ISSN: 2633-4828 Vol. 5 No.3, September, 2023

International Journal of Applied Engineering & Technology

EXAMINING MACHINE LEARNING'S POTENTIAL FOR MEDICAL DIAGNOSIS AND TREATMENT

[1]Mr. Ketankumar C. Patel and [2]Dr. Satish Narayan Gurjar
[1]Ph.D. Scholar, Department of Computer Engineering, University of Technology-Jaipur, Rajasthan
[2]Principal, Samarth College of Engineering, Belhe, Pune, Maharashtra, India

ABSTRACT

There is a lot of promise for machine learning in the field of medical diagnosis and therapy. Large data sets can be analysed by it, and it can spot patterns that humans might find difficult to see. The revolutionary potential of machine learning (ML) in medical diagnosis and treatment is examined in this research. It draws attention to the advantages, difficulties, and potential uses of ML in healthcare today. The paper seeks to provide a thorough grasp of how machine learning (ML) is transforming medical practices and increasing patient outcomes by evaluating current breakthroughs and case examples. Through the investigation of privacy issues, data security, and possible biases in algorithmic decision-making, the research can aid in the creation of policies and procedures that guarantee the ethical and responsible application of machine learning algorithms. This can reduce any hazards related to the use of machine learning in healthcare settings and foster trust between patients, healthcare professionals, and legislators.

Keywords – Machine Learning (ML), Medical Diagnosis, Medical Treatment, Healthcare, Decision making; Disease Diagnosis, Medical disciplines Artificial Intelligence (AI), etc.

I. INTRODUCTION

A new age of innovation has been ushered in by machine learning (ML), with one of the most potential applications being in the healthcare industry. As a branch of artificial intelligence (AI), machine learning uses statistical models and algorithms to teach computers to carry out certain tasks without explicit instructions, instead relying on patterns and inference. The progress in technology has led to notable advancements in medical diagnosis and treatment, with the potential to increase healthcare services' precision, efficacy, and customization. Fundamental to a medicine's function in society, diagnosis is a means of categorizing it. In the medical system, it is essential. It classifies illness by defining treatment options, projecting outcomes, and providing a useful tool [1]. A complete diagnosis is often necessary for appropriate and successful treatment. The entire diagnostic procedure has undoubtedly improved as a result of the concomitant advancements in diagnostic imaging and testing. However, even in this age of swift technological advancement, the human technique of scientific judgment leading to accurate diagnosis is still essential to provide high-quality and healthful medical services [2]. Nonetheless, patients are harmed by diagnostic errors on a regular basis. Diagnostic errors are often the result of several variables, including both perceptual and system-related issues. A number of common causes include misinterpreting data, misjudging the significance of observations, making mistakes when applying heuristics, and making judgment calls, especially when developing and evaluating diagnostic hypotheses [3-5].

Medical practitioners' knowledge and experience have always been crucial to medical diagnosis and treatment. The enormous amount of medical data that physicians must process, human error, and skill level fluctuation are some of the drawbacks of this technique, despite the fact that it has produced significant improvement. By using enormous datasets to find patterns and insights that might not be immediately obvious to human observers, machine learning offers a potent answer to these problems. ML can evaluate medical imaging, forecast disease outcomes, create individualized treatment regimens, and even help with drug discovery with to complex algorithms. The likelihood of misdiagnosing an ailment that is easily treatable is significantly increased due to the increasing efficiency and cost of treatment alternatives. Improved patient treatment is therefore lost [6]. By reducing these diagnostic errors, healthcare services could be enhanced through the use of methods like machine learning (ML) and fuzzy logic [7]. When it comes to treating patients, the type of analytics a doctor can obtain with ML can give them more knowledge and, consequently, better care [8]. The question of how to create systems that continuously learn from experience is one that machine learning attempts to address. Situated at the intersection of analytics and computers, and at the core of artificial intelligence (AI) and data science, it is regarded as one of the fastest-growing technical fields in existence today [9]. Markets that have been able to access such data and hire the labour required to transform their products remain the main beneficiaries of the 21st-century growth in big data, machine learning, and data science. Given that physicians use electronic health records (EHRs) on a regular basis, the algorithms developed for and around these markets provide significant promise for advancing clinical and medical research. Healthcare machine learning integration is already showing revolutionary results. For example, ML algorithms in radiology have been created to help with the remarkably precise analysis of medical pictures to aid in the early diagnosis of diseases like cancer. Personalized therapy and early intervention are made possible by the use of machine learning (ML) in genomics to find genetic markers for disease. Furthermore, proactive and preventive care is made possible by the use of predictive analytics driven by machine learning to foresee disease outbreaks and track patient health in real-time [9-10]. The use of machine learning in healthcare is not without difficulties, despite its enormous promise. There are many obstacles to overcome, including those pertaining to data quality, privacy, ethical issues, and integrating ML systems into the current healthcare infrastructures. Furthermore, in order to fully utilize the potential of machine learning, further research and development are required due to the complexity of medical data and the requirement for interdisciplinary collaboration [11-12]. The purpose of this study is to present a thorough analysis of machine learning's potential for medical diagnosis and therapy. It will examine recent implementations, draw attention to advantages, tackle difficulties, and go over potential future developments. This study aims to provide insights into how machine learning is transforming medical practices and enhancing patient outcomes by examining current developments and case examples. We hope to demonstrate through our investigation how vital it is to carry out more study and work together in order to fully utilize machine learning in the healthcare industry.

There is a lot of promise for machine learning in the field of medical diagnosis and therapy. Large data sets can be analysed by it, and it can spot patterns that humans might find difficult to see. The medical industry is using machine learning in the following important areas:

- **Medical Imaging:** To aid in the identification of different ailments, machine learning algorithms may analyse medical pictures including MRIs, CT scans, and X-rays. Specifically, deep learning systems have demonstrated potential in precisely identifying anomalies and supporting radiologists in their interpretation. The effectiveness and precision of image-based diagnosis can be enhanced with the aid of these algorithms.

- **Disease Diagnosis:** By evaluating patient data, such as symptoms, medical history, test findings, and genetic information, machine learning can help with disease diagnosis. Machine learning algorithms are able to recognize patterns and risk factors linked to certain diseases by use of training on extensive datasets. This can help medical experts diagnose patients more accurately, forecast how a disease will progress, and suggest the best course of action.

- **Drug Discovery & Development:** The process of creating new medications is difficult and time-consuming. Large-scale biological and chemical data can be more effectively analysed using machine learning to find possible medication candidates. Additionally, it can forecast a medication's efficacy and possible adverse effects, which aids researchers in selecting the most promising candidates for additional testing.

- **Personalized Medicine:** Personalized Medicine is made possible by machine learning, which takes into account the unique characteristics of each patient, such as their medical history, lifestyle choices, and genetic information. Machine learning models can assist in identifying the best course of action, appropriate dosage, and possible side effects for each patient by examining these variables.

- **Predictive Analytics:** By utilizing patient data, machine learning can be applied to forecast illness outcomes and patient risks. Machine learning algorithms can find patterns and risk factors linked to certain diseases by examining historical data, such as medical histories and patient records. Healthcare professionals can use this information to identify individuals who are at high risk and to put preventive measures in place.

