Aalto University
School of Business

DETECTING AND ANALYZING BOTS ON FINNISH POLITICAL TWITTER

Master’s Thesis
Sippo Rossi
Aalto University School of Business
Information and Service Management
Spring 2019

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Sippo Rossi		
Title of thesis Detecting and analyzing bots on Finnish political twitter		
Degree Master of Science in Economics and Business Administration		
Degree programme Information and Service Management		
Thesis advisor(s) Bikesh raj Upreti, Yong Liu		
Year of approval 2019	**Number of pages** 56	**Language** English

Abstract

This master's thesis develops a machine learning model for detecting Twitter bots and applying the model to assess if bots were used to influence the 2019 Finnish parliamentary election. The aim of the thesis is to contribute to the growing information systems science literature on the use of social media and information systems to influence voters as well as to increase the general awareness in Finland of the effects of bots on Twitter.

The thesis relies primarily on quantitative analysis of a dataset consisting of 550,000 unique Twitter accounts. The data was collected from Twitter during March 2019. The accounts in the dataset belong to humans and bots that were following 14 prominent Finnish politicians on Twitter. To determine which accounts are bots and to assess the feasibility of a new method for Twitter bot detection, a machine learning model that utilizes metadata-based features for classifying Twitter accounts as bots or humans is developed and tested on the dataset.

The findings of this thesis indicate that a metadata-based approach is suitable for detecting bots and that there are several large botnets in the Finnish Twittersphere. Over 30% of the 550,000 accounts are labeled as bots by the model, which implies that the prevalence of bots is much higher than previously suggested by Twitter's official estimates. Furthermore, a majority of the accounts seem inactive and either no longer being used or dormant and waiting for activation. The purpose of most of the bot accounts is obscure, and it is not certain how many of them are following and inflating the politicians' popularity on purpose. Although the bots clearly increase the visibility of certain politicians, the effects of the bots on Finnish political Twitter are deemed negligible.

Keywords Twitter, botnet, bot detection, social bots, political big data, network analysis

Aalto-yliopisto
Kauppakorkeakoulu

Aalto-yliopisto, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Maisterintutkinnon tutkielman tiivistelmä

Tekijä Sippo Rossi		
Työn nimi Detecting and analyzing bots on Finnish political twitter		
Tutkinto Kauppatieteiden maisteri		
Koulutusohjelma Tieto- ja palvelujohtaminen		
Työn ohjaaja(t) Bikesh raj Upreti, Yong Liu		
Hyväksymisvuosi 2019	**Sivumäärä** 56	**Kieli** Englanti

Tiivistelmä

Tässä pro gradu -tutkielmassa keskitytään kehittämään koneoppimismallia Twitter-bottien havaitsemiseen ja tutkimaan mallilla pyrittiinkö koneella luotujen tilien avulla vaikuttamaan vuoden 2019 eduskuntavaalien tuloksiin. Tutkielman tarkoituksena on myötävaikuttaa kasvavaan tietojärjestelmätieteen tutkimukseen, jossa käsitellään sosiaalisen median ja tietojärjestelmien hyödyntämistä äänestäjiin vaikuttamisessa. Lisäksi tavoitteena on lisätä yleistä tietämystä bottien vaikutuksesta Twitterissä.

Tutkielman keskiössä on 550,000 Twitter-tilistä koostuvan datasetin analysointi kvantitatiivisin menetelmin. Tutkimuksessa käytetty data louhittiin Twitteristä vuoden 2019 maaliskuun aikana. Datasetti koostuu 14 tunnetun suomalaisen poliitikon Twitter-seuraajista, johon kuuluu sekä botteja, että ihmisiä. Jotta datasetistä saadaan määritettyä mitkä tilit kuuluvat boteille ja mahdollistaakseen uuden bottienhavaitsemismenetelmän toimivuuden arvioimisen, tutkielmassa kehitetään koneoppimiseen perustuva malli, joka käyttää Twitterin metadataa määrittääkseen onko Twitter-tili ihminen vai botti.

Tutkielman löydökset viittaavat, että metadataan perustuva malli kykenee tunnistamaan botit ja, että suomenkielisestä Twitteristä löytyy useita suuria bottiverkkoja. Mallin mukaan yli 30% datasetin 550,000 tileistä kuuluvat boteille, mikä viittaa bottien määrän olevan merkittävästi suurempi kuin mitä Twitter on virallisesti arvioinut. Lisäksi suurin osa näistä tileistä on epäaktiivisia mikä viittaisi niiden jo täyttäneen tarkoituksensa tai yhä odottavan aktivointia. Botti-tilien tarkoitus jää suurimmassa osassa tapauksista epäselväksi ja täten on vaikea määrittää mikä osuus niistä on tietoisesti luotu kasvattamaan poliitikkojen seuraajien määrä. Siitä huolimatta, että botit selvästi vahvistavat tiettyjen poliitikkojen näkyvyyttä, bottien olemassaolon vaikutus Suomen poliittiseen ympäristöön Twitterissä arvioidaan vähäiseksi.

Avainsanat Twitter, bottiverkko, bottien havaitseminen, sosiaaliset botit, poliittinen massadata, verkostoanalyysi

Acknowledgements

First, I would like to thank my primary thesis advisor Bikesh raj Upreti. Thanks to Bikesh's encouragement, I ended up picking this challenging topic. Along the way he taught me a lot about conducting quantitative research as well as machine learning and ultimately helped me achieve something I would have not even dreamt of being able to do before.

I would also like to thank Yong Liu for acting as the second advisor and for challenging me at the right times to push the quality and rigor of the thesis further.

A big thanks goes to Andy Patel, whose research and online guides on Twitter bots inspired me to choose this topic. Furthermore, his comments and suggestions helped shape the findings of the thesis.

Additionally, I would like to thank my academic mentor Nenad Jukić for providing feedback and support.

I would like to also thank my father Matti for his guidance and comments as well as my mother Tuuli for her support during the process.

Last but not least, I would like to acknowledge the people at the slack group "Suhina" for giving tips on machine learning and feature engineering as well as for introducing me to Andy Patel.

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Background and motivation

Governments, organizations and individual people are actively attempting to persuade or influence the users of social media by spreading propaganda with large networks of bot accounts (Twitter, 2018; Nickerson & Rogers, 2014). There are multiple examples of bots being used to distort political discussions on Twitter in the past few years according to academic research, government institutions and the social media site itself. One of the most notable cases is the 2016 US presidential election, where an organization linked to the Russian government has been accused of attempting to influence the elections by using Twitter accounts to spread fake news or otherwise biased content (Bessi and Ferrara, 2016; Twitter, 2018).

The number of academic papers investigating Twitter botnets is growing rapidly. Recent publications have found evidence of bots being used to influence opinions in countries such the United States (Bessi & Ferrara, 2016), Japan (Schäfer, Evert & Heinrich, 2017), Brazil (Salge & Karahanna, 2018) and Russia (Stukal et al, 2017). Similar studies have not been conducted in Finland but there is already evidence of at least one large but inactive Finnish Twitter botnet according to a researcher at F-Secure (Gallagher, 2018; Patel, 2018). Since the 2019 Finnish parliamentary election is approaching, further research in this timely topic can bear interesting results.

The aim of the thesis is to investigate how bots are being used to influence Finnish politics on Twitter. This thesis contributes to the growing information systems science and political data science literature on the use of bots and information systems to influence voters. Furthermore, if an active botnet is identified and the results are distributed for example, through an online article or post, the thesis will also have a practical contribution as it can potentially increase the awareness of Finnish Twitter users.

1.2 Objectives and scope

1.2.1 Research problem and questions

The primary objective and research problem of this thesis is to find evidence of a Finnish botnet operating in Twitter and to determine whether it is used to influence or blur political discourse in Finland by developing and implementing a new bot detection model. The research questions and sub-questions are the following:

- Is there a Twitter botnet that can be found from analyzing the followers of popular Finnish politicians?
    - Do the bots tend to follow accounts that support a certain party or ideology?
    - What is the structure of the network?
- Can a metadata-based bot detection model classify accounts with sufficient accuracy when compared to models that utilize both metadata and tweeting behavior?

1.2.2 Scope

The scope of the thesis is limited to the development of a model that utilizes Twitter metadata for bot detection and using it to analyze the number of bots that are following and interacting with Finnish politicians. Additionally, the networks that the bots form will be visualized, and basic network analysis is be applied to examine the structure. The thesis will not evaluate the content distributed by the bots or who or what are the individuals or organizations creating and operating the bots.

1.3 Structure of the thesis

The thesis starts by reviewing relevant research from the past several years on social bots, Twitter bots, Twitter bot detection methods and the use of Twitter for political influencing around the world. How the data was collected, the bot detection model developed, and the botnet analyzed is described in the third chapter. This will be followed by a chapter that describes how the model can be deployed and the results reproduced. Chapter five presents the results and describes the characteristics of the botnet and assesses what kind of an impact they have on Finnish political Twitter. Lastly, the final chapter summarizes the most important findings regarding the use of bots in the Finnish Twittersphere as well as the primary contributions and limitations of the thesis.

2 Literature review

The literature review is divided into three parts. The first part will look into how previous research has classified bots and provides a clear definition for certain ambiguous, but important terms and concepts. The second part analyzes different methods that have been used to detect bots in Twitter related research and will provide a background and benchmarks for the bot detection model proposed in the methodology chapter of the thesis. The third and last part covers a range of literature on the use of bots in political influencing during the recent years to support the findings and assumptions made in the later chapters.

2.1 Terminology and the definition of a bot

The definition of a bot can vary quite significantly in research, news articles and other reports. Currently there is no clear standard for either classifying bots or categorizing different types of bots. Characteristics that are often used in determining whether an account can be labeled a bot or not include the level of automation as well as the method used in creating the account's content. (Grimme et al., 2017)

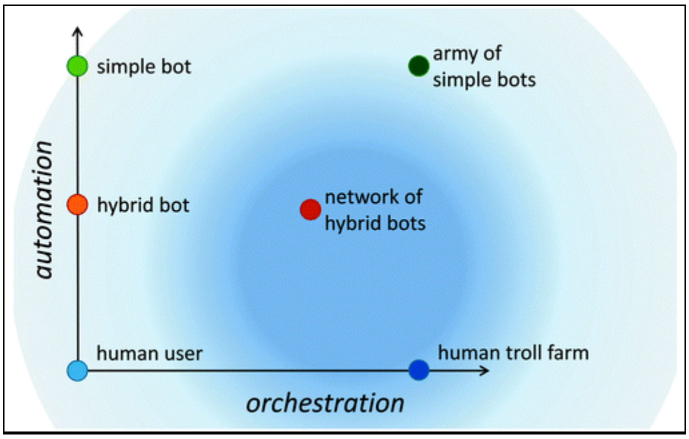


Figure 1. Grimme et al.'s qualitative classification of humans and bots and their potential influence in social media (2017)

Figure 1 provides a crude way to define the difference between a human user and a bot and illustrates how difficult it is to exactly draw a line between the two categories. In its most basic form, a bot is an account that is purely controlled by an unmonitored program. Examples of these simple bots include spambots that automatically spread content based on a script as well as bots belonging to like farms that are used on social media to increase the number of followers of an account or likes of a particular post. However, these are prone to detection and thus deletion. More advanced bots adjust their content dynamically based on the behavior of other accounts making them more difficult to detect for both humans and bot detection algorithms even if the bot is still operated solely by a program. The most sophisticated bots are such that humans control parts of their activities, such as content creation, which blurs the line between the bot and human user. When properly operated, these hybrid bots are almost invisible to automatic detection mechanisms according to Grimme et al. (2017).

2.1.1 Twitter bots

On Twitter, bots can be divided into two types: benign and malicious (Oentaryo et al., 2016; Chu et al., 2012). Both of these types can include bots ranging from simple content sharing accounts to human-like social bots that participate in discussions and create original content. The primary difference is that benign bots adhere to the social media site’s rules and guidelines and are clearly distinguishable from human accounts usually by name and or a description. Conversely, malicious bots participate in activities that are not permitted by Twitter and rarely disclose the fact that they are operated by a program. Typical use cases include artificially boosting the number of followers, likes or retweets and directing or blurring discussions as well as spreading spam or content that supports a certain cause.

A relevant subtype of Twitter bots are political bots. As the name implies, a political bot is specifically designed to participate in political discourse and to promote a certain ideology, organization or individual (Woolley 2016). In the context of political influencing, in most cases it can be assumed that a bot will either not have any references to it being a bot and it may even attempt to mimic human behavior in order to avoid detection and to more efficiently influence other users.

2.1.2 Social bots

One of the most prevalent terms in social media related bot research is the term social bot (Grimme et al., 2017; Ferrara et al, 2016; Woolley, 2016). There are various and occasionally contradictory definitions for social bots (Grimme et al., 2017). The phrase most commonly refers to a bot that is meant to mimic human behavior (Woolley, 2016) and to communicate and interact with human users (Davis et al., 2016). In contrast, a regular bot operates in a mechanical manner without dynamically producing content or engaging in discussions. Similar to the practice with Twitter bots, social bots also are classified as either benign or malicious.

2.1.3 The definition of bots in this thesis

There are several reasons why it is important to have a specific definition for what can be counted as a bot and which types of bots a piece of research focuses on. First, a bot's type can determine whether it is relevant for answering the research questions. Secondly, the bot detection algorithm described in the methodology chapter will rely on a set of assumptions when classifying accounts and these assumptions are based on what is the definition of a bot.

In this thesis, a bot is defined as account that is either fully or partially operated by a program and thus, at least some parts of the account's activities are automated. This includes inactive accounts, also known as sleeper bots (Woolley & Howard, 2017), which can be assumed to have been created automatically and that are followers of politicians or have liked political content. Therefore, the bots can also be classified as political bots. Furthermore, in most cases the term social bot may also applicable, but for the sake of clarity and brevity the word social is omitted. In the later sections the word bot will refer to the type of bots that are described in this paragraph.

2.2 Detecting bots on Twitter

In this section, several recent bot detection models and their features are discussed and evaluated based on their applicability as a basis or benchmark for the model proposed in the methodology part of the thesis.

2.2.1 Simple versus complex models

As bot detection algorithms become more advanced, so do the algorithms that control bots. In literature, the models range from the very simplest ones that are based on analyzing one piece of metadata to those that use ensemble methods and have large feature sets that use a mix of metadata and tweeting behavior and content data.

Surprisingly accurate results can be obtained even with simplistic models and as an example Beskow and Carley (2019) managed to identify bot accounts based on a single piece of metadata, the profile name, with approximately 95%-99% accuracy depending on the algorithm used. This type of an approach results in a very narrow use and the aforementioned model could only detect bot accounts that have an account name consisting of a randomly generated string of 15 characters. However, as Beskow and Carley (2019) propose, a tool-box approach where multiple different models are combined can make even the simple models an important contribution to more advanced bot detection models that combine methods from multiple papers.

On the other end of the spectrum, in terms of sophistication and complexity, are the models that look into various characteristics of accounts combining metadata and behavior features to identify bots (Varol et al., 2017; Minnich et al., 2017). However, one notable issue exists that hinders the reusability of these models. Many of the models rely on some form on natural language processing, sentiment analysis (Davies et al, 2017) and a list of keywords (Minnich et al., 2017; Stukal et al., 2017; Fernquist, Kaati, & Schroeder, 2018), which restricts their use to a particular language and region as well as a time period or event such as an election due to certain themes and hashtags being important only in that specific context. Thus, time and effort have to be put into altering and retraining the model to make it accurate and applicable to research for example focusing on analyzing bot behavior in a country where a different language is used.

2.2.2 Feature space selection

Almost all models that were inspected and cited in this literature review utilize machine learning methods in identifying bots. Therefore, one very essential aspect of designing the bot detection model is determining the optimal feature space. There are two main considerations when selecting the features that are included. First, the features should be added only if they improve the accuracy of bot detection and secondly the features must not make the data collection phase overly time-consuming, since Twitter's API has strict rate limits.

