

Research Article

Text Messaging-Based Medical Diagnosis Using Natural Language Processing and Fuzzy Logic

Nicholas A. I. Omoregbe,¹ Israel O. Ndaman,¹ Sanjay Misra,^{1,2} Olusola O. Abayomi-Alli,¹ and Robertas Damaševičius^{3,4} 

¹Center of ICT/ICE Research, CUCRID Building, Covenant University, Ota, Nigeria

²Department of Computer Engineering, Atilim University, Incek, Ankara, Turkey

³Department of Applied Informatics, Vytautas Magnus University, Kaunas, Lithuania

⁴Faculty of Applied Mathematics, Silesian University of Technology, Gliwice, Poland

Correspondence should be addressed to Robertas Damaševičius; robertas.damasevicius@polsl.pl

Received 3 May 2020; Revised 9 September 2020; Accepted 18 September 2020; Published 29 September 2020

Academic Editor: Ayush Dogra

Copyright © 2020 Nicholas A. I. Omoregbe et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The use of natural language processing (NLP) methods and their application to developing conversational systems for health diagnosis increases patients' access to medical knowledge. In this study, a chatbot service was developed for the Covenant University Doctor (CUDoctor) telehealth system based on fuzzy logic rules and fuzzy inference. The service focuses on assessing the symptoms of tropical diseases in Nigeria. Telegram Bot Application Programming Interface (API) was used to create the interconnection between the chatbot and the system, while Twilio API was used for interconnectivity between the system and a short messaging service (SMS) subscriber. The service uses the knowledge base consisting of known facts on diseases and symptoms acquired from medical ontologies. A fuzzy support vector machine (SVM) is used to effectively predict the disease based on the symptoms inputted. The inputs of the users are recognized by NLP and are forwarded to the CUDoctor for decision support. Finally, a notification message displaying the end of the diagnosis process is sent to the user. The result is a medical diagnosis system which provides a personalized diagnosis utilizing self-input from users to effectively diagnose diseases. The usability of the developed system was evaluated using the system usability scale (SUS), yielding a mean SUS score of 80.4, which indicates the overall positive evaluation.

1. Introduction

Remote diagnosis systems are becoming increasingly popular and accurate, with enormous advantages such as cost-effectiveness, fast and reliable decision support for medical diagnostics, and treatment and prevention of disease, illness, injury, and other physical and mental damages in human beings. The rise in remote health services (or telehealth) offered by healthcare institutions coincided with the evolution of assisted living systems and environments, aiming to widen the possibility for older and disadvantaged people to access appropriate healthcare services and thus improve their health status and clinical outcome [1]. With the increase in the innovation of medical technologies, there is a

need to adopt medical expert systems that will oversee and control diagnosis and treatment processes [2]. Medical diagnostic processes carried out with the aid of computer-related technology which is on the rise daily have improved the experience and capabilities of physicians to make an effective diagnosis of diseases while employing novel signal processing techniques for analysis of patients' physiological data [3, 4] and deep neural networks for decision support [5]. With the rise of the artificial intelligence (AI) techniques, the chatbots have appeared as a promising direction in streamlining the communication between doctors and patients [6]. Such chats are becoming increasingly popular as remote health interventions are implemented in the form of the synchronous text-based dialogue systems [7]. Patients

with chronic diseases could make the most advantage from the use of chatbots which can continuously monitor their condition, provide reliable up-to-date information, and remind of taking medication [8]. For the effective use of chatbots in the healthcare domain, the chatbot technology need advanced reasoning capabilities based on the formalization of medical knowledge (semantics) and health state of patients coupled with language vocabularies and dialogue engines [9].

The natural language processing (NLP) technology can serve as an interaction between computers and humans using linguistic analysis and deep learning methods to obtain knowledge from an unstructured free text [10]. The NLP systems have shown their uniqueness and importance in the areas of information retrieval mostly in the retrieval and processing of large amount of unstructured clinical records and return structured information by user-defined queries. In general, the NLP system is aimed at representing explicitly the knowledge that is expressed by the text written in a natural language. There are few applications of the NLP techniques in diagnosing diseases despite the enormous amount of text-based information, which can be retrieved from patients' self-narrations [11]. The main challenges addressed by the application of NLP for medical records are flexible formatting, structure without sentences, missing expected words and punctuation, unusual parts of speech (POS), medical jargon, and misspellings [12]. Linguistic structures such as coreferences make medical texts difficult to be interpreted [13]. Moreover, unique linguistic entities such as medical abbreviations make the inference of knowledge from medical texts much harder [14].

This study introduces the use of the NLP model via an SMS and a chatbot platform to improve the health self-assessment and decision support in digital healthcare systems. The extraction of knowledge from the electronic health record (EHR) is a growing area of interest in medicine, and the use of electronic medical records (EMRs) at the healthcare center and on the cloud [15] has provided a vast amount of data to be analyzed. An EMR is a digital record of health-related information that is created, collected, and managed by medical experts [16]. Compilation of existing and available medical data complications includes integrating NLP into multiple EMRs, ensuring privacy and security of patients' data [17] and clinical validation of a tool. All these can be overwhelming to medical research for improving patient care. However, the application of NLP techniques to screen patients and assist medical experts in their diagnosis would serve as a boost in successfully improving healthcare services through effective analysis of narrative text of symptoms provided by a patient.

For example, Langer et al. [18] used several NLP tools along with classification methods to process the drug-related questions. They developed a natural language-based interface that enables the users to phrase their queries and get an accurate result up to 81% in classifying drug-related questions. Pendyala et al. [19] presented an application which allows machines to take on the function of life support. The focus of the study was based on medical diagnosis, and an experiment was conducted to show the relationship of

information retrieval and text mining to the medical diagnosis problem. The study concludes that the proposed system would help in improving the goals of providing a ubiquitous medical diagnosis. Fernandez-Millan et al. [20] presented a rule-based expert system using the list of likely diseases regarding laboratory test results for diagnosis. The authors concluded that the proposed system clinically gave a better accuracy and speed, thereby improving the efficiency and quality of service. Atutxa et al. [21] used deep learning models to extract the International Classification of Diseases (ICD) codes from the death certificates written in a regular natural language, obtaining an F-score of 0.963, 0.838, and 0.952 for Hungarian, French, and Italian, respectively. Combi et al. [22] proposed an NLP method for the trans-coding of natural text descriptions of adverse drug reactions into MedDRA standard terms, reaching an average precision and recall of 91.8% and 86.9%. Evans et al. [23] analyzed patient safety incident reports written in free text to categorize incident type and the severity of outcome, reaching an accuracy of 0.891 and 0.708, respectively. Kloehn et al. [24] generate explanations for complex medical terms in Spanish and English using WordNet synonyms and summaries, as well as the word embedding vector as a source of knowledge. Sarker et al. [25] used fuzzy logic and set theory-based methods for learning from a limited number of annotated examples of unstructured answers in health examinations for recognizing correct concepts with an average F1-measure of 0.89. Zhou et al. [26] used deep learning models pre-trained on the general text sources to learn knowledge for information extraction from the medical texts. Lauraitis et al. [27] used text input acquired by a smartphone app for evaluation of cognitive and motor deficits for people showing symptoms of central nervous system (CNS) disorders as a part of self-administered cognitive testing (SAGE) tests.

