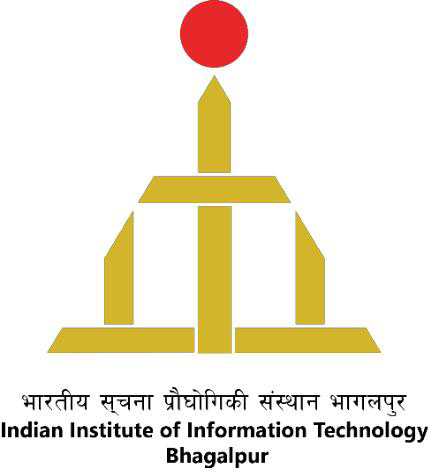
INDIAN INSTITUTE OF INFORMATION

TECHNOLOGY BHAGALPUR



**PROJECT ON IMAGE PROCESSING**

**TITLE – Fruit quality detection using deep learning**

# Aim-

Given a dataset containing rotten and fresh apples we have train a model which will classify a given sample fruit image into two categories “fresh” or “rotten”

## 

## Dataset –

<https://www.kaggle.com/sriramr/fruits-fresh-and-rotten-for-classification>

An example of rotten fruit image An example of fresh fruit image

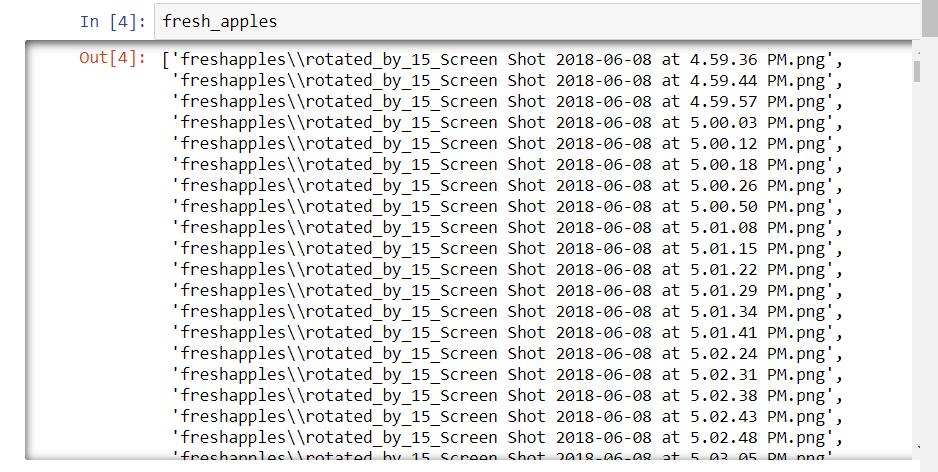
## Approach

1. Our first step is to make a working dataset from folders.
2. Then converting to Gray scale and resize them to specific shape.
3. Then we made a training and test data set for training of the model
4. After making training and test data set we trained a classic SGD linear classifier to get a base accuracy.
5. After tweaking some parameters and more classical ML models we made a decision to move to the deep learning approach .
6. In deep learning approach we used only sequential api models with some tweaks.
7. Some pre-processing is done to get better results.

