An Exploration of Political Facebook Posts and Their Public's Reactions.

Capstone Project

March, 2019

Submitted by Paola Morrone

Email: morronep@mcmaster.ca

Phone: 289. 208.7784

Submitted to: Dr. Alexandre Sevigny, Dr. Philip Savage

Course: MCM 740 Capstone Project

Department of Communication Studies and Multimedia

Faculty of Humanities, McMaster University

Abstract

This research explored how and to what extent do political Facebook posts engage their followers; to what extent does the Facebook post's content and features influence the follower's response; and to what extent can a model predict the followers' responses to a particular post. The research was based on all the Facebook posts published on the Liberal MP's official Facebook pages including their associated follower responses over the course of one year. The data was classified using quantitative methods; post and response messages were subjected to content analysis to identify topics and sentiment; and predictive modeling was used to predict engagement levels and response sentiment. The results showed that the Facebook messages posted by the Liberal members of parliament over the lapse of one non-elections year, neglected, for the most part, to engage citizens in either voicing their opinions, sharing the messages or sharing their emotions regarding the topics at hand. Although the correlations between each individual Facebook post features and the followers' response were weak, the study found that shorter messages received more positive follower engagement; posts with higher positive connotation received more positive comments as well. Similarly, videos and links receive more positive engagement. The results also showed that it is possible to predict follower engagement and response sentiment within a 30% error with the most significant predictors being the post message length, the time of publication, and the post sentiment. The study recommends further research using other social media venues and different political parties as well as incorporating demographic data as predictors.

Keywords: Political social media, political engagement, citizen engagement, machine learning, sentiment analysis.

Background

START WITH POLITICIAN SOCIAL MEDIA BLUNDERS Among the multiple critiques on this incident, was the use of social media to conduct foreign policy as opposed to using more formal methods to engage and discuss delicate matters. The Twitter post not only triggered disruption to the relationship between both countries, it inevitably affected the trust of Canadians in the capabilities of the country's leadership.

This research explored how and to what extent machine learning and predictive analytics can be used to predict the public's reactions to governmental social media posts. The research focused on the Facebook posts from a North American political party and the associated comments and reactions from the public. The applicability of machine learning and predictive models to social media posts will aid in proactively evaluating whether the communications strategy is meeting its objectives and in mitigating the inadvertent causation of a PR crisis.

Literature Review

In the last two decades, social networks have given the mass population unparalleled access to share their opinions publicly. In Canada alone, the internet has a penetration rate of 92.7% (Statista, 2018), making it the second mass communication medium after television. It is not surprising then that governments have included the internet as part of their communication strategies. Ever since, researchers have studied this phenomenon in order to understand how governments can best make use of new technologies within their strategic communications campaigns, how the data produced on social networks can be utilised to measure the outputs and outcomes of their communications efforts and, furthermore, how to predict future public political engagement.

Impacts of data science in modern politics

Before the dawn of the social media era, electoral campaigns were restricted not only by economic resources, but by time and, most importantly, by the limited methods available to gauge the emotions of potential voters. Political campaigns were focussed on persuading the masses to vote and the political message was geared to the overall perception of the population's needs.

The evolution of technology has shifted consumer behaviours, increasing drastically the amount of personal data stored which, in turn, has enabled the creation of algorithms that provide valuable insights that have changed the landscape of political campaigning worldwide. Data science provides the means to maximise the effectiveness of campaigns by identifying different public segments based on a multitude of factors (Bennet & Bayley, 2018; Issenberg, 2012; Siegel, 2013) such as micro-targeted messages, and identifying undecided voters, and those who are most likely to convert with minimal polling and canvassing effort (Siegel, 2013; Bennet & Bayley, 2018). With these insights at hand, the campaign team can then determine those tailored messages and distribution channels with the highest probability of success. During the campaign, canvassers gather additional information that is entered into the systems to be used to keep improving the machine learning algorithms (Bennett & Bayley, 2018).

Data-driven campaigns have allowed political parties to involve the voters in the campaign process rather than only receiving a political message. Social media has been instrumental in increasing reach since voters themselves are sharing the message (Issenberg, 2012). Additionally, machine learning algorithms are evolving to identify the sentiment and preferences of past voters and further forecast the behaviour of future voters, hence predicting the results of elections (Issenberg, 2012). Data analytics have also enabled campaign teams to identify which population segment hasn't made it to the polling stations and to create specific messages to persuade them to actually cast a vote (Siegel, 2013; Bennett & Bayley, 2018).

Although insights from the use of big data in political campaigns is still in its infancy, and on-going research continues to find new political applications of data science, current privacy concerns may limit the kind of analysis allowed by political parties. After the Cambridge Analytica scandal early in 2018, privacy concerns have escalated world-wide triggering the review of privacy laws and regulations.

Strategic use of social media to cultivate citizen-government relationships

Initially, government websites consisted of listings of policies, bylaws and services. With the communications shift caused by social networks, political communications have been adapting by adding more interactivity to websites with the purpose of engaging the public and improving government-citizen relationships. The shift in e-government is still in its infancy and studies have shown that it has not kept pace with the growth of social media or with meeting public expectations (Sandoval-Almazán & Gil-García, 2012; Welch et al, 2004).

Social networks and digital media have proven powerful in facilitating the creation of different publics worldwide, as they are conduits for people motivated by their life situations to find each other and share opinions in regards to a common interest, in line with the situational theory of publics (Grunig, 1997). Social networks have enabled publics to control messages - as they always have regardless of the medium (Kim as cited by Grunig, 2013). These opinions and perceptions will directly affect the organisation's reputation, image and brand (Grunig & Huang, 2000). Furthermore, messages in social media express the different cognitions and interpretations publics perceived by individuals when they are exposed to different messages disseminated by organisations and governments (Grunig, 2009).

"Digital media have dialogical, interactive, relational, and global properties that make them perfectly suited for a strategic management paradigm of public relations" (Grunig, 2009, p.6), enabling organisations to create strategic, interactive and symmetrical communication programmes which research has shown to be more successful in fostering relationships (Grunig & Huang, 2000). However, government is still using digital technologies mainly to disseminate messages in a unidirectional manner. Although e-government sites have improved government services, limited opportunities have been taken to utilise the interactive capabilities of social networks in order to increase political engagement and improve political trust (Welch, Hinnant & Moon, 2004).

Trust is an essential component of quality relationships (Grunig & Huang,1999), which has proven true not only for organisation-public relationships but also for government-citizen associations (Lendingham, 2001). Although members of the public are generally satisfied with the perception of greater political transparency brought about by an increase in the information disseminated by their government, they are dissatisfied with the lack of interactivity of current government sites (Welch et al, 2004; Bonsón et al. 2012, Bonsón, 2015).

In order to improve citizen's trust as expressed by the individual's viewpoint of overall responsiveness, and the perception that government attends to the citizen's needs, government needs to do much more than create and maintain a good website, and offer quality customer service. It needs to engage individuals with high pre-existing levels of trust if it wishes its online efforts to succeed (Parent et al., 2005).

Use of social network data to predict electoral results

In 2008, the successful Obama election campaign demonstrated an innovative and strategic use of technology where social media took a central role in engaging thousands of citizens across the country and in obtaining an unprecedented degree of active participation throughout the election campaign process. Obama's campaign team "combined more features into the campaign strategy, such as SMS, distributed media, phone tools, and Web capacity to support campaign activities such as donating money, organizing meetings and media events, distributing news, and offering actualities and feeds" (Talbot as cited by Cogburn & Espinoza-Vasquez, 2011, p.198). This electoral campaign used social media to build communities, mobilise supporter networks, facilitate targeted messages and hosted meetings, and raise funds (Cogburn & Espinoza-Vasquez, 2011).

Since the successful use of social media by the Obama campaign in 2008, many studies have focused on the use of social networks in electoral campaigns. Furthermore, studies have researched the use of the data collected from different digital platforms to predict election results (Ceron, Curini & Iacus, 2015; Tumasjan et al. 2010; Franch 2013). While using different approaches and methodologies, research has shown that as long as the demographic of social media users is a representative sample of the total voter population, the data extracted from social media enables the accurate prediction of election results. For instance, Ceron et al. (2015) performed a supervised content analysis of different social networks and web pages to measure voter satisfaction (or dissatisfaction) with each candidate. Their prediction model, in contrast to survey polls, could capture the general sentiment of internet users who may not be predisposed to answer poll surveys, giving this methodology a more accurate prediction capability. Similarly, Franch (2013) used statistical predictive models based on social network data to predict election results, thereby outperforming the polls.

These studies represent only a portion of research found in the literature that demonstrates the value of using social media data to gauge political sentiment during electoral campaigns and to predict election results in an accurate and inexpensive manner.

Using social network data to measure citizen engagement

Further to predicting election results, exploratory research has been employed in assessing the impact of political digital campaigning on overall elections (Hong & Nadler, 2012; Franch, 2013; Jungherr, Jurgens & Schoen, 2012). In other words, how and to what extent can the use of social networks and websites by the political party and by the politician during the electoral campaign influence voters' opinions. However, the mixed results indicate that the use of digital media and e-government is still in its infancy and specific predictors and factors are still to be discovered (Quintelier & Theocharis, 2013).

The literature also highlights the impact that data collection and analysis methods have in utilising social network data to predict elections. For instance, Tumasjan et al (2010) concluded that the quantification of messages in reference to each candidate mirrored the electoral results. However, Jungherr et al. (2012)

countered this conclusion, indicating that their data collection methods were limited without justification from the original authors.

> Outputs of a communication programme measure whether the target audience groups actually received the messages directed at them ... paid attention to them ... understood the messages ... and retained those messages in any shape or form. They also measure whether the communications materials and messages that were disseminated have resulted in any opinion, attitude and/or behavior changes on the part of those targeted publics to whom the messages were directed. (Hon & Grunig 1999, p. 2)

Even though there are many outputs that can be measured based on social media posts, this study is focused on measuring the resulting opinions and follower engagement following a Facebook post which in turn will aid in gauging if the post is meeting its main objective - or not, and in evaluating if the effects of the post in terms of sentiment and engagement correlate with the attributes of good relationship (Hon & Grunig, 1999).

