

ISSN(O): 2319-8753

ISSN(P): 2347-6710



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



**Impact Factor: 8.699** 

Volume 14, Issue 10, October 2025





|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 10, October 2025

|DOI: 10.15680/IJIRSET.2025.1410046|

### Chronic Disease Detection and Prediction: Integrating Machine Learning with Pathophysiology and Ethical Frameworks

Neelambari D, Jerrit D, Jeffrid D

Department of Computer Applications, Arunachala College of Engineering for Women, Manavilai, Nagercoil,
Tamil Nadu, India

ABSTRACT--The need for better efforts in early diagnosis and prediction is highlighted by the fact that chronic diseases, including diabetes, cancer, and heart disease, are the main causes of death globally. Emerging molecular biology fields such as proteomics, glycomics, lipidomics, transcriptomics, and genomics have shown promise in understanding the pathogenesis and treatment of chronic diseases. Traditional statistical methods have difficulties in analysing and interpreting the complex biomarker and mechanistic data from these "omics" research. Methods of machine learning (ML) hold great potential for opening up new avenues for data-driven risk assessment and prognosis of chronic diseases. This paper provides a comprehensive overview of the latest applications of machine learning algorithms in identifying and predicting chronic diseases across various datasets, including wearable technology, genomics, medical imaging, and electronic health records. In particular, we evaluate and summarise important research using essential machine learning algorithms, spanning from contemporary deep learning neural network topologies to more conventional methods, such as logistic regression and random forests. To summarize key trends and trajectories that could guide future research and clinical translation initiatives in this expanding field, we compile the existing body of literature on machine learning for chronic illness prediction. We highlight the significant advancements and achievements in this field, while also acknowledging the substantial obstacles and problems that remain to be addressed. We conclude by talking about future directions for scalable, fair, and clinically applicable machine learning systems that will revolutionize the screening and prevention of chronic diseases.

KEYWORDS: Big data, OMICs data, machine learning algorithms, chronic diseases, and disease prediction

#### I. INTRODUCTION

According to the WHO, millions of people worldwide suffer from chronic diseases such as type II diabetes (T2DM), cardiovascular disorders, cancer, and chronic respiratory ailments, which lower quality of life and result in over 41 million deaths yearly. People who suffer from these illnesses bear a heavy financial burden, particularly in low-resource nations where 77% of fatalities from chronic illnesses take place. Current treatments are not curative, and the illness burden is still increasing, even though medicine and lifestyle modifications can help manage chronic conditions. Healthcare has changed from reactive to proactive methods to address this, utilising data-driven risk classification and personalised prevention. However, the majority of this data is unstructured and unlabelled, which limits its immediate utility. The development of "omics" technologies (such as proteomics, lipidomics, and genomics) has led to the creation of massive datasets that provide important new information about the mechanisms underlying disease. But the majority of this data is unstructured and unlabelled, which limits its ability to be used right away. Despite being expensive and time-consuming, effective data labelling is necessary to allow artificial intelligence (AI) and machine learning (ML) systems to effectively analyse these complex datasets. Training trustworthy AI models that can assist physicians in early diagnosis and monitoring is based on properly labelled data.

Machine learning finds patterns in a variety of data sources to provide effective solutions for chronic disease detection and prediction. Personalised medicine is being advanced by the use of techniques like random forests and neural networks for tasks like disease risk prediction and feature selection. The increasing accessibility of medical data, driven by cloud computing and digital technologies, has





|www.ijirset.com | A Monthly, Peer Reviewed & Refereed Journal | e-ISSN: 2319-8753 | p-ISSN: 2347-6710 |