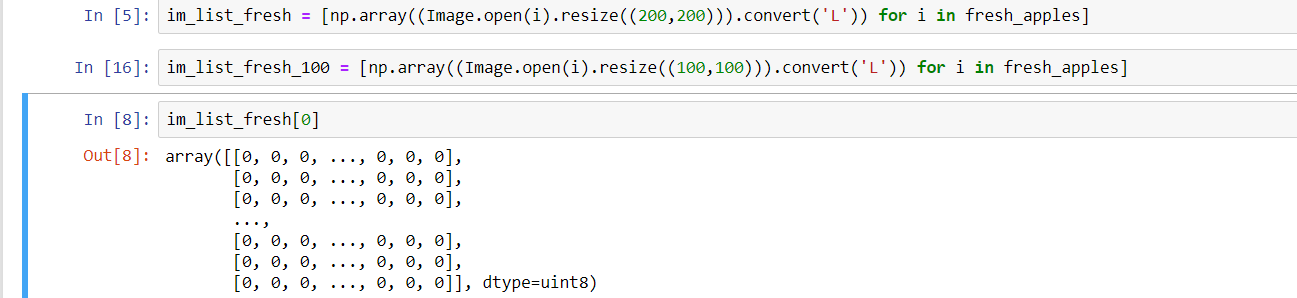
### working dataset from folders

1. Importing images from folder



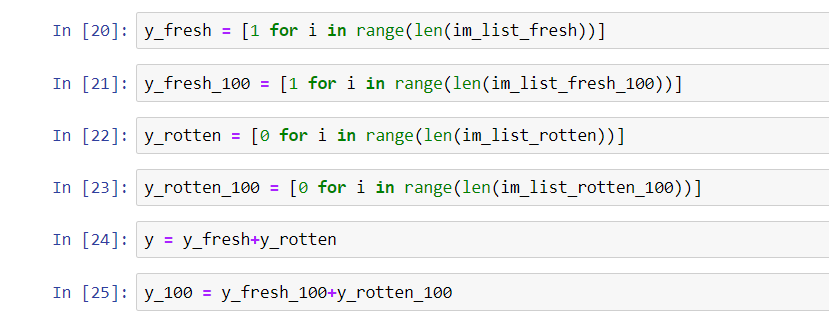


1. Resizing into 100X100 and 200X200 size and Gray Scale

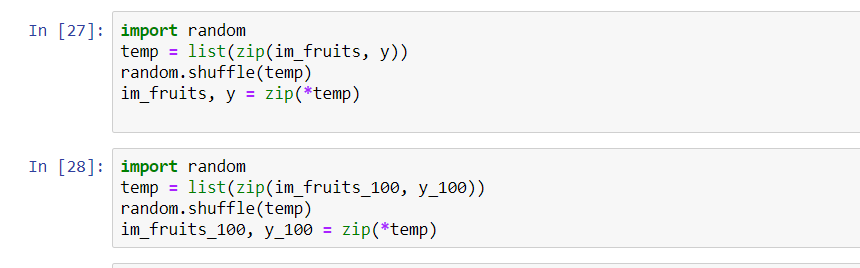




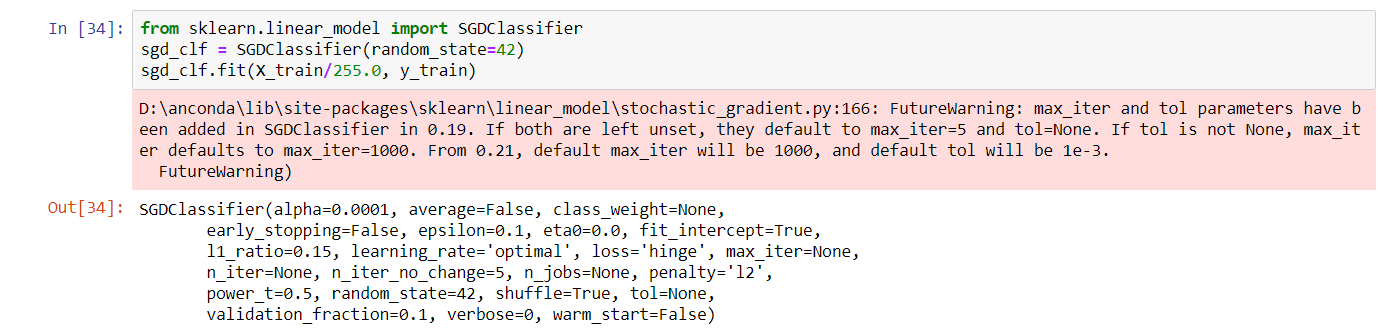
1. Creating data labels (Y set )