The development of the medical domain-oriented conversational chatbots has been addressed by several researchers. Such conversational agents powered by AI techniques may serve patients with minor health concerns while allowing medical doctors to allocate more time to treat more serious patients [28] or find suitable donors [29]. A chatbot-powered healthcare service can promptly respond to the problems that arise in daily life and to the health state changes of people with chronic diseases such as obesity, diabetes, or hypertension [30]. For example, Ahmad et al. [31] developed a chatbot that is able to advice on the kind of drugs to be taken based on the data submitted by the user. Avila et al. [32] developed a chatbot to find the best prices for medicines and suggest their best possible substitutes. Bao et al. [33] suggested a hybrid model composed of a knowledge graph and a textual similarity model to construct a system for responding to medical questions using Hierarchical BiLSTM Attention Model (HBAM). Chaix et al. [34] developed a chatbot for patients with breast cancer to provide support and answers to their concerns on their disease as well as to remind taking the prescribed drugs. Denecke et al. [35] developed a mobile app with a chatbot that uses the elements of the cognitive behavior therapy to support mentally ill people in addressing their psychological

problems. Harilal et al. [36] developed a chatbot app aimed at supporting empathetic conversations, sensing the associated emotions, and extending medical advice for people with depression. Huang et al. [37] developed an AI-powered chatbot for promoting healthy lifestyle and providing advices for weight management. Hussain and Athula [38] developed a chatbot that uses Media Wiki API to extract information from Wikipedia to supplement the chatbot's knowledge for advising diabetic patients on diabetes management. Kökciyan et al. [39] integrated the data from commercial health sensors, EHR, and clinical guidelines with a conversational chatbot that provides further explanations about their overall well-being based on the argumentation-based dialogue. Ni et al. [40] suggested a knowledge-driven primary care chatbot system that has an analytic NLP-based engine for understanding the descriptions of patients' symptom, a reasoner for mapping symptoms to possible causes, and a question generator for creating further dialogue questions. On the other hand, Zini et al. [41] developed a deep learning framework-based conversational agent to represent a virtual patient that can be used for teaching medical students on patient examination.

Machine learning algorithms, especially SVM, have shown promising results in classifying free text such as Georgian language in medical records [42]. SVM with a polykernel was used for classifying primary care patient safety incident report content and severity [23]. However, authors claimed that improving definitions and increased training samples of select categories will further improve performance of the system. Deep learning methods were proposed by Zhou et al. [26]. The authors presented transfer learning methods based on the traditional multilayer neural network (NN) model to develop a clinical information extraction system. Other methods in existing studies include an interactive NLP tool for identifying incidents in radiology reports presented by Trivedi et al. [43]. The authors implemented and assessed usability based on an open-ended questionnaire and the system usability scale (SUS). The summary of the selected literature is depicted in Table 1.

Medical chatbot has been designed and implemented in various clinical areas for developing conversational tools with wide access to medical knowledge and healthcare issues. Existing chatbots are designed for either generic or specific disease purposes. A novel approach based on the AI method was proposed by Madhu et al. [46] for designing a simple and interactive medical assistance chatbot for medicine dosage intake considering the age and weight of patients. Mandy chatbot system was designed to assist healthcare staff in automating patient intake process [40]. The proposed chatbot is based on three sections which are the mobile app front for patient interaction, the diagnosis section, and the doctor's interface for assessing patient's records. The chatbots combine NLP with knowledge-driven diagnosis abilities. Similarly, Siangchin and Samanchuen [47] developed a chatbot application using the auxiliary NLP library. The system was further compared with traditional ICD-10 application based on analytic hierarchy process (AHP) for analyzing, selecting, and classifying diabetes mellitus, trauma, and external causes. The integration of

NLP and machine learning algorithms has also played a key role in creating chatbot application for disease prediction and treatment recommendation [48]. Deep learning framework was proposed by Zini et al. [41] for enhancing virtual patients' conversational skills. The authors integrated long-short-term memory networks and CNN for sentence embeddings in a given QA script. Other methods include a study by Roca et al. [8] which introduced the chatbots-patient interaction system for specific chronic disease psoriasis. Further study from Rosruen and Samanchuen [49] also developed an intent-based approach chatbot known as MedBot using Dialogflow for medical consultant services. The authors claimed that the proposed system was able to maximize user's convenience, increase service capability, and reduce operational cost.

The successful adoption of chatbot technology from Table 2 has shown effective interaction between users and machines especially in various domains within the healthcare system. However, there are some limitations with some of the methods proposed in the literature such as challenges associated with the static local knowledge based in chatbots and time consumption during training especially for a specific domain [38]. Therefore, there is a need for a future study to develop chatbot software with more scalability, increased data sharing and reusability, and improved standard conversation model [8].

The continuous growth of mobile technology has affected every facet of human life around the globe as its support of healthcare objectives through telemedicine, telehealth, and m-health [52] has helped to diagnose and treat patients at low cost especially in the developing countries, where there are limited options of diagnosis and treatment. Out of various communication media available on mobile devices, short messaging service (SMS) has proven to be unique and reliable due to its low cost, reliable delivery, personal to users, and not Internet-oriented service [53, 54]. Considering the need to provide good medical care to everyone including rural dwellers with poor electricity and slow Internet connections, it is therefore important to integrate SMS with a medical diagnosis system, thus establishing an SMS-medical diagnosis system to best meet the needs of a common man [55]. Considering the overall progress and research efforts made by researchers in improving e-health systems and designing decision support systems (DSS) [56–58], there is still much work to be done for effective understanding and identifying key features based on NLP for enhancing diagnosis, thus improving good health and well-being of the global society at large.

Summarizing the existing medical diagnosis systems (MDS) often adopts poor decisions due to interpretation of the text-based input provided by the patient. Therefore, there is a need to automate MDS for efficient diagnosis of diseases and support their decisions based on the severity of symptoms. Moreover, the medical experts need a platform to keep track of large text-based chunk of knowledge narrated by patients in a natural language, hence improving healthcare delivery for remote patients.

The contribution of this paper is as follows: (1) we have developed a text-based medical diagnosis system which

TABLE 1: Summary of related work on the medical diagnosis system.

References	Methods	Contributions	Limitations
Ayush et al. [44]	Developed integral model including probability and fuzzy models for determination of human constitutional types	Proposed MDES system creates and supports decision system to users via providing reliable information about disease manifestation	It has not enough practical evidence for effectiveness and efficiency
Korenevskiy [45]	Synthesis of fuzzy decision rules	Simple to calculate with high possibility of diagnosis and predetermined level of reliability	It requires larger training samples
Atutxa et al. [21]	ICD-10 encoding based on neural networks	Multilingual ICD-10 coding. The method is interpretable and it outperforms alternative approaches.	Worse performance was detected on larger datasets
Combi et al. [22]	MagiCoder, an NLP algorithm	Simple, efficient in terms of computational complexity for Italian pharmacovigilance language	Inability to handle negations in textual medical records
Lu et al. [14]	Combined classic enhanced sequential inference model (ESIM) and BiLSTM network	Achieved higher accuracy compared to existing methods without knowledge enhanced	Challenges of concepts with multiple definition was not addressed
Kloehn et al. [24]	Proposed a novel algorithm SubSimplify	Improved quality in English and Spanish by providing multiword explanations for difficult terms	There is a possibility of the proposed model generating incomplete explanations
Sarker et al. [25]	Combination of fuzzy matching and intersection	Increased accuracy against human annotations	Inability to detect negations in expressions

TABLE 2: Summary of the literature based on medical chatbots.

References	Method	Contributions	Domain
Bao et al. [33]	Hybrid model chatbot that combines knowledge graph and a text similarity model	Proposed method was able to identify and reduce similarity in large QA dataset	Generic
Harilal et al. [36]	Developed CARO, a chatbot app, which is capable of performing empathetic conversations and providing medical advice	Proposed method has the ability to sense the conversational context, intent, and associated emotions	Depression
Bibault et al. [28]	Vik chatbot: blind, randomized controlled noninferiority trial	Proposed method was able to improve conversation between chatbot and physician	Breast cancer
Bali et al. [50]	Ensemble learning using advanced NLU techniques	Improved accuracy in diabetes prediction when compared to generic health prediction	Diabetes and generic
Cameron et al. [51]	Proposed iHelper using questionnaire developed by chatbottest	Recommendations to increase the usability of a chatbot for mental healthcare	Mental healthcare
Chaix et al. [34]	Vik chatbot	Evaluated a yearlong of conversations between patients with breast cancer and a chatbot	Breast cancer
Chung and Park [30]	Chatbot-based healthcare service framework in cloud	Provides a smooth human-machine interaction for the chatbot healthcare service	Accident response
Hussein and Athula [38]	Virtual Diabetes Management System (VDMS) using modified open source AIML web-based chatbot	Proposed method provides a more robust knowledge using Wikipedia knowledge	Generic and diabetes patients
Huang et al. [37]	AI-based health chatbot known as "Smart Wireless Interactive Healthcare System" (SWITCHes)	Proposed system can provide advice to user on food intake based on calorie order, advice on physical activities, etc.	Weight control and health promotion
Ahmad et al. [31]	NLP	Ability to give adequate advice on the right type of medication based on information provided	Pharmacy

provides a personalized diagnosis utilizing self-input from users to effectively diagnose diseases. (2) The proposed system combines the NLP and machine learning algorithms for SMS and Telegram bot. (3) The system is able to diagnose using a direct approach of the question and answering technique to suggest a medical diagnosis.

