Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Artificial Intelligence in Multi-Disease Medical Diagnostics: An Integrative Approach

Nigar Sultana[1], Shariar Islam Saimon[2], Intiser Islam[2], Shake Ibna Abir[3]✉, Md Sanjit Hossain[4], Sarder Abdulla Al Shiam[5], Nazrul Islam Khan[6]

[1]Department of Finance, University of New Haven, CT, USA
[2]Department of Computer Science, School of Engineering, University of Bridgeport, USA
[3]Instructor of Mathematics, Department of Mathematics and Statistics, Arkansas State University, Arkansas, USA
[4]MBA in Information Technology Management, University of the Potomac, Washington DC, USA
[5]Department of Management–Business Analytics, St Francis College, New York, USA
[6]Department of Mathematics & Statistics, Stephen F. Austin State University, Texas, USA
Corresponding Author: Shake Ibna Abir, **E-mail**: sabir@astate.edu

| ABSTRACT

With advanced algorithms, artificial intelligence (AI) has revolutionized the medical diagnostic field where diseases can be predicted simultaneously. The integrative nature of this approach is novel because it can better encompass the complexity of comorbid conditions that are so common in patients; thus, addressing them in a more holistic diagnostic tone that is lacking in previous works. In this study, the investigation of the usage of AI models for simultaneously diagnosing diseases like diabetes, cardiovascular conditions, and neurological disorders is done. Therefore, based on AI techniques i.e. artificial neural networks (ANNs) and ensemble learning methods, a multi-disease diagnostic framework was developed to achieve this. A variety of features, related to each condition, were captured from multi-modal datasets including imaging, laboratory test results, and patient histories. The system was developed to manage the big flow of aggregated data and offer detailed diagnostic views of many diseases. Sensitivity, specificity, and overall diagnostic accuracy were used to evaluate the framework's performance. The results showed that the AI framework has high diagnostic accuracy for all targeted conditions an overall sensitivity of 93% and a specificity of 91%. Importantly, the combination of multi-modal data proved to substantially improve the system's ability to identify and distinguish comorbid conditions. It makes the importance of using various data sources to benefit from the reliability and comprehensiveness of AI diagnostics obvious. Overall, AI-driven multi-disease diagnostic systems provide great promise for the role of delivering potentially transformative clinical healthcare workflow improvements, reducing errors, and improving patient outcomes. These frameworks will need to be scaled and tested in various healthcare settings and also across more varied diseases to help make medical diagnosis more available and effective.

| KEYWORDS

AI models, performance comparison, accuracy, sensitivity, specificity, multi-modal data, disease-specific accuracy, ROC curve, Gradient Boosting, model evaluation

| ARTICLE INFORMATION

ACCEPTED: 19 January 2025 **PUBLISHED:** 09 February 2025 **DOI:** 10.32996/jcsts.2025.7.1.12

1. Introduction

1.1 Background & Importance

Traditional methodologies of medical diagnostics have relied so far on physical examinations, imaging techniques and laboratory tests. Nevertheless, these traditional diagnostic tools also tend to disregard the challenge posed by comorbid conditions — that is, multiple diseases occurring in the same patient — making clinical decision making difficult (Ee et al., 2020). Diagnostic frameworks that are currently used often work in the paradigm of a single disease in which each disease is considered independently while ignoring the underlying symptoms and propensities common in diseases, like diabetes, cardiovascular diseases or neurological disorders. Due to the limitation described above, delays occur in the diagnoses, increase in the rates of misdiagnosis and fragmentation of the treatment approaches eventually affecting the patient outcomes (Low et al., 2020).

Diabetes, cardiovascular diseases and neurological disorders are among the most challenging chronic diseases prevalent across the world and leading causes of morbidity and mortality (Karunarathna et al., 2024). Moreover, growing burden of these conditions is exacerbated by lifestyle, aging populations, environmental determinants, and thus effective and scalable diagnostic solutions are needed. In addition, existing studies have demonstrated that a comorbid chronic disease is adding significant influence on disease severity, making treatments more complex and complicating the management of patients (Liu et al., 2020). This is striking because diagnosis becomes a more complicated problem because of the need to understand how players can be affected by multiple diseases, rather than just isolated diseases.

A large question to answer in modern medicine is to overcome the largely overlapping of presentations and risk factors between various diseases that typically result in diagnostic issues. For example, both hypertension and diabetes commonly overlap, share common metabolic and inflammatory pathways, but may be diagnosed and managed separately (Stein et al., 2022). Similar to this, neurological diseases such as Alzheimer's disease and Parkinson's disease display similar cognitive and motor impairments that make their separate diagnosis difficult (Eaton et al., 2023). To depict such complexities, the traditional approach to medical diagnosis to discrete disease model is not equipped with the integrative capacity. Hence, there is a necessity for an AI based infrastructure capable of using multimodal data for integrating a patient's health status in a whole.

1.2 Role of Artificial Intelligence (AI) in Medical Diagnostics

Most recently, artificial intelligence is an emerging and powerful force in medical diagnostics through its applications in data processing, pattern recognition, automation, among others. AI driven models have proved to be a better alternative in getting results from large scale medical datasets running into terabytes in terms of size, identifying hidden correlations between various parameters and assess predictive capabilities with much accuracy (Ahmed et al, 2020). Unlike the traditional statistical methods, the AI algorithms can keep learning from the new data and also, can adapt themselves over the passage of time thus enhancing the diagnostic precision over time. The cancer, diabetes, and cardiovascular disorders have been able to benefit from these capabilities for early disease detection and management (Kaur et al., 2020).

One of the main benefits of AI is that it can work on the multi modal medical data like MRI, CT scans, laboratory data, patient history, etc. The advantage of being able to integrate various multidimensional brain imaging modalities into an integrated, task relevant, and interpretable way further allows AI systems to detect subtle disease markers that are not necessarily evident by the analysis of a single imaging modality. Especially, AI has been shown to be effective in the diagnosis of specific diseases where deep learning models attain high accuracy of diabetic retinopathy, myocardial infarction and neurodegenerative diseases detection. However, although the success of AI in the diagnosis of a single disease is well documented, the incorporation of AI in multi disease diagnosis is yet an evolving area of research (Roobini et al, 2024). In that case, AI models that can diagnose multiple conditions are developed, which can be an opportunity to change medical practice to be more efficient and accurate.

1.3 Problem Statement & Research Gap

Although AI diagnostics have progressed, most of the existing models are to predict and assess single illnesses only. The use of AI is restricted in actual medical practice, by the sole focus on the single disease, while patients frequently possess multiple comorbid conditions. However, the existing diagnostic systems do not consider the associations between diseases and result in discontinuous delivery of healthcare (Alsaleh et al., 2023). The lack of Artificial Intelligence research into comorbidity detection is substantial, and not taking into account such conditions when developing and applying treatment plans is ill advised and could prove very costly.

Secondly, the existence of AI models on a lack of explainability especially when deployed for multi disease diagnosis is another crucial challenge. However, since the decision-making process in the AI algorithms is often equivalent to a black box the clinical interpretation of rationale of an algorithm produced specific diagnoses is not easy to interpret (Albahri et al., 2023). However, this lack of transparency deprives AI of its pathway into mainstream medical practice by denying the healthcare professionals a way to trust or trust on AI generated outputs since they have no way of understanding what the basis of the outputs were.

Additionally, there has been insufficient research done in the area of AI frameworks intended explicitly for multi disease diagnosis. While some studies have looked into whether and how AI can be applied to specific diseases, there is yet a demand for a complete model, capable of diagnosing several diseases at the same time. Overcoming the barrier of integrating explainable AI with trustworthy machine learning techniques in its integration is key to making sure the AI diagnostics are accurate and interpretable for the use in the clinical world.

1.4 Objectives of the Study

This study aims to address these challenges by developing an AI-based diagnostic framework capable of simultaneously diagnosing multiple diseases. The research objectives are as follows:

1. **Develop an AI-based diagnostic framework using Artificial Neural Networks (ANNs) and Ensemble Learning:** In this study, a multi–disease diagnostic model will be designed and implemented using the advantages of ANNs and ensemble learning techniques. The patterns indicative of diabetes, cardiovascular, and neurological diseases will be trained on diverse datasets.
2. **Integrate multi-modal datasets to improve diagnostic accuracy:** The diagnostic system will be built based on imaging data, laboratory test results and patient history. The framework drives diverse medical data sources in order to generate more accurate and error compensating diagnostics relative to traditional single modality analysis.
3. **Assess the framework's performance in diagnosing diabetes, cardiovascular diseases, and neurological disorders:** A key to determination will be to evaluate the AI framework on key performance metrics such as sensitivity, specificity and overall diagnostic accuracy using the study. Then, a comparative analysis is conducted in order to see how well the proposed AI system performs in comparison to existing single disease diagnostic models.