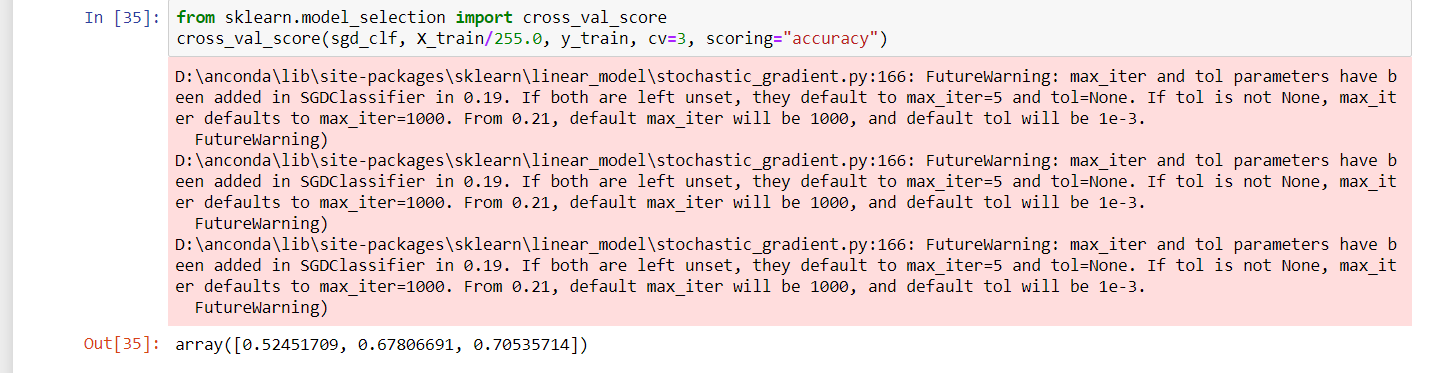
1. Randomizing the dataset



### Training on Classical ML models and Results



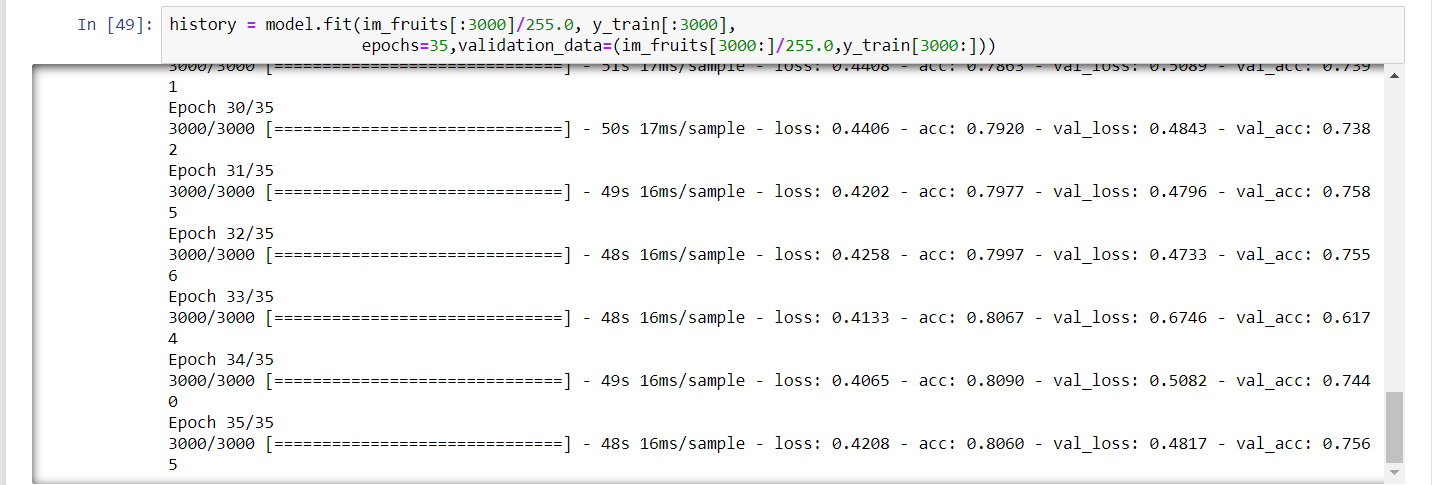
