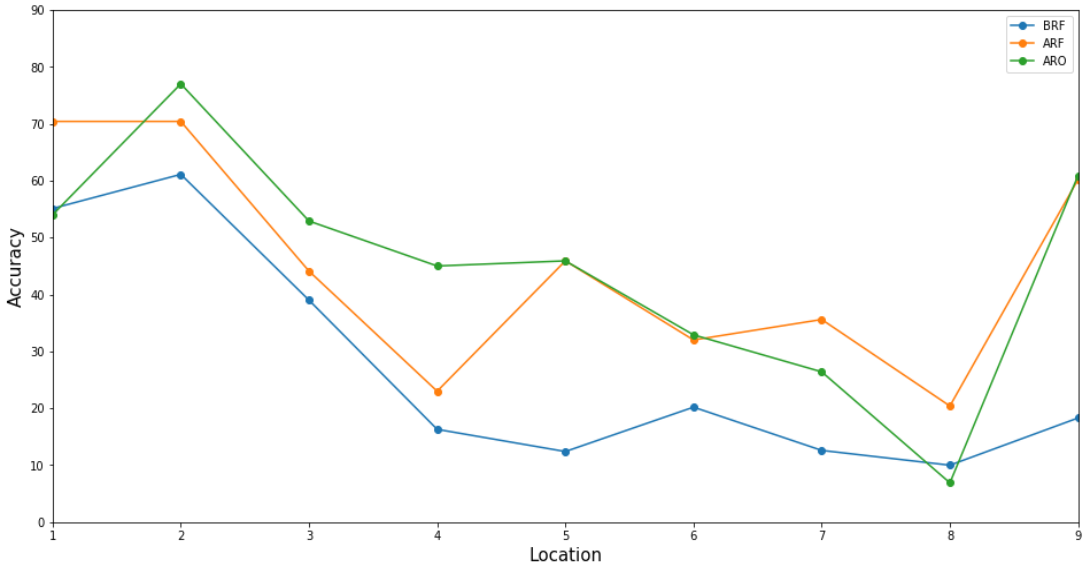


Figure 5.10: BRF, ARF, ARO accuracy results using the SVM algorithm

We create an MLPN model using input, output, and 2 hidden layers to predict the locations using the BRF, ARF, and ARO datasets. Table 5.9 and Figure 5.11 show the accuracy results for the MLPN with 3 hidden layers, 3500 neurons for each layer for 200 iterations.

Table 5.9: BRF, ARF, ARO accuracy results using MLPNN algorithm.

	BRF	ARF	ARO
Palm of hand	87.4%	89.5%	91.3%
Back of hand	86%	86.3%	87.2%
Fingers	75%	75.1%	79%
Inner wrist middle	82.5%	84.9%	85.1%
Inner wrist side	64.3%	61.3%	71.7%
Outer wrist	70.2%	71.5%	71.7%
Volar side	83.2%	83%	82.8%
Dorsal Surface	77%	77.5%	78.2%
Elbow	83%	85.6%	86.5%

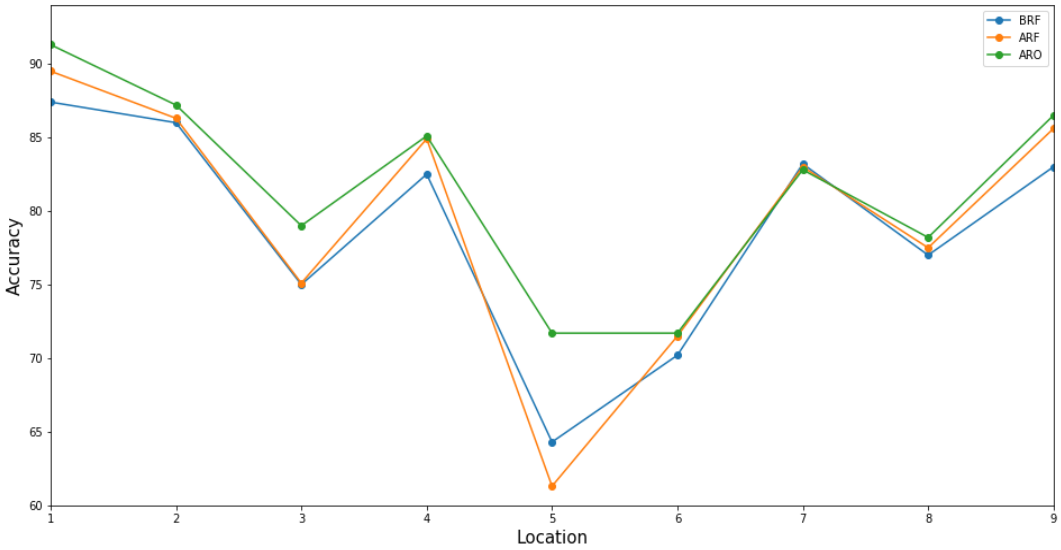


Figure 5.11: BRF, ARF, ARO accuracy results using MPNN algorithm

Table 5.6 and Figure 5.8 show the accuracy results while predicting the emissivity measures for the 9 different locations and for the different machine learning algorithms using dataset before removing highly correlated features, dataset after removing highly correlated features, and dataset after removing highly correlated features and outliers. The highest accuracy measures were recorded while using MLPNN dataset after removing highly correlated features and outliers, Figure 5.12 shows a comparison between the different machine learning algorithms using data after removing outliers and highly correlated features. Some locations such as the palm of the hand and back of the hand achieved higher accuracy records than other locations, because of the high correlation between these locations, so the palm when used to predict the back of the hand will give accurate results because of the high correlation between the palm and the back of the hand.

