International Journal of Advanced Technology in Engineering and Science
Vol. No. 10, Issue No. 07, July 2022
www.ijates.com
ijates
ISSN 2348 - 7550

An Interactive Medical Chatbot for the prediction of diseases using Hidden Markov Models

Mahesh Kamble, Guided by Prof. Nagaraju Bogiri

Mahesh Kamble, Dept of Computer Engineering, K J College of Engineering and Management Research, India

Prof. Nagaraju Bogiri, Dept of Computer Engineering, K J College of Engineering and Management Research, India

Abstract

One of the most underrated values of leading a fulfilling lifestyle is having a healthy body and mind. There are medical organizations and institutions that have been focused on research in making the people need a healthy life and keep them free of diseases and other ailments. There has been countless amounts of research that have been performed over centuries which has led to a significant improvement in the medical platform which we see nowadays. What do you do the fast page lifestyle nowadays images cities across the world this has led to a lot of individuals not getting enough time to have regular checkups and other interactions with their doctors. With the pandemic loom enlarge there are a lot of individuals that are scared going to the hospital which leads to a lot of symptoms and other deadly diseases going unchecked. Therefore it is suggested to improve the current diagnostic procedures through the implementation of technology leading towards remote diagnosis through the use of chatbots. For this purpose the proposed methodology in this research article outlines and effective medical chatbot that it implements deep learning for the purpose of disease identification and remedy suggestion. The presented approach utilizes Pearson Correlation, K Nearest Neighbor and Linear Regression along with Hidden Markov Models and Decision Tree to achieve accurate predictions for diseases and remedial suggestions. Extensive experiment analysis of the proposed approach has resulted in highly satisfactory outcomes indicating precise prediction capability.

Keywords: *Decision Tree, Hidden Markov Model, K nearest Neighbors, Linear Regression, Pearson Correlation.*

1. INTRODUCTION

The healthcare sector is currently one of the most sophisticated and rapidly advancing paradigms in the modern world. However, certain aspects continue to be observed as a direct result of various factors. With the process of urbanization as well as the lack of professional balance, there has been an increase in illnesses all over the world. This way of life causes distress and a complete disregard for one's health. The health-care system is one of the motivating factors that may influence a group's capability to make use of some of the resources available.

Medical management could be extremely beneficial in a multitude of settings, remarkably efficient technological innovations and a substantial improvement in rising living standards.

As a result, medical enterprises and the biomedical sector have been identified as among the most essential and significant subjects in the discipline. Large-scale developments have generally concentrated on medical advancements that have been demonstrated to be successful in expanding and improving human average lifespan. Only through consistently sustaining human health and reducing severely destructive diseases and major disorders has this been done. A person's health is vitally important to their general well-being and the enhancement of their day-to-day activities such as work. Being healthy helps a person to operate to their full ability, which may lead to a more fulfilling experience.

Maintaining one's health is a difficult and conscientious activity that necessitates a highly healthy lifestyle. However, owing to the operations of the contemporary world, a significant portion of the population does not live a particularly healthy life and has a standard of living which does not permit for proper health management [1]. This has resulted in a significant increase in diseases and other afflictions that may have been avoided by living a healthier lifestyle. The developments were and still are extremely beneficial in sustaining human health and welfare in the face of highly dangerous illnesses.

Because of the emergence of the digital platform, which has made important contributions to the literature of study, scientists can now assist learning and cooperate on innumerable innovative medicines and other preventative measures in a very short period of time. The idea that perhaps the World Wide Web has elicited an essentially omnipresent reaction from consumers, with the vast proportion of the population linked to this worldwide network infrastructure, has far-reaching ramifications. Because the communications infrastructure can be utilized for efficient communication, this link has encouraged more participation among individuals. As a result, a variety of chatbots as well as other electronic medical support portals have been developed to assist the general public with their health concerns.

With so many people who suffer from persistent illnesses all over the world, the healthcare practice is under increasing pressure to provide adequate health care to this rising population. Medical practitioners have been under intense pressure to diagnose and treat these patients as soon as feasible [2]. Specialists are under a huge amount of pressure because of their heavy workload, which can lead to a lot of user error and other problems. Add to it the fact that practically of people with physical disabilities cannot travel dozens of kilometers to their physician for a conclusive diagnosis and preventative care. As a consequence, this method has been valuable in understanding previous attempts for healthcare chatbot implementations, including in achieving our plan, which makes use of deep learning deployments.

