A Neural Network-based Diabetes Self-Management Chatbot System

Chieh-Yuan Tsai*, Hen-Yi Jen, and Tai-Jung Chiang

Department of Industrial Engineering and Management, Yuan Ze University, Taiwan
Email: cytsai@saturn.yzu.edu.tw (C.-Y. T.); henyi@saturn.yzu.edu.tw (H.-Y. J.); s1075415@mail.yzu.edu.tw (T.-J. C.)
*Corresponding Author
Manuscript received March 11, 2024; revised June 3, 2024; accepted June 26, 2024; published August 28, 2024

***Abstract*—The advances in mobile technology and natural language processing have made chatbots suitable for personal health care management. When the world's population is aging, diabetes becomes one of the most common chronic diseases in the world. After being discharged from a hospital, diabetes patients must conduct personal health care management such as monitoring blood glucose, professional diet advice, and regular exercise reminders to control the disease. Unfortunately, it has been found that there are few chatbots designated for diabetes patients, especially in the Chinese language. To fulfill the need, this research proposes a diabetes self-management chatbot to assist patients in recording their blood glucose, exercise, and diet through conversation. The proposed chatbot system consists of four main components: a dialog controller, a neural network, a personal database, and a diabetic management rule base. Iterated Dilated Convolutional Neural Network with Conditional Random Field (ID-CNN-CRF) is applied for Named Entity Recognition (NER) to achieve high chatting quality. The experiments show that the ID-CNN-CRF outperforms the other three popular CNN, LSTM, and Bi-LSTM-CRF models regarding intention prediction accuracy and response time. In addition, the chatbot can advise people with diabetes on proper diet and exercise. The feedback from the diabetes caregivers and patients shows that the proposed chatbot is recognized as a convenient self-management tool.**

***Keywords*—diabetes, self-management, chatbot, neural networks**

I. Introduction

As the world's population ages, chronic diseases become prevalent worldwide. About 14 million people aged 30 to 70 die yearly from chronic diseases [1]. Typically, chronic diseases include cardiovascular diseases (such as heart disease and stroke), chronic respiratory diseases (such as chronic lung obstruction and asthma), and diabetes. Among them, diabetes is one of the leading causes of death globally [2]. To battle the disease, diabetes patients must receive strong and consistent healthcare management after being discharged from the hospital. Personal healthcare management includes constant blood glucose level monitoring, diet control, and regular exercise. Without careful healthcare management, the danger caused by blood glucose fluctuation might permanently hurt diabetes patients [3].

Recently, conversational agents such as chatbots have shown potential in telemedicine and healthcare since conversational agents provide more user-friendly operations than conversational interfaces [4]. Chatbots can communicate with human users using spoken, written, and visual languages. For example, a Smart Wireless Interactive Healthcare System (SWITCHes) allows users to quickly get their diet and exercise records in a communication application and solve the health problems caused by obesity [5]. Chung and Park [6] proposed a chatbot-based healthcare service with a knowledge base for cloud computing. The proposed method is a mobile health service in the form of a chatbot to provide immediate treatment in response to accidents that may occur in everyday life and response to changes in the conditions of patients with chronic diseases. Dong, Chowdhury *et al*. [7] found that chatbots can provide users with information more quickly, especially when organizing information from different sources. *Laranjo* et al. [8] showed that patients receiving chatbot care are expected to have improved adherence to medication regimen, health-promoting behaviors, and control of symptoms as their self-efficacy increases (e.g., blood pressure control, blood glucose control, and cancer fatigue).

Chatbots can play a leading role by embodying the function of virtual assistance and bridging the gap between patients and clinicians. Unfortunately, few such healthcare management chatbots are designed for diabetes patients. Typically, the time between two clinical visits for diabetes patients could last several months. Patients may forget to record their blood glucose levels or living conditions during this period. However, failing to monitor the living conditions or blood glucose level may result in abnormal blood glucose fluctuations and increase the incidence of cardiovascular disease [3]. In addition, regular and moderate exercise can help people with diabetes stabilize their blood glucose levels. Therefore, providing a friendly tool for recording daily blood glucose levels, diet, and exercise will be very important to control the disease continuously.

An essential issue in building a high-quality chatbot is precisely recognizing the keywords in conversation sentences. One of the most common methods to achieve this goal is Named Entity Recognition (NER). Recently, many neural network models have been proposed to solve NER problems. X. Dong, Chowdhury *et al*. [9] constructed bidirectional RNNs to extract medical knowledge from medical conversations and disease records. In their work, the first step is to train a shallow bi-directional RNN in the general domain. The second step is to transfer knowledge from the general domain to train a deeper bi-directional RNN for recognizing medical concepts from Chinese electronic medical records. Athavale, Bharadwaj *et al*. [10] described an end-to-end Named Entity Recognition neural model for Hindi based on bidirectional RNN-LSTM. While these neural network models are expressive and accurate, they fail to exploit the parallelism opportunities of GPU [11]. This leads to longer response time and an unsatisfied user experience.

doi: 10.18178/ijml.2024.14.3.1163

The contribution of this study is summarized as follows:

- This research proposes a chatbot system specifically designed for diabetes patients, particularly for the Chinese language, which addresses a gap in the current availability of such tools.
- The use of Iterated Dilated Convolutional Neural Network with Conditional Random Field (ID-CNN-CRF) for NER enhances the chatbot's ability to understand and respond to user inputs accurately and efficiently.
- The proposed chatbot can provide personalized diet and exercise advice to diabetes patients based on their specific needs and health data, which is a significant feature for effective self-management.
- The feedback from diabetes caregivers and patients indicates that the proposed chatbot is recognized as a convenient self-management tool, which suggests that it has the potential to improve the overall user experience and health outcomes for diabetes patients.

