The Role of Artificial Intelligence (AI) in the Detection and Diagnosis of Chronic Diseases – A Review

By

Mumtahina Zaman
17146007

A thesis submitted to the Department of Pharmacy in partial fulfillment of the requirements for the degree of
Bachelor of Pharmacy (Hons.)

The Department of Pharmacy
Brac University
January, 2022

Declaration

It is here by declared that

1. The thesis submitted is my own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. I have acknowledged all main sources of help.

Student's Full Name & Signature:

Mumtahina Zaman

Mumtahina Zaman
17146007

Approval

The project titled “Role of Artificial Intelligence (AI) in Diagnosis of Chronic Diseases” submitted by Mumtahina Zaman (17146007) of Spring, 2017 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Pharmacy (Hons.) on January 2022.

Examining Committee:

Namara 22/01/2022

Supervisor:
(Member)

____
Namara Mariam Chowdhury
Lecturer, Department of Pharmacy
Brac University

Program Coordinator:
(Member)

____
Dr. Zara Sheikh
Assistant Professor, Department of Pharmacy
Brac University

Departmental Head:
(Chair)

____
Dr. Eva Rahman Kabir
Professor, Department of Pharmacy
Brac University

Ethics Statement

This study does not involve any human or animal trial.

Abstract

The diagnosis of diseases is decisive for planning proper treatment and ensuring the well-being of patients. Human error hinders accurate diagnostics, as interpreting medical information is a complex and cognitively challenging task. The application of artificial intelligence (AI) can improve the level of diagnostic accuracy and efficiency. While the current literature has examined various approaches to diagnosing various diseases, an overview of fields in which AI has been applied, including their performance aiming to identify emergent digitalized healthcare services, has not yet been adequately realized in extant research. By conducting a critical review, we portray the AI landscape in diagnostics. Practitioners understand the extent to which AI improves diagnostics and how healthcare benefits from it. However, several issues need to be addressed before successful application of AI in disease diagnostics can be achieved.

Keywords: Artificial Intelligence, Chronic Diseases, Diagnosis, Machines Learning, Algorithm, Health Care

Dedication

Dedicated to my parents

Acknowledgement

This is a review project under the supervision of Namara Mariam Chowdhury, Lecturer, Department of Pharmacy, Brac University.

I would like to convey my earnest respect to my supervisor, Namara Mariam Chowdhury (Lecturer, Department of Pharmacy, Brac University) without whom working on this project would not have been possible. I would also like to offer my heartfelt appreciation and gratefulness for her valuable guidelines and recommendations, which helped my project work to be more presentable. Her generous motivation, insightful discussion and continuous support shaped the accomplishment of this project. I would like to convey my gratitude towards her for giving me this brilliant opportunity to work on this project and for constant guidance. I am truly thankful for the support; she has provided me, with her valuable time, assistance and significant suggestions.

My acknowledgement would be incomplete without thanking my parents for their constant support and encouragement, which helped me to keep patience throughout all the difficulties.

Table of Contents

Declaration i

Approval ii

Ethics Statement iii

Abstract iv

Dedication v

Acknowledgement vi

Table of Contents vii

List of Tables ix

List of Figures x

Chapter 1: Introduction 1

- 1.1 Artificial Intelligence 1
- 1.2 Chronic Diseases 2
- 1.3 Types of Chronic Diseases 2
- 1.4 Pathophysiology of Diseases: 2
- 1.5 Impact of AI in Chronic Diseases 8
- 1.6 Aims and Objectives of the Study 9

Chapter 2: Methodology 10

Chapter 3: Discussion 11

- 3.1 Different Types of Machine Learning Algorithm 12
- 3.2 Use of Machine Learning Algorithm in Diagnosis of Chronic Diseases: 14
- 3.3 Machine Learning in Heart Disease: 15
- 3.4 Machine Learning in Diabetes: 19
- 3.5 Machine Learning in Liver Diseases: 21

3.6 Machine Learning in Hepatitis: 25
3.7 Machine Learning in Dengue: 27
Chapter 4: Observation 31
Chapter 5: Limitations of AI in Diagnosis of Diseases 33
5.1 Software Security: 33
5.2 Human Factors: 33
5.3 Protection of Human Genetic Resources 34
Chapter 6: Future Prospects 35
6.1 Advancements and Explicability 35
6.2 Corroboration and Portability 35
6.3 Integration and Collaboration 36
Chapter 7: Conclusion 38
References 39

List of Tables

Table 1: Types of Heart Diseases 16

List of Figures

Figure 1: Impacts of AI in Chronic Diseases 8

Figure 2: Types of Machine Learning Techniques 12

Figure 3: Uses of Machine Learning Algorithm in Diagnosis of Chronic Diseases 15

Figure 4: Machine Learning Algorithm's Accuracy to Detect Heart Disease 18

Figure 5: Accuracy of Machine Learning Algorithm to Detect Diabetes 21

Figure 6: Accuracy of Machine Learning Algortihm to Detect Liver Disease 24

Figure 7: Accuracy of Machine Learning Algorithm to Detect Hepatitis 27

Figure 8: Accuracy of Machine learning Algorithm to Detect Dengue 29

Chapter 1: Introduction

1.1 Artificial Intelligence

Artificial intelligence (AI) is defined as machine intelligence in compared to the intelligence of humans or other living organism. (Minsky, 1961; Weng et al., 2001) AI can alternatively be described as the study of "intelligent agents"—that is, any agent or system that can observe and comprehend its surroundings and afterwards act accordingly to maximize its possibility of success (Wooldridge & Jennings, n.d.).

AI also refers to cases in which machines can resemble human minds in learning and analysis, allowing them to resolve issues. This type of intelligence is also known as (ML) machine learning (Huang et al., 2015). The ability of a machine to perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting with the environment, problem solving, decision-making, and even demonstrating creativity, is defined as "the ability of a machine to perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting with the environment, problem solving, decision-making, and even demonstrating creativity." There is no universal definition for AI (Mirbabaie et al., 2021). Artificial intelligence (AI) is a broad term that refers to a variety of scientific fields, including natural language processing and robots. Artificial intelligence (AI) has been utilized in several clinical studies over the last decade to improve the healthcare of patients with gastroenterological diseases, to effectively identify polyps and early cancer lesions, to assess the accuracy of inflammatory lesions, to analyse liver fibrosis, and to anticipate the response and clinical outcomes of medications (Hajjar & Rey, 2020).

1.2 Chronic Diseases

According to WHO: “Cardiovascular disorders (such as heart attacks and strokes), cancer, chronic respiratory diseases (such as chronic obstructive pulmonary disease and asthma), and diabetes are the four primary categories of chronic diseases. Chronic diseases are the major cause of mortality worldwide, accounting for more than 60% of all fatalities each year. In 2008, chronic illnesses caused 36 million deaths worldwide, the most of which were cardiovascular diseases, diabetes, chronic lung diseases, and malignancies (Alwan & World Health Organization., 2011).” Around a quarter of all chronic disease-related fatalities in the world occurred before the age of 60. Low- and middle-income nations accounted for 80% of all chronic illness mortality. Lower-income nations, populations, and communities are withstanding the worst of chronic disease burden, which is expected to rise dramatically over the next twenty years (Grover & Joshi, 2015.)

1.3 Types of Chronic Diseases

1. Heart Disease
2. Diabetes
3. Liver Disease
4. Dengue
5. Hepatitis

1.4 Pathophysiology of Diseases:

1.4.1 Pathophysiology of Heart Disease:

Heart failure is a worldwide epidemic that affects 1% to 2% of the population. The etiology and phenotype of heart failure varies greatly. Following a heart injury (e.g., myocardial infarction, increased preload or afterload), cellular, structural, and neurohumoral modulations occur that influence the phenotypic. These mechanisms have an impact on cell function, both intra- and intercellularly. As a result, the sympathoadrenergic and renin-angiotensin-aldosterone systems are activated, resulting in adaptive processes that are accompanied by volume overload, tachycardia, dyspnoea, and further impairment of cellular function (vicious circle). There is no unique clinical sign for heart failure; the clinical symptoms reveal increasing worsening either abruptly or persistently. The levels of neurohormones (norepinephrine) and natriuretic peptides (e.g., NT-pro BNP) rise because of cellular dysfunction. Non-invasive (echocardiography, NMR, NT-proBNP) and invasive (heart catheterization, biopsy) diagnostic methods are used to determine heart failure. Modulation of the activated systems with ß-blockers, ACE inhibitors, and ARNI improves outcome and symptoms in heart failure patients with left ventricular dysfunction. Interventional and surgical therapies may also be used. Understanding the underlying pathophysiology of heart failure is critical for initiating the appropriate therapy choice for each patient. Furthermore, lowering the risk of heart failure requires the avoidance of cardiovascular risk factors (Schwinger, 2021).

1.4.2 Pathophysiology of Diabetes:

Diabetes mellitus is a chronic metabolic illness with a multifactorial pathophysiology. It is characterized by high blood glucose levels, also known as hyperglycaemia, which is caused by problems in insulin production, insulin action, or both. Hyperglycaemia can express itself in a variety of ways, resulting in carbohydrate, lipid, and protein metabolic dysfunctions. Long-term hyperglycaemia frequently results in a variety of microvascular and macro vascular diabetic complications, which are the primary cause of diabetes-related morbidity and death. Hyperglycaemia is also used as a major indicator for diabetes diagnosis (Banday et al., 2020). Hyperglycaemia is caused by abnormalities in insulin production, insulin action, or both, and presents as carbohydrate, lipid, and protein metabolic dysfunctions in a chronic and varied way. Diabetes has a gradual course with a complicated etiology and a wide range of symptoms. DM is a heterogeneous metabolic illness that manifests as a condition with overlapping and difficult-to-distinguish symptoms. Each of these diabetes kinds has its own therapy and management regimen, but they all have a lot in common, just like the illness. All of this underlines the significance of accurate and quick identification of each of these diabetes kinds, as well as the crucial role of knowing their pathophysiology. This is critical to protect diabetics from the potentially harmful consequences of inefficient, unnecessary, or avoidable pharmacological therapies, which frequently delay the desired prognosis and lengthen hyperglycaemic exposures. Long-term hyperglycaemia, in turn, has been linked to an increased risk of microvascular and macro vascular diabetic complications both of which have a negative impact on quality of life and contribute significantly to diabetes-related morbidity and death (Gonzalez et al., 2011).

Insulin resistance and -cell dysfunction are the two primary insulin-related abnormalities in this kind of diabetes. Insulin resistance occurs when numerous cellular pathways are disrupted, resulting in a reduced response or sensitivity of cells in peripheral tissues, including muscle, liver, and adipose tissue. Reduced insulin sensitivity causes -cells to hyperactive function in order to achieve a compensatory increase in insulin production to maintain normoglycemia in the early stages of the disease. Although insulin secretion is reduced, it is usually enough to prevent DKA from occurring. DKA, on the other hand, can develop under extreme stress situations, such as those caused by infections or other pathophysiological events. Drugs including sodium-glucose co-transporter-2 (SGLT2) inhibitors, corticosteroids, and atypical antipsychotics can all cause DKA (second-generation antipsychotic drugs). In the absence of acute physiological stress, persons with T2DM seldom require insulin treatment at the outset of the illness or later in life. T2DM progresses slowly and asymptomatically, with even mild hyperglycaemia developing over time, and as T2DM has become more commonly associated with increasing age, obesity, family history of diabetes, physical inactivity, and adoption of modern lifestyles: with prior GDM in women, and with pathophysiological conditions such as hypertension and dyslipidaemia. T2DM's high prevalence in the aforementioned racial or ethnic groupings, as well as its significant link to first-degree relatives. T2DM's high prevalence among the aforementioned racial or ethnic groupings, as well as its significant link to first-degree blood relatives, strongly suggest that genetic factors play a role in the disease's genesis, but these aspects are complicated and largely unknown. Unlike T1DM, however, no link has been found between this disease and genes involved in the immune response, including autoimmunity, and as a result, there is no immune-mediated pancreatic regulation of systemic glucose due to its influence on the development of insulin resistance through its effect on tissue insulin sensitivity, and as a result, most, but not all, patients with T2DM are overweight or obese. Obesity is associated with an increase in body fat content, which is a major risk factor for cardiovascular disease. In comparison to increasing gluteal/subcutaneous fat or peripheral obesity, increased abdominal fat or visceral obesity has been commonly linked to this kind of diabetes. Patients with T2DM frequently present with cardiovascular risk factors such as hypertension and lipoprotein metabolic abnormalities defined by raised triglycerides and low levels of high-density lipoproteins due to its significant relationship with increased body fat content or obesity (HDLs)(Banday et al., 2020).

