Al - Saeed University Journal of Applied Sciences
journal@alsaeeduni.edu.ye
Vol (9), No(1), Jun., 2026
ISSN: 3104 - 8978 (Print) ISSN: 3104-8986 (Online)

Advances in Machine Learning for Chronic Disease Prediction: A Comprehensive Review

Ansab Naif
Computer Sciences Department, Faculty of Applied Sciences, Taiz University, Taiz, Yemen
ansab.naif@taiz.edu.ye

Ahmed Naef
Computer Sciences Department, Faculty of Applied Sciences, Taiz University, Taiz, Yemen
ahmednaef@taiz.edu.ye

Received: 18/3/2026 **Accepted: 20/5/2026**

Journal Website: https://journal.alsaeeduni.edu.ye

Advances in Machine Learning for Chronic Disease Prediction: A Comprehensive Review

Ansab Naif
Computer Sciences Department, Faculty of Applied Sciences, Taiz University, Taiz, Yemen

Ahmed Naef
Computer Sciences Department, Faculty of Applied Sciences, Taiz University, Taiz, Yemen

Abstract

Chronic conditions including diabetes, cardiovascular disorders, kidney disease, and Alzheimer's contribute increasingly to the global health burden, highlighting the critical need for early diagnosis and timely intervention to enhance patient care and lower medical expenses. Recent advances in Machine Learning (ML) have demonstrated significant potential in predicting the onset and progression of these conditions by analyzing large-scale medical datasets and uncovering complex patterns often missed by traditional diagnostic methods. These technologies enable faster, more accurate and cost-effective assessments, benefiting both clinicians and patients. This review comprehensively examines the application of ML techniques to the prediction of these four major chronic diseases, highlighting their transformative role in early diagnosis. Deep learning models, such as stacked Artificial Neural Network (stack-ANN) and hybrid Convolutional Neural Network Long Short-Term Memory (CNNLSTM) architectures, have achieved high predictive performance, with reported accuracies reaching up to 99.51% and AUC scores of 1.00 in specific contexts. Boosting algorithms, including XGBoost and LightGBM, also deliver robust results, frequently exceeding 98% of accuracy. The review emphasizes the crucial role of data preprocessing and feature selection in enhancing model interpretability and performance. Ensemble methods, such as bagging and voting classifiers, further contribute to improved predictive outcomes. Despite these advancements, the generalizability of many models remains limited due to heavy reliance on single-source datasets.

Keywords: Chronic disease prediction, Machine learning, Healthcare analytics, Literature review.

Introduction

Chronic diseases represent a significant challenge in the global healthcare sector, arising from metabolic syndrome — a condition influenced by lifestyle choices, genetic predisposition, and environmental determinants (C. Lee et al., 2022). Despite significant advancements in treatment, these conditions continue to be primary contributors to morbidity, mortality, and economic burden worldwide, (A et al., 2022); (Mph, 2024) and (Bloom et al., 2020).

Examples of chronic diseases with significant impact include: diabetic retinopathy (DR), a leading cause of recently diagnosed visual impairment, particularly among individuals of working age (Tsao et al., 2018), cancer, a major global mortality contributor responsible for approximately 13% of all global deaths(O. O. A & O.K, 2020); Alzheimer disease, a neurodegenerative disorder primarily affecting individuals aged 65 and older, progressively damages brain tissue; chronic kidney disease, characterized by a gradual and long-term loss of kidney function, continues to be asymptomatic until the kidneys have lost about 25% of their normal function (Pujianto et al., 2018); and cardiovascular diseases which also represent a leading cause of death, especially in developed countries (Atallah & Al-Mousa, 2019) and (Terrada & Bouattane, 2019). Consequently, addressing the risk factors for these diseases has become crucial (Alanazi, 2022) and (Battineni et al., 2020).

Though diverse, these risk factors for chronic diseases include inactivity, smoking, alcohol consumption, poor nutrition, overall lifestyle behaviors, and lifestyle. Environmental factors, such as prolonged exposure to sun, pollutants, and carcinogens also play a crucial role (A et al., 2022); (Mph, 2024) and (Alanazi, 2022). Global changes in agricultural practices, technological advancements, and cultural shifts have significantly impacted both behaviors and lifestyles worldwide. As a result, there has been an increasing risk of chronic diseases, including diabetes, cardiovascular disease, chronic kidney disease, and Alzheimer's disease (Culberson, 2017).

Chronic diseases pose several critical challenges, such as high treatment costs, late diagnoses, and poor prognoses, which can decrease survival rates, reduce quality of life, and, in some instances, lead to fatal outcomes (A et al., 2022). These diseases may result in various complications or necessitate long-term treatment (C. Lee et al., 2022). Hence, it is important to implement preemptive measures for the prevention of metabolic syndrome and develop prediction models for these diseases to reduce the risk of complications and medical costs (C. Lee et al., 2022).

Predicting chronic diseases using traditional methods presents several challenges that can restrict patient access to healthcare and hospital treatment (H. Wang et al., 2021). Relying solely on the established treatment guidelines may prove insufficient due to various factors, including complications and mental conditions (H. Wang et al., 2021). Furthermore, conventional techniques often struggle to effectively analyze the vast amounts of information related to chronic diseases for the reason that they are time-consuming and require extensive effort (Battineni et al., 2019).

