A picture containing room

Description automatically generated

**Faculty of Engineering and Environment**

**Research Proposal**

|  |  |  |
| --- | --- | --- |
| Module | : | KF7029 – MSc Computer Science and Digital Technologies Project |
| Module Tutor | : | Rebecca Strachan |
| Title of MSc Research Project | : | User Interaction Prediction in Web Applications: An Evaluation of Machine Learning Models |
| Student Name | : | Akhil SARIKONDA |
| Student ID | : | W19048844 |
| Programme of Study | : | MSc Advanced Computer Science with Advanced Practice |
| Name of supervisor | : | Abbas Mehrabidavoodabadi |
| Name of second marker | : | Jill Bradnum |
| Word | : | 1500 |

**Contents**

[1. Aim 3](#_Toc87345959)

[2. Individual Research Objectives 3](#_Toc87345960)

[3. Background and Literature Review 3](#_Toc87345961)

[4. Scope, Objectives, and Approach 5](#_Toc87345962)

[4.1. Scope 5](#_Toc87345963)

[4.2. Objectives 6](#_Toc87345964)

[4.3. Approach or Methodology 6](#_Toc87345965)

[5. Timelines and Timescales 7](#_Toc87345966)

[5.1. Project Break-down structure 7](#_Toc87345967)

[5.2. Task schedule plan 7](#_Toc87345968)

[5.3. Gantt Chart 8](#_Toc87345969)

[6. Ethics, Legal, Social, Security and Professional Issues 8](#_Toc87345970)

[6.1. Ethical issues 8](#_Toc87345971)

[6.2. Security issues 8](#_Toc87345972)

[6.3. Legal issues 9](#_Toc87345973)

[6.4. Professional issues 9](#_Toc87345974)

[6.5. Social issues 9](#_Toc87345975)

[References 10](#_Toc87345976)

# Aim

The overall aim of this research is to evaluate the usage of machine learning models to predict the user interaction in web applications

# Individual Research Objectives

Within the context of user interaction prediction, the specific objectives of this research are to do the following:

* Identify why user interaction prediction is required and how it improves the overall user experience
* Evaluate critically the machine learning models and frameworks relevant to supporting user interaction prediction in coping with user experience improvement.
* Examine the issues in currently existing machine learning models
* Formulate recommendations upon practical implementation work

# Background and Literature Review

**Background:**

User Interaction Prediction plays a major role in improving the user’s Web browsing quality of experience (QoE) [1, 3].

Web browsing pattern consists of a number of series of clicks triggered by users [1] and navigations to other applications. If the user interaction data is collected and provided, then a web browser can prefetch the resources (website content) to be accessed next in background, to make it immediately available when the user requests that content. Prefetching enables Web applications involving several multimedia resources or subsidiary resources from web applications to reduce the page load time to download. Google Chrome, for example, offers a predictive tool for loading pages faster by initiating acts such as DNS prefetching, TCP and TLS pre-connection, and preloading of Web applications (Gray J, 2009).

There are two main motivations for why I chose this proposal. Firstly, this project needs to have an in-depth experience in Web Development and good grasp of JavaScript Language and its practical implementation needs to be known. I have been learning JavaScript and Web Development for more than two years now and built several projects using it. I always wanted to build a Web Application which uses Machine Learning techniques in it.

JavaScript has its own frameworks and libraries like tensorflow.js (Smilkov, Daniel, et al, 2019) to build and develop machine learning models. Secondly, I want pursue my career as a Full stack JavaScript Developer after the completion of master’s course. This project experience will add more value to my portfolio and it adds to my learning curve as well.

**Literature Review:**

Despite the importance of user interface profiling and prediction for general Web applications, only a few research efforts have focused extensively on these topics and most of the current studies have to tackle their own limitations.

With the development of Network architectures and protocols for dynamic and complex Web applications, there have been several attempts to enhance Web QoE (Quality of Experience);

In fact, as one of the most widely used metrics for Web QoE perception (Quality of Experience) is latency, which affects the page load time [15], there were several efforts made to prefetch Web content utilising user interaction information for better Web QoE.

There are many machine learning models, which were developed in the past for predicting the user interaction on various platforms including web, mobile and desktop applications. I have studied and understood the results of the following Machine Learning Models - GRU (Gated Recurrent Unit) [4], RNN (Recurrent Neural Network) [2, 5], LSTM (Long short-term memory) [3, 6], SVM (Support Vector Machine) [7], LR (Learning Rate) [2,6] and Web-profiler [1].

Suykens et al. (1999) first proposed SVM matched filtering method which uses least squares support vector machine classifiers. Due to equality type constraints in the formulation, the solution follows quadratic programming for classical prediction. Expected links of the web pages are mentioned in source code of the websites by developers in advance, which are typically generated by using heuristic rules [3, 11] based on the semantic information of document mark-up, along with the navigation history, and user-dependent data (e.g., device model, screen resolution, used and available memory resources, Browser used, network connectivity, user settings).

