Detecting User Engagement Using Mouse Tracking Data:
Project Specification

David Saunders (910995)

April 2020

Abstract

This project specification reviews relevant materials and the background of my project. The motivation and aims of the project are explained, and a comprehensive plan of work for the summer is present.

Contents

1 Motivation

Crowd-sourcing marketplaces like Amazon's Mechanical Turk are a popular service that provides a way of gathering data from real participants for studies, and human intelligence tasks [1]. The level of user engagement, attention, and low quality responses can all be issues when gathering data from participants in such a distributed way [2]. Research has been conducted on how to increase the accuracy and attention of Turk tasks. Methods such as offering financial incentives [3] and engaging a users curiosity [4] have been found to motivate workers into performing better crowdsourced tasks. However no work has been done to calculate if crowdsourced users are paying attention during tasks.

2 Introduction

This project uses data from a previous study where participants were asked to perform a simple repetitive tasks [5]. Data was gathered both with a closely monitored lab study, and using a crowd-sourcing website. We assume that participants in the lab study were paying attention, and that crowd sourced participants may or may not be paying attention. The task of this project is to classify crowd sourced participants to determine who were paying attention during the task. We hypothesise that a participants level of attention can be measured from their mouse movement data. The project will propose methods of identifying and quantifying user engagement by using machine learning and visual analytics techniques.

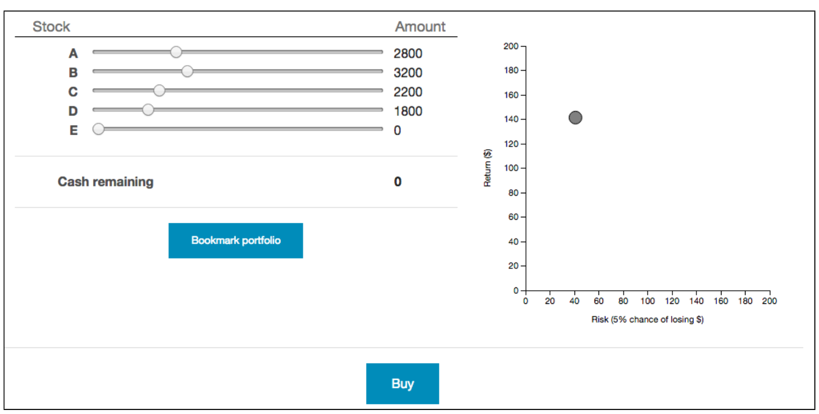

Figure 1: Interface for the study [5].

Fig 1 is the design of the interface participants of the study interacted with. The aim of the task is to maximise the prophet made over a series of tasks. Participants may interact with the stocks sliders to select how much of a stock they would like to buy. The return and risk is shown on the plot on the right of the figure, this plot updates when the sliders are changed in an attempt to inform participants of the stocks volatility. It is hypothesised that participants paying more attention will interact with the sliders more attempting to find the best combination, something their mouse data may reveal.

2.1 Aims of project

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

- Visualise, analyse and understand the data.
- Use the data to train machine learning models to classify users by their attention level.
- Determine if there is a link between a users attention and task performance.
- Combine the data and methods from the study data with other datasets to create a more robust model.
- Publicly share all data and any findings of this project.

Here I detail how I will achieve each aim, and describe the components of the project that I will need to complete.

To visualise and understand the data I will use tools like tableau and plot basic plots such as histograms. If any interesting correlation or information is found in the data at this stage it can be more fully explored in depth later in the project.

Semi supervised learning techniques can be used to classify which of the crowdsourced participants were paying attention. The idea of semi supervised is combining labelled and unlabelled data to improve the learning behaviour of a machine learning technique [6]. This will be used in this project as only a subset of the data is labelled, where the lab study results are labelled as paying attention, and crowdsourced results are unlabelled.