To address these challenges, various machine learning techniques and algorithms are being developed. These methods have been extensively utilized in healthcare research to predict disease outcomes and have shown promising results. Selecting the appropriate algorithm is a complex task influenced by factors, such as data volume, the type of information available, and specific industry requirements regarding the desired outcomes (Battineni et al., 2019). Consequently, it is crucial to identify significant features and apply machine learning techniques to enhance chronic disease prediction performance (Rashid et al., 2022).

The present review is structured as follows: Section 2, providing a detailed survey of machine learning techniques, is organized into four main subsections, each dedicated to the prediction of a specific chronic disease such as Diabetes, Cardiovascular Diseases, Kidney Disease, and Alzheimer Disease. Further, each subsection outlines the machine learning methodologies employed, the datasets utilized, and the results obtained. Section 3 evaluates metrics, highlights the significant findings, and discusses the effectiveness of the techniques used for predicting chronic diseases. Finally, section 4 summarizes the conclusions drawn from this review, with an emphasis on its implications and significant potential for shaping future research.

Machine Learning in Chronic Disease Prediction

This section presents a comprehensive overview of recent studies that explore the use of machine learning techniques in the prediction of various chronic illnesses, such as diabetes, heart disease, kidney disorders, and

Alzheimer's disease. Relevant studies were identified through searches of reliable electronic databases, focusing on recent advancements in machine learning models designed for the prediction of these diseases. Subsections that follow explore the diverse machine learning methods employed in these studies to improve the predictive accuracy and support early diagnosis of chronic diseases.

Machine learning has been widely utilized in the prediction of several major chronic diseases such as diabetes, cardiovascular disease, chronic kidney disease, and Alzheimer's disease. Diabetes, one of the fastest-growing chronic illnesses worldwide, requires timely diagnosis and accurate prediction to reduce its severe impact on health, with ML techniques increasingly used to forecast critical events, such as hypo- and hyperglycemia (Khan et al., 2021) and (Hossain, 2024). Similarly, cardiovascular diseases continue to be leading causes of global mortality, where ML-based approaches support early detection, reduce risks of severe outcomes, such as heart attacks and strokes, and improve patient management (Pal et al., 2022) and (Panjiyar et al., 2023). Chronic kidney disease, affecting more than 800 million people globally, can progress silently over time, and ML models play an instrumental role in its early identification and monitoring by leveraging diverse clinical indicators (Khan et al., 2021) and (Thongprayoon et al., 2020). Alzheimer's disease, a progressive neurodegenerative disorder affecting approximately 50 million people worldwide, has also been the focus of predictive ML research, particularly through the integration of genetic, clinical, and neuroimaging data to allow earlier diagnosis and intervention (Porteri et al., 2017), (Bari Antor et al., 2021) and (Q. Wang et al., 2019). A concise summary of these studies, including applied ML methods, datasets, major contributions, and key findings, is presented in Table 1.

Table 1: Summary of ML Algorithms used in Chronic Diseases Prediction

	ML Method(s)	Key Contribution	Dataset	Best Result	Reference
Diabetes Disease	ANN, LSTM, CNN, Stack-ANN, Stack-LSTM, and Stack-CNN.	Ensemble deep learning model.	PIMA, DDFH-G, IDPD-I	99.51% accuracy	(Al Reshan et al., 2024a)
	DL	Early detection, handles class imbalance.	DPD	94% accuracy	(Shaheen et al., 2024)
	LightGB, KNN, SVM, NB, Bagging, RF, and XGB	Early diabetes diagnosis.	ZMHDD	98.1% accuracy	(Rufo et al., 2021)
	SVM + Ontology	Combined ontology with ML.	PIDD	Highest accuracy	(Massari et al., 2022)
	KNN, SVM, LR, and GB	Accurate detection model.		81.25% accuracy	(Panda et al., 2022)
	DT, KNN, RF, NB, AB, LR, SVM, and NN	Evaluated 7 ML algorithms		88.6% accuracy	(Khanam & Foo, 2021)
	KNN, DT, RF, AB, NB, and XGB	Handles outliers & missing values		95% accuracy	(Hasan et al., 2020)
	RF (with DMP_MI)	Handles missing values & imbalance using ADASYN + Naïve Bayes.		Outperformed SVM, NB, DT	(Q. Wang et al., 2019)
	NB, j48, LR, and RF	Compared 4 ML techniques for diabetes prediction.		Logistic Regression: 77% accuracy	(Battineni et al., 2019)
	DT, SVM, and NB	Compared 4 ML techniques.		Naïve Bayes 76.30 % accuracy	(Sisodia & Sisodia, 2018)
	SVM, DT, ANN, and LR	Predict diabetic retinopathy using biomedical feature.	Taiwan" DM Shared Care"	SVM: 79.5% accuracy, 0.839 AUC	(Tsao et al., 2018)
Cardiovascular Disease	KNN, SVM, LR, CNN, GB, XGB, RF	Tested multiple classifiers with imbalance handling.	Heart Disease datasets (UCI + IEEE)	XGBoost (fine-tuned): 98.50% accuracy	(Ogunpola et al., 2024)
	Self-Supervised Learning (SSL)	Two-phase SSL for cardiovascular event prediction.	Pretext + Downstream (483 patients)	Increased accuracy from 63% → 90%	(Chen et al., 2024)
	SVM, LR	Compared SVM & LR for heart disease prediction.	UCI Heart Disease dataset	97.25% accuracy (outperformed CNN, XGB, RF)	(Kumar et al., 2023)

