Ian Tai

# Learning the Stock Market: Deep Learning and Sentiment Analysis-Based Stock Price Prediction

# Proforma

| | |
|---|---|
| Name: | **Ian Tai** |
| College: | **Trinity College** |
| Project Title: | **Learning the Stock Market: Deep Learning and Sentiment Analysis-Based Stock Price Prediction** |
| Examination: | **Computer Science Tripos Part II, May 2018** |
| Word Count: | [1] |
| Project Originator: | **Ian Tai** |
| Supervisors: | **Dr Sean Holden** |
| | **Prof Stephen Satchell** |

## Original Aims of the Project

Deep Learning has increasingly been applied to many fields of industry. Among Finance, the applications of Deep Learning in predicting stock market prices is an increasingly popular research field. This dissertation proposes a method of stock price prediction using a combination of Long Short-Term Memory Recurrent Neural Networks and a variety of Sentiment Analysis techniques. The project aims to apply this method to collected news headlines from selected news agencies on Twitter and stock price data from the latter two quarters of 2017.

## Work Completed

This project has been successful; all success criteria have been met. I collected financial data from various sources, parsing and converting the data. I built a data collection and processing system for the Twitter dataset. I implemented a Gaussian Naive Bayes Classifier, a Multinomial Naive Bayes Classifier, and used the Scikit-Learn library for implementing Multi-Class Support Vector Machines and Semi-Supervised Support Vector Machines for Twitter sentiment analysis. Finally, I built a Long Short-Term Memory Recursive Neural Network for stock price prediciton.

---

[1]This word count was computed by TeXcount

## Special Difficulties

Finding appropriate Twitter datasets proved to be more difficult than was foreseen, which led to trying various sources of data, different models of classification, and ultimately, manual classification of the data for train/test purposes.

# Declaration

I, Ian Tai of Trinity College, being a candidate for Part II of the Computer Science Tripos, hereby declare that this dissertation and the work described in it are my own work, unaided except as may be specified below, and that the dissertation does not contain material that has already been used to any substantial extent for a comparable purpose.

Signed: Ian Tai

Date: April 25, 2018

# Contents

# List of Figures

# Acknowledgements

I would like to thank the following people for the help they have given me:
    TODO: COMPLETE AFTERWARDS

# Chapter 1

# Introduction

TODO: Starting sentence

## 1.1 Motivation & Aims

# Chapter 2

# Preparation

This chapter is empty!

