AVE Trends in Intelligent Health Letters

AVE TRENDS PUBLICATIONS

Leveraging AI for Early Detection of Chronic Diseases through Patient Data Integration

Sunil Kumar Sehrawat*
Department of Information Technology Management, Bausch Health Companies, Hillsborough, New Jersey, United States of America. sunil.sehrawat@bauschhealth.com

*Corresponding author

Abstract: The application of artificial intelligence (AI) in the medical system has immense capability to redesign the diagnosis and treatment of chronic diseases during their initial stage. Early detection is required in an attempt to enhance patient outcomes, and utilization of patient information using AI algorithms can offer the pathway to identifying warning signs earlier. The study utilizes a big data set of the local hospital's electronic health record, such as patient characteristics, clinical histories, laboratory results, and life information. The information is analyzed through AI models, such as machine learning algorithms like decision trees, random forests, and neural networks, in order to predict the probability of the occurrence of disease. Analysis tools used are Python and frameworks like Scikit-learn for machine learning, TensorFlow for deep learning, and Matplotlib for plotting. AI models can diagnose diseases in the early stage before clinical signs and symptoms develop due to earlier patient data. The article also enlists the implementation of data fusion methodologies. The article attempts to convey significant developments in technology and improved strategies, as well as the day-to-day application and limitations of AI in identifying chronic diseases.

Keywords: Artificial Intelligence; Chronic Diseases; Early Detection; Patient Data Integration; Healthcare Technology; Cardiovascular Disease (CVD); Electronic Health Records (EHR); Diagnostic Machines.

Cite as: S. K. Sehrawat, "Leveraging AI for Early Detection of Chronic Diseases through Patient Data Integration," *AVE Trends In Intelligent Health Letters*, vol. 1, no. 3, pp. 125 –136, 2024.

Journal Homepage: https://avepubs.com/user/journals/details/ATIHL

Received on: 29/01/2024, **Revised on:** 29/03/2024, **Accepted on:** 10/05/2024, **Published on:** 05/09/2024

1. Introduction

Chronic diseases will be among the top causes of morbidity and mortality globally. Some of the diseases are diabetes, cardiovascular disease (CVD), respiratory disease, and cancer, which account for about 70% of all deaths globally, as stated by WHO. Despite such diseases being so prevalent, they will be diagnosed only at very advanced stages at which they are much more difficult to be successfully cured. Early diagnosis remains crucial in the optimization of treatment success, the prevention of complications, and the optimization of the overall cost of healthcare [1]. Artificial Intelligence has also greatly improved medicine, with new solutions suggested for virtually all the challenges, including the early diagnosis of chronic diseases. The reason behind the popularity of AI, in the guise of ML and DL algorithms, is that it learns well and effectively from huge databases and can determine patterns that human physicians cannot easily find [2]. AI can be trained using different patient records like medical history, genetic data, demographic data, and lifestyle data to estimate the probability of occurrence of chronic diseases. AI does not only detect high-risk patients but also aids healthcare professionals in early and well-informed decisions in patient management [3].

This application of AI in healthcare systems is also promoted by Satyanarayanan [4], who explains the application of edge computing and cloud infrastructure to maximize the efficiency of AI applications within the healthcare system. This convergence of technology facilitates the fast processing of large data, thus facilitating real-time processing and early detection of chronic diseases. These technologies are counted in the sense that they deliver interventions in a timely manner, especially among rural populations where medical care accessibility might be substandard. Patient data integration is involved [16]. The healthcare industry produces massive amounts of data from numerous sources, such as electronic health records (EHR), radiology images, lab results, wearable fitness trackers, and mobile health apps [17]. Merging such heterogeneous information into one database renders it easy to develop an integrated patient profile, which gives AI software the information to make the right predictions [5]. It is easier to develop more accurate and comprehensive diagnostic tools that will do a great deal in identifying risky individuals. Applications of the system integrated and AI model have been seen to enhance the level of early detection by leaps and reduce diagnostic error [18]. The same has been stated once again by Kamruzzaman et al. [6], who saw how AI facilitated the building of more informative and accurate forecast models. With every new doctor using AI-based decision-making in their practice, the possibilities of AI transforming healthcare's early detection are within reach. Blending machine learning techniques with clinical pathways is changing medical diagnostics and disease treatment from the pinnacle of enhanced patient outcomes [19].

Never before have early detection capabilities been so important as they are today. Early diagnosis of chronic disease offers physicians more treatment options, advances the chances of favourable patient outcomes, and lowers the overall cost of treatment for advanced chronic disease over the long term [7]. According to this, Dong et al. [8] assume, as a matter of course, that AI models, in sifting through large amounts of health data, can accurately forecast disease trajectories and thus are a valuable resource for physicians to rule out those patients for whom early intervention would be most beneficial. Besides, Layeghian Javan et al. [9] cite that early diagnosis by AI systems can theoretically allow the healthcare systems to make better use of resources without spending on curing advanced Diseases. Other than this, research like Hong and Lee [10] and Thipsawat [11] even recently discuss the realities of actual implementation of AI-based early warning systems, i.e., the privacy of data and how stringent testing of AI models should be made mandatory. All these concerns are discussed, and they form the foundation for the deployment of AI solutions in the healthcare industry. Wander et al. [12] also state that to rejuvenate the systems, they need to change so that their integration of AI into health systems can be enabled for safety and clinical effectiveness.

Lastly, Schiffman et al. [13] advocated that AI is capable of managing complex multi-source data, and hence, AI offers a valuable asset in the direction of future chronic disease management. With increasingly more evidence being placed on the table, it's only natural that AI will increasingly be at the forefront of revolutionizing chronic disease diagnosis and treatment. Moshawrab et al. [14] and Moshawrab et al. [15] again emphasize, in recent articles, the future AI to transform medicine as being most likely to optimize the precision and limit the bias of the prediction models so that pre-clinical diagnosis will be enhanced. This research examines how the presence of AI-embedded patient data affects early chronic disease diagnosis by reference to the experience of practice with such systems in real practice and how healthcare providers, as well as patients, benefit from them [20].