	ML Method(s)	Key Contribution	Dataset	Best Result	Reference
	HRFLM (RF + DT + NN + SVM)	Hybrid model combining multiple classifiers.	Cleveland Heart Disease	88.7% accuracy	(Mohan et al., 2019)
	SVM, LR, ANN, KNN, NB, DT (with FCMIM)	Proposed FCMIM for feature selection.		FCMIM-SVM: 92.37% accuracy	(Li et al., 2020)
	ANN, CNN, LSTM, Hybrid CNN-LSTM	HDNN hybrid model for heart disease prediction.	Cleveland + Combined (5 datasets)	98.86% accuracy, AUC=1.00	(Reshan et al., 2023)
	NB, SVM, Voting, XGB, AB, Bagging, DT, KNN, RF, LR	Compared 10 ML algorithms with feature selection.	Cleveland + Private dataset	XGBoost: 97.57% accuracy	(El-Sofany, 2024)
	MLP, SVM, RF, L-GB	CVD diagnosis with risk analysis, handling imbalance.	KNHANES (Korea)	All models: AUC > 0.853	(Oh et al., 2022)
Kidney Disease	LR, DT, SVM	Evaluated 3 ML techniques for CKD prediction,	UCI ML Repository	DT: 95% accuracy → 97% with Bagging	(Kaur et al., 2023)
	RF, SVM, DT, XGB (with RFECV)	Evaluated 4 ML techniques with RFECV and UFS.	St. Paulo's Hospital, Ethiopia	SVM: 99.8% accuracy (binary); XGBoost: 82.56% (5-class)	(Debal & Sitote, 2022)
	DBN	Deep learning model for early CKD detection.	UCI CKD Dataset	98.5% accuracy, 87.5% sensitivity	(Elkholy et al., 2021)
	ANN, C5.0, CHAID, LR, LSVM, RT	Evaluated 7 algorithms with feature selection.		DNN: 99.6% accuracy; LSVM (L2): 98.86%	(Chittora et al., 2021)
	LR, RF, SVM, KNN, NB, NN	Handled missing values with KNN, evaluated 6 techniques.		LR + RF: 99.83% accuracy	(Qin et al., 2020)
	LR, SVM, KNN	Evaluated 4 ML techniques		SVM: 99.25% accuracy	(Gudeti et al., 2020)
	SVM, AB, LDA, GB	Evaluated 4 ML techniques		Gradient Boosting: 99.80% accuracy	(Ghosh et al., 2020)

	ML Method(s)	Key Contribution	Dataset	Best Result	Reference
	Deep Learning (Stacked Autoencoder + Softmax)	Deep learning model using stacked autoencoders for accurate classification.		100% classification accuracy	(Khamparia et al., 2019)
	DT, SVM, RF	Evaluated 3 ML techniques for CKD prediction.		Random Forest: 99.16% accuracy	(Revathy et al., 2019)
	LR	Predict diabetic kidney disease progression using EMR.	EMR dataset (64,059 patients)	71% accuracy	(Makino et al., 2019)
Alzheimer Disease	SVM, LR, DT, RF	Evaluated 4 ML techniques for Alzheimer's prediction	OASIS dataset	SVM: 92% accuracy	(Bari Antor et al., 2021)
	IVM, RELM, SVM	Used MCI data and sMRI for Alzheimer's recognition	ADNI dataset	RELM outperformed SVM & IVM	(Sudharsan & Thailambal, 2021b)
	KNN, DT, Rule Induction, NB, GLM, DL	Assessed 6 ML models for Alzheimer's prediction	ADNI (TADPOLE dataset)	GLM: 88.24% accuracy	(Shahbaz et al., 2019)
	SVM, NB, NN, RF	ML framework integrating gene expression from 6 brain regions	NCBI GEO (GSE5281)	Identified 44 genes as potential Alzheimer's biomarkers	(L. Wang & Liu, 2019)
	SVM, RF, LR, AB	4-stage ML framework: preprocessing, gene selection, classification, prediction	Combined Alzheimer dataset	SVM achieved highest performance	(El-Gawady et al., 2022)
	LR, L1-GLM, RF, SVM, DNN	Feature selection with ML to identify Alzheimer's-associated genes	ADNI, ANM1, ANM2	Best AUC: ANM1=0.874, ANM2=0.804, ADNI=0.657	(T. Lee & Lee, 2020)
	GaussianNB, DT, RF, XGB, Voting Classifier, GB	Innovative ML approach for early Alzheimer's detection	OASIS longitudinal dataset	Voting Classifier: 96% accuracy	(Uddin et al., 2023a)

Evaluation Metrics and Findings

Analysis of Performance Metrics in Chronic Disease Prediction

Recent studies emphasize the importance of selecting evaluation metrics specifically designed to address the unique challenges involved in predicting chronic diseases. While accuracy is a commonly reported metric for all diseases and provides a general indication of model performance, it has limitations when dealing with missing values and imbalanced datasets. Therefore, it is essential to consider additional metrics, such as Precision, Recall, F1-Score, Specificity, and AUC.

