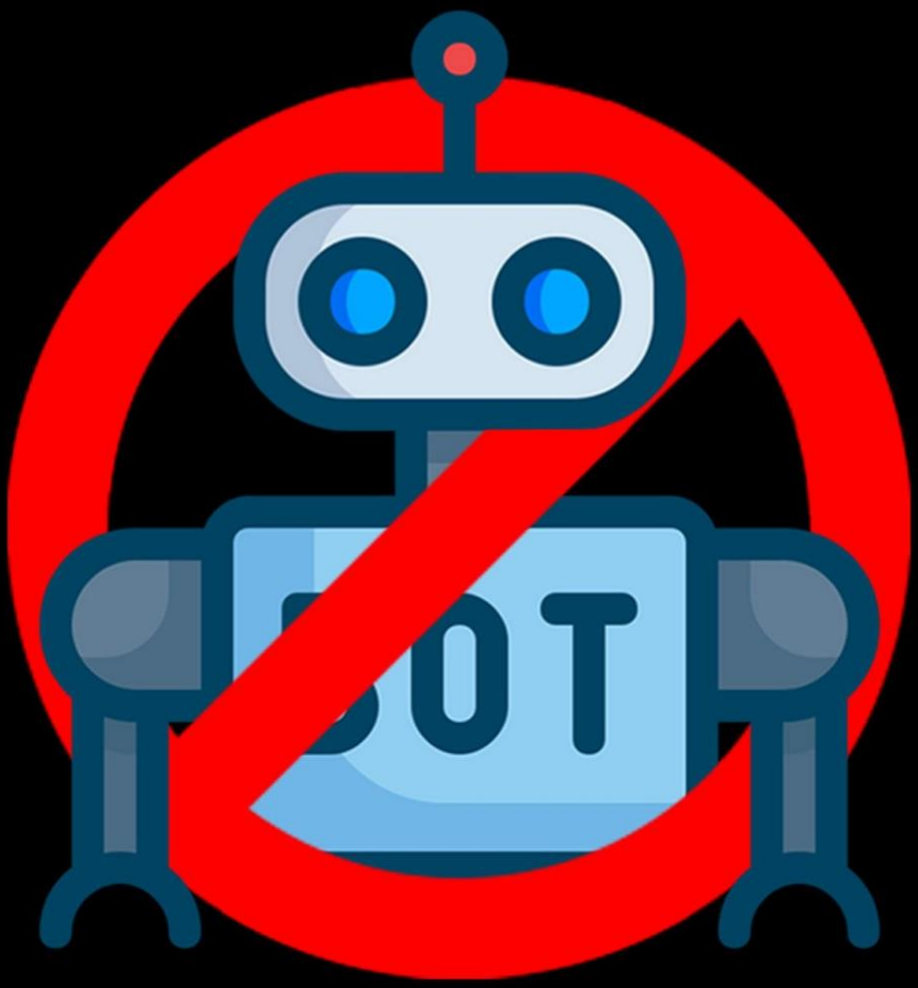
41004 AI/ANALYTICS CAPSTONE PROJECT - AUTUMN 2021

ASSIGNMENT 1: PLAN & PROPOSAL



GROUP 35:

ASHLEY NGUYEN DENZEL MOK

MITCHELL VALENTINUS PATRICK SONGCO

BOT BUSTERS

PROJECT 24: EXPLAINABLE BOT ACCOUNTS DETECTION IN SOCIAL NETWORKS WITH GRAPH MINING TECHNIQUES

SUPERVISOR: DR. HONGXU CHEN

Table of Contents

1. Bot Busters

1.1. About Our Company

As companies of all sizes begin to recognise the tremendous potential for data, many are struggling to draw actionable insights that improve business results. Founded in 2021, Bot Busters set out with the goal to solve the biggest problem of the century. We are excited to simplify and provide data analytics for everyone and to be a leader in bringing innovation and change through our expertise in data analytics.

1.2. Mission Statement

Our mission is to provide leading edge technology and analytics solutions to empower every person and organisation for success.