The approach deploys the use of Pearson correlation for extracting the correlation from the user input into the system as well as the datasets that are being used for the purpose of achieving the prediction of the diseases. The resultant correlations are effectively clustered using the K Nearest Neighbors, the clusters are then effectively utilized for the regression analysis through the use of Linear Regression. This resultant regression list is then

provided to the Hidden Markov Model for the purpose of achieving the probability scores for the disease prediction. The obtained values of the achieved probability scores need to be classified, for this purpose the decision tree classifier is utilized which effectively classifies the predictions and provides the suggestion to the patient. This approach is further elaborated in the next sections of this research paper.

This research paper dedicates section 2 for analysis of past work as literature survey, section 3 deeply elaborates the proposed technique and whereas section 4 evaluates the performance of the system and finally section 5 concludes the paper with traces of future enhancement.

2. RELATED WORKS

Shaikh Farhad Hossain [3] explores that almost all of the prescription drugs as well as other appropriate compounds that are implemented for the goal of accomplishing rehabilitation of some of these ailments have indeed been obtained from plant derived source materials. These components are immensely important because they can easily help the patient with severely compromised body systems. The majority of these plants are hugely beneficial and are cultivated in various parts of the world. The plants provide a variety of medicines as well as other appropriate excipients for a variety of prescription drugs in the Unani methodology. The majority of these plants are not grown with innovative improvement engagement for the purpose of predicting disease outbreaks in these crops. As a result, the authors propose using the DPclusO algorithm to effectively predict plant disease relationships.

Prakhar Srivastava [4] establishes the concept of using an automated chatbot system to improve the healthcare and medical platform. Most medical approaches have made significant progress in recent years. This has contributed in more accurate diagnosis and better treatments, which can be extremely beneficial in providing better healthcare to a huge number of participants. This results in a considerable advances that are presented to individuals, which can be extremely beneficial in significantly lowering their pain and misery. However, as the population has grown, there seems to be an increase in hospital overcrowding, as well as an increase in patients who may find it difficult to travel long distances to the hospital. As a result, the authors of this methodology have provided an important healthcare chatbot that is structured to efficiently make a diagnosis patients' manifestations and provide therapeutic compounds.

Sai Zhang [5] elaborates that there is now a noticeable difference in the healthcare framework in the current history. These advancements have been extremely beneficial in achieving effective advancements in the healthcare framework for the purposes of disease diagnosis and treatment. These techniques have proven to be effective in realizing useful implementations that have the potential to dramatically enhance the conceptual framework of emergency treatment and its applicability around the world. As a result, the authors of this approach proposed and effectively used MiRNA for the intention of proximity-based prediction of higher-order associations.

Sayali Ambekar [6] articulates the definition of health care and maintaining a healthy lifestyle as an extremely valuable and interesting aspect of living a fulfilling life. There is a need for the effective implementation of a medical approach that really can minimize disease incidence while also effectively treating a specific ailment. A wide range of approaches have been developed to aid in the screening and prevention of degenerative illnesses. This is due to a dramatic increase in the percentage of persistent illnesses around the world, which have been incredibly painful and excruciating for patients. As a result, the author suggested a disease risk reduction methodology that accurately simulates the ailment to accomplish extremely effective results. The authors proposed using a convolutional algorithm to determine heart disease effectively.

Anjan Nikhil Repaka [7] introduces the phenomenon of medical groups, which have been providing consistent experience to the health care system sector for many years. These institutions have been extremely efficient in implementing their stated goals of effective management and therapy of a wide range of diseases. These diseases have claimed many lives and have different ways of interacting with different people's bodies. This results in a variety of diverse specifications for the goal of accomplishing an accurate diagnosis, effective treatment, and pain relief for the individual. As a result, the authors proposed the implementation of an effective technique that employs the naive Bayes algorithm for reasonably precise predicting heart disease.

Pahulpreet Singh Kohli [8] describes that what a lot of professional care regions such as hospitals or medical organizations have already been attempting to deal with a massive influx of maladies and massively complicated care recipients with comorbidities. These clients are in an amount of discomfort and have maladies that necessitate a thorough diagnosis and appropriate therapeutic interventions in a short period of time. The diagnosis of these chronic diseases necessitates extensive testing and the elimination of numerous variables. This procedure can take a long time, causing serious difficulties for the participant. As a result, the authors presented an improved usage deep learning for disease prediction. The authors used this methodology for the Pima Indians diabetes data source, as well as breast cancer detection and prediction as well as heart disease monitoring and forecasting.

