

# Machine Learning Application In The Realm Of Health Care

Mrs. J Padmavathi, (Assistant professor),

Bejjanki Nikhitha, Kodela Praveen, Ganesh Loka, Rajashekar Rondi,

Department of CSE,

MALLA REDDY INSTITUTE OF TECHNOLOGY AND SCIENCE, telangana, Hyderabad.

## ABSTRACT—

*In recent years, several applications have surfaced in the fields of artificial intelligence and machine learning, which have been the center of study. Not only is it something near to our lives, but it also exists as a kind of intellectual frontier. The integration of AI with healthcare is becoming more tight in this trend. Additionally, the current state of unequal medical distribution and resource pressure was significantly reduced by the introduction of its central concept. In this paper, we will take a look at how machine learning and auxiliary tumor treatment have been integrated into medical resource allocation. We will also discuss new ways that AI can bring these applications closer to people's lives and examine a mutually beneficial situation where the computer industry and the medical industry work together.*

## I. INTRODUCTION

### I. The state of machine learning generally

The goal of the scientific field known as machine learning (ML) is to program computers to learn new tasks. Alpha Go, Google's go program, and Go player Li Sedol's world-famous 2015 match, which ended in a 4:1 tie, brought machine learning back into the spotlight. Not only has this incident sparked heated controversy in adjacent fields, but it has also increased machine learning's visibility outside the realm of computer science. Machine learning may be a relatively new topic in artificial intelligence, but it has been around for a while. The capacity to learn from training data and make predictions about incoming data based on learning outcomes is at the heart of machine learning (ML), which is often described as the application of certain computer

algorithms to a collection of data known to the event outcomes. Instead of deductive reasoning, it relies on induction and summarization. Samuel, an American

computer scientist, created a chess software that could learn from its own play in the early to mid-1950s. At the same time that this program revealed machines' abilities to humans for the first time, it also brought the surprising learning capability of machines into the public eye. But machine learning entered a cooling down phase as research progressed. Its recovery was sluggish and did not occur until the 1970s. Throughout this time of ongoing study and development, machine learning has emerged as a significant field, including data mining, pattern recognition, NLP, and more. Additionally, it is now fundamental to AI. At this point, supervised, semi-supervised, and unsupervised learning are the primary categories into which machine learning falls. Among them, Nearest Neighbor, Naïve Bayes, Artificial Neural Network, and Support Vector Machine classifiers are some of the most popular. Among these algorithms, ANN has recently gained a lot of attention from researchers. Machine learning (ML) is an interdisciplinary area that encompasses many domains and can be studied in many different fields, including computer science, medicine, finance, and many more. ML can train itself to become better at what it does while also mimicking the way humans learn.

Issues with medical treatment, such as the uneven and inadequate distribution of available resources, have recently risen to the forefront of public discourse. Applying ML has become the current trend in medical care development in this case. Researchers from the United Kingdom's University of Leeds have been attempting to utilize ANN algorithms to diagnose stomach discomfort since 1972. A growing number of academics are devoting their time and energy to studying how ML and healthcare might work together. Over time, ML-based approaches to pathological diagnosis of cancers, lung cancer, etc. have made their way into the public eye. Alibaba, Amazon, and Baidu are just a few of the corporations that have formed in-house research teams. The use of ML in healthcare has opened up new avenues for individuals to get medical treatment, which has improved people's quality of life while simultaneously preserving precious medical resources. Simultaneously, human need offers fresh

..ISSN: 2040-0748

Vol-12 Issue-02 NOV 2023

impulses for ML R&D, which in turn encourages its ongoing enhancement.

accomplish the goal of learning, nonlinear data is processed using the input of the training set to train the neural network, which in turn modifies various weight values.

**Decision tree based methods are the primary algorithm of machine learning.**

Using the training set and current data, the algorithm of a decision tree may construct a decision tree. One way to categorize the test data is by looking at the decision tree's development direction. Which feature is best, how many branches may be formed, and when to cease splitting are the key ideas of a decision tree. During this procedure, it may be ascertained using a mathematical approach and a variable known as impurity. On the other hand, decision trees aren't always able to identify the optimal tree since they're greedy.