A review of previous research reveals that performance metrics are commonly used in chronic disease prediction, with variations observed across different diseases and studies. For instance, in predicting diabetes and Alzheimer disease, Precision, which is particularly important to reduce false positives, helps minimize unnecessary follow-up tests and treatments (Al Reshan et al., 2024b), (Shaheen et al., 2024), (Panda et al., 2022), (Massari et al., 2022), (Sudharsan & Thailambal, 2021a), (Uddin et al., 2023a), (El-Gawady et al., 2022) and (T. Lee & Lee, 2020). Notably, diabetes studies utilize all evaluation metrics except that for Specificity, highlighting the versatility of these criteria for this condition. In contrast, Precision is utilized in only a few Alzheimer studies, indicating its limited adoption despite its significant potential to identify true positive cases accurately. This observation from the literature suggests the need for further exploration to better leverage Precision’s ability to allocate resources efficiently to individuals requiring more diagnostic evaluation and care.

In the prediction of cardiovascular and kidney diseases, Recall is prioritized to ensure that high-risk cases are accurately identified, which is crucial for early intervention. However, as noted in previous studies, only a few studies on these diseases have incorporated Recall, indicating that its application remains limited (Ogunpola et al., 2024), (Chen et al., 2024), (El-Sofany, 2024), (Kaur et al., 2023), and (Debal & Sitote, 2022). This selective use underscores an opportunity to more widely adopt Recall, potentially improving the identification of at-risk individuals and enhancing the effectiveness of early intervention strategies.

Alzheimer disease prediction poses additional challenges, particularly in detecting early-stage and subtle symptoms. In such complex classification settings, the F1-Score is often employed to balance the trade-off between false positives and false negatives. This metric ensures reliable performance, where both types of errors can have significant consequences. Despite its advantages, only two studies identified in this review have utilized the F1-Score for Alzheimer prediction (Uddin et al., 2023b) and (Atallah, 2019). This limited adoption underscores the need for further exploration of the F1-Score's utility, especially in handling imbalanced datasets and reducing diagnostic uncertainty during early disease progression.

In summary, the variation in metric selection downplays the importance of adapting evaluation approaches to the unique challenges of each chronic disease. This approach is essential for ensuring both the clinical relevance and the effectiveness of predictive models in real-world healthcare applications.

Comparative Analysis of ML Methods for Chronic Disease Prediction

This review included studies that used various algorithms, such as deep learning techniques (Al Reshan et al., 2024a), (Shaheen et al., 2024), (Reshan et al., 2023), (Elkholy et al., 2021); (Khamparia et al., 2019), (T. Lee & Lee, 2020), SVM (Rufo et al., 2021), (Massari et al., 2022), (Panda et al., 2022), (Ogunpola et al., 2024), (Kumar et al., 2023), DT, KNN, and LF (Massari et al., 2022), (Khanam & Foo, 2021), (El-Sofany, 2024), RF (Bari Antor et al., 2021), (Q. Wang et al., 2019) and (T. Lee & Lee, 2020). These studies provided valuable insights into the relevant variables and risk factors and contributed to disease prediction and guided our feature selection process. Our focus was on the four major chronic conditions: diabetes, kidney disease, cardiovascular disease, and Alzheimer.

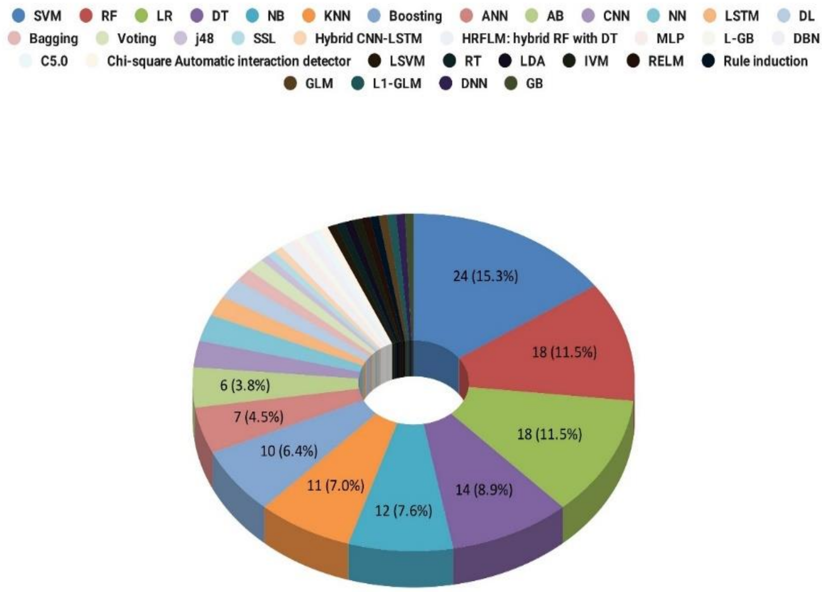

Figure 1: Application of ML Algorithms Frequency

Table 2: Best Results of Chronic Disease Prediction in Studies, Datasets, and Algorithms

Study	Disease	Dataset	ML Algorithm	Best Accuracy (%)
(Al Reshan et al., 2024a)	Diabetes	PIMA-IDD-I, DDFH-G, and IDPD-I	stack-ANN	99.51%
(Reshan et al., 2023)	Cardiovascular	Data from (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA)	CNN-LSTM	98.86%
(Qin et al., 2020)	Kidney	St. UCI CKD Dataset	LR+RF	99.83%
(Uddin et al., 2023a)	Alzheimer	Imaging Studies (OA-SIS) data	Voting Classifier	96%

Figure (1) illustrates the frequency of machine learning algorithms applied across the reviewed studies. It shows that Ensemble and deep learning techniques demonstrated high accuracy in predicting chronic diseases like diabetes, cardiovascular disease, kidney disease, and Alzheimer. For instance, the Ensemble Deep Learning model achieved an accuracy of 99.51% in diabetes prediction.

