From Greetings to Corruption: Politicians, Political Parties, and Tweeting in India

Lia Bozarth
University of Michigan

Anmol Panda
Microsoft Research, India

Joyojeet Pal
University of Michigan

Abstract. We present a study of social media discourse in India, examining diverging theories put forth by past work that social media dilutes or centralizes the quality of political discourse, particularly in Global South settings. Classifying 1711 Indian political elites on Twitter and studying their feeds of about 4.6 million messages, we find that politicians are indeed more keen on establishing personal branding through low-substance, positive-focused messaging rather than broadcasting hard policies stands. Moreover, compared to the party in power, opposition politicians collectively post more detailed tweets and demonstrate higher negativity, especially regarding corruption, on social media. Finally, through contextual human examination of the most retweeted messages from two key leaders - the prime minister and the leader of the opposition, we find that there are qualitatively important distinctions between both the styles of key politicians, and its affective outcomes on the public, suggesting that a healthy contrarian, albeit abusive, space for democratic discourse exists online for politicians in India.

Keywords: technology and development; social media; political communications; India politics; Indian policies

1. Introduction

Social media has brought to center-stage facets of collaborative communications as political elites and advocacy groups alike have moved to direct outreach online, buoyed by supporters and private communications teams (Pal et al., 2018; Kou et al., 2017; Mascaro and Goggins, 2010; Baumer et al., 2011). This change has three important ramifications. First, social media has allowed for new forms of rhetorical construction and outreach, including the personalization of communication to focus on individual politicians' "personality" and "lifestyle" rather than policy-relevant discussions (McGregor, 2018; Enli, 2013; Meeks, 2016). Second, for substantive issues, political actors can now rely on private resources and networks to provide their own narratives, (Kellner, 2018; Enli, 2017; Casero-Ripollés et al., 2017; Fuller et al., 2018), bypassing the editorial constraints of traditional news media. Finally, social media enable networked collaboration among political elites belonging the same party or of common interests to strengthen and amplify their social media presence and influence, or to build and advance shared policy agendas (Chadwick, 2017; Hemphill et al., 2013).

Prior research, largely US and Euro-centric, have focused on examining politicians' social media strategies when cultivating individual brand or when collectively championing for or against salient legislative issues (Chadwick, 2017; Hemphill et al., 2013). As with elsewhere in the world, politicians in India have also taken to social media aggressively in recent years. In this paper, we compare and contrast tweets by Indian political elites around policy-relevant substantive topics with non-issue tweets [1].

We examine these differences from 3 distinct dimensions: i) frequency as determined by the absolute and normalized percentage of tweet contributions by politicians from ruling and opposition parties to each topic; ii) collaboration patterns of politicians with each other and their engagement behavior with news media on Twitter when discussing these different topics; iii) the text complexity and sentiment differences in tweets belonging to distinct topics and parties.

Our paper makes the following contributions:

- To the best of our knowledge, this is the first large scale quantitative study of social media language and reach of a substantial number of politicians in a Global South setting with respect to a selection of the topics they discuss. In this, we describe our classifier, which involves elements of natural language processing as well as classification through studying the collaboration or networks between accounts on Twitter.

- We show that politicians' messages are more frequently on non-issue subjects such as greetings compared to issue or policy-relevant topics such as technology, development, poverty, or welfare.

  This finding suggests that politicians see greater value in signaling to their followers and the online community a positive-focused, personability rather than a purely political-business centric message, which aligns with prior findings using qualitative coding of tweets (Pal, 2015).

- In addition, we observe that when politicians do talk about substantive issues, they show a stronger preference for engaging with news media comparing to when they are tweeting about celebrations, birthday wishes, et cetera. For non-issue tweets, we see that politicians are more likely to engage with each other. This implies that when showing policy expertise, politicians strategically involve news media to position their views on specifics.

[1] For our purposes, non-issue tweets are defined as regards of various kinds such as birthday or festival greetings, condolences etc. While messaging on such 'non-issue' tweets serves an important symbolic and political purpose and contributes to personal branding, we argue that these are less likely to be immediately policy-relevant.

– Finally, we show that issue-focused tweets tend to possess higher language complexity. Furthermore, non-ruling parties depict more negative sentiments in their tweets. This pattern suggests that non-ruling party elites, to some extent, have a shared affinity to utilize Twitter to criticize the policies established by the ruling party.

2. Related Work

We review three areas of research most relevant to our work. First, we discuss the role of mainstream media and traditional journalism as well as factors, both internal and external, that contribute to news quality with an emphasis on Indian development context. Next, focusing on elected government officials and candidates, we survey literature that study their communication strategies on social media for both personal brand-building as well as advancing individual and party platforms. Finally, we report on existing work that examine large-scale online collaborative actions taken by various political actors with a focus on the role and importance of the Web and social media.

2.1. Social Media, Mainstream News Media and Journalism within the Indian Development Context

The news media - whether print, radio, television, or indeed social media - has broadly been seen as playing the role of checks and balances against the political establishment. However, the ability of the media to play this role is predicated upon both the state of the free press and the traditions of public commentary by political actors. The growth of social media in the Global South has offered a new set of continuities to this work on development communications, which has long argued that media studies needs to be de-westernized and seen as posing distinct challenges in such settings (Park, 2000). In part, the ability of social media to play an arbiter of accountability is problematised by issues such as colonial legacies of mainstream media as a propaganda tool (Agrawal, 2006) or histories of political control of public discourse through state media ownership (Rahman, 2014).

Indian news media has had its own issues with challenging politicians on the news. Work has shown that, in the era of urban-centric media, political journalists treated public officials with a colonial legacy of deference, a trend that had started to reverse with the increase of vernacular journalism but indeed that this critical journalism coincided with an advertising-centric capitalization and converging corporate ownership of mainstream media (Rao, 2010). The resulting effect on the media has been well an outsize influence of moneyed interests (Rajagopal, 2017). In a development context, where the institutional mechanisms to track or curb media capture or corruption leads to a risk of what Saima Saeed has referred to as "phantom journalism" in which a nexus between politicians, profit-maximizing news, and concealed media ownership poses a new set of challenges for democratic discourse (Saeed, 2015).

Commentators have noted the emergence of a sensationalist media environment, where investigative or probing journalism has been pushed towards dramaturgic events such as sting operations (Sundaram, 2015). Issues of public importance such as corruption and development are mediated through the performance of outrage by celebrities at these events on social media (Rodrigues, 2014). With politicians setting up their direct channels of communication on social media, their ability to be heard can be decoupled from the substantive matters of the issues they are willing to address online.

Studies have argued that the weakened professional press corps has diluted the quality of discussion on issues of core policy importance - research on the Indian 2014 elections showed that the notion of development is increasingly discussed in the abstract, as a brand, rather than in specifics of issues such as poverty or inequality (Mudgal, 2015). This work aligns closely with what we are doing in this paper in terms of how issues such as corruption, development and poverty get talked about on Twitter, and how these in turn, have resonance in the public.

Furthermore, the growth of social media as a source of news has further exacerbated the challenges with quality news access, creating a heteronomic effect on journalistic practices, essentially enabling political hegemons to dominate the media discourse (Maheshwari, 2018). We find elements of this trend in the domination of social media political conversation by a few political parties in this research. More importantly, other work has also shown that politicians' social media messaging reverberate on both television (Kumar, 2015) and print (Chakraborty et al., 2018), which suggest that politicians can dramatically reduce their non-social media output and still find voice on traditional media. Indeed, several politicians have turned social media into the primary driver of their signature campaigns which not only replaces the traditional print and television channels but also blurs the boundaries between what is overt advertising and what is regular campaign outreach (Rodrigues, 2017).

2.2. Social Media and Personalized Communication by Political Elites

Prior studies have observed an increased effort from established politicians and candidates in utilizing social media for personal branding. While a lot of these discussion around social media and politics has focused on the experiences of a few key politicians in parts of the West, there is increasingly a body of work that examines the role of social media in the communications of politicians in the Global South (Lufkens, 2016; Pal, 2016; Ahmed, 2014; Lee, 2017; Lee, 2017; Ahmed et al., 2016; Correa, 2017). Within the context of India, Jaffrelot (Jaffrelot, 2015), for instance, note that "the second largest independent PR firm in America, charged over US$25,000 a month to manage Modi's [media] account", underscoring the importance of digital branding to Prime Minister Narendra Modi. Additionally, Pal et al. (Pal et al., 2016) observe politicians' selective engagement with topics in Tweeting, depending on the stage of the electoral cycle. In the west, Libby Hemphill et al (Hemphill et al., 2013), investigating the US Congress's use of social media, found that officials frequently use Twitter to advertise their political positions. Furthermore, researchers such as Howard, Bennett, Kreiss, and Bimber (Howard, 2006; Bennett, 2012; Kreiss, 2016; Bimber, 2014) show that political candidates increasingly rely on data mining techniques to classify and divide potential supporters into granular groups, and then personalize their campaign messages targeted at these groups accordingly both online and offline. Moreover, campaign messages are becoming increasingly focused on "lifestyle" messages instead of being policy and issue-based. Researcher including Bennett and Pfetsch (Bennett and Pfetsch, 2018) stress the importance of examining and understanding the changes in political discourse within the new media environment, and how these changes affect democratic institutions at large. Luckily, the digital prints, such as policy tweets on Twitter, left by these political entities, as suggested by Wilkerson and Casas (Wilkerson and Casas, 2017), provide interested parties opportunities to study individual politicians at a unprecedented scale and depth.

2.3. Social Media and Large-scale Political Collaborative Network

By providing a digitally enhanced public sphere, social media present the specter of a new form of collaborative people to power communication (Dutton, 2009; Saeed et al., 2009; Ludwig et al., 2016; Baumer et al., 2011; Handler and Conill, 2016; Niederer and Priester, 2016). It has been proposed that these enable participatory politics through communicative action (Loader, 2011; Mascaro and Goggins, 2010). Within the literature of CSCW, Baumer et al. (Baumer et al., 2011), analyzing the interaction between political bloggers and blog readers, highlighted social media's facilitating role in online community building. Mascaro and Goggins (Mascaro and Goggins, 2010), studying collaborative information seeking behavior of Tea Party activists, arrived at a similar conclusion. Feltwell et al (Feltwell et al., 2017), focusing on the specific issue of poverty and welfare in the U.K., argued that social media affordances enabled minority groups to band together and collectively challenge the negative tone set by mainstream news media.

