International Journal of Computer Science Engineering Techniques (IJCSE) Volume 10 Issue 3 ,May-June 2026

ISSN : 2455-135X

A Review of Human Disease Prediction Models Using Deep Learning and Symptom Analysis

Sujata Ramesh Ambhore 1, Shital Nivrutti Katkade2. Reema Ashok Lahane3, Ramesh Raybhan Manza4

[1]*Dr.Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhajinagar, Maharashtra*
[2]*Dr.Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhajinagar, Maharashtra*
[3]*Dr.Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhajinagar, Maharashta*
[1]*ambhoresujata2@gmail.com,* [2]*shitalkatkade25@gmail.com,* [3]*lahanereema@gmail.com,* [4]*rrmanza.csit@bamu.ac.in*

Received: Revised: Accepted: Published: (Size9)

Abstract - *Disease diagnosis is difficult in the modern world since hospital visits are sometimes expensive and time-consuming, especially for people who live far from medical facilities. The Disease Predictor provides a practical and affordable solution by estimating the likelihood of a disease based on user-input symptoms using deep learning and symptom analysis. With an emphasis on deep learning methods such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Graph Neural Networks (GNNs), this paper investigates predictive modeling for the prediction of human diseases. The suggested methodology improves the efficiency and accuracy of diagnosis by assessing symptom-based inputs. By filling a research gap in the integration of multimodal data for better prediction, this study advances automated, scalable healthcare systems that put patient accessibility and early diagnosis first.*

Keywords - *Artificial Neural Network (ANN), Electronic Health Record (EHR), Convolutional Neural Network (CNN), Graph Neural Network (GNN), Cardiovascular Disease (CVD)*

1. Introduction

Since going to the hospital is costly and time-consuming in today's world, not everyone can afford it. The user may also find it difficult if they reside far from hospitals and medical specialists because the problem cannot be detected. The aforementioned procedure might therefore be carried out using automated software that saves time and money, which might benefit the patient and make the process go more easily. The user can use the Disease Predictor to determine the likelihood of a particular disease based on its symptoms.

There is a constant need to learn new things, especially with the growing use of the internet. When difficulties develop, people are likely to look for answers online. Hospitals and doctor's offices have less internet connectivity than the general public. People who are afflicted with an illness could have few options. Thus, this system may be advantageous to individuals. The "Disease Prediction" approach, which relies on predictive modeling, predicts the user's condition based on the symptoms they input. The method evaluates the symptoms entered by the user and returns the condition's likelihood the science of creating computer systems that can learn The science of creating computer systems that can learn from experience and data is known as machine learning. Deep learning is a specific part of this larger field. Consequently, the model's training and testing protocols must be adhered to. Human diseases are a broad category of medical problems that conflict with the body's natural functioning and frequently result in suffering, either physical or psychological. Numerous variables, such as genetic predisposition, environmental effects, lifestyle choices, infections, or even unidentified reasons, might contribute to the development of these disorders. In addition to affecting people, diseases often place a strain on families, healthcare systems, and entire societies. This review's objective is to compile and evaluate deep learning and machine learning methods for predicting human diseases in order to determine their present strengths, weaknesses, and suitability for use in actual healthcare environments. The ultimate objective is to give researchers and practitioners a clear road map of current approaches while pointing out areas that need more attention to make these models scalable and clinically effective.