2. Our Team

2.1. Meet the Team

2.1.1. Ashley Nguyen

Ashley is currently doing a double major in Data Analytics and Enterprise System Development at University of Technology Sydney. Ashley is competent in data collection and analysis as well as conducting detailed and coherent reports, and she is also improving her skills with other fields of data science. Ashley hopes to bring her skill sets and enthusiasm about the study of data as well as the innovations in Machine Learning to this project as the Project Manager.

2.1.2. Denzel Mok

Currently completing a double degree majoring in Data Analytics and Accounting at University of Technology Sydney. Denzel is passionate about technology and innovation, hoping to combine knowledge from both disciplines to draw actionable insights to improve business processes.

2.1.3. Mitchell Valentinus Liauw

Mitchell is a data analytics student at University of Technology Sydney. Eager, team-spirited, adaptive individual who likes challenges and enthusiastic to always learn more. Highly motivated to gain and enhance his skills within the industry and bring a positive impact on the company and organization he works for. He has a great interest in the data analytics field, but also currently improving his skills in the scientific aspect of data.

Mitchell hopes to enrich his career in the data analytics field in the future, and will be the Lead Data Gatherer/Supervisor for this project.

2.1.4. Patrick Songco

Patrick is currently studying a Computer Science degree at University of Technology Sydney. Passionate about the latest innovations in Machine Learning and Deep Learning, Patrick will bring this past knowledge and experience to help implement the graph-based neural network model in this project.

2.2. Our Roles

The structure of our team will have a strong focus on flexibility and fluidity, with shared and overlapping roles and responsibilities, however due to the nature of the project, well-defined and specific roles will be assigned to members to facilitate and ensure that tasks can be completed in a timely and efficient manner. The roles and responsibilities will be assessed and revised in weekly meetings to reflect on the current progress and tasks which need to be completed. Although assigned specific roles, members are by no means limited to their stated roles and will provide support and assistance to others as required.

Ashley will serve as the team manager of this project, primarily responsible for managing and overseeing the overall project progress as well as directing discussion between key stakeholders and team members to meet the project requirements and deadlines. To ensure that everyone understands and can meet the objectives of the project, her main duties will include developing and implementing realistic and practical targets and guidelines as well as establishing clear communication between members of the team.

The data collection responsibilities will be shared by all members, however Mitchell will supervise and oversee the process to ensure that the quality and integrity of the data collected is relevant and of an appropriate standard to the project. Furthermore, his duties will involve collating and sorting the relevant data to be compiled and used for further analysis, visualisation and modelling. The accuracy and precision of the data is a critical aspect of this project and will serve an important function in this project, requiring careful and meticulous scrutiny to provide the best results.

Denzel will be responsible for managing and creating data visualisations of the project, working together with the supervisor of data analysis to understand and convey findings to other members and key stakeholders through developing simple and comprehensible visualisations.

The data analysis portion of the project will be led and directed by Patrick. He will be responsible for analysing and drawing insights from the data to form conclusions and machine learning models that can solve and address the data mining problem of this project.

3. The Project

3.1. Our Client

Although the first social media site, Six Degrees, was created in 1997, online social networks (OSNs) have only started to become globally popular around the mid-2000s. In recent years, OSNs such as Facebook, Twitter and Instagram have become increasingly popular, as

Emarsys's study in 2019 points out that there are approximately 3.5 billion active users on social media daily, accounting for roughly 45% of the current population, on which an average user spends three hours per day (Globalwebindex, 2019). However, due to the immense number of users, the convenience and the highly engagingly outreach nature of OSNs, their users have unfortunately been subjected to pernicious conducts such as public opinion manipulation, unwanted advertising, malicious links and fake news, all of which was generated via the use of anomaly accounts (or also defined in this project as social bots or spambots).

A group of these social bots forms a social botnet that is managed and programmed to act in an organised manner (Latah, 2020), with an alarmingly large amount of which is used for spreading malicious contents. Although efforts have been put into the study and implementation of bot detections, spambots have become more advanced and complex, thus making them more difficult to detect and prevent. As Cresci *et al*. (2017) presented in their analysis in 2018, 71% of retweeting users on Twitter were classified as bots, and to the point of the study, only 37% of which had been suspended.