1.4.3 Pathophysiology of Chronic Liver Disease

Alcohol is ingested into the bloodstream through the gastrointestinal system and thus metabolized by hepatocytes in the liver (Teschke, 2019). Within the hepatocytes, there are three primary enzymatic metabolic pathways that are responsible for alcohol metabolism (Buchanan & Sinclair, 2021). Hepatocyte cytoplasmic alcohol dehydrogenase (ADH), which employs nicotinamide adenine dinucleotide (NAD+) as a co-factor to oxidize ethanol to acetaldehyde, which is highly toxic and impairs DNA synthesis, is the first and most important process (Meroni et al., 2018). The second pathway the microsomal ethanol-oxidizing system (MEOS) in the smooth endoplasmic reticulum uses the cytochrome P450 2E1 (CYP2E1) enzyme to convert ethanol to acetaldehyde, which produces reactive oxygen species (ROS) and causes oxidative stress and inflammation (Jiang et al., 2020; Meroni et al., 2018). It is worth noting that under normal physiological settings, CYP2E1 only converts around 10% of ethanol to acetaldehyde; nevertheless, due to increased CYP2E1 expression, it becomes more prominent in chronic alcohol usage (Kong et al., 2019; Meroni et al., 2018). The third, and minimal, process involves the peroxisomes' heme-containing catalase, which may also convert ethanol to acetaldehyde (Ceni et al., 2014). The enzyme aldehyde dehydrogenase (ALDH) is found in the mitochondria of hepatocytes and oxidizes acetaldehyde to acetate, which is then discharged into the bloodstream and oxidized to carbon dioxide in various extrahepatic tissues (Kong et al., 2019).

1.4.4 Pathophysiology of Dengue:

The majority of people infected with any of the four serotypes of dengue virus (DENV) are asymptomatic. The clinical spectrum of symptomatic infection includes everything from undifferentiated fever (viral syndrome), dengue fever (DF), and dengue haemorrhagic fever (DHF), to enlarged dengue syndrome with localized organopathy (unusual manifestations). DF might be asymptomatic, have atypical bleeding, whereas DHF can be asymptomatic, or have shock, resulting in dengue shock syndrome. The presence of acute and continuous fever for 2 to 7 days, haemorrhagic manifestations associated with thrombocytopenia (100,000 cells/c.mm or less), and haemoconcentration (haematocrit >20 % from baseline of patient or population of same age) are all required by the WHO criteria for the clinical diagnosis of DHF. Mucosal and cutaneous signs, as well as a positive tourniquet test, are all possible haemorrhagic manifestations. Hepatomegaly develops at some point during DHF and often precedes plasma leakage, making it a useful early indicator of plasma leakage. The pleural and peritoneal surfaces are the only places where plasma leaks. There is no vasculitis in DHF, therefore there is no damage to the vessel walls, and plasma leakage is caused by an increase in vascular permeability caused by cytokines. Further fluid loss from the intravascular compartment is facilitated by the migration of albumin and the resulting drop in intravascular oncotic pressure. The underlying Starling principle, which is based on the balance of oncotic and hydrostatic pressures, remains true in describing microvascular ultrafiltration. The glycocalyx, a gelatinous layer lining the vascular endothelium, is also involved in regulating fluid mobility by allowing albumin molecules to cling to its matrix, which when damaged allows albumin to leak into the extravascular compartment (Sellahewa, 2013) .

1.4.5 Pathophysiology of Hepatitis

The hepatitis B virus is made up of three parts: an exterior capsule holding HBsAg (hepatitis B surface antigen), an inner core containing HBcAg (HBV core antigen), and an inner core containing HBeAg (hepatitis B envelope antigen) (hepatitis B antigen). When HBV is detected in the blood, the body produces a cell-mediated immune response, sending cytotoxic T cells and natural killer cells to the virus and releasing inflammatory cytokines. The bigger the immunological reaction, the more likely the infection will be defeated. Because the HBsAg infiltrates the cytoplasm of the hepatocytes, they seem to have a "ground glass" appearance under histological examination — this is a distinguishing feature of HBV against other kinds of hepatitis (Hepatitis B: Pathophysiology, Protection, and Patients, 2013).

1.5 Impact of AI in Chronic Diseases

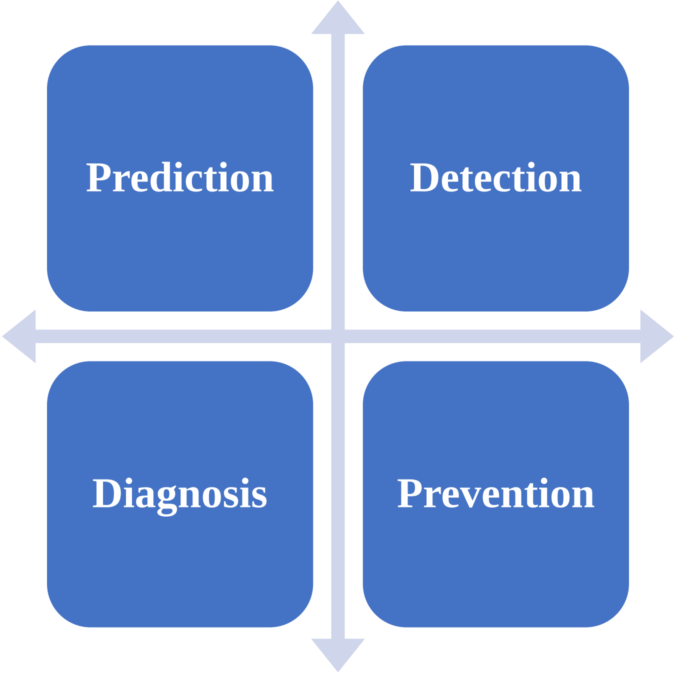

Figure 1: Impacts of AI in Chronic Diseases

In the past decade, the clinical approach to autoimmune diseases has evolved considerably and has been driven mainly by initiatives in the fields of accuracy, precise health and precision in public health. To comprehend and pragmatically adopt these techniques, the motivations, gaps and limits of precision medicine need to be understood. It is essential to build public trust and patients, but to realize that technologies like the artificial intelligence and machine learn continue to require contexts, which can be supplied only through human input or so-called advanced machine learning. The significance of genomics (such as autoantibody studies), microbiome and proteomics requires further optimization via research and pragmatic methods in its use in the sphere of precision medicine (Conrad et al., 2020). In recent years, artificial intelligence (AI) has developed fast across software, hardware and applied algorithms in a wide range of fields. The current advancements in biomedical AI applications, including illness diagnosis, live care, biomedical information treatment, and biomedical research, are being brought into light (Rong et al., 2020). Chronic, multi-factorial disorders are autoimmune diseases. With Machine Learning (ML), an artificial intelligence discipline, patterns may be extracted from patient data, and these patterns can be used to anticipate patient outcomes in order to enhance clinical care (Stafford et al., 2020).

1.6 Aims and Objectives of the Study

This analysis contains evaluation of several machine-learning algorithms for the diagnosis of various chronic diseases, such as it draws attention to the array of machine learning algorithms and techniques that are utilized for disease identification and decision-making. During this research, the student will collaborate directly with clinical workers, and use huge data sets to build chronic diagnostic and clinical-compatible machine-study algorithms. An essential element of this is the measurement of uncertainty, which allows for certain diagnoses. Automatic relevance determination will also be utilized to discover key diagnostic drivers capable of providing new biomarkers for illness.

Chapter 2: Methodology

Thorough literature review was done to obtain all the information used in this review paper. The information was collected from various credible sources, including different peer reviewed journals, online scholarly database, books and newspapers. Following are the list of the many databases that were search extensively for the present study.

- Journal Database
- Library Catalogue
- Subject Specific Professional Websites
- Newspaper Databases

On quest to collect as much relevant information regarding the role of AI in diagnosis of different chronic diseases thorough search of various journals, research papers and review articles from official sites and research databases was performed. Renowned and reliable databases such as PubMed, SCOPUS and Science Direct helped to gather the information of the role of AI in diagnosis of different chronic diseases for this review paper. Appropriate key terms, such as Artificial Intelligence, Chronic Diseases, Diagnosis, Machines Learning, Algorithm, Health Care; was used to collect relevant articles. Since this is a fairly niche topic, the number of articles is minimal to say the least. Depending on the title and key words, almost 150 articles have been screened through. Then, 100 articles were narrowed down after reading abstract. Finally, 87 articles were selected and methodically analysed to write this review paper. Mendeley software was used for proper and just referencing to as to be respectful with the work or the original authors.

Chapter 3: Discussion

For a great many years, data analysis and categorization have been beneficial. Humans have a remarkable ability to sense their surroundings (Sharma & Kaur, 2013).They take action in response to what they see in the surroundings. The interdisciplinary joint effort of machine learning, databases, and statistics transforms big data into chunks (Ricardo & dela Cruz, 2021). In today's medical world, illness diagnosis is a difficult undertaking. It is critical to know the precise diagnosis of patients based on clinical examination and evaluation. Decision support systems based on computers may play a critical role in successful diagnosis and cost-efficient management. The health-care industry generates many data on clinical evaluations and reports. Computer Aided Diagnosis (CAD) is an active area of study in medical imaging that is continuously developing. Since, mistakes in medical diagnostic systems might lead to significantly misleading medical treatments, as a result, major efforts to improve computer-assisted diagnosis applications have been conducted in recent years. In Computer Aided Diagnosis, machine learning is vital. Organs, for example, might not be appropriately identified after employing a simple calculation. As a result, learning from examples is essential to pattern recognition. Pattern recognition and machine learning have the potential to increase illness perception and diagnostic accuracy in the biomedical industry. They also help to ensure that the decision-making process is objective. Machine learning is a beneficial strategy for building sophisticated and efficient algorithms for the analysis of high-dimensional and multimodal bio-medical data. The effectiveness of several machine-learning algorithms for diagnosing diseases such as heart disease, diabetes, liver disease, dengue disease, and hepatitis disease is studied. It draws attention to a set of machine learning techniques and technologies that are used to analyse diseases and make decisions based on that information (Fatima & Pasha, 2017).

