

# Common Skin Disease Diagnosis and Prediction

## A Project Synopsis

*Submitted by*

Group ID:15

Team Members:

Ruchika Kiran Gaikwad                      Roll No.2243022

Samruddhi Mahesh Gaikwad              Roll No.2243020

Atharva Pravin Gurav                      Roll No.2243059

Deep Rajesh Khadke                      Roll No.2243025

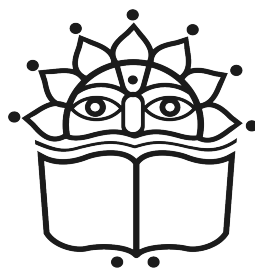
*In partial fulfilment for the requirement of the degree*

**BE( Information Technology )**

Under the guidance of

**Mr. L.J. Deokate**

Assistant Professor in IT Deptt



Department of Information Technology

Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering  
and Technology Bhigawan Road, Baramati

August 2022

**Project Title :** Common Skin Disease Diagnosis and Prediction

**Area of Project** Deep Learning and Image Processing

## Abstract

The wide adaptation of computer technology in the healthcare sector has emerged in the accumulation of electronic data. Skin is an extraordinary human structure. It frequently suffered from many known and unknown disease. Therefore, diagnosis of human skin diseases is the most uncertain and complicated branch of science. However, machine learning algorithms have showcased significant capability in overshadowing standard systems for skin disease diagnosis and supporting medical experts in the early detection of high-risk skin diseases.

The goal is to recognize trends across different types of ML models in skin disease detection by examining performance metrics. We used Deep Learning to train our model, Deep Learning is a part of machine learning in which unlike machine learning it uses large data sets and therefore the number of classifiers is substantially reduced. The machine learns itself and divides the data provided into prediction levels and provides accurate results in a very short time, thus supporting and promoting the development of dermatology. The algorithm we used is Convolutional Neural Network (CNN) because it is one of the most preferred algorithms for image classification.

**Keywords:** Skin Disease, Machine Learning, Deep Learning, Convolutional Neural Network(CNN)

# Problem Statement

The patient provides an image of the infected area of the skin as an input to the system. Image processing techniques and machine learning techniques are performed on this image and the detected disease is displayed at the output.

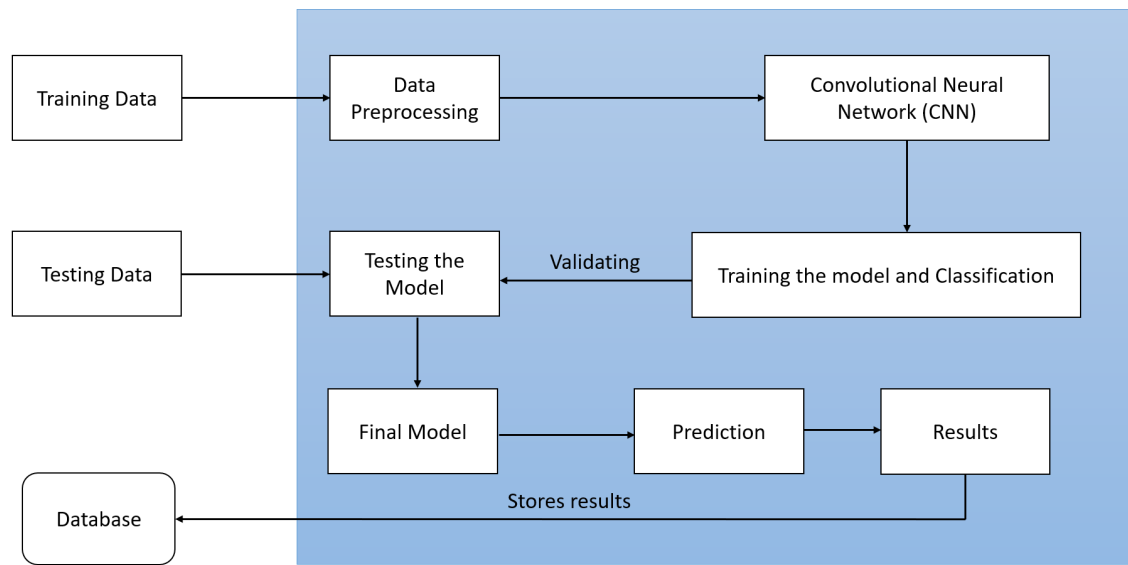


Figure 1: System Architecture

# Introduction

Dermatology remains the most uncertain and complex scientific field, because it is complicated in the procedures of diagnosing diseases of hair, skin, and nails. Differences in these diseases can be observed due to many environmental and geographical variations in factors. Usually, it's caused by factors such as different cells of the organism, different diets and internal and external factors such as the hierarchical genetic group of cells, hormones and immune system conditions. These factors may act together or in sequence in skin disease. There are chronic and incurable diseases, such as eczema and psoriasis, and malignant diseases such as malignant melanoma. Recent researchers have found the availability of cures for these diseases if they are detected in the early stages.

