IEEE International Conference on Communication information and Computing Technology (ICCICT), June 25-27, 2021, Mumbai, India

Krushi – The Farmer Chatbot

Mihir Momaya[1]
Department of Electronics and Telecommunication Engineering
Mukesh Patel School Technology Management and Engineering, NMIMS
Mumbai, India
mihir.momaya55@nmims.edu.in

Anjnya Khanna[2]
Department of Electronics and Telecommunication Engineering
Mukesh Patel School Technology Management and Engineering, NMIMS
Mumbai, India
anjnya.khanna38@nmims.edu.in

Jessica Sadavarte[3]
Department of Electronics and Telecommunication Engineering
Mukesh Patel School Technology Management and Engineering, NMIMS
Mumbai, India
jessica.sadavarte69@nmims.edu.in

Manoj Sankhe[4]
Professor in Electronics and Telecommunication Engineering
Mukesh Patel School Technology Management and Engineering, NMIMS
Mumbai, India
manoj.sankhe@nmims.edu

2021 International Conference on Communication information and Computing Technology (ICCICT) | 978-1-6654-0430-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCICT50803.2021.9510040

Abstract— **As per the reports of November 2020, around 58% of India's population earns a primary source of living from agriculture. But, close to 10,000 farmers every year succumb to the harsh conditions in the agricultural sector. These conditions arise when the product of crops is not as desired, spoilt crops, which leads to large loans on the farmers.**

In this paper, we present an Artificial Intelligence (AI) chatbot that assists the farmers by providing solutions to agricultural queries. Some of the questions are concerning crop sowing, crop diseases, seasons related issues. Thus, benefit the farmers in making the right decisions regarding their crops, thereby increasing their yield.

Krushi - The Farmer Chatbot is an end-to-end trainable learning model to create a conversational system with minimum error and answer questions about current conditions. The chatbot is build using Artificial Intelligence (AI) and Machine Learning (ML) techniques. The dataset for the chatbot is used from Kisan Call Centre (KCC). The proposed system, answers queries related to weather, plant protection, animal husbandry, market price, fertilizer uses, government schemes, soil testing with an overall accuracy of 96.1% using RASA X.

Keywords— Agriculture, Artificial Intelligence (AI), End-To-End Trainable Learning Model, Kisan Call Center (KCC), Machine Learning (ML), Natural Language Processing (NLP), Neural Networks, RASA.

I. INTRODUCTION

Agriculture is one of the major sources of employment for a large number of people in the world. However, there are millions of small scale and marginal farmers who have a low level of awareness as they live in remote areas. Conventionally the field officers visit the fields and interact with farmers to provide supply training and advisory on practices that the farmers can use in farming and aspects of agriculture. The farmers that are probably main users of the traditional agricultural data about rainfall and crop production which is collected by the government in its raw form are unable to utilize it. To make it useful for the farmers this raw data has to be analyzed and fed to a system that would provide relational trends. The agricultural field is growing at a rapid pace not only in the technological aspects but also in the ways of production of crops. Several software are being developed to educate and instruct the farmers with this innovative technology. Basic information about farming is provided, they require a sizable amount of research to generate accurate information. The proposed framework defeats the downsides by giving a User Interface, where the farmers or different clients can communicate adequately to get desired responses with a lesser number of steps.

The database for the user interface has been created by the collection of data from Kisan Call Centre (KCC) followed by Data Preparation and Cleaning. Firstly, the data was Formatted into a single readable file by Eliminating Repeated Questions and Eliminating Questions with Improper Answers. Followed by combining all the regional datafiles{9 districts–108 files (9x12)} into a single CSV file which is further converted into RASA X readable YAML file. All the weather related queries were eliminated from the database to answer them in real time using OpenWeatherMap API. Further Data Analysis was performed on (a) Query Type Questions Analysis (b) Regional Analysis. After the above steps have been performed the Data obtained is then interfaced with RASA X Software. The chatbot is deployed on WhatsApp using third party applications like Twilio – SandBox and ngROK.