3.1 Different Types of Machine Learning Algorithm

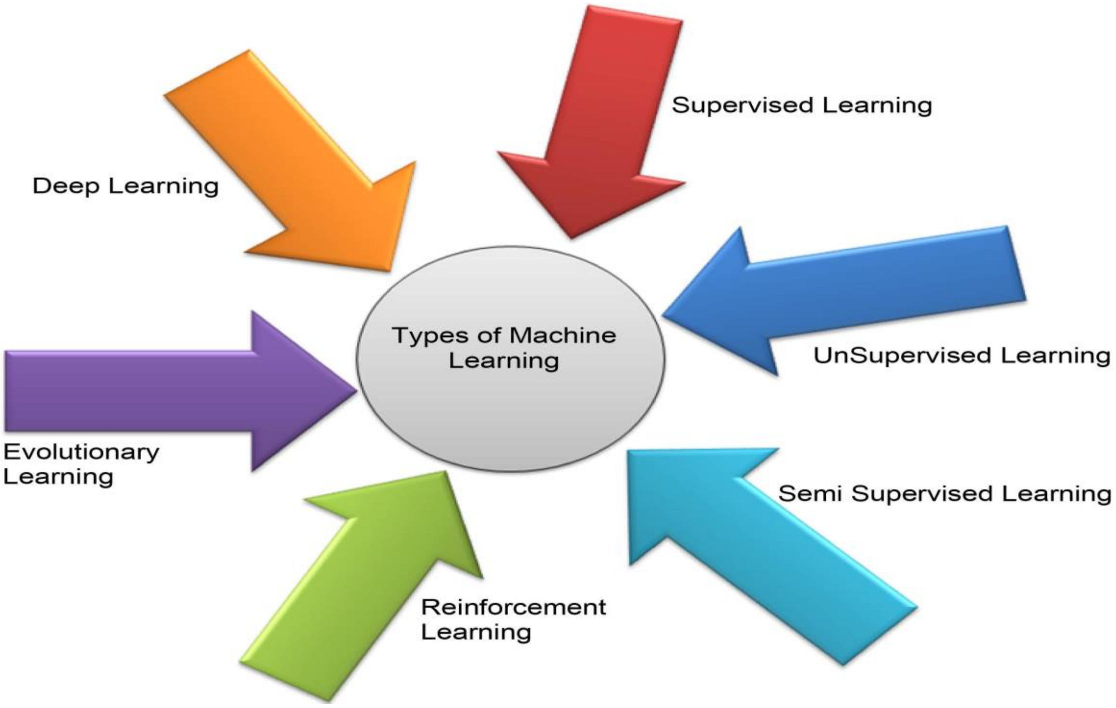

Figure 2: Types of Machine Learning Techniques

1) **Supervised learning:** Depending upon a set of examples of training it offers significant targets through which the algoritms respond correctly to all the feasible inputs. It can also be called Learning from Exemplers. Supervised learning is consisted of Classification and Regression.

Classification: A direct positive or negative feedback is found in this type of Supervised Learning. For example: ‘Yes’ or ‘No’.

Regression: A qualitative or quantitative result it obtained from regression.

2) Unsupervised learning: There are no appropriate solutions or objectives specified. Unsupervised learning aims to detect similarities between the input data and classifies the data based on these similarities. Density estimation is another name for this. Clustering is a feature of unsupervised learning. Clustering is what, that creates groups based on similarities.

3) Semi supervised learning: A supervised learning approach that is semi-supervised is a type of supervised learning technique. Unlabelled data was also employed for training purposes in this study (generally a minimum amount of labelled-data with a huge amount of unlabelled-data). Between unsupervised (unlabelled data) and supervised learning is semi-supervised learning (labelled-data).

4) Reinforcement learning: Behaviourist psychology encourages reinforcement learning. The algorithm is notified when the output is erroneous, but it is not notified how to rectify it. It must investigate and test numerous options until it discovers the correct solution. Learning with a critic is another term for it. It makes no suggestions for improvement. Reinforcement learning differs from supervised learning in that it does not provide correct or suboptimal input and output sets. The actions are well defined. Furthermore, it concentrates on online performance.

5) Evolutionary Learning: This biological evolution learning may be regarded of as a learning process: biological organisms are modified to improve their survival rates and chances of reproducing. We may utilize this model in a computer by employing the concept of fitness to determine how accurate the answer is.

6) Deep learning: This field of machine learning is built on a group of algorithms. These learning algorithms delivers unique abstraction in data. It employs a deep graph with several processing layers comprised of numerous linear and nonlinear transformation (Marsland.Ml-Alg-Perspect.09, n.d.).

3.2 Use of Machine Learning Algorithm in Diagnosis of Chronic Diseases:

The use of AI in the diagnosis process to assist medical practitioners might be extremely beneficial to the healthcare industry and the general health of individuals. The integration of artificial intelligence (AI) into current technical infrastructure speeds up the discovery of essential medical data from a variety of sources that is personalized to the patient's needs and the treatment process. Simultaneously, AI breaks down silos, allowing for cross-departmental knowledge sharing since data from all involved departments is taken into consideration. Additionally, AI provides findings based on a bigger population rather than subjective, personal experiences, gets comparable outcomes when using identical medical data, and is unaffected by situations, emotions, or time of day. Since the development of more realistic construction systems, the relevance of AI as a component of the diagnostic process has continuously grown. AI continues to elicit excitement and anticipation, with both researchers and practitioners examining the concept from many angles (Mirbabaie et al., 2021).

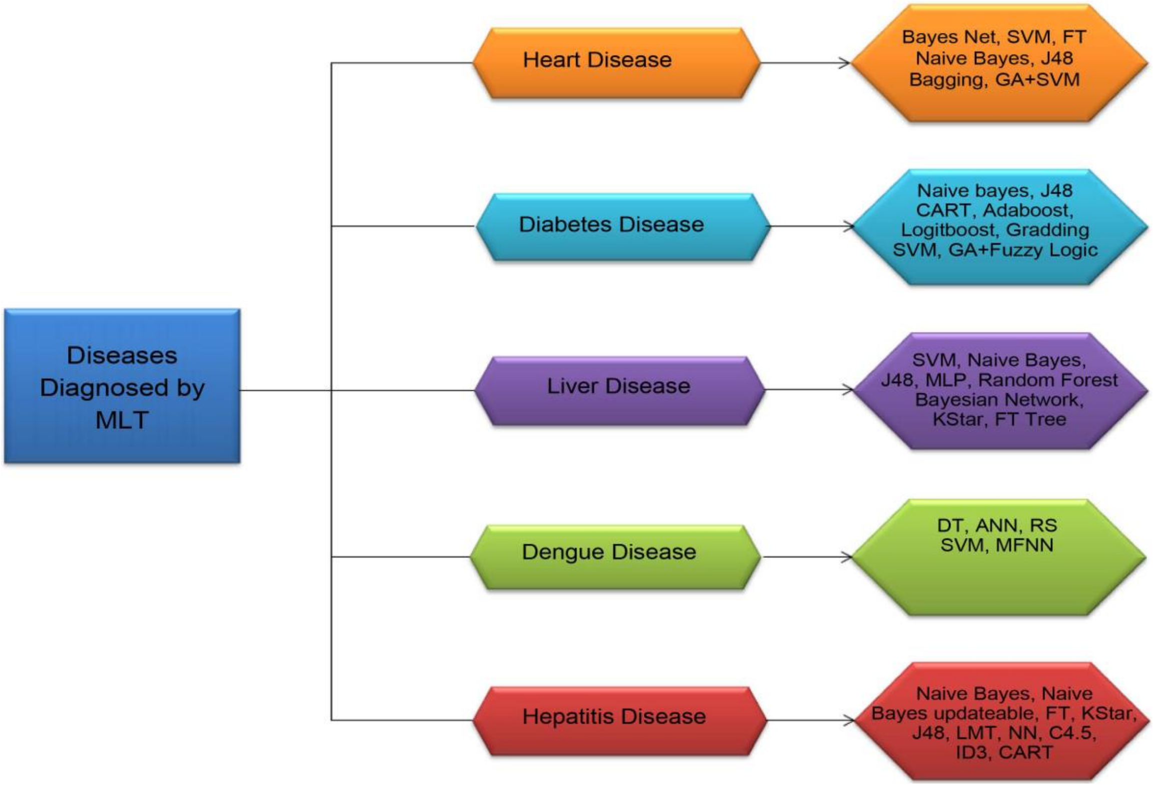

*Figure 3***:** *Uses of Machine Learning Algorithm in Diagnosis of Chronic Diseases*

3.3 Machine Learning in Heart Disease:

Hypertension, high blood pressure, old age, excessive cholesterol, and other conditions all contribute to the development of heart disease. In order to construct and refine our prediction models, we leverage the Cleveland heart disease dataset from the University of California, Irvine (uci) repository (Pandita, 2021). According to the WHO; there are different types of cardiovascular diseases that affect people**.** Majority of the related deaths, close to 85**%**, are caused by heart attack and strokes, while a third of the deaths occur prematurely in people under 70 years of age**.** Other types of heart diseases include the following**:**

Table 1: Types of Heart Diseases

Disease	**Description**
Coronary Heart Disease	Disease of the blood vessel supplying the heart muscle
Cerebrovascular Disease	Disease of the blood vessel supplying the brain
Peripheral Arterial Disease	Disease of the blood vessel supplying arms and legs
Rheumatic Heart Disease	Damage to the heart muscle and heart valves from rheumatic fever caused by the streptococcal bacteria
Congenital heart Disease	Malformations of heart structure existing at birth
Deep Vein Thrombosis and Pulmonary Embolism	Blood clots in the leg veins which can dislodge and move to the heart and lungs

This suggested technology detects and monitors coronary artery disease**.** UCI provided the Cleveland heart data set**.** There are 303 examples in this data collection, with 76 attributes/features**.** Out of the 76 features, 13 are utilized**.** For detection, two tests are run using three algorithms**:**

Bayes Net, Support Vector Machine, and Functional Trees FT**.** For detection, the WEKA tool is employed **(**Otoom et al**.**, **2015).** Using SVM approach, 88**.**3**%** accuracy is achieved after experimenting with the Holdout test**.** SVM and Bayes net both exhibit 83**.**8 **%** in the Cross Validation test**.** After employing FT, you will have an accuracy of 81**.**5 **%.** Using the Best First selection technique, the top seven characteristics are selected**.** A cross validation test is used for validation**.** Bayes Net achieved 84**.**5 **%** correctness, SVM 85**.**1 percent accuracy, and FT classification 84**.**5 **%** properly by executing the test on 7 best specified features (Fatima & Pasha, 2017; Vembandasamyp et al., 2015) In Naive Bayes, Bayes' theorem is applied**.** As a result, the independence assumption in Naive Bayes is quite strong**.** One of the premier diabetes research institutes in Chennai provided the data set used in this study**.** 500 patients make up the data set**.** Weka is used as a tool for categorization, and it uses a 70**%** Percentage Split to do so**.** With 86**.**419 percent accuracy, Naive Bayes is a good choice **(**Vembandasamyp et al**.**, **2015).** For data mining, the WEKA data mining tool is utilized, which offers a collection of machine learning algorithms**.** This viewpoint employs Naive Bayes, J48, and bagging**.** The UCI machine**-**learning lab offers a heart disease data collection with 76 variables**.** For prediction, only 11 characteristics are used**.** The accuracy of naive Bayes is 82**.**31 **%.** J48 has an accuracy rate of 84**.**35 **%.** Bagging achieves an accuracy of 85**.**03 **%.** On this data set, bagging provides a higher categorization rate **(**Chaurasia, **2012).** WEKA is used to apply Naive Bayes and SVM algorithms. A data collection of 500 patients was obtained from the Chennai Research Institute. There are 142 persons with the condition, and 358 patients do not have it. The Naive Bayes Algorithm achieves a 74 percent accuracy rate. SVM has the highest accuracy, with a score of 94.60 (Tan et al., 2009). The Genetic Algorithm (G.A) and Support Vector Machine (SVM) are two machine-learning algorithms that are combined successfully utilizing the wrapper method in the hybrid technique. In this study, the LIBSVM and

WEKA data mining tools were employed. This experiment uses five data sets from the UC Irvine machine-learning repository: iris, diabetes, breast cancer disease, heart disease, and hepatitis disease. For cardiac illness, an accuracy of 84.07% was achieved using a hybrid GA and SVM technique. The accuracy for the diabetes data set is 78.26%. Breast cancer accuracy is 76.20%. Hepatitis illness has an accuracy rate of 86.12%.

