Using Jointly Learned Covariate Specific Word Embeddings to train a Bidirectional LSTM based Language Model

TODO

Jonathan Magbadelo

Candidate Number: 148708
Supervisor: Dr. Julie Weeds

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

Joe Bloggs

UNIVERSITY OF SUSSEX

JOE BLOGGS, DOCTOR OF PHILOSOPHY

MY THEORY OF EVERYTHING

HOW IT ALL WORKS

SUMMARY

Acknowledgements

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This is the introduction to the thesis.[1] The conclusion is in Chapter 3 on page 3.

1.1   Overview

Songwriting is a form of composition through which songs are created, often drawing from personal experiences. The most important aspects of songwriting are structure and content, with the latter being a key influence on the listeners ability to relate to a song. The main issue songwriters face is word choice through which they can express their ideas concisely and create clear contexts for the listener. The choice of words will either enhance or hinder the ability of the lyrics to effectively impact a listener.

---

[1]And this is a footnote.

Chapter 2

Professional Considerations

Throughout the development of this project professional and ethical considerations will be considered: including those highlighted in the British Computing Societies (BCS) Code of Conduct 4 and Code of Good Practice. The following sections outline relevant areas in the specified documents as well as how this project will adhere to them:

2.1   Code of Conduct

2.2   Good Practices

2.3   Ethical Considerations

The success of this project relies heavily on the availability of correctly labelled lyric data, and as stated before, there is no central repository where lyrics, along with the required metadata for the project, are stored. Consequently, data collection techniques such as web scraping may be used to gather additional data. Though web scraping is common practice, it is important to be a "good citizen" of the web and as such the following will be adhered to: ● robots.txt will be obeyed ● requests to sites will be done politely and if stated in the robots.txt file any request rate limit will be upheld ● the user agent string of the request will be made explicit for the webmaster to see Due to the nature of the project, appropriate user testing/research will be carried out during development. In accordance with section 1.a of the BCS Code of Conduct, any private data will be stored securely and if possible anonymised. The project utilises textual data which may have explicit content within it. After the training of models, which will not be filtered to allow artistic freedom into the model, a filtering option will be implemented in order to prevent potential users from seeing explicit content.

Chapter 3

Conclusion

I was right all along.

3.1　What was I right about?

I was right about the following things.

3.1.1　Previous theories were wrong

People thought they understood, but they didn't.

3.1.2　My new idea is right

Of course.

Bibliography

Einstein, A. (1905). Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig? *Annalen der Physik*, 323:639–641.

Appendix A

Code

```
10 PRINT "HELLO WORLD"
```