In this study, a neural network-based diabetes self-management chatbot system is proposed to solve the above difficulties. The proposed chatbot can record patients' bio-medical data, exercise, and diet through a human-like conversation. We apply Iterated Dilated Convolutional Neural Networks (ID-CNNs) with Conditional Random Field (CRF) in the proposed system to make a shorter response time. ID-CNNs use filters simultaneously across the entire sequence, reducing the operation time [12]. In addition, with a dilated filter, ID-CNNs can be more effective in integrating context and providing higher accuracy. The remainder of the paper is structured as follows. Section II discusses previous work related to this research. In Section III, the framework of the proposed system is detailed. Section IV describes the system implementation and discusses the effectiveness of the system adoption. Finally, the limitations and suggestions for further directions are summarized in Section V.

II. Literature Reviews

A chatbot is a conversational agent that interacts with users using natural language sentences in a specific domain or topic. Recently, chatbots have been used in many fields, such as services, commerce, and medicine. Al-Zubaide and Issa [13] proposed a new ontology-based approach to operate chatbots (OntBot). OntBot uses an appropriate mapping technique to transform ontologies and knowledge into a relational database and then use that knowledge to drive its chats. Ludwig [14] developed a new adversarial learning method for generative conversational agents (GCA). A GCA generates a short sentence and lets another model predict whether the sentence is machine-generated or spoken by humans. Hussain and Athula [15] proposed a chatbot that extends beyond the local knowledge base. This chatbot connects to Wikipedia and uses the Media Wiki API to gain more knowledge, regardless of the local knowledge base. Angara [7] proposed a chatbot based on IBM's Watson Assistant that can perform complex natural language processing, including intent analysis, entity recognition, and more. The system can provide user recipes and adjust them according to user preferences.

Many chatbots have been proposed for healthcare and health management. Huang, Yang *et al.* [5] proposed a Smart Wireless Interactive Healthcare System (SWITCHes) to address the health problems caused by obesity. The mobile application is used to track diet, activity, and weight. Users can manage their calorie intake and consumption to achieve weight management. Divya, Indumathi *et al.* [16] constructed a text-to-text diagnosis bot to engage patients in conversation about their medical issues and provide a personalized diagnosis based on their symptoms. Fadhil [17] presented a chatbot that dynamically interacts with the aging population to gather information, monitor health, and provide support. This system is based on Dialogflow, which allows the elderly, patients, or residents of rural areas to set up telemedicine. Bibault *et al.* [18] proposed a chatbot to provide information for patients with breast cancer. The authors demonstrated that the effects of the information offered by the chatbot were similar to those of the information provided by doctors, and the information provided by the chatbot could reduce the number of patients with mild health concerns seeking medical attention at hospitals, thereby conserving medical resources. Medical consultation chatbots also enable physicians to spend more time treating patients with the highest needs.

Ye *et al.* [19] developed a chatbot program for follow-up management of the workers' general health examination (WGHE). Furthermore, this study investigated the effectiveness of the chatbot program to healthcare providers who have undergone the WGHE program and the applicability of the chatbot program to nurses in the particular agency of occupational health management. Castilla, Escoba *et al.* [20] presented HIGEA, a digital system based on a conversational agent to help to detect caregiver burden. The conversational agent naturally embeds psychological test questions into informal conversations to increase adherence to use and avoid user bias. Preliminary results show the system is valuable and practical. Hong, Piao *et al.* [21] aimed to assess the efficacy of a child vaccination chatbot based on changes in variables such as vaccination information, motivation, self-efficacy, and vaccination behavioral intention. It has been found that chatbots are valuable tools to encourage immunization through the provision of reminders and real-time consultation messenger services during the global health crisis and beyond.

In the field of diabetes, many diabetes-related mobile apps have been proposed, mainly focusing on diabetes self-management, lifestyle modification, and medication adherence motivation [22]. Intelligent Diabetes Management (IDM) is an app developed by the University of Alberta for patients with type 1 diabetes. This application includes a glucose and meal tracker, providing a detailed record in a diary format [23]. Diabetes Manager is an all-in-one app with an insulin calculator, a carbohydrate database, a favorites database, and a diary. It was developed to assist patients with type 1 diabetes mellitus in calculating pre-meal insulin doses [24]. Diabetes Pal is an app designed for patients with type 2 diabetes. It enables the user to monitor their blood glucose measurements directly from their glucose meter via Bluetooth or manually by inserting the glucose measurements into the app. The app also enables healthcare professionals to track the patients' glucose measurements

[25]. BlueStar was the first app in the USA to be given FDA approval as a mobile prescription therapy [26]. The BlueStar app lets patients enter their blood glucose levels and receive real-time coaching. They can also organize their medication plan and get advice on their lifestyle and diet. The mySugr app was designed to support patients in the diabetes self-management areas [27]. Data from self-monitoring of blood glucose (SMBG) and continuous glucose monitoring devices can be uploaded automatically and synced between devices via a cloud-based service. Insulin data can be entered manually into the app. Abbott Freestyle Libre (FSL) is a form of 'wearable' glucose sensor that monitors blood glucose levels without finger-prick testing; FSL has been shown to safely and significantly reduce HbA1c levels in people with type 1 diabetes. The FreeStyle Libre app can be connected to the FSL and monitor glucose through a smartphone [28]. Although the above apps have been successfully applied in monitoring glucose and providing appropriate advice for diabetes patients, they are not designed to interact with patients through natural language conversation.

