Nutrition Management using Disease Prediction with Machine Learning

Pratibha Dohare, B.Tech (IT&MI) VII Semester (pratibhadohare@gmail.com)
Cluster Innovation Centre , University of Delhi, Delhi-110002, India.
Mentor: Prof. Shobha Rai, Cluster Innovation Centre, University of Delhi, Delhi-110007, India

INTRODUCTION

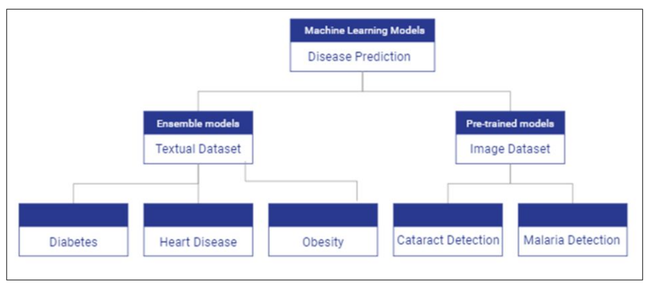
The aim of this project is to detect diseases at an earlier stage using Machine Learning and Deep Learning for Nutrition Management with high accuracy. The diseases include Cataract, Malaria, Obesity, Diabetes and Heart Disease. As getting the high accuracy is a concern , ensemble machine learning models and pre-trained models were used in this work. The results of all disease prediction models gave accuracy of more than 90%. Cataract Prediction gives an accuracy of 96% using VGG19 pre-trained model. Malaria Detection gives an accuracy of 95% using NASNet pre-trained model.

EXPERIMENTAL DETAILS

- The methodology is shown below:

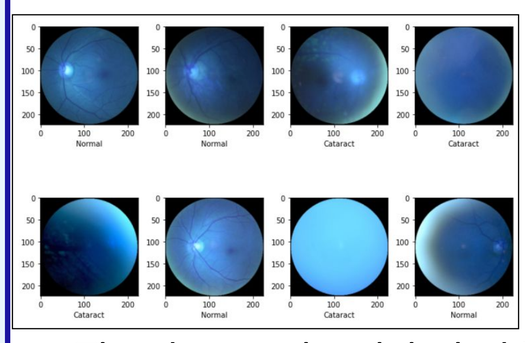
❏ Data collection
❏ Data Pre-Processing
❏ Data Splitting
❏ Model Building
❏ Accuracy metrics

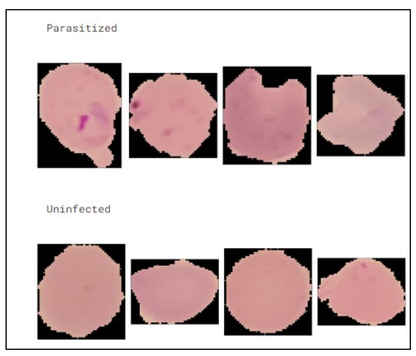
Two types of models are used in this project. Pre-trained and ensemble models are used here. The diseases and their model are shown below:



METHODOLOGY

Figure 1: Dataset for Cataract Prediction
Figure 2: Dataset for Malaria Prediction





The dataset has labeled images as normal and affected. Data is small so Deep Learning models did not give high accuracy. Pre-trained models are required here.

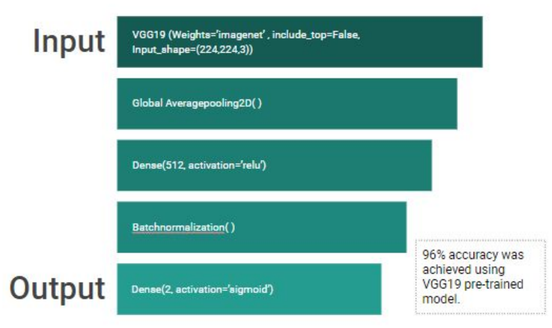


Figure 3: Pre-trained model VGG19 and NASNet are used here. VGG19 for Cataract Prediction and NASNet forMalaria Prediction. The images were resized to the dimensions required for pre-trained models.

RESULTS AND DISCUSSION

Figure 4: Accuracy and Loss for Cataract Prediction using VGG19 pre-trained model. 96% accuracy has been achieved using VGG19 model. Within 10 epochs only, accuracy has been increased. Loss is also minimized

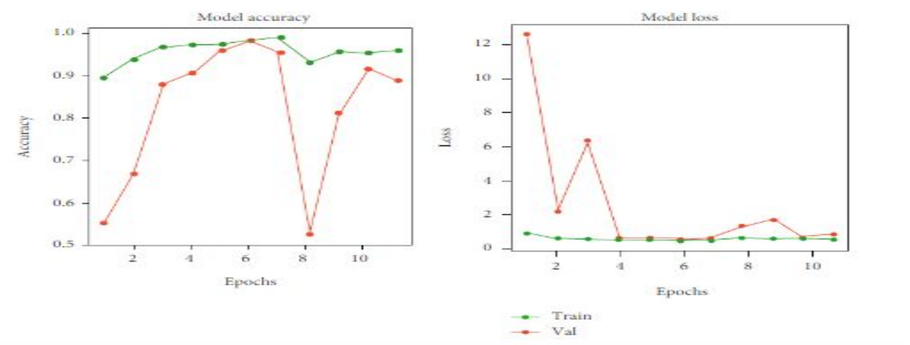
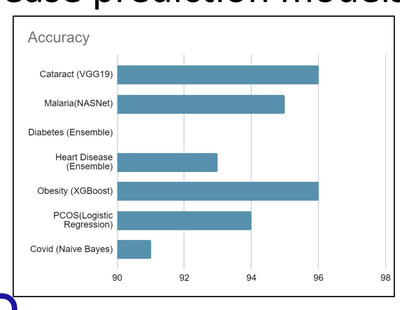


Figure 5: The ensemble model is used to predict obesity. The preidictions are made as insufficient weight, normal, overweight and then obese. XGBoost ensemble model is used here.

	precision	recall	f1-score	support
Insufficient Weight	1.00	0.95	0.97	20
Normal Weight	0.91	1.00	0.95	29
Obesity Type I	0.89	1.00	0.94	31
Obesity Type II	0.97	0.94	0.95	33
Obesity Type III	1.00	0.98	0.99	45
Overweight Level I	1.00	0.88	0.94	26
Overweight Level II	0.96	0.93	0.95	28
accuracy			0.96	212
macro avg	0.96	0.95	0.96	212
weighted avg	0.96	0.96	0.96	212

❏ The results of all disease prediction models gave accuracy of more than 90%.



CONCLUSION

- The main aim of this project (i.e. disease prediction with high accuracy) has been completed.
- Due to the lack of datasets, some predictions models were hard to build.
- Pre-trained models have found to be very useful in the image classification task.