The structure of the remaining parts of the paper is as follows. We present and discuss our methodology and the algorithm used in Section 2. In Section 3, we evaluate and discuss the results. We present conclusions and outline future work in Section 4.

2. Design Methodology

2.1. Outline of the Architecture. The study assesses the clinical data needs and requirements in diagnosing the tropical diseases in Nigeria and assesses the patients' clinical data found in EHRs or manual records. The architecture of the proposed text-based medical diagnosis system is depicted in Figure 1.

The steps involved in the proposed text-based medical diagnosis system are as follows: (1) description of the knowledge base; (2) preprocessing of text-based documents;

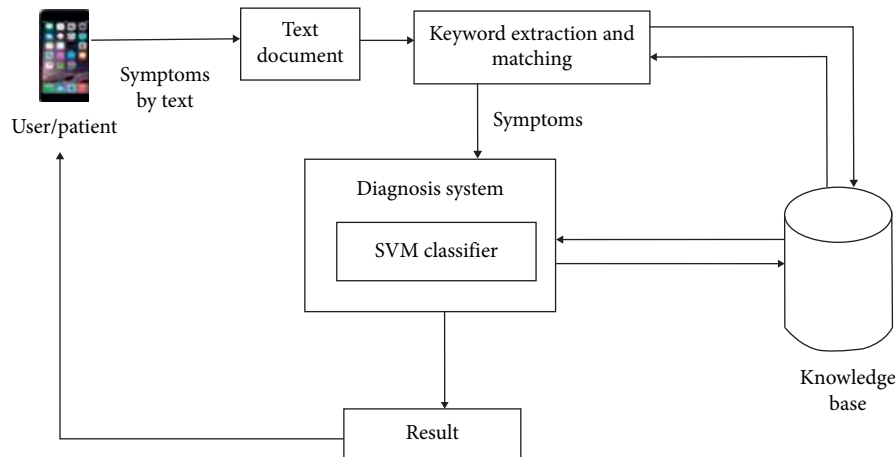


FIGURE 1: The architecture of the proposed system.

(3) tagging of document; (4) extraction of answer; and (5) ranking of candidate answers.

The implementation of the diagnosis system framework was done using the *Python* language due to the following functionalities: cross-platform and high availability of third-party libraries for tasks relating to machine learning and NLP. The system uses *Python* library packages to access the machine learning functions and NLP needed for categorization.

2.2. Knowledge Base. The knowledge base is the principal source of data in a question and answering system, and it could be in the structured or unstructured format. To develop the knowledge base, the information from a medical database system was collected and divided into categories which are referred to as context knowledge of the disease.

The main sources of information in the knowledge base are as follows: (1) WordNet [59], which provides a lexical database and defines the relationships of words and phrases; (2) YAGO (Yet Another Great Ontology) [60], an open source knowledge base which we use to construct a knowledge graph of common knowledge entities; (3) UMLS (Unified Medical Language System), which combines many medical vocabularies, including ICD-10-CM and SNOMED CT, and is used to link medical terms and extract medical concepts, relationships, or knowledge; UMLS is recognized as a comprehensive knowledge source in the healthcare domain [61]; and (4) Disease Ontology (DO) [62], which contains a knowledge base of over 10,000 human diseases.

We used WordNet as WordNet is used in most of the question-answering systems, and it has proved to be useful when dealing with words. An access to WordNet is implemented via WordNet HTTP API. An access to YAGO is implemented using a SPARQL [63] query engine, which sends queries to the SPAQL endpoint and returns the semantic fact triples (subject-predicate-object). An access to UMLs is implemented via UMLS REST API using the *Python* language. An access to DO metadata for a specific DO term is by using the REST metadata API by constructing a HTTP request. The knowledge base is specified using

eXtensible Meta Language (XML), which ensures a common way to specify and share structurally organized data that are not dependent of an application.

For knowledge representation, we have adopted a three-layer model of disease-symptom-property (DSP) originally suggested in [64] as shown in Figure 2. The knowledge in the knowledge database is stored as resource description framework (RDF) triples (property, symptom, and disease), while the computational model is adopted from the disease compass [65], which allows us to query the causal chains of diseases.

Our knowledge base includes 71 instances of diseases (mainly tropical ones) and 542 pieces of information. The model (see Figure 3) uses the knowledge from the knowledge database and applies the fuzzy rules described in Section 2.6.

2.3. Communication System. The communication system was implemented based on the knowledge base for the efficient communication of users through Telegram or SMS with the medical doctor using question-answer rules. Each diagnosis question has certain features and attributes that give additional information about the request. The different attributes of a request/question can be as follows:

- (1) Diagnosis question: the actual diagnosis question that will be sent across to the user
- (2) Response: the list of responses that will be shown to the user that denote the answers that he/she can send to the system via the Telegram GUI or SMS
- (3) Serial id: the order in which the question should be asked

The types of questions are as follows.

2.3.1. Basic Data Questions. These are the preliminary questions asked once a communication has been established by the user and the system. These questions are basic information about the users, some of which include information such as gender, age, height, and weight. The sample question sequence is demonstrated in Algorithm 1.

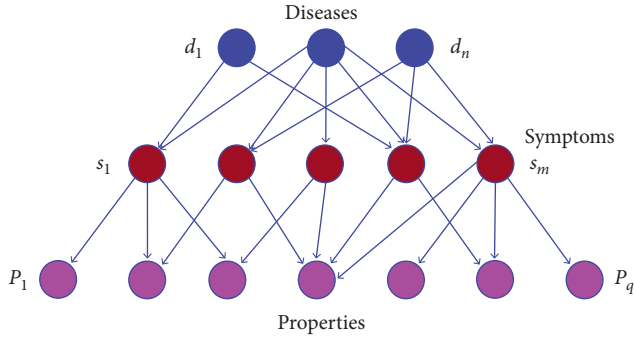


FIGURE 2: Three-layered “disease-symptom-property” model of knowledge representation.

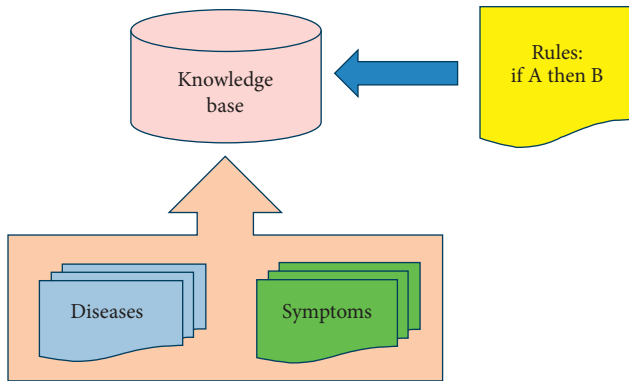


FIGURE 3: Integration of the knowledge representation model with fuzzy logic rules.

2.3.2. Symptom-Related Questions. These questions are the questions that will be asked from users to confirm whether they are showing signs of a symptom or not. They can gather specific responses based on predetermined constraints or either binary (yes/no) responses. These questions are subdivided into two types:

- (i) Target diagnosis questions: these are questions that confirm if a symptom exists or not
- (ii) Linked diagnosis questions: these are question used for asking more information about a symptom in case the user replies affirmatively to the target question of that symptom

Each of these questions is designed such that they give an affirmative response of the existing symptoms or non-existing symptoms. For instance, a top diagnosis question that will ask the users of the presence of certain symptoms is represented in Algorithm 2.

