THE ROLE OF MACHINE LEARNING MODELS FOR HEALTHCARE APPLICATIONS

Dr.P.Sujatha[1], T.Ravishankar[2]

[1]Professor, Department of Information Technology, Vels Institute of Science,
Technology & Advanced Studies, Pallavaram, Chennai, India.sujathap.research@gmail.com

[2]Research Scholar, Vels University, Asst Professor, Dept of IT,
Guru Nanak College (Autonomous), Chennai, India, trsravi2016@yahoo.com

ABSTRACT

Machine learning is a key to enabling AI driven healthcare, finding patterns in huge data sets to facilitate decision-making in clinical practice. The effective use of machine learning will help healthcare specialists and organisations to extract the insights locked in large repositories of data from numerous sources. In the medical level, machine learning applications may help specialists detect and treat disease and health conditions more efficiently. Its algorithm process huge amounts of patients data to make reasonable conclusion faster and more exactly than the human brain.Various algorithms have been designed, such as SVM algorithm, random forest algorithm, decision tree algorithm, artificial neural network algorithm, BP algorithm. Boosting and Bagging algorithm and many programming languages have been used to for implementation. The main goal of ML in the healthcare industry is to improve processes and outcomes.

Keywords: Machine learning, HealthCare, Machine learning algorithms, Machine learning techniques and applications.

1. INTRODUCTION

Machine learning generally refers to the modifications in systems that perform tasks related with artificial intelligence. [1],[2] Such process involves diagnosis, prediction, robot control, planning, recognition, etc. Machine Learning is a division in the field of computer science that works on data and algorithms. It replicate how human brain learns, continuously improving its accuracy. The area of applications in the healthcare industry is quite broad.
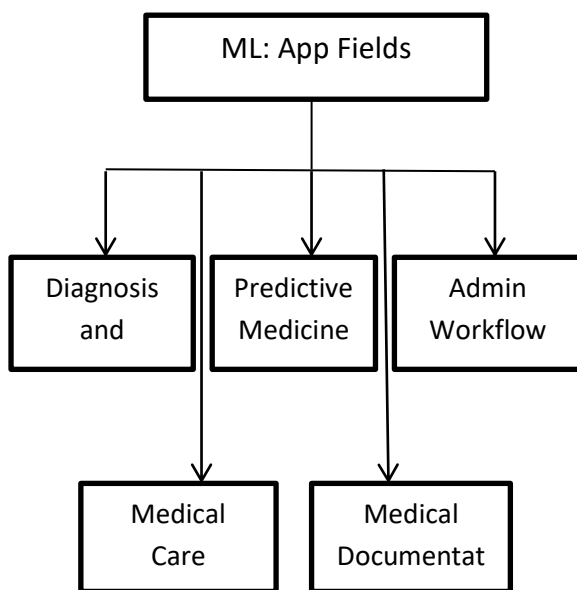

Table 1: Application Fields

Machine learning is also used in different industries such as marketing, finance, logistics, etc. Many companies use advanced algorithms to predict what kind of goods we are likely to buy and what content we would prefer to consume[3]. Different algorithms are applied for the healthcare industry to bring valuable outcomes for the patients.

2. LITERATURE REVIEW

R. Karwa and V. Honmane[2019] stated that healthcare is developing now days, investigators are focusing on types of statistics used for prediction. Machine leaning algorithms plays a dynamic role in many applications, e.g. disease diagnostics, data mining, image detection, and natural language processing. It also offers possible solutions for all these domains. For analysing different types

of data in healthcare it has been concluded in the survey that various machine learning algorithms and feature removal methods are proposed.

K.B.Cohen and L. Hunter [2006], for processing and handling of biomedical words "Biomedical Language Processing" are involved. By using the natural language it takes the name of disease and gives the resolution which has been kept in record of that disease by parsing user statement but it does not do any analysis of the disease.

Jiang Na, Yang Haiyan [2019] quoted that Machine learning supports the specialists to compare previous patient's record with a new one so that they can assess the disease in detail earlier. From the documented information, concludes that machine learning supports doctors to compute the symptoms of disease and diagnose.

Raheja K, Dubey A, Chawda R. [2018] mentioned that any hereditary and cancer diseases are not easy to identify or recognize, but machine learning can handle them. Healthcare tools will now depend on machine learning because, with this the identification and diagnosis of diseases become easier and error-free. ML techniques are limited to the basic methods of Naïve Bayes, genetic programming, Bayesian networks, random forest, neural networks,   and classification and regression tree (CART).