Result



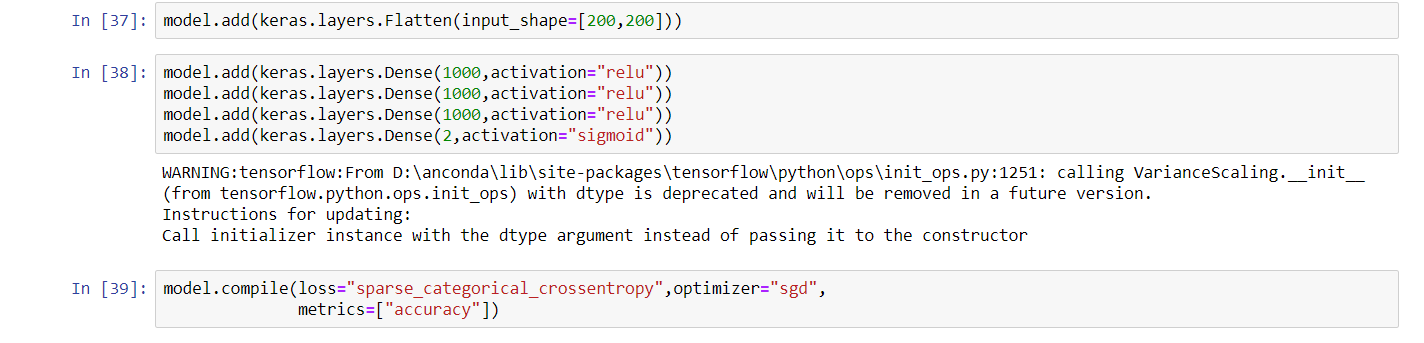
Also we tried SVM ,Logistic Regression but nothing exceeded the test accuracy of 75%.