From the models reviewed for this thesis, the most common classes of features used in classification of bots are metadata-based features and tweeting characteristics-based features (Stukal et al., 2017; Fernquist, Kaati, & Schroeder, 2018; Patel, 2018). Each user profile as well as tweet provides a large amount of metadata and while tweets are limited to a certain number of characters (280), the amount of analyses that can be performed on them is vast. Although other classes of features are described and used in Twitter related research (Varol et al., 2017; Davis et al., 2016; Wang et al., 2015), this literature review does not evaluate them as the metadata and tweeting characteristics classes are most relevant for the thesis.

Since metadata can be obtained from both profiles and tweets, metadata-based features can be respectively divided into two different branches. Intuitively, metadata extracted from a profile gives more information on the account, while metadata from tweets gives a combination of information from the profile posting it as well as the tweet itself (Wang et al., 2015). Most models seem to use only one of these, which is most likely to streamline the data mining process and to avoid having to combine two datasets with overlapping data.

Examples of metadata that can be extracted from Twitter include all basic profile information such as name, description and number of friends. Simple and commonly used features include checks on whether a list of different pieces of profile information are blank or at default resulting in binary features such as whether or not the profile picture has been added (Stukal et al., 2017; Patel, 2018), with the more fields that are left at default, the more likely it is that the profile is a bot. Data on the number of users that the profile is following, the number of followers and ratios of these are also often used (Stukal et al., 2017; Fernquist, Kaati, & Schroeder, 2018; Patel, 2018; Wang et al., 2015) and an example of a suspicious profile is such that it has none or only a few followers, but it follows many. Lastly, the contents of the textual metadata can be analyzed and used to classify bots for instance by inspecting the length or frequency of certain keywords in the description or name (Beskow and Carley, 2019; Fernquist, Kaati, & Schroeder, 2018; Patel, 2018).

Features based on tweeting characteristics are more difficult to compare and contrast than those based on metadata due to the variations in approaches used in different models. Furthermore, they include more event or theme related aspects, particularly when assessing the use of keywords that are related to campaigns or social media phenomena (Stukal et al., 2017; Fernquist, Kaati, & Schroeder, 2018). However, a certain group of features that belong to this class can be found in multiple models and these include the number hashtags in tweets, number of URLs in tweets and number of retweets (Stukal et al., 2017; Wang et al., 2015; Chen et al., 2017).

Earlier findings suggest that a combination of both metadata and content features yields optimal results (Fernquist, Kaati, & Schroeder, 2018, Bessi and Ferrara 2016). There are hundreds of different features that can be derived from Twitter's metadata and content data, making it a matter of preference on which ones to choose as various combinations have resulted in highly accurate results. Furthermore, some of the analyses used to create content-based features can be applied to metadata-based features as well to create new ones. Examples of this include counting the number of hashtags, URLs and instances of specified keywords in the name or description of an account.

The model proposed in this thesis will utilize metadata-based features only and therefore, they are examined more thoroughly than content-based and other types of features. Table 1 illustrates some of the features that have been used in the papers mentioned in the literature review (Fernquist et al., 2018, Stukal et al., 2017; Varol et al., 2017; Wang et al., 2015). Unsurprisingly, the most common features are the ones that are directly related to how Twitter functions, with the number of followers, friends, tweets and retweets being examples of these. There are no standard naming conventions and consequently the same features may have slightly different names depending on the author's preferences.

Table 1: Metadata-based feature types and examples

Binary features	**Profile information features**	**Ratio features**	**Metadata content features**
Defaults: - Profile image - Background image - No user description **Other:** - Profile verified - Location specified - No friends - No tweets	**General:** - Number of followers - Number of friends - Number of tweets - Number of likes - Age of account - Account language **Length:** - Profile name - Profile description	**Activity:** - Ratio of following and followers (FE/FI) - Reputation (FE/(FI + FE)) - Given likes per friend - Given likes per follower **Account age:** -Friends/Account age - Following rate (FI/AU)	**Bot check:** - Name contains bot - Description contains bot **Other content:** - Number of # in description - Keywords in description - URL(s) in description

* FI = Number of following, FE = Number of followers, AU = Age of account

In Table 1, the features are grouped into four different types. Some of the most commonly used features are found in the first group that is labeled as binary features. Based on the popularity, it can be assumed that they are appropriate for the detection of bot accounts despite of their simplicity. More specifically, binary features designed to check which profile customization options such as the profile image and background image, are left at default are found in many models (Fernquist et al., 2018, Stukal et al., 2017; Varol et al., 2017). This seems to imply that leaving the profile settings to default and personalization to minimal is a common practice despite of it making the bot easier to detect by humans and machines.

The second group also contains many of the most prevalent features in bot detection models. These are mostly numerical characteristics, many of which are related to how popular a Twitter user is and how actively it is used. Particularly, the number of followers, friends, tweets, retweets and likes are frequently used (Fernquist et al., 2018; Varol et al., 2017; Wang et al., 2015). Another commonly used feature is the length of the description, which cannot be obtained directly from Twitter, but can be calculated easily from the metadata (Stukal et al., 2017; Varol et al., 2017; Wang et al., 2015).

The third group of features are ratios that can be obtained from the same metadata as many of the profile information features of the 2nd group. When compared to the two previous groups, the ratio features contain more variety as they are not based on Twitter's built-in attributes. A common ratio features that can be found in many papers is the followers-to-friends ratio (Stukal et al., 2017; Fernquist et al., 2018; Wang et al., 2015). In the model created by Fernquist, Kaati, and Schroeder (2018) the top 10 features include multiple of ratios, with examples being given likes per friend (#1), followers-friends ratio (#2) and number of likes per followers (#10).

The last group consists of the features that are based on the contents of the different attributes. Features in this group are less commonly used and vary significantly per model. Two of the features in this list simply check whether or not an account is a bot according to the profile description or name (Beskow and Carley, 2019). The rest of the features are more complex, and they are used for example to look for URLS, hashtags or other keywords (Wang et al., 2015) that can be linked to bot accounts based on heuristics or previous research.

Since the performance of each individual feature is not always included in articles related to Twitter bot detection, evaluating the best alternatives requires testing. Models can perform well with both small and large numbers of features (Beskow and Carley, 2019; Fernquist et al., 2018). Since ratio features were proven to be among the best performing features in some cases (Fernquist et al., 2018) and they are widely used (Stukal et al., 2017; Fernquist et al., 2018; Wang et al., 2015), several of them will be included in the model introduced in the methodology chapter. The reason for the ratio features' popularity is probably a result of the fact that they essentially capture more information than the underlying attributes and the possible features created from them. Selecting the other metadata features is more difficult as there is considerable variation by article on what are the best performing ones. Therefore, initially as many as possible should be included and tested and the results reported before removing the redundant features.

2.2.3 Classification methods

Bots have been evolving rapidly during the past few years up to the point that they may be difficult even for a human to distinguish them from real users (Ferrara et al., 2016). Previously, simple supervised machine learning models could identify accounts with unnatural tweeting volumes or robotic tweeting schedules and were sufficient for keeping the number of bot accounts low (Ferrara et al., 2016). Because Twitter, like most of social media sites, attempts to actively detect and disable bot accounts, the creators of bots have responded by making bots behave more like humans. Consequently, the selection of features as well as preparing the training data has become more demanding and for a model to stay up to date, feature engineering and adding new training datasets is needed (Yang, Varol, Davis, Ferrara, Flammini, & Menczer, 2019).

During the past few years both supervised (Stukal et al., 2017; Beskow and Carley, 2019) and unsupervised (Minnich et al, 2017; Chavoshi, Hamooni, & Mueen, 2016) machine learning models have been used in bot detection research. The drawback of supervised learning has been that creating a labeled dataset for training the model either requires a large amount of manual labeling (Stukal et al., 2017) or using a prelabeled dataset, which may limit the capabilities of the model as the datasets most likely represent only a fraction of the possible behavior of bot accounts in Twitter. Another argument that supports the use of unsupervised learning is that these models can detect novel bot behavior that may get past a supervised model (Minnich et al., 2017), which can only detect bots that are similar enough to the dataset that was used to train it. However, the results of unsupervised models are more difficult to validate due to the absence of labeled data.

One of the reasons for supervised models being used is that they are better suited for analyzing topical datasets that are collected from Twitter’s streaming API (Stukal et al., 2017). Twitter’s API allows performing searches and collecting the data on tweets that contain for certain keywords or hashtags, which is particularly useful when analyzing political discourse that is related to a specific topic, such as an election (Fernquist, Kaati, & Schroeder, 2018; Neuder, Kollanyi, & Howard, 2017). Since campaigns, political parties, candidates and users use hashtags to make their tweets visible when commenting on specific topic, it is more efficient to mine data on a topical level with the keyword search instead of first collecting a large dataset of Twitter accounts and then analyzing the content of their tweets.

Table 2 contains information on the algorithms used, numbers of features and performances of a collection of the models discussed in the literature review. There are considerable differences between the models particularly in the number of features used, as well as in the algorithm of choice. In most cases the random forest algorithm is recommended (Fernquist, Kaati, & Schroeder, 2018), but based on this sample other algorithms should be tested as well. It should be noted that the models cannot be directly compared with each other since there are differences in the use cases.

Table 2: Models, features, algorithms and performance

Model by	Algorithm	Features	Accuracy	Precision	Recall	F1
Beskow & Carley (2019)	Log. regression	1	0.996	0.994	0.997	0.996
Davis et al. (2016)	Custom	1150	0.734	0.471	0.208	0.288
Fernquist et al. (2018)	Random forest	140	0.957	0.941	0.976	0.958
Stukal et al. (2017)	Ridge log. reg.	42	-	0.99	0.77	0.866
	SVM		-	0.92	0.87	0.894
	Ensemble		-	0.99	0.77	0.866
Minnich et al. (2017)	Custom	130	-	0.90	-	-

* If information is unavailable the field is marked "-"

2.2.4 Detecting the right bots

As stated in the previous section of the literature review, not all bots are malicious or relevant for this type of research. Twitter allows the use of bots under certain limitations, which has resulted in a large number of non-malicious bots being used for curating and distributing information by organizations as well as individuals (Oentaryo et al., 2016). A practice used by several models to ignore these accounts is to ignore Twitter users which use the word bot either in their name of profile.

Furthermore, not all malicious bots that like or retweet political posts or follow politicians are political bots by design. As an example, Gallagher (2018) points out that the Finnish botnet discovered by Patel (2018) seems to be following individuals on Twitter based on the recommended profiles that the platform offers when creating an account. Since politicians are often popular and visible users in Twitter, these bots received them as recommendations. As a result, these bots will become false positives in a model that is designed to look for political bots.

2.3 Use of bots in political influencing

2.3.1 Evidence and examples

The paper closest to this thesis geographically and timewise is a study done on political bots in the Swedish general election in September 2018. Although the amount of bot activity was negligible at only 5% of all election related content during the monitoring period, it is still unsettling and suggests further research may be valuable. The paper does not discuss what are the most likely culprits behind the bots, but it does draw the conclusion that the bots were more actively participating in discussions with right-wing users and the most followed bots were also skewed towards the right. (Fernquist, Kaati, & Schroeder, 2018)

The results from two papers that monitored bot activity in Germany during a state parliament as well as federal presidential election (Neudert, 2017) and federal election during 2017 (Mortatter et al., 2018) are similar to those from Sweden's general election of 2018. Bots represented around 7 - 11% of the accounts and bot-driven content represented 7.4 – 9% of all traffic during the German elections (Mortatter et al., 2018; Fernquist, Kaati, & Schroeder, 2018). These are modest numbers and in line with Twitter's estimate of bots accounting for approximately 10% of activity. The main reason for concern is that the bot activity was skewed towards supporting the alt-right movement and was possibly produced by accounts outside of Germany (Mortatter et al., 2018).

Stukal, Sanovich, Bonneua and Tucker's (2017) findings on bots in Russian political Twitter are far beyond those of any other papers discussed in this thesis in terms of the magnitude of bot created content. According to the article, up to 85% of the daily tweets containing political keywords were posted by bot accounts during 2014-2015. This is not overly surprising as the target country, Russia, differs from the other examples that are all from Western democracies.

Based on the findings of multiple papers and studies, there is evidence of Twitter-based computational propaganda being used across the world by organizations and governments (Woolley & Howard, 2017). There are regional differences in both the prevalence of bot accounts as well as in their activeness (Woolley & Howard, 2017), and based on this sample it could be assumed that if Finland is similar to other geopolitically unpowerful and small states, there will be low levels of bot activity. However, the current growth of Euroscepticism and rise of right-wing political movements alongside Finland's historical relations and proximity to Russia may provide a more fertile foundation for bot activity than for example Sweden.

2.3.2 Goals of political bots

Measuring the successfulness of political bots is difficult as it is hard to quantify the impact that they have had for example on voting behavior (Bjola, 2018). This is partially due to the difficulty of determining causality as voters are constantly influenced by multiple different sources. Nevertheless, the prevalence of computational propaganda campaigns would suggest that they are viewed as a functional tool that does have an effect on the target audience (Bjola, 2018).

There are several different hypothesized goals that the creators of bots are trying to reach. These range from high level goals such as increasing the partisanship of a population or advancing a cause that the creator of the bots supports. Bjola (2018) suggests that one of the reasons for governments to participate in creation and distribution of digital propaganda is that “it is an effective non-military means for achieving political and strategic goals”.

More measurable and easily achievable targets include manipulating the popularity and visibility of tweets by liking, following and retweeting content with a botnet. These methods can cause a particular hashtag to trend thus pushing it higher into the feeds of other Twitter users. Other goals may be to make an opinion seem more popular than it actually is or to bury actual discussions or factual information making it difficult to follow. Concrete examples of use cases include spamming pro-government tweets or flooding search results related to protests with meaningless content making it more difficult for human users to find and participate in discussions with each other (Suárez-Serrato et al., 2016).

3 Methodology

This chapter describes the methodology used to answer the research questions. The chapter is divided into three main sections which describe the collection of the data, the creation of the bot detection algorithm and how the network analysis was conducted. In each of the sections, the tools and activities related to the task are described.

Figure 2 below shows the main steps of the process and the tools used in each phase. The first step of the process is to select a sample of politicians with active Twitter accounts that have as many followers as possible. The second step is to download all metadata from the profiles of these politicians' followers. This creates a large dataset consisting of Twitter accounts and their metadata. The third step is to clean and reshape the data for the machine learning algorithm. Between the third and fourth step various features and algorithms are tested in order to develop the bot detection model. In the fourth step, the finished bot detection model is used on the dataset to create a new dataset consisting only of bot accounts. In the fourth step this data set is analyzed in various ways until finally in the last step it is reformatted and visualized.

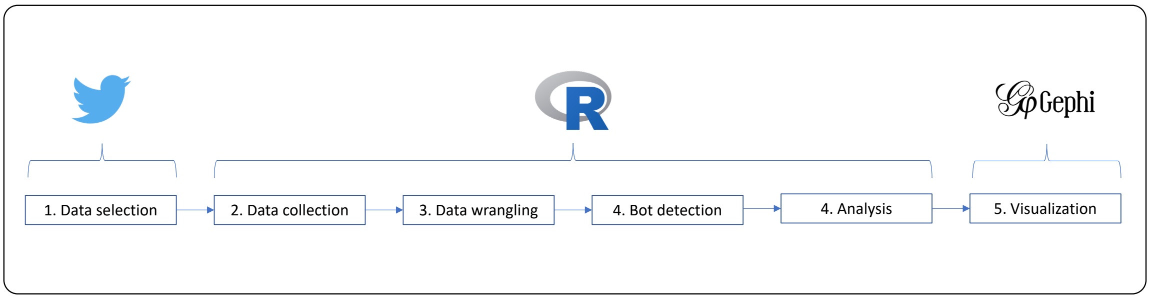


Figure 2. Process map

3.1 Collecting the data from Twitter

This thesis tests an unorthodox approach to collecting Twitter data as the dataset is compiled from individual accounts' followers instead of the more common method where all tweets (and associated account metadata) that use specific hashtags or keywords are gathered through Twitter's Streaming API. This method is appropriate since the proposed bot detection model only requires metadata. The benefits of this approach include that it allows detecting both dormant bots as well as those that do not use specific hashtags or words that the streaming method queries for. The primary drawback is that analyzing the intentions or goals of the botnet is mainly limited to identifying amplification attempts, or in other words measuring the amount of bot followers that specific accounts have and seeing how much content these bots are producing.