Effective feature selection and data preprocessing were essential for improving prediction accuracy. Techniques like Proximity-Weighted Synthetic Oversampling (ProWSyn) and Recursive Feature Elimination with Cross Validation (RFECV) played a key role in enhancing model performance by refining the input data. Feature importance evaluations uncovered a set of dominant attributes that exerted substantial influence on the predictive outcomes across models. In diabetes prediction, attributes including glucose level, age, and BMI were identified as the most influential. Similarly, for cardiovascular disease, critical features included cholesterol levels (chol), maximum heart rate (thalach), and blood pressure (oldpeak). The XGBoost model, in particular, could accurately rank these features and further optimizing its predictive power.

The performance of various algorithms varied depending on the disease. For instance, Random Forest and XGBoost achieved high accuracy in predicting cardiovascular disease, while support vector machine proved to be particularly effective in the detection of Alzheimer disease.

The present study evaluated the performance of various machine learning algorithms in predicting chronic diseases, such as diabetes, heart disease, and kidney disease. The tested models included classical machine learning algorithms (e.g., Logistic Regression, k-Nearest Neighbors, Support Vector Machines), and more advanced techniques including XGBoost, LightGBM, Convolutional Neural Networks.

For diabetes prediction, the LightGBM model demonstrated superior performance compared to traditional algorithms. It achieved an accuracy of 98.1%, an AUC (Area Under the Curve) of 98.1%, sensitivity of 99.9%, and specificity of 96.3%. These results outperformed models like KNN, SVM, and XGBoost, which, despite their effectiveness, had slightly decreased accuracy and precision.

In cardiovascular disease, the XGBoost model was determined to be the most accurate, achieving an accuracy rate of 98.50%, thereby outperforming models such as RF and SVM. The high performance of XGBoost can be attributed to its depth-first tree-building approach, which enables it to effectively capture complex feature interactions. This renders it particularly suitable for heart disease prediction, where numerous interdependent risk factors (e.g., cholesterol levels, age, and blood pressure) need to be considered.

For chronic kidney disease prediction, a combination of Logistic Regression (LR) and Random Forest achieved an accuracy of 99.83%, thereby outperforming models such as SVM, KNN, and Neural Networks (NN). This integrated approach proved to be highly effective, leveraging both the linear modeling of LR and the ensemble learning of RF to enhance predictive accuracy.

Table (2) shows a summary of the best results obtained for each chronic disease, including the specific datasets used and the algorithms that yielded the highest performance across the reviewed studies.

Conclusion

Machine learning has emerged as a powerful tool in the prediction of chronic diseases, offering significant improvements in diagnostic accuracy, efficiency, and cost-effectiveness over traditional methods. This review concluded that advanced ML models, such as deep learning techniques and boosting algorithms, consistently outperform conventional approaches, achieving exceptional accuracies reaching 99.51% in chronic disease prediction. Furthermore, data preprocessing and feature selection were shown to be essential in optimizing the performance of the ML models. However, the generalizability of these models continues to be limited due to the use of single and specific datasets, which may not represent diverse populations.

References

A, D. P., A, S. G., A, E. K., A, A. K., Mohammed S. Alqahtani b c, Mohamed Abbas d e, G. S. f, & Ajaikumar B. Kunnumakkara. (2022). *An overview of the pharmacological activities of scopoletin against different chronic diseases*. https://www.sciencedirect.com/science/article/abs/pii/S1043661822001475

A, O. O., & O.K, B. (2020). Prediction of Prostate Cancer using Ensemble of Machine Learning Techniques. International Journal of Advanced Computer Science and Applications, 11(3). https://doi.org/10.14569/ijacsa.2020.0110318

Al Reshan, M. S., Amin, S., Zeb, M. A., Sulaiman, A., Alshahrani, H., Shaikh, A., & Elmagzoub, M. A. (2024a). An Innovative Ensemble Deep Learning Clinical Decision Support System for Diabetes Prediction. *IEEE Access*, *12*(March), 106193–106210. https://doi.org/10.1109/ACCESS.2024.3436641

Al Reshan, M. S., Amin, S., Zeb, M. A., Sulaiman, A., Alshahrani, H., Shaikh, A., & Elmagzoub, M. A. (2024b). An Innovative Ensemble Deep Learning Clinical Decision Support System for Diabetes Prediction. *IEEE Access*, *12*(May), 106193–106210. https://doi.org/10.1109/ACCESS.2024.3436641

Alanazi, R. (2022). Ide ntification and Prediction of Chronic Diseases Using Machine Learning Approach. *Journal of Healthcare Engineering*, *2022*. https://doi.org/10.1155/2022/2826127

Atallah, R. (2019). Heart Disease Detection Using Machine Learning Majority Voting Ensemble Method. *2019 2nd International Conference on New Trends in Computing Sciences (ICTCS)*, 1–6.