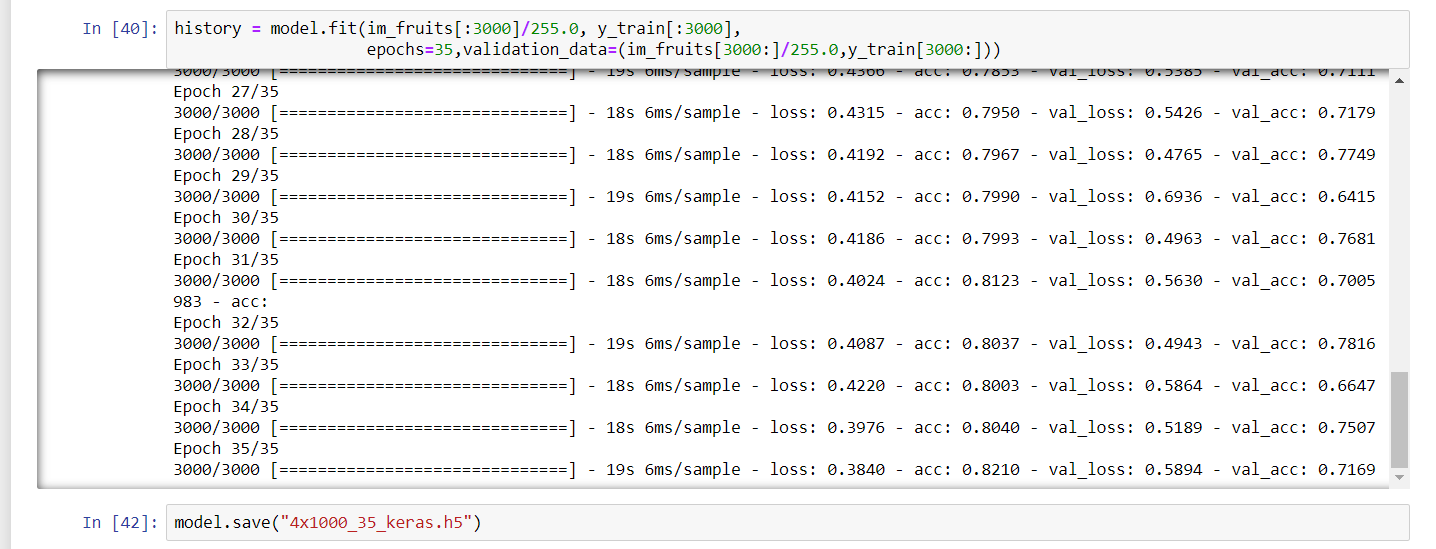
### Deep Learning Approach





Best Validation Accuracy -> 75.6%



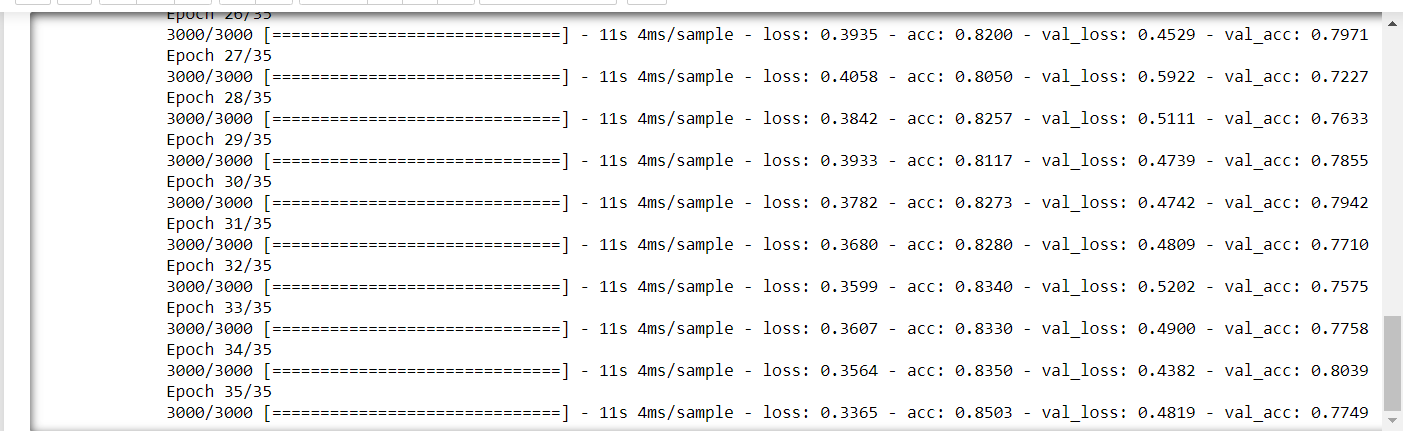


Best Validation Accuracy -> 78.7%

1. Histogram Normalization

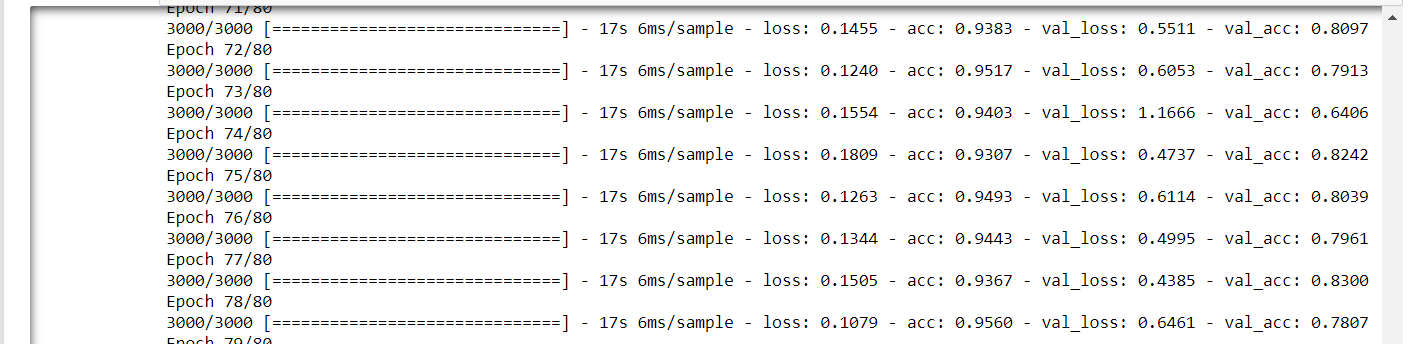


