An Implementation of Machine Learning Based Healthcare Chabot for Disease Prediction (MIBOT)

Gaurav Ganguly
Department of Computer Science & Engineering
University of Engineering & Management Jaipur
Jaipur, India.
gauravgngly@gmail.com

Sourav Kumar
Department of Computer Science & Engineering
University of Engineering & Management Jaipur
Jaipur, India.
sourav.kumar03021996@gmail.com

2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) | 979-8-3503-0611-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCCIS60361.2023.10425125

***Abstract*— In an era where quality healthcare is pivotal for a successful life journey, challenges arise when seeking medical attention amid the ongoing pandemic. The traditional approach of consulting doctors in person or visiting hospitals is hindered, emphasizing the necessity for innovative solutions. Harnessing the capabilities of Natural Language Processing (NLP) and machine learning, this study introduces an advanced chatbot application. Built upon supervised machine learning, the proposed chatbot system not only offers disease diagnosis and comprehensive treatment insights but also does so preemptively, prior to consulting a physician. This system boasts a user-friendly GUI-based text assistant, facilitating seamless interaction with the chatbot. Unlike existing models, this approach intelligently presents user symptoms and corresponding risk factors, tailored to the user's ailment, while also offering optimal recommendations. Moreover, the chatbot clarifies the appropriate instances for in-person medical consultations. By addressing the underutilization highlighted in previous research, this study signifies the chatbot's potential advantages. By embracing this cost-free application, individuals can circumvent cumbersome hospital visits. The comparison with the reference model demonstrates the enhanced efficacy of the proposed approach in addressing healthcare challenges, ensuring convenient and accessible medical guidance.**

Keywords*— *Chatbot; Healthcare; Consultant System; Artificial Intelligence; Machine Learning; NLP; TF-IDF

I. INTRODUCTION

Most people are unaware the increasing health information from internet. When people search for information on health, they are influenced by a variety of factors. In their busy life schedule, it becomes very difficult for people to be aware and careful about their health issues. Most working-class people claim that their hectic schedules do not allow them to consult their physician on a regular basis and they ignore any discomfort they feel until it becomes too intense. The use of reputable medical information including diseases, symptoms and treatments is important before visiting a doctor or a medical center or shop for assistant with a common illness. However, less computer knowledge of users leads to access difficulties. A number of health applications like “Doctor Me”, “MedBot”, and “MedChat” are available to inform all. Even so, multiple steps are required to reach desired information.

This proposed system involves the development of an intelligent agent (MIBOT) to facilitate the interaction between users and a chatbot that returns the diagnosis of different diseases based on various symptoms that users provide. Through the Chabot’s interface, this system is capable of detecting symptoms by providing input from the user. Using NLP [1] automatic translation the text and classify text into categories after these extracted symptoms, chatbot diagnoses and prescribes a treatment for the disease. Instant Messaging (IM) applications can easily accommodate the chatbot. SVM [2] classifying is a machine learning algorithm [3] that determines decision boundaries in the range of the problem by using hyper-planes that define decision boundaries. A neural network is an example of support vector machines (SVM’s) [4]. The use of neural networks (NNs) and Bayesian networks can be used in static imperfection prediction. The TF-IDF [5] is vectorized and the cosine similarity measure [6] is used to generate similarities between texts. Medical Chatbots can diagnose patients with simple symptom analysis, proving that they can somewhat accurately diagnose patients and natural language processing is used to create a conversational approach.

Medical chatbots have a substantial impact on healthcare culture, enhancing reliability while reducing human errors. Today, individuals often prioritize online platforms over their health, leading to untreated mild illnesses that can escalate into major diseases. To address this, we propose a free, accessible chatbot solution. The user-friendly nature and widespread accessibility of chatbots lower the cost of consulting expert healthcare professionals. This approach aims to raise awareness about health maintenance and encourage timely action. By simplifying the healthcare process, individuals can easily access chatbots, ensuring minimal disruption to their daily routines. This increased health consciousness and accessibility have the potential to benefit both individuals and the healthcare industry.

The paper is structured as follows: Part II includes a literature review. Part III explains the proposed method with examples, a demonstration of the algorithm, and a system diagram. In Part IV of the article, performance and experimental findings are illustrated. In Part V, a performance analysis is offered, and in Part VI, the report is concluded with a discussion of the shortcomings of our suggested system and future plans.