3. MACHINE LEARNING TECHIQUES FOR BIO-MEDICAL APPLICATIONS

In ML past data is given as input to the computer to solve the problem based on the parameters applied for the training data set. ML is used in various domain applications to make better, efficient and accurate solutions than human with the help of cheaper configuration computer with inexpensive cloud storage, we can efficiently analyse the high dimensional data's and easily find out the hidden patterns and correlations among large data set. Based on the algorithms applied for the input training data ML perform and provides an accurate outcome in efficient manner enables machine to make intelligent decision and prediction.

DL is a sub-field of ML, however, an additional innovative approach which supports computers to extract, evaluate and understand valuable information from the raw documents by imitating how humans think and absorb. [10] Exactly it can define that, deep learning is a group of methods that is neural data driven and based on automatic feature engineering processes. The technique of extracting the features on which the classifier works on plays the main difference of ML and DL.
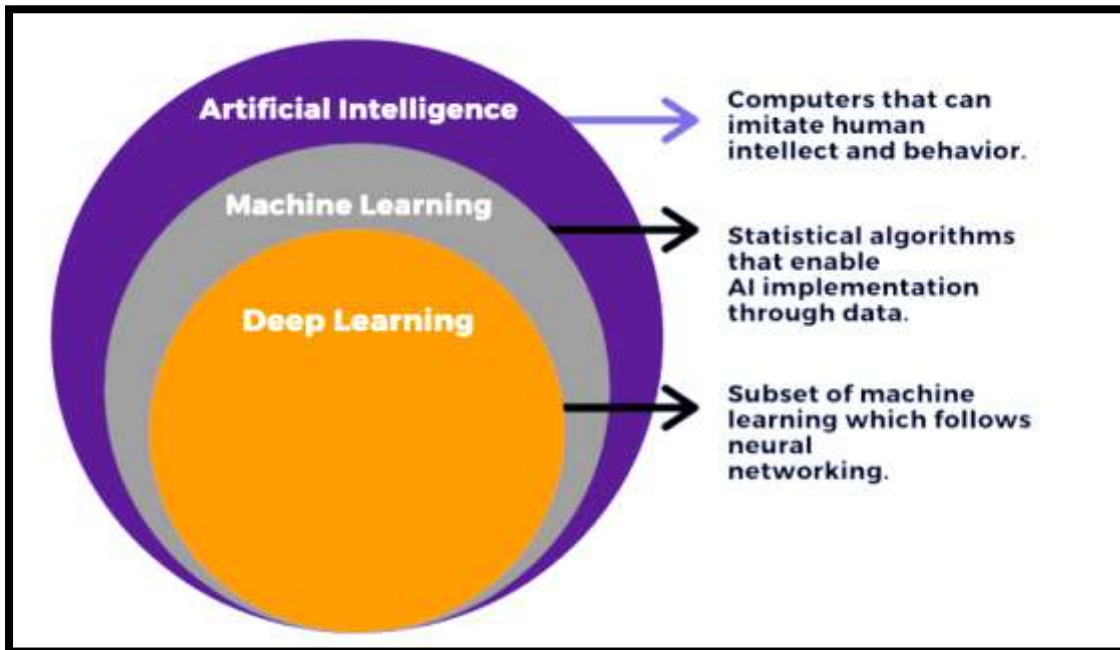
FIG1: Comparison

The main branches of machine learning are listed below

I. Supervised Learning

Here we train the machine using data which are labelled. The data used are split into two i) training and ii) test dataset. Using the training dataset we train our model and with testing dataset function as new data to get accuracy. We are giving the pre-labelled data to make the model learn from it.Based on different parameters the function helps in dividing the dataset into classes which is called as classification. To detect patterns and calculate the probability of unavoidable outcomes the regression method is used. Many algorithms, Decision Trees Linear Regression, Random Forest, Logistical Regression, and Naive Bayes, are based on Supervised Learning Algorithm.

Classification: It is a supervised learning (Labelled data).To predicts two classes we using classification problem, another name is binary classification .If we want to classify more than two classes, then it is said to be a multi-classification.

Regression: A target value is modelled by independent predictors. Regression method mostly used to find out prominent relationship among variables, regression method varies based on the individual variables and type of relationship between dependent and independent variables.

II. Unsupervised Learning

Unsupervised learning is machine learning method in which model are not supervised using training dataset. In the dataset data used to train are not labelled. When you give input to the dataset, it will automatically work on it, storing the classified information and finding patterns and relationships as the trained data set. This process will continue until the test data is exhausted. To find hidden patterns or grouping in data the cluster analysis method is used for exploratory data analysis. The main mission of the machine is to group similar types of data together. An example of an unsupervised learning technique is clustering method.

Clustering: Unsupervised Learning (Unlabelled data), based on the similarity on the data will group the data together and mathematical concepts are used to measure the similarity distance between data points.