3.1.1 Tools

The primary tools used in data collection and formatting phase were the statistical programming language R and its “rtweet” package, which is a “R client for accessing Twitter’s REST and stream APIs” (Kearney, 2018). Additionally, the “tidyverse” library and packages included in it were used in cleaning the data.

3.1.2 Selecting the data

Selecting which politicians’ followers to inspect for bot activity is a critical choice for the results of the thesis, since the percentage of bots compared to real users as well as interpretations on whether the bots are influencing the discussions under right-wing or left-wing politicians are highly dependent on the dataset that is analyzed. In other words, if the sample that is taken happens to have a particularly disproportionate amount of bot activity or is otherwise skewed, consequently the results will most likely also be misleading. As a further complication, many Finnish politicians do not own a Twitter account and those that do, usually have only a modest number of followers, which makes it difficult to obtain equal amounts of data on all parties. These limitations will be accounted for by taking an as comprehensive as possible sample and by not making overly broad conclusions or inferences regarding the whole Finnish political Twittersphere in the findings and conclusions chapters.

It should be noted that this method results in the collected dataset being essentially a snapshot from March 2019. Consequently, the results are likely to be different if the study is repeated at a later point in time, since there is a possibility that the bot activity intensifies as the election day gets closer. Additionally, some of the users studied now may be identified by Twitter as bots and deleted. This creates a possibility for further studies as it can be monitored how effectively Twitter identifies the bots detected by this thesis and removes them.

The profiles were selected based on several heuristically chosen criteria to ensure that as many political parties as possible were represented and that a sufficient amount of data was collected. At least one member of parliament was taken from each of the current coalition parties as well as from all parties that have a support of over 5%, however with a maximum of two per party. Furthermore, only accounts with over five thousand followers were picked. Lastly, some prominent politicians with over 10k followers were selected even if they do not match the other criteria as several influential political figures would otherwise be excluded. Table 3 shows the politicians that were selected and whose followers' data was mined.

Table 3: Selected politicians

Name	Political party	Username	Followers (K)*
Alexander Stubb	National Coalition Party	@alexstubb	370
Sauli Niinistö	National Coalition Party	@niinisto	159
Juha Sipilä	Centre Party	@juhasipila	126
Anne Berner	Centre Party	@AnneBerner	21.7
Pekka Haavisto	Green League	@Haavisto	130
Ville Niinistö	Green League	@VilleNiinisto	84.3
Paavo Arhinmäki	Left Alliance	@paavoarhinmaki	109
Li Andersson	Left Alliance	@liandersson	76.5
Antti Rinne	Social Democratic Party	@AnttiRinnepj	25.6
Sanna Marin	Social Democratic Party	@MarinSanna	14.3
Jussi Halla-aho	Finns Party	@Halla_aho	14.5
Laura Huhtasaari	Finns Party	@LauraHuhtasaari	13.7
Sampo Terho	Blue Reform	@SampoTerho	7.6
Paavo Väyrynen	Seven Star Movement	@kokokansanpaavo	10

*Number of followers at March 2019

The sample consists of 14 different politicians from 8 different parties ranging from liberal to conservative and left-wing to right-wing. These include the current president and three ministers as well as 6 party leaders. Many of the small parties were left out by this approach, but simultaneously their politicians did not have accounts or had much fewer followers, which reduced the impact that their exclusion could have had on the results. The number of Twitter accounts whose data was collected totaled over 1.1 million, but it was reduced down to a bit above 550000 after duplicates were removed. The duplicates were a result of the fact that many Twitter users were following multiple selected politicians.

3.1.3 Data wrangling

The rtweet package's lookup_users method returns 88 variables, which includes both metadata and other data, such as information related to the most recent tweet made by the user. Most of these variables are unnecessary for the model and after removing the unused and non-metadata related columns, the table is left with less than 15 variables. Additionally, some of the columns were renamed according to the features that they represent to add clarity. Other changes that were made to the datasets are described in the next section as they are more closely related to the feature engineering phase. The used scripts and a sample view of the collected data can be found in appendices A and B.

3.2 Implementation of the bot detection algorithm on the dataset

This section starts by briefly describing how the dataset was further formatted and what were the packages used in the machine learning model. The second part focuses on the process used to create the bot detection model, which was divided into two steps with both of them containing a different version of the model. The goal of the first step was to assess the overall feasibility of using metadata to classify bots and to support the creation of a new training dataset for the second version of the model. The second version of the model was refined by adding new features and by being trained with the new training data. Lastly, final part of this section describes how the second version was used on the 550,000 accounts to create a final dataset consisting only of bots.

3.2.1 Preparing the data

Further manipulations to the dataset were made to create new attributes to support the binary and ratio features. The binary features were calculated from corresponding attributes where a setting left at default or blank equals 1. The ratio features were created similarly by calculating the values from the profile information metadata and then placed into new columns. Ratio calculations that resulted in NaN (not a number) or Inf (infinite), were replaced with a zero. Lastly, redundant attributes were removed.

3.2.2 Tools

The bot detection model was built with the statistical programming language R by using caret (Kuhn, 2008), a machine learning library. Caret was chosen since it allows experimenting with machine learning algorithms from different packages and provides a unified process and syntax for all tasks from data preparation to model evaluation.

3.2.3 First version of the bot detection model

Based on the findings of previous research and datasets available for training the model, a selection of 11 features were picked for testing the first version of the model. The feature space consists of four binary features, four profile information features and three ratio features that can be seen in Table 4.

Table 4: Features included in the first version of the model

Binary features	Profile information features	Ratio features
Default profile image No banner No user description No location	Number of followers Number of friends Number of tweets Number of likes	Followers / Friends Likes / Followers Likes / Friends

The model was trained with the cresci-2017 dataset (Cresci et al., 2017), which contains over 13000 labeled accounts divided into groups of social spambots, traditional spambots, fake followers and genuine accounts. The dataset was formatted to have the same attributes as described in Table 4. Lastly, the training dataset was balanced including 3000 randomly sampled bot accounts and 3000 randomly sampled genuine accounts.

To find a suitable algorithm for the model, the LDA, CART, KNN, SVM and Random Forest algorithms were tested. Out of these Random Forest performed the best, although there were signs of either the training data not representing the variety of real data or that the model being overfitted as the accuracy was over 97% or 98% on most runs. This issue was ignored as the model was deemed sufficiently accurate for the first phase where the goal was mainly to make the manual validation of the results quicker by creating a list of potential bot accounts. The model was then tested on a sample of 5000 accounts from the dataset that was collected for this thesis.

After manually inspecting on Twitter the accounts that the model labeled as bots, it was evident that the model had difficulties distinguishing bots and genuine accounts. Particularly accounts which were apparently created by people trying out Twitter without becoming active users were prone to being labeled as bots due to the behavior being similar. In most cases, the easily distinguishable bots were following approximately 20-100 accounts, had 0-2 followers and little to no tweets, retweets or likes.

3.2.4 Creating a new training dataset

Based on the performance of the first version of the model, it was apparent that the cresci-2017 dataset was unsuitable for training a model that could accurately distinguish bots from humans based on metadata. A possible explanation is that the training data had only very clear examples of bots and genuine accounts, where the behavior in terms of tweets, retweets, likes and ratios of followers and following differed widely depending on whether the account was a bot or not. This does not reflect the actual behavior of accounts where in some cases even with quantitative and qualitative assessment it is difficult to label an account accurately as either a bot or a human.

By manually labeling a set of accounts from the dataset consisting of followers of the Finnish politicians, a new training dataset that represents the actual distribution and behavior of the accounts of the target dataset was created. This was done by checking and verifying the accuracy of 2000 accounts predicted to be bots by the first model. The results were that out of these accounts 1336 were accurately labeled as bots, as they were either bots or accounts exhibiting extremely bot like behavior while 664 were actually humans or accounts which were impossible to determine as belonging to either group.

A qualitative approach was employed for classifying the accounts as either bots or humans. The process used in labeling the accounts is illustrated in Figure 3. The classification started by inspecting the profile information of the account. Common signs of a bot were the name or description of the account, which often included Russian or Arabic and or a seemingly random string of characters and numbers coupled with the account following 21 other Twitter users, which is the default number of recommended users to follow given by Twitter when creating a new account. Other possible predictors included in this step are the profile image and banner as well as the age of the account. As a second step the tweets and retweets were checked when available to see what kind of activity the account has and what other accounts it interacts with. As the third step, the accounts that the possible bot was following were inspected to find discrepancies. For example, a user following mainly seemingly random foreign accounts coupled with one Finnish politician or if it was following exactly 21 very popular Finnish accounts were usually the best predictors of an accurate classification as a bot even though the machine learning model could not look for these.

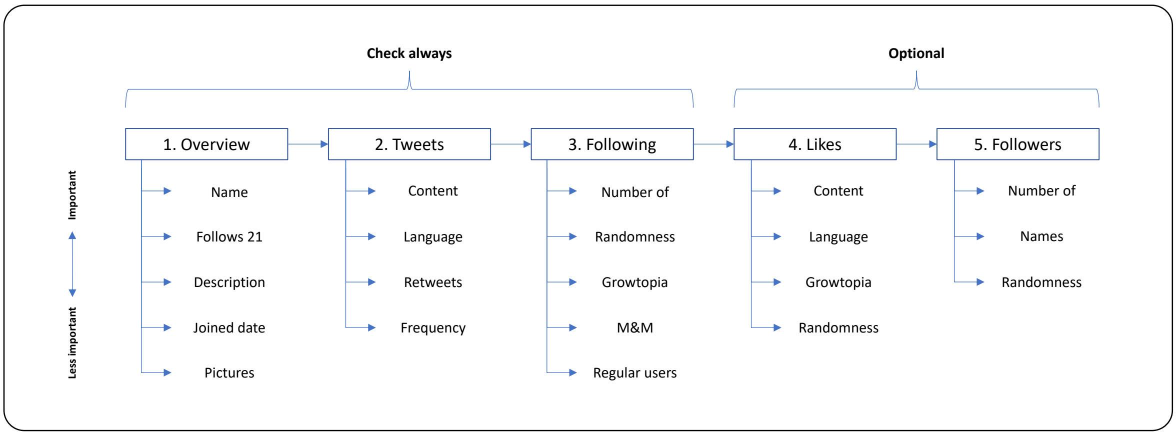


Figure 3. Framework for qualitative classification of bots

If after the three first steps the account was still too ambiguous for classification, the likes and followers were checked for bot like behavior. In Figure 3, randomness implies the entropy of some behavior. An example of this would be an account that follows only Finnish content but likes content that is in a foreign language and not at all in line with the other characteristics of the account.

During this process several interesting findings were made, which can be used later in the analysis of the whole dataset. Firstly, most of the bot accounts were dormant as well as possibly a part of follower boosting operation. Secondly, most of the bots were difficult to label as political bots as it is not sure whether they were created to boost the followers of a particular politician or if it followed them by coincidence based on Twitter's recommendations. Common shared characteristics among bots included that they barely engaged with content or interacted with other users and that they followed a random group of 21 accounts, which most likely are those suggested by Twitter during the creation of the account (Patel, 2018; Gallagher 2018). Peculiar accounts that they often followed included less well-known US politicians, an obscure game called Growtopia and a niche Finnish newspaper called Markkintointi and Mainonta. These findings and their implications to the model will be analyzed and discussed further in later parts of the thesis.

3.2.5 Improving the bot detection model

The second version of the bot detection model differed from the previous one mainly in how the splitting of the training and validation data was done, what parameters and algorithms were used as well as how many features were included.

New features could be added to the second version of the model as the training data was no longer a limiting factor. By including the age of the account, and two ratio features derived from comparing the profile information to the age of the account, the number of features was increased from 11 to 14. The new feature set is shown in Table 5.

Table 5: Features included in the second version of the model

Binary features	**Profile information features**	**Ratio features**
Default profile image No banner No user description No location	Number of followers Number of friends Number of tweets Number of likes Age of account	Followers / Following Likes / Followers Likes / Following Following / Age of account Likes / Age of account

Several variants of the Random Forest algorithm were tested, but the standard version still performed optimally and was selected for the final model. The model was trained with a randomly sampled set of 500 bots and 500 humans from the new manually labeled dataset. The remaining 1000 were used in the validation of the performance. The final version of the bot detection model has an accuracy of 83% with only slight changes after multiple runs and small variations in parameter settings. Table 6 below lists the most important statistics for assessing the performance.

Table 6: Performance of the bot detection model

Metric	Value
Accuracy	0.837
Recall	0.846
Specificity	0.793

In terms of feature importance, the top features were a mix of profile information and ratio features, while the binary features were all in the bottom half of the feature ranking. The model gives much weight to the number of accounts that an account is following, since the two top features are related to the following attribute. This is somewhat problematic for the overall goal of the thesis as it implies that the model is best at detecting dormant bots and bots belonging to follower farms. These accounts can be political bots, but in many cases determining if they are following politicians on purpose or by coincidence is difficult. This is due to the fact that politicians often appear on the top of the recommended accounts to follow in Finland, which makes them prone to attracting bots.

Table 7: Features ranked

Rank	Feature	Importance
1.	Following	100.00
2.	Following to age of account	61.00
3.	Age of account	56.57
4.	Followers to following	22.87
5.	Likes to following	15.68
6.	Tweets	15.18
7.	Likes to age of account	11.95
8.	Followers	10.05
9.	Default profile image	7.11
10.	Likes	6.34
11.	No description	4.72
12.	No location	3.61
13.	No banner	2.56
14.	Likes to followers	0.00

3.2.6 Implementation

To implement the bot detection algorithm, the dataset consisting of the 558,983 followers of the 14 Finnish politicians was formatted to match the training dataset. The model predicted that out of the dataset approximately 36.6% are bots. Since the model's accuracy is 83%, out of the 204,426 accounts classified as bots it can be assumed that 169,673 should be the real number of bots when not taking into consideration the accounts labeled as humans that in reality are bots.

To ensure that the analysis is conducted on accounts that are true positives, or in this case bots labeled as bots, an additional measure is taken to subset the 204,426 accounts. During the manual validation phase several initial findings were made regarding the characteristics of typical bots within the dataset. By querying the dataset containing the prediction results with additional search criteria, such as if the account is following Growtopia or exactly 21 other accounts, new groups which have a much higher likely hood of being true positives could be formed. Although these accounts may not be purely political bots, their existence in the Finnish political Twitter sphere merits investigating them.

Based on the validation phase, four main groups of bots were identified. Three of these groups are related to a bot's behavior in following other accounts, with the first being the bots following 21 other accounts. The second group is bots following Growtopia and the third is bots following Markkinointi & Mainonta (M&M). The fourth group is bots that have Cyrillic or Arabic writing in the profile, which also may indicate where their creators are located. The groups are not mutually exclusive and there were occasional examples of bots belonging to two or three of them. Appendix C contains examples of each bot type. Later in the thesis the primary focus will be on analyzing these new groups of bots and the fifth chapter will describe the main findings and results.

3.3 Network analysis

3.3.1 Tools

The network analysis was completed with R and Gephi, an open-source network exploration and visualization tool. In R, the bot dataset was split into four separate data tables based on the groups they belong to. In the four new datasets, each row represents an edge and contains the name of the bot account and the Twitter name of the politician that it is following. The bot networks in each dataset were then analyzed, clustered and graphed in Gephi. In the network graph, each node is an account that belongs either to a politician or a bot. Edges have a weight of one and indicate which politicians a bot node is following. To highlight the politicians' accounts in the network graph, their nodes have titles with a font size related to how many edges are directed towards the node.

3.3.2 Analysis

To numerically analyze the networks and to cluster them, the modularity of the networks and weighted degree of the nodes were calculated in Gephi. Table 8 shows the characteristics and modularity of the networks as well as the average weighted degree of nodes within the network.

