

# CONTRIBUTION OF MACHINE LEARNING TECHNIQUES TO DETECT DISEASE IN-PATIENTS: A COMPREHENSIVE ANALYSIS OF CLASSIFICATION TECHNIQUES

SWARN AVINASH KUMAR<sup>\*</sup>, HARSH KUMAR<sup>\*\*</sup>,  
VISHAL DUTT<sup>\*\*\*</sup>, HIMANSHU SWARNKAR<sup>\*\*\*\*</sup>

## ABSTRACTS

Technology is expanding rapidly. A lot of major advancements have been brought by overcoming the challenges in various fields using technology. The same is true for the medical domain where AI is augmenting a lot of existing methodologies such as predicting disease or guiding the doctors to reach efficient treatment procedures. In this study we are going to focus on various applications or relevant investigations where Machine Learning is playing a major role in health care in predicting diseases.

**KEYWORDS:** Data Mining, Image Processing, IoT, Machine Learning, Artificial Intelligence, Detection, Disease.

## INTRODUCTION

In today's time, the possibility of physical exercise is less in most jobs. Either the person's work is computer based or the sitting based. As a result, the people are sitting in front of computers for a work purpose and most of the people due to daily busy schedule, are skipping the exercises and not focusing on health, this result leads to getting various severe problems such as obesity, heart diseases, cancer, ulcers, etc. Many of the domains that came into the existence such as Data Mining, Image Processing, IoT, Machine Learning, Artificial Intelligence, etc., were the domains playing a vital role. Many of the algorithms were designed for the health domain such as Fuzzy

Techniques, hidden Markov methods, Neural Networks, SVM classification, Genetic Algorithm, Hybrid algorithms, Feature Selection, etc., these are the various methods being used for detecting the disease [1].

Even though there are a lot of techniques in Artificial Intelligence and Machine Learning the main thing is making all the algorithms should be available on a single roof. The concept of Machine Learning is used in Manufacture sector, industry, medical domain; business, etc., the use of Machine Learning is not limited. It is used in almost all domain applications.

---

<sup>\*</sup>IIIT Allahabad, UP, India.

<sup>\*\*</sup>Peoples' Friendship University of Russia, Moscow, Russia.

<sup>\*\*\*</sup>Department of Computer Science, Arybhata College, Ajmer, India.

<sup>\*\*\*\*</sup>Department of Computer Science Engineering, Engineering College, Banswara, India.

**Correspondence E-mail Id:** editor@eurekajournals.com

The field of improvement can be observed on large scale and cannot be limited. The assisting of the input and the terms were limited to the number of terms and conditions of the particular domain. There is a great expectation that the use of Machine Learning in the medical field is going to raise the standard and increase the use in medical concepts. As an increase in the medical domain the advantages that we can get like flexibility, reliability, ease of use, elasticity, helps in making quick decisions, etc.

## LITERATURE WORK

Many of the authors, reviewers, scientists were working on health care in many aspects like identifying the disease, predicting the disease, analysis of the disease, treatment suggestions of the disease to doctors. By linking various domains such as IoT, Machine Learning, Deep Learning and Image Processing into the health care domain which made the task easy for disease prediction, analysis, diagnosis and medication.

### ARTICLE 1

**Title:** "A novel machine learning approach for early detection of hepatocellular carcinoma patients" [3]

### CONCEPT

In this paper, the author has focused on Hepatocellular Carcinoma (HCC) which is one of the types of liver cancer. This particular cancer has a very high number of deaths every year in the overall world. The author has worked regarding this cancer and surveyed over 165 patients. The author has worked on over 10 various Machine Learning algorithms. The proposed algorithm in this paper was applied during the initiation of the task i.e., during pre-processing step. Here the author used SVM classifier technique for simplifying the given input. The algorithm which was designed for

detecting various cancer symptoms of body based on reports so its termed as C SVC. The Type C SVC is a two-level approach mechanism in the genetic optimizer. This Type C SVC algorithm was sampled over 165 cases and got positive results with 0.9958 and 0.9873 respectively.

### ARTICLE 2

**Title:** "Machine Learning for Predicting Cognitive Diseases: Methods, Data Sources and Risk Factors". [4]

### CONCEPT

Cognitive is a type of disorder that comes under the category of mental health issues. This disorder affects the memory of a person, Perception of a person as well as problem-solving of various issues which include amnesia, dementia as well as delirium. The author worked on Alzheimer's disease, mild cognitive impairment, and Parkinson's disease. The author worked on various concerns such as detection, diagnosis, practice, and therapy planning. Here the author used Magnetic Resource Imaging (MRI) scan. The author has designed the algorithm by taking the disease six major symptoms such as memory test, concentration test, functional test, dementia rating scale, Multimodal status exam, personal behaviour. Here the author has proposed the Classifier mechanism.