To design the navigation prediction model, gated recurrent unit (GRU) [6] are most commonly adopted, which was proposed by Tang, Duyu, and Ting (2015), one of the variants of the recurrent neural network (RNN), a popular deep learning technique suitable for the handling of Web interaction data in time series.

Most of the previous authors have optimized the performance of GRU-based predictions by analysing various hyper-parameters [8, 10] and in other machine learning models like Web Profiler, the authors implemented Web embedding [1,6] and URL grouping [1] to increase the performance.

# Scope, Objectives, and Approach

## Scope

* Understand Machine Learning (ML).
* Critically evaluate User Interaction prediction.
* Analyse how web applications can benefit from User Interaction Prediction.
* Investigate the drawbacks and issues of currently existing models.
* Implement practical work and try to develop a better ML model.
* Analyse practical research work outcomes and suggest recommendations for future work

## Objectives

* Identify why user interaction prediction is required and how it improves the overall user experience
* Evaluate critically the machine learning models and frameworks relevant to supporting user interaction prediction in coping with user experience improvement.
* Examine the issues in currently existing machine learning models
* Formulate recommendations upon practical implementation work

## Approach or Methodology

The Approach to be adapted is to investigate the existing literature and evaluate the proposed solutions by conducting quantitative analysis and focus on common numerical measurement factors like Accuracy or Time taken. Various literature sources including reports, books, conference proceedings, journals, and particularly library databases like ACM Digital Library will be used for the purpose of literature review. (Dudovskiy, 2011)

* Evaluation of existing machine learning models
  + Existing papers, journals, books, conference papers etc
  + Database libraries like Northumbria University Library portal, ACM Digital Library
* Analysis
  + Qualitative data analysis (Considering the common metrics/measurement factors like Accuracy, Time Taken)
  + Practical implementation possibilities
* Practical Work
  + Conduct practical research work based upon the previously done literature review analysis
  + Collect the observed results, analyse, conclude and formulate recommendations

# Timelines and Timescales

## Project Break-down structure

Project Break-down structure

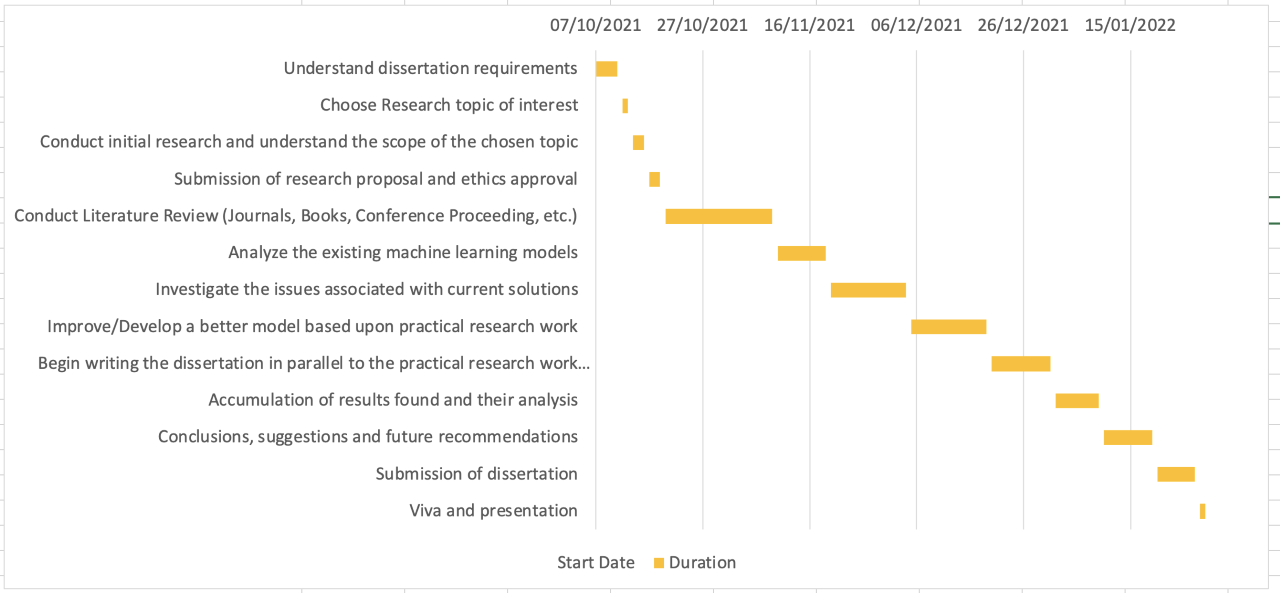
## Task schedule plan

Bellow table show overall research activities schedule and duration.

