A Study On Cardiac Disorders Using Machine Learning

Pradeep Kumar Verma , Research Scholar, Department of Computer Science & Engineering, Kalinga University Raipur (C.G.)

Dr. Nidhi Mishra, Department of Computer Science & Engineering, Kalinga University Raipur (C.G.)

Dr. Brijesh Kumar Bharadwaj, Associate Professor MCA, Dr. Rammanohar Lohia Avadh University Ayodhya U.P.

Abstract

Cardiac disorders are especially vulnerable if not diagnosed early on due to their greater criticality. The use of medical sensors and wearable devices has led to the development of new methods, such as electronic health records and body area networks, for the continuous monitoring and diagnosis of patients' health issues. Machine learning algorithms are employed for efficient health data classification processes due to the continuous and enormous number of data provided by body area networks. The classification of health data, however, presents the greatest difficulty because of the importance of precision and the requirement for early detection of cardiovascular disorders. Consequently, a major shortcoming of current systems is their inability to diagnose diseases early and with greater accuracy. The primary goal of this study was to suggest a framework that would decrease the dimensionality of medical datasets by omitting superfluous or redundant characteristics, hence increasing classification accuracy while decreasing computational time.

Index Terms

I Introduction

Disease can be clinically classified based on a thorough assessment of the data and then selected based on the sensitivity and specificity of the diagnostic features. The clinical, pathologic, and nosology research based on the frequency of occurrence, features, and risk factors in patients necessitate prospective with clinic-pathologic studies in a representative population of patients exhibiting Disease. Classifiers' accuracy in disease diagnosis performance has traditionally been measured using neural networks, D-M-neural, regression, and decision trees[1]. Damage to the voice from illness impairs not only communication but also motor skills and other cognitive processes. Using voice measurement for telemonitoring of the disease is a useful tool in early disease diagnosis. To evaluate the importance and relevance of the disease-attribute relationships, a predictive model can be constructed using Support Vector Machine (SVM) instead of the more traditional bootstrapping or leave-one-out validation approaches. Multiple-systems Parkinson's disease can be diagnosed thanks to the diagnostic and prognostic usefulness of a wide range of clinical characteristics. DISEASE's diagnosis model relies on accuracies such as ACC, KE, and AUC, calculated using K-Star and I-Bk as its basis classifiers. Diagnosing diseases and other medical conditions is a major area of use for medical biometrics[2]. Separating healthy participants from those with sickness using medical decision boundaries demonstrates the efficacy and computational efficiency of the mechanism of class borders. There have been predictions made using WEKA and MATLAB v7 about the accuracy that can be achieved by training a probabilistic neural network with the Disease dataset. There has been a lot of focus on using neural networks to make medical diagnoses and disease predictions in the last decade. There is a growing need for shrewdness and precision as the world's technology enters a purple phase. People in the modern era are likely hooked on the internet and care little about their physical wellbeing. People tend to put off going to the doctor or hospital, even for relatively minor issues, until they have reached a critical stage[3]. Experts hope to use this

expanding technology to create a system that can forecast many diseases based on the symptoms patients record without them having to go to the hospital or see a doctor. Machine learning, a branch of artificial intelligence, is concerned with learning algorithms, or programs that get smarter the more data and experience they are exposed to. There are two stages to machine learning, known as training and testing. The medical profession may take advantage of Machine Learning's effective platform to address a wide range of healthcare challenges at a much higher clip. Both supervised and unsupervised methods of machine learning exist. In supervised learning, we build a model with the help of labeled data. Contrarily, unsupervised learning models acquire knowledge from data that has not been labeled. The goal here is to infer a useful Machine Learning system that can reliably and accurately forecast the occurrence of disease[4].

AI AND MACHINE LEARNING

The process of selecting a choice, or decision-making, is referred to as a strategy. It's a way of thinking and making choices in response to specific situations. It's a laborious procedure that requires careful planning. Methods such as gathering data, analyzing it, making assessments, and drawing conclusions are all part of this process. The study of computer simulations to evaluate human thought processes and intelligent behavior is known as artificial intelligence (AI) (such as learning, reasoning, thinking, planning, etc.). Computer intelligence is the primary foundation of AI, which aims to give machines capabilities analogous to those of human brains. The finest decisions can be made with the help of AI and ML.