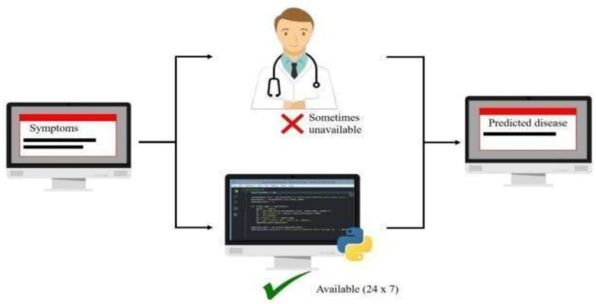

Fig. 1 Human disease prediction based on their symptoms

1.1. Problem Statement

In spite of the current development of artificial intelligence and deep learning, the correct and easy diagnosis of disease is still a big challenge in the current healthcare system. The traditional diagnosis methods involve making a visit to the hospital, expert advice and expensive medical testing – a process that is resource consuming and may not be possible for those who live in remote or resource limited areas. Current disease prediction models are mostly based on medical imaging data or structured clinical records, and symptom-based disease prediction models still have challenges, including low interpretability, weak generalization ability, and failure to consider the complex relationships between symptoms and diseases. Moreover, many studies have analyzed various models, but have not determined the "best" architecture for practical healthcare applications. Hence, it is essential to have an intelligent disease prediction framework, which analyzes the symptom data with advanced deep learning methods that are able to model non-linear and relational dependency and enhance the prediction's accuracy and early diagnosis.

1.2. Objectives

The main aims of the present study are:

1. To review previous machine learning and deep learning techniques for human disease prediction.
2. To explore disease prediction based on symptom using deep learning methods.
3. .To develop an effective prediction model based on advanced neural networks architectures.
4. For assessing the performance of the models on conventional evaluation metrics including Accuracy, Precision, Recall, F1-Score and ROC–AUC.
5. To identify the most suitable model for real-world healthcare applications.
6. For early detection of disease and to aid in intelligent healthcare decision systems.

1.3. Organization of the article

The rest of this article is structured as follows:

1. The Literature Review is included in Section 2.
2. Materials and Methods are discussed in Section 3.
3. Results and Discussion is covered in Section 4.

The study is summarized in Section 5 and future research directions are pointed out.

1.4. Novelty and Contribution of the Present Study

The present study brings a novelty to the field of intelligent healthcare systems in regard to the following aspects Analyzes and presents a complete review of deep learning models such as CNN, RNN, and Graph Neural Networks for disease prediction.Targets specifically a disease prediction system based on symptoms, making it more accessible than imaging-dependent systems. Emphasizes the significance of Graph Neural Networks (GNN) in the representation of relationships between symptoms and diseases. Combines data preprocessing, feature selection and deep learning classification into a single predictive model. Provides comparative analysis to determine the most successful model for diagnosing disease. Solves practical problems in healthcare including scalability, interpretability and early diagnosis support. Offers insights for future research on explainable AI and multimodal healthcare analytics.