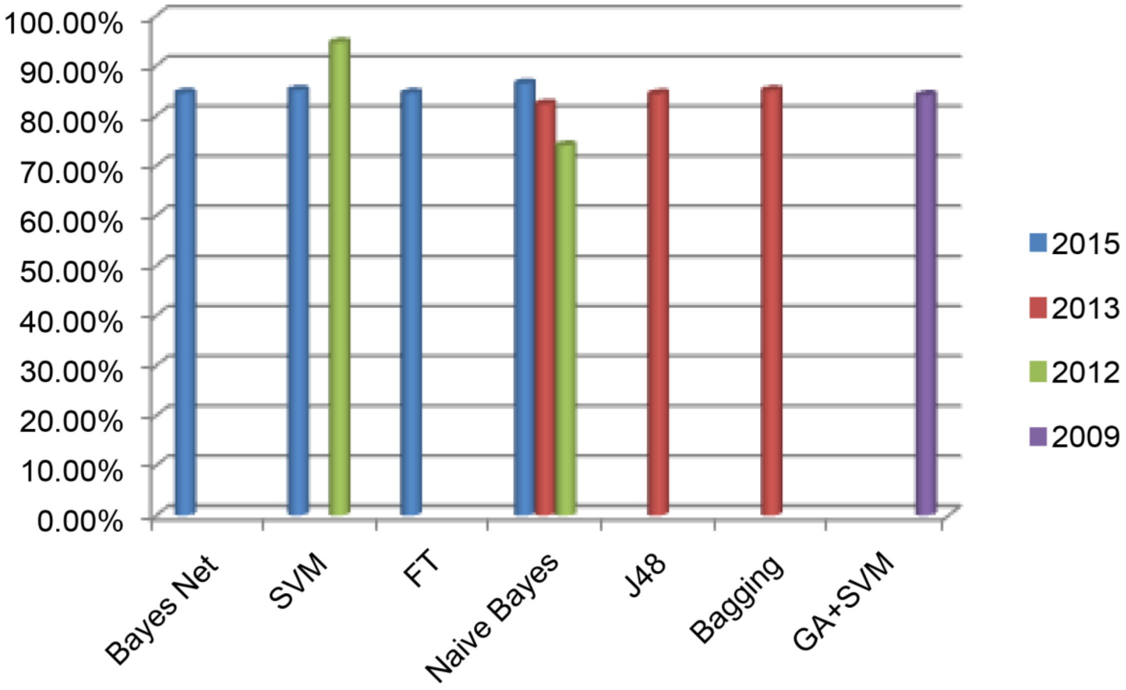

Figure 4: Machine Learning Algorithm's Accuracy to Detect Heart Disease

3.3.1 Analysis:

According to the available research, SVM has the best accuracy of 94.60 percent. SVM has shown to be effective in a variety of applications. SVM properly responds to attributes or characteristics were utilized (Parthiban & Srivatsa, 2012). A SVM variation was employed named SMO (Otoom et al., **2015**). It also uses the FS approach to locate the greatest characteristics. SVM reacts to these characteristics and has an accuracy of 85.1 percent, which is lower than in 2012. The data sets used for training and testing are distinct, as are the data kinds (Fatima & Pasha, 2017).

3.3.2 SVM Advantages and Disadvantages:

Advantages: Construct robust classifiers with less overfitting and accurate classifiers.

Disadvantages: It is a binary classifier, which has several drawbacks. It is used to classify multi-classes. Pair-wise categorization is an option. It has a high computational cost, so it is slow (Karamizadeh et al., **2014).**

3.4 Machine Learning in Diabetes:

Using Decision Tree and Naves Bayes diabetes disease can be predicted according the author. When insulin production is inadequate or insulin is used incorrectly, diseases develop. The Pima Indian diabetes data collection was used in this study. The WEKA data-mining program was used to run a number of experiments. Cross validation predicts better than percentage split (70:30) in this data set. Using Cross Validation and Percentage Split, J48 achieves accuracy of 74.8698 percent and 76.9565 percent, respectively. Using PS, Naive Bayes shows 79.5652 percent accuracy. Using a percentage split test, algorithms demonstrate the best accuracy (Iyer et al., **2015).** A brief description of Meta learning algorithm for diagnosis of diabetes is stated. The data set used is Pima Indian diabetes, which was obtained from the UCI Machine Learning lab. For analysis, WEKA is employed. To determine if a patient has diabetes or not, CART, Adaboost, Logiboost, and grading learning algorithms are employed. Based on right or erroneous categorization, the experimental findings are compared. CART has an accuracy rate of 78.646**%.** The Adaboost achieves a precision of 77.864 percent. The accuracy of Logiboost is 77.479 percent. The proper categorization rate for grading is 66.406 percent. When compared to other approaches, CART has the highest accuracy of 78.646 percent and the lowest misclassification rate of 21.354 percent (Tech & Rathore, n.d.). SVM is the machine learning technology utilized by the scientist in this experiment. In SVM, the RBF kernel is utilized to classify the data. The University of California, Irvine's machine learning lab has released the Pima Indian diabetes data collection. To conduct the experiment, MATLAB 2010a is utilized. SVM provides 78 percent accuracy (Anuja Kumari & Chitra, n.d.).

There are three forms of diabetes. Kind-1 diabetes is the most common type, followed by Type-2 diabetes and gestational diabetes. Insulin resistance causes type 2 diabetes to develop. The data set consists of 415 instances, and the data was collected from various segments of society in India for variety's sake. The model was created using MATLAB and a SQL server. Naive Bayes is capable of predicting 95**%** of the time (Sarwar & Sharma, 2012). The proposed model incorporates both GA and fuzzy logic. It is used for selecting the optimal subset of characteristics as well as improving classification accuracy. For this experiment, a dataset with 8 characteristics and 769 examples was obtained from the UCI Machine Learning Laboratory. The implementation is done in MATLAB. Only the three top traits / qualities are chosen using a genetic algorithm. The fuzzy logic classifier uses these three features to give 87 percent accuracy. The cost is about half of what it was originally (Ephzibah, **2011).**

3.4.1 Analysis:

Diabetes illness may be diagnosed using a Naive Bayes-based method. In 2012, Naive Bayes had the greatest accuracy of 95 percent. The results suggest that this system can make accurate predictions with a small margin of error, and that this approach is useful in diagnosing diabetes. However, in 2015, the accuracy of Naive Bayes is poor. It has an accuracy of 79.5652 percent or 79.57 percent. For the design and testing of this suggested model for detecting Diabetes illness, more training data would be required. Figure 4 depicts the Accuracy graph of Algorithms for Diabetes Disease Diagnosis through Time (Fatima & Pasha, 2017)

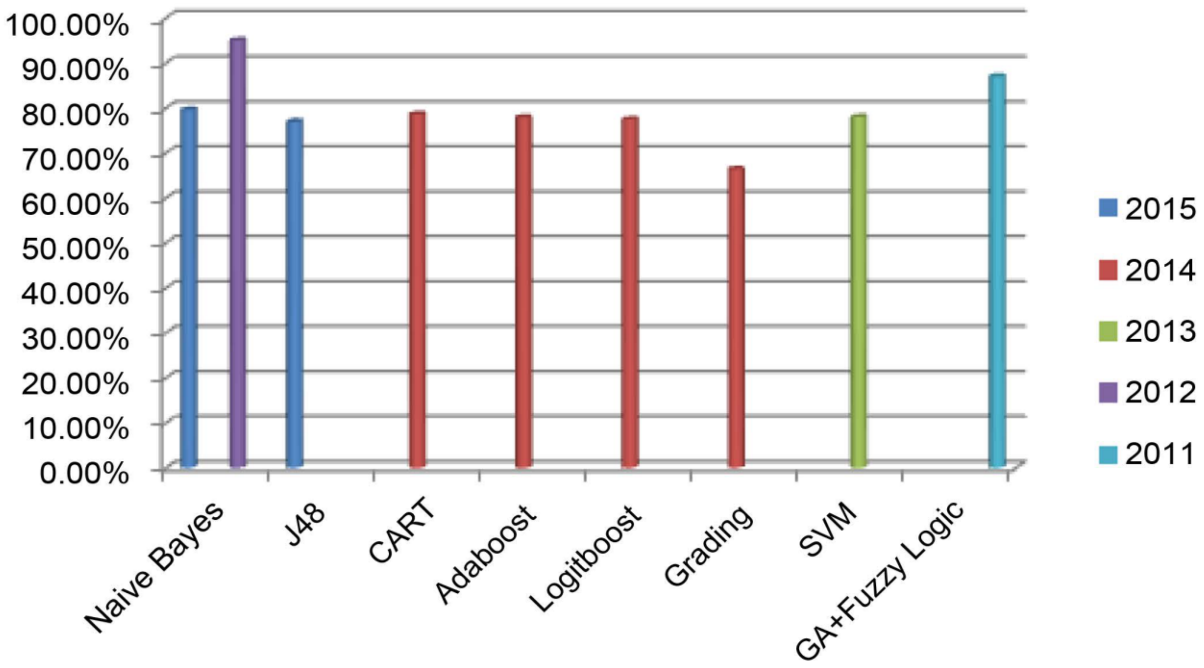

Figure 5: Accuracy of Machine Learning Algorithm to Detect Diabetes

3.4.2 Advantages and Disadvantages of Naive Bayes:

Advantages: It improves classification performance by removing irrelevant features**.** It performs admirably**.** It takes less time to compute**.**

Disadvantages: To achieve solid outputs, this approach requires a big quantity of data**.** They are layabouts because they save every training dataset (Archana & Elangovan, 2014)

3.5 Machine Learning in Liver Diseases:

The hepatic system is the second largest internal organ in the human body, and it plays an important part in metabolism as well as various other crucial processes, such as red blood cell decomposition **(**Karthik et al**.**, **2011).** It weighs around three pounds**.** The liver is responsible for numerous vital activities in the body, including digestion, metabolism, immunity, and nutrition storage**.** Because of these duties, the liver is a vital organ; without it, bodily tissues would soon perish due to a lack of energy and nutrition**.** A number of variables contribute to the development of liver disease**.** Some mostly mentioned are**:**

- Family history of liver disease
- Smoking
- Consumption of alcohol
- Intake of contaminated food
- Obesity
- Diabetes

(Mohan, **2015).** Three major liver diseases such as Liver cancer, Cirrhosis and Hepatitis with the help of distinct symptoms can be predicted. For disease prediction, the Nave Bayes and FT Tree algorithms has been employed**.** Based on their classification accuracy measure, these two methods were compared**.** Based on the results of the experiments, it determined that the Nave Bayes algorithm was the superior method for predicting diseases with the highest prediction performance **(**Dhamodharan, **2014).** The Principal Component Analysis technique is used to extract feature vectors**.** Because the LS**-**SVM technique is sensitive to changes in parameter values, the Modified**-**PSO Algorithm was utilized to find the best LS**-**SVM parameter values in less iterations**.** On the benchmark HCV data set from the UCI collection of machine learning databases, the suggested system was developed and assessed**.** Another classification method that used PCA and LS**-**SVM was compared to it**.** According to the testing data, the suggested system outperformed the other systems in terms of classification accuracy (Soliman & Elhamd, 2014). Liver disease can be predicted by using Support vector machine and Naive Bayes Classification algorithms. UCI provided the ILPD dataset. 560 instances and 10 characteristics make up the data set. The precision and timeliness of the execution are compared. In 1670.00 milliseconds, the Naive Bayes algorithm achieves 61.28 percent accuracy. In 3210.00 ms, SVM achieves 79.66 percent accuracy. MATLAB is utilized for implementation. When compared to Naive Bayes, SVM has the greatest accuracy in predicting liver illness. Naives Bayes is faster than SVM in terms of execution time (Mohan, 2015). UCI provided the used data set. 5 sophisticated strategies and the WEKA data-mining tool. This experiment employs the classifiers J48, MLP, Random Forest, SVM, and Bayesian Network. All methods are applied to the original data set in the first phase, and the percentage of accuracy is calculated. In the second phase, a feature selection approach is used to the whole data set to extract a meaningful subset of liver patients, and all of these methods are then utilized to evaluate that subset. They compare the results before and after selecting features in the third stage. Following FS, algorithms with the best accuracy are J48 (70.669%), MLP (70.8405%), SVM (71.3551%), Random Forest (71.8696%), and Bayes Net (69.1252%) (Gulia et al., n.d.). The data set comes from UCI and includes 345 occurrences and 7 characteristics. Using the WEKA tool, ten cross validation tests are performed. In 0 seconds, Naive Bayes provides 96.52 percent correctness. In 0.2 seconds, the FT tree achieves 97.10 percent accuracy. In 0 seconds, the K star method correctly classifies the instances by 83.47 percent. In comparison to other data mining methods, FT tree provides the greatest classification accuracy on the liver disease dataset (Rajeswari & Reena, 2010).

