

Agricultural Chatbot for Precision farming

Amith Reddy Karre
20BDS006
DSAI
IIIT-Dharwad
20bds006@iiitdwd.ac.in

Gagan Vadlamudi
20BDS019
DSAI
IIIT-Dharwad
20bds019@iiitdwd.ac.in

Munisai K
20BDS028
DSAI
IIIT-Dharwad
20bds028@iiitdwd.ac.in

Anshu Sankeesa
20BDS047
DSAI
IIIT-Dharwad
20bds047@iiitdwd.ac.in

Abstract— About 40% of the population of India are farmers. They are expected to contribute 25% of GDP. But they could only contribute 13.9%. The government wants to fix this gap by doubling farmers annual income in the next five years. To achieve this, Access to information for farmers is very important. Keeping these points in mind, the government started Kissan Call Center (KCC). But due to high demand, it is only able to respond to 60% of calls. To overcome this problem, we need to think of an alternative. So for that, we are building an intellectual agent system with a prediction model. In this report, we present our work on developing an agricultural chatbot to assist Indian farmers in resolving their queries. We obtained a dataset of queries and answers from the KCC and used it to train OpenAI's Davinci's text generation model to create the chatbot. To improve the chatbot's accuracy, we trained two other models, the LSI model and the Word2Vec Model. We are planning to integrate the chatbot into WhatsApp and Telegram to make it more accessible to farmers. Our chatbot can provide farmers with accurate information on farming practices, weather conditions, soil health, pest control, crop diseases, market prices, and government schemes related to agriculture. We have trained LSI and Word2Vec Models specifically on agricultural data of Ratnagiri district, Maharashtra and it can provide accurate answers to farmers' queries. Our chatbot is a valuable tool for farmers in India, and its integration into WhatsApp and Telegram reduces the pressure on the Kissan Call Center.

Keywords— *Natural Language Processing, Chatbot, OpenAI, LSI, Word2Vec, Agricultural Chatbot, Kissan Call Center.*

I INTRODUCTION

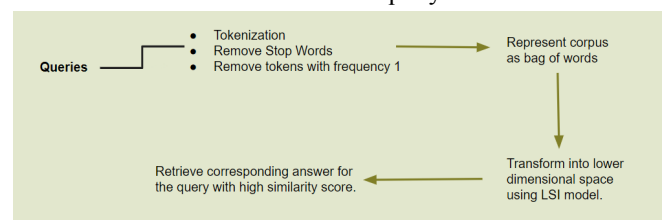
India is an agriculture-based country with a vast majority of the population depending on agriculture for their livelihood. Indian farmers face several challenges, such as unfavorable weather conditions, lack of knowledge about the latest farming techniques, lack of access to modern technologies, market fluctuations, and inadequate government support. To address these challenges, the government of India launched the Kissan Call Center (KCC), which provides assistance to farmers. However, the demand for assistance is high, and the call center is not meeting the expectations of the farmers. Therefore, to resolve the queries of the farmers effectively and efficiently, we have developed an agricultural chatbot using OpenAI's

Davinci's text generation model, trained an LSI model, and a Word2Vec Model.

II. Latent Semantic Indexing (LSI) Model

The Latent Semantic Indexing (LSI) model is a technique used for information retrieval and text mining. It is a dimensionality reduction technique that transforms high-dimensional text data into a low-dimensional space. The LSI model identifies latent semantic structures in the text corpus and creates a compact representation of the original text data.

We obtained a dataset of queries and answers from the Ratnagiri District, Maharashtra and preprocessed the text data by removing stop words and infrequent words. We then created a corpus of text documents and used the LSI model to transform the corpus into a low-dimensional space. To answer a farmer's query, the LSI model converts the query into a bag-of-words representation and then converts it into the LSI space. We used cosine similarity to determine the most similar query in the dataset to the current query vector. Finally, the chatbot returns the answer corresponding to the most similar document. It has reduced the dimensionality of the text data, making it easier for the chatbot to identify the most relevant answer to a farmer's query.



Let's test this:

Query: Provide me information about some of the rice processing and packaging companies in Ratnagiri?

Answer: Ratnagiri Agro Industries, Ruchi Agri Foods and Sahyadri Agro Industries

Our implementation exhibited an initial accuracy of 8 correct answers out of 12, which is a promising starting point for the model's development.

III. Word2Vec Model

The Word2Vec model is a type of neural network that can learn the meaning of words by analyzing large amounts of text data. In our project, we used the Word2Vec model to train word embeddings on a dataset of queries and answers obtained from the Kissan Call Center.