Dhiraj Dahiwade [9] analyzes how promotion of sustainable in the technology industry have empowered impactful significant improvement in a variety of methodologies across multiple frameworks. The pace of innovation has drastically enhanced the standard of living and also the work performed that are accomplished every day. These technological advancements have enabled a more important and successful moment of realization of an increase in user convenience. This also increases the ease of access for many people, making their lives more fulfilling. As a result, there is a need to incorporate technology into disease diagnosis and health care centers that can be used for impactful forecasting through the use of classification techniques.

Mohammed Jawwad Ali Junaid [10] presents the idea of heart disease diagnosis among the most interesting and valuable constituents for further developing an effective medical framework. Because of their inherent complexity, heart diseases are extremely difficult to comprehend and diagnose. These diseases are not detected

until it is too late, making it highly likely that they will cause significant damage without first being treated. Heart disease diagnosis is extremely limited because it can contribute to higher troublesome circumstances which can be comprehensive and fatal. The authors of this methodology have presented an improved methodology for predicting heart disease using data science for this intention. For their prediction methodology, the authors proposed using an artificial neural network, support vector machines, and a hybrid naive Bayes algorithm.

Nudtaporn Rosruen [11] describes that now the affordable healthcare and appropriate care and management of ailments is amongst the most basic need of the human being. Illnesses and other diseases can be extremely painful and cause a great deal of suffering for the individual. Throughout the world, medical institutes and healthcare approaches have attempted to augment the difficulties related to various ailments of the patient through in-depth prognosis and individualized treatments. However, most of these patients may or may not be able to easily access this type of healthcare, which can result in increased in their difficulties and appropriate pain, causing a slew of issues. To improve this approach and make it more accessible to a larger number of people, the researchers suggested a healthcare consultant framework that uses a chatbot to collect symptoms from the participant and operate natural language processing to identify and diagnose any disease.

Dipesh Kadariya [12] affirms that the bronchitis illness was among the most controversial and downright dangerous ailments that can be immensely potentially lethal. Conventional methods to asthma identification and diagnosis are highly archaic and require marked enhancement to accomplish their intended results. Traditional techniques are less efficacious in decreasing effective asthma patient monitoring. The researchers but rather suggested an efficacious chatbot that is individualized for the intention of asthma management in children for this reason.

Lekha Athota [13] expresses that a virtual assistant is a powerful resource for discussion. Using this approach, an implementation can be created that provides high-quality explanations in a small amount of time. It relieves the answer supplier of the responsibility by straightforwardly providing the response to the consumer via an intelligent system. The project was created for the consumer that would save them time when consulting practitioners or specialists for medical technology. In this application, we used the N-gram and TF-IDF to extract the search term out from user request. Each important word is decided to weigh in order to find the best answer to the query. The Web Framework is designed for users to enter queries. The application is enhanced with confidentiality and relevance improvements that ensure promote resilience and images while retrieving answers to questions.

Rohit Binu Mathew [14] discusses that the healthcare staff and medical facilities are almost always used with the intention of diagnosis and treatment of diseases and appropriate therapy of the other. The average person typically employs these methodologies for this purpose, such as maintaining health and efficient and effective manner comprehension health status. There is an increment in the patients that come at medical institutes owing

to an appropriate increment in the amount of individual citizens and it's not a substantial spike in the proportion of clinics and other health organizations. This causes overcrowding in the systems, which can then be improved dramatically with today's technological advancements. The authors developed an appropriate chatbot for medication and pathogen diagnostics based on neural network models for this intention.

3. PROPOSED METHODOLOGY

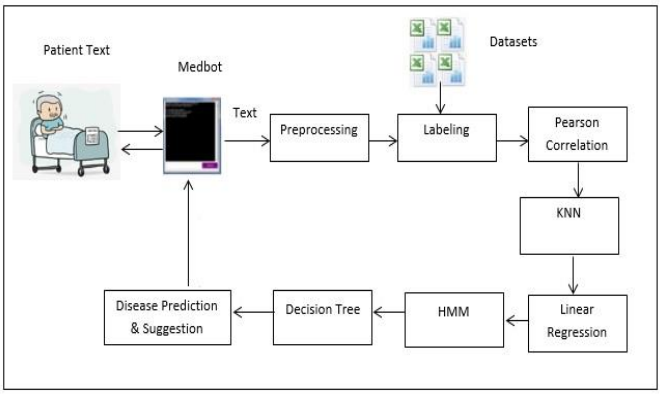


Figure 1: Proposed model System Overview

The proposed methodology for a medical chatbot that performs disease detection and suggestion using deep learning has been described in the system overview given in figure 1. The steps utilized for achieving the presented technique has been detailed in the steps given below.