Bali, Mohanty *et al.* [29] developed a DIAgnostic chatBOT, which provides personalized prediction using the general health dataset and based on the various symptoms sought from the patient. The authors explore ensemble learning, which combines multiple machine learning models to improve the overall prediction accuracy. Naim, Singh *et al.* [30] proposed a combination of K-Neighbors Classifier, Voting Classifier, and Light GBM Classifier for the diabetes prediction model. Their developed model shows the prediction of diabetic patients and suggests them appropriate action accordingly. Thongyoo, Anantapanya *et al.* [31] developed a chatbot to recommend a diet suitable for individuals with diabetes and build a cooperative health society. The chatbot's content, design, and implementation are evaluated to determine the user's degree of satisfaction.

Khan, Agarwalm *et al.* [32] mentioned that ChatGPT can potentially increase diabetes patients' access to information and assistance, but more research is required to comprehend its benefits and drawbacks fully. In [33], authors focused on analyzing methods for vectorizing textual information, selecting a model for scientific paper classification, and training the linguistic model BERT on the domain of scientific texts. They present the results of experiments to train scientific article classification models at the first and second levels of the Russian State Rubricator of Scientific and Technical Information (SRSTI). Tang, Yang *et al.* [34] studied the relationship between the symptoms of diabetes and the treatment of drugs in the clinical text of Chinese medicine. They used the Jieba word segmentation algorithm and the BERT model to study the relationship between symptoms and corresponding drugs in traditional Chinese medicine clinical texts. Ahne, Khetan *et al.* [35] applied deep learning and natural language processing methods to focus on tweets with personal and emotional content. A CRF model with BERT-based features outperformed a fine-tuned BERT model for cause-effect detection. Although BERT and GPT models have recently received much attention, both are primarily designed for tasks such as natural language processing, language translation, and text generation, which differ from the paper's specific requirements.

III. Methodology

In this study, a diabetes self-management chatbot is developed to record diabetes patients' daily blood glucose through a human-like dialog. The chatbot will give users proper diet and exercise recommendations based on their updated biomedical data. In addition, the chatbot will show a warning message if the users' biomedical data is abnormal. The development framework of the proposed chatbot is visually illustrated in Fig. 1. The major steps are:

1) Identify the objectives of the diabetes self-management chatbot by analyzing patients' needs and clarifying the intent for achieving the goals.
2) Collect utterances that users might talk with the chatbot and identify representative sentences.
3) Identify the entities of different intents and tag sentences with specific words.
4) Construct the diabetes self-management chatbot, DiabeteCare, where Iterated Dilated Convolutional Neural Networks (ID-CNNs) with Conditional Random Field (CRF) are adopted to identify entities in sentences.
5) Test the usability of the proposed chatbot by the subjects of diabetes patients and diabetes caregivers.

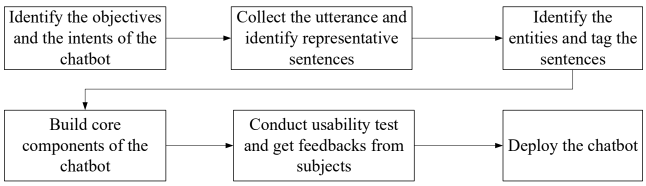


Fig. 1. The development framework for the diabetes self-management chatbot system.

A. Identification of Objectives and Intents

This study uses the chatbot development process suggested by [7] and [17] as a guideline. First, to provide better health management and telecare service for diabetes patients, the goals of the proposed chatbot are defined as follows:

- Recording the biomedical data of diabetes patients.
- Providing diet and exercise advice based on users' biomedical data.
- Providing warnings according to the blood glucose level.

For each of the high-level goals, the intents, which are fine-grained descriptions of the goal, are subsequently determined. Note that all intents defined for the goals should be mutually exclusive.

B. Utterance Collection and Representative Sentence Identification

The utterances are sentences that users use to convey a particular intent. However, the same intent might be expressed in very different ways. For example, "I want to record my blood glucose" and "My morning blood glucose is 99" express the same intent "Record blood glucose." Therefore, we collect variant utterances for each intent through different conversation scenarios. Table I shows representative sentences from target users in four different intent types. Note that the Chinese sentence "我要記錄我的血糖" means "I want to record my blood glucose" in English.

Table 1. The representative utterances for each intent

Intent	Representative sentences
Record blood glucose	“我要記錄我的血糖” (I want to record my blood glucose.) “我早上的血糖是 99” (My morning blood glucose is 99.)
Check the record of blood glucose	“我要看我的血糖紀錄” (I want to review my blood glucose record.) “我昨天的血糖是多少?” (What was my blood glucose yesterday?)
Ask for exercise advice	“我需要運動建議” (I need exercise advice) “我今天可以做什麼運動? “ (What exercise can I do today?)

C. Entities Identification and Sentence Tagging

For each utterance, we need to identify its keywords manually. The core idea of keyword recognition is that if the chatbot’s response needs to be different due to the particular words in the sentence, those words should be identified as entities. Table II shows the corresponding entities for each intent.

