Multiple Disease Prediction Using Machine Learning with Chatbot and Doctor-Patient Appointment System

Aleena Raichel George
Dept of Computer Science & Engg.
MITS
Ernakulam, India
20cs026@mgits.ac.in

Jeswin Eldho Saji
Dept of Computer Science & Engg.
MITS
Ernakulam, India
20cs045@mgits.ac.in

Juliya Francis
Dept of Computer Science & Engg.
MITS
Ernakulam, India
20cs047@mgits.ac.in

Laven Shibu
Dept of Computer Science & Engg.
MITS
Ernakulam, India
20cs199@mgits.ac.in

Haritha H (Asst.Proff)
Dept of Computer Science & Engg.
MITS
Ernakulam, India
haritha@mgits.ac.in

***Abstract*—The paper analyzes the potential of machine learning in disease prediction and its integration into healthcare systems, offering a proactive healthcare strategy. ML-driven disease prediction combines powerful algorithms and a range of healthcare information. By carefully examining this information, machine learning models show patterns and relationships that are sometimes hard for human therapists to discover. These models could predict the chance of patients having numerous diseases, such as brain tumors, breast cancer, CKD,and heart disease, reaching beyond chronic ailments. This project contains three key components. Firstly, it contains a robust illness detection system that employs machine learning methods in Convolutional Neural Networks (CNN), Support Vector Machine(SVM) and Random Forest(RF). Our proposed approach exhibits amazing accuracy and faster convergence compared to standard machine-learning disease risk prediction techniques. It reliably predicts a spectrum of illnesses by evaluating huge datasets, exceeding existing disease risk prediction approaches. Secondly, the platform provides a chatbot that uses natural language processing (NLP) to assist with medical enquiries. Additionally, it features an efficient appointment system for live video consultations with doctors through WebRTC. In summary, this initiative represents an innovative mix of machine learning and healthcare services, delivering a holistic and proactive approach to healthcare.**

***Index Terms*—Machine Learning, SVM, CNN, Random Forest, SVM, Disease Detection, Healthcare Integration.**

I. Introduction

In the realm of healthcare, professionals grapple with the challenge of early disease diagnosis amidst a deluge of medical data. While existing solutions often focus on individual ailments, our project takes a comprehensive approach. Our app leverages advanced Machine Learning techniques, including K-Nearest Neighbors (KNN), Random Forest, and Convolutional Neural Networks (CNN), to predict a diverse range of diseases accurately. By integrating these sophisticated algorithms, we aim to empower users with timely and precise insights into their health status, facilitating proactive healthcare management and improved patient outcomes.

Central to our app's functionality is its intuitive interface, which provides users with easy access to a multitude of disease predictions within a single platform. Alongside predictive capabilities, we have implemented a doctor-patient appointment system, ensuring seamless access to healthcare professionals for consultations and treatments. This integration streamlines the healthcare journey, encouraging early intervention and fostering a culture of proactive health management among users.

Moreover, our app incorporates Natural Language Processing (NLP) technology to power a chatbot feature, enhancing user engagement and accessibility. The NLP-driven chatbot enables users to ask medical queries in natural language, receiving prompt and informative responses. Combined with the predictive capabilities of KNN, Random Forest, and CNN, the chatbot provides users with comprehensive support, from disease prediction to real-time medical assistance. By seamlessly integrating advanced Machine Learning techniques and NLP capabilities, our app strives to revolutionize disease prediction and healthcare accessibility, ultimately empowering individuals to take control of their health journey.

II. Related Work

Anusha Ampavathi have explored different applications that demonstrates the significance of the ML methods in various areas [1]. They proposed that big data plays a pivotal role in healthcare, particularly in disease prediction based on patient

symptoms, ensuring timely intervention. This research focuses on multi-disease prediction, encompassing Diabetes, Hepatitis, Lung cancer, Liver tumor, Heart disease, Parkinson's, and Alzheimer's. The study employs a comprehensive approach involving three main phases: data normalization, weighted feature extraction utilizing JA-MVO algorithm, and prediction using hybrid Deep Belief Network (DBN) and Recurrent Neural Network (RNN). The proposed model demonstrates superior performance compared to conventional ML algorithms. For instance, in the case of Parkinson's disease, the accuracy of JA-MVO-RNN + DBN surpasses SVM, KNN, RNN, DBN, and RNN + DBN by significant margins (10% , 1.4%, 28% , 13%, and 44%, respectively). This analysis underscores the effectiveness of the developed method in multi-disease prediction, showcasing its potential for enhancing healthcare decision-making.

