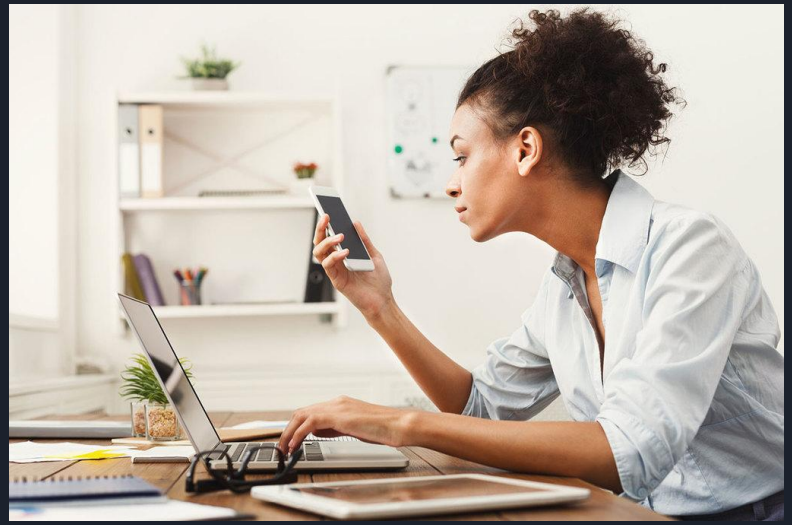
Detecting User Engagement Using Mouse Tracking Data

David Saunders - 910995
Supervisor: Dr Thomas Torsney-Weir

Motivation



[1]

Question: Are you all paying attention?

People are lazy.

Often don't pay much attention

Is there any way of measuring people's attention?

Presentation Overview

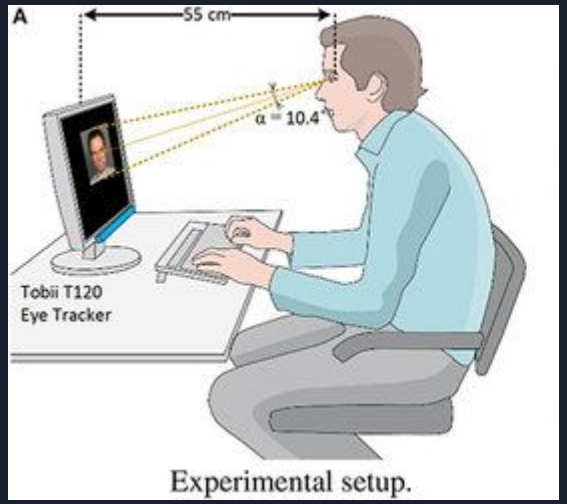
Table of Contents

Existing methods of attention monitoring

Eye tracking

Speech?

Mouse tracking.



Experimental setup.

[2]

Why mouse tracking?

Strongly correlated with eye position

Cheaper than eye tracking

Easier

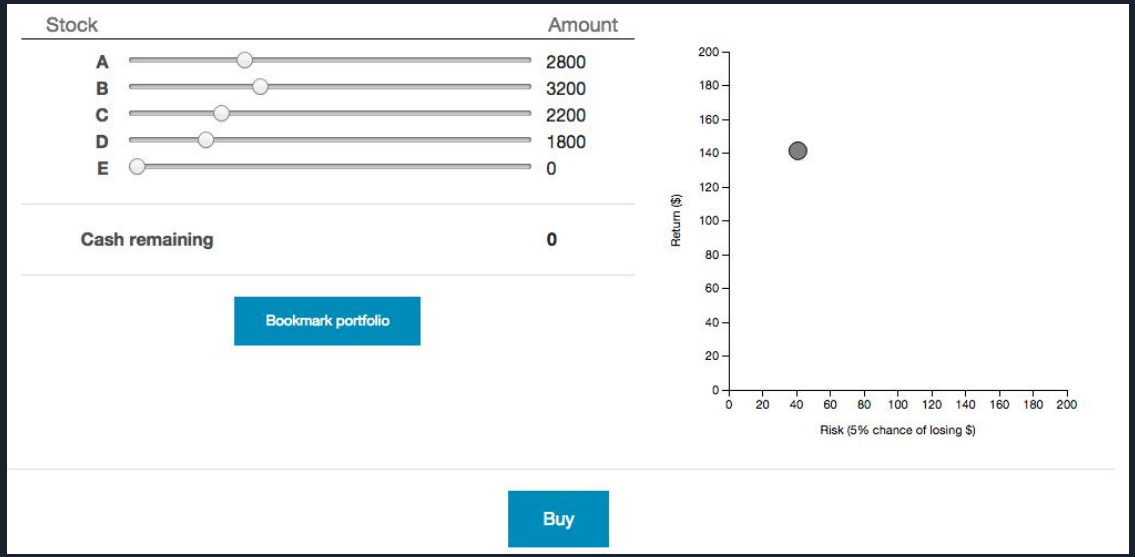
Less obtrusive

Background

Study where participants were asked to perform a simple task.

Asked to maximise return of stock portfolio, minimising risk

5 stages



Data

- Lab Study
  - Actually conducted in an Austrian dance hall
  - 13 participants

- Online Data
  - Used Amazon's Mechanical Turk
  - Crowdsourcing website.
  - Crowdworkers / Turkers
  - Problems?