Step 1: Dataset collection, preprocessing and Labeling – The system for the medbot comprising of disease prediction and suggestion requires an input of 3 datasets, consisting of kidney disease, Covid-19, and Heart Disease. The kidney disease dataset is downloaded from the URL - https://www.kaggle.com/mansoordaku/ckdisease, the Covid-19 dataset from URL - https://data.gov.il/dataset/covid-19, and Heart Disease from URL -https://www.kaggle.com/ronitf/heart-disease-uci .

The datasets are extracted and provided to the presented technique as an input which initially performs the labeling procedure. This is performed by reading the datasets in the form of a double dimension list through the use of the JXL API. This process indexes the dataset which converts it into a labeled dataset. The user input is also grabbed using the interactive user interface that is designed in the Java programming language using the Swings Framework. This user interface prompts the user with the most common symptom of the 3 diseases. Once the user selects this disease the parameters related to the particular disease are selected from the patient as an input and preprocesses it before providing it to the system for further processing.

Step 2: Pearson Correlation – The outputs achieved in the previous step have been taken as an input into this step for the purpose of achieving the correlation between the two entities. The two entities in this approach are the attributes of the dataset and the user input. The correlation between these two entities is calculated using the Pearson Correlation approach. The Pearson Correlation allows for the realization of the correlation coefficient

which is achieved for each of the rows of the dataset. This results in a correlation list which is then provided to the next step for the clustering. The equation for the Pearson correlation has been given in the equation 1 below:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \text{ ------- (1)}$$

Step 3: K Nearest Neighbor Clustering – The labeled list and the user input is provided as an input to this step of the system for performing the clustering operation. The 3 datasets, namely, kidney disease, Covid-19 and heart disease, are in the form of a double dimension list. The K -Nearest Neighbors are used to perform the segregation of input data into clusters which are useful in determining the semantic groups. The clusters are obtained through the use of the following steps.

Distance Evaluation – The Euclidean Distance is used to determine the distance in the selected attributes of the input double dimension list in comparison with the user input for the same. The distance evaluated for the selected attributes is appended to the end of the particular row as the row distance R_D. This is performed for all the rows in the list and the respective row distances are appended accurately through the evaluation using the equation 1 given below. These row distances are also subjected for the average row distance evaluation which is then stored appropriately.

ED=√([[∑(ATi-ATj)]]^2) _____ (1)

Where,

ED=Euclidian Distance

ATi=Attribute at index i

ATj= Attribute at index j

Centroid Estimation – The output from the previous step of distance calculation is provided as an input here. The list containing the distances added to the end of the rows is used for the purpose of centroid estimation. For achieving the centroid, the list is first sorted into the ascending order of the row distances. This sorted list is subjected to data point selection randomly. These data points are nothing but row distances that are k in number. These row distances are then used to determine the boundaries through the use of the Average row distance acquired previously.

The selected row distances from the data points and the average row distance is then used to form the minimum and maximum values by addition of both the values and subtraction of the same respectively. These boundaries are highly useful for the formation of the clusters in the next step.

Cluster Formation – The k boundaries attained in the previous step are used as a major aspect in this step for the purpose of cluster formation. The row distances in the double dimension list are subjected to scrutiny based on the boundaries attained in the previous step. The clusters are then formed by the row distances that abide by these boundaries which are then stored as a cluster list and transferred to the next step of the system. The entire process can be illustrated through the algorithm 1 given below.