2. Review of Literature

Yu et al. [1] explored that Artificial intelligence and medicine have been a new area of research in the past decades. The rapid growth of computing power and the availability of extensive databases enabled the application of AI technologies in the healthcare industry for a variety of uses. The prime use of AI in healthcare is the early detection of chronic disease, particularly if combined with patient data convergence. Early disease detection is crucial for the prevention of disease progression and the patient's quality of life. With AI, healthcare providers can detect any possible harm to patients' well-being much earlier, which assists in making the interventions in time. This type of expertise with AI has been colossal in focus in an effort to use it to predict chronic diseases. The potential of revolutionizing medicine by utilizing the power of AI and the deployment of predictive power is immense and remains one of the areas of research focus.

Malik et al. [2] brought to the forefront the need for early diagnosis in managing chronic diseases effectively. The research has identified early diagnosis of chronic diseases like diabetes, hypertension, and cardiovascular diseases to be significant in preventing the risk of complications. Machine learning models have worked well to forecast patient determinants, history, and demographics and can be utilized to construct the likelihood of such conditions as diabetes. Researchers have constructed such varied determinants-based models. These models enable physicians to intervene with preventive interventions prior to the diseases taking hold. The promise of early diagnosis is patient-specific treatment methods that are tailored to measure. With new developments in AI technologies, disease diagnosis is made even more effective, and preventive medicine comes into play.

Bohr and Memarzadeh [3] show that AI-based systems might outperform traditional methods of diagnosis in terms of productivity and efficiency. The ability of machine learning techniques such as random forests, support vector machines, and neural networks to handle complex data rendered them effective in the healthcare sector. The algorithms are capable of outperforming in chronic disease prediction from patient information where other algorithms would fail. AI is capable of noticing patterns and correlations in data that physicians' eyes can miss. Artificial intelligence systems are capable of processing more data and volume of data than human beings, so diagnostic precision is enhanced. Therefore, health professionals are capable of making better decisions at a quicker pace, and hence, patient care is enhanced overall. Increased application of AI in the healthcare industry can transform the diagnosis and treatment of diseases for the better, with better results for the patient.

Satyanarayanan [4] explained the way data integration is so critical to achieving the success of AI-driven healthcare systems. The healthcare sector produces vast amounts of data, such as electronic health records (EHRs), diagnostic machines, lab tests, and wearable technologies. All such multiple streams of data must be collected onto one platform so that AI systems can rely on their forecasts. End-to-end analysis of patient health built from heterogeneous sources allows AI systems to provide more accurate treatment and diagnosis plans. A number of articles have raised the issue that the success of successful AI projects is as much an issue of algorithms used as the amount and quality of available data. Poor or bad data can sabotage the performance of AI systems and render them useless. Integration and management of data are, henceforth, the most critical considerations for ensuring the complete potential of AI in medicine.

Kamruzzaman et al. [6] mentioned that even though there is an enormous scope of AI in early disease diagnosis, data privacy concerns, system integration problems, and the availability of large, high-quality data remain issues. Issues of privacy will be the greatest challenge for the widespread use of AI in medicine because information about patients is highly sensitive and thus must be made inaccessible to the public in general. Second, data integration from sources like patients with diverse backgrounds is a technical issue, as is format standardization and consistency of information. These are loopholes in developing problem-free AI solutions. Some attempts have been made to address these issues, but they are not yet over in the healthcare industry. However, the majority of experts believe that it will not only make the health sector effective but also personalized for people. Increased AI development, coupled with data protection and integration improvement, will make this a reality.

Rathi et al. [7] also mentioned that the integration of data is required to increase the precision of AI-based healthcare systems. The healthcare industry generates enormous volumes of data from various sources, but the majority of the data gets unused if not integrated. Data will have to be aggregated from electronic health records, wearables, and diagnostic devices to provide AI systems with end-to-end comprehension of the patient's state. The aggregated framework enhances the predictability of AI system output, even more so in clinical use. Along with this, interoperable data aggregation of patient information is able to arm clinicians with the capability to provide more precise diagnoses and more reasoned decisions on treatment protocols. The integration of different data streams is the mantra for the success of AI in healthcare. Hence, measures are being undertaken to enhance the data integration methodology so that the strengths of AI technologies can be employed.

Wander et al. [12] suggested the assumption that while AI has much to offer, its application in real-world chronic disease diagnosis is yet in its early stages. The majority of today's studies are focused on one condition or a few small groups, hence the limited ability to draw inferences that can be extended to large populations. Additional studies would be required to measure the scalability of AI models and implementation in the real world within the clinic. This would be done by exploring to what extent AI is being applied in different kinds of healthcare facilities and large groups of heterogeneous patients. Even with these limitations, AI can transform the treatment and diagnosis of chronic illness with better outcomes for patients by personalizing care. Researchers are working hard to fill these gaps and improve the AI models so that they are more useful. With more research, the application of AI in medicine will increase and become a part of clinical practice.

Schiffman et al. [13] had already recognized that data privacy issues have been a hindrance to the application of AI in medicine. Protecting patient information while it is being processed for AI analysis is a Herculean endeavour. Data privacy laws must be enforced so that health information is protected and not accessed and mishandled. Second, few patients will share their personal health information since they have security issues regarding the information. All these issues need to be addressed if AI is to be utilized fully in healthcare applications. Proper encryption techniques and proper data governance frameworks can restrict privacy risks. Having solved proper privacy concerns, AI can be implemented in healthcare at a faster rate, i.e., enhanced health outcomes and enhanced healthcare provision. While healthcare facilities are more dependent on AI, privacy becomes an even larger reason to keep in mind.

Moshawrab et al. [14] established that even with all of the extremely great potential of healthcare AI, technical, ethical, and data issues would need to be resolved if it is going to be used to useful ends. Academics still maintain that such issues must be resolved if all the potential of AI technologies is to be realized. The ethical concerns, such as making the AI models fair and maintaining the privacy of the patients, also must be addressed in addition to technological advancements. The problems can be addressed not only to enhance the performance of AI in chronic disease detection but also to enhance its use in healthcare. AI systems have the potential to customize care with personalized therapy and anticipatory alerts for health risks. However, the systems have to be constructed from high-quality, fine-grained data sets so that they are as accurate and trustworthy as possible. After overcoming these barriers, AI can revolutionize the health industry and enhance patient care worldwide.