II. LITERATURE REVIEW

Innovations in healthcare and education have been sparked by the emergence of artificial intelligence (AI) and natural

ISBN: 979-8-3503-0611-8/23/$31.00 ©2023 IEEE

language processing (NLP) technology. Various studies have explored the potential of chatbots, powered by AI, to revolutionize medical consultation, education, and healthcare management.

The work by Athota et al. [7] emphasizes the significance of healthcare accessibility and cost reduction. Their chatbot utilizes AI to diagnose diseases and offer preliminary medical information, aiding users before consulting a doctor. This aligns with the broader trend of AI-driven healthcare systems enhancing patient engagement and knowledge dissemination.

Oniani and Wang's research [8] delves into COVID-19-related information dissemination challenges. They leverage advanced language models like GPT-2 to develop chatbots that automatically address consumer questions. Such applications bridge the gap between rapidly evolving health concerns and on-demand, accurate information, illustrating the potential for NLP in the healthcare domain.

Kandpal et al. [9] extend the concept of chatbots beyond information provision. They explore context-based chatbots, evolving from menu-based to keyword-based, with a focus on healthcare. Integrating deep learning and natural language processing, they envision chatbots reshaping healthcare delivery, offering predictive diagnoses and support, thus revolutionizing patient care.

Mellado-Silva, Faúndez-Ugalde, and Lobos [10] underscore the transformative impact of chatbots on business processes and education. Their study explores using decision tree-based chatbots to teach complex subjects. The research reveals promising results in enhancing student learning and experience, exemplifying the versatility of chatbots across various sectors.

Srivastava and Singh [11] offer insights into the potential of medical chatbots. By building a diagnosis bot that engages patients in conversation, they demonstrate that AI-driven chatbots can offer accurate disease predictions. Their work contributes to the ongoing dialogue on the integration of AI in healthcare services, focusing on symptom analysis and personalized medical guidance.

Mohan et al.'s study [2] highlights the role of support vector machines (SVM) in data mining and pattern recognition. SVM, a robust classification tool, contributes to accurate data analysis and classification tasks, which is especially valuable in fields like healthcare where precise categorization is crucial.

The study by Divya et al. [12] underscores the importance of healthcare and introduces an AI-powered medical chatbot for disease diagnosis and information dissemination. This chatbot, capable of personalized text-to-text diagnosis based on symptoms, aims to enhance healthcare accessibility, knowledge dissemination, and early intervention.

The work of other researchers, including Mathew et al. [13], Rosruen and Samanchuen [14], Setiaji and Wibowo [15], and Rahman et al. [16], collectively underline the expansive potential of AI-powered chatbots in healthcare. From disease prediction to medical consultation, these studies emphasize AI's role in enhancing healthcare accessibility, diagnosis accuracy, and patient engagement.

In summary, the reviewed papers collectively contribute to the growing body of literature exploring AI-driven chatbots' potential in healthcare and education. From preliminary diagnosis to symptom analysis, these studies illuminate how AI and NLP can reshape medical services, bridge information gaps, and foster more efficient and effective patient care.

III. PROPOSED METHOD

In this proposed system, "MIBOT" designed in this project serves as conversational agent that facilitates the discussion of health concerns based on the symptoms provided. Using a user interface, chatbot can identify different diseases based on input from the user. The NLP technology used in the Chabot's diagnosis and treatment of the disease relies on automatic translation and categorization of the text based on the extracted symptoms. Written commands are input into the proposed system because it is text-based approach. At the start of the process, bot ready to prompted for user's query like

You: 'Hi' or 'how are you' or 'is anyone there?'
MIBOT: 'Hello' or 'Hi there, how can I help?'

Chatbot face many challenges during the input natural language by the user. Input from the user can take any form of organization and any structure. Input can be provided by a user in a variety of ways at different times and it is possible for different users to provide input in many ways.

Such as: 'what is your name', 'what should I call you', 'what's your name?'