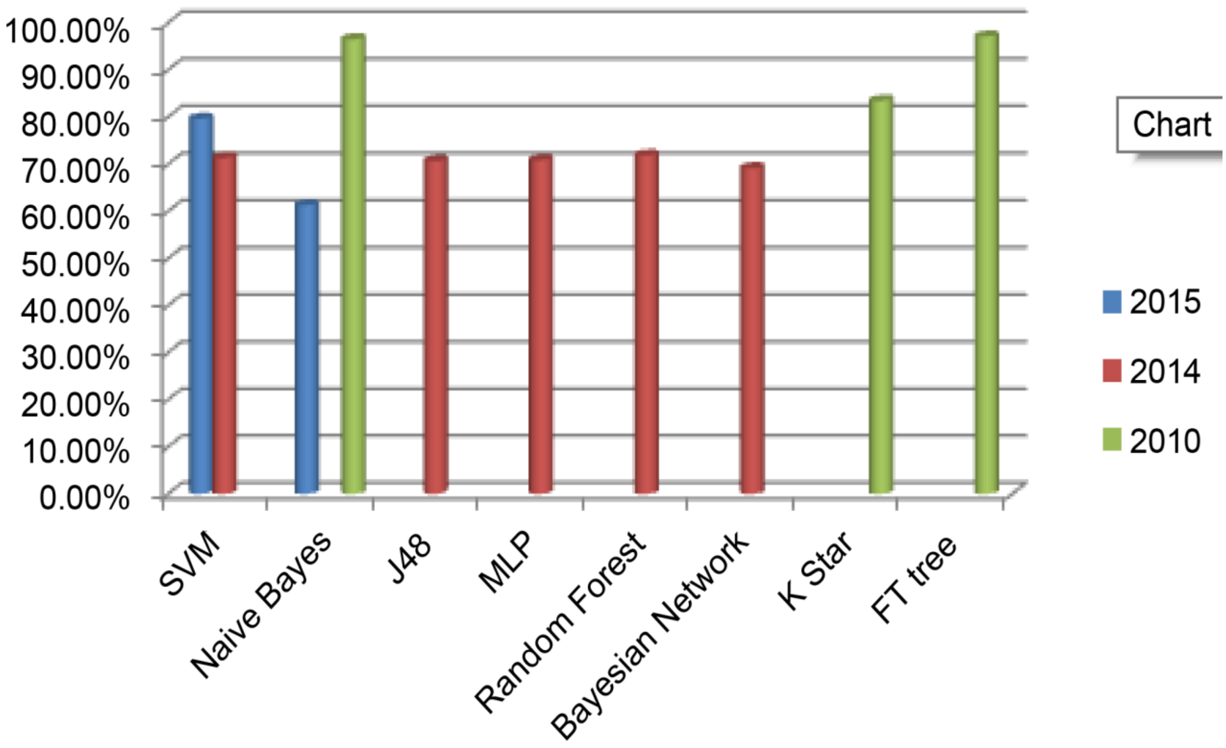

Figure 6: Accuracy of Machine Learning Algortihm to Detect Liver Disease

3.5.1 Analysis:

In comparison to the other algorithms, the FT Tree Algorithm provides the best results for diagnosing liver disease. When the FT tree algorithm is used on a dataset of liver disease, the time it takes to get a result or build a model is much faster than with other algorithms. It demonstrates improved performance based on its attribute. This algorithm correctly classified all of the attributes and has a 97.10 percent accuracy rate. As a consequence of the findings, this Algorithm plays a critical part in defining improved data set categorization accuracy (Fatima & Pasha, 2017).

3.5.2 Advantages and Disadvantages:

Advantage: Convenient to perceive and comprehend; speedy forecast

Disadvantage: Calculations are difficult, especially when values are unclear or several outcomes are related.

3.6 Machine Learning in Hepatitis:

Naive Bayes, Naive Bayes updatable, FT Tree, K Star, J48, LMT, and NN are data mining techniques used to diagnose hepatitis illness. The data set for hepatitis illness was obtained from the UCI Machine Learning library. The accuracy and timeliness of the classification results are measured. Using neural networks and the WEKA data-mining tool, a comparative analysis is performed. The results obtained by employing neural connections are inferior to those obtained by using WEKA's techniques. The second approach employed in this analysis of hepatitis illness detection is rough set theory, which is implemented using WEKA. Rough set technique outperforms NN, especially when it comes to medical data processing. In 0 seconds, Naive Bayes delivers an accuracy of 96.52 percent. A whopping 84 percent The Naive Bayes Updateable method achieves accuracy in 0 seconds. FT Tree has an accuracy of 87.10 percent in 0.2 seconds. K star has an accuracy rate of 83.47 percent. The K star algorithm takes 0 seconds. J48 achieves an accuracy of 83 percent with a classification time of 0.03 seconds. LMT has an accuracy of 83.6 percent and takes 0.6 seconds. The accuracy of the neural network is 70.41 percent. In the rough set technique, the best classification method is Naive Bayes. It provides great precision in a short amount of time (Mutaher Ba-Alwi & Hintaya, 2013). The fact that these classifiers only require a little quantity of data for classification is a major benefit. "A, B, C, D, and E" are the several types of hepatitis. Several hepatitis viruses cause these. In this study, the free source program Rapid Miner was employed. UCI provided the hepatitis data set. There are 20 characteristics and 155 occurrences in the data set. In this experiment, 15 characteristics are employed. The accuracy of the Naive Bayes classifier is 97 percent. Three-layered feedforward NNs are utilized and they are trained using the Back propagation approach with 155 examples. A 98 percent accuracy rate is achieved (Karlik, **2011).** The UCI hepatitis patient data collection is used in this investigation. In this analysis, the WEKA tool is employed. CART has performed admirably when dealing with missing values. As a result, the CART algorithm has a classification accuracy of 83.2 percent. The ID3 Algorithm has a 64.8 percent accuracy rate. The C4.5 algorithm achieves a 71.4 percent success rate. The CART method generates a binary decision tree (DT) with just two or no children. C4.5 and ID3 create a DT that can have two or more offspring. In terms of accuracy and time complexity, the CART algorithm excels (Sathyadevi, **2011).**

3.6.1 Analysis:

For the diagnosis of various illnesses, a variety of algorithms has been utilized. Table 5 depicts the overall picture. Feed forward neural networks with back propagation had the maximum accuracy of 98 percent in detecting Hepatitis illness. Because this model uses an error, back propagation approach to train a three-layered feed forward neural network. Back propagation training using the rule of delta learning is an iterative gradient approach that aims to reduce the RMSE (root mean square error) between the real and intended outputs of multi-layered feed-forward neural networks. Every layer is linked to the one before it and has no additional connections. Naive Bayes provides the second-best result. However, when compared to neural networks, Naive Bayes takes less time to develop a model (Fatima & Pasha, 2017).

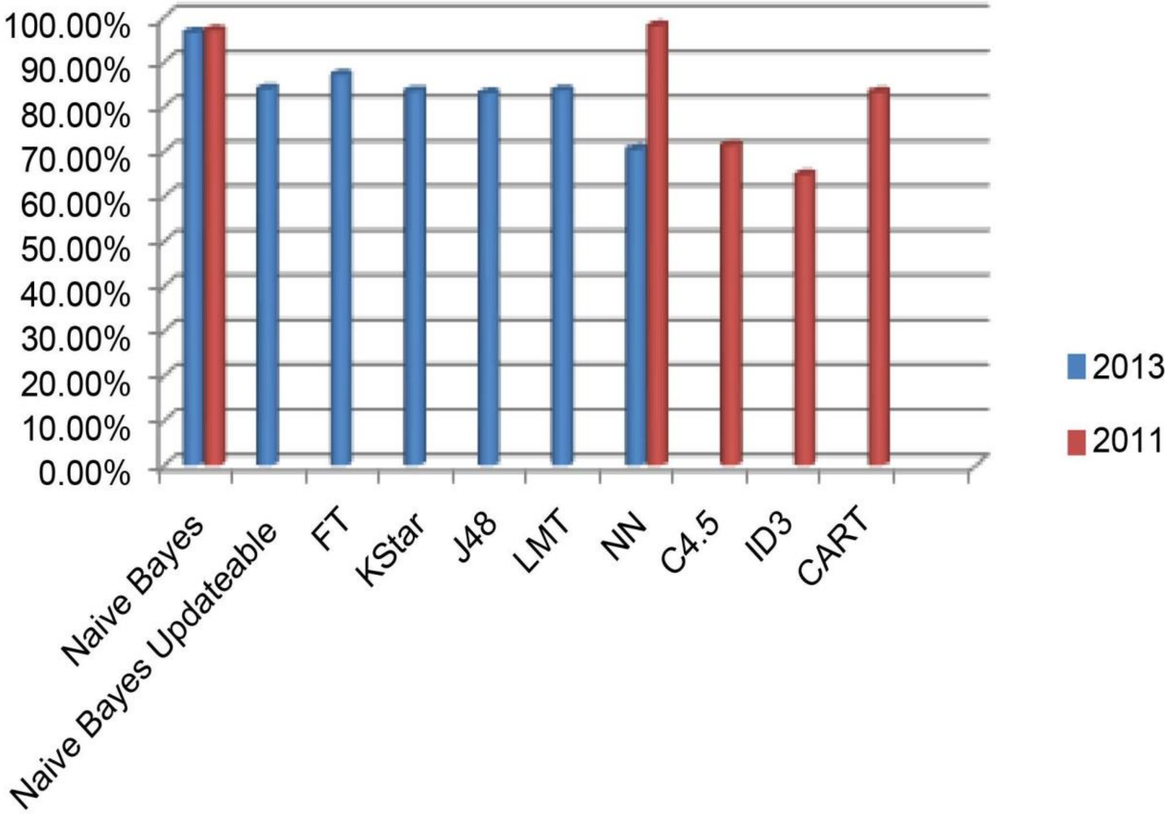

Figure 7: Accuracy of Machine Learning Algorithm to Detect Hepatitis

3.6.2 Advantages and Disadvantages of NN:

Advantages: Adaptive Learning, Self-Organization, Real Time Operation Fault
Tolerance via Redundant Information Coding.

Disadvantages: Less overfitting necessitates a significant amount of computing work. A big sample size is required. It takes a long time. Engineering judgment does not develop the relationships between the input and output variables, resulting in a black box model.