Table 8: Network characteristics

Network	Nodes	Edges	Modularity	Avg. Weighted degree
Following 21	2827	6749	0.245	2.518
Following M&M	36851	120952	0.222	3.282
Following Growtopia	11051	21391	0.288	2.387
Russian & Arabic language bots	13623	34299	0.364	1.936

Further analysis was done in R, where for example the percentage of bots in each politicians' list of followers was calculated and the relationships between the number of bots and different characteristics of the politicians' accounts were scrutinized.

3.3.3 Visualizations

The network visualizations were created using Gephi. To create the layout of the graphs, the OpenOrd and ForceAtlas2 algorithms were applied consecutively. By applying the algorithms to the network graphs, they became easier to inspect and evaluate visually. The graphs were colored based on the community structures detected via the modularity measure. It should be noted that the colors are not related to the political party or ideological leaning of the politicians. The goal of the visualization phase was to help visual analysis and to find interesting structures, which could later be evaluated numerically. The four network graphs can be seen in Figures 4-7 in the following pages. The findings of the analysis are discussed in the fifth chapter of the thesis. Larger and higher definition versions of the network graphs can be found by following the link in appendix D.

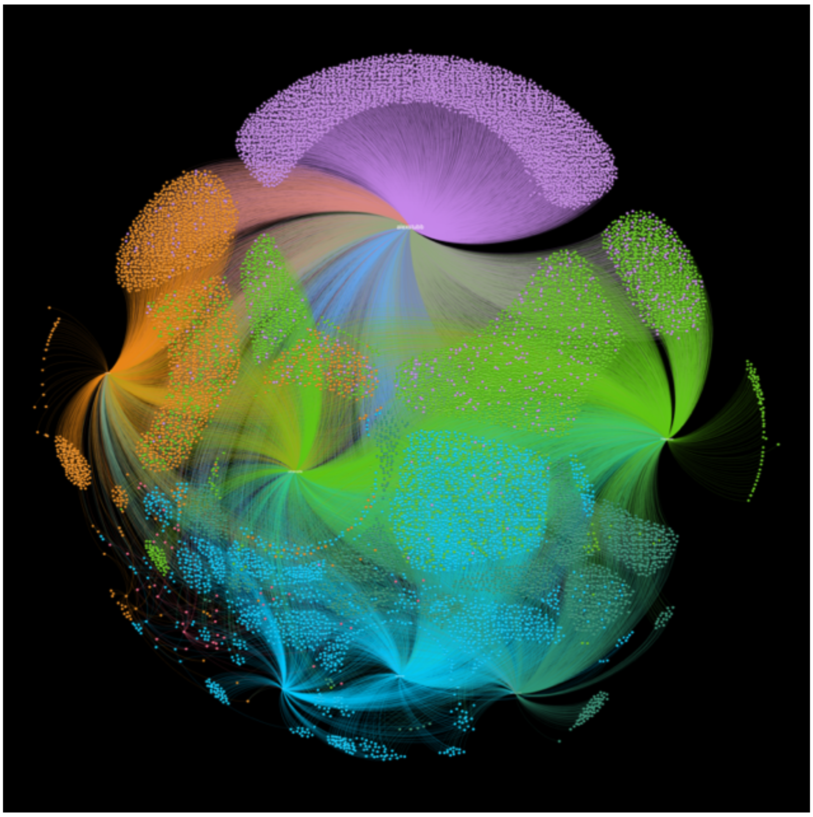


Figure 4. Network graph of the bots following 21 other accounts

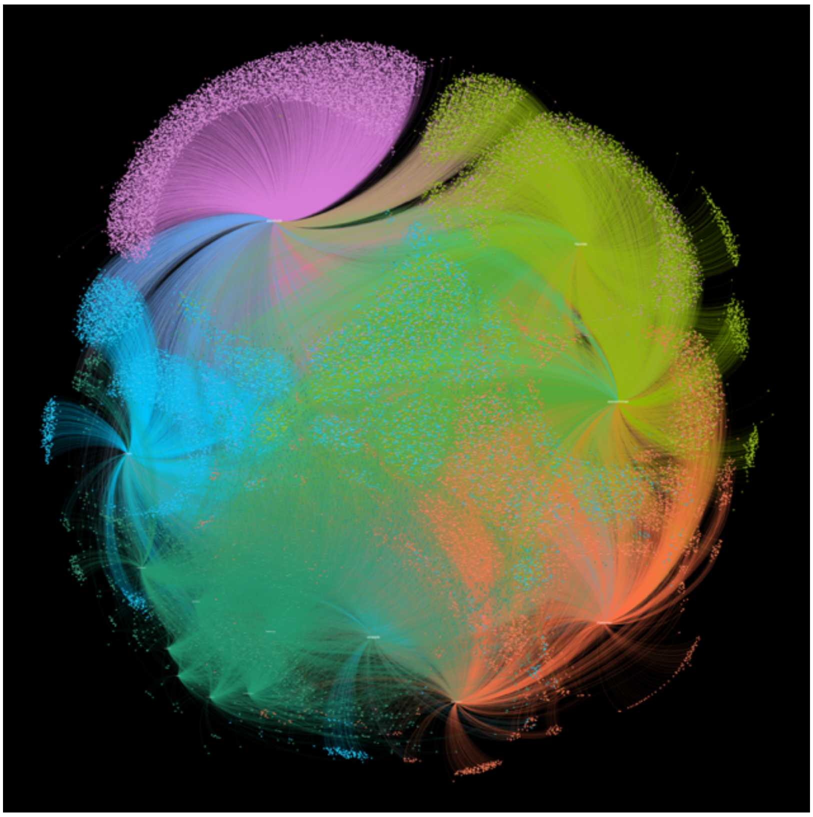


Figure 5. Network graph of the bots following Markkinointi & Mainonta

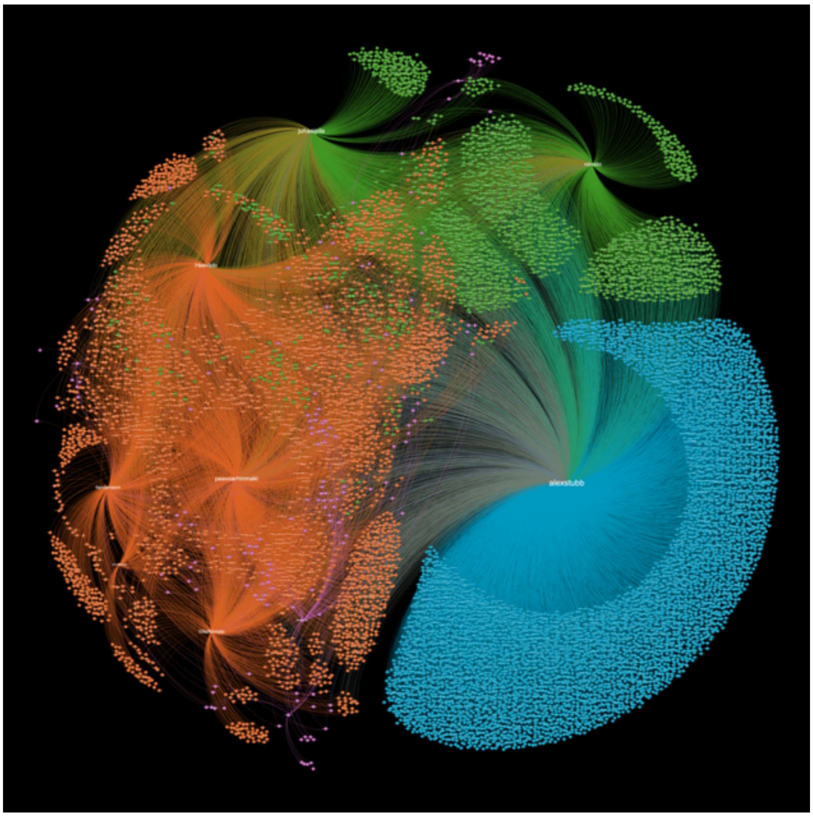


Figure 6. Network graph of the bots following Russian and Arabic accounts

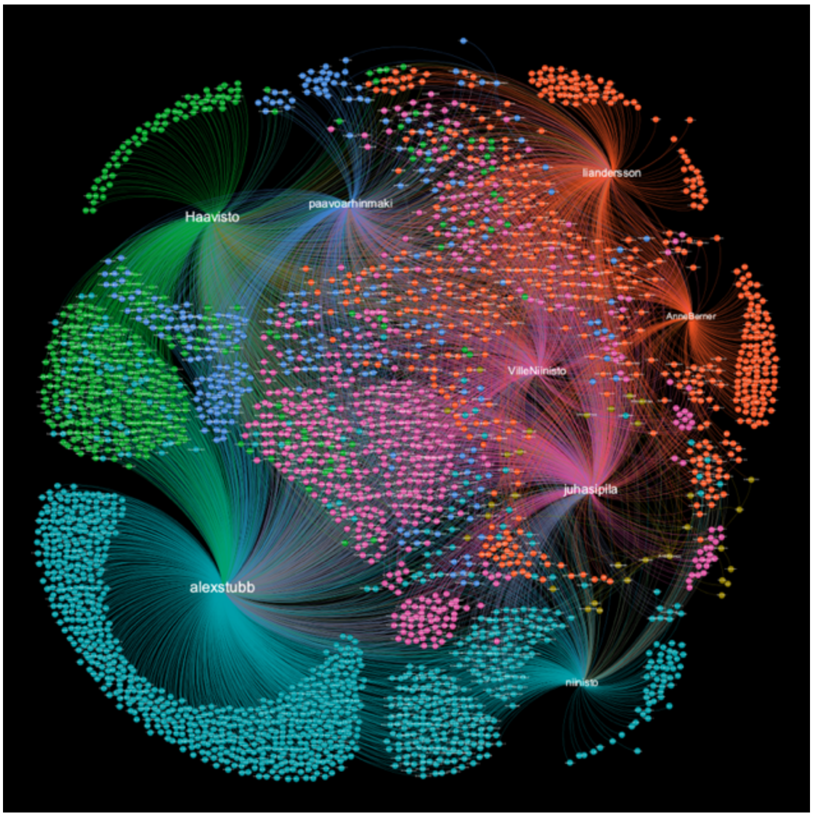


Figure 7. Network graph of the bots following Growtopia

4 Experiment

This chapter describes the setup of the experiment used to obtain the results of the thesis and can be used to test and replicate the findings. More specifically, this chapter lists the model parameters and the correct steps required to format and clean data so that the model can be used with the provided or any other similar datasets. The code and links to the datasets and R script can be found in appendices E and F.

4.1 Required libraries

In order to run the code, the caret, tidyverse and naniar packages in R are required. Additionally, in the supporting scripts the libraries zoo, e1071 and gtools were utilized or tested, but they are not required when running only the bot detection model.

4.2 Formatting data

The data is stored in data tables and the correct format for each column can be found from Table 9 below. Furthermore, when adding new features, binary data should be turned into factors in order for the model to correctly evaluate them while all other data can be stored in the numeric format whether they are integers or floats. The only exception of is the screen name column that contains characters and is removed before entering data into the model and returned after running the algorithm.

Table 9: Formatting columns

Feature:	**followers**	**following**	**tweets**	**likes**	**age of account**
Type:	numeric	numeric	numeric	numeric	numeric
Feature:	**followers to following**	**likes to following**	**likes to followers**	**likes to age of account**	**friends to age of account**
Type:	numeric	numeric	numeric	numeric	numeric
Feature:	**no banner**	**no location**	**no description**	**screen name**	**bot**
Type:	factor	factor	factor	factor	factor

4.3 Model parameters

All parameters described refer to the corresponding ones in the caret package. The metric used in the bot detection model is “Accuracy”. The modes and values used in the train function are the following. The resampling method is “repeatedcv” and the 10-fold cross validation is repeated 10 times, meaning that the data is split into ten parts and the model is trained on 9/10 of the data, while the last 1/10 is used for validation. This process is then repeated ten times with different ways of splitting the data in each iteration. Lastly, the search argument is set at “grid”, which is the default setting.

5 Results and discussion

In this section the key findings of the thesis as well as the results of the different analyses performed on the data are discussed. The chapter is divided into two parts starting with the findings related to the groups of bots followed by an assessment of the botnets' potential to impact Finnish political Twitter.

5.1 Findings

5.1.1 Bots in Finnish political Twitter

Overall, the findings of the thesis do not support the notion that Finland and Finnish politics would be the target of internal or external bot influencing campaign. There are two possible explanations for this. The first alternative is that the metadata-based model is inefficient or incapable of detecting social bots, which is a category to which most modern political bots belong to. The other explanation is that there is no error in the design of the model and/or that there are not many political bots in the Finnish Twittersphere. The latter is line with a very recent announcement made by Supo, the Finnish Security Intelligence Service, which stated that it has not found evidence of foreign entities attempting to influence the elections (Simojoki, 2019).

Although this announcement omits commenting on the possibility of internal attempts to influence political discussions online, it can still be seen as probable that local organizations or individuals would not utilize Twitter when attempting to influence elections. This is based on the fact that the social media site is not very popular or widely used and some estimates suggest that only about five to ten percent of the Finnish have at some point created a Twitter account. This would mean that there are only 275,000 to 550,000 Finnish Twitter profiles assuming that each individual has created only one account (Statcounter, 2019; Isotalus et al., 2018). Furthermore, of these accounts it has been estimated that only 50,000 are active users (Isotalus et al., 2018), which suggests that Twitter is not a viable platform when trying to reach and influence wide audiences in Finland.

Despite of there being few clear political bots, there are possibly over 150,000 identified bot accounts following Finnish politicians on Twitter according to the bot detection model. As was stated in the methodology chapter, four primary groups of bots were identified and analyzed. Although these bot accounts do not interact much with other accounts, they still help the politicians that they follow by two ways. Firstly, they artificially inflate the number of followers a politician has making them possibly seem more popular than they actually are. Secondly, they help increase the visibility of politicians, since being followed by many promotes an account over other less popular accounts in Twitter's who to follow suggestions. Consequently, bot accounts that are created for an entirely different purpose may unintentionally help politicians when they follow their accounts based on Twitter's recommendations. Since Twitter's recommendations are regional, this effect is more pronounced in an area like Finland that has a small population.

The four main groups of bots which were introduced in the methodology section and that were inspected more closely in this thesis can be seen in Figure 8. The figure also illustrates how the groups overlap, as for example some of the Russian and Arabic bots have similar following patterns as bots from other groups.

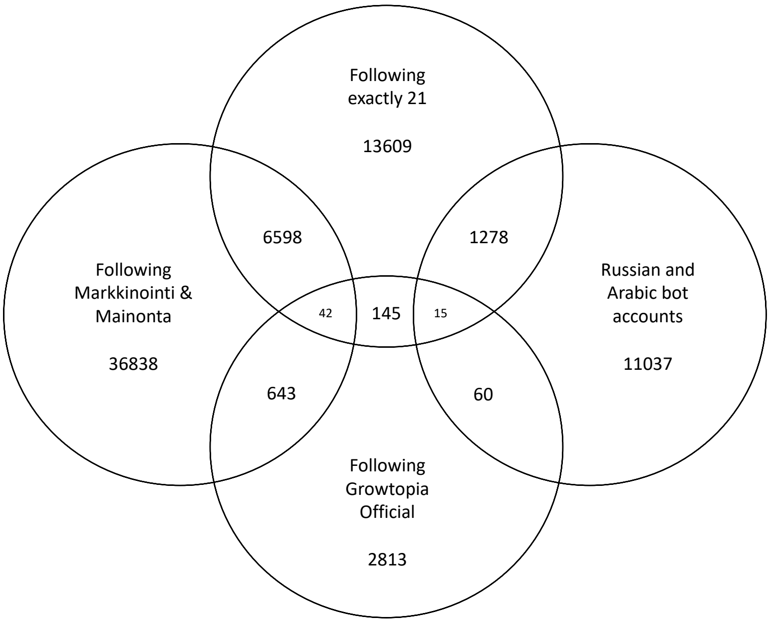

Following exactly 21
13609
Following Markkinointi & Mainonta
36838
Russian and Arabic bot accounts
11037
Following Growtopia Official
2813
6598
1278
42
145
15
643
60

Figure 8. Venn digram describing four most common types of bots found in the dataset

The first and largest group of bots are those that follow the Finnish paper Markkinointi & Mainonta. The group consists of 36838 bots and they represent 18% of all the accounts classified as bots by the model. There are several possible explanations for a seemingly random paper such as M&M attracting bots, which represent 49% of its 74937 followers on Twitter. One is that the paper has ended up being among the accounts that Twitter often recommends when registering to the social media site from a Finnish IP address. The other theory is that the paper has purchased followers to improve its visibility and popularity on Twitter, but this would seem unlikely to happen in Finland due to the beforementioned reasons. It can be assumed that these are simply unsophisticated bots that are either waiting to be activated for some future purpose or have already fulfilled their purpose as part of a follower farm. It is difficult if not impossible to categorize them as political or unpolitical with the information currently available.