*Association***:** This mining technique finds most continuously occurring data items or inter-related elements.

Summarization: Relevant data set is summarized from different resources and result reports aggregates as a smaller set of data.

III.   Reinforcement learning

For training the model we do not use the datasets. Instead, the machine takes certain steps on its own, analyses the feedback, and then tries to improve its next step to get the best outcome. The model develops its performance and figures out how to deliberate its faults and to get the correct outcome through testing various prospects and assessment based on its association.

4. APPLICATIONS IN HEALTHCARE SECTOR

- **Health Records**

There is a rise in requiring personalized medicines and treatments, there's also a growing need for smarter health records. Using machine learning, healthcare facilities can keep up with health records without having to waste so much time to just keep them up-to-date. ML tries to save people time, money, and effort; the market size for it has been increasing over the years in the healthcare industry.

- **Diagnosis Diseases at an Earlier Stage**

In major predictions of health condition such as diabetes, heart disease and other viruses machine learning plays an vital role. To monitor the health of a person many AI based wearable are being developed and display any warnings when the devices observe something unusual. Example: Apple watch and Fitbit. These devices monitor a person's sleep cycle, heart rate level, activity level,

breathing rate level, and blood pressure, so on.

FIG2: Diagnosis

- **Epidemic Outbreak Prediction**

    Artificial Intelligence and Machine Learning tools are being useful to observe besides predict epidemic outbreaks around the world, based on data collected from satellites, ancient information, real-time social media, and other source. Here the artificial neural networks and Support vector machines have been used, for example, to predict malaria outbreaks, taking into account data such as average periodic rainfall, temperature, overall total positive cases, and other data points

- **Healthcare for Surgeries**

    For patient and doctor the current technological innovations continuously strive to improve the healthcare situation in all aspects. Machines focus on improving the performance of operations, and they can help doctors by using surgical robots. Helps the specialists with high definition imagery and extended flexibility to extent out in areas that are critical for a surgeon. Machine learning has several other applications in numerous fields that try to improve human life.
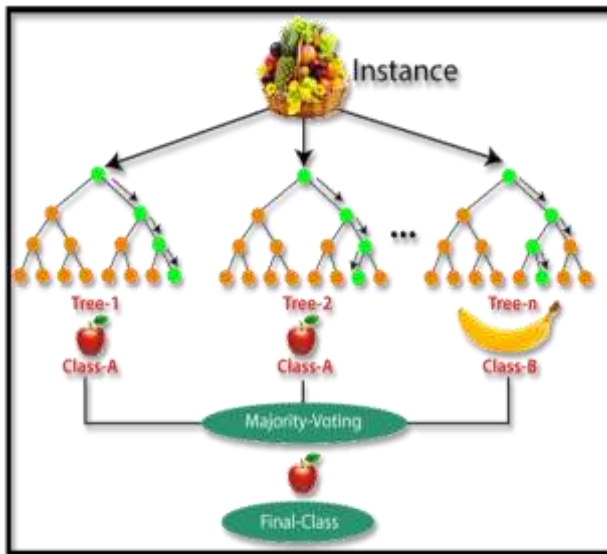
FIG: 3: Advanced Technology in Healthcare

5. EXISTING MACHINE LEARNING METHODS

- **Random Forest**

Random forest is a form of supervised machine learning algorithm. In both Classification and Regression problems this algorithm can be applied. It uses an ensemble learning method known as 'bagging' (Bootstrap Aggregation) which is a process of joining multiple classifiers to solve a composite problem. In the bootstrap aggregation each tree will do its own random check break-up of the information within the random wooded area and the samples are no longer protected are called the out-of-bag samples. Additionally, every tree will do specific bagging at every node-branch split to reduce the results of a characteristic mostly correlated with the response. While splitting a node Random forest searches for the most important feature and build a better model.