The enthusiastic embrace of a "transformative" social media that followed collaborative communications between citizens during events such as the 'Arab Spring' (Lotan et al., 2011; Morozov, 2009; Tusa, 2013) were tempered by the possibility that social media actually enhanced political actors' ability to censor and monitor citizens (Schäfer, 2016; Zittrain et al., 2017; Gayo-Avello, 2017; David, 2018), or feed citizens a sanitized version of events (Ohm, 2015; Curato, 2017). More recently, controversy over voter influencing enabled by centralized control over user profiles at major social networking sites like Facebook have further dampened expectations of social media presenting organic, collaborative action by common citizens and highlighted the ability of institutions or politicians to guide both the discourse on social media. Finally, recent work by Pal et al. (Pal et al., 2018) demonstrates populists' ability to leverage social media to attack and discredit their political opponents and erode the goodwill and opportunities of partnership between citizens from opposing parties.

This rich body of work shows how a range of disciplines has approached the questions of what politicians do and say on social media and what that may mean for the democratic process. In this work, we dive deep into the what and how of politicians' speech online. We first examine how Indian politicians prioritize and tweet about issues related to technologies, development, corruption, inflation, and poverty and welfare comparing to non-issues such as celebrations and birthday wishes by assess how frequently politician talk about these topics. We then examine to what extent do political elites engage with each other and with news media on Twitter when tweeting about these different topics and policies. Finally, we analyze the linguistic differences of tweets belonging to these topics.

Unlike prior studies that focused on a small set of key politicians or political actors without an official role (e.g. bloggers or activists), we leverage Twitter collaborative networks to classify a substantial large number of influential Indian politicians both at national and regional levels. Our study has a dataset spanning 4 years and we use both qualitative and quantitative methods to examine social media strategies of these political entities.

3. Data

Our dataset consists of 4.6 million tweets in both English and Hindi (We filter out the small subset of tweets written in regional dialects such as Kannada, Malayalam etc.), from Jan, 2014 to October, 2018, contributed by 1711 distinct Twitter accounts of Indian politicians from both state and national levels. Approximately 2.94M of the tweets are in English, while about 1.66M are in Hindi. We observe that 779 accounts are ruling BJP politicians who supplied 2.34M tweets (or, 51.0% of total tweets); 417 accounts are from the main national opposition INC politicians with aggregated sum of 1.25M tweets (or, 26.9% of total tweets); the remaining 515 politicians from other parties posted 1.02M tweets (or, 22.1% of total tweets). The data is consequently inordinately skewed towards the ruling BJP politicians, but at the same time underlines the party's dominance of social media, partly led by the party leadership's diktat that only politicians with over a certain threshold of followers on social media could be offered party seats for elections (Bureau, 2018).

3.1. Twitter Account Classification

The Indian Twitter universe consists of over thirty million active users (sta, 2018) who belong to different identity groups, divided most notably by language but also by caste and religious divisions. Varying educational backgrounds, employment profiles, socio-economic and cultural differences between cities and villages and between states produce distinct social media profiles and discourses. Given this scale of volume and variety, it is not possible to manually accumulate all political accounts one at a time. While the names of members of parliament (about 800 at any given time) can be acquired and their Twitter accounts can be collected manually, such data is not easily accessible for unelected individuals such as party functionaries, secretaries, *karyakartas* and celebrities who support or oppose certain parties. Thus, our goal was to first build a `classifier` that can reliably identify political Twitter handles with limited human intervention. Figure 1 details this process. We first i) manually collect a substantial set of political twitter handles as groundtruth, we then ii) conduct feature engineering and iii) build supervised classifiers using groundtruth labeling, and finally, we iv) batch classify additional Twitter accounts into political and non-political accounts.

3.1.1. *Groundtruth labels*

We cover politicians from 20 different parties (both national and regional parties) including BJP, INC, Shiv Sena, CPI(M), AAD, etc. Politicians were added by searching through the list of elected representatives of both houses of parliament, searching the key parties for major post-holders such as party presidents, general secretaries etc, chief ministers of states, and finally state legislatures for current lists of elected politicians. We filtered out all that had less than 1000 followers or have posted less than 1000 tweets. This procedure gives us a total of 1002 manually collected political handles which constitute to the positive class of the `training` set of our classifier. In addition, we also collect over 800 top retweeted Indian accounts as the negative class in the `training` set.

3.1.2. *Feature engineering*

To featurize Twitter accounts, we must leverage the limited information about Twitter users that is exposed by the Twitter API, including the user's name, screen name, description, number of followers, number of accounts followed, number of statuses (tweets), location, profile picture, etc. Among these, we find the `user description` (also known as profile text) to be most relevant to identifying politically affiliated Twitter accounts. Table III lists a sample of political account profiles. The keywords that identify potential political accounts are in bold.

We use a `Bag of Words` model to represent the `user description` and rely on the `CountVectorizer` package provided by `scikit learn` to vectorize the description. Table IV shows the most important features for the positive and negative classes and their corresponding weights with the `SVM classifier` using a `linear` kernel. As shown, the terms “mp", “minist", “parliamentari", et cetera have the largest positive weights, while the words "super", "andhra" and "columnist" have the largest negative weights. The weight results make intuitive sense given the former terms are closely associated with governing whereas the later ones are not.

3.1.3. *Classifier performance metrics and selection*

We experimented with five different binary classifiers: `NaiveBayes`, `RandomForest`, `StochiasticGradientDescent`, `LogisticRegression`, and `SupportVectorMachine` (SVM with `linear`, `rbf` and `poly` kernels). For each classifier, we apply `GridSearchCV` to find the optimal parameters. Tables VI and VII show the parameter grids used in `GridSearchCV`. The figures in bold indicate the values that `GridSearchCV` returned for the best estimator. Table VII details the performance of all the classifiers. We find that the `SVM` classifier with an `rbf` kernel has the highest precision rate at 0.930. It has an accuracy score of 0.898, recall score of 0.847 and an F1 score of 0.887.

3.1.4. *Account selection and classification*

The next step in the process is to select a subset of Twitter accounts from the universal set that can be fed as input to the classifier. As iterating on all Twitter accounts is neither feasible nor necessary, we choose a novel approach to locate potential political handles. We collect all accounts that are followed by the politicians in the training set by iterating on their `friends` list. This method is motivated by the expectation that politicians follow at least their party colleagues and possibly even major politicians from other parties. This also allows us to reduce the number of political handles before classification. From the 1002 accounts in the `training` set, we gather over 140,000 accounts that these politicians follow. As we had already collected most members of parliament and popular members of state assemblies manually and the remaining politicians, especially local leaders, are expected to have fewer followers, we reduce the threshold of number of tweets from 1000 to 100. This set is fed as input to the classifier which outputs a little over 9500 political handles.

Finally, we manually annotate each politician in this list to remove false positives.

3.2. Tweet Classification

Prior to examining how politicians tweet about social, economic, and other policy-related issues, we first classify tweets into the following 7 categories: 1) information and communication technologies (`Technology`) related tweets, 2) poverty & welfare (`Poverty`) related tweets, 3) economics and development (`Development`), and 4) corruption, scams, and bribery (`Corruption`), 5) inflation and fuel price surge (`Inflation`), 6) greetings and holiday celebrations (`Greetings`), and 7) other (`Other`).

The goal of selecting these categories is to compare and contrast the prevalence of tweets on substantive topics (e.g. technology, poverty, corruption) with that of non-substantive tweets (e.g. greetings). Additionally, we aim to not only consider the extent to which key topics of democratic relevance are part of the political communications of Indian politicians, but also to examine the affective relationship these tweets have with the public, as measured through retweets.

Keywords and/or Cosine similarity measurement based document clustering has been used extensively by many prior work (Conover et al., 2011; Adamic and Glance, 5 08; Stieglitz and Dang-Xuan, 2013) focused on political communication on social media. In this paper, we use a similar approach to cluster tweets into different issue-based and personal-appeal-based categories.

Classification procedure: We use word2vec and cosine similarity scoring to generate keywords for each category of non-baseline tweets, we then use these keywords to assign each tweet into the matching category. To be more specific, i) we first use gensim, an open source natural language processing (NLP) toolkit in python, to produce a 300 dimensional vector space representation of our corpus (here, each tweet is a single document). Within this space, each unique word is assigned a numeric vector. For instance, the word "govt" has the corresponding 300d vector [-1.9078784 1.654422 1.8524699...]. ii) Given that word vectors are positioned such that words that share similar contexts are located closer to each other in the vector space (i.e. the cosine of the angle between the vectors of a pair of semantically similar words are smaller), we are able to manually select a handful keywords related to `Technology`, such as "technology", "digital", et cetera, and then use cosine similarity scoring to discover additional keywords that share comparable semantic meanings to the selected words. iii) Using this approach, we generate 157 keywords for `Technology`, 91 for `Poverty`, 131 for `Development`, 151 for `Corruption`, 37 for `Inflation`, and 275 for `Greetings`. iv) For each tweet, we categorize it as `Technology` if it contains 1 or more `Technology` keywords (similar for `Poverty`, `Development`, et cereta). If a tweet has no matching keywords, it's assigned to `Other`. Additionally, a tweet can belong to multiple categories. The complete list of keywords can be found in the Appendix Table V.

Using this approach, we label 187.7K or 3.4% tweets as `Technology`, 188.1K or 3.4%, as `Poverty`, 568.1K or 10.2%, as `Greetings`, 327.3K or 5.9% as `Development`, 171.0K or 3.1% as `Corruption`, 63.2K or 1.1% as `Inflation`, and the reminding `Other` tweets. This suggests that politicians generally tweet more about holiday celebrations and greetings in comparison to actually policies relation to technology and development. In order to assess the performance of our classification, we randomly select 100 tweets from the subset of tweets labeled as `Technology` and manually assess whether each tweet indeed focuses on information and technology related topics. We observe that 93 of the 100 do, resulting in 93% accuracy. Similarly, we see a 89% accuracy for `Poverty` tweets, 86% for `Development`, 92% for `Corruption`, 81% for `Inflation`, and 92% for `Greetings` tweets. Furthermore, we randomly select 100 sample from `Other` tweets to assess our recall, 3 out of the 100 samples are related to `Development`, 3 for `Greetings`, 2 for `Inflation`, 1 each for `Technology` and `Corruption`, the remaining 90 tweets are true negatives.