Atallah, R., & Al-Mousa, A. (2019). Heart Disease Detection Using Machine Learning Majority Voting Ensemble Method. *2019 2nd International Conference on New Trends in Computing Sciences, ICTCS 2019 - Proceedings*, *October 2019*. https://doi.org/10.1109/ICTCS.2019.8923053

Bari Antor, M., Jamil, A. H. M. S., Mamtaz, M., Monirujjaman Khan, M., Aljahdali, S., Kaur, M., Singh, P., & Masud, M. (2021). A Comparative Analysis of Machine Learning Algorithms to Predict Alzheimer's Disease. *Journal of Healthcare Engineering*, *2021*. https://doi.org/10.1155/2021/9917919

Battineni, G., Chintalapudi, N., & Amenta, F. (2019). Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM). *Informatics in Medicine Unlocked*, *16*(June), 100200. https://doi.org/10.1016/j.imu.2019.100200

Battineni, G., Sagaro, G. G., Chinatalapudi, N., & Amenta, F. (2020). Applications of machine learning predictive models in the chronic disease diagnosis. *Journal of Personalized Medicine*, *10*(2). https://doi.org/10.3390/jpm10020021

Bloom, D. E., Chen, S., Kuhn, M., McGovern, M. E., Oxley, L., & Prettner, K. (2020). The economic burden of chronic diseases: Estimates and projections for China, Japan, and South Korea. *Journal of the Economics of Ageing*, *17*. https://doi.org/10.1016/j.jeoa.2018.09.002

Chen, L. C., Hung, K. H., Tseng, Y. J., Wang, H. Y., Lu, T. M., Huang, W. C., & Tsao, Y. (2024). Self-Supervised Learning-Based General Laboratory Progress Pretrained Model for Cardiovascular Event Detection. *IEEE Journal of Translational Engineering in Health and Medicine*, *12*(March 2023), 43–55. https://doi.org/10.1109/JTEHM.2023.3307794

Chittora, P., Chaurasia, S., Chakrabarti, P., Kumawat, G., Chakrabarti, T., Leonowicz, Z., Jasinski, M., Jasinski, L., Gono, R., Jasinska, E., & Bolshev, V. (2021). Prediction of Chronic Kidney Disease - A Machine Learning Perspective. *IEEE Access*, *9*, 17312–17334. https://doi.org/10.1109/ACCESS.2021.3053763

Culberson, J. W. (2017). *Clinical Aspects of Glucose Metabolism and Chronic Disease* (pp. 1–11). https://doi.org/10.1016/bs.pmbts.2016.12.011

Debal, D. A., & Sitote, T. M. (2022). Chronic kidney disease prediction using machine learning techniques. *Journal of Big Data*, *9*(1). https://doi.org/10.1186/s40537-022-00657-5

El-Gawady, A., Makhlouf, M. A., Tawfik, B. S., & Nassar, H. (2022). Machine Learning Framework for the Prediction of Alzheimer's Disease Using Gene Expression Data Based on Efficient Gene Selection. *Symmetry*, *14*(3). https://doi.org/10.3390/sym14030491

El-Sofany, H. F. (2024). Predicting Heart Diseases Using Machine Learning and Different Data Classification Techniques. *IEEE Access*, *12*(July), 106146–106160. https://doi.org/10.1109/ACCESS.2024.3437181

Elkholy, S. M. M., Rezk, A., & Saleh, A. A. E. F. (2021). Early Prediction of Chronic Kidney Disease Using Deep Belief Network. *IEEE Access*, *9*, 135542–135549. https://doi.org/10.1109/ACCESS.2021.3114306

Ghosh, P., Javed Mehedi Shamrat, F. M., Shultana, S., Afrin, S., Anjum, A. A., & Khan, A. A. (2020). Optimization of Prediction Method of Chronic Kidney Disease Using Machine Learning Algorithm. *Proceedings - 2020 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing, ISAI-NLP 2020*. https://doi.org/10.1109/iSAI-NLP51646.2020.9376787

Gudeti, B., Mishra, S., Malik, S., Fernandez, T. F., Tyagi, A. K., & Kumari, S. (2020). A Novel Approach to Predict Chronic Kidney Disease using Machine Learning Algorithms. *Proceedings of the 4th International*

Conference on Electronics, Communication and Aerospace Technology, ICECA 2020, 1630–1635. https://doi.org/10.1109/ICECA49313.2020.9297392

Hasan, M. K., Alam, M. A., Das, D., Hossain, E., & Hasan, M. (2020). Diabetes prediction using ensembling of different machine learning classifiers. *IEEE Access*, *8*, 76516–76531. https://doi.org/10.1109/ACCESS.2020.2989857

Hossain, J. (2024). *Diabetes mellitus, the fastest growing global public health concern: Early detection should be focused*. 5–9. https://doi.org/10.1002/hsr2.2004

Kaur, C., Kumar, M. S., Anjum, A., Binda, M. B., Mallu, M. R., & Ansari, M. S. Al. (2023). Chronic Kidney Disease Prediction Using Machine Learning. *Journal of Advances in Information Technology*, *14*(2), 384–391. https://doi.org/10.12720/jait.14.2.384-391

Khamparia, A., Saini, G., Pandey, B., & Tiwari, S. (2019). *KDSAE: Chronic kidney disease classification with multimedia data learning using deep stacked autoencoder network.*