II. LITERATURE SURVEY

A few conversational systems, crop disease detection applications, and weather forecast-based advisory systems can be seen in references. In almost every paper referred, we could analyze that the approach adopted was finding word vectors using the given data, then performing optimization and eventually prediction by finding the most similar question from the database as performed in the base paper. Various techniques have been employed for query analysis which includes CNN (Convolutional Neural Networks) algorithms namely Sequence to Sequence RNN models, Random Forest, Decision trees but apart from them we found Google's Word2Vec model to be the most promising with accuracy reaching as high as 91% on scores based on Modified Lesk technique. In every paper that we went through a rarity was observed which led us to a research gap that suggested us to include a speech input query that would be implemented using Google APIs to make the system user-friendly. During the conversation, one important thing to be maintained is the intent of the conversation. The empty fields are populated by retrieving data from the domain-specific knowledge base database and eventually answering the question.

The first paper we considered was written by Mohit Jain.[1] The key takeaway from the paper was • Data collection was done from the online government site for KCC. • Compares two interfaces of FarmChat: 1. Audio-only - Input: speech; output: audio 2. Audio + Text - Input: speech, button; output: audio, text, image.

The second paper we considered was written by Naman Jain.[2] The key takeaway from the paper was 1. Data collection was done from the online government site for KCC. • Data analysis was performed, since the same question could have different answers in different states. • One of the important features of the paper was the use of the Sen2Vec model which was used for sentence embedding. • All the weather questions were eliminated and weather API was used to find real time weather data. • Jaccard and Lesk scores were used as metrics to find the results.

The third paper we considered was written by Divya Sawant.[3] The key takeaway from the paper was •The proposed system uses KNN (K-Nearest Neighbors) algorithm. • This system Agribot is a web application (app) • Interactive chatbot uses DialogFlow API.• The chatbot can also accept voice input.

The fourth paper we considered was written by Prashant Y.Niranjan.[4] The key takeaway from the paper • The proposed system uses RNN (Recurrent Neural Networks) deep learning algorithm. • Dataset is created in the XML format. • NLP (Natural Language Processing) is used for entity extraction and then input to the RNN sequence which determines the result by considering previous actions

The fifth paper we considered was written by Kanakamedala Deepika. [5] The key takeaway from the paper • The proposed system uses RASA X for the development of a chatbot that is deployed on the telegram app. • Dataset is created in the YAML format.

III. METHODOLOGY

A. Data Collection

The data for the database is collected from the official website of Kisan Call Centre. The data here is collected for the 9 severely affected districts in Maharashtra. The districts that we considered are Ratnagiri, Satara, Ahmednagar, Beed, Aurangabad, Washim, Amravati, Wardha, and Gondia. We have taken files for 12 months for each district which is 108 files. These files are then combined into one single CSV file to form a database for the chatbot.

B. Data Preparations

The data is prepared by removing repeated questions that were asked. The questions with improper answers are also eliminated from the file. The questions are now combined region wise for better analysis. All the questions in the correct format are kept and the rest are discarded. Thus, the data is furthermore compact, filtered, and easy to understand.

C. Data Analysis

The data is analyzed in two types (1) Query Type Analysis and (2) Region Analysis

Query analysis is important for query processing it improves the overall performance of the database. Thus, help speed up the process. Query analysis helps optimize the database to further optimize queries for performance.

We could also infer from Figure 1 that the queries asked regarding weather were nearly 46%. Since there was a large number of queries asked pertaining to weather, the query analysis for such queries is performed differently. We chose to eliminate questions regarding weather to give real time answers using OpenWeatherMap API.

Figure 1 gives Query type analysis of Maharashtra. The pie-chart shows the percentage of queries that were asked for sectors like weather, plant protection, etc., in Maharashtra.

