Detecting User Engagement Using Mouse Tracking Data:
Project Specification

David Saunders (910995)

April 2020

Abstract

Write abstract here

Contents

1 Literature review

In this section I will review the literature on how to monitor attention.

2 Background Research

Anything I've looked at with help for mouse data classification algorithms?

3 Motivation and Aims of project

Can copy from presentation slides but fill in so they're more wordy.

3.1 Motivation

People are lazy. Often don't pay much attention Is there any way of measuring people's attention?

Why mouse data? Mouse cursor position is strongly correlated with eye position. One paper calls it a "poor man's eye-tracker" [find] Bulky expensive equipment for eye tracking is expensive and very obtrusive. Hawthorn / observer effect - People react differently when being observed. Less obtrusive mouse tracking can make people feel less tracked and act more naturally. Could even not tell them (legal ethical repercussions)!

3.2 Aims of project

The aims of what I want to achieve in the project will be as follows:

- Visualise, analyse and understand the data results.
- Use the data to train machine learning models to classify users between 2 groups.
- Combine the data and methods from the study data with other datasets to create a more robust model.
- Stretch goal? Test methods and models developed with other applications?

Talk here about how I will achieve each aim, then describe the components of the project that I will need to complete. Try and link each component of the project to an aim.

Machine learning methods SVM Natural Language Processing N-Grams LSTM Neural Networks Markov models Deal with Imbalances in classes Sampling Oversampling, Undersampling Other mouse data sources

Applications A good system developed could be used for other tasks to monitor attention - E.g. Survey Monika made us do. Not just for joes ice-cream Have to decide on the trade off between a good narrow (is this the right word) classifier between attention or not and a more generalised model that can work on any task. What I mean by that is I can model the html elements / sliders to see how users interacted to see the stock prices, or I can generalise to any such task involving mouse data.

4 Project plan

4.1 Development methodology

Discuss software life cycle methodologies with Jacques. An agile methodology such as scrum would probably be best but am I constrained by this specification document?

Figure 1: A Gantt chart showing the planned milestones of the project. OR A Gantt chart created in the GanttProject free software

5 Risk Analysis

When creating a project there is always potential risks that the project might encounter and hinder its chances of success. In order to prepare and to hopefully avoid these risks I will now list and analyse the most likely, and most devastating risks to my project. By analysing each risk individually I will be prepared in case I come across any of the potential risks and I will have developed a plan of action of what to do and how to manage myself in case of encountering them. I will initially list the risks in a table where I will briefly examine them, then I will go into each one in more detail. Below I have listed and analysed the risks and have ordered them from potentially the most dangerous to least dangerous.

Caption A risk analysis table DAAAAAAAAAAAAAAAAAVE

6 Conclusion

Measuring user engagement is challenging Mouse data can help us solve that issue by showing user attention Data Science techniques could be used to help classify the data (Not SVM)

Table 1: The top association rules between individual items.

Risk	Probability	Impact	Combined Risk	Mitigation Plan	Contingency Plan
Unrealistic time plan and poor time management.	High	High	High	Create work schedule and stick to it.	If I am unable to stick to my work schedule, I must adapt my approach to work and create an undated, more realistic schedule.
Coronavirus affects me or a close family member, negatively effecting my work.	Medium	High	High	Stay safe during the quarantine to keep everyone safe and mitigate any risks of me catching anything.	Inform the University as soon as a situation develops so we can arrange something.
Coronavirus has a greater impact on Swansea University and effects the available support and deadlines.	Medium	High	Medium	Keep informed with the University College of Science and supervisor to any news effecting the University.	Keep my options open? Keep updated?
No correlation between attention and mouse tracking data can be found.	Low	High	Medium	Attempt as many different methods of classification early before writing in depth about them.	If no insights can be gained from the given dataset, I will attempt to find correlations in other datasets.