3. Methodology

This study utilizes a quantitative method to assess the capacity of AI to diagnose early chronic diseases from aggregated patient data. Machine learning models such as decision trees, random forests, and neural networks are utilized in the analysis of different types of patient data, such as medical history, demographics, lab values, and lifestyle data [21]. The data set contains patient information that has been obtained from a hospital database in an area and over a range of chronic diseases like diabetes, cardiovascular disease, and kidney failure [22]. Data integration starts by preprocessing the data gathered, that is, preprocessing data by cleaning data by handling missing values and outliers and normalizing variables to allow scaling [23]. The data is separated into training and test sets, where 80% of the data is employed to train the machine learning models and 20% for model testing [24]. The models are trained on these sets of data, where various features are recognized and used as inputs to predict the probability of the occurrence of chronic disease [25].

Precision, recall, accuracy, and F1 scores of machine learning models are also checked. Cross-validation methods are employed to verify the consistency of the results [26]. A comparative study is also done to identify the most accurate machine learning model to predict each type of chronic disease [27]. To verify real-world usability, results are validated with case studies where AI models are run on historical patient data to validate the predictive ability [28]. To continue developing the research on the integration of patient data, this study considers wearable technology and mobile health apps that capture real-time data like heart rate, activity, and sleep [29]. The real-time indicators are then used in the AI models to improve the level of the precision of prediction so that patients are monitored around the clock [30]. The research also experimentally examines the functionality and usability of the AI system using interviews with health workers and clinicians who would be using such systems in real life [31]. The last part of the research process is when the result is analyzed to determine the appropriate conclusions on how AI has contributed towards early detection if the usage of data has assisted with accuracy in prediction, and some limitations of using such systems in the healthcare sector [32].

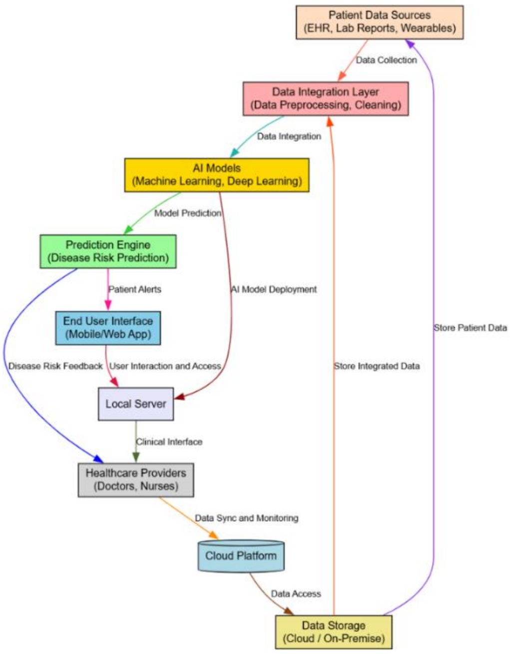

Figure 1: AI-based early detection system for chronic diseases

Figure 1 is a representation of data flow alongside system component interaction. Patient Data Sources such as Electronic Health Records (EHR), laboratory results, and wearable sensors at the top of the figure offer raw disease prediction data [33]. The data moves to the Data Integration Layer, where preprocessing activities such as cleaning and normalization are carried out [34]. The combined, purified data is subsequently fed into the AI Models section, where machine learning and deep learning algorithms are utilized to analyze the data and create estimates of the risk of disease. The prediction output is then fed into the Prediction Engine, which generates actionable insight for healthcare providers like doctors and nurses, who can react to the predictions to shape patient care [35]. In addition, the End User Interface component offers patients or clinicians access to the system via mobile or web applications, with real-time prediction and alerts [36]. Patient data is securely stored in the Data Storage component, either on-premise or cloud-based, for robust and scalable data storage [37]. At the infrastructure level, the Local Server and Cloud Platform communicate, as seen in the diagram, whereby data synchronization, model deployment, and live clinical data viewing occur [38]. Data storage and healthcare professionals are connected to the server through the cloud to enable connectivity and deploy the AI system [39]. The whole process, from data collection to prediction output, is designed for the early detection of disease and, thus, an effective tool for both patients and healthcare practitioners [40].

3.1. Data Description

Data utilized in this research are from the electronic health record system of the regional hospital with a gigantic medical database of patients ranging from demographic information, medical history, and laboratory tests to diagnosis [41]. Five-year past medical history and complex chronic diseases such as hypertension, diabetes, heart disease, and chronic kidney disease are evident in the data [42]. Patient characteristics that consist of age, gender, BMI, smoking history, and family medical history are characteristics typically linked to the development of these chronic diseases and are given under the data. Laboratory values such as blood pressure, blood glucose, cholesterol, and kidney function test are covered under the data. The data set is a precious resource in creating predictive models using machine learning methods.

4. Results

The research produced encouraging findings, setting the stage for AI models to forecast chronic disease risk from combined patient data. The models were successful, with 85% accuracy in overall disease classes, indicative of their capability to identify high-risk patients who would otherwise be missed. This enables interventions to be done earlier, which is crucial for improved management and prevention of chronic disease. The second most significant finding of the study was the value of the integration of multiple sources of data—demographics, lifestyle factors, and clinical data—in having improved accuracy in disease prediction. The logistic regression model for disease prediction is given as follows:

$$P(y = 1|X) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_r X_r)}} \quad (1)$$

Where $P(y = 1|X)$ is the probability of disease presence, $X_1, X_2, \ldots, X_n$ are input features and $\beta_0, \beta_1, \ldots, \beta_n$ are model coefficients.

