

Design and Implementation of a Rule Based Afaan Oromoo Conversational Chatbots

Fikadu Wayesa

Wachemo University, School of Computing, Information Technology Department, Ethiopia

ABSTRACT

Chatbot technologies are very popular one in the field of NLP and AI to process basic expressions of interaction between humans (end users) and devices as Question Answering System (QAS). They are used in many ranges like in industry to provide customer services based on their need to reach on their desires, transportation, health, entertainment, education, psychology to monitor human emotions and in many. Messaging service is a popular means of communication as text based messaging (SMS) or through messaging applications. Chatbots are virtual assistant technologies that are designed and developed based on the user's query and to their language. The aim of this paper is to develop and implement a rule based conversational chatbot for Afaan Oromo Language. The approach gets a simple Afaan Oromoo query as input and retrieves answers from a set of predefined rules or responses. The performance of our model is evaluated through human judgment.

Keywords: Afaan Oromoo, Chatbot, NLP, Conversation, Rule-Base

I. INTRODUCTION

In the digital era abundant amount of end user services interactions are helping us to solve our daily activities. In the field of Artificial Intelligence (AI), Chatbots are technologies designed for making conversations to deliver answers quickly to user's need in short amount of time as human-human interaction to assist them to accomplish a task (Martin 2018).

These systems are used in everywhere to engage customer service such as deliver flight information, customer support or in psychological counseling. It assists the user by asking series questions and guide them to what the user is looking for.

Basically, there are three forms of Chatbots: Rule-based (fixed rule) chatbot, Self-Learning and Generative-based chatbots. Self-learning is also called AI-Based Chatbots (Martin 2018). We can combine these forms of chatbots and we can implement an approach called Hybrid.

The rule-based methods depend on a list of simple predefined queries and possible resultant answers. It does not need any machine learning approach and language processing is not mandatory. They are intended for simple queries and they may fail for more complicated questions since they can't produce their own responses. This form of chatbots has one-to-one input and replies. So a bot would be trained according to the defined rules (Soufyane Ayanouz April 2020).

Self-Learning chatbots depend on machine learning and word vectorization. Words are vectorized, because machine learning methods use numbers for prediction. These bots follow the machine learning approach which is rather more efficient and is further divided into two more categories Retrieval-Based Models (corpus-based) and Generative Models (O'Reilly 2017).

Retrieval-based chatbot uses empirical to select a response from predefined responses. It can handle context of the query and selects the best answer from predefined answer.

Generative-based chatbots are the most complex compared to the two previous approaches. It also called cognitive based chatbots. It requires a much bigger data set and use deep learning methods for training. They have cognitive ability to generate different answers (O'Reilly 2017)

II. RELATED WORKS

ELIZA chatbot dialog system was developed to stimulate a Rogerian psychologist, based on a branch of clinical psychology whose methods involve drawing the patient out by reflecting patient's statements back at them based on the predefined rule (Weizenbaum 1996).

(D'ÁVILA 2018) Designates a rule based chatbot development for long conversations using KINO dataset. The goal of the paper was to match between a response candidates in a given conversational speeches. They proposed a framework in which the possible to verify that the positively associated words were part of the information in which the bot had been programmed to search in its knowledge base. A manual classification was necessary for the words negatively related to understand and identify the

types of issues in the identification of the user's intentions.

(Jagdish Singh 2019) Proposed a Rule-based chatbot for Asian Pacific University (APU) students to answer their enquiries from their administrative offices. The designed chatbot permits end users to observe information and interact with the chatbot regarding their enquiries. The developed system was compatible with bot's app page on Facebook for access via laptop or computer devices.

There is no work attempted work for Afaan Oromo Language in the area of chatbot system. This language is rich in its structure, grammar and semantics.

Here, this research work focused on the following questions.

1. What are the possible methodologies used in the ChatBot implementation system?
2. What challenges are there in the developing rule based chatbot?

The Model Architecture

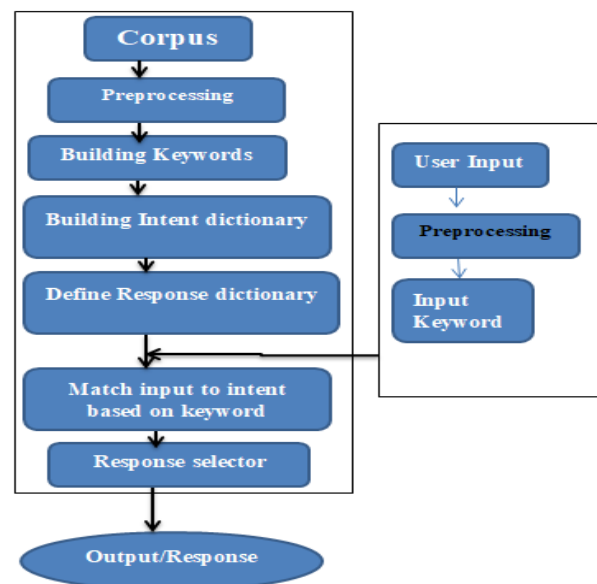


Figure 1 Proposed System Architecture

III. Experimental Results

A. Components of the Model

Corpus

We have collected a lists of query with their possible response that are used in Afaan Oromo Conversational System.

Preprocessing

Data should be preprocessed first in order to design any application, the first thing we need is data, to attain a particular format to make machine understandable. To get the possible answer for the query a number of NLP tasks have been applied to our corpus and to the query. Generally, the following tasks have been applied.

1. The corpus and the query have converted to lower cases. Unnecessary special characters like punctuation marks have been removed.
2. Tokenization has been applied to both the corpus and user query.

Building Keywords

Once the corpus has cleaned, list of keywords that our model will look for is built.

Building Intent Keywords

After the keywords list is completed, a dictionary that matches the user's query keywords has constructed to the required intents. We reformat the keywords in a special syntax that helps to get the required answer. We tried to design a common intent for a conversation that use common approach.

Defining Response Dictionary

We defined responses for each intent type as a key.

```
QUERY={ 'nagaa': [Akkam? Nagaa
ooltee? Fayyaadhaa? Akkam bulte?'],
'yeroo': [yeroon meeqa?, sahatiin
meeqa? sa'atii meeqatti dhufu?'],
'Walii gala': [Maqaan kee eenyu?,
Eenyutu si uume?, Na jaallattaa?'],
}
```

Match Input to intent based on keyword and Response Selector

An input from the user is matched with the keywords dictionary using RegEx Search function.

The current intent gets selected if the match is found and the key to the responses has selected.

```
const reply = [{"nagaa!", "fayyaa!", "galata
waaqayyoo!", "Galata
rabbii!", "jirra", "nagumaa"}, [ "sa'a 6", "sa'a
booda", "boru" ]]
```

Figure 2. A Sample response

```
function Chatbot(user input) returns response
Find the word i in input that has the intent

if i exists
    Choose the response that matches input
    response Apply the transform in r to input
else (no keyword applies)
    response Apply the transform for the NONE keyword to input

return(response)
```

Figure 3. Rule based Algorithm

B. The Model Results

Chatbot system has significant applications in transforming the human being's lives. For the evaluation purpose, we collected corpus during the experiment phase, the Afaan Oromo language conversations of users of the total of 114 conversations and 30 users, 20 conversations were selected for the evaluation purpose. The model performs a good result by imitating a character user and system from the sample conversations.

Right responses are given to user to to make a small conversation. The work is analysed by human evaluators, ten (10) evaluators were selected to

evaluate the performance of the model and the average score being evaluated was 88%.

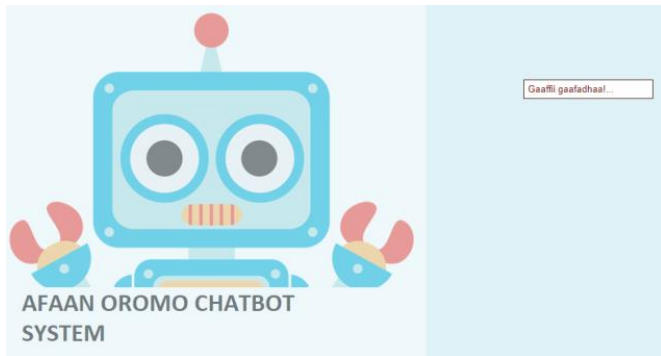


Figure 4. Afaan Oromo Rule Based Chatbot system User Interface

The above figure denotes the physical interface where a user asks a system to get his /her request.

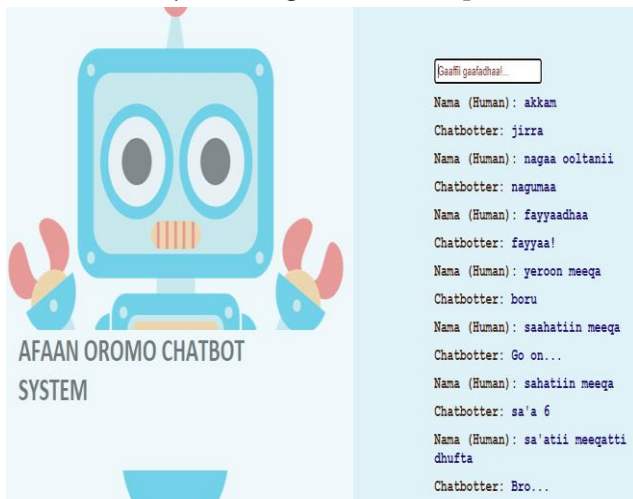


Figure 5 A Sample Conversations

A user is denoted by Nama (Human) and Here the user is denoted by “Human” and the chatbot is represented by “Chatbotter”.

IV. CONCLUSION AND FUTURE WORK

In any language, chatbots are user-friendly for any user who has an awareness of typing in the language they can understand on the desktop application and in the mobile. In this work, we designed a simple rule based Chatbot for Afaan Oromo Language. We only tested our model with 20 intents in this work for the

evaluation purpose. From our result we saw designing and implementing a rule-based chatbot is a very challenging task since a rule based chatbot requires more intents.

In addition, the chatbot would severely be limited in terms of its conversational capabilities as it is near impossible to describe exactly how a user will interact with the bot. since Afaan Oromo has very complex structure, defining all rule was very challenging one. Another challenge was grammar error cannot be recognized.

We can make improvements to the developed chatbot to further enhance its affectivity since rule based has many limitations. First, AI-based Chatbots are a more practical solution for solving the limitations of rule based scenarios. So, we’ll build an AI-based Chatbot for this language.

V. REFERENCES

- [1]. D’ÁVILA, THIAGO CARVALHO. KINO: AN APPROACH FOR RULE-BASED CHATBOT DEVELOPMENT, MONITORING ANDEVALUATION. 2018.
- [2]. Jagdish Singh, Minnu Helen Joesph and Khurshid Begum Abdul Jabbar. "Rule-based chabot for student enquiries." International conference on computer vision and machine learning, 2019.
- [3]. Martin, Daniel Jurafsky & James H. Speech and Language Processing. 2018.
- [4]. O’Reilly, Amir Shevat. "Designing Bots." 2017.
- [5]. Soufyane Ayanouz, Soufyane Ayanouz, Mohammed Benhmed. "A Smart Chatbot Architecture based NLP and Machine learning for health care assistance." Research Gate, April 2020.
- [6]. Weizenbaum, J. "ELIZA – A computer program forthe study of natural language communication between manand machine." Communications of the ACM, 1996: 36–45.