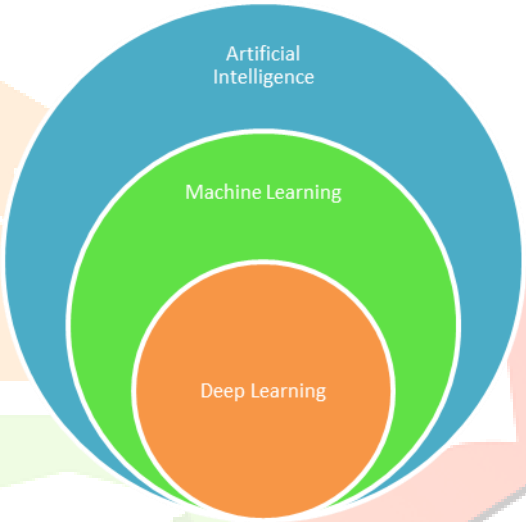

Fig 1 Differences

According to Simon, learning is "the process through which a person's behavior is modified and improved over time as a result of exposure to new information." The term "machine learning" is used to describe situations where computers are doing the "learning" described in the definition. During the machine learning process, the term "enhancement" refers to the act of producing the optimal answer in light of previously collected data[5]. The term "big data" was coined as a result of breakthroughs in computer and network technologies. Big data is not a new notion; it refers to massive, ever-growing collections of unstructured information that exceed the capacity of conventional database management systems. Internet applications, automated teller machines, credit card readers, and other similar devices generate massive amounts of data. Data obtained in this fashion is currently awaiting analysis. Analysis of data from various domains has varying goals depending on the type of company conducting the research. Natural language processing, image processing and computer vision, speech and handwriting recognition, automotive, aviation, production, energy generation, computational finance, and biological sciences are just a few of the industries that use machine learning applications. However, the guiding idea is founded on the analysis and interpretation of historical facts. Machine learning approaches and algorithms have been created to analyze and interpret this data because humans simply can't do it[6]. The paper covers the development of machine learning, its current methodologies and algorithms, and its current and potential future uses. Finally, a

summary of the research done up to this point is presented. Genetic algorithms, cellular automata, and agent-based models are the most popular AI and learning machine tools for decision making. Applications of AI and machine learning-based decision making can be found across many disciplines, including IT, engineering, physics, math, and medicine. Mathematical system theory, fuzzy logic, and fuzzy rules are examples of this mode of decision-making. Decisions are powered by AI and ML, allowing professionals in a wide range of domains to make accurate predictions and take informed actions. Artificial intelligence (AI) and machine learning (ML) play pivotal roles in a wide range of industries and market niches.

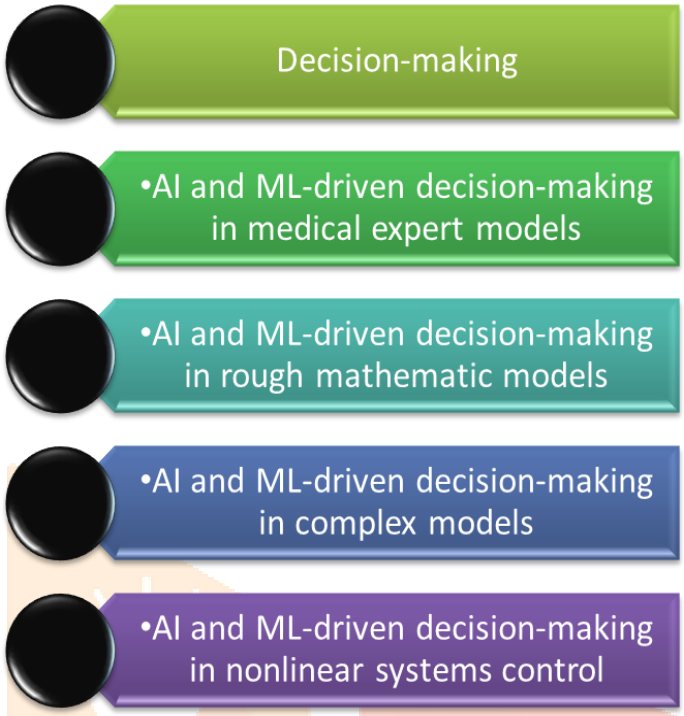

Fig 2: AI and ML behaviour for various domains

Machine Learning Algorithms

Machine learning is used to train computers to perform better using data. Even after looking at the data, we may be unable to deduce any useful information from it. Machine learning aims to achieve this through acquiring knowledge through observation of data[7]. The topic of teaching machines to learn on their own has been the focus of a lot of research. Numerous mathematicians and computer scientists employ various strategies in order to solve this puzzle. Fig. 2 depicts an example of one such method.