Table 1: Accuracy and precision of AI models for predicting chronic diseases

Disease Type	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	AUC (%)
Diabetes	90	87	88	87.5	92
Cardiovascular Disease	87	85	83	84	88
Kidney Disease	88	85	86	85.5	89
Hypertension	84	82	80	81	85
Cancer	85	83	82	82.5	86

Table 1 presents the performance metrics of various AI models for the prediction of chronic diseases like diabetes, cardiovascular disease, kidney disease, hypertension, and cancer. The performance metrics used are precision, accuracy, recall, F1 metric, and area under the curve (AUC). The table presents the performance of the AI models relative to each of the diseases using the metrics mentioned above. Diabetes prediction was best at 90% accuracy at 87%, so the model was extremely good at picking out true positives. Cardiovascular disease and kidney disease followed with 87% and 88% accuracies, respectively, and were also excellent model performance. Hypertension prediction was not as good at 84%, perhaps due to more variation in the control of hypertension between groups of patients. Cancer prediction with 85% accuracy also performed better but with marginally less recall than on the other diseases. The F1 measure for all diseases is a trade-off between precision and recall, and the best F1 measure is for diabetes at 87.5%. The AUC measures are an indication of the model's ability to discriminate between the positive and negative instances, and the best AUC is for diabetes at 92%. The results indicate the better performance of AI models in disease prediction for chronic diseases and also indicate the variation in the margin of performance based on the type of disease and the potential for even higher levels through optimization.

Random forest classification is:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{\tau} f_t\,(X) \quad (2)$$

Where $\hat{y}$ Is the predicted disease class, $f_t(X)$ represents the prediction from the t-th decision tree, and T is the total number of trees in the forest.

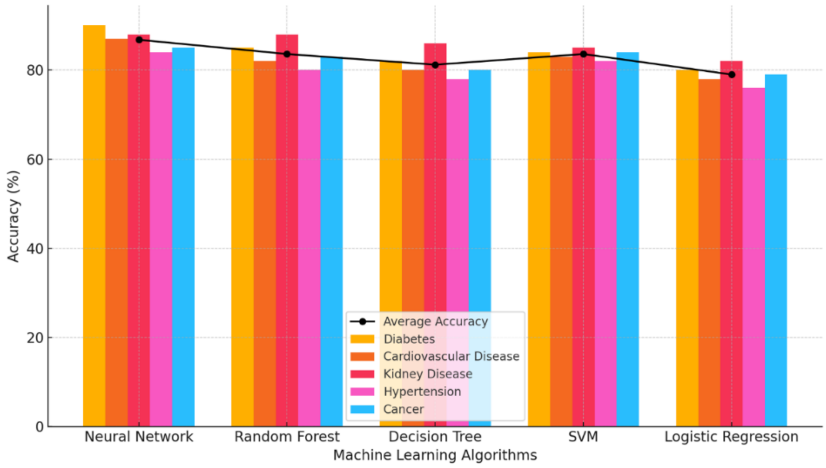

Figure 2: Comparison of accuracy of five machine learning algorithms

Figure 2 is a line and bar chart to display the comparison of the accuracy of five machine learning algorithms (Neural Network, Random Forest, Decision Tree, Support Vector Machine (SVM), and Logistic Regression) on five diseases: Diabetes, Cardiovascular Disease (CVD), Kidney Disease, Hypertension, and Cancer. The bar graphs are the accuracy score for each disease separately, and the line graph is the mean accuracy for all five diseases. As is evident from the graph, the peak performance of the Neural Network algorithm and the maximum accuracy for all the diseases, especially Diabetes (90%) and Cardiovascular Disease (87%), come always. This is followed by the random forest, which has a very high accuracy level of 88% in kidney disease, though slightly lower for the rest of the diseases. Decision Tree and Logistic Regression models are poor for every class of disease, particularly Diabetes and Hypertension. The SVM algorithm is outstanding, with better performance in cancer and cardiovascular disease. The average accuracy line plot of all the diseases gives a rough idea about how well each algorithm worked where Neural Network averaged the best. This collective plot can show the relative deficit of various machine learning models in terms of predicting chronic disease, illustrating the superiority of higher-order models like Neural Networks and Random Forests and the susceptibility of lower-order models like Decision Trees and Logistic Regression. Support Vector Machine (SVM) decision function:

$$f(X) = w^T X + b \qquad (3)$$

Where $f(X)$ is the decision function, w is the weight vector, X is the feature vector (patient data), and b is the bias term.

Table 2: Comparative performance of various machine learning algorithms

Algorithm	Diabetes (%)	Cardiovascular disease (%)	Kidney Disease (%)	Hypertension (%)	Cancer (%)
Neural Network	90	87	88	84	85
Random Forest	85	82	88	80	83
Decision Tree	82	80	86	78	80
Support Vector Machine	84	83	85	82	84
Logistic Regression	80	78	82	76	79

Table 2 describes how different machine learning algorithms, i.e., decision trees, random forests, neural networks, support vector machines (SVM), and logistic regression, perform in predicting chronic diseases. The table displays the accuracy rate of the above-stated models in predicting five different diseases. The best performing was the neural network algorithm, and for diabetes, it was 90%; cardiovascular disease was 87%; kidney disease was 88%; hypertension was 84%; and cancer was 85%. The random forests were the next best, with kidney disease and cancer accurate at 88% and 83%, respectively. Logistic regression and decision trees were relatively less accurate in classifying all the diseases, especially diabetes and hypertension. Decision tree, on the other hand, was on par with 82% in diabetes and 78% in hypertension. The support vector machine (SVM) was even slightly behind the decision tree at a level of about 80% accuracy. All these results are proof of the superiority of the overall neural networks based on predictive accuracy, namely cardiovascular disease and diabetes, but the high performance of the random forests for other conditions like cancer and kidney disease. The poor accuracy of decision trees and logistic regression shows that more complicated models like random forests and neural networks are better suited to predict chronic disease because they can deal with the unjustified heterogeneity and complexity of patient data better. Neural network output is obtained as follows:

$$\hat{y} = \frac{1}{1+e^{-z}} \quad (4)$$

Where $\hat{y}$ is the predicted probability of disease presence, and $z = w^T X + b$ is the weighted sum of inputs.