Deep learning is a part of the broader family of machine learning wherein the learning can be supervised, unsupervised or semi supervised. Deep learning unlike machine learning uses a large dataset for the learning process and the number of classifiers used gets reduced substantially. The training time for the deep learning algorithm increases because of the usage of the very large dataset. Deep learning algorithm chooses its own features unlike the machine learning making the prediction process easier for the end user as it does not use much of pre-processing.

A CNN is a types of artificial neural network used in major recognition. CNNs are a category of neural networks that have proven very effective in areas such as image recognition and classification. Convolutional networks have been successful in identifying faces, objects and traffic signs apart from powering vision in robots and self-driving cars. CNNs are important tools for most machine learning practitioners today. CNN is a deep learning neural network designed for processing structured arrays of data such as images.

Skin diseases occur in almost all age groups of people. The rate of skin diseases has increased due to lifestyle and changing environments. Today, skin diseases are becoming a more common problem in human life. Most of these diseases are dangerous and harmful, especially if not treated in the initial stage. People do not take skin diseases seriously.

The most common diseases in the dataset mainly include:

- MelanocyticNevi

- Melanoma
- Benign keratosis-like lesions
- Basal cell carcinoma
- ActinicKeratoses
- Vascular lesions
- Dermatofibroma, etc

Despite being common its diagnosis is extremely difficult because of the complexities of skin tone, color, and presence of hair. We used a convolutional neural network (CNN), SVM, Naive Bayes, and Statistical Analysis to detect these skin diseases. A convolutional neural network is a category of deep neural networks where the machine learns on its own and divides the provided data into prediction levels and in a very short time it gives accurate results.

# System Requirements

## **Software Requirements:**

1. Operating System: Windows 10/11, 64bit
2. Technology : Python
3. IDE: Jupyter Notebook/PyCharm
4. Database: MySQL

## **Hardware Requirements:**

1. Hardware: Intel Core i3/i5 processor
2. RAM: 8GB

# Project Design

To develop any ML-AI based system, be it this system following steps are to be followed.

1. Data Gathering -The proposed system has been assessed on dermatoscopic images which is collected from publicly available dataset based on Skin- Cancer-MNIST (Modified National Institute of Standards and Technology Database)-HAM10000. The number of options is endless. To save time and effort one can use publicly available data.
2. Data Pre-processing- Dirty data can cause confusion and results in unreliable and poor output. Hence first step in Data Pre-processing is Data Cleaning. Cleaning of data is done by filling in missing values, smoothing noisy data by identifying and/or removing outliers, and removing inconsistencies. In pre-processing the input image data to convert it into meaningful floating-point tensors for feeding into Convolutional Neural Networks.
3. Data Transformation- Data Transformation involves converting data from one format into another. It involves transforming actual values from one representation to the target representation.
4. Model Building.- We used a convolutional neural network (CNN). A convolutional neural network (CNN) is a category of deep neural networks where the machine learns itself and divides the provided data into prediction levels and provides accurate results in a very short time. A convolutional neural network (CNN) is a deep learning algorithm that consists of a combination of convolutional and pooling layers in sequence and then followed by fully connected layers at the end as a multilayer neural network. CNN excels among all alternative algorithms in image classification. Sparse connectivity, shared weights, and pooling functions are critical features to get the best features. Also, the use of graphics processing units (GPUs) reduced the training time of deep learning methods. huge databases of labeled data and pre-trained networks are now publicly available.

5. Model Evaluation- More the accuracy, better is the model. Every model is evaluated based on the accuracy achieved and the loss obtained. There are two accuracies involved: Validation accuracy And Test accuracy. Before this Validation set is different from Train set i.e. Validation set is independent from the Train set, Validation set is used for selecting parameters. Just for an instance if your model has 90 percent train accuracy and 89 percent validation accuracy then your model is expected to have 89 percent accuracy on new data.

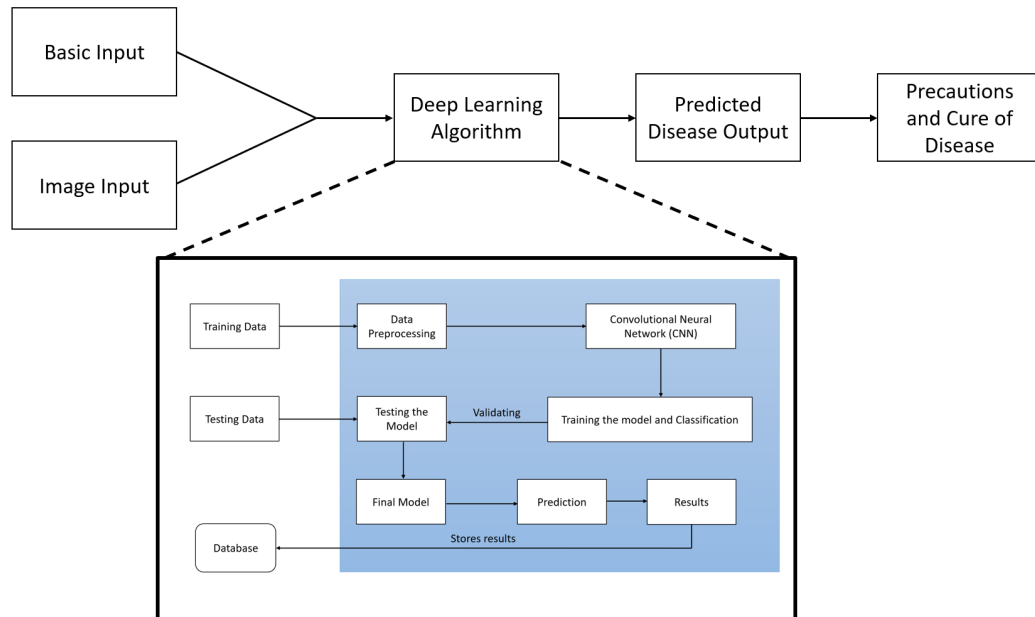


Figure 2: System Design



# Project Plan 1.0

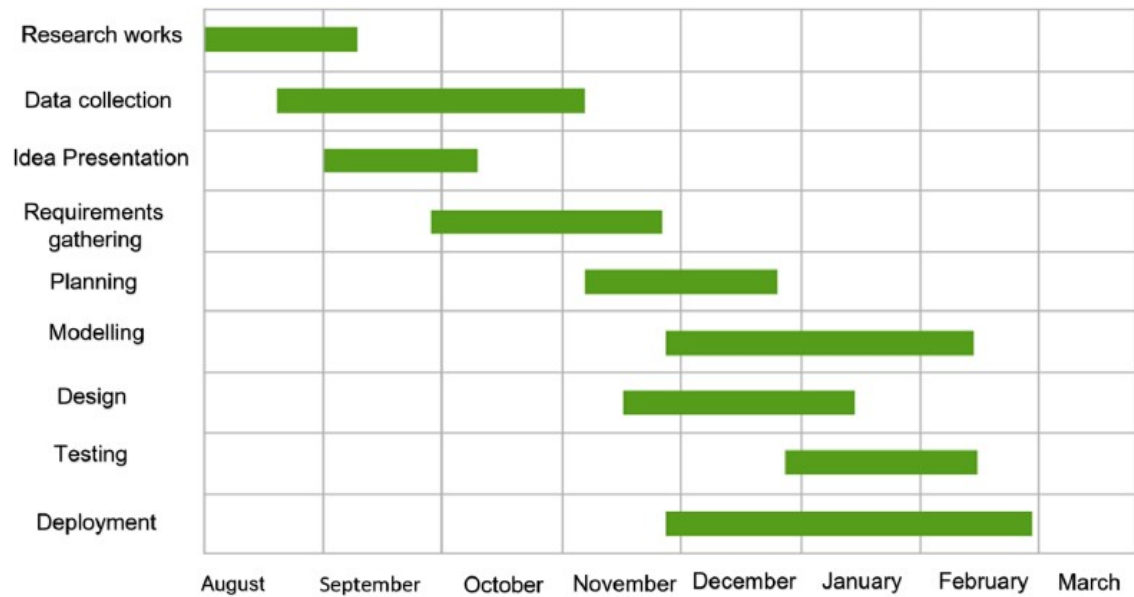


Figure 3: Gantt Chart

# References

- [1] Kritika Rao ,Pooja Yelkar, Omkar Pise and Dr. Swapna Borde,and , "Skin Disease Detection using Machine Learnings", 2021.
- [2] Ahmed A. Elngar,Rishabh Kumar,Amber Hayat,Prathamesh Churi,"Intelligent System for Skin Disease Prediction using Machine Learning" ,August 2021

Roll No	Name	Sign
---------	------	------

2243022	Ruchika Kiran Gaikwad	
---------	-----------------------	--

2243020	Samruddhi Mahesh Gaikwad	
---------	--------------------------	--

2243059	Atharva Pravin Gurav	
---------	----------------------	--

2243025	Deep Rajesh Khadke	
---------	--------------------	--

Project Guide  
(Mr. L.J. Deokate)

Project Coordinator  
( Mr. P. M. Patil )

Head of Department  
( Dr. S. A. Takale )

Principal  
( Dr. R. S. Bichkar)