- **Treatment Optimization:** By examining how patients react to different therapies and drugs, machine learning can assist in improving treatment regimens. Machine learning models can offer recommendations for individualized treatment techniques that improve patient outcomes and minimize side effects by taking into account unique patient features and treatment outcomes from similar patients.
- **Remote Monitoring and Telemedicine:** Telemedicine and remote patient monitoring are possible with machine learning algorithms, which can identify irregularities in patient-reported data or vital signs. This makes it possible to identify health problems early and take appropriate action. Additionally, machine learning can help telemedicine platforms by aiding in patient triage, offering preliminary diagnosis, and recommending suitable course of action.

II. METHODS OF MEDICAL DIAGNOSIS AND TREATMENT

We carried out a systematic literature review (SLR), which entailed a thorough search for pertinent medical research articles and survey studies. A systematic literature review (SLR) is a dispassionate and tried-and-true procedure for locating, evaluating, and interpreting all relevant material pertaining to a field that is, at best, somewhat unpredictable [1]. Prior to completing the evaluation, it is necessary to have a thorough understanding of the methodology [4-5]. The search and selection procedures that were employed to extract the papers for this survey paper adhered to predetermined standard criteria, which are detailed in the subsections that follow. Medical diagnosis and therapy have found a new and powerful tool in machine learning (ML). The concepts and strategies used to incorporate machine learning into healthcare are examined in this study.

A comprehensive diagnosis is typically required for appropriate and successful treatment. The diagnostic procedure as a whole has undoubtedly improved due to the concomitant advancements in imaging and diagnostic testing. Even in this period of rapid technological development, however, the human approach of scientific judgment leading to proper diagnosis remains critical to high quality and healthy medical services [2]. Nonetheless, patient-harming diagnostic errors do occur often. Diagnostic errors are typically caused by a variety of variables, including both perceptual and system-related issues. Misjudging the significance of observations, interpreting data incorrectly, making mistakes when applying heuristics, and making judgment calls, especially when developing and evaluating diagnostic hypotheses, are some common causes [3-5]. There's a significant financial and health risk associated with misdiagnosing an easily treatable condition because treatment alternatives are growing more effective and costly. As a result, better patient care is lost [6]. Healthcare services could be improved by reducing these diagnostic errors through the use of methods like machine learning (ML) and fuzzy logic [7]. When treating a patient, a clinician can use machine learning (ML) to obtain analytics that give them additional knowledge and improve the quality of their care [8]. The question of how these systems may be created to continually evolve with experience is addressed by machine learning. Being at the core of artificial intelligence (AI) and data science, as well as the intersection of computers and analytics, it is regarded as one of the technical fields with the quickest rate of growth in existence today [9]. Until far, the main beneficiaries of the 21st-century surge in big data, machine learning, and data science have been the markets that have been able to access such data and hire the labor required to transform their products. The algorithms developed for these markets have a great deal of promise to advance clinical and medical research, especially if physicians are utilizing electronic health records (EHRs) on a large scale. Two fields that benefit from the application of ML approaches in the healthcare industry are diagnosis and outcome estimation [10].

In addition to handling different raw data combinations and applying context weighting, machine learning (ML) can assess the prediction ability of every potential factor combination to identify diagnostic and prognostic components [11]. For instance, ML models may diagnose aphasia speech type [12], urinary tract infections [13], and even breast cancer [14] based on clinical data, helping physicians with "second opinions." the capacity to handle massive data sets considerably larger than what is humanly possible, and then to efficiently transform that data into clinical knowledge that helps medical professionals plan and administer therapy, ultimately improving outcomes, lowering costs, and raising patient satisfaction. Machine learning (ML) is presently driving the development of guidelines for precision medication, treatment recommendations, and disease diagnosis [8]. The healthcare internet of things (H-IoT) is one application where machine learning (ML) capabilities are being used to assess and handle vast amounts of sensor-generated healthcare data [15]. As a result, a great deal of study has been done to determine its utility in the context of treating particular disorders. Therefore, the primary goal of this work is to conduct a systematic analysis of the trials involving the application of machine learning techniques to various medical fields and illnesses in order to identify any patterns or potential applications for these techniques in disease detection. According to **Myszczynska et al. (2020)**, machine learning can help with early diagnosis, medical picture interpretation, and the identification and creation of novel treatments. The automatic extraction of useful insights from a variety of high-dimensional data sources—each of which offers a unique perspective on illness—is a common thread among the various machine learning applications. In **Jamshidi et al. (2020)**, This article developed an artificial intelligence (AI) reaction to the pathogen. To achieve this, a few Deep Learning (DL) techniques have been demonstrated, such as Long/Short Term Memory (LSTM), Extreme Learning Machine (ELM), and Generative Adversarial Networks (GANs). It outlines an integrated approach to bioinformatics whereby various informational components from a range of structured and unstructured data sources are combined to create easily navigable platforms for researchers and doctors. **Alafif and others (2021)**, They provide an overview of the performance, tools, datasets, and AI-based ML and DL techniques. This survey provides a thorough summary of the most recent state-of-the-art approaches available to ML and DL researchers as well as the larger medical community. It also includes information on how data and ML can be used to improve COVID-19 status and encourage further research to prevent COVID-19 outbreaks. Future directions and problems are also described in detail. Machine learning (ML) and deep learning (DL) have a lot of success stories and are widely used in many aspects of daily life. **Shaheen (2021)** examined how the healthcare sector has traditionally embraced technological innovations early and profited greatly from doing so. Many areas of health, such as the development of novel medical therapies, the administration of patient data and records, and the management of chronic illnesses, are utilizing machine learning, a subset of artificial intelligence. The identification and diagnosis of illnesses and ailments that are otherwise challenging to diagnose is one of the most significant applications of machine learning in healthcare.

Machine Learning (ML) Techniques

- **Supervised Learning**
  - **Classification Algorithms:** These algorithms classify input data into specified classifications in order to diagnose diseases.
  - **Support Vector Machines (SVM):** Good at classifying complicated datasets in high-dimensional domains.
  - **Decision Trees:** Easy-to-understand models for classification applications.
  - **Random Forests**: An ensemble technique that combines several decision trees to increase accuracy.
  - **Neural Network:** Deep learning models that can handle big datasets with intricate patterns are called neural networks.
- **Regression Algorithms:** Predict continuous outcomes like the course of a disease or the length of time it takes for a patient to recover.
  - **Linear Regression:** The fundamental approach for forecasting results based on linear relationships is called linear regression.
  - **Logistic Regression:** When doing binary classification tasks, like forecasting the existence or absence of a disease, logistic regression is utilized.