To see if a users attention influences their task performance relies on me being able to extract user attention information from the mouse data. If I am unable to do so I could attempt to use just the lab study data, however there is only a small number of datapoints, and not enough to make a significant

Other datasets can be researched and explored. Similar datasets do exist online and are publicly available [7]. Overfitting is a big issue with any machine learning technique, especially when there is not a large amount of data for training [8]. In an attempt to avoid this I can increase the quantity of data available to me by incorporating other sources and hopefully create a more robust model.

To publicly share the result of this project I will upload any work completed for it in a GitHub repository. Sharing data and results with the Data Science community is an important part of scientific research as it allow others to peer review and reproduce my results [9].

3 Background Research

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

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

3.1 User attention and characteristic.

3.2 Eye tracking

As mentioned previously in the report non-verbal information can be used to detect a user's level of engagement, eye tracking is a prime example of this [10]. Vision is one of the most powerful human senses so it may give a good measure of user engagement. The methodology of eye tracking is that we move our eyes to focus on particular areas that we want to see in more detail, and divert our attention to that area [11]. Thus tracking a user's gaze can provide insight into which part of a system they're engaged with, and how much so.

Figure 2: Heatmap showing the popular locations of users eyes on a webpage [12].

Eye tracking data can be used to show user interface elements that users focus their attention on as shown in Fig 2. From this researchers were able to predict the amount of attention elements of the page would receive. By observing what parts of an interface users are interacting with we can determine what a user is engaging with [12]. Eye-tracking has been used, and found success novel applications such as recording the engagement of users when playing a game. Tracking users eye movements helped game designers understand how users can recognise interactable game objects and could be used to investigate problematic game design issues [13].

3.3 Mouse cursor tracking

Eye tracking has historically had its limitations. To track subjects eyes with a good degree of accuracy required the use of expensive, intrusive equipment that frequently needed recalibrating [14]. In contrast mouse movement data can be collected without the drawbacks of eye tracking and with more automatic methods, meaning more data can be collected, and on a larger scale [15]. Research has also found that there is a correlation between a user's gaze and their cursor position. The position can be considered a "poor man's eye tracker" as it has been found that eye gaze match mouse position 69% of the time [16]. Therefore it can be said that mouse data can be used as a good alternative to eye tracking data.

Mouse activity can be used as input to a neural network and output a quantifiable level of activity for a webpage. By using mouse data it is possible to unobtrusively record a user's normal use of a web browser without disturbing their experience [17]. It has also been found that users tend to follow the text they're reading with the mouse cursor [18], and similarly scientists were able to determine what paragraph of a page was being read with an accuracy of 79% by using mouse cursor data [19]. Methodologies mentioned above explore ways of classifying user engagement from eye and cursor data, however it is also possible to predict users attention and user frustration in complex webpages [20]. In contrast to the above literature some studies disagree that mouse cursor is always a good approximation for eye data. Hauger et al found that distinct cursor behaviour exists depending on the task, and that the relationship between eye gaze and mouse position is more nuanced than measuring only mouse data [21].

3.4 Machine learning techniques

I've given a high level overview of the existing literature, here I will be more in depth as to the algorithms and techniques relevant to this project.

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

Table 1: The first 5 records of results of the crowdsourced task.

event_type	target	time	x	y	step	turkId
mousedown	alloc-slider-1	0	477	405	1	A35YFAFWP33C70
mouseup	alloc-slider-1	0.111	478	405	1	A35YFAFWP33C70
click	alloc-slider-1	0.111	478	405	1	A35YFAFWP33C70
mousedown	alloc-slider-1	1.516	479	405	1	A35YFAFWP33C70
mousedirchange	alloc-slider-1	2.395	543	403	1	A35YFAFWP33C70
mousedirchange	alloc-slider-1	3.161	594	402	1	A35YFAFWP33C70
mouseup	alloc-slider-1	5.048	514	407	1	A35YFAFWP33C70
click	alloc-slider-1	5.048	514	407	1	A35YFAFWP33C70
mousedown	alloc-slider-2	5.461	494	441	1	A35YFAFWP33C70
mouseup	alloc-slider-2	5.513	494	441	1	A35YFAFWP33C70

4 Description of the project

4.1 Components of project

Each of the component's cannot be fully separated into unique sections as there is an order that they must be completed in.

Now that the background for the project has been researched the more technical programming side of the project can begin. The data must first be extracted from the JSON format that its in to a more tabular format. After this the main bulk of the project can begin. This will consist of researching different algorithms and methods and attempting to implement them on the project data. Any results from this will then be analysed, visualised, and evaluated. These stages can be repeated multiple times as different methods are explored. After multiple different approaches have been explored the results should be properly documented and compiled in my dissertation.

4.2 Description of data

The first component of the project has been completed, and data has been extracted form JSON format to a csv format.

Table 1 shows us the features of the data. The lab study data and the crowdsourced data have the same schema, except lab results have a different ID field.

Target shows us which element in Fig 1 a participant is interacting with, and event_type details the type of interaction. The time field shows us the time taken in seconds since the first recorded mouse event. We can hypothesise that participants with a shorter time may have paid less attention than a participant who took much longer, thinking about their actions more. The x and y fields show the location of the mouse and step shows which stage of the task, from 1 to 5, a participant was in.

5 Project plan

The different components of the work have been explained above. This section will specify the timeframe and order in which the modules will be carried out.

5.1 Development methodology

I will be using an agile methodology as it will allow flexibility of my project and the iterative nature should help me to constantly improve it [22].

Scrum will be used as the short scrum periods will encourage bursts of development over the long summer period [23].

Fig 3 shows the sprints I will be undertaking. The Gantt Chart was created with the free software Gantt Project [24]. Each sprint starts with a supervisor meeting where the previous work, and the plan for the next sprint will be discussed. There will be three sprints in total, with each sprint being a fortnight long. Each sprint will consist of researching a method I may be able to use in my project. Then I will spend time modifying and implementing the method so that it may be used in the context of this project. The latter half of a sprint will consist of analysing and visualising the results of the method and writing these down roughly in the dissertation document. The effectiveness of the method will be evaluated, and I will attempt to fix any failures or shortcomings of the method. For example I may come across a particular variant of an artificial neural network in my research that I believe may be useful in the project. The sprint plan would allow me to spend time implementing this network with my data and experiment with the parameters.

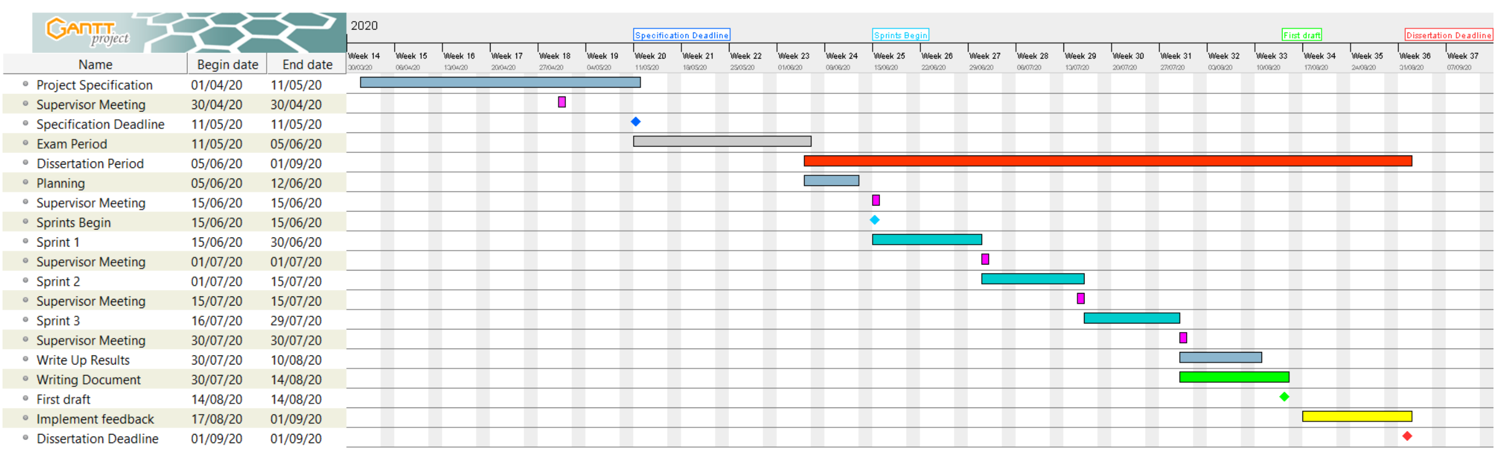

Figure 3: Gantt chart showing the planned timeline and milestones of the project.

