Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/170221

Developing A Primary Healthcare Diagnosis System Based on Natural Language Processing

Ahmad A. Alzahrani

Department of Information Systems, College of Computers and Information Systems, Umm Al-Qura University, P.O.Box 8XH2+XVP, Mecca 24382, Saudi Arabia

Received: 1 Dec. 2022, Revised: 11 Feb. 2023, Accepted: 15 Feb. 2023
Published online: 1 Mar. 2023

Abstract: A machine learning (ML) algorithm is a complex algorithm that automates based on experience. Artificial intelligence (AI) is a subfield of machine learning (ML) that builds computer networks that can execute activities that would normally need cognitive functions. Artificial intelligence and chatbots are two revolutionary technologies that have altered the way patients and clinicians view healthcare. A diagnostic is needed to kind the system of healthcare is more participatory. Chabot was created utilizing the most up-to-date machine learning methods, including the decision tree algorithm, to assist users in making a diagnosis of their disease based on their symptoms. The application of natural language processing (NLP) approaches to constructing conversational schemes for the health diagnosis that improves patients' access to the medical information. Based on the symptoms entered, a fuzzy support vector machine (SVM) is employed to accurately forecast the disease. The service evaluates tropical clinical manifestations. To interface the chatbots to the service, the Telegram Bot Application Programming Interface (API) has been used, whereas the Twilio API was being used to attach the scheme to an SMS user. As a result, a medical diagnosis system has been developed that gives a personalized diagnosis based on user input to successfully diagnose disorders. The system usability scale (SUS) was used to assess the usability of the developed system, obtaining an average SUS score, indicating a positive overall assessment.

Keywords: Machine Learning; Natural Language Processing; Medical diagnostic; Health care systems.

1 Introduction

Remote diagnosis systems become more frequent and efficient, with enormous benefits such as financial impact, rapid and trustworthy data integration for clinical diagnosis, and treatment and prevention of disease, illness, accident, and other developmental disabilities in humans [1]. Chatbots, also known as chatter-bots, are computer programs that were considered to start a discussion with an individual and make yourself as real as feasible. Text or speech can be used to conduct the dialogue. Chatbots can be built with a specific role in mind, such as a medical Chatbot that exclusively responds to healthcare-related questions. The capacity of such a smart system to pass the turning test can be used to assess its efficiency [2].

The continued advancement of mobile technology has had an impact on every aspect of humanity all over the world, as its supporting of healthcare goals via telecommunication, telemedicine, and m-health is helped with in diagnosis and management of diseases at a minimal cost, particularly in underdeveloped nations where testing and therapy options were restricted [3]. Because of its low cost, dependable delivery, customization for users, and lack of Internet-oriented service, short message service (SMS) has shown to be distinctive and reliable as among different communication mediums available on mobile devices. It generally means that system is educated or taught with pre-labeled information. To put it differently, this signifies that the data has already been tagged with the right answer [4]. This training data is analyzed by the computer, which then creates a function that connects the data flow.

Natural language processing (NLP) technique may be used to connect with computers and humans by extracting knowledge from unstructured free text using the analysis of linguistic and deep learning approaches. The systems of NLP have demonstrated on their distinctiveness and significance in the field of data recovery, particularly in the recovery and giving out of vast amounts of the unorganized healthcare accounts and the delivery of

* Corresponding author e-mail: aalzahrani@uqu.edu.sa

organizing data via user-defined queries [5]. In general, the NLP system aims to explicitly reflect the knowledge provided by document written in a natural language. This research proposes the use of a natural language processing paradigm using an SMS and a chatbot interface to promote health identity and decision support in digital medical systems [6]. Extracting useful information from electronic health records (EHR) is a growing issue of interest in healthcare, including the use of EMRs there at healthcare center as well as in the internet has provided a large amount of data to also be analyzed. A digital record of wellbeing data generated, gathered, and maintained by medical practitioners is known as an electronic medical record (EMR). Gathering of current and accessible health information challenges include merging NLP into numerous EMRs and assuring patient data privacy and security [7].

Machine learning methods, particularly SVM, have demonstrated good results in categorizing free text in medical records, such as Georgian. SVM using a polykernel was used to classify the content and severity of the primary care patient security occurrence reports [8]. However, the authors claim that strengthening descriptions and increasing the training models of specific groups will advance the system's presentation even further. To design a clinical information extraction system, The authors created pretrained techniques that are based on previous multilayered neural network (NN) structure. The researchers utilized a wide question as well as the numeric rating scale to deploy and analyze usability (SUS).