### ARTICLE 3

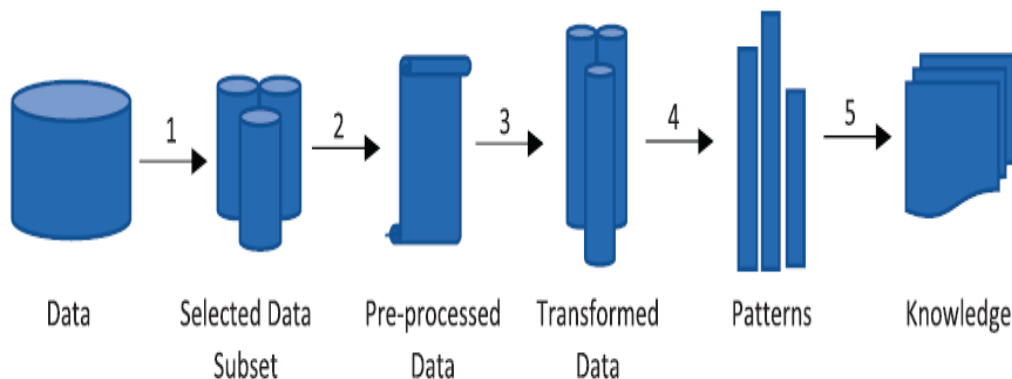
**Title:** "Machine learning and artificial intelligence-based Diabetes Mellitus detection and self-management: A systematic review" [5].

### CONCEPT

Diabetes Mellitus is a conditional state which is caused by diabetes. Diabetes having the capacity to make the failure of the body

functions if it goes severely. AI and ML are used in early identification as well as medication of Artificial Intelligence and Machine Learning was also used for detecting diabetes at early stages. For implementing this mechanism we have to focus on various terms such as Datasets, Pre-Processing, Patients diabetes reports, medication as well as treatment. Here these six factors were taken based on the study of 177

relevant concepts. Here the self-management personalization is also discussed. Here they have taken multiple patient's information and created the dataset and named it Knowledge Exploration Dataset (KED). Here they have taken various learning mechanisms like Supervised Learning, Unsupervised Learning, and Reinforcement Learning.



**Figure 1. Basic Phases of Knowledge Extraction Datasets**

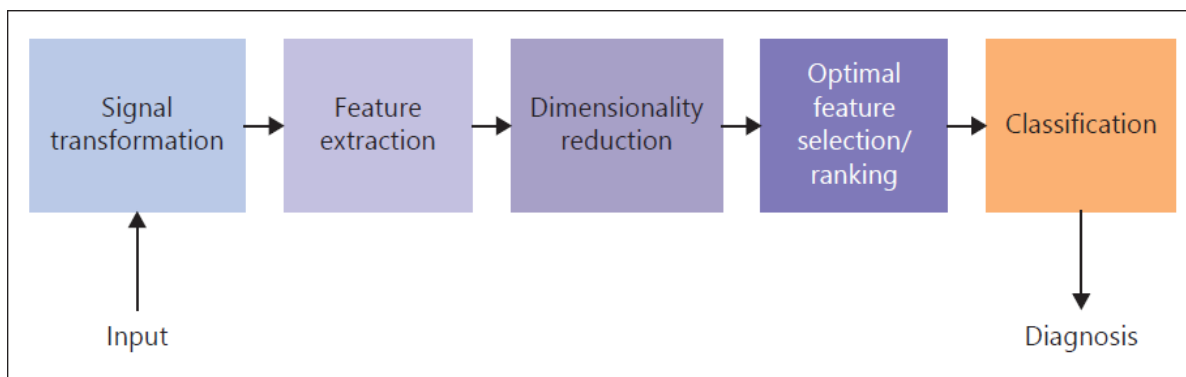
**ARTICLE 4**

**Title:** "Artificial Intelligence Techniques for Automated Diagnosis of Neurological Disorders." [6]

**CONCEPT**

The authors have designed the concept based on a computer-aided diagnosis. Here patients were trained based on the psychological signals as well as advanced intelligence signals based on Artificial Intelligence and Machine Learning

techniques. Here these techniques help neurologists, neurosurgeons and radiologists help in taking wise decisions in the medical field. This paper gives the best review of the various neurological simulations of the body. Here the author has designed based on the CAD systems for early and easy way of identifying various Neurological Disorders and also by implementing CAD mechanism the medication and analysis will be easy. In this paper the author has focused on Dimensionality Reduction, Optimal Feature selection, and Ranking and feature classification.



