DALIA

CONVERSATIONAL AI CHATBOT ON DIABETES

THE DATA PLAYER'S TEAM

1. Felix Nyagah
2. Samson Kamunyu
3. Wendy Mwiti
4. John Mungai
5. Magembe Brian Nyambane
6. Sophia Mbataru
7. Calvince Ochieng

SUPERVISORS:

NIKITA NJOROGE

WILLIAM OKOMBA

LUCILLE KALEHA

Moringa School

Data Science Core

Cohort:DSF-FT02

1. Business Understanding

1.1 Overview

According to figures from International Diabetes Federation, an estimated 44.7% of adults living with diabetes (240 million people) across the world are undiagnosed – with the overwhelming majority having type 2 diabetes.

Diabetes Mellitus also called diabetes, refers to a chronic medical condition characterized by high levels of glucose in the blood. The body's ability to produce or respond to the hormone insulin which regulates blood sugar is impaired. Diabetes comes in different forms; Prediabetes, type 1 diabetes or insulin-dependent diabetes, Type 2 diabetes or non-insulin-dependent or adult-onset diabetes, and Gestational Diabetes. Research also shows that in 1% to 5% of people who have diabetes, other conditions might be the cause. These include diseases of the pancreas, certain surgeries and medications, and infections.

People having diabetes have a high risk of complications like heart and kidney disease, stroke, eye problems, nerve damage, etc. Current practice in hospitals is to collect the required information for diabetes diagnosis through various tests i.e glucose, BMI, age, and insulin. They then obtain the appropriate treatment based on the diagnosis.

Unfortunately, there is still inadequate awareness of this problem among the general public. There is also a lack of awareness about the existing interventions for preventing diabetes and for the management of complications. According to a study by Urban Health Centre, one of the barriers in seeking health care advice is the misconceptions about the disease, which revolve around all the aspects of diabetes; its prevention, control, and treatment. Studies have indicated that misconceptions and inadequate knowledge are significant barriers to the way of proper management of diabetes.

Knowledge is the greatest weapon in the fight against diabetes mellitus. Information can help people assess their risk of diabetes, motivate them to seek proper treatment, and care, and inspire them to take charge of their health. It is therefore in our interest as Data Players to design and develop a comprehensive health educative system for diabetes mellitus and its related risk factors.

Dalia, a conversational chatbot for diabetes, will provide information on managing diabetes, including suggestions for nutrition and exercise, and help those suffering from diabetes better understand their condition and take action to manage it more efficiently. For people with diabetes, this may result in better health outcomes and higher quality of life. This chatbot will attempt to clarify the various misconceptions that are prevalent among the population by providing improved, accessible, and reliable diabetes information.

1.2 Problem Statement

A major contribution of the prevalent diabetes cases has been a lack of awareness and support in managing and controlling the situation; resulting in immediate and long-term health, emotional and financial complications that could have been avoided were the condition detected early. Our goal, therefore, is to develop a conversational chatbot that can provide individuals with information and support on a wide range of diabetes-related topics. This will lead to better health outcomes and quality of life for those suffering from diabetes and improved accessibility to diabetes information to the public.

1.3 Objective

To develop a conversational AI chatbot that provides relevant, personalized information about questions regarding diabetes.

2. Data Understanding

The data was sourced from reliable websites through web scraping from different sources.

The websites the data was obtained from include:

https://agamatrix.com/blog/diabetes-nutrition-myths/

https://diabetesaction.org/questions-general-information

https://www.wisdomjobs.com/e-university/diabetes-interview-questions.html

https://gleneagles.com.my/articles/7-most-commonly-asked-diabetes-questions

https://www.msm.edu/blog/2019/six-questions-to-ask-your-doctor-about-diabetes.php

https://www.eatingwell.com/article/290863/top-common-questions-about-diabetes/

https://www.everydayhealth.com/hs/type-1-diabetes/type-1-diabetes-questions-answered/

http://ndl.ethernet.edu.et/bitstream/123456789/42448/1/Michael%20Bryer%20Ash.pdf

3. Data Preparation

3.1 steps for rasa data preparation

(i) Gather and organize the training data:

The data were collected and organized into intents, and entities, and used to train models.

(ii) Annotate the data:

The data was annotated to identify intents and entities. This was done using a tool such as Rasa NLU Trainer.

(iii) Pre-processing: tokenization and lemmatization

These are common steps in natural language processing (NLP) that are used to prepare text data for further analysis. Pre-processing refers to the cleaning and normalization of text data. This can include tasks such as removing special characters, converting text to lowercase, and removing stopwords.

Tokenization is the process of breaking down a piece of text into individual words or phrases. This is done so that the text can be more easily analyzed and processed by NLP algorithms.

Lemmatization is the process of reducing a word to its base form. This is done so that words that have the same meaning but different forms, such as "running" and "ran", are treated as the same word.

Featurization is the process of feature engineering in user messages. In RASA, we used pre-trained word embeddings and supervised embeddings.

Rasa is an open-source framework for building conversational AI that provides pre-processing, tokenization, and lemmatization functionalities. Rasa provided a python package rasa_nlu which contains the tools to do pre-processing, tokenization, and lemmatization.