Meanwhile, the chatbot awaits a command. The commands in our system can be divided into two types,

1. Command type: Disease classification
2. Command type: General

The core task of the system is the disease classification command. For example,

1. I need help
2. I have some problem
3. Can you help me?
4. Tell me

Symptoms will be asked by the system. At a time, it will take one symptom of a disease. On the other hand, inserting too many symptoms at a time by the user is also unexpected. As many symptoms as possible are encouraged to be entered into the system. When there are more symptoms, the actual disease can be predicted more accurately. Our bot, after a user inputs a symptom, instantly identifies the pattern using Cosine Similarity to measure our training and TF-IDF for vectorization using user-inputted symptoms and our training set features. According to our system, based on both the train and test sets, the SVM classifier gives the highest accuracy among all the symptoms. On the basis of the classification result, the system generates appropriate suggestions. On the other hand, the general commands are analyzed for keywords and whether only medical information is required. Such as:

1. Yellowish white part of eye
2. Bleeding from the bottom (rectal bleeding)
3. I am suffering from anxiety
4. I am suffering from sneezing and an itchy

An enormous collection of disease information is available in our enriched knowledge base. Several authentic internet sources are used to collect this information. The bot retrieves information about diseases from our knowledge base whenever a command asks about them based on similarity measure (Cosine Similarity).

One of the most important tasks of the bot is the creation of test and training datasets. It is necessary to rely on a reliable source of information for us to be able to real-life symptoms into categories that correspond to real-life illnesses. Our test datasets contain independent tests based, on the most commonly used disease classification train [17]. Data were collected from various doctors and hospitals directly by the authors.

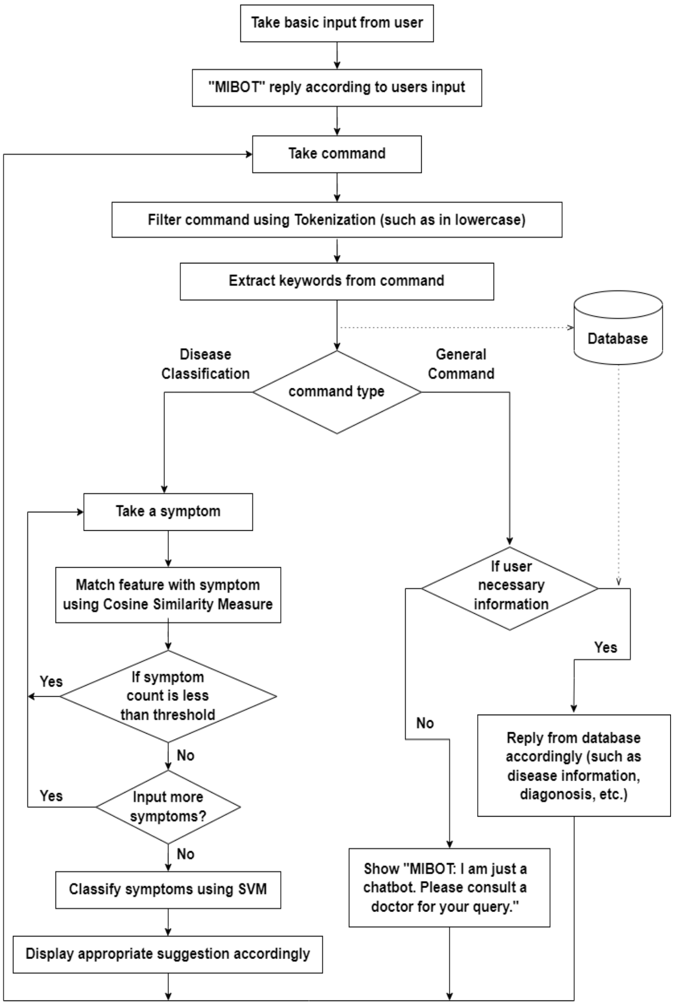

Fig. 1. Schematic diagram of the proposed system

One of the most important tasks of the system is the creation of test and training datasets. Symptoms in real life of real-life diseases must be classified by a reliable source of information. There are also independent test datasets in [18] which used the most popular disease classification trains. Datasets were gathered by contacting doctors and hospitals directly and collecting data from them. Table I summarizes the source datasets.

