Functional Analysis of the 2020 U.S. Elections on Twitter and Facebook using Machine Learning

Saud Alashri
King Abdulaziz City for Science and Technology
Riyadh, Saudi Arabia
salashri@kacst.edu.sa

Turki Alalola
Prince Sultan University
Riyadh, Saudi Arabia
alola.turki@gmail.com

***Abstract*—Social Networking Sites (SNS), such as Facebook and Twitter, are important tools for political campaigns. A line of related work analyzed political campaigns online. The initial efforts in analyzing campaign discourse functions relied on human analysis, which is time consuming and does not scale well with big data. To address these gaps, we propose a model to detect the type of campaign topics: Policy vs. Character, and how the public (commentators) responded to these messages. The proposed model yielded an accuracy of 78% (F-measure) in detecting post type. Moreover, experimental results show the analysis of commentators linguistic and psychological characteristics.**

***Index Terms*—Social Networking Sites, Political Discourse, US Elections, Big Data, Machