BINDURA UNIVERSITY OF SCIENCE EDUCATION

FACULTY OF SCIENCE AND ENGINEERING

COMPUTER SCIENCE DEPARTMENT

AgriBot: An intelligent Assistant for Zimbabwe Virginia Tobacco farmers using natural language processing and deep learning Algorithms

By

Precious Mafundu

B192058B

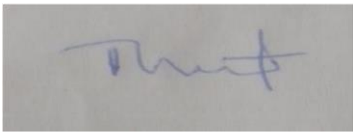
SUPERVISOR: Mr Muzurura

A RESEARCH PROJECT SUBMITTED TO THE COMPUTER SCIENCE DEPARTMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE BACHELOR OF SCIENCE HONOURS DEGREE IN COMPUTER SCIENCE.

APPROVAL FORM

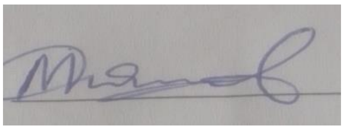
The undersigned certify that they have supervised the student Divine Precious Mafundu dissertation entitled, "AGRIBOT: AN INTELLIGENT ASSISTANT FOR ZIMBABWE VIRGINIA TOBACCO FARMERS USING NATURAL LANGUAGE PROCESSING AND DEEP LEARNING ALGORITHMS" submitted in partial fulfillment of the requirements for a Bachelor of Computer Science Honors Degree at Bindura University of Science Education.

STUDENT: **DATE:**

 02/10/24

………………………………………………… …………………………..

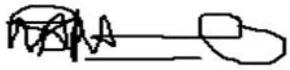
SUPERVISOR: **DATE:**



02/10/24

………………………………………………… …………………………..

CHAIRPERSON: **DATE:**



02/10/24

………………………………………………… …………………………..

EXTERNAL EXAMINER: **DATE:**

………………………………………………… …………………………..

Copyright

Dedication

This dissertation is dedicated to my husband and son Taona for their unconditional support.

DEDICATION

I dedicate this dissertation to my mother, Miriam Mutambaneshiri, my father George Mafundu and my husband Perseverence John who did well in their ability to get me in the path I have chosen and supported each one of my dreams in every way that was possible to them. I know they waited for so long time to see this achievement.

ACKNOWLEDGEMENTS

Above all, I want to thank the Lord Almighty for life. Special thanks go to the entire student's and lecturers at BUSE, especially Mr. O. Muzurura the project supervisor, for the author would not have gone that far had it not been for their tutoring and guidance.

I am particularly indebted to BUSE for taking me in, giving me more than I expected, their patience and cooperation in my learning process has been extraordinary. Of mention are Mr. O. Muzurura my supervisor, Mr. D. Ndumiyana (Industrial attachment supervisor), Mr P. Muyambo, Mr P. Chaka, Mr C. Zano, Dr Sakala and the rest of the Lecturers in the Faculty Science.

Finally, my big thank goes to my parents and family members for helping me and supporting me to succeed in this event. There are also other people that I was unable to mention by their names, I am grateful for all your support and help.

Abstract

This research focuses on the development and implementation of an intelligent interactive virtual assistant (Chabot) specifically designed for Virginia Tobacco farmers in Zimbabwe. The assistant leverages natural language processing and deep learning algorithms to enhance decision-making processes, optimize crop yield, and improve overall efficiency in the Virginia Tobacco farming systems for small to large scale growers in A1, A2 and commercial farming systems of Zimbabwe. This assistant aims to provide timely and accurate information to farmers, enabling them to make informed decisions regarding Virginia Tobacco production, disease management and pest control. The agricultural sector plays a crucial role in the economy of Zimbabwe, with tobacco farming being one of the significant contributors. However, tobacco farming faces numerous challenges, including limited access to expert advice, changing climatic conditions and minimal management. To address these challenges, the integration of intelligent interactive assistants powered by natural language processing and deep learning algorithms has emerged as a promising solution which can help in answering and advising farmers with expert knowledge regarding the best farming practices to attain higher

yields. Due to continuous miserable harvests obtained by new Virginia flue cured tobacco farmers and smallholder tobacco farmers which is being attributed to lack of information, ignorance and improper farming practices here in Zimbabwe, there is a great need for more convenient, fast, reliable and sustainable technological driven mitigating strategies for assisting these farmers with real-time consultation on how to conduct proper farming practices 24/7 as recommended. The development and implementation of a Natural Language Processing (NLP) and deep learning based Chatbot has been considered as such a solution to answer these farmers' questions about best growing practices, recommended seeds, best time to plant, different ways of applying correct fertilizers, different watering practices, best recommended ways to control pests and diseases, as well as to help farmers troubleshoot problems with their crops and to recommend relevant solutions in real time. The tobacco farmer needs to know about all this, and hence the value of a Chabot to act as an Agricultural Extension Worker (AEW) and available 24/7. However, the author will build the NLP based Chatbot with Rasa and Python that can reach out to thousands of tobacco farmers across Zimbabwe to disseminate valuable information on tobacco farming techniques and recommended agricultural practices among the tobacco farmers.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Introduction

The tobacco industry plays a crucial role in Zimbabwe's economy, with Virginia tobacco being one of the key crops cultivated by farmers. However, tobacco farming is a complex and labour-intensive process that requires continuous monitoring, decision-making, and adherence to best practices, in order to achieve the best yields. To support Zimbabwean Virginia tobacco farmers and enhance their productivity, this research aims to develop an intelligent interactive assistant using natural language processing and deep learning algorithms. The Virginia tobacco industry faces numerous challenges, including changing seasonal climates, increased competition, and environmental concerns. To remain competitive, farmers need access to accurate and timely information to make informed decisions about their crops so that best crop management practices and yields can be obtained. An intelligent interactive assistant powered by natural language processing and deep learning algorithms can provide valuable support to Virginia tobacco farmers by offering:

- Personalized recommendations: The assistant can analyze data from various sources, including weather forecasts, soil conditions, and market trends, to provide farmers with customized recommendations on planting, fertilization, irrigation, disease and pest management.

- Disease and pest detection: The assistant can use natural language processing algorithms to help identify potential diseases or pests at an early stage, allowing for timely intervention and minimizing crop damage.

- Access to information: The assistant can serve as a central repository of information on best practices, regulations, and research findings, providing farmers with easy access to the knowledge they need.

1.2 Background and Motivation

Tobacco is one of Zimbabwe's primary agricultural exports, with Virginia tobacco being a major cash crop for many smallholder farmers in the country. However, these farmers often face significant challenges in maximizing their yields and profitability due to a variety of factors, including unpredictable weather patterns, pests and diseases, and a lack of access to timely and relevant agricultural information and advisory services.

In this context, the development of an intelligent interactive assistant powered by advanced natural language processing and deep learning algorithms could provide a valuable tool to support Zimbabwe's Virginia tobacco farmers. By leveraging the latest advancements in artificial intelligence, this system could offer personalized, real-time guidance and recommendations to help farmers optimize their production practices, improve their decisionmaking, and ultimately enhance their livelihoods.

1.3 Problem statement

Due to continuous poor yields by new joining farmers and smallholder tobacco farmers because of lack of information, and poor farming practices in Zimbabwe, there a is need for a fast, reliable and sustainable technological way of helping farmers on how to conduct their farming practices using latest techniques that will assist them from the beginning up to the end of their entire production. The aim of this work is to develop an intelligent and sustainable chatbot that can reach out to thousands of tobacco farmers across Zimbabwe using and machine learning. Such a chatbot is a vital tool to disseminate valuable information (Abdul-Kader & Woods, 2015) on tobacco farming techniques and agricultural practices among the tobacco farmers.

The intelligent and sustainable chatbot can assist farmers by providing better and up to date farming techniques resulting in high farming production and output that will encourage most people to join the tobacco farming sector and at the end the country's Gross Domestic Product (GDP) will rise (Abu Shawar and Atwell, 2007).

Despite being a significant contributor to Zimbabwean economy, Virginia Tobacco farming is facing numerous challenges including:

- Low yield and decreased quality due to climate change, pests, diseases and poor crop management practises.

- Limited access to timely and accurate information on best farming practices, weather, and market trends.

- Inefficient use of resources, leading to reduces profitability and environmental sustainability.

Current Virginia Tobacco farming practices rely heavily on manual observations, trial and error approaches, and limited extension services, resulting in delayed decision-making process and ultimately reduced productivity. There is a need for an intelligent assistant that can provide Virginia Tobacco farmers in Zimbabwe with:

- Real-time guidance on optimal Virginia Tobacco crop management, pests and disease control, and weather resilience.

- Personalised recommendations on bets farming practices, input applications and market access.

- Data driven insights to inform decision-making, reduce uncertainty and improve overall efficiency and profitability.

The development and deployment of such an intelligent interactive assistant, leverage natural language processing and deep learning techniques, has the potential to transform the Virginia Tobacco farming industry in Zimbabwe. However, there is lack of research on the design, training, and implementation of such systems in this specific context, highlighting the need for investigation into the following research questions:

- How can natural language processing and deep learning algorithms be trained and validated using local data to provide accurate and relevant guidance to Virginia Tobacco farmers in Zimbabwe?

- What are the key factors influencing the adoption and effectiveness of intelligent interactive assistant among the Virginia Tobacco farmers and how can these be addressed through design and implementation?

- How can the intelligent interactive assistant for Virginia Tobacco farmers be integrated with existing farming practices, extension services, and market structures to maximise its impact and sustainability?

1.4 Research Objectives

The main objectives of this research are as follows:

1. To develop an intelligent assistant that can provide real-time recommendations and insights to Zimbabwe Virginia tobacco farmers.

2. To integrate into the Chabot natural language processing and deep learning algorithms as well as train and validate using local data to provide accurate and relevant guidance to Virginia Tobacco farmers in Zimbabwe

3. To evaluate the effectiveness and user-friendliness of the intelligent assistant through field trials and user feedback.

1.5 Justification of Virginia Tobacco Farming Assistant

With the development and deployment of an intelligent interactive Assistant for Zimbabwe Virginia Tobacco farmers using natural language processing and deep learning Algorithms, the following outcomes will be obtained.

Expected Outcomes

- An intelligent interactive assistant specifically tailored for Zimbabwe Virginia tobacco farmers, capable of providing personalized recommendations and insights.

- Enhanced decision-making capabilities for farmers through real-time data analysis and integration with agricultural knowledge and best practices.

- Improved productivity, yield, and profitability for Virginia tobacco farmers in Zimbabwe.

- A comprehensive evaluation of the intelligent assistant's impact on farmers' performance and feedback for further improvements.

The Chatbot shall also be used to connect farmers with other farmers and experts in the tobacco industry. This shall be done through online forums, chat rooms, or video conferencing. By connecting with other farmers, farmers can share information and experiences, learn from each other, and build relationships.

Finally, the Chatbot shall be used to advocate for the interests of tobacco farmers. This shall be done by contacting policymakers, writing letters to the editor, or participating in protests or demonstrations. By advocating for the interests of tobacco farmers, the Chatbot can help to ensure that farmers have a voice in the political process.

The use of a natural language processing Chatbot for Virginia tobacco farmers shall have a number of benefits, including:

- Increased access to information and resources
- Improved crop management
- Increased networking opportunities
- Increased political representation

The Chatbot shall be a valuable tool for Virginia tobacco farmers. It shall help farmers to be more successful in their businesses and to have a greater voice in the political process.

1.6 Conclusion

The development and implementation of a Natural Language Processing (NLP) and deep learning based Chatbot has been considered as such a solution to answer these farmers' questions about best growing practices, recommended seed, best time to plant, different ways of applying correct fertilizers, different watering practices, best recommended ways to control pests and diseases, as well as to help farmers troubleshoot problems with their crops and to recommend relevant solutions in time. The tobacco farmer needs to know about all this, and hence the value of a Chabot to act as an Agricultural Extension Worker (AEW) and available 24/7. However, the author will build the intelligent Chatbot with Rasa and Python 3.10.3 that can reach out to thousands of tobacco farmers across Zimbabwe to disseminate valuable information on tobacco farming techniques and recommended agricultural practices among the tobacco farmers.

Overally, the intelligent Chatbot for Virginia tobacco farmers shall be a valuable tool for providing farmers with information and assistance. The Chatbot shall help farmers to improve their yields, reduce their costs, and improve their communication with extension agents.

Here are some specific examples of how a natural language processing Chatbot shall be used to help Virginia tobacco farmers:

- The Chatbot shall be used to answer farmers' questions about best practices for growing tobacco. For example, the Chatbot shall provide information about the best time to plant tobacco, the amount of fertilizer to apply, and the best way to control pests.
- The Chatbot shall be used to help farmers troubleshoot problems with their crops. For example, the Chatbot shall help farmers to identify the cause of a disease and to recommend a course of treatment.
- The Chatbot shall be used to collect data from farmers about their crops. This data shall then be used to improve the Chatbot's responses and to better understand the challenges facing tobacco farmers.
- The Chatbot shall be used to improve communication between farmers and extension agents. For example, the Chatbot shall be used to collect questions from farmers and to deliver those questions to extension agents. This shall help to ensure that farmers are getting the information and assistance they need.

A natural language processing Chatbot for Virginia tobacco farmers shall be a valuable tool for providing farmers with information and assistance. The Chatbot shall help farmers to improve their yields, reduce their costs, and improve their communication with extension agents.

Chapter 2: Literature Review

2.1 Introduction

This literature review aims to explore the existing research and developments in the field of intelligent assistants for Zimbabwe Virginia tobacco farmers, focusing on the utilization of natural language processing and deep learning algorithms.

Zimbabwe is known for its thriving tobacco industry, with Virginia tobacco being one of the primary cash crops. However, tobacco farming comes with its challenges, including the need for accurate decision-making, pest and disease management, and optimizing crop yield. In this research, the integration of natural language processing (NLP) and deep learning (DL) algorithms in the form of intelligent interactive assistants will demonstrate great potential in supporting farmers. This literature review aims to explore the applications and benefits of ML and DL algorithms in developing an intelligent interactive assistant specifically tailored for Zimbabwe Virginia tobacco farmers.

2.2 Overview of Natural language processing and Deep Learning

Natural language processing is a subset of artificial intelligence that focuses on developing algorithms and models that enable computers to learn and make predictions or decisions without being explicitly programmed. Deep learning, on the other hand, is a subset of natural language processing that utilizes artificial neural networks to process and analyse complex data patterns. Deep learning algorithms have demonstrated exceptional performance in various domains, including image recognition, natural language processing, and data analysis.

2.3 Applications of ML and DL in Agriculture

The adoption of ML and DL algorithms in agriculture has the potential to revolutionize farming practices and enhance crop production. In the context of Zimbabwe Virginia tobacco farming, an intelligent interactive assistant powered by ML and DL algorithms can provide valuable support to farmers in several key areas.

2.3.1 Crop Monitoring and Yield Prediction

ML and DL algorithms can analyze various data sources, including satellite imagery, weather data, and soil conditions, to monitor crop health, predict yield, and identify potential issues. By leveraging historical data and real-time information, the intelligent assistant can provide farmers with accurate predictions and proactive recommendations for optimal crop management.

2.3.2 Pest and Disease Management

Early detection and effective management of pests and diseases are critical for maintaining healthy tobacco crops. ML and DL algorithms can analyze images of leaves, stems, and other plant parts to identify signs of pests and diseases. The intelligent assistant can alert farmers about potential threats, recommend appropriate control measures, and facilitate timely intervention, thus minimizing crop losses.

2.3.3 Decision Support System

ML and DL algorithms can analyze vast amounts of data, including historical crop performance, market trends, and agronomic practices, to provide personalized recommendations and decision support to farmers. The intelligent assistant can assist farmers in making informed choices regarding seed selection, fertilization, irrigation, and other critical aspects of tobacco farming.

2.3.4 Knowledge Sharing and Training

The intelligent interactive assistant can act as a platform for knowledge sharing and training. By incorporating expert knowledge and best practices into the algorithms, the assistant can provide farmers with valuable insights, tutorials, and training materials. This empowers farmers with up-to-date information and skills, promoting continuous learning and improvement.

2.3.5 Artificial Intelligence and Robots in Agriculture

The agricultural industry plays a crucial role in global food production, job creation, and economic development. With the world's population projected to reach 10 billion by 2050, there is a pressing need to enhance agricultural productivity and efficiency. In recent years, the integration of artificial intelligence (AI) and robotics technologies in agriculture has gained significant attention. This literature review aims to provide an overview of the applications and benefits of AI and robots in agriculture.

2.3.6 Emergence of AI and Robotics in Agriculture

The adoption of advanced technologies, including AI and robotics, has the potential to revolutionize the agricultural sector. AI enables machines to replicate human intelligence, allowing them to analyze data, make decisions, and perform tasks with precision. Robotics, on the other hand, involves the use of autonomous or semi-autonomous machines to carry out agricultural operations. These technologies offer numerous advantages, such as increased productivity, reduced labor requirements, improved resource management, and enhanced crop quality.

2.3.7 Applications of AI and Robotics in Agriculture

AI and robots have found applications in various aspects of agriculture. One prominent area is crop management, where AI-based systems can provide real-time monitoring of crop conditions, weather forecasting, and pest detection. This enables farmers to make informed decisions regarding irrigation, nutrient management, and pest control, leading to higher yields and reduced losses. Additionally, robots equipped with sensors and cameras can perform tasks like planting, harvesting, and weed control with precision and efficiency.

- Another significant application of AI and robotics in agriculture is smart irrigation. By integrating AI algorithms and sensor technologies, farmers can optimize water usage by providing the right amount of water to crops based on their specific needs. This not only conserves water resources but also prevents over-irrigation, which can lead to soil degradation and water pollution.

- Furthermore, AI and robotics play a crucial role in crop protection. AI-powered systems can detect diseases, pests, and weed infestations at an early stage, enabling timely intervention and targeted treatment. Robots equipped with spraying mechanisms can apply pesticides or herbicides precisely, minimizing chemical usage and reducing environmental impact.

2.4 Benefits and Future Prospects

The adoption of AI and robotics in agriculture offers several benefits. These technologies improve productivity by automating labor-intensive tasks, reducing human error, and enabling round-the-clock operations. They also enhance resource efficiency by optimizing inputs like water, fertilizers, and pesticides. Moreover, AI-based systems provide valuable insights and recommendations for decision-making, empowering farmers to make informed choices and improve overall farm management.

The integration of ML and DL algorithms in an intelligent interactive assistant for Zimbabwe Virginia tobacco farmers offers several benefits. It enhances productivity by optimizing crop management practices, reducing losses due to pests and diseases, and improving yield predictions. The intelligent assistant also contributes to sustainable farming practices by minimizing the use of pesticides and fertilizers through targeted recommendations.

Looking ahead, there are significant opportunities for further advancements in this field. Ongoing research in ML and DL algorithms can lead to more accurate predictions, improved decision support, and enhanced performance of the intelligent assistant. Integration with other emerging technologies, such as Internet of Things (IoT) devices and cloud computing, can expand the capabilities and scalability of the assistant.

2.5 Related Work

According to Lin (2023), Colace et.al. (2018) and Muzurura et.al (2023) various AI chatbots have been developed in various fields, such as healthcare, marketing, education, supporting systems, cultural heritage, entertainment and many other industries. Chatbots are systems designed for extended conversations, set up to mimic the unstructured conversational or

‘chats’ characteristic of human-human interaction and some are focusing on a particular tasks. Such bots can talk about general topics and respond appropriately for example ChatGPT (developed by Google), it uses natural language processing (ML) concept and learns from previous experiences. According to Colace et.al. (2018), majority of big companies have designed and developed their own chatbots both for business solutions and for exploration on their business ideas. Apple Siri, Microsoft Cortana, Facebook Messenger and IBM Watson are good and notable famous systems in such category. They also claimed that there are also wide ranges of a less famous chatbots that have a greater relevance for research and for their applications Colace et al. (2018).

Currently, various tobacco grower programs as well as companies and the Zimbabwean government have employed extension workers and field officers. Their main duty is to visit the tobacco fields and interact with tobacco farmers in rural areas, A1 and A2 farms in a bid to provide them with training, give advice on best tobacco farming practices and other crucial aspects of good tobacco farming that yields recommended quality of tobacco yield. They also monitor and inspect the crop on all stages, which is; from seed bed to curing, such that good quality can be achieved. Due to the distances long to travelled and the limited means to travel from one farmer to another, vital information is delayed in time of critical need and this have led to poor quality yield in most cases and hence a need for a Chabot to assume the role advisory and information center for the farmers in need.

Information & Communication Technology (ICT) is now being actively used in agriculture farming communications and with a notable and significant adoption of the use of smart phones by farmers for communication with their field agents and extension workers.

Farmers' Portal (which is a central repository of all the information required by farmers at all stages at one location), Call Centers (where a toll free number is available for farmers to query) , sms system(which is used to send sms advisories by Agriculture Officers and Experts to farmers and other registered users) and mobile applications (mostly available on Google Play) have been developed to provide information that a farmer would need from sowing stage to post harvesting including market prices and weather alerts on their mobile. This have worked and still working to many tobacco farmers but due to limited time on the device and millions of search results, a Chatbot becomes the best option as it refines an answer to a farmer's request in real time without delay.

2.6 Chatbot building methods

Agreeing to Lin C.C. et al (2023), researchers are developing contemporary conversational chatbots for either opened or closed domains using artificial intelligence (AI) techniques, such as NLP and artificial neural networks, to understand and respond to users' input. Stretching from casual and open-domain to more domain-specific and fact-based, these chatbots are developed by means of a number of deep learning models, such as RNN (Recurrent Neural Network), Seq2Seq (Sequence to Sequence), LSTM (Long Short-Term

Memory), BERT (Bi-directional Encoder Representation from Transformers), GPT (Generative Pre-trained Transformer)), as well as enforcing various training techniques, such as reinforcement learning or transfer learning, in order to improve the performance of NLP algorithms and chatbots. Lin C.C et al (2023) also indicated that ChatGPT is the latest inspiring, popular and encouraging case that have received a great amount of attention and inspired researchers to create new concepts concerning Chatbot applications.

2.7 Natural Language Processing

In Bafna S. et al. (2021) NLP is described as a combination of two components, Natural Language Understanding (NLU) and Natural Language Generation (NLG). NLU recognizes the meaning of the user's input, in actual fact; it gets the Chatbot to comprehend what a body of text means. The NLU recognizes a body of text input given and classifying it into proper intents. The NLU involves:

- Natural Language Inference (NLI) and paraphrasing
- Dialogue agent
- Semantic parsing
- Question answering
- Sentiment analysis
- Summarization of users' input

Bafna S. et al. (2021) also defined NLG as software that produces understandable texts in human languages. They also claimed that NLG techniques provide ideas on how to build systems that can take advantage of the knowledge and capabilities of both humans and machines. NLG involves:

- Content determination

- Discourse planning
- Sentence aggregation
- Lexicalisation
- Referring expression generation

2.8 Chatbot Dataset

Lin C.C. et al (2023) claimed that an application-specific dataset is normally used by closed domain chatbots because these kinds of chatbots have a very clear goal when a conversation is happening. A good example is SupeRAgent (Cui, L. et al. 2017) and a customer contact system (Pawlik, L et al 2022.); they leverage pre-defined or historic system-generated data to train the system since it is unique to the Chatbot. Another example is when a Chatbot is built as a learning companion to improve users' reading comprehension skills (Liu, C.C et al. 2022 and Hollander, J. 2022); the dataset used to train the system includes the books the Chatbot will use, which means that the conversation scope is fixed. In this scenario, the researcher will focus on a fixed dataset for the tobacco famers Chatbot based on web scrapping, books, journals, government database etc.

2.9 Conclusion

The literature review concludes by summarizing the findings and emphasizing the significance of intelligent interactive assistants powered by natural language processing and deep learning algorithms in the context of Zimbabwe Virginia tobacco farming. It highlights the potential of these technologies to enhance agricultural productivity, improve resource management, and support sustainable farming practices. The review encourages further research and development in this area to address the specific needs and challenges of tobacco farmers in Zimbabwe.

The development of an intelligent interactive assistant for Zimbabwe Virginia tobacco farmers using ML and DL algorithms holds great promise in addressing the challenges faced by farmers in decision-making, pest and disease management, and optimizing crop yield. By leveraging the power of data analysis and pattern recognition, the assistant can provide personalized recommendations, training, and real-time insights. It is crucial to continue research, collaboration, and implementation efforts to ensure the successful adoption and widespread accessibility of this technology, ultimately benefiting Zimbabwe's tobacco farming community.

Chapter 3: Methodology

3.1 Introduction

This chapter presents the methodology for developing the natural language processing Chatbot for Virginia tobacco farmers. The first task is to identify the needs of tobacco farmers. In this case, the author looks at what information the tobacco farmers need as well as what problems the farmers need solutions with. The second exercise is data collection. In this case, data is be collected from a variety of sources such as farmers' surveys, extension service reports, Agricultural reports, Tobacco Industry and Marketing Boards as well as from the government database. This data is be used to create a dataset as well as to train the Chatbot. The third task is to train the Chatbot. In this case, the Chatbot is trained on a variety of natural language processing algorithms, such as NLP and machine translation with training data created from the previous exercise. The final task is to test the Chatbot. The Chatbot will be tested with a variety of farmers to ensure that it is accurate, reliable, and easy to use before it can be it can be deployed on a variety of platforms, such as websites, mobile apps, and social media for final use.

This chapter outlines the methodology employed in the development of an intelligent interactive assistant for Zimbabwe Virginia tobacco farmers using natural language processing (ML) and deep learning (DL) algorithms. The chapter describes the data collection process, the ML and DL techniques used, and the overall system architecture. The goal is to provide a comprehensive understanding of the approach and techniques used to create the intelligent assistant.

3.2 Data Collection

The first step in developing the intelligent assistant is to collect relevant data for training and validation. Data sources include historical crop performance records, weather data, pest and disease databases, and expert knowledge retrieved from TIMB. The data cover a representative sample of Virginia tobacco farmers, Extension officers, and field agents for contract farming companies in Zimbabwe, capturing various environmental conditions, crop management practices, and pest and disease occurrences. An experimental campaign was conducted during the 2023-2024 tobacco farming season. In particular, the aim of the experiment was the evaluation of the Chatbot system's effectiveness in the recognition of the tobacco farmers' requests. Additionally, the usability of the Chatbot system was evaluated. During the 2023-2024 tobacco farming season, all the Agricultural field officers and all the tobacco farmers used the Chatbot platform which was developed for Android mobile phones and the experimental analysis has been conducted. The users who liked the Chatbots at the first call were 135 farmers and 38 Agricultural officers. At the end of the season, an email with a questionnaire has been sent to the tobacco farmers and Agricultural officers who liked the Chatbot, and in a month about 130 farmers and 35 Agricultural officers replied with the completed questionnaire. The experimental analysis was conducted on these ones. First of all, the effectiveness of the Chatbot in providing the correct recommendations to the user has been evaluated.

To ensure data accuracy and quality, data cleaning and pre-processing techniques were employed to try to filter the data. This involved removing outliers, handling missing values, and normalizing the data to ensure consistency across different variables. The cleaned and pre-processed data forms the foundation for training and evaluating the ML and DL models.

3.3 ML and DL Techniques

Various ML and DL techniques are employed to develop the intelligent assistant. The specific algorithms chosen depend on the nature of the problem and the data available. Some commonly used techniques in this context include:

3.3.1 Supervised Learning

Supervised learning algorithms are utilized for tasks such as crop yield prediction and pest and disease classification. Algorithms like decision trees, random forests, support vector machines

(SVM), and deep neural networks are trained on labelled data to learn patterns and make predictions. The labelled data consists of input features (e.g., weather parameters, soil conditions) and corresponding output labels (e.g., crop yield, pest or disease presence).

3.3.2 Unsupervised Learning

Unsupervised learning algorithms are employed for tasks like clustering similar farms or identifying patterns in historical crop performance data. Algorithms such as k-means clustering, hierarchical clustering, and principal component analysis (PCA) are used to discover hidden structures and relationships in the data without the need for labelled examples.

3.3.3 Deep Learning

Deep learning algorithms, particularly convolutional neural networks (CNNs), are well-suited for image analysis tasks, such as leaf disease detection or pest identification. CNNs can learn hierarchical representations of images and automatically extract relevant features. Transfer learning, where pre-trained models on large-scale image datasets are fine-tuned, can be utilized to leverage existing knowledge and accelerate model training.

3.3.4 System Architecture

The intelligent interactive assistant is designed to provide a user-friendly interface for farmers to interact with the system. The system architecture consists of several modules, including:

3.3.4.1 Data Integration and Pre-processing

This module handles the integration of various data sources, performs data cleaning and preprocessing, and prepares the data for model training and inference.

3.3.4.2 ML (NPL) and DL Techniques

A NLP and DL model was developed and trained. The NLP and DL model was trained using the collected and pre-processed data. The model was implemented using appropriate nltk libraries and RASA over python 3.10.0 for the NLP. Tensorflow version 2.12.2 with keras library where used to support the CNN Model whose hyper parameters were tuned with cross-validation to optimize performance.

3.3.4.3 Inference and Decision Support

Once trained, the models are deployed in the intelligent assistant for real-time inference and decision support. Given new input data, the models generate predictions and recommendations based on the learned patterns and rules.

3.3.4.4 User Interface

The user interface module provides an intuitive interface for farmers to interact with the intelligent assistant. It may include features such as data input, visualization of predictions, alerts for pest or disease outbreaks, and access to educational resources.

3.3.4.5 Evaluation and Validation

The developed intelligent assistant is evaluated and validated to assess its performance and effectiveness. Evaluation metrics depend on the specific tasks, such as accuracy, precision, recall, or mean squared error. The system is tested using both historical data and real-time scenarios to ensure its robustness and reliability.

3.3.5 System Architecture

The architecture of our natural language processing Chatbot for Virginia tobacco farmers' model is composed of:

- Front-End
- Back-Office
- Knowledge Base Module
- Chatbot Module

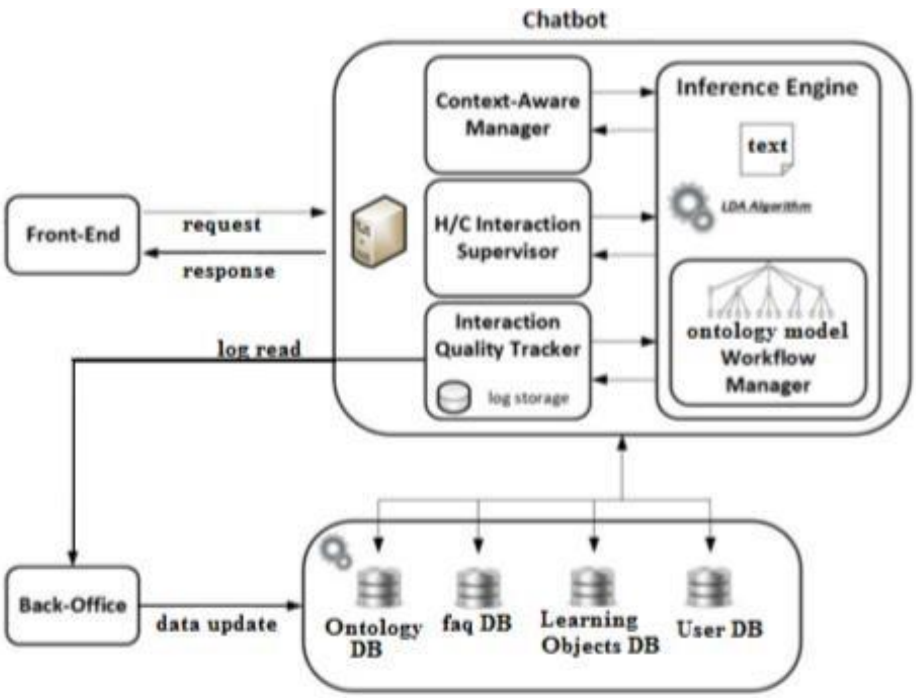


Figure 1: Chatbot Architecture by Colace F. et al. (2018)

Figure :1 above is of the system architecture as adopted from Colace F. et al. (2018), designed for their work on "Chatbot for E-Learning". The researcher adopted a similar architecture to model the chatbot. In the figure above, the user interacts with the system via the presentation layer (front end) which provides a userfriendly mobile interface that can be accessed through various kinds of different kinds of device like tablets and smartphones. The Back-Office manages back-end operations to better satisfy user demand such as handling logic and data storage, and knowledge base transactions. The Knowledge Base Module is the special type of database, where data is processed by a server, for the management of knowledge and information. The Chatbot Module in the Figure 2: above is the main engine of the proposed system whose roles includes; Interaction Quality Tracker (which monitors interactions between users and Chatbot, evaluating conversation logs based on quality indicators and highlighting critical aspects of Human-machine interactions), Human Computer Interaction Supervisor (which supervises dialogue, tracks interaction times, identifies ambiguous questions, recognizes non-convergent interaction sessions) , ContextAware Information Manager (which allows to drive the dialogue based on contextual) and the Inference Engine(which provide right answer to the user through Latent Dirichlet Allocation Algorithm and Workflow Manager which have definition of an ontology for the description of a certain knowledge domain and definition of a workflow navigation module as described by Casillo M. et al. (2016) and in Colace F. et al (2017).

Colace F. et al (2018) claimed that according to the conversation (through word analysis), there is need to surf the ontology and select the more appropriate sentences and that will require a way for the description of the workflow which made the Petri Net effective way to do so. When we adopt a Petri Net, each phase of the conversation is modelled as a node of a Petri Net while the transaction is obtained so that the right intent/request of the end user identified. By using ontological representation, we will be able to infer a user's intention, even though the user may not know what he/she wants exactly.

The following Figure 2: is the proposed propose a framework that applies the ontology technology to the Chatbot environment such that the Chatbot systems can be more intelligent, powerful and adaptive.

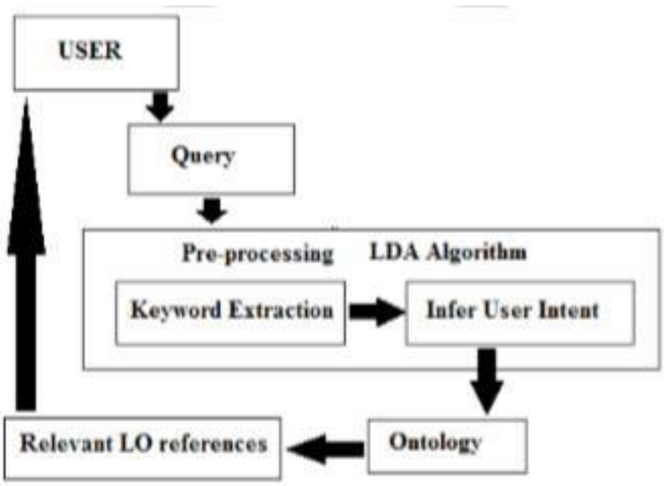


Figure 2: Reference framework

The main purpose of using ontological representation is to gain the ability of inferring a user's intention, even though the user may not know what he/she wants exactly.

3.3.6 Implementation

The researcher implemented the chatbot with Python 3.10.0 and Rasa 3.6.0 to build the required Chatbot system. We choose Rasa because Rasa has a lot of advantages such as: highly customisable with various pipelines and can be employed to process user dialogues, can be run as a simple http server or can be used from python using APIs. It has the Rasa-nlu, when run on a server; it can mimic other commercial NLP platforms such as LUIS or wit.ai.

The implementation was categorized into; installation, using dataset to train the model, custom actions/forms and domain data/ integrating with API and Database. Testing and evaluation was done at the end.

1. Installation of Python and Rasa, setting-up and programming: In this step all the necessary installation, configuration and library imports were done. The following are the basic steps:

    - Install Python 3.10
    - Create the "Virtual Environment"
    - Setup and activate the virtual environment using the below CLI commands.
      python -m venv ./venv .\venv\Scripts\activate
    - Created a project directory **mkdir farmerChatbot**
    - Install Rasa open-source with the below command **pip install Rasa**

      NB: the drive must be online since the command download and install, Rasa will automatically start downloading and installing Rasa current updated version.
    - Chatbot Programming
    - The following are the actual parts that were required and needed to be programmed for Chatbot using Rasa Framework.

a. The Rasa Server: This is the actual software that runs and interprets the user's input and handles the responses based on a trained model (more on that below).

b. The Action Sever: This handles some of the complex form logic, and/or API calls.

c. The chat interface - Rasa has connectors for this major services to allow us to interact with the bot. There are also open source and commercial chat interfaces. These are not part of the normal Rasa framework. Facebook Messenger, Slack, Telegram or other service are good examples of such interfaces.

2. Using Dataset to train the model: In this step we use our dataset to train the model. We enter the training data in the form of Entities and Intent in specific file called NLU.md. We also use user conversations examples that help the chat-bot. These are mentioned in the stories.md file.
3. Custom actions/forms and Domain data: In this step we use python to create custom actions. These custom actions will help our chat-bot to respond in certain way to a certain user input. We mention these custom actions in a file called actions.py. We have a domain.yml file in the project directory. This is used mention the domain and domain related metadata for the chat-bot.
4. Integrating with API and Database When query or input is received by Rasa from the end user; it will predict the values of entities and intents from the message, all this handling is done by RASA NLU unit.
5. Testing and evaluating the model: This is the step where we test our model with our test data. Testing can be done using CLI. In testing we just take sample data and give it as input to the model to see if we get expected output and with what accuracy.

3.3.7 How RASA works

Rasa's technology can understand natural language and decide about the next best action based on the context of the conversation using Machine Learning

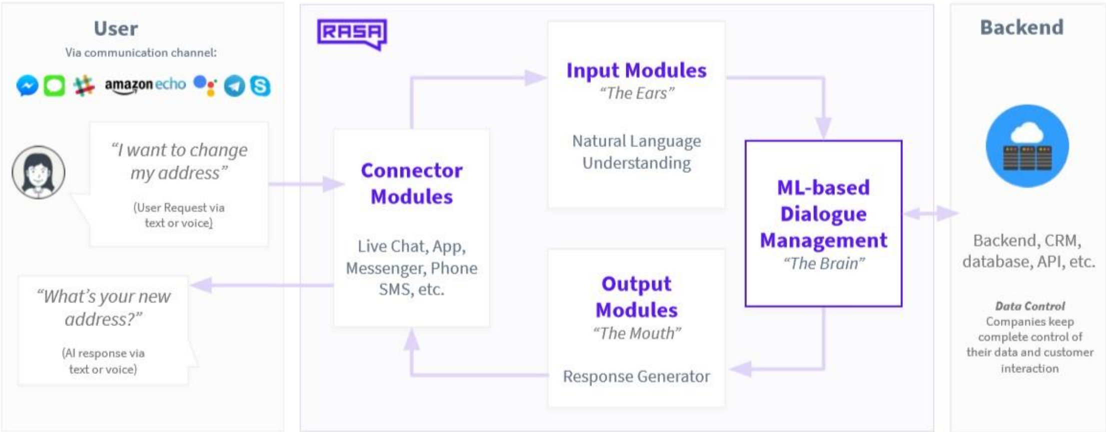


Figure 3: Chatbot interacts with Rasa NLU and Rasa Core

The figure above is a context illustration of how the user chatbot interacts with RASA NLU and Rasa Core during a natural language processing based chat session. it uses Supervised Word Vectors (SWV) from scratch to extract the structured data from unstructured user inputs and generate appropriate responses to the user. A user input can have more than one intention thus enabling the assistant to understand them, which leads to more natural conversations.

3.3.8 Why Dialogue Handling with Core?

With the aid of Rasa Core, the system can learn from real conversational data instead of enforcing rules. The core leverage the power of ML to build assistants that scale in production as it hold back-and-forth conversations that remember context. the Rasa core adds business logic when needed The Core also learns to Converse from real conversational data.

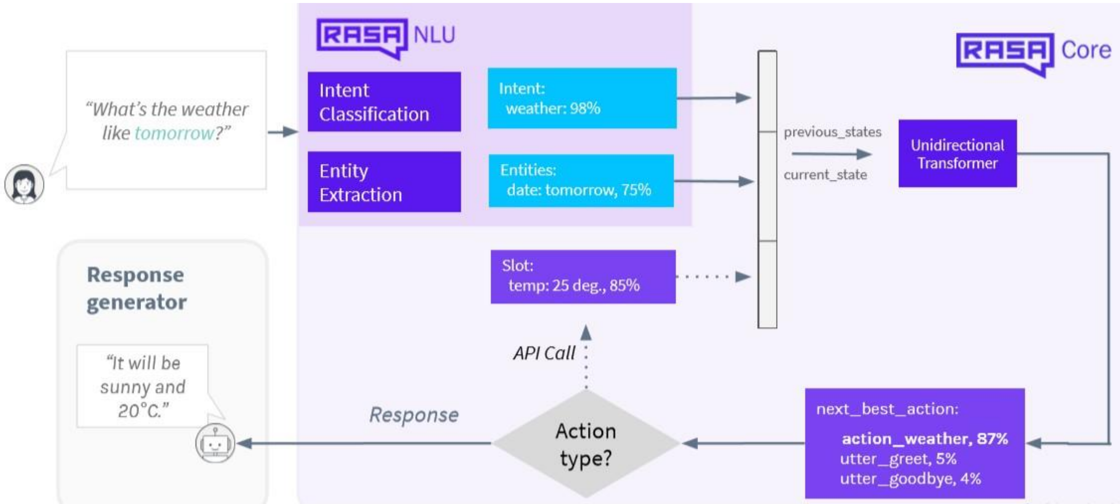


Figure 4: Rasa Core: Dialogue Handling

Rasa Core: Dialogue Handling

3.4 Population and Sample

A model Chatbot was deployed and implemented in the vicinity of the Bindura University and Shamva in Zimbabwe to cover local tobacco farmers for the 2022 -2023 tobacco season in the Mashonaland central province of Zimbabwe. The researcher considered 2 focus groups: Agricultural field officers and the Tobacco farmers (new and residual). The first focus group was comprised of 40 Agricultural field officers and the second one by 150 farmers.

3.5 Conclusion

This chapter presented the methodology employed in the development of an intelligent interactive assistant for Zimbabwe Virginia tobacco farmers using ML and DL algorithms. It covered the data collection process, ML and DL techniques utilized, and the overall system architecture. The next chapter will focus on the implementation details and present the results of the developed intelligent assistant.

Chapter 4 Results

4.1 Introduction

This chapter focus on the presentation of the research findings obtained from the research practice. In order to assess the effectiveness of the proposed natural language processing based Chatbot system built from Rasa framework and Python, an experimental campaign was conducted during the 2022-2023 tobacco farming season. In particular, the aim of the experiment was the evaluation of the Chatbot system's effectiveness in the recognition of the tobacco farmers' requests. Additionally, the usability of the Chatbot system was evaluated. During the 2022-2023 tobacco farming season, all the Agricultural field officers and all the tobacco farmers used the Chatbot platform which was developed for Android mobile phones and the experimental analysis has been conducted. The users who liked the Chatbots at the first call were 135 farmers and 38 Agricultural officers. At the end of the season, an email with a questionnaire has been sent to the tobacco farmers and Agricultural officers who liked the Chatbot, and in a month about 130 farmers and 35 Agricultural officers replied with the completed questionnaire. The experimental analysis was conducted on these ones.

4.2 AgriBot, the Farmers' Chatbot

Successful implementation of the natural language processing and deep learning based system indicated a major step in this research project. The following illustration is the screenshot of the final implementation on the desktop version of the system.

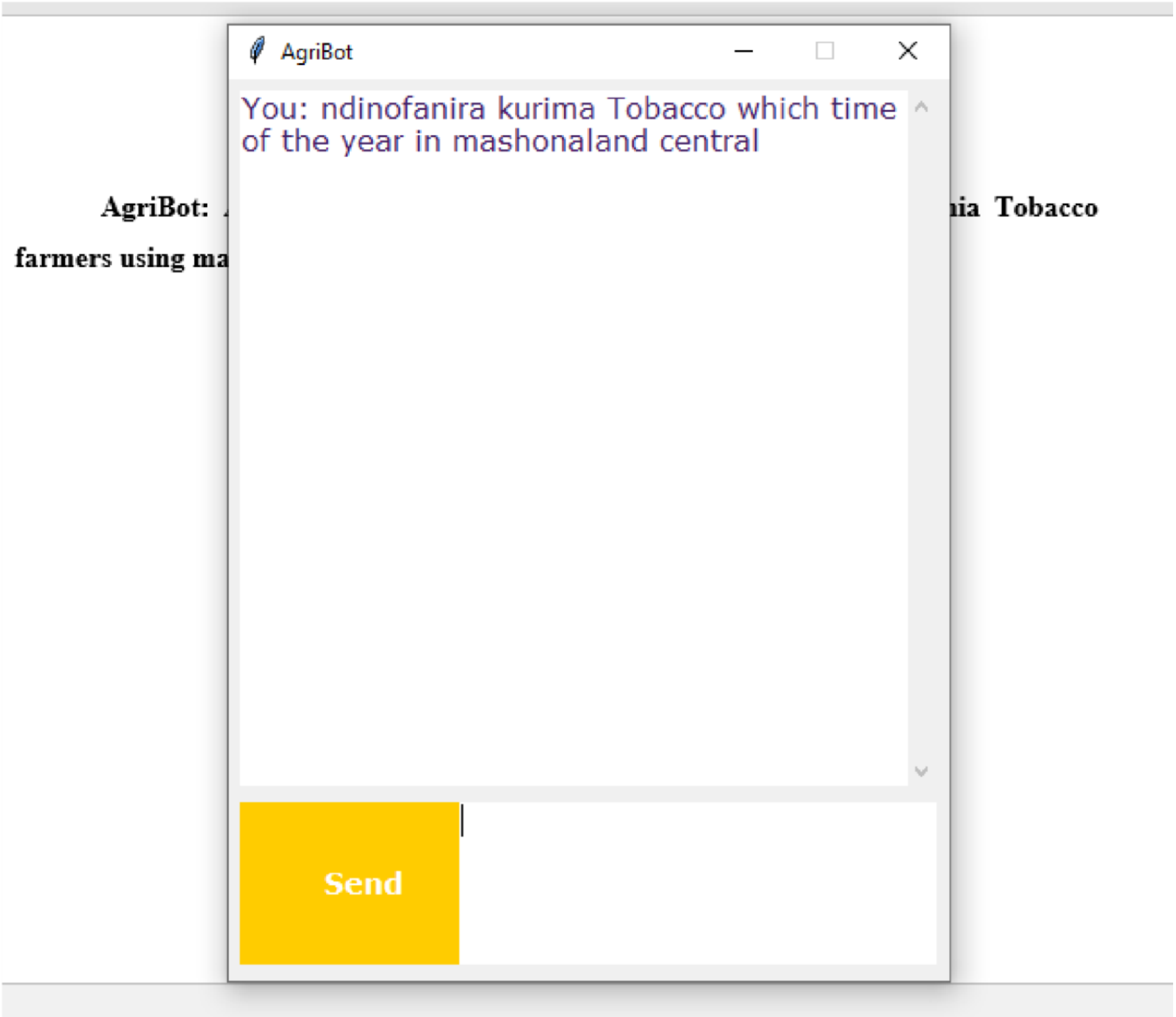


Figure 5: Chatbot Interface

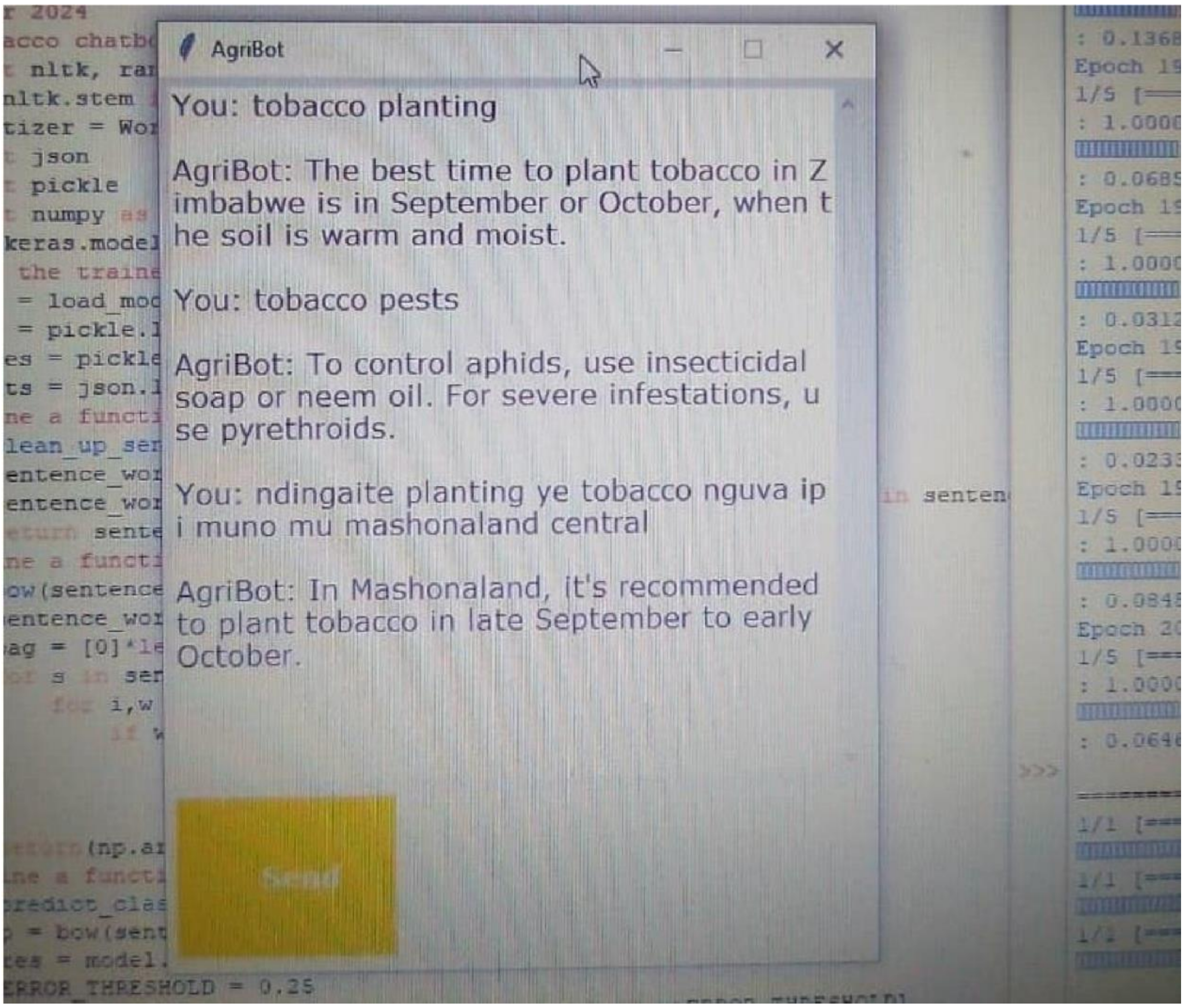


Figure 6; Testing the Chatbot

The deep learning neural network model

```
# Create model - 3 layers. First layer 128 neurons, second layer 64 neurons and
# equal to number of intents to predict output intent with softmax
# Sequential() allows you to create models layer-by-layer
model = Sequential()
# Dense layer is a regular layer of neurons in a neural network
model.add(Dense(128, input_shape=(len(train_x[0]), ), activation='relu'))
# Dropout is used for prevent overfitting.
# Dropout works by randomly setting the outgoing edges of hidden units (neurons
model.add(Dropout(0.5))
model.add(Dense(64, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(len(train_y[0]), activation='softmax'))

# Compile model. Stochastic gradient descent with Nesterov accelerated gradient
# nesterov is an optimal method (in terms of oracle complexity) for smooth conve
sgd = SGD(lr=0.01, decay=1e-6, momentum=0.9, nesterov=True)
# loss is the loss function
model.compile(loss='categorical_crossentropy',
              optimizer=sgd,
              metrics=['accuracy'])

# fitting and saving the model
# By setting verbose 0, 1 or 2 you just say how do you want to 'see' the trainin
# The batch size defines the number of samples that will be propagated through t
hist = model.fit(np.array(train_x),
                 np.array(train_y),
                 epochs=200,
                 batch_size=5,
                 verbose=1)
model.save('agribot_model.h5', hist)
```

Figure 7;The deep learning neural network model

The figure above illustrates the definition of deep learning neural network which generate the responses to the user request. The words.pkl and classes.pkl files are generated from the " intents.json" file and then used for training. The figure bellow illustrate the contents of the words and classes file used for training the deep learning model called agribot_model.h5.

```
, 'prevent', 'disease', 'infestations', 'on', 'my', 'tobacco', 'farm', '?'], 'di
seasemanagement'), (['What', 'are', 'the', 'symptoms', 'of', 'tobacco', 'disease
s', '?'], 'diseasemanagement'), (['When', 'is', 'the', 'best', 'time', 'to', 'ha
rvest', 'tobacco', 'in', 'Zimbabwe', '?'], 'harvesting'), (['How', 'do', 'I', 'c
ure', 'tobacco', 'after', 'harvesting', '?'], 'harvesting'), (['What', 'are', 't
he', 'different', 'types', 'of', 'tobacco', 'harvesting', 'methods', '?'], 'harv
esting'), (['Where', 'can', 'I', 'sell', 'my', 'tobacco', 'in', 'Zimbabwe', '?']
, 'marketing'), (['What', 'is', 'the', 'current', 'market', 'price', 'for', 'tob
acco', 'in', 'Zimbabwe', '?'], 'marketing'), (['How', 'do', 'I', 'market', 'my',
 'tobacco', 'to', 'potential', 'buyers', '?'], 'marketing'), (['What', 'are', 't
he', 'regulations', 'for', 'tobacco', 'farming', 'in', 'Zimbabwe', '?'], 'regula
tions'), (['Do', 'I', 'need', 'a', 'license', 'to', 'grow', 'tobacco', 'in', 'Zi
mbabwe', '?'], 'regulations'), (['What', 'are', 'the', 'laws', 'regarding', 'tob
acco', 'marketing', 'in', 'Zimbabwe', '?'], 'regulations')]
24 documents
8 classes ['diseasemanagement', 'fertilization', 'harvesting', 'irrigation', 'ma
rketing', 'pestmanagement', 'planting', 'regulations']
70 unique lemmatized words ['a', 'affect', 'after', 'aphid', 'apply', 'are', 'be
st', 'buyer', 'can', 'common', 'control', 'crop', 'cure', 'current', 'different'
, 'disease', 'do', 'farm', 'farming', 'fertilize', 'fertilizer', 'for', 'grow',
'harvest', 'harvesting', 'how', 'i', 'in', 'infestation', 'irrigate', 'irrigatio
n', 'is', 'law', 'license', 'market', 'marketing', 'mashonaland', 'method', 'muc
h', 'my', 'need', 'of', 'often', 'on', 'pest', 'plant', 'planting', 'potential',
 'prepare', 'prevent', 'price', 'recommended', 'regarding', 'regulation', 'sell'
, 'should', 'soil', 'symptom', 'system', 'that', 'the', 'time', 'to', 'tobacco',
 'type', 'water', 'what', 'when', 'where', 'zimbabwe']
```

Figure 8; Words and Classes file

Creating training data and training the model

```
# create the training data
training = []
# create an empty array for our output
output_empty = [0] * len(classes)
# training set, bag of words for each sentence
for doc in documents:
  # initialize our bag of words
  bag = []
  # list of tokenized words for the pattern
  pattern_words = doc[0]
  # lemmatize each word - create base word, in attempt to represent related word
  pattern_words = [
      lemmatizer.lemmatize(word.lower()) for word in pattern_words
  ]
  # create our bag of words array with 1, if word match found in current pattern
  for word in words:
    bag.append(1) if word in pattern_words else bag.append(0)

  # output is a '0' for each tag and '1' for current tag (for each pattern)
  output_row = list(output_empty)
  output_row[classes.index(doc[1])] = 1

  training.append([bag, output_row])

# shuffle our features and turn into np.array
random.shuffle(training)
training = np.array(training)

# create train and test lists. X - patterns, Y - intents
train_x = list(training[:, 0])
train_y = list(training[:, 1])
print("Training data created")
```

Figure 9;Creating training data and training the model

Machine learning model to handle user input synthesis and determine what to respond to

```
#Define a function to clean up user input:
def clean_up_sentence(sentence):
    sentence_words = nltk.word_tokenize(sentence)
    sentence_words = [lemmatizer.lemmatize(word.lower()) for word in sentence_wo
    return sentence_words
#Define a function to create a bag of words from user input:
def bow(sentence, words, show_details=True):
    sentence_words = clean_up_sentence(sentence)
    bag = [0]*len(words)
    for s in sentence_words:
        for i,w in enumerate(words):
            if w == s:
                bag[i] = 1
                if show_details:
                    print ("found in bag: %s" % w)
    return(np.array(bag))
#Define a function to predict the user's intent:
def predict_class(sentence, model):
    p = bow(sentence, words,show_details=False)
    res = model.predict(np.array([p]))[0]
    ERROR_THRESHOLD = 0.25
    results = [[i,r] for i,r in enumerate(res) if r>ERROR_THRESHOLD]
    results.sort(key=lambda x: x[1], reverse=True)
    return_list = []
    for r in results:
        return_list.append({"intent": classes[r[0]], "probability": str(r[1])})
    return return_list
```

Figure 10; Machine learning model to handle user input

4.3 Chatbot evaluation

First of all, the effectiveness of the Chatbot in providing the correct recommendations to the user has been evaluated. Specifically, three dissimilar circumstances where deliberated on:

- Chatbot providing the correct recommendations
- Chatbot providing the correct recommendations, but it does not fit with the real needs of the farmer / field officer
- Chatbot providing the wrong recommendations

From the 130 farmers and 35 Agricultural Field officers who completed the questionnaire, the following results were obtained:

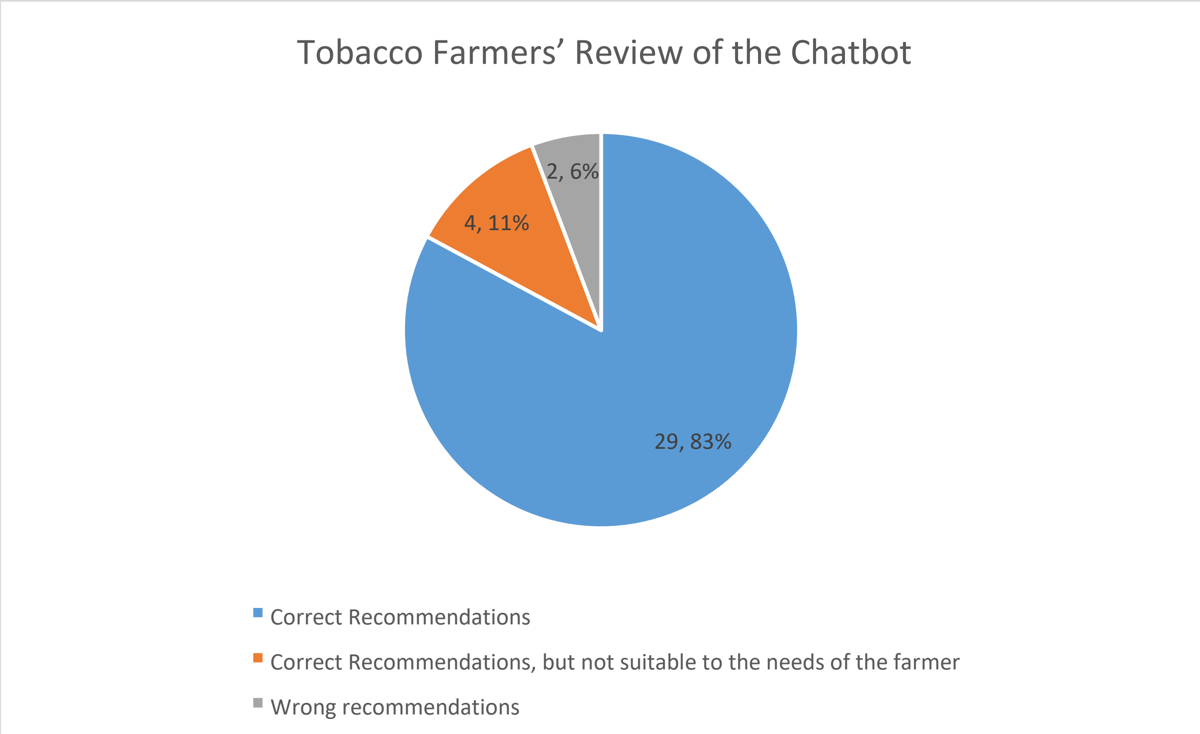


Table 1: Tobacco Farmers’ Review of the Chatbot

	Tobacco Farmers
Correct Recommendations	123
Correct Recommendations, but not suitable to the needs of the farmer	5
Wrong recommendations	2

Figure 11: Tobacco Farmers’ Review of the Chatbot Table 2: Agricultural Field Officers’ Review of the Chatbot

	Agricultural Field Officers
Correct Recommendations	**29**
Correct Recommendations, but not suitable to the needs of the farmer	**4**

Wrong recommendations	**2**

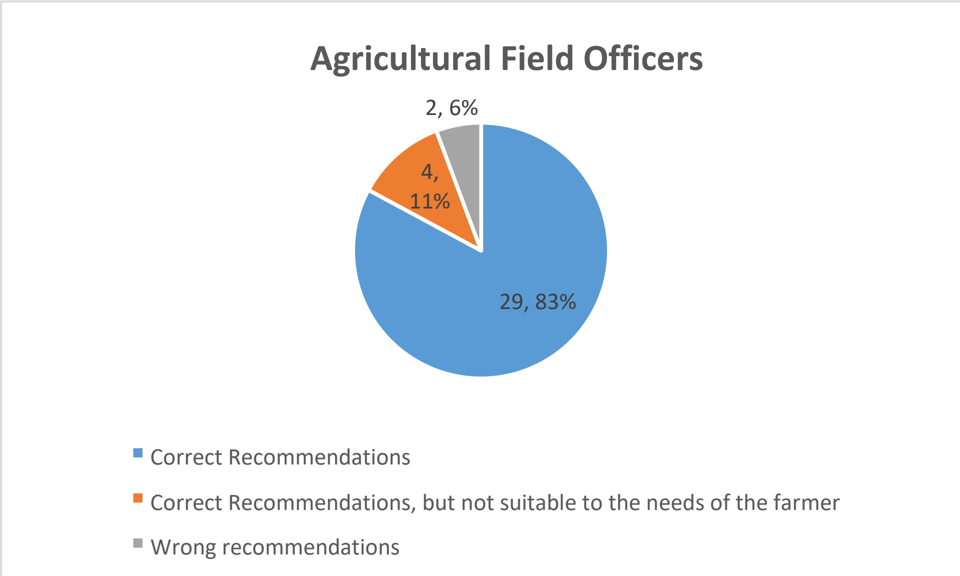


Figure 12: Agricultural Field Officers' Review of the Chatbot

4.4 Fall Back Rates (FBR)

Most Chatbots have fallback answers, programmed to suit the user if he "explores" areas that are still unknown to your robot. Usually, the VA says: "Not sure of what you mean, Please Try ChatGPT for further assistance!"

Monitoring incidents of this type of response was crucial, as this may mean the need for dataset update and re-training or simply the identification of new intentions and entities not currently covered in the bot design.

If we divide the number of times the Chatbot has had to use a fallback response by the total number of messages, we will have the rate of confusion.

Confusion rate = number of fallback answers / total answers offered

The Confusion metric was developed where a total of 60 questions on the Chatbot from the 30 random participants, with two questions from each participant were crafted. From the 60 questions the number of fallback answers was 4.

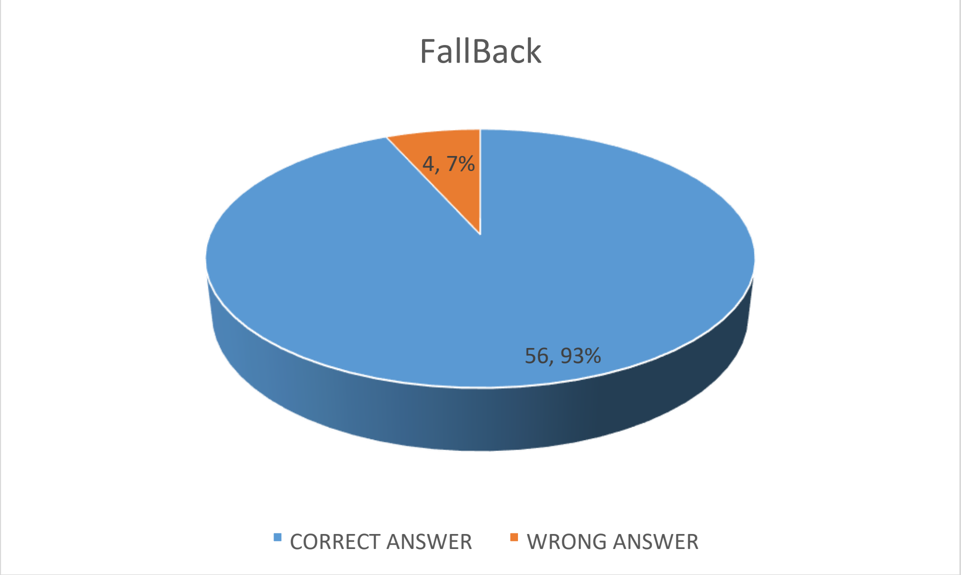


Figure 13: Chatbot Fallback

Therefore, the confusion rate =4/60

=0.0667

=6.67% fallback rate

4.5 Goal Completion

The VA, did not respond well to 4 questions from the 60 questions given which gives the success rate of 93.3%. The 6.67% fallbacks mean that the questions missed where outside the scope of the VA. This means that the VA requires an update on dataset and re-training since technology changes. There is usually new knowledge generated every day, hence the need for re-training to update the system knowledge base.

