Capstone Project Weekly Progress Report

Semester	Fall 2022
Course Code	AML 2404
Section	Section 2
Project Title	Skin Diseases Classification using Deep Learning
Group Name	G
Student names/Student IDs	Tomson George (C0857730) Praveen Mahaulpatha (C0860583) Thulana Abeywardana (C0861333) Jaskaran Singh Moti (C0860026)
Reporting Week	Week 3 (25 September 2022 to 30 September 2022)
Faculty Supervisor	William Pourmajidi

1. **Tasks Outlined in Previous Weekly Progress Report**

    - Understanding dataset and pre-processing.
    - Build, Train and Validate Initial Neural Network Model
    - Finding the feasibility of deploying the code in either Heroku or AWS.
    - Finding natural remedies for common skin diseases and documenting them.

2. **Progress Made in Reporting Week**

    Task 1: Praveen: Model Implementation, Training and Validation

    The images corresponding to 2 classes were separated into two folders for training and validation purposes. Images were loaded and preprocessed up to a certain degree for training efficiency. Later, Data augmentation was carried out on the training data, so the model learns to generalize its learning. Thirdly the model was created, compiled, and fit to finally plot Accuracy and Losses against training and validation datasets. Below shows the initial structure of the neural network:

```
_____
Layer (type)                Output Shape          Param #
================================================================
conv2d (Conv2D)             (None, 224, 224, 32)  896

max_pooling2d (MaxPooling2D (None, 112, 112, 32)  0
)

conv2d_1 (Conv2D)           (None, 112, 112, 32)  9248

max_pooling2d_1 (MaxPooling (None, 56, 56, 32)    0
2D)

conv2d_2 (Conv2D)           (None, 56, 56, 64)    18496

max_pooling2d_2 (MaxPooling (None, 28, 28, 64)    0
2D)

dropout (Dropout)           (None, 28, 28, 64)    0

flatten (Flatten)           (None, 50176)         0

dense (Dense)               (None, 128)           6422656

dense_1 (Dense)             (None, 2)             258

================================================================
Total params: 6,451,554
Trainable params: 6,451,554
Non-trainable params: 0
```

Result 1:

Learning Rate	Loss Function	Epochs
0.01	Sparse Categorical Cross Entropy	20
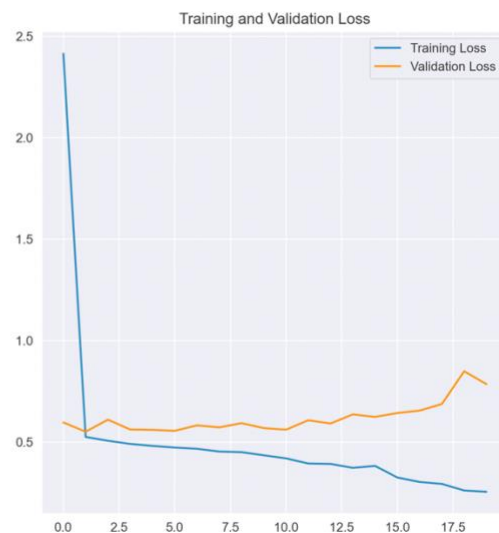
Result 2:

Learning Rate	Loss Function	Epochs
0.01	Sparse Categorical Cross Entropy	10
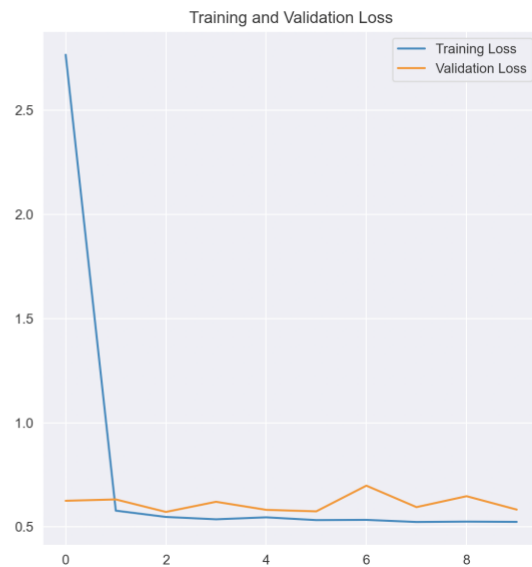
Result 3:

Learning Rate	Loss Function	Epochs
0.01	KLDivergence	10
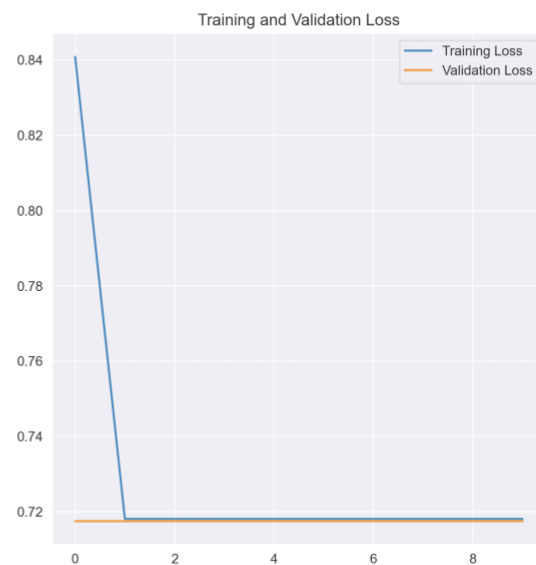
Task 2: Thulana: Model Implementation, Training and Validation

A sample model was built based on the dataset; the dataset was categorized into 4 main types of diseases and each has approximately 1500 to 1700 images.

Original images are scaled to 0 -255 and those were scaled down to 0 – 1 in the data preprocessing stage.

Dataset was divided into Training, Testing, and Validation dataset as follows:

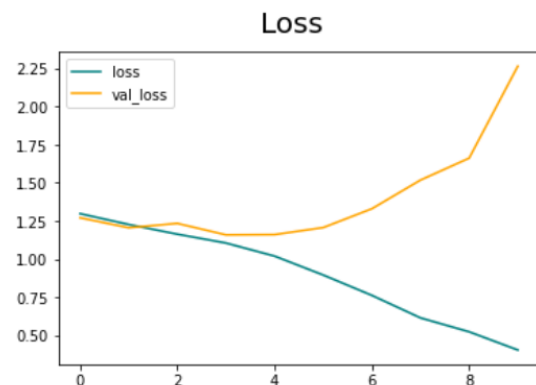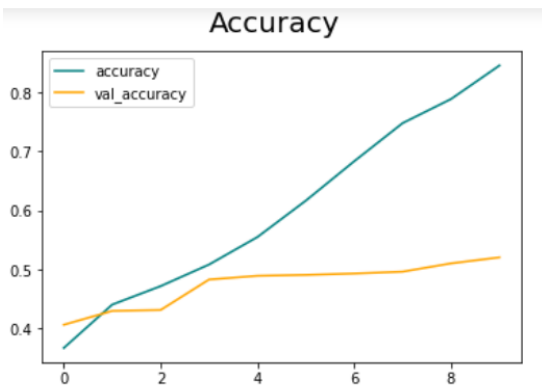- Training – 70%
- Validation – 20%
- Testing – 10%

Further, training, validation and testing data were selected as following sequence from the dataset:

Training data

	Training
	Validation
	Testing

Result 1:

Initially, the model has trained up to 10 epochs and the results are as bellows:

As the validation accuracy and validation error was not improving to the expected level, the model was simplified by dropping some convolution layers.
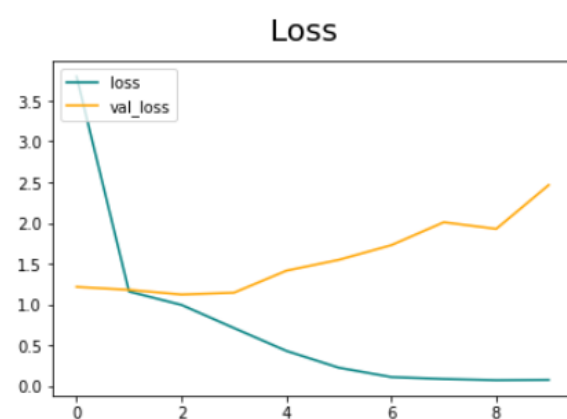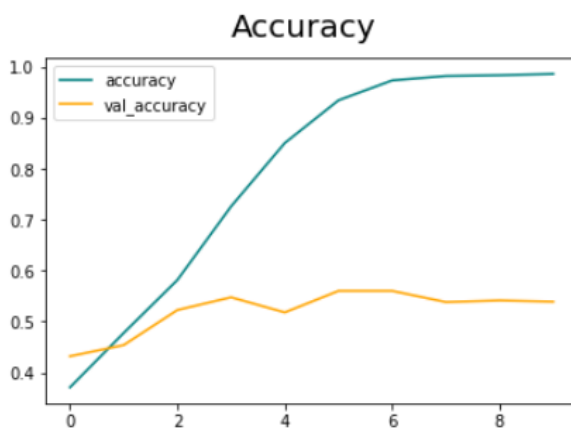
```
Model: "sequential_13"

Layer (type)                Output Shape              Param #
=================================================================
conv2d_27 (Conv2D)          (None, 254, 254, 32)      896

max_pooling2d_22 (MaxPoolin (None, 127, 127, 32)      0
g2D)

flatten_7 (Flatten)         (None, 516128)            0

dense_10 (Dense)            (None, 127)               65548383

dense_11 (Dense)            (None, 4)                 512

=================================================================
Total params: 65,549,791
Trainable params: 65,549,791
Non-trainable params: 0
```
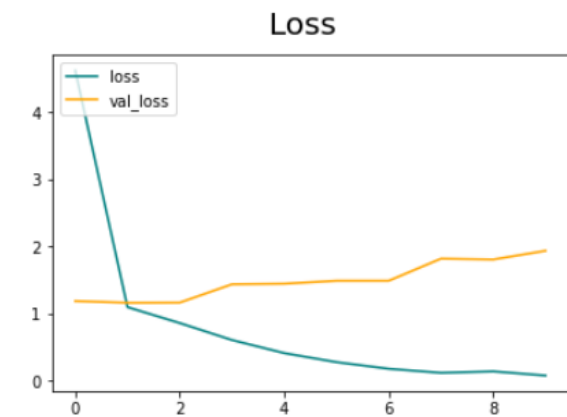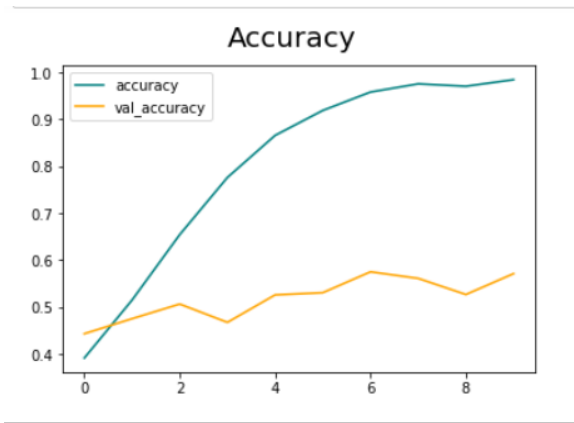

Result 3:

Then another Dense layer was added.

Task 3: Tomson: Research on cloud platforms

We found out that it's possible to host the model in an EC2 instance which is free of cost. Although AWS provides an easy solution called Sagemaker, we found out it's better to go with manually setting up a server and make it live for the first phase.

Task 4: Jaskaran: Shortlisting recommendations

Worked to find home remedies for shortlisted skin problems which will help to train the program. Along with that for the diseases which cannot be treated at home suggestions to see a dermatologist will be given by the app what is the closest appointment date available and if the patient wants to confirm it or not. Also Worked on some antiseptics and medicines which can be used for curing diseases without any prescriptions. Furthermore, found some precautions according to different skin diseases.

3. **Difficulties Encountered in Reporting Week**

   a. 1st model shows a low increase rate of validation accuracy.
      Solution: Tweaking parameters of the model / Preprocessing dataset further

   b. 2nd model shows an increase in validation loss.
      Solution: Tweaking parameters of the model / Preprocessing dataset further

   c. While experimenting with AWS, we faced difficulties while connecting to the instances using putty and understanding various terminologies like security group, VPC etc.

   d. The symptoms of some diseases are similar but the disease is different. And to suggest any remedies specifically is going to be challenging.

4. **Tasks to Be Completed in Next Week**

Tasks	Responsible
Creating server. Understand how to implement the model in cloud and make it live.	Tomson
Optimizing Model 1 for Better Validation Accuracy / Validation Loss	Praveen
Optimizing Model 2 for Better Validation Accuracy / Validation Loss	Thulana
Plotting the validation results using Plotly	Jaskaran