#### Volume 14, Issue 10, October 2025

#### |DOI: 10.15680/IJIRSET.2025.1410046|

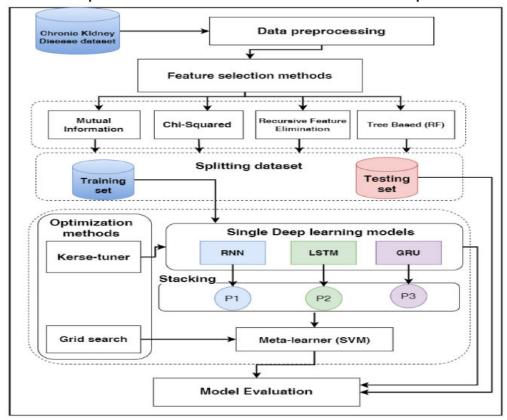


Fig. 1. Laboratory techniques and machine learning applications for diagnosing and predicting diseases

Necessitated collaboration between healthcare professionals, experts, and data scientists. Even though machine learning has demonstrated potential in enhancing diagnosis, prognosis, and treatment planning, there are still moral and practical issues with it. Key issues include bias, accountability for clinical outcomes, data privacy, and transparency in algorithm decisions. Careful consideration of these issues is necessary for ethical application while optimising the advantages of technology. Fairness, beneficence, and non-maleficence are ethical concepts that must direct the development and implementation of machine learning in order to ensure its appropriate use in healthcare, maintain patient trust, and assist in clinical decision-making.

#### II. SYNOPSIS OF MACHINE LEARNING ALGORITHM

Machine learning (ML), a branch of artificial intelligence, is the process by which computers adapt, learn from experience, and create models without prior guidance or oversight. After receiving data, the computer makes predictions using mathematical formulae. ML algorithms fall into three general categories: reinforcement learning, unsupervised learning, and supervised learning.

#### A. Supervised learning:

Supervised learning methods use labeled input-output pairs in the training data to train models to make predictions. Decision trees (DTs), logistic regression (LR), support vector machines (SVMs), RFs, NNs, and linear regression are examples of popular supervised learning techniques. LR classifies based on logistic sigmoid outputs, whereas linear regression predicts continuous values. To classify data points, SVMs identify the best decision boundaries. Twelve DTs conduct sequential binary splits on the features, and RFs enhance performance by averaging forecasts from several DTs. NNs are made up of interconnected layers of nodes that are modeled after biological neurons and may learn to change inputs in nonlinear ways. The benefits and drawbacks of these machine learning methods have been thoroughly examined in research.





|www.ijirset.com | A Monthly, Peer Reviewed & Refereed Journal | e-ISSN: 2319-8753 | p-ISSN: 2347-6710 |

#### Volume 14, Issue 10, October 2025

|DOI: 10.15680/IJIRSET.2025.1410046|

#### B. Unsupervised learning:

Unsupervised learning looks for hidden patterns by analysing unlabelled data. Clustering algorithms like k-means, hierarchical clustering, and Gaussian mixture models—which group data points according to similarity—are important unsupervised learning techniques. Using dimension reduction techniques like principal component analysis (PCA) and t-distributed stochastic neighbour embedding (t-SNE), significant dimensions that capture the majority of the variation are identified.

#### C. Reinforcement Learning:

Models are trained to optimise behaviours using rewards and penalties through reinforcement learning. The model learns by trial-and-error interactions with a dynamic environment rather than precise training labels. Q-learning and policy gradient approaches are two popular reinforcement learning strategies.