Government social networks seek to engage citizens so that they and government officials interact in online communities to discuss local policy issues, service delivery and regulation with the objective of having a more informed, innovative and citizen-centric government (Bonsón & Ratkai 2017).

The citizen engagement of a social network post can be measured quantitatively by aggregating the popularity (number of likes per post), commitment (number of comments per post) and virality (number of shares per post) of each post (Bonsón, Royo & Ratkai 2015). In addition to the actions taken by the public in response to a post, proposed social network marketing considers the media type of the original post (text only, image, or video), the content category, as well as the date and time that the post was published. These factors influence the levels of engagement and enable the Facebook page administrator to identify the characteristics that result in increased public engagement (Cvijikj, Spiegler, & Michahelles, 2013).

Although there is a vast amount of literature evaluating engagement, these metrics are taken after the fact. There is very little literature regarding the prediction of public engagement based on a Facebook post. Using similar metrics, Moro, Rita & Vala (2016) proposed a predictive public engagement model using text mining processes on a set of posts and their corresponding responses. Their model would predict public engagement based on the original post's content category, the post's date and time, whether or not the post was paid advertisement and the post's reactive engagement from the public. Moro et al. (2016) focused their research on a brand's Facebook page. This research will try to expand the application of the predictive model to government sites. Further, this study will look into predicting public sentiment.

Using technology to determine public sentiment in unstructured texts.

> Sentiment analysis is the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. (Liu, 2012, p. 7)

Sentiment analysis has been the subject of technical, social science and communications research for over a decade. The application of sentiment analysis can be extended to: organisations seeking consumer opinions regarding products they produce or sell and the services they provide; consumers searching online reviews of products, services, or even public opinion on politics or local issues; online social media advertising, where an organisation may monitor consumer reaction and place an advertisement in response to a favourable review of a product; and for general searches of opinions (Feldman, 2013).

While social science and communications research focuses on the application of sentiment analysis, technical research focuses on proposing different methods and frameworks to perform sentiment analysis within three different levels (Collomb, Costea, Joyeux, Hasan & Brunie 2014; Liu, 2012; Medhat, Hassan

& Korashy 2014):

(1) The document level which focuses on one subject being positive or negative;

(2) The sentence level which evaluates each sentence in a document as positive, neutral or negative and then compounds a sentiment based on all the sentences within the document; and

(3) The aspect, or word, level which looks at the opinion being positive or negative and identifies the target or subject of the opinion.

Lexicon-based methods, rule-based methods, and statistical methods are amongst those mostly used for sentiment analysis. The lexicon-based method involves calculating a text's sentiment by using sentiment lexicons in the form of "opinions and phrases" (Medhat et al.,2014). A lexicon can be compiled either manually, by employing a corpus (Medhat et al., 2014), or by using a "Bag of Words (BoW) with ensemble classifiers which are much faster than a supervised approach to sentiment classification while yielding similar accuracy." (Augustyniak et al., 2014). The rule-based method searches for opinion words and then labels the text as positive or negative using classification rules such as dictionary polarity, negation words, booster words and idioms (Collomb et al., 2014). Statistical methods, on the other hand, seek patterns and clusters of sentiment words based on their frequency in the text (Medhat et al, 2014; Collomb et al, 2014).

Research Problem

Although digital media is a widespread method of communication around the world, digital communication campaigns are being monitored and evaluated after the messages are published. On many occasions, even with the best of intentions, public reactions are the opposite of those expected, as was the case with the Dove and RAM commercials. Government digital campaigns are no different. If their digital messages are taken the wrong way by a section of the public, misconceptions can escalate into crises that will inevitably have a negative effect on the government's reputation and the public's trust in government. This research explores correlations between multiple factors of Facebook posts and the public's general reaction; further the research seeks to find a method of proactively evaluating digital campaign messages using machine learning methods, prior to publishing them. The research objective is to provide a method by which, within a certain margin of error, practitioners will be able to predict the public's reactions to their messages. If the model does not return the desired outcome, the practitioner will then have the opportunity to make the necessary amendments in order to increase the probability of the desired outcome. The predictive model can also serve to improve messages that initially have an acceptable outcome, and transform them into messages that increase the public's engagement, and furthermore improve the public's opinion and trust in government agencies.

Research Questions

RQ1: How and to what extent do political Facebook posts engage their followers?

The objective of this research question is to explore if and how are politicians leveraging Facebook's capability of opening 2-way communications with thee multiple publics, or if the medium is instead used as a one-way communication tool no different from traditional media.

RQ2: How and to what extent do the characteristics of a political Facebook posts influence the public's engagement and sentiment?

The purpose of this research question is to find correlations between several Facebook post factors such as post type, posted date, post sentiment, and post content and the public's engagement in terms of comments, reaction and shares. This research question also looks into the correlation between the mentioned post factors and the overall public sentiment (positive, neutral or negative) in response to the post. The purpose of this research question is to explore if the text content and the sentiment of a Facebook post can be correlated to the public's engagement in terms of count of comments, reaction and shares; as well as to the public's sentiment (positive, neutral or negative).

RQ3: How and to what extent can machine learning methods predict the public's reactions to governmental social media posts?

The purpose of the research question is to explore the applicability of machine learning for communication practitioners to evaluate the social media messages prior to posting publicly.

Research Methods

The methodology followed will be a mixed method approach: content analysis of a sample of Facebook posts and their corresponding responses, quantitative analysis of engagement on each post, qualitative analysis of the posts and responses in the form of sentiment analysis of comments and emoticons used to respond.

For the purposes of this study, and in line with the research by Bonsón et al. (2015), the definition of quantitative engagement measurement is the aggregation of the post's popularity - number of likes per post -, commitment – number of comments per post -, and virality – number of shares per post. Furthermore, the definition of engaged follower percentage is measured as the engagement of a Facebook post divided by the number of page followers. The engaged follower percentage metric seeks to put the total engagement in proportion to the number of followers each official Facebook page has and therefore measure what is the proportion of followers that actually respond to a post.

Data Collection Method

This research will use secondary data from publicly available Facebook posts and their corresponding comments. The researcher received anonymized secondary data with the extract of 12 months' worth of Facebook posts –and their related responses - from the official Facebook pages of all Liberal Party Members of Parliament.

For each Facebook post, the data contained a unique anonymized post id, an anonymized author id, post date and time, type of post – video, link, photo, etc. – post message content, number of likes, number of sad reactions, number of "Ha ha" reactions, number of "wow" reactions, number of "love" reactions, number of "angry reactions", amount of unique impressions, post engaged users, and number of page followers. For each response comment collected, the file contained the associated post unique identifier, an anonymized comment identifier, the date and time the comment was posted, and the comment message. Subsequent replies and reactions to the response comments were not collected as part of this research.

Data Analysis Technique

This research followed the data science process starting with identifying the appropriate independent variables were identified based on the data at had and the research questions; followed by performed data and text mining techniques to create a clean structured data set on which to perform quantitative, content, and predictive analysis.

Variables

To analyze the Facebook posts this study included as independent variables, (1) an anonymized identifier for the author of the post, (2) the date of the post, (3) the type of post identifying if the post was text, video, image, a link to a different webpage or an event; (4) Sentiment category of the Facebook book indicating if the content had a positive, negative or neutral content; (5) the sentiment value indicating the level of negativity or positivity of the post message ; and (6) the post topic providing a context of the message.

Pre-processing of post messages

Due to the unstructured nature of the data, this research relied on text mining methods to allow finding patterns, classify and organize the posts' content into a structured dataset that can be fed to predictive models. The unstructured data was subject to a tokenization method that separated – or broke apart – the text of each post into either individual words creating a bag of words – or corpus – to further analyze in terms of the usage of each word (Weiss et al., 2010). To clean the data further, punctuation is eliminated as well as words that do not provide context – referred to as stop words (Weiss et al. 2010). Once the tokens were identified they were lemmatized, which converted each token to its standard form and reduced the number of distinct words consequently increasing the frequency of usage of words with the same meaning (Weiss et al., 2010).

After the post and comment messages were tokenized and lemmatized, a vectorization process is conducted in order to generate a structured data set in which each token can be quantified based on usage in each message. This structured dataset is called a vector and was be the input to the predictive models (Weiss et al, 2010).

Furthermore, the text was categorised and applied to machine learning processes using the random forest classifier, support vector machine methods (SVM), and regression model to evaluate which method best predicts engagement and response sentiment based on the context of the post.

Quantitative Analysis

The data collected was analyzed using statistical qualification methods that classified the Facebook posts by their different characteristics – post type, month of posting, day of week, length of post message, sentiment of post message, and topic of post message.

Furthermore, statistical methods were used to normalize the data. For optimal predictive modeling, it is necessary that the data is distributed normally, therefore the data were normalized using power transformation.

Content Analysis of Post

After the posts were represented in a structured data set, content analysis was performed in order to categorize the posts by type, sentiment, and keywords, topics, post's time of day, post's day of week, and post's author.

The possible types of posts were categorized as either photo, video, link, status, note or event. The possible sentiment categories were either positive, neutral or negative. Sentiment analysis was only applied to the post and comment text messages due to current software limitations. The Vader sentiment analysis software (Hutto & Gilbert, 2014) was used to determine sentiment. A sub-sample was analyzed by the researcher to determine the margin of error.

The post keywords were determined by applying lemmatization logic to the posts which eliminated neutral words and kept action words. Secondly, a quantitative analysis was carried out on the remaining words to find the words mostly used.

Furthermore, correlation between keywords and post reactions was analyzed.

Content Analysis of Comments

The second phase was to perform a content analysis of the responses to the sampled posts in order to categorize the responses by type, sentiment, and keywords.

The possible types of response will be either comments, shares, or reactions which subsequently were categorized as likes, "ha ha", sad, angry, and wow. The possible sentiment of comments were: positive, neutral or negative.

Quantitative analysis was performed on the reaction types and correlation analysis was performed between the reactions and the type, sentiment, as well as with the keywords.

Sentiment analysis was applied to reactions and text comments. This research did not look into links added in the comment statements. As done with the post messages, the Vader sentiment analysis software (Hutto & Gilbert, 2014) was be used to determine sentiment. A sub-sample was analysed by the researcher to determine the margin of error of the automatic tool.

The comment keywords were determined by applying lemmatisation logic to the posts which eliminated neutral words and kept action words. Secondly a quantitative analysis was performed on the remaining words to find the words or topics mostly used. Correlation between response sentiment and post type, sentiment and keywords was performed.

Machine Learning

Machine learning can only be applied to numeric values, therefore the post messages and their corresponding responses needed to be coded and translated into numerical values using text mining techniques. Once the posts and responses were coded, the data were separated into three samples, a training sample and two test samples. The second test sample enabled the researcher to refine the predictive models and retest. Three predictive models were evaluated in this research, Support Vector Machine classifier, Random Forest classifier and XGBoost Regressor. The researcher used the predictive models from the SciKit Learn library for predictive analytics in Python. The results from all models were evaluated for accuracy, precision, and recall.

Results

The most significant finding is the meagreness of follower reactions to the Party's Facebook posts. The descriptive statistics show that the 75% (N=45,410) of the posts have less than 6 shares, less than 5 comments and less than 50 likes. When we put these figures into perspective of the number of people that follow the Facebook page, we observe that 75% of the posts received less than 0.08 shares for every 100 page followers, less than 0.054 comments for every 100 page followers and less than 0.75 likes for every 100 page followers.

The 25% of posts with higher amount of response show an extensive range caused by a small amount of posts that had an exceptional amount of responses - outliers. The maximum number of shares obtained by one of the collected posts was 278,789; the maximum amount of likes received by one post was 402,678; and the maximum amount of comments received by one post was 31,486. These exceptional cases are statistical outliers that were comprised of 158 posts or 0.35% of the total amount of collected Facebook posts.

RQ1: To what extent do political Facebook posts engage their followers?

The objective of this research question is to explore if and how are politicians leveraging Facebook's capability of opening 2-way communications with their Facebook followers, or if the medium is instead used as a one-way communication tool no different from traditional media.

The quantitative analysis of over 45,000 Facebook posts, showed that in total the Party's Facebook posts received a total of 0.29% of follower engagement, based on over 9.9 million shares, comments and likes received and around 3.4 billion no-unique followers in total. Although in one particular extreme case, one post generated as much as 208% follower engagement, only seven percent of all the collected posts produced a follower engagement of two percent or more. In other words, over 93% (N=45,410) of the posts collected, received less than two percent of follower engagement. The analysis showed that out of 159 official caucus Facebook pages, 91%(n=145) had at least one post with over two percent follower engagement in the last twelve months. As shown in **Error! Reference source not found.**, over 2,142 of the posts with most engagement were photo posts, 679 were videos, 308 were links to other websites, 69 were status updates and 2 were notes.

Most of the posts with higher follower engagement were posted in the months of October and November 2018 mostly related to the new NAFTA negotiation that were occurring during that time, remembrance day commemorations, and the introduction of the Viola Desmond banknote that occurred in November 2018. Even though 29% (n=3,200) of the higher engaged posts had a positive sentiment with a compound sentiment value over 0.5, only 15% of the post response comments had a positive sentiment, with the other 85 percent of comments having a neutral sentiment. Interestingly, however, there was only one post with an overall negative response.

Facebook Posts with High Engagement by Type

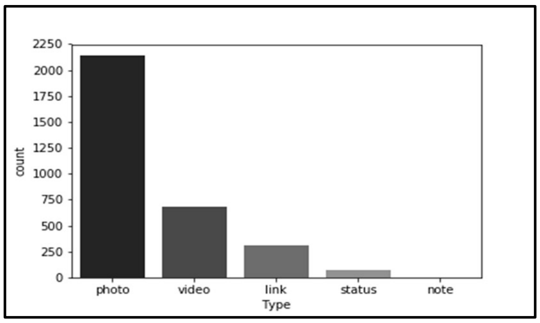

Figure 1: Posts with higher follower engagement

Follower Engagement Breakdown

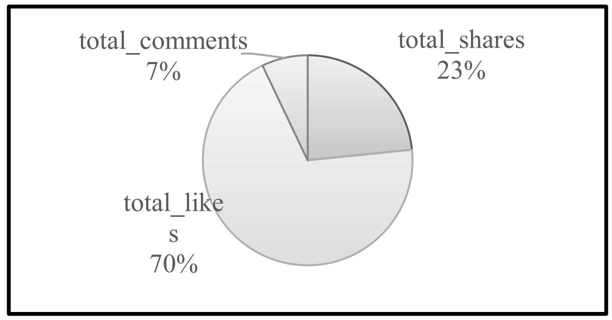

Figure 2: Engagement Breakdown

A significant communication outcome measures whether the message resulted in any opinion (Hon & Grunig, 1999) and Facebook comments provide a venue for followers to provide their opinions and engage in an open two-way communication. However, the results show that the comments received by the collected posts are the lowest contribution to the overall follower engagement suggesting that even though the Facebook infrastructure does have dialogical, interactive, relational, and global properties that would make it well suited for a strategic management paradigm of public relations (Grunig, 2009), overall the Liberal MPs' Facebook posts do not engage their followers in two-way communication.

RQ2: How and to what extent do Facebook post characteristics influence engagement?

To respond to this research question, it is important to understand the follower post engagement for the sample of Facebook posts. The qualifying statistical analysis showed that out of the total amount of posts in the sample (N=45,410), 99% received at least 1 like, 74% of posts were shared at least once, and 63% of posts received at least 1 comment. These preliminary results indicate that at the majority of posts are being looked at. Furthermore, 66% of the total posts collected received at least one reaction emoji. The median measure indicates that in average each post received 24 likes, two shares and one comment. Table 1 contains the descriptive statistics of all the follower reactions to the collected Facebook posts.

Table 1
Statistics of follower responses

Measure	Shares	Comments	Like	Sad	Haha	Wow	Love	Angry
Count	33,452	28,665	45,008	3,431	4,653	5,959	27,952	3,361
Mean	30	16	171	2	2	1	23	1
Std	1,588	250	2,909	140	56	27	693	42
Median	2	1	24	0	0	0	1	0
Max	287,789	31,486	402,678	21,416	10,278	3,628	110,391	7,906

These results indicate that most of the follower reactions to a post consist in liking the post, indicating, according to Bonsón et al. (2015) definition, that these posts are more popular than viral or committed.

Furthermore, the statistics show that the likes are strongly correlated to shares and comments, meaning that the more likes a post attracts, the more likely that post is to be shared and/or commented on. As shown in Table 2, follower likes are also strongly correlated to the love reaction, suggesting that a like reaction can be interpreted as a positive reaction.

Table 2

Correlation between page follower's response

Response type	Shares	Comments	Like	Sad	Haha	Wow	Love	Angry
Shares	1.00	0.80	0.84	0.03	0.15	0.77	0.93	0.20
Comments	0.80	1.00	0.90	0.14	0.24	0.80	0.89	0.40
Like	0.84	0.90	1.00	0.15	0.28	0.86	0.96	0.26
Sad	0.03	0.14	0.15	1.00	0.02	0.09	0.13	0.08
Haha	0.15	0.24	0.28	0.02	1.00	0.22	0.24	0.08
Wow	0.77	0.80	0.86	0.09	0.22	1.00	0.83	0.33
Love	0.93	0.89	0.96	0.13	0.24	0.83	1.00	0.21
Angry	0.20	0.40	0.26	0.08	0.08	0.33	0.21	1.00

Given the low percentage of emoji reactions in the sample, follower engagement metric used for the rest of the research will only include the sum of likes, shares and comments. The Facebook posts collected were categorized by the amount of engagement received. The first category labeled "No Engagement" comprised all the Facebook posts that did not receive any engagement, the second category, labeled "Low Engagement" is comprised by all the posts that received a follower engagement of up to two percent; the third category labeled "High Engagement" contains the Facebook posts that had a follower engagement between two and ten percent; and lastly a category labeled "Extraordinary" containing the Facebook posts with a follower engagement of over 10 percent.

To better understand how each Facebook post characteristic impacts the follower engagement and response sentiment, analysis was performed on each of the independent variables.

How and to what extent does the posted month impact the follower engagement and comment sentiment.

Although the calculated correlation between the Month and the follower engagement is not significant (0.03), the qualifying statistics showed that most of the Facebook posts with no engagement were published towards the beginning of the year, as opposed to most of the posts with high engagement were published in the last quarter of the year. The data also shows that there are less extraordinary posts during the summer months and the low engagement posts are similarly distributed throughout the year.

Distribution of follower engagement by month

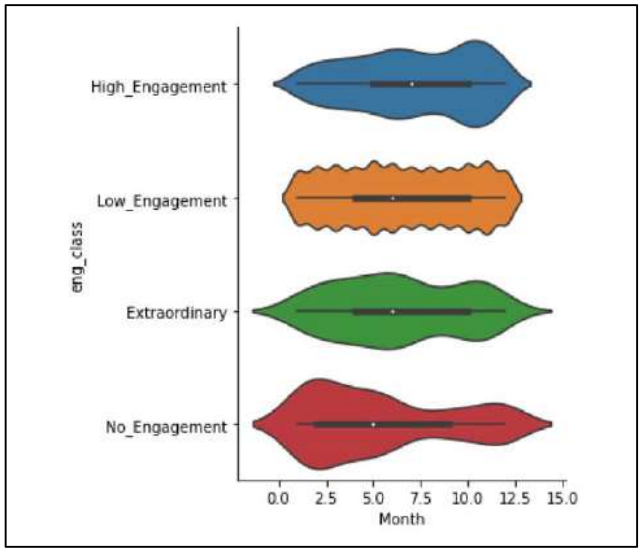

Figure 3: Monthly Follower Engagement

The violin plot depicted in Figure *3*: highlights that the majority of posts with no engagement occurred in the first quarter of the year whereas the majority of posts with extraordinary engagement occurred in the latter months of the year.

Furthermore, when breaking down the data by the post sentiment, we observe that there is only one extraordinary post with a negative sentiment, all other extraordinary posts are either neutral or positive. The data also shows that negative posts have higher follower engagement during the third and fourth quarter of the year.

Monthly Follower Engagement Level grouped by Post Sentiment

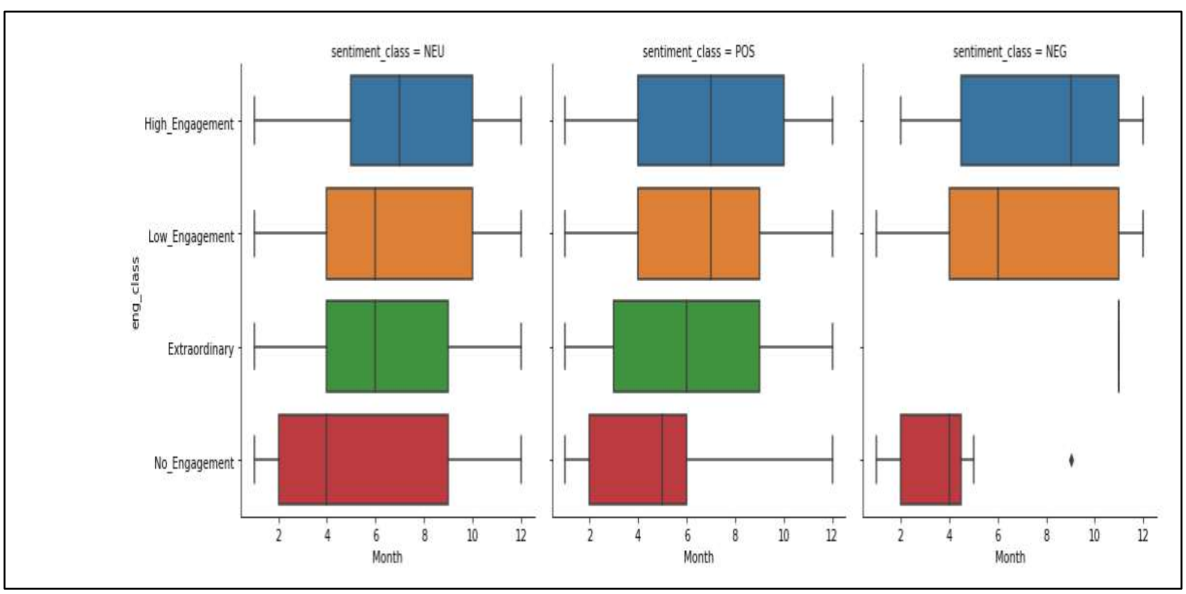

Figure 4: Monthly Follower Engagement grouped by Post Sentiment

Although the calculated correlation between the posted month and the response sentiment is not a significant amount, it is a negative value of -0.01 indicating the little correlation there is, would be an inverse relationship. This is shown to be true in Figure 4 where the violin plot highlights that during the higher months of the year there is a concentration of posts with negative response.

Monthly Response Sentiment

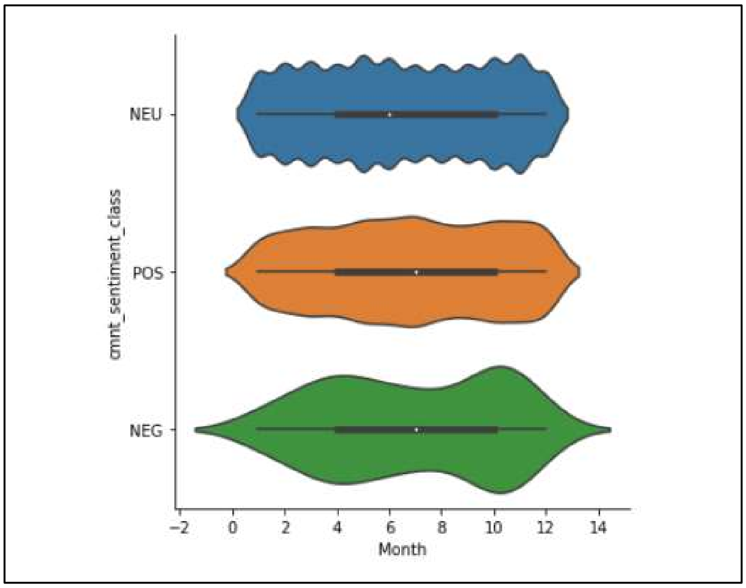

Figure 5: Monthly Response Sentiment Distribution

How and to what extent does the sentiment of the post impact the follower engagement and comment sentiment.

When comparing the level of follower engagement based on post sentiment the data shows that the engagement level distribution is very similar for neutral and positive posts. However, for posts with a negative connotation, there is an insignificant percentage of posts with extraordinary follower engagement; and the higher the level of negativity (or lower the sentiment value), the least level of follower engagement.

Follower Engagement Level grouped by Post Sentiment

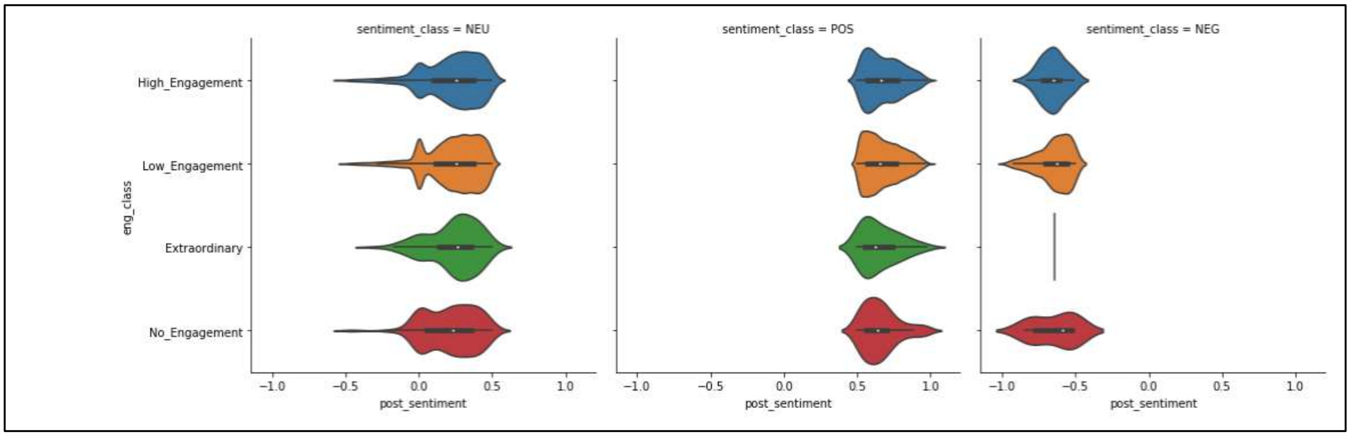

Figure 6: Follower Engagement Level grouped by Post Sentiment

Between the post message sentiment and the response sentiment, there is a calculated correlation of 0.2, indicating a small positive correlation which is mostly due to the quantity of posts with sentiment value less than zero with response sentiment greater than zero as shown below.

Positive Response Sentiment in relation to Post Sentiment

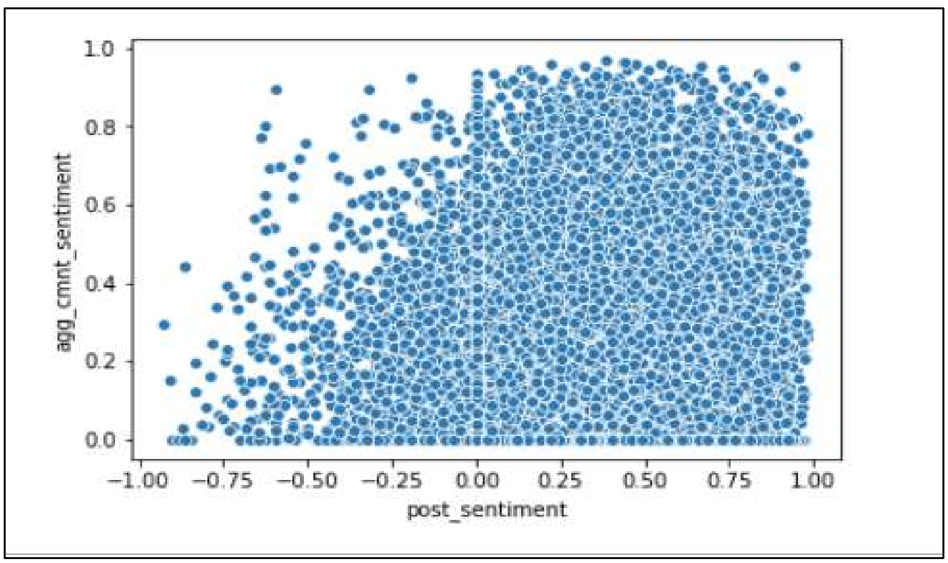

Figure 7: Scatter plot of Post sentiment in relation to the positive response sentiment

Negative Response Sentiment in relation to Post Sentiment

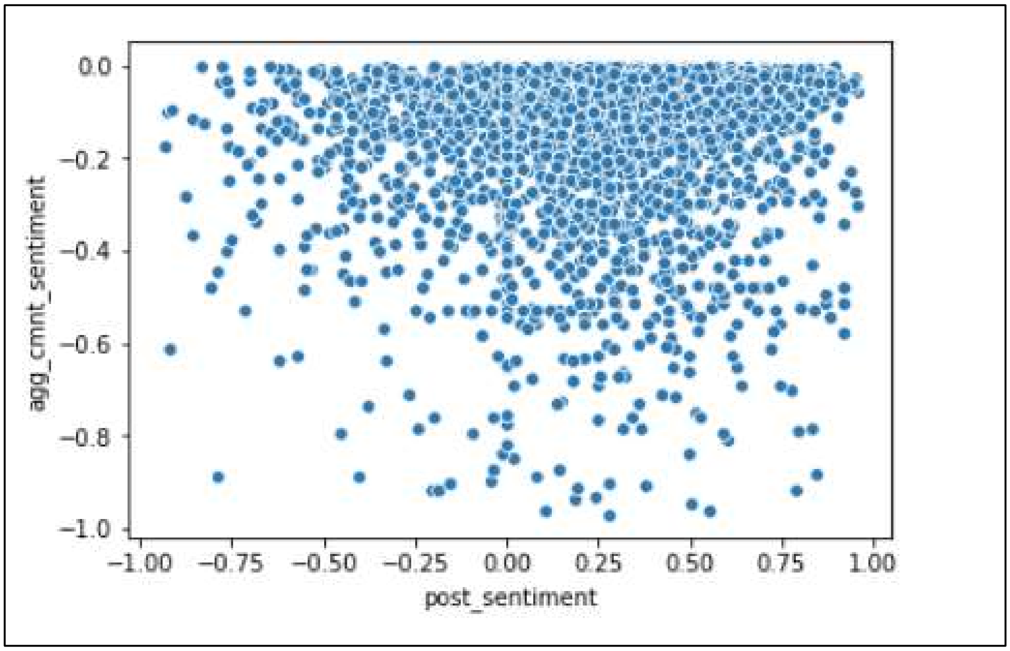

Figure 8: Scatter plot of Post sentiment in relation to the negative response sentiment

How and to what extent does the post message length impact the follower engagement and response sentiment.