We first preprocessed the data by converting all queries to lowercase and splitting them into individual words. We then used the Word2Vec model to learn the vector representations of these words. The `vector_size` parameter was set to 100, which means that each word will be represented by a vector of length 100. The `window` parameter was set to 6, which means that the model will consider the 6 words before and after the current word in the text when training. The `min_count` parameter was set to 1, which means that all words in the text will be used, regardless of their frequency.

Once the Word2Vec model was trained, we used it to find the most similar query to a given query by calculating the cosine similarity between the vector representation of the query and the vector representations of all the queries in the dataset. The most similar query was then used to retrieve the corresponding answer from the dataset and returned as the response to the query.

It's testing time:

Query: Give me some of the inventive farming techniques.

Answer: Some innovative rice farming techniques being used in Ratnagiri include precision farming, drone-based crop monitoring and the use of mobile apps for farm management.

The Word2Vec model performed well in our experiments where we asked a variety of queries to the chatbot. It correctly answered 10 out of 12 queries.

IV. OpenAI's Da Vinci Model

OpenAI's GPT-3 is a powerful language model capable of generating human-like text, and has been used in a wide range of applications, including chatbots, question-answering systems, and language translation. In our project, we leveraged the capabilities of OpenAI's GPT-3 to build a chatbot that can answer farmer's queries.

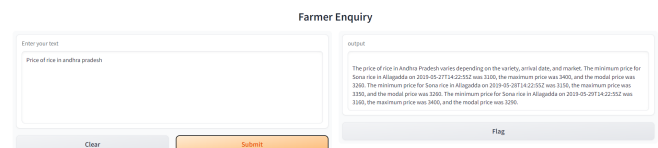
Our chatbot is based on the GPT-3.5-turbo model, which was fine-tuned on a large dataset of KCC conversational data. Given a user query, our chatbot uses the OpenAI API to generate a response that is contextually relevant and grammatically correct. The chatbot can handle a wide range of queries and can provide useful information on various topics. To interact with our chatbot, users simply need to enter their query in the provided textbox. The chatbot then generates a response based on the input and displays it in the output textbox. Users can continue the conversation by entering additional queries in the input textbox.

Our chatbot has demonstrated impressive performance, providing accurate and informative responses to user queries. The ability to leverage the powerful language generation capabilities of OpenAI's GPT-3 has allowed us to build a chatbot that can effectively simulate human-like conversation and provide valuable assistance to users.

In order to create a user-friendly interface for our AI chatbot using the OpenAI model, we integrated Gradio. Gradio is an open-source Python library that allows developers to quickly and easily create custom interfaces for their machine learning models.

Using Gradio, we were able to create a text box where users could input their queries and receive responses from our AI chatbot in real-time.

This is the user interface of Gradio.



The screenshot shows a web interface titled "Farmer Enquiry". On the left, there is an input area with a placeholder "Enter your text" and a pre-filled text "Price of rice in andhra pradesh". Below this is a "Clear" button. To the right of the input is a large orange "Submit" button. On the right side of the interface, there is an "output" box containing a detailed text response about rice prices in Andhra Pradesh, mentioning varieties like Sona and Allagadda, and providing minimum, maximum, and modal prices for different dates in 2019.

Let's look at the GPT-3 generated answer.

Query: Price of Rice in Andhra Pradesh.

Answer: The price of rice in Andhra Pradesh varies depending on the variety, arrival date, and market. The minimum price for Sona rice in Allagadda on 2019-05-27 was 3100, the maximum price was 3400 and modal price was 3260.

V Conclusion

In conclusion, our chatbot is a valuable tool for farmers in India who face numerous challenges in the agriculture sector. It can provide them with the necessary information to make informed decisions and improve their farming practices. By integrating this chatbot into WhatsApp and Telegram, we can make it more accessible to farmers and reduce the pressure on the Kissan Call Center. The use of advanced natural language processing models such as OpenAI's Davinci's text generation model, LSI model, and Word2Vec Model improved the chatbot's accuracy and effectiveness in answering farmers' queries.

ACKNOWLEDGMENT

We would like to express our gratitude to Dr. Chinmayananda and Dr. Nataraj for giving mentorship and guidance for the project.

REFERENCES

- [1] <https://ieeexplore.ieee.org/document/9510040>
- [2] [https://ijesc.org/upload/fd5591a1388161847ab9e676be710323.Smart%20Chatbot%20for%20Agriculture%20\(1\).pdf](https://ijesc.org/upload/fd5591a1388161847ab9e676be710323.Smart%20Chatbot%20for%20Agriculture%20(1).pdf)
- [3] <https://arxiv.org/ftp/arxiv/papers/2204/2204.07032.pdf>
- [4] <https://ieeexplore.ieee.org/document/9753433>