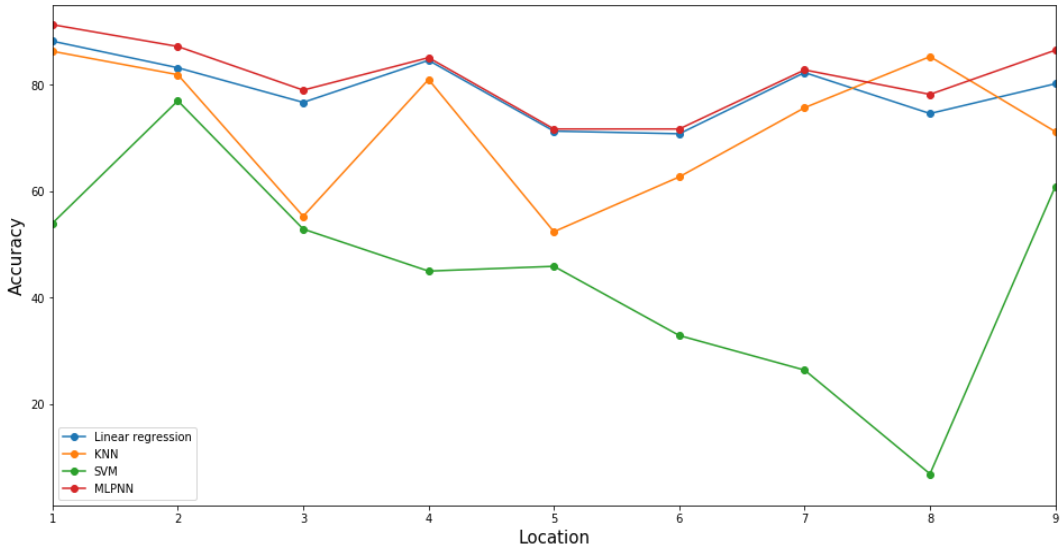


Figure 5.12: Comparison between the different machine learning algorithms using ARO data

Figures 5.13 – 5.21 show the actual emissivity values vs the predicted emissivity values for the 9 locations using the highest accuracy condition while using MLPNN with the dataset after removing the highly correlated.

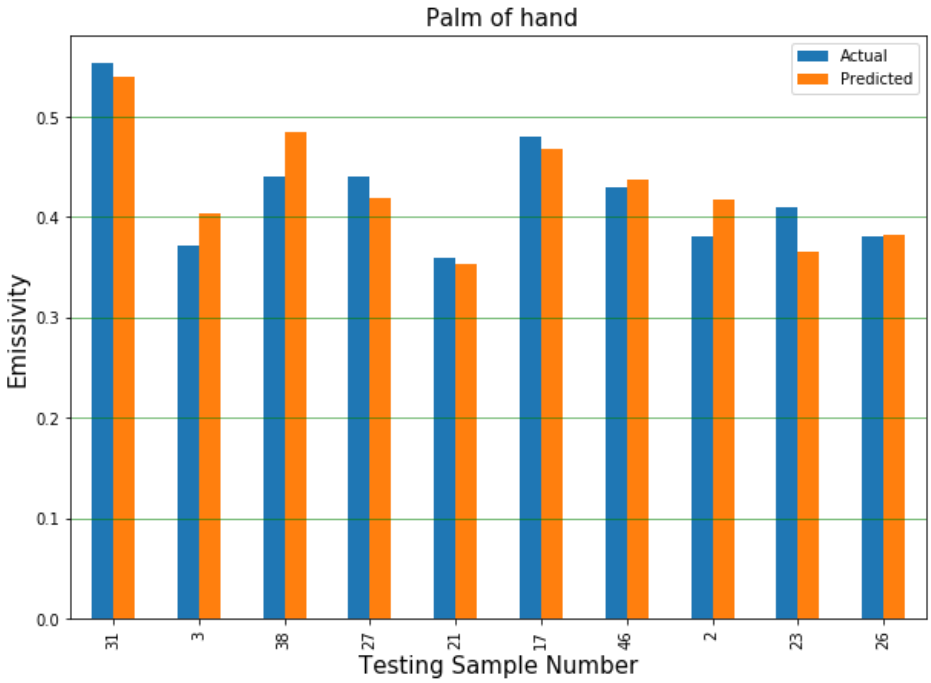


Figure 5.13: Predicted vs actual emissivity for the palm of the hand.

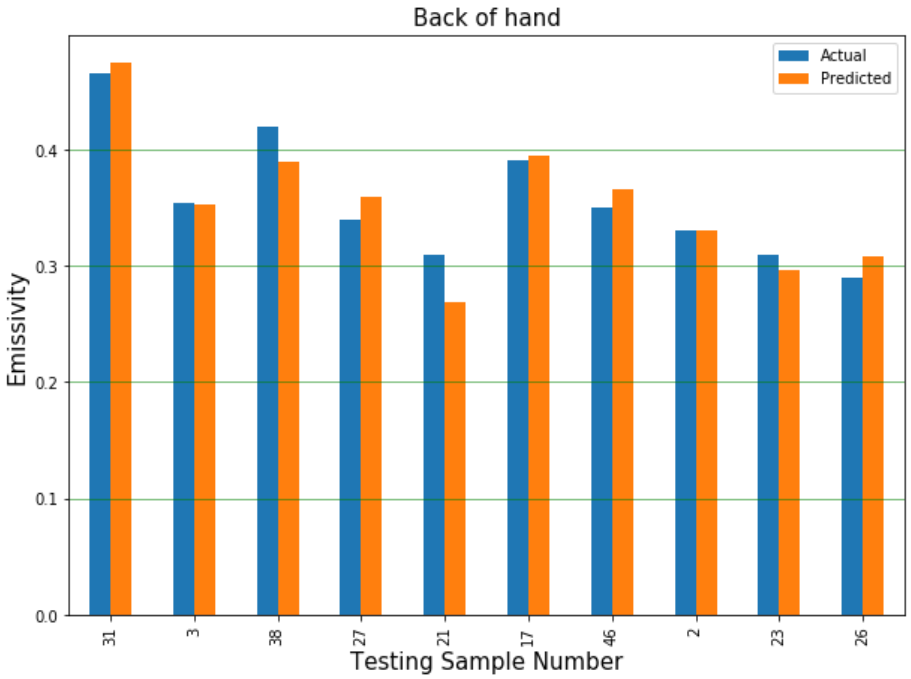


Figure 5.14: Predicted vs actual emissivity for the back of the hand.