## Types of Machine Learning

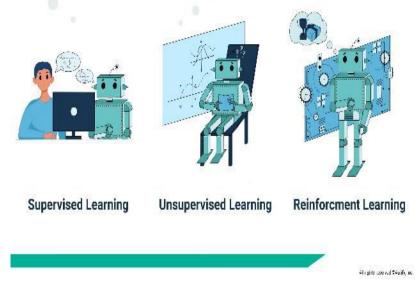


Fig.2. Types of Machine Learning

### III. COMPREHENDING CHRONIC DISEASE PATHOPHYSIOLOGICAL PROCEDURE USING ML ALGORITHMS

Diabetes, heart disease, and cancer are examples of chronic diseases that typically take years to manifest. A person's lifestyle choices (diet, exercise, smoking), genetics, and environmental factors are some of the causes. Type 2 diabetes, for instance, results from a person's genes and bad lifestyle choices, causing the body to cease using insulin as it should. In a similar vein, heart disease is a gradual condition brought on by issues like smoking, high blood pressure, high cholesterol, and diabetes. These conditions cause numerous, often long-term, minute alterations in the body at the molecular level. Comprehending these complex shifts is crucial. Knowing how diseases begin and progress can help us identify them early and treat them according to each patient's needs. This is where machine learning (ML) can be useful. Computers are used in machine learning (ML) to identify patterns in large amounts of health data, such as genes, blood tests, medical imaging, etc. In some cases, this helps physicians detect disease symptoms before they manifest. As an illustration:

In heart illness, machine learning can identify individuals at high risk of heart attacks by analyzing genes, proteins, medical scans, and other data.





|www.ijirset.com | A Monthly, Peer Reviewed & Refereed Journal | e-ISSN: 2319-8753 | p-ISSN: 2347-6710 |

#### Volume 14, Issue 10, October 2025

#### |DOI: 10.15680/IJIRSET.2025.1410046|

To forecast which tumours may recur or spread, machine learning can compare tumour data with patient outcomes. A unique form of machine learning called deep learning may combine many data sources, such as genetic information, bodily data, and medical records, to enhance predictions. According to the article, ML is being used to treat a variety of conditions, including diabetes, kidney disease, heart issues, cancer, lung diseases, and inflammation. It demonstrates how improved understanding, earlier disease prediction, and more individualized patient care are all made possible by the use of intelligent algorithms with large & complicated data.

#### IV. T2DM PATHOPHYSIOLOGY

T2DM results from the body's improper metabolism of glucose, which causes blood glucose levels to rise. According to the International Diabetes Federation, more than a million individuals passed away from the condition in 2015 alone. The intricacy and variability of type 2 diabetes continue to obstruct a solution, even though many patients' longevity and quality of life have increased due to medication and lifestyle modifications. Insulin resistance in tissues and insufficient insulin production from the pancreatic beta ( $\beta$ ) cells are the two main causes of type 2 diabetes. Beta cells first make up for this by secreting more insulin. However, as the  $\beta$  cells age and die, insulin production gradually decreases.

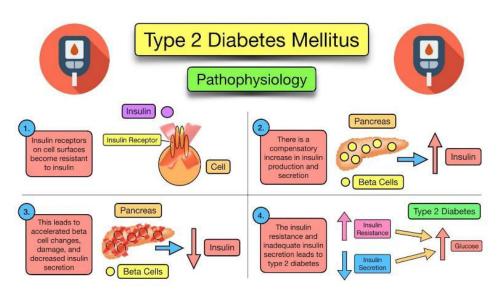


Fig.3.Type II diabetes's pathophysiology The pancreatic beta cells secrete glucagon, while the alpha cells secrete insulin

Insulin-producing beta cells in the pancreas are diminished in individuals with Type 2 Diabetes (T2DM) and impaired fasting glucose. Research demonstrates that individuals with poor fasting glucose see a 40% reduction in beta cell volume. T2DM sufferers lose 63%. The primary cause of these beta cells' demise is apoptosis, or cell death, a normal process. Although the primary focus is on beta cells, as the disease progresses, alpha cells—which also aid in blood sugar regulation—are also impacted. Genetics accounts for almost 30% of occurrences of type 2 diabetes. More than 176 genes, including CDKAL1, SLC30A8, HHEX, and others, have been connected by researchers to type 2 diabetes. Genes that produce or metabolize insulin are impacted by some mutations. For instance, the KCNQ1 gene influences the secretion of insulin. The body develops insulin resistance in type 2 diabetes, which implies:

- Glucose cannot be adequately absorbed by muscles.
- When glucose cannot be reabsorbed by the kidneys, it is lost in the urine.
- Additionally, this results in water loss, which makes urinating more frequent.