Best Validation Accuracy -> 78.7%

 Best Validation Accuracy -> 80%

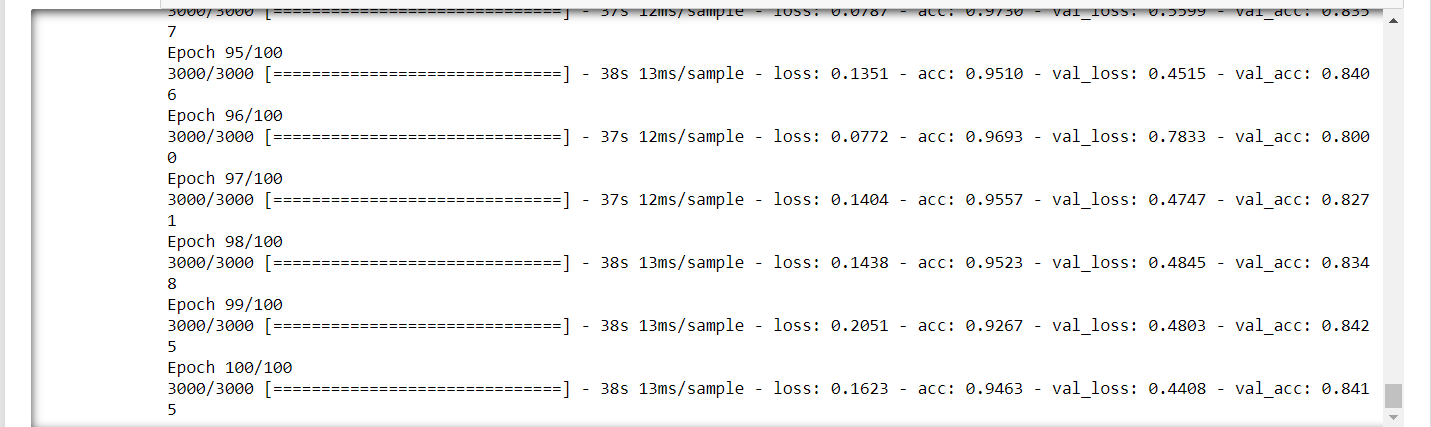
1. Using 100X100 image size to train the model (Surprisingly along with HistNorm. Technique it gave us better results rather than 200X200)