In summary, present medical diagnosis systems (MDS) frequently make erroneous conclusions as a result of incorrect understanding about the patient's text-based contribution. As a result, there is a need to automate MDS for quick disease detection and to support clinicians' judgments depending on the severity of symptoms [9, 10]. Furthermore, medical professionals seek a system to keep track of enormous amounts of message facts pertaining by individuals in natural language, thus improving health care for telehealth. The following is the paper's contribution: (1) This resulted in the development of a text-based clinical diagnosis technique that enables individualized diagnosis by utilizing user identity to efficiently help diagnose [10]. (2) For SMS and Telegram bots, the proposed solution integrates NLP and machine learning methods. (3) By using an aggressive approach of an inquiry and response technique, the system may make a clinical condition.

2 Related Work

Exams termed "Objective Structured Clinical Examinations" (OSCEs) are given to medical students to measure their medical competence in clinical assignments [11]. To increase the virtual patient's communication skills, create a deep learning framework. For starts, deep neural networks acquired domain-specific word vectors. Then, until a convolutional neural network architecture chose an answer from such a script, long short-term memory network generated phrase embeddings. On a domestic corpus, empirical data showed that this framework beat previous techniques, with an accuracy of 81 percent.

This paper gives a semiotic study of the communicability of the relaxed boundaries (such as chatbots) [12]. A chatbot is a software system that uses a conversational interface to replicate genuine discussions between gadgets and users (CI). Thus, while exploring the communicability aspects of a CI, a user test was done to compare traditional and chatbot interfaces. This study compares the communicability advantages of a chatbot to a standard GUI for boosting the efficiency and effectiveness of communications among users of the system, especially for people who really are resistant to technology. In specifically, the communicability of two models that might be used to execute simple tasks was evaluated in order to foster user involvement, especially among persons with little or no expertise with technology.

This research describes a solution that fills this void by allowing Patients will be able to complete their treatment at home [13]. A psychologist is just engaged online to monitor progress and provide encouragement. With this technology, patients can replay their experiences in a digital journal and reproduce them in a three-Dimensional Creator. These questions were found to be beneficial for memory retrieval in usability research with former PTSD sufferers ($n = 4$). Furthermore, the overall system's usability was graded favorably. This technique offers great ability to be a critical contribution to the PTSD therapy landscape, allowing for a novel type of home treatment with the assistance of a chatbot.

An architectural concept for a chatbot is act as a virtual diabetes physician was proposed [14]. Diabetic patients will be able to get diabetes control/management guidance from this chatbot without having to go to the hospital. A broad history of chatbots is given, as well as a brief overview of each chatbot. This recommended the creation of a novel technique that will be deployed as a crucial component in this chatbot's ability to act as a diabetes physician. The discussion path will be remembered by the chatbot using this approach via a parameter called Vpath. Because Vpath is created to be a virtual diabetes physician, it will allow the chatbot to deliver an answer that is mainly appropriate for the entire conversation.

In today's world, where most items are a click left, many reflect by visiting a hospital to be the more active, dependable, and expedient option to get disease diagnosis. Machine Learning is played as a significant influence in the healthcare field in recent years [15]. The proposed method or strategy emphases on developing an alternative by using the Decision Tree Algorithm in which individuals will communicate with a chatbot that would recognize further symptoms, anticipate the condition, and recommend a particular doctor. By using the approach outlined overhead can assist individuals but both a time and the money.

3 Methodology

The planned system will be function as medical application. The person can create a profile on the website then communicate with a doctor in real time. In the event that the doctor is not present, the site includes an initial diagnosis a chatbot. Then the user can enter on his or her symptoms by using either text. The chatbot that will use natural language processing that interpret the person enquiry [16]. Once the bot acknowledges the early symptoms, this offers follow-up queries this attempt to develop an analysis based on the feedback provided by the user. To support generate a precise analysis, system employs top-down strategy and decision tree algorithm. The first symptoms provided will serve as the decision tree's root.

The network will include several modules, but the three basic components are as follows: The result is a text-to-speech and speech-to-text translation of the user inquiry [17]. NLP is used to comprehend user intent by processing the entered response. Create a decision tree algorithm and browse until you reach an accurate node.

The research evaluates the clinical information expectations and demands for detecting diseases, as well as the clinical information of patients discovered in EHRs or physical records. Figure 1 that shows the structure of the planned text-based clinical analysis organization. The following phases are performed in the described text-based clinical diagnosis organization: specification base of knowledge; text-based preprocessing data; document labeling; strategy employed; and candidate answer rating [18]. The Python programming language was used to construct the diagnosis system framework because of the following features: Third-party frameworks for the machine learning and NLP are cross-platform and widely available. The system makes use of Python language modules to retrieve the machine learning functions and NLP required for classification.

3.1 Natural Language Processing (NLP)

The user's information is analyzed via a number of phases in order to grasp how each user is attempting to communicate Natural language processing (NLP) is the capacity of a software to comprehend the meaning of human-spoken natural language. The study of a mathematical handling of natural (human) language is known as NLP [19]. NLP development is difficult since machines are accustomed to receiving properly organized input, while the natural language is exceedingly complicated and imprecise, with a various language pattern and a complex variable. NLP has a different phase explained below.

