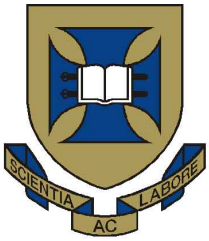

THE UNIVERSITY
OF QUEENSLAND

AUSTRALIA

Undergraduate Engineering Honours Thesis

Predicting Responses to Spaced Repetition Flash Cards with Machine Learning Techniques

---

Author:
Jordan West

Supervisor:
Dr. Mark Schulz

5 November, 2012

Contents

1 Introduction

test

2 Background

2.1 Machine Learning

2.1.1 Neural Networks

2.1.2 Support Vector Machines

2.2 Spaced Repetition

Supermemo 2 Algorithm (SM2)

$$I(1) := 1 \tag{1}$$

$$I(2) := 6 \tag{2}$$

$$I(n) := I(n-1) \times EF \tag{3}$$

$$EF := EF + (0.1 - (5-q) \times (0.08 + (5-q) * 0.02)) \tag{4}$$

2.3 Similar Projects

Memrise

Memrise is a private company which produces web-based flashcard software.

The Mnemosyne Project

Mnemosyne is open source spaced repetition software collecting anonymised data from its many users in order to evaluate the effectiveness of the implemented spaced repetition algorithm [2]. Mnemosyne uses a modified version of the Supermemo algorithm. The project does not appear to have produced any papers or research publications at this time.

Anki

3 Goals

Develop an online spaced repetition learning environment

Visualise student usage of the learning environment

Figure 1: Screenshot of Mnemosyne in use

Predict student responses to flash cards

4 Methods and Materials

4.1 Ethical Clearance

Data Considerations

Application Details

4.2 Software Design

4.2.1 Requirements

4.2.2 Tools

Git and Github (`http://git-scm.com/`), (`http://www.github.com/`)

Git is a distributed version control system (VCS) which tracks changes to source code (often amongst multiple developers) and keeps a complete history of changes. This is invaluable when a change in code occurs that results in a critical bug. Versions can be compared to find the change that introduced the bug, and production code can be reverted if need be [3].

Git can be hosted anywhere, however Github is a popular

Ruby on Rails (`http://www.rubyonrails.org/`)

Ruby on Rails is a popular open source framework for developing web applications[1].

Heroku (`http://www.heroku.com/`)

Highcharts (`http://www.highcharts.com/`)

Highcharts is a Javascript

Twitter Bootstrap (`http://twitter.github.com/bootstrap`)

Twitter Bootstrap is a set of default styles for websites and web applications, provided as open-source by Twitter. Using Twitter Bootstrap rapidly speeds up theming of a web application with default looks for navigation, buttons, text and layout.

Compare the following pages with and without Twitter Bootstrap default styles added:

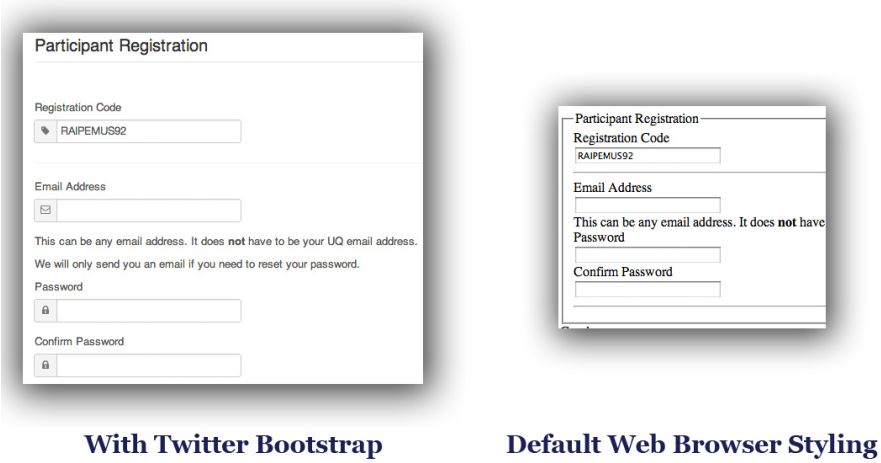

Figure 2: Comparison of a page with no styling and Twitter Bootstrap default styling

More significantly, Twitter Bootstrap offers a 'responsive' layout system which provides a reduced screen size (ie. smartphone) layout with little to no extra work on the part of the developer. This means a smartphone version of the web application could be designed at the same time. Twitter Bootstrap was also chosen for this reason.

4.2.3 Data Entry

4.2.4 Screen Mockups

4.2.5 Spaced Repetition Algorithm

4.2.6 Data storage, formatting and output

4.3 Data Analysis and Prediction

4.3.1 R Programming Language

4.3.2 Usage Data

4.3.3 Forgetting Curves

4.3.4 Prediction of Recall

5 Results

5.1 Online Learning Environment (Membit)

5.1.1 Overview

5.1.2 Bugs and Issues

5.2 Usage Statistics

5.3 Forgetting Curves

Generated from Recorded Reviews

Generated from Machine Learning Algorithms

5.4 Prediction of Recall

Support Vector Machines

6 Discussion

6.1 Evaluation

6.2 Potential Future Work

References

[1] M. Bachle and P. Kirchberg. Ruby on rails. *Software, IEEE*, 24(6):105 –108, December 2007.

[2] Peter Bienstman. Principles | the mnemosyne project. http://www.mnemosyne-proj.org/principles.php, 2012.

[3] Scott Chacon. *Pro Git*. Apress, 1st edition, 2009.

7 Acknowledgments

This thesis would not have been possible without the support of my supervisor Dr. Mark Schulz whose input and guidance has been invaluable for the project.

I would like to thank Dr Yuriko Nagata for her assistance with this and for allowing me to introduce the software to her students.

Dr Michael Harrington of the School of Languages and Comparative Cultural Studies for his advice with the online learning software, particularly the suggestion to record the time students take to review individual words.

I would also like to thank the UQ CEIT team for their help with this project - particularly Alan Cody and Phil Long.

Last but not least I want to thank my friends for their encouragement and for putting up with me throughout the project, and Mum and Dad for supporting me