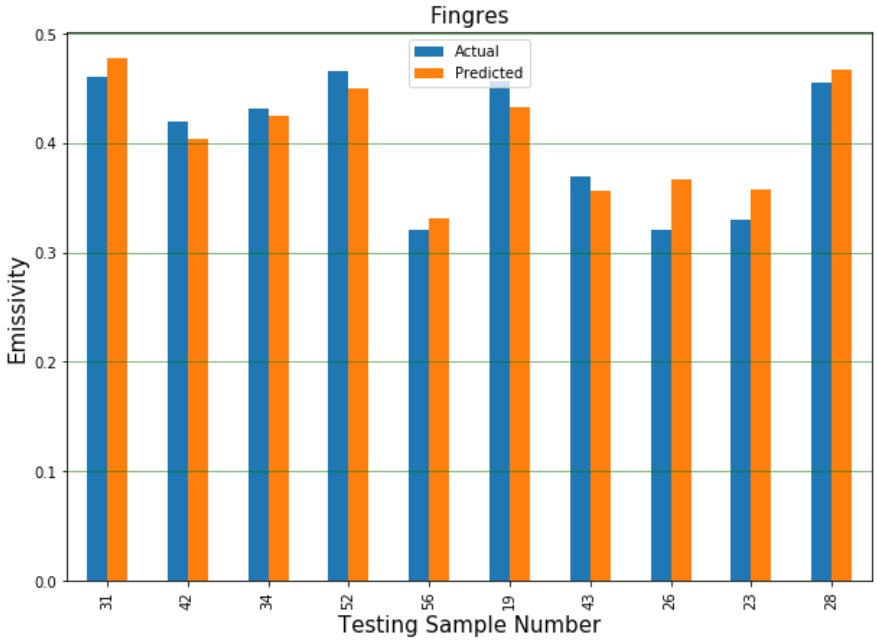


Figure 5.15: Predicted vs actual emissivity for fingers.

Figure 5.16: Predicted vs actual emissivity for inner wrist middle.

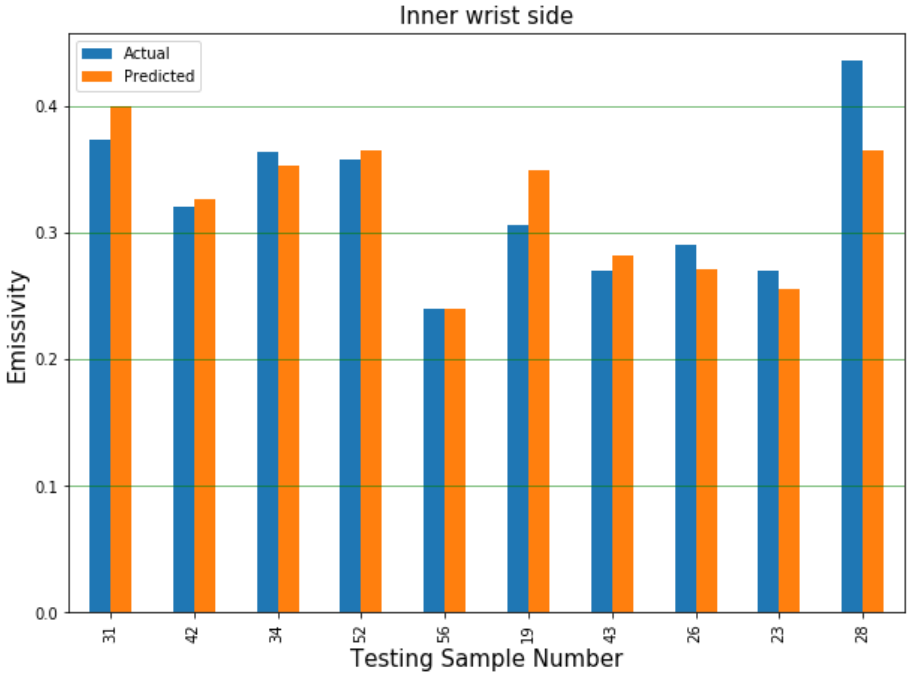


Figure 5.17: Predicted vs actual emissivity for inner wrist side.

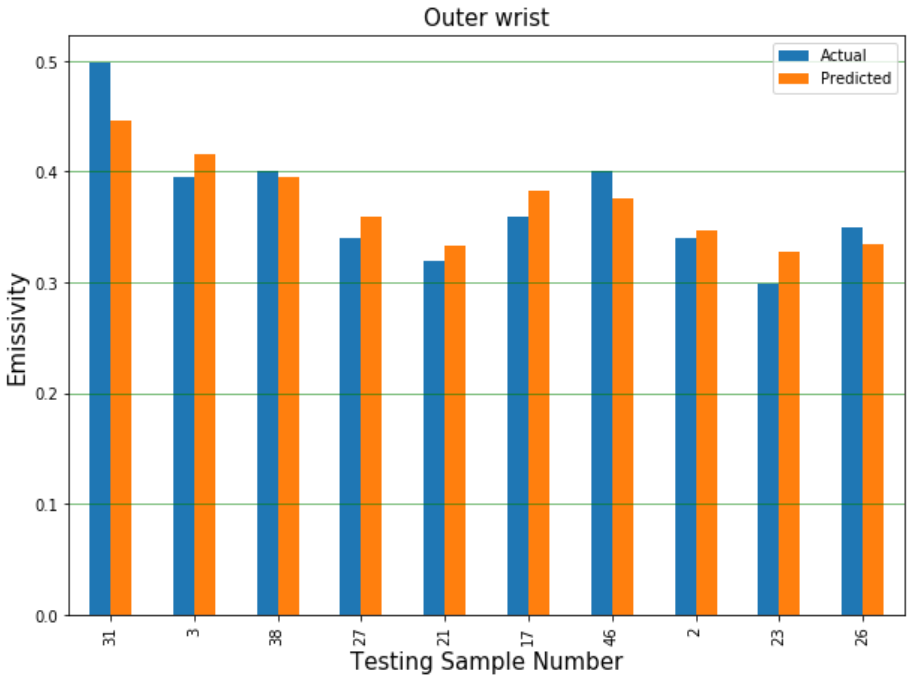


Figure 5.18: Predicted vs actual emissivity for the outer wrist.

