Samrat Sahoo

Mr. Speice

Independent Study Mentorship 3A

2 December 2020

<div align="center">Analyzing Eliza: A Retrieval Chatbot</div>

<div align="center">Assessment 9 - Research Assessment</div>

Date: 2 December 2020

Subject: Retrieval Language Generation Algorithms

Works Cited:

Brainerd, Wade. "Wadetb/Eliza." GitHub, Github , 8 Feb. 2018, github.com/wadetb/eliza.

Assessment:

Language generation models can often be split into two categories: retrieval-based and data-based. Each model type has a set of usages, advantages, and disadvantages and it is important to understand how each model functions in order to evaluate the best model for your uses. Data-driven models rely on large amounts of conversational data and have responses that are far more versatile whereas retrieval-based models are far more specialized and do not require large amounts of data. In order to evaluate the effectiveness of a retrieval model in a mental health scene, it is important to analyze and understand the algorithm behind one of the most famous retrieval mental health chatbots: Eliza.

To start, the chatbot uses a decomposition algorithm. This is important to note because it helps understand the effectiveness of the model, more specifically how well it can understand different types of sentences. Since the decomposition model uses a weightage on each word and sorts these weights, it can be concluded that the Eliza algorithm is especially effective in identifying keywords and generating responses based on the responses. Additionally, the decomposition model has a feature where it can generate different responses based on the type of

subject that is being talked about. This is important to note because it allows the bot to be more versatile in its nature by being able to respond to different topics in a more appropriate fashion, overcoming one of the main barriers of retrieval chatbots: the rigid nature. This also ensures that the bot holds with one of the fundamental principles of natural language processing: the fluidity of human language.

Continuing through the algorithm, the next part entailed a reassembly algorithm. This algorithm used one of the most important features of computer science: the idea of randomness. Using randomness is instrumental in providing a more human-like conversation, one of the primary goals of natural language processing bots. Additionally, this factor of randomness will make the user experience of the person receiving therapy a far more effective experience because they will feel like that this is not just a bot that repeats the same responses over and over again. Additionally, the reassembly algorithm takes into account the keywords in the sentence and generates a new sentence based on the keywords. This is fundamental to the algorithm because it ensures that the conversation stays on track and therefore is able to feel more human-like. In essence, through reassembly algorithms, the bot is able to produce a more natural-sounding conversational, an incumbent part of artificial intelligence-based therapy.

To conclude, a large part of human language is the idea of keeping consistent and accurate grammar so that the user can actually understand what is happening with the text. To do this, regular expression manipulation allows this. This is important to note because regular expression manipulation can not only be used in this algorithm but in other parts of all-natural language processing projects to attain the ideal language algorithms throughout projects. In synthesis, through the decomposition algorithm, reassembly algorithm, and regular expressions, a human-like and effective therapy chatbot can be produced.

Annotations