Khan, F. A., Zeb, K., Al-Rakhami, M., Derhab, A., & Bukhari, S. A. C. (2021). Detection and Prediction of Diabetes Using Data Mining: A Comprehensive Review. *IEEE Access*, *9*, 43711–43735. https://doi.org/10.1109/ACCESS.2021.3059343

Khanam, J. J., & Foo, S. Y. (2021). A comparison of machine learning algorithms for diabetes prediction. *ICT Express*, *7*(4), 432–439. https://doi.org/10.1016/j.icte.2021.02.004

Kumar, A., Singh, K. U., & Kumar, M. (2023). A Clinical Data Analysis Based Diagnostic Systems for Heart Disease Prediction Using Ensemble Method. *Big Data Mining and Analytics*, *6*(4), 513–525. https://doi.org/10.26599/BDMA.2022.9020052

Lee, C., Jo, B., Woo, H., Im, Y., Park, R. W., & Park, C. (2022). Chronic Disease Prediction Using the Common Data Model: Development Study. *JMIR AI*, *1*(1), e41030. https://doi.org/10.2196/41030

Lee, T., & Lee, H. (2020). Prediction of Alzheimer' s disease using blood gene expression data. *Scientific Reports*, 1–13. https://doi.org/10.1038/s41598-020-60595-1

Li, J. P., Haq, A. U., Din, S. U., Khan, J., Khan, A., & Saboor, A. (2020). Heart Disease Identification Method Using Machine Learning

Classification in E-Healthcare. *IEEE Access*, *8*(Ml), 107562–107582. https://doi.org/10.1109/ACCESS.2020.3001149

Makino, M., Yoshimoto, R., Ono, M., Itoko, T., Katsuki, T., Koseki, A., Kudo, M., Haida, K., Kuroda, J., Yanagiya, R., Saitoh, E., Hoshinaga, K., Yuzawa, Y., & Suzuki, A. (2019). Artificial intelligence predicts the progression of diabetic kidney disease using big data machine learning. *Scientific Reports 2019 9:1*, *9*(1), 1–9. https://doi.org/10.1038/s41598-019-48263-5

Massari, H. El, Sabouri, Z., Mhammedi, S., & Gherabi, N. (2022). Diabetes Prediction Using Machine Learning Algorithms and Ontology. *Journal of ICT Standardization*. https://doi.org/10.13052/jicts2245-800X.10212

Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access*, *7*, 81542–81554. https://doi.org/10.1109/ACCESS.2019.2923707

Mph, K. H. (2024). The Burden of Chronic Disease. *Mayo Clinic Proceedings: Innovations, Quality & Outcomes*, *8*(1), 112–119. https://doi.org/10.1016/j.mayocpiqo.2023.08.005

Ogunpola, A., Saeed, F., Basurra, S., Albarrak, A. M., & Qasem, S. N. (2024). Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases. *Diagnostics*, *14*(2). https://doi.org/10.3390/DIAGNOSTICS14020144

Oh, T., Kim, D., Lee, S., Won, C., Kim, S., Yang, J. soo, Yu, J., Kim, B., & Lee, J. (2022). Machine learning-based diagnosis and risk factor analysis of cardiocerebrovascular disease based on KNHANES. *Scientific Reports 2022 12:1*, *12*(1), 1–11. https://doi.org/10.1038/s41598-022-06333-1

Pal, M., Parija, S., Panda, G., Dhama, K., & Mohapatra, R. K. (2022). Risk prediction of cardiovascular disease using machine learning classifiers. *Open Medicine (Poland)*, *17*(1), 1100–1113. https://doi.org/10.1515/med-2022-0508

Panda, M., Mishra, D. P., Patro, S. M., & Salkuti, S. R. (2022). Prediction of diabetes disease using machine learning algorithms. *IAES International Journal of Artificial Intelligence*, *11*(1), 284–290. https://doi.org/10.11591/ijai.v11.i1.pp284-290

Panjiyar, B. K., Davydov, G., Nashat, H., Ghali, S., Afifi, S., Suryadevara, V., Habab, Y., Hutcheson, A., & Arcia Franchini, A. P. (2023). A

Systematic Review: Do the Use of Machine Learning, Deep Learning, and Artificial Intelligence Improve Patient Outcomes in Acute Myocardial Ischemia Compared to Clinician-Only Approaches? *Cureus*, *15*(8), 1–11. https://doi.org/10.7759/cureus.43003

Porteri, C., Albanese, E., Scerri, C., Carrillo, M. C., Snyder, H. M., Martensson, B., Baker, M., Giacobini, E., Boccardi, M., Winblad, B., Frisoni, G. B., & Hurst, S. (2017). The biomarker-based diagnosis of Alzheimer's disease. 1—ethical and societal issues. *Neurobiology of Aging*, *52*, 132–140. https://doi.org/https://doi.org/10.1016/j.neurobiolaging.2016.07.011

Pujianto, U., Ramadhani, N. A. yun., & Wibawa, A. P. (2018). Support Vector Machine with Purified K-Means Clusters for Chronic Kidney Disease Detection. *Proceedings - 2nd East Indonesia Conference on Computer and Information Technology: Internet of Things for Industry, EIConCIT 2018*, 56–60. https://doi.org/10.1109/EIConCIT.2018.8878511