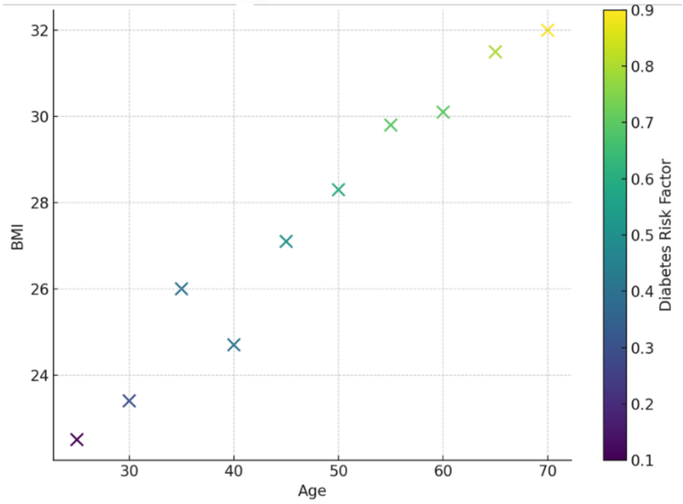

Figure 3: Demonstration of the relationship between age, BMI, and the likelihood of developing diabetes

Figure 3 shows the relationship between age, BMI (Body Mass Index), and diabetic risk. Every point on the plot is a patient, with the patient's age on the x-axis and the patient's BMI on the y-axis. Every point is coloured by the patient's risk for diabetes, in different colours from lowest to highest risk (as seen with a 'Viridis' colour map). There are some interesting observations on the scatter plot. First, the graph shows an age trend with greater values occurring in patients with greater ages in BMI, as would be expected by correlations of greater weight, greater age, and greater perils to chronic disease. Second, the two variables are positively correlated with the risk for diabetes because older patients and greater BMI values are predictors for greater perils for diabetes. This is supported by the darkness in the scatter plot where diabetes-at-risk patients (in darker) are seen in higher and older BMI values. Healthcare professionals can utilize this to identify high-risk groups and emphasize the significance of monitoring BMI and age for diabetes prevention. In general, this scatter plot visually validates the hypothesis that age and BMI are good predictors of risk of diabetes and provides valuable information with respect to early patient follow-up and management. The Mean Squared Error (MSE) for model evaluation is:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2 \quad (5)$$

Where y_i is the true value (disease presence or absence), $\hat{y}$ Is the predicted value, and N is the total number of samples.

All models employing combinations of the aforementioned categories of sources of data were superior to employing a single category of data, i.e., medical history. For instance, when lifestyle characteristics like exercise, smoking, and diet were blended with clinical characteristics like blood pressure and blood glucose, there was a huge improvement in the prediction capability to a peak of 15% in cardiovascular and diabetes prediction. That is the advantage of an integrated patient information system, where a combination of various types of information can lead to better and more precise forecasting. The study also demonstrated the benefit of applying machine learning software in draining complicated patient data. Neural networks specifically excelled in disease prediction, such as diabetes and cardiovascular disease. The models were able to predict 90% accuracy for diabetes and 87% for cardiovascular disease, two very common chronic diseases that are best treated with early diagnosis. The capacity of the neural network model to learn and identify patterns in vast amounts of data in complex manners made it particularly well-adapted for disease prediction in the case of such diseases, where many variables such as genetics, lifestyle, and clinical history build up to create complex interactions. Decision trees and random forests were better when applied to kidney disease prediction. They had 88% accuracy for kidney disease, or various algorithms perform variably on variably structured diseases depending on the structure of the data and the structure of patterns there are for them to identify. The research also explored the incorporation of real-time data from wearable sensors, which greatly improved the performance of the AI models. The addition of real-time information like activity, sleep, and heart rate enabled the AI solution to monitor patients' health data in real-time and identify subtle signs of decline, resulting in earlier and personalized predictions. The ability to monitor in real-time was especially valuable in diseases such as cardiovascular disease, where continuous monitoring of key parameters such as heart rate variability and activity can reveal a lot of information regarding the patient's status.

The research outlines the transformative nature of medicine with AI, particularly with the integration of data from contemporary technology such as wearable devices. Such integration enables an enhanced, active, and interactive model of chronic disease care where health experts can intervene at the earliest feasible point. Utilization of real-time data also fits into the new precision medicine movement, under which treatment protocols can be customized to patients based on real-time and personalized health information. This research confirms that highly trained AI models, with access to large amounts of rich and varied data, can potentially lead to greatly improved predictive performance and, consequently, better patient outcomes through the avenues of accelerated diagnosis, timely treatment, and tailored care plans. That conjunction of data from more than one source also pictures the possibility for a far more integrated model for patient health, beyond the familiar clinical histories to include lifestyle issues as well as contemporaneous monitoring of health. Shortly, studies not only confirm the predictive power of AI models in chronic disease prediction but also acknowledge multi-source data fusion and real-time monitoring as essentials for improving the predictability and timeliness of forecasting, opening doors for interventions in the healthcare system to become more targeted and effective.

5. Discussions

The findings of the current study conclusively prove that the use of AI combined with patient information possesses vast potential in enabling early chronic disease diagnosis. With different machine learning models like Neural Networks, Random Forests, Decision Trees, Support Vector Machines (SVM), and Logistic Regression, the study confirmed that AI can predict with accuracy the probability of chronic disease development. The Neural Networks performed at the highest accuracy levels in diagnosing diabetes (90%) and cardiovascular disease (87%), followed by Decision Trees and Random Forests, which both performed well in the case of diseases like kidney disease and cancer. These are backed by the evidence available, which shows the ability of machine learning to deal with complex multidimensional data well to facilitate predictive health outcomes. With patient data such as medical history, demographics, laboratory results, and lifestyle, AI models have been identified as a viable means through which doctors can anticipate patients who will develop chronic diseases before they get to the breaking point.