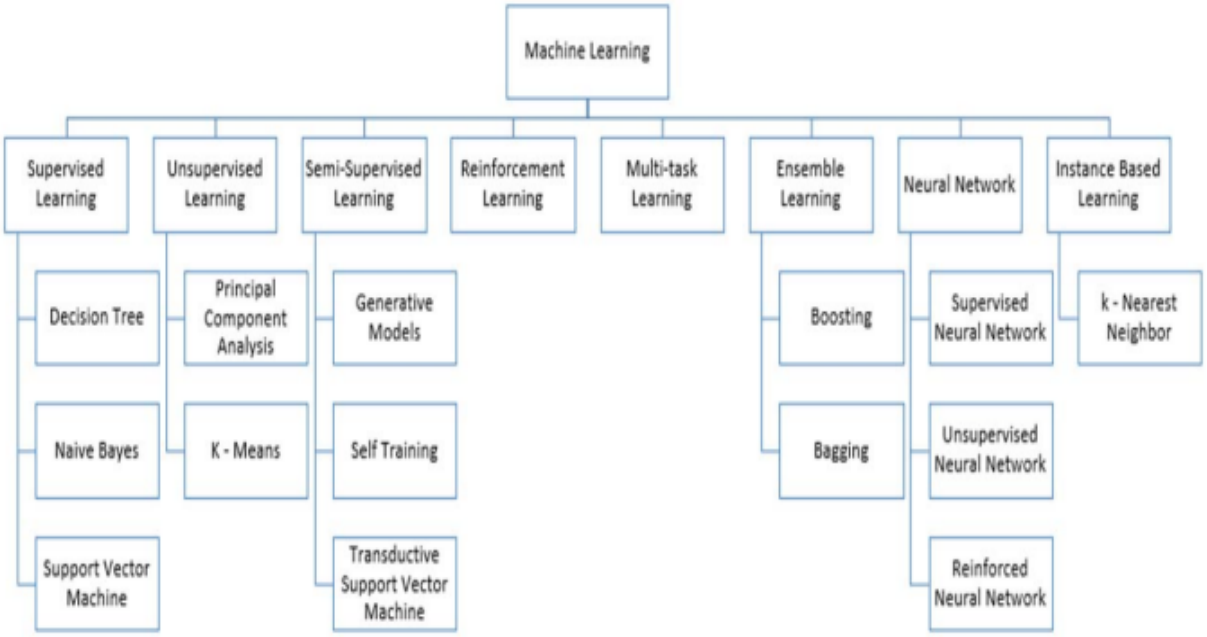

Fig 2 Machine Algorithms for Prediction

Supervised Learning

Learning a function that maps an input to an output is called supervised learning, and it is a machine learning activity. It does this by inferring a function from a set of training examples that have been labeled. A subset of machine learning algorithms, supervised algorithms rely on outside supervision to get results. There are two types of input datasets: the train dataset and the test dataset. It is necessary to make predictions or categorize data based on an output variable found in the train dataset. To make predictions or classify data, all algorithms first learn patterns from a training dataset.

Unsupervised Learning

Unsupervised learning refers to the aforementioned scenarios in which neither a right response nor an instructor is provided. The onus is on the algorithms themselves to unearth and convey the data's intriguing structure[8]. Most data features are ignored by unsupervised learning algorithms. In order to determine the category of newly introduced data, it draws on previously learned features. Clustering and feature reduction are two of its most common applications.