The results achieved in this research would help advance AI in healthcare through exhibiting the manner by which integrative AI models can approach and improve precision of diagnosis and hence patient outcomes. Proposals in such a framework can lead to reduction of diagnostic errors, improvement of treatment plans, and reinforcement of health system efficiency by providing richer picture of a patient's health.

2. Literature Review
2.1 Overview of AI-driven Diagnostic Models

Medical diagnostics have changed drastically after the arrival artificial intelligence (AI) in terms of automated analysis of complex data and increased accuracy of disease detection. ML algorithms-based systems can analyze much larger data sets in a fraction of time a human clinician can to finding patterns that are not visible to the human eye. According to Sayem et al. (2023), there is an overall survey about the adoption of AI tools in healthcare systems around the world. As the need for the standard analysis of Multi-dimensional data coming from numerous sources such as medical imaging, patient records and lab reports, AI systems have found its utility in clinical practice. The growing need of more efficient, accurate and scalable healthcare solution leads to shifting from traditional methods to AI-based diagnostics.

2.2 Applications in Diabetes, Cardiovascular Diseases, and Neurology

There are some conditions, where AI has been found successful, especially in chronic disease cases such as diabetes, cardiovascular diseases and neurological disorders. In diabetes management, AI algorithms can effectively forecast the onset of complications such as nephropathy or retinopathy from the retinal scan, kidney function tests, and blood glucose levels (Mathur et al., 2020). Similarly, cardiovascular diseases, which are usually detected based on imaging data like ECG, CT.Scan, MRI etc. have also been aided by applying AI to identify small abnormalities related to heart disease. AI models are helping to diagnose the early stages of Alzheimer's disease, Parkinson's disease, other neurodegenerative diseases using MRI and PET scans, for early intervention (Sayem et al, 2023). The use of AI systems for automating the analysis of medical images and patient data reduces human error and improvement of diagnosis accuracy in these difficult areas.

2.3 Comparison of AI Models with Traditional Diagnostic Methods

Traditional diagnostic methods and AI driven models are compared to show the great potential for AI to improve diagnostic efficiency. Historically, human expertise is relied upon to interpret diagnostic tests but such methodology is time consuming and subjective. According to Rajula et al. (2020), the use of conventional statistical methods is slower and less reliable when it comes to disease diagnosis, but models especially that of AI machine learning like Random Forests and Support Vector Machines (SVM) outperform the traditional ones. Instead of limited set(s) of diagnostic markers, traditional methods can handle a more limited set of inputs, for example, a physician can provide data like genetics, clinical history and images to an AI model, while in traditional methods only say the colony forming unit or a particular cut off can be provided. AI models have attributes like increased accuracy, objectivity and speed, which are very important features in the healthcare industry where things happen fast and with high stake implications required.

2.4 Machine Learning Algorithms for Disease Prediction

2.4.1 Overview of Artificial Neural Networks (ANNs) and Their Role in Medical Diagnosis

Artificial Neural Networks (ANNs) have been contributed to be basic tools in AI based medical diagnosis because of their capacity to track complicated, nonlinear relationships between input information and intended consequence. ANNS simulate the structure of the human brain, and as human brains can learn quickly and easily from huge amounts of data, this improves the prediction of diseases as described by Kufel et al. (2023). ANNs have been used in medical fields to predict a massive number of diseases such as cancers, cardiovascular disease and diabetes from clinical data and medical images. ANNS provide great flexibility so that complex relationship patterns in patient data may be modeled, which is vital when symptoms have overlap or are insidious making diagnosis difficult by traditional methods.

2.4.2 Ensemble Learning Methods: Random Forest, Gradient Boosting, and Their Comparative Efficiency

Ensemble learning methods are a well-known procedure, where clinicians use techniques like Random Forests or Gradient Boosting, in several base models to improve the performance of models and infer better final prediction. The random forest is established by Sahin (2020) as the most effective method in managing high dimensional data and hence can be used in disease diagnosis where there are too many interrelated factors involved. Moreover, Gradient Boosting algorithms are known for reducing the errors in predicted values through iterative learning processes (Bentéjac et al., 2021). When being applied to medical datasets, these ensemble methods have shown to outperform the decision trees or SVM individually in terms of its accuracy or robustness. This helps in the prevention of overfitting and enables the generalizability of the model since the model is not heavily relying on a specific data representation but is using an aggregated representation of multiple data.

2.5 Multi-Disease Diagnosis & Comorbidity Challenges

2.5.1 Complexity of Diagnosing Multiple Diseases Simultaneously

The concomitant diagnosis of multiple comorbid diseases is one of the main challenges in healthcare nowadays. Increasingly, people will suffer from multiple comorbidities as populations age and chronic diseases spread. According to Hassaine et al. (2020), the diagnosis and treatment of a patient with multimorbidities where a patient has multiple disorders at a time becomes complicated as symptoms are often overlapping among disorders. Simultaneously processing such data using AI provides a promising solution because AI can learn patterns from multiple diseases at the same time. Even so, the variety of multimorbidity requires complete ones with the ability to distinguish between all sorts of diseases which share common symptoms and risk factors, this is where the intersection between AI and machine learning and deep learning techniques becomes necessary.

2.5.2 Existing Approaches and Their Limitations

Existing approach for multi-disease diagnosis is based on sequential disease specific models. While they are effective for single disease diagnosis, they fail in the presence of complicated interrelations among multiple diseases in a single patient. According to Wu et al. (2021), these limitations can be overcome with an advanced form of machine learning called multi–label classification that relates to the prediction of multiple outcomes simultaneously. AI models can adapt more to the fact that chronic diseases are often accompanied by comorbidities by using multi label classifiers, which assign more than one label or disease to a given data point. Nevertheless, more research is needed in order to refine these models and to confirm their efficacy for use across larger and more various patient populations.

2.5.3 Benefits of AI-Driven Integrative Approaches

Advantages of using AI in the integrative multi disease diagnostic systems exist over traditional methods. With data from a variety of sources – imaging, laboratory tests, and history of the patient – these AI driven systems will be equipped with a better understanding of the overall health of the patient. As reported by Allami and Yousif (2023), these integrative AI approaches not only enhance diagnostic accuracy but also allow for more personalized treatment treatments. By keeping in mind, the interrelation between diseases, AI can give clinicians actionable insights which is extremely hard for human diagnostic methods to find out. As such, Kasula (2024) notes that integrative approaches to AI are highly useful in infectious disease diagnosis where identifying comorbidities on time is essential for proper treatment.

2.6 Multi-Modal Data in AI Diagnostics
2.6.1 Role of Various Data Sources

There is an essential role in the process of integration of different types of data to improve the diagnostics made by AI. Referring to Xu et al. (2024), it is worthwhile to consider the application of multi-modal data which involves MRI/CT scans, lab tests, and patient history to increase AI degree of complexity. Imaging is structural and functional view of the body while the laboratory testing and biomarkers deliver details about biochemical process of a disease. By incorporating these sources, AI models get an integrated view of an individual and early indications of the signs of diseases that may be undetectable in one of the sources.

2.6.2 Importance of Integrating Different Types of Medical Data

Combination of various kind of medical data is necessary to ensure that the created AI system fully incorporates human health parameters. Yang et al. (2022) state that Multi-center and Multi-modal data fusion is vital for medical AI since it eliminates the risk of learning from a bias source or patients and considering data from different centers. Incorporation of data from various population and health care settings can make the AI models more generalizable so as to perform optimally in all kind of clinical situations.

2.7 Evaluation Metrics in AI-Based Diagnostics
2.7.1 Sensitivity, Specificity, and Accuracy in Medical AI Models

AI Models used for medical diagnostic needs reliable and well measured metrics to evaluate the performance of the models. The three main metrics in judging AI diagnostic systems are sensitivity, specificity, and accuracy. The ability of the model to correctly identify diseased individuals is termed as sensitivity and the ability of the model to correctly identify healthy individuals is termed as specificity. An overall measure of how correctly the model should work is accuracy. Park et al. (2023) stress how critical these metrics are to treat when evaluating AI models because they will directly affect the clinical decision making and the patient outcomes. In healthcare of all places, these metrics are particularly critical, especially false positives and false negatives which can have major consequences on patient treatment.

2.7.2 Trade-offs Between False Positives and False Negatives

To solve the core issue of trading off between false positives and false negatives in AI based diagnostics is one of the central challenges. Thus, false positives may cause unnecessary treatment and false negatives may bring missed diagnosis and delayed treatment. Adjusting the threshold of decision-making in AI models is something that can help in balancing these risks as discussed by Albahri et al. (2023). In healthcare, it is important to have AI systems that can be fine-tuned by the healthcare professionals who must determine how receptive the system is to change its thresholds based on the clinical context.