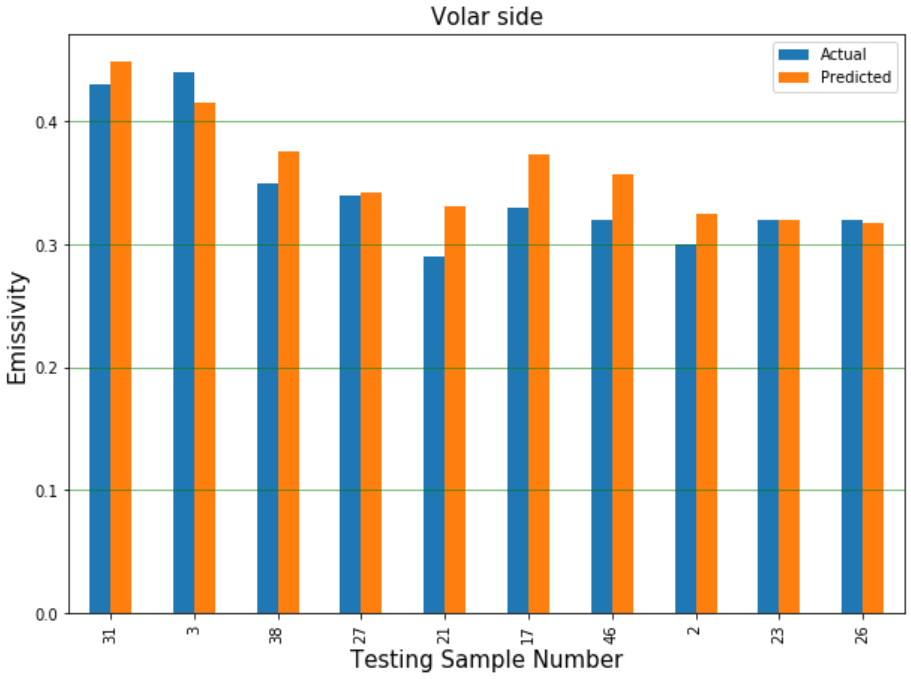


Figure 5.19: Predicted vs actual emissivity for the volar side.

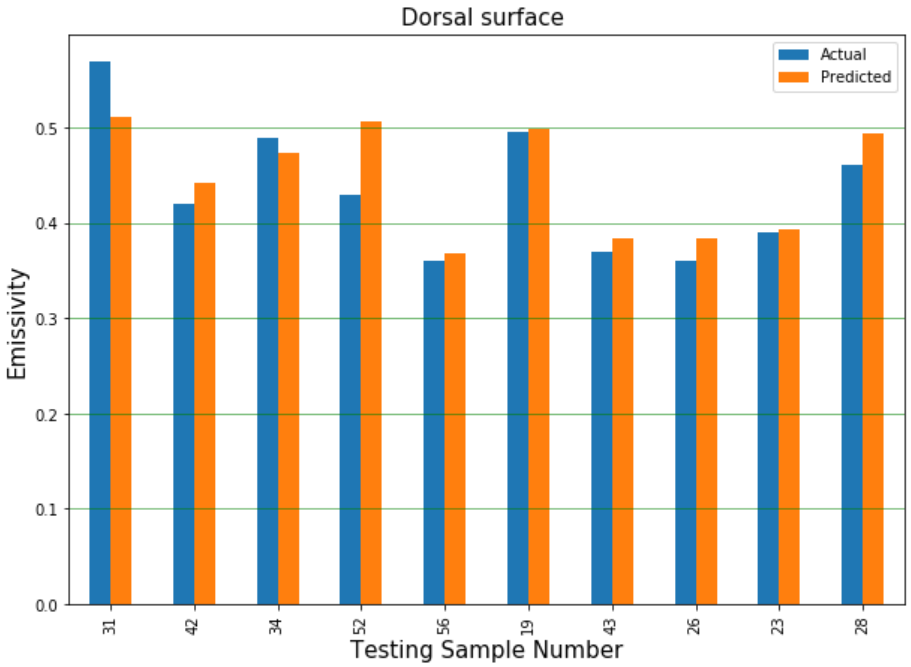


Figure 5.20: Predicted vs actual emissivity for the dorsal surface.

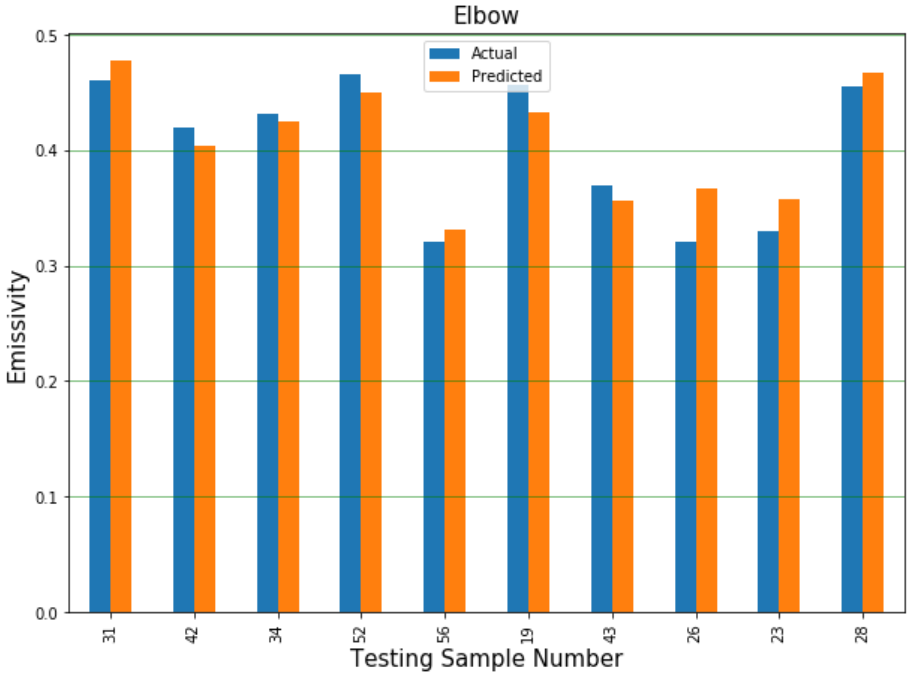


Figure 5.21: Predicted vs actual emissivity for the elbow.