- **Un-Supervised Learning**
   - **Clustering Algorithms:** Algorithms for clustering data are used to find patterns and put comparable data points together without the need for labels.
   - **K-means Clustering:** Data partitioning into k clusters using an easy-to-use and effective technique.
   - **Hierarchical Clustering:** Using hierarchical clustering, one may comprehend the structure of the data by creating a tree of clusters.
   - **Dimensionality Reduction:** Methods for cutting features without sacrificing important details.
   - **Image Recognition:** The Principal Component Analysis (PCA) method converts data into a space with fewer dimensions.
   - **T-Distributed Stochastic Neighbour Embedding (t-SNE):** Provides two or three-dimensional high-dimensional data visualization.
- **Deep Learning**
   - **Convolutional Neural Networks (CNN):** CNNs are excellent in identifying tumours and segmenting organs from medical pictures, such as MRIs and X-rays.
   - **Recurrent Neural Networks (RNN):** RNNs, or recurrent neural networks, are utilized in sequence data analysis, including time-series data processing from wearable technology and electronic health records (EHRs).
   - **Long Short-Term Memory (LSTM):** This kind of RNN is useful for modelling the evolution of a disease or forecasting patient outcomes because it can identify long-term dependencies in sequential data.
- **Reinforcement Learning**
   - By using patient feedback and data, reinforcement learning can be utilized to optimize treatment plans. It can help with tailored treatment plans or dose modifications.
- **Transfer Learning**
   - Transfer learning applies huge datasets to pre-trained models to refine them for particular medical purposes. When medical datasets are scarce, this method helps by enabling models to leverage insights from other fields.
- **Uses in Medical Diagnostics**
   - **Diagnostic Imaging Image Recognition:** This section describes how to use convolutional neural networks (CNNs) to examine medical images and identify abnormalities including lesions, cancers, and fractures.
   - **Predictive Analytics:** Segmentation is the process of locating and separating particular areas of interest in a picture, including organ borders or aberrant growths. Analysis that Predicts Disease prediction is the process of applying predictive models to estimate a patient's risk of contracting a particular disease based on patient data. Early detection is seeing any health problems before symptoms appear so that therapy and intervention can begin as soon as possible.
   - **Analysis of Genomes:** Variant calling is the process of using sequencing data analysis to find genetic variants linked to certain diseases. Understanding gene activity and how diseases are correlated with it through gene expression analysis.

Many areas of medical diagnosis and treatment, such as image analysis, disease classification, risk prediction, treatment planning, drug development, and patient monitoring, have made use of these methods and algorithms. It's crucial to remember that in order to guarantee patient safety and therapeutic relevance, the use of machine learning in healthcare necessitates rigorous validation, ethical concerns, and cooperation with medical personnel.

By evaluating how patients react to different therapies and drugs, machine learning can assist in improving treatment regimens. Machine learning models can offer recommendations for individualized treatment techniques that improve patient outcomes and minimize side effects by taking into account unique patient features and treatment outcomes from similar patients. Algorithms for machine learning can be used to remotely monitor patients and identify irregularities in patient-reported data or vital signs. This makes it possible to identify health problems early and take appropriate action. Additionally, machine learning can help telemedicine platforms by aiding in patient triage, offering preliminary diagnosis, and recommending suitable course of action.

Applications in Medical Treatment

- **Personalized Medicine**
  - **Treatment Recommendation:** Recommending individualized treatment recommendations based on prediction models and patient profiles is known as personalized medicine treatment recommendation.
  - **Dosage optimization:** figuring out how much medication is best taken by each patient in order to maximize effectiveness and reduce negative effects.
- **Medication Discovery**
  - **Molecule Screening:** Using machine learning methods, a large chemical library is screened to find possible drug candidates.
  - **Predicting Drug Interactions:** evaluating possible interactions between various medications in order to avert negative consequences.
- **Robotic Surgery**
  - **Surgical Assistance:** Using machine learning algorithms to process real-time data, this approach improves control and precision in robotically assisted surgeries.
  - **Procedure Optimization:** Procedure optimization is the planning and enhancement of surgical procedures with the use of predictive models.

III. IMPLEMENTATION OF ML IN MEDICAL DIAGNOSIS AND TREATMENT

Artificial intelligence (AI) has become a potent instrument in the medical field in recent years, revolutionizing the methods by which doctors identify, treat, and forecast patient outcomes. By incorporating AI approaches into healthcare procedures, the suggested solution seeks to improve decision-making accuracy and efficiency. AI is capable of offering insightful analysis, supporting diagnosis, proposing therapy options, and forecasting the course of disease by evaluating complicated medical data. [10] The gathering and combining of many sources of healthcare data is a key element of the suggested system. This comprises patient-reported outcomes, genomic data, wearable device data, medical imaging data, and electronic health records (EHRs). The system will collect extensive patient data for analysis using interoperable and secure data sharing techniques.

Use of Machine Learning Algorithms for the Detection and Prediction of Diseases

Numerous machine learning-based methods are used to identify the disorders. Most of them are listed from this review research below:

Support Vector Machine (SVM): - For further assistance, use a support vector machine (SVM). SVM has shown to be beneficial in a variety of classification scenarios. It counts the nodes located on the boundary between two classes in an attempt to find the best hyperplane. The difference in values between two groups is called a margin. A higher margin of safety increases a classification system's success rate. Support vectors are used to represent boundary information. Both regression and classification issues can be handled with SVM [4-5]. This method works well with both linear and nonlinear data sets. Several distinct kernel types are employed by the SVM method in a prediction model. These consist of sigmoid, polynomial, and linear radial basis function (RBF) kernels. To divide data points into two groups, SVM uses a high-dimensional space for qualities and chooses the best hyperplane. It works well with both small and large datasets, which are typically difficult to analyse. A very basic illustration of support vector machine (SVM) analysis of hyperplane data for diabetes detection can be found in Fig.1 [5]. The process of classifying data points into different classes using a Support Vector Machine (SVM) is shown in Fig.1. This machine learning model creates a hyperplane or series of hyperplanes in a high-dimensional environment. The idea of maximizing the margin between classes—which is essential to the SVM's efficiency in categorizing complicated datasets—is graphically represented in the figure.

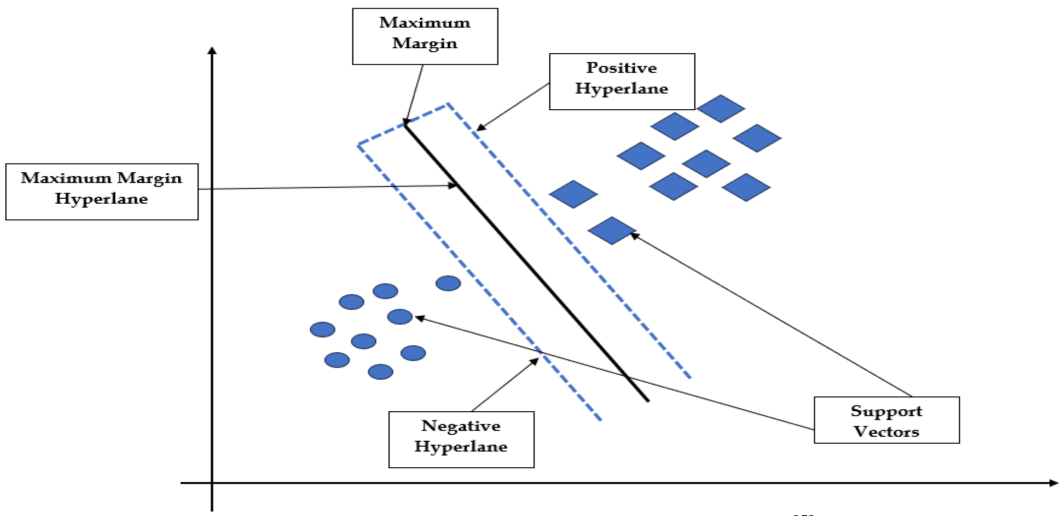

Fig. 1- Classification using SVM classifier [5]