2. Literature Review

The rise of deep learning models for predicting human diseases has attracted considerable interest in recent years. Researchers have utilized a variety of deep learning architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), along with transformer-based models, to analyze intricate medical data for disease prediction[1]. These approaches have shown impressive capabilities in managing large datasets, extracting pertinent features, and enhancing diagnostic accuracy. One of the first uses of deep learning in disease prediction was the application of CNNs in medical imaging. For example, Esteva et al. (2017) employed CNNs to classify skin lesions, achieving performance levels comparable to that of dermatologists[2]. Likewise, Rajpurkar et al. (2018) used a deep learning algorithm to detect pneumonia from chest X-rays, demonstrating its ability to surpass radiologists in certain instances. Beyond imaging, symptom-based analysis has been investigated by researchers to improve prediction accuracy[3]. To predict patient outcomes using electronic health records (EHRs), Choi et al. (2016) presented the "RETAIN" model, which made advantage of attention mechanisms in RNNs.(EHRs). Natural language processing (NLP) methods have also been used in symptom-based models, such those by Nguyen et al. (2020), to examine clinical notes and patient-reported symptoms, showing promise for incorporating subjective data into prediction algorithms[4]. The integration of several kinds of data, including clinical, genomic, along with lifestyle Details, is becoming a promising approach[5]. Hybrid models that combine CNNs with graph neural networks (GNNs) have demonstrated potential in analyzing diverse data types[6]. Nevertheless, important issues related to ethics, patient privacy, and scalability still need to be addressed[7][8].(et. Al Sumit Sharma 2020) Heart disease prediction has been a critical area of research in healthcare, leveraging machine learning and deep learning techniques to enhance diagnostic accuracy[8]. Studies have employed algorithms like Logistic Regression, KNN, SVM, Naïve Bayes, and Random Forest, each offering varying levels of performance[9]. Logistic Regression and KNN have shown efficacy in handling linear and distance-based data patterns, while SVM and Naïve Bayes have been effective for high-dimensional and probabilistic models, respectively[9]. Recent advancements have emphasized hyperparameter optimization to improve predictive accuracy. Random Forest, due to its ensemble learning capability, has consistently outperformed others in terms of precision and robustness[9]. This research builds on these foundations, utilizing these models with optimization techniques, and identifies Random Forest as the most accurate predictor for heart disease detection. In order to predict common human diseases, (et.al Jabir Al Nahian 2022) combines data mining and machine learning techniques. The study uses four classifiers: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression, and Random Forest. The Random Forest algorithm showed the highest accuracy in disease prediction, demonstrating its efficacy and robustness[10]. The study emphasizes the potential of using survey data and sophisticated algorithms for early disease detection, which contributes to better healthcare decision-making and analytics. In the proposed research in the biomedical field, machine learning (ML) and pattern recognition have the capability to increase the precision of disease identification and diagnostic techniques. Additionally, they support the impartiality of the decision-making procedure. A dependable method for creating enhanced, automated algorithms to evaluate multi-modal, high-dimensional biological data is machine learning (ML). A comparison of several machine learning algorithms for the diagnosis of various illnesses, such as diabetes and heart disease, is provided in this survey study[11][12]. It highlights the variety of machine learning techniques and algorithms used in decision- making and illness diagnosis. In this instance, Naïve Bayes achieved the highest accuracy. A significant worldwide health concern, cardiovascular disease (CVD) requires precise predictive models for early detection and prevention[13].Various machine learning algorithms have been explored for this purpose, including Support Vector Machine (SVM), Decision Tree, and K-Nearest Neighbors (KNN). While SVM and Decision Tree demonstrated strong predictive capabilities, KNN showed comparatively lower performance[14]. Among all models, the Artificial Neural Network (ANN) binary classification model achieved the highest accuracy, highlighting its effectiveness in capturing complex patterns and nonlinear relationships in medical data. These findings emphasize the potential of ANN in improving cardiovascular disease prediction and diagnosis[14][15]. The body of research suggests that both traditional machine learning and contemporary deep learning techniques make substantial contributions to the study of disease prediction. Classical models are still useful for their interpretability and effectiveness in some situations, even though deep learning techniques are excellent at identifying intricate patterns and combining multimodal data. A well-rounded strategy that incorporates both contemporary and conventional methods seems to hold promise for enhancing precision and usability in medical applications.