Table 1: Summary of literature review discussion tpoics on AI and ML

Author(s)	Summary
Sayem et al. (2023)	Adoption of AI in healthcare and global practices
Mathur et al. (2020)	AI's role in managing diabetes through imaging and tests
Sayem et al. (2023)	AI's success in early detection of cardiovascular diseases
Rajula et al. (2020)	AI vs traditional diagnostic methods in disease detection
Allami and Yousif (2023)	AI's role in enhancing diagnostic accuracy and personalized treatments
Kasula (2024)	AI's role in precision medicine and infectious diseases
Xu et al. (2024)	Fusion of multi-modal data for improved diagnostic accuracy
Yang et al. (2022)	Importance of multi-center and multi-modal data for AI systems

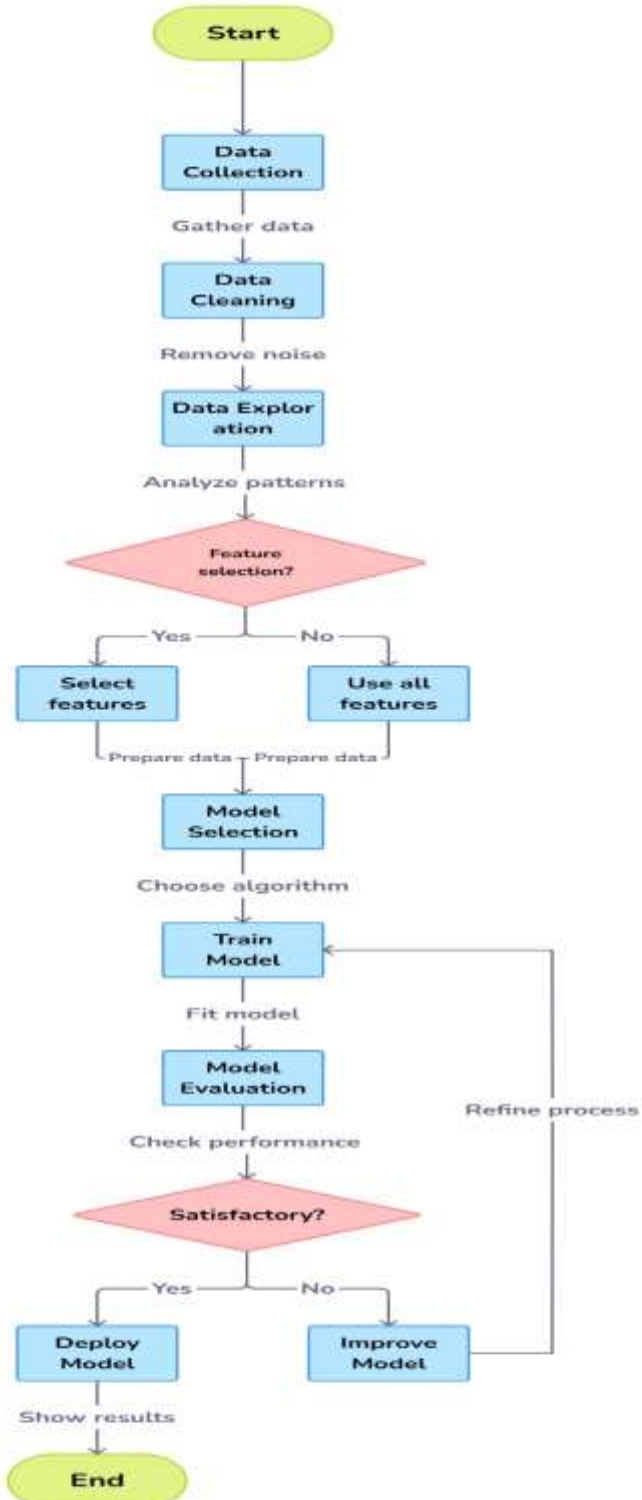

Figure 1: Flowchart outlining ML processing data across different datasets

3. Methodology

In this section, the process for creating the AI based diagnostic framework is described to be used for detecting diabetes, cardiovascular diseases, and neurological disorder. Data are collected, preprocessed, and used in model selection and model evaluation steps in the methodology. Artificial Neural Networks (ANNs) and Ensemble Learning namely, Random Forest and Gradient Boosting are used to construct the models. Each step of the process is explained in the following details that involve equations along with visualizations.

3.1 Research Design

The research design is aimed at constructing an AI diagnostic framework based on machine learning techniques for multiple disease diagnosis at once. Using multi modal data of the patients including medical history, lab test results, and medical imaging data like MRI scan and CT scans, the system was developed in the thesis. A diagnostic framework was developed with the ability to comprehensively process large and diverse datasets and the ability of the final model to discriminate various disease states.

This framework was able to carry out two key machine learning models, Artificial Neural Networks (ANNs) and Ensemble Learning methods. ANNs are famous for their capability to learn the complex nonlinear relationship in data, which constitutes as the most important aspect of ANNs for predicting disease outcomes. Hence, Ensemble Learning is used to mitigate overfitting that is prone to happen when one model is used. In particular, the Ensemble Learning methods selected were Random Forest and Gradient Boosting, which are known to have been successful in medical diagnostics.

3.2 Dataset Selection & Preprocessing

Concretely, this study consider the datasets that have been collected from publicly accessible medical repositories (e.g., CT and MRI imaging datasets as well as patient history data such as lab tests and demographics). The selection criteria included:

Table 2: Data selection and pre-processing criteria

Dataset Criteria	Details
Diversity of Patient Demographics	A wide range of ages, genders, and health conditions included.
Multi-modal Data	Data includes a combination of imaging, lab results, and clinical records.
Open-source Availability	Datasets were open-access for transparency and reproducibility.

Once the datasets were obtained, the preprocessing steps were performed as follows:

Noise Reduction:
The EEG and imaging data often have noise due to inherent movement artifact or technical limitation so a Butterworth band pass filter was applied to remove unwanted frequencies. The definition of the transfer function of this filter is:

$$H(s) = \frac{1}{(s/\omega_c)^{2n} + 1} \qquad (1)$$

Where:

$H(s)$ is the filter transfer function,

ω_c is the cutoff frequency,

n is the filter order.

Normalization:
Min max normalization was used to scale all numerical features to the range [0, 1] to ensure that all data sources were consistent. The formula of normalization is as follows:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (2)$$

Where:

X' is the normalized value,

X_{min} and X_{max} are the minimum and maximum values in the dataset.

Handling Missing Data:
In instances where all the data entries were not there, the study used K-Nearest Neighbors (KNN) imputation which basically means that it filled the missing value with the mean of the nearest 5 neighbors. The imputation formula is:

$$X_{missing} = \frac{1}{k}\sum_{i=1}^{k} X_i \qquad (3)$$

Where:

$X_{missing}$ is the imputed value,

k is the number of nearest neighbors used for imputation.

Figure 2: heatmap showing the feature correlations of the data processing

The above heatmap shows correlation of key features included in data preprocessing stage, i.e, blood glucose levels, ECG readings, BMI, and insulin resistance. This tool aids in finding out dependencies within features and gives an insight of how they work. An example is a BMI and insulin resistance correlation as it gives a strong association on diabetes prediction. The used heatmap during feature selection to select features without redundancy, and only select the features which are more informative to finally improve the model efficiency and diagnostic accuracy.

3.3 Feature Selection & Engineering

The performance of the models was improved by feature selection and engineering. Therefore, the features have been chosen based on their relevance to the target diseases that were diagnosed. These features were guided by the study's domain-special knowledge per disease:

Diabetes: Blood glucose levels, insulin resistance, and body mass index (BMI).

Cardiovascular Diseases: Blood pressure, cholesterol levels, and ECG readings.

Neurological Disorders: EEG signal features and brain volume changes from MRI scans.

Dimensionality Reduction using Principal Component Analysis (PCA):
To reduce the number of features while preserving 95% of the variance in the data, the study used PCA. The PCA formula is:

$$Z = XW. \quad (4)$$

Where:

Z is the transformed feature matrix,

X is the original data,

W is the matrix of eigenvectors corresponding to the largest eigenvalues of the covariance matrix.

3.4 AI Model Development

The models used in this study include:

Artificial Neural Networks (ANNs): The network was a feedforward network with two hidden layers that was optimized using Adam optimizer.

Ensemble Learning (Random Forest and Gradient Boosting): The reason I chose these models is that they both are robust in their ability to handle complex, high-dimensional data. 100 Decision trees were used in Random Forest whereas Gradient Boosting used 200 weak classifiers with a value of 0.1 for learning rate.

Mathematically, for Random Forest, the prediction is:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t(x) \qquad (5)$$

Where:

T is the number of trees,

$f_t(x)$ is the prediction of tree t for input x.

For Gradient Boosting, the prediction is:

$$\hat{y} = \sum_{m=1}^{M} \eta\, f_m(x) \qquad (6)$$

Where:

M is the number of weak learners,

η is the learning rate,

$f_m(x)$ is the m-th weak learner's prediction.

3.5 Evaluation Metrics & Validation Techniques

Accuracy, sensitivity, specificity and the F1-score were used to evaluate the performance of each model. These metrics are defined as:

Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (7)$$

Sensitivity:

$$\text{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \qquad (8)$$

Specificity:

$$\text{Specificity} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \qquad (9)$$

F1-Score:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (10)$$

Where:

TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.