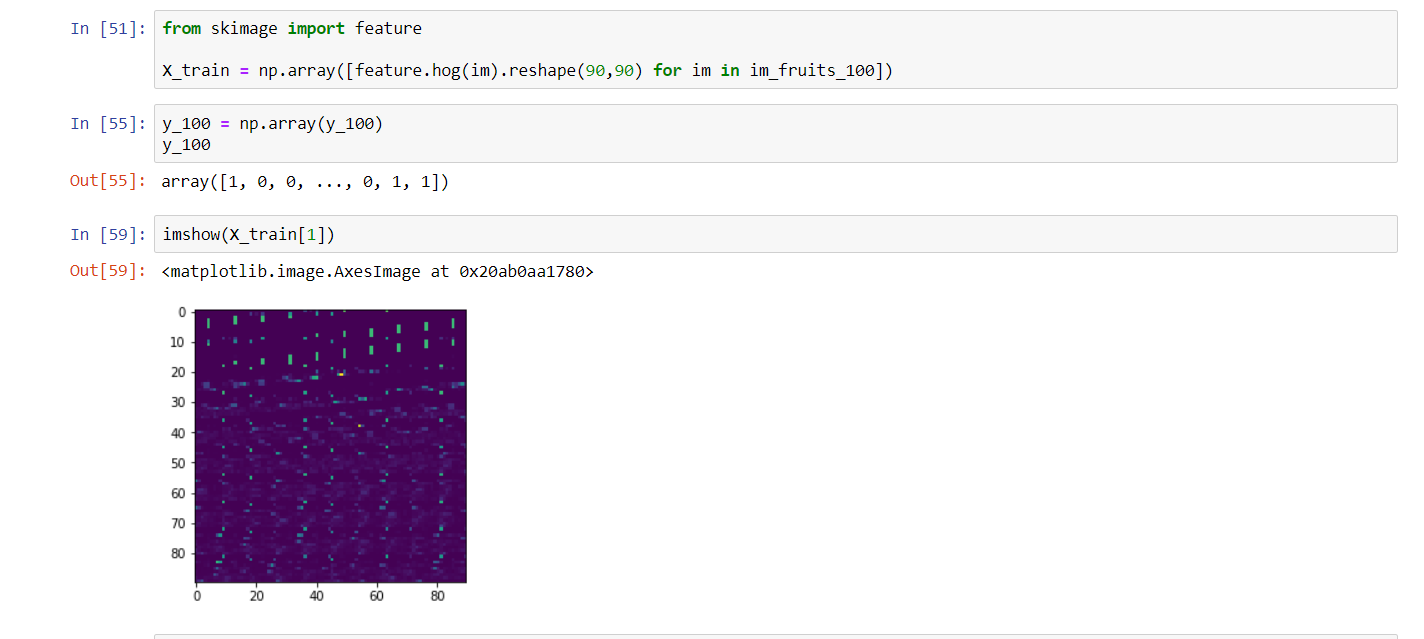
Best Validation Accuracy -> 83%

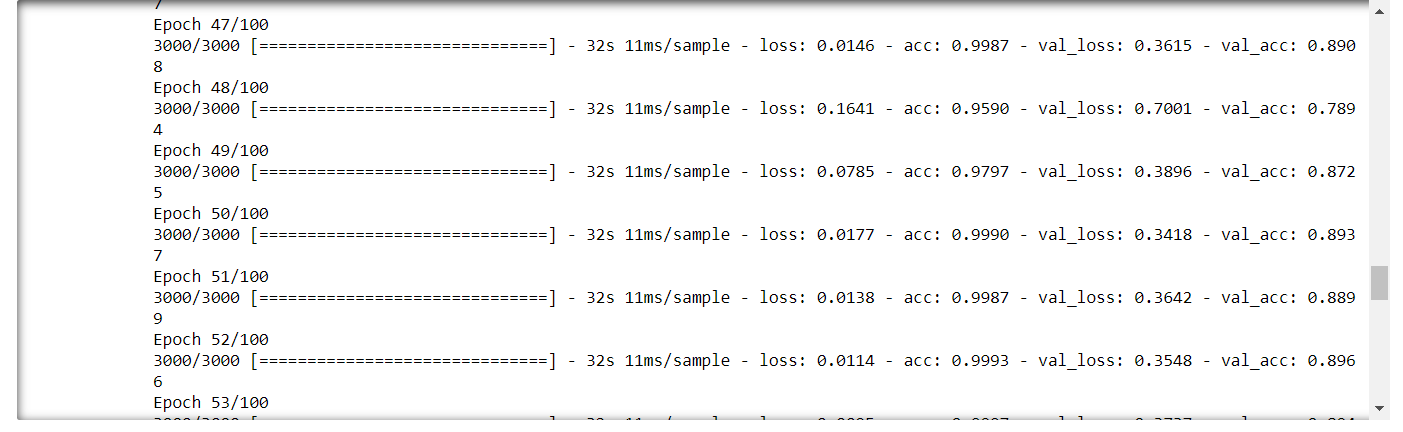
Fine tuning some parameters gave us slightly better results



Best Validation Accuracy -> 85%

1. Using Hog Feature Extraction Technique

 Best Validation Accuracy -> 90%



1. Future Improvement

There is a Huge scope for further improvement in this Solution

Some are ->

1. Convolutional neural network

A CNN can solve complex problems and also it is used for automatic images tagging and other image classification problems.

1. Better image feature extraction Techniques

Canny edge algorithm , Pixel Differential Method , PCA , MDR can be applied to find better results.

1. Using Different api and optimizers

Different Keras API s like Functional , Subclassing API and different dynamic models can be used to find better results.