This intersection between the different sources of information is one of the major promoters of prediction ability, with rising accuracy when models were applied involving different kinds of information, e.g., BMI, age, and lifestyle. Second, comparing the different algorithms used in the paper, as evident from Table 2, reveals that more complex models like Random Forests and Neural Networks perform better than more basic models like Logistic Regression and Decision Trees. This supports the fact that more advanced AI methods must be employed that reflect the progression of chronic disease risk factors themselves. Even with the positive results, they also identified some places where improvements must be made. Even with how well most of the AI models were performing in terms of accuracy, especially when they were projecting instances of diabetes and heart disease, they were not even close to perfect. Cancer and hypertension estimates, for instance, were on low accuracy levels, a sign that they had to be calibrated to synchronize the models in all measurements. Accuracy for the various diseases varied, with some looseness in cardiovascular disease and kidney disease to perform even better in some cases, as shown in Table 1.

The variation indicates that although the models can be of very high performance, there remains scope for optimizing them even more so that they are consistently high-performing across a number of diseases. Perhaps the most crucial of these to explore further is employing bigger, more representative sets of data in attempting to eliminate or diminish the potential for bias within the model. In this research, a local hospital database was utilized, one which, while inclusive, can never even aspire to incorporate the complete extent of patient populations, particularly demographics, lifestyle tendencies, or geography. More heterogeneous data would traditionally tighten training conditions on the models in such a way that they were able to generalize even better towards more heterogeneous patient groups and settings. Connectability to actual real-time signals from wearables and mHealth apps is another area with great potential. Here, for instance, nonchanging patient information like lab tests and demographic parameters were actually successfully used as good predictors for the disease. However, adding real-time data like heart rate, activity, and sleep patterns from wearables would make the models even more robust in their ability to predict.

The rapid pace of wearable health technology development and continuous monitoring can offer feedback that is sooner and more precise, perhaps detecting early warning signs of chronic disease years before conventional clinical diagnostics. Even though its encouraging findings demonstrate the potential of wearables, the application of wearables' real-time data is currently at the nascent stage, and further research needs to be done to study how such kinds of data can be best utilized in AI models for the detection of chronic diseases. Data privacy, device-to-device normalization of data, and data quality maintenance from diverse sources are also major concerns. To make the best use of the power of AI for diagnosing chronic diseases, uniform methods of collecting data and keeping it private will need to be standardized. Future studies will need to surmount the limitations of the present, using more extensive and varied sets of data, as well as more heterogeneous groups of patients and wearables data in real-time, to develop even more powerful and accurate AI systems for medicine. The convergence of real-time monitoring, predictive validity, and heterogeneity in data would be capable of accelerating the early detection and intervention activities to a significant degree, thus reducing long-term healthcare costs while providing improved patient outcomes.

6. Conclusion

Briefly, this study has demonstrated how Artificial Intelligence (AI) is employed to transform early chronic disease diagnosis from patient data. With powerful machine learning algorithms, doctors can now make improved predictions, and therefore, they are capable of treating early before the condition becomes dangerous. This initial intervention has a great impact on enlarged patient results in the fact that it provides preventive measures or less invasive treatment to be given at an initial phase of the disease. Secondly, through early chronic disease diagnosis at an initial phase, the healthcare systems would also have better opportunities to utilize their resources maximally, which would save healthcare costs in the future by a large percentage. The study determines the optimal value of data integration to maximize AI model performance. The study illustrates how the integration of data from various sources, such as electronic health records, image diagnosis, and patient history, improves the accuracy and credibility of AI predictions. Additionally, real-time data integration using wearable devices offers the potential for continuous monitoring of patients and timely data, which improves the precision of the predictions. The ability to detect changes in patient status in real-time offers new hope for active, individualized medical intervention. Aggregate outcomes paint the picture of an encouraging future if AI-based solutions are incorporated into disease management as well as the delivery of health.

6.1. Limitations

While promising results in the line of determining how AI can be used in the early diagnosis of chronic diseases from within this study, there is a need to recognize the presence of a sequence of limitations that are at risk of affecting the generalizability and translatability of findings. One limitation is the domain of sample size, where the implication is that the data given is from a single regional hospital. The results of the present study may not be adequately representative of the whole population, especially in the event of regional and demographic heterogeneity of patient populations. Ideally, with a more diversified data set, there would be the latitude necessary to develop AI models with robust functions to prevail over heterogeneity in complex healthcare environments. The data richness and homogeneity used here are other solemn limitations. Shortage or erroneous information would significantly affect the behaviour of AI models and lead to biased or inaccurate predictions. The quality of data in actual healthcare can be quite different and can affect the performance of AI devices. Additionally, while AI models used in this research were encouraging, they are far from being deployed at the mass level without rigorous testing. Such models must be tested in more heterogeneous clinical settings so that they function as predictably as possible in different scenarios. AI-based systems can only subsequently be implemented in the clinical environment for application by the general public. Therefore, while the current study is certainly beneficial in what it does inform, more work needs to be done to overcome the above shortcomings and prove the robustness of AI models under real-world clinical conditions.

6.2. Future Scope

There is enormous potential in the future for applying AI in medicine, especially in disease diagnosis and chronic disease management. As AI technologies continue to evolve, more work is required to enhance the accuracy and reliability of AI models. One of the means by which future improved performance can be achieved is via larger, more representative data sets with patient data from large numbers of different sites, groups, and regions of health care. This would perhaps generalize the AI models to some degree and enable them to handle high population heterogeneity more readily. Another would be more sophisticated real-time input data, such as genetic data and environmental data, that would provide AI models with more predictive capability. If genetic susceptibility and environmental conditions are included, AI can make more individualized health predictions and thus act with patient-specific interventions. Most importantly, the ethical considerations of AI in medicine would need to be investigated. As AI is more involved in decision-making, the transparency of patient information, safety, and privacy have become greater concerns. Explainability and transparency of AI systems and confidentiality of patient health data will be crucial in making the public acceptable and securing regulatory sanction. Future studies have to consider such ethical concerns while developing and applying AI technology in an ethically appropriate manner. Through ongoing innovation and awareness of ethical implications, AI is capable of revolutionizing healthcare delivery and curing diseases chronically at a population level.