4. Analyses

4.1. Overview

A look at the overall descriptive statistics from the sample gives a sense of the landscape of political twitter in India. Figure 1 shows that Prime Minister Narendra Modi is clearly the dominant figure both in terms of his following, and the extent to which his tweets get retweeted. However, that there are a number of other political leaders with a significant presence and influence online, including the main leader of the opposition, Rahul Gandhi, of the rival Indian National Congress Party (INC), whose messages, on average, get more retweets than the prime minister, who currently has several times as many followers. We also see a trend that the key leader in several parties has an outsized influence in terms of retweets. In fact, all of the leaders whose median retweets were above 500 retweets per tweet were leaders of specific parties, even if they weren't necessarily the most followed leaders from within the overall sample. Thus alongside Modi and Gandhi, we see Akhilesh Yadav of Samajwadi Party, Laloo Prasad Yadav of Rashtriya Janata Dal, Arvind Kejriwal of Aam Aadmi Party, and YS Jagan Reddy of YSR Congress Party, each had high retweets to their tweets, suggesting a centralizing tendency for social media communications, or coalescing of discourse around a specific leader, for each of these parties.

4.2. Distribution of Direct Contributions

We first examine the distribution of contributions for each category of original tweets (i.e not including retweets). Descriptive statistics for party based contributions are summarized on Figure 2. We observe that the ruling BJP contributed 348.9K `Greetings` tweets, accounting for 12.8% of its total tweets. This is much higher than the opposition INC which only contributed 124.3K or 8.1% of its total tweets to `Greetings` (similar observation for politicians in 3rd parties). These numbers suggests that the ruling party members are more frequently engaged in non-issue-based greetings (over 50 per cent more than the main opposition party), which have a signaling purpose rather than a policy commentary. Moreover, BJP also demonstrates a higher preference to discuss `Development` related issues such as economic development, GDP, et cetera. INC, on the other hand, shows a higher affinity in discussing `Corruption` and `Inflation`: 83.5K or 5.4% of its total tweets are focused on `Corruption` while the number is only 2.2% for both BJP and 3rd-party politicians. The broad suggestion here is that opposition parties use social media more extensively for critical messaging aimed at the ruling government.

The top 10 contributors for each category of tweets determined by the absolute number of tweets are shown on Table I. The first item of significance is the extent to which the BJP dominates the conversation in topics including `Greeting`, `Development`, and `Technology`, whereas INC and 3rd-party politicians lead in `Corruption` and `Inflation`.

Politicians such as the ministers of Electronics (RS Prasad), Science and Technology (Harsh Vardhan), Commerce (Suresh Prabhu), and Finance (Jayant Sinha), Skill and Entrepreneurship (Dharmendra Pradhan) contribute significantly to the Twitter conversation on technology, as well as Andhra Pradesh Chief Minister Chandrababu Nadu, and INC member and tech billionaire Nandan Nilekani.

The discussions on development are interesting in that we see a coming together of the discussion of technology and poverty - leaders from both lists are featured in the development discourse. Thus Raghubar Das and Radha Mohan Singh, who are significant in discussing poverty are present in this list, but also Dharmendra Pradhan, Piyush Goyal, and Suresh Prabhu, who feature in the technology list. The suggestion here is that technology and development are closely co-occurring themes, an idea that has much purchase in a significant body of work that proposes technology as fundamentally tied to the idea of development in the Global South.

The greetings selection was primarily in place to examine the role of casual and non-policy tweeting by politicians. Prime Minister Narendra Modi emerges as the greeter-in-chief, and has the most greeting-themed tweets. Remarkably, all the politicians in this set are from the BJP. This aligns with previous research that suggests that Narendra Modi's online presence is inordinately aimed at positive messaging and presenting a selective image of national harmony (Kaur, 2015).

The discussion by key leaders on corruption is a mixed bag - opposition figures including key spokespersons from both the ruling and opposition sides - Sanjay Jha, Amit Malviya, Sanjay Singh and RS Surjewala each figure on the list, suggesting that corruption is an issue that the official mouthpieces of parties must address. Corruption is an important subject for opposition politicians overall as seen in Figure 3, not a single elected BJP member figures in this list. The BJP accounts in this list are BJPDelhi (where it is in the opposition), as well as Subramanian Swamy and Sambit Patra, both unelected politicians.

The discussion on inflation is dominated by non-BJP politicians, as we see in figure 3. However, the list of biggest contributors to the subject is led by Dharmendra Pradhan, who is a false positive due to his position as petroleum minister (he dominates the list by repeatedly tweeting about petroleum which is otherwise typically an inflation-related subject - as we see, inflation was the category with the lowest accuracy, partly due to the stem term 'petrol' which is typically associated with inflation, but not necessarily), and Sanju Verma, a BJP economist whose tweets are largely about defending the party against inflation-related charges.

The discussion on poverty is the fairly diverse in terms of the actors involved - there is both a significant presence of opposition figures (Arun Yadav, Sitaram Yerchuri, Ashok Gehlot), and high contributions by the central ministers who deal more centrally with issues of poverty and rural development as part of their work of Agriculture (Radha Mohan Singh), Rural Development (Ram Kripal Yadav), and Social Justice (TC Gehlot), as well as Chief Ministers in states with high poverty rates such as Madhya Pradesh (Shivraj Chouhan) and Jharkhand (Raghubar Das).

Our overall findings suggest that politicians from all parties preferentially tweet about soft topics comparing to difficult policies and issues. The topical preference of certain politicians is an element both of their individual branding and their specific political roles.

4.3. Engagement with News Media

Theoretically, the affordances provided by social media such as Twitter allow politicians to strategically communicate with their constituents directly without relying on media to spread their agenda. Nevertheless, politicians can and do selectively engage with news media on Twitter for agenda-setting purposes (Chadwick, 2017; Jungherr, 2014). In this section, we use social network analysis to examine the difference in importance of news media when politicians are discussing policy-related issues comparing to when they are not.

Social network analysis is one of the most commonly used method to assess an actor's social role within a network consist of many distinct actors (such as authority over others within the network). Briefly, a social network consists of 2 components: nodes (individual actors) and edges (relationship between pairs of nodes). A network is directed if the relationship is unidirectional and the edges are called directed edges. For instance, John follows CNN on Twitter, a directed edge exists from John to CNN, often written as $e_{John,CNN}$. However, CNN does not follow John on Twitter. A network is undirected if the relationship is bidirectional and edges are called undirected edges. For instance, John is married to Jan; therefore, Jan is also married to John.

Here, we measure the importance of news media using 2 distinct social network metrics: indegree centrality and pagerank centrality. Indegree measures the number of ties, or edges, directed at a participant and captures a participant's popularity or reputation (Bonacich, 1987). Pagerank takes into account of both the number of ties as well as the quality of nodes themselves (Kwak et al., 0 04).

4.3.1. *Generate News Media Accounts*

To examine politicians' intensity and pattern of engagement with news media within different context, we first aggregate a list of 605 influential news media Twitter accounts using the following procedure: i) We first determine the entire list of accounts being retweeted or mentioned by politicians, and filter out all the accounts with less than 1K followers. ii) We then generate a list of keywords that are associated with news media and careers relating to journalism such as "news", "reporter", "blogger". iii) We obtain a sublist of account that contain at least 1 of the keywords in the user profile description field. iv) Finally, we manually examine each account and remove the non-media accounts. A similar approach was used by Priante et al. (Priante et al., 2016) in identifying the occupations of Twitter users. This method helps craft a picture of the extent to which Indian politicians discuss certain topics, and how in turn these relate to the politicians' engagement with news media.

4.3.2. *Analyze Politicians' News Media Engagement*

Using our data, we then build directed retweeting and mentioning networks of politicians and news media on Twitter for each category of tweets, resulting in 10 distinct networks. Let directed graph $G_{retweet,technology} = \{V,E\}$ where V is the list of politician and news media accounts. A directed edge $e_{u,v}$ exists in E if politician u retweeted one of v's `Technology` tweets on Twitter. Note that while u has to be a politician, v can either be a politician or news media. Likewise, we build $G_{retweet,development}$, $G_{retweet,poverty}$ and $G_{mention,poverty}$, et cetera.

We then applied networkx, a python network analysis library, to generate the indegree and pagerank centrality scores of each node for all 10 networks. Both centrality scores are summarized in Figures 4 and 5. As shown, the median indegree centrality of news media accounts in graph $G_{mention,technology}$ (this is the network where politicians mention each other or news media using '@' when tweeting about technology) is 2, which is actually higher than the median indegree centralities for politicians (median 1). In other words, an average news media account is mentioned by 2 different politicians whereas an average politician account is only mentioned by 1 other politician when the topic of technology is under discussion. News media also garners higher indegree centrality in `Corruption` and `Inflation`. News media accounts' indegree centrality scores are comparable to politicians' for `Development` and `Poverty`, but are much lower for `Greetings`. That is, politicians mention news media just as much as they do other politicians when discussing policies and issues related to technology, development, and social welfare. But, politicians seldom mention news media when they talk about celebrations, birthday wishes, et cetera.

While indegree centrality is the direct measurement of the number of mentions and retweets an Twitter account gets, an account's pagerank is high only if the account is being retweeted or mentioned by accounts that are themselves retweeted or mentioned by many. To elaborate, an account that is retweeted or mentioned by x number of politicians who are not significant figures (e.g. less known state officials who don't get retweeted or mentioned) has a lower pagerank than another account who may be retweeted or mentioned by fewer than x politicians, but the ones who did retweet or mention him/her are high profile (e.g. being retweeted by both Narendra Modi and Sushma Swaraj). As shown in Figure 4 and 5, the median pagerank scores for news media accounts are actually slightly higher than or comparable to politicians' for substantive topics including `Inflation`, `Corruption`, `Technology`, `Development`, and `Poverty`, suggesting that politicians preferentially retweet or mention news media when discussing important policy-related issues. Both indegree centrality measurements and pagerank scores for new media in the `Greetings` tweet categories are, however, noticeably smaller.

Related work observe increased use of social media by politicians for personal branding and political narratives of their own (McGregor, 2018; Enli, 2013; Meeks, 2016). Our results expands on these studies by demonstrating that when actual policies are in discussion, politicians are strategically choosing to rely on news media either via retweets or mentions instead of talking about the topic themselves. This indeed may also be interpreted as a proxy for relying on relatively expert information on these subjects coming from professional journalists, rather than trying to talk about them directly. In fact, when politician tweet about soft, "lifestyle" topics, they rarely engage with the media.

4.4. Text Complexity and Sentiment Appeal

In this section, we go beyond analyzing what categories of issues politicians tweet about and the frequency of tweets by also assessing how they talk about these issues. i) We first determine the difference in language complexity of tweets of substantive topics, and `Greetings` tweets. ii) We then conduct sentiment analysis on each category of tweets to evaluate politicians' strategic use of emotional appeals.

Prior to performing any NLP related analysis, we first remove all non-English tweets and all retweets from our dataset, result in 1.07M total tweets. We then apply standard text-preprocessing steps to remove hyperlinks and reserved keywords, such as "@" from tweets.