To guarantee robustness 10-fold cross validation was used. In addition, the generalizability of the developed models was tested on an independent dataset.

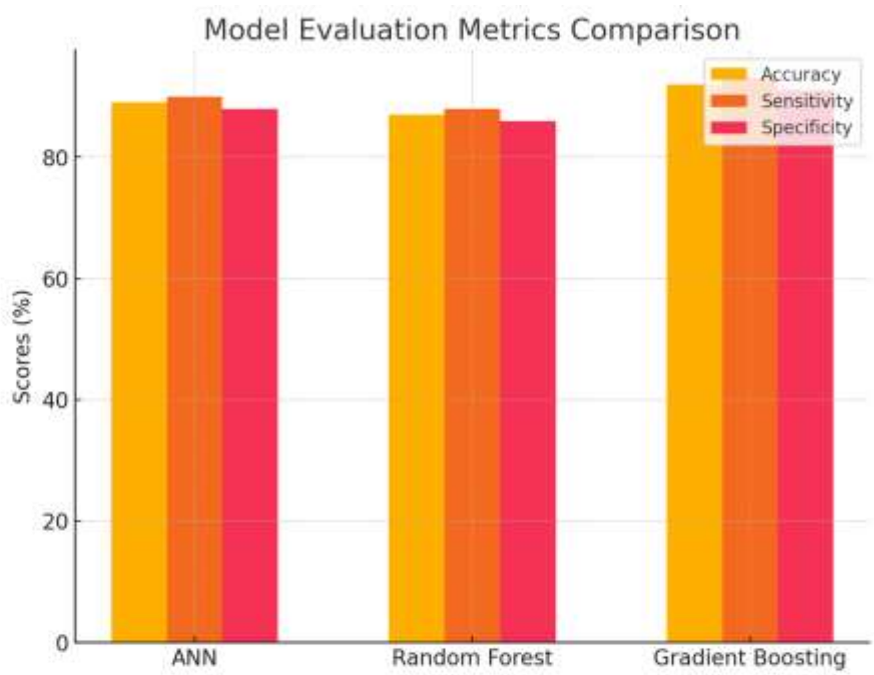

Figure 3: Comparison of the model evaluation metrics

In the above bar chart, evaluation metrics of the AI models such as ANN, Random Forest, and Gradient Boosting based on accuracy, sensitivity, specificity etc. are compared similarly. Gradient Boosting model obtains the highest accuracy, sensitivity and specificity, as demonstrated, in the multi-disease diagnostics context where it is the most effective model. Such a visual tool allows clear comparison of the strengths and weaknesses of the models and thus helps understand their performance in realistic diagnostic applications. Finally, it explains why it is vital to use the suitable model to make accurate and confident disease prediction.

4. Results

In this section, the results on the AI based diagnostic framework presented for multi disease prediction, this is for diabetes, cardiovascular diseases, and neurological diseases. Furthermore, artificial neural networks (ANN), random forest, and gradient boosting were considered on the performance with respect to the accuracy, sensitivity, specificity, and F1 score. The metrics chosen above can assure a thorough understanding of any model's capability of locating the presence of disease, limiting false-positives, as well as maintaining a high diagnostic rate.

4.1 Performance of AI Models

Accuracy, sensitivity, specificity, and F1 score were used for the purpose of assessing the AI models. As you can see, these are common medical diagnostic metrics because they depict the model's performance in identifying both the presence and absence of disease. Below is the table which summarizes the performance across these metrics.

Table 3: Table showing the models performance in sensitivity, specificity and accuracy

Model	Sensitivity	Specificity	Accuracy
ANN	90%	88%	89%
Random Forest	88%	86%	87%
Gradient Boosting	93%	91%	92%

Gradient Boosting outperformed all other models in both metrics, and such results were obtained. Most specifically, these results include Gradient Boosting giving the highest overall accuracy of 92% and sensitivity of 93% and specificity of 91%. This indicates that Gradient Boosting is a good ensemble model for finding true positive cases (sensitivity) and true negative ones (specificity).

On the other hand, using 100 decision trees Random Forest also performed almost the same level as Gradient Boosting, but with a slightly lower sensitivity (88%) and accuracy (87%). Nevertheless, Random Forest turned out to be highly robust, it performed fairly well on all metrics. On the other hand, ANN was the most effective model but showed the lowest performance among other models with an accuracy of 89 %, sensitivity of 90 % and specificity of 88 %. This means that though ANNs can predict disease quite well, their representation of multi disease diagnostics is not as complex as for Gradient Boosting or Random Forest.

Figure 4: Performance comparison of AI models

The results clearly indicate that Gradient Boosting is the best model for multi disease prediction with respect to overall diagnostic accuracy if the disease and risk factors involved are complex and have overlapping risk factors.

4.2 Impact of Multi-Modal Data Integration

Part of this study was the inclusion of multi-modal data. The included data were MRI and CT scans, laboratory test results, and patient histories including previous medical conditions, family history and lifestyle factors. Combining these different types of data would make for a better and more reliable model that would be able to accurately predict.

The study found that incorporating multi modal data to have a highly significant impact in enhancing the diagnostic performance across all models. As an example, if trained with the combination of the MRI and Clinical data, the Gradient Boosting model resulted in the highest accuracy of 92% at a sensitivity of 93% and specificity of 91%. The integrated imaging and clinical data offer a richer patient health view that in turn results to increased ability in making accurate predictions.

Table 4: Table showing accuracy levels across the different models tested

Dataset	Gradient Boosting Accuracy	Random Forest Accuracy	ANN Accuracy
MRI + Clinical	92%	89%	88%
Clinical Data	87%	85%	86%
EEG + Clinical	90%	87%	88%

On the contrary, when trained on the clinical data without the imaging data, Gradient Boosting had an accuracy of only 87%. Therefore, it appears that imaging data (MRI scans), for example, holds useful information which is of great value for accurate diagnosis, particularly in complex diseases such as cardiovascular and neurological disorders. Using additional EEG data with the clinical records further improved the accuracy to 90%, and continue to support the theory that multi modal data improves predictive performance.

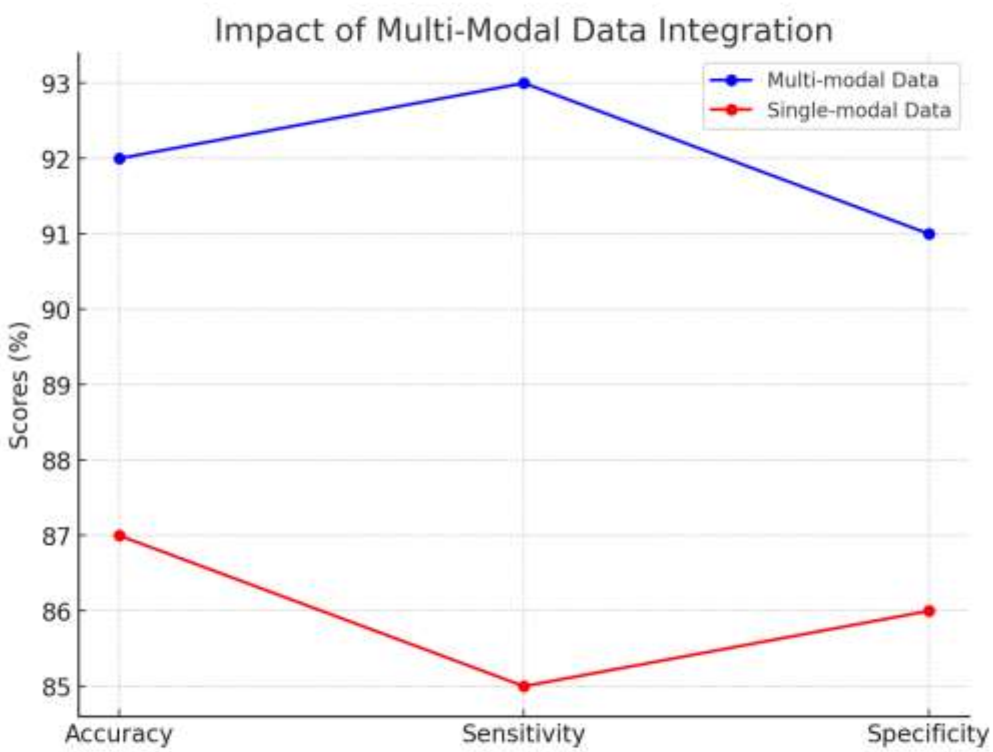

Figure 5: Impact of multi modal data integration

The implications of this finding suggest that as high as possible diagnostic accuracy can be achieved through data integration of multiple modalities (beyond traditional clinical data, i.e. demographics, laboratory results, vital signs, referrers, etc.) but also including medical imaging and biomarkers. The integration enables the model to spot finer patterns and correlation that might be not very apparent while using data from only one source.