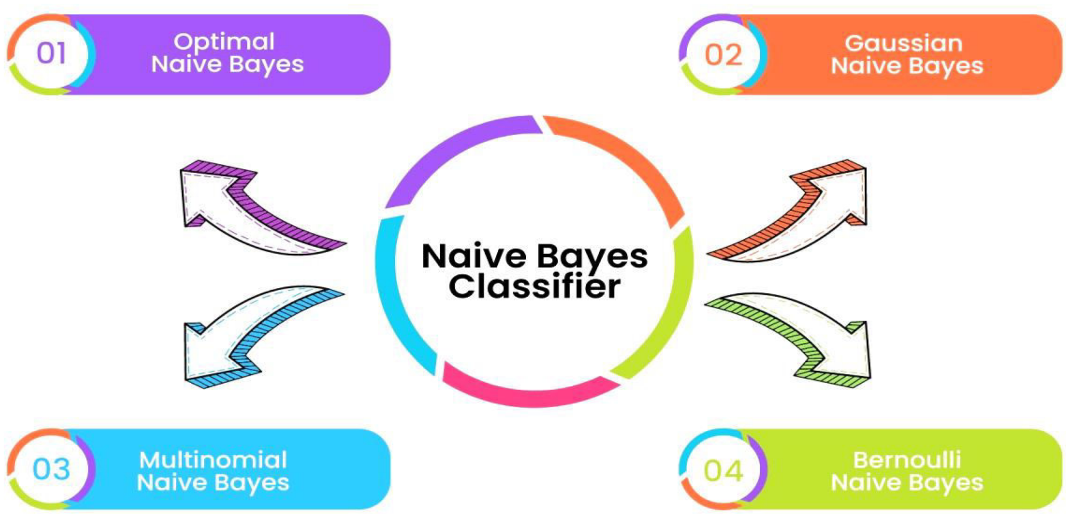

Fig.2- Naïve bayes classifier approach [5]

Naive Bayes Algorithm:- As shown in fig.2 for the Naive Bayes (NB) classification process used in probability and statistics. The technique is widely used in machine learning applications due to its user-friendliness and ability to weigh all relevant aspects equally before making a final decision. The simplicity and adaptability of the NB technique stem from its high processing efficiency. Three fundamental ideas underpin the NB classification: class conditional probability, posterior, and prior [5]. Apart from its numerous advantages, this technique works especially well with large datasets and is simple to use. It could be used to solve multi-class and binary classification issues. Both discrete and continuous data can be used, and less data is needed for training. By using this method, unwanted mail could be removed and papers could be categorized [4-5]. The Naïve Bayes Classifier method, a probabilistic machine learning model that uses the assumption of predictor independence and Bayes' theorem to predict the category of a given sample, as demonstrated in Fig.2. The process by which this classifier determines the conditional probability of each characteristic within those classes as well as the probability of each class is graphically depicted in the image.

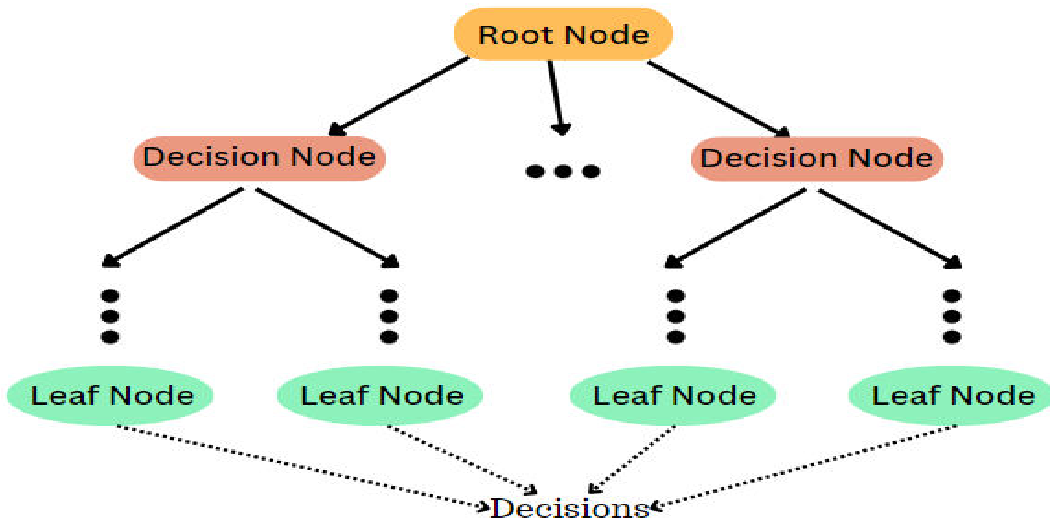

Fig. 3- Classification using a simple decision tree approach [5]

Decision Trees Algorithm: - Periodically dividing data according to a specified variable is how the Decision Tree (DT) addresses problems with regression and classification. Data structures are represented by nodes, while evaluations are represented by the leaf of the tree (Fig. 3). Using training data, decision trees are used to build models for forecasting dependent variables by teaching them basic decision rules [5]. By building the tree from the training set of data, the tree is trained. One type of non-leaf node is a decision node, where the class name is kept. The decision tree can be used to process both category and numerical data. The nonlinear relationship between arguments has no effect on the tree's efficacy [3,4,5]. There is no need to prepare the data beforehand. Overfitting could arise from repeatedly building the tree. The decision tree seen in Figure (3) is made up of the parent node of the tree, one child node, and one leaf node. The medical industry has made use of decision trees. For example, if we utilize DT to diagnose breast cancer, we might label each leaf node on the tree as benign or malignant. By creating rules among the required data set components, lump thickness (CT) will be utilized to determine whether or not the tumour is malignant. The DT algorithm is demonstrated in a breast cancer diagnostic situation in Fig. (3) [5]. The classification process utilizing a basic decision tree approach is demonstrated in Fig. (3), which emphasizes how this strategy divides data into subsets according to the input feature values, resulting in a decision tree-like model. The picture walks the user through the logical phases of data categorization by showing the hierarchical structure of decision nodes and leaf nodes, which stand in for the queries and results, respectively.

Classification by Random Forest: - Another application of the random forest ensemble model is the creation of a predictor based on the closest neighbours. The idea behind ensemble approaches is that combining numerous models produces a more robust model than any of the individual models. Apart from the random forest, the ensemble also consists of a decision tree, which is a conventional machine learning technique. After receiving a starting input, this process divides the data into ever-finer subgroups. This idea is expanded upon by the random forest, which combines the notion of an ensemble with trees. The advantages of a random forest classifier are its inherent inequality, speed of execution, and flexibility in handling missing data [5-6]. The freshly created subtrees in a random forest move with the new dataset or the testing data. The categorization of the dataset can be determined using any decision subtree in the forest. The model will choose the alternative that receives the greatest support after a vote. Fig. (4) [5] illustrates the basic concept underlying the random forest method for the identification of heart disease. Random Forest is an ensemble learning method that builds many decision trees during training and outputs a class that is the mode of the classes of individual trees. Fig. (4) shows how this method is used to categorize datasets. This graphic illustrates how predictions from different decision trees are combined to improve classification accuracy and reduce overfitting, which is a drawback of the Random Forest method.