Table 2. The corresponding entities of each intent

Intents	Entities
Record blood glucose	Records (REC), Blood glucose (BSG), Value of blood glucose (VBS)
Check the record of blood glucose	Blood glucose record (BSR)
Record diet	Records (REC), Food (FOD), Portion size of food (PSF)
Check the record of the diet	Diet record (DIR)
Record exercise	Records (REC), Exercise (EXE)
Check the record of exercise	Exercise record (EXR)
Ask for dietary advice	Food (FOD), Advice (ADV)
Ask for exercise advice	Exercise (EXE), Advice (ADV)

Next, sentence tagging will be conducted for all collected sentences. Each word in a sentence is labeled with an appropriate tag. This study applies the character-based “IOB” tagging approach widely used in Chinese word segmentation for tagging sentences [36]. The symbol O indicates that the word is not the desired entity. The symbol B suggests that the word is the first word of the entity. The symbol I indicates that the word is the word after the first word of the entity. Two tagging examples are shown as follows:

[我 要 記 錄 血 糖]
[O O B-REC I-REC B-BSG I-BSG]
(I want to record blood glucose.)

[下 午 打 籃 球 和 跑 步]
[O O O B-EXE I-EXE O B-EXE I-EXE]
(Playing basketball and jogging this afternoon.)

D. Diabetes self-Management Chatbot

The major components of DiabeteCare include the dialogue controller, the neural network, the database, and the diabetic management rule base. The dialogue controller is designed to control the conversation flow, record the users’ data in the database, and provide a response to users. The neural network is designed to recognize keywords (entities) based on user sentences. The database records the user’s age, body weight, blood glucose level, diet, and exercise activity. Based on the user’s biomedical data in the database, the dialogue controller will warn when the user’s biomedical value is abnormal. To provide diet and exercise advice, the dialogue controller will regularly check pre-defined rules provided by medical institutions in the diabetic management rule. Fig. 2 shows the relationship among major components in DiabeteCare.

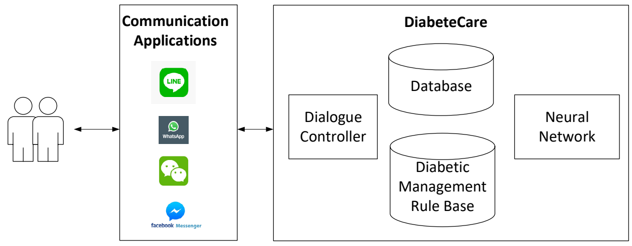


Fig. 2. Major components in the diabetes self-management chatbot system.

In the Neural Network component, Iterated Dilated Convolutional Neural Networks (ID-CNNs) with Conditional Random Field (CRF) are adopted to identify entities in sentences [12]. First, ID-CNN generates a logit for each word in the input sentence. Then, the CRF model decodes the logits into the labeled result with the Viterbi algorithm and predicts the token’s tag by its context.

In a dilated CNN [37], a CNN filter with a dilation width is applied to a continuous position of the input matrix. When moving, the filter of dilated CNN will skip the input data in the middle of all the dilation widths. While the size of the filter remains the same, the filter gets the data from the wider input matrix, which looks dilated. The ID-CNN architecture repeatedly applies the same block of dilated convolutions to token-wise representations. The output of each iteration is used as the input to the next iteration, and the same parameters are iteratively repeated. In this way, the parameters of CNN will be shared between iterations to prevent overfitting and provide opportunities to inject supervision on intermediate network activations. The architecture of the ID-CNNs with CRF is shown in Fig. 3.

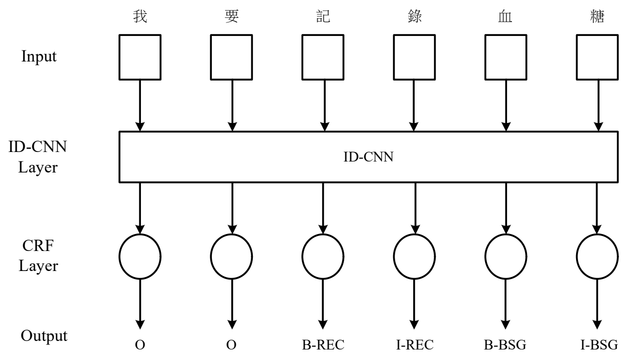


Fig. 3. The architecture of the ID-CNNs with CRF.

A CRF is a conditional probability distribution model for a sequence of output random variables given a sequence of input random variables. It is characterized by assuming that the output random variables constitute a Markov random field. The CRF prediction explicitly reasons about interactions among neighboring output tags. In this study, the CRF uses the logits output by ID-CNN as input and predicts the tag of each token according to the context. Given a conditional model $P(y \mid x)$, the most likely y is predicted as:

$$P(y|x) = \frac{1}{Z_x}\prod_{t=1}^{T}\psi_t\left(y_t\middle|h_t^{(L_b)}\right)\psi_p(y_t, y_{t-1}) \quad (1)$$

where $h_t^{(L_b)}$ is the logit of x_t, $\psi_t\left(y_t\middle|h_t^{(L_b)}\right) = \frac{Count(y_t \to h_t^{(L_b)})}{Count(y_t)}$ is the emission probability, $Count(y_t)$ is the number of times tag y_t occurs in the training sentences and $Count(y_t \to h_t^{(L_b)})$ is the number of times where $h_t^{(L_b)}$ maps to the tag y_t. ψ_p is a pairwise factor that scores consecutive tags, $\psi_p(y_t, y_{t-1}) = \begin{cases} 1, & y_t = y_{t-1} \\ \alpha, & \text{elsewhere} \end{cases}$, usually $\alpha < 1$ [38]. To avoid overfitting, ψ_p does not depend on t and input x in the model. Minimize the average loss for applying of the block B:

$$\frac{1}{L_b}\sum_{k=1}^{L_b}\frac{1}{T}\sum_{t=1}^{T} \log P\left(y_t|h_t^{(k)}\right) \quad (2)$$