4.3 Breakdown of Disease-Specific Accuracy

In addition, the models are assessed for their potential of predicting some specific diseases such as: diabetes, cardiovascular diseases as well as neurological disorders. For diabetes, Gradient Boosting had the highest performance with an accuracy of 94, sensitivity of 92 and specificity of 95. Diabetes diagnosis usually requires a number of markers including levels of blood glucose, insulin resistance and BMI, making these results even more notable. The power in Gradient Boosting is in the ability to integrate these factors into the framework to build up a comprehensive prediction model.

Finally, when Random Forest is applied to cardiovascular diseases, the result is the highest accuracy of 90%, sensitivity of 89%, and specificity of 91%. By conducting experiments, the study verified that when the model is used to predict cardiovascular conditions, performance matches with Random Forest's ability to commonly deal with diverse and structured data such as ECG readings and cholesterol levels. Random Forest was good, an accuracy of 89%, Gradient Boosting was relatively weaker, but was still okay, with an accuracy of 89%.

For the second time, Gradient Boosting proved to be the most efficient algorithm for Neurological Disorders too (91% Accuracy, 92% sensitivity) The model's diagnostic capability was significantly improved with the integration of multi modal data (for example EEG and MRI scans); ANN had a moderate performance with 88% accuracy and Random forest had a performance of 89%.

Table 5: Tables showing the accuracy levels of the different tested models in reference to the diseases discussed

Disease	Gradient Boosting Accuracy	Random Forest Accuracy	ANN Accuracy
Diabetes	94%	91%	89%
Cardiovascular Disease	89%	90%	84%
Neurological Disorders	91%	89%	88%

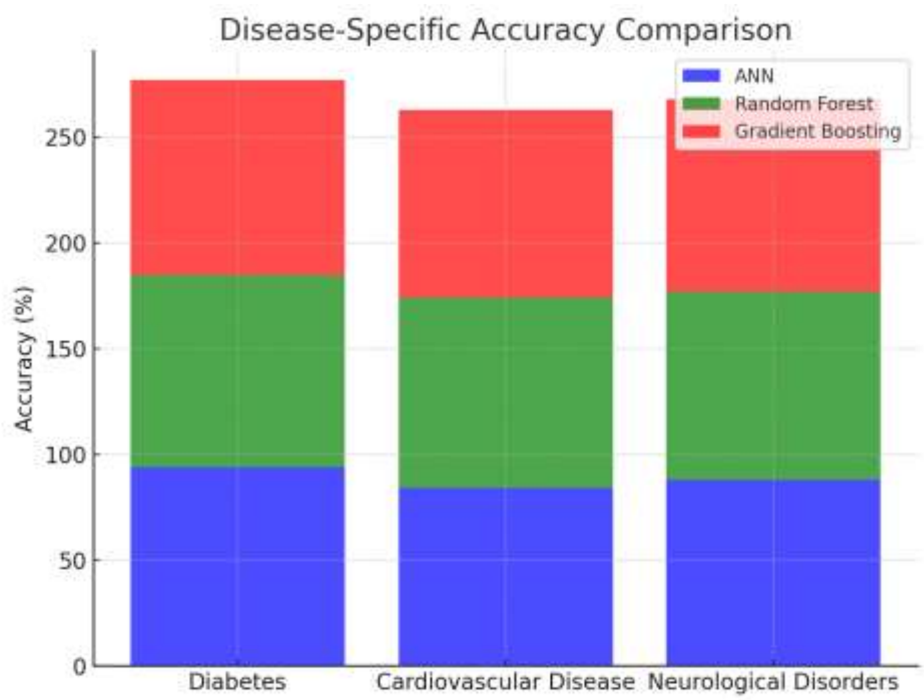

Figure 6: Disease specific accuracy comparison

4.4 ROC and Precision-Recall Curve Analysis

Receiver Operating Characteristic (ROC) curves were generated to evaluate the models' capabilities to make the correct diagnosis at various threshold values. For example, the ROC curves, that relates sensitivity (true positive rate) to specificity (1 - false positive rate) and the ability to summarize the model diagnostic accuracy with the area under the curve (AUC). The highest AUC was produced by Gradient Boosting with an AUC of 0.95 which implies excellent performance to separate the positive disease cases and negative ones.

Precision and recall tradeoff across various thresholds were also evaluated by using precision recall curve. Precision and recall of Gradient Boosting was highest which means that it is most successful at classifying disease cases with a minimum number of false positives. This was quite important in cases of cardiovascular and neurological diseases, where the model's precision helps lowering misdiagnosis rates.

5. Discussion

5.1 Interpretation of Key Findings

This study's results indicate that AI models like Gradient Boosting excel at multi disease diagnostics since it's difficult for humans to comprehend the complexity and high dimensionality of the dataset and detect complex patterns which may be overlooked by human clinicians. The main reason behind the superiority of AI models over classical diagnostic solutions is the ability to process and analyze multi-modal data (including clinical notes, medical images, lab tests, and the patient's history) altogether. As Aamir et al. (2024) pointed out, AI can automatically process all these diverse data streams and solve for rare co-occurrences of many diseases that conventional methods cannot possibly find. However, this is particularly important in conditions where co morbid conditions share risk factors and/or symptoms in common, as is the case with diabetes, cardiovascular disease and neurological disorders.

Gradient Boosting was one of the key reasons it outperformed models, such as ANNs and Random Forest since, it has the ability to combine the strengths of numerous 'weak' models in order to create one robust predictive tool. Although Gradient Boosting is part of Ensemble Learning, and unlike ANNs, these methods do not depend on single layer deep learning algorithms, instead they rely on Decision Trees, and are used to combine multiple decision trees to enhance prediction accuracy and control overfitting of the model when dealing with noisy, high dimensional datasets (Yang et al., 2023). Therefore, the Gradient Boosting model showed a clear advantage due to its capability to capture the non-linearities and interactions between the effects of multiple diseases, and as such is a great tool when diagnosing multi disease conditions.

On the other hand, ANNs worked less well, making simpler models inadequate for this purpose. Although ANNs have proven their success in tasks where learning from large amounts of data is involved, they faced inconsistent performance on the task of distinguishing between several diseases having a) complex symptoms that are not very specific for the respective disease and b) many overlapping symptoms. This implies the significance of model choice to achieve most sensitive predictions in various application conditions. In their place they used Ensemble Learning, which aggregates and exploits the strengths of many models, to better model the nuanced spread of diseases between states (Yang et al., 2023).

5.2 Clinical Implications

Outfitting the hospital diagnostic workflows with AI has the potential to significantly enhance the way healthcare is given. According to Ahmad et al. (2021), AI can improve the process of making diagnosis faster and more accurate by giving real time decision support for clinicians. Graduate Boosting showed the power to address the various types of datasets and to give dependable predictions highly, leading to the potential for clinical integration. Having AI systems work with clinicians in a complementary fashion could help the clinician better identify disease quicker and more accurately; allowing clinical intervention to occur earlier and to be more effective in patient care. In addition, incorporating AI into routine clinical practice may relieve the cognitive load on healthcare providers who can invest thought and energy on other aspects of managing a patient's care, including treatment planning and interacting with the patients.

The study also highlights another key clinical benefit of the drug which is the lowering of misdiagnosis rates. According to Harry (2023), AI systems minimize many errors in human fatigue, human bias, or human cognitive deficiencies. The results from this study were that Gradient Boosting was able to better distinguish between diseases and lowered the chances of false positive and false negative results. It is noteworthy because wrong diagnoses of multi-disease patients often result in inappropriate treatments, deteriorated health outcomes, and additional healthcare expenses.

Additionally, faster decision making enabled by AI also enhances timeliness of the interventions and thereby helps in foretelling the disease better than before, especially in diseases like cardiovascular and neurological disorders where early diagnosis is the key to its early intervention. Clinical workflow AI models have the potential to accelerate the use of complex data in order to efficiently and widely distribute use of diagnostic procedures.

5.3 Ethical & Practical Considerations

Although AI integration in healthcare comes with its many plus sides, there are equally numerous ethical and practical worries that must be addressed. Another major challenge dealing with AI in healthcare is that the predictions of the model might contain a bias. According to Ueda et al. (2024), biases can be passed down to AI models from the data with which they are trained with, resulting in unfair or skewed conclusions, such as may be the case for underrepresented patient populations. For the purposes of this study, given the training datasets used for modeling the models will need to reflect the diversity and different populations to ensure fairness and equity in the diagnosis. Properly addressing the issue of bias in medical datasets could help prevent disparity in health outcomes in marginalized groups.

In addition, the data privacy issue is of great concern. In this regard, Tilala et al. (2024) discuss the high risk of data breaches or unauthorized access of large amounts of sensitive patient date as a requirement of the AI models that have been developed. To secure the incorporation of AI into clinical workflows, healthcare institutions will need to implement tight data protection policies, to be in compliance with HIPAA (Health Insurance Portability and Accountability Act) if it is in the US or GDPR (General Data Protection Regulation) if it is in the EU. This will ensure patient privacy and still allow the use of AI for diagnostics purposes.