#### A. T2DM Complications:

People on insulin or other medications who exercise excessively may get hypoglycemia or low blood sugar. Less frequent is diabetic ketoacidosis (DKA). In the absence of glucose, the body breaks down protein and fat for energy. This produces ketones (such as beta-hydroxybutyric and acetoacetic acids), which can result in symptoms including





|www.ijirset.com | A Monthly, Peer Reviewed & Refereed Journal | e-ISSN: 2319-8753 | p-ISSN: 2347-6710 |

#### Volume 14, Issue 10, October 2025

|DOI: 10.15680/IJIRSET.2025.1410046|

vomiting and stomach pain. HHS, or hyperglycaemic hyperosmolar state, is caused by low insulin and extremely elevated blood sugar. Dehydration is the result of the body losing salts (electrolytes) and water through urine. By encouraging the breakdown of fat, hormones such as cortisol and adrenaline exacerbate this issue. Other chronic issues associated with type 2 diabetes include:

- Heart attacks and strokes
- Retinopathy, or eye issues
- Nephropathy (injury to the kidneys)
- Damage to the nerves (neuropathy)
- Liver fatty condition

#### V. PATHOPHYSIOLOGY OF CKD

In 2017, chronic kidney disease (CKD) caused approximately 35.8 million disability-adjusted life years (DALYs) and 1.2 million deaths, making it a significant worldwide health concern. End-stage renal disease (ESRD) patients sometimes have to pay hefty prices for dialysis and drugs, and many of them pass away before receiving the right care. To stop CKD from getting worse, risk factors such as diabetes, high blood pressure, obesity, and excessive salt consumption must be identified early.

#### VI. ML ALGORITHMS' USE IN T2DM IDENTIFICATION AND PROGNOSIS

Type 2 Diabetes Mellitus (T2DM) often develops silently with no noticeable symptoms, but early detection is essential for effective treatment and preventing complications. Using large health datasets, machine learning (ML) has been widely used to identify high-risk patients and predict diabetes risk. Many machine learning algorithms have shown effectiveness in detecting early signs of type 2 diabetes. Polat and Günes, for example, used Principal Component Analysis (PCA) to distinguish between diabetics and healthy individuals, while Farran et al. employed the K-Nearest Neighbour (KNN) method, achieving 75% accuracy in diabetes prediction and 94% accuracy in hypertension prediction. Lai et al. evaluated methods including Logistic Regression (LR), Random Forest (RF), Decision Trees (DT), and Gradient Boosting Machine (GBM), with GBM and LR producing the best results. Perveen et al. predicted who might develop diabetes within eight years using a Hidden Markov Model, while Ravaut forecasted diabetes complications over three years using a gradient-boosting decision tree model. The performance of these models has also been compared through research. According to Wang et al., Artificial Neural Networks (ANN) outperformed standard logistic regression in predicting type 2 diabetes. In another study, Random Forest was found to be more accurate than Decision Trees, Neural Networks, Logistic Regression, and Naïve Bayes. ML has also been valuable for identifying biological markers. Dias-Audibert, for example, used Random Forest to find biomarkers linked to diabetes and weight gain, while Pedinti employed feature selection techniques to identify key metabolites associated with type 2 diabetes. Sisodia's research indicated that Naïve Bayes outperformed other algorithms like Decision Trees and Support Vector Machines (SVM) in terms of accuracy.

Numerous important predictors of type 2 diabetes have been found through research. Oh et al. found that impaired fasting glucose, high Blood pressure, and high cholesterol all increase the risk of diabetes. Marcos predicted childhood obesity based on BMI-related variables using Random Forest and GBM. Dinh emphasised age as important, but Zhang considered urine glucose to be a powerful risk signal. The leading predictor, according to Kopita, was hyperglycemia, which was followed by age, HDL, triglycerides, exercise, and medical history. All things considered, machine learning methods provide insightful information for early diabetes prediction and individualised treatment. The research under discussion shows how ML techniques like LR, RF, and NNs can be used to identify people who are at risk for diabetes and forecast future complications. Risk stratification, patient condition evaluation, and complications prediction utilising time series data are examples of machine learning advancements for diabetes. The research under discussion shows how ML techniques like LR, RF, and NNs can be used to identify people who are at risk for diabetes and forecast future complications. Risk stratification, patient condition evaluation, and complications prediction utilising time series data are examples of machine learning advancements for diabetes. These computer analyses' primary predictors are consistent with established pathophysiological elements that influence the development of diabetes.