FIG 4: Random Forest

- **Decision Tree**

A decision tree is an upside-down tree that makes decision based on the condition present in the data. To decrease the variance of a decision tree the bootstrap aggregation is used. Here we produce several subsets of information from training sample chosen randomly. Based on the observation aggregation reduces these sample datasets into summary statistics and combines them .The random state was kept as 0 and then the training model was fitted with the decision tree model.
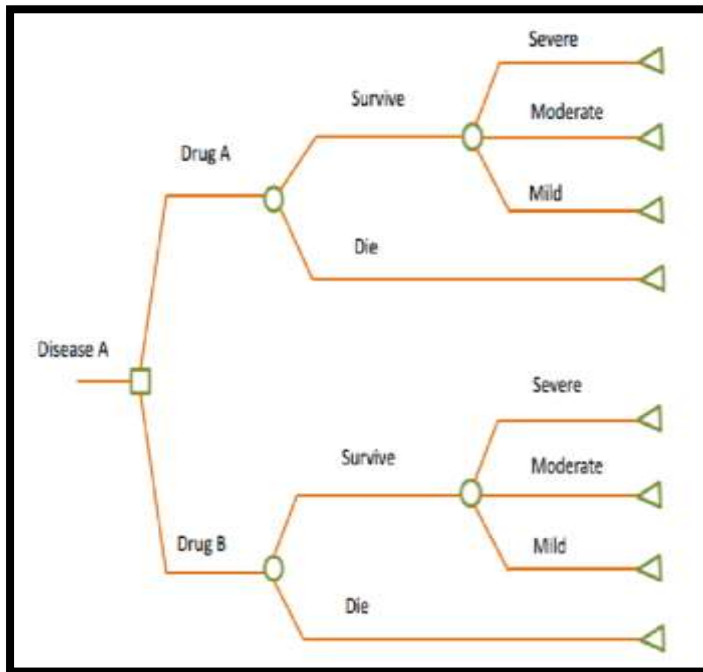
FIG: 6: Decision Trees for Disease

- **Confusion Matrix**

A confusion matrix is a precis of prediction outcomes on a category problem. The huge sort of correct and misguided predictions are summarized with rely values and broken values thru each elegance. By computing confusion matrix can offer you the style of errors it is making and with a better idea of what your elegance model is getting right.

6. EFFICIENT TOOLS FOR MACHINE LEARNING MODELS- PROGRAMMING LANGUAGES

- **Python**

Python is the most popular language for machine learning even though its considered and is used as a general-purpose programming language. Many researchers favour Python because it's has a huge source of libraries for specific purposes.

- **R   Programming**

Many data analysts and statisticians preferred R is statistical computing languages, who are making their mode into the domain of ML. The effective machine learning solutions, can be delivered with its heavy computing capabilities

- **Matplotlib**

Matplotlib is a python multi-platform data visualization library used to create 2Dimension graphs and plots by using python scripts. It can knob a variety of plots such as line, image, scatter, polar, contour, histogram and 3D.

- **SQL**

SQL is one of the best data collective tools for researchers apply for extracting information from relational and non-relational databases. Data scientist or a data analyst, are able to write structured queries to remove data is an essential skill.

- **Hadoop**

It's a group of software services that support in the scattered handling of enormous dimensions of data, through clusters. It breakdowns figures into files and allocates them across nodes in a cluster room which is controlled by the Hadoop Distributed File System (HDFS) and the handling by MapReduce.

7. CONCLUSION

In the modern years Machines plays a big role in our life. A lot of data collected in every part of our lives and these records are growing day by day. This paper is about smart health care system using machine learning where a system monitors our regular health data. Extra pressure over doctors and hospitals are reduced. It is also essential to set up appropriate machine algorithms for the growth of domestic enterprises, and provide minimal support for the financial growth of the industry.

REFERENCES

1. D. B. Calne "Is idiopathic parkinsonism the consequence of an event or a process" Neurology, Vol. 44, no. 15, pp. 5–5, 1994.

2 . J. Parkinson "An Essay on Shaking Palsy. London: Whittingham and Rowland Printing 1817".

3. Arwinder Dhillon*, Ashima Singh "Machine Learning in Healthcare Data Analysis: A Survey "

4, OPEN ACCESS:" Research on Machine Learning and Its Algorithms and Development"

5. R. Karwa and V. Honmane, "Building Search Engine Using Machine Learning Technique," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 1061- 1064.

6. Jiang Na, Yang Haiyan, Gu Qingchuan, Huang Jiya. Machine learning and its algorithm and development analysis [J]. Information and Computer Science (Theoretical Edition), 2019.

7. Alpaydin E. Introduction to Machine Learning. MIT press;2009.

8. Raheja K, Dubey A, Chawda R. Data analysis and its importance in health care. Int Computer Trends and Technology J. 2018; 48:176-180.

9. Karlik, B. (2011) Hepatitis Disease Diagnosis Using Back Propagation and the Naive Bayes Classifiers. Journal of Science and Technology, 1, 49-62.

10. L. Hunter And K.B. Cohen, "Biomedical Language Processing:What's Beyond Pubmed?" Molecular Cell, Vol. 21-5, Pp. 589-594,2006.

11. Parthiban, G. and Srivatsa, S.K. (2012) Applying Machine Learning Methods in Diagnosing Heart Disease for Diabetic Patients. International Journal of Applied Information Systems (IJAIS), 3, 25-30.