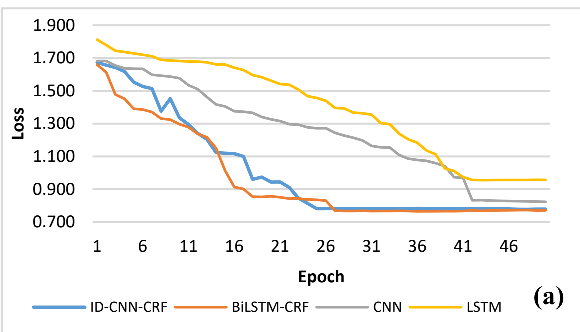
The model is trained by rewarding correct predictions after using the block. The output of a block in each iteration is used as the input of the block in the next iteration. The prediction from the previous block is used to improve the performance of the block after [14].

IV. Implementation and Validation

The proposed chatbot is implemented on the LINE communication app, one of Asia's most popular communication applications. After discussion with diabetes patients and diabetes caregivers, we identified eight intents for our proposed chatbot. In addition, 249 representative sentences are collected and used to train the Neural Network model in DiabeteCare. Each Chinese character in the sentence is considered a token and labeled with an appropriate tag. Next, a word embedding (word2vector) approach is applied to transform each token into a 300-dimensional vector. By default, the corpus used for training the word2vector model comes from diabetes medical documents.

A. Comparison among Different Neural Network Models

To validate the suitability of the ID-CNN-CRF model, three popular neural network models, CNN, LSTM, and BiLSTM-CRF [39], are compared through the 8-fold cross-validation. Fig. 4(a) shows the loss within 50 epochs for the four models. It is found that the ID-CNN-CRF is the fastest converged model. When the epoch is greater than 43, the loss for the four models is stable. In addition, Fig. 4(b) shows the training time of the four models at different epochs. It shows that the training time for the BiLSTM-CRF model is significantly longer than others, while ID-CNN-CRF and LSTM have similar training time.



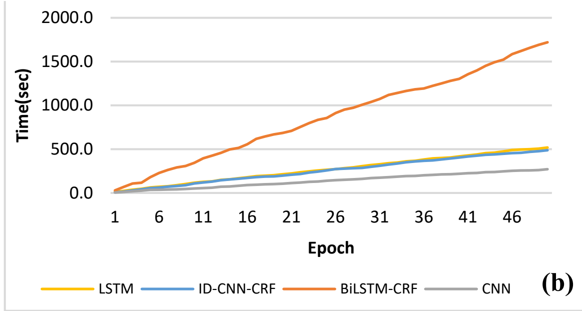


Fig. 4. (a) The loss of the four models, and (b) the training time of the four models.

After the four models have been built, their prediction accuracy is evaluated. In this study, the prediction is correct if the predicted tag outputted by a model is the same as the tag it should be. On the other hand, if the predicted tag is not the same as the tag should be, the prediction is incorrect. Fig. 5(a) shows the average prediction accuracy of the four models. The prediction accuracy for the ID-CNN-CRF and BiLSTM-CRF is over 85%. Although the prediction accuracy of BiLSTM-CRF is slightly higher than that of ID-CNN-CRF, the average response time of BiLSTM-CRF is much longer than that of ID-CNN-CRF, as shown in Fig. 5(b). Therefore, we conclude that ID-CNN-CRF is more suitable for the proposed chatbot.

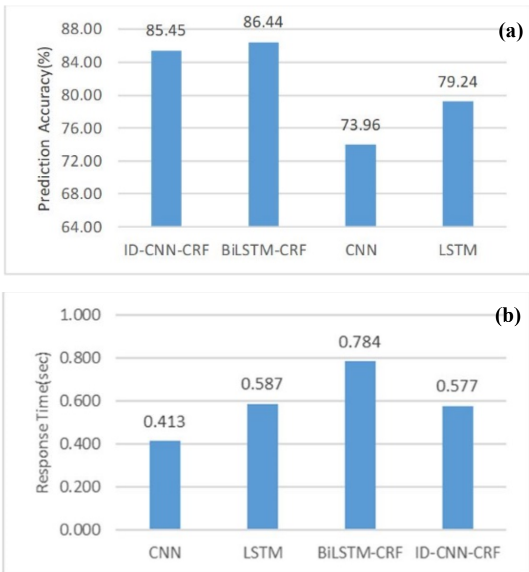


Fig. 5. (a) Average prediction accuracy and (b) average response time for the four models.

B. Comparison with Different Chinese Word Segmentation Methods and Word2vector Models

In this study, the Chinese language is used in DiabeteCare. Different Chinese word segmentation methods might affect the performance of the chatbot. In the previous experiment, each Chinese character in the sentence is considered a token. We call this "segmentation by character" (S.C.) method. Another segmentation method is to take multiple Chinese characters with meaningful concept as a token. We call this "segmentation by term" (S.T.) method. For example, the sentence "我要記錄血糖" is segmented as {"我", "要", "記", "錄", "血", "糖"} using the S.C. method, while segmented as {"我", "要", "記錄", "血糖"} by the S.T. method. Except for the segmentation method, different corpus sources might result in other word vectors. Therefore, two various corpus sources are tested. In this experiment, Q indicates the word vectors generated from the corpus of collected sentences, while M indicates the word vectors generated from the corpus of diabetes medical documents.