- Tokenization (linguistic research), also known as segmentation, is the process of dividing a paragraph or article into symbols, numerals, or meaningful phrases. Tokens are little parts, similar to how a term is a symbolic in a phrase and a phrase is a symbol in a sentence. Word borders were used to split the words. Because English is a space-limited language, word borders are the spaces between the end of one word and the beginning of another.
- Syntactic evaluation is analyzing words for grammar and arranging them together in a way that demonstrates their connection. This could be accomplished through the use of a dataset such as syntax tree. The grammatical rules of the languages are used to build the trees. If the information will be generated by using the syntax tree, it was considered to have a correct syntax.
- Semantic analysis examines the semantic meanings of words in order to determine the true meaning of a statement. It is the translation of sentence patterns to the actual or message definition of words. Strings such as ”hot winter” will be ignored.
- Pragmatic research controls outside word knowledge, which means understanding outer surface to the archiving and furthermore queries. Pragmatics study centered on what was displayed reconstructed it truly indicated, concluding the various aspects of the language that a necessitate type of the learning.

Figure 2. represents the text image pre-processing process. To placed it another way, a following satisfied is divided into an individual phrases for labelling with a linguistic type's markings based on their places and a neighbor in the phrase. Various types of phrases construction will be used in this stage to be chunk the separate identified words and the frame statements. By excluding unfavorable terms in chinking operations, keywords can be deleted from these statements. These slogans can be rectified if they are incorrect.

3.2 Decision Tree

These are the most preferred special for classification and guess since they are not only simple to grasp but also extremely controlling. Then the decision tree takes its designation from its tree-like form, in which the nodes represent an assessment, in this case an indication, the division represents the conclusion, and the tree structure represent a probable diagnosis. In platform, the chatbots must make a choice depending on each user input [20]. For an example, if a person arrives that he or she has a fever, how should the chatbot has respond? will the person be asked, and will get at a diagnostic test? This function is aided by the decision tree algorithm. The state's major element is decision making. Decision -

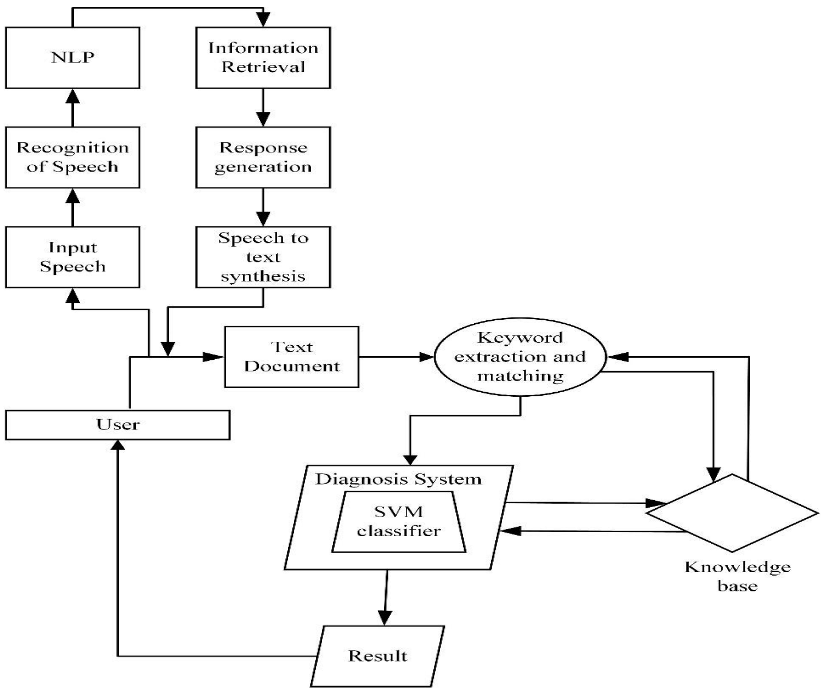

Fig. 1: System model.

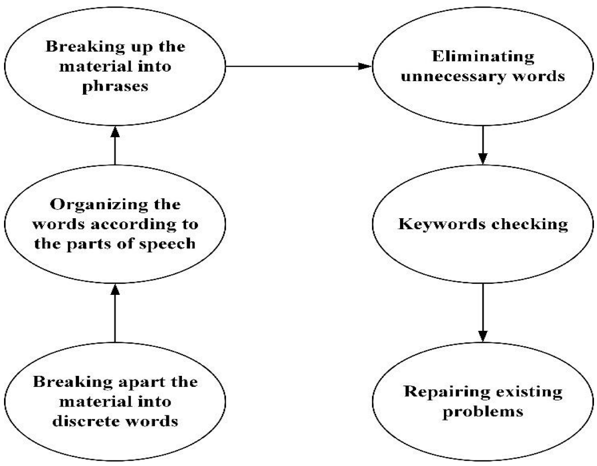

Fig. 2: Text image Pre-processing.

making process is critical for the system's operation as well as the accurateness of the outcomes. The system will be neutral, depending on human input at every stage. By comparing the input from the user with the conditions for each layer, the decision tree algorithm will assist in traversing to discover a remedy. When a no match is made, the algorithm will be continued till end of the tree is achieved.