ALGORITHM 1: KNN Classified Cluster Formation

//Input : Sorted Distance List $SD_{L,}$

//Output:Cluster List K_{CL}

1: ***Start***

2: $I_L = \emptyset$ [Inner Layer] $O_L = \emptyset$ [Outer Layer], $K_{CL} = \emptyset$

3: MIN= 0 , MAX=SDL_{SIZE-1}

4: K= (MAX-MIN) /2

5: K=MIN+K

6: ***for*** i=0 to Size of SD_L

7: R= $SD_{L\,[i]}$

8: ***if***(i<=K), ***then***

9: $I_{L=}\ I_{L+R}$

10: ***else***

11: $O_L = O_{L+R}$

12: ***end for***

13: $K_{CL[0]} = I_L$

14: $K_{CL[1]} = O_L$

15: return K_{CL}

16: ***Stop***

Step 4: Linear Regression –The linear regression procedure performs the regression between the user input and the cluster attributes that are formed in the previous step. These values are provided as input to this step for the regression analysis through the use of Linear Regression. The regression analysis performed through the linear regression determines the change between two different variables and quantifies it. The lists are the x [] and y[ ] out of which the x [] is the independent list and the y []is the dependent list. The equation for the same is given in equation 1 below.

$Y=Mx + B$ _____ (2)

The regression is measured through the equation given above for which the values of the slope given as m and the value of the intercept given as b are unknown. These values are achieved by the evaluation of the equation 3 and 4 given below. The values of x [ ] in these are the user attributes and the value Y [ ] is the clusters values achieved in the previous step. These values are added to the equation to achieve the required values of m and b.

_____ (3)

_____ (4)

Where:

x = Independent variable

y = Dependent variable

M = Slope or Gradient

B = the Y Intercept

N= Size of the array

Y=Intercept value

The attained values of m and b from the above calculations are then used in the equation 2 above to attain the dependent variable values. Here an independent value from X[ ] is used in equation 2, and its obtained values are averaged to get the mean regression value for each of the rows in the cluster. The regression of the values of x [] and y[ ] allow for a greater understanding of the relationship between the two variables. These are the regression values from the linear regression analysis which are then aggregated in the form of a list which is given to the next step as an input.

Step 5: Hidden Markov Model – The Hidden Markov Model is highly effective and useful model for the purpose of identification and detection based problems. The Hidden Markov Model utilizes time series for the purpose of detection of the hidden or the unobservable states. For providing the input to this step of the methodology the clusters achieved in the previous steps are being utilized as an input to this step of the procedure.

The various attributes and collection of values relating to these attributes along with the regression list is taken as an input and added to a double dimension array list along with the corresponding user information entered by the patient. These values are being used for the purpose of extracting the probability of the instance of a particular disease out of the three diseases being identified in the presented approach.

In this type of the approach the probability is realized through the effective calculation of the time taken for the attributes to transform from one state to another. This procedure for the calculation of the probability values is performed recursively for each of the rows of all the attributes taken as an input. The termination equation is given by the equation 5 given below.

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i) \text{--- (5)}$$

Where,

$P(O|\lambda)$ is the observational probability of the sequence O based on the λ (Hidden Markov Model) for a summation of all the variables in time T.

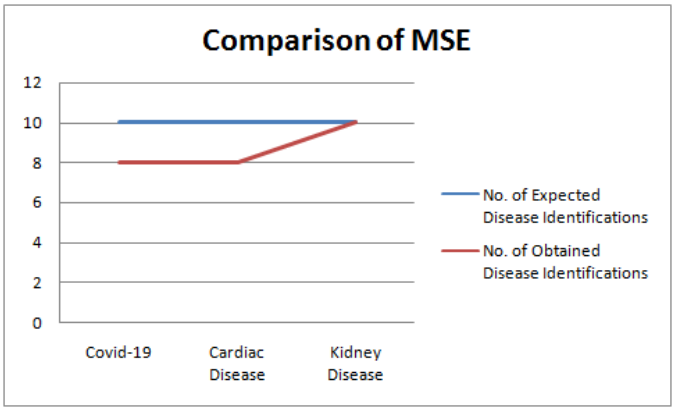
The achieved values of the probabilities are utilized by adding all of these values into a list that is then sorted in the descending order of the probability scores. This list is effectively shorten by selection of 50% of the data size and then provided to the next step of the procedure.

Step 6: Decision Tree – The probability values generated from the Hidden Markov Model stage of the technique has taken as an input in this step of the approach. This step involved the categorization of probability values using the decision tree methodology. This categorization is accomplished through the application of if-then rules, which properly select the right result for disease diagnosis and recommendation realization. The categorization can also help to limit the amount of false positives in the system. The system's ideas are utilized to populate the medbot user interface.