4.4.1. *Language Complexity*

We first assess the complexity difference between different categories of tweets using 2 distinct metrics: 1) word count and 2) readability. In part our goal was to understand computationally the complexity of messages and its relationship to the topics discussed. We observe that the average word count for `Technology` tweets is 14.7 (median 13), `Poverty` is 16.07 (median 14), `Development` 15.3 (median 14), `Corruption` 16.6 (median 15), `Inflation` 16.8 (median 15) and finally, `Greeting`, which has the lowest average word count of 13.4 (median 12).

Next we use 3 measurements of readability to indicate how difficult a tweet is to understand: Flesch-Kincaid Reading grade, Smog Index, and Gunning Fog Index. For all 3 metrics, a higher score indicates that a reading material is more complex. For instance, a Flesch-Kincaid reading grade of 10.0 suggests that the text corresponds to 10th grade reading-level. These scales are used by many prior studies focused on studying the complexity of political communications (Ott, 2017; Flaounas et al., 2013). For each tweet i, we first remove tweets that have less than 2 word tokens, we then use the ReadCal python library to generate all three scores and denote as the value as $readability_{i,flesch}$, $readability_{i,smog}$, $readability_{i,gunning}$. Then, we run OLS treating *readability* as the dependent variable, and $category_i$, and $party_i$ as the independent variables. As shown in Table II, comparing to `Greetings`, all other categories of tweets except for `Inflation` are correlated with higher language complexity (e.g. `Development` tweets have roughly 2-grades higher reading-level than `Greetings`).

Both word count and reading ease measurements suggest that tweets about substantive issues such as corruption, development, and poverty and welfare are associated with higher language complexity. For inflation-related tweets, we see a higher average word count but comparable readability scores comparing to greetings. This suggests that inflation tweets have simpler vocabularies. Finally, we see that INC and 3rd party politicians contribute tweets with higher reading-levels.

4.4.2. *Sentiment Appeal*

Next we go beyond measuring text complexity by further examining the sentiment component of tweets. Sentimental appeal in political communication, as observed by prior research (Soroka et al., 2015; Sharma and Moh, 6 12) have important consequences (e.g. altering voter behavior) and predictive power (e.g. forecast which candidate is like to win). Moreover, past studies also suggest that political tweets with higher emotional appeal are more likely to be propagated (Stieglitz and Dang-Xuan, 2013) . We first use VADER (Valence Aware Dictionary and sEntiment Reasoner), a sentiment analysis library (Hutto and Gilbert, 4 06) specifically attuned to sentiments expressed in social media, to generate positive and negative emotional scores of each tweet and denote them as pos and neg. We also obtain the compound affect score which indicates a text's overall valence. Looking at the visual representation of the differences between each category of tweets in Figure 6, we can see that median affect scores are more positive for both `Greetings` and `Technology`, indicating that politicians, as a whole, tend to utilize more positive emotional appeals when discussing celebrations or information technology. Furthermore, we also see that INC demonstrates noticeable negativity in their tweets about corruption, inflation, and poverty comparing to INC and 3rd party politicians.

Next, we provide a more granular analysis by applying ordinary linear regression with pos_i and neg_i as the dependent variables, and $type_i$ and $party_i$ as the independent variables for each tweet i. As shown on Table II, all substantive topics are associated with lower positive and higher negative sentiment when comparing to `Greetings` tweets. Additionally, INC and other party politicians demonstrate lower positivity and higher negativity. One possible explanation is that non-ruling parties are more like to use social media to criticize the policies of the current majority rule party (i.e. the BJP).

Finally, we qualitatively analyze the sentiment appeal of tweets belong to key politicians. We considered the text of the most retweeted messages in each of the categories studied. Out of the top 200 tweets for each, Modi alone accounted for 41% of the total tweets. When Rahul Gandhi's tweets were added to the list, it covered 78% of the most viral tweets. This highlights the centrality of these two key leaders in social media universe. However, a look at a sampling of their tweets from among the most retweeted messages also shows how their messaging styles are very distinct. We took one example of a viral tweet from each of the four categories from the two key leaders to compare differences.

`Development:` On development, the distinct styles of the two politicians is evident. As demonstrated in Figure 8, while Modi talks about development in a tweet with positive valence, presenting a positive vision of India in a tweet that congratulates the average Indian citizen, Rahul Gandhi includes a link to a news story alongside a confrontational tweet attacking the government for a development-related fiscal policy.

`Technology:` On Technology, a viral tweet from Prime Minister Modi from 2016 urges young Indians to use more digital technology, shown in Figure 12, specifically in financial transactions, in the immediate aftermath of the demonetization move by the Indian government. The tone is optimistic, the images used in the tweet show technology and its users in a positive light. In contrast, the viral tweet from Rahul Gandhi, shown in Figure 12, includes a link to a story that claims to expose the prime minister's role in a data breach that exposed the personal information of millions of Indians. Unlike the positive tone on technology that Modi has, Gandhi presents a dystopian view of technology, of companies tracking citizens through their mobile devices.

`Poverty:` On poverty, the two tweets from Figure 10 highlight the distinct styles of the leaders. While Modi highlights the role of his supporters in a political victory for the poor, the viral tweet from Rahul Gandhi in the subject of poverty uses wordplay to attack on the Modi government in late 2017. The tweet that links to a longer address the leader made at a Chamber of Commerce address.

`Inflation:` On inflation, we found no tweets from Narendra Modi that were relatively more retweeted than the median - in fact, he barely ever discusses inflation. His most retweeted message that is related to obliquely related to prices is a 2016 tweet offering discounts on petrol for cashless buyers, which is arguably more about demonetization than about the price of gas. Rahul Gandhi on the other hand tweets frequently and aggressively about inflation, referring to Modi as a "king of misinformation", using excerpts from Modi's own speech to discredit his claims on inflation.

`Corruption:` While Modi tweeted significantly on corruption while an opposition politician, being in power makes that difficult since one would presumably be responsible for the corruption. His highly retweeted corruption-related message, in Figure 18, was a quasi-greeting to citizens and had a congratulatory tone marking the year anniversary of demonetization. Rahul

Gandhi on the other hand referred to Modi as "Supreme Leader" on corruption-related tweets, suggesting the prime minister had centralizing, dictatorial tendencies. Following the arms deal scandal, referred to as the Rafale scandal, corruption has been a major topic of tweeting for Rahul Gandhi, and the tweet highlighted here is reflective both of the frontal, insulting tone towards Modi, and of Gandhi's own appropriation of the soldier metaphor, borrowing from his opponent the long-successful affective value of bringing patriotism into an antagonistic exchange.

Greetings: Lastly, the two leaders' tweets that are viral in the greetings category also differ, though there is most similarity in the phrasing of these tweets in this grouping. Modi's tweets are overwhelmingly positive and usually around festivals or key moments such as congratulations or condolences for events related to individuals or groups. While Rahul Gandhi has similar tweets, he has also used congratulations with sarcasm. From our data, the two following tweets point to the differences in the two leaders approaches.

Modi's tweet, Figure 14, from early 2018 congratulates the blind cricket team for its victory, but rather than discussing their skill on the ground, his tweet uses upbeat language to present the disabled sportspersons as having a great attitude and being a source of inspiration for their fellow citizens. In contrast, Rahul Gandhi uses combative language and irony in congratulating Narendra Modi's chief ally, the BJP party president Amit Shah, specifically calling out a financial concern of his, and implying that the party has benefited from an event that had destructive outcomes for some Indians (Figure 14).

As we see in all of these tweets, the common thread is a tone of combativeness and sarcasm from Rahul Gandhi, compared to the positive-toned tweets from the Prime Minister. It is important though that all of the viral tweets were from a period when Modi has been in power, these roles may well be reversed for the two individuals. It is significant however that the affective value of sarcasm and challenge rests with the opposition leader, suggesting there is hope for contrarian discourse in the democratic process.

5. Discussion

In this paper, we first classified a large list of Indian politicians using collaborative networks on Twitter. We then presented a picture of social media behavior of Indian politicians, using their strategy on and outcomes with tweeting about policy-relevant topics. We presented these – technology, development, poverty, corruption, and inflation as policy-centric, whereas the comparative category of greetings is relatively less policy-centric. Building on previous research, part of this comparison was to contextualize past research that has suggested a dramatic shift in the media environment for politicians in terms of both their choice of subjects of discussion, as well as their evolving relationship with each other and the professional mainstream media corps. We show that politicians in opposition are keen to focus on the more antagonistic discussions, while politicians in power lean towards the less controversial topics. The messaging of the prime minister - in particular his avoidance of the inflation topic altogether suggests also that the movement away from communicating with a professional journalist corps may present serious challenges to the checks and media-led balances process.

We assessed 3 distinct dimensions: i) frequency as determined by the absolute number and normalized percentage of tweets that politicians of different parties contributed to each type of topics; ii) politicians' engagement with each other and with news media on Twitter when discussing salient issues comparing to when they are not; iii) text complexity and sentiment appeal of tweets belonging to different topics. To contextualize these, we examined a selection of the top tweets in each of the categories to provide a qualitative illustration on what some of the differences in online discursive style are.

We observed that politicians do indeed more frequently tweet about greeting and celebrations as compared to about corruption, development, and poverty. On the surface, this confirmed past research that politicians are focused more on building soft, personal ties rather than broadcasting their policies stands, viewpoints, or accomplishments/failures. However, when we look deeper at the qualitative findings, we see both that tweets about what would seem to be "policy" topics such as technology, or development can sometimes a priori be examples of political performance, or confrontation, rather than substantive discussion about policy. Likewise, we also see that greetings may not necessarily be benign or free of policy implications, but rather be deeply political or commentaries about policy issues.

We also observed that when politicians did talk about actual issues, they demonstrated a higher preference of engaging with news media comparing when they are tweeting about greetings. This implies that politicians are strategically involving news media when discussing actual policies and setting public policy agenda. Thus, rather than professional politics disengaging from the mainstream news process, which has been one of the concerns about the implications of social media use in Global South settings, we are seeing that politicians are still engaging the mainstream media in very specific ways.

We illustrated that policy and issue-focused tweets tend to have a much higher language complexity. This too has important implications, since it reconfirms prior research, some specific in India, of social media oversimplifying the political discourse, or moving it in the direction of banal, rather than politically substantive topics.

The mapping of specific politicians with repeat engagement in or thought leadership around the specific topic areas including corruption, poverty, and development intuitively suggest that active engagement in these areas is related to the domains of practice of these specific politicians. However, what we also see is both the clear dominance of the BJP in setting the agenda with many of these topical areas, as well as the relative difference in valence between parties. Non-ruling parties collectively depicted more negative sentiments on Twitter, especially on the topic of corruption. This observation reflected existing research in computer-supported collaborative work (CSCW) that focus on examining the role of party-based confrontations and attacks on social media and the impact these actions have on political good-will, compromises, and collaboration.

Finally, a deeper look at the tweets of the two most important leaders in our sample in terms of following and retweet activity - Narendra Modi and Rahul Gandhi - is deeply instructive on both their individual styles, but also on the differences in sentiment between ruling and opposition parties. While the ruling party politician steers clear of controversy and gets widely retweeted for his non-combative, positive tweeting, the opposition politician is rewarded for being aggressive and confrontational on social media.

There are several caveats in our study worth note here. First of all, while we provided in-depth assessment of politicians' Twitter priorities and strategies, our paper did not cover to what extent politicians are successful when utilizing these different strategies. Future work should focus on building normative measurements of politicians' success on Twitter. One possible metric could be how many new followers, retweets or mentions politicians gain by, for instance, providing higher emotional appeal in their tweets or engaging more with prominent news media account on Twitter for different type of issues. Furthermore, we also did not delve deep into politicians' social media strategies within different time periods (e.g. before, during, and after elections). Future work that makes this distinction can provide valuable insights into how politicians and politically inclined individuals shift their personal branding within different political environments.

6. Figures

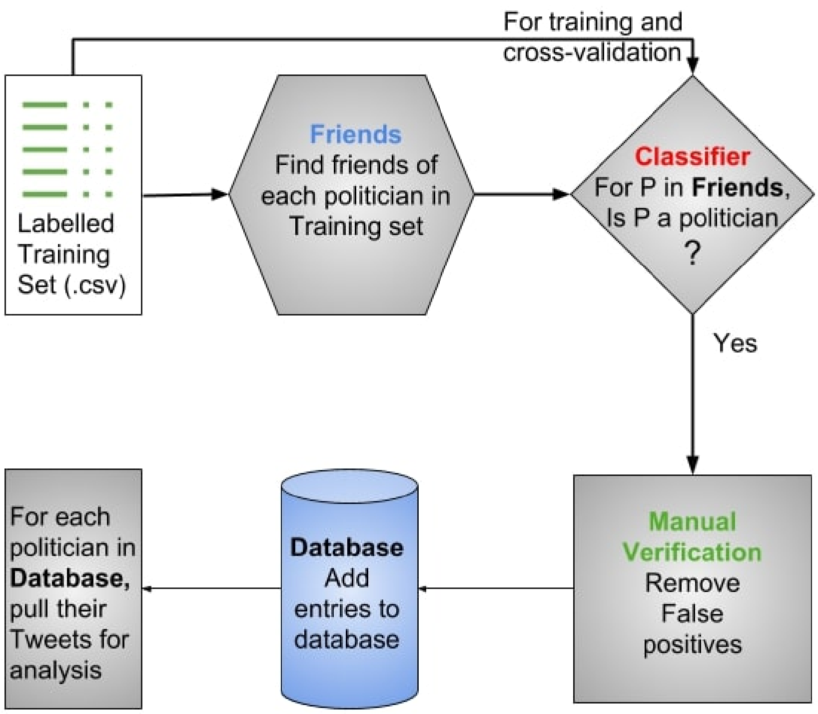

Figure 1. Flow chart of classification process

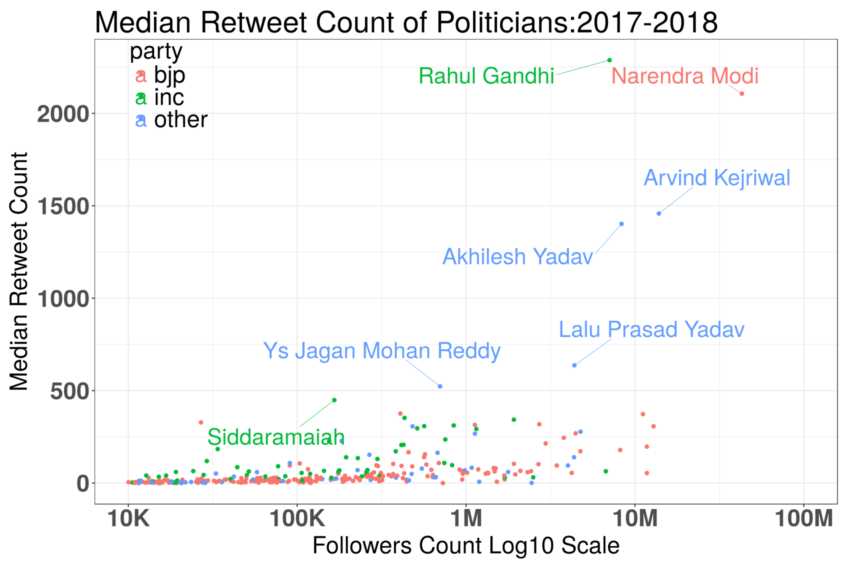

Figure 2. Median Retweet Count for Individual Political Account

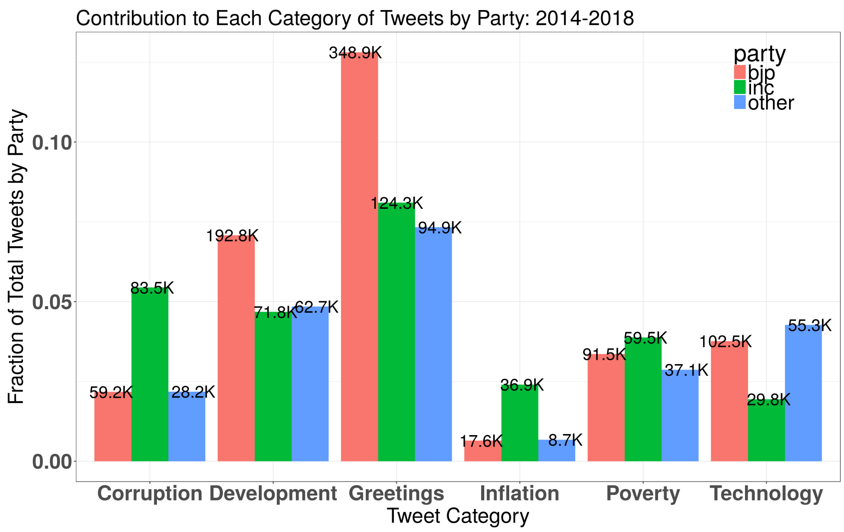

Figure 3. Party Contribution Distribution by Tweet Categories

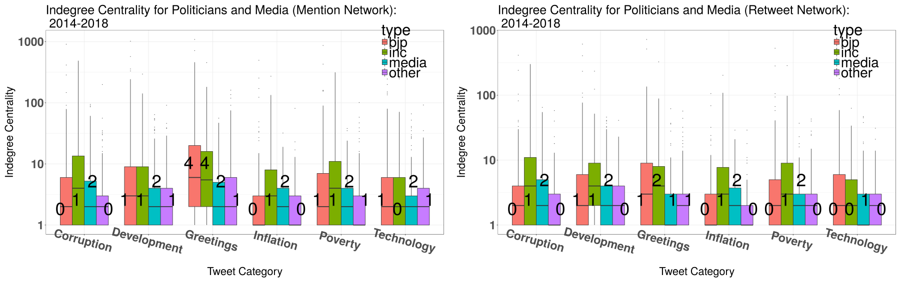

Figure 4. Indegree Mention Network and Indegree Retweet network

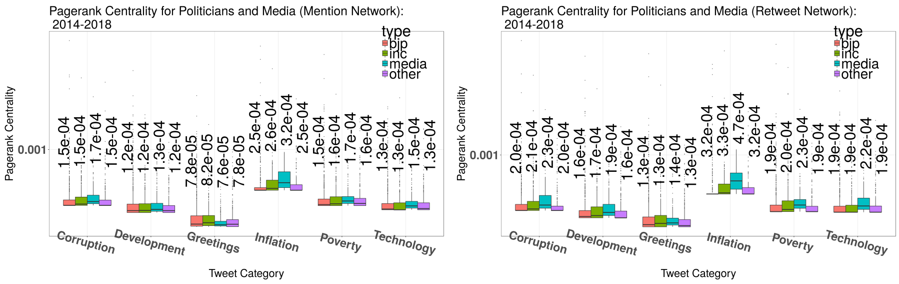

Figure 5. Pagerank Mention Network and Pagerank Retweet Network

Figure 6. Overall Compound Sentiment Score for Each Category of Tweets

Figure 8. Narendra Modi and Rahul Gandhi on `development`

Figure 10. Narendra Modi and Rahul Gandhi on `poverty`

Figure 12. Narendra Modi and Rahul Gandhi on `technology`

Figure 14. Narendra Modi and Rahul Gandhi on `greeting`

Figure 16. Narendra Modi and Rahul Gandhi on `inflation`

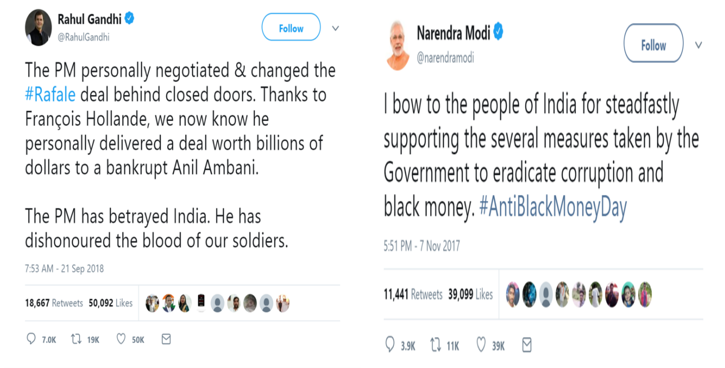

Figure 18. Narendra Modi and Rahul Gandhi on `corruption`

7. Tables

Table I. Top 10 Contributors (Followers>=100K) by Absolute Tweet Count for Each Category of Tweets

