Stroke prediction using decision trees in artificial intelligence

Aishwarya Roy
royaishwarya11@gmail.com
National Institute of Engineering, Mysuru, Karnataka

Anwesh Kumar
anwesh.rv@gmail.com
National Institute of Engineering, Mysuru, Karnataka

Navin Kumar Singh
navinkrsingh28@gmail.com
National Institute of Engineering, Mysuru, Karnataka

Shashank D
shashank.nie@gmail.com
National Institute of Engineering, Mysuru, Karnataka

ABSTRACT

Artificial Intelligence has become a hot topic in the present tech-driven world. Artificial Intelligence is one of the promising technology that has been greatly evolved from the past years. Artificial intelligence (AI) aims to serve as human cognitive functions. It will be of great importance to healthcare, p by increasing availability of healthcare data and rapid progress of analytics techniques. We survey the current status of AI applications in healthcare and discuss its applications in future. AI can be applied to various types of structured and unstructured healthcare data. There are various popular AI techniques that include machine learning methods for both structured data and unstructured data, such as support vector machine, neural networks, natural language processing respectively. Disease like cancer, neurology and stokes can be easily detected by AI. We conducted a survey about AI applications in stroke, in major areas that include detection and diagnosis, treatment, and lastly outcome prediction and prognosis evaluation. Our project is based on how we can make accurate predictions of stroke occurrence that can be of great help for the doctors. This can be time saving. It can also serve as a helping hand for the new practitioners. The predictive algorithm that will be used will increase the efficiency of stroke prevention that will surely improve the patients' health through early detection and treatment. The objective of this project is to have a system that can make accurate predictions on stroke so that it can be cured as early as possible. For this we need some predictive algorithms and parameters that includes patient's characteristics like age, gender, weight, BMW, height etc. We will have a data model that will analyse all these parameters. After having this survey, whenever a new patient will come this trained data model will compare the new parameters with it surveyed parameter with the help of learning algorithm.

Keywords: *Artificial Intelligence, Decision Tree, Heat Srtoke detection.*

1. INTRODUCTION

In this modern world, the advancement in technology has become an important part of everyone's daily life .Nowadays Artificial Intelligence has become a promosing technology in world of Science AI is defined as the study of "agents that are intelligent". Artificial Intelligence is a device that perceives its environment and takes actions according to environment to maximize the chance of success of achieving its goals. Artificial Intelligence is applied when a machine tries to be or act as or perform cognitive functions that humans associate with other human minds. Cognitive Function refers to learning and problem solving.

AI can perform various functions healthcare, even fuelling an active discussion of whether AI doctors will eventually replace human physicians in the future. We believe that human physicians will not be replaced by machines in the foreseeable future, but AI can definitely assist physicians to make better clinical decisions or even replace human judgement in certain functional areas of healthcare. There is sudden increase in hospital data along with rapid development of big data analytic methods which made possible the recent successful applications of AI in healthcare. AI techniques can unlock clinically relevant information hidden in the massive amount of data, which in turn can assist clinical decision making. There are varieties in movies like educational, thriller, entertainment movies .It can also be classified by their genre like comedy, horror, thriller etc. Due to this large no of options, choosing the favorable one is a troublesome job.

Machine Learning and Artificial Intelligence are the first choice for data mining and big data. They provide several applications which include techniques such as Neural Network modeling, Simulation models, DNA computing and Quantum computing and several others. Because of the application of AI in Biomedical field, the problem randomness in handling such type of data has reduced significantly. Many technological advancements have helped AI techniques to evolve in such a way that promotes handling such data effectively and conveniently. This implies that machine learning and AI models, also serve as advanced data analytics and optimisation approaches such as Drug Designing and Analysis, Medical imaging, biologically inspired learning and adaption for analytics, etc

Heat stroke also called sun stoke occurs when human body temperature is more than 40 degree. Heat stroke happens when there is sudden increase in external temperatures or due to physical exertion. Heat waves, high humidity and also some drugs such as diuretics or alcohol increases the risk of heat stroke. Skin disease and heart attack can also be considered as risk factor. Heat stroke occurs when heat regulation is disbalanced due to exertion or excessive hot climate, and insufficient heat loss, resulting in an abnormally high body temperature and lack of sweating. So, early prediction of heat stroke can prevent many ill-effects on people. In this application we are trying to predict the symptoms of heat stroke at a very early stage so that the consequences on patients should be minimum. We will have a valid set of data which will be analysed thoroughly and then based on that we will have our application compare the that set we have with the new data that have been given by the patients and will generate a report accordingly. We will have a training model that will take parameter as input like patient's age, sex, blood pressure etc. and will compare with the datasets that we have using decision tree and then with the help of collaborative learning algorithm it can suggest the patients report how much he is susceptible to heat stroke and will also send it to their email. This application will be of a great help for the newly practitioner who have to consult to their senior doctors for any kind of recommendation or suggestions to their patients.