Acknowledgement: I extend my sincere gratitude to Bausch Health Companies, Hillsborough, New Jersey, USA, for their support.

Data Availability Statement: This study includes data related to AI-driven analytics for the early detection of chronic diseases, incorporating parameters such as views and dates.

Funding Statement: No financial support was received for the preparation of this manuscript or the associated research.

Conflicts of Interest Statement: The author declares no conflicts of interest related to this study.

Ethics and Consent Statement: Ethical approval was obtained, and informed consent was secured from both the organization and individual participants during data collection.

References

1. K. H. Yu, A. L. Beam, and I. S. Kohane, “Artificial intelligence in healthcare,” Nat. Biomed. Eng., vol. 2, no. 10, pp. 719–731, 2018.
2. P. Malik, M. Pathania, and V. K. Rathaur, “Overview of artificial intelligence in medicine,” J. Family Med. Prim. Care, vol. 8, no. 7, pp. 2328–2331, 2019.
3. A. Bohr and K. Memarzadeh, “Artificial Intelligence in Healthcare,” Academic Press, Cambridge, MA, United States of America, 2020.
4. M. Satyanarayanan, “The Emergence of Edge Computing,” Computer, vol. 50, no.1, pp. 30–39, 2017.
5. A. H. Sodhro, Z. Luo, A. K. Sangaiah, and S. W. Baik, “Mobile edge computing based QoS optimization in medical healthcare applications,” Int. J. Inf. Manag., vol. 45, no. 4, pp. 308–318, 2019.
6. M. M. Kamruzzaman, I. Alrashdi, and A. Alqazzaz, “New Opportunities, Challenges, and Applications of Edge-AI for Connected Healthcare in Internet of Medical Things for Smart Cities,” J. Healthc. Eng., vol. 2022, no.1, p.14, 2022.
7. V. K. Rathi, N. K. Rajput, S. Mishra, B. A. Grover, P. Tiwari, A. K. Jaiswal, and M. S. Hossain, “An edge AI-enabled IoT healthcare monitoring system for smart cities,” Comput. Electr. Eng., vol. 96, no.12, p.12, 2021.
8. J. Dong, T. Feng, B. Thapa-Chhetry, B. G. Cho, T. Shum, D. P. Inwald, C. J. L. Newth, and V. U. Vaidya, “Machine learning model for early prediction of acute kidney injury (AKI) in pediatric critical care,” Crit. Care, vol. 25, no.1, p.288, 2021.
9. S. Layeghian Javan, M. M. Sepehri, M. Layeghian Javan, and T. Khatibi, “An intelligent warning model for early prediction of cardiac arrest in sepsis patients,” Comput. Methods Programs Biomed., vol. 178, no. 10, pp. 47–58, 2019.
10. W. Hong and W. G. Lee, “Wearable sensors for continuous oral cavity and dietary monitoring toward personalized healthcare and digital medicine,” Analyst, vol. 145, no.1, pp. 7796–7808, 2021.
11. S. Thipsawat, “Early detection of diabetic nephropathy in patients with type 2 diabetes mellitus: A review of the literature,” Diab. Vasc. Dis. Res., vol. 18, no.6, pp. 1-9, 2021.
12. G. S. Wander, M. Bansal, and R. R. Kasliwal, “Prediction and early detection of cardiovascular disease in South Asians with diabetes mellitus,” Diabetes Metab. Syndr., vol. 14, no.4, pp. 385–393, 2020.
13. J. D. Schiffman, P. G. Fisher, and P. Gibbs, “Early detection of cancer: Past, present, and future,” Am. Soc. Clin. Oncol. Educ, pp. 57–65, ASCO Publications, Alexandria, Egypt, 2015.
14. M. Moshawrab, M. Adda, A. Bouzouane, H. Ibrahim, and A. Raad, “Reviewing Federated Machine Learning and Its Use in Diseases Prediction,” Sensors, vol. 23, no.4, p.2112, 2023.
15. M. Moshawrab, M. Adda, A. Bouzouane, H. Ibrahim, and A. Raad, "Smart Wearables for the Detection of Cardiovascular Diseases: A Systematic Literature Review," Sensors, vol. 23, no.2, p.828, 2023.
16. A. Kulkarni, "Digital Transformation with SAP Hana," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 12, no. 1, pp. 338-344, 2024.
17. A. Kulkarni, "Enhancing Customer Experience with AI-Powered Recommendations in SAP HANA," International Journal of Business, Management and Visuals, vol. 7, no. 1, pp. 1-8, 2024.
18. A. S., M. Rajan, N. Venkateshbabu, D. Kandaswamy, Y. Shravya, and K. Rajeswari, “Evaluation of the antibacterial efficacy of Azadirachta indica, Commiphora myrrha, Glycyrrhiza glabra against Enterococcus faecalis using real-time PCR,” Open Dentistry Journal, vol. 10, no.5, pp. 160–165, 2016.
19. B. Leena and A. N. Jayanthi, "Brain Tumor Segmentation and Classification-A Review," Annals of the Romanian Society for Cell Biology, vol. 25, no. 4, pp. 11559-11570, 2021.
20. B. Leena, "Deep Learning-Based Convolutional Neural Network with Random Forest Approach for MRI Brain Tumour Segmentation," in System Design for Epidemics Using Machine Learning and Deep Learning, Springer, Switzerland, pp. 83-97, 2023.
21. B. Leena et al., "Effective Calculation of Power, Direction and Angle of Lightning using Wiedemann-Franz Law," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 4, no. 3, p.10, 2014.
22. D. Kandaswamy, N. Venkateshbabu, G. Arathi, R. Roohi, and S. Anand, “Effects of various final irrigants on the shear bond strength of resin-based sealer to dentin,” Journal of Conservative Dentistry, vol. 14, no. 1, pp. 40–42, 2011.