**Figure 2. Architecture of CAD**

**ARTICLE 5**

**Title:** "Artificial intelligence for brain diseases: A systematic review". [7]

**CONCEPT**

Artificial Intelligence is playing a great role in the health industry. It is showing the best result in the all body parts diagnosis such as heart, lungs, etc., Artificial Intelligence is playing a very major role in brain problems diagnosis. It succeeds very great because of correct planning, implementation, and outcome prediction. By Artificial Intelligence the

approach of disease identifying in the brain, curing the disease, etc., AI and ML played very major role in medical domain. Based on the concept of this article the author performed the task with 162 persons and with 1562 test cases. Various clinical issues for the task implementation and development were proposed. Here they have taken structural connectivity, micro connectivity, and radio connectivity was designed. Here for development, they have taken clustering mechanism, reinforcement mechanism, and decision tree mechanism was implemented. Here the author has designed the PRISMA flow diagram for designing the technique.

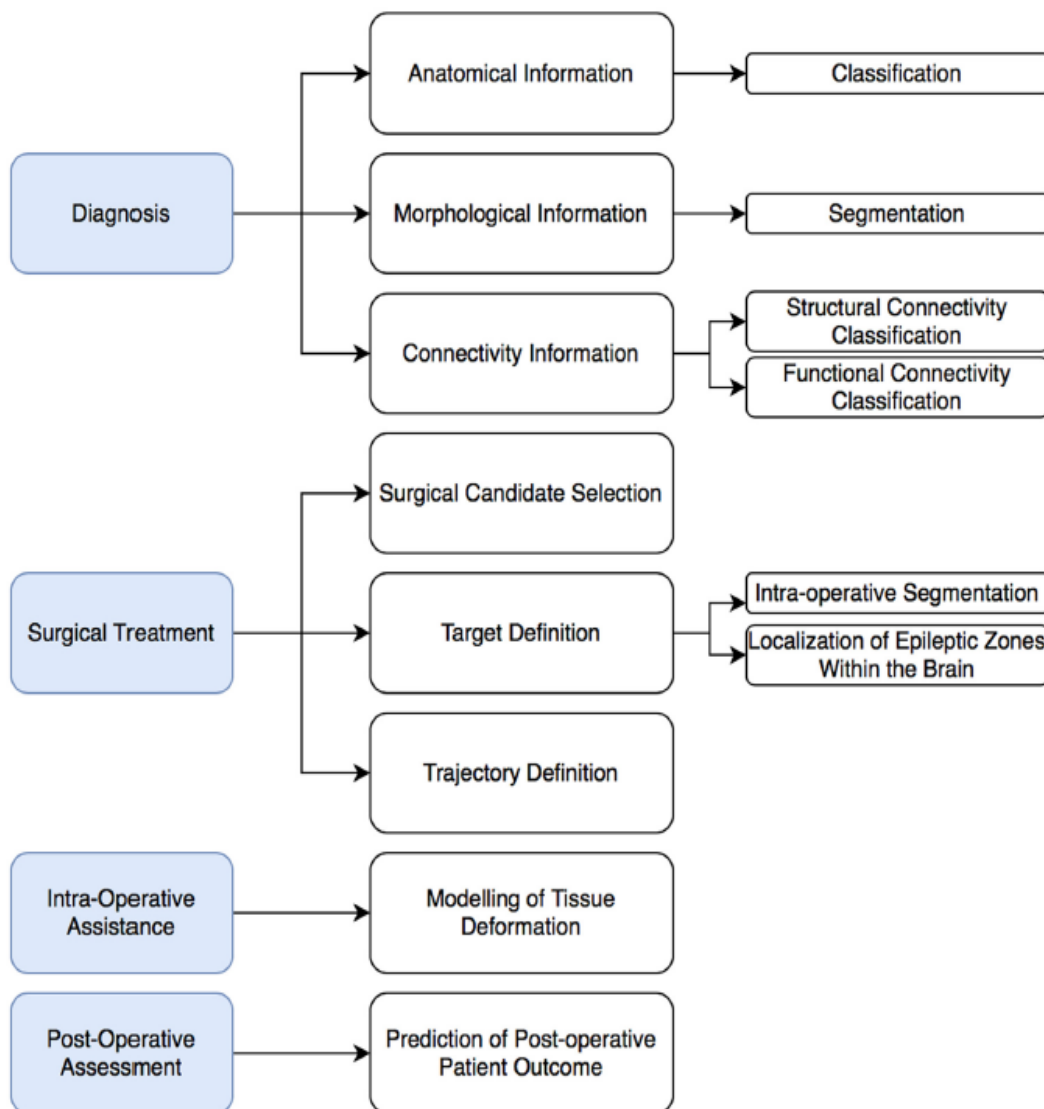


Figure 3. Architecture in Detecting Brain Care

## ARTICLE 6

**Title:** "Leveraging Deep Learning Techniques for Malaria Parasite Detection Using Mobile Application". [8]

### CONCEPT

A lot of research was going on in detecting malaria using Machine Learning. Many of the people were suffering a lot from malaria. Many of the mechanisms came into the existence but not up to the mark. The author has created the solution to solve the problems of the malaria using Machine Learning technique. Most of the research is going on in the medical domain for reliability and ease of use and for many more advantages. Here the techniques such as convolution neural networks (CNN) and Recurrent Neural Network (RNN) were the proposed things that came into the existence. There is another model called a Deep Neural Network (DNN) was designed for the better efficiency and privacy of the algorithm. There is an approach named as Extreme Machine Learning approach was designed. The segmentation is called Deep Aware CNN. Deep Aware CNN is a mechanism that produces the task at a segmentation level and analysis level. This algorithm has the following characteristics as

- Design and Evaluation
- Use of Evaluation and
- Deployment for better performance.

## ARTICLE 7

**Title:** "A Survey on various Machine Learning Approaches for ECG Analysis". [9]

### CONCEPT

Electrocardiogram (ECG) is a mechanism that is used in determining the proper functioning of the heart. Multiple numbers of features were graded for calculating the correct pressure of