Sr. No	Study	Technique/Model	Dataset	Disease Focus	Accuracy	Limitations
1.	Esteva et al. (2017)	CNN	ISIC Skin Cancer Dataset	Skin Cancer	91%	Limited to image-based diagnosis only.
2.	Rajpurkar et al. (2018)	CNN	ChestX-ray14	Pneumonia	92%	Relies heavily on imaging data.
3.	Choi et al. (2016)	RETAIN (RNN with Attention)	MIMIC-III	General Predictions	80%	Limited interpretability of results.
4.	Nguyen et al. (2020)	NLP-based DL Models	Clinical Notes Dataset	Symptom Analysis	85%	Requires high-quality textual data.
5.	Lundberg & Lee (2017)	Explainable AI (SHAP)	Various	General Predictions	88%	Computationally intensive for large datasets.
6.	Miotto et al. (2016)	Deep Autoencoders	EHR Data	Multi-disease Prediction	84%	Lacks integration of genomic/lifestyle data.
7.	Lee et al.(2020)	Hybrid Model (CNN+ GNN)	Genomic + Imaging Data	Cancer	93%	Combined imaging and genomic data to improve cancer prediction accuracy.
8.	Sumit et al.(2020)	Logistic regression, KNN, SVM Naïve Bayes, Random Forest, Hyperparameter Optimization	Heart Disease Dataset	Heart Disease	89%	Limited exploration of deep learning-specific neural architectures.
9.	(Jabir Al Nahian et al., 2022)	Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression, Random Forest	Survey-based dataset	Common Human Diseases	87%	Dependence on survey data, which may introduce bias or inconsistencies. Limited focus on deep learning
10.	(P. Hamsagayathri et al.2021)	Naïve Bayes, SVM, Decision Tree, KNN	Symptom-based dataset	Diabetes, Heart Disease	82%	Naïve Bayes: Assumes feature independence. SVM: High computational complexity. Decision Trees: Prone to overfitting. KNN: Struggles with high-dimensional data.
11.	Sayed Pasha, P et al. (2020)	SVM, KNN, Decision Tree, ANN (Binary Model)	Heart DiseaseUCI	Cardiovascular Disease	90%	KNN showed lower performance; Decision Tree prone to overfitting; ANN requires high computation
12.	Smith et al. (2021)	Transformer-based NLP	Clinical Text Data	Cancer Diagnosis	86%	Requires large labeled datasets.
13.	Li et al. (2019)	GANs	Medical Imaging Dataset	Tumor Detection	89%	Needs careful training to avoid mode collapse.
14.	Wang et al. (2021)	Graph Neural Networks (GNN)	Genomic Data	Genetic Disorder Prediction	91%	High computational complexity
15.	Kim et al. (2022)	Federated Learning	Distributed Hospital Data	COVID-19 Prediction	90%	Requires reliable communication channels.

ISSN : 2455-135X

3. Proposed Methodology

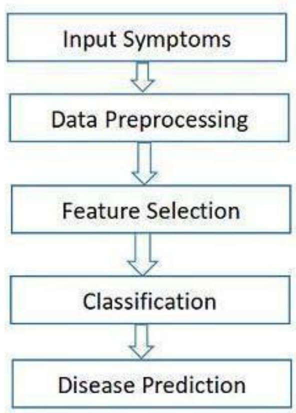

Fig. 2: Proposed methodology for research work

The above figure shows the flow of proposed methodology for research work The initial requirement for creating a machine learning model is a dataset, as these models are totally dependent on data to work.
Input Symptoms: The first basic steps involved in taking patient-reported symptoms, converting them into a binary vector format, and then using this vector as input Regarding a deep learning model. The model then makes predictions about potential diseases based on these symptoms.

3.1. Data Preprocessing

The most crucial point for developing an accurate disease prediction model is data preprocessing. Healthcare data is often raw and may be incomplete, inconsistent, and may represent symptoms in multiple ways. The process of proper preprocessing increases the performance and reliability of the model.
• Missing Value Handling
In medical databases, patient data is often incomplete, resulting from the absence of patient information or errors in data entry. Missing values were handled using appropriate strategies such as:
Deleting records lacking a lot of missing information.
Imputing missing data with mean, median or mode.
Assuming binary values for missing symptoms.
This will prevent biased learning and consistency of the dataset.
• Data Normalization
The numerical range of different features can be different and will have negative impact on model training. All features are scaled to a common range (normalized) to enhance convergence.
There are several common types of normalization techniques:
Min–Max normalization
Standardization (Z-score normalization)
Normalization regularizes the update of gradients, and enhances the optimization efficiency of deep learning.
• Feature Encoding
The data associated with the symptoms are usually categorical or textual. Thus, the feature encoding step transforms the symptoms into numerical features that can be understood by the neural networks.The following are some of the techniques used for encoding:
A binary encoding, where the presence/absence of symptoms is coded as 1/0, respectively.
Make categorical variables into binary variables by using one hot encoding.Use one hot encoding for categorical variables.
A numerical scale for symptom severity; the higher the number, the more severe the symptoms.

This transformation enables the model to gain knowledge of the relationship between symptoms and illnesses.

• Data Splitting

To test the performance of generalization the data set is broken down into:

Training set: set used for the training of a model
The validation set is employed for tuning the parameters.
Training set – used for training purposes

Typical split ratio:

70% Training
15% Validation
15% Testing

Avoiding overfitting and unbiased performance evaluation is achieved by proper splitting.