# Chapter 3

# Implementation

## 3.1 Verbatim text

Verbatim text can be included using \begin{verbatim} and \end{verbatim}. I normally use a slightly smaller font and often squeeze the lines a little closer together, as in:

```
GET "libhdr"

GLOBAL { count:200; all  }

LET try(ld, row, rd) BE TEST row=all
                        THEN count := count + 1
                        ELSE { LET poss = all & ~(ld | row | rd)
                               UNTIL poss=0 DO
                               { LET p = poss & -poss
                                 poss := poss - p
                                 try(ld+p << 1, row+p, rd+p >> 1)
                               }
                             }
LET start() = VALOF
{ all := 1
  FOR i = 1 TO 12 DO
  { count := 0
    try(0, 0, 0)
    writef("Number of solutions to %i2-queens is %i5*n", i, count)
    all := 2*all + 1
  }
  RESULTIS 0
}
```
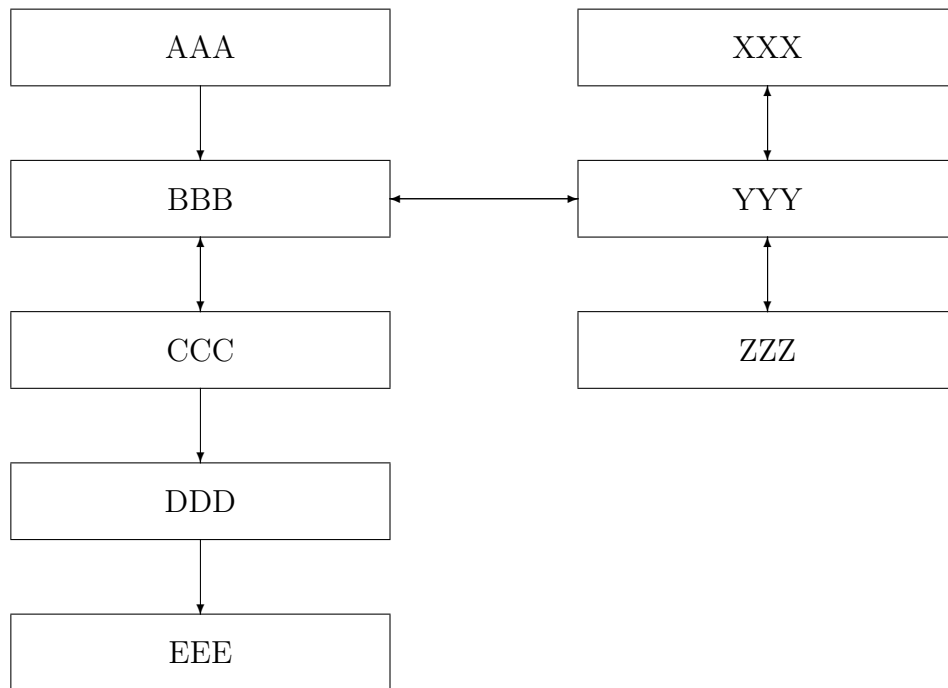
```
┌─────────────────┐              ┌─────────────────┐
│       AAA       │              │       XXX       │
└─────────────────┘              └─────────────────┘
         │                                 ↕
┌─────────────────┐              ┌─────────────────┐
│       BBB       │ ←──────────→ │       YYY       │
└─────────────────┘              └─────────────────┘
         │                                 ↕
┌─────────────────┐              ┌─────────────────┐
│       CCC       │              │       ZZZ       │
└─────────────────┘              └─────────────────┘
         │
┌─────────────────┐
│       DDD       │
└─────────────────┘
         │
┌─────────────────┐
│       EEE       │
└─────────────────┘
```

Figure 3.1: A picture composed of boxes and vectors.

## 3.2   Tables

Here is a simple example[1] of a table.

| Left Justified | Centred | Right Justified |
|---|---|---|
| First | A | XXX |
| Second | AA | XX |
| Last | AAA | X |

There is another example table in the proforma.

## 3.3   Simple diagrams

Simple diagrams can be written directly in LaTeX. For example, see figure ?? on page ?? and see figure ?? on page ??.

## 3.4   Adding more complicated graphics

The use of LaTeX format can be tedious and it is often better to use encapsulated postscript (EPS) or PDF to represent complicated graphics. This is my recommended way of drawing
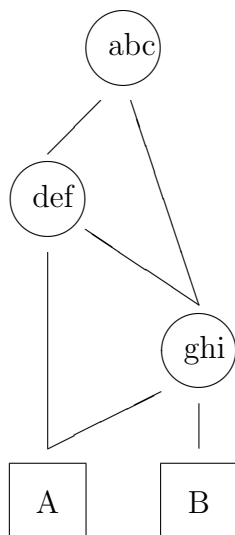
---

[1]A footnote

Figure 3.2: A diagram composed of circles, lines and boxes.

all diagrams.

Figure 3.3: Example figure where a picture can be pasted in

# Chapter 4

# Evaluation

## 4.1 Printing and binding

Use a "duplex" laser printer that can print on both sides to print two copies of your dissertation. Then bind them, for example using the comb binder in the Computer Laboratory Library.

## 4.2 Further information

See the Unix Tools notes at

```
http://www.cl.cam.ac.uk/teaching/current-1/UnixTools/materials.html
```

# Chapter 5

# Conclusion

I hope that this rough guide to writing a dissertation is LaTeX has been helpful and saved you time.

# Appendix A

# Latex source

# Appendix B

# Project Proposal