The authors of [5] proposed a methodology for efficient query resolution in a hospital management system with an integrated AI chatbot. Leveraging advanced machine learning techniques, particularly Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN), the chatbot demonstrates robust language comprehension. The LSTM, as a type of RNN, incorporates a sophisticated memory mechanism with gates, enabling effective sequential data processing and contextual information utilization for improved understanding and response generation. Training the model with Stochastic Gradient Descent (SGD) optimization algorithm iteratively fine-tunes parameters to minimize the cost function. This integration of advanced neural network architectures and optimization techniques empowers the chatbot with proficient query resolution, symptom diagnosis, and provision of relevant healthcare information, achieving an accuracy of 96.4%.

In the study by authors of [7], Nisha M.Luniya introduced a Virtual Treatment and Consultation System using Data Mining-symptom analyzer and WEBRTC signaling. Despite lacking appoint schedule algorithms, the system facilitated virtual consultations and appointment services, enhancing accessibility amid pandemic challenges. It aimed to streamline internet applications for healthcare, improving operational efficiency and minimizing errors. The system allowed for efficient appointment management, patient record keeping, and prescription delivery. It prevented duplicate appointments, offered doctors access to schedules and patient information, and provided an administrator section for centralized management.

III. LITERATURE REVIEW

In paper [1] introduces a novel methodology combining Convolutional Neural Networks (CNN), Deep Belief Networks (DBN), and Recurrent Neural Networks (RNN) for disease prediction. Advantages include RNN's text generation capability and DBN's ability to mitigate overfitting, while complexities and the vanishing gradient problem are noted.

In paper [2], Shuxuan Xie, Zengchen Yu, and Zhihan utilize a CNN model for brain tumor detection. This approach boasts efficient image processing and high accuracy. However, challenges include high computational demands and limitations with small datasets. Nonetheless, the method achieves notable success in accurately detecting brain tumors.

In paper [3], Debal, D.A., and Sitote, T.M. employed Random Forest, Support Vector Machine, and Decision Tree methods for disease prediction. Their approach excels in data handling and preprocessing, yet may lack localized context. Nevertheless, the method achieves a commendable accuracy rate of 94.6 percent in disease prediction tasks.

In paper [4], Rabiei R, Ayyoubzadeh SM, Sohrabei S, Esmaeili M, and Atashi A employed Random Forest, Gradient Boosting Trees, Multi-Layer Perceptron, and Genetic Algorithm for disease prediction. Leveraging mammographic features alongside other factors enhanced model performance, although the absence of genetic data was a limitation. Random Forest notably contributed to achieving higher accuracy.

In paper [5], the methodology highlights an advanced integration of Natural Language Processing (NLP) and Machine Learning (ML) algorithms, particularly focusing on recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) networks. By leveraging NLP techniques, the system effectively extracts relevant keywords from user input, enhancing its understanding of complex medical queries

In paper [6], the emphasis is on advanced Machine Learning techniques, particularly Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN), to enhance the hospital management system's AI chatbot. By leveraging LSTM's sophisticated memory mechanism and RNN's sequential data processing, the chatbot efficiently resolves user queries and provides prompt responses, thereby improving overall user experience.

In paper [7], Nisha M.Luniya utilized Data Mining-symptom analyzer and WEBRTC signalling for virtual consultation. Offering convenience and avoiding physical contact, the system lacked appoint schedule algorithms. However, it successfully implemented virtual consultation and appointment services, enhancing accessibility and healthcare delivery amid pandemic challenges.

In paper [8], Venkatesh Rallapalli employed the Waterfall Model for developing an appointment/registration system. The model streamlined scheduling, reducing patient waiting time. However, technical issues could disrupt appointment booking. Nonetheless, the study successfully implemented a basic system, improving patient experience and efficiency in healthcare service delivery.

IV. METHODOLOGY

A. RANDOM FOREST IN PREDICTING HEART DISEASE

Random Forest, a robust machine learning algorithm, predicts heart disease by constructing multiple decision trees. It evaluates feature importance, including age, blood pressure, and cholesterol levels. Employing bootstrap sampling and random feature selection enhances model diversity and prevents overfitting. Each decision tree contributes to the final prediction through a voting mechanism, ensuring robustness and accuracy. By aggregating individual tree predictions, Random Forest effectively identifies heart disease risk factors and aids

in timely intervention. Its ability to handle complex datasets and mitigate overfitting makes it a valuable tool for precise heart disease prediction in healthcare settings.