Category	Top10.Contributors
Technology	DIPP India(other 2073); Ravi Shankar Prasad(bjp 1977); Amitabh Kant(other 1696); Dr. Harsh Vardhan(bjp 1550); Suresh Prabhu(bjp 1252); Dharmendra Pradhan(bjp 1075); Jayant Sinha(bjp 908); Narendra Modi(bjp 896); Manoj Sinha(bjp 861); data.gov.in(other 796)
Poverty	Agriculture INDIA(other 1991); Yogendra Yadav(other 1297); Radha Mohan Singh(bjp 1182); Jitu Patwari(inc 877); ShivrajSingh Chouhan(bjp 828); Raghubar Das(bjp 785); Narendra Modi(bjp 573); UP Congress(inc 565); Narendra Singh Tomar(bjp 557); BJP Delhi(bjp 536)
Greetings	Narendra Modi(bjp 2762); Kailash Vijayvargiya(bjp 2647); Sudarsan Pattnaik(other 2125); Raghubar Das(bjp 2091); Mukhtar Abbas Naqvi(bjp 2012); Sidharth Nath Singh(bjp 1784); Nand Kishore Yadav(bjp 1776); Ashwini Kr. Choubey(bjp 1741); Vasundhara Raje(bjp 1687); Yogi Adityanath(bjp 1617)
Development	Raghubar Das(bjp 2180); Suresh Prabhu(bjp 1922); Agriculture INDIA(other 1852); Dharmendra Pradhan(bjp 1485); Radha Mohan Singh(bjp 1308); Narendra Modi(bjp 1259); Vasundhara Raje(bjp 1229); CMO Chhattisgarh(bjp 1228); Piyush Goyal(bjp 1118); Dr. Pankaj Shukla(bjp 1078)
Corruption	Sanjay Jha(inc 1277); Tejashwi Yadav(other 833); Subramanian Swamy(bjp 765); BJP Delhi(bjp 679); Jitu Patwari(inc 655); Sanjay Singh AAP(other 644); Sanjay Nirupam(inc 638); Sambit Patra(bjp 607); Amit Malviya(bjp 576); Randeep Singh Surjewala(inc 552)
Inflation	Dharmendra Pradhan(bjp 839); Jitu Patwari(inc 307); Sanju Verma(bjp 277); Sanjay Nirupam(inc 244); Alka Lamba(other 238); Randeep Singh Surjewala(inc 229); Harish Rawat(inc 215); Sanjay Jha(inc 202); Sitaram Yechury(other 159); Yogendra Yadav(other 154)

Table II. Regression Results for Readability and Sentiment with Respect to Party and Tweet Categorgy. Readability is Measured by Flesch, Gunning, and Smog Index. Sentiment is Measured Using VADER.

	Dependent variable:				
	Flesch	Smog	Gunning	Pos Emo	Neg Emo
	(1)	(2)	(3)	(4)	(5)
partyinc	0.620***	0.371***	0.367***	−0.030***	0.043***
	(0.085)	(0.067)	(0.101)	(0.003)	(0.002)
partyother	0.115	0.111*	0.075	−0.027***	0.020***
	(0.083)	(0.065)	(0.098)	(0.003)	(0.002)
categoryCorruption	0.724***	0.735***	1.144***	−0.299***	0.109***
	(0.116)	(0.092)	(0.138)	(0.004)	(0.003)
categoryDevelopment	2.039***	1.616***	2.506***	−0.259***	0.022***
	(0.112)	(0.089)	(0.134)	(0.004)	(0.003)
categoryInflation	−0.322***	−0.106	−0.169	−0.303***	0.072***
	(0.124)	(0.098)	(0.147)	(0.004)	(0.003)
categoryPoverty	1.343***	1.033***	1.377***	−0.286***	0.104***
	(0.114)	(0.090)	(0.136)	(0.004)	(0.003)
categoryTechnology	1.593***	1.183***	1.939***	−0.252***	0.016***
	(0.112)	(0.089)	(0.134)	(0.004)	(0.003)
friends_count	−0.286***				
	(0.060)				
followers_count	0.260***	0.336***	0.469***	0.013***	−0.007***
	(0.047)	(0.036)	(0.054)	(0.002)	(0.001)
statuses_count	0.108	−0.009	−0.276***	−0.018***	0.006***
	(0.066)	(0.046)	(0.069)	(0.002)	(0.002)
Constant	11.193***	10.819***	13.199***	0.449***	−0.00001
	(0.282)	(0.214)	(0.323)	(0.010)	(0.007)
Observations	8,116	8,116	8,116	8,116	8,116
R^2	0.080	0.072	0.073	0.519	0.254
Adjusted R^2	0.079	0.071	0.072	0.518	0.253
Residual Std. Error	3.077 (df = 8105)	2.431 (df = 8106)	3.668 (df = 8106)	0.109 (df = 8106)	0.082 (df = 8106)

Note: *p<0.1; **p<0.05; ***p<0.01

Table III. User descriptions of select twitter accounts

Screen name	User descirption (profile text)
narendramodi	Prime **Minister** of **India**
AmitShah	**National President** of **Bharatiya Janata Party / President**, Gujarat Cricket Association
RahulGandhi	This is the official account of Rahul Gandhi | **Member of Parliament** | President, **Indian National Congress**
ShashiTharoor	**MP** for Thiruvananthapuram. Author of 17 books. Former Minister of **State,Govt.**of India. Former UnderSecretaryGeneral,UnitedNations. RTs do not imply endorsement
VarshabenDoshi	**BJP Karykarta** | **Ex.M.L.A.**-Wadhwan **Constituency**, Surendranagar (Gujarat) (2007-2017) | Best National awardee teacher -2004 | Best State awardee teacher - 2001

Table IV. Most important features for positive and negative classes with linear SVM classifier and their feature weights

Positive class features	Weight	Negative class features	Weight
mp	1.67	super	-0.81
ministri	1.33	andhra	-0.8
spokesperson	1.18	tune	-0.67
parliamentari	1.17	columnist	-0.65
young	1.14	mani	-0.62
shri	1.13	legal	-0.6
mla	0.95	nadu	-0.56
vice	0.81	joke	-0.51
polit	0.81	liber	-0.51

7.1. KEYWORDS

Complete list of keywords used to generated each category of tweets are shown in Table V. Note that Hindi tokens were translated into their respective English alphabetic representation.

Table V.: List of Keywords and Terms Used to Generate Each Category of Tweets. Note: all Hindi Tokens were Translated.

Category	Keywords
Technology	comput, internet, kiosk, scienc, scient, technolog, telecom, satellite, digital, engineer, innovati, biotech, start up, software, startup, wifi , fintech, pslv , computer, device, multimodal, biometric, msme , electrical, 4g , aviation, electronic, cisf , app , battery, isro , automobile, portal, bhimapp, digitalindia, engineering, hotspot, wi-fi, web , multi-modal, innovate, iot , technological, real-time, automatic, dashboard, automated, csir , brahmos, software, digitisation, nano , tracking, rocket, smartphones, invention, test-firing, drdo , digital, smartphone, gsat-19, bsnl , sme , start-up, world-class, electronics, msmes, biotech, gps , mfg , digitalpayments, cashless, internet, jio , railwire, digitalpayment, petrochemical, e-commerce, robotics, technology, apps , virtual, technologist, mobileapp, entrepreneur, wifi , hardware, drone, prototype, resourcesat-2a, ip , broadband, ict , scientist, orbit, makeinindia, entrepreneurship, eco friendly, tech , e-wallets, android, technical, innovation, upi , state-of-the-art, evinindia, cng , startup, google, multi-purpose, pioneer, umang, cipet, srcitizen science, nobelprize, cartosat-2, mangalyaan, scientific, petroleum, startupindia, hydrocarbon, gsat , digita, tablet, digiti, ecommerce, techno, android, aeps , choupal, pslvc38, cartosat-2, dispensaryautomotive, kampyootar, vigyaan, paaribhaashikee, vaigyaanik, kampyootar, shilpkala, injeeniyar, praudyogikee, phesabuk, tvitar, intaranet, instraagraam, vaigyaanikon, aadhunik, ilektroniks, fesabuk, tvitar, atyaadhunik, kampyootar, vaigyaanik, ilektronik, takaneekee, intaranet, brodabaind, vigyaanan, vaeephaee, sophtaveyar, teknolojee, vebasait, sikyoritee, laeev , tveetar, injeeniyar, ee-mitr, vigyaan, klastar, phesabuk, post, kampyootar, googal

Poverty	accessib, poor , alleviation, malnutrition, financialinclusion, low-income, homeless, unemploy, handicap, amenity, prostitution, marginalized, neo-middle, mortality, worldfooddaytrapped, destitute, rightsnotcharity, casteism, neglected, exploited, inequality, backwardness, starving, illiteracy, deprivation, foodwater, robbed, upliftment, struggling, downtrodd, shortage, benefitting, govtthatcares, child labourvulnerable, untouchab, deprived, under-privileged, poorest, underserved, emancipation, casteless, downtrodden, poverty, marginali, unprivileged, exploiting, underprivileged, jobless, nrega, slumdisadvantaged, childlabour, labourer, hard-working, morbidity, survival, globalhungerindex, assisted-living, elderly, unskilled, exploitation, affordab, mathrupoorna, dwindling, remotest, hardshiplivelihood, divyang, welfare, needy, universalhealthcoverage, rural, garib, houseless, daridr, niraashrit, besahaara, nirdhan, bekas , mufali, abhaav , gareeb , nirbal, muhataaj, zaroorataman d, bhookh , peedit, shoshit, kisaan, krsh , laachaar, bebas , peedit, laachaar
Development	self-reliance, iip , developmental, economy, cleanenergy, socio-economic, indiameansbusiness, gdpgrowth, gva , patenting, employability, growth, consumption, growth, urbanisation, infrastructure, year-on-year, budgetforbetterindia, double-digit, agricultural, macro-economic, globalization, trajectory, fiscal, q3 , deflation, growt, industrialization, q2 , modernity, competitiveness, expanding, infrastructural, dvpt , urbanization, macroeconomic, newindiatakeoff, accelerating, development, liquidity, budget2017, wpi , infra, devpt, psbs , investment, modernising, expenditure, productivity, mfg , accessibility, disparity, gdp , infrastruct, connectivity, dvlp , devt , devl , productivity, globalisation, protectionism, fy17 , q4 , production, upliftment, economicsurvey2018, shrinking, efficiency, crisil, inflow, slowdown, logistics, grwth, connectivity, q1 , macro, oilseed, transport, transportation, output, yoy , diversification, equality, steep, borrowing, deficit, profitability, maize, sectoral, indianeconomy, stimulate, export, krishikalyanabhiyan, ikaas, vrddhi, sanvrddhi, utpaadan, unnati, samrddhi, baagavaanee, badhotaree, petrol-deejal, khetihar, phal-sabjiyon, vrddhi, vikaas, mahangaee, badhottaree, mahangaee, pashupaalakon, utthaan, svarozagaar, utpaadakata, kushalata, keematon, vikaas., karz , pratishat, utpaadan, vrddhi., sinchaee, arthavyavastha, utpaadan, badhottaree, khetee , badhottaree, paidaavaar, unnayan, vrddhi, rojagaar, udyaanikee