5.2 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 risks of 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. Each risk is explained with the likelihood of the risk occurring and impact to the project the risk would have. A mitigation plan is created in an attempt to prevent the risk from happening, and a contingency plan is made so I can be prepared if the risk does occur. Below I have listed and analysed the risks and have ordered them from potentially the most dangerous to least dangerous.

1. **Risk:** *Unrealistic time plan and poor time management.*

   Likelihood and Impact: Medium likelihood, Medium Impact

   Explaination If my time is spent poorly then I could not have a piece of work finished for the submission deadline, or the work may not represent the best of my abilities.

   Mitigation: Create work schedule and stick to it. A work schedule and plan for the summer has been created in this document which I aim to follow.

   Contingency: If I am unable to stick to my work schedule, I must adapt my approach to work and create an undated, more realistic schedule.

2. **Risk:** *Coronavirus affects me or a close family member, negatively effecting my work.*

   Likelihood and Impact: Medium likelihood, High Risk

   Explaination Coronavirus is very contagious. in spite of protective measures it is still likely that the I may become infected with the virus.

   Mitigation: Stay safe indoors during the quarantine to keep everyone safe and mitigate any risks of me becoming infected.

   Contingency: Inform the University as soon as any negatively situation develops so that alternative assessments can be organised.

3. **Risk:** *No correlation between attention and mouse tracking data can be found.*

   Likelihood and Impact: Medium likelihood, High impact

   Explaination: The project will involve the use of many methods to find a link between mouse tracking data and user attention. However the number of total data samples are in the hundreds so there is not a massive amount of information to draw conclusions from. It is possible that after all methods have been exhausted no correlation is ever discovered, or simply doesn't exist.

Mitigation: Attempt as many different methods of classification early before writing in depth about them.

Contingency: If no insights can be gained from the given dataset, I will explore other similar datasets and attempt to find correlations there. I will then attempt to apply findings from other datasets to the original dataset.

4. **Risk:** *Coronavirus has a greater impact on Swansea University and effects my available support and deadlines.*

   Likelihood and Impact: Low likelihood, Low impact

   Explaination: The virus has already shut down in person teaching and with the UK in lockdown it is unlikely the situation will become vastly different.

   Mitigation: Keep informed with the University College of Science and supervisor to any news effecting the University.

   Contingency: Keep updated with the situation and follow whatever advice is recommended from the university.

6 Conclusion

This report has given an introduction to the work that will be conducted over the next few months. The background of the project has been presented, and existing methods researched. A detailed plan of the project details the timeline, and components of the project.

References

[1] Gabriele Paolacci, Jesse Chandler and Panagiotis G Ipeirotis. "Running experiments on amazon mechanical turk". In: *Judgment and Decision making* 5.5 (2010), pp. 411–419.

[2] Panagiotis G Ipeirotis, Foster Provost and Jing Wang. "Quality management on amazon mechanical turk". In: *Proceedings of the ACM SIGKDD workshop on human computation*. 2010, pp. 64–67.

[3] Chien-Ju Ho et al. "Incentivizing high quality crowdwork". In: *Proceedings of the 24th International Conference on World Wide Web*. 2015, pp. 419–429.

[4] Edith Law et al. "Curiosity killed the cat, but makes crowdwork better". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016, pp. 4098–4110.