This project will address the issue of social media bots and propose an appropriate solution to tackle and improve the current bot detection efforts. Since this project serves as a practical research study, it is not aimed at a specific client.

3.2. How We Can Help

The goal of this project is to analyse current progress in social bot detections, visualise and study the possible application of graph-based algorithms to the detection procedure, as well as to propose a novel methodology in explainable social bot detections, as well as the initial trial implementation of this methodology to the Twitter bot datasets. The reason we choose Twitter is that it is the third most popular worldwide OSN with a rather straightforward API (Antonakaki, Fragopoulou, and Loannidis, 2021) that can be efficient and scalable in a structure implementation (Kantrowitz, 2018; Kwak *et al*., 2010). We believe, through this research project, we will be able to provide a comprehensive view on the current progress of bot detection efforts as well as a suggestion on the application of graph-mining techniques as a bot detection method.

4. The Data Mining Problem

The data used for this research will be collected from online academic sources as well as crawled from Twitter itself. As we will be using Twitter datasets for this project, there are some possible issues that might occur during the process. For instance, there might be some missing data regarding followers and following lists (which are required to build the nodes and edges for graph-mining methods), text processing (for tweets data), as well as some datasets found online might have been outdated, especially with suspended accounts. We anticipate that these issues can be addressed and resolved in the data pre-processing phase.

This data will require further process and analytics using data mining techniques such as text mining, graph mining, algorithms and visualisation. Noticeably, graph-based metrics can also be applied to the bot detection process by using users as nodes and their connections (followers - following lists) as edges. Useful information will be extracted from the result of

these techniques for model building and training. Some of these graph features could be the clustering coefficient or the betweenness centrality, which are metrics often used for graph analysis (Kouvela, 2020).

The DARPA Twitter Bot Challenge result in 2016 suggested that the bot-detection system needs to be semi-supervised, meaning that the accuracy of the bot detection process requires human judgements (Latah, 2020). Therefore, in order to improve the accuracy for further applied models, we will explore and attempt some novel graph-based methodologies to detect and explain the bot accounts and their behaviours as they will provide a better understanding and easier visualisation to assist the decision of bot identification and suspension.

5. Project Proposal

We anticipate that the Twitter dataset will be the primary data for our research project proposal. The dataset includes extensive and detailed data for each Twitter user and their attributes, such as user ID, date of user creation, number of followings, number of followers, number of tweets, length of screen name, length of description in user profile, and their corresponding tweets. As our research project is based on bot detections, we believe the data from Twitter API itself would be the most appropriate and reliable enough for us to serve as our primary data source. Throughout our research project, we will continue to assess additional data sources or scoping into smaller datasets based on Twitter platform, and integrate them if needed and would provide additional values to our research.

We anticipate that we will be using Python as our primary coding language for data modelling and analysis. TensorFlow will also be needed to run the machine learning aspect of our research project.

We will also require Gephi, NetworkX, and other additional data visualization tools to illustrate our findings and show the insights of our research.

For our research project, we will analyse each of the Twitter users attributes, such as their length of screen name and length of description on their profiles. The Tweets that are being put out from each user will also be one of the parameters in our research as we can learn the type of contents the organic and the non-organic Twitter users generate. The correlation between the accounts will also be our area of research as we believe the edges based on the following and followers network will be able to provide a resourceful insight for our project. These are all the primary interesting areas that we are focusing on and intend to explore them fully to enrich our research.

Our research project strives to get a functional bot detection training model that can be used and allow our clients to support their cause and generate insightful content for their publications.

6. Project Plan

6.1. CRISP-DM Methodology

To achieve the goals of this project, we intend to use the robust and well-proven Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology as it provides a structured approach to planning the data mining project. We intend to complete tasks in the phases below (Figure 1). It should be noted that ideally, the process will be carried out in sequential order, although, in practice, many of the tasks can be completed in a different order and it will often be necessary to backtrack to previous tasks and repeat certain actions.