2.4. Content Extraction and Text Preprocessing. The content package is responsible for extracting the content knowledge from an SMS. For convenience, the package provides several different content extractors that specialize on extracting different sets of data from an SMS. The content extractors perform text processing using the NLP package. When a patient sends an SMS of his or her symptoms, the SMS

receiver of the system receives it and passes the text (SMS body) to the NLP module, which looks at the text, makes necessary corrections if needed, and extracts important keywords.

The text processing operations include three major steps: noise removal, tokenization of document sentences, and splitting of sentences.

2.4.1. Noise Removal. Text contains a sequence of characters, both relevant and irrelevant characters. Therefore, noise is removed from the raw documents leaving only relevant content which is related to subjects for further processing.

2.4.2. Tokenization. This involves the fragmentation of strings of characters into its lexical elements. In this context, sentence splitting process was used for splitting of the text into a separate sentence. Here, we used the natural language toolkit (NLTK) tokenizer.

2.4.3. Tagging of Document. Useful information from the knowledge source is tagged for identifying quality information from the specific document. The names of diseases were used as labels for documents, and tools such as parser and WordNet were used for document tagging.

2.4.4. Parser. Stanford Parser was used as a tool for generating parts of speech (POS) of each word inputted by the user query, and, in addition, the candidate answers selected from the knowledge database. The WordNet was used to discover the relationship between the words of the user query and the data source. Words are grouped into nouns, verbs, adjectives, and adverbs.

2.4.5. Term Matching. Then, the system performs querying of the knowledge base to match the extracted words with information stored in the knowledge base.

2.5. Feature Selection and Extraction. The extracted collection of important words then was transformed into a feature vector suitable for use with the machine learning algorithm. For transformation into the feature vector, we used the word embedding [66] technique. Word embeddings are the representation of words in the semantic multidimensional space. For our system, we have adopted an available word embedding Glove [67], which is based on Twitter data, because it provides a good approximation to common English for an informal communication channel. For an effective training of data, the text messages were pre-processed and converted into feature vectors. This was achieved based on feature extraction using the word embeddings. The feature vector of the unlabeled document was then given to the classifier’s decision function to return a category for the unlabeled document.

```

def question_data ():
    return {
        'user_age': {
            'diagnosis_question': "What is your age?",
            'diagnosis_response': ['15-25', '25-40', '40-50', '>50'],
            'serial': 1
        },
        'user_weight': {
            'diagnosis_question': "What is your weight?",
            'diagnosis_response': ['40-50 kg', '50-70 kg', '70-90 kg', '>90 kg'],
            'serial': 2
        },
        'user_height': {
            'diagnosis_question': "What is your height?",
            'diagnosis_response': ['4-5 ft', '5-6 ft', '6-7 ft'],
            'serial': 1
        },
        'user_gender': {
            'diagnosis_question': "What is your gender?",
            'diagnosis_response': ['Male', 'Female', 'Unspecified'],
            'serial': 3
        }
    }

```

ALGORITHM 1: Algorithm showing the samples of basic data questions.

```

def diagnosis_top_data ():
    return {
        'fever': {
            'diagnosis_question': "Do you have a fever?",
            'diagnosis_response': ['Yes, High (>103 F)', 'Yes, Mild (101-103 F)',
                                   'Yes, Very Mild (99 - 101 F)', no],
        },
        'head_ache': {
            'diagnosis_question': "Do you have a head ache?",
            'diagnosis_response': [yes, no],
        },
        'body_chills': {
            'diagnosis_question': "Are you experiencing body chills?",
            'diagnosis_response': [yes, no],
        },
        'diarrhea': {
            'question': "Are you having very frequent loose motions?",
            'response': [yes, no],
        },
        'extreme_weakness': {
            'diagnosis_question': "Do you experience extreme weakness?",
            'diagnosis_response': [yes, no],
        },
    }

```

ALGORITHM 2: Algorithm showing the samples of top diagnosis questions.

2.6. Fuzzy Reasoning Module. Its primary purpose is to use fuzzy logic-based algorithms [68] to read and interpret the responses from the user, track and monitor all the symptoms that the user has already responded to, and to administer questions to the user that are most relevant based on the

dataset of diseases that is maintained. Each disease is modeled as a bucket, where each bucket is associated with a symptom. The appropriate fuzzy rules are formulated to deal with multiple symptomatic diseases. The algorithms employed by CUDoctor read the state of these buckets and

send the most relevant question to the user. This helps us narrow down the number of questions that would be asked by the system to reach a diagnosis.

The weighted fuzzy logic rule system is used, where each fuzzy rule has a weight assigned based on the historic data. The Mamdani fuzzy logic model of fuzzy inference was used, in which the IF-THEN statement represents each rule. The fuzzy rules were specified as “If x_1 is A_1 and y_1 is B_1 then z_1 is C_1 ,” where A_1 , B_1 , and C_1 are fuzzy sets. The fuzzy rules are weighted by an assessment of the level of contribution of the properties and symptoms to diagnosing the disease. The crisp dataset D is described by n features and k samples $[F_1, F_1, \dots, F_n]$, and the n -dimensional tuple $T_i = [a_1, a_1, \dots, a_n]$ is represented as a kn -dimensional feature vector:

$$T_i = [\langle \mu_{FT1}(a_1), \mu_{FT2}(a_1), \dots, \mu_{FTk}(a_1) \rangle, \dots, \langle \mu_{FT1}(a_n), \mu_{FT2}(a_n), \dots, \mu_{FTk}(a_n) \rangle], \quad (1)$$

where $\mu_{FTk}(a_i)$ is the membership degree of the fuzzy term FTk of feature F_i ($F_i = a_i$). If the fuzzy variable F_n has k fuzzy terms, FT_1, FT_2, \dots, FT_k , then for each value v of F_n , the fuzzy value is computed as $\max\{\mu_{FT1}(v), \mu_{FT2}(v), \dots, \mu_{FTk}(v)\}$.

The max-min operators were used for implication. In order to obtain the crisp output, the center-of-gravity (centroid) defuzzification was employed, where the weights are expressed by the degree of membership of the value x_i with the concept modeled by the fuzzy set A .

The process is implemented (Figure 4) as follows:

- (1) Fuzzification: transforms the crisp inputs into fuzzy values. Expert judgement is used for defining the degree of membership function. During fuzzification, a fuzzy rule controller receives input data (fuzzy variable) and analyzes it according to membership functions.
- (2) Knowledge base: it comprises a fuzzy definition database and a fuzzy IF-THEN rule base. The rule base describes the diseases for each combination of crisp input variables.
- (3) Inference engine: applies the appropriate fuzzy rules on the input data.
- (4) Defuzzification: produces the crisp output values from the fuzzy values as the results.

The final decision is made by selecting the fuzzy rule that achieved the highest score. The system architecture, which was adopted from [64], is presented in Figure 5.

2.7. Classification Module. In this study, the choice of our classifier was dependent on the practical requirements of the proposed application and the need for a classifier with better results based on the established literature for document classification. However, there were specific requirements in the proposed application which influence the selection of our classifier such as the computational complexity for the training and/or testing phase.

Machine learning was used to provide category prediction on the text messages using the fuzzy SVM classifier

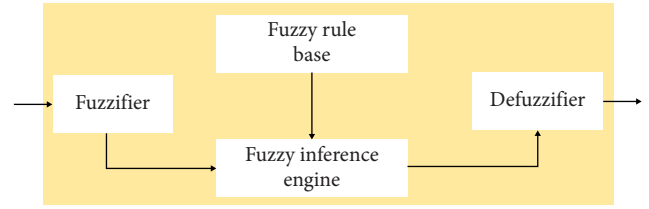


FIGURE 4: Structure of the fuzzy logic-based reasoning module.