The second group is bots that are following 21 other accounts and there are 13609 of them. As explained by Patel (2018) the most probable explanation for the prevalence of bots with this characteristic is that they are programmed to follow all the twenty-one accounts that Twitter suggests upon creating a new account. Similar to the Markkinointi & Mainonta group, the bots in this group cannot be labeled as being political with the current amount of information. Furthermore, the fact that these accounts are most likely following accounts purely based on Twitter's recommendation supports the theory that they are meant for something completely different and thus they do not intentionally influence Finnish politics.

The third group consists of both accounts assumed to be originating from Russian or Arabic speaking countries. Originally these accounts were identified during the validation phase by the use of non-ASCII characters, which in this case is Arabic and Cyrillic writing. Currently there does not exist any highly accurate or easily usable library for determining whether or not characters in a string are Cyrillic or Arabic. Therefore, as a proxy for this the language setting of the Twitter account was used instead to query for these accounts. This gives us a very rough estimate as it ignores accounts with Russian and Arabic text that have a different language such as English in the settings. The group has 8235 accounts with Russian and 2802 with Arabic set as the language of Twitter. Based on the characteristics of the accounts, they can be mainly categorized as belonging to follower farms or as being spambots that are spreading varying content.

The fourth and smallest group of bots are those that are following Growtopia Official, the Twitter page of an online game that has existed since 2013. These 2813 accounts are unique when compared to the bots in other groups, since their purpose is clear as the creators have not put much effort into concealing it. The Growtopia bots not only boost the number of followers the page has, but also often have one or two tweets solely dedicated to promoting the game. An example of this behavior can be seen in Appendix D under the Twitter screen name kivilahti_miro. The reason why these bots end up following Finnish politicians is most likely the same as with the other groups, or in other words by coincidence and as a result of Twitter's recommendation system.

Based on the network analysis, the modularity of the networks is low and thus there are only few clusters within the four different groups. With the current data it is not possible to determine whether or not these clusters actually have relevance or if they can be used in identifying bots that have been created by the same individual or organization. One possible reason for bot datasets being dividable into clusters is that Twitter's suggestions on which accounts to follow changes periodically resulting in the bots following different politicians depending on their date of creation (Gallagher, 2018).

5.1.2 The proposed bot detection model

The bot detection model proposed in this thesis suggests that metadata alone is sufficient for classifying at least spambots and bots that belong to follower farms. However, the lack of more specific data on tweeting behavior and the contents of the tweets renders the model unable to accurately identify advanced social bots based on the 2000 accounts reviewed during the validation phase of the model building process.

The primary benefit of a model based on metadata is that the data collection is much quicker as 90,000 accounts' information can be retrieved every 15 minutes. Therefore, a model that uses metadata works particularly well when studying countries that have a small population, since then even the most popular Twitter users are likely to have a manageable number of followers when compared to global standards. In other words, due to the limited number of users in these countries, it is possible to gather comprehensive datasets for analysis in short periods of time. Furthermore, analyzing entire populations instead of samples is feasible with a purely metadata-based model, contrary to models that use tweet data, where the number of accounts to analyze is restricted by Twitter's streaming API's rate limits.

Regarding the selection of the feature space and algorithm, most of the results were in line with the reviewed literature, although some of the results were surprising. Random forest was the optimal classification algorithm, which was the result in several other models as well (Fernquist et al., 2018). While ratio features had high feature importance as suggested by previous research, the binary features did not despite of their popularity in earlier models. Overall, the performance of the model was below most of those listed in the literature review, but as stated earlier direct comparison is difficult due to the differences in the goals of the models.

As a more theoretical contribution, the framework in Figure 9 was developed. Based on the different steps of the research process during the thesis as well as the backgrounds of other studies (Salge & Karahanna, 2018), it seems apparent that due to the vast variety and volume of Twitter data, interesting findings tend to be found as a byproduct of attempting to solve or find answers to an unrelated research question. In the case of this thesis, the original goal was to detect political bots, but due to the ambiguity of the results and lack of a clear presence as well as the surfacing of different types of bots the focus was pivoted to studying them instead of attempting to refine the model to answer the initial research question.

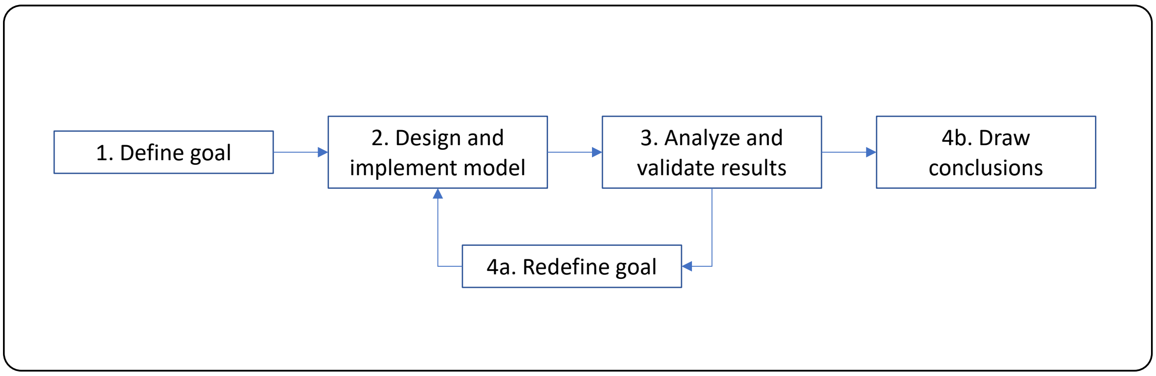

1. Define goal
2. Design and implement model
3. Analyze and validate results
4b. Draw conclusions
4a. Redefine goal

Figure 9. Framework for building bot detection models

The framework suggests that bot detection model research can and perhaps should be viewed as an iterative process where the problem being solved can be adjusted or changed dynamically based on interim results. Using this thesis as an example, in the first step the goal was set to finding evidence of political bots in the Finnish Twittersphere. To reach this goal the first version of the bot detection model was developed. After analyzing and validating the results produced by the first model, it was clear that the model did not work entirely as intended. However, it produced certain interesting findings as it managed to classify correctly a large number of bots that were then divided into four groups. Instead of proceeding to drawing conclusions, the goal was redefined, and the model redesigned as illustrated in the framework in steps 4a and 2. After analyzing the outcome of the second model, the findings were deemed strong enough to allow proceeding to the final step, which is drawing conclusions.

5.2 Potential impact on Finnish political Twitter

5.2.1 Overview

The primary impact that the bots have on Finnish political Twitter is related to the visibility and perceivable popularity of the politicians' accounts. Considering Twitter's low utilization as a medium for political debate in Finland, the possible effects the bots may have had on voters can be considered negligible. Nevertheless, one metric for measuring a politician's popularity that can be used to predict election results is how many followers they have on different platforms and how much their audience engages with them (DiGrazia et al., 2013). Therefore, even if the impact on actual voting behavior is minimal, the presence of bots may manipulate perceptions, influence predictions and damage the validity of social media engagement as an indicator of actual popularity.

5.2.2 Empirical evidence

Although at first glance the network graphs would suggest that the bots tend to form clusters around certain politicians, no correlation was found regarding party preference of the bots that belong to any of the four groups. This would imply that the politicians were followed most likely due to their presence in Twitter's suggestions. This is further strengthened by the fact that the number of followers an account has correlates very strongly (0.97) with the number of bots. This suggests that there may be a vicious cycle where popular accounts are likely to attract increasing numbers of bot followers, thus further boosting their visibility at the expense of less followed politicians.

Table 10 lists each of the fourteen politicians whose data was mined and shows how many bot followers from the four groups they have as well as the percentage of such bots in their follower base. It should be noted that this data represents a snapshot of the moment when the data was collected, which was in March 2019. Although the numbers are likely to have remained in the same magnitude, there may be some differences in the number of followers since an increased interest in politics and their Twitter profiles is expectable particularly during April, since the election date is on the 14th. Furthermore, Twitter removes accounts that it has identified as bots and this has already happened to some of the accounts belonging to the collected dataset.

Table 10: Percentage of bot[*] followers by politician

Twitter	Party	Followers	Bots	Bot %
@alexstubb	National Coalition Party	370000	49096	13%
@niinisto	National Coalition Party	151000	13107	9%
@juhasipila	Centre Party	124000	19949	16%
@AnneBerner	Centre Party	21700	2791	13%
@Haavisto	Green League	128000	24507	19%
@VilleNiinisto	Green League	83100	13891	17%
@paavoarhinmaki	Left Alliance	108000	16989	16%
@liandersson	Left Alliance	73700	7136	10%
@AnttiRinnepj	Social Democratic Party	24000	2529	11%
@MarinSanna	Social Democratic Party	14300	1555	11%
@Halla_aho	Finns Party	11900	634	5%
@LauraHuhtasaari	Finns Party	12300	815	7%
@SampoTerho	Blue Reform	7000	903	13%
@kokokansanpaavo	Seven Star Movement	10000	1115	11%

* Includes only bots that belong to the four groups. Percentages are higher if taking into consideration all potential bots

When inspecting the scores of individual politicians, Pekka Haavisto had the highest percentage of bot followers in the sample, and the percentage is beyond Twitter's own estimates of typical rates of bot followers. The strong bot presence in Haavisto's Twitter follower base was subject to debate already in 2017 during his presidential election campaign (Yle, 2017). Previous analysis attributed the bot followers to be a result of a sudden increase in Growtopia bots and Twitter's recommendations boosting Haavisto, which is similar to the findings of this thesis.

The other notable example of a politician benefitting from the added visibility is Alexander Stubb who has acquired the largest absolute number of bot followers. Many of the bots were not following any other politicians besides Stubb, which is likely due to his strong presence in Twitter as the 3rd most followed account in Finland. This is clearly visible in the network graphs, where typically Stubb is the only politician to have a large and separate cluster of bot followers.

Contrary to findings elsewhere (Schäfer et al., 2017; Morstatter et al. 2018), the candidates most likely to be linked to the Finnish alt-right movement Laura Huhtasaari and Jussi Halla-aho had the lowest percentage of bot followers. However, this is not surprising when taking into consideration the other findings and that they also have the lowest number of followers, which means that bots are less likely to follow them by coincidence.

6 Conclusions

This chapter paraphrases the main implications, contribution and limitations of the thesis and provides suggestions for further research.

6.1 Implications

The main implications of the thesis are that there is no clear evidence of Twitter being used as a platform for influencing Finnish politics during the 2019 parliamentary election and while a significant portion of the Finnish Twittersphere consists of bots, they are possibly non-political by nature. The bots can be described as malicious rather than benign and many of them are either dormant and waiting for activation or abandoned for having already completed their purpose. Furthermore, it is evident that Twitter's official estimates on the share of bot users are too conservative at least based on the data used in the thesis.

6.2 Contributions

The thesis has both practical and theoretical contributions to Twitter related research and the development of bot detection algorithms. The primary practical contribution is the bot detection model itself, which adds to the existing knowledge of how well metadata-based models can perform. Furthermore, it can be implemented in its current form or built upon in future research.

The development of the bot detection algorithm resulted in an unintentional, but additional methodological contribution for the development of future bot detection models, as a high quality and accurately labeled training dataset was created. Datasets with accurately classified Twitter accounts are not abundant and they lose value as they age due to the constant development of bots. If the dataset is accepted to the bot repository of Indiana University Network Science Institute, a portal which stores and distributes datasets and tools related to bot research, or otherwise shared online, it can be utilized in future research.

The theoretical contributions are the two qualitative frameworks used in the thesis. The first framework described the process used to qualitatively classify accounts as bots or humans and was introduced in the methodology chapter in Figure 3. Although currently is designed specifically for the dataset used in this thesis, with minor alterations it can be used in other contexts as well. The second framework is a higher-level and more conceptual tool for building bot detection models. It is used in the findings chapter and shown in Figure 9. The purpose of the framework is to propose an iterative and more open-ended approach to bot detection research.

6.3 Limitations

The limitations of this research are primarily related to methodology. More specifically, the approach used in the selection of politicians and data collection phase as well as choice of features in the machine learning model introduced some constraints to the analyses that could be performed. Although it was possible to determine if an account is a bot based on metadata, the collected data did not enable examining the content that they interacted with or spread via tweets, retweets and likes. Therefore, one of the major limitations is that the data collection method employed in this thesis allows only determining the number of bot followers that the selected politicians have and based on that estimating if the botnets are supporting certain parties or candidates. However, the impact of this was significantly reduced by the findings of the thesis, which suggest that a most of the bots detected are not actively creating or distributing content. Lastly, there is a limitation related to the selection of politicians. It is possible that politicians with much higher or lower percentages of bot followers may have been omitted from the sample.

6.4 Suggestions for further research

6.4.1 Future elections

In order to further understand the use of bots in the Finnish Twittersphere, the model could be reused during the upcoming 2019 European Parliament election by collecting a new dataset. This would be particularly interesting due to the Finnish Security Intelligence Service’s suggestion that the EU elections are likely to be a more attractive target for external influencing attempts than the Finnish parliamentary election (Simojoki, 2019).

6.4.2 Model development

The amount of testing and adjustments made to the model during the thesis was limited. Therefore, additional research could help identify new and more powerful features and thus improve the model, providing further support for metadata-based models. By introducing new features and testing alternative classification algorithms with different parameters, it is likely that all key performance metrics such as accuracy and recall could be improved and results closer to the models discussed in the literature review achieved.

6.4.3 Development of the botnet

To analyze the efficiency of Twitter's own bot detection and removal practices, the rate at which accounts labeled as bots get removed from the social media site can be followed. In addition, changes in the activity of the bots can be monitored by inspecting how the attributes such as number of tweets and likes changes overtime. Especially interesting would be to find evidence if some of the accounts were sleeper bots waiting for activation.

6.5 Summary

The goals of this thesis were to investigate if Twitter bots were used to influence the 2019 Finnish parliamentary election and to test a new approach to Twitter bot detection. In the thesis a new supervised machine learning bot detection model was developed and then used to determine the number of bot followers that a sample of the most popular Finnish politicians have in their follower base. To study the networks that the bots form more closely as well as to highlight certain interesting characteristics of these botnets, more in-depth analysis was conducted on four smaller groups of bots.

The dataset used in the thesis consisted of 550,000 unique accounts out of which roughly 169,600 were potentially bots according to the model's predictions. The metadata-based model was found to be feasible for classifying bots on Twitter and the predictions of the model were used to assess if bots were utilized during the 2019 Finnish parliamentary election. Based on the findings it was concluded that there was no evidence of attempts to influence the elections with Twitter bots. Although the bots increased the visibility of some politicians and made them seem more popular, the bots are unlikely to have had much effect due to their passive behavior coupled with the low usage of Twitter in Finland.