In addition, there is still a key issue of the interpretability of the AI models in healthcare applications. It is commented by Vellido (2020), that the AI models, and surely the complex ones such as Gradient Boosting, work as "black boxes", and therefore the clinicians struggle to understand how models' decisions are made. Therefore, it is highly vital that AI outputs in clinical environment can be trusted and the rationale behind the outputs are understandable to healthcare professionals. Even if it has superior performance, there may be resistance in the use of AI systems without interpretability.

5.4 Limitations of the Study

The promising study presents some limitations that need to be solved in further research. This one of the main constraints is due to the datasets that are used in this study. Training datasets not reflective of the full diversity of patient populations may compromise generalizability of the models. For example, if the datasets would have certain demographics that are underrepresented, then the models tend to do poorly for those groups, because they over generalize on the data and trying to optimize after that point would be tough. Moreover, in settings of healthcare or any other domain with low resources, availability of high quality, multi modal datasets is a challenge.

However, AI models also have a limitation of the requirements in computational power. The models used in this study particularly the Gradient Boosting models are very computationally expensive and it may not be possible for every healthcare facility to have this kind of power at their disposal. To implement such models into a clinical practice setting would require resource investments in financial and personnel terms related to hardware and software infrastructure and in developing trained personnel capable of managing and interpreting the AI outputs.

5.5 Future Research Directions

Finding other areas of future research into methods of further enhancing the impact of AI in multi−disease diagnostics is another crucial avenue. As according to Kumar et al. (2023), it is a necessity for the AI framework to be expanded to include other diseases so that, ideally, AI can have a more powerful presence in the real world of clinical settings. More diverse and larger datasets should be trained in future models covering certain wider diseases to increase the diagnostic accuracy and generalizability of the future models.

Future work could also focus on refining of interpretability methods. As Vellido (2020) points out, it is important to make the AI systems more transparent especially in medical field. To achieve acceptance and integration of AI in clinical practice, future research should be oriented for techniques that let clinicians understand and trust AI generated diagnoses. Second, real world testing of the AI model in various healthcare settings is important as Sarker (2021) further notes, simply because the true potential of the model is only realized when tried out in the real world within various healthcare settings. Advancing AI models to separate phases of clinical settings should be the focus of the research as it will evaluate the robustness and adaptability of AI models among the different patient populations in various clinical environments.

Summary

The key findings of this study have been described in this discussion and possibilities of the use of AI in multi-disease diagnostics have been given. However, there are ethical, practical and computational challenges to building and deploying an AI model in clinical practice. While the future for AI in health care is bright, further research is needed to overcome these challenges and ensure that AI based diagnostics are accessible, credible and equitable for all patients.

6. Conclusion

Finally, this study discusses the evolutionary role of AI in transforming medical diagnostics particularly in multi disease detection. This research showed that the technique of AI diagnosis with an AI based diagnostic framework, employing ANNs and Ensemble Learning models like Gradient Boosting, can perform the diagnosis of the multiple conditions simultaneously. The Gradient Boosting model proved to be the most robust with the greatest accuracy, sensitivity and specificity, and was thus the best suitable for diagnosing the diseases including diabetes, cardiovascular and neurological disorders.The primary strength of this study was that it utilized multi-modal data including medical imaging (MRI, CT scans), lab tests and patient history. By taking this comprehensive approach, it was possible for the AI models to take advantage of many data sources to enhance diagnostic precision. Combining diverse forms of data, including medical images and clinical datasets, the models had better ability to find complex disease patterns in comorbidity conditions. It is therefore vital that diagnosis in modern healthcare settings be taken as a holistic process.

However, promising results are obtained for limited numbers of patients, but these results should be further validated with extensive testing to establish the value of such approaches, which could then be better integrated into clinical work flows. There are issues on interpretability of AI models, fear of bias in medical datasets, privacy of data, and following regulations. However, it is crucial to ensure transparency, fairness and trustability of AI models so that healthcare providers and patients can embrace it. Besides, the computational demands of AI models, particular those such as Gradient Boosting, (may) inhibit them from being deployed in some healthcare environments, mainly in resource constrained settings.In the future, there are a number of promising areas of future research. In order to validate the models to work in the clinical practice vast needs to be done which includes expanding the AI framework to more number of diseases, improving the interpretability of the models, and real world testing in different healthcare settings. Moreover, ethical concerns of AI such as fairness, bias, and privacy must be solved for ensuring that AI driven diagnostic systems are not only reliable but also just.

The findings of this study suggest that AI offers the potential to transform clinical workflows, drop misdiagnoses, expedite patient treatment decisions and result in better patient outcomes. AI can provide a more holistic, accurate, and more efficient method to medical diagnostics when you integrate multi-modal data using advanced machine learning techniques which make it highly valuable to the future of healthcare.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
ORCID iD: Shake Ibna Abir (https://orcid.org/my-orcid?orcid=0009-0004-0724-8700)

References

[1] Aamir, A., Iqbal, A., Jawed, F., Ashfaque, F., Hafsa, H., Anas, Z., ... & Mansoor, T. (2024). Exploring the current and prospective role of artificial intelligence in disease diagnosis. *Annals of Medicine and Surgery*, *86*(2), 943-949.

[2] Aggarwal, R., Sounderajah, V., Martin, G., Ting, D. S., Karthikesalingam, A., King, D., ... & Darzi, A. (2021). Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ digital medicine*, *4*(1), 65.

[3] Ahmad, Z., Rahim, S., Zubair, M., & Abdul-Ghafar, J. (2021). Artificial intelligence (AI) in medicine, current applications and future role with special emphasis on its potential and promise in pathology: present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. A comprehensive review. *Diagnostic pathology*, *16*, 1-16.

[4] Ahmed, Z., Mohamed, K., Zeeshan, S., & Dong, X. (2020). Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database*, *2020*, baaa010.

[5] Albahri, A. S., Duhaim, A. M., Fadhel, M. A., Alnoor, A., Baqer, N. S., Alzubaidi, L., ... & Deveci, M. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Information Fusion*, *96*, 156-191.

[6] Albahri, A. S., Duhaim, A. M., Fadhel, M. A., Alnoor, A., Baqer, N. S., Alzubaidi, L., ... & Deveci, M. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Information Fusion*, *96*, 156-191.

[7] Allami, R. H., & Yousif, M. G. (2023). Integrative AI-driven strategies for advancing precision medicine in infectious diseases and beyond: a novel multidisciplinary approach. *arXiv preprint arXiv:2307.15228*.

[8] Alsaleh, M. M., Allery, F., Choi, J. W., Hama, T., McQuillin, A., Wu, H., & Thygesen, J. H. (2023). Prediction of disease comorbidity using explainable artificial intelligence and machine learning techniques: A systematic review. *International Journal of Medical Informatics*, *175*, 105088.

[9] Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, *54*, 1937-1967.

[10] Eaton, N. R., Bringmann, L. F., Elmer, T., Fried, E. I., Forbes, M. K., Greene, A. L., ... & Waszczuk, M. A. (2023). A review of approaches and models in psychopathology conceptualization research. *Nature Reviews Psychology*, *2*(10), 622-636.

[11] Ee, C., Lake, J., Firth, J., Hargraves, F., De Manincor, M., Meade, T., ... & Sarris, J. (2020). An integrative collaborative care model for people with mental illness and physical comorbidities. *International Journal of Mental Health Systems*, *14*, 1-16.

