A Natural Language Processing Framework for Training a Neural Network Chatbot

Kamila Michel

Maciej Majchrzak

B.Sc.(Hons) in Software Development

March 11, 2019

Final Year Project

Advised by: Gerard Harrison

Department of Computer Science and Applied Physics
Galway-Mayo Institute of Technology (GMIT)

Contents

About this project

Abstract A brief description of what the project is, in about two-hundred and fifty words.

Authors Explain here who the authors are.

Chapter 1

Introduction

Artificial intelligence is now a vast and progressively developing field of computer science. The number of new products, tools and services using that mechanisms has been growing rapidly. AI technologies are becoming increasingly available as practical solutions in the real world. According to Wikipedia, the term "artificial intelligence" was first used by John McCarthy in 1956 at a scientific conference in Dartmouth, where was defined as the ability of machines to exhibit manifestations of human-like intelligence. The basis of AI are algorithms that allow computers to learn and make decisions. The most advanced programs have been inspired by the human understanding of how our brain works - the connections between neurons. One of the techniques used to build AI systems is machine learning. ML algorithms are not based on rules manually entered into the system by humans, but they automatically build models based on training data and a defined goal. Another method is Deep Learning. DL models are extremely complex structures containing millions of parameters and capable of analysing complicated tasks such as human speech, images, video streams or non-trivial patterns in business data. Deep Learning models usually require a large amount of training data and computational power necessary to process them - that's why we have been observing their effects only for several years. But it is the deep models behind the recent successes of AI and machine learning.

The introduction should be about three to five pages long. Make sure you use references [1]

Chapter 2

Context

- Provide a context for your project.
- Set out the objectives of the project
- Briefly list each chapter / section and provide a 1-2 line description of what each section contains.
- List the resource URL (GitHub address) for the project and provide a brief list of the main elements at the URL.

2.1 Objectives

2.2 Chapter Review

2.2.1 Methodology

2.2.2 Technology Review

2.2.3 System Design

2.2.4 System Evaluation

2.2.5 Conclusion

2.3 Background Research

Chapter 3

Methodology

3.1 Agile Development

3.2 Version Control

3.3 Sprints

3.3.1 Sprint 1

3.3.2 Sprint 2

3.3.3 Sprint 3

3.3.4 Sprint 4

3.3.5 Sprint 5

3.3.6 Sprint 6

3.3.7 Sprint 7

3.3.8 Sprint 8

3.3.9 Sprint 9

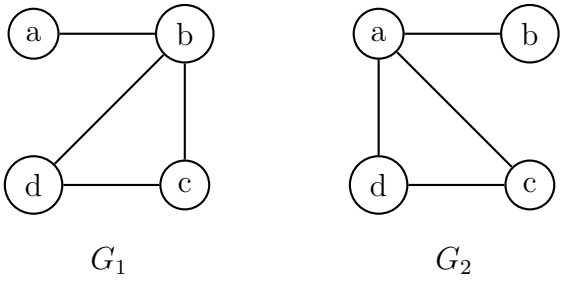
3.3.10 Sprint 10

3.3.11 Sprint 11

3.3.12 Sprint 12

3.4 Testing

About one to two pages. Describe the way you went about your project:



G_1 G_2

Figure 3.1: Nice pictures

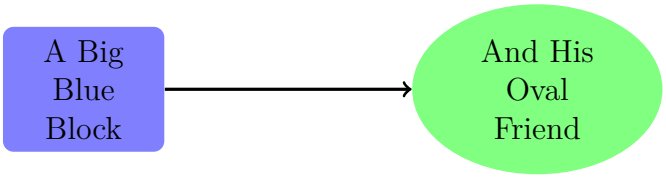


Figure 3.2: Nice pictures

- Agile / incremental and iterative approach to development. Planning, meetings.
- What about validation and testing? Junit or some other framework.
- If team based, did you use GitHub during the development process.
- Selection criteria for algorithms, languages, platforms and technologies.

Check out the nice graphs in Figure 3.2, and the nice diagram in Figure **??**.

Chapter 4

Technology Review

About seven to ten pages.

- Describe each of the technologies you used at a conceptual level. Standards, Database Model (e.g. MongoDB, CouchDB), XMl, WSDL, JSON, JAXP.
- Use references (IEEE format, e.g. [1]), Books, Papers, URLs (timestamp) – sources should be authoritative.

4.1 Visual Studio Code

4.2 GitHub

4.3 Python

4.4 JSON

4.5 Google Cloud Platform

4.6 AIY Voice Kit from Google

4.7 Natural Language Processing

4.8 TensorFlow

4.9 SQLite3

4.10 MySQL

4.11 Tkinter

4.12 CSV

4.13 LaTex

4.14 TeXstudio

Chapter 5

System Design

5.1 Architecture

5.2 Data Design

5.2.1 Dataset Generation

5.2.2 JSON

5.2.3 CSV

5.3 Component Design

5.3.1 Artificial Neural Networks

5.3.2 Pattern Matchers

5.3.3 NLP

5.4 GUI

As many pages as needed.

- Architecture, UML etc. An overview of the different components of the system. Diagrams etc... Screen shots etc.

Column 1	Column 2
Rows 2.1	Row 2.2

Table 5.1: A table.

Chapter 6

System Evaluation

As many pages as needed.

- Prove that your software is robust. How? Testing etc.
- Use performance benchmarks (space and time) if algorithmic.
- Measure the outcomes / outputs of your system / software against the objectives from the Introduction.
- Highlight any limitations or opportuni-ties in your approach or technologies used.

Chapter 7

Conclusion

About three pages.

- Briefly summarise your context and ob-jectives (a few lines).
- Highlight your findings from the evalua-tion section / chapter and any opportuni-ties identified.

Bibliography

[1] A. Einstein, "Zur Elektrodynamik bewegter Körper. (German) [On the electrodynamics of moving bodies]," *Annalen der Physik*, vol. 322, no. 10, pp. 891–921, 1905.