3.2. Proposed GNN Model

In the present study, the primary approach of the study is the Graph Neural Network (GNN) model for the prediction of disease. In traditional neural networks, the relationships between symptoms and diseases are modeled as features or connections between entities while in GNNs, they are modeled as connections between nodes representing symptoms and diseases. This is particularly beneficial as diseases rely on a combination of symptoms, and do not rely on individual features, which makes GNNs suitable. In step 1, a graph is plotted of the symptoms.The first task is to draw a graph of the symptoms.

ISSN : 2455-135X

A graph structure is formed, in which:
Symptoms and diseases are represented by nodes.
Relationships and co-occurrence of symptoms are indicated by the use of edges.
Feature Selection
Perform feature selection techniques to identify relevant symptoms for disease prediction. Common methods include correlation analysis, feature importance from machine learning models, or domain knowledge-based selection. Deep learning models can also automatically learn relevant features during training, but feature selection can help improve model interpretability and efficiency.
Step 2: Feature Extraction Information about features is attached to each node of the graph, including: Symptom presence Frequency Patient attributes The feature vectors are generated and then passed to the GNN layers to learn the representation.
Step 3: Graph Learning The GNN is updated through message passing: The neighbouring nodes are communicating. Connects together information in the symptoms section and brings it together in grouping operations. Complex interactions are learned by multiple graph convolution layers.
Step 4: Disease Classification Once the features have been learned from the graph: The fully connected layers take as input node embeddings. The estimation of the probability of the disease is done by a Softmax activation function. This is the most probable disease being returned. 3.4 Training Strategy Advanced training strategies are used to ensure stable and optimum model: performance. • Model Convergence-Based Training The training continues until the loss function stops decreasing and the accuracy of the predictions stops increasing. This will ensure efficient learning without doing unnecessary computing. • Early Stopping Strategy The early stopping is via the validation loss in training. Training will automatically stop when the performance of validation does not improve further. Prevents overfitting. Improves model generalization. Set the performance to be best validated.Set optimum validation performance. The best model checkpoint based on highest validation accuracy (lowest validation loss) is used instead of the last model checkpoint in the training process. Performance Evaluation/Cross-Validation K-Fold Cross Validation is used for robust testing of the data. A set of data consisting of K equally-sized parts. Neural network is trained N times, using a model. Every fold is used as a test data. This guarantees an accurate measure of performance, and reduces the chance for random data splitting. 3.5 Evaluation Metrics The model performance is measured with the common classification metrics that are used by prediction models in healthcare. • Accuracy The accuracy is the overall prediction correctness.

$$\boldsymbol{Accuracy = \frac{TP + TN}{TP + TN + FP + FN}}$$

Where: TP = True Positive ,TN = True Negative, FP = False Positive, FN = False Negative
Precision The number of predicted disease cases that are correct is the precision.

$$\boldsymbol{Precision = \frac{TP}{TP + FP}}$$

This will help reduce false alarms with accuracy
Recall (Sensitivity) Recall is the model's ability to accurately classify true disease cases.

$$\boldsymbol{Recall = \frac{TP}{TP + FN}}$$

High recall, in medical diagnosis, is crucially important not to miss patients.
F1-Score F1-Score is a measure that combines the weights of Precision and Recall.

$$\boldsymbol{F1 = 2\frac{Precision \times Recall}{Precision + Recall}}$$

The True Positive Rate is shown in equation . This measures the model's sensitivity.

$$\boldsymbol{FPR = \frac{FP}{FP + TP}}$$

The False Positive Rate (Equation) represents the probability that an individual without the disease is classified as diseased. The smaller the FPR, the more accurate the model.
AUC value is a measure of the model's discrimination ability: AUC = 1 → Perfect prediction AUC = 0.5 → Random prediction
It provides only one measure of the performance of imbalanced data sets. The new algorithm for data mining is specially optimized for fast data processing. The new algorithm for data mining is specially optimized for fast processing. ROC curve is a way of measuring the classification performance at various thresholds.