3.7 Machine Learning in Dengue:

Dengue fever is evolving into a highly infectious illness. It causes problems in nations with humid climates, such as Thailand, Indonesia, and Malaysia. The classification algorithms used in this study to predict dengue disease are Decision Tree (DT), Artificial Neural Network (ANN), and

Rough Set Theory (RS). The data was obtained from the Selangor State Public Health Department. The WEKA data-mining program is utilized, along with two tests (10 Cross-fold Validation and Percentage Split). DT has a 99.95 percent accuracy, ANN has 99.98 percent correctness, and RS has 100 percent accuracy utilizing 10-Cross fold validation. Both the Decision Tree and the Artificial Neural Network are 99.92 percent correct after using PS. RS has a 99.72 percent accuracy rate (Tarmizi et al., 2013). Support Vector Machine is a data-mining algorithm utilized by these researchers. The data for this study came from the King Institute of Preventive Medicine, as well as surveys of various hospitals and laboratories in Chennai and Tirunelveli, India. There are 29 qualities and 5000 samples in it. R project version 2.12.2 is used to evaluate the data. SVM achieves a 0.9042 accuracy level (Shameem Fathima & Manimeglai, 2012). For detection, the suggested approach relies solely on clinical indications and symptoms. The information was acquired from 252 hospitalized patients, 4 of whom had Dengue fever and 248 of whom had DHF (Dengue haemorrhagic fever) (dengue haemorrhagic fever). The neural network toolkit in MATLAB is utilized. In this experiment, the Multilayer feed-forward neural network (MFNN) algorithm is applied. In DF and DHF, MFNN reliably predicts the day of effervescence of fever with a 90% accuracy rate (Ibrahim et al., 2005).

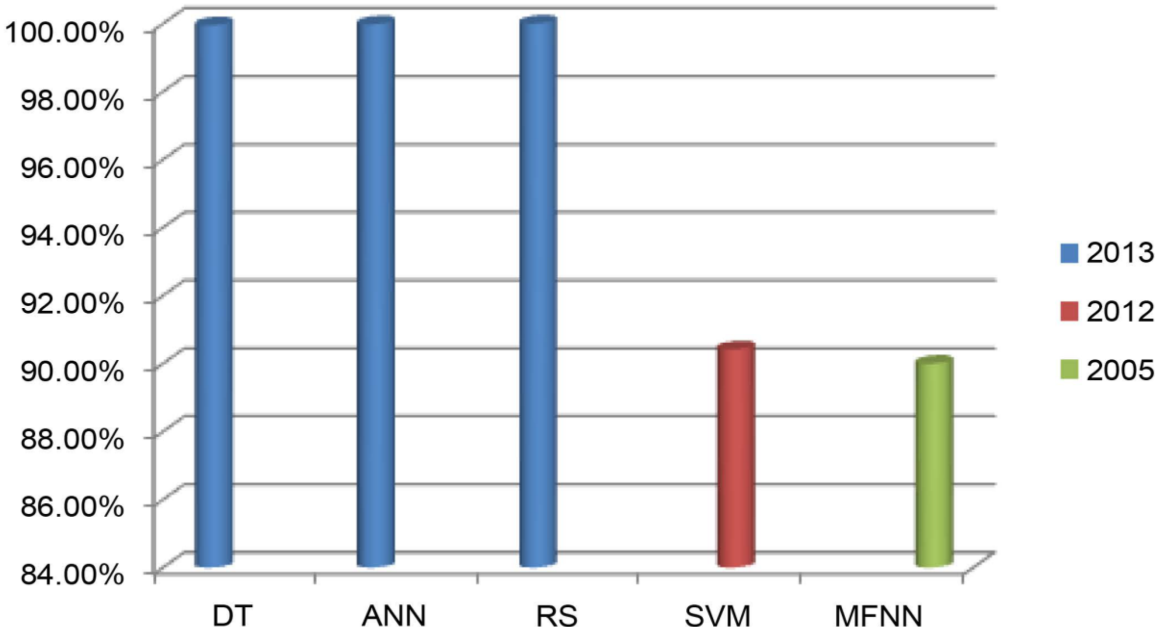

Figure 8: Accuracy of Machine learning Algorithm to Detect Dengue

3.7.1 Analysis:

Dengue illness is diagnosed using a variety of machine learning approaches. Dengue fever is one of the most dangerous communicable illnesses. As seen in Table 4, when it comes to detecting dengue fever, the RS theory outperforms the other methods. Researchers employed many algorithms between 2005 and 2012 but could not get the best results or improvements. The use of RS in 2013 enhanced accuracy. It can deal with ambiguity, noise, and missing data. The Rough set theory is used to classify the data in the developed RS classifier. The choice of attribute allows the classifier to outperform the competition. RS is a promising rule-based technique that provides useful data. In terms of time, RS outperforms neural networks. Building a model with NN takes a long time. The DT algorithm is both sophisticated and expensive. RS does not require any initial or extra data input, but Decision Tree requires (Fatima & Pasha, 2017).

3.7.2 Advantages and Disadvantages:

Advantage: It is simple to comprehend and offers a clear picture of the achieved product**.** It assesses the relevance of data**.** It may be used to analyse both qualitative and quantitative data**.** It uncovers the patterns that have been buried for a long time**.** It also locates the smallest amount of data**.** It can uncover relationships that statistical approaches are unable to detect**.**

Disadvantages: Despite the fact that it has few drawbacks, it is not commonly utilized**.**

(Fatima & Pasha, 2017).

Chapter 4: Observation

Several machine-learning algorithms do exceptionally well in the diagnosis of heart, diabetes, liver, dengue, and hepatitis disorders. The Naive Bayes Algorithm and SVM are commonly used algorithms for illness identification, according to the literature. In comparison to other algorithms, both approaches are more accurate. Prediction can also benefit from the deployment of an artificial neural network. It also displays the maximum output; however, it takes longer than other methods. Trees algorithms are also utilized, but because to their complexity, they have not gained widespread popularity. They also showed improved accuracy when responding properly to the data set's properties. Although the RS hypothesis is not commonly employed, it provides the best results. Statistical models for estimating that are incapable of producing good performance outcomes have saturated the assessment field. Statistical models are unable to handle categorical data, missing values, or big data points. All of these factors contribute to the significance of MLT. Many applications, such as image identification, data mining, natural language processing, and illness diagnosis, rely on machine learning. ML can help in all of these areas. This study examines several machine-learning algorithms for diagnosing ailments such as heart disease, diabetes, liver disease, dengue fever, and hepatitis. Many algorithms have produced positive outcomes because they properly detect the property. According to a recent study, SVM has a 94.60 percent accuracy rate for detecting heart problems. Naive Bayes accurately diagnoses diabetes. It has a classification accuracy of 95%. For the diagnosis of liver illness, FT has a 97.10 percent accuracy rate. The RS hypothesis achieves 100 percent accuracy in detecting dengue illness. The hepatitis illness is appropriately classified by the feed forward neural network, which has a 98 percent accuracy rate. The benefits and drawbacks of various algorithms are highlighted in this survey. The improvement graphs of machine learning algorithms for illness prediction are detailed. Based on the results of the investigation, it is obvious that these algorithms improve the accuracy of various illnesses. This survey document also includes a set of tools created by the AI community. These tools are quite valuable for analysing such issues and give opportunities for better decision-making.

Chapter 5: Limitations of AI in Diagnosis of Diseases

5.1 Software Security:

Under design assaults, even strong algorithm programs are susceptible (Mozaffari-Kermani et al., 2015; Mygdalis et al., 2020; Nguyen et al., 2019; Yuan et al., 2019). Despite its strong performance in the first design examination, the AI system's performance in a targeted design confrontation is frequently inadequate. In fact, provided the attacker knows everything about the trained neural network model, all phases of the AI algorithm generation process will be targeted (training data, model architecture, hyper parameters, number of layers, activation function, and model weights) (Yuan et al., 2019). A false-positive attack can be used to create a negative sample, or a false-negative assault can be used to generate a positive sample, resulting in system classification confusion. Attacks can be launched even if the target model's structure and parameters, as well as the training data set, are unknown (Yuan et al., 2019; Zhang & Wu, 2019). Errors will arise in the system even if no external disturbance is present. Because to changes in illness patterns, missing data, and autonomous update mistakes, the initial algorithm will eventually stray from the proper path (Davis et al., 2017; Dawson & Abkes, n.d.; Nemati et al., 2016).

5.2 Human Factors:

Software codes are used to implement AI nowadays. Engineers will unavoidably make blunders while dealing with thousands of codes. Following patches and upgrades, an AI system can be enhanced. However, with AI algorithms employed in the medical industry, such mistakes may put patients' health at risk. The efficacy of the AI system, not the security, is frequently the end aim to which developers pay attention (Belard et al., 2017).

Doctors are human, and when the AI system confirms their diagnosis, busy professionals stop looking for other options (Carter et al., 2020; Parasuraman & Manzey, 2010).

5.3 Protection of Human Genetic Resources

To accomplish the mapping of the human genome, all human chromosomal DNA sequences have been mapped, and a common information system has been built. The link between function and genes, particularly the association between genes and illnesses, is the emphasis of this information system (S. Lander et al., 2001). Over the last three decades, scientists have successfully mapped the completely human genome, allowing for considerably more thorough genetic and cellular study into clinical illnesses. Human genetic resources, on the one hand, are extremely useful in medical diagnosis and therapy (E. S. Lander, 2011).

However, if unlawfully employed, such resources have the potential to be devastating for all human beings. The People's Republic of China's Ministry of National Defence has announced the development of a biologic weapon known as a "gene weapon," which comprises of bacteria, insects, or microorganisms bearing disease-causing genes via genetic engineering changes with deadly military consequences. The study of genetic traits might be used to create genetic mutations in certain ethnic groups; as a result, safeguarding human genomic resources is critical. In addition, as the Internet and gene testing technology advance, a growing number of individuals are receiving genetic testing to aid in diagnosis and treatment, with the findings being documented in hospitals or testing companies. This puts data at danger of being exposed to the wrong people, and it might lead to insurance or job discrimination. As a result, legal safeguards for patient privacy in DNA collection, transmission, and storage should be strengthened (L. Jiang et al., 2021).

Chapter 6: Future Prospects

6.1 Advancements and Explicability

The construction of a specialized algorithm for the identification of a certain disease has been the focus of recent research (Parisi et al., 2018; R, 2019). Furthermore, certain algorithms have received more attention than others. Future study should look at prospective enhancements by merging several existing algorithms to get better results than observing a single method in isolation (Gil & Johnson, n.d.; Yu et al., 2016). More study in the field of deep learning for illness diagnostics is urgently recommended so that enormous volumes of medical data may be handled faster and more gratifying findings can be achieved. The fact that the outputs of AI remain a black box to humans is an important technical limitation of the more sophisticated yet performant deep learning systems (Goodfellow et al., n.d.; Jin et al., 2020; Shah et al., 2018). The results are not always apparent, which makes it difficult to grasp from the AI's actions on the one hand and difficult to create confidence in the system on the other. Future research should concentrate on making AI-derived results more understandable and explainable. A transparent prediction-making method establishes a trusting relationship between AI and medical professionals (Maghooli et al., 2016).

6.2 Corroboration and Portability

AI algorithms aimed at assisting in the diagnosis process have recently been developed, and they usually evaluate a single dataset based on textual input. There have been various studies that have shown positive findings (Devikanniga & Raj, 2018). However, there is a danger that the results will not be relevant to other domains and will only provide appropriate results for a single application. These findings will need to be confirmed in a larger patient group in the future. This may be accomplished by combining diverse and bigger datasets (i.e., N > 1,000 samples) with a variety of formats, including X-ray pictures and ultrasounds, which are currently underutilized. Additionally, massive datasets should be divided using crossvalidation to mimic external validity and get better results (Li et al., 2019; Mazzocco & Hussain, 2012; Shrivastava et al., 2015). Simultaneously, discoveries must be applied to other diseases, such as other forms of cancer, as well as to other clinical applications. Furthermore, one must inquire as to why scientific data has not yet been widely integrated into disease diagnosis, such as in hospitals or other clinical settings. The concern is whether AI methods will hold up in real-world circumstances. As a result, mobility is a critical component in AI's future in medical diagnosis (Alvi et al., 2021; Catto et al., 2010; Shrivastava et al., 2015).