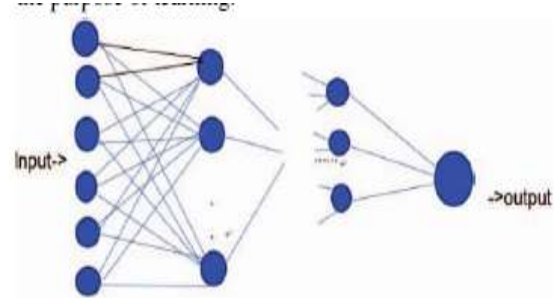


Figure 1. Artificial Neural Network (ANN) model (f)Support Vector Machine (SVM)

**Bayesian Belief Networks and Naïve Bayes**

By converting input space into high-dimensional space, support vector machines (SVMs) play a significant role in statistical learning theory. For linear indivisible problems, the approach may be employed to segment the hyper surface with feature space; in linear classification, the hyper plane and loss function are built to solve the minimum of the agent's loss. Medical condition analysis and tumor classification using support vector machines (SVM) is common, however SVM's potential need to calculate high-order matrices makes it challenging to apply to large-scale training sets.

Naïve Bayes is a machine learning approach grounded on probability theory. It computes the target object's prior and posterior probabilities, supposing that vents are independent. What follows is the formula.

$$P(Y|X)= \frac{P(X|Y)P(Y)}{P(X)} \quad (1)$$

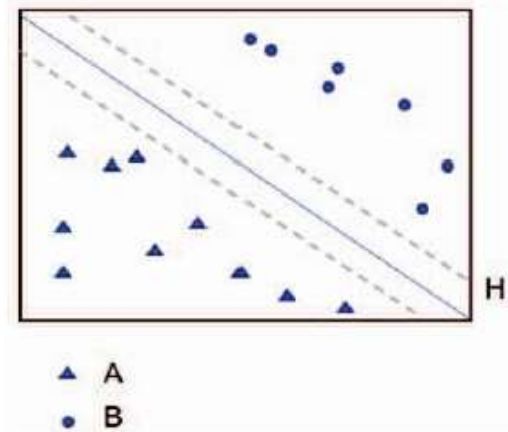


Table 2. Support Vector Machine model

Despite the stringent requirements, this approach continues to see extensive application in domains such as natural language recognition because to its relatively simple premise and high prediction efficiency.

**The k-means distribution**

**Part II: Getting Started with Machine Learning and Its Medical Care Applications**

Here, the scenario dictates the value of the variable k. Next, it takes the distance between each item and the cluster centers (k objects in total), sorts them into the nearest cluster based on that distance, and finally, assigns each object to a cluster. The centroid of each cluster will be recalculated until all items have been allocated. When the centroids remain unchanged, the procedure will be repeated. The programming is easier, and the method is iterative.

**ANN, or Artificial Neural Network**

**A. Diagnosing tumors with assistance**

An program that uses interconnected "neurons," or nodes, to mimic the way the human brain learns is known as an artificial neural network (ANN). The activation functions are represented by each node. Every connection between two nodes has a weight. Neural networks are categorized into two types: forward networks and feedback networks. To

..ISSN: 2040-0748

Vol-12 Issue-02 NOV 2023

Research into ML for tumor prediction, follow-up therapy, and similar purposes has increased with the growth of AI in healthcare. Cancers of the breast, lungs, and skin have recently been the focus of important research. Additional malignancies are also the subject of active investigation. As an example, support vector machines (SVMs) are often used in breast cancer therapies. SVMs classify tumor types as benign or malignant and then map them in three-dimensional space, simplifying difficult analysis in two-dimensional space. It selects the optimal hyper plane to divide the data every time. Using data mining on 1150 chromosomal pictures using SVM, CHEN et al.[1] discovered that breast cancer pathogenesis is influenced by cancer nest characteristics, cancer nest cell density characteristic, and stromal cell structure characteristic. When it comes to nuclear data, Chen and others have the best accuracy rate at 99%. After going through 3121 breast tissues, Jiang[2] used super hashing based on a joint kernel to bridge the semantic gap between low-level and high-level diagnosis. It identified benign instances in massive databases with an accuracy of 88.1% in retrieval and 91.3% in classification, all in 16.6 milliseconds of query time. When it comes to skin cancer, Jaworek[3] suggested a better way to differentiate melanoma. To prepare the skin pictures for analysis, he employed the dermoscopy technique to exclude any elements unrelated to the skin problems. He proceeded to extract and segment photos before analyzing color attributes.

At the same time, he rendered textures using gray-level co-occurrence matrices. An accuracy rate of 92% was achieved while evaluating the extent of skin cancer. There are also a lot of studies looking at it from many sides, including clinical viewpoints and DNA viral data. With ML, we can make predictions with an accuracy of about 85%, and we get really good outcomes.

### Part B: Uses in medical imaging

Most individuals aren't satisfied with the efficiency of medical imaging examinations and the outcomes that are relevant in this day and age when medical resources are limited. What this implies is that medical imaging may benefit substantially from the adoption of ML in terms of efficiency and labor savings. Medical professionals have increasingly turned to ML for help with medical image processing tasks such as CT segmentation and MRI interpretation. Regarding ultrasound detection, Zhu[4] suggested in his study that a neural network method may be used to distinguish between thyroid nodules that are benign and those that are cancerous.