Figures 5.12- 5.20 show 10 samples of the testing data, the blue bar shows the actual measured emissivity value while the orange bar shows the predicted value by the machine-learning algorithm. The minimum difference between the actual bar and the predicted bar means higher accuracy and the big difference between the actual and predicted emissivity value means that the machine-learning algorithm did not predict the emissivity probably.

5.3.2 Mean Absolute Error (MAE)

Mean absolute error was recorded for all skin locations using the different machine learning algorithms, Table 5.10 and Figure 5.22 show a comparison of the mean absolute error score for the 9 locations while using linear regression, KNN, SVM, and MLPNN.

Table 5.10: MAE for different locations using different machine learning algorithms.

	Linear regression	KNN	SVM	MLPNN
Palm of hand	0.0183	0.0338	0.0427	0.0185
Back of hand	0.025	0.0262	0.0378	0.0197
Fingers	0.0544	0.0683	0.0914	0.0486
Inner wrist middle	0.0182	0.0194	0.0529	0.021
Inner wrist side	0.0244	0.0294	0.0437	0.0219
Outer wrist	0.0254	0.0292	0.0453	0.0248
Volar side	0.0217	0.0214	0.0583	0.0212
Dorsal Surface	0.0301	0.0228	0.0704	0.0245
Elbow	0.0226	0.0227	0.0483	0.0184

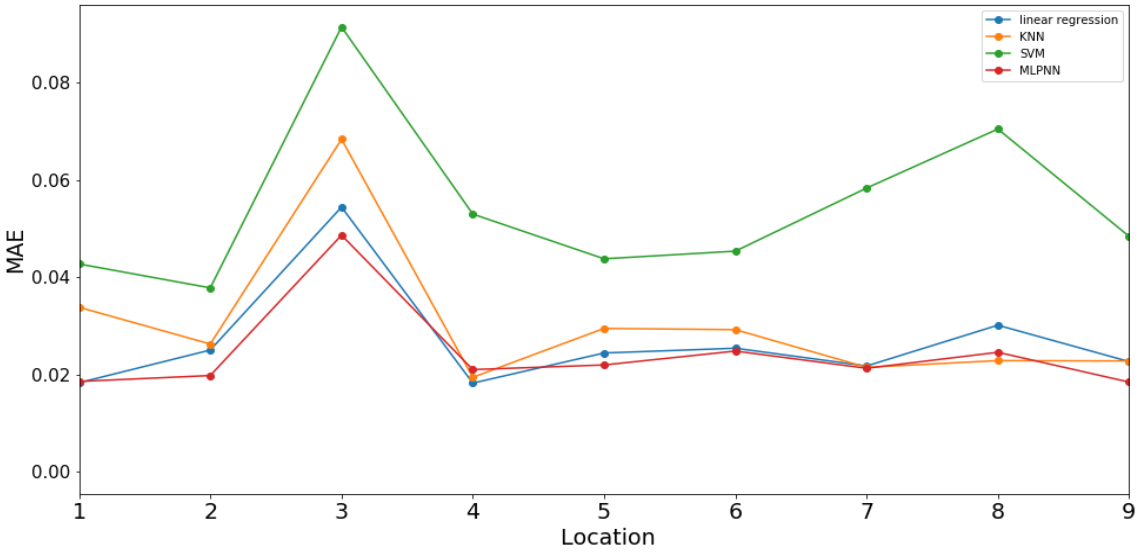


Figure 5.22: MAE for different algorithms used in predicting different locations' emissivity.

MAE records while using MLPNN are better than using the other algorithms for 6 locations which are the back of the hand, fingers, inner wrist side, outer wrist, volar side, and elbow. Using linear regression MAE is lower for 2 locations which are the palm of a hand and inner wrist middle with a small difference than MLPNN, the KNN achieved the lowest MAE while predicting dorsal surface also with little difference than MLPNN.

5.4 Conclusion

In this chapter, we present and discuss the results of the proposed methods for classification and regression using the different evaluation metrics. For classification, we used accuracy, confusion matrix, precision, recall, and F-score. For the regression, we calculate the accuracy and MAE. MLPNN for classification and regression achieved the best score almost for all conditions, for classification the highest accuracy was 91.6% while using MLPNN with 4 hidden layers, 200 hidden neurons, and 100 epochs,

the best confusion matrix was achieved while using MLPNN with 46% TP and 43% TN. For recall, precision and F measuring MLPNN achieved the best scores with 93.3%, 87.5%, and 90.3% sequentially.

For regression, the highest accuracy for almost all locations was achieved while using MLPNN, for the error metrics, also MLPNN achieved a better error score almost for all locations and the other locations' lowest error was achieved using linear regression or SVM with a small difference than MLPNN.

Chapter Six

Conclusion and Future Work

Machine Learning Techniques (MLT) in the fields of prediction and classifications of skin locations emissivity have excellent effectiveness in this research and have shown excellent results that can be applied in the real life.

Human skin emissivity has valuable information about the skin status, and this information could be used in different life applications such as the healthcare field, this research problem is how to classify the skin conditions and regions using a machine-learning algorithm through emissivity measurements, also how to predict the skin location emissivity measure based on other related locations emissivity.

We made dataset analyses using python libraries and find that there is a sufficient correlation between some human skin locations with each other, also for the same volunteer, there is noticed a pattern for skin emissivity. We also plot the box plot for the male samples and the female samples separately and for each ethnicity alone, we find for most locations the female emissivity range is smaller than the male range, and the white emissivity is higher than other ethnicities.