The information used in tabulated form and contains information on roughly 50 diseases as well as associated symptoms [21]. Because it was challenging to find an appropriate diagnosis information, the type of data has already been selected to guarantee that only general diseases with appropriate indications are listed to enhance productivity in disease identification.

The system's operation is better explained by the use graph in the Fig. 3. shown below.

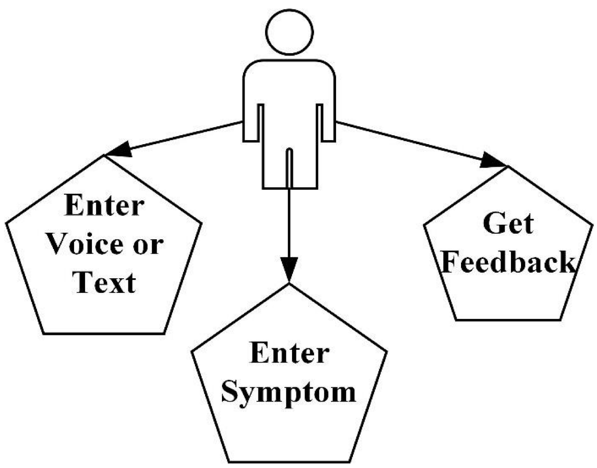

Fig. 3: Text image Pre-processing.

3.3 Knowledge Base

A base of knowledge was the primary source of information in an inquiry technology, that can be supervised or unsupervised. To create a base of knowledge, information from the healthcare database management scheme is gathered that classified into a category known as illness context knowledge.

3.4 Communication System

A communication system is created on a solid foundation of information. to allow consumers to communicate with the medical practitioner via Telegram or SMS using question-and-answer rules [22]. Each diagnostics inquiry includes specific characteristics and features that provide further information about the inquiry. The following are the many characteristics of a request / inquiry:

- Diagnosis inquiry: This is the real diagnosis issue which will be delivered to the person.
- Reaction: a list of replies that will be displayed to the person, denoting the inquiry that he or she will provide to a network through the Telegram GUI or Text.
- Serial id: A sequence where queries will be requested is indicated by serial id.

3.5 Extraction of content and Pre-processing text

The content module is in charge of collecting learning material from SMS messages. For benefit, the module includes a number of text extraction methods that specialize in extracting various types of data from such an SMS. The information concentrators use the NLP software to structural analysis. When a sender submits an SMS explaining his or her issues, the program's SMS nodes receive it and forwards the text to the NLP modules, which examines the message, performs required edits, and extracts essential keywords.

Three major steps are involved in text processing techniques: Noise reduction, tokenization of text sentences, and sentence breaking

3.5.1 Noise Reduction

Text is made up of a series of elements, both significant and irrelevant [23]. As a result, noise from raw texts is removed, leaving only useful content connected to topics for the further analysis.

3.5.2 Tokenization

This requires breaking down substring into lexical pieces. In this case, the sentence splitting technique was utilized to separate the material into discrete sentences [24]. This used as the natural language toolkit (NLTK) tokenizer in this case.

3.5.3 Document tagging

To identify quality information from a single text, valuable data from the knowledge basis is get tagged. Disease terms were employed as document labels, and the tools including parser and WordNet have been used for the text labeling.

3.5.4 Parser

A standard Parser is utilized as a tool for producing the parts of speech (POS) for every expression entered by the person query, as well as potential solutions chosen from a data base. The WordNet were utilized by determine the association among both the person query's terms and the foundation of data. Nouns, verbs, adjectives, and adverbs are the four types of words.

3.5.5 Term Matching

The system then queries the base of knowledge in order to the extracted terms with an information is kept in the base of knowledge.

3.6 Feature Extraction and selection

A retrieved set of the key phrases was converted into a feature space appropriate for use with a machine learning approach. they were using the word embedding technique to turn the data into the feature vector. Embeddings are word representations in a semantic multidimensional space [25]. They used an accessible embeddings Glove based on Twitter for approach since it offers a fair representation to normal English for just an informal channel of communication. Messages was normalized and turned into extracted features for efficient information learning [26]. This was accomplished using extraction of features utilizing word embeddings. The unlabeled document's extracted features were then passed to the classifier's judgement purpose, which returned a classification for unlabeled text.