|www.ijirset.com | A Monthly, Peer Reviewed & Refereed Journal | e-ISSN: 2319-8753 | p-ISSN: 2347-6710 |

#### Volume 14, Issue 10, October 2025

#### |DOI: 10.15680/IJIRSET.2025.1410046|

#### VII. ML ALGORITHM APPLICATION IN CKP DETECTION AND PREDICTION

Dovgan and colleagues (54) investigated the use of machine learning methods, including RF, SVM, LR, k-NN, XG Boost, and NN, to forecast when a patient with chronic kidney disease (CKD) should begin renal replacement therapy (RRT), which includes dialysis or renal transplantation. In another study, an NN was used to predict which patients with acute renal injury will require dialysis within 48 hours. ML algorithms, such as k-NN, SVM, NN, ridge regression, lasso regression, and XG Boost, produced a sensitivity of 0.8.56 and an average area under the curve (AUC) of 0.87 for proteomics data. Employed machine learning methods to distinguish between CKD brought on by immunological mechanisms and CKD brought on by other causes. 57 Using the UCI repository data set for chronic renal failure (CRF), Almansour et al.58 used SVM, k-NN, and Soft Independent Modelling of Class Analogy, and they were able to discriminate CKD patients from healthy people with 93% accuracy.

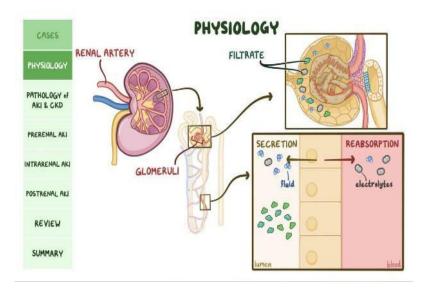


Fig.4. Pathophysiology of chronic kidney disease

Chronic kidney disease (CKD) development and progression have been predicted using a variety of machine learning (ML) models. The best performance was achieved by Decision Trees (DT), ANN, NB, and random subspace classifiers. While Bueno et al. utilised CNNs to classify glomeruli with 98% accuracy, Lee et al. predicted CKD outcomes with excellent accuracy (AUC 0.93) using an unsupervised model on kidney biopsy pictures. Models such as SVM, Random Forest (RF), XG Boost, and Light GBM proved more effective than logistic regression in forecasting acute renal damage following surgery. Using deep learning and machine learning models, Kim et al. discovered that XG Boost was the most successful in evaluating dialysis adequacy. The explanation of CKD pathways using these methods holds promise for more prompt and focused prevention and treatment. Nevertheless, when using algorithms in a therapeutic setting, careful consideration of model interpretability and causality is necessary.

#### VIII. PATHOPHYSIOLOGY OF CVDs

Heart disease, stroke, coronary artery disease, congenital defects, and rheumatic heart disease are among the world's top causes of death. Coronary artery calcium (CAC) scoring, electrocardiography (ECG), coronary CT angiography (CCTA), and blood lipid tests are common diagnostic procedures. Through the study of the heart's electrical signals (P-wave, QRS complex, and T-wave), an ECG can identify problems with heart rhythm. Machine learning (ML) models such as support vector machines (SVMs) and artificial neural networks (ANNs) can enhance ECG analysis. To assess arterial calcification and forecast the risk of CVD, CAC scoring employs CT scans. Despite providing precise images of coronary plaques and artery architecture, CCTA is frequently constrained by variations in the interpretation of the results by medical professionals. Using advanced machine learning to analyse CCTA scans could increase the precision and reliability of CVD diagnosis.