To use these functionalities in Rasa:

Rasa_nlu package was installed and then the necessary classes and functions were imported. Then, text data was passed through the pre-processing, tokenization, and lemmatization steps, and then the resulting data was used for further training.

(iv) Pipeline configuration.

The NLU pipeline was configured in the Rasa config file to specify which components should be used to process the data.

(v) Training the NLU model:

annotated data was used to train the NLU model. This will be used to extract intents and entities from user input. This will be used to determine how the Chatbot should respond to different user inputs.

4. Chatbot Architecture

4.1 Steps involved in chatbot architecture

(i) User Input:

The user interacts with the chatbot through the user interface, which can be a website, mobile app, or messaging platform. The user's input is sent to the Rasa chatbot in the form of a text message.

(ii) Rasa NLU

Is responsible for understanding the intents and entities in a user's message where it takes in the user's message as inputs and outputs the intent and entities in the message.

Intents refer to the goal or purpose of a user's message and entities refer to specific useful pieces of information that can be extracted from users' input.

Rasa NLU uses machine learning to train models on annotated training data which is used to predict the intent and entities in new and unseen messages.

(iii) Rasa Core

Is responsible for managing the conversation flow and making decisions about what to do next in the conversation. It takes in the output of Rasa NLU as input and decides what action to take next.

Rasa Core uses a dialogue management model called a policy to make decisions about what to do next in the conversation. The policy is trained on examples of dialogs and uses machine learning to predict the next action.

(iv) Utter Response

Here the bot will respond to the user and if the user has any other questions or wants more clarification their new message will go through the same cycle.

Together, RASA NLU and RASA Core form a complete conversational AI system that can understand user inputs and make decisions about how to respond. The two components can be used separately or together, depending on the use case.

4.2 Evaluation Metrics

There are several ways we measured the accuracy of our bot including:

Intention accuracy: Dalia was able to understand and fulfill the user's intent.

Response accuracy: Our bot’s response aligned with the expected or correct answer.

Task completion rate: Dalia was able to carry out a large percentage of tasks successfully

User satisfaction: This can be measured by using surveys or questionnaires to gather feedback from users about their experience interacting with the bot.

Error analysis: From tasks, we have been giving Dalia we are in the process of identifying and analyzing errors made by the bot to understand where it struggles and improve upon it.

6. Conclusions and Recommendations

6.1 Conclusion of the project

A significant tool that can aid someone with diabetes in better managing their condition is the Rasa diabetes chatbot. The chatbot can understand and respond to user input in a way that mimics speaking to a real expert thanks to its natural language processing abilities. It can enhance diabetes self-management by giving individualized information because it is readily available, accessible, and simple to use.

6.2 Recommendations

(i) Multi-language support

Increased accessibility: Multi-language support allows chatbots to reach and serve a broader audience, particularly in multilingual countries or regions where people speak different languages.

Improved customer service: Multi-language support allows chatbots to communicate with customers in their preferred language, which can improve the customer service experience and build trust and loyalty.

Increased efficiency: Multi-language support allows chatbots to handle customer inquiries and requests in multiple languages, which can reduce the need for human intervention and increase efficiency.

Better data analysis: Multi-language support allows chatbots to collect data from a wider range of users, which can improve the accuracy and effectiveness of data analysis and decision-making.

(ii) Integration with other systems

The chatbot can be integrated with other apps and services, such as fitness trackers, to provide users with a more comprehensive view of their health and The chatbot can also be integrated with smart home devices, such as smart thermostats, to provide users with more control over their environment and reduce the risk of hypoglycemia.

The chatbot can be integrated with e-prescription systems, allowing users to easily refill their prescriptions and get reminders when it's time to take their medication.

(iii) Personalization

The chatbot can use machine learning to create personalized treatment plans for each user, based on their individual needs and goals.

(iv) Predictive capabilities

The chatbot can use advanced predictive modeling techniques to predict the risk of diabetes-related complications, such as heart disease or kidney failure, and provide users with personalized recommendations to prevent them.

(v) Integrated care

Diabetes chatbots can provide personalized care based on a patient's specific needs, medical history, and treatment plan. They can offer tailored advice and support, such as personalized nutrition and exercise plans, and help patients to better manage their diabetes.

(vi) Remote monitoring

Diabetes chatbots can provide remote care, allowing patients to access diabetes-related information and support from anywhere, at any time. This can be especially beneficial for patients who live in rural areas, or who have mobility or transportation issues.

(vii) Cost-effective

Diabetes chatbots can reduce costs associated with diabetes care, by automating routine tasks, reducing the need for face-to-face visits, and helping patients to better manage their diabetes. This can help to reduce overall healthcare costs, and can improve the affordability of diabetes care.

(viii) Improved patient engagement

Diabetes chatbots can improve communication between patients and healthcare providers, allowing for more frequent and convenient check-ins, and reducing the need for face-to-face visits. This can improve patient's access to care, and can help healthcare providers to more effectively manage patients' diabetes.