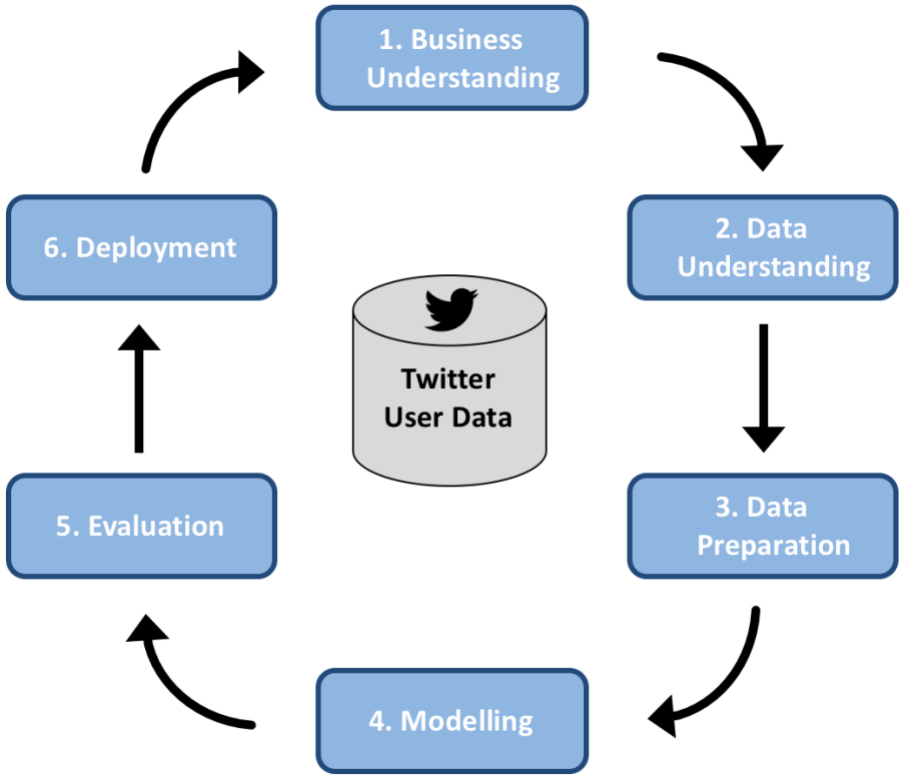


Figure 1: CRISP-DM Methodology

6.1.1. Business Understanding

During this first phase, we focus on the project objectives and requirements set out by our project supervisor. The goal of this stage is to uncover important factors that could influence the outcome of the project.

This stage is currently being completed by the team. After rounds of planning and preliminary data gathering, our objective for the project is to implement an explainable graph-based model for detecting malicious bots in Twitter social network data. Additionally, a visualization of the graph network will be completed.

Implementation of the model with an accuracy of malicious bot detection of 65% or greater will be considered the criteria for success in this project. This criterion was chosen because

current state of the art graph-based malicious bot detection models achieve comparable results.

6.1.2. Data Understanding

This second stage involves the collection and deeper exploration of the data. It is an important phase as it dictates what inputs will be available for the model.

Preliminary data gathering tasks in the Business Understanding phase have led us to the conclusion that although Twitter user data is widely available, the available datasets do not contain social network information such as user-follower relations. Furthermore, to our knowledge, labelled malicious bot Twitter social network data is scarce/unavailable. To this end, we attempt to data mine or collect the Twitter data ourselves using the Twitter API. This requires developer access and the request for this access is currently pending. An alternative to this is to use Twitter scraping tools available here.

Next, we will use the method proposed by Kouvela *et al*. (2020) to extract labels for each Twitter ID to determine whether they are a bot or a human account. This requires use of the Botometer service that provides a score for each Twitter account. A higher score means more bot-like activity. This will allow us to train the model in a semi-supervised/supervised fashion.