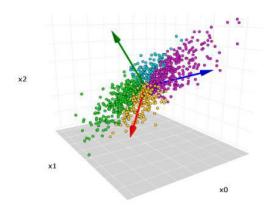
|www.ijirset.com | A Monthly, Peer Reviewed & Refereed Journal | e-ISSN: 2319-8753 | p-ISSN: 2347-6710 |

#### Volume 14, Issue 10, October 2025

|DOI: 10.15680/IJIRSET.2025.1410046|

#### IX. ML ALGORITHM APPLICATION IN CVD'S DETECTION AND PREDICTION

Cardiovascular disease (CVD) detection and prediction have been improved with machine learning (ML). ML models such as SVM and k-nearest neighbours have improved the precision and reliability of coronary CT angiography (CCTA) interpretation in imaging. Similar clustering techniques assisted in identifying distinct aortic valve phenotypes, while Bruse et al. utilised cardiac MRI with clustering analysis to distinguish between healthy people and those with congenital heart disease. Models such as Random Forest (RF), Gradient Boosting Machine (GBM), XG Boost, and Logistic Regression (LR) fared better than conventional risk score techniques in predicting atherosclerotic CVD. RF is frequently more successful than CART, AdaBoost, Bagged Trees, and Naïve Bayes, according to studies. Risk prediction has been significantly enhanced by tools such as AutoPrognosis in comparison to conventional models like the Framingham Score. Yu et al. employed PCA and RF on DNA methylation data to identify important biomarkers in congenital heart defect studies, with great success. In general, the analysis of ECGs, medical pictures, and electronic health data is becoming more automated and better thanks to machine learning techniques like RF, SVMs, and clustering. These methods allow for the earlier and more precise identification of heart conditions such as congenital abnormalities, atherosclerosis, and arrhythmias.



Analysis of CVD

#### X. PATHOPHYSIOLOGY OF CANCER

Cancer is brought on by genetic mutations and epigenetic modifications brought on by inherited characteristics or environmental influences like chemicals and radiation. These alterations impact important genes, resulting in unchecked cell proliferation and organ-wide dissemination. This aberrant growth is supported by growth factors and signalling mechanisms. Histone modifications and DNA methylation are examples of epigenetic alterations that also mute crucial tumor-suppressing genes. Collectively, these disturbances propel the advancement of cancer. Improving targeted treatments and cancer therapy requires an understanding of these pathways.

#### XI. ML ALGORITHM APPLICATION IN CANCER DETECTION AND PREDICTION

Cancer diagnosis is difficult and frequently depends on biopsies, which can omit crucial information. Although MRI aids in tumour location and severity assessment, various physicians may interpret it differently. In order to find cancer biomarkers, lab methods such as PCR, NGS, and DNA microarrays provide massive datasets that can be further analysed using machine learning (ML). ML finds patterns in genomic and imaging data to enhance cancer detection. It assists in classifying cancer kinds and forecasting results by analysing gene sequences, expression levels, histone alterations, and RNA data. Even with rare disorders, CNNs and other algorithms can identify DNA-RNA interactions and forecast genetic variations linked to the condition. Overall, ML enhances cancer detection, diagnosis, and treatment by automating complex data analysis. It improves MRI interpretation, identifies gene expression patterns, and uncovers new cancer biomarkers, offering insights into how tumors start, grow, and spread.





|www.ijirset.com | A Monthly, Peer Reviewed & Refereed Journal | e-ISSN: 2319-8753 | p-ISSN: 2347-6710 |