the body. In this article the author has designed the protocol for the diagnosing Myocardial Infarction which is also known as heart attack. This design also helps in identifying various symptoms such as thickness of the heart muscle, the variation of the heart as well as enlargement of the heart. Moreover, there are various Machine Learning methods for performing the multiple tasks such as heart monitoring, alerting doctors and higher authority if the condition of the patient turns critical. Multiple test cases and medical reports of the patients were taken for creating datasets. These datasets were bits of help in analyzing the future patient's reports. So that with the help of the reports doctors can take quick decisions and make tasks efficient and simplified manner. Here the authors have taken the neural network methods for performing the task. Here the Fuzzy Logic mechanism and Artificial Neural Network method, SVM classifier, Normal Sinus Rhythm (NSR) was playing a major role in identifying things most efficiently.

In this concept, ECG is designed in a Multilayer Perception manner. The complete information is stored in the NSR database. These all algorithms were merged and termed as Back Propagation Algorithm (BPA). Based on these tests and results it is found that 94.12 percent of efficient results were produced. Finally in this article it is proved that by using the concepts of Artificial Intelligence and Machine Learning we will get more efficient and better results.

## ARTICLE 8

**Title:** "Implementation of machine learning algorithms to create diabetic patients-admission profiles". [10]

### CONCEPT

In this paper, the author has proposed the dataset for diabetic patients using Artificial

Intelligence and Machine Learning. In this paper, the authors have studied various patterns and multiple diabetic patients' information into the set of the dataset. Here in this paper, the author used various algorithms like Linear Regression, Linear Discriminate Analysis, Random forest, K- Nearest Neighbor, Naïve Bayes, J48 as well as Super Vector Machine for maintaining patients profiles.

Based on these results by testing 10,000 people 47% of people came to know that they were affected by diabetes for the first time and went to the hospital for confirmation. The confirmation came true. Moreover, most of the research was going on to take the test as a positive response from people. Here the authors for performing the task have used the concept of ROC efficiency analysis as well as a convolution confusion matrix. The testing was done by using various strategies.

### **ARTICLE 9**

**Title:** "Image analysis and machine learning for detecting malaria". [11]

### **CONCEPT**

**By** Malaria disease more than 4 lakhs of deaths occurring every year. Many of the modern mechanisms came into the existence as the remedy for the problem. Bio-Medical research played a prominent role in disease detection. Many of the mechanisms were in processing in the domains such as Image Processing, Machine Learning, Artificial Intelligence, etc., here the authors have worked out on various algorithms such as Rule-Based Segmentation, Fuzzy Based Segmentation, Neural Network, Watershed

Algorithm, etc., were designed in the development of the work. Many methods were designed with better results with 83.16 and 76.92.