The figure above, does indicates that the VA, did not respond well to 4 questions from the 60 questions given which gives the success rate of 93.3%. The 6.67% fallbacks mean that the questions missed where outside the scope of the VA. Most Chatbots have fallback answers, programmed to suit the user if he “explores” areas that are still unknown to your robot. In our case, the VA says: “Not sure of what you mean, Please Try ChatGPT for further assistance!” The 93.3% indicates that majority of the responses from the Chatbot where relevant, while the 6.67% indicates that some responses from the Chatbot where either wrong or correct but not the users really wanted as the correct response.

4.6 Conclusion

A natural language processing Chatbot for Virginia tobacco farmers shall have a number of positive results. First, the Chatbot shall provide farmers with access to information and assistance 24/7. This shall help farmers to make better decisions about their crops and to troubleshoot problems more quickly.

Second, the Chatbot shall help farmers to save time by answering their questions and providing recommendations. This shall free up farmers to focus on other tasks, such as growing and harvesting their crops.

Third, the Chatbot shall help farmers to reduce their costs by providing them with information about best practices and by helping them to troubleshoot problems. This shall save farmers money on things like pesticides, fertilizers, and crop insurance.
Fourth, the Chatbot shall help farmers to improve their communication with extension agents. The Chatbot shall be used to collect data from farmers, which shall then be used to improve the Chatbot's responses and to better understand the challenges facing tobacco farmers. This information shall also be used to improve communication between farmers and extension agents.

Overall, a natural language processing Chatbot for Virginia tobacco farmers shall be a valuable tool for providing farmers with information and assistance. The Chatbot shall help farmers to improve their yields, reduce their costs, and improve their communication with extension agents.

Chapter 5 Analysis and Discussion

5.1 Introduction

The first focus group was comprised of 40 Agricultural field officers and the second one by 150 farmers. The users who liked the Chatbots at the first call were 38 Agricultural officers which was and 135 farmers which constituted 95% and 90% respectively. At the end of the season, an email with a questionnaire has been sent to the tobacco farmers and Agricultural officers who liked the Chatbot, and in a month about 130 farmers and 35 Agricultural Field Officers replied with the completed questionnaire.

5.2 Chatbot providing the correct recommendations

From the 130 farmers who completed the questionnaire, 95% farmers indicated that the Chatbot provided them with correct recommendations, 4% of the farmers indicated that the Chatbot provided them with correct recommendations but not fit for their real needs and 1% indicated that the Chatbot provided them with wrong recommendations. From the 35 Agricultural Field Officers who completed the questionnaire, 83% farmers indicated that the Chatbot provided them with correct recommendations, 11% of the farmers indicated that the Chatbot provided them with correct recommendations but not fit for their real needs and 6% indicated that the Chatbot provided them with wrong recommendations.

5.3 Analyzing the Wrong recommendations

It is possible to see that the system fails to reply the user with a correct answer when tobacco farmers or Agricultural field officers gives a subject that have various meanings because the Chatbot suggests a reply that was not what the farmer or Agricultural was looking for. Another dreadful scenario occurs when the Chatbot does not apprehend what kind of language the farmer or Agricultural Officer was using for communication: it occurs, for example, when it is not indistinct if the farmer or officer was using English or Shona or Ndebele in communication. In the scenario of correct recommendations, but not fit for the needs of the farmer or Agricultural Officer, the main tricky is the identification of the real user needs, for example when the farmer ask a question about a subject on a specific pest or disease but the Chatbot fails to pinpoint the exact one. From the point of view of the Chatbot usability, a questionnaire about the users' interaction with the Chatbot has been submitted to each farmer: generally, they find the Chatbot easy to use and user friendly. Comparing it with other Chatbots, both the tobacco farmers and Agricultural field officers says that the Chatbot was more modest and effective in assisting with relevant recommendations.

The results of the study on the use of a natural language processing Chatbot for Virginia tobacco farmers are promising. The Chatbot was able to provide farmers with accurate and reliable information, help them troubleshoot problems, and collect data that shall be used to improve the Chatbot's responses and to better understand the challenges facing tobacco farmers.

The Chatbot was also able to improve communication between farmers and extension agents. Farmers were able to get the information and assistance they needed more quickly and easily, and extension agents were able to better understand the challenges facing farmers and develop more effective programs.

5.4 Recommendations and future work

Here are some specific examples of how a natural language processing Chatbot shall be used to help Virginia tobacco farmers:

- The Chatbot shall be used to answer farmers' questions about best practices for growing tobacco. For example, the Chatbot shall provide information about the best time to plant tobacco, the amount of fertilizer to apply, and the best way to control pests.
- The Chatbot shall be used to provide farmers with recommendations for pest management. For example, the Chatbot shall recommend specific pesticides to use and the best way to apply them.

- The Chatbot shall be used to help farmers troubleshoot problems with their crops. For example, the Chatbot shall help farmers to identify the cause of a disease and to recommend a course of treatment.

- The Chatbot shall be used to collect data from farmers about their crops. This data shall then be used to improve the Chatbot's responses and to better understand the challenges facing tobacco farmers.

- The Chatbot shall be used to improve communication between farmers and extension agents. For example, the Chatbot shall be used to collect questions from farmers and to deliver those questions to extension agents. This shall help to ensure that farmers are getting the information and assistance they need.

A natural language processing Chatbot for Virginia tobacco farmers shall be a valuable tool for providing farmers with information and assistance. The Chatbot shall help farmers to improve their yields, reduce their costs, and improve their communication with extension agents.

REFERENCES

[1] A framework for Chatbot evaluation. 2023. Retrieved from https://isquared.wordpress.com/2017/01/24/aframework-for-chatbot-evaluation/

[2] Abdul-Kader, S., & Woods, J. 2015. Survey on Chatbot Design Techniques in Speech Conversation Systems. International Journal of Advanced Computer Science and Applications, 6(7). http://doi.org/10.14569/ijacsa.2015.060712

[3] Abu Shawar, B. and Atwell, E. 2002. A Comparison Between Alice and Elizabeth Chatbot Systems', School of Computing Research Report Series. Available at:http://eprints.whiterose.ac.uk/81930/1/AComparisonBetweenAliceElizabeth.pdf. [4] Abu Shawar, B. and Atwell, E. 2007.Chatbots: are they really useful?, LDV-Forum

[5] Bafna, S., Kakade, A. and Upasani, D. 2021. International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 06 , June

[6] Casillo, M., Cerullo, L., Colace, F., Lemma, S., Lombardi, M. and Pietrosanto, A. 2016. An adaptive context aware app for the tourism," The 3rd Multidisciplinary International Social Networks Conference (MISNC)

[7] Casillo, M., Colace, F., Santo, M. D., Lemma, S., Lombardi, M., Pietrosanto, A. 2016. An Ontological Approach to Digital Storytelling, The 3rd Multidisciplinary International Social Networks Conference (MISNC),

[8] Colace, F., De Santo, M., Lemma, S. , Lombardi, M., Pascale, F. 2017. BotWheels: a Petri Net based Chatbot for Recommending Tires, in Proc. the 3rd edition of Special Session on Knowledge Discovery meets Information Systems: Applications of Big Data Analytics and BI - methodologies, techniques and tools (KomIS), within the 6th International Conference on Data Science, Technology and Applications (DATA),

[9] Colace, F., Santo, M. D., Greco, L. and Napoletano, P. 2015. Weighted word pairs for query expansion, Information Processing and Management, vol. 51, no. 1, pp. 179-193,

[10] Colace, F., Santo, M. D., Marco, L., Francesco, P. and Antonio, P. S. 2018. Chatbot for E-Learning: A Case of Study, International Journal of Mechanical Engineering and Robotics Research

[11] Colace, F., Santo, M.D., Greco, L. and Napoletano, P. 2015. Improving relevance feedback-based query expansion by the use of a weighted word pairs approach, Journal of the Association for Information Science and Technology, vol. 66, no. 11, pp. 2223- 2234,

[12] Cui, L., Huang, S.,Wei, F., Tan, C., Duan, C. and Zhou, M. 2017. Superagent: A customer service chatbot for e-commerce websites. In Proceedings of the ACL 2017, System Demonstrations, Vancouver, BC, Canada, 30 July–4 August; pp. 97–102.

[13] Hollander, J., Sabatini, J. and Graesser, A. 2022. How Item and Learner Characteristics Matter in Intelligent Tutoring Systems Data. In Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks,

Practitioners' and Doctoral Consortium: 23rd International Conference, AIED 2022, Durham, UK, 27–31 July 2022, Proceedings, Part II; Springer International Publishing: Cham, Switzerland,; pp. 520–523.

[14] Liao, M.G., Chang, C.H. and Lin, H.M. 2022. An analysis of children 'interaction with an AI chatbot and its impact on their interest in reading. Comput. Educ., 189, 104576. [CrossRef]

[15] Liu, C.C., Huang, A.Y.Q and Yang, S.J.H. 2023. A Review of AI-Driven Conversational Chatbots Implementation Methodologies and Challenges (1999–2022) Sustainability, 15, 4012. https://doi.org/ 10.3390/su15054012 Academic Editors: Andreas Kanavos

[16] Muzurura , O., Mzikamwi, T., Rebanowako, T.G and Mpini. D. 2023. Application of Artificial Intelligence for Virtual Teaching Assistance (Case Study: Introduction to

Information Technology) International Research Journal of Engineering and Technology (IRJET), Vol. 10 Issue: 09, Sep, pp., 276283

[17] Pawlik, L., Płaza, M., Deniziak, S. and Boksa, E. 2022. A method for improving bot effectiveness by recognising implicit customer intent in contact center conversations. Speech Commun. 143, 33–45. [CrossRef]

[18] Yan, M., Castro, P., Cheng, P. and Ishakian, V. 2016. Building a chatbot with server less Computing, in Proc. the 1st International Workshop on Mashups of Things and APIs Article No. 5.