[69]. It has only one hyperparameter C , which is a regularization parameter. The value of the hyperparameter was tuned by grid search. In a fuzzy SVM, a fuzzy membership values are used for each data point of SVM, and the SVM is reformulated such that different input points can make different contributions to the learning of the decision surface. Here, the fuzzy membership values are calculated using an algorithm suggested by Le et al. [70]. First, we use clustering techniques to find clusters of data. Fuzzy membership values of data points belonging to clusters are set to 1, and fuzzy memberships of other data points are determined by their distance to the closest cluster, respectively.

The classifier was trained using a set of training documents which have been processed by the NLP package and converted into word embedding used as feature vectors. The length of the feature vector is 300 as suggested by Mikolov et al. [66]. Finally, the feature vectors extracted from the user answers are passed to the fuzzy SVM model, which suggests the diagnosis by performing the classification on those important words contained in the SMS and then sends the result to the patient via an SMS.

2.8. Graphical User Interface (GUI). The system initiates a communication/conversation with the user/patient to obtain more insights about their basic personal data such as gender, age, height, and weight. Once the basic data are acquired, the CUDoctor moves to the second stage and proceeds to query the patient for symptoms based on the above algorithms. The Telegram API was used for the GUI design with an additional custom keyboard that is also provided by the Telegram API. The screenshot of a question-answer subsystem is shown in Figure 6.

The SMS text request and response were integrated for communication since the application does not require an Internet connection, and it is compatible with all mobile devices. The python-telegram-API library was used as the *Python* wrapper that executes communication with the Telegram API. It enables easy setting method hooks that are triggered whenever a function is executed on the part of the Telegram chatbot. In the same way, the application sends a request to the Twilio communication API for the SMS text formatting server for incoming requests and passes responses to the logical layer for actual processing and presentation of the result to the client users. The python-Twilio-communication-SMS-API library was used as the *Python* wrapper that executes communication with the Twilio communication SMS API. The SMS interface for diagnosis conversation is represented in Figure 7.

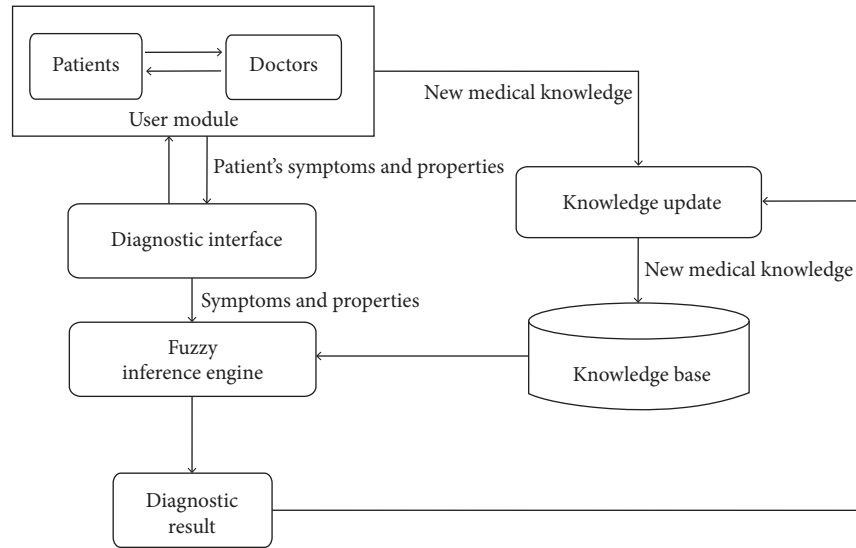


FIGURE 5: Architecture of the fuzzy logic inference system.



FIGURE 6: Telegram graphical user interface.

3. Evaluation and Results

3.1. Data Collection. The data used in this study were collected from a medical database, and an interview was conducted for extraction of text content from experts and individuals with knowledge about the various diseases. The extracted text content was then stored on the local file of the system.

3.1.1. Selection of Participants. Selection of respondents was based on snowball sampling. The inclusion criteria require participants who have recently been diagnosed of the disease

we are working on, and/or specialist who is engaged in clinical research, or having an indebt experience with the diseases by coming in contact with or treating patients who in recent times have been diagnosed. To be included for the questionnaire, the participants must be actively working in the hospital, responsible for treating patients who have been to the hospital for treatment on the various selected diseases. Individuals who do not match any of these inclusion criteria were excluded. Information extracted for the implementation of the system was the selected result from some individuals' experience about the disease (individuals who can provide expert information about the selected diseases and their varying symptoms and individuals who have recently been diagnosed or hosted in the hospital for the selected diseases were recruited).

3.2. Evaluation of Results. To evaluate the performance of the developed service, we used the Bilingual Evaluation Understudy (BLEU) score [71], which has become a typical metric for evaluating chatbot services [72, 73]. BLEU scores an output response from the service as compared to the reference, where a BLEU score ranges from 0 to 1. Here, we used BLEU-2, which is based on unigram and bigram matches between the generated and the reference sentences. We also used the Recall-Oriented Understudy for Gisting Evaluation (ROUGE-L) metric [74], which is based on the longest common subsequence (LCS). The results are 25.29 for BLEU-2 and 31.56 for ROUGE-L.

3.3. Usability Testing. To perform the usability testing of the developed system, we followed the recommendations outlined by Cameron et al. [51]. We used the system usability scale (SUS), which is a questionnaire to evaluate the ease of use of a system using a Likert scale (five-point scale that varies from strongly disagree to strongly agree).

The SUS score is able to evaluate usability performance in terms of effectiveness, efficiency, and overall ease of use.

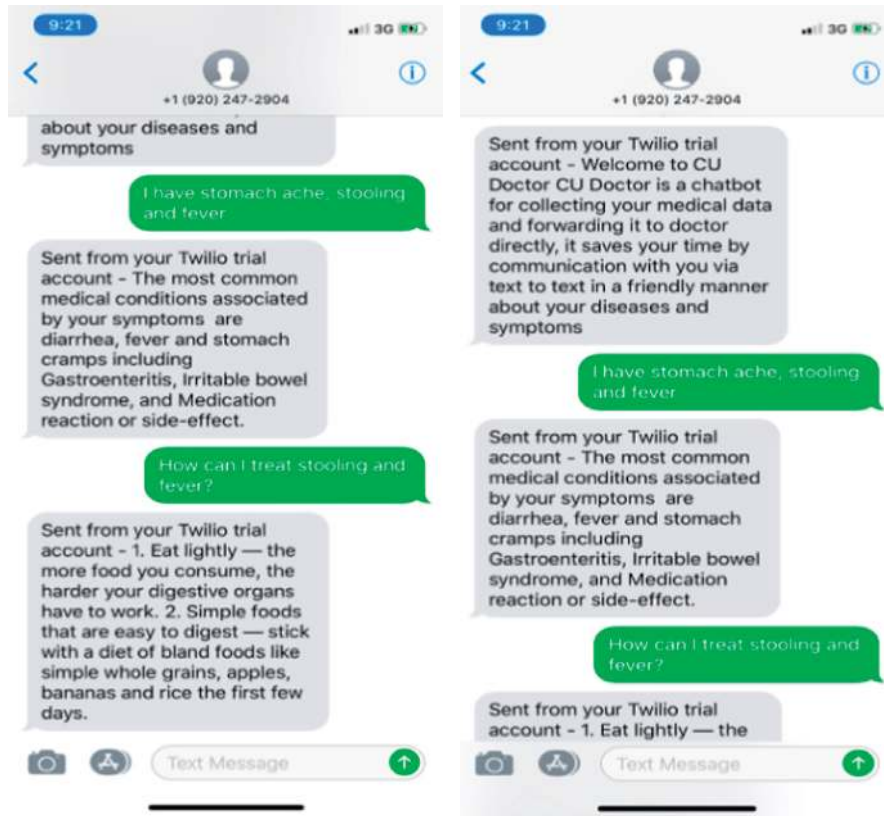


FIGURE 7: An example of SMS diagnosis conversation in natural language.