The message length measured in number of characters is not limited in Facebook messages. Any type of message can have a message associated with it and this study wanted to explore if the length of the message is an indicator of the followers' engagement. The calculated correlation of 0.02 indicates an insignificant relationship between the message length and the follower engagement. However, it is important to note that for posts with high level of follower engagement, there is a visual downward trend between the message length and the follower engagement, indicating that for this subset of the sample, longer messages receive less follower engagement.

Follower Engagement in relation to Message Length

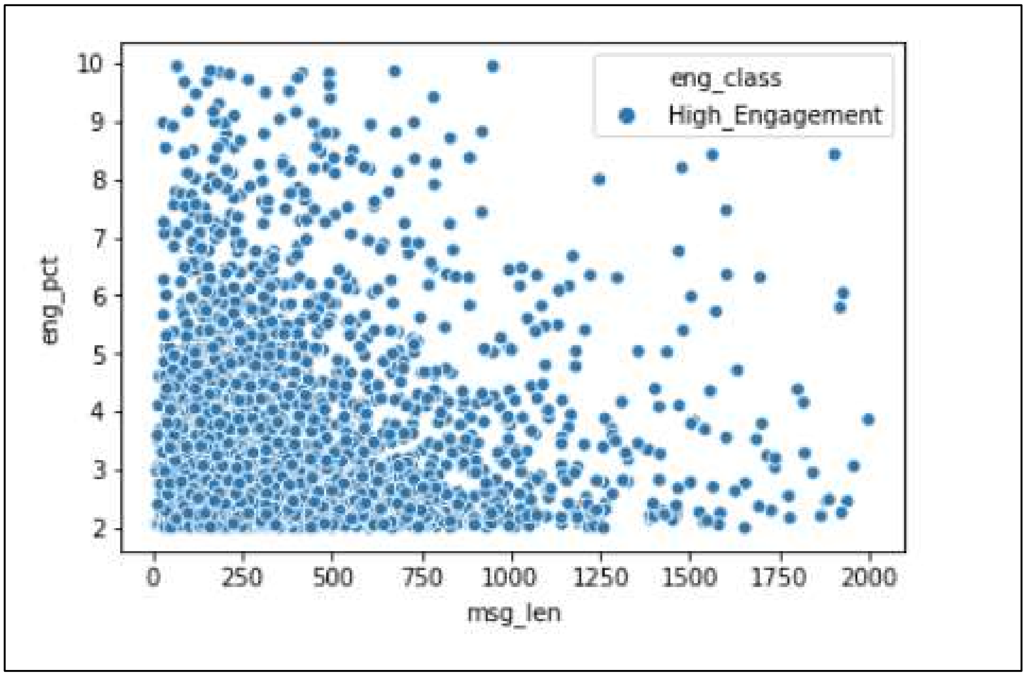

Figure 9: Scatter plot of Follower Engagement in relation to Facebook Post Message Length

Contrarily, the response sentiment improves with the length of the message, especially for the responses with lower sentiment values. The data shows a positive correlation trend for posts with response sentiment lower than zero, however, for posts with respnse sentiment value higher than zero, the correlation trend is inversly proportionate, in other words, the longer the message, the lesser the response sentiment value. According to the collected data, the most concentration of posts with positive response sentiment occurs with messages under 400 characters in length.

Response Sentiment in Relation to Message Length

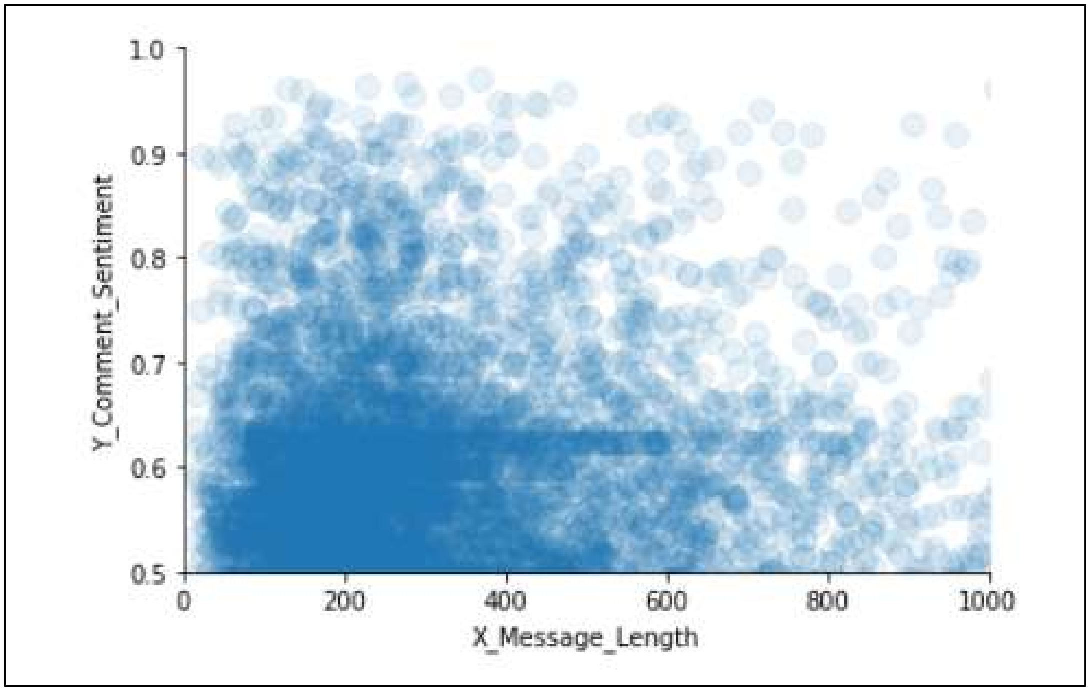

Figure 10: Concentration Facebook posts based on message length and response sentiment.

How and to what extent does the post type impact the follower engagement and response sentiment.

Overall the data shows that 65% (n=21,933) of posts are photo posts followed by video posts comprising a 17%(n=7,943) of the total posts and link posts adding up to a 14% (n=6,538) of the total posts in the sample. However, when comparing which post type received the most follower engagement, the data showed that video posts received the highest follower engagement followed by the photo posts. Nevertheless, photo posts received the highest response sentiment value, indicating that photo posts received more positive response than any other post type.

Follower Engagement by Post Type Response Sentiment by Post Type

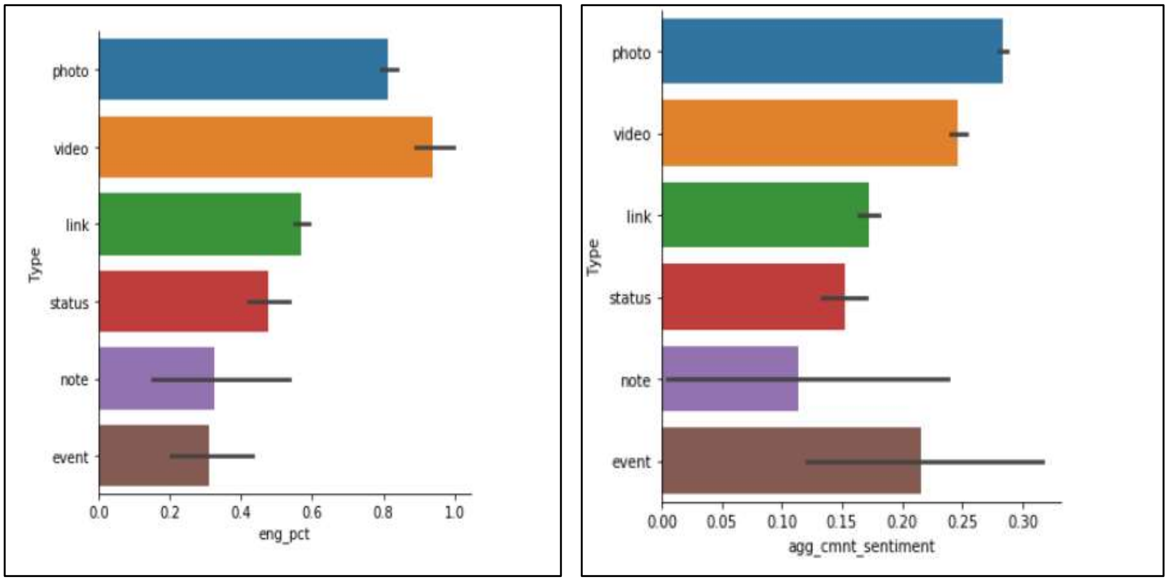

Figure 11: Follower Engagement and Response Sentiment by Post Type

Interestingly, the event post type received higher response sentiment than links, status and notes.