3.7 Reasoning Fuzzy Module

The main goal is employed fuzzy logic-based techniques to read and understand the patient's answers, control and measure all of the indications that a user takes previously reacted, and give queries to the users which are greatest relevant depend on the disease information that is kept. Each condition is represented by a container, with each container representing a symptom. To deal with multiple symptom disorders, concept of fuzzy rules is developed [27]. CUDoctor's systems analyze the state of such categories and send a most appropriate inquiry to the user. This allows us to reduce the series of questions that the system would ask in order to arrive at a diagnosis [28]. The weighted fuzzy inference system rule system is used, with each fuzzy inference system allocated a value based on statistics. The Mamdani fuzzy inference system model of fuzzy system was employed, with each rule represented by an IF-THEN expression.

The rules of fuzzy are specified as "If $i1$ is $P1$ and $j1$ is $Q1$ then $k1$ is $R1$," where $p1$, $q1$, $r1$ are the fuzzy sets. The weighting of the fuzzy rules is determined by estimation of the amount of the importance of the features and signs to disease diagnosis. C is a crisp dataset with n topographies and k models $[F1, F1, \ldots, Fn]$ and then n-dimensional tuple as $T_i = [p1, p1, \ldots, pn]$ is characterized as the kn-dimensional feature path:

$$Ti = [<\mu FT1(p1), \mu FT2(p1), \ldots, \mu FTk(p1)>, \ldots, \\ <\mu FT1(pn), \mu FT2(pn), \ldots, \mu FTk(pn)>], \quad (1)$$

Where $\mu FT1(pi)$ is membership of the fuzzy degree period feature FTk as $Fi(Fi = pi)$. If fuzzy variable Fn has a k fuzzy term as, $FT1, FT2, \ldots, FTK$, then for the individually value x of Fn, the rate of fuzzy is computed as $max\{\mu FT1(x), \mu FT2(x), \ldots, \mu FTk(x)\}$. For implications, the limit operators were utilized. The central fuzzy rules were used to create the proper result, where the weight is represented by the grade of association of the rate xi with the notion modelled by the fuzzy system A.

The procedure is carried out as follows (Figure 4.):

1. Fuzzification is the development of changing sharp inputs into a fuzzy value. The degree of membership function is defined using expert judgement. A fuzzification controller gets data input (fuzzy variable) and evaluates it using algorithms during fuzzifier.
2. It consists of a fuzzy term collection and a fuzzy IF-THEN rule basis. The disease is described in the classification model for each combination of crisp input parameters.
3. The inference mechanism applies relevant fuzzy rules to the incoming data.
4. As a result of fuzzy output, the crisp measurement results are produced from a fuzzy value. The fuzzy rule with the maximum score is chosen as the final decision.

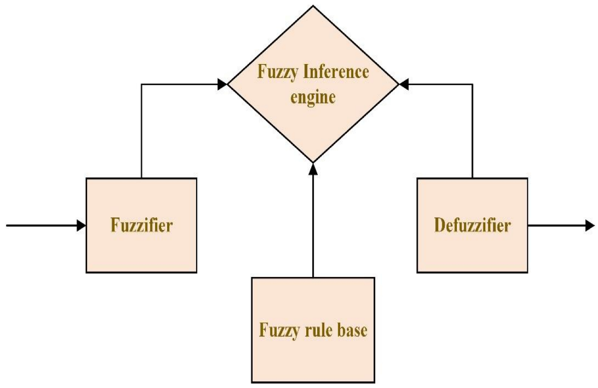

Fig. 4: The fuzzy logic-based reasoning module's structure.

3.8 Module of classification

The choosing of the classification in this research was formulated on the basis needs of the planned application as well as the essential for a classifier with superior outcomes following the design literature for text categorization. Furthermore, there were unique constraints in the new application that influenced classifier decision, such as the computing difficulty for the learning and/or testing stages. The fuzzy SVM classifier is already provide classification predictions on messages utilizing machine learning [29]. It just has one hyper - parameter C, a regularization parameter. Grid search was used to fine-tune the hyperparameter's value. In a fuzzy SVM, a fuzzification value is utilized for every information point of the SVM, then the SVM was recast so that various input points can contribute to the knowledge of the decision function in various ways. To begin, employ clustering techniques to identify data groupings. Fuzzification values of cluster datasets are set to a 1, while fuzzification numbers of the other information points were calculated by the location to a closest group, respectively. The classifiers are skilled to use a collection of experimental documents that were analyzed by the NLP software and turned into word embeddings that were then utilized as extracted features [30]. Finally, the extracted features collected from user responses are fed into the fuzzy SVM classifier, which offers a diagnosis by associated with the chosen on the important aspect of life in the SMS and then communicates the results to the client via Text [31].

3.9 Graphical User Interface (GUI)

The organization begins information exchange with the consumer in order to learn further about their basic private information, including such height, weight, gender, and age. When the fundamental information is collected, the CUDoctor advances to the another phase and enquiries client for the symptoms using aforementioned algorithm. The Telegram API was used in the GUI plan, with a bespoke keyboard supplied by the Telegram API.