SUS is considered as a reliable tool for measuring the usability, and it allows to evaluate a wide variety of products and services. SUS has become an industry standard, with references in over 1300 articles and publications [75], which also includes medical chatbots and other NLP-based medical diagnostics systems such as presented by Tielman et al. [76] and Valtolina et al. [77]. SUS is designed to support assessment and comparison of the user experience when interacting with different tools and is recommended to be included in any evaluation of health chatbots [44].

SUS has only 10 questions. The results of an SUS fall between 0 and 100, while a score of 68 is considered average. SUS has already been applied before to assess the usability of chatbots [78, 79]. The SUS questions are easily adaptable for use with different types of systems; therefore, it was adapted to be used for our study. The SUS questions we asked are listed in Figure 8.

The participants of the usability test comprised 27 participants, including 11 females and 16 males, 13 were aged between 25 and 34, 9 aged between 35 and 44, and 6 aged between 45 and 53. The study participants have provided an informed consent to take part in this research. Usability study lasted less than 45 min per session. Information booklets informing about the study performed were distributed to the participants, and the participants were requested to communicate with the chatbot by using a mobile phone. All received answers were anonymized.

The SUS scores were computed using the procedure provided by Brooke [80]: subtracting 1 from the score for

questions 1, 3, 5, 7, and 9, whereas subtracting the score from 5 for questions 2, 4, 6, 8, and 10 and multiplying the total of the scores by 2.5 to obtain the final evaluation score.

The SUS score is assessed as follows: >80.3 (excellent), 68–80.3 (good), 68 (okay), 51–68 (poor), and <51 (bad). As a result of evaluation, the mean SUS score obtained for CUDoctor was 80.4, which is above the threshold of 68, which means that the overall evaluation was excellent. The results for all SUS questions are presented in Figure 9. They show that the user provided the worst evaluation for Q2 (“I found CUDoctor unnecessarily complex”) giving a score of 72.7, but still above the threshold of 68. The best evaluated feature was integration (Q5) with a score of 91.6.

3.4. Comparison with Other NLP-Based Services. Several chatbots with medical-related applications are provided on social networking platforms such as Facebook. For example, the FLORENCE bot that reminds the users when to take their medication and monitors their weight and moods. SMOKEY warns the users on bad air quality. HealthTap provides answers using a database of knowledge that contains similar questions. Google provides the Dialogflow Application Programming Interface (API) for the integration of NLP to the target applications. Woebot provides a cognitive behavior therapy service for patients with and has been tested with depression [81]. It allowed to reduce their symptoms of depression as evaluated by the depression questionnaire PHQ-9. XiaoIce is a social chatbot that

1. I think that I would like to use CUDoctor frequently.
2. I found CUDoctor unnecessarily complex.
3. I thought CUDoctor was easy to use.
4. I think that I would need the support of a technical person to be able to use CUDoctor.
5. I found the various functions in CUDoctor were well integrated.
6. I thought there was too much inconsistency in CUDoctor.
7. I imagine that most people would learn to use CUDoctor very quickly.
8. I found CUDoctor very cumbersome to use.
9. I felt very confident while using CUDoctor.
10. I needed to learn many things before I could use CUDoctor.

FIGURE 8: The questionnaire form for evaluation of CUDoctor.

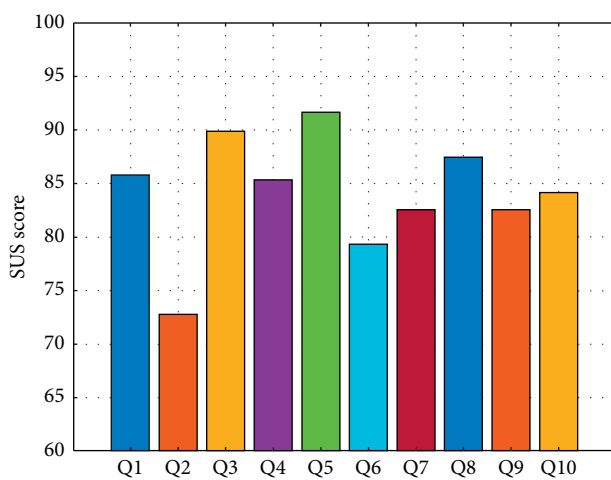


FIGURE 9: Results from usability evaluation of CUDoctor using SUS.

emphasizes emotional connection [82], while using deep learning for meaningful response dialogue tasks. Chatbots are also used in suicide prevention and cognitive behavioral therapy, aiming at risk groups such as HARR-E and Wysa [83]. The main difference of the system described in this paper is that the service is delivered over the SMS rather than social networks, which require very good Internet connectivity often unavailable in remote rural regions of developing countries. Moreover, the described solution focuses on the niche domain of tropical disease symptom assessment, and we are not aware of any other NLP-based systems focusing on this domain of application.

4. Conclusion and Future Work

The timely access to healthcare avoiding unnecessary time wastage of patients is a major issue in sub-Saharan Africa. However, considering the exponential growth of mobile users and the need for a real-time medical diagnosis assistance tool, it is therefore important to explore the need for a cost-effective telehealthcare platform, which allows the earlier detection of diseases and effective communication with patients (users) to a diagnosis system (remote doctor at

proxy). Based on the highlighted needs, this study was able to successfully build a text-based medical diagnosis system, which provides a personalized diagnosis utilizing self-input response from users to effectively suggest a disease diagnosis. The proposed system was able to combine NLP and machine learning algorithm for SMS and Telegram bot. The system was able to suggest a diagnosis using a direct approach of the question and answering technique to offer a diagnosis. The limitation of the system is that it is not secure against the false-positive cases of falsely suggesting the disease; therefore, a final diagnosis still must be confirmed by the medical doctor.

The future recommendations include the automation of this medical diagnosis system to easily recognize diseases, recommend treatments, prescribe a medication, and perform medication adherence. Audio interaction will be incorporated to make the system more interactive. These improvements will serve towards reducing cost and mortality rate, thereby reducing the workload burden on medical doctors in underdeveloped regions.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The authors gratefully acknowledge the support and sponsorship of Covenant University through the Centre for Research, Innovation, and Discovery (CUCRID).

References

- [1] R. Maskeliunas, R. Damaševičius, and S. Segal, "A review of internet of things technologies for ambient assisted living environments," *Future Internet*, vol. 11, no. 12, 2019.
- [2] A. Keleş, "Expert doctor verdis: integrated medical expert system," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 22, no. 4, pp. 1032–1043, 2014.