Table 3. The average response time and average prediction with different word segmentation methods and word vectors

(Segmentation, W2V) / Metrics	(S.T., M)	(S.C., M)	(S.T., Q)	(S.C., Q)
Average prediction accuracy (%)	81.95	81.59	82.47	82.85
Average response time (sec.)	0.663	0.671	0.648	0.727

Table III shows the ID-CNN-CRF model's average prediction accuracy and average response time for the four settings with 8-fold cross-validation. It is found that when the segmentation method is S.C., and the corpus source is Q, the average prediction accuracy is the highest. The "segmentation by character" method and corpus from "collected sentences" achieve the best prediction accuracy. However, the response time of configuration (S.C., Q) is slightly longer than others.

C. Case Demonstration

Fig. 6(a) shows the scenario where a user wants to record blood glucose step-by-step by inputting "我要記錄我今天的血糖" (I want to record my current blood glucose). In contrast, Fig. 6(b) shows another scenario where the user directly inputs his/her blood glucose as "我剛剛的血糖178" (my current blood glucose is 178). After the user inputs the blood glucose level, the chatbot will warn the user if the value is not within the normal range. Fig. 6(c) shows the warning message sent to the users.

Fig. 6. Dialogue example.

To help users keep track of their diet records, the chatbot returns the detail of blood glucose, diet, and exercise with graphical visualization for the past 30 days. Fig. 7(a) displays the blood glucose levels of the user; Fig. 7(b) shows the food and protein the user has consumed; Fig. 7(c) summarizes the exercise the user has taken.

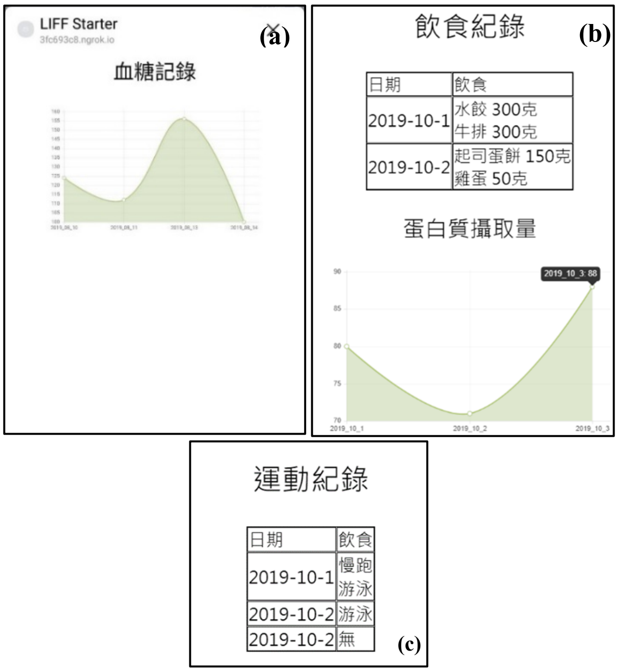


Fig. 7. Personal information retrieval for the last 30 days in the chatbot.

D. Usability Test

A preliminary usability test is performed to determine the proposed chatbot's feasibility. Based on the guidelines provided by [40–42], the following six questions were proposed:

Q1: Can you understand the functions of the chatbot?

Q2: Is the chatbot easy to use?

Q3: Compared with manual recording, is it more convenient to record blood glucose using the chatbot?

Q4: Compared with manual recording, is it more convenient to record diet using the chatbot?

Q5: Compared with manual recording, is it more convenient to record exercise activities using the chatbot?

Q6: Is the suggestion or warning from the chatbot helpful?

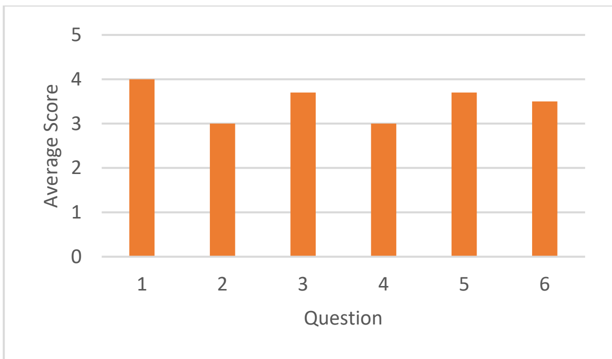


Fig. 8. The average score of each question.

Five diabetes patients and one diabetes caregiver are invited to operate the chatbot and then evaluate the six questions using a five-level Likert item evaluation approach where five indicates very satisfied, and one indicates very unsatisfied. Fig. 8 shows the average score of each question. The highest score of Q1 reveals that the chatbot functions can be understood clearly. In addition, most users found that making recordings with the chatbot is more convenient than doing it manually. However, some respondents reported that adding button functions in the chatbot would be helpful since they did not know how to start the chatbot initially. This might be true since most of our subjects are older persons who need more time to get familiar with the novel conversational chatbot. Instead of a button, adding a welcome message with basic instructions on how to use the chatbot would be helpful.