B. RANDOM FOREST IN PREDICTING KIDNEY DISEASE

Random Forest, a potent machine learning technique, excels in predicting Chronic Kidney Disease (CKD). It leverages ensemble learning, where multiple decision trees are built and their outputs combined for final predictions. By analyzing features like blood pressure and serum creatinine levels, it identifies crucial CKD risk factors. Through bootstrap sampling and random feature selection, it ensures model robustness and prevents overfitting. During prediction, the algorithm aggregates individual tree outputs via a voting mechanism, yielding accurate CKD predictions. Its versatility and ability to handle diverse datasets make Random Forest an invaluable tool for early CKD detection, enabling timely interventions for improved patient outcomes.

C. CNN IN PREDICTING BRAIN TUMOR

Convolutional Neural Networks (CNNs) excel in predicting brain tumors [12] by analyzing MRI or CT scans. Through image analysis, CNNs extract intricate patterns and features indicative of tumor presence. Trained on labeled datasets, they learn to differentiate normal brain tissue from tumor regions. This allows CNNs to provide accurate predictions, aiding in early tumor detection. Their ability to extract and analyze complex features enables precise diagnosis, facilitating timely medical interventions. CNNs play a crucial role in improving patient outcomes by enabling early detection and treatment planning for brain tumors, ultimately enhancing the effectiveness of neuro-oncological care.

D. CNN IN PREDICTING BREAST CANCER

Convolutional Neural Networks (CNNs) are pivotal in predicting breast cancer through mammographic image analysis. CNNs excel in discerning subtle abnormalities indicative of malignancies. Trained on annotated datasets, they extract intricate features from mammograms, including mass shapes and microcalcifications. Through iterative optimization, CNNs learn to differentiate between benign and malignant lesions. Once trained, they accurately predict the presence of breast cancer in new mammograms. Early detection facilitated by CNNs enables timely interventions such as biopsy or further imaging studies, improving patient prognosis. CNNs revolutionize breast cancer screening by enhancing diagnostic accuracy and facilitating prompt treatment, ultimately saving lives through early detection and intervention.

V. PROPOSED METHODOLOGY

Our proposed system introduces a groundbreaking approach to healthcare delivery by offering multiple disease prediction using Machine Learning, augmented by a chatbot interface and a doctor-patient appointment system. By harnessing the power of advanced Machine Learning algorithms, our system provides accurate predictions for specific conditions such as brain tumor, breast cancer, heart disease, and kidney disease, empowering users with proactive healthcare management. The integration of a chatbot facilitates real-time interaction, enabling users to seek medical advice and information effortlessly. Furthermore, the doctor-patient appointment system streamlines access to healthcare professionals, ensuring timely consultations and treatments. This comprehensive approach not only enhances the efficiency of healthcare delivery but also promotes early intervention, ultimately leading to improved patient outcomes and overall well-being.

A. MULTIPLE DISEASE PREDICTION

Our innovative healthcare system not only excels in precise illness prediction but also offers a user-friendly interface for seamless navigation. By incorporating cutting-edge machine learning algorithms such as K-Nearest Neighbors (KNN), Convolutional Neural Networks (CNN), and Random Forest, our system ensures unparalleled accuracy and faster convergence, setting a new standard in healthcare prediction models. This advanced approach enables the effective analysis and interpretation of diverse medical data, allowing for reliable predictions across various diseases [10], including brain tumors and heart disease. The use of CNN is particularly beneficial for detecting intricate patterns in medical imaging data, making it well-suited for tasks like brain tumor detection. Additionally, Random Forest excels in handling complex data structures and interactions, making it ideal for predicting conditions [11] such as heart disease, which often involve multiple risk factors and variables. Moreover, our system's integration of these diverse algorithms enhances its adaptability and scalability, catering to the evolving needs of healthcare consumers and professionals alike. Through providing users with reliable healthcare support and guidance, our system not only fosters early intervention but also promotes a proactive approach to healthcare management, ultimately leading to improved patient outcomes and overall well-being.

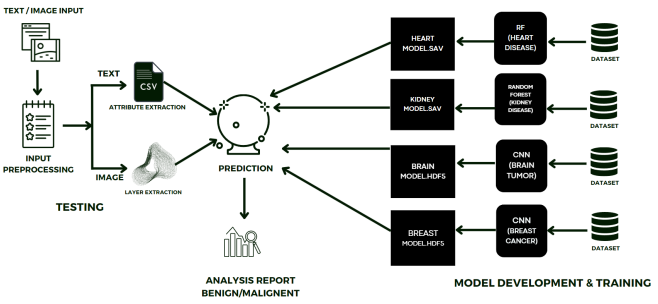
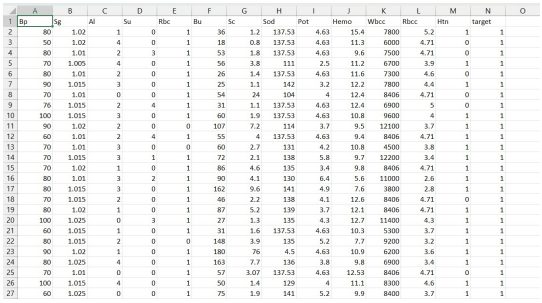