Many algorithms have been used in the prediction phase of skin locations emissivity including linear regression, KNN, SVM, and MLPNN, the study proved that the MLPNN has the lowest mean absolute error of 0.0184 while predicting Elbow emissivity and the highest accuracy while predicting palm of hand emissivity with 0.913% score.

In the process of classifying the skin status for wet or dry, many algorithms were used to give accurate results, including KNN, Random Forest, Decision Tree, and MLPNN and it has been noticed that the MLPNNs were the best algorithm with an accuracy of about 91.6%, Recall 93.3%, precision 87.5% and F-measure 90.3%.

A hybrid system that integrates Machine Learning Techniques like thinking of integrating Naive Bayes (NB) with Multilayer Perceptron Neural networks will be created to have an NB-MLP hybrid system. Another hybrid system that can contain a particle Swarm Optimization algorithm with Multilayer Perceptron Neural networks and a Genetic algorithm can be made. Moreover, another hybrid system that consists of a Genetic algorithm with Particle Swarm Optimization namely GA-PSO can be made so that it shows better results and accuracy than using single techniques. Also, a more comprehensive dataset for all human body emissivity measures could be used to determine exactly the high correlated human body locations, this will help in focusing on specific locations to predict the skin emissivity.

We recommended creating a comprehensive emissivity dataset for all human skin and to generate the normal ranges of emissivity for each location. Also creating a full diagnosing system consists of radiometer hardware and the computer software supported with machine learning capabilities to immediately classify the human skin for wet or dry, because of that, the medical staff can determine if the patient skin is normal or not within minutes.

References

[1] S. A. U, S. Badiger and N. k. Siddappa, "Pattern of common skin conditions among school children in an urban," *International Journal Of Community Medicine And Public Health,* vol. 4, p. 2901, 2017.

[2] F. Xu and T. Lu, "Skin Structure and Skin Blood Flow," in *Introduction to Skin Biothermomechanics and Thermal Pain*, pringer Berlin Heidelberg, 2011, pp. 7-19.

[3] H. Yousef, M. Alhajj and S. Sharma, Anatomy, Skin (Integument), Epidermis, StatPearls Publishing, Treasure Island (FL), 2020.

[4] R. Jainesh, W. Vishal, S. Aniruddh and B. Praseniit, "Diagnosis of skin diseases using Convolutional Neural Networks," in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, 2018.

[5] K. M. Perreira and E. E. Telles, "The color of health: skin color, ethnoracial classification, and discrimination in the health of Latin Americans," *Social Science & Medicine,* vol. 116, pp. 241-250, 2014.

[6] Rathod, Jainesh; Wazhmode, Vishal; Sodha, Aniruddh; Bhavathankar, Praseniit, "Diagnosis of skin diseases using Convolutional Neural Networks," in *IEEE*, Coimbatore, 2018.

[7] D. L. Cummins, J. M. Cummins, H. Pantle, M. A. Silverman, A. L. Leonard and A. Chanmugam, Cutaneous Malignant Melanoma, chicago : Mayo Clinic Proceedings, 2006, pp. 500-507.

[8] C. Griffiths, J. Barker, T. Bleiker, R. Chalmers and D. Creamer, Rook's textbook of dermatology, Chichester, West Sussex ; Hoboken, NJ, 2016.

[9] A. Y. Owda, N.-D. Rezgui and N. Salmon, "Signatures of human skin in the millimetre wave band (80-100) GHz," in *Millimetre Wave and Terahertz Sensors and Technology X*, Warsaw, SPIE, 2017, pp. 16-22.

[10] S. Dua, U. R. Acharya and P. Dua, Machine Learning in Healthcare Informatics, Springer, Berlin, Heidelberg, 2014.

[11] S. Perveen, M. Shahbaz, A. Guergachi and K. Keshavjee, "Performance Analysis

of Data Mining Classification Techniques to Predict Diabetes," *Procedia Computer Science,* pp. 115-121, 2016.

[12] R. Venkatesh, C. Balasubramanian and M. Kaliappan, "Development of Big Data Predictive Analytics Model for Disease Prediction Using Machine Learning Technique," *J. Med. Syst.,* vol. 43, p. 1–8, 2019.

[13] F. F. Ting, Y. J. Tan and K. S. Sim, "Convolutional neural network improvement for breast cancer classification," *Expert Systems with Applications,* vol. 120, pp. 103-115, 2019.

[14] A. Mir and S. N. Dhage, "Diabetes Disease Prediction Using Machine Learning on Big Data of Healthcare," in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, IEEE, 2018, pp. 1-6.

[15] D. Shen, G. Wu and H.-I. Suk, "Deep Learning in Medical Image Analysis," *Annual Review of Biomedical Engineering,* vol. 19, pp. 221-248, 2017.

[16] A. Fourcade and R. Khonsari, "Deep learning in medical image analysis: A third eye for doctors," *Journal of Stomatology, Oral and Maxillofacial Surgery,* Vols. 279-288, no. 4, pp. 279-288, 2019.

[17] Y. C. Zhang and A. C. Kagen, "Machine Learning Interface for Medical Image Analysis," *Journal of Digital Imaging,* vol. 30, 2017.

[18] S. Mohan, C. Thirumalai and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," *IEEE Access,* vol. 7, no. 4, pp. 81542-81554, 2019.

[19] S. Uddin, A. Khan, M. E. Hossain and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," *BMC Medical Informatics and Decision Making,* vol. 19, 2019.

[20] A. Abdelaziz, M. Elhoseny, A. S. Salama and A. Riad, "A machine learning model for improving healthcare services on cloud computing environment," *Measurement,* vol. 119, pp. 117-128, 2018.

[21] A. J. Hung, J. Chen and I. S. Gill, "Automated Performance Metrics and Machine Learning Algorithms to Measure Surgeon Performance and Anticipate Clinical Outcomes in Robotic Surgery," *JAMA Surgery,* vol. 153, no. 14, pp. 770-771,

2018.

[22] B. Zhao, R. S. Waterman, R. D. Urman and R. A. Gabriel, " A Machine Learning Approach to Predicting Case Duration for Robot-Assisted Surgery," *Journal of Medical Systems,* vol. 43, no. 2, 2019.