This study included 618 participants with 689 thyroid nodules; all participants had a clean medical history free of thyroid illness, neck radiation, and ultrasounds. A neural network was built using the extracted morphology, margin, echo, internal combination, calcification, halo sign, and color Doppler vascular properties of each nodule. Numerical values of 0 and 1 were utilized to denote benign and malignant nodules, respectively. The research limited the aforementioned six eigenvalues and associated attributes for a total of 561 nodules in an effort to prevent overtraining. The interneurons' weight is determined by the discrepancy between the predicted and actual values of the output.

Using a gray forecast model to segment MRI images more accurately is something that HUANG[5] suggests. The following is the formula he used to construct the G(1,1) model and generate the original sequence using the data he gathered:

$$dX_1(i)dt + \frac{1}{\alpha} X_1(i) = \tilde{u} \quad (2)$$

The control coefficient, expansion factor, and starting projected value are all variables that can be solved using the least squares approach.

$$X_1(i) = \left[ X_0(1) - \frac{\beta}{\alpha} \right] e^{-\alpha(i-1)} + \frac{\beta}{\alpha} \quad (3)$$

The experiment shown that this technique can easily detect the tumor location, decrease false-positives, and increase the accuracy of tumor segmentation while also reducing the unpredictability of the original data. Using a convolution neural network (CNN) and the structure of LeNet-5, Sarraf[6] evaluated 62335 photos; 52507 of these images were associated with Alzheimer's disease, and the accuracy reached 98.85%. Dou[7] used a three-dimensional full convolution neural network (FCN) to identify cerebral micro bleeds (CBMs). With a sensitivity of up to 93.16 percent, this finding may drastically cut down on unnecessary computations and dramatically increase the detection speed. It is possible to apply this strategy to other medical tests in theory as well. Numerous studies on cancer MRI and liver fiber CT have used ML, in addition to the aforementioned uses. Clustering techniques, ANN, and support vector machines are among the most popular approaches. Which are produced from the primary ML algorithms also progressively broaden up the possibilities of researchers.

### III. APPLICATIONS OF MACHINE LEARNING IN PRE-DIVISION AND REFERRAL

#### A. Conventional and alternative approaches to healthcare

At the moment, most studies investigating the use of ML in healthcare are concentrating on symptom assessment and the enhancement of associated medical metrics. Reducing medical resource investment and avoiding subjective mistake caused by human assessment are both possible outcomes of relevant research. The current standard of care, taking into account both new knowledge and the results of previous studies, involves registering with the hospital and then visiting the appropriate clinic, where one must undergo the questions and examinations of the department physicians in order to have a first assessment. Changing departments is still an option in many circumstances.

As a result, we may want to consider a fresh approach that uses less medical resources overall, as well as one that uses machine learning to categorize patients based on their pre-treatment data. Patients will go straight to the appropriate department after the examination, rather than consulting their initial doctor for recommendations. Here are the differences and similarities between intelligent medical procedures and the conventional ones: become different. In the end, the percentage was 84.3%.

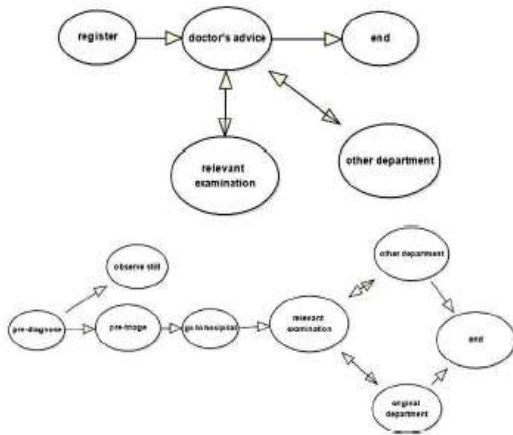


Figure 3: Old and modern approaches to healthcare

**B. Handling Data** In order to guarantee the most accurate diagnosis, it is necessary to build a database that contains all relevant patient records, define the target data set, preprocess it (this step can consume 60% of the energy), convert the data to identify valuable features, and finally mine the data.