V. Conclusion and Future Work

This study proposes a self-management chatbot to help diabetes patients record their blood glucose level, exercise, and diet through natural language conversations. This chatbot consists of four components: the dialog controller, neural network, database, and diabetes management rule base. We apply Iterated Dilated Convolutional Neural Networks with Conditional Random Field (ID-CNN-CRF) to achieve high chatting quality for Named Entity Recognition (NER). The experiments show that the ID-CNN-CRF outperforms the other three popular CNN, LSTM, and Bi-LSTM-CRF models regarding prediction accuracy and response time. In addition, two Chinese word segmentation methods and two corpus sources for building word2vector models are tested. It was found that using the "segmentation by character" method and corpus from "collected sentences" achieves the best prediction accuracy. In addition, the proposed chatbot can record personal records and provide proper blood glucose warnings, exercise suggestions, and diet advice.

The proposed diabetes self-management chatbot, DiabeteCare, demonstrates significant potential in assisting users with diabetes management. However, there are some weaknesses and limitations that need to be addressed in future research. One major limitation is the reliance on a predefined set of intents and entities, which may not be comprehensive enough to cover all possible user inputs. This could lead to situations where the chatbot is unable to understand or respond appropriately to certain user queries. Additionally, the model's performance may degrade when dealing with out-of-vocabulary words or slang terms commonly used in informal conversations. Second, to improve the chatbot's capabilities, future research could focus on incorporating more advanced natural language processing (NLP) techniques, such as BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer) models. This would enable the chatbot to better understand user inputs and provide more accurate and personalized responses. Third, another area of improvement is the incorporation of more diverse and extensive training data. The current dataset is limited to a specific set of sentences and may not be representative of the broader range of conversations that users may have with the chatbot. Expanding the dataset to include more varied and realistic user inputs could significantly improve the chatbot's performance and accuracy.

References

[1] World Health Organization. (2014). World Health Organization: Global Status Report on Noncommunicable diseases, No. WHO/NMH/NVI/15.1., [Online]. Available: https://www.who.int/nmh/publications/ncd-status-report-2014/en/

[2] A. Al-Mawali, "Non-communicable diseases: Shining a light on cardiovascular disease, Oman's biggest killer," *Oman Med. J.*, vol. 30, no. 4, pp. 227–228, 2015.

[3] L. Monnier *et al.*, "Activation of oxidative stress by acute glucose fluctuations compared with sustained chronic hyperglycemia in patients with type 2 diabetes," *JAMA*, vol. 295, no. 14, pp. 1681–1687, 2006.

[4] S. Garg, N. L. Williams, A. Ip, and A. P. Dicker, "Clinical integration of digital solutions in health care: An overview of the current landscape of digital technologies in cancer care," *JCO Clin. Cancer Inform.*, vol. 2, no. 2, pp. 1–9, 2018.

[5] C.-Y. Huang, M.-C. Yang, C.-Y. Huang, Y.-J. Chen, M.-L. Wu, and K.-W. Chen, "A chatbot-supported smart wireless interactive healthcare system for weight control and health promotion," in *Proc. 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2018.

[6] K. Chung and R. C. Park, "Chatbot-based heathcare service with a knowledge base for cloud computing," *Cluster Comput.*, vol. 22, no. S1, pp. 1925–1937, 2019.

[7] P. P. Angara, "Towards a deeper understanding of current conversational frameworks through the design and development of a cognitive agent," Doctoral dissertation, 2018.

[8] L. Laranjo *et al.*, "Conversational agents in healthcare: A systematic review," *J. Am. Med. Inform. Assoc.*, vol. 25, no. 9, pp. 1248–1258, 2018.

[9] X. Dong, S. Chowdhury, L. Qian, Y. Guan, J. Yang, and Q. Yu, "Transfer bi-directional LSTM RNN for named entity recognition in Chinese electronic medical records," in *Proc. 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services* (*Healthcom*), 2017.

[10] V. Athavale, S. Bharadwaj, M. Pamecha, A. Prabhu, and M. Shrivastava, "Towards deep learning in Hindi N.E.R.: An approach to tackle the labelled data sparsity," in *Proc. the 13th International Conference on Natural Language Processing*, vol. 13, Varanasi, India, 2016.

[11] Y. Kim, "Convolutional Neural Networks for Sentence Classification," in *Proc. the 2014 Conference on Empirical Methods in Natural Language Processing* (*EMNLP*), 2014.

[12] E. Strubell, P. Verga, D. Belanger, and A. McCallum, "Fast and accurate entity recognition with iterated dilated convolutions," in *Proc. the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017.

[13] H. Al-Zubaide and A. A. Issa, "OntBot: Ontology based chatbot," in *Proc. International Symposium on Innovations in Information and Communications Technology*, 2011.

[14] O. Ludwig, "End-to-end adversarial learning for generative conversational agents," arXiv [cs.CL], 2017.

[15] S. Hussain and G. Athula, "Extending a conventional chatbot knowledge base to external knowledge source and introducing user based sessions for diabetes education," in *Proc. 2018 32nd International Conference on Advanced Information Networking and Applications Workshops* (*WAINA*), 2018.

[16] S. Divya, V. Indumathi, S. Ishwarya, M. Priyasankari, and S. K. Devi, "A self-diagnosis medical chatbot using artificial intelligence," *Journal of Web Development and Web Designing*, vol. 3, no. 1, pp. 1–7, 2018.

[17] A. Fadhil, "Beyond patient monitoring: Conversational agents role in telemedicine & healthcare support for home-living elderly individuals," arXiv [cs.CY], 2018.

[18] J.-E. Bibault *et al.*, "A chatbot versus physicians to provide information for patients with breast cancer: Blind, randomized controlled noninferiority trial," *J. Med. Internet Res.*, vol. 21, no. 11, p. e15787, 2019.