Using natural language inputs from users, the dataset complexity is reduced without losing originality. As mentioned earlier, users indicate that symptoms usually vary and are fewer. Additionally, the datasets have a high number of features with ‘0s’ indicating negatives and ‘1s’ indicating positives. Here is an example of abstract data of the “Dehydration”

TABLE I. ABSTRACT OVERVIEW DATA OF THE DEHYDRATION

Laziness	**Itching**	**Light headed-ness**	**Dry mouth**	**Mild fever**	
1	0	1	1	0	Dehydration

The first step is to remove all the negative features and keep only the positive ones. After that, in step 2, we have placed the positive feature names where they belong.

Laziness	**Light headed-ness**	**Dry mouth**	**Dehydration**

Table II illustrates the relationship between prognosis and specialty.

TABLE II. DIAGNOSIS AND SPECIALTY MAPPING EXAMPLE

Diagnosis	**Medical specialty**
Dehydration	Gastroenterology
Blisters and Pustules	Dermatology
Excessive Urination	Endocrinology/Urology

A. *Dataset:*

Disease Prediction Using Machine Learning with GUI [17] is the name of the dataset that was crawled from Kaggle for this project. The training dataset has 4920 records and 649572 symptoms, and the testing dataset contains 41 logs and 5544 occurrences of symptoms. The attribute's mean value fills in the gaps left by normalizing the data in the range [0, 1]. For testing our MIBOT algorithm for human illness identification, this dataset supplied 41 different distinct ailments.

B. *Algorithm:*

Once the user has categorized, the bot shows suggestions. The mapping is used to identify the illness, and the user is then sent to the right clinician. The suggested system's operation is described in Algorithm 1.

Algorithm 1: The MIBOT Algorithm

01: Take basic input from user
02: "MIBOT": Reply According to users Input
03: *command* ← Take command input
04: *command* ← Filter *command* using Tokenization (such as in Lowercase)
05: *keyword* ← Extract keywords from *command*
06: *type* ← Type of detect from *keyword*

07: *tally* ← 0
08: **if** command *type* is Disease Classification, **then**
09: *symptom* ← Ask the user to provide certain signs
10: *tally* ← *tally* + 1
11: *indicator* ← utilizing the Cosine Similarity Measure, match the attribute with the *symptom*.
12: **if** *tally* is less than the threshold **then**
13: **repeat** from 9
14: **else**
15: **if** user wants to set additional symptoms, **then**
16: **repeat** from 9
17: **else**
18: Arrange symptoms based on *indicator* using the SVM algorithm
19: show suitable proposition consequently
20: **end**
21: **end**
22: **else**
23: **if** user necessary information, **then**
24: Reply from database accordingly (Predict Disease Information, Diagnosis, etc.)
25: **else**
26: Show "MIBOT: I am just a chatbot Please consult a doctor for your query."
27: **end**
28: **end**

IV. EXPERIMENTAL RESULTS

The new "MIBOT" application has been tested with sick people. They used Chatbots to check the state of their health. The Chatbot is designed to integrate with a web browser. The presence of a cold fever is characterized by symptoms such as coughing, headaches, body aches. On the basis of those symptoms, the MIBOT was able to correctly predict cold fever based on the dataset. The dataset and training requirements vary between different algorithms. Different algorithms have different accuracy levels. This is done with the SVM algorithm. SVM is the algorithm that was better at our experiment because it had the highest accuracy of 98.58 percent, which is the good performance among the entire algorithms. It is clear, therefore, that SVM should be the system central classifier.