|  |  |  |  |
| --- | --- | --- | --- |
| **S. No.** | **Activity** | **Start Date** | **No. of days** |
| 1 | Understand dissertation requirements | 10/07/2021 | 5 |
| 2 | Choose Research topic of interest | 10/12/2021 | 2 |
| 3 | Conduct initial research and understand the scope of the chosen topic | 10/14/2021 | 3 |
| 4 | Submission of research proposal and ethics approval | 10/17/2021 | 3 |
| 5 | Conduct Literature Review (Journals, Books, Conference Proceeding, etc.) | 10/20/2021 | 20 |
| 6 | Analyze the existing machine learning models | 11/10/2021 | 10 |
| 7 | Investigate the issues associated with current solutions | 11/20/2021 | 16 |
| 8 | Improve/Develop a better model based upon practical research work | 12/05/2021 | 15 |
| 9 | Begin writing the dissertation in parallel to the practical research work happening | 12/20/2021 | 11 |
| 10 | Accumulation of results found and their analysis | 01/01/2022 | 9 |
| 11 | Conclusions, suggestions and future recommendations | 01/10/2022 | 10 |
| 12 | Submission of dissertation | 01/20/2022 | 1 |
| 13 | Viva and presentation | Around 2 weeks after submission | 1 |
|  |

Table 1: Task schedule plan

## Gantt Chart



# Ethics, Legal, Social, Security and Professional Issues

## Ethical issues

The practical research work involved in this dissertation is completely software-based and will not have any human/animal indulgence in it as no human participation is required. This does not create any harmful or dangerous impact on any person and will strictly follow the ethical rules

## Security issues

The security issues are less to none due to lack of usage of any user’s personal information or confidential data. Dummy data will be used for practical research work which does not require any security measures

## Legal issues

This project is completely legal as all the web information required for this project will be created by my-self and/or used from open-source websites (like, Wikipedia). Any kind of data, including information, images, videos and other work will not be taken from any commercial websites or organizations. The project will be adhered to the legal laws of the government and follows the university rules.

## Professional issues

I commit to be professional throughout this project and follow the design patterns, programming standards, and data security while handling the data. I understand that I’m responsible for this whole project and work towards my goals honestly and sincerely

## Social issues

All the data or information to be used in this project will be created by me. No real user’s data will be used for user-interaction prediction and does not have any kind of human involvement. Also, this research is environment friendly as it does not impact the environment in any way as all the research is computer-based.

# References

* 1. M. Joo and W. Lee, "WebProfiler: User Interaction Prediction Framework for Web Applications," in IEEE Access, vol. 7, pp. 154946-154958, 2019.
  2. I. Grigorik. Resource hints. W3C Working Draft. Accessed: Jul. 2, 2019. [Online]. Available: https://www.w3.org/TR/resource-hints/
  3. R. B. Mobasher, J. Srivastava, and R. Cooley, ‘‘Automatic personalization based on Web usage mining,’’ Commun. ACM, vol. 43, no. 8, pp. 142–151, Aug. 2000
  4. GG. Wang, T. Konolige, C. Wilson, X. Wang, H. Zheng, and B. Y. Zhao, ‘‘You are how you click: Clickstream analysis for sybil detection,’’ in Proc. SEC, Washington, DC, USA, 2013, pp. 241–255.
  5. C. Li, Y. Lu, Q. Mei, D. Wang, and S. Pandey, ‘‘Click-through prediction for advertising in Twitter timeline,’’ in Proc. KDD, Sydney, NSW, Australia, 2015, pp. 1959–1968.
  6. H. Obendorf, H. Weinreich, E. Herder, and M. Mayer, ‘‘Web page revisitation revisited: Implications of a long-term click-stream study of browser usage,’’ in Proc. CHI, San Jose, CA, USA, 2007, pp. 597–606.
  7. Y. Li, ‘‘Reflection: Enabling event prediction as an on-device service for mobile I. interaction,’’ in Proc. UIST, Honolulu, HI, USA, 2014, pp. 689–698
  8. S. Lee, R. Ha, and H. Cha, ‘‘Click sequence prediction in Android mobile applications,’’ IEEE Trans. Human-Mach. Syst., vol. 49, no. 3, pp. 278–289, Jun. 2019.
  9. F. Guo, C. Liu, A. Kannan, T. Minka, M. Taylor, Y.-M. Wang, and C. Faloutsos, ‘‘Click chain model in Web search,’’ in Proc. WWW, Madrid, Spain, 2009, pp. 11–20
  10. S. Jain et al., “Applying big data technologies to manage QoS in an SDN” in *Network and Service Management (CNSM) 12th International Conference,* Montreal Canada, Conf., 2016, pp. 302-306.
  11. T. Cheng et al., “QoS-guaranteed controller placement in SDN” in *Global*

*Communications Conference (GLOBECOM), IEEE,* San Diego USA, Conf., 2015, pp. 1-6.

* 1. Rabinovich, M. and Spatscheck, O., 2002. *Web caching and replication* (Vol. 67). Boston, USA: Addison-Wesley.
  2. Trnka, K., McCaw, J., Yarrington, D., McCoy, K.F. and Pennington, C., 2009. User interaction with word prediction: The effects of prediction quality. *ACM Transactions on Accessible Computing (TACCESS)*, *1*(3), pp.1-34.