4. RESULTS AND DISCUSSIONS

The proposed methodology for medbot a medical chatbot for disease identification and recommendation by deep learning has been demonstrated in the Java programming language. To make the technique more effectively executed, the NetBeans IDE has indeed been utilized. The graphical user interface was created using the Swings Framework. The system was implemented on a developer computer with an Intel Core i5 CPU, 8 GB of RAM, and 1 TB of hard disk space as its setup.

The experimental evaluation of the approach is an essential necessity to determine the performance of the presented system. This determines the amount of error achieved by the approach as well as stipulates if the deep learning methodologies have been accurately implemented in the system properly. For the purpose of evaluation, the RMSE evaluation mechanism is being utilized. The system provides appropriate detection of the error achieved by the methodology for the detection or identification of the disease.



Performance Evaluation through Root Mean Square Approach

Several tests were conducted to establish the error produced by the suggested approach, the process for disease detection employing Convolutional neural networks. The decreased accuracy reached by the approaches attributable to the Disease Detection component's predisposition for error may be utilized to define the predefined threshold.

The Root Mean Square Error, or RMSE, is used to calculate the error caused by the specified approach. The presence of any type of imprecision in the suggested strategy for disease identification through HMM indicates the identification accuracy of the proposed methodology. The RMSE approach makes it easier to calculate

errors between two continuously connected parameters. The metrics tested in this technique are the expected disease identification and the achieved disease identification. Equation 6 is used to compute the error estimates.

_ (6)

Where,

$\sum$ - Summation

$(x_1 - x_2)^2$ - Differences Squared for the summation in between the expected disease identification and the achieved disease identification counts.

n - Number of Trails

These two properties were measured on 3 distinct diseases based on the user inputs on the medical chatbot standalone application. The outcomes of these assessments are depicted in table 1 below.

Disease	No. of Expected Disease Identifications	No. of Obtained Disease Identifications	MSE
Covid-19	10	8	4
Cardiac Disease	10	8	4
Kidney Disease	10	10	0

Table 1: Mean Square Error measurement

Figure 3: Comparison of MSE in No. of expected disease identification and the No. of achieved disease identification

The experimental analysis of the concept's outputs for 10 trials for each disease have made it simpler to comprehend the error rate visually, as illustrated in figure 3. The graph shows the system's error rate in guessing the Disease that the patient is suffering from.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_{1,i} - x_{2,i})^2}{n}}$$

The RMSE scores analyzed for disease identification confirmed the efficacy of the proposed approach in considerable detail. The suggested method outperformed the disease detection approach powered by LSTM published in [15]. Our method achieves an RMSE of 1.632 Table 2 below shows a tabular comparison of the LSTM approach of [15] and the proposed HMM technique.

Performance Metric	Our Approach (HMM)	LSTM based Disease Identification approach [15]
RMSE	1.632	3.657
MAE	2.667	2.496

Table 2: Comparison with [15]

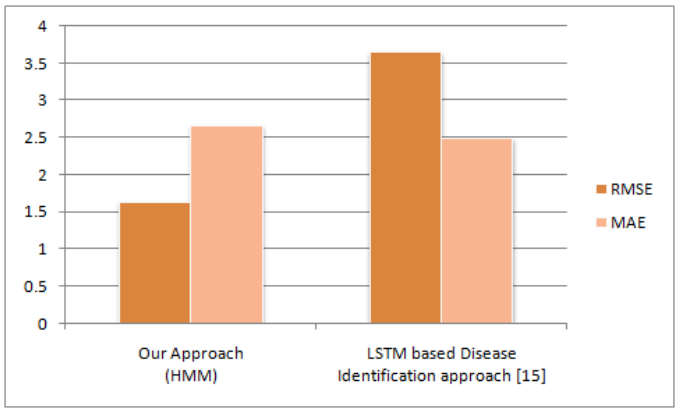


Figure 4: Comparison with LSTM based Disease identification technique in [15]

The deep learning algorithm suggested in this research work dependably surpasses the LSTM based disease identification approach described in [15], as seen in Figure 4. This is owing to the Hidden Markov Model's extremely accurate implementation, which dramatically enhances recognition capability. The Decision tree technique further improves the outcomes, as evident by MSE and RMSE values of 2.667 and 1.632, respectively. This evaluation reveals the proper and effective implementation of HMM for the disease identification on the proposed MedBot: A Medical chatbot.