References

Beskow, D., & Carley, K. (2019). Its all in a name: detecting and labeling bots by their name. *Computational And Mathematical Organization Theory*, 25(1), 24-35. doi: 10.1007/s10588-018-09290-1

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday*, *21*(11). doi: 10.5210/fm.v21i11.7090

Chavoshi, N., Hamooni, H., & Mueen, A. (2016). DeBot: Twitter Bot Detection via Warped Correlation. In *ICDM,* 817-822. doi: 10.1109/ICDM.2016.0096

Chen, Z., Tanash, R., Stoll, R., & Subramanian, D. (2017). Hunting Malicious Bots on Twitter: An Unsupervised Approach. In *International Conference on Social Informatics*, 501-510. doi: 10.1007/978-3-319-67256-4_40

Bjola, C. (2017). Propaganda in the digital age, *Global Affairs*, 3(3), 189-191. doi: 10.1080/23340460.2017.1427694

Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., & Tesconi, M. (2017). The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race. In *Proceedings of the 26th International Conference on World Wide Web Companion,* 963-972.

Davis, C., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. (2016). BotOrNot: A System to Evaluate Social Bots. In *Proceedings Of The 25Th International Conference Companion On World Wide Web - WWW '16 Companion*. 274-274. doi: 10.1145/2872518.2889302

DiGrazia, J., McKelvey, K., Bollen, J., & Rojas, F. (2013). More Tweets, More Votes: Social Media as a Quantitative Indicator of Political Behavior. *PloS one*, 8(11). doi: 10.2139/ssrn.2235423

Fernquist, J., Kaati, L., & Schroeder, R. (2018). Political Bots and the Swedish General Election. In *2018 IEEE International Conference On Intelligence And Security Informatics (ISI)*, 124-129. doi: 10.1109/isi.2018.8587347

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications Of The ACM*, *59*(7), 96-104. doi: 10.1145/2818717

Gallagher, E. (2018). Visualizations of the Finnish-themed Twitter botnet. Retrieved from https://medium.com/@erin_gallagher/visualizations-of-the-finnish-themed-twitter-botnet-bfc70c6f4576

Isotalus, P., Jussila, J., & Matikainen, J. (2018). Twitter viestintänä - Ilmiöt ja verkostot (1st ed.). Tampere: Vastapaino.

Kearney, M. (2018). rtweet: Collecting Twitter Data. R package version 0.6.7. Retrieved from https://cran.r-project.org/package=rtweet

Kollanyi, B., & Howard, P. (2017). Junk News and Bots during the German Federal Presidency Election: What Were German Voters Sharing Over Twitter?. Technical report, Data Memo 2017.2. *Project on Computational Propaganda*. Retrieved from http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/03/What-Were-German-Voters-Sharing-Over-Twitter-v6-1.pdf

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. Journal of Statistical Software, 28(5), 1 - 26. doi: 10.18637/jss.v028.i05

Grimme, C., Preuss, M., Adam, L., & Trautmann, H. (2017). Social Bots: Human-Like by Means of Human Control?. *Big Data*, *5*(4), 279-293. doi: 10.1089/big.2017.0044

Minnich, A., Chavoshi, N., Koutra, D., & Mueen, A. (2017). BotWalk. *Proceedings Of The 2017 IEEE/ACM International Conference On Advances In Social Networks Analysis And Mining 2017 - ASONAM '17*, 467-474. doi: 10.1145/3110025.3110163

Morstatter, F., Shao, Y., Galstyan, A., & Karunasekera, S. (2018). From alt-right to alt-rechts: Twitter analysis of the 2017 german federal election. In *Companion of the The Web Conference 2018 on The Web Conference 2018*, 621-628. doi: 10.1145/3184558.3188733

Neudert, L. M. N. (2017). Computational propaganda in Germany: A cautionary tale. *Computational Propaganda Research Project*. Retrieved from http://blogs.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Comprop-Germany.pdf

Oentaryo, R., Murdopo, A., Prasetyo, P., & Lim, E. (2016). On Profiling Bots in Social Media. In *8th International Conference in Social informatics*, 92-109. doi: 10.1007/978-3-319-47880-7_6

Patel, A. (2018). Someone Is Building A Finnish-Themed Twitter Botnet. Retrieved from https://labsblog.f-secure.com/2018/01/11/someone-is-building-a-finnish-themed-twitter-botnet/

Salge, C., & Karahanna, E. (2018). Protesting Corruption on Twitter: Is It a Bot or Is It a Person?. *Academy Of Management Discoveries*, *4*(1), 32-49. doi: 10.5465/amd.2015.0121

Schäfer, F., Evert, S., & Heinrich, P. (2017). Japan's 2014 General Election: Political Bots, Right-Wing Internet Activism, and Prime Minister Shinzō Abe's Hidden Nationalist Agenda. Big data, 5(4), 294-309.

Simojoki, N., (2019). ”Ei se ulkopuolelta suunnatulta kampanjalta vaikuta” – Asiantuntijat: EU-vaalit houkuttelevampi vaikutusyritysten kohde. *Demokraatti*. Retrieved from https://demokraatti.fi/ei-se-ulkopuolelta-suunnatulta-kampanjalta-vaikuta-asiantuntijat-eu-vaalit-houkuttelevampi-vaikutusyritysten-kohde/

Suárez-Serrato, P., Roberts, M. E., Davis, C., & Menczer, F. (2016). On the influence of social bots in online protests. In *International Conference on Social Informatics*, 269-278.

Statcounter, (2019). Social Media Stats Finland. Retrieved from http://gs.statcounter.com/social-media-stats/all/finland

Stukal, D., Sanovich, S., Bonneau, R., & Tucker, J. (2017). Detecting Bots on Russian Political Twitter. *Big Data*, *5*(4), 310-324. doi: 10.1089/big.2017.0038

Twitter, (2018). Update on Twitter's review of the 2016 US election. Retrieved from: https://blog.twitter.com/official/en_us/topics/company/2018/2016-election-update.html

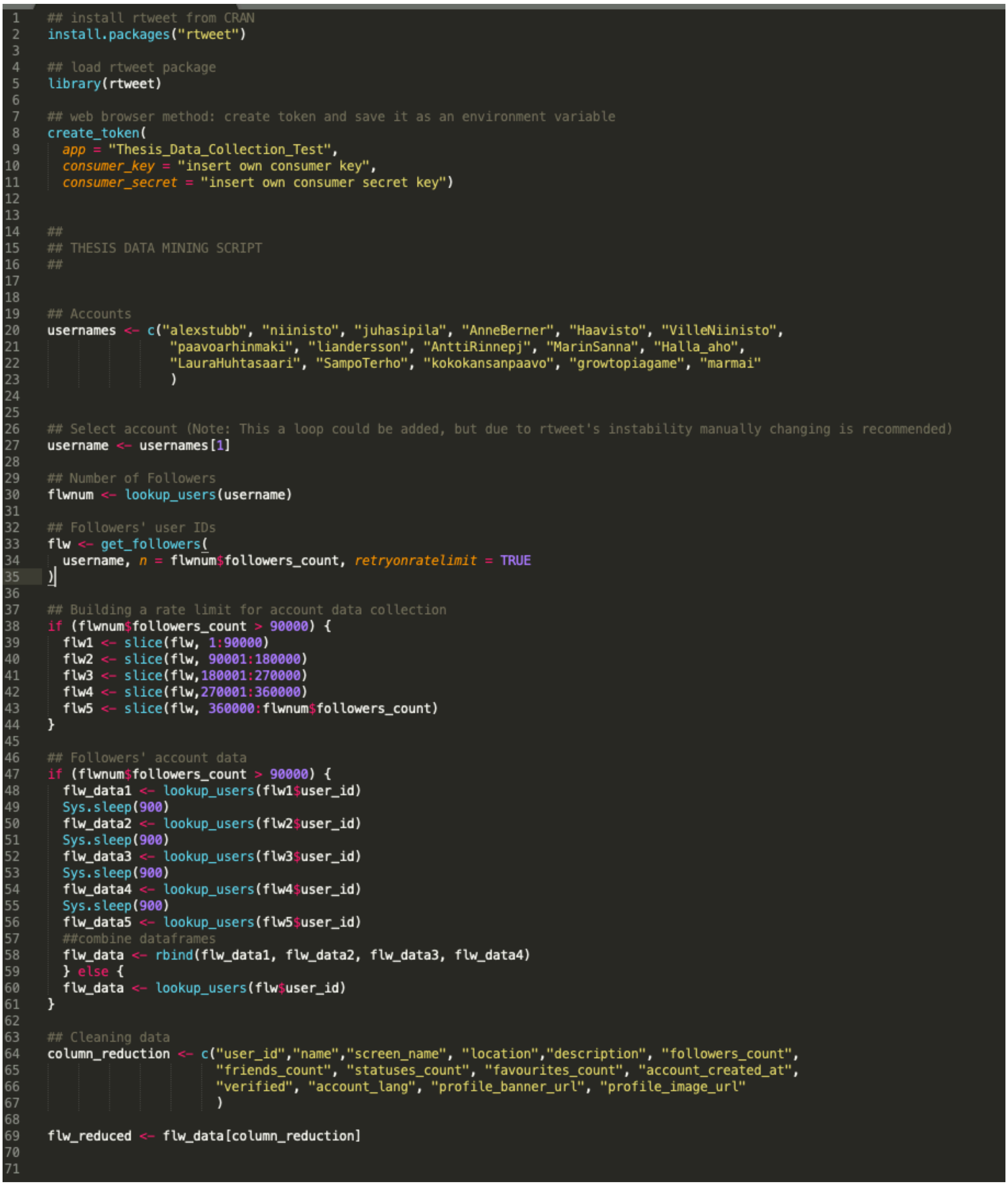
Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization. In *Eleventh international AAAI conference on web and social media*.

Wang, B., Zubiaga, A., Liakata, M., & Procter, R. (2015). Making the most of tweet-inherent features for social spam detection on Twitter. In *5th Workshop on Making Sense of Microposts,* 10-16.

Woolley, S. C., & Howard, P. N. (2017). Computational propaganda worldwide: Executive summary. *Computational Propaganda Research Project*. Retrieved from http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Casestudies-ExecutiveSummary.pdf

Yle. (2017). Pekka Haavisto campaign concerned over suspected Twitter bots. Retrieved from https://yle.fi/uutiset/osasto/news/pekka_haavisto_campaign_concerned_over_suspected_twitter_bots/9988551

Appendix A: Twitter data mining script

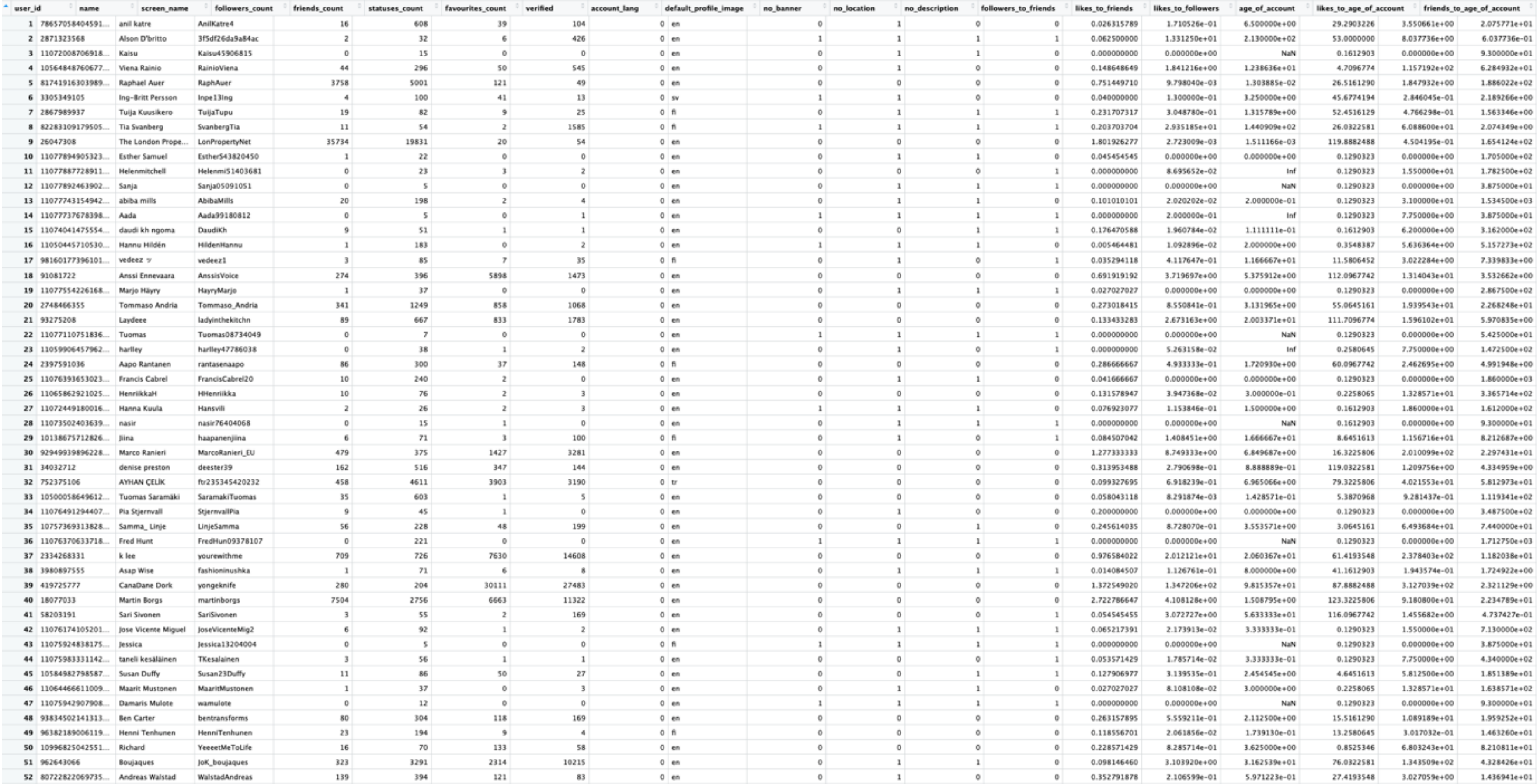


```r
## install rtweet from CRAN
install.packages("rtweet")

## load rtweet package
library(rtweet)

## web browser method: create token and save it as an environment variable
create_token(
  app = "Thesis_Data_Collection_Test",
  consumer_key = "insert own consumer key",
  consumer_secret = "insert own consumer secret key")


##
## THESIS DATA MINING SCRIPT
##


## Accounts
usernames <- c("alexstubb", "niinisto", "juhasipila", "AnneBerner", "Haavisto", "VilleNiinisto",
                "paavoarhinmaki", "liandersson", "AnttiRinnepj", "MarinSanna", "Halla_aho",
                "LauraHuhtasaari", "SampoTerho", "kokokansanpaavo", "growtopiagame", "marmai"
                )


## Select account (Note: This a loop could be added, but due to rtweet's instability manually changing is recommended)
username <- usernames[1]

## Number of Followers
flwnum <- lookup_users(username)

## Followers' user IDs
flw <- get_followers(
  username, n = flwnum$followers_count, retryonratelimit = TRUE
)

## Building a rate limit for account data collection
if (flwnum$followers_count > 90000) {
  flw1 <- slice(flw, 1:90000)
  flw2 <- slice(flw, 90001:180000)
  flw3 <- slice(flw,180001:270000)
  flw4 <- slice(flw,270001:360000)
  flw5 <- slice(flw, 360000:flwnum$followers_count)
}

## Followers' account data
if (flwnum$followers_count > 90000) {
  flw_data1 <- lookup_users(flw1$user_id)
  Sys.sleep(900)
  flw_data2 <- lookup_users(flw2$user_id)
  Sys.sleep(900)
  flw_data3 <- lookup_users(flw3$user_id)
  Sys.sleep(900)
  flw_data4 <- lookup_users(flw4$user_id)
  Sys.sleep(900)
  flw_data5 <- lookup_users(flw5$user_id)
  ##combine dataframes
  flw_data <- rbind(flw_data1, flw_data2, flw_data3, flw_data4)
  } else {
  flw_data <- lookup_users(flw$user_id)
}

## Cleaning data
column_reduction <- c("user_id","name","screen_name", "location","description", "followers_count",
                      "friends_count", "statuses_count", "favourites_count", "account_created_at",
                      "verified", "account_lang", "profile_banner_url", "profile_image_url"
                      )

flw_reduced <- flw_data[column_reduction]