Finally, a deeper analysis of the complete dataset will be completed to ensure the feasibility of using the dataset for our project.

6.1.3. Data Preparation

In this phase, we will assess how much of the collected dataset will be used for analysis. The criteria that will be assessed will be the relevance of the data to our data mining problem, the quality of the data and the technical constraints such as available computation. Once chosen, the data will be cleaned, transformed and pre-processed as required for input into the model.

6.1.4. Modelling

The modelling phase involves selecting the modelling technique that we will use to analyse the dataset. While it may be too early to predict exactly what model/approach will be used to solve the problem, we have gauged the feasibility of various models in the Business Understanding phase and have determined that a graph-based neural network architecture is a likely candidate. A more concrete methodology will be provided in the mid-project update. This phase will involve:

- Generating test designs of the model
- Building the model
- Training the model
- Testing the model

6.1.5. Evaluation

In this phase we evaluate the results of the model in terms of the business success criteria. We assess whether changes need to be made in the previous phases and action them accordingly. Once the success criteria is met, we will determine the next steps for deployment.

6.1.6. Deployment

The final phase, deployment, may not be applicable to our project since this project serves as a practical research study and is not aimed at a specific client. Therefore, deployment will be defined as finalizing the model, the final report and the presentation for submission.

6.2. Work Plan

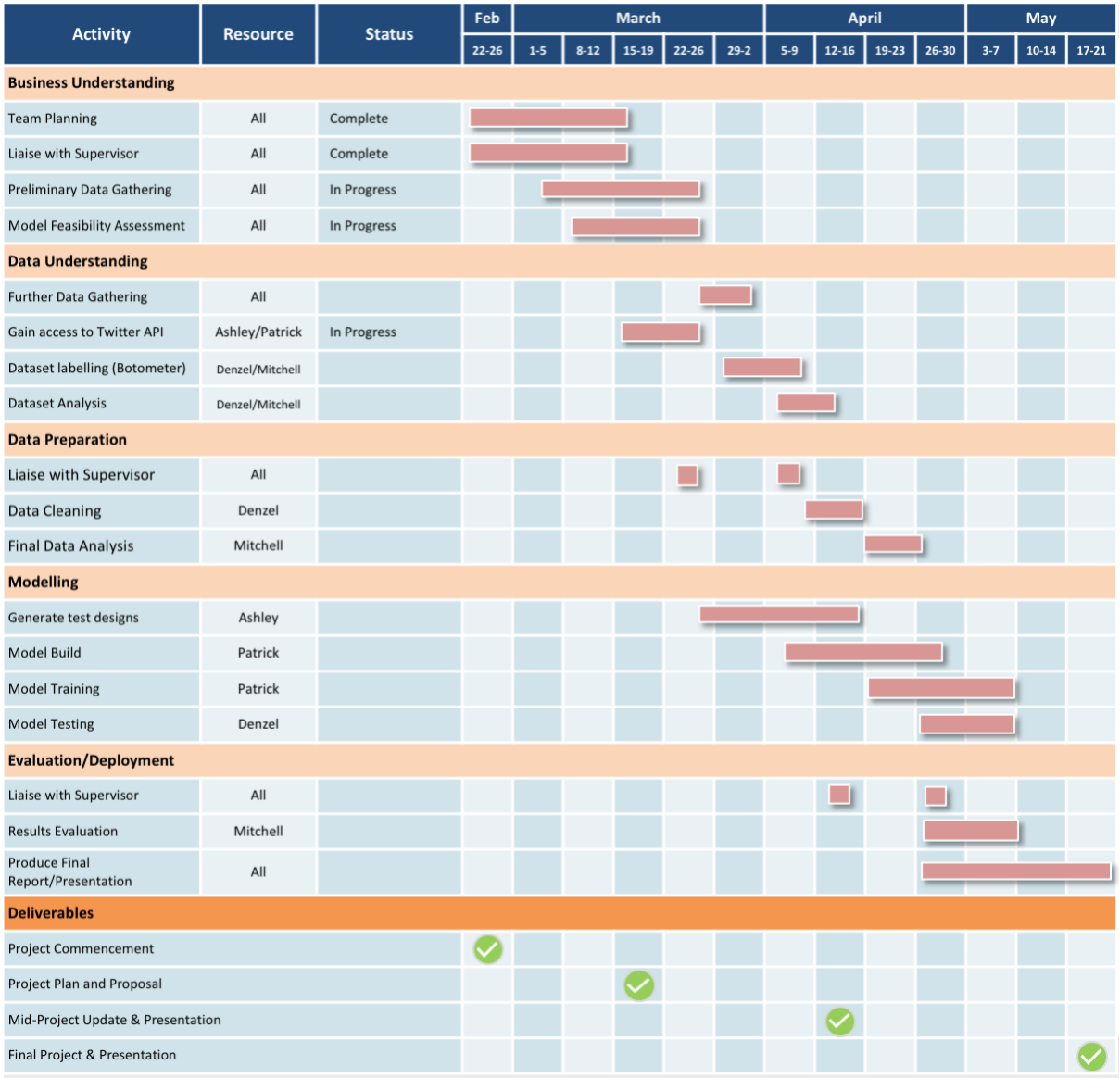


Figure 2: Gantt Chart of Project

A preliminary project plan (Figure 2 above) was created to track the progress of the project. It outlines the individual duties of each team member, projected timelines, key dates and project milestones.

6.3. Milestones

There will be four milestones/project deliverables throughout the duration of this project. These milestones track the progression of the project and provide an update to our stakeholders.

6.3.1. Project Commencement

22nd February 2021
This marks the start date of the project. During this time the group was formed, supervisor was introduced, and the project scope was identified

6.3.2. Plan Proposal

19th March 2021