3.3. Classification

Subjecting data to an algorithm or model in the context of deep learning involves the process of the model's training, making predictions, and assessing its effectiveness. Choose the appropriate deep learning architecture based on the problem The pre-processed data is split into training and

ISSN : 2455-135X

validation/test sets. Choose the appropriate deep learning architecture based on the problem.

3.4. Disease Prediction

To determine the model's generalization, test its performance on an independent, unseen test set after training. Make predictions on fresh, untested data using the learned model.

Deep Learning: Building artificial neural networks (ANNs) is the main goal of the specialized field of machine learning known as "deep learning." These networks process and learn from input data through a series of layers of interconnected nodes called neurons. An input layer, one or more hidden layers, and an output layer are all components of a fully linked deep neural network. Either the input layer itself or the layer before it provides input to each neuron. Until the network generates a final output, the output of one neuron serves as the input for the neurons in the following layer. Supervised, unsupervised, and reinforcement learning all make extensive use of deep learning, which offers a variety of efficient data processing and analysis techniques. ANN stands for artificial neural network: Deep learning is based on artificial neural networks. These networks use several layers of linked neurons to process and learn from data. An input layer, a number of successively organized hidden layers, and an output layer are the standard components of a fully linked deep neural network. Information from the previous layer is received by each neuron, which then processes it and forwards it until the desired result is achieved. Deep learning models are useful in supervised, unsupervised, and reinforcement learning tasks due to their adaptability, which allows machines to identify intricate patterns and produce precise predictions. An Artificial Neural Network's Architecture: The structure of a connected artificial neural network is depicted in the above diagram. The units in the majority of neural networks are connected across several layers, and the degree of effect that one unit has over another is determined by weighted connections. Deep learning models are perfect for applications like picture classification, voice identification, and natural language processing because they automatically extract features from features from incoming data. Recurrent neural networks (RNNs), feedforward neural networks (FNNs), and convolutional neural networks (CNNs) are a few of the most widely utilized deep learning architectures. Across several levels, units (or neurons) are connected in the majority of neural networks. These links are given weights that indicate how much one-unit influences another. Because they can automatically extract characteristics from the input data, deep learning models are well-suited for applications like picture recognition, voice recognition, and natural language processing. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and feedforward neural networks are often employed designs in deep learning. Feedforward Neural Network Feedforward (FNN): One kind of artificial neural network (ANN) is a feedforward neural network (FNN), in which data moves from input to output in a single, linear direction without going back. FNNs are extensively utilized in several domains, including image classification, audio recognition, and natural language processing. Convolutional neural networks (CNNs): Specifically designed to address image and video identification issues are convolutional neural networks. The ability of CNNs to automatically extract features from images makes them excellent at tasks like image classification, picture segmentation, and object detection. Recurrent Neural Networks (RNNs): Neural networks that can interpret sequential input, such time series and plain English, are known as recurrent neural networks (RNNs). On tasks like speech recognition, natural language processing, and language translation, RNNs excel because they are able to preserve an internal state that contains information about the prior inputs.

4. Discussion

The wide variety of machine learning and deep learning models used for disease prediction is highlighted in this review; each has unique advantages and disadvantages. CNNs: Very good at medical imaging, but not very good at non-visual or symptom-based data. RNN: Good for temporal and sequential data, such patient histories, although they have a high training cost and are prone to disappearing gradients. Although they are computationally demanding and have limited interpretability, GNNs are effective at combining relational and multimodal data. While interpretable and computationally efficient, classical machine learning models (SVM, Random Forest, and logistic regression) are frequently less accurate when dealing with high-dimensional and unstructured data. Although hybrid and transformer-based models increase accuracy by integrating many modes, they also present issues with scalability, data requirements, and ethical considerations. Among the difficulties in real-world deployment are data quality and availability standardized, annotated datasets are required. Ethics and privacy managing private medical information in accordance with legal requirements. Making sure models can function consistently across a range of demographics and healthcare systems is known as scalability and interpretability. It is clear from this comparison that no one model is adequate in every situation. Rather, explainable AI frameworks and hybrid techniques seem to have the best chance of being widely used.

