Research Proposal

Nirbhay P. Tandon
968675
Email: N.P.Tandon@2021.ljmu.ac.uk
Project Supervisor: Mr. Ankit Jha

Abstract

Attention based Transformer architectures have become the norm of modern day Natural Language Processing. Google began this trend back in 2017 with their paper *Attention Is All You Need*[1], by introducing the Transformer architecture that works solely on attention mechanisms. The purpose of our work will be to explore a new kind of Transformer architecture. Compare & contrast its performance against the SQuAD 2.0 Dataset[2] based other architectures such as BERT[3], RoBERTa[4], ALBERT.

Contents

1 Introduction

2 Background & Related Research

2.1 Background

2.2 Related Research

3 Research Questions

Highlighted below are some of the questions that we will answer through our research.

4 Aims & Objectives

Through our research we aim to establish the efficiency of our new Transformer architecture. We shall implement the existing models that are available via libraries such as HuggingFace, PyTorch & Tensorflow, run the SQuAD 2.0[2], to obtain benchmark scores & then compare the results with our proposed architecture. We hope to establish our proposed transformer architecture as a competent enough contender to be used within both industry & academia.

5 Research Methodology

6 Expected Outcomes

7 Requirements & Resources

8 Research Plan

References

[1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, Aidan N, Kaiser, L. and Polosukhin, I. (2017). Attention Is All You Need. [online] arXiv.org. Available at: https://arxiv.org/abs/1706.03762.

[2] Rajpurkar, P., Jia, R. and Liang, P., 2018. Know what you don't know: Unanswerable questions for SQuAD. arXiv preprint arXiv:1806.03822.

[3] Devlin, J., Chang, M.W., Lee, K. and Toutanova, K., 2018. Bert: Pretraining of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

[4] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. and Stoyanov, V., 2019. Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.