- [3] H. Zhao, D. Li, W. Deng, and X. Yang, "Research on vibration suppression method of alternating current motor based on fractional order control strategy," *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, vol. 231, no. 4, pp. 786–799, 2016.
- [4] H. Zhao, S. Zuo, M. Hou et al., "A novel adaptive signal processing method based on enhanced empirical wavelet transform technology," *Sensors*, vol. 18, no. 10, p. 3323, 2018.
- [5] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, "An improved quantum-inspired differential evolution algorithm for deep belief network," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, 2020.
- [6] A. Greene, C. C. Greene, and C. Greene, "Artificial intelligence, chatbots, and the future of medicine," *The Lancet Oncology*, vol. 20, no. 4, pp. 481–482, 2019.
- [7] S. Hoermann, K. L. McCabe, D. N. Milne, and R. A. Calvo, "Application of synchronous text-based dialogue systems in mental health interventions: systematic review," *Journal of Medical Internet Research*, vol. 19, no. 8, p. e267, 2017.
- [8] S. Roca, J. Sancho, J. García, and Á. Alesanco, "Microservice chatbot architecture for chronic patient support," *Journal of Biomedical Informatics*, vol. 102, p. 103305, 2020.
- [9] A. Sheth, H. Y. Yip, and S. Shekarpour, "Extending patient-chatbot experience with internet-of-things and background knowledge: case studies with healthcare applications," *IEEE Intelligent Systems*, vol. 34, no. 4, pp. 24–30.
- [10] S. Doan, C. K. Maehara, J. D. Chaparro et al., "A natural language processing tool to identify patients with high clinical suspicion for Kawasaki disease from emergency department notes," *Academic Emergency Medicine*, vol. 23, no. 5, pp. 628–636.
- [11] R. G. Jackson, R. Patel, N. Jayatilleke et al., "Natural language processing to extract symptoms of severe mental illness from clinical text: the Clinical Record Interactive Search Comprehensive Data Extraction (CRIS-CODE) project," *BMJ Open*, vol. 7, no. 1, 2017.
- [12] R. Leaman, R. Khare, and Z. Lu, "Challenges in clinical natural language processing for automated disorder normalization," *Journal of Biomedical Informatics*, vol. 57, pp. 28–37, 2015.
- [13] V. Žitkus, R. Butkiene, R. Butleris, R. Maskeliūnas, R. Damaševičius, and M. Woźniak, "Minimalistic approach to coreference resolution in Lithuanian medical records," *Computational and Mathematical Methods in Medicine*, vol. 7, 2019.
- [14] M. Lu, Y. Fang, F. Yan, and M. Li, "Incorporating domain knowledge into natural language inference on clinical texts," *IEEE Access*, vol. 7, pp. 57623–57632, 2019.
- [15] N. Ouerhani, A. Maalel, and H. Ben Ghézela, "SPeCECA: a smart pervasive chatbot for emergency case assistance based on cloud computing," *Cluster Computing*, vol. 8, 2019.
- [16] M. V. Nectoux, M. F. Chiang, M. C. Lim et al., "Adoption of electronic health records and preparations for demonstrating meaningful use," *Ophthalmology*, vol. 120, no. 8, pp. 1702–1710, 2013.
- [17] L. Hang, E. Choi, and D.-H. Kim, "A novel EMR integrity management based on a medical blockchain platform in hospital," *Electronics*, vol. 8, no. 4, p. 467, 2019.
- [18] A. Langer, R. Banga, A. Mittal, L. V. Subramaniam, and P. Sondhi, "A text based drug query system for mobile phones," *International Journal of Mobile Communications*, vol. 12, no. 4, pp. 411–429, 2014.
- [19] V. S. Pendyala, Y. Fang, J. Holliday, and A. Zalzal, "A text mining approach to automated healthcare for the masses," *GHTC*, vol. 8, pp. 28–35, 2014.
- [20] R. Carson-Stevens, J.-A. Medina-Merodio, R. Plata, J.-J. Martinez-Herraiz, and J.-M. Gutierrez-Martinez, "A laboratory test expert system for clinical diagnosis support in primary health care," *Applied Sciences*, vol. 5, no. 3, pp. 222–240, 2015.
- [21] A. Atutxa, A. D. de Ilaraza, K. Gojenola, M. Oronoz, and O. Perez-de-Viñaspre, "Interpretable deep learning to map diagnostic texts to ICD-10 codes," *International Journal of Medical Informatics*, vol. 129, pp. 49–59, 2019.
- [22] C. Combi, M. Zorzi, G. Pozzani, U. Moretti, and E. Arzenton, "From narrative descriptions to MedDRA: automatically encoding adverse drug reactions," *Journal of Biomedical Informatics*, vol. 84, pp. 184–199, 2018.
- [23] H. P. Kanegaye, A. Anastasiou, A. Edwards et al., "Automated classification of primary care patient safety incident report content and severity using supervised machine learning (ML) approaches," *Health Informatics Journal*, vol. 23.
- [24] N. Kloeckner, G. Leroy, D. Kauchak et al., "Improving consumer understanding of medical text: development and validation of a new subsimplify algorithm to automatically generate term explanations in English and Spanish," *Journal of Medical Internet Research*, vol. 20, no. 8, p. e10779, 2018.
- [25] A. Sarker, A. Z. Klein, J. Mee, P. Harik, and G. Gonzalez-Hernandez, "An interpretable natural language processing system for written medical examination assessment," *Journal of Biomedical Informatics*, vol. 98, 2019.
- [26] L. Zhou, H. Suominen, and T. Gedeon, "Adapting state-of-the-art deep language models to clinical information extraction systems: potentials, challenges, and solutions," *JMIR Medical Informatics*, vol. 7, no. 2, 2019.
- [27] A. Lauraitis, R. Maskeliūnas, R. Damaševičius, and T. Krilavičius, "A mobile application for smart computer-aided self-administered testing of cognition, speech, and motor impairment," *Sensors*, vol. 20, no. 11, p. 3236, 2020.
- [28] J.-E. Bibault, B. Chaix, A. Guillemassé et al., "Chatbot versus physicians to provide information for patients with breast cancer: blind, randomized controlled noninferiority trial," *Journal of Medical Internet Research*, vol. 21, no. 11.
- [29] B. Ayeni, O. Y. Sowunmi, S. Misra, R. Maskeliūnas, R. Damaševičius, and R. Ahuja, "A web based system for the Discovery of blood banks and donors in emergencies," in *Advances in Intelligent Systems and Computing*, pp. 592–600, Springer, Berlin, Germany, 2020.
- [30] K. Chung and R. C. Park, "Chatbot-based healthcare service with a knowledge base for cloud computing," *Cluster Computing*, vol. 22, no. S1, pp. 1925–1937, 2019.
- [31] N. S. Ahmad, M. H. Sanusi, M. H. Abd Wahab, A. Mustapha, Z. A. Sayadi, and M. Z. Saringat in *Proceedings of the 2018 IEEE Conference on Open Systems Conversational Bot for Pharmacy: A Natural Language Approach ICOS*, pp. 76–79, New York, NY, USA, 2018.
- [32] C. V. S. Avila, A. B. Calixto, T. V. Rolim et al., "Medibot: an ontology based chatbot for Portuguese speakers drug's users," *ICEIS 2019-21st International Conference on Enterprise Information Systems*, vol. 1, pp. 25–36, 2019.
- [33] Q. Bao, L. Ni, and J. Liu, "HHH: an online medical chatbot system based on knowledge graph and hierarchical bi-directional attention," *Proceedings of the Australasian Computer Science Week Multiconference*, vol. 7, 2020.
- [34] B. Chaix, J.-E. Bibault, A. Pienkowski et al., "When chatbots meet patients: one-year prospective study of conversations between patients with breast cancer and a chatbot," *JMIR Cancer*, vol. 5, no. 1, 2019.

