

Develop New Transformer Architecture For Question and Answering(QandA)

Nirbhay P. Tandon

June, 2021

Contents

List of Figures	3
List of Tables	4
1 Introduction	5
1.1 Background Of The Study	5
1.2 Aims And Objectives	5
1.3 Scope Of The Study	5
1.4 Significance Of The Study	5
1.5 Structure Of The Study	5
2 Literature Review	6
2.1 Question Answering Using Neural Nets	6
2.2 Question Answering Using LSTMs	6
2.3 Question Answering Using Transformers	6
2.4 Comparison Of Techniques	7
2.5 Summary	7
3 Research Methodology	8
3.1 Data Selection	8
3.2 Data Pre-processing And Transformation	8
3.3 Existing Models And Benchmarks	8
4 Architecture Creation	9
4.1 Drawbacks Of Current Architectures	9
4.2 Proposed Architecture Improvements	9
4.3 Architecture Refinement	9

Appendices	10
Bibliography	12

List of Figures

List of Tables

Chapter 1

Introduction

?<c1)? **1.1 Background Of The Study**

?<11)? **1.2 Aims And Objectives**

?<12)? **1.3 Scope Of The Study**

?<13)? **1.4 Significance Of The Study**

?<14)? **1.5 Structure Of The Study**

?<15)?

Chapter 2

Literature Review

?<c2>? 2.1 Question Answering Using Neural Nets

?<21>? 2.2 Question Answering Using LSTMs

- ?<22>?
1. Hochreiter et al. (2001)
 2. Schmidhuber and Hochreiter (1997)
 3. Schmidhuber (2015)
 4. Tan et al. (2015)
 5. Wang and Jiang (2016)
 6. Gennaro et al. (2020)
 7. Weissenborn et al. (2017)

2.3 Question Answering Using Transformers

- ?<23>?
1. Vaswani et al. (2017)
 2. Devlin et al. (2018)
 3. Lan et al. (2019)
 4. Liu et al. (2019)

5. Sanh et al. (2019)

6. Rajpurkar

7.

8.

2.4 Comparison Of Techniques

?⁽²⁴⁾? 2.5 Summary

?⁽²⁵⁾? (Corsair, 2021)

Chapter 3

Research Methodology

?<c3)? DONT DO ANY ANALYSIS HERE

3.1 Data Selection

?<c31)? 3.2 Data Pre-processing And Transformation

?<c32)? 3.3 Existing Models And Benchmarks

?<c33)?

Chapter 4

Architecture Creation

?<c4>? 4.1 Drawbacks Of Current Architectures

?<c41>? 4.2 Proposed Architecture Improvements

?<c42>? 4.3 Architecture Refinement

?<c43>?

Appendices

Bibliography

- roberta** Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019.
- albert** Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut. Albert: A lite bert for self-supervised learning of language representations. *arXiv preprint arXiv:1909.11942*, 2019.
- haighextractive** M. Tan, B. Xiang, and B. Zhou. Lstm-based deep learning models for non-factoid answer selection. *CoRR*, abs/1511.04108, 2015. URL <http://arxiv.org/abs/1511.04108>.
- mhu2016question** S. Wang and J. Jiang. Machine comprehension using match-lstm and answer pointer. *CoRR*, abs/1608.07905, 2016. URL <http://arxiv.org/abs/1608.07905>.
- distil** V. Sanh, L. Debut, J. Chaumond, and T. Wolf. Distilbert, a distilled version of BERT: smaller, faster, cheaper and lighter. *CoRR*, abs/1910.01108, 2019. URL <http://arxiv.org/abs/1910.01108>.
- bert** J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- RAM** Corsair. Vengeance® lpx 8gb (1 x 8gb) ddr4 dram 2400mhz c14 memory kit - black, 2021. URL <https://www.corsair.com/uk/en/Categories/Products/Memory/VENGEANCE-LPX/p/CMK8GX4M1A2400C14>. Accessed: 2021-04-16.

atayl	A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need. <i>arXiv preprint arXiv:1706.03762</i> , 2017.
lstmintent	G. D. Gennaro, A. Buonanno, A. D. Girolamo, A. Ospedale, and F. A. N. Palmieri. Intent classification in question-answering using LSTM architectures. <i>CoRR</i> , abs/2001.09330, 2020. URL https://arxiv.org/abs/2001.09330 .
squad	P. Rajpurkar. Squad2.0. URL https://rajpurkar.github.io/SQuAD-explorer/ .
lstm	S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber, et al. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies, 2001.
schmid	J. Schmidhuber. Deep learning in neural networks: An overview. <i>Neural networks</i> , 61:85–117, 2015.
fastQA	D. Weissenborn, G. Wiese, and L. Seiffe. Fastqa: A simple and efficient neural architecture for question answering. <i>CoRR</i> , abs/1703.04816, 2017. URL http://arxiv.org/abs/1703.04816 .
lstmoriginal	J. Schmidhuber and S. Hochreiter. Long short-term memory. <i>Neural Comput</i> , 9(8):1735–1780, 1997.