Greetings	harvest, vijayadashami, sankranti, mubarakmay, b’day, republic day, wonderful, gurupurnima, utkaladibasa, christmas, cheiraoba, celebrated, tranquility, ramadan, gratitude, eiduladha, new year, warmth, celebrating, eidmubarak, festivity, xmas , mahashivaratri, makarsankranti, birthday, long life, colourful, compassion, festival, childrensday, holi , prosper, id-e-milad, bday , loving, greetngs, fantastic, peacefulness, joyful, prosperi, wishing, wish , brotherhood, birthday, happybirthday, fabulous, greeting, joy , baisakhi, bless, heartiest, gudipadwa, peace, oneness, merry, beloved, bathukamma, longlife, eidaladha, akshayatritiya, happyholi, eidulfitr, bohagbihu, joyous, harmonious, blessed, heartfull, boishakh, congradulations, congradul, congratulation, congratulat, happydiwali, delightful, reopens, puthandu, holiday, dhanteras, onam , easter, greeted, happiness, auspicious, janmashtami, sankranthi, bliss, mahavirjayanti, greets, abundance, samvatsari, colorful, praying, carnival, heartwarming, harmony, heartfelt, prosperity, cherish, contentment, blissful, pongal, terrific, jesus, divine, diwali, gudhi padwa, good health, independence day, sincerest, harvest, vijayadashami, sankranti, mubarakmay, b’day, republic day, wonderful, gurupurnima, utkaladibasa, christmas, cheiraoba, celebrated, tranquility, ramadan, gratitude, eiduladha, new year, warmth, celebrating, eidmubarak, festivity, xmas , mahashivaratri, makarsankranti, birthday, long life, colourful, compassion, festival, childrensday, holi , prosper, id-e-milad, bday , loving, greetngs, fantastic, peacefulness, joyful, prosperi, wishing, wish , brotherhood, birthday, happybirthday, fabulous, greeting, joy , baisakhi, bless, heartiest, gudipadwa, peace, oneness, merry, beloved, bathukamma, longlife, eidaladha, akshayatritiya, happyholi, eidulfitr, bohagbihu, joyous, harmonious, blessed, heartfull, boishakh, congradulations, happydiwali, delightful, reopens, puthandu, holiday, dhanteras, onam , easter, greeted, happiness, auspicious, janmashtami, sankranthi, bliss, mahavirjayanti, greets, abundance, samvatsari, colorful, praying, carnival, heartwarming, harmony, heartfelt, prosperity, cherish, contentment, blissful, pongal, terrific, jesus, divine, diwali, gudhi padwa, good health, independence day, haardik, shubh , janmadin, chhuttee ka din, avasar par, shubhakaamanaen, holee , naya saal, tyohaar, gaandhee jayantee, mahaaveer jayantee, buddh poornima, lohadee, janmadivas, raamanavamee, janmadin, baisaakhee, janamadin, pongal, punyatithi, shubhakaamanaayen, holika, shubhakaamanae, shubhakaamanaen, deepaavalee,jeevanaparv, navaraatr, shubhakaamanaen, navaraatree, shubhakaamanaen., shubh , jayantee, shubhakaamanaayen, shubhakaamanaaye, janmaashtamee, krisamas, dhanateras, shubhakaamanaen, mubaarakabaad, mahaaparv, makar , vijayadashamee, badhaiyaan, jayantee, janm , eed , shubhakaamana, janmotsav, mahaashivaraatri, janmadeen, tyauhaar, utsav, parv , makarasankraanti, shubhakaamana, shubhakaam, badhaiyaan, paavan , vijayaadashamee, rakshaabandhan, phitar, janyatee, mubaarak, mangalamay

Corruption	rafale, corruption, deal , scam , rafalescam, rafaledeal, cag , mumbaiagainstrafalescam, middleman, rafalescamsilence, corrupt, anothercongressdefencescam, congressdefencescam, hal , crony, defaulters, deals, doublestandards, scandal, cheating, transparency, quattrocchi, truthofrafale, bofors, 130000crorerafalescam, scams, dubious, exposed, iafbacksrafale, rafalerobbery, modiscandals, blackmoneycrackdown, siphoning, rafalescammonday30th, hawala, defaulter, proberafalescam, modiscam, vyapam, modiambanirafaleblockbuster, thegreatrafalemystery, kickback, blackmoney, augusta, westland, bribe, chowkidarnahibhagidar, fraudster, rafalechargesheet, jpc , embezzlement, corruption-free, fraudulent, rahulgandhiwithhal, bribery, rafalepejhoothbandkaro, agusta, agustawestland, mehulchoksi, modiambanirafalescam, rafaledeal, favouritism, aircel-maxis, pnbfraudcase, rafalescamexpose, pnbscam, modighotala, choksi, pnbfraud, dealmeinkuchkalahai, niravmodilootsindia, niravmodi, robert vadra, mehul choksi, adarsh scam, national herald, rahulkapurakhandanchor, 2g spectrum, antigua, vijay mallya, puricongresschorhai, raaphel, raphel , raphaal , raafel, bhrasht, bhrashtaachaar, jhoot , sauda , saude , agasta, vestalaind, vestalend, vestalaind, dhokha , dhoka , dhokhaadhadee, dhokaadhadee, dhokebaaj, dhokhebaaz, vyaapam, vyaapaman, vyaapam, khaoonga, ophaset, ghotaalebaajon, ghotaalon, ghotaalo, ghotaale, ghotaala, ghotaalebaaj, dalaalee, dalaal , kameeshanakhoree, kameeshan, kamishan, kameeshanakhor, kamishanakhor, kamishanakhoree, kameeshanabaazee, lokapaal, janalokapaal, poonjeepati, kaaledhan, kaalaadhan, aadarshaghotaala_sangharsh_yaatra, rishvat, rishavat, rishvatakhor, rishvatakhoree, rishvatabaazee, jhaansa, karapshan, robart, vaadra, vaadra, aadarsh ghotaala, 2g , bophors, bofors, vijay maalya, maalaya, aanteegva, enteegua, entigva, aanteegua, ghotalapradeshmp, puricongresschorhai, ghotala, rafale, vyapam
Inflation	petrol, diesel, bharatbandh, pricing, loot , fuelloot, lpg , inflation, cylinder, highdiesel, fuelhikeprice, bharatbandhinflation, petrolpricehike, mehngaikiaag, rocketed, fuelonfire, crude, fuellootbysuitboot, wpi , petrolchormodi, inflation, petroldieselpricehike, daam , silendar, petrol, deezal, petrol-deezal, mahangaee, silendar, silindar, daamon, petroladeejal, paitrol, rasoee , elapeejee, manhagaee, mahangaee

Table VI. Parameter Grid for `RandomForest` classifier

Parameters	Range of values
n_estimators	[5, 6 , **7**, 8, 9]
min_samples_leaf	[**40**, 45, 50, 55]
max_features	[**None**, "sqrt", "log2"]
n_jobs	[**4**, 8]
scoring parameter	F1

Table VII. Parameter Grid for `SVM` classifiers with `liner`, `rbf` and `poly` kernels

Parameters	Range of values		
	SVM (Linear kernel)	SVM (RBF kernel)	SVM (poly kernel)
C	[0.1, 0.3, **1**, 3, 10, 30, 100, 1000]	[0.1, 0.3, **1**, 3, 10, 30, 100, 1000]	[0.1, 0.3, **1**, 3, 10, 30, 100, 1000]
gamma	[**10**, 1, 0.01, 0.1, 0.001]	[10, 1, 0.01, **0.1**, 0.001]	[10, **1**, 0.01, 0.1, 0.001]
kernel	linaer	rbf	poly
degree	-	-	[**2**,3,4]
scoring parameter	F1	F1	F1

Table VIII. Performance metrics of different classifiers

Classifier	Accuracy	Precision	Recall	F1 score
RandomForest	0.805	0.901	0.673	0.770
NaiveBayes	0.817	0.737	**0.969**	0.837
LogisticRegression	0.898	0.918	0.860	0.887
StochiasticGD	0.894	0.910	0.860	0.882
SVM (linear)	**0.907**	0.904	0.898	**0.901**
SVM (rbf)	0.898	**0.930**	0.847	0.887
SVM (poly)	0.898	0.924	0.853	0.887

References

(2018), India: number of Twitter users 2019 | Statistic.

Adamic, L.A.; and N. Glance (2005-08). The political blogosphere and the 2004 US election: divided they blog. , 2005-08. p. 3643.

Agrawal, B.C.and S.Raghaviah (2006). India Public Service Broadcasting and Changing Perspectives. Public service broadcasting in the age of globalization. : , *I. Banerjee and K. Seneviratne* pp. 149164.

Ahmed, S.; ; and K.Jaidka J. Cho (2016). The 2014 Indian elections on Twitter: A comparison of campaign strategies of political parties. *Telematics and Informatics*, vol. 33, no. 4, pp. 10711087.

Ahmed, S.and M.M.Skoric (2014). My name is Khan: the use of Twitter in the campaign for 2013 Pakistan General Election. *47th Hawaii International Conference on System Sciences (HICSS), IEEE.*

Baumer, Eric PS; Mark Sueyoshi; and Bill Tomlinson (2011). Bloggers and readers blogging together: Collaborative co-creation of political blogs. *Computer Supported Cooperative Work (CSCW)*, vol. 20, no. 1-2, pp. 1–36.

Bennett, W.L. (2012). The personalization of politics: Political identity, social media, and changing patterns of participation. *The ANNALS of the American Academy of Political and Social Science*, vol. 644, no. 1, pp. 2039.

Bennett, W.L.; and B. Pfetsch (2018). Rethinking political communication in a time of disrupted public spheres. *Journal of Communication*, vol. 68, no. 2, pp. 243253.

Bimber, B. (2014). Digital media in the Obama campaigns of 2008 and 2012: Adaptation to the personalized political communication environment. *Journal of Information Technology and Politics*, vol. 11, no. 2, pp. 130150.

Bonacich, P. (1987). Power and centrality: A family of measures. *American journal of sociology*, vol. 92, no. 5, pp. 11701182.

Bureau, Z.M. (2018), PM Narendra Modi tells BJP MPs to have at least 3 lakh followers on Twitter.

Casero-Ripollés, A.; ; and M.Sintes-Olivella P. Franch (2017). "The populist political communication style in action: Podemos's issues and functions on Twitter during the. *American Behavioral Scientist*, vol. 61, no. 9, pp. 9861001.

Chadwick, A. (2017). *The hybrid media system: Politics and power*. Oxford University Press.

Chakraborty, S.; J. Pal; ; and P.Chandra D. M. Romero (2018). Political Tweets and Mainstream News Impact in India: A Mixed Methods Investigation into Political Outreach. , 2018.

