

Research Assessment

Kavan Mehta

Ms. Whitcomb

ISM 2

13 November 2022

Kavan Mehta

Ms. Whitcomb

ISM 2

13 November 2022

Research Assessment #11 (Original Work Related)

Date: 13 November 2022

Subject: Transformers Applications in Image Recognition

MLA citation(s):

Dosovitskiy, Alexey, et al. "An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale." *ArXiv.org*, ArXiv, 3 June 2021, <https://doi.org/10.48550/arXiv.2010.11929>.

Assessment:

I have started to create my Original Work projects in ISM 2 with an understanding of basic algorithms and the foundation of natural language processing (NLP). Hence, I decided to focus on the applications of transformers, an efficient NLP algorithm that has the power to understand relationships between sequences in text. To learn more about the applications of transformers, I found an academic journal article, "An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale," which went over the concept of utilizing transformers for image recognition with more efficiently pre-trained models that can use word sequencing methods that provide lower computational cost than traditional state-of-the-art Convolutional Neural Networks.

Through the academic journal article, I first learned about the use of natural language processing transformers with self-attention mechanisms into computer vision with functions such

as image processing, neural analysis, and object detection. Moreover, they “split images into patches and use them in sequences” in order to analyze the pixels of each image as tokens of a text sequence (Dosovitskiy et al 1). This is something that is primarily used by Convolutional Neural Networks for the best processing of images. This helps connect to my understanding from last year about Convolutional Neural Networks (CNNs) as they are generally used for image processing rather than other efficient models. Furthermore, I believe that the capability of transformers in image processing as shown in the journal article could also serve as one of the points that I could use to improve the accuracy and efficiency of my final product from ISM 1. This helps directly connect to my goals in ISM 2 to work on NLP transformers and apply them to different fields such as computer vision. I think this has helped me clarify my understanding and I am excited to understand how this works on a deeper level in terms of the calculus that is used in transformers! I was able to grasp a deeper understanding of other NLP models too that could possibly work on computer vision! With their possible uses of transformers and other deployment services for understanding natural language, I will explore possible options with them to create a fully functional deployed product in ISM effectively.

With this step in understanding the applications of transformers, I am one step closer to implementing transformers in a new environment in my Original Work. I have already started to learn the Hugging Face library and the many functions that I could utilize in my projects. As I use my foundational understanding of transformers and other natural language processing algorithms, I hope to continue to develop projects in my Original Work with NLP technology. I also want to continue to learn more about the algorithms as a whole on a mathematical level to get more intuition about the mechanism process through these scholarly articles and research interviews with professionals in the field. Moreover, with a strong theoretical foundation, I will

shift my focus on the implementation of these algorithms in ISM to create my Original Work that could help the community.