#### Volume 14, Issue 10, October 2025

#### |DOI: 10.15680/IJIRSET.2025.1410046|

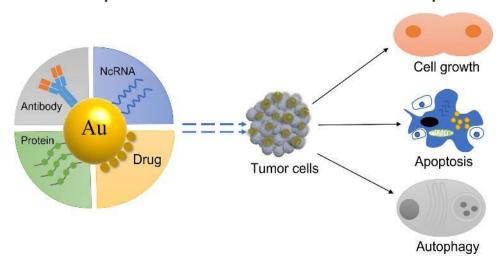


Fig.6. Diagnosis of Cancer

#### XII. PATHOPHYSIOLOGY OF CHRONIC RESPIRATORY DISEASES

Chronic respiratory disorders are long-term lung conditions that impair breathing by obstructing airflow, such as asthma, pulmonary hypertension, and COPD (which includes emphysema and chronic bronchitis). If left untreated, these illnesses can be lethal and result in irreversible harm. Asthma alone killed more than 46,000 people in 2019 and afflicted almost 262 million people. Factors including smoking, infections, and chronic inflammation frequently cause or exacerbate them. Cigarette smoke, for instance, can release chemicals and enzymes that harm lung tissue, cause inflammation, and constrict airways. Spirometry and other pulmonary function tests measure the amount of air a person can exhale, which aids in evaluating lung health. The FEV1/FVC ratio is an important diagnostic metric; values less than 70% indicate obstructive illnesses such as COPD. A precise diagnosis is essential since both overdiagnosis and underdiagnosis are prevalent and can result in inadequate care or needless drug side effects.

#### XIII. ML ALGORITHM APPLICATION IN RESPIRATORY DISEASE DETECTION AND PREDICTION

The diagnosis of interstitial lung illnesses is now aided by bronchoalveolar lavage, high-resolution CT scans, and, where required, surgical lung biopsies. By identifying patterns in photos, gene expression data, and lung function tests that physicians would overlook, machine-learning (ML) algorithms go one step further. Algorithms like AdaBoost, RF, and SVM have identified smoking-linked genes (e.g., PRKAR2B, SLITRK6) associated with COPD; studies suggest that K-means clustering can categorise advanced COPD with 94% accuracy; and ensemble models perform better than pulmonologists in predicting exacerbations. Convolutional neural networks accurately grade fibrotic illness, identify new emphysema subtypes, and assist radiologists in identifying malignant lung nodules. ML systems are also used to evaluate inflammation in Crohn's disease and triage the severity of COVID-19 pneumonia with up to 96% accuracy. Neural networks, SVMs, and clustering techniques work together to provide earlier, more accurate respiratory disease diagnosis, subtype identification, and risk prediction; when verified in conjunction with pulmonologists, these methods hold promise for more individualised treatment.

### XIV. IMPLICATIONS OF ML APPLICATIONS IN PREDICTION AND DETECTION OF CHRONIC DISEASES

There are currently very few ML-based systems for predicting and detecting chronic diseases that have received international approval. Notable examples include Kardio Screen TM, a CE-marked machine learning tool that uses patient data to forecast the risk of cardiovascular disease, and the FDA-approved ID x-DR, which uses retinal pictures to identify diabetic retinopathy. Medical AI regulatory systems are still developing in spite of this promise. Model transparency, causation, bias, validation, data quality, and privacy are some of the main obstacles. Clinical trust is hampered by deep learning models' tendency to function as "black boxes," which makes it difficult to understand their judgements. Although there are post-hoc interpretability tools, they only provide an approximation of how models





|www.ijirset.com | A Monthly, Peer Reviewed & Refereed Journal | e-ISSN: 2319-8753 | p-ISSN: 2347-6710 |

#### Volume 14, Issue 10, October 2025

#### |DOI: 10.15680/IJIRSET.2025.1410046|

function. While ML is excellent at identifying correlations, it has trouble proving causation; this requires experimental validation and integration with biological knowledge. Unbalanced datasets can lead to bias, which disadvantages underrepresented groups. Fair sampling, diverse teams, and evaluation that prioritises equity are necessary to address this. Generalisability depends on validation, but it is hindered by a lack of high-quality data. It is crucial to conduct thorough testing on a variety of datasets. Although useful, EHRs are frequently disjointed and loud, which limits their usefulness; transfer learning and GANs provide some improvement. Furthermore, extensive data utilisation creates privacy issues that could erode patient-doctor confidence. For ethical ML integration in healthcare, explainable AI, transparent data governance, and clinician monitoring are essential. In the end, ML ought to supplement clinical judgement rather than take its place.