2. LITERATURE SURVEY

Many system have been developed in recent times to for stroke predictions in recent times all differing by small factors.

M. Sheetal Singh *et al*.[1] used the heat stroke prediction is based on Cardiovascular Health Study dataset. Decision tree is used to select form the given condition on each stage. Principal component analysis is done to work on simpler dimension.

Mhd Saeed Sharif*et al*.[2] stated that because of the increase of the strokes there are many premature death nowadays. We can predict the the symptoms of stroke through medical data analytics.These medical data will be analysed by the Nao Robots that will give us an automated approach and also encourages robotics real time interaction with human body.

Aviv Yuniar Rahman *et al*. [3] survey that there are so many online health videos that guide people about various disease. On the other hand we came to know that these and comparative results were an improvement in the quality, accuracy and capability of the movie recommender as compared to traditional approaches. Information about diseases are not correct. These will create negative emotions on the patients' health that will increase the patients' blood pressure which will result in stroke. Handling such patients begins with proper information about each and every aspect of stroke symptoms and cure. For this we need early detection of stroke and stroke rehabilitation automatically. So, we will make a stroke extraction video that will help the patients to be properly aware of the symptoms of stroke. The video object extraction includes segmentation process that helps in future stroke prediction. Segmentation is used to maximize the prediction process of healing and rehabilitation of stroke patients in subsequent studies. This study uses a minor stroke extracted video object using LVQ which has been modified.

Chen-Ying Hung *et al*.[4] suggested that electronic medical claim is used for higher precision and accuracy to predict the disease. Electronic medical claim is a database on which various predictive algorithm is applied to give results of higher accuracy. We use deep learning i.e deep neural network to simplify all the complexities.

Asma Rashed Al Taleb *et al*.[5] surveyed that for proper management of hospital resources it is necessary to plan the resources according to the average length of the disease. Data Mining is used here. The information is selected on the basis of a selection process from the large set of data. Then the predictions results that will be done by predictive algorithm will have certain comparisions . The better and accurate one will be used.

3. PROPOSED SYSTEM DESIGN

SOFTWARE REQUIREMENTS AND SPECIFICATIONS

A software requirements specification is a description of a software system to be developed, describing all the functional and non-functional requirements which also include a set of use cases that describe interactions the users. Software requirements specification establishes the basis for an agreement between customers and contractors or suppliers (in market-driven projects, these roles may be played by the marketing and development divisions) on what the software product is to do as well as what it is not expected to do. Software requirements specification permits a rigorous assessment of requirements before design can begin and reduces later redesign.

The software requirements specification document conatins enough and all the necessary requirements that are required for the project development. To have the requirements we need to have thorough understanding of the products being developed. Software.

Functional Requirements

A functional requirement is related to the functionality of a system or of its subsystems. It also depends upon the type of software, the type of system where the software is used and the users.

Functional Requirement of our system is listed below:

Admin:

- Admin will view the registrations

- Admin will decide whether to accept the registration or else will reject the registration

- Admin will verify all department (Doctor, Patient, and Nurse).

Doctor:

- Doctor should register and he will login through user id and the password.

- Doctor will view the patient details and receive the report from nurse

- After doctor will check the patient report details and predict the disease based on history of data (using Decision tree technique) and give suggestion to patient

Patient:

- Patient should register and login through user id and the password.

- Patient enter the all details.

- After consulted doctor he go for tests

- Patient will view the final reports details.