### **ARTICLE 10**

**Title:** "Secure and Robust Machine Learning for Healthcare: A Survey". [12]

### **CONCEPT**

In recent days the concept of Machine Learning and Artificial Intelligence is booming a lot in the medical domain. ML and AI were playing the greater role because it is making quick decisions, flexible task, more efficient, reliable, etc., here the signals were generated based on the computer-aided diagnosis (CAD) [13]. The concepts such as pathology, radiology, and clinicology were evolved a lot. In this paper, many of the related papers were studied and designed the mechanism for the implementation. Various contributions were done in this paper namely:

- Here the authors have formulated the various methods in developing various Machine Learning algorithms such as Rule-Based Segmentation, Fuzzy Based Segmentation, Neural Network, Watershed Algorithm, etc.,
- Identifies the various types of vulnerabilities of the problem.
- Highlighting various security patches and privacy concerns of the work.
- Developed and robustly presented the solutions.
- Finally brought out the final research issues which are yet to be solved.



**Figure 4. Applications of Health care system**

## APPLICATIONS

Various applications such as

- Electronic health records [14]
- ML in Medical Image Analysis
  - Enhancement
  - Detection
  - Classification
  - Segmentation and
  - Reconstruction
  - Image Registration
  - Retrieval
- Image Interpretation
- Health monitoring
- Disease Prediction and diagnosis
- Computer-aided detection and diagnosis
- Clinical Reinforcement Learning
- Clinical Time series data

- Clinical Natural Language Processing
- Clinical speech and audio processing

Many of the algorithms or mechanisms were developed in the designing of the Machine Learning techniques [15]. Few of them are

- Unsupervised Learning
- Supervised Learning
- Semi-supervised Learning and
- Reinforcement Learning

The working mechanism will be different for various applications are as follows:

## UNSUPERVISED LEARNING

The general examples of unsupervised learning such as clustering etc., here the data from the above values will be taken [16].

## **SUPERVISED LEARNING**

Here in Supervised Learning the build or mapping of the training data will be taken for the estimation. All the data will be characterized into a single data. The resultant of all the data is considered as the single procedure for processing the task. Various pathology, geographical and analytical results will be considered of the processing [17].

## **SEMI-SUPERVISED LEARNING**

This method is used for both the labeled as well as unlabelled data for processing. Here multiple amounts of sampling are available for Labeled data as well as unlabelled data. For the applications such as health care applications, more amounts of datasets, sampling, and training are needed for processing [18]. Here in semi-supervised learning, the time taken for processing the task is more than the usual. Here the data recognition will be the more fast and secure manner in image data sets compared to other data sets.

## **REINFORCEMENT LEARNING**

Here each consideration is treated as policy. For each function, there are a set of defined functions, predefined values, terms, conditions, observations, reward points, etc., which were all considered as reinforcement learning. Here context-aware mechanism, disease identification and disease diagnosis became more simplified manner. Moreover, Reinforcement Learning is the mechanism that is more adequate and uses in very reliable applications which are giving more adequate results with direct interaction to the user.

## **RESEARCH GOING ON IN THIS DOMAIN**

Many of the Machine Learning and Artificial Intelligence challenges were proposed for future works. Most of the research is going

on. But many of the problems were leftover. Few of the challenges were

- Source of vulnerabilities in the pipeline.
- Vulnerability due to data annotation.
- Limited and imbalanced datasets.
- Class imbalance and Bias.
- Scarcity
- Vulnerabilities in model training
- Vulnerabilities in Deployment Phase

The other more challenges such as vulnerabilities, security violation, Attack specificity, poison attack, evasion attack etc., many of the challenges such as safety challenges, privacy challenges, Ethical challenges, Causality challenging, Regulatory challenging, availability of good quality data, Lack of data standardization, exchange, Distribution shifts and Updated hospital infrastructure became difficult. Cryptographic encryption, Homographic encryption, Garbled Encryption, Secret sharing, Secure Processors are the various methods for data privacy.