```

Full script available at:

https://drive.google.com/open?id=19pCe5mLat9L34BqJznoYPk5aBVQW683i

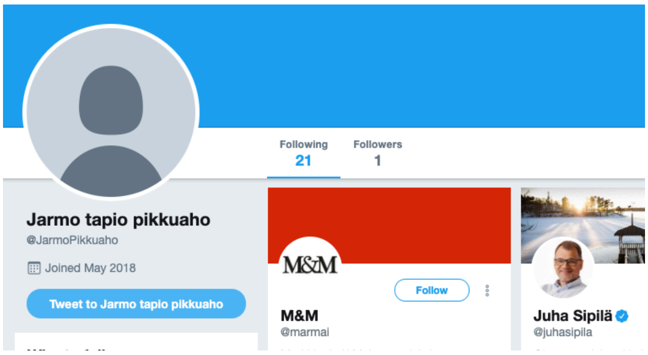
Appendix B: Sample view of the collected raw data



	user_id	name	screen_name	followers_count	friends_count	statuses_count	favourites_count	verified	account_lang	default_profile_image	no_banner	no_location	no_description	followers_to_friends	likes_to_friends	likes_to_followers	age_of_account	likes_to_age_of_account	friends_to_age_of_account
1	78657058404591...	anil katre	AnilKatre4	16	608	39	104	0	en	0	1	0	0	0.026315789	1.710526e-01	6.500000e+00	29.2903226	3.550661e+00	2.075771e+01
2	2871323568	Alson D'britto	3f5df26da9a84ac	2	32	6	426	0	en	1	1	1	1	0.062500000	1.331250e+01	2.130000e+02	53.0000000	8.037736e+00	6.037736e-01
3	11072008706918...	Kaisu	Kaisu45906815	0	15	0	0	0	en	0	1	1	0	0.000000000	0.000000e+00	NaN	0.1612903	0.000000e+00	9.300000e+01
4	10564848760677...	Viena Rainio	RainioViena	44	296	50	545	0	en	0	1	0	0	0.148648649	1.841216e+00	1.238636e+01	4.7096774	1.157192e+02	6.284932e+01
5	81741916303989...	Raphael Auer	RaphAuer	3758	5001	121	49	0	en	0	0	0	0	0.751449710	9.798040e-03	1.303885e-02	26.5161290	1.847932e+00	1.886022e+02
6	3305349105	Ing-Britt Persson	Inpe13Ing	4	100	41	13	0	sv	1	1	0	0	0.040000000	1.300000e-01	3.250000e+00	45.6774194	2.846045e-01	2.189266e+00
7	2867989937	Tuija Kuusikero	TuijaTupu	19	82	9	25	0	fi	0	1	1	1	0.231707317	3.048780e-01	1.315789e+00	52.4516129	4.766298e-01	1.563346e+00
8	82283109179505...	Tia Svanberg	SvanbergTia	11	54	2	1585	0	fi	1	1	1	1	0.203703704	2.935185e+01	1.440909e+02	26.0322581	6.088600e+01	2.074349e+00
9	26047308	The London Prope...	LonPropertyNet	35734	19831	20	54	0	en	0	0	0	0	1.801926277	2.723009e-03	1.511166e-03	119.8882488	4.504195e-01	1.654124e+02
10	11077894905323...	Esther Samuel	Esther543820450	1	22	0	0	0	en	0	1	1	0	0.045454545	0.000000e+00	0.000000e+00	0.1290323	0.000000e+00	1.705000e+02
11	11077887728911...	Helenmitchell	Helenmi51403681	0	23	3	2	0	en	0	0	0	1	0.000000000	8.695652e-02	Inf	0.1290323	1.550000e+01	1.782500e+02
12	11077892463902...	Sanja	Sanja05091051	0	5	0	0	0	en	0	1	1	1	0.000000000	0.000000e+00	NaN	0.1290323	0.000000e+00	3.875000e+01
13	11077743154942...	abiba mills	AbibaMills	20	198	2	4	0	en	0	1	1	0	0.101010101	2.020202e-02	2.000000e-01	0.1290323	3.100000e+01	1.534500e+03
14	11077737678398...	Aada	Aada99180812	0	5	0	1	0	en	1	1	1	1	0.000000000	2.000000e-01	Inf	0.1290323	7.750000e+00	3.875000e+01
15	11074041475554...	daudi kh ngoma	DaudiKh	9	51	1	1	0	en	0	0	1	1	0.176470588	1.960784e-02	1.111111e-01	0.1612903	6.200000e+00	3.162000e+02
16	11050445710530...	Hannu Hildén	HildenHannu	1	183	0	2	0	en	1	1	1	0	0.005464481	1.092896e-02	2.000000e+00	0.3548387	5.636364e+00	5.157273e+02
17	98160177396101...	vedeez ♉	vedeez1	3	85	7	35	0	fi	0	1	0	1	0.035294118	4.117647e-01	1.166667e+01	11.5806452	3.022284e+00	7.339833e+00
18	91081722	Anssi Ennevaara	AnssisVoice	274	396	5898	1473	0	en	0	0	0	0	0.691919192	3.719697e+00	5.375912e+00	112.0967742	1.314043e+01	3.532662e+00
19	11077554226168...	Marjo Häyry	HayryMarjo	1	37	0	0	0	en	0	1	1	1	0.027027027	0.000000e+00	0.000000e+00	0.1290323	0.000000e+00	2.867500e+02
20	2748466355	Tommaso Andria	Tommaso_Andria	341	1249	858	1068	0	en	0	0	0	0	0.273018415	8.550841e-01	3.131965e+00	55.0645161	1.939543e+01	2.268248e+01
21	93275208	Laydeee	ladyinthekitchn	89	667	833	1783	0	en	0	0	0	0	0.133433283	2.673163e+00	2.003371e+01	111.7096774	1.596102e+01	5.970835e+00
22	11077110751836...	Tuomas	Tuomas08734049	0	7	0	0	0	en	1	1	1	1	0.000000000	0.000000e+00	NaN	0.1290323	0.000000e+00	5.425000e+01
23	11059906457962...	harlley	harlley47786038	0	38	1	2	0	en	0	1	0	1	0.000000000	5.263158e-02	Inf	0.2580645	7.750000e+00	1.472500e+02
24	2397591036	Aapo Rantanen	rantasenaapo	86	300	37	148	0	fi	0	0	0	0	0.286666667	4.933333e-01	1.720930e+00	60.0967742	2.462695e+00	4.991948e+00
25	11076393653023...	Francis Cabrel	FrancisCabrel20	10	240	2	0	0	en	0	1	1	0	0.041666667	0.000000e+00	0.000000e+00	0.1290323	0.000000e+00	1.860000e+03
26	11065862921025...	HenriikkaH	HHenriikka	10	76	2	3	0	en	0	0	0	0	0.131578947	3.947368e-02	3.000000e-01	0.2258065	1.328571e+01	3.365714e+02
27	11072449180016...	Hanna Kuula	Hansvili	2	26	2	3	0	en	1	1	1	0	0.076923077	1.153846e-01	1.500000e+00	0.1612903	1.860000e+01	1.612000e+02
28	11073502403639...	nasir	nasir76404068	0	15	1	0	0	en	0	1	1	1	0.000000000	0.000000e+00	NaN	0.1612903	0.000000e+00	9.300000e+01
29	10138675712826...	Jiina	haapanenjiina	6	71	3	100	0	fi	0	1	0	1	0.084507042	1.408451e+00	1.666667e+01	8.6451613	1.156716e+01	8.212687e+00
30	92949939896228...	Marco Ranieri	MarcoRanieri_EU	479	375	1427	3281	0	en	0	0	0	0	1.277333333	8.749333e+00	6.849687e+00	16.3225806	2.010099e+02	2.297431e+01
31	34032712	denise preston	deester39	162	516	347	144	0	en	0	0	0	0	0.313953488	2.790698e-01	8.888889e-01	119.0322581	1.209756e+00	4.334959e+00
32	752375106	AYHAN ÇELİK	ftr235345420232	458	4611	3903	3190	0	tr	0	0	0	0	0.099327695	6.918239e-01	6.965066e+00	79.3225806	4.021553e+01	5.812973e+01
33	10500058649612...	Tuomas Saramäki	SaramakiTuomas	35	603	1	5	0	en	0	0	0	0	0.058043118	8.291874e-03	1.428571e-01	5.3870968	9.281437e-01	1.119341e+02
34	11076491294407...	Pia Stjernvall	StjernvallPia	9	45	1	0	0	en	0	1	0	0	0.200000000	0.000000e+00	0.000000e+00	0.1290323	0.000000e+00	3.487500e+02
35	10757369313828...	Samma_ Linje	LinjeSamma	56	228	48	199	0	en	0	0	1	0	0.245614035	8.728070e-01	3.553571e+00	3.0645161	6.493684e+01	7.440000e+01
36	11076370633718...	Fred Hunt	FredHun09378107	0	221	0	0	0	en	1	1	1	1	0.000000000	0.000000e+00	NaN	0.1290323	0.000000e+00	1.712750e+03
37	2334268331	k lee	yourewithme	709	726	7630	14608	0	en	0	0	0	0	0.976584022	2.012121e+01	2.060367e+01	61.4193548	2.378403e+02	1.182038e+01
38	3980897555	Asap Wise	fashioninushka	1	71	6	8	0	en	0	1	1	1	0.014084507	1.126761e-01	8.000000e+00	41.1612903	1.943574e-01	1.724922e+00
39	419725777	CanaDane Dork	yongeknife	280	204	30111	27483	0	en	0	0	0	0	1.372549020	1.347206e+02	9.815357e+01	87.8882488	3.127039e+02	2.321129e+00
40	18077033	Martin Borgs	martinborgs	7504	2756	6663	11322	0	en	0	0	0	0	2.722786647	4.108128e+00	1.508795e+00	123.3225806	9.180800e+01	2.234789e+01
41	58203191	Sari Sivonen	SariSivonen	3	55	2	169	0	en	0	0	0	1	0.054545455	3.072727e+00	5.633333e+01	116.0967742	1.455682e+00	4.737427e-01
42	11076174105201...	Jose Vicente Miguel	JoseVicenteMig2	6	92	1	2	0	en	0	1	1	1	0.065217391	2.173913e-02	3.333333e-01	0.1290323	1.550000e+01	7.130000e+02
43	11075924838175...	Jessica	Jessica13204004	0	5	0	0	0	fi	1	1	1	1	0.000000000	0.000000e+00	NaN	0.1290323	0.000000e+00	3.875000e+01
44	11075983331142...	taneli kesäläinen	TKesalainen	3	56	1	1	0	en	0	0	0	1	0.053571429	1.785714e-02	3.333333e-01	0.1290323	7.750000e+00	4.340000e+02
45	10584982798587...	Susan Duffy	Susan23Duffy	11	86	50	27	0	en	0	0	1	1	0.127906977	3.139535e-01	2.454545e+00	4.6451613	5.812500e+00	1.851389e+01
46	11064466611009...	Maarit Mustonen	MaaritMustonen	1	37	0	3	0	en	0	1	1	0	0.027027027	8.108108e-02	3.000000e+00	0.2258065	1.328571e+01	1.638571e+02
47	11075942907908...	Damaris Mulote	wamulote	0	12	0	0	0	en	1	1	1	1	0.000000000	0.000000e+00	NaN	0.1290323	0.000000e+00	9.300000e+01
48	93834502141313...	Ben Carter	bentransforms	80	304	118	169	0	en	0	0	0	0	0.263157895	5.559211e-01	2.112500e+00	15.5161290	1.089189e+01	1.959252e+01
49	96382189006119...	Henni Tenhunen	HenniTenhunen	23	194	9	4	0	fi	0	0	0	0	0.118556701	2.061856e-02	1.739130e-01	13.2580645	3.017032e-01	1.463260e+01
50	10996825042551...	Richard	YeeeetMeToLife	16	70	133	58	0	en	0	0	0	0	0.228571429	8.285714e-01	3.625000e+00	0.8525346	6.803243e+01	8.210811e+01
51	962643066	Boujaques	JoK_boujaques	323	3291	2314	10215	0	en	0	1	1	0	0.098146460	3.103920e+00	3.162539e+01	76.0322581	1.343509e+02	4.328426e+01
52	80722822069735...	Andreas Walstad	WalstadAndreas	139	394	121	83	0	en	0	1	0	0	0.352791878	2.106599e-01	5.971223e-01	27.4193548	3.027059e+00	1.436941e+01

Appendix C: Examples of bots detected by the model

@JarmoPikkuaho is an account in the M&M and following 21 accounts groups.



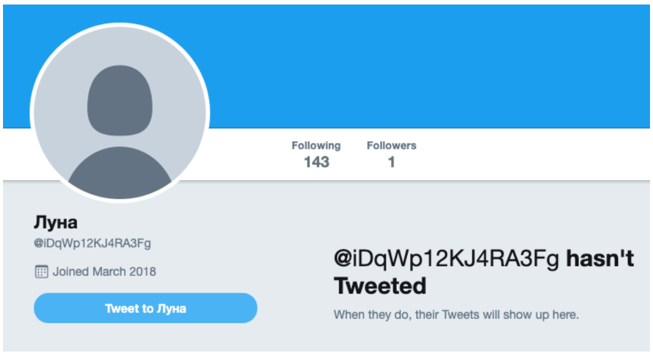
Following 21 Followers 1

Jarmo tapio pikkuaho
@JarmoPikkuaho
Joined May 2018

M&M
@marmai

Juha Sipilä
@juhasipila

@iDqWp12KJ4RA3Fg and @oL3Oe5OxNFJ86tL are accounts belonging to the group of foreign language bots.

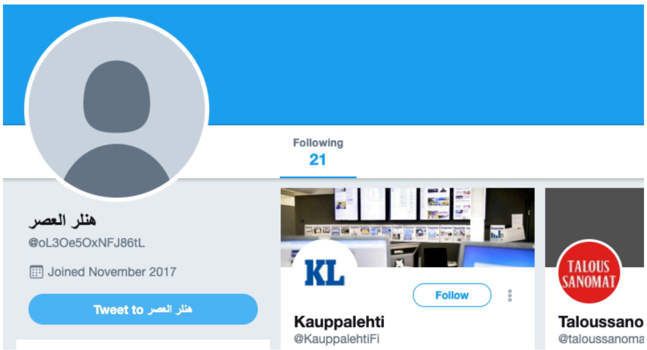


Following 143 Followers 1

Луна
@iDqWp12KJ4RA3Fg
Joined March 2018

@iDqWp12KJ4RA3Fg hasn't Tweeted

When they do, their Tweets will show up here.



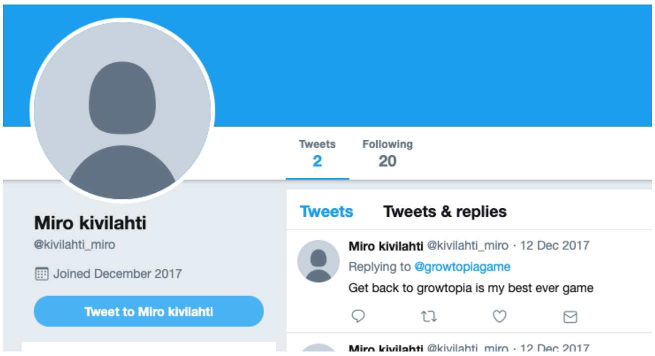
Following 21

هلا العصر
@oL3Oe5OxNFJ86tL
Joined November 2017

Kauppalehti
@KauppalehtiFi

Taloussano
@taloussanoma

@kivilahti_miro is an account in the group of Growtopia bots.