Use Case Diagram

The use case diagram will give us the graphical representation of all the main actors and their functionality of the system like doctor, nurse, patients and how they are related and what actions or functions they perform in the system.
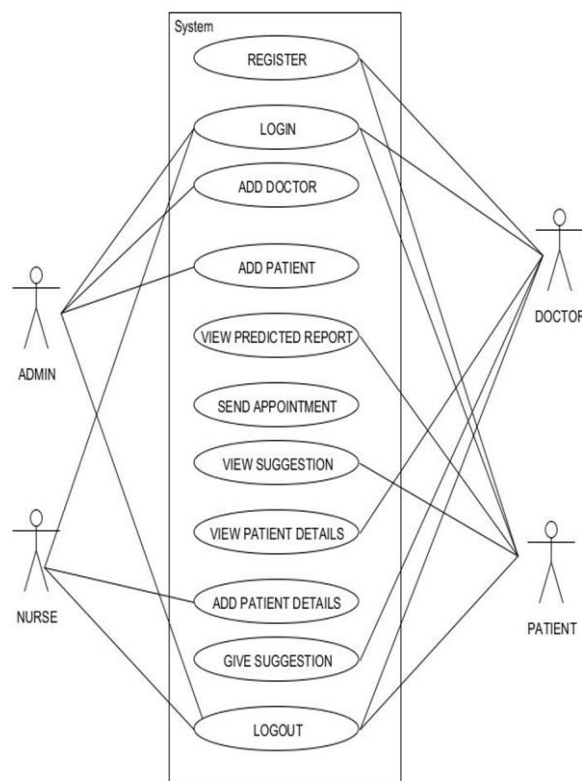

Fig. 1: Use case diagram of the system

Sequence diagram of system operation

A sequence diagram is a diagram that shows the interaction between all the major actors like doctor, nurse, patients. This will also

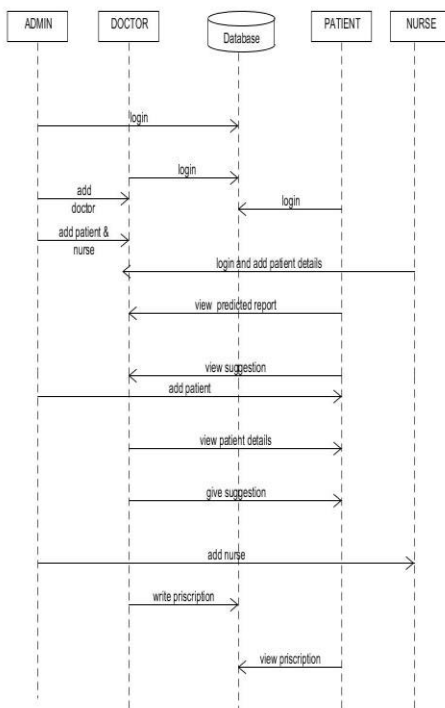tell us how one processed operate with another and also tell us the sequence of operations.

Fig.2: The sequence diagram

Data Flow Diagram of the system

A data-flow diagram is a representation that how data flow through a system. It is used for data processing which show us how data from external to internal process.
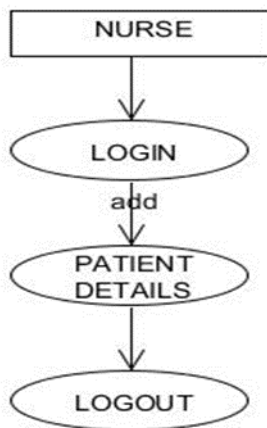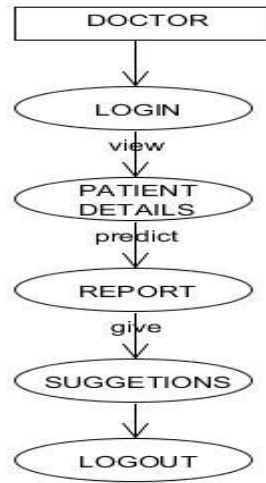