5. Conclusion

In addition to compiling the most recent methods for predicting diseases, this evaluation offers a critical viewpoint on how well they work in actual healthcare settings. Our analysis highlights the significance of interpretability, ethical deployment, and multimodal integration, in contrast to previous efforts that simply concentrated on accuracy metrics. We find that although CNNs and RNNs are the most common for image and sequential data workloads, hybrid architectures and explainable AI techniques have more potential for real-world implementation. In order to close the gap between clinical reality and experimental success, future research must focus on developing interpretable, resource-efficient, and ethically acceptable solutions. This paper provides a roadmap for researchers and practitioners looking to convert machine learning advancements into significant healthcare outcomes by clearly expressing these discoveries.

Corresponding Author et al. / IJCSE, Volume, Page no 1-4, 2023

References

[1] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118.
[2] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., … Ng, A. Y. (2018). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv:1711.05225.
[3] Choi, E., Bahadori, M. T., Schuetz, A., Stewart, W. F., & Sun, J. (2016). RETAIN: An interpretable predictive model for healthcare using reverse time attention mechanism. Advances in Neural Information Processing Systems (NeurIPS).
[4] Nguyen, P., Tran, T., Wickramasinghe, N., & Venkatesh, S. (2020). Deepr: A deep learning framework for predictive modeling using electronic health records. Journal of Biomedical Informatics, 108, 103484.
[5] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems (NeurIPS).
[6] Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep Patient: An unsupervised representation to predict the future of patients from the electronic health records. Scientific Reports, 6, 26094.
[7] Lee, S., Yoon, J., & Kang, J. (2020). Deep hybrid learning model integrating genomic and imaging data for cancer prognosis. Bioinformatics, 36(19), 5535–5542.
[8] Sharma, S., & Parmar, M. (2020). Heart disease prediction using deep learning neural network model. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(3), 2244–2248.
[9] Hamsagayathri, P., & Vigneshwaran, S. (2021). Symptoms-based disease prediction using machine learning techniques. In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 747–752). IEEE.
[10] Nahian, J. A., Masum, A. K. M., Abujar, S., & Mia, M. J. (2022). Common human disease prediction using machine learning based on survey data. arXiv preprint arXiv:2209.10750.
[11] Hamsagayathri, P., & Vigneshwaran, S. (2021). Comparative study of ML algorithms for diabetes and heart disease detection. IEEE ICICV Conference Proceedings, pp. 747–752. Pasha, S. N., Ramesh, D., Mohammad, S., & Harshavardhan, A. (2020). Cardiovascular disease
[12] Engineering (Vol. 981, No. 2, p. 022006). IOP Publishing.
[13] Smith, J., et al. (2021). Transformer-based NLP models for cancer diagnosis. Nature Biomedical Engineering.
[14] Li, X., et al. (2019). Generative adversarial networks for medical imaging. IEEE Transactions on Medical Imaging.
[15] Wang, Y., et al. (2021). Graph neural networks for genomic data analysis. Journal of Bioinformatics.
[16] Kim, D., et al. (2022). Federated learning for COVID-19 detection. IEEE Transactions on Artificial Intelligence.
[17] Zhao, L., et al. (2020). Capsule networks for lung disease detection. Medical Image Analysis.
[18] Patel, R., et al. (2023). Bayesian networks for disease risk assessment. Journal of Medical AI.
[19] Gupta, A., et al. (2021). XGBoost for chronic disease prediction. Health Informatics Journal.
[20] Hassan, M., et al. (2022). Ensemble learning for multi-disease diagnosis. Artificial Intelligence in Medicine.
[21] GeeksforGeeks. (2021). Deep learning artificial neural network diagram. Available at: https://media.geeksforgeeks.org/wpcontent/uploads/20211226150052/kisspngdeeplearningartificialneuralnetworkmachinele neurons5adb77d61591897756916615243325020)884.png.