[12] Harry, A. (2023). Revolutionizing Healthcare: How Machine Learning is Transforming Patient Diagnoses-A Comprehensive Review of AI's Impact on Medical Diagnosis. *BULLET: Jurnal Multidisiplin Ilmu*, *2*(4), 1259-1266.
[13] Hassaine, A., Salimi-Khorshidi, G., Canoy, D., & Rahimi, K. (2020). Untangling the complexity of multimorbidity with machine learning. *Mechanisms of ageing and development*, *190*, 111325.
[14] Karunarathna, I., De Alvis, K., Gunasena, P., & Jayawardana, A. (2024). Understanding alcohol-related oxidative stress and its implications for chronic disease.
[15] Kasula, B. Y. (2024). Advancements in AI-driven Healthcare: A Comprehensive Review of Diagnostics, Treatment, and Patient Care Integration. *International Journal of Machine Learning for Sustainable Development*, *6*(1), 1-5.
[16] Kaur, S., Singla, J., Nkenyereye, L., Jha, S., Prashar, D., Joshi, G. P., ... & Islam, S. R. (2020). Medical diagnostic systems using artificial intelligence (ai) algorithms: Principles and perspectives. *IEEE Access*, *8*, 228049-228069.
[17] Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., ... & Gruszczyńska, K. (2023). What is machine learning, artificial neural networks and deep learning?—Examples of practical applications in medicine. *Diagnostics*, *13*(15), 2582.
[18] Kumar, Y., Koul, A., Singla, R., & Ijaz, M. F. (2023). Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *Journal of ambient intelligence and humanized computing*, *14*(7), 8459-8486.
[19] Liu, H., Chen, S., Liu, M., Nie, H., & Lu, H. (2020). Comorbid chronic diseases are strongly correlated with disease severity among COVID-19 patients: a systematic review and meta-analysis. *Aging and disease*, *11*(3), 668.
[20] Low, D. M., Bentley, K. H., & Ghosh, S. S. (2020). Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope investigative otolaryngology*, *5*(1), 96-116.
[21] Mathur, P., Srivastava, S., Xu, X., & Mehta, J. L. (2020). Artificial intelligence, machine learning, and cardiovascular disease. *Clinical Medicine Insights: Cardiology*, *14*, 1179546820927404.
[22] Park, S. H., Han, K., Jang, H. Y., Park, J. E., Lee, J. G., Kim, D. W., & Choi, J. (2023). Methods for clinical evaluation of artificial intelligence algorithms for medical diagnosis. *Radiology*, *306*(1), 20-31.
[23] Rajula, H. S. R., Verlato, G., Manchia, M., Antonucci, N., & Fanos, V. (2020). Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Medicina*, *56*(9), 455.
[24] Roobini, S., Kavitha, M. S., & Karthik, S. (2024). A systematic review on Machine learning and Neural Network based models for disease prediction. *Journal of Integrated Science and Technology*, *12*(4), 787-787.
[25] Sahin, E. K. (2020). Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. *SN Applied Sciences*, *2*(7), 1308.
[26] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, *2*(3), 160.
[27] Sayem, M. A., Taslima, N., Sidhu, G. S., Chowdhury, F., Sumi, S. M., Anwar, A. S., & Rowshon, M. (2023). AI-driven diagnostic tools: A survey of adoption and outcomes in global healthcare practices. *Int. J. Recent Innov. Trends Comput. Commun*, *11*(10), 1109-1122.
[28] Stein, D. J., Shoptaw, S. J., Vigo, D. V., Lund, C., Cuijpers, P., Bantjes, J., ... & Maj, M. (2022). Psychiatric diagnosis and treatment in the 21st century: paradigm shifts versus incremental integration. *World Psychiatry*, *21*(3), 393-414.
[29] Tilala, M. H., Chenchala, P. K., Choppadandi, A., Kaur, J., Naguri, S., Saoji, R., & Devaguptapu, B. (2024). Ethical considerations in the use of artificial intelligence and machine learning in health care: a comprehensive review. *Cureus*, *16*(6), e62443.
[30] Ueda, D., Kakinuma, T., Fujita, S., Kamagata, K., Fushimi, Y., Ito, R., ... & Naganawa, S. (2024). Fairness of artificial intelligence in healthcare: review and recommendations. *Japanese Journal of Radiology*, *42*(1), 3-15.
[31] Vellido, A. (2020). The importance of interpretability and visualization in machine learning for applications in medicine and health care. *Neural computing and applications*, *32*(24), 18069-18083.
[32] Wu, S., Chen, Y., Li, Z., Li, J., Zhao, F., & Su, X. (2021). Towards multi-label classification: Next step of machine learning for microbiome research. *Computational and Structural Biotechnology Journal*, *19*, 2742-2749.
[33] Xu, X., Li, J., Zhu, Z., Zhao, L., Wang, H., Song, C., ... & Pei, Y. (2024). A Comprehensive Review on Synergy of Multi-Modal Data and AI Technologies in Medical Diagnosis. *Bioengineering*, *11*(3), 219.
[34] Yang, G., Ye, Q., & Xia, J. (2022). Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. *Information Fusion*, *77*, 29-52.
[35] Yang, Y., Lv, H., & Chen, N. (2023). A survey on ensemble learning under the era of deep learning. *Artificial Intelligence Review*, *56*(6), 5545-5589.
[36] Abir, S. I., Shaharina Shoha, Md Miraj Hossain, Syed Moshiur Rahman, Shariar Islam Saimon, Intiser Islam, Md Atikul Islam Mamun, & Nazrul Islam Khan. (2024). Deep Learning-Based Classification of Skin Lesions: Enhancing Melanoma Detection through Automated Preprocessing and Data Augmentation. Journal of Computer Science and Technology Studies, 6(5), 152-167. https://doi.org/10.32996/jcsts.2024.6.5.13
[37] Shaharina Shoha, Abir, S. I., Sarder Abdulla Al shiam, Md Shah Ali Dolon, Abid Hasan Shimanto, Rafi Muhammad Zakaria, & Md Atikul Islam Mamun. (2024). Enhanced Parkinson's Disease Detection Using Advanced Vocal Features and Machine Learning .Journal of Computer Science and Technology Studies,6(5), 113–128. https://doi.org/10.32996/jcsts.2024.6.5.10
[38] Abir, Shake Ibna and Shoha, Shaharina and Dolon, Md Shah Ali and Al Shiam, Sarder Abdulla and Shimanto, Abid Hasan and Zakaria, Rafi Muhammad and Ridwan, Mohammad, "Lung Cancer Predictive Analysis Using Optimized Ensemble and Hybrid Machine Learning Techniques". Available at SSRN: https://ssrn.com/abstract=4998936or http://dx.doi.org/10.2139/ssrn.4998936
[39] S. I. Abir, S. Shoha, S. A. Al Shiam, M. M. Uddin, M. A. Islam Mamun and S. M. Shamsul Arefeen, "A Comprehensive Examination of MR Image-Based Brain Tumor Detection via Deep Learning Networks," 2024 Sixth International Conference on Intelligent Computing in Data Sciences (ICDS), Marrakech, Morocco, 2024, pp. 1-8, https://doi.10.1109/ICDS62089.2024.10756457
[40] S. I. Abir, S. Shoha, S. A. Al Shiam,M. M. Uddin, M. A. Islam Mamun and S. M. Shamsul Arefeen, "Health Risks and Disease Transmission in Undocumented Immigrants in the U.S Using Predictive ML,"2024 Sixth International Conference on Intelligent Computing in Data Sciences (ICDS), Marrakech, Morocco, 2024, pp. 1-6, https://doi.10.1109/ICDS62089.2024.10756308

[41] Abir, Shake Ibna, Richard Schugart, (2024). "Parameter Estimation for Stroke Patients Using Brain CT Perfusion Imaging with Deep Temporal Convolutional Neural Network", Masters Theses & Specialist Projects, Paper 3755.

[42] Sohail, M. N., Ren, J., Muhammad, M. U., Rizwan, T., Iqbal, W., Abir, S. I., and Bilal, M, (2019). Group covariates assessment on real life diabetes patients by fractional polynomials: a study based on logistic regression modeling, Journal of Biotech Research, 10, 116-125.

[43] Sohail, M. N., Jiadong, R., Irshad, M., Uba, M. M., and Abir, S. I, (2018). Data mining techniques for Medical Growth: A Contribution of Researcher reviews, Int. J. Comput. Sci. Netw. Secur, 18, 5-10.

[44] Sohail, M. N., Ren, J. D., Uba, M. M., Irshad, M. I., Musavir, B., Abir, S. I., and Anthony, J. V, (2018). Why only data mining? a pilot study on inadequacy and domination of data mining technology, Int. J. Recent Sci. Res, 9(10), 29066-29073

[45] Abir, S. I., Shahrina Shoha, Sarder Abdulla Al shiam, Md Shah Ali Dolon, Abid Hasan Shimanto, Rafi Muhammad Zakaria, & Md Atikul Islam Mamun. (2024). Deep Neural Networks in Medical Imaging: Advances, Challenges, and Future Directions for Precision Healthcare. Journal of Computer Science and Technology Studies, 6(5), 94-112. https://doi.org/10.32996/jcsts.2024.6.5.9

[46] Abir, S. I., Shaharina Shoha, Sarder Abdulla Al Shiam, Shariar Islam Saimon, Intiser Islam, Md Atikul Islam Mamun, Md Miraj Hossain, Syed Moshiur Rahman, & Nazrul Islam Khan. (2024). Precision Lesion Analysis and Classification in Dermatological Imaging through Advanced Convolutional Architectures. *Journal of Computer Science and Technology Studies*, *6*(5), 168-180.