Fig. 1. Architecture of Multiple Disease Prediction System

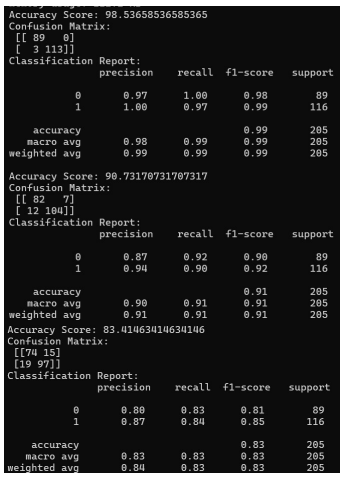
Heart Disease Detection with Random Forest (98.53% accuracy): Leveraging the Random Forest algorithm, our system adeptly analyzes the intricate interplay of various risk factors, including cholesterol levels, blood pressure, and lifestyle habits, in predicting heart disease. Demonstrating an impressive accuracy rate of 98.53%, it furnishes dependable

insights crucial for early intervention and preventive measures. By assimilating multidimensional data encompassing both clinical and lifestyle variables, our system furnishes a comprehensive evaluation of cardiovascular health. This facilitates the implementation of tailored interventions and personalized preventive strategies, thereby enhancing heart health outcomes. Through proactive management and early detection facilitated by Random Forest, our system effectively alleviates the burden of cardiovascular morbidity and mortality. By empowering individuals to prioritize their cardiovascular health, it promotes a proactive approach to healthcare management, contributing to better heart health outcomes and overall well-being. With its ability to provide accurate predictions and personalized recommendations, our system represents a significant advancement in cardiovascular care, enabling individuals to take proactive steps towards improving their heart health and overall quality of life.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Bp	Sg	Al	Su	Rbc	Bu	Sc	Sod	Pot	Hemo	Wbcc	Rbcc	Htn	target	
2		80	1.02	1	0	1	36	1.2	137.53	4.63	15.4	7800	5.2	1	1
3		50	1.02	4	0	1	18	0.8	137.53	4.63	11.3	6000	4.71	0	1
4		80	1.01	2	3	1	53	1.8	137.53	4.63	9.6	7500	4.71	0	1
5		70	1.005	4	0	1	56	3.8	111	2.5	11.2	6700	3.9	1	1
6		80	1.01	2	0	1	26	1.4	137.53	4.63	11.6	7300	4.6	0	1
7		90	1.015	3	0	1	25	1.1	142	3.2	12.2	7800	4.4	1	1
8		70	1.01	0	0	1	54	24	104	4	12.4	8406	4.71	0	1
9		76	1.015	2	4	1	31	1.1	137.53	4.63	12.4	6900	5	0	1
10		100	1.015	3	0	1	60	1.9	137.53	4.63	10.8	9600	4	1	1
11		90	1.02	2	0	0	107	7.2	114	3.7	9.5	12100	3.7	1	1
12		60	1.01	2	4	1	55	4	137.53	4.63	9.4	8406	4.71	1	1
13		70	1.01	3	0	0	60	2.7	131	4.2	10.8	4500	3.8	1	1
14		70	1.015	3	1	1	72	2.1	138	5.8	9.7	12200	3.4	1	1
15		70	1.02	1	0	1	86	4.6	135	3.4	9.8	8406	4.71	1	1
16		80	1.01	3	2	1	90	4.1	130	6.4	5.6	11000	2.6	1	1
17		80	1.015	3	0	1	162	9.6	141	4.9	7.6	3800	2.8	1	1
18		70	1.015	2	0	1	46	2.2	138	4.1	12.6	8406	4.71	0	1
19		80	1.02	1	0	1	87	5.2	139	3.7	12.1	8406	4.71	1	1
20		100	1.025	0	3	1	27	1.3	135	4.3	12.7	11400	4.3	1	1
21		60	1.015	1	0	1	31	1.6	137.53	4.63	10.3	5300	3.7	1	1
22		80	1.015	2	0	0	148	3.9	135	5.2	7.7	9200	3.2	1	1
23		90	1.02	1	0	1	180	76	4.5	4.63	10.9	6200	3.6	1	1
24		80	1.025	4	0	1	163	7.7	136	3.8	9.8	6900	3.4	1	1
25		70	1.01	0	0	1	57	3.07	137.53	4.63	12.53	8406	4.71	0	1
26		100	1.015	4	0	1	50	1.4	129	4	11.1	8300	4.6	1	1
27		60	1.025	0	0	1	75	1.9	141	5.2	9.9	8400	3.7	1	1