5. CONCLUSION AND FUTURESCOPE

The implementation of a medical chatbot that provides disease diagnosis and useful suggestions to the particular condition has contributed in the process of productive enhancement in the healthcare framework. The symptoms supplied by the user are used as input by the system. These symptoms are efficiently preprocessed to provide a light-weight query through the user inputs provided to the chatbot that may be readily processed in order to obtain the diagnosis. The query is then labeled using a variety of datasets comprising illnesses and symptoms. This labeling enables the system to efficiently restrict the disease range. The tagged text is then sent via Pearson Correlation to generate the correlation, followed by K Nearest Neighbors to generate appropriate clusters. The obtained clusters are provided to the next step of the approach that extracts the regression of the input symptoms through the use of Linear Regression. The regression list is then utilized for the subsequent step that performs the evaluation of the disease and provides prediction through the use of Hidden Markov Model. The Hidden Markov Model provides the Disease Prediction and suggestion which is then effectively classified using if-then rules of the Decision Tree Module. The results are then displayed to the user through the graphical user interface. The approach is effectively evaluated for its performance through RMSE and the resultant error is compared with the predominant disease identification mechanism that has proven the superiority of the proposed technique.

For the enhancement of this approach in the future, this technique can be realized in the form of a mobile application that is much more accessible to healthcare professionals and the patients.

REFERENCES

[1]N. Siangchin et al, "Chatbot Implementation for ICD-10 Recommendation System", International Conference on Engineering, Science, and Industrial Applications (ICESI), 2019.

[2]T. -H. Hwang, J. Lee, S. -M. Hyun and K. Lee, "Implementation of interactive healthcare advisor model using chatbot and visualization," 2020 International Conference on Information and Communication Technology Convergence (ICTC), 2020, pp. 452-455, doi: 10.1109/ICTC49870.2020.9289621.

[3] S. Hossain et al, "Prediction of Plant-Disease Relations Based on Unani Formulas by Network Analysis", IEEE 18th International Conference on Bioinformatics and Bioengineering, 2018.

[4]P. Srivastava and N. Singh, "Automatized Medical Chatbot (Medibot)," 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC), 2020, pp. 351-354, doi: 10.1109/PARC49193.2020.236624.

[5]S. Zhang et al, "Higher-Order Proximity-based MiRNA-Disease Associations Prediction", ACM Transactions on Computational Biology and Bioinformatics, 2020.

[6]S. Ambekar et al, "Disease Risk Prediction by Using Convolutional Neural Network", Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018.

[7]A. Repaka et al, "Design and Implementing Heart Disease Prediction Using Naive Bayesian", Proceedings of the Third International Conference on Trends in Electronics and Informatics, ICOEI 2019.

[8]P. Kohli et al, "Application of Machine Learning in Disease Prediction", 4th International Conference on Computing Communication and Automation (ICCCA), 2018.

[9]D. Dahiwade et al, "Designing Disease Prediction Model Using Machine Learning Approach", Proceedings of the Third International Conference on Computing Methodologies and Communication, ICCMC 2019.

[10]M. Junaid et al, "Data Science and Its Application in Heart Disease Prediction", International Conference on Intelligent Engineering and Management (ICIEM), 2020.

[11]N. Rosruen and T. Samanchuen, "Chatbot Utilization for Medical Consultant System," 2018 3rd Technology Innovation Management and Engineering Science International Conference (TIMES-iCON), 2018, pp. 1-5, doi: 10.1109/TIMES-iCON.2018.8621678.

[12]D. Kadariya et al, "kBot: Knowledge-enabled Personalized Chatbot for Asthma Self-Management", IEEE International Conference on Smart Computing (SMARTCOMP), 2019.

[13] L. Athota, V. K. Shukla, N. Pandey and A. Rana, "Chatbot for Healthcare System Using Artificial Intelligence," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020, pp. 619-622, doi: 10.1109/ICRITO48877.2020.9197833.

[14]R. Mathew et al, "Chatbot for Disease Prediction and Treatment Recommendation using Machine Learning", Proceedings of the Third International Conference on Trends in Electronics and Informatics, ICOEI 2019.

[15] H. D. Park, Y. Han and J. H. Choi, "Frequency-Aware Attention based LSTM Networks for Cardiovascular Disease," 2018 International Conference on Information and Communication Technology Convergence (ICTC), 2018, pp. 1503-1505, doi: 10.1109/ICTC.2018.8539509.