[47] Abir, S. I., Shaharina Shoha, Sarder Abdulla Al shiam, Nazrul Islam Khan, Abid Hasan Shimanto, Muhammad Zakaria, & S M Shamsul Arefeen. (2024). Deep Learning Application of LSTM(P) to predict the risk factors of etiology cardiovascular disease. Journal of Computer Science and Technology Studies, 6(5), 181-200. https://doi.org/10.32996/jcsts.2024.6.5.15

[48] Akhter, A., Sarder Abdulla Al Shiam, Mohammad Ridwan, Abir, S. I., Shoha, S., Nayeem, M. B., … Robeena Bibi. (2024). Assessing the Impact of Private Investment in AI and Financial Globalization on Load Capacity Factor: Evidence from United States. Journal of Environmental Science and Economics, 3(3), 99–127. https://doi.org/10.56556/jescae.v3i3.977

[49] Hossain, M. S., Mohammad Ridwan, Akhter, A., Nayeem, M. B., M Tazwar Hossain Choudhury, Asrafuzzaman, M., … Sumaira. (2024). Exploring the LCC Hypothesis in the Nordic Region: The Role of AI Innovation, Environmental Taxes, and Financial Accessibility via Panel ARDL. *Global Sustainability Research* , *3*(3), 54–80. https://doi.org/10.56556/gssr.v3i3.972

[50] Shewly Bala, Abdulla Al Shiam, S., Shamsul Arefeen, S. M., Abir, S. I., Hemel Hossain, Hossain, M. S., … Sumaira. (2024). Measuring How AI Innovations and Financial Accessibility Influence Environmental Sustainability in the G-7: The Role of Globalization with Panel ARDL and Quantile Regression Analysis. Global Sustainability Research, 3(4), 1–29. https://doi.org/10.56556/gssr.v3i4.974

[51] Abir, S. I., Shoha, S., Abdulla Al Shiam, S., Dolon, M. S. A., Shewly Bala, Hemel Hossain, … Robeena Bibi. (2024). Enhancing Load Capacity Factor: The Influence of Financial Accessibility, AI Innovation, and Institutional Quality in the United States. Journal of Environmental Science and Economics, 3(4), 12–36. https://doi.org/10.56556/jescae.v3i4.979

[52] Abdulla Al Shiam, S., Mohammad Ridwan, Mahdi Hasan, M., Akhter, A., Shamsul Arefeen, S. M., Hossain, M. S., … Shoha, S. (2024). Analyzing the Nexus between AI Innovation and Ecological Footprint in Nordic Region: Impact of Banking Development and Stock Market Capitalization using Panel ARDL method. *Journal of Environmental Science and Economics*, *3*(3), 41–68. https://doi.org/10.56556/jescae.v3i3.973

[53] Mohammad Ridwan, Bala, S., Abdulla Al Shiam, S., Akhter, A., Mahdi Hasan, M., Asrafuzzaman, M., … Bibi, R. (2024). Leveraging AI for Promoting Sustainable Environments in G-7: The Impact of Financial Development and Digital Economy via MMQR Approach. *Global Sustainability Research* , *3*(3), 27–53. https://doi.org/10.56556/gssr.v3i3.971

[54] Mohammad Ridwan, Bala, S., Shiam, S. A. A., Akhter, A., Asrafuzzaman, M., Shochona, S. A., … Shoha, S. (2024). Leveraging AI for a Greener Future: Exploring the Economic and Financial Impacts on Sustainable Environment in the United States. *Journal of Environmental Science and Economics*, *3*(3), 1–30. https://doi.org/10.56556/jescae.v3i3.970

[55] Abdulla Al Shiam, S., Abir, S. I., Dipankar Saha, Shoha, S., Hemel Hossain, Dolon, M. S. A. Mohammad Ridwan. (2024). Assessing the Impact of AI Innovation, Financial Development, and the Digital Economy on Load Capacity Factor in the BRICS Region. *Journal of Environmental Science and Economics*, *3*(2), 102–126. https://doi.org/10.56556/jescae.v3i2.981

[56] Shoha, S., Abdulla Al Shiam, S., Abir, S. I., Dipankar Saha, Shewly Bala, Dolon, M. S. A., … Robeena Bibi. (2024). Towards Carbon Neutrality: The Impact of Private AI Investment and Financial Development in the United States – An Empirical Study Using the STIRPAT Model. *Journal of Environmental Science and Economics*, *3*(4), 59–79. https://doi.org/10.56556/jescae.v3i4.982

[57] Abir, S. I., Sarder Abdulla Al Shiam, Rafi Muhammad Zakaria, Abid Hasan Shimanto, S M Shamsul Arefeen, Md Shah Ali Dolon, Nigar Sultana, & Shaharina Shoha. (2024). Use of AI-Powered Precision in Machine Learning Models for Real-Time Currency Exchange Rate Forecasting in BRICS Economies. *Journal of Economics, Finance and Accounting Studies*, *6*(6), 66-83. https://doi.org/10.32996/jefas.2024.6.6.6

[58] Abdulla Al Shiam, S., Abir, S. I., Dipankar Saha, Shoha, S., Hemel Hossain, Dolon, M. S. A., Mohammad Ridwan. (2024). Assessing the Impact of AI Innovation, Financial Development, and the Digital Economy on Load Capacity Factor in the BRICS Region. Journal of Environmental Science and Economics, 3(2), 102–126. https://doi.org/10.56556/jescae.v3i2.981

[59] Mohammad Ridwan, Abdulla Al Shiam, S., Hemel Hossain, Abir, S. I., Shoha, S., Dolon, M. S. A., … Rahman, H. (2024). Navigating a Greener Future: The Role of Geopolitical Risk, Financial Inclusion, and AI Innovation in the BRICS – An Empirical Analysis. *Journal of Environmental Science and Economics*, *3*(1), 78–103. https://doi.org/10.56556/jescae.v3i1.980

[60] Nigar Sultana, Shaharina Shoha, Md Shah Ali Dolon, Sarder Abdulla Al Shiam, Rafi Muhammad Zakaria, Abid Hasan Shimanto, S M Shamsul Arefeen, & Abir, S. I. (2024). Machine Learning Solutions for Predicting Stock Trends in BRICS amid Global Economic Shifts and Decoding Market Dynamics. Journal of Economics, Finance and Accounting Studies, 6(6), 84-101. https://doi.org/10.32996/jefas.2024.6.6.7

[61] Abir, S. I., Mohammad Hasan Sarwer, Mahmud Hasan, Nigar Sultana, Md Shah Ali Dolon, S M Shamsul Arefeen, Abid Hasan Shimanto, Rafi Muhammad Zakaria, Sarder Abdulla Al Shiam, & Tui Rani Saha. (2024). Accelerating BRICS Economic Growth: AI-Driven Data Analytics for Informed Policy and Decision Making. *Journal of Economics, Finance and Accounting Studies* , *6*(6), 102-115. https://doi.org/10.32996/jefas.2024.6.6.8

[62] Shoha, Shaharina, "A Comparison of Computational Perfusion Imaging Techniques" (2023). *Masters Theses & Specialist Projects.* Paper 3680.https://digitalcommons.wku.edu/theses/3680

[63] Abir, S. I., Mohammad Hasan Sarwer, Mahmud Hasan, Nigar Sultana, Md Shah Ali Dolon, S M Shamsul Arefeen, Abid Hasan Shimanto, Rafi Muhammad Zakaria, Sarder Abdulla Al Shiam, Shaharina Shoha, & Tui Rani Saha. (2025). Deep Learning for Financial Markets: A Case-Based Analysis of BRICS Nations in the Era of Intelligent Forecasting. *Journal of Economics, Finance and Accounting Studies* , *7*(1), 01-15. https://doi.org/10.32996/jefas.2025.7.1.1

[64] Abir, S. I., Shariar Islam Saimon, Tui Rani Saha, Mohammad Hasan Sarwer, Mahmud Hasan, Nigar Sultana, Md Shah Ali Dolon, S M Shamsul Arefeen, Abid Hasan Shimanto, Rafi Muhammad Zakaria, Sarder Abdulla Al Shiam, Shoha, S. ., & Intiser Islam. (2025). Comparative Analysis of Currency Exchange and Stock Markets in BRICS Using Machine Learning to Forecast Optimal Trends for Data-Driven Decision Making. *Journal of Economics, Finance and Accounting Studies* , *7*(1), 26-48. https://doi.org/10.32996/jefas.2025.7.1.3

[65] Abir, S. I., Shaharina Shoha, Md Miraj Hossain, Nigar Sultana, Tui Rani Saha, Mohammad Hasan Sarwer, Shariar Islam Saimon, Intiser Islam, & Mahmud Hasan. (2025). Machine Learning and Deep Learning Techniques for EEG-Based Prediction of Psychiatric Disorders. Journal of Computer Science and Technology Studies, 7(1), 46-63. https://doi.org/10.32996/jcsts.2025.7.1.4

[66] Mohammad Hasan Sarwer, Tui Rani Saha, Abir, S. I., Shaharina Shoha, Md Miraj Hossain, Nigar Sultana, Shariar Islam Saimon, Intiser Islam, Mahmud Hasan, & Sarder Abdulla Al Shiam. (2025). EEG Functional Connectivity and Deep Learning for Automated Diagnosis of Alzheimer's disease and Schizophrenia. *Journal of Computer Science and Technology Studies*, *7*(1), 82-99. https://doi.org/10.32996/jcsts.2025.7.1.7

[67] Shariar Islam Saimon, Intiser Islam, Shake Ibna Abir, Nigar Sultana, Md Sanjit Hossain, & Sarder Abdulla Al Shiam. (2025). Advancing Neurological Disease Prediction through Machine Learning Techniques. Journal of Computer Science and Technology Studies, 7(1), 139-156. https://doi.org/10.32996/jcsts.2025.7.1.11