Logistic Regression

It is widely used as a supervised machine learning algorithm. It is put to use in order to foretell the various possible outcomes regarding a specified variable. There would be only two categories for a target's attributes. Logistic regression is a type of regression analysis where the results are either 1 (indicating success) or 0 (indicating failure) and the data is represented by a simple 1 or 0. As a mathematical function of X, this model makes the prediction of Predict(Y=1). As a result of the ease with which it can forecast a dependent variable of interest, It's been put to good use in diagnosing and treating everything from cancer to lung failure. Predictions using the binary logistic regression with dependent variables can be made using two additional categories: binomial and multinomial, as well as ordinal. It also lays out guidelines for how to achieve accurate predictions, such as using a sizable training set and a meaningful independent variable.

Decision Tree

It's a predictive modeling method for making decisions that can be used in many contexts. A decision-building strategy, which makes it simple to partition the dataset in several rational ways for making predictions under various scenarios, is a viable option for this purpose. As a supervised learning model, it is considered "one of the most powerful algorithms." The decision tree algorithm can be further broken down into two sun-like branches: classification tasks and regression tasks. In the context of decision tree modeling, the points at which information leaves and branches off in different directions according to predetermined

criteria are known as decision nodes[9]. Using the Gini index to implement a decision-tree method that generates splits.

Classification decision tree

In a non-regular Regression decision tree, the split occurs according to a decision variable that is given. The decision is based on the target variable, which is a continuous. Consistently engaged in physical activity Simple Bayes (SB): As a widely used classification method, it relies heavily on a key assumption of Bayes' theorem: namely, that all predictors and algorithms can be treated as completely separate from one another. In Bayesian theory, the posterior probabilities are the most sought after category. With the help of the Nave Bayes algorithm, we can quantitatively display the feature that it is stated that a probability of a label provided with some observed nature. The formula for calculating the probability of a certain outcome in a given set of data is as follows: Predict(L| f) = Predict(L)(f | L) / Predict (f) With the scikit learn library imported into Python, creating a naive bayes model has become more adaptable. Traditional uses for naive bayes include real-time prediction, multi-class prediction, text classification, and recommendation. Related complex-structures-reinforcement-platform-system.

Support Vector Machine (SVM)

In addition to being one of the most well-known supervised learning algorithms, it is also put to use in solving classification and regression issues. Machine learning practitioners typically employ SVM to tackle classification problems. The purpose of a support vector machine is to divide a space into classes using a single line of best fit, or a decision boundary. Since the precision is time-lined, we can confidently assign the newly acquired data point to the appropriate time period. Additionally, the term "hyperplane" is used to describe the limit at which a choice must be made. As a means of producing a hyperplane, this technique selects the extreme vectors/data points. Support vector machine is widely used for face detection, image classification, and text classification thanks to its ability to create these types of classes[10].

Proposed Steps

We propose a system with these traits: it will be fast and easy to use, saving both time and money. By facilitating communication between medical professionals and patients, our suggested method benefits everyone involved. Diseases are predicted using this system based on symptoms.

The evolution of machine learning and its medical applications show how new methods and approaches have arisen to do complex data analysis with very elementary programming of machine learning algorithms. In this section, we give the results of an exhaustive analysis of algorithms' performance on a medical record, with some of them achieving an accuracy of 95% or higher. The results are evaluated using a confusion matrix and an accuracy score. Due to the massive amounts of data that can be collected and saved with today's technology, AI will play an increasingly crucial role in data analysis in the near future[11]. In figure 3, we describe a system for prediction that employs a machine learning technique.

In order to extract useful pieces of information from the specified data in healthcare communities, biomedical fields, etc., machine learning environments have begun to employ a number of well-known machine learning techniques that have been developed and exploited in a wide variety of real-world application areas (such as industry, healthcare, and bio science). Correctly analyzing medical databases helps in early disease prediction, better care for patients, and expanded community resources. Disease prediction is only one area where machine learning techniques have been successfully implemented.

