

Effect of Training Large Language Models For KG Generation on their General Language Abilities: The Cost of Specialization?

by

Sharyar Memon, Navid Rezai, Marek Reformat

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Engineering

Department of Electrical and Computer Engineering

University of Alberta

Abstract

This project focused on understanding the side effects of "specializing" a large language model such as GPT2-XL on the task of Knowledge Graph Generation. It was theorized that this kind of "specialization" may lead to a decrease in the performance of the models in other language modelling and understanding tasks such as multiple choice selection, text entailment, word sense disambiguation and other similar tasks. This comparison was conducted by taking a baseline (non-specialized) GPT2-XL model and comparing its performance with a "specialized" GPT2-XL model used for Knowledge Graph Generation.

Preface

A preface is required if you need to describe how parts of your thesis were published or co-authored, and what your contributions to these sections were. Also mention if you intend to publish parts of your thesis, or have submitted them for publication. It is also required if ethics approval was needed for any part of the thesis.

Otherwise it is optional.

See the FGSR requirements for examples of how this can look.

To the Count

For teaching me everything I need to know about math.

I think there is a world market for maybe five computers.

– Thomas J. Watson, IBM Chairman, 1943.

Acknowledgements

Put any acknowledgements here, such as to your supervisor, and supervisory committee. Remember to list funding bodies, and external scholarships. The acknowledgements can't be more than 2 pages in length.

Acknowledgements are optional, but are recommended by the FGSR.

Contents

1	Introduction	1
1.1	Cross-Referencing	1
2	Conclusion	3
	References	4
	Appendix A Background Material	5

List of Tables

1.1 A sample table 2

List of Figures

A.1 A supporting figure	6
-----------------------------------	---

Glossary

A Sample Acronym (ASA)

A sample acronym description

Glossary Entry

This is a sample glossary entry.

Glossary entry descriptions can span multiple paragraphs.

Remember that glossaries are optional. The glossary implementation in this template is intended to be simple, and makes use of only one package, **glossaries**. There are more flexible, and fully-featured methods for creating glossaries than the one used here.

Chapter 1

Introduction

Here is a test reference [1]. These additional lines have been added just to demonstrate the spacing for the rest of the document. Spacing will differ between the typeset main document, and typeset individual documents, as the commands to change spacing for the body of the thesis are only in the main document.

1.1 Cross-Referencing

Cross-references between child documents are possible using the **zref** package.

Text on a new page, to test top margin size.

A sample equation (1.1) follows:

$$y = \frac{1}{x^2} + 4 \tag{1.1}$$

A sample table, Table 1.1:

Non-wrapping column	Wrapping column
This is an ordinary column	This is a balanced-width column, where text will wrap

Table 1.1: A sample table created using the `tabularray` package

If there are many acronyms, such as A Sample Acronym (ASA), and specialized technical terms, consider adding a glossary. Sample glossary entry, and acroynm (ASA) descriptions are provided above.

Chapter 2

Conclusion

Referring back to the introduction (Section 1.1), we see that cross-references between files are correctly handled when the files are compiled separately, and when the main document is compiled. When the main document is compiled, cross-references are hyperlinked. The values of the cross-references will change between the two compilation scenarios, however. (Each chapter, compiled on its own, becomes “Chapter 1”.)

Caution: For cross-references to work, when files are compiled separately, the referenced file must be compiled at least once before the referring file is compiled.

References

- [1] D. E. Knuth, *The Art of Computer Programming, Volume I: Fundamental Algorithms*. Addison-Wesley, 1968.

Appendix A

Background Material

Material in an appendix.

We plot an equation in figure A.1.

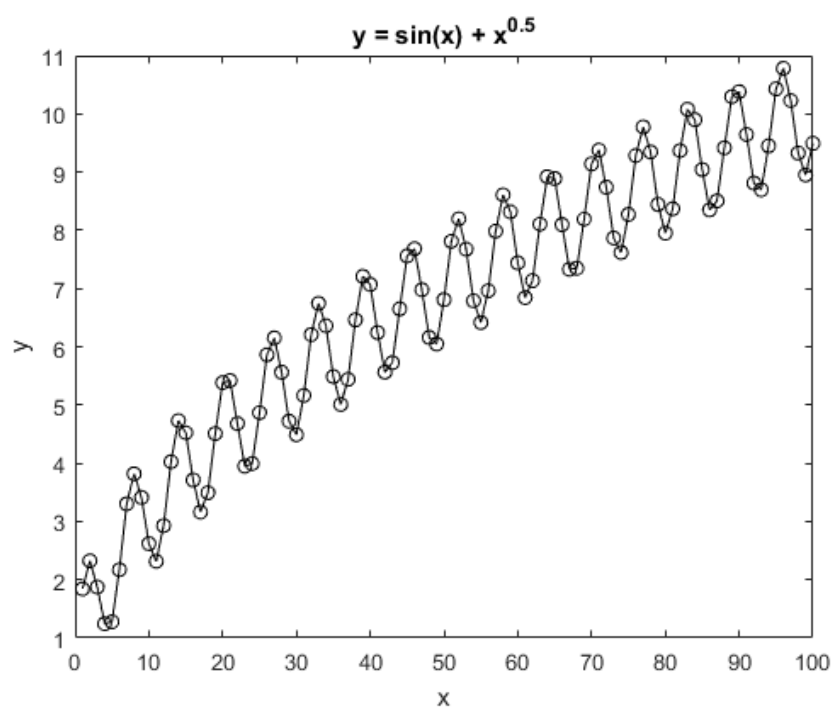


Figure A.1: A graph of $y = \sin(x) + \sqrt{x}$