Fig. 2. Output Snapshot of MIBOT

Fig. 3. Output Snapshot of MIBOT

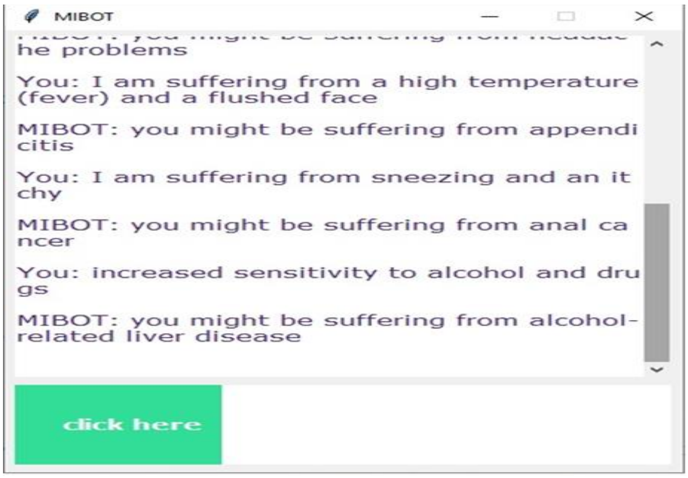

Fig. 4. Output Snapshot of MIBOT

V. PERFORMANCE ANALYSIS

TABLE III. COMPARISON OF OLD AND NEW MODEL

	MedBot [4]	**DoctorMe [4]**	**Disha [9]**	**MIBOT**
Chatbot Types	AI Based	Rule Based	AI Based	AI Based
Installation Required	No	Yes	No	No
Knowledge Required	No	Yes	No	No
Device Compatibility	Computer, Laptop, Tablet	Mobile, Tablet	Computer, Laptop, Mobile, Tablet	Computer, Laptop, Mobile, Tablet
Symptoms Covered	16 symptoms	Various Symptoms	Various symptoms	Various symptoms
Medium	Thai	Thai	Bangla	English

Experimental results were performed on a system with 32 GB RAM, Intel Processor with 8GB graphics card. The experiment was conducted with Python 3.7. The classifier has been validated using an independent test set as well as K fold cross validation with K-fold CV, the complete dataset is randomly subdivided into equal sized samples. As training data, K-1 subsamples are employed, while one subsample is used as test data. Afterward, K Cross-Validation runs were performed. In our experiment we used a 20-fold CV. Table IV summarizes the average 20-Fold score.

A. Algorithm Comparison:

There are two types of datasets that were used in this medical chatbot: training dataset with 75% and testing dataset with 25% respectively.

For simple testing Decision Tree, Random Forest, Multinomial NB, KNN, SVM, AdaBoost algorithms are basically used. The majority of the time, SVM provided perfect results and can handle large datasets as well as work faster. Table III holds a comparison between the old and new models, whereas, the K-Fold average score summary for MIBOT is given in Table IV.

Fig. 5 showcases a graphical comparison between the different algorithms employed for Disha [9] as well as our proposed MIBOT system. Among the six supervised machine learning algorithms, Support Vector Machine (SVM) demonstrated superior effectiveness in disease prediction.

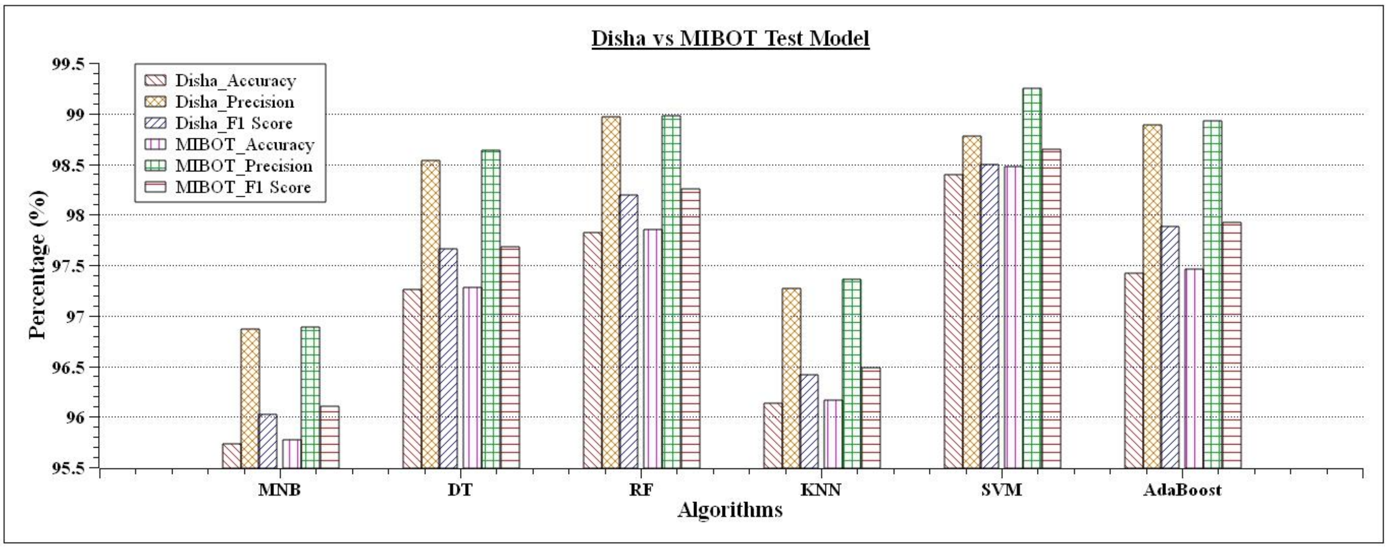

Fig. 5. Graph representation of Test Accuracy for Disha vs. MIBOT.