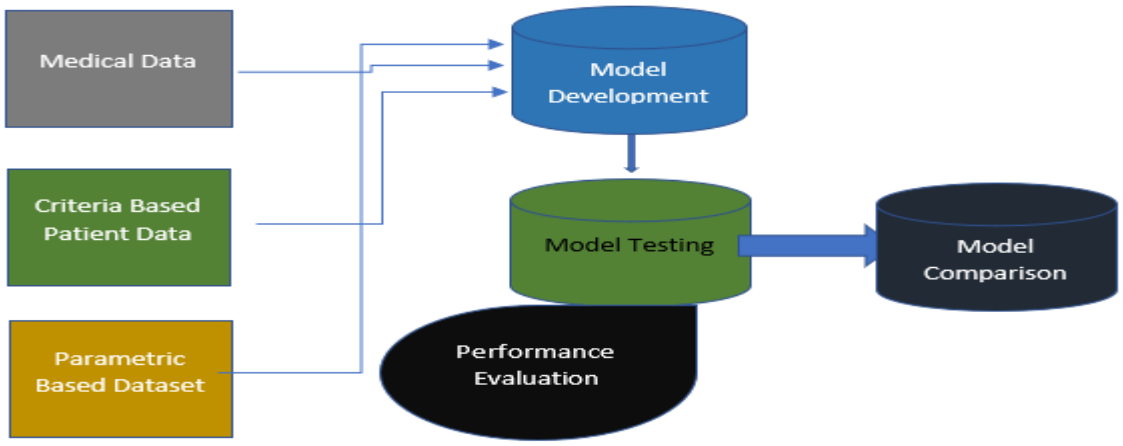

Fig 3: Prediction Concept for Evaluations

Conclusion

The process of disease diagnosis and severity evaluation poses enormous obstacles to the health care system, which in turn affects millions of people on an annual basis. Early detection is still a difficult effort, despite the great progress that has been made in medical science, which is why the disease cannot be prevented or treated in a timely manner. It is vital to create speedier and more efficient ways due to the complexity of the problem, the rising expense of the diagnostic procedures, and the lack of availability of suitable resources in certain locations, particularly rural ones. The creation of a machine learning model for the diagnosis could assist medical professionals in refining their knowledge for the stratification of risk of diseases concerning obscure information and classifying patients into different health patterns for the purpose of developing more effective treatments.

References

1. Sırmaçek, B. (2007). Fpga İle Mobil Robot İçin Öğrenme Algoritması Modellenmesi. Yüksek lisans tezi, Yıldız Teknik Üniversitesi, İstanbul.
2. Albayrak, A. S. ve Yılmaz Koltan, Ş. (2009). Veri Madenciliği: Karar Ağacı Algoritmaları ve İMKB Verileri Üzerine Bir Uygulama. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 14(1), 31-52.
3. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, Mahendiran T, Moraes G, Shamdas M, Kern C, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Dig Health. 2019;1(6):e271.
4. Panayides AS, Amini A, Filipovic ND, Sharma A, Tsaftaris SA, Young A, Foran D, Do N, Golemati S, Kurc T, et al. AI in medical imaging informatics: current challenges and future directions. IEEE J Biomed Health Inf. 2020;24(7):1837.
5. Oktay O, Ferrante E, Kamnitsas K, Heinrich M, Bai W, Caballero J, Cook SA, De Marvao A, Dawes T, Oregan DP, et al. Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation. IEEE Trans Med Imaging. 2017;37(2):384.
6. Jeyaraj PR, Nadar ERS. Deep Boltzmann machine algorithm for accurate medical image analysis for classification of cancerous region. Cogn Comput Syst. 2019;1(3):85. doi: 10.1049/ccs.2019.0004.

7. Harmon SA, Sanford TH, Xu S, Turkbey EB, Roth H, Xu Z, Yang D, Myronenko A, Anderson V, Amalou A, et al. A systematic review of antibody mediated immunity to coronaviruses: kinetics, correlates of protection, and association with severity. Nature Commun. 2020;11(1):1.
8. 18. Zheng Y, Jiang Z, Zhang H, Xie F, Shi J, Xue C. Adaptive color deconvolution for histological WSI normalization. Comput Methods Progr Biomed. 2019;170:107.
9. 19. Hoffman RA, Kothari S, Wang MD. In: 2014 36th annual international conference of the IEEE engineering in medicine and biology society, IEEE; 2014. pp. 194–7
10. Currie G, Hawk KE, Rohren E, Vial A, Klein R. Machine learning and deep learning in medical imaging: intelligent imaging. J Med Imaging Radiat Sci. 2019;50(4):477. doi: 10.1016/j.jmir.2019.09.005.
11. Xia W, Chen EC, Peters T. Endoscopic image enhancement with noise suppression. Healthcare Technol Lett. 2018;5(5):154.