Tweets 2 Following 20

Miro kivilahti
@kivilahti_miro
Joined December 2017

Miro kivilahti @kivilahti_miro · 12 Dec 2017
Replying to @growtopiagame
Get back to growtopia is my best ever game

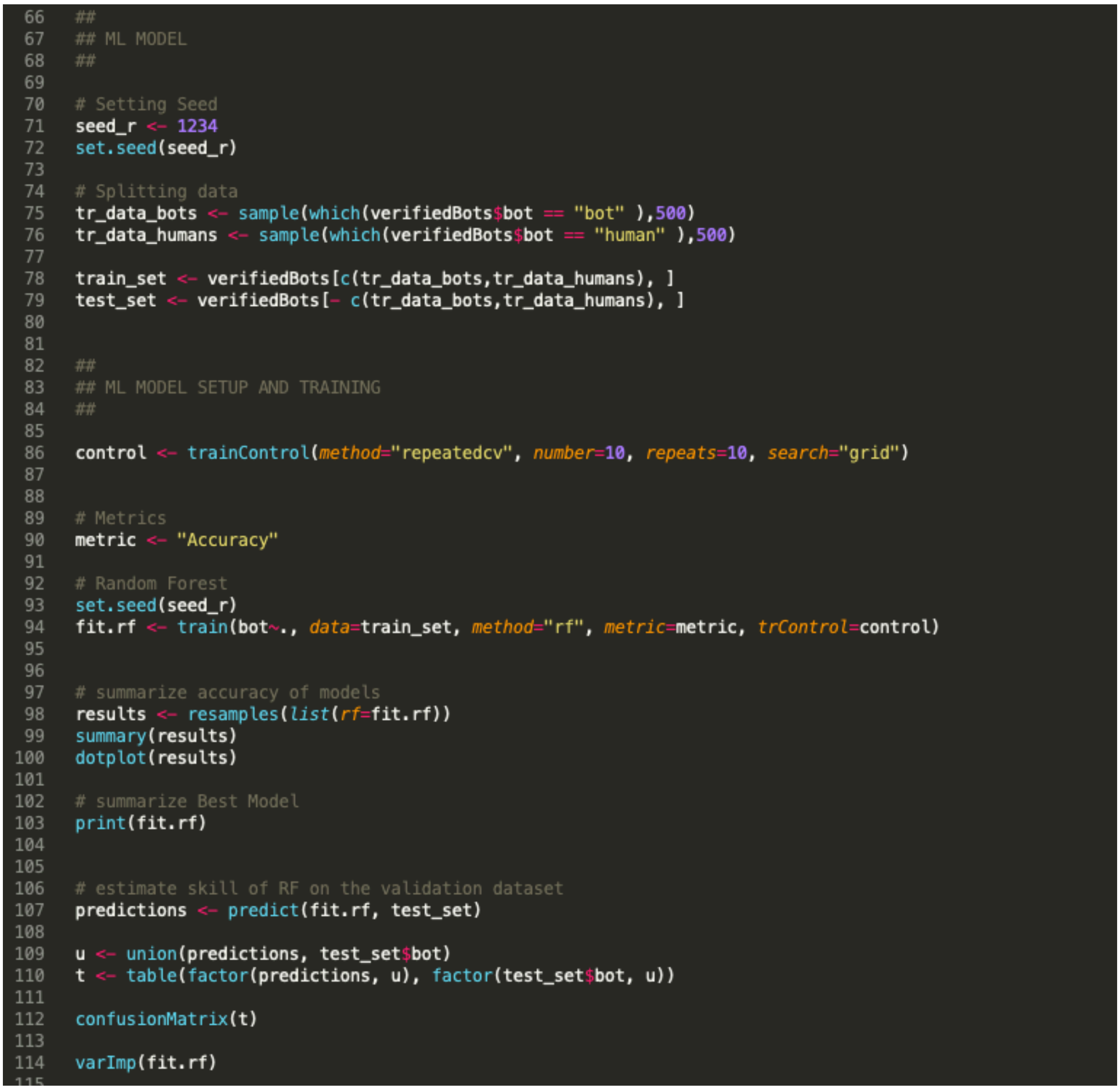
Miro kivilahti @kivilahti_miro · 12 Dec 2017

Appendix D: Network graphs

The following link leads to higher resolution online version of the network graphs where the account names are readable:

https://drive.google.com/open?id=1Vao2JneZHU5lTNulZmGgqG1QDLaz7Fzp

Appendix E: Bot detection model script

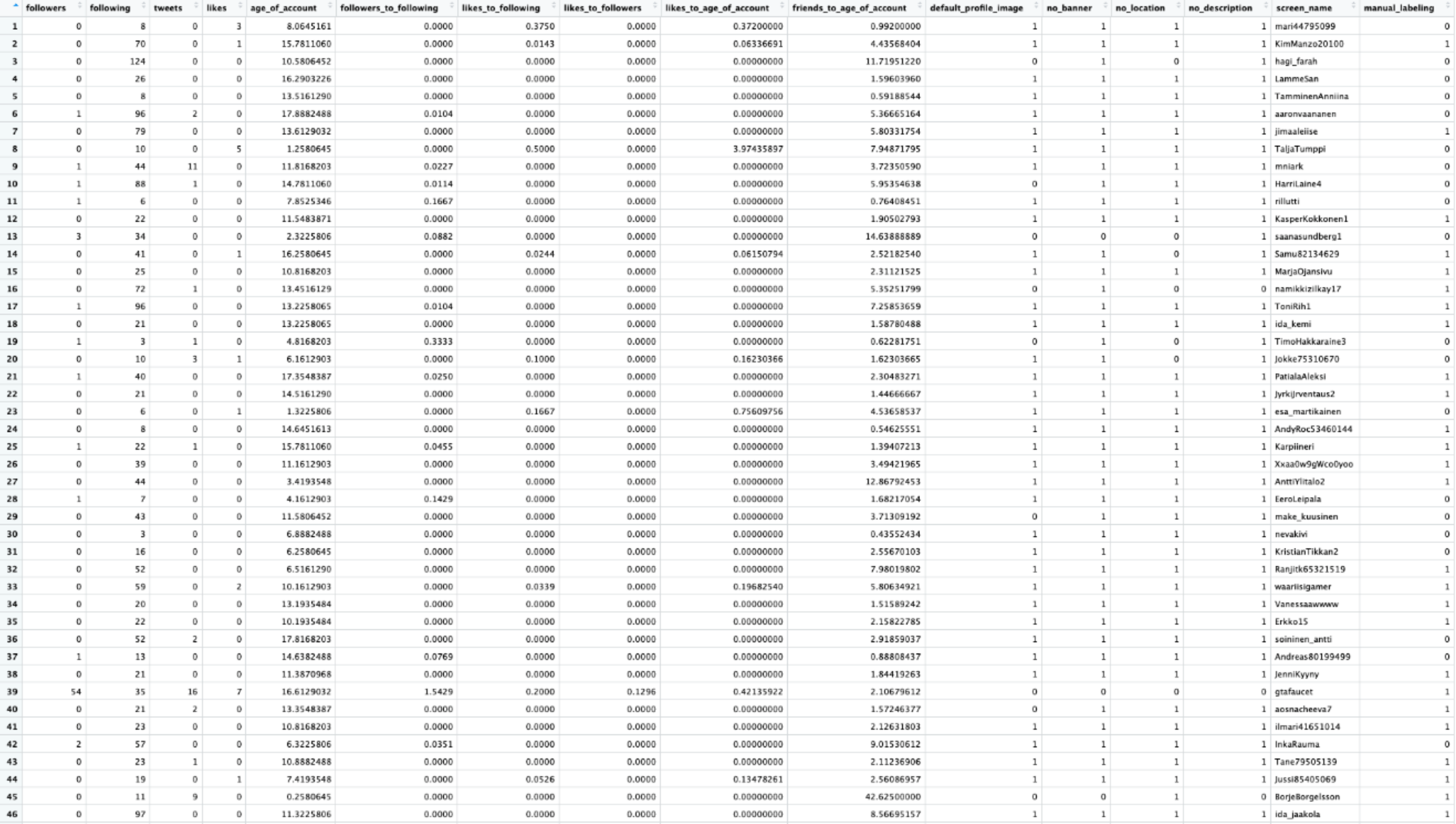


```
##
## ML MODEL
##

# Setting Seed
seed_r <- 1234
set.seed(seed_r)

# Splitting data
tr_data_bots <- sample(which(verifiedBots$bot == "bot" ),500)
tr_data_humans <- sample(which(verifiedBots$bot == "human" ),500)

train_set <- verifiedBots[c(tr_data_bots,tr_data_humans), ]
test_set <- verifiedBots[- c(tr_data_bots,tr_data_humans), ]


##
## ML MODEL SETUP AND TRAINING
##

control <- trainControl(method="repeatedcv", number=10, repeats=10, search="grid")


# Metrics
metric <- "Accuracy"

# Random Forest
set.seed(seed_r)
fit.rf <- train(bot~., data=train_set, method="rf", metric=metric, trControl=control)


# summarize accuracy of models
results <- resamples(list(rf=fit.rf))
summary(results)
dotplot(results)

# summarize Best Model
print(fit.rf)


# estimate skill of RF on the validation dataset
predictions <- predict(fit.rf, test_set)

u <- union(predictions, test_set$bot)
t <- table(factor(predictions, u), factor(test_set$bot, u))

confusionMatrix(t)

varImp(fit.rf)
```

Full script available at:

https://drive.google.com/open?id=19pCe5mLat9L34BqJznoYPk5aBVQW683i

Appendix F: Sample view of the training data



	followers	following	tweets	likes	age_of_account	followers_to_following	likes_to_following	likes_to_followers	likes_to_age_of_account	friends_to_age_of_account	default_profile_image	no_banner	no_location	no_description	screen_name	manual_labeling
1	0	8	0	3	8.0645161	0.0000	0.3750	0.0000	0.37200000	0.99200000	1	1	1	1	mari44795099	0
2	0	70	0	1	15.7811060	0.0000	0.0143	0.0000	0.06336691	4.43568404	1	1	1	1	KimManzo20100	1
3	0	124	0	0	10.5806452	0.0000	0.0000	0.0000	0.00000000	11.71951220	0	1	0	1	hagi_farah	0
4	0	26	0	0	16.2903226	0.0000	0.0000	0.0000	0.00000000	1.59603960	1	1	1	1	LammeSan	0
5	0	8	0	0	13.5161290	0.0000	0.0000	0.0000	0.00000000	0.59188544	1	1	1	1	TamminenAnniina	0
6	1	96	2	0	17.8882488	0.0104	0.0000	0.0000	0.00000000	5.36665164	1	1	1	1	aaronvaananen	0
7	0	79	0	0	13.6129032	0.0000	0.0000	0.0000	0.00000000	5.80331754	1	1	1	1	jimaaleiise	1
8	0	10	0	5	1.2580645	0.0000	0.5000	0.0000	3.97435897	7.94871795	1	1	1	1	TaljaTumppi	0
9	1	44	11	0	11.8168203	0.0227	0.0000	0.0000	0.00000000	3.72350590	1	1	1	1	mniark	0
10	1	88	1	0	14.7811060	0.0114	0.0000	0.0000	0.00000000	5.95354638	0	1	1	1	HarriLaine4	0
11	1	6	0	0	7.8525346	0.1667	0.0000	0.0000	0.00000000	0.76408451	1	1	1	1	rillutti	0
12	0	22	0	0	11.5483871	0.0000	0.0000	0.0000	0.00000000	1.90502793	1	1	1	1	KasperKokkonen1	1
13	3	34	0	0	2.3225806	0.0882	0.0000	0.0000	0.00000000	14.63888889	0	0	0	1	saanasundberg1	0
14	0	41	0	1	16.2580645	0.0000	0.0244	0.0000	0.06150794	2.52182540	1	1	0	1	Samu82134629	1
15	0	25	0	0	10.8168203	0.0000	0.0000	0.0000	0.00000000	2.31121525	1	1	1	1	MarjaOjansivu	1
16	0	72	1	0	13.4516129	0.0000	0.0000	0.0000	0.00000000	5.35251799	0	1	0	0	namikkizilkay17	1
17	1	96	0	0	13.2258065	0.0104	0.0000	0.0000	0.00000000	7.25853659	1	1	1	1	ToniRih1	1
18	0	21	0	0	13.2258065	0.0000	0.0000	0.0000	0.00000000	1.58780488	1	1	1	1	ida_kemi	1
19	1	3	1	0	4.8168203	0.3333	0.0000	0.0000	0.00000000	0.62281751	0	1	0	1	TimoHakkaraine3	0
20	0	10	3	1	6.1612903	0.0000	0.1000	0.0000	0.16230366	1.62303665	1	1	0	1	Jokke75310670	0
21	1	40	0	0	17.3548387	0.0250	0.0000	0.0000	0.00000000	2.30483271	1	1	1	1	PatialaAleksi	1
22	0	21	0	0	14.5161290	0.0000	0.0000	0.0000	0.00000000	1.44666667	1	1	1	1	JyrkiJrventaus2	1
23	0	6	0	1	1.3225806	0.0000	0.1667	0.0000	0.75609756	4.53658537	1	1	1	1	esa_martikainen	0
24	0	8	0	0	14.6451613	0.0000	0.0000	0.0000	0.00000000	0.54625551	1	1	1	1	AndyRoc53460144	1
25	1	22	1	0	15.7811060	0.0455	0.0000	0.0000	0.00000000	1.39407213	1	1	1	1	Karpiineri	1
26	0	39	0	0	11.1612903	0.0000	0.0000	0.0000	0.00000000	3.49421965	1	1	1	1	Xxaa0w9gWco0yoo	1
27	0	44	0	0	3.4193548	0.0000	0.0000	0.0000	0.00000000	12.86792453	1	1	1	1	AnttiYlitalo2	1
28	1	7	0	0	4.1612903	0.1429	0.0000	0.0000	0.00000000	1.68217054	1	1	1	1	EeroLeipala	0
29	0	43	0	0	11.5806452	0.0000	0.0000	0.0000	0.00000000	3.71309192	0	1	1	1	make_kuusinen	0
30	0	3	0	0	6.8882488	0.0000	0.0000	0.0000	0.00000000	0.43552434	1	1	1	1	nevakivi	0
31	0	16	0	0	6.2580645	0.0000	0.0000	0.0000	0.00000000	2.55670103	1	1	1	1	KristianTikkan2	0
32	0	52	0	0	6.5161290	0.0000	0.0000	0.0000	0.00000000	7.98019802	1	1	1	1	Ranjitk65321519	1
33	0	59	0	2	10.1612903	0.0000	0.0339	0.0000	0.19682540	5.80634921	1	1	1	1	waarlisigamer	1
34	0	20	0	0	13.1935484	0.0000	0.0000	0.0000	0.00000000	1.51589242	1	1	1	1	Vanessaawwww	1
35	0	22	0	0	10.1935484	0.0000	0.0000	0.0000	0.00000000	2.15822785	1	1	1	1	Erkko15	1
36	0	52	2	0	17.8168203	0.0000	0.0000	0.0000	0.00000000	2.91859037	1	1	1	1	soininen_antti	0
37	1	13	0	0	14.6382488	0.0769	0.0000	0.0000	0.00000000	0.88808437	1	1	1	1	Andreas80199499	0
38	0	21	0	0	11.3870968	0.0000	0.0000	0.0000	0.00000000	1.84419263	1	1	1	1	JenniKyyny	1
39	54	35	16	7	16.6129032	1.5429	0.2000	0.1296	0.42135922	2.10679612	0	0	0	0	gtafaucet	1
40	0	21	2	0	13.3548387	0.0000	0.0000	0.0000	0.00000000	1.57246377	0	1	1	1	aosnacheeva7	1
41	0	23	0	0	10.8168203	0.0000	0.0000	0.0000	0.00000000	2.12631803	1	1	1	1	ilmari41651014	1
42	2	57	0	0	6.3225806	0.0351	0.0000	0.0000	0.00000000	9.01530612	1	1	1	1	InkaRauma	0
43	0	23	1	0	10.8882488	0.0000	0.0000	0.0000	0.00000000	2.11236906	1	1	1	1	Tane79505139	1
44	0	19	0	1	7.4193548	0.0000	0.0526	0.0000	0.13478261	2.56086957	1	1	1	1	Jussi85405069	1
45	0	11	9	0	0.2580645	0.0000	0.0000	0.0000	0.00000000	42.62500000	0	0	1	0	BorjeBorgelsson	1
46	0	97	0	0	11.3225806	0.0000	0.0000	0.0000	0.00000000	8.56695157	1	1	1	1	ida_jaakola	1