Detecting User Engagement Using Mouse Tracking Data:
Project Specification

David Saunders (910995)

April 2020

Abstract

Write abstract here

Contents

Mark scheme

This coursework contributes 50% of the mark for the module. The size is approximately 5000 words (excluding references) – due on Wednesday 29 April 2020 (11:00 am).

This report should give a literature review over your project and describe any background research that you have carried out. You should state the motivation and aims of the project. It should include a complete specification of your project. It should describe the project clearly and the components of the work which need to be developed. An outline project plan for the summer should be included. This plan should take into account the development methodology

being used. You should provide a risk analysis for the project. You should view this document as providing the plan for the work you expect to carry out over the summer.

1 Literature review

Write section 1 here [1] and talk about figure 1.

This can pretty much be the review I did for the first assignment.

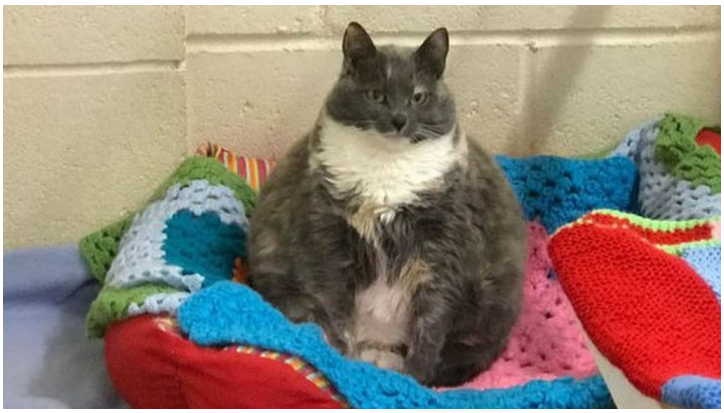

Figure 1: This will be a figure showcasing some of my work

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?

5 Risk Analysis

Copy this from my project last year.

Table 1: A risk analysis table

Risk	Likelihood	Impact
		Low High

Figure 2: A Gantt chart showing the planned milestones of the project.

Risks: - BIGGEST we find that there is no correlation between attendtion and mouse data and nothing is prooved. - COID19 effecting the UK more. - I was considering running more lab studies to get more data but the shutdown has stopped that ambishion. It has already effected, anything worse like close family and elderly parents getting ill so supervisor or me would be a risk. Plan ahead, wash hands.

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)

References

[1] Thomas Torsney-Weir et al. "Tuner: Principled parameter finding for image segmentation algorithms using visual response surface exploration". In: *IEEE Transactions on Visualization and Computer Graphics* 17.12 (2011), pp. 1892–1901.