Fig. 3: Nurse
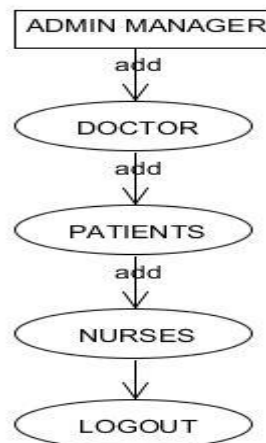
Fig. 4: Doctor

Fig. 5: Admin

Fig. 6: Patients

System Architecture Design

This represents a workflow of a system showing all the steps diagrammatically how you initiate a process one after the another.
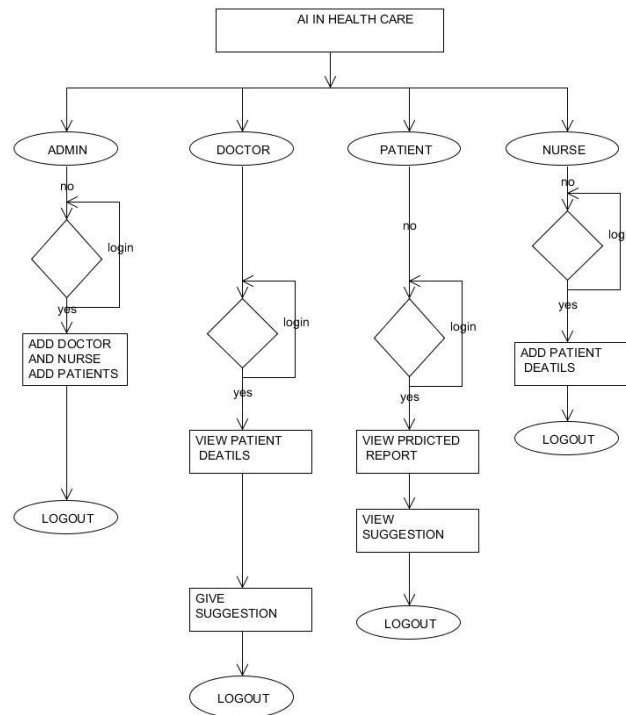
Fig. 7: Basic Architecture of the model

Hardware Requirements

- Processor   -   Pentium III

- Speed   -   1.1 Ghz

- Hard Disk   -   20 GB

- Key Board   -   Standard Windows Keyboard

4. RESULT ANALYSIS

The result is analyzed using Confusion Matrix and the accuracy is calculated using the following rule:

$$Accuracy = (TP+TN) / (TP+TN+FP+FN)$$

Where, TP = Actual results is true and obtained result is also true.

TN = Actual result is true and obtained result is false.

FN = Actual result is false and obtained result is true.

Table 2 shows the experimental result of different methods and proposed method applied on our selected dataset. Here, the proposed method gives better accuracy rate as compared to the other two methods.

Dataset	Methods and Accuracy		
	NN [1]	*PCA and NN [2]*	*DT, PCA and NN (Proposed)*
STCL	82.5%	92.5%	95.0%

STTIA	78.6%	92.9%	95.2%
STAN	90.9%	95.5%	97.7%

5. CONCLUSION

The project will have web application through which the doctor or even the patient can fill the details. Details will consist of patient's characteristics which includes person's age, sex, weight, height etc. We will have the training model that helps to compare the newly feed data with the surveyed one and will generate the report on the basis of the comparison. This report will be sent to the person email id.

There is huge contribution of AI in healthcare by being a helping hand to the newly practitioneers. This agent can help them to assit in suggesting the prescription or some health related advices to the patients that will decrease the time taken to generate the report

or to take advices from the senior doctors.

6. REFERENCES

[1] M.Sheetal Singh,Prakash Choudhary "Stroke Prediction Using Artificial Intelligence" 2017 IEEE International Conference.

[2] Mhd Saeed Sharif, Mohammed Hayyan Alsibai "Medical Data Analysis Based on Nao Robot: An Automated Approach towards Robotic Real-Time Interaction with Human-Body" 2017 IEEE Interntional Conference.

[3] Aviv Yuniar Rahman, Surya Sumpeno, Mauridhi Hery Purnomo "Video Minor Stroke Extraction Using Learning Vector Quantization" 2017 Fifth International Conference on Information and Communication Technology (ICoICT).

[4]Chen-Ying Hung, Wei-Chen Chen, Po-Tsun Lai, Ching-Heng Lin, and Chi-Chun Lee "Comparing Deep Neural Network and Other Machine Learning Algorithms for Stroke Prediction in a Large-Scale Population-Based Electronic Medical Claims Database" 2017 IEEE International Conference

[5] Asma Rashed Al Taleb, Mozaherul Hoque Abul Hasanat, Muhammad Badruddin Khan "Application of Data Mining Techniques to Predict Length of Stay of Stroke Patients" 2017 IEEE International Conference