4 Result and discussion

Users was using the Bilingual Evaluation Understudy (BLEU) score to monitor the efficiency of the built product, and it has become a standard method for assessing chatbot solutions. BLEU evaluates a provider's output responsiveness in comparison to the position, with a BLEU value ranging from 0 to 1. The BLEU-2 was employed in this case, which would be dependent on unigram and bigram similarities among both the produced and situation phrases. The Recall-Oriented Understudy for Gisting Evaluation (ROUGE-L) metric is based on the common patterns subgraph, was also employed (LCS). The values for BLEU-2 are 26.28 and 32.55 for ROUGE-L, respectively.

4.1 Testing usability

This followed the instructions in order to do usability testing on the developed solution. By employing the system usability scale (SUS), a form designed to assess the ease of use of the system by using a Likert scale. The SUS score can assess usability effectiveness in terms of usefulness, efficiency, and navigation application. SUS is regarded as a valid instrument for measuring utility, and it provides for the evaluation of a widespread range of the products and services. SUS becomes an accepted practice, with an over 1300 references in articles and journals, including healthcare chatbot as well as other NLP-based advanced medical systems. SUS is intended to aid in the evaluation and comparison of user experiences while dealing with various tools, and this is advised that it should be included in assessment of healthcare chatbots.

Let consider usability test had 25 volunteers, 11 females and 14 males, with 12 being between the ages of 23 and 35, 7 being between the ages of 36 and 45, and 6 being between the ages of 46 and 55. The study subjects gave their informed consent before participating in the investigation. Each session of the analysis process was less than 45 minutes. The participants were given information papers well about study research to interact with the chatbots via their mobile phones. Anonymization was applied to all of the responses received. The SUS score is divided into four categories: ¿80.3 (outstanding), 68–80.3 (excellent), 68 (acceptable), 51–68 (bad), and 51–68 (very bad). As a result of the analysis, CUDoctor received an average SUS score of 80.4, which is higher than the threshold of 68, indicating that the overall assessment was extremely good. Figure 5 shows the findings for all of the SUS questions.

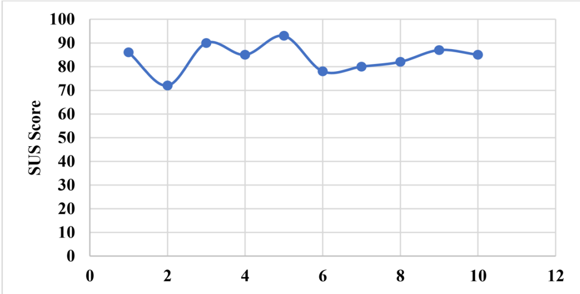

Fig. 5: Evaluation usability of CUDoctor using SUS.

4.2 Comparison with other NLP-based services

On social networking sites like Facebook, there are many chatbot with medical-related applications. For instance, the FLORENCE bot effects appear when to participate in their care and checks their health and moods. SMOKEY alerts users to poor air quality. Health Tap gives responses by utilizing a knowledge library that contains comparable queries. Google supplies the Dialog flow Application Program Interface (API) for integrating natural language processing (NLP) into predefined objectives. Woebot offers and it has been evaluated with cognitive behaviour therapy for individuals suffering from depression. It enabled them to minimize their mood disorders as measured by a depression inquiry PHQ-9.

XiaoIce is a common chatbot which focuses on expressive relationship while employing a deep learning to do expressive answer discussion responsibilities. The chatbots are also utilized in the stoppage strategies and the cognitive communication treatment, specifically targeting high-risk populations including such HARR-E and Wysa. The key difference between the scheme as labeled in this article and the social networks is the service was supplied via SMS is rather than the social networking sites, which demand extremely high Internet speed, which is sometimes inaccessible in rural regions of poor countries. Furthermore, the given approach concentrates on the unique area of tropical diseases symptoms evaluation, but not know of any NLP-based algorithms that do so. The suggested system is a functional prototype that diagnoses roughly 150 symptoms and 50 diseases utilizing different layers of decision tree. The procedure generates a structure of tree from a specific sign and utilizes to cross-examine the person.

- Round Trip Time (RTT): The program's reaction rate was discovered to be 10ms to 20ms based on the intensity of the symptoms from Fig. 6.
- Accuracy: For common ailments, the suggested method provided 76 correct replies out of 100 queries. The efficiency was determined to be 76% from Fig. 7.