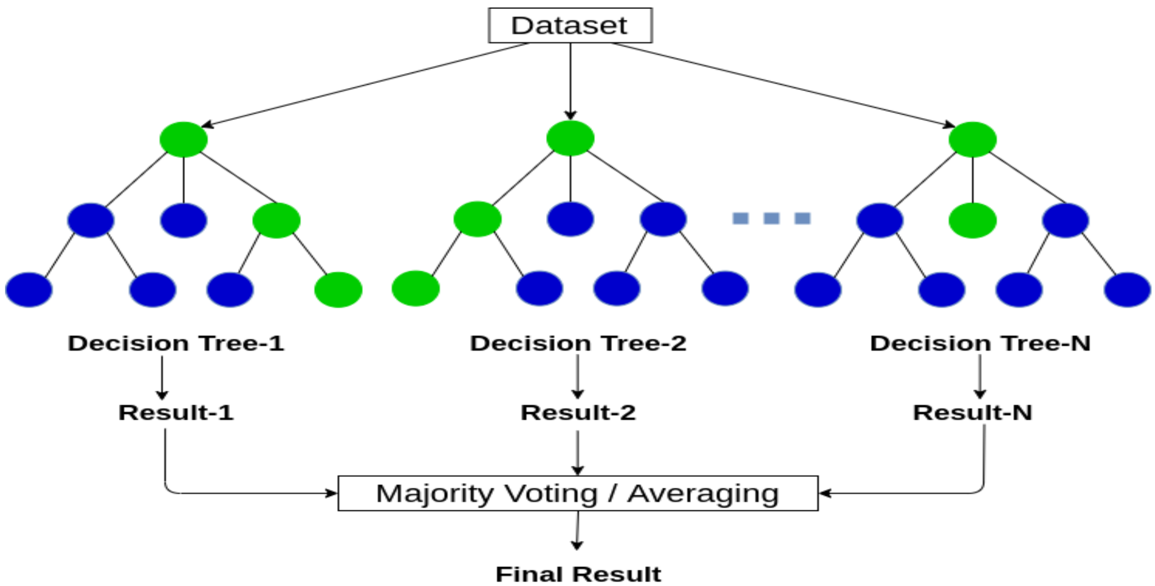

Fig.4- Use of random forest for categorizing datasets [5]

Logistic Regression: - A controlled learning process called logistic regression is employed to address double classification problems. Science uses the logistic relapse and logistic capability, two concepts with more complex extensions, to demonstrate a twofold order. In its most basic form, logistic relapse is a type of relapse model that forecasts the likelihood that a specific piece of information or data point belongs in a specific class [5]. The sigmoid function (Fig.-5) is used in logistic regression to model the data. Three key characteristics of logistic regression are its ease of application, computational economy, and regularization simplicity. It is not necessary to scale input features. On the other hand, overfitting and the capacity to address nonlinear issues [5].

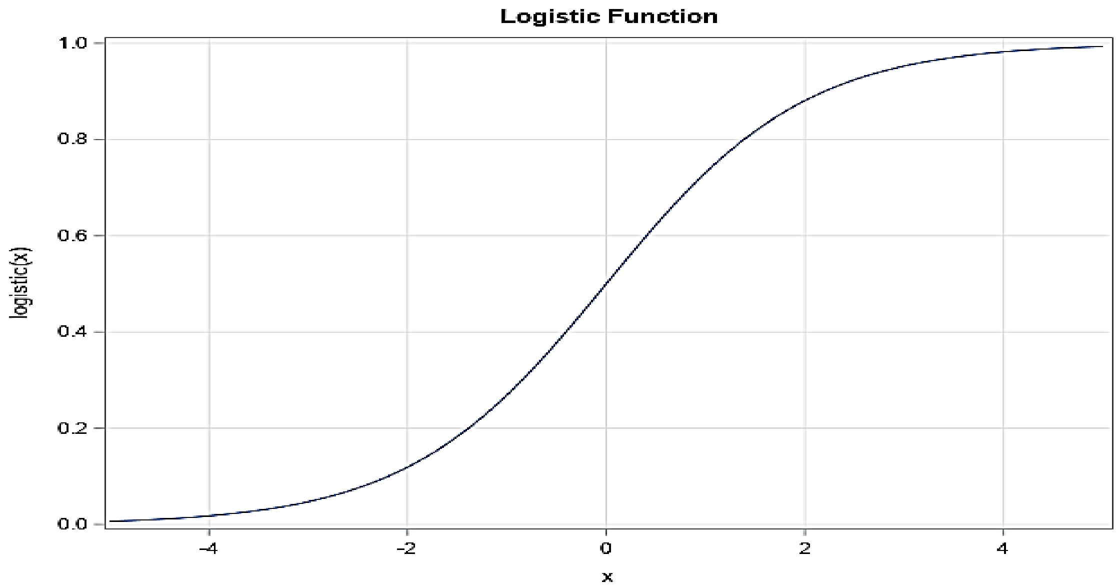

Fig.5- Logistic function [5]

Deep Learning: - Deep learning is a branch of machine learning that stands out for using multi-level layered nonlinear transformations. The input slices, which are essentially sequential data, can be used by the Recurrent Neural Network (RNN) to infer associations between slices [5]. Convolutional neural networks (CNNs) are widely used for complex classification tasks and data analysis in deep learning applications. Convolutional, max-pooling, completely connected, and output layers make up its total of four layers. The CNN model is fed a 32x32 pixel segment of the original medical image. Convolutional layers in deep neural networks generate feature maps, which are then transferred to pooling layers in the subsequent layers to minimize their size. As stated in a different study, the fully connected layer ultimately produces a prediction for the right class.

IV. PROPOSED APPROACH

Metrics for Assessing Machine Learning (ML) Models

There are a number of widely used evaluation measures when assessing machine learning models for medical diagnosis and treatment. The particular task and the type of data being used will determine which metrics are best. The following assessment metrics are often applied in medical diagnosis and treatment:

- **Accuracy:** By dividing the number of correctly classified examples by the total number of instances, accuracy determines how accurate the model's predictions are overall. On the other hand, in datasets with unequal distributions of classes, accuracy might not be adequate.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{N}$$

- **Precision:** Out of all cases projected as positive, precision determines the percentage of correctly anticipated positive instances. When reducing false positives is the goal, like in the case of cancer diagnosis, it is helpful since incorrectly diagnosing a patient as having cancer when they do not might have grave repercussions.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

- **Recall (Sensitivity):** The percentage of accurately predicted positive cases among all positive instances that actually occur is measured by recall. It is helpful in situations when reducing false negatives is the main objective, like in disease detection, where it can be harmful to overlook a positive case.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

- **F1 Score:** The harmonic mean of recall and precision is the F1 score. It is helpful when both false positives and false negatives need to be taken into account and offers a balance between the two measures.