To better understand the impact of the post type on the response sentiment and follower engagement, heat maps were created to find the major concentration of each type of post. The heat maps below corroborate that video posts receive more follower engagement, however, it also highlights peaks of follower engagement on other types such as the high engagement on the event post in February, as well as the high engagement received on the note posts in August and October.

Monthly Average Follower Engagement by Post Type

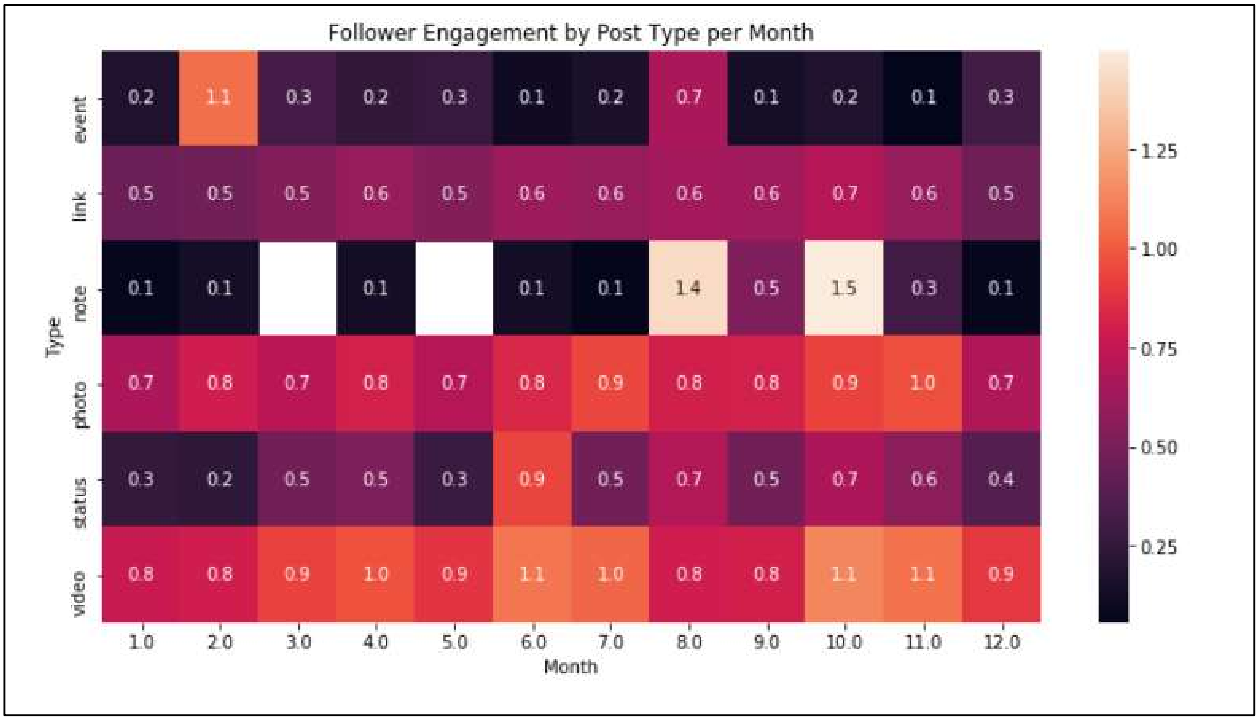

Figure 12: Monthly Average Follower Engagement by Post Type.

To better understand how the post type affects the response sentiment, two heat maps were created, one considering only posts with positive response sentiment – or response sentiment value over 0.5 – and the other heat map considering only the posts that received a negative response sentiment – or response sentiment value less than -0.5. The results show high volatility of response sentiment on status posts, as well as the high variability of negative response sentiment on video posts. These high variable results do not show any trend or pattern to indicate that the post type impacts the follower engagement or response sentiment.

Monthly Posts with Negative Response by Post Type

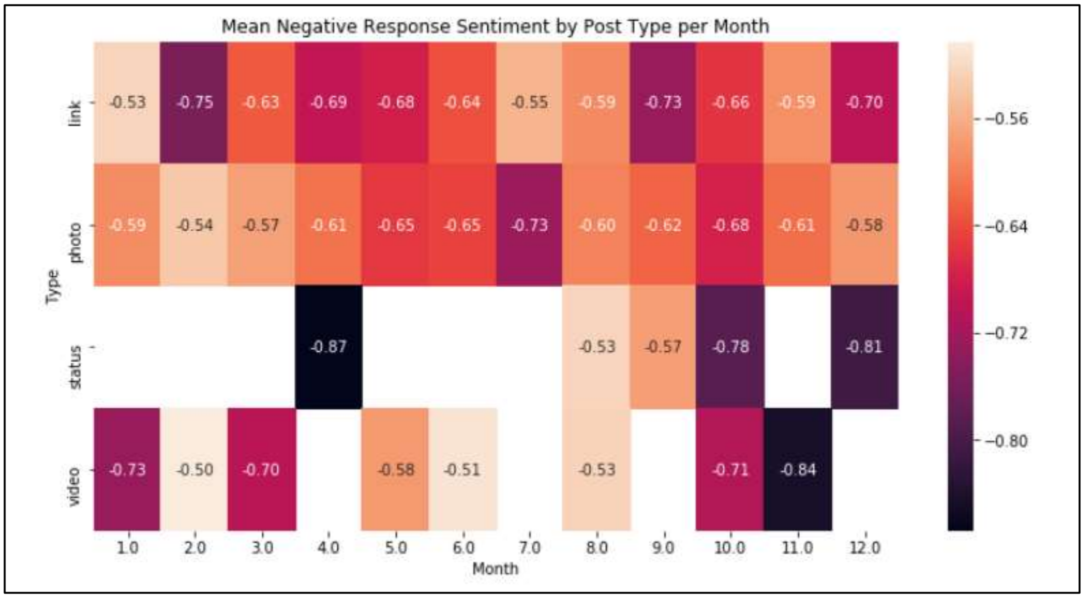

Figure 13: Monthly average response sentiment value by post type of all posts with negative response.

Monthly Posts with Positive Response by Post Type

Figure 14: Monthly average response sentiment value by post type of all posts with positive response.

How and to what extent does the Post content impact the follower engagement and response sentiment?

Beyond looking at the characteristics of the Facebook posts, this study looked into the content of each post message in order to explore if there is any relationship between the content and the overall follower engagement or a relationship between the content and the response sentiment. The content analysis was performed on posts with no engagement, posts with low follower engagement, posts with high engagement, and posts with extraordinary engagement separately to best understand which where the top 20 topics for each category of posts.

There are several topics that overlap amongst posts with different engagement levels, however, posts discussing Canada Child Benefits show up in posts across all engagement levels. Interestingly, almost 11% of posts with extraordinary follower engagement mention Canada's Prime Minister Justin Trudeau range compared to 3% of posts with high follower engagement and 1.2% of posts with low engagement; however, every post mentioning the prime Minister received some form of follower engagement. Although there are other intersections between topics in posts with different engagement levels, for the most part, the topics that receive higher follower engagement are not included in the topics that have lower or no follower engagement. Appendix A contains a detailed list of the top 20 topics for each follower engagement category.

On the other hand, when comparing the top 20 topics of posts that received positive response with the top 20 topics that received a negative response, there is no overlap. It is important to state however that only 0.34% (n=153) of the total posts received negative response in comparison to 10% (n=4,715) of the total posts received a positive response. Appendix B contains the detailed list of top 20 topics that generated either a positive or a negative response.

RQ3: How and to what extent can machine learning methods predict the public's reactions to governmental social media posts?

To answer this research question, the study looked into classification and regression predicting models utilizing the Facebook posts characteristics as well as the top 50 topics as predicting indicators. The classification models were trained to predict either the follower engagement class (extraordinary, high engagement, low engagement or no engagement) and the response sentiment class (positive, neutral or negative). The regression models were trained to predict the follower engagement percentage and the response sentiment value (ranging between -1 and 1).

Random Forest Classifier

The random forest classifier can only predict binary classes, therefore two additional classes were created based on the follower engagement and the response sentiment values. For the follower engagement the classes were defined as "Met engagement goal" which included all the Facebook posts with follower engagement over 0.43 percent and all other posts were categorized as "Not engagement goal". Similarly, a binary class was created for response sentiment, class one included all the posts with positive response sentiment and class zero for all other posts.

When the random forest model was applied to predict the follower engagement class, the result yielded up to a 60% accuracy with a recall of 64% which means that each time the model runs, there is a 64% probability of the model making the same predictions consecutively. The accuracy of predicting the response sentiment class as either positive or anything else, yielded an 89% accuracy, with a recall of 7%.

However, when looking into the results, the accuracy of the Random Forest prediction seemed to be correlated to the percentage of posts within the target class. For instance, approximately 10% of the posts belonged to sentiment class one (positive sentiment) and the other 90% of the posts belonged to the sentiment class zero (neutral and negative sentiment) the algorithm predicted most of the posts in class zero, therefore always receiving an accuracy score of around 90%. In other words, if we have one positive post out of ten posts, the algorithm predicted all the posts as being "not positive" obtaining a 90% accuracy.

Multivariate Support Vector Machine

By classifying the training model into our five classes based on engagement level where class zero (0) are posts with no engagement and class four (4) are posts with extraordinary engagement, the multivariate SVM model returned a poor accuracy of 26%. The matrix shows the predicted values versus the true values. The results indicate that, with the training data available, this model was not the optimal.