[23] P. Forbrig and A.-N. Bundea, "Modelling the Collaboration of a Patient and an Assisting Humanoid Robot During Training Tasks," in *Human-Computer Interaction. Multimodal and Natural Interaction*, Cham, Springer International Publishing, 2020, pp. 592--602.

[24] A. Y. Owda, N. Salmon, N. D. Rezgui and S. Shylo, "Millimetre wave radiometers for medical diagnostics of human skin," in *IEEE*, Glasgow, 2017.

[25] A. Y. Owda, N. Salmon and R. nacer-Ddine, "Electromagnetic Signatures of Human Skin in the Millimeter Wave Band 80-100 GHz," *Progress In Electromagnetics Research,* vol. 80, pp. 79-99, 2018.

[26] A. Y. Owda, N. Salmon, S. Shylo and M. Owda, "Assessment of Bandaged Burn Wounds Using Porcine Skin and Millimetric Radiometry.," *Sensors (Basel),* vol. 19, 2019.

[27] H. Wu, S. Yang, Z. Huang and J. He, "Type 2 diabetes mellitus prediction model based on data mining," *Informatics in Medicine Unlocked,* vol. 10, pp. 100-107, 2018.

[28] X.-H. Meng, Y.-X. Huang, D.-P. Rao and Q. Zhang, "Comparison of three data mining models for predicting diabetes or prediabetes by risk factors," *The Kaohsiung journal of medical sciences,* vol. 29, no. 5, p. 93–99, 2013.

[29] J. Velasco, J. Rojas, J. P. M. Ramos and H. M. Muaña, "Health Evaluation Device Using Tongue Analysis Based on Sequential Image Analysis," *International Journal of Advanced Trends in Computer Science and Engineering,* vol. 8, pp. 451-457, 2019.

[30] D. Sisodia and D. S. Sisodia, "Prediction of Diabetes using Classification Algorithms," in *International Conference on Computational Intelligence and Data Science*, Gurugram, 2018.

[31] H. Arghandabi and P. Shams, "A Comparative Study of Machine Learning

Algorithms for the Prediction of Heart Disease," *International Journal for Research in Applied Science and Engineering Technology,* vol. 8, no. 4, pp. 677-683, 2020.

[32] A. Nawal, "A Method Of Skin Disease Detection Using Image Processing And Machine Learning," in *16th Learning and Technology Conference 2019Artificial Intelligence and Machine Learning: Embedding the Intelligence*, Jeddah, 2019.

[33] V. B. Kumar, S. S. Kumar and V. Saboo, "Dermatological disease detection using image processing and machine learning," in *Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR)*, Lodz, 2016.

[34] D. Gavrilov, A. V. Melerzanov, N. Shchelkunov and e. .. Zakirov, "Use of Neural Network-Based Deep Learning Techniques for the Diagnostics of Skin Diseases," *Biomedical Engineering*, vol. 52, no. 5, 2019.

[35] X. Dai, I. Spasic, B. Meyer and S. Chapman, "Machine Learning on Mobile: An On-device Inference App for Skin Cancer Detection," in *2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC,* IEEE, 2019, pp. 301-305.

[36] T. Philipp, C. Rosendahl and H. Kittler, "The HAM10000," *A Large Collection of Multi-Source Dermatoscopic Images,* 2018.

[37] O. I. Ali and K. Murat, "Skin Lesion Classification using Machine Learning Algorithms," in *International Journal of Intelligent Systems and Applications in Engineering*, Texas, 2017.

[38] J. Velasco, C. Pascion, J. W. G. Alberio and J. Apuang, "A Smartphone-Based Skin Disease Classification Using MobileNet CNN," *International Journal of Advanced Trends in Computer Science and Engineering,* vol. 8, pp. 2632-2637, 2019.

[39] S. Li, M. Ardabilian and A. Zine, "Quantitative Analysis of Skin using Diffuse Reflectance for Non-invasive," in *16th International Conference on Computer Vision Theory and Applications,* 2021.

[40] R. Gunaratne, I. Monteath, J. Goncalves and R. Sheh, "Machine learning classification of human joint tissue from diffuse reflectance spectroscopy data," *Biomed. Opt. Express,* vol. 10, no. 8, pp. 3889-3898, 2019.

[41] K. Karadağ, M. e. Tenekeci, R. Taşaltın and A. Bilgili, "Detection Of Pepper Fusarium Disease Using Machine Learning Algorithms Based On Spectral Reflectance," *Sustainable Computing: Informatics and System,* vol. 28, no. 8, 2019.

[42] G. Asuero, Agustin, A. Sayago and G. González, "The Correlation Coefficient: An Overview," *Critical Reviews in Analytical Chemistry - CRIT REV ANAL CHEM*, vol. 36, pp. 41-59, 2006.

[43] C. C. Aggarwal, data mining "the textbook", new york: IBM T.J. Watson Research Center, 2015.

[44] A. Jain, K. N, akumar and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognition,* vol. 38, no. 12, pp. 2270-2285, 2005.

[45] A. Burkov, The Hundred-Page Machine Learning Book, Quebec : Andriy Burkov, 2019.

[46] G. Gan and M. K.-P. Ng, "k-means clustering with outlier removal," *Pattern Recognition Letters,* vol. 90, no. 3, pp. 8-14, 2017.

[47] K. S. Kannan, K. Manoj and S. Arumugam, "Labeling Methods for Identifying Outliers," *International Journal of Statistics and Systems,* vol. 10, pp. 231-238, 2015.

[48] E. Fix and J. Hodges, Discriminatory Analysis - Nonparametric Discrimination: Consistency Properties, CALIFORNIA: International Statistical Review, 1989.

[49] S. Theodoridis and K. Koutroumbas, Pattern Recognition (Fourth Edition), Boston: Academic Press, 2009.

[50] Y.-Y. Song and Y. Lu, " Decision tree methods: applications for classification and prediction," *Shanghai Arch Psychiatry*, vol. 27, 2015.

[51] S. Misra, H. Li and J. He, "Chapter 9 - Noninvasive fracture characterization based on the classification of sonic wave travel times," in *Machine Learning for Subsurface Characterization*, Gulf Professional Publishing, 2020, pp. 243-287.

