**Yoofi Brown-Pobee 49142019**

**Chapter 1: Introduction**

Many businesses in Ghana make use of customer service call centers to respond to the queries of customers and resolve any issues they may have. A simple search on-line would produce .a list of jobs available said call centers. In principle, these centers are very instrumental to the business operations as they serve as a contact point to customers who are revenue drivers for the business. However, practicality dictates otherwise. In some cases, the numbers provided for customer service are no longer in service (some have been changed but not updated), in other cases the call experience is marred by the emotional state of calling customers or the receiving agents at a given point in time. Even beyond these issues exist a larger problem of cost. For companies, as they scale up and acquire more customers, there is a need to hire more customer service personnel to cater to the needs of customers.

An already existing solution to this problem is the use of automated conversational systems to respond to the queries of customers. Intelligent systems exist that can quickly serve numerous customers at any given time before deciding if human intervention is needed. An example of a working solution is MTN’s Facebook chatbot. Querying this chatbot exposed some flaws as its responses as it could not properly understand what was being asked hence disbursing unsuitable information for the given domain. A domain in this case is simply the scope of questions, statements and classes that all communication can be categorized under.

This projects main contribution is to create an intelligent chatbot that can give suitable responses to a wide range of customer queries in a given domain. Chalkboard Education, a startup that hosts an education platform would be used as the host company. Chalkboard education allows schools to make all their content digital and enrolls their students unto a platform to access it. This results in many schools and individual students constantly calling and direct messaging Chalkboard’s customer service support line for help. The company currently have only one person available for this job. By the end of this project, the goal is to build full-functioning chat bot to handle all of Chalkboard’s support issues.

**Background & Motivation**

The use of conversational agents and dialog systems can be traced as far back as 1966 with the development of ELIZA, an early natural language processing computer program to present day Siri and Alexa of Apple and Amazon respectively. Natural Language Processing and its related fields have provided a means for us humans to communicate which machines and machine systems in ways similar to how we do with each other. There are many systems with differing implementations and nuances but majority generally fall in two classes: Task oriented dialog agents and Chatbots .

The former is predominantly concerned with specific tasks and designed to have as very little interactions as possible to accomplish the task [Jurafsky and Martin, 2014]. Examples include digital assistants like the aforementioned Siri and Alexa. This class of dialogue systems are mostly employed by companies on their websites and products to enable customers and users address problems and answer questions. Key benefits of these systems especially to businesses and their customers include faster and more convenient query handling as well as lower costs relative to hiring of human agents. They are not designed to have prolonged conversations, however, as chatbots, the second class of dialogue systems, are designed for that.

Chatbots are set up to mimic the natural conversational characteristics of human beings. They are designed to have more casual and less directed conversations than task oriented dialog agents. Cleverbot is a chatbot capable of carrying on prolonged conversations with humans and exists mainly for entertainment purposes [Carpenter, n.d.]. These are more suited for social uses such as for psychological analysis and entertainment purposes. A form of evaluation for many chatbots is a test called the Turing test developed by Alan Turing in 1950. A Turing Test is a method of inquiry for determining whether or not a computer/system is capable of thinking like a human being[2]. A human should be able to find the two systems indistinguishable in order to pass the Turing test. A chatbot that is able to pass this test can be said to be a good one.

The existence of the above systems in the field of natural language processing have made it possible for the application of knowledge to different scenarios and domains. Chalkboard Education’s particular situation presents an opportunity to apply this knowledge in a local context. The startup has been running for close to three years and current has over 4000 students enrolled on its platfrom. Customer queries are handled mainly through WhatsApp and voice phone calls and require employees to be present to respond at all times. The creation of a dialogue system on Chalkboard specific information and queries appears a potential solution to solve the problem by removing the need for the physical presence of a human and saving cost for the startup. As the startup scales, the system does not have to scale proportionally hence cost of customer care can be minimal for a given number of users.

Related Work

ELIZA

ELIZA is an early natural language processing computer program created from 1964 to 1966 [Ireland,

2012]. It was created to demonstrate communication between humans and machines using natural

language and was created by Joseph Weizenbaum at the MIT Artificial Intelligence Laboratory. It

works by breaking down a sentence and ranking its keywords and transforming the users sentence into

a suitable response using rules from its preprogrammed learning script. The system was modelled after

a Rogerian psychotherapist in that it allowed the program to not need to have knowledge base of the

topic the human was discussing [Weizenbaum, 1966]. By simply transforming the sentences based on

key words in the users sentence, the program can appear to be engaging in conversation regardless of

the topic. There were a number of reports of human test subjects developing an emotional connection

to the system; a testament of how well it was able to sustain a regular conversation. This system formed

the basis for the creation of other conversational agents and improvements in the field of natural

language processing.

Cleverbot

Cleverbot is a web based chat bot that uses artificial intelligence to have unsupervised conversations

with humans and was created by Rollo Carpenter. It holds the distinction of having passed the Turing

test. As mentioned earlier, the idea of the test is for a machine to pretend to be a human and will only

pass if this pretense is found to be convincing [Gehl, 2013]. Cleverbot has performed quite well in

Turing Test competitions, giving it some credibility [Gehl, 2013]. Unlike ELIZA, Cleverbot actually

learns how to have a conversation as it interacts with more humans. It’s responses at any given time are

as a result of an analysis of previous conversations it has had with other humans. This contributes to its

ability to perform well in the Turing Test. Despite the difference in how they are implemented, they are

both capable of holding a conversation with a human for a considerable length of time.

Dialogue System Evaluations

Dialogue System Evaluations are crucial to determining the quality of a dialog system by assessing its

responses against human generated supervised signals such as user satisfaction scores [Walker et al.,

1997]. Dialogue system models which are optimized for supervised objectives are known as supervised

dialogue models and those that are not known as unsupervised dialogue models [Liu et al., n.d.]. The

former requires the large scale collection of labels which can be often expensive while the latter’s

major concern is how to perform the evaluation automatically. N-gram based measures like BLEU exist

but are found to be a poor fit for dialogue models because two different responses may have no overlap

but would be good replies to a given statement [Kannan and Vinyals, 2017]. There exist many models

to evaluation dialogue models like ROUGE and adversarial models to suggest the quality of a dialogue

system.

References

[1] Daniel Jurafsky and James H. Martin. 2014. *Speech and language processing*, Harlow: Pearson.

Carpenter. Cleverbot. Retrieved October 9, 2018 from <http://www.cleverbot.com/>

[2] Anon. What is Turing test? - Definition from WhatIs.com. Retrieved October 10, 2018 from

<https://searchenterpriseai.techtarget.com/definition/Turing-test>

[3] Corydon Ireland. 2012. Alan Turing at 100. (September 2012). Retrieved October 10, 2018 from

<https://news.harvard.edu/gazette/story/2012/09/alan-turing-at-100/>

[4]Robert Gehl. Teaching to the Turing Test with Cleverbot. *The Journal of Inclusive Scholarship and*

*Pedagogy*24, 1-2.

Anjuli Kannan and Oriol Vinyals. 2017. Adversarial Evaluation of Dialogue Models. (2017).