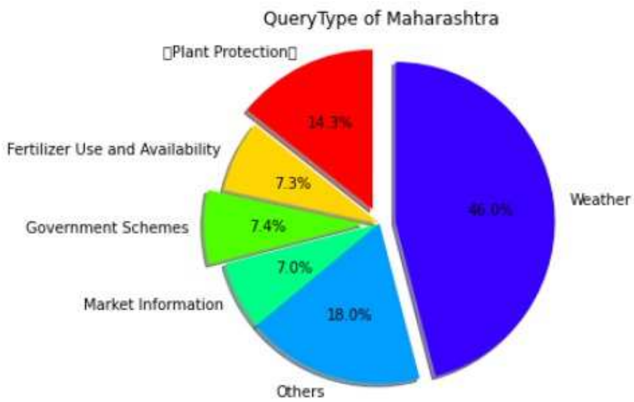

Figure 1: Query Type Analysis of Maharashtra

The region type data analysis gives a good picture of the agricultural landscape of Maharashtra regarding which crops are popular in which region, what kind of queries are most commonly asked, and the different sectors about which the queries are related. From figure 2 we can see that the maximum number of queries asked regarding the crops were related to cotton.

Figure 2 gives Crop wise Query Analysis. It shows the number of queries that were asked for various crops in Maharashtra in the form of a bar graph

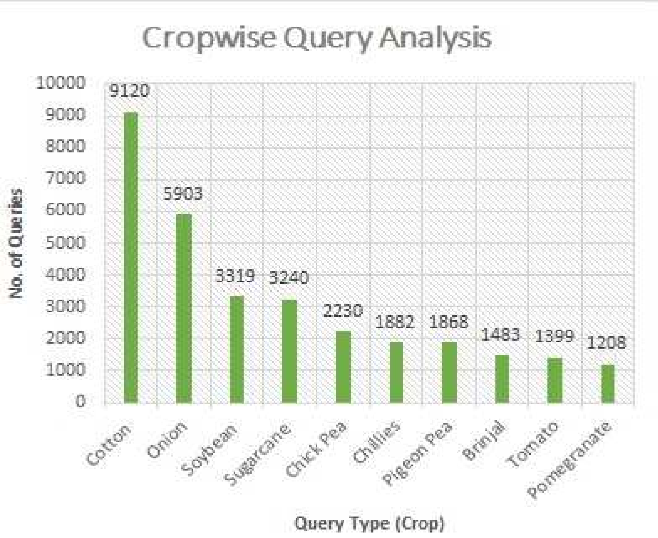

Figure 2: Crop wise Query Analysis

Table I shows the regional analysis for the major crops in the 9 districts of Maharashtra that we have taken into account.

Table I : REGIONAL ANALYSIS FOR MAJOR CROPS

District Name	Major Crops	Value (%)
Ahmednagar	Onion	12.14
Amravati	Cotton	11.04
Aurangabad	Cotton	9.9
Beed	Cotton	10.16
Gondia	Paddy	13.48
Ratnagiri	Mango	10.01
Satara	Sugarcane	13.17
Wardha	Cotton	14.64
Washim	Soyabean	8

D. Flow Chart

RASA X is an open-source machine learning framework for automated text and voice-based conversations, understand messages, hold conversations, and connect to messaging channels and APIs. Rasa X is a tool for Conversation-Driven Development (CDD), the process of listening to your users and using those insights to improve your AI assistant.