Fig. 2. Heart Disease Dataset [17]



```
Accuracy Score: 98.53658536585365
Confusion Matrix:
 [[ 89   0]
 [  3 113]]
Classification Report:
              precision    recall  f1-score   support

           0       0.97      1.00      0.98        89
           1       1.00      0.97      0.99       116

    accuracy                           0.99       205
   macro avg       0.98      0.99      0.99       205
weighted avg       0.99      0.99      0.99       205

Accuracy Score: 90.73170731707317
Confusion Matrix:
 [[ 82   7]
 [ 12 104]]
Classification Report:
              precision    recall  f1-score   support

           0       0.87      0.92      0.90        89
           1       0.94      0.90      0.92       116

    accuracy                           0.91       205
   macro avg       0.90      0.91      0.91       205
weighted avg       0.91      0.91      0.91       205

Accuracy Score: 83.41463414634146
Confusion Matrix:
 [[74 15]
 [19 97]]
Classification Report:
              precision    recall  f1-score   support

           0       0.80      0.83      0.81        89
           1       0.87      0.84      0.85       116

    accuracy                           0.83       205
   macro avg       0.83      0.83      0.83       205
weighted avg       0.84      0.83      0.83       205
```

Fig. 3. Comparison between different algorithm

Brain Tumor Detection with Convolutional Neural Networks(CNN) (98.17% accuracy): By harnessing CNN, our system excels in precise brain tumor detection through the analysis of medical imaging data. It discerns subtle abnormalities in MRI or CT scans, enhancing diagnostic accuracy for early detection and intervention. Achieving an impressive accuracy rate of 98.17%, it empowers healthcare providers to swiftly make treatment decisions, potentially improving patient prognosis and enhancing quality of life. Through its advanced image analysis capabilities, our system revolutionizes brain tumor detection, facilitating timely interventions that optimize patient care and outcomes. Prioritizing accuracy and efficiency in diagnosis, CNN enhances healthcare professionals' ability to deliver prompt and effective treatment, ultimately leading to improved brain tumor management and enhanced patient well-being. This represents a significant advancement in brain tumor diagnosis and care, offering hope for better outcomes and quality of life for individuals affected by this condition. By leveraging CNN's capabilities, our system contributes to the early detection and treatment of brain tumors, providing a critical tool in the fight against this challenging disease.

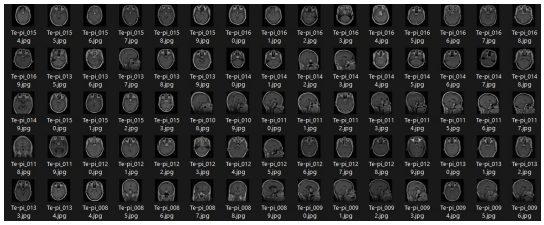


Fig. 4. Brain Tumor Dataset [21]

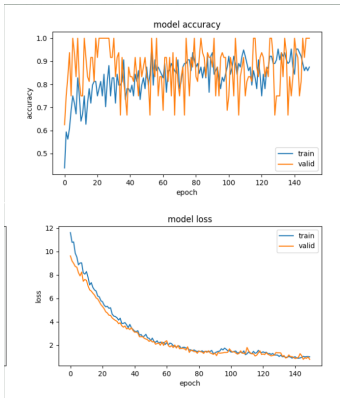


Fig. 5. Graph showing model accuracy and model loss of Brain Tumor

Breast Cancer Detection with CNN (95.13% accuracy): In breast cancer screening, CNN's feature extraction proficiency enhances our system's ability to detect suspicious areas in mammographic images. Achieving 95.13% accuracy, it aids in early breast cancer identification, enabling timely interventions like biopsy. Prioritizing early detection, our system improves breast cancer outcomes and patient survival rates. CNN's precise mammogram analysis revolutionizes screening, facilitating proactive measures for enhanced patient prognosis. Through advanced imaging analysis, our system provides reliable tools for early diagnosis, leading to better breast cancer management and outcomes. Leveraging CNN's capabilities, our system signifies a significant advancement in breast cancer screening, promising improved detection rates and patient care in combating this challenging disease.