## **CONCLUSION**

The usage of Machine Learning, Artificial Intelligence, and Deep Learning is increasing very widely. Many of the encrypted algorithms came into use to give a guarantee regarding the privacy of the patient's sensitive information. A lot of health service delivery models came into the force. Moreover to provide security many of the encrypted methods were done in the clinical domains. A lot of privacy is enabled for various privacy and confidentiality of the data to overcome the address of the particular challenge. Here in this paper, we have studied various disease treatments in Machine Learning applications. Here various Machine Learning techniques were using in disease detection, Disease identification, and disease treatment, precautions of the disease, and final diagnosis of the disease. Here we have defined and identified various Machine Learning



mechanisms for the development of the work. Here multiple resources of vulnerabilities were identified. Here we have designed various solutions for privacy and preserving mechanisms for the development of various protocols. Many of the security and the critical applications were got solved when it was designed in the concept of Machine Learning. Finally, the main agenda of our survey research is to know what are the various Machine Learning concepts are using in patients treatments in multiple fields such as cancer detection, diabetes treatment, brain problems, heart-related issues, various types of cancer detection and prediction system, etc., and also used in the Machine Learning.

## REFERENCES

- [1]. S. Latif, J. Qadir, S. Farooq, and M. Imran, "How 5G wireless (and concomitant technologies) will revolutionize healthcare?" *Future Internet*, vol. 9, no. 4, p. 93, 2017.
- [2]. Z. Yan, Y. Zhan, Z. Peng, S. Liao, Y. Shinagawa, S. Zhang, D. N. Metaxas, and X. S. Zhou, "Multi-instance deep learning: Discover discriminative local anatomies for body part recognition," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1332-1343, 2016.
- [3]. A.Kumar, D.Hung, D. Le, T.Sairam "Smart Cities Management Using Deep Learning And lot With Cloud Computing ", "Global Journal on Innovation, Opportunities, and Challenges in Applied Artificial Intelligence and Machine Learning Vol. 3, Issue 1-2019.
- [4]. Wojciech Książek, Moloud Abdar, U. Rajendra Acharya, Paweł Pławiak, "A novel machine learning approach for early detection of hepatocellular carcinoma patients, *Cognitive Systems Research*, Volume 54,2019, Pages116-127, ISSN 1389-0417.
- [5]. Bratić, B., Kurbalija, V., Ivanović, M. et al. Machine Learning for Predicting Cognitive Diseases: Methods, Data Sources, and Risk Factors. *J Med Syst* 42, 243 (2018).
- [6]. A.Kumar, Surbhi Sharma and Sammah Rasheed "A Shape Mechanism along Text-Based Advanced Face Detection Methodology Using EFM Model", "IJIRCCE Vol 7, Issue 2- 2019, 23209798.
- [7]. J. Chaki, S. Thillai Ganesh, S. K. Cidham, et al., Machine learning and artificial intelligence-based Diabetes Mellitus detection and self-management: A systematic review, *Journal of King Saud University-Computer and Information Sciences*.
- [8]. Artificial Intelligence Techniques for Automated Diagnosis of Neurological Disorders by U. Raghavendra.
- [9]. A.Dubey, S.Narang, A.Kumar, S.Sasubilli, V.Díaz (2020), "Performance estimation of machine learning algorithms in the factor analysis of COVID-19 dataset" *Computers, Materials & Continua*, ISSN: 1546-2226.
- [10]. Artificial intelligence for brain diseases: A systematic review by Alice legato.
- [11]. Leveraging Deep Learning Techniques for Malaria Parasite Detection Using Mobile Application by Mehedi Masud.
- [12]. S. Sasubilli, K. Attanguddi and A. Kumar "APSO based feature extraction approach and SVM based classification for sentimental analysis", "Global Journal on Innovation, Opportunities, and Challenges in Applied Artificial Intelligence and Machine Learning Vol. 2, Issue 1-2018.
- [13]. A Survey on various Machine Learning Approaches for ECG Analysis by C K Roopa.
- [14]. Implementation of machine learning algorithms to create diabetic patients-admission profiles by allophane.

- [15]. S. Chandrasekaran and A. Kumar  
Implementing Medical Data Processing  
with Ann with Hybrid Approach of  
Implementation Journal of Advanced  
Research in Dynamical and Control  
Systems-JARDCS issue 10, vol.10, page  
45-52, ISSN-1943-023X. 2018/09/15.
- [16]. Image analysis and machine learning for  
detecting malaria by Bethesda.
- [17]. Secure and Robust Machine Learning for  
Healthcare: A Survey by Adnan Qayyum.
- [18]. R. Raturi and A. Kumar " An Analytical  
Approach for Health Data Analysis and  
finding the Correlations of attributes  
using Decision Tree and W-Logistic Modal  
Process", 2019, "IJRCCE", Vol 7, Issue 6,  
ISSN: 2320-9801.