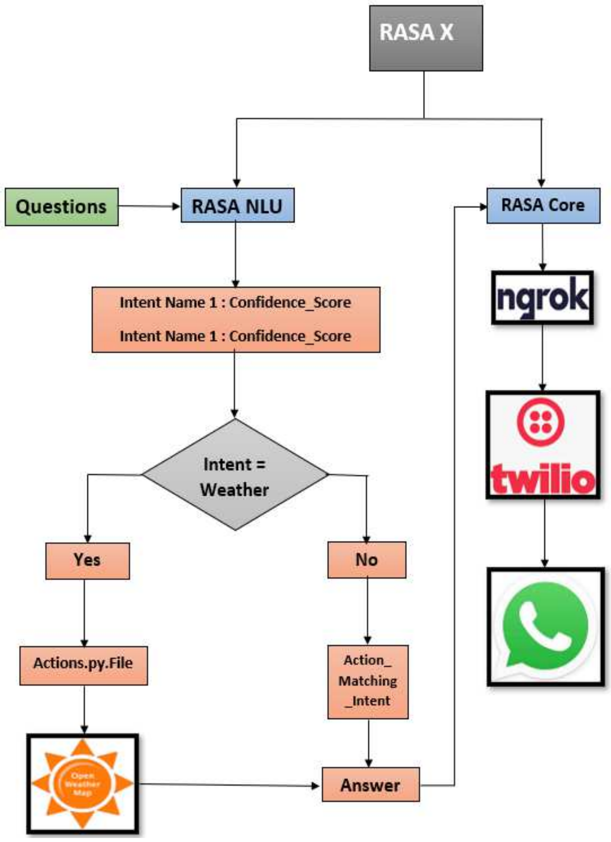

Figure 3 : Flow of the Chatbot

RASA X is further divided into two parts RASA NLU (Natural Language Understanding) and RASA Core. RASA NLU handles intent classification, entity extraction, and response retrieval. RASA Core decides/guesses which is the next probable state (again just an intent) of the chatbot conversation. It's off-line trained with a RASA specialty: stories. These are possible sequences of intents, following examples of conversations that developers submit in the training phase.

Figure 3 shows the flow of Krushi – The Farmer Chatbot. We propose a system where a question is given as an input to RASA NLU which then classifies it into intents, identifies entities and finds out the matching intent with the highest confidence score. Further, it is checked if the intent corresponds to the weather query or not.

If the intent corresponds to the weather query then it is passed to the Actions.py.file where it connects to the OpenWeatherMap Key for real time answers. Once the answer is generated it is sent to RASA Core which after analysis sends it to ngROK which acts as a sever helping the chatbot go online which can be accessed globally and the HTTPS link is then configured with Twilio. Twilio -SandBox acts as a console that helps configure messages on WhatsApp.

If the intent does not correspond to the weather query then its Matching Intent is found from the database and the corresponding action takes place as can be seen in Figure 4. The intents check the stories of the query (explained in Figure 5). The system then throws the corresponding action from the matching story to the RASA Core. The process from Rasa Core to ngROK to Twilio to WhatsApp remains the same irrespective of the intent.

```
- story: happy weather
  steps:
  - intent: greet
  - action: utter_greet
  - intent: weather
  - action: utter_city
  - intent: city
  - action: action_weather_api

- story: animal husbandry
  steps:
  - intent: animal_husbandry_contact_number
  - action: utter_contactnumber_animalhusbandry
```

Figure 4 : Specimen Stories File

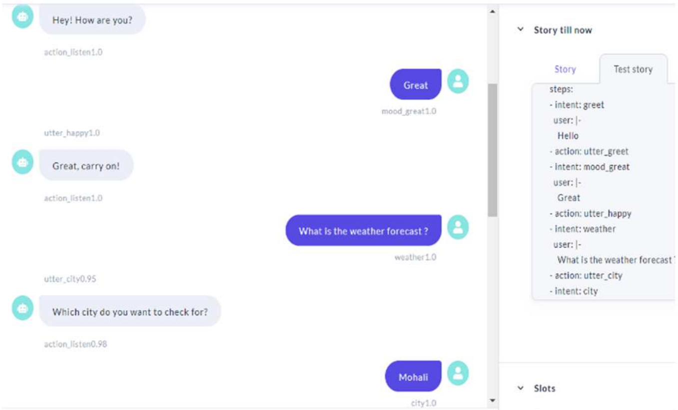

Figure 5: Intent Checking

IV. RESULTS AND DISCUSSION

The proposed chatbot will be able to solve farmer solutions through pre-existing answers given by KCC experts, thereby reducing the load on the helpline number. This system is connected to WhatsApp via Twilio, thereby making it convenient and accessible to the users 24X7. The model was evaluated using data from the KCC and we found the evaluated parameters as follows (mentioned in Table II) :

Table II : CLASSIFICATION REPORT

PARAMETERS	SCORE
Intent Accuracy	95.67
Story Accuracy	94.285
F1 Score	97.90
Precision	98.46
Accuracy	96.1

To evaluate the robustness of the system parameters like intent accuracy and story accuracy are used.