This is the first submission of the project. It outlines the project goals, how we plan to achieve these goals, and the anticipated timeframe for completing the project. It also allows the team to receive feedback from our supervisor on the proposed methods and adjust accordingly.

6.3.3. Mid-project Update

16th April 2021
The Mid-Project update and presentation communicates to our stakeholders the current state of the project. It allows the team to assess the progression towards the project goals and raise any concerns or unforeseen complications. Additionally, it provides the team and our supervisor a chance to change any project requirements if necessary.

6.3.4. Final Report

21st May 2021
The final deliverable will be the completed report and presentation. The Final Report will outline the entire project from start to finish. It will provide our stakeholders visibility around the project goals, model evaluation metrics and a discussion about the achievements/limitations of the project.

7. References

Antonakaki, D., Fragopoulou, P. and Ioannidis, S. (2021). A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks. Expert Systems with Applications, 164, p.114006.

Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A. and Tesconi, M. (2017). Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling. IEEE Transactions on Dependable and Secure Computing, 15(4), pp.561-576.

Emarsys. (2021). *Top 5 Social Media Predictions for 2019* | *Emarsys*. [online] Available at: <https://emarsys.com/learn/blog/top-5-social-media-predictions-2019/> [Accessed 16 March 2021].

GWI. (2021). *2019 in Review: Social Media is Changing, and It's Not a Bad Thing. - GWI*. [online] Available at: <https://blog.globalwebindex.com/trends/2019-in-review-social-media/> [Accessed 16 March 2021].

Kantrowitz, A. (2018). How Twitter Made the Tech World's Most Unlikely Comeback. BuzzFeed News.

Kouvela, M., Dimitriadis, I., & Vakali, A. (2020). Bot-Detective: An explainable Twitter bot detection service with crowdsourcing functionalities. In Proceedings of the 12th International Conference on Management of Digital Ecosystems (MEDES '20). Association for Computing Machinery, New York, NY, USA, 55-63. https://doi.org/10.1145/3415958.3444075

Kwak, H., Lee, C., Park, H. and Moon, S. (2010). What is Twitter, a social network or a news media?. In Proceedings of the 19th international conference on World wide web (pp. 591-600).

Latah, M., (2020). Detection of malicious social bots: A survey and a refined taxonomy. *Expert Systems with Applications*, *151*, p.113383.