- [35] K. Denecke, S. Vaaheesan, and A. Arulnathan, "A mental health chatbot for regulating emotions (SERMO)-concept and usability test," *IEEE Transactions on Emerging Topics in Computing*, vol. 1, 2020.
- [36] N. Harilal, R. Shah, S. Sharma, and V. Bhutani, "Caro," *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, vol. 8, pp. 349–350, 2020.
- [37] C.-Y. Huang, M.-C. Yang, C.-Y. Huang, Y.-J. Chen, and M.-L. Wu, "A chatbot-supported smart wireless interactive healthcare system for weight control and health promotion," *2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, vol. 8, 2018.
- [38] S. Chen and G. Athula, "Extending a conventional chatbot knowledge base to external knowledge source and introducing user based sessions for diabetes education," *WAINA*, vol. 8, 2018.
- [39] N. Kõkciyan, M. Chapman, P. Balatsoukas et al., "A Collaborative Decision Support Tool for Managing Chronic Conditions," 2019.
- [40] L. Ni, C. Lu, N. Liu, and J. Liu, "MANDY: towards a smart primary care chatbot application," in *Knowledge and Systems Sciences*, vol. 780, Singapore, Springer, 2017.
- [41] J. E. Zini, Y. Rizk, M. Awad, and J. Antoun, "Towards A deep learning question-answering specialized chatbot for objective structured clinical examinations," *International Joint Conference on Neural Networks*, vol. 7, 2019.
- [42] M. Khachidze, M. Tsintsadze, and M. Archuadze, "Natural language processing based instrument for classification of free text medical records," *BioMed Research International*, vol. 8, 2016.
- [43] G. Trivedi, E. R. Dadashzadeh, R. M. Handzel, W. W. Chapman, S. Visweswaran, and H. Hochheiser, "Interactive NLP in clinical care: identifying incidental findings in radiology reports," *Applied Clinical Informatics*, vol. 10, no. 4, pp. 655–669, 2019.
- [44] A. Abd-Alrazaq, Z. Safi, M. Alajlani, J. Warren, M. Househ, and K. Denecke, "Technical metrics used to evaluate health care chatbots: scoping review," *Journal of Medical Internet Research*, vol. 22, no. 6, 2020.
- [45] N. A. Korenevskiy, "Application of fuzzy logic for decision-making in medical expert systems," *Biomedical Engineering*, vol. 49, no. 1, pp. 46–49, 2015.
- [46] D. Madhu, C. J. N. Jain, E. Sebastain, S. Shaji, and A. Ajayakumar, "A novel approach for medical assistance using trained chatbot," *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, vol. 2017, pp. 243–246, 2017.
- [47] N. Sheth and T. Samanchuen, "Chatbot implementation for ICD-10 recommendation system," *ICESI*, vol. 8, 2019.
- [48] 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)*, vol. 2019, pp. 851–856, 2019.
- [49] N. Rosruen and T. Samanchuen, "Chatbot utilization for medical consultant system," *TIMES-iCON*, vol. 8, 2019.
- [50] M. Bali, S. Mohanty, S. Chatterjee, M. Sarma, and R. Puravankara, "Diabot: a predictive medical chatbot using ensemble learning," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 6334–6340, 2019.
- [51] G. Cameron, D. Cameron, G. Megaw et al., "Assessing the Usability of a Chatbot for Mental Health Care," 2019.
- [52] E. Krupinski and R. Weinstein, "Telemedicine, telehealth and m-health: new frontiers in medical practice," *Healthcare*, vol. 2, no. 2, pp. 250–252, 2014.
- [53] J. Park, T.-O. Kim, N.-Y. Lee et al., "The effectiveness of short message service to assure the preparation-to-colonoscopy interval before bowel preparation for colonoscopy," *Gastroenterology Research and Practice*, vol. 2015, pp. 1–8, 2015.
- [54] B. Townsend, J. Abawajy, and T.-H. Kim, "SMS-based medical diagnostic telemetry data transmission protocol for medical sensors," *Sensors*, vol. 11, no. 4, pp. 4231–4243, 2011.
- [55] X. Miaoju, Z. Jianke, and T. Xiaoqi, "Intelligent fish disease diagnostic system based on SMS platform," in *Proceedings of the 2013 Third International Conference on Intelligent System Design and Engineering Applications (ISDEA)*, IEEE, New York, NY, USA, pp. 897–900, 2013.
- [56] O. M. Alade, O. Y. Sowunmi, S. Misra, R. Maskeliūnas, and R. Damaševičius, "A neural network based expert system for the diagnosis of diabetes mellitus," *Advances in Intelligent Systems and Computing*, vol. 7, pp. 14–22, 2018.
- [57] N. A. Azeez, T. Towolawi, C. Van der Vyver et al., "A fuzzy expert system for diagnosing and analyzing human diseases," *AISC*, vol. 939, pp. 1–11, 2019.
- [58] A. Lauraitis, R. Maskeliūnas, and R. Damaševičius, "ANN and fuzzy logic based model to evaluate huntington disease symptoms," *Journal of Healthcare Engineering*, vol. 2018, p. 1, 2018.
- [59] G. A. Miller, "WordNet," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.
- [60] F. M. Suchanek, G. Kasneci, and G. Weikum, "YAGO: a large ontology from wikipedia and WordNet," *Journal of Web Semantics*, vol. 6, no. 3, pp. 203–217, 2008.
- [61] D. Chen, R. Zhang, K. Liu, and L. Hou, "Knowledge Discovery from posts in online health communities using unified Medical Language system," *International Journal of Environmental Research and Public Health*, vol. 15, no. 6, p. 1291, 2018.
- [62] L. M. Schriml, C. Arze, S. Nadendla et al., "Disease Ontology: a backbone for disease semantic integration," *Nucleic Acids Research*, vol. 40, no. D1, pp. D940–D946, 2012.
- [63] C. P. De Laborda and S. Conrad, "Bringing relational data into the semanticweb using SPARQL and relational OWL," *ICDEW*, vol. 84, 2006.
- [64] Y. Jiang, B. Qiu, C. Xu, and C. Li, "The research of clinical decision support system based on three-layer knowledge base model," *Journal of Healthcare Engineering*, vol. 2017, pp. 1–8, 2017.
- [65] K. Kozaki, Y. Yamagata, R. Mizoguchi, T. Imai, and K. Ohe, "Disease Compass- a navigation system for disease knowledge based on ontology and linked data techniques," *Journal of Biomedical Semantics*, vol. 8, no. 1, p. 22, 2017.
- [66] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," 2013.
- [67] J. Pennington, R. Socher, and C. Manning, "Glove: global vectors for word representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543, Doha, Qatar, October 2014.
- [68] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [69] C. F. Lin and S. D. Wang, "Fuzzy support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 464–471, 2002.
- [70] T. Le, D. Tran, W. Ma, and D. Sharma, "A new fuzzy membership computation method for fuzzy support vector

- machines,” *Third International Conference on Communications and Electronics (ICCE 2010)*, vol. 8, 2010.
- [71] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, “Bleu: a method for automatic evaluation of machine translation,” *Association for Computational Linguistics*, vol. 8, pp. 311–318, 2002.
 - [72] J. A. Kapočiūtė-Dzikienė, “Domain-specific generative chatbot trained from little data,” *Applied Science*, vol. 10, p. 2221, 2020.
 - [73] J. Sedoc, D. Ippolito, A. Kirubarajan, J. Thirani, L. Ungar, and C. Callison-Burch, “ChatEval: a tool for chatbot evaluation,” *North American Chapter of the Association for Computational Linguistics (Demonstrations)*, vol. 8, pp. 60–65, 2019.
 - [74] C. Y. Lin and F. J. Och, “Automatic evaluation of machine translation quality using longest common subsequence and skip-bigram statistics,” in *Proceedings of the 42nd Annual Conference of the Association for Computational Linguistics*, pp. 605–612, Barcelona, Spain, July.
 - [75] J. Brooke, “SUS: a retrospective,” *Journal of Usability Studies*, vol. 8, no. 2, pp. 29–40, 2013.
 - [76] M. L. Tielman, M. A. Neerincx, R. Bidarra et al., “A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories,” *Journal of Medical System*, vol. 41, no. 8, p. 125, 2017.
 - [77] S. Valtolina, B. R. Barricelli, and S. Di Gaetano, “Communicability of traditional interfaces VS chatbots in healthcare and smart home domains,” *Behaviour & Information Technology*, vol. 39, no. 1, pp. 108–132, 2020.
 - [78] K. Coperich, E. Cudney, and H. Nembhard, “Continuous improvement study of chatbot technologies using a human factors methodology,” 2017.
 - [79] D. Kadariya, R. Venkataramanan, H. Y. Yip, M. Kalra, K. Thirunarayanan, and A. Sheth, “kBot: knowledge-enabled personalized chatbot for asthma self-management,” *2019 IEEE International Conference on Smart Computing (SMARTCOMP)*, vol. 8, 2019.
 - [80] J. Brooke, “SUS: a “quick and dirty” usability scale,” in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds., pp. 189–194, Taylor & Francis, London, UK, 1996.
 - [81] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, “Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial,” *JMIR Mental Health*, vol. 4, no. 2, 2017.
 - [82] L. Zhou, J. Gao, D. Li, and H.-Y. Shum, “The design and implementation of XiaoIce, an empathetic social chatbot,” *Computational Linguistics*, vol. 46, no. 1, pp. 53–93, 2020 Mar.
 - [83] A. N. Vaidyam, H. Wisniewski, J. D. Halamka, M. S. Kashavan, and J. B. Torous, “Chatbots and conversational agents in mental health: a review of the psychiatric landscape,” *The Canadian Journal of Psychiatry*, vol. 64, no. 7, 2019.