[52] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity. Bulletin of mathematical biophysics," *Journal of Symbolic Logic*, vol. 5, p. 115–133, 1943.

[53] M. Gardner and S. Dorling, "Artificial neural networks (the multilayer

perceptron)—a review of applications in the atmospheric sciences," *Atmospheric Environment,* vol. 32, no. 14, pp. 2627-2636, 1998.

[54] C. Kadılar and H. a. C¸ıngı, "Improvement in Variance Estimation in Simple Random Sampling," *Hacettepe Journal of Mathematics and Statistics,* vol. 35, no. 11, p. 111 – 115, 2006.

[55] F. A. Gers, J. Schmidhuber and F. Cummins, "Learning to Forget: Continual Prediction with LSTM," in *1999 Ninth International Conference on Artificial Neural Networks ICANN 99. (Conf. Publ. No. 470,* Edinburgh, 2000.

[56] F. Khademi, M. Akbari, S. M. Jamal and M. Nikoo, " Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete," *Frontiers of Structural and Civil Engineering,* vol. 11, 2017.

[57] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," *Association for Computing Machinery,* vol. 2, no. 3, 2011.

[58] P.-Y. Hao, "New support vector algorithms with parametric insensitive/margin model," *Neural Networks,* vol. 23, no. 1, pp. 60-73, 2010.

[59] K. Yan and C. Shi, "Prediction of elastic modulus of normal and high strength concrete by support vector machine," *Construction and Building Materials,* pp. 1479-1485, 2010.

[60] G. Nguyen, A. Bouzerdoum and S. Phung, "Learning Pattern Classification Tasks with Imbalanced Data Sets," *Pattern Recognition,* 2009.

[61] Al-Khowarizmi, R. Syah, M. K. M. Nasution and M. Elveny, "Sensitivity of MAPE using detection rate for big data," *International Journal of Electrical and Computer Engineering,* vol. 11, pp. 2696-2703, 2021.

[62] B. Hu, M. Palta and J. Shao, "Properties of R-2 statistics for logistic regression," *Statistics in medicine,* vol. 25, no. 2, pp. 1383-95, 2006.

Appendix

This section contains the abstract of “Skin location emissivity prediction using emissivity records for other locations.” Paper, which will be submitted for publication purposes.

Abstract

The motive of this study is to design a regression model that can predict the emissivity measure for a specific skin location by using measures from other different locations collected from the same human. Predicting measurement location emissivity will help to determine the normal emissivity value for the needed location; the implication of having this is the non-invasive diagnosis of diseased skin as the location of the skin may be infected with one of the skin diseases or thermal burns or may be covered with a hidden object, which will affect the emissivity. Regression models were used to predict the emissivity of the skin from nine locations using Linear Regression, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Multiple-Layer Perceptron Neural Network (MLPNN). Experimental results show that MLPNN gives the highest accuracy for most measurements locations, and the palm of hand achieved the highest accuracy of 91.3%. In addition, the lowest mean squared error of 0.00055 for the fingers location using MLPNN.

الملخص

في هذه الرسالة، تم استخدام تقنيات التعلم الآلي للتنبؤ بانبعاثيه جلد الإنسان بالإضافة إلى تصنيف حالة الجلد بناءَ على انبعاثيته إلى جلد جاف أو رطب، توقع انبعاثية الجلد تمت من خلال اختيار موقع معين للجلد واستخدام مواقع جلد أخرى لنفس الشخص لتوقع انبعاثية هذا الموقع، إن توقع انبعاثية موقع معين سيساعد في معرفة قيمة الإنبعاثية الطبيعية لهذا الموقع. سيساعد هذا في التشخيص غير الجراحي لأمراض الجلد حيث أن هذا الموقع قد يكون مصاباً بأحد الأمراض الجلدية أو الحروق أو قد يكون مغطىَ بجسم غريب والذي سيؤثر على الأنبعاثية. تم بناء نموذج للتنبؤ بإنبعاثية جلد الإنسان باستخدام أربعة تقنيات تعلم آلي وهي خوارزمية الانحدار الخطي، أقرب الجيران، آلة المتجهات الداعمة والشبكات العصبية متعددة الطبقات، أظهرت النتائج بأن الشبكات العصبية متعددة الطبقات حققت أفضل النتائج لغالبية أماكن القياس ولراحة اليد حققت كفاءة بمقدار 91.3% بالإضافة إلى أقل نسبة خطأ عند توقع انبعاثية كوع يد الإنسان بمقدار 0.0184 .

تم إنشاء مجموعة البيانات المستخدمة في الانحدار من دراسات منشورة مسبقاً عن انبعاثية الجلد وتحتوي على 540 عيّنة من 60 متطوّع لتسعة أماكن مختلفة لكلا الجنسين. تحتوي هذه البيانات على خمسة سمات مختلفة وهي الموقع، العرق، العمر، الإنبعاثية والجنس . في الجزء الآخر من العمل وهو نموذج التصنيف، تم استخدام أربعة خوارزميات لتصنيف حالة الجلد بناءً على انبعاثيته لجاف أو رطب وهذه الخوارزميات هي أقرب الجيران، الغابة العشوائية، شجرة القرار والشبكات العصبية متعددة الطبقات. تم إنشاء مجموعة البيانات المستخدمة للتصنيف من دراسات منشورة مسبقاً من خلال استخدام جهاز قياس الإشعاع ل 120 متطوع ل تسعة أماكن مختلفة في حالتين وهي الجلد الرطب والجلد الطبيعي وتحتوي مجموعة البيانات هذه على ستة سمات وهي الموقع، العرق، العمر، حالة الجلد، الإنبعاثية والجنس. وقد أظهرت النتائج أن الشبكات العصبية متعددة الطبقات حققت أفضل نسبة كفاءة بمقدار ,93.3% و recall بمقدار 91.6% و f measuring بمقدار 90.3 و precision بمقدار % 87.5 سيتم مناقشة النتائج بالتفصيل في القسم الخامس.