#### XV. CONCLUSION AND PROSPECTS

Traditional analysis approaches are no longer able to keep up with the increasing development of biological data from sources such as genomics, imaging, EHRs, and wearables. ML provides strong capabilities for processing this high-dimensional, complex data, identifying trends, and producing precise forecasts. Although ML's full potential in managing chronic diseases is still being discovered, it is already revolutionising industries like imaging and genomics. Nonetheless, it is necessary to address practical issues, including data quality, bias, model interpretability, validation, and clinical integration. Incorporating causal biological information is also crucial for understanding illness causes and creating therapeutics, going beyond pattern recognition. Promising developments in explainable AI, hybrid models, and deep learning for multimodal analysis are part of ongoing research to improve machine learning techniques appropriate for biomedical data. Ultimately, using ML to improve the diagnosis, treatment, and decision-making of chronic diseases will require responsible development and close cooperation between data scientists and doctors.

#### XVI. ACKNOWLEDGMENT

Our work was greatly aided by the insightful knowledge and experience of our colleagues from the School of Science, Edith Cowan University, Perth, Australia, and the Rural Clinical School, Medicine and Health, University of New South Wales, Wagga Wagga, Australia. Their assistance has been crucial in determining the focus and calibre of our study. We also like to thank the anonymous reviewers for their insightful and helpful criticism, which has greatly raised the calibre of this work.

#### XVII. STATEMENT OF CONFLICT OF INTEREST

No conflicts of interest are disclosed by the writers

#### XVIII. STATEMENT OF DATA AVAILABILITY

Not relevant.

#### REFERENCES

- 1. World Health Organisation (WHO). Non-Communicable Diseases, WHO; 2021. Accessed January 2, 2024. https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases
- 2. Adua E, Roberts P, Sakyi SA, et al. Profiling of cardio-metabolic risk factors and medication utilisation among type II diabetes patients in Ghana: a prospective cohort study. Clin Transl Med. 2017;6:e32. doi:10.1186/s40169-017-0162-5
- 3. Adua E, Kolog EA, Afrifa-Yamoah E, et al. Predictive model and feature importance for early detection of type II diabetes mellitus. Transl Med Commun. 2021;6:17. doi:10.1186/s41231-021-00096-z
- 4. Adua E, Afrifa-Yamoah E, Kolog EA. Leveraging supervised machine learning for determining the link between suboptimal health status and the prognosis of chronic diseases. In: Wang W, ed., All Around Suboptimal Health. Advances in Predictive, Preventive and Personalised Medicine. Springer; 2024:18. doi:10. 1007/978-3-031-46891-99
- 5. Golubnitschaja O, Baban B, Boniolo G, et al. Medicine in the early twenty-first century: paradigm and anticipation position paper 2016. EPMA J. 2016;7:23. doi:10.1186/s13167-016 0072-4





www.ijirset.com | A Monthly, Peer Reviewed & Refereed Journal | e-ISSN: 2319-8753 | p-ISSN: 2347-6710 |

#### Volume 14, Issue 10, October 2025

#### |DOI: 10.15680/IJIRSET.2025.1410046|

- 6. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform. 2018;19(6):1236-1246. doi:10.1093/bib/bbx044
- 7. Clifton DA, Niehaus KE, Charlton P, Colopy GW. Health informatics via machine learning for the clinical management of patients. Yearb Med Inform. 2015;24:38-43. doi:10.15265/IY 2015-014
- 8. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med. 2019;380:1347-1358. doi:10.1056/NEJMra1814259
- 9. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media; 2009.
- 10. Delpino, F. M., da Rosa, T. H., Paiva, H., & Silva, L. A. (2022). Machine learning for predicting chronic diseases: A systematic review. Computers in Biology and Medicine, 147, 105632











# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY







9940 572 462 🔯 6381 907 438 🔀 ijirset@gmail.com