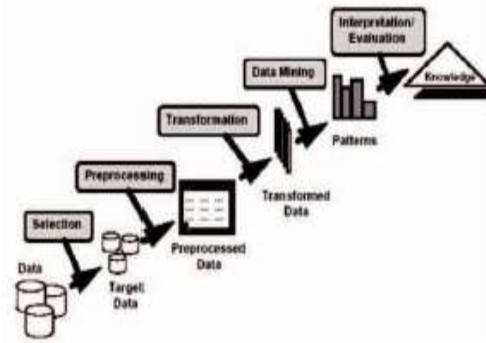


Figure 4: Data Mining Process Steps [8]

The first step in data pre-processing is to add historical symptoms and cases into a database. To account for the possibility of case defects, pre-processing uses a decision tree approach that is insensitive to missing data. Using an impurity-based version of the entropy formula, choose which node to separate:

$$E(t) = - \sum_j p(j|t) \log_2 p(j|t) \tag{4}$$

$$G_{split} = E(p) - \sum_{i=1}^k \frac{n_i}{n} E(i) \tag{5}$$

The node that is now separated from the others has the highest G value of all the nodes. The best node may be located and categorized based on important features using this traversal. It may sort patients into different categories throughout the referral and assessment processes according to how similar they are, using a high-dimensional data space to quantify this similarity. In this section, the Jaccard coefficient is presented as a measuring standard. The Jaccard similarity coefficient shows how much overlap there is between two sets relative to how much overlap there is between the two sets.

**A preliminary cutoff is established for the first categorization.**

Let me give you the formula:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{6}$$

Following this, we compute the Jaccard and finally, we create a similarity coefficient matrix from the similarity coefficient. The similarity matrix coefficient is added up for every row, the highest total is saved as a variable called max, and thereafter, the element on the line and any element bigger than are sought for. Keep track of the element's row number and the column in which it appears; then, divide the two into the new matrix's rows and columns.

You have to build a new matrix, calculate the row again, and then evaluate it.

Otherwise, it begins from the sum of the rows in the second search matrix unless all of the elements are bigger than , in which case it returns to the starting state.

$$sim(i,j) = \frac{|A \cap B|}{|A \cap B| + |A \cap \bar{B}| + |\bar{A} \cap B|} \quad (7)$$

$$\begin{pmatrix} sim_{11} & \dots & sim_{1n} \\ \vdots & \ddots & \vdots \\ sim_{n1} & \dots & sim_{nm} \end{pmatrix} \quad (8)$$

In order to categorize the patient's disease, the similarity matrix analysis technique is utilized. The pertinent data is then communicated back to the patient and medical staff. This allows them to successfully execute intermediate referrals, which significantly reduces the investment in personnel.

**IV. PROBLEMS AND CHALLENGES**

The field of machine learning has been booming in popularity as of late, and its medical applications have been expanding at a steady clip. There are, however, bound to be a number of constraints. Here are some of the most prevalent issues:

- (a) Medical image analysis still has a ways to go before it can fully eliminate human error, which is a major drawback when dealing with sensitive patient data;
- (b) Despite machine learning's promising track record in cancer treatment, there is a significant lack of resources both financial and human to fully develop and implement this technology on a large scale. It is still not up to snuff when it comes to this criterion;
- (c) While there have been advancements in pathological analysis, these advancements are not relevant to the most prevalent illnesses that

necessitate human resources, such fever and analgesia.

**V. CONCLUSION REMARKS**

Following an overview of machine learning's background and contemporary uses in healthcare, this article delves into its primary methodologies and outlines a number of illustrative examples. Here are the most common concepts and algorithms outlined. Also offered is an approach to visitation that makes use of machine learning for enhancement. Nevertheless, ML is far from flawless. There are issues with it on many fronts, including technology, ethics, and legality. Technicians and legal staff are needed to resolve these issues. Collaborating and finding the right mix of human and mechanical labor is an issue that we all have to deal with.

**REFERENCES**

[1] G. Eason, B. Noble, and I.N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529-551, April 1955. (references)

[2] Jiang M, Zhang S, Huang J, et al. Scalable histopathological image analysis via supervised hashing with multiple features[J]. *Medical Image Analysis*, 2016, 34:3-12.

[3] Joanna J K, Pawel K . Automatic Classification of Specific Melanocytic Lesions Using Artificial Intelligence[J]. *BioMed Research International*, 2016, 2016:1-17.

[4] Lu-Cheng, Zhu,Yun-Liang, Ye,Wen-Hua, Luo,Meng, Su, Hang-Ping, Wei,Xue-Bang, Zhang,Juan, Wei,Chang-Lin, Zou.A model to discriminate malignant from benign thyroid nodules using artificial neural network.[J].*PloS one*,2013,8(12):e82211.

[5] Huang W C , Chang C P . Automatic Nasal Tumor Detection by grey prediction and Fuzzy C-Means clustering[C]// *IEEE International Conference on Systems. IEEE*, 2006.M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science, 1989.

[6] Sarraf S , Tofighi G . Classification of Alzheimer's Disease using fMRI Data and Deep Learning Convolutional Neural Networks[J]. 2016.