Conover, M.; J. Ratkiewicz; M.R. Francisco; B. Gonçalves; F. Menczer; and A. Flammini (2011). Political polarization on twitter. *Icwsm*, vol. 133 pp. 8996.

Correa, J.C.and J.E.Camargo (2017). "Ideological consumerism in colombian elections. : , *Cyberpsychology, Behavior, and Social Networking*.

Curato, N. (2017). Flirting with authoritarian fantasies? Rodrigo Duterte and the new terms of Philippine populism. *Journal of Contemporary Asia*, vol. 47, no. 1, pp. 142153.

David, Y.and C.Baden (2018). Reframing community boundaries: the erosive power of new media spaces in authoritarian societies. : , *Information, Communication and Society* pp. 118.

Dutton, W.H. (2009). The fifth estate emerging through the network of networks. *Prometheus*, vol. 27, no. 1, pp. 115.

Enli, G. (2017). "Twitter as arena for the authentic outsider: exploring the social media campaigns of Trump and Clinton in the. *European Journal of Communication*, vol. 32, no. 1, pp. 5061.

Enli, G.S.and E.Skogerbø (2013). Personalized campaigns in party-centred politics: Twitter and Facebook as arenas for political communication. *Information, Communication and Society*, vol. 16, no. 5, pp. 757774.

Feltwell, Tom; John Vines; Karen Salt; Mark Blythe; Ben Kirman; Julie Barnett; Phillip Brooker; and Shaun Lawson (2017). Counter-Discourse Activism on Social Media: The Case of Challenging Poverty Porn Television. *Computer Supported Cooperative Work (CSCW)*, vol. 26, no. 3, pp. 345–385.

Flaounas, I.; O. Ali; T. Lansdall-Welfare; T. De Bie; N. Mosdell; J. Lewis; and N. Cristianini (2013). Research methods in the age of digital journalism: Massive-scale automated analysis of news-contenttopics, style and gender. *Digital Journalism*, vol. 1, no. 1, pp. 102116.

Fuller, G.; ; and A.Jolly C. Fisher (2018). Malcolm Turnbull's conversational career on Twitter: the case of the Australian Prime Minister and the NBN. *Media International Australia*, vol. 167, no. 1, pp. 88104.

Gayo-Avello, D. (2017). Social Media Won't Free Us. *IEEE Internet Computing*, vol. 21, no. 4, pp. 98101.

Handler, Reinhard A; and Raul Ferrer Conill (2016). Open Data, Crowdsourcing and Game Mechanics. A case study on civic participation in the digital age. *Computer Supported Cooperative Work (CSCW)*, vol. 25, no. 2-3, pp. 153–166.

Hemphill, Libby; Jahna Otterbacher; and Matthew Shapiro (2013). What's congress doing on twitter? , 2013. pp. 877–886.

Howard, P.N. (2006). *New media campaigns and the managed citizen*. Cambridge University Press.

Hutto, C.J.; and E.E. Gilbert (2014-06). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*. Ann Arbor, MI: .

Jaffrelot, C. (2015). Narendra Modi and the power of television in Gujarat. : , *Television and New Media*.

Jungherr, A. (2014). The logic of political coverage on Twitter: Temporal dynamics and content. *Journal of Communication*, vol. 64, no. 2, pp. 239259.

Kaur, Ravinder (2015). Good times, brought to you by brand Modi. *Television & New Media*, vol. 16, no. 4, pp. 323–330.

Kellner, D. (2018). Donald Trump and the War on the Media: From Election '16 into the Trump Presidency. *The Trump Presidency, Journalism, and Democracy, Routledge*. p. 1938.

Kou, Yubo; Yong Ming Kow; Xinning Gui; and Waikuen Cheng (2017). One Social Movement, Two Social Media Sites: A Comparative Study of Public Discourses. *Computer Supported Cooperative Work (CSCW)*, vol. 26, no. 4-6, pp. 807–836.

Kreiss, D. (2016). Seizing the moment: The presidential campaigns' use of Twitter during the 2012 electoral cycle. *New Media and Society*, vol. 18, no. 8, pp. 14731490.

Kumar, S. (2015). Contagious memes, viral videos and subversive parody: The grammar of contention on the Indian web. : , *International Communication Gazette*.

Kwak, H.; C. Lee; H. Park; and S. Moon (2010-04). What is Twitter, a social network or a news media? , 2010-04. p. 591600.

Lee, C. (2017). Facebooking to Power: The Social Media Presence of Malaysian Politicians. *ISEAS Perspectives*, vol. 74.

Loader, B.D.and D.Mercea (2011). Networking democracy? Social media innovations and participatory politics. *Information, Communication and Society*, vol. 14, no. 6, pp. 757769.

Lotan, G.; E. Graeff; M. Ananny; ; and D.Gaffney I. Pearce (2011). "The Arab Springl the revolutions were tweeted: Information flows during the. *International journal of communication*, vol. 5 pp. 31.

Ludwig, Thomas; Christian Reuter; and Volkmar Pipek (2016). From publics to communities: Researching the path of shared issues through ICT. *Computer Supported Cooperative Work (CSCW)*, vol. 25, no. 2-3, pp. 193–225.

Lufkens, M (2016). Twiplomacy Study 2016. : , *Burson-Marsteller, Editor, Geneva*.

Maheshwari, S.and C.Sparks (2018). Political elites and journalistic practices in India: A case of institutionalized heteronomy. : , *Journalism*.

Mascaro, Christopher M; and S Goggins (2010). Collaborative information seeking in an online political group environment. , 2010.

McGregor, S.C. (2018). Personalization, social media, and voting: Effects of candidate self-personalization on vote intention. *new media and society*, vol. 20, no. 3, pp. 11391160.

Meeks, L. (2016). Gendered styles, gendered differences: Candidates' use of personalization and interactivity on Twitter. *Journal of Information Technology and Politics*, vol. 13, no. 4, pp. 295310.

Morozov, E. (2009). Iran: Downside to the "twitter revolution". *Dissent*, vol. 56, no. 4, pp. 1014.

Mudgal, V. (2015). Framing the 2014 elections: The curious absence of development. *Television and New Media*, vol. 16, no. 4, pp. 354360.

Niederer, Sabine; and Ruurd Priester (2016). Smart citizens: Exploring the tools of the urban bottom-up movement. *Computer Supported Cooperative Work (CSCW)*, vol. 25, no. 2-3, pp. 137–152.

Ohm, B. (2015). Organizing Popular Discourse with and against the Media: Notes on the Making of Narendra Modi and Recep Tayyip Erdogan as Leaders-without-Alternative. *Television and New Media*, vol. 16, no. 4, pp. 370377.

Ott, B.L. (2017). The age of Twitter: Donald J. Trump and the politics of debasement. Critical Studies in Media. *Communication*, vol. 34, no. 1, pp. 5968.

Pal, J.; P. Chandra; and V.V. Vydiswaran (2016). Twitter and the rebranding of Narendra Modi. *Economic and Political Weekly*, vol. 51, no. 8, pp. 5260.

Pal, J.and A.Gonawela (2016). *Political social media in the global South. Conference on e-Business, e-Services and e-Society*. Springer.

Pal, Joyojeet (2015). Banalities turned viral: Narendra Modi and the political tweet. *Television and New Media*, vol. 16, no. 4, pp. 378–387.

Pal, Joyojeet; Udit Thawani; Elmer van der Vlugt; Wim Out; Priyank Chandra; et al. (2018). Speaking their Mind: Populist Style and Antagonistic Messaging in the Tweets of Donald Trump, Narendra Modi, Nigel Farage, and Geert Wilders. : , *Computer Supported Cooperative Work (CSCW)* pp. 1–34.

Park, M.-J.and J.Curran (2000). *De-Westernizing media studies*. Psychology Press.

Priante, A.; D. Hiemstra; T. van den Broek; A. Saeed; M. Ehrenhard; and A. Need (2016). #WhoAmI in 160 characters? Classifying social identities based on twitter profile descriptions. , 2016. p. 5565.

Rahman, A. (2014). The problems with reimagining public media in the context of global South. *Stream: Inspiring Critical Thought*, vol. 6, no. 1, pp. 5665.

Rajagopal, A. (2017). On Media and Politics in India: An Interview with Paranjoy Guha Thakurta. *South Asia: Journal of South Asian Studies*, vol. 40, no. 1, pp. 175190.

Rao, U. (2010). *News as culture: Journalistic practices and the remaking of Indian leadership traditions*. Berghahn Books.

Rodrigues, U.M. (2014). Social media's impact on journalism: a study of media's coverage of anti-corruption protests in India. *Global media journal: Australian edition*, vol. 8, no. 1, pp. 110.

Rodrigues, U.M.and M.Niemann (2017). Social Media as a Platform for Incessant Political Communication: A Case Study of Modi's Clean India Campaign. *International Journal of Communication*, vol. 11 pp. 23.

Saeed, S. (2015). Phantom Journalism: Governing India's proxy media owners. *Journalism Studies*, vol. 16, no. 5, pp. 663679.

Saeed, Saqib; Markus Rohde; and Volker Wulf (2009). Technologies within transnational social activist communities: an ethnographic study of the european social forum. , 2009. pp. 85–94.

Schäfer, Mirko Tobias (2016). Challenging Citizenship: Social Media and Big Data. *Computer Supported Cooperative Work (CSCW)*, vol. 25, no. 2-3, pp. 111–113.

Sharma, P.; and T.S. Moh (2016-12). Prediction of indian election using sentiment analysis on hindi twitter. *2016 IEEE International Conference on Big Data (Big Data)*. IEEE, p. 19661971.

Soroka, S.; L. Young; and M. Balmas (2015). Bad news or mad news? Sentiment scoring of negativity, fear, and anger in news content. *The ANNALS of the American Academy of Political and Social Science*, vol. 659, no. 1, pp. 108121.

Stieglitz, Stefan; and Linh Dang-Xuan (2013). Social media and political communication: a social media analytics framework. *Social Network Analysis and Mining*, vol. 3, no. 4, pp. 1277–1291.

Sundaram, R. (2015). Publicity, transparency, and the circulation engine: the media sting in India. *Current Anthropology*, vol. 56, no. 12, pp. 297305.

Tusa, F. (2013). "How social media can shape a protest movement: The cases of Egypt in. *Arab Media and Society*, vol. 17 pp. 119.

Wilkerson, J.; and A. Casas (2017). Large-scale computerized text analysis in political science: Opportunities and challenges. *Annual Review of Political Science*, vol. 20 pp. 529544.

Zittrain, J.L.; R. Faris; H. Noman; J. Clark; ; and C.Tilton R. Morrison-Westphal (2017), The Shifting Landscape of Global Internet Censorship.