6.3 Integration and Collaboration

The diagnosis of various diseases is a highly subjective and cognitively hard undertaking that is influenced by the clinician's personal experience and varies depending on the clinician's emotions and mental state medical specialists are aided by AI in the diagnosing process, perhaps leading to better results (Putatunda, 2019; Sawarkar & Pande, n.d.). However, present AI in disease diagnostics research has focused solely on technical implementations, rather than how AI may be incorporated into existing technical infrastructure. Recent AI advancements may provide significant findings for health diagnosis; however, how data will be presented to medical practitioners is yet uncertain. Human-AI cooperation is still required for diagnostics (Akhavan Darabi et al., 2014). This necessitates a greater focus on integration strategies, particularly the creation of user-friendly interfaces for different devices (Devi et al., 2019; Singh et al., 2021). To improve patient outcomes, researchers and healthcare practitioners should work together to create AI. Instead of focusing just on detecting a certain ailment, scientific attempts may go even farther by establishing a system that could aid in the complete diagnostic process (Mishra et al., 2018).

Furthermore, as we have stated, human-AI collaboration can provide superior results. Due to a paucity of practical examples, future research should focus on the collaborative features of human-AI collaboration in the diagnosis process. It's possible that virtual human–AI collaborations surpass people operating alone (Mani et al., 2014; Tenório et al., 2011).

Chapter 7: Conclusion

Illustrating the intensity of studies, highly correlated area and an overview of unexplored research is helpful for future development of AI in the diagnosis of diseases**.** On a practical level, practitioners understand the extent to which AI improves the diagnostic process and how the overall healthcare system benefits from it**.** Medical experts can realise how AI can be applied in diagsoning diseases, which could result in having suitable suggewstions for further developing AI approaches**.** Moreover, medical professionals comprehend which challenges are still needed to be triggered before diseases are diagnosed in collaboration with AI**.** In terms of implications of society, readers realise that AI is likely to be used in healthcare to diagnos diseases or atleast to assist in the process**.** Nevertheless, the application of AI as a component of the diagnostic process provides opportunities for innovative digital health and is simultaneously able to ensure enhanced patient outcomes**.**

References

Akhavan Darabi, S., Teimourpour, B., Zolnoori, M., & heydarnejad, D. (2014). Case-Based-Reasoning System for Feature Selection and Diagnosing Disease; Case Study: Asthma. 5(5). www.iiste.org

Alvi, R. H., Rahman, M. H., Khan, A. A. S., & Rahman, R. M. (2021). Deep learning approach on tabular data to predict early-onset neonatal sepsis. Journal of Information and Telecommunication, 5(2), 226–246. https://doi.org/10.1080/24751839.2020.1843121

Alwan, Ala., & World Health Organization. (2011). Global status report on noncommunicable diseases 2010. World Health Organization.

Anuja Kumari, V., & Chitra, R. (n.d.). Classification Of Diabetes Disease Using Support Vector Machine. International Journal of Engineering Research and Applications (IJERA) Www.Ijera.Com, 3, 1797–1801. www.ijera.com

Archana, S., & Elangovan, K. (2014). Survey of Classification Techniques in Data Mining. International Journal of Computer Science and Mobile Applications, 2, 65–71. www.ijcsma.com

Banday, M. Z., Sameer, A. S., & Nissar, S. (2020). Pathophysiology of diabetes: An overview. Avicenna Journal of Medicine, 10(04), 174–188. https://doi.org/10.4103/ajm.ajm_53_20

Belard, A., Buchman, T., Forsberg, J., Potter, B. K., Dente, C. J., Kirk, A., & Elster, E. (2017). Precision diagnosis: a view of the clinical decision support systems (CDSS) landscape through the lens of critical care. In Journal of Clinical Monitoring and Computing (Vol. 31, Issue 2, pp. 261–271). Springer Netherlands. https://doi.org/10.1007/s10877-016-9849-1

Buchanan, R., & Sinclair, J. M. A. (2021). Alcohol use disorder and the liver. Addiction, 116(5), 1270–1278. https://doi.org/10.1111/add.15204

Carter, S. M., Rogers, W., Win, K. T., Frazer, H., Richards, B., & Houssami, N. (2020). The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. Breast, 49, 25–32. https://doi.org/10.1016/j.breast.2019.10.001

Catto, J. W. F., Abbod, M. F., Wild, P. J., Linkens, D. A., Pilarsky, C., Rehman, I., Rosario, D. J., Denzinger, S., Burger, M., Stoehr, R., Knuechel, R., Hartmann, A., & Hamdy, F. C. (2010). The Application of Artificial Intelligence to Microarray Data: Identification of a Novel Gene Signature to Identify Bladder Cancer Progression. European Urology, 57(3), 398–406. https://doi.org/10.1016/j.eururo.2009.10.029

Ceni, E., Mello, T., & Galli, A. (2014). Pathogenesis of alcoholic liver disease: Role of oxidative metabolism NECROTIZING ENTEROCOLITIS View project HCC target therapy View project. Article in World Journal of Gastroenterology. https://doi.org/10.3748/wjg.v20.i47.00000

Chaurasia, V. (2012). Data Mining Approach to Detect Heart Diseases Application of Data Mining Techniques with special reference to academic Performance Monitoring and evaluation in higher education View project Disease Prediction using Data mining Techniques View project. https://www.researchgate.net/publication/258374505

Conrad, K., Shoenfeld, Y., & Fritzler, M. J. (2020). Precision health: A pragmatic approach to understanding and addressing key factors in autoimmune diseases. In Autoimmunity Reviews (Vol. 19, Issue 5). Elsevier B.V. https://doi.org/10.1016/j.autrev.2020.102508

Davis, S. E., Lasko, T. A., Chen, G., Siew, E. D., & Matheny, M. E. (2017). Calibration drift in regression and machine learning models for acute kidney injury. Journal of the American Medical Informatics Association, 24(6), 1052–1061. https://doi.org/10.1093/jamia/ocx030

Dawson, N. v, & Abkes, H. R. (n.d.). Systematic Errors in Medical Decision Making: Judgment Limitations.

Devi, D., Biswas, S. K., & Purkayastha, B. (2019). Learning in presence of class imbalance and class overlapping by using one-class SVM and undersampling technique. Connection Science, 31(2), 105–142. https://doi.org/10.1080/09540091.2018.1560394

Devikanniga, D., & Raj, R. J. S. (2018). Classification of osteoporosis by artificial neural network based on monarch butterfly optimisation algorithm. Healthcare Technology Letters, 5, 70–75.

Dhamodharan, S. (2014). Liver Disease Prediction Using Bayesian Classification. http://www.cs.waikato.ac.nz/ml/weka.

Ephzibah, E. P. (2011). Cost Effective Approach on Feature Selection Using Genetic Algorithms and Fuzzy logic for Diabetes Diagnosis. International Journal on Soft Computing, 2(1), 1–10. https://doi.org/10.5121/ijsc.2011.2101

Fatima, M., & Pasha, M. (2017). Survey of Machine Learning Algorithms for Disease Diagnostic. Journal of Intelligent Learning Systems and Applications, 09(01), 1–16. https://doi.org/10.4236/jilsa.2017.91001

Gil, D., & Johnson, M. (n.d.). P a g e | 63.

Goodfellow, I., Bengio, Y., & Courville, A. (n.d.). Deep Learning.

Gonzalez, C. D., Lee, M. S., Marchetti, P., Pietropaolo, M., Towns, R., Vaccaro, M. I., Watada, H., & Wiley, J. W. (2011). The emerging role of autophagy in the pathophysiology of diabetes mellitus. In Autophagy (Vol. 7, Issue 1, pp. 2–11). Taylor and Francis Inc. https://doi.org/10.4161/auto.7.1.13044

Grover, A., & Joshi, A. (2015). An overview of chronic disease models: a systematic literature review. In Global journal of health science (Vol. 7, Issue 2, pp. 210–227). https://doi.org/10.5539/gjhs.v7n2p210

Gulia, A., Vohra, R., & Rani, P. (n.d.). Liver Patient Classification Using Intelligent Techniques. http://archive.ics.uci.edu/ml/

Hajjar, A. el, & Rey, J. F. (2020). Artificial intelligence in gastrointestinal endoscopy: General overview. In Chinese Medical Journal (Vol. 133, Issue 3, pp. 326–334). Lippincott Williams and Wilkins. https://doi.org/10.1097/CM9.0000000000000623

Hepatitis B: Pathophysiology, Protection, and Patients. (2013).

Huang, G., Huang, G. bin, Song, S., & You, K. (2015). Trends in extreme learning machines: A review. In Neural Networks (Vol. 61, pp. 32–48). Elsevier Ltd. https://doi.org/10.1016/j.neunet.2014.10.001

Ibrahim, F., Taib, M. N., Abas, W. A. B. W., Guan, C. C., & Sulaiman, S. (2005). A novel dengue fever (DF) and dengue haemorrhagic fever (DHF) analysis using artificial neural network (ANN). Computer Methods and Programs in Biomedicine, 79(3), 273–281. https://doi.org/10.1016/j.cmpb.2005.04.002

Iyer, A., S, J., & Sumbaly, R. (2015). Diagnosis of Diabetes Using Classification Mining Techniques. International Journal of Data Mining & Knowledge Management Process, 5(1), 01–14. https://doi.org/10.5121/ijdkp.2015.5101

Jiang, L., Wu, Z., Xu, X., Zhan, Y., Jin, X., Wang, L., & Qiu, Y. (2021). Opportunities and challenges of artificial intelligence in the medical field: current application, emerging problems, and problem-solving strategies. The Journal of International Medical Research, 49(3), 3000605211000157–3000605211000157. https://doi.org/10.1177/03000605211000157

Jiang, Y., Zhang, T., Kusumanchi, P., Han, S., Yang, Z., & Liangpunsakul, S. (2020). Alcohol metabolizing enzymes, microsomal ethanol oxidizing system, cytochrome P450 2E1, catalase, and aldehyde dehydrogenase in alcohol-associated liver disease. In Biomedicines (Vol. 8, Issue 3). MDPI AG. https://doi.org/10.3390/biomedicines8030050

Jin, Y., Qin, C., Huang, Y., Zhao, W., & Liu, C. (2020). Multi-domain modeling of atrial fibrillation detection with twin attentional convolutional long short-term memory neural networks. Knowledge-Based Systems, 193, 105460. https://doi.org/10.1016/j.knosys.2019.105460

Karamizadeh, S., Abdullah, S. M., Halimi, M., Shayan, J., & javad Rajabi, M. (2014). Advantage and drawback of support vector machine functionality. 2014 International Conference on Computer, Communications, and Control Technology (I4CT), 63–65.

Karlik, B. (2011). Hepatitis Disease Diagnosis Using Backpropagation and the Naive Bayes Classifiers Intelligent Control Systems View project Machine Learning and Deep Learning Applications in Plasma and Thermodynamics View project Hepatitis Disease Diagnosis Using Backpropagation and the Naive Bayes Classifiers. In Journal of Science and Technology. https://www.researchgate.net/publication/265946801

Karthik, S. A., Priyadarishini, A., Anuradha, J., & Tripathy, B. (2011). Classification and Rule Extraction using Rough Set for Diagnosis of Liver Disease and its Types.

