Machine learning Applications in Healthcare

Submitted in partial fulfilment of the requirements

of the degree of

BACHELOR OF ENGINEERING
in
INFORMATION TECHNOLGY
(A.Y. 2018-2019)

by

Darpan Mehra (Roll No.: 42)
Varun Mishra (Roll No.: 47)

Under the Guidance of

Mr. Shridhar Kamble

Designation, I.T Department, TCET

University of Mumbai

CERTIFICATE

This is to certify that Mr. Darpan Mehra, Mr. Varun Mishra, are bonafide students of Information Technology Department, Thakur College of Engineering and Technology, Mumbai. They have satisfactorily completed the requirements of PROJECT-I as prescribed by the University of Mumbai, while working on "Project Title".

Signature :--------------------------
Name   : Mr. Shridhar Kamble
                 Assistant Professor

Signature :------------------------
Name   : Dr. Rajesh S. Bansode
                 HOD-IT

Signature: ---
Name   : Dr. B. K. Mishra
              Principal,
              Thakur College of Engineering and Technology.

Internal Examiner: **External Examiner:**

Signature :-------------------------------   Signature :------------------------------

Name   :   Name   :

Thakur College of Engineering and Technology, Kandivali(East) Mumbai.

Date:

Place:

DECLARATION

I/we declare that this written submission represents my/our ideas in my/our own words and where others ideas or words have been included, I/we have adequately cited and referenced the original sources. I/we also declare that I/we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my/our submission. I/we understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Signature)

1.Darpan Mehra (Roll no. 42)
2.Varun Mishra (Roll no. 47)

Date:

ACKNOWLEDGEMENT

We sincerely thank our principal, Dr. B. K. Mishra, Vice Principal, Dr. Deven Shah, Dean Dr. Kamal Shah, HOD, Dr. Rajesh Bansode for always encouraging us to do our best. We are highly indebt to our guide Mr. Shridhar Kamble who supported and constantly supervised us through this project and helped us not only in completing this project but also to provide us with ample amount of knowledge which was indeed beneficial to us.

1.Darpan Mehra (42)

2.Varun Mishra (47)

ABSTRACT

Though machine learning is not a new technology, there are new developments coming up faster than most people can learn them. It is a cutting-edge technology in every sense of the word. In such a case, it is necessary that we handle it with care. It could very well throw up results at an accuracy rate that is infeasible for medical applications. The capabilities required for the project are fairly in the feasible range. Most of the planned feature hinge around processing datasets. Due to the above, a challenge also visible here is the problem of dealing with huge amounts data and processing it thoroughly to get proper results.

The central parameter of success is the accuracy at which the machine learning algorithm predicts the diseases for which it is trained. It is expected to be above 96% - 98% to make a real difference, as medical diagnosis has very rigorous standards [2]. The secondary parameters are the response and loading time of each modules. A high value of module latency can make any user impatient. Besides these 2, the other parameter is the number of features that work without any bugs. This parameter should be ideally above 90% as well, or it can make users frustrated.

Although machine learning is a nascent specialization, we found an abundance of algorithms, not just in the proposed state, but rigorously scrutinized in research papers. Convolutional neural networks were suggested for their high accuracy. Chinese researchers had used Convolutional neural networks in a large-scale analysis of medical records of hospitals in China and reported high accuracy in prediction through symptoms [2]. Besides this, many researchers had tested SVM and ANNs, which were conventionally image classification algorithms. They also reported a high value of accuracy [4].

Until now, our work has focused on the training of our algorithm, with the use of medical data available openly. The algorithm has not shown the required accuracy and will need to be trained further. The other module we have built is the GUI. We built a web-based GUI in flask, an open source python framework. These two modules are the endpoints of

the project and building them first allows the ones in the middle to be built with convenience.

The comparison between existing results and actual results show a gap in accuracy that has to be closed. However, we were able to achieve more versatility, which was probably the reason that the algorithm was less accurate in the first place.

CONTENTS

I. List of figures

Figure number	Name	Page no.
Figure 3.4	Timeline chart	21
Figure 3.5	Bar chart	22
Figure 4.1	Workflow of bot	24
Figure 4.2	UML Diagram	25
Figure 4.3.1	Web GUI	26
Figure 4.3.2	Command line interface	26
Figure 4.4.1 & 4.4.2	AIML Database	27
Figure 4.4.3	Binary storage (.brn Database file)	28
Figure 5.1.1	Command line working	29
Figure 5.1.2	Working of web implementation	30

II. List of tables

Table number	Name	Page no.
Table 2.2	Literature survey table	7
Table 3.4	Timeline chart table	20
Table 3.5	Gantt chart table	21