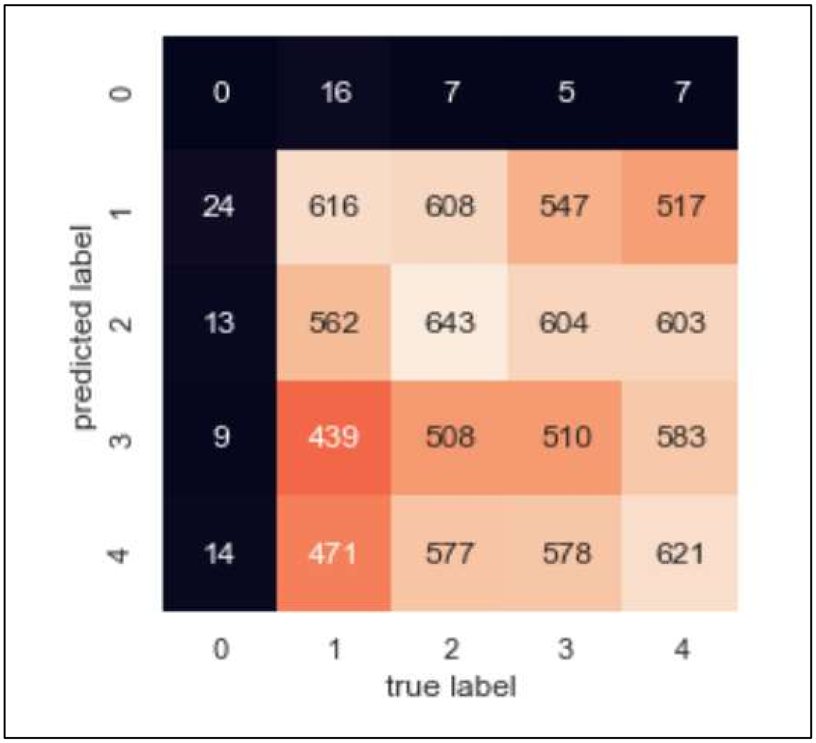

Figure 15: Predicted follower engagement level versus true follower engagement level

The multivariate support vector machine model run on response sentiment class returned an accuracy level of 87% which is in line with the random forest results. However, looking at the confusion matrix, it becomes apparent that the model is predicting most of the values into one category. By looking at the accuracy per category, there are two out of three categories that are very poorly predicted.

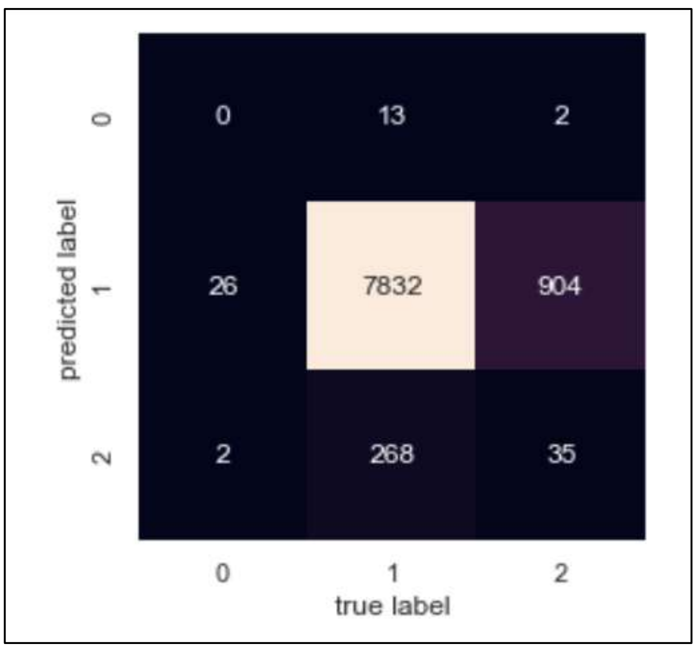

Figure 16: Predicted sentiment category versus true sentiment category

Linear regression Boosting model

From the previous results, it became apparent that the classes created are not discrete enough for a model to distinguish between one class and another without overlapping. Therefore, linear regression model was used in order to predict the numeric values of follower engagement, as well as the numeric value of response sentiment. XGBoost linear regression modeling was used for this purpose. By training the model using the normalized follower engagement values, the model returned a root mean square error of 29%, indicating that each predicted value could be 29% off the real value. The most significant features to predict the follower engagement were the message length, the time of day of the post, the day of week of the post, the post sentiment, month, type of post. The most significant topics were the mention of the Party's leader, posts related to job offers, and particular special campaigns.

Following the same model to predict the response sentiment value, the model returned predicted values with a root mean square error of 24%. Similar to predicting engagement, the most significant features are the post's characteristics. The most significant topic features however are messages with the terms "Look forward to see", job postings, and the mention of the Party's leader.

Discussion

This research generated many insights from a relatively small sample of data. As stated by many researchers, including Issenberg (2012) and Siegel (2013), data science is providing the means to identify different insights of the political landscape based on a multitude of factors, without the need to formally canvassing the different publics. The tools available are increasingly rapidly and can be incorporated into the campaigning tasks to reduce costs and focus on the most meaningful tipping points (Siegel, 2013).

An important aspect of data-driven campaigning is involving the voters in the process rather than only receiving a political message (Issenberg, 2012), however the extremely low average follower engagement found in this study suggests that the Party members are using their Facebook pages to send political messages rather than seeking the involvement of the citizens and therefore missing the opportunity of using Facebook as a tool for maintaining a permanent campaign (Bennett & Bayley, 2018). These results also suggest that the dialogical, interactive and relational aspects of Facebook (Grunig, 2009) are being underutilized. In line with Sandoval-Almazán & Gil-García (2012) and with Welch et al. (2014), this study also suggests that the political use of social media has not met the public expectation. Furthermore, the Facebook comments collected, are not providing enough opinion or perceptions to extrapolate to the overall voters' perceptions about the Party's Caucus, nor to understand the current government-citizen relationship status (Lendingham, 2001). For government's online efforts to succeed, it needs to engage individuals with pre-existing high-levels of trust (Parent et al., 2005); the results however indicate that there is much work to be done in order to increase the levels of engagement and the use of the framework used in this study may be a starting point in creating more engaging posts and proactively measuring engagement and sentiment.

Although one of the major resources that social media provides is the personal information of voters (Bennet & Bayley, 2018); this research found that many valuable insights can still be drawn without the use of personal information, demographics, or consumer preferences. Further research is necessary to validate how and to what extent adding demographic and consumer behaviour data will improve the predictive models.

Many studies have been done to predict elections during the election year, however Bennett and Bayley (2018) suggest that the political landscape is shifting towards the permanent campaign, the results of this research provide a framework to continuously gauge engagement and sentiment at all times and not limited to election campaigning without utilising massive resources.

This study focused on measuring the resulting sentiment and follower engagement - based on the posts virality, popularity and commitment- associated to a Facebook post in order to proactively gauge if the post is meeting its main objective - or not, and in evaluating if the effects of the post in terms of sentiment and engagement correlate with the attributes of good relationship (Hon & Grunig, 1999). The research findings highlight that although the characteristics of the Facebook post weakly correlate to the follower engagement and response sentiment, the combination of these characteristics provide strong correlations that influence the levels of engagement (Cvijikj et al.,2013), such combinations applied to a predictive algorithm can predict within an acceptable error range the follower engagement and response sentiment.

Limitations and Further Considerations

This study was limited by the secondary data taken from the Liberal Research Bureau which excluded any demographic information from the authors of the posts as well as any personal or demographic information from the Facebook page followers that responded to the post; future research may look at how adding demographic features such as geographic location that the MP represents, geographic location of followers, follower gender and age group influence on the follower engagement and response sentiment.

The posts were limited to those published between December 2017 and November 2018 in any of the official Liberal Party of Canada Members of Parliament Facebook pages. Repeating the same study with samples from other Canadian political parties will provide perspective on the use of the official Facebook pages by different parties as well as to understand if there is one political party that significantly engages more followers than the other political parties. Additionally, performing the same analysis on different social media networks such as Twitter and Instagram will provide insights on the citizen's social media behaviour in regards to political messages, for instance do citizens engage more on one social network than another – or not.

Another limitation of the data was the static sample of posts. Further studies can use a dynamic set of data that may allow for computational training sets to evolve with the political climate and adapt follower engagement and response sentiment accordingly.

Communicators in charge of digital campaigns need to be up to date in available tools and frameworks in order to make use of them in the best possible way. This study utilized pre-existing algorithms and predictive models to obtain valuable insights based on one static dataset. By integrating several tools and creating a research framework, it can be applied to evolving datasets and produce current results. The major libraries used were (1) Vader rule-based sentiment analysis model which returns a compound sentiment value ranging between negative one and one; (2) SciKit Learn Random Forest classifier model which provided a training sample with a binary target can then be used to predict results; (3) SciKit Learn multivariate SVM classifier which classifies the sample into multivariate categories; and (4) SciKit Learn XGBoost Regression model which predicts numeric values using linear regression. There is a plethora of libraries and models available free and at a cost that can be used in practice and in further studies.

Conclusions

Although Facebook provides the infrastructure to increase two-way communication between government and citizens, this research found that the Facebook messages posted by the Liberal members of parliament over the lapse of one non-elections year, neglected, for the most part, to engage citizens in either voicing their opinions, sharing the messages or sharing their emotions regarding the topics at hand.

This study found a strong correlation between the post's popularity (likes) and its virality (shares), indicating that publishing likeable posts will increase the spread of the message. However, there is also a small correlation between comments and the angry emotion, indicating that the more comments a post receives increases the likelihood of negative emotions. Although the correlation between each Facebook post characteristic and the follower engagement and response sentiment are weak, the study highlights that shorter messages tend to receive more positive follower engagement; posts with higher positive connotation receive more positive comments as well. Similarly, videos and links receive more positive engagement. Even though this study found some relation between post topics and follower engagement and the response sentiment, it is important to note that these relations will vary everyday. This study shows however, that there are tools available for communicators to rapidly and accurately gauge the followers' sentiment towards a topic in any given time.

In order to understand what kind of posts engage citizens the most, communicators need to find the best combination of message length, time of day an day of week of posting, the type of post as well as the sentiment connotation of the message. The linear regression prediction models utilized in this study show that it is possible to predict follower engagement with a 30% error with the most significant predictors being the post message length, the time of publication, and the post sentiment. Similarly, the study found that a regression predictive model could predict the follower sentiment with a 25% error which means that predicting a false positive or a false negative sentiment is very unlikely; the most significant predictors for follower sentiment are also the post message length, the sentiment of the post message, and the time of publication. The predictive models provide communicators a good idea of how the post will be received by the followers, and if need be, they have the opportunity to modify the post messages to increase the probability of a positive sentiment as well as increase the virality, popularity and commitment of the post; providing the means to be used by communicators in order to proactively measure the outcomes of a Facebook post before publishing.