The number of intents correctly identified by the system is given in equation (1)

$$\text{Intent Accuracy} = \frac{\textit{Number Intents Correctly Identified}}{\textit{Total Number of Intents}} \quad (1)$$

The number of correct flows of interaction identified by the system is given in equation (2)

$$\text{Stories Accuracy} = \frac{\textit{Number Stories Correctly Identified}}{\textit{Total Number of Stories}} \quad (2)$$

Figure 6 shows our conversation with Krushi – The Farmer Chatbot with respect to the crops. The figure shows Crop Query Results that were answered by the chatbot.

Figure 7 shows our conversation with Krushi – The Farmer Chatbot with respect to weather analysis. The chatbot was able to answer questions with validity and abreast with the intent of the dialogue.

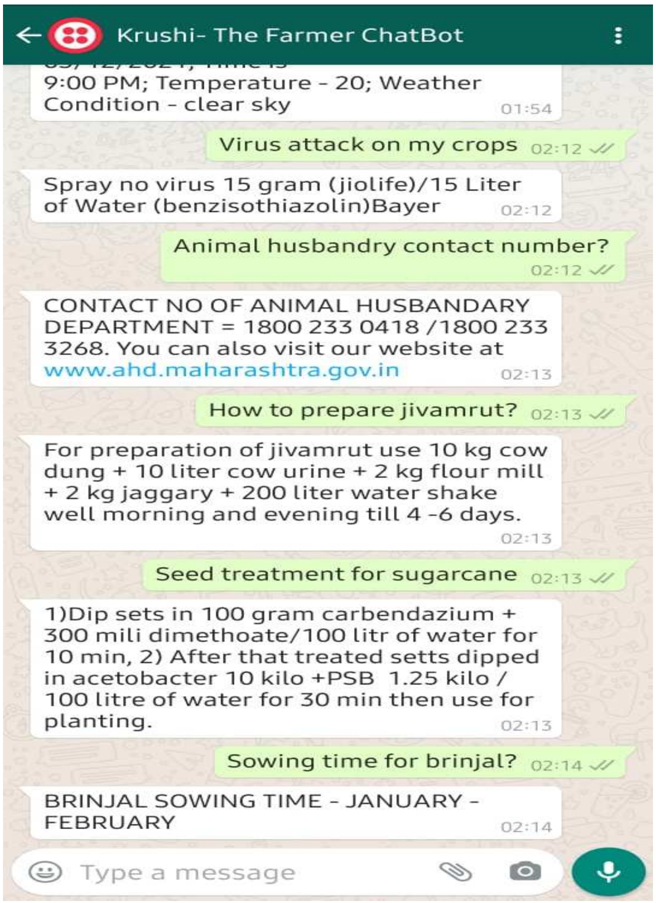

Figure 6: Crop Query Results

Figure 7 : Weather Query Results

V. CONCLUSION

Krushi- The Farmer Chatbot satisfies the information need of the farmers with high accuracy and also alleviates the burden on the KCC website. In June 2018 KCC received 1.11 million calls out of which 40% went unanswered. Krushi- The Framer Chatbot will help in better utilization of human resources and avoid unnecessary costs in setting up new call centers. The chatbot can positively impact underserved communities by solving queries related to agriculture, horticulture, animal husbandry, fisheries, government schemes, market prices, etc. using NLP (Natural Language Processing).

Krushi – The Farmer Chatbot is implemented in RASA X and deployed on WhatsApp. Experimental results show that Krushi will be able to answer 70% of the overall queries that are asked on the KCC helpline number, during any time of the day with an overall accuracy of 96.1% and an increase in the training data will only lead increase in accuracy.

The farmer will be able to receive agricultural information as well as weather forecast through WhatsApp. Our system would enable the farmer to have KCC verified expert answers anytime and from anywhere, which will in turn help in spreading the modern farming technology faster and to a large number of farmers. The scope of the Chatbot is not only limited to the farmers but can be used by anyone who wishes to utilize it for their knowledge and research purpose.