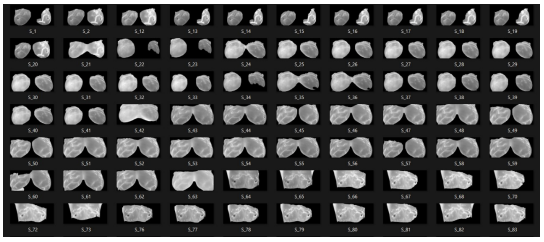


Fig. 6. Breast Cancer Dataset [20]

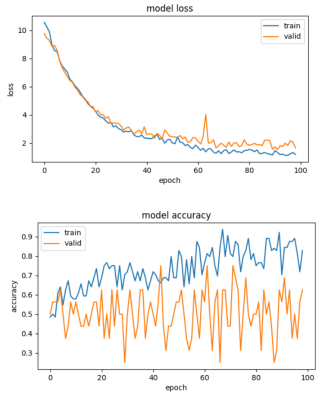


Fig. 7. Graph showing model accuracy and model loss of Breast Cancer

Chronic Kidney Disease (CKD) Prediction with Random Forest(97.25% accuracy): Our system utilizes Random Forest to predict Chronic Kidney Disease (CKD) risk with 97.25% accuracy. It equips healthcare professionals with insights for preventive measures and personalized management. By identifying CKD risk factors early, it mitigates disease burden, improves outcomes, and enhances overall well-being. Through optimized care and timely interventions, it significantly enhances CKD management, reducing complications. The proactive approach to risk factors and tailored interventions facilitates early detection and management, leading to better outcomes and improved quality of life for CKD patients. Leveraging advanced data analysis and predictive modeling, our system signifies a pivotal advancement in healthcare, empowering providers with proactive and personalized care for CKD patients.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target	
2	52	1	0	125	212	0	1	168	0	1	2	2	3	0	
3	53	1	0	140	203	1	0	155	1	3.1	0	0	3	0	
4	70	1	0	145	174	0	1	125	1	2.6	0	0	3	0	
5	61	1	0	148	203	0	1	161	0	0	2	1	3	0	
6	62	0	0	138	294	1	1	106	0	1.9	1	3	2	0	
7	58	0	0	100	248	0	0	122	0	1	1	0	2	1	
8	58	1	0	114	318	0	2	140	0	4.4	0	3	1	0	
9	55	1	0	160	289	0	0	145	1	0.8	1	1	3	0	
10	46	1	0	120	249	0	0	144	0	0.8	2	0	3	0	
11	54	1	0	122	286	0	0	116	1	3.2	1	2	2	0	
12	71	0	0	112	149	0	1	125	0	1.6	1	0	2	1	
13	43	0	0	132	341	1	0	136	1	3	1	0	3	0	
14	34	0	1	118	210	0	1	192	0	0.7	2	0	2	1	
15	51	1	0	140	298	0	1	122	1	4.2	1	3	3	0	
16	52	1	0	128	204	1	1	156	1	1	1	0	0	0	
17	34	0	1	118	210	0	1	192	0	0.7	2	0	2	1	
18	51	0	2	140	308	0	0	142	0	1.5	2	1	2	1	
19	54	1	0	124	266	0	0	109	1	2.2	1	1	3	0	
20	50	0	1	120	244	0	1	162	0	1.1	2	0	2	1	
21	58	1	2	140	211	1	0	165	0	0	2	0	2	1	
22	60	1	2	140	185	0	0	155	0	3	1	0	2	0	
23	67	0	0	106	223	0	1	142	0	0.3	2	2	2	1	
24	45	1	0	104	208	0	0	148	1	3	1	0	2	1	
25	63	0	2	135	252	0	0	172	0	0	2	0	2	1	
26	42	0	2	120	209	0	1	173	0	0	1	0	2	1	
27	61	0	0	145	307	0	0	146	1	1	1	0	3	0	

Fig. 8. CKD Dataset [18]

```
Accuracy Score: 97.5
Confusion Matrix:
 [[29  1]
 [ 1 49]]
Classification Report:
              precision    recall  f1-score   support

           0       0.97      0.97      0.97        30
           1       0.98      0.98      0.98        50

    accuracy                           0.97        80
   macro avg       0.97      0.97      0.97        80
weighted avg       0.97      0.97      0.97        80

Accuracy Score: 96.25
Confusion Matrix:
 [[29  1]
 [ 2 48]]
Classification Report:
              precision    recall  f1-score   support

           0       0.94      0.97      0.95        30
           1       0.98      0.96      0.97        50

    accuracy                           0.96        80
   macro avg       0.96      0.96      0.96        80
weighted avg       0.96      0.96      0.96        80

Accuracy Score: 96.25
Confusion Matrix:
 [[29  1]
 [ 2 48]]
Classification Report:
              precision    recall  f1-score   support

           0       0.94      0.97      0.95        30
           1       0.98      0.96      0.97        50

    accuracy                           0.96        80
   macro avg       0.96      0.96      0.96        80
weighted avg       0.96      0.96      0.96        80
```