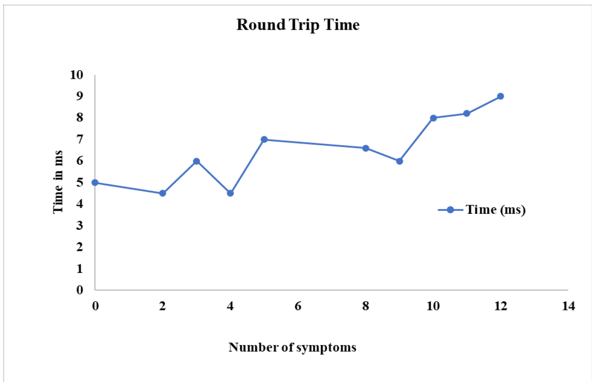

Fig. 6: Round Trip Time.

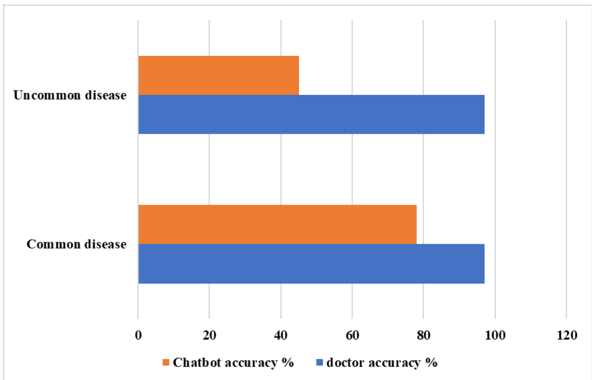

Fig. 7: Accuracy.

5 Conclusion

Today's healthcare landscape has been altered by artificial intelligence. The suggested system intends to further close the gap between healthcare and patients. To construct an efficient diagnostic chatbot, AI techniques such as decision tree and the NLP are utilized to retrieve a knowledge from medical database including roughly 150 ailments. The chatbot asks the customer queries in the style of a doctor-patient consultation. This research was capable of creating a text-based medical diagnostic system based on the highlighted needs, it provides a customized diagnostic by using user identity reaction to successfully suggest a diagnosis of diseases Bots with SMS and Telegram, the proposed solution was able to mix NLP and machine learning algorithms. By employing a direct approach to the issue and response procedure, the system was able to provide a diagnosis. The system's shortcoming is that it is not safe from illegal acts of disease suggestion; so, a definitive diagnosis must be verified by a medical expert. To make the program extra interactive, sound interaction will be implemented. These advancements will help to reduce costs and death rates, lowering the great difficulties on medical officers in developing countries.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

References

[1] B. Bhattacharya, S. Mohapatra, A. P. Mukhopadhyay and S. Sah, Remote cardiovascular health monitoring system with auto-diagnosis, in *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*, IEEE2019, pp. 1–5.

[2] B. Kidwai and R. Nadesh, Design and development of diagnostic chabot for supporting primary health care systems, *Procedia Computer Science* **167** (2020) 75–84.

[3] N. A. Omoregbe, I. O. Ndaman, S. Misra, O. O. Abayomi-Alli, R. Damaševičius and A. Dogra, Text messaging-based medical diagnosis using natural language processing and fuzzy logic, *Journal of Healthcare Engineering* **2020** (2020) 1–14.

[4] D. J. Hongoro, The cost and cost-effectiveness of a text-messaging based intervention to support management of hypertension in south africa, Master's thesis, University of Cape Town (2017).

[5] C. Friedman, G. Hripcsak *et al.*, Natural language processing and its future in medicine, *Acad Med* **74**(8) (1999) 890–5.

[6] S.-i. Toyabe, Detecting inpatient falls by using natural language processing of electronic medical records, *BMC health services research* **12**(1) (2012) 1–8.

[7] T. Cai, A. A. Giannopoulos, S. Yu, T. Kelil, B. Ripley, K. K. Kumamaru, F. J. Rybicki and D. Mitsouras, Natural language processing technologies in radiology research and clinical applications, *Radiographics* **36**(1) (2016) 176–191.

[8] W.-H. Weng, K. B. Wagholikar, A. T. McCray, P. Szolovits and H. C. Chueh, Medical subdomain classification of clinical notes using a machine learning-based natural language processing approach, *BMC medical informatics and decision making* **17**(1) (2017) 1–13.

[9] S. Harispe, S. Ranwez, S. Janaqi and J. Montmain, Semantic similarity from natural language and ontology analysis, *Synthesis Lectures on Human Language Technologies* **8**(1) (2015) 1–254.

[10] A. Elhadad, F. Alanazi, A. I. Taloba and A. Abozeid, Fog computing service in the healthcare monitoring system for managing the real-time notification, *Journal of Healthcare Engineering* **2022** (2022).

[11] J. El Zini, Y. Rizk, M. Awad and J. Antoun, Towards a deep learning question-answering specialized chatbot for objective structured clinical examinations, in *2019 International Joint Conference on Neural Networks (IJCNN)*, IEEE2019, pp. 1–9.