Qin, J., Chen, L., Liu, Y., Liu, C., Feng, C., & Chen, B. (2020). A machine learning methodology for diagnosing chronic kidney disease. *IEEE Access*, *8*, 20991–21002. https://doi.org/10.1109/ACCESS.2019.2963053

Rashid, J., Batool, S., Kim, J., Nisar, M. W., & Hussain, A. (2022). *An Augmented Artificial Intelligence Approach for Chronic Diseases Prediction*. *10*(March), 1–20. https://doi.org/10.3389/fpubh.2022.860396

Reshan, M. S. Al, Amin, S., Zeb, M. A., Sulaiman, A., Alshahrani, H., & Shaikh, A. (2023). A Robust Heart Disease Prediction System Using Hybrid Deep Neural Networks. *IEEE Access*, *11*(November), 121574–121591. https://doi.org/10.1109/ACCESS.2023.3328909

Revathy, S., Bharathi, B., Jeyanthi, P., & Ramesh, M. (2019). Chronic kidney disease prediction using machine learning models. *International Journal of Engineering and Advanced Technology*, *9*(1), 6364–6367. https://doi.org/10.35940/ijeat.A2213.109119

Rufo, D. D., Debelee, T. G., Ibenthal, A., & Negera, W. G. (2021). Diagnosis of diabetes mellitus using gradient boosting machine (Lightgbm). *Diagnostics*, *11*(9), 1–14. https://doi.org/10.3390/diagnostics11091714

Shahbaz, M., Ali, S., Guergachi, A., Niazi, A., & Umer, A. (2019). Classification of Alzheimer's disease using machine learning techniques. *DATA 2019 - Proceedings of the 8th International Conference on Data Science, Technology and Applications*, *Data*, 296–303. https://doi.org/10.5220/0007949902960303

Shaheen, I., Javaid, N., Alrajeh, N., Asim, Y., & Aslam, S. (2024). Hi-Le and HiTCLe: Ensemble Learning Approaches for Early Diabetes Detection Using Deep Learning and Explainable Artificial Intelligence. *IEEE Access*, *12*(May), 66516–66538. https://doi.org/10.1109/ACCESS.2024.3398198

Sisodia, D., & Sisodia, D. S. (2018). ScienceDirect Prediction of Diabetes using Classification Algorithms. *Procedia Computer Science*, *132*(Iccids), 1578–1585. https://doi.org/10.1016/j.procs.2018.05.122

Sudharsan, M., & Thailambal, G. (2021a). Alzheimer's disease prediction using machine learning techniques and principal component analysis (PCA). *Materials Today: Proceedings*, *81*(2), 182–190. https://doi.org/10.1016/j.matpr.2021.03.061

Sudharsan, M., & Thailambal, G. (2021b). *Materials Today: Proceedings Alzheimer' s disease prediction using machine learning techniques and principal component analysis (PCA)*. *xxxx*.

Terrada, O., & Bouattane, O. (2019). Classification and Prediction of atherosclerosis diseases using machine learning algorithms. *2019 5th International Conference on Optimization and Applications (ICOA)*, 1–5.

Thongprayoon, C., Kaewput, W., Kovvuru, K., Hansrivijit, P., Kanduri, S. R., Bathini, T., Chewcharat, A., Leeaphorn, N., Gonzalez-Suarez, M. L., & Cheungpasitporn, W. (2020). Promises of big data and artificial intelligence in nephrology and transplantation. *Journal of Clinical Medicine*, *9*(4). https://doi.org/10.3390/jcm9041107

Tsao, H.-Y., Chan, P.-Y., & Su, E. C.-Y. (2018). Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms. *BMC Bioinformatics*, *19*(S9), 283. https://doi.org/10.1186/s12859-018-2277-0

Uddin, K. M. M., Alam, M. J., Jannat-E-Anawar, Uddin, M. A., & Aryal, S. (2023a). A Novel Approach Utilizing Machine Learning for the Early Diagnosis of Alzheimer's Disease. *Biomedical Materials and Devices*, *1*(2), 882–898. https://doi.org/10.1007/s44174-023-00078-9

Uddin, K. M. M., Alam, M. J., Jannat-E-Anawar, Uddin, M. A., & Aryal, S. (2023b). A Novel Approach Utilizing Machine Learning for the Early Diagnosis of Alzheimer's Disease. *Biomedical Materials & Devices*, *1*(2), 882–898. https://doi.org/10.1007/s44174-023-00078-9

Wang, H., Yuan, X., & Wang, J. (2021). Telemedicine maybe an effective solution for management of chronic disease during the COVID-19 epidemic. *Primary Health Care Research and Development*, *22*, 1231–1235. https://doi.org/10.1017/S1463423621000517

Wang, L., & Liu, Z. P. (2019). Detecting diagnostic biomarkers of Alzheimer's disease by integrating gene expression data in six brain regions. *Frontiers in Genetics*, *10*(MAR). https://doi.org/10.3389/fgene.2019.00157

Wang, Q., Cao, W., Guo, J., Ren, J., Cheng, Y., & Davis, D. N. (2019). DMP_MI: An effective diabetes mellitus classification algorithm on imbalanced data with missing values. *IEEE Access*, *7*, 102232–102238. https://doi.org/10.1109/ACCESS.2019.2929866