Fig. 9. Comparison between different algorithm

B. AI CHATBOT

In our proposed methodology, we integrate advanced Natural Language Processing (NLP) techniques to develop an interactive chatbot feature [9], a crucial component of our healthcare system. Leveraging Long Short-Term Memory (LSTM) networks, known for their sequence modeling capabilities, our chatbot can comprehend and generate contextually relevant responses to user queries, ensuring a seamless conversational experience. Additionally, we harness the power of the Natural Language Toolkit (NLTK), enhancing our chatbot's language processing capabilities for accurate and efficient responses across a wide range of medical inquiries. By combining LSTM and NLTK, our chatbot adeptly addresses user concerns, offering valuable support in navigating healthcare-related questions [14] within a user-friendly interface. This integration of advanced NLP techniques enables us to augment user engagement, enhance accessibility to medical information, and foster proactive healthcare management within our system.

C. DOCTOR PATIENT APPOINTMENT SYSTEM

Our proposed methodology introduces a comprehensive doctor-patient appointment system using PHP, HTML, CSS, and MySQL database [13]. This system offers a user-friendly interface, allowing patients to effortlessly browse appointment slots, choose preferred timings, and book appointments with desired healthcare providers. Integrated with Google Meet for secure video conferencing, it enables convenient remote consultations from patients' homes. Automated appointment reminders and notifications enhance communication efficiency [15], reducing the likelihood of missed appointments. Additionally, administrative functionalities empower healthcare providers to efficiently manage their schedules, update availability in real-time, and handle appointment rescheduling or cancellations with ease. By integrating these technologies and functionalities, our system aims to optimize healthcare delivery, minimize wait times, and enhance patient satisfaction by providing accessible, convenient, and personalized medical

care tailored to the needs of modern healthcare consumers, ultimately improving overall patient outcomes and experiences [16].

VI. RESULTS

At the end of our experiment, distinct algorithmic performances were evident across various medical conditions. For kidney and heart disease diagnosis, Random Forest (RF) exhibited unparalleled accuracy, achieving 97.25% and 98.53% respectively, surpassing Support Vector Machine (SVM), k-Nearest Neighbors (KNN) and Naive Bayes. Conversely, Convolutional Neural Networks (CNN) independently showcased remarkable effectiveness, achieving 98.17% accuracy for brain conditions and 95.13 % for breast conditions. These findings highlight the necessity of tailored algorithm selection for specific medical domains, ensuring enhanced diagnostic precision and improved patient care outcomes.

PERFORMANCE MEASURE OF MODELS

Model	Accuracy (%)	Medical Condition
Random Forest	97.25	Kidney
Random Forest	98.53	Heart
Convolutional Neural Network	98.17	Brain
Convolutional Neural Network	95.13	Breast

VII. CONCLUSION

Through this research we have attempted to evaluate machine learning algorithms and determine a specific person will develop a disease or not given a variety of individual attributes and indicators.Our project's main focus was on examining the accuracy and examining the factors that contribute to the variations in various algorithms.For Heart disease, we have used four databases: Cleveland, Hungary, Switzerland, and Long Beach V. It contains 76 attributes, including the predicted attribute and used percent split to divide the data into two sections which are training and testing datasets. For Brain Tumor, we have used dataset brain tumor classification MRI dataset, consists of two classes, YES for those having brain tumor and NO for those not having brain tumor. For Breast Cancer, Images are divided into two categories : Healthy and Sick . For training, both categories contain 700 MRI scan images of both healthy and sick patients. For validation, both categories contain 40 MRI scan images of both healthy and sick patients. For Kidney Disease, we have used dataset from UCI Repository, the disease is predicted based on certain diagnostic measurements included in the dataset. By the end of the implementation part, we have discovered that Random Forest is giving the maximum accuracy level for both Kidney and Heart Disease which is 97.25% and 98.53% respectively and CNN is giving the maximum accuracy level for Brain and Breast Cancer which is 98.17% and 95.31% respectively. Additionally, if we increase the quantity of training data, there's a chance we'll obtain more accurate results, but processing time will increase and the system's speed will decrease relative to now because it will be dealing more complex and confusing data. Thus, taking into account these possible outcomes, we made this decision, which will be easier for us to work with.

REFERENCES

[1] Anusha Ampavathi, T. Vijaya Saradhi (2021): Multi disease-prediction framework using hybrid deep learning: an optimal prediction model, Computer Methods in Biomechanics and Biomedical Engineering, DOI: 10.1080/10255842.2020.1869726.