[5] Thomas Torsney-Weir et al. “Risk Fixers and Sweet Spotters: a Study of the Different Approaches to Using Visual Sensitivity Analysis in an Investment Scenario”. In: *EuroVis 2018 - Short Papers*. Ed. by Jimmy Johansson, Filip Sadlo and Tobias Schreck. The Eurographics Association, 2018. ISBN: 978-3-03868-060-4. DOI: 10.2312/eurovisshort.20181089.

[6] Xiaojin Zhu and Andrew B Goldberg. “Introduction to semi-supervised learning”. In: *Synthesis lectures on artificial intelligence and machine learning* 3.1 (2009), pp. 1–130.

[7] Human Computation. *Workers Browser Activity in CrowdFlower Tasks*. Oct. 2017. URL: https://www.kaggle.com/humancomp/worker-activity-crowdflower (visited on 03/05/2020).

[8] Tom Dietterich. “Overfitting and undercomputing in machine learning”. In: *ACM computing surveys (CSUR)* 27.3 (1995), pp. 326–327.

[9] Jeremy P. Birnholtz and Matthew J. Bietz. “Data at Work: Supporting Sharing in Science and Engineering”. In: *Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work*. GROUP ’03. Sanibel Island, Florida, USA: Association for Computing Machinery, 2003, pp. 339–348. ISBN: 1581136935. DOI: 10.1145/958160.958215. URL: https://doi.org/10.1145/958160.958215.

[10] Divesh Lala et al. “Detection of social signals for recognizing engagement in human-robot interaction”. In: *arXiv preprint arXiv:1709.10257* (2017).

[11] Andrew T Duchowski. “Eye tracking methodology”. In: *Theory and practice* 328.614 (2007), pp. 2–3.

[12] Georg Buscher, Edward Cutrell and Meredith Ringel Morris. “What do you see when you’re surfing? Using eye tracking to predict salient regions of web pages”. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2009, pp. 21–30.

[13] Tony Renshaw, Richard Stevens and Paul D Denton. “Towards understanding engagement in games: an eye-tracking study”. In: *On the Horizon* (2009).

[14] Daniel C Richardson and Michael J Spivey. “Eye tracking: Characteristics and methods”. In: *Encyclopedia of biomaterials and biomedical engineering* 3 (2004), pp. 1028–1042.

[15] Urška Demšar and Arzu Çöltekin. “Quantifying gaze and mouse interactions on spatial visual interfaces with a new movement analytics methodology”. In: *PloS one* 12.8 (2017).

[16] Lynne Cooke. “Is the Mouse a” Poor Man’s Eye Tracker”?” In: *Annual Conference-Society for Technical Communication*. Vol. 53. 2006, p. 252.

[17] Jeremy Goecks and Jude Shavlik. “Learning users’ interests by unobtrusively observing their normal behavior”. In: *Proceedings of the 5th international conference on Intelligent user interfaces*. 2000, pp. 129–132.

[18] Chen-Chung Liu and Chen-Wei Chung. "Detecting mouse movement with repeated visit patterns for retrieving noticed knowledge components on web pages". In: *IEICE transactions on information and systems* 90.10 (2007), pp. 1687–1696.

[19] David Hauger, Alexandros Paramythis and Stephan Weibelzahl. "Using browser interaction data to determine page reading behavior". In: *International Conference on User Modeling, Adaptation, and Personalization.* Springer. 2011, pp. 147–158.

[20] Vidhya Navalpakkam and Elizabeth Churchill. "Mouse tracking: measuring and predicting users' experience of web-based content". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 2012, pp. 2963–2972.

[21] Jeff Huang, Ryen White and Georg Buscher. "User see, user point: gaze and cursor alignment in web search". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 2012, pp. 1341–1350.

[22] Kent Beck et al. "Manifesto for agile software development". In: (2001).

[23] Ken Schwaber. "Scrum development process". In: *Business object design and implementation.* Springer, 1997, pp. 117–134.

[24] GanttProject Team. *GanttProject - Free project scheduling and management app for Windows, OSX and Linux.* Apr. 2020. URL: `https://www.ganttproject.biz/` (visited on 03/05/2020).