REPORT - GROUP 31

Members:
- Mayur Shegokar
- Shreya Dutta
- Srijani Bheemarapu
- Ayan Khan
- Suraj Wamne
- Yadav S Nair

Topic:
Chest cancer detection using machine learning and deep leaning (CNN)

Synopsis:

The CNN Model we have used here is VGG-16. It stands for visual geometry group with a 16 layer depth. It is an object detection and classification algorithm which is able to classify 1000 images of 1000 different categories with 92.7% accuracy. It also won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) of 2014.

Given data contains 3 chest cancer types which are
- adenocarcinoma
- large cell carcinoma
- squamous cell carcinoma

and 1 folder for the normal cell.

The given dataset is divided as follows:
training set - 70%, testing set - 20%, validation set - 10%

The dataset was first imported from Kaggle and unzipped in the source directory of Google Colab.

Model:

Base model used was VGG-16. A series of flatten, batch normalization, dense and relu activation layers were added to it. Dropouts of 0.5 were added in between the sequence.

Accuracy: An accuracy of upto 70% was achieved on training the compiled model over 20 epochs.