[2] Shuxuan Xie, Zengchen Yu and Zhihan Lv(2021), Multi-Disease Prediction Based on Deep Learning: A Survey,CMES,DOI:10.32604/cmes.2021.016728.

[3] Debal, D.A., Sitote, T.M. Chronic kidney disease prediction using machine learning techniques. J Big Data 9, 109 (2022).

[4] Rabiei R, Ayyoubzadeh SM, Sohrabei S, Esmaeili M, Atashi A. Prediction of Breast Cancer using Machine Learning Approaches. J Biomed Phys Eng. 2022 Jun 1;12(3):297-308. doi: 10.31661/jbpe.v0i0.2109-1403. PMID: 35698545; PMCID: PMC9175124.

[5] Chatbot for hospital management using artificial intelligence technology by K.Mannivannan, Laathishavran, G.Gokul, F.Johan 2021 Volume 2, Issue 1:1-81.

[6] Nashif, S., Raihan,Md.R., Islam, Md.R. and Imam, M.H.(2018) Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System. World Journal of Engineering and Technology, 6, 854-873. https://doi.org/10.4236/wjet.2018.64057.

[7] Virtual Treatment and Consultation System Nisha M. Luniya , Rucha R. Sali, Sanjana S. Deshmukh , Arti N. Pandharpure , Prof. Rohit K. Bamane Department of Computer Engineering, Dr. D. Y. Patil Institute of Technology Pimpri, Pune. Savitribai Phule Pune University .

[8] Online Doctor Appointment System Venkatesh Rallapalli, Dipti Menghani, Hema Gallani, Gaytri Aasija Dr. Dashrath Mane Department of Computer Engineering, Vivekanand Education Society's Institute of Technology, Chembur, Mumbai, Marashtra. .

[9] Chatbot for Hospital Management System Srinivasa Rao Dammavalami, Chandana Nukala" Rajeswar Rao Thakkallapallys, LahariI Anegama, Manish Kumar Ravikantis Volume 5, Issue 8, August 2022.

[10] Palle Pramod Reddy, Dirisinala, MadhuBabu,HardeepKumar and Dr.ShiviSharma,Disease Prediction using Machine Learning,Volume9,Issue5May2021—ISSN:2320-2882 .

[11] Assistant Professor Dr.Kishore Raju, Hari Priya, Mohana Supraja, Lakshmi Sowjanya, Sai Pravallika, Multiple disease prediction using machine learning, Volume 10, Issue4.

[12] ZainEldin H, Gamel SA, El-Kenawy EM, Alharbi AH, Khafaga DS, Ibrahim A, Talaat FM. Brain Tumor Detection and Classification Using Deep Learning and Sine-Cosine Fitness Grey Wolf Optimization. Bioengineering (Basel). 2022 Dec 22;10(1):18. doi: 10.3390/bioengineering10010018. PMID: 36671591; PMCID: PMC9854739.

[13] Vijay Ingawat, Dinesh Bartekez, Shrikant Virkar. Sagar Chavan, Smart Hospital Chatbot-Virtual doctor consultation and appointment, Volume 2, Issue 1, February 2022.

[14] A Kumares Angappan, N Jazeem Khan, M Sharan and S Fazil Ahamed ,AI based healthcare chatbot system, Volume 5, Issue March 31, 2021.

[15] Luona Yin , Aiqing Zhang, Xinrong Ye and Xiaojuan, Security-Aware Department Matching and Doctor Searching for Online Appointment Registration System , Wuhu 241002, China 3Anhui Provincial Engineering Laboratory on Information Fusion and Control of Intelligent Robot, Wuhu 241002, China.

[16] E-Health Appointment Solution, A Web based approach Cristian Cola1 , Affiliation 2: Technical University of Cluj Napoca 26-28, Baritiu St., 3400, Cluj-Napoca.

[17] Andras Janosi, M.D, William Steinbrunn, M.D, Matthias Pfisterer ,M.D, Robert Detrano, M.D., Ph.D., https://archive.ics.uci.edu/ml/datasets/Heart+Disease: year=2023.

[18] Rubini L, Soundarapandian P, and Eswaran P (2015). Chronic Kidney Disease. UCI Machine Learning Repository. https://doi.org/10.24432/C5G020.

[19] https://www.researchgate.net/publication/350936156-Brain-Tumor - Detection - using - Deep-Learning-and-Image-Processing

[20] https://www.researchgate.net/publication/323644876-Breast - cancer - detection - using - convolutional-neural-networks-for-mammogram-imaging-system