Samrat Sahoo

Mr. Speice

Independent Study Mentorship 3A

3 September 2020

An Introduction to Natural Language Processing

Assessment 1 - Research Assessment

Date: 3 September 2020

Subject: Natural Language Processing Introduction

Works Cited:

Torfi, Amirsina, et al. "Natural Language Processing Advancements By Deep Learning: A

Survey." ArXiv:2003.01200 [Cs], June 2020. arXiv.org, http://arxiv.org/abs/2003.01200

Assessment:

Computers best understand numerical data. Combining computer science with language, an aspect of human culture that cannot easily be quantified into numbers, is one of the most difficult challenges computer scientists around the world face today. Despite the challenges, humanity has persevered and made great innovations with deep learning in the world of natural language processing: the way computers understand human language. Combining the power of computers and the complexity of human language, we have been able to generate text, analyze sentiments, categorize documents, and perform many other tasks.

One of the various ideas explored is applying image processing techniques, such as convolutional neural networks or generative adversarial networks, to text. In the past, image processing techniques have been extremely successful with some image classification algorithms reaching accuracies greater than human accuracies. Reaching results similar to these could prove to result in a breakthrough in natural language processing which would be unimaginably

beneficial. Additionally, there are many resources available to continue further research within image processing fields and their applications to natural language processing, making future research and development a smoother process. Applying image processing techniques to natural language definitely seems to be an important first step to creating a successful product.

Natural language processing has many applications including sentiment analysis, text generation, text classification, and information extraction. These four categories of natural language processing seem to be the most prevalent and so it is important to keep potential applications of these in mind when brainstorming a product idea. In addition to the primary categories, current advancements in natural language processing allow for tasks such as --but not limited to-- part of speech tagging and keyword extraction. While there have been other advancements in natural language processing, many require intensive computational resources, something that is not ideal for consumer-based products. Knowing the limitations and abilities of current natural language algorithms will be helpful in order to determine if a project is feasible for consumer use while also being useful for society.

Finally, rigid natural language techniques such as Rule-based Question Answering were explored. While on the surface, these ideas may be easy to implement, they overlook one significant aspect of the English language: fluidity. The English language is ever-changing, has many forms, and can often be ambiguous. However, with Rule-based systems, this fluidity does not hold. From this, it should be noted that any projects should not have any major rule systems. Additionally, any artificial intelligence created should be able to adapt to changing ideas presented by humans. Having insight into the lack of versatility of rule-based systems is insightful and helps avoid poor practices during the development phase.

In synthesis, by analyzing possible methods, applications, and practices of natural language processing, it is clear that the next steps should involve applying prior technology in image processing while avoiding rigid systems and keeping in mind potential applications to develop an ideal product. Additionally, through this research, there are new architectures and technologies presented that could smoothen the research or development processes for the product. Collectively speaking, the analysis of all aspects of natural language processing has paved the path for new ideas, further research, and a more straightforward experience.

Annotations