TABLE IV. K-FOLD AVERAGE SCORE SUMMARY

Algorithms	K-Fold average score
Multinomial Naive Bayes	0.9625
Decision Trees	0.9685
Random Forest	0.9753
KNN	0.9713
SVM	**0.9856**
AdaBoost	0.9687

Our own dataset of independent tests, however, is also used to test the model. Three performance metrics are used in our experiment: F1- Score (Weighted average), Precision (Weighted average), and Accuracy.

TABLE V. CONFUSION MATRIX

Actual Class	Predicted Class: Positive	Predicted Class: Negative
Positive	TP	FN
Negative	FP	TN

Test set contains a range of samples from each class accuracy, precision and F1-Score are weighted according to the number of samples from each class. According to the definition, metrics are defined as follows:

The accuracy (Acc.) measures how many samples were effectively classified out of all the samples, and it measures how well the classification process worked overall. The computation looks like this:

$$\text{Acc.}=\frac{TP+TN}{TP+TN+FP+TN}\times 100\ \% \tag{1}$$

The precision (Prec.) is the ratio of positively classed samples that are successfully classified to all positively projected samples. It is computed as follows and serves as a measure of how precisely a positive sample is identified as such:

$$\text{Prec.}=\frac{TP}{TP+FP}\times 100\ \% \tag{2}$$

The sensitivity (Sens.) is the proportion of positively categorized samples to all positive samples. Thus, false-negative results will be reduced as sensitivity increases. This is how it is calculated:

$$\text{Sens.}=\frac{TP}{TP+FN}\times 100\ \% \tag{3}$$

The F1 Score (F_1) is referred to as the harmonic median of sensitivity and accuracy. It reveals how accurate and reliable the classifier is. The F1 Score is established as follows, and elevated values correlate to high categorization performance:

$$F_1 = \frac{2 \times Prec. \times Sens.}{Prec. + Sens.} \times 100\ \% \quad (4)$$

TABLE VI. OVERVIEW OF EXPERIMENTAL RESULTS

Algorithms	Accuracy	Precision (Weighted avg.)	F1- Score (Weighted avg.)
Multinomial Naive Bayes	95.77	96.89	96.10
Decision Trees	97.28	98.64	97.68
Random Forest	97.85	98.98	98.25
KNN	96.16	97.36	96.48
SVM	**98.58**	**99.25**	**98.65**
AdaBoost	97.46	98.89	97.92

Although the entire algorithms the SVM algorithm is the most accurate and best performance at 98.58 percent, even though all algorithms have shown high performance with high accuracy.

Among the other models, Random Forest, meanwhile, shows the best performance in this case with 97.85 percent accuracy rate, whereas Multinomial Naive Bayes offers the worst result with 95.77 percent accuracy rates. SVM is clearly the most suitable classifier for the system.

VI. CONCLUSION AND FUTURE WORK

Our research successfully deployed MIBOT, a machine learning-based healthcare chatbot for disease prediction. We systematically outlined the chatbot's construction process, which included the creation of custom datasets. Among six supervised machine learning algorithms, Support Vector Machine (SVM) demonstrated superior effectiveness in disease prediction. However, minor shortcomings persist, primarily stemming from the challenge of capturing comprehensive symptom profiles from users, potentially leading to occasional incorrect predictions. We are actively exploring strategies to address this limitation. Our next research phase will focus on enhancing MIBOT's robustness through Deep Learning (DL) integration. This will enable the chatbot to accommodate a broader range of symptoms and patterns. In forthcoming iterations of our research, we will provide a comprehensive elucidation of the specific factors that contributed to SVM's superior performance in comparison to alternative machine learning methodologies. Additionally, the viability of chatbots in healthcare and their potential integration with doctor consultation is pivotal. We are committed to addressing these aspects in our ongoing research, ensuring that MIBOT evolves into a valuable tool for disease prediction and seamless collaboration with healthcare professionals.