[19] B. J. Ye *et al.*, "Development of a chatbot program for follow-up management of workers' general health examinations in Korea: A pilot study," *Int. J. Environ. Res. Public Health*, vol. 18, no. 4, p. 2170, 2021.

[20] E. Castilla, J. J. Escobar, C. Villalonga, and O. Banos, "HIGEA: An intelligent conversational agent to detect caregiver burden," *Int. J. Environ. Res. Public Health*, vol. 19, no. 23, p. 16019, 2022.

[21] Y.-J. Hong, M. Piao, J. Kim, and J.-H. Lee, "Development and evaluation of a child vaccination chatbot real-time consultation messenger service during the COVID-19 pandemic," *Appl. Sci.*, vol. 11, no. 24, p. 12142, 2021.

[22] J. Doupis, G. Festas, C. Tsilivigos, V. Efthymiou, and A. Kokkinos, "Smartphone-based technology in diabetes management," *Diabetes Ther.*, vol. 11, no. 3, pp. 607–619, 2020.

[23] E. A. Ryan *et al.*, "Improved A1C levels in type 1 diabetes with smartphone app use," *Can. J. Diabetes*, vol. 41, no. 1, pp. 33–40, 2017.

[24] S. K. Garg, V. N. Shah, H. K. Akturk, C. Beatson, and J. K. Snell-Bergeon, "Role of mobile technology to improve diabetes care in

adults with type 1 diabetes: The remote-T1D study iBGStar® in type 1 diabetes management," *Diabetes Ther.*, vol. 8, no. 4, pp. 811–819, 2017.
[25] N. T. Ersotelos, A. N. Margioris, X. Zhang, and F. Dong, "Review of mobile applications for optimizing the follow-up care of patients with diabetes," *Hormones*, vol. 17, no. 4, pp. 541–550, 2018.
[26] P. Agarwal *et al*., "Mobile app for improved self-management of type 2 diabetes: Multicenter pragmatic randomized controlled trial," *JMIR MHealth UHealth*, vol. 7, no. 1, p. e10321, 2019.
[27] F. Debong, H. Mayer, and J. Kober, "Real-world assessments of mySugr mobile health app," *Diabetes Technol. Ther.*, vol. 21, no. S2, pp. S235–S240, 2019.
[28] C. Gordon, "Real world impact and experience of Abbott Freestyle Libre Glucose Monitoring," in *Proc. SAD - RAPID Conference: Research Accelerating Psychosocial Innovations in Diabetes*, 2022.
[29] 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.
[30] I. Naim, A. R. Singh, A. Sen, A. Sharma, and D. Mishra, "Healthcare CHATBOT for diabetic patients using classification," in *Proc. Soft Computing: Theories and Applications: SoCTA 2021*, pp. 427-437, Singapore: Springer Nature Singapore, 2022.
[31] P. Thongyoo, P. Anantapanya, P. Jamsri, and S. Chotipant, "A personalized food recommendation chatbot system for diabetes patients," in *Proc. Cooperative Design, Visualization, and Engineering: 17th International Conference, CDVE 2020*, Bangkok, Thailand, October 25–28, 2020, pp. 19-28, Springer International Publishing.
[32] I. Khan and R. Agarwalm, "Can ChatGPT help in the awareness of diabetes?" *Annals of Biomedical Engineering*, vol. 51, no. 10, pp. 2125-2129, 2023.
[33] I. K. Kusakin, O. V. Fedorets, and A. Y. Romanov, "Classification of Short Scientific Texts," in *Proc. Sci. Tech. Inf.* 50, pp. 176–183, 2023.
[34] Z. Tang, B. Yang, L. Li, and J. Hao, "BERT and word segmentation algorithm research on the relationship between diabetes symptoms and traditional chinese medicine," in *Proc. the 6th International Conference on Intelligent Autonomous Systems (ICoIAS)*, pp. 167-172, 2023.
[35] A. Ahne, V. Khetan, X. Tannier, M. I. H. Rizvi, T. Czernichow, F. Orchard, and G. Fagherazzi, "Identifying causal relations in tweets using deep learning: Use case on diabetes-related tweets from 2017-2021," 2021, arXiv preprint arXiv:2111.01225.
[36] H. Tseng, P. Chang, G. Andrew, D. Jurafsky, and C. Manning, "A conditional random field word segmenter for sighan bakeoff 2005," in *Proc. the Fourth SIGHAN Workshop on Chinese Language Processing*, 2005.
[37] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," arXiv [cs.CV], 2015.
[38] J. Lafferty, A. Mccallum, and F. C. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proc. the Eighteenth International Conference on Machine Learning*, 2001, pp. 282–289.
[39] Z. Huang, W. Xu, and K. Yu, "Bidirectional LSTM-CRF models for sequence tagging," arXiv [cs.CL], 2015.
[40] K. O'Shea, "An approach to conversational agent design using semantic sentence similarity," *Appl. Intell.*, vol. 37, no. 4, pp. 558–568, 2012.
[41] D. Elmasri and A. Maeder, "A conversational agent for an online mental health intervention," *Brain Informatics and Health, Cham: Springer International Publishing*, 2016, pp. 243–251.
[42] M. Jain, P. Kumar, R. Kota, and S. N. Patel, "Evaluating and informing the design of chatbots," in *Proc. the 2018 Designing Interactive Systems Conference*, 2018.

Copyright © 2024 by the authors. 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 (CC BY 4.0).