

Mr. Speice

Independent Study Mentorship 3A

5 November 2020

Implementations and Text Generation

Assessment 7 - Interview Assessment

Name of Professional: Abhiramon Rajasekharan

Title: Graduate Student

Company: University of Texas at Dallas

Date of Interview: 5 November 2020

Works Cited:

Abhiramon Rajasekharan, Informational Interview. 5 Nov. 2020

Assessment:

Deep learning has become so ubiquitous in the field of machine learning and related fields that it has nearly ruled out some of the more traditional natural language processing techniques. One of the largest developments in natural language processing is the idea of deep learning-based text generation with transformer networks. Additionally, other developments related to deep learning include embeddings like BERT which is able to understand the text on a sentence-level. During an informational interview with Mr. Abhiramon Rajasekharan, I was given the opportunity to take a deeper look into why deep learning has become so common within natural language processing.

We initiated the conversation by discussing an education pathway that is less common than the traditional software developer pathway: the pathway of a graduate student. The rewarding and challenging aspects of being a graduate student, particularly one in natural language processing are insightful in the sense that it was a new perspective that is usually not

mentioned: the perspective of having a rewarding experience because of the potential of natural language processing. This serves to be beneficial for the future because it delivers the idea that natural language processing is a field of possibilities and research is a potential career field that should be considered. This has the potential to drastically change educational plans within the future and therefore has been instrumental within the interview process.

The conversation continued with the discussion of different types of embeddings for classification networks and how embeddings have progressed. Through this discussion, there was a valuable obstacle of natural language processing that was overcome: the obstacle of context. Older embeddings like Char2Vec and Word2Vec do not have the potential to understand sentences with context however embeddings like Doc2Vec and BERT do. This is not only helpful in recognizing the embeddings that should be used in any future natural language processing projects but also helps to identify the underlying reasons for why the older embeddings perform at a lower success rate which is instrumental during the presentation process when explaining embeddings. Because of this conversation, the original work will be further refined because both an obstacle will be overcome and the type of embeddings will enable for the deep learning classification models to perform at higher accuracies.

The conversation concluded with the discussion of deep-learning-based generative models. The first stream of discussion was how to get data. From this conversation, one of the largest challenges was overcome: obtaining conversational therapy data. Because of this conversation, it is now known that sources like Reddit can be scraped to get the correct type of data needed for transformer networks. From this, the data collection process for the second part of the project has become much easier and now there is additional knowledge on how transformer networks work will enable implementations of the networks to become more robust.

In conclusion, the topics discussed and new information learned will be invaluable going forward. This conversation served as an excellent stepping point to bridging the gap in knowledge with text generation, expanding the information in text classification, and exploring a new potential career pathway. It goes without saying that the information gained from Mr. Rajasekharan will aid in the further development of the Twitter Bots For Social Good project.

[Interview Notes](#)