[12] S. Valtolina, B. R. Barricelli and S. Di Gaetano, Communicability of traditional interfaces vs chatbots in healthcare and smart home domains, *Behaviour & Information Technology* **39**(1) (2020) 108–132.

[13] M. L. Tielman, M. A. Neerincx, R. Bidarra, B. Kybartas and W.-P. Brinkman, A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories, *Journal of medical systems* **41** (2017) 1–10.
[14] A. S. Lokman, J. M. Zain, F. Komputer and K. Perisian, Designing a chatbot for diabetic patients, in *International Conference on Software Engineering & Computer Systems (ICSECS'09)*, 2009, pp. 19–21.
[15] R. Pradhan, J. Shukla and M. Bansal, 'k-bot'knowledge enabled personalized healthcare chatbot, in *IOP Conference Series: Materials Science and Engineering*, **1116**(1), IOP Publishing2021, p. 012185.
[16] R. Shah, S. Lahoti and K. Lavanya, An intelligent chat-bot using natural language processing, *International Journal of Engineering Research* **6**(5) (2017) 281–286.
[17] M. U. H. Al-Rasyid, S. Sukaridhoto, M. I. Dzulqornain and A. Rifai, Integration of iot and chatbot for aquaculture with natural language processing, *TELKOMNIKA (Telecommunication Computing Electronics and Control)* **18**(2) (2020) 640–648.
[18] A. Jiao, An intelligent chatbot system based on entity extraction using rasa nlu and neural network, in *Journal of physics: conference series*, **1487**(1), IOP Publishing2020, p. 012014.
[19] T. Beysolow II and T. Beysolow II, What is natural language processing?, *Applied Natural Language Processing with Python: Implementing Machine Learning and Deep Learning Algorithms for Natural Language Processing* (2018) 1–12.
[20] P. Rani, R. Kumar, N. M. S. Ahmed and A. Jain, A decision support system for heart disease prediction based upon machine learning, *Journal of Reliable Intelligent Environments* **7**(3) (2021) 263–275.
[21] M. Shouman, T. Turner and R. Stocker, Integrating decision tree and k-means clustering with different initial centroid selection methods in the diagnosis of heart disease patients, in *Proceedings of the International Conference on Data Science (ICDATA)*, The Steering Committee of The World Congress in Computer Science, Computer . . . 2012, p. 1.
[22] R. B. Mathew, S. Varghese, S. E. Joy and S. S. Alex, Chatbot for disease prediction and treatment recommendation using machine learning, in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, IEEE2019, pp. 851–856.
[23] K. Sivaraj, K. Jeyabalasuntharam, H. Ganeshan, K. Nagendran, J. Alosious and J. Tharmaseelan, Medibot: End to end voice based ai medical chatbot with a smart watch, *International Journal of Creative Research Thoughts* **9**(1) (2021).
[24] J. J. Sophia, D. A. Kumar, M. Arutselvan and S. B. Ram, A survey on chatbot implementation in health care using nltk, *Int. J. Comput. Sci. Mob. Comput* **9** (2020).
[25] W. M. A. W. Zaki, M. F. M. Shakhih, M. H. Ramlee and A. Ab Wahab, Smart medical chatbot with integrated contactless vital sign monitor, in *Journal of Physics: Conference Series*, **1372**(1), IOP Publishing2019, p. 012025.
[26] A. El-Komy, O. R. Shahin, R. M. Abd El-Aziz and A. I. Taloba, Integration of computer vision and natural language processing in multimedia robotics application, *Inf. Sci* **7**(6) (2022).
[27] S. Tandra, D. Gupta, J. Amudha and K. Sharma, A fuzzy-neuro-based clinical decision support system for disease diagnosis using symptom severity, in *Soft Computing and Signal Processing: Proceedings of 2nd ICSCSP 2019 2*, Springer2020, pp. 81–98.
[28] R. Vineeth, R. Rithish, D. Sai Varma and B. Ajay Prakash, Smart health care chatbot for prognosis of treatments and disease diagnosis using machine learning, *Journal of Computational and Theoretical Nanoscience* **17**(9-10) (2020) 3947–3951.
[29] M. H. Khalaf, Z. A. Azim, W. H. Elkhateeb, O. R. Shahin and A. I. Taloba, Explore the e-learning management system lower usage during covid-19 pandemic, *Information Sciences Letters* **11**(2) (2022) 537–548.
[30] P. Sharma, M. Ved, J. Chintalapti and S. N. Pal, Big data analytics and machine learning technologies for hpc applications, in *Evolving Technologies for Computing, Communication and Smart World: Proceedings of ETCCS 2020*, Springer2021, pp. 411–424.
[31] M. A. Fouly, T. H. A. Soliman and A. I. Taloba, Developing an efficient secure query processing algorithm for unstructured data on encrypted databases, in *2022 10th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC)*, IEEE2022, pp. 252–257.