As part of the outcome metrics of the political party's Facebook strategy, communicators should incorporate follower engagement and response sentiment that best align to the digital strategy's primary objective. The Ministers of Parliament from the Liberal Party of Canada currently have low follower engagement on its official Facebook pages and efforts should be made to create posts that result in increased positive follower engagement. Predictive analytics can allow communicators to proactively –and inexpensively - gauge the public's response to any type of post prior to its publication.

References

Augustyniak, L., Kajdanowicz, T., Szymański, P., Tuligłowicz, W., Kazienko, P., Alhajj, R., & Szymanski, B. (2014). Simpler is better?: lexicon-based ensemble sentiment classification beats supervised methods (pp. 924–929). Presented at the Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, IEEE Press.

Bennett, C. J., & Bayley, R. M. (n.d.). The Influence Industry: The Global Business of Using Your Data in Elections. Retrieved February 27, 2019, from https://ourdataourselves.tacticaltech.org/posts/influence-industry/

Bonsón, E., Royo, S., & Ratkai, M. (2015). Citizens' engagement on local governments' Facebook sites. An empirical analysis: The impact of different media and content types in Western Europe. *Government Information Quarterly*, *32*(1), 52–62.

Bonsón, E., Royo, S., & Ratkai, M. (2017). Facebook practices in Western European municipalities: An empirical analysis of activity and citizens' engagement. *Administration & Society*, *49*(3), 320–347.

Bonsón, E., Torres, L., Royo, S., & Flores, F. (2012). Local e-government 2.0: Social media and corporate transparency in municipalities. *Government Information Quarterly*, *29*(2), 123–132.

Canada: internet penetration 2017 | Statistic. (2018). Retrieved February 27, 2019, from https://www-statista-com.libaccess.lib.mcmaster.ca/statistics/209104/number-of-internet-users-per-100-inhabitants-in-canada-since-2000/

Ceron, A., Curini, L., & Iacus, S. M. (2015). Using sentiment analysis to monitor electoral campaigns: Method matters—evidence from the United States and Italy. *Social Science Computer Review*, *33*(1), 3–20.

Cogburn, D. L., & Espinoza-Vasquez, F. K. (2011). From networked nominee to networked nation: Examining the impact of Web 2.0 and social media on political participation and civic engagement in the 2008 Obama campaign. *Journal of Political Marketing*, *10*(1–2), 189–213.

Collomb, A., Costea, C., Joyeux, D., Hasan, O., & Brunie, L. (2014). A study and comparison of sentiment analysis methods for reputation evaluation. *Rapport de Recherche RR-LIRIS-2014-002*.

Cvijikj, I. P., Spiegler, E. D., & Michahelles, F. (2013). Evaluation framework for social media brand presence. *Social Network Analysis and Mining*, *3*(4), 1325–1349.

Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, *56*(4), 82–89.

Franch, F. (2013). (Wisdom of the Crowds) 2: 2010 UK election prediction with social media. *Journal of Information Technology & Politics*, *10*(1), 57–71.

Grunig, J. E. (1997). A situational theory of publics: Conceptual history, recent challenges and new research. *Public Relations Research: An International Perspective*, *3*, 48.

Grunig, J. E. (2006). Furnishing the edifice: Ongoing research on public relations as a strategic management function. *Journal of Public Relations Research*, *18*(2), 151–176.

Grunig, J. E. (2009). Paradigms of global public relations in an age of digitalisation. *PRism*, *6*(2), 1–19.

Grunig, J. E., & Huang, Y.-H. (2000). From organizational effectiveness to relationship indicators: Antecedents of relationships, public relations strategies, and relationship outcomes. *Public Relations as Relationship Management: A Relational Approach to the Study and Practice of Public Relations*, 23–53.

Hon, L. C., & Grunig, J. E. (1999). Guidelines for measuring relationships in public relations.

Hong, S., & Nadler, D. (2012). Which candidates do the public discuss online in an election campaign?: The use of social media by 2012 presidential candidates and its impact on candidate salience. *Government Information Quarterly*, *29*(4), 455–461.

Hutto, C. J., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. Presented at the Eighth international AAAI conference on weblogs and social media.

Issenberg, S. (2012). How Obama's team used Big Data to rally voters. *Línea]. MIT Technology Review. Disponible En< Https://Www. Technologyreview. Com/s/509026/How-Obamas-Team-Used-Big-Data-to-Rally-Voters>[Última Consulta: 10 de Marzo de 2016]*.

Jungherr, A., Jürgens, P., & Schoen, H. (2012). Why the pirate party won the german election of 2009 or the trouble with predictions: A response to tumasjan, a., sprenger, to, sander, pg, & welpe, im "predicting elections with twitter: What 140 characters reveal about political sentiment." *Social Science Computer Review*, *30*(2), 229–234.

Ledingham, J. A. (2001). Government-community relationships: Extending the relational theory of public relations. *Public Relations Review*, *27*(3), 285–295.

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, *5*(1), 1–167.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113.

Moro, S., Rita, P., & Vala, B. (2016). Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach. *Journal of Business Research*, *69*(9), 3341–3351.

Parent, M., Vandebeek, C. A., & Gemino, A. C. (2005). Building citizen trust through e-government. *Government Information Quarterly*, *22*(4), 720–736.

Quintelier, E., & Theocharis, Y. (2013). Online political engagement, Facebook, and personality traits. *Social Science Computer Review*, *31*(3), 280–290.

Sandoval-Almazán, R., & Gil-García, J. R. (2012). Are government internet portals evolving towards more interaction, participation, and collaboration? Revisiting the rhetoric of e-government among municipalities. *Government Information Quarterly*, *29*, S72–S81.

Siegel, E. (2013). *Predictive analytics: The power to predict who will click, buy, lie, or die*. John Wiley & Sons.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. Presented at the Fourth international AAAI conference on weblogs and social media.

Weiss, S. M., Indurkhya, N., Zhang, T., & Damerau, F. (2010). *Text mining: predictive methods for analyzing unstructured information*. Springer Science & Business Media.

Welch, E. W., Hinnant, C. C., & Moon, M. J. (2004). Linking citizen satisfaction with e-government and trust in government. *Journal of Public Administration Research and Theory*, *15*(3), 371–391.

Appendix A: Follower Engagement Level by Topic

Table 3

Top 20 Post Topics and its related engagement level

Topics	Extraordinary	High Engagement	Low Engagement	No Engagement
prime minist justin	10.76%	3.12%	1.23%	
tamil heritag month	5.06%			
canada child benefit	3.80%	0.97%	1.32%	2.01%
parliamentari secretari minist	3.80%	0.75%		
rise hous common	3.80%	0.53%		
thank prime minist	3.16%			
look forward work	2.53%	1.00%	0.47%	
separ children parent	2.53%			
appoint parliamentari secretari	2.53%			
welcom prime minist	2.53%			
member parliament brampton	2.53%			
justin trudeau confid	1.90%			
keep cost live	1.90%			
emerg room physician	1.90%			
member parliament scarborough	1.90%			
happi canada day	1.90%	1.15%		
celebr thai pongal	1.90%			
thai pongal tamil	1.90%			
nation pharmacar program	1.90%			
canada summer job		0.87%	2.00%	1.67%
nation hous strategi			0.75%	
summer job program			0.68%	
new horizon senior			0.63%	
canadian arm forc			0.59%	
liber parti canada		0.62%	0.57%	
horizon senior program			0.57%	
look forward see			0.55%	
constitu youth council			0.54%	
parti ral canada		0.53%	0.51%	
coast coast coast		0.59%	0.47%	

Topics	Extraordinary	High Engagement	Low Engagement	No Engagement
work hard join			0.46%	
summer job student			0.44%	
parti canada parti			0.44%	
canada parti ral			0.43%	
ginett petitpa taylor		0.62%	0.41%	
remain onlin fill				2.68%
canadian across countri				1.34%
rose hous common		0.59%		
justin trudeau justintrudeau				2.01%
parent share benefit				1.00%
keep great work				1.34%
review applic begin				1.67%
learn child care				1.00%
januari posit remain				1.00%
innov scienc econom				1.00%
look forward continu		0.69%		
canadian affect mental				1.00%
canadian feder student				2.01%
posit remain onlin				2.68%
would like thank		0.50%		
progress agreement transpacif				1.00%
year ago today		0.53%		
retweet justin trudeau				2.01%
minist bill morneau				1.00%
secur dignifi retir		0.50%		
minist chrystia freeland		0.62%		
fight climat chang				1.34%
morn nation park		0.56%		
gro morn nation		0.56%		
chief execut offic				1.34%
affect mental ill				1.00%

Appendix B: Top 20 Topics by Response Sentiment

Table 4

Top 20 topics that received either positive or negative response.

Top 20 Topics with Positive Response	**% positive responses (n=4,715)**	**Top 20 Topics with Negative Response**	**% negative responses (n=153)**
prime minist justin	1.68%	remain onlin fill	7.84%
minist justin trudeau	1.65%	begin march posit	2.61%
canada summer job	1.34%	onlin fill chairperson	1.31%
happi new year	1.08%	chief execut offic	1.96%
canada child benefit	0.91%	posit remain onlin	7.84%
happi canada day	0.91%	climat action fund	1.96%
merri christma happi	0.74%	applic begin march	2.61%
look forward work	0.72%	onlin fill chief	1.31%
look forward see	0.70%	march posit remain	2.61%
liber parti canada	0.68%	honor labradorian distinct	1.96%
new horizon senior	0.64%	fight climat chang	1.96%
parti ral canada	0.62%	ukrainian nation feder	1.31%
look forward continu	0.59%	begin januari posit	1.96%
ginett petitpa taylor	0.55%	review applic began	1.96%
parti canada parti	0.53%	review applic begin	5.88%
canada parti ral	0.53%	canada review applic	1.96%
constitu youth council	0.53%	intern develop global	1.31%
summer job program	0.53%	applic begin januari	1.96%
horizon senior program	0.51%	januari posit remain	1.96%
nation hous strategi	0.51%		