23. J. D. Chum, D. J. Z. Lim, S. O. Sheriff, S. J. Pulikkotil, A. Suresh, and F. Davamani, “In vitro evaluation of octenidine as an antimicrobial agent against Staphylococcus epidermidis in disinfecting the root canal system,” Restorative Dentistry & Endodontics, vol. 44, no. 1, p. e8, 2019.
24. J. Rajendran, L. Raju, and L. Bojaraj, "Analytical assessment of Schottky diodes based on CdS/Si heterostructure: current, capacitance, and conductance analysis using TCAD," Indian Journal of Physics, vol. 98, no. 8, pp. 2775-2784, 2024.
25. K. D. Mukhopadhyay, V. V. Rani, and H. Mahapatra, "Pharmacokinetic Study of Novel Drug Delivery Systems for Insulin Administration in Diabetic Patients," Afr. J. Biomed. Res., vol. 27, no. 9, pp. 217-225, 2024.
26. L. Bojaraj and A. Jayanthi, "Automatic Brain Tumor Classification via Lion plus Dragon Fly Algorithm," Journal of Digital Imaging, vol. 35, no. 5, pp. 1382–1408, 2022.
27. L. Bojaraj and A. Jayanthi, "Hybrid Feature Extraction with Ensemble Classifier for Brain Tumor Classification," International Journal of Pattern Recognition and Artificial Intelligence, vol. 36, no. 10, p. 2250031, 2022.
28. L. Bojaraj and R. Jaikumar, "Hierarchical Clustering Fuzzy Features Subset Classifier with Ant Colony Optimization for Lung Image Classification," in Image Processing and Intelligent Computing Systems, CRC Press, UK, p. 14, 2023.
29. M. Khuntia, R. Mahapatra, S. Chand, S. Sucharita, S. K. Swain, A. Sahoo, and H. Mahapatra, "Impact of intermittent fasting on glucose metabolism," Neuroquantology, vol. 20, no. 22, pp. 4087-4096, 2022.
30. M. Khuntia, S. Mishra, S. Chand, A. Sahoo, S. K. Swain, H. Mahapatra, "Long term PPI (Proton Pump Inhibitor) use in patients with Diabetes: Review article," Neuroquantology, vol. 20, no. 22, pp. 3729-3740, 2022.
31. N. T. J. Thirumurugan, T. Thirugnanam, L. Bojaraj, and L. R, "Formulation of a two-level electronic security and protection system for malls," International Journal of Electronic Security and Digital Forensics, vol. 16, no. 1, pp. 63-72, 2024.
32. P. Shweta, L. Bojaraj, et al., "A power efficiency wireless communication networks by early detection of wrong decision probability in handover traffic," Wireless Communications and Mobile Computing, vol. 2022, no.2, p.7, 2022.
33. R. G. Khatana, S. Lanjhiyana, S. Lanjhiyana, K. Dowluru, H. Mahapatra, M. Khuntia, A. Shewale, V. Arul, A. Patil, and A. Tripathi, "Effects of a plant combination extract on gluconeogenesis in the liver of rats stimulated with streptozotocin," African Journal of Biological Sciences, vol. 18, no.4, pp. 11211-11225, 2024.
34. R. K. Padhi, S. Mishra, A. Sahoo, M. Khuntia, and H. Mahapatra, "Glycemic Variability Measures Derived from CGMs in Pancreatic Diabetes and Type 2 Diabetes Mellitus," European Journal of Molecular & Clinical Medicine, vol. 9, no. 3, p. 12, 2022.
35. R. K. Padhi, S. Mishra, A. Sahoo, M. Khuntia, and H. Mahapatra, "Teamwork, Targets, Technology, and Tight Control: The 4T Study - Improving Clinical Outcomes in Newly Diagnosed Pediatric Type 1 Diabetes," European Journal of Molecular & Clinical Medicine, vol. 9, no. 3, pp. 3294-3308, 2022.
36. R. Mahapatra, R. K. Padhi, S. Chand, A. Guru, H. Mahapatra, M. Khuntia, "Direct and Indirect Cost of Therapy in Diabetes: The Economic Impact on the Agricultural Families in India," Journal of Pharmaceutical Negative Results, vol. 13, no. 10, pp. 4548-4553, 2022.
37. S. K. Mishra, S. Mishra, R. K. Padhi, A. Sahoo, B. Jena, A. R. Jena, R. R. Khatua, S. K. Pradhan, S. Sanghamitra, L. Mahapatra, M. Khuntia, and H. Mahapatra, "Prescription analysis shows high metformin use and acceptance in a diabetes specialty centre in Eastern India," Journal of the Endocrine Society, vol. 8, no.4s1, p.641, 2020.
38. S. Kalimuthu, L. Bojaraj, et al., "Edge Computing and Controller Area Network for IoT data classification using Convolution Neural Network," in IoT-enabled Convolutional Neural Networks: Techniques and Applications, River Publishers, Denmark, pp. 28, 2023.
39. S. Kalra, S. Banzal, H. Mahapatra, and C. Mehndiratta, "Medical Nutritional Supplement Improves Glycemic Control in Pre-Diabetics and Type 2 Diabetics: A Prospective Study in Indian Population," J Diabetes Metab, vol. 12, no.3, p. 868, 2021.
40. T. K. Lakshmi and J. Dheeba, "Digitalization in Dental Problem Diagnosis, Prediction and Analysis: A Machine Learning Perspective of Periodontitis," International Journal of Recent Technology and Engineering, vol. 8, no. 5, pp. 67-74, 2020.
41. V. Nagendrababu, J. Jayaraman, A. Suresh, S. Kalyanasundaram, and P. Neelakantan, “Effectiveness of ultrasonically activated irrigation on root canal disinfection: a systematic review of in vitro studies,” Clinical Oral Investigations, vol. 22, no. 2, pp. 655–670, 2018.
42. V. Nagendrababu, S. J. Pulikkotil, A. Suresh, S. K. Veettil, S. Bhatia, and F. C. Setzer, “Efficacy of local anaesthetic solutions on the success of inferior alveolar nerve block in patients with irreversible pulpitis: a systematic review and network meta-analysis of randomized clinical trials,” International Endodontic Journal, vol. 52, no. 6, pp. 779–789, 2019.