[3]

Data - sample

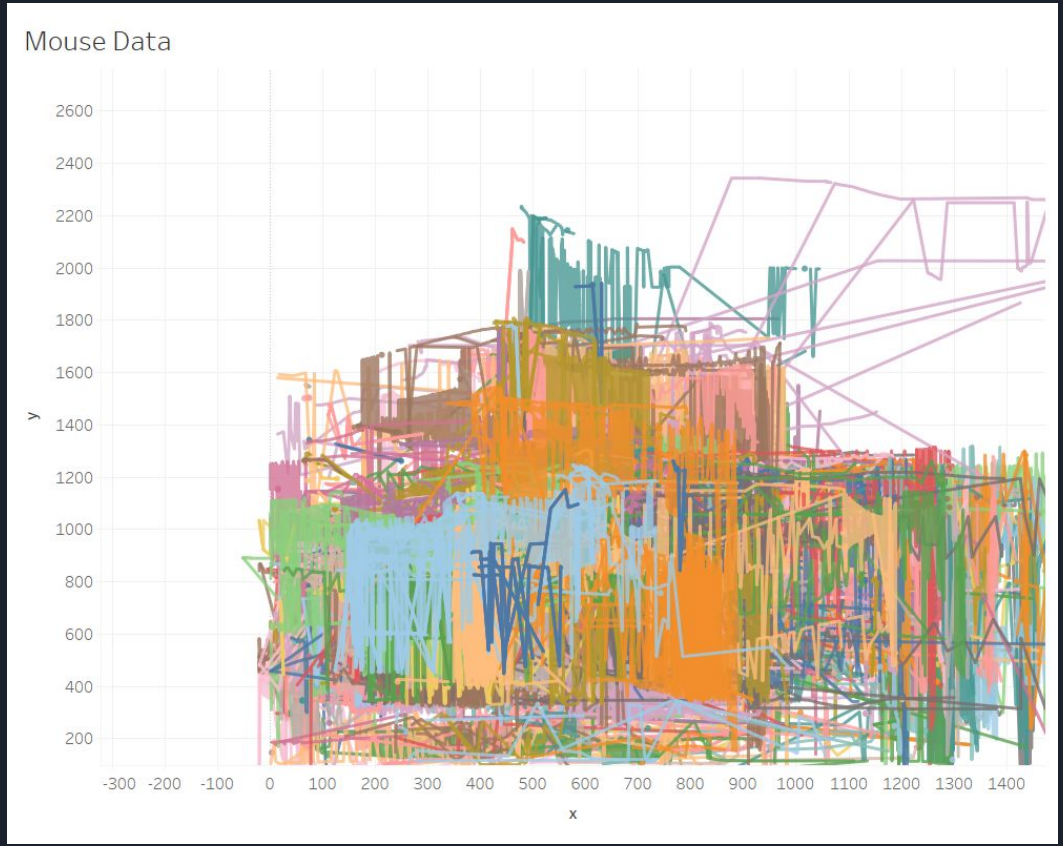
event_type	target	time	x	y	step	turkId
mousedown	alloc-slider-1	0	477	405	1	A35YFAFWP33C70
mouseup	alloc-slider-1	0.111	478	405	1	A35YFAFWP33C70
click	alloc-slider-1	0.111	478	405	1	A35YFAFWP33C70
mousedown	alloc-slider-1	1.516	479	405	1	A35YFAFWP33C70

AIMS of project

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.



Problem - Is this showing us user attention? - Too much data 1.3 millions rows of data for 461 users

Work so far

[4]

Tidy and preprocess Data

Extracted data from megabytes (maybe gigabyte with lab data) of json files (painful) to tabular.

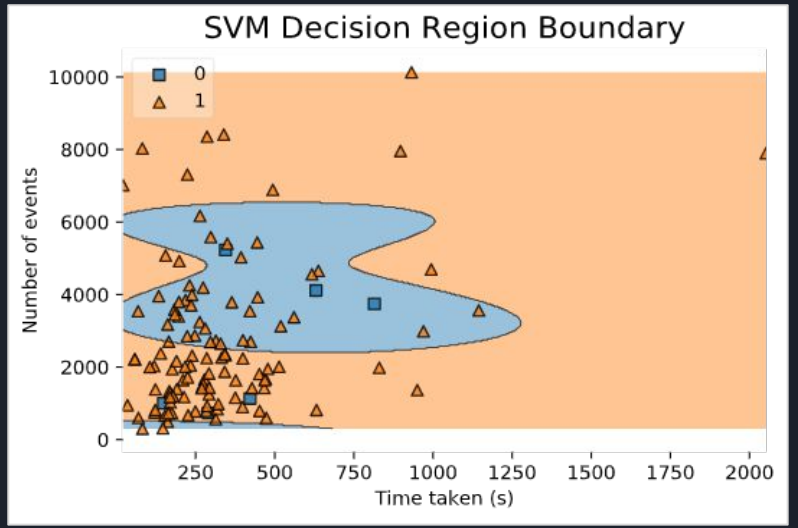
Doing everything in python, in-built json extension not overly useful because of the specifics of the task.

Data Visualisation

Initial visualisations in tableau

Work to do

- 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 have other applications.

Trade off between generalised and specific model.

Summary

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)

Thank you for listening!

Any Questions?

References

[1] Distracted woman at PC, Retrieved Match 2020.
https://static.seattletimes.com/wp-content/uploads/2018/10/web-tips-Distracted-photo.jpg

[2] Eye tracking experiment, Retrieved Match 2020.
https://www.researchgate.net/publication/316083826_Gaze_Behavior_Consistency_among_Older_and_Younger_Adults_When_Looking_at_Emotional_Faces/figures?lo=1&utm_source=google&utm_medium=organic

[3] Amazon Turk logo, Retrieved Match 2020.
https://d2q79iu7y748jz.cloudfront.net/s/_squarelogo/212fed405f4047eceee0efbe6fb858d1

[4] JSON logo, Retrieved Match 2020.
https://miro.medium.com/max/512/1*MjsyjENCtzfGKzT13aZJcg.png