REFERENCES

[1] M. C. Surabhi, "Natural language processing future," 2013 International Conference on Optical Imaging Sensor and Security (ICOSS), Coimbatore, India, 2013, pp. 1-3, doi: 10.1109/ICOISS.2013.6678407.

[2] L. Mohan, J. Pant, P. Suyal and A. Kumar, "Support Vector Machine Accuracy Improvement with Classification," 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), Bhimtal, India, 2020, pp. 477-481, doi: 10.1109/CICN49253.2020.9242572.

[3] Batta Mahesh, "Machine Learning Algorithms-A Review", International Journal of Science and Research (IJSR) ISSN: 2319-7064 Research Gate Impact Factor (2018): 0.28| SJIF (2018): 7.426.

[4] Evgeniou, Theodoros & Pontil, Massimiliano, (2001), Support Vector Machines: Theory and Applications, 2049, 249-257, 10.1007/3-540-44673-7_12.

[5] Qaiser, Shahzad, and Ramsha Ali. "Text mining: use of TF-IDF to examine the relevance of words to documents." International Journal of Computer Applications 181, no. 1 (2018): 25-29.

[6] Rahutomo, Faisal, Teruaki Kitasuka, and Masayoshi Aritsugi. "Semantic cosine similarity." In The 7th international student conference on advanced science and technology ICAST, vol. 4, no. 1, p. 1. 2012.

[7] 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), Noida, India, 2020, pp. 619-622, doi: 10.1109/ICRITO48877.2020.9197833.

[8] David Oniani and Yanshan Wang. 2020. A Qualitative Evaluation of Language Models on Automatic Question-Answering for COVID-19. In Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (BCB '20). Association for Computing Machinery, New York, NY, USA, Article 33, 1–9. https://doi.org/10.1145/3388440.3412413.

[9] P. Kandpal, K. Jasnani, R. Raut and S. Bhorge, "Contextual Chatbot for Healthcare Purposes (using Deep Learning)," 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), London, UK, 2020, pp. 625-634, doi: 10.1109/WorldS450073.2020.9210351.

[10] R. Mellado-Silva, A. Faúndez-Ugalde and M. B. Lobos, "Learning tax regulations through rules-based chatbots using decision trees: a case study at the time of COVID-19," 2020 39th International Conference of the Chilean Computer Science Society (SCCC), Coquimbo, Chile, 2020, pp. 1-8, doi: 10.1109/SCCC51225.2020.9281267.

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

[12] Divya, S., V. Indumathi, S. Ishwarya, M. Priyasankari, and S. Kalpana Devi. "A self-diagnosis medical chatbot using artificial intelligence." Journal of Web Development and Web Designing 3, no. 1 (2018): 1-7.

[13] R. B. Mathew, S. Varghese, S. E. Joy and S. S. Alex, "Chatbot for Disease Prediction and Treatment Recommendation using Machine Learning," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019, pp. 851-856, doi: 10.1109/ICOEI.2019.8862707.

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

[15] B. Setiaji and F. W. Wibowo, "Chatbot Using a Knowledge in Database: Human-to-Machine Conversation Modeling," 2016 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), Bangkok, Thailand, 2016, pp. 72-77, doi: 10.1109/ISMS.2016.53.

[16] M. M. Rahman, R. Amin, M. N. Khan Liton and N. Hossain, "Disha: An Implementation of Machine Learning Based Bangla Healthcare Chatbot," 2019 22nd International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2019, pp. 1-6, doi: 10.1109/ICCIT48885.2019.9038579.

[17] Neelima, "Disease Prediction Using Machine Learning with GUI", 2019. [Online].Available: https://www.kaggle.com/datasets/neelima98/disease-prediction-using-machine-learning. [Accessed: May 7, 2023].

[18] Tin Kam Ho, Random Decision Forests, In Document Analysis and Recognition, 1995., proceedings of the 3rd International Conference on, volume 1, pages 278-282, IEEE, 1995.