Kong, L. Z., Chandimali, N., Han, Y. H., Lee, D. H., Kim, J. S., Kim, S. U., Kim, T. D., Jeong, D. K., Sun, H. N., Lee, D. S., & Kwon, T. (2019). Pathogenesis, early diagnosis, and therapeutic management of alcoholic liver disease. In International Journal of Molecular Sciences (Vol. 20, Issue 11). MDPI AG. https://doi.org/10.3390/ijms20112712

Lander, E. S. (2011). Initial impact of the sequencing of the human genome. Nature, 470(7333), 187–197. https://doi.org/10.1038/nature09792

Lander, S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., … Yeh, R.-F. (2001). Initial sequencing and analysis of the human genome International Human Genome Sequencing Consortium* The Sanger Centre: Beijing Genomics Institute/Human Genome Center. In NATURE (Vol. 409). www.nature.com

le Berre, C., Sandborn, W. J., Aridhi, S., Devignes, M. D., Fournier, L., Smaïl-Tabbone, M., Danese, S., & Peyrin-Biroulet, L. (2020). Application of Artificial Intelligence to Gastroenterology and Hepatology. In Gastroenterology (Vol. 158, Issue 1, pp. 76-94.e2). W.B. Saunders. https://doi.org/10.1053/j.gastro.2019.08.058

Li, H., Wang, X., Liu, C., Wang, Y., Li, P., Tang, H., Yao, L., & Zhang, H. (2019). Dual-input neural network integrating feature extraction and deep learning for coronary artery disease detection using electrocardiogram and phonocardiogram. IEEE Access, 7, 146457–146469. https://doi.org/10.1109/ACCESS.2019.2943197

Maghooli, K., Langarizadeh, M., Shahmoradi, L., Habibi-Koolaee, M., Jebraeily, M., & Bouraghi, H. (2016). Differential Diagnosis of Erythmato-Squamous Diseases Using Classification and Regression Tree. Acta Informatica Medica : AIM : Journal of the Society for Medical Informatics of Bosnia & Herzegovina : Casopis Drustva Za Medicinsku Informatiku BiH, 24(5), 338–342. https://doi.org/10.5455/aim.2016.24.338-342

Mani, S., Ozdas, A., Aliferis, C., Varol, H. A., Chen, Q., Carnevale, R., Chen, Y., Romano-Keeler, J., Nian, H., & Weitkamp, J. H. (2014). Medical decision support using machine learning for early detection of late-onset neonatal sepsis. Journal of the American Medical Informatics Association, 21(2), 326–336. https://doi.org/10.1136/amiajnl-2013-001854

marsland.ml-alg-perspect.09. (n.d.).

Mazzocco, T., & Hussain, A. (2012). Novel logistic regression models to aid the diagnosis of dementia. Expert Systems with Applications, 39(3), 3356–3361. https://doi.org/10.1016/j.eswa.2011.09.023

Meroni, M., Longo, M., Rametta, R., & Dongiovanni, P. (2018). Genetic and epigenetic modifiers of alcoholic liver disease. In International Journal of Molecular Sciences (Vol. 19, Issue 12). MDPI AG. https://doi.org/10.3390/ijms19123857

Minsky, M. (1961). Steps Toward Artificial Intelligence. Proceedings of the IRE, 49(1), 8–30. https://doi.org/10.1109/JRPROC.1961.287775

Mishra, S., Yamasaki, T., & Imaizumi, H. (2018). Supervised classification of Dermatological diseases by Deep learning. http://arxiv.org/abs/1802.03752

Mohan, V. (2015). Liver Disease Prediction using SVM and Naïve Bayes Algorithms Privacy Preserving Data Mining View project. In International Journal of Science, Engineering and Technology Research (IJSETR) (Vol. 4, Issue 4). https://www.researchgate.net/publication/339551659

Mutaher Ba-Alwi, F., & Hintaya, H. M. (2013). Comparative Study for Analysis the Prognostic in Hepatitis Data: Data Mining Approach. International Journal of Scientific & Engineering Research, 4(8). http://www.ijser.org

Mozaffari-Kermani, M., Sur-Kolay, S., Raghunathan, A., & Jha, N. K. (2015). Systematic poisoning attacks on and defenses for machine learning in healthcare. IEEE Journal of Biomedical and Health Informatics, 19(6), 1893–1905. https://doi.org/10.1109/JBHI.2014.2344095

Mygdalis, V., Tefas, A., & Pitas, I. (2020). K-Anonymity inspired adversarial attack and multiple one-class classification defense. Neural Networks, 124, 296–307. https://doi.org/10.1016/j.neunet.2020.01.015

Nemati, S., Ghassemi, M. M., & Clifford, G. D. (2016). Optimal medication dosing from suboptimal clinical examples: A deep reinforcement learning approach. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2016-October, 2978–2981. https://doi.org/10.1109/EMBC.2016.7591355

Nguyen, D. T., Pham, T. D., Lee, M. B., & Park, K. R. (2019). Visible-light camera sensor-based presentation attack detection for face recognition by combining spatial and temporal information. Sensors (Switzerland), 19(2). https://doi.org/10.3390/s19020410

Otoom, A. F., Abdallah, E. E., Kilani, Y., Kefaye, A., & Ashour, M. (2015). Effective diagnosis and monitoring of heart disease. International Journal of Software Engineering and Its Applications, 9(1), 143–156. https://doi.org/10.14257/ijseia.2015.9.1.12

Pandita, A. (2021). Prediction of Heart Disease using Machine Learning Algorithms. International Journal for Research in Applied Science and Engineering Technology, 9(VI), 2422–2429. https://doi.org/10.22214/ijraset.2021.3412

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. Human Factors, 52(3), 381–410. https://doi.org/10.1177/0018720810376055

Parisi, L., RaviChandran, N., & Manaog, M. L. (2018). Feature-driven machine learning to improve early diagnosis of parKinson's disease. Expert Systems with Applications, 110, 182–190. https://doi.org/10.1016/j.eswa.2018.06.003

Parthiban, G., & Srivatsa, S. K. (2012). ISSN : 2249-0868 Foundation of Computer Science FCS. In International Journal of Applied Information Systems Parthiban, G., & Srivatsa, S. K. (2012). ISSN : 2249-0868 Foundation of Computer Science FCS. In International Journal of Applied Information Systems (IJAIS) (Vol. 3, Issue 7). www.ijais.org

Putatunda, S. (2019). A Hybrid Deep Learning Approach for Diagnosis of the Erythemato-Squamous Disease.

R, V. (2019). Detection and Classification of Kidney Disorders using Deep Learning Method. JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES, 14(2). https://doi.org/10.26782/jmcms.2019.04.00021

Rajeswari, P., & Reena, G. S. (2010). Analysis of Liver Disorder Using Data mining Algorithm. In Global Journal of Computer Science and Technology (Vol. 10). (IJAIS) (Vol. 3, Issue 7). www.ijais.org

Ricardo, C., & dela Cruz, N. N. (2021). International Journal of Advanced Trends in Computer Science and Engineering. International Journal of Advanced Trends in Computer Science and Engineering Research, 9(3), 3178–3183. https://doi.org/10.13140/RG.2.2.27464.24322

Rong, G., Mendez, A., Bou Assi, E., Zhao, B., & Sawan, M. (2020). Artificial Intelligence in Healthcare: Review and Prediction Case Studies. Engineering, 6(3), 291–301. https://doi.org/10.1016/J.ENG.2019.08.015

Sarwar, A., & Sharma, V. (2012). Intelligent Naïve Bayes Approach to Diagnose Diabetes Type-2.

Sathyadevi, G. (2011). Application of CART algorithm in hepatitis disease diagnosis. 2011 International Conference on Recent Trends in Information Technology (ICRTIT), 1283–1287.

Sawarkar, S. D., & Pande, A. P. (n.d.). Neural Network Aided Breast Cancer Detection and Diagnosis Using Support Vector Machine.

Schwinger, R. H. G. (2021). Pathophysiology of heart failure. In Cardiovascular Diagnosis and Therapy (Vol. 11, Issue 1). AME Publishing Company. https://doi.org/10.21037/CDT-20-302

Sellahewa, K. H. (2013). Pathogenesis of Dengue Haemorrhagic Fever and Its Impact on Case Management. ISRN Infectious Diseases, 2013, 1–6. https://doi.org/10.5402/2013/571646

Shah, P. M., Zeb, A., Shafi, U., Zaidi, S. F. A., & Shah, M. A. (2018). Detection of Parkinson Disease in Brain MRI using Convolutional Neural Network. 2018 24th International Conference on Automation and Computing (ICAC), 1–6.

Shameem Fathima, A., & Manimeglai, D. (2012). Predictive Analysis for the Arbovirus-Dengue using SVM Classification. International Journal of Engineering and Technology, 2(3).

Sharma, P., & Kaur, M. (2013). Classification in Pattern Recognition: A Review. In International Journal of Advanced Research in Computer Science and Software Engineering (Vol. 3, Issue 4). www.ijarcsse.com

Shrivastava, V. K., Londhe, N. D., Sonawane, R. S., & Suri, J. S. (2015). Reliable and Accurate Psoriasis Disease Classification in Dermatology Images Using Comprehensive Feature Space in Machine Learning Paradigm. Expert Syst. Appl., 42(15), 6184–6195. https://doi.org/10.1016/j.eswa.2015.03.014

Singh, A., Mehta, J. C., Anand, D., Nath, P., Pandey, B., & Khamparia, A. (2021). An intelligent hybrid approach for hepatitis disease diagnosis: Combining enhanced k-means clustering and improved ensemble learning. Expert Systems, 38(1). https://doi.org/10.1111/exsy.12526

Soliman, O. S., & Elhamd, A. (2014). Classification of Hepatitis C Virus using Modified Particle Swarm Optimization and Least Squares Support Vector Machine. International Journal of Scientific & Engineering Research, 5(3). http://www.ijser.org

Stafford, I. S., Kellermann, M., Mossotto, E., Beattie, R. M., MacArthur, B. D., & Ennis, S. (2020). A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. In npj Digital Medicine (Vol. 3, Issue 1). Nature Research. https://doi.org/10.1038/s41746-020-0229-3

Tan, K. C., Teoh, E. J., Yu, Q., & Goh, K. C. (2009). A hybrid evolutionary algorithm for attribute selection in data mining. Expert Systems with Applications, 36(4), 8616–8630. https://doi.org/10.1016/j.eswa.2008.10.013

Tarmizi, N. D. A., Jamaluddin, F., Abu Bakar, A., Othman, Z. A., Zainudin, S., & Hamdan, A. R. (2013). Malaysia dengue outbreak detection using data mining models. Journal of Next Generation Information Technology (JNIT), 4(6), 96–107.

Tech, M., & Rathore, Y. K. (n.d.). A Survey: Detection and Prediction of Diabetes Using Machine Learning Techniques. www.ijert.org

Teschke, R. (2019). Alcoholic liver disease: Current mechanistic aspects with focus on their clinical relevance. In Biomedicines (Vol. 7, Issue 3, pp. 1–12). MDPI AG. https://doi.org/10.3390/BIOMEDICINES7030068

Tenório, J. M., Hummel, A. D., Cohrs, F. M., Sdepanian, V. L., Pisa, I. T., & Marin, H. F. (2011). Artificial intelligence techniques applied to the development of a decision-support system for diagnosing celiac disease. International Journal of Medical Informatics, 80 11, 793–802.

Vembandasamyp, K., Sasipriyap, R. R., & Deepap, E. (2015). Heart Diseases Detection Using Naive Bayes Algorithm. In IJISET-International Journal of Innovative Science, Engineering & Technology (Vol. 2, Issue 9). www.ijiset.com

Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M., & Thelen, E. (2001). Autonomous mental development by robots and animals. In Science (Vol. 291, Issue 5504, pp. 599–600). https://doi.org/10.1126/science.291.5504.599

Wooldridge, M., & Jennings, N. R. (n.d.). Intelligent Agents: Theory and Practice.

Yu, X., Ye, C., & Xiang, L. (2016). Application of Artificial Neural Network in the Diagnostic System of Osteoporosis. Neurocomput., 214(C), 376–381. https://doi.org/10.1016/j.neucom.2016.06.023

Yuan, X., He, P., Zhu, Q., & Li, X. (2019). Adversarial Examples: Attacks and Defenses for Deep Learning. IEEE Transactions on Neural Networks and Learning Systems, 30, 2805–2824.

Zhang, X., & Wu, D. (2019). On the Vulnerability of CNN Classifiers in EEG-Based BCIs. http://arxiv.org/abs/1904.01002