VI. FUTURE SCOPE

In the future, we hope to make the Chatbot compliant with local languages using voice input, including KCC data pan India and integrate it with WhatsApp, thereby reaching out to people from every corner of the country. Real time data would be made available to the users by using web scrapping technology and questions which could not be answered by the Chatbot would be stored, updated and the model would be made to retrain, hence broadening the spectrum of our Chatbot.

REFERENCES

[1] Mohit Jain, Pratyush Kumar, Ishita Bhansali, Q. Vera Liao, Khai Truong, and Shwetak Patel. 2018. FarmChat: A Conversational Agent to Answer Farmer Queries. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 2, 4, Article 170 (December 2018), 22 pages. DOI:https://doi.org/10.1145/3287048.

[2] Naman; Jain, Pranjali; Kayal, Pratik; Sahit, Jayakrishna; Pachpande, Soham; Choudhary, Jayesh and Singh, Mayank, “AgriBot: agriculture-specific question answer system”, in the International Conference on Science, Technology, Engineering and Mathematics, Education Department, Department of Science and Technology Government of Gujarat, IN, Jan. 17, 2019.

[3] D. Sawant, A. Jaiswal, J. Singh and P. Shah, "AgriBot - An intelligent interactive interface to assist farmers in agricultural activities," 2019 IEEE Bombay Section Signature Conference (IBSSC), Mumbai, India, 2019, pp. 1- 6, doi: 10.1109/IBSSC47189.2019.

[4] P. Y. Niranjan, V. S. Rajpurohit and R. Malgi, "A Survey on Chat-Bot system for Agriculture Domain," 2019 1st International Conference on Advances in Information Technology (ICAIT), Chikmagalur, India, 2019, pp. 99-103, doi: 10.1109/ICAIT47043.2019.8987429.

[5] K. Deepika, V. Tilekya, J. Mamatha and T. Subetha, "Jollity Chatbot- A contextual AI Assistant," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 1196-1200, doi: 10.1109/ICSSIT48917.2020.9214076.

[6] Prof. Yashaswini. D. K 1 , Hemalatha. R2 , Niveditha. G3,” Smart Chatbot for Agriculture”, in the International Journal of Engineering Science and Computing, May 2019

[7] J. Vijayalakshmi, K. PandiMeena , “Agriculture TalkBot Using AI”, in the International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277 – 3878, Volume-8, Issue-2S5, July 2019.

[8] B. Arora, D. S. Chaudhary, M. Satsangi, M. Yadav, L. Singh and P. S. Sudhish, "Agribot: A Natural Language Generative Neural Networks Engine for Agricultural Applications," 2020 International Conference on Contemporary Computing and Applications (IC3A), Lucknow, India, 2020, pp. 28-33, doi: 10.1109/IC3A48958.2020.233263.

[9] H. Du, P. Jones, E. L. Segarra, C. F. Bandera, “Development Of A REST API For Obtaining Site-Specific Historical And Near-Future

Weather Data In EPW Format," 4th Building Simulation and Optimization Conference, Cambridge, sept. 2018.

[10] Ekanayake, Jayalath & Saputhanthri, Luckshitha. (2020). E-AGRO: Intelligent ChatBot. IoT and Artificial Intelligence to Enhance Farming Industry. Agris on-line Papers in Economics and Informatics. 12. 15-21. 10.7160/aol.2020.120102.

[11] Mostaço, Gustavo & Campos, Leonardo & Souza, Ícaro & Cugnasca, Carlos. (2018). AgronomoBot: a smart answering Chatbot applied to agricultural sensor networks.

[12] B. Gamage, R. Pushpananda and R. Weerasinghe, "The impact of using pre-trained word embeddings in Sinhala chatbots," 2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer), Colombo, Sri Lanka, 2020, pp. 161- 165, doi: 10.1109/ICTer51097.2020.9325440.