$$F = 2 * \frac{Precision_p * Recall_p}{Precision_p + Recall_p}$$

Case Study of Machine Learning and Lung Cancer Diagnosis

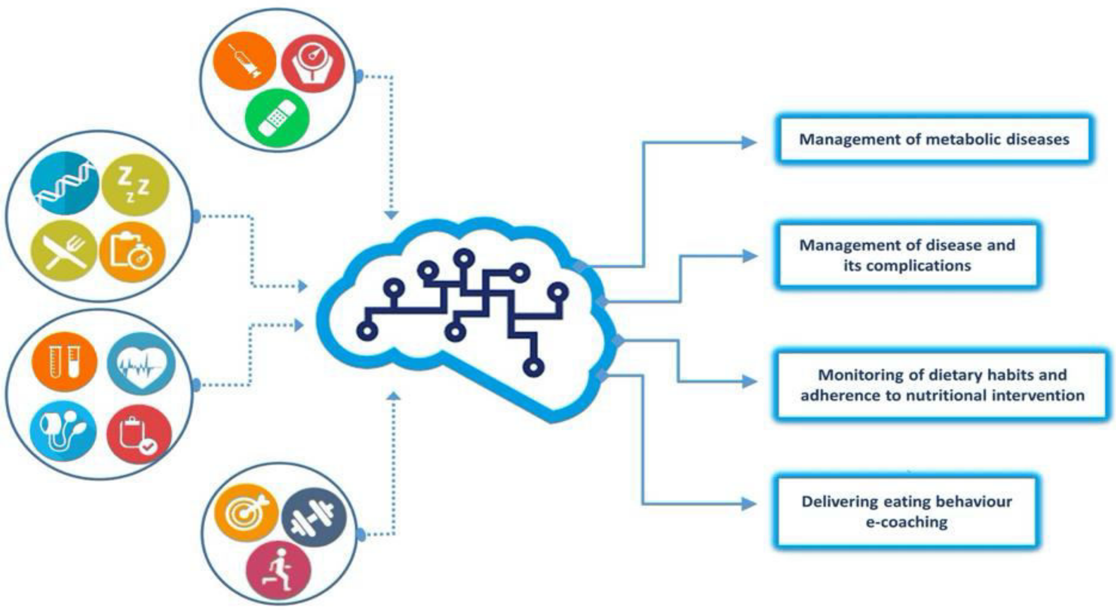

Fig. 6 Machine Learning and Lung Cancer Diagnosis [Source: Wikimedia Commons]

The article emphasizes how, akin to the early days of the Internet, numerous computer science and software sectors, including robots, augmented reality, and machine learning, have been garnering a lot of attention. Particularly machine learning has attracted a lot of attention because of its wide range of applications in several fields, such as home security, autonomous vehicles, and gaming. In the medical industry, machine learning has the potential to save lives in addition to improving daily living. Excited people are using it for drug research, illness diagnosis from medical imaging, and patient care. The application of machine learning algorithms to the early diagnosis of lung cancer, a disease with a high death rate, is one particular example. Three-dimensional neural networks have been employed by the engineering firm Draper to increase the precision of lung cancer diagnosis and detection. Although the results are encouraging, greater advancements can be achieved by adding patient referral data and bigger, more varied datasets. Drug discovery procedures can also be accelerated by machine learning. Using machine learning algorithms, Exscientia, a British drug development business, and Sumitomo Dainippon Pharma are working to produce a medication for the treatment of obsessive-compulsive disorder. The partnership produced a novel medication in twelve months, which is much faster than the usual five-year drug development schedule. Furthermore, machine learning has been used in blood glucose monitoring in an effort to provide people a less intrusive choice. Researchers have looked on detecting low blood glucose levels using electrocardiogram (ECG) readouts. Abnormal glucose levels can be identified by using patient ECG measurements to train an algorithm that looks for distinctive patterns. With its minimally intrusive blood glucose monitoring option, this technology offers potential benefits to both non-diabetics and diabetics.

V. SIMULATION AND RESULT DISCUSSION

This section presents the findings from the testing of the FLS-SVM based algorithm utilizing the absolute slope, entropy, and combined entropy criterion. Using data-set 3 of the CSE multi-lead measurement library, the suggested algorithm for QRS-detection in simultaneously recorded ECG signal is carried out [6-7]. Original 12-lead simultaneous ECG recordings from 125 patients, representing a range of pathological scenarios, are included in this data set. If the algorithm properly recognizes the QRS-complex, the detection is referred to as true positive (TP); if not, it is referred to as false negative (FN). When a non-QRS wave is identified as a QRS complex, false positive (FP) detections are made. The following examples show how well the FLS-SVM based algorithm performs when signal absolute slope is used as a feature.

1) QRS-Detection using Absolute Slope as a Feature

When the algorithm is tested with the optimal parameter values ($\gamma = 10$, $\sigma = 4$), the detection rate is 99.82%. False positive detection rates are 1.61% and false negative detection rates are 0.18%. In certain instances, the strong absolute slope of the P and T waves is the primary cause of false positive detections. The 125, 12-lead simultaneously recorded original ECG recordings of the standard CSE ECG data-set 3 contain a total of 1488 QRS-complexes. Of the MO1_075 record, only three QRS-complexes are missed by the suggested algorithm. Any additional effort to locate or eliminate this false negative by modifying the FLS-SVM's parameters lowers the algorithm's overall detection rate.

The following examples show how well the FLS-SVM based algorithm performs when signal absolute slope is used as a feature.

- **Example 1:** The 12-lead ECG signal of record MO1_108 from the CSE ECG data-set 3 is displayed in Fig. 7, along with a square wave that indicates the positions of the QRS-complexes found by the FLS-SVM. Since the morphology of the QRS-complexes in each of the ECG signal's leads is obviously consistent, the FLS-SVM has been effective in detecting each and every one of the QRS-complexes.

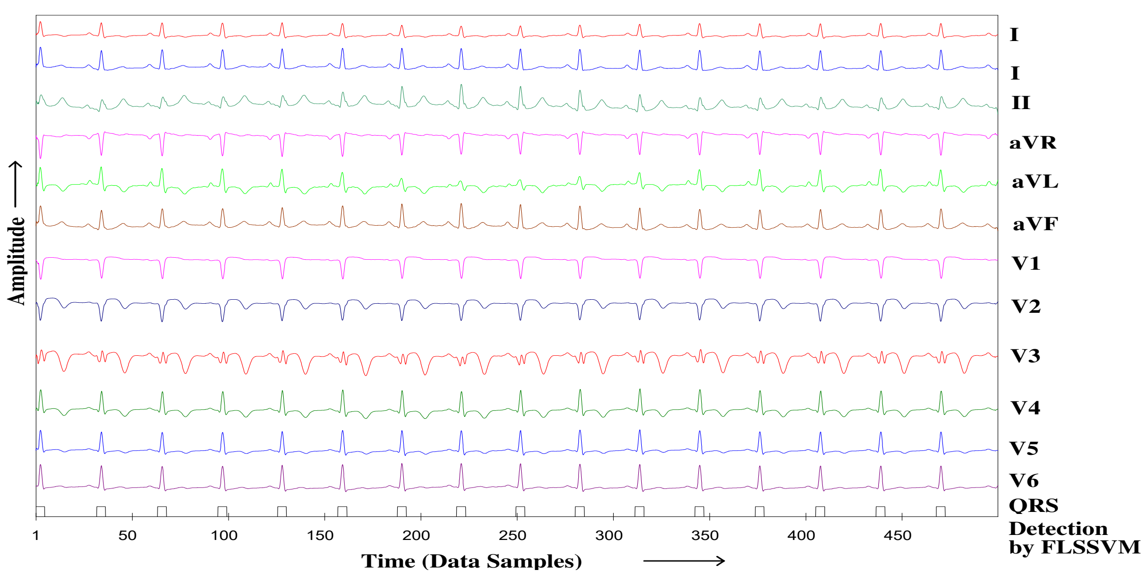

Fig. 7 Detection of QRS-complexes in record MO1_108 using absolute slope as feature

- **Example 2:** Record MO1 075's QRS detection is displayed in Fig. 8. As seen in this picture, the algorithm is unable to identify the second, seventh, and eleventh QRS-complexes out of thirteen because of their lower absolute slope in comparison to the other QRS-complexes. Out of 125 ECGs in data-set 3, this 12-lead ECG is the only one where three false negatives (FN) are found, highlighting FLS-SVM's potency.

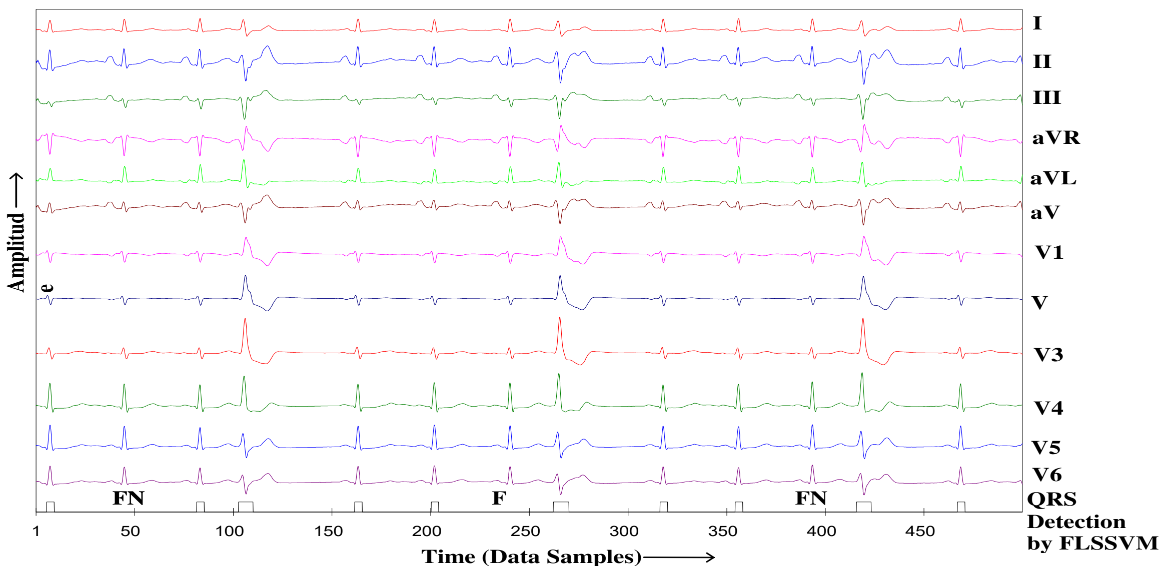

Fig. 8 Detection of QRS-complexes in record MO1_075 using absolute slope as feature

2) QRS-Detection using Entropy as a Feature

Utilizing entropy criteria, the QRS-detection algorithm yields a 99.93% detection rate when tested with optimal parameter values (γ = 10 and σ = 4). There is a decrease in the proportion of false positive and false negative detection to 0.87% and 0.06%, respectively. Peaky P and T-wave are sometimes the cause of false positive detections. The 125, 12-lead simultaneously recorded original ECG recordings of the standard CSE ECG data-set 3 contain a total of 1488 QRS-complexes. The record MO1_045 has only one QRS-complex that the suggested method is unable to identify. Higher QRS-entropies and lower non-QRS-entropies in the QRS-region of the majority of the leads are the cause of this false negative detection. Any additional efforts to identify this specific QRS-complex by modifying the FLS-SVM parameter reduce the detection rate overall. As previously discussed, the three QRS-complexes in record MO1_075 that gave false negative detections when detected using absolute slope criteria are correctly identified when detected using entropy criteria, demonstrating the superiority of entropy over absolute slope feature.

The following examples show how the FLS-SVM based technique, which uses signal entropy as a feature, might be useful.

- **Example 1:** The 12-lead ECG signal of record MO1_020 is displayed in Fig. 9, along with a square wave that indicates the positions of the QRS-complexes found by the FLS-SVM. Since the morphology of the QRS-complexes in each of the ECG signal's leads is obviously consistent, the FLS-SVM has been effective in detecting each and every one of the QRS-complexes.

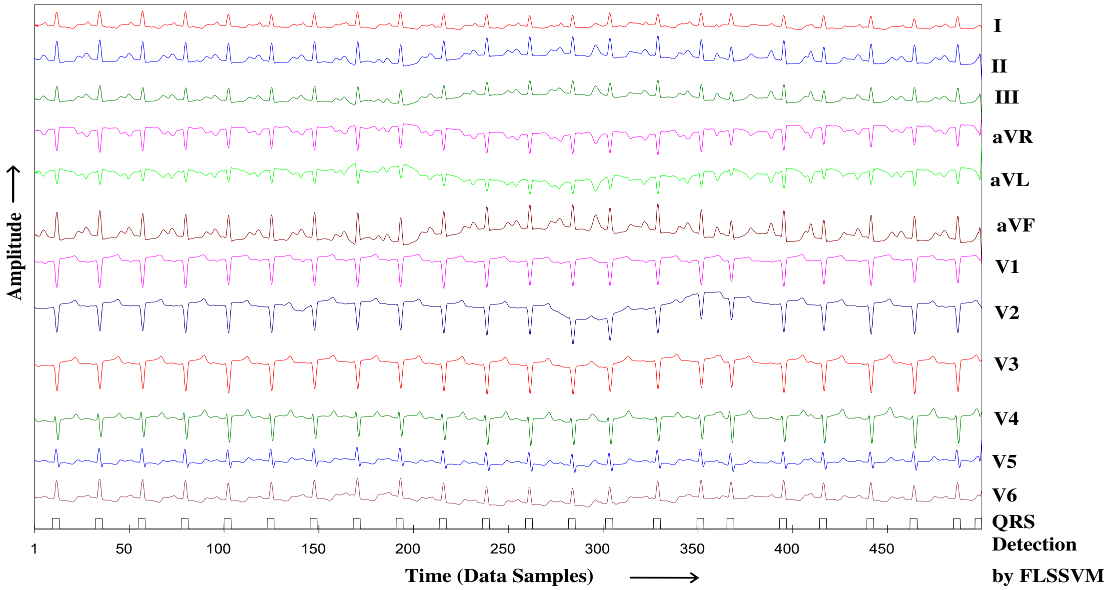

Fig. 9 Detection of QRS-complexes in record MO1_020 using entropy as feature

- **Example 2:** Fig. 10 shows the QRS-detection of record MO1_075. In comparison to other QRS-complexes, the amplitude of the fourth, eighth, and twelve QRS-complexes is greater in the majority of the leads. The algorithm correctly identifies all of these QRS-complexes and the other very small amplitude QRS-complexes, demonstrating its efficacy.

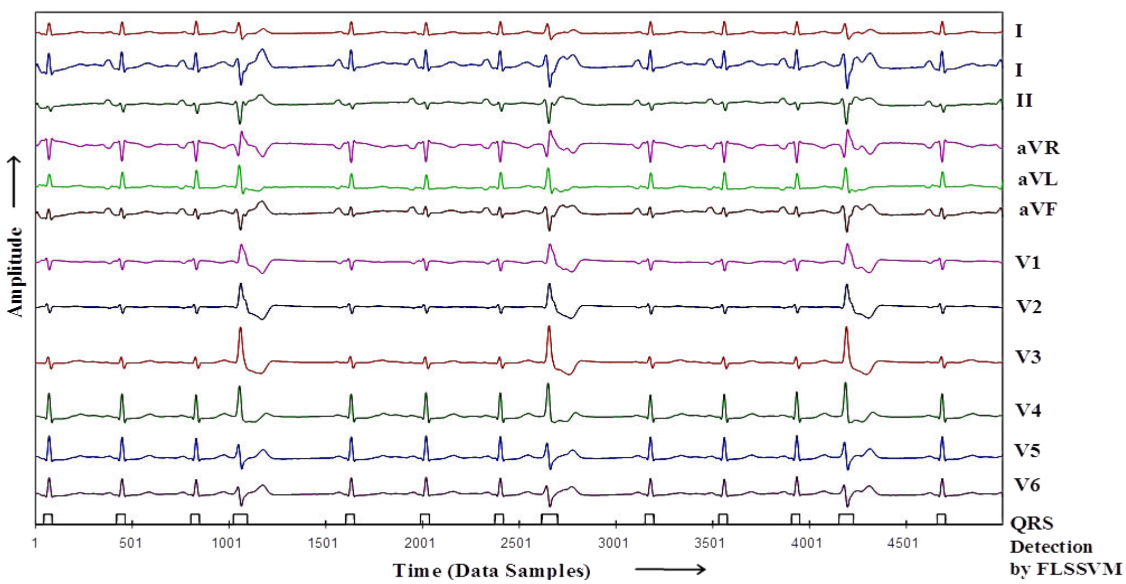

Fig. 10 Detection of QRS-complexes in record MO1_075 using entropy as feature

VI. CONCLUSION

Machine learning techniques are being utilized in medical research to learn, analyse, and extrapolate details as a result of the rise of large data. In recent years, a wide range of medical data and applications have been using machine learning (ML) techniques, such as support vector machines, K-means clustering, decision trees, random forests, Naïve Bayes, K nearest neighbours, neural networks, and convolution neural networks. This survey offers a thorough overview of these techniques. In this chapter, the FLS-SVM technique has been successfully applied to identify QRS-complexes in a 12-lead ECG signal that was concurrently recorded. The performance of the FLS-SVM based algorithm is found to be better compared to the approaches reported in the literature, according to simulation results of the algorithm utilizing data-set 3 of the CSE multi-lead measurement library. The detection of cardiac complexes, particularly QRS-complexes, is crucial to the accuracy and dependability of computerized ECG processing systems' performance. The algorithms based on FLS-SVM represent a step in this direction.

REFERENCES

[1] Kaustubh Arun Bhavsar, Jimmy Singla, Yasser D. Al-Otaibi, Oh-Young Song, Yousaf Bin Zikria and Ali Kashif Bashir, “Medical Diagnosis Using Machine Learning: A Statistical Review”, Computers, Materials & Continua · January 2021, Tech Science Press.

[2] Arunim Garg, Vijay Mago, “Role of machine learning in medical research: A survey”, Computer Science Review 40 (2021), 1574-0137/© 2021 Elsevier

[3] Samin Poudel, “A Study of Disease Diagnosis Using Machine Learning”, Med. Sci. Forum 2022 © 2022 Licensee MDPI, Basel, Switzerland

[4] Md Manjurul Ahsan, Shahana Akter Luna and Zahed Siddique, “Machine-Learning-Based Disease Diagnosis: A Comprehensive Review”, Healthcare 2022, 10, 541, © 2022 Licensee MDPI, Basel, Switzerland.

[5] Basu Dev Shivahare, Jagendra Singh, Vinayakumar Ravi, Radha Raman Chandan, Tahani Jaser Alahmadi, Prabhishek Singh and Manoj Diwakar, “Delving into Machine Learning's Influence on Disease Diagnosis and Prediction”, The Open Public Health Journal, 2024, Vol. 17 © 2024 Published by Bentham Open.

[6] Kharibam Jilenkumari Devi, Wajdi Alghamdi, Divya N, Ahmed Alkhayyat, Artikbaeva Sayyora T.Sathish, “Artificial Intelligence in Healthcare: Diagnosis, Treatment, and Prediction”, E3S Web of Conferences 399, 04043 (2023) © published by EDP Sciences ICONNECT-2023.

[7] S.M. Atikur Rahman, Sifat Ibtisum, Ehsan Bazgir and Tumpa Barai, “The Significance of Machine Learning in Clinical Disease Diagnosis: A Review”, International Journal of Computer Applications (0975 – 8887) Volume 185 – No. 36, October 2023.

[8] Deepak Kumar, Punet Kumar, Iftekhar Ahmed, Sangam Singh, “Integrating Artificial Intelligence in Disease Diagnosis, Treatment, And Formulation Development: A Review”, Asian J Pharm Clin Res, Vol 16, Issue 11, 2023, 1-8.

[9] Alafif, T., Tehame, A. M., Bajaba, S., Barnawi, A., & Zia, S. (2021). Machine and deep learning towards COVID-19 diagnosis and treatment: survey, challenges, and future directions. International journal of environmental research and public health, 18(3), 1117.

[10] Jamshidi, M., Lalbakhsh, A., Talla, J., Peroutka, Z., Hadjilooei, F., Lalbakhsh, P., ... & Mohyuddin, W. (2020). Artificial intelligence and COVID-19: deep learning approaches for diagnosis and treatment. Ieee Access, 8, 109581-109595.

[11] Myszczynska, M. A., Ojamies, P. N., Lacoste, A. M., Neil, D., Saffari, A., Mead, R., ... & Ferraiuolo, L. (2020). Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nature Reviews Neurology, 16(8), 440-456.

[12] Shaheen, M. Y. (2021). Adoption of machine learning for medical diagnosis. ScienceOpen Preprints.

[13] Singh, P., Singh, N., Singh, K. K., & Singh, A. (2021). Diagnosing of disease using machine learning. In Machine learning and the internet of medical things in healthcare (pp. 89-111). Academic Press.

[14] Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. New England Journal of Medicine, 375(13), 1216-1219.

[15] Topol, E. J. (2019). Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. Basic Books.

[16] Doshi-Velez, F., & Waldo, J. (Eds.). (2020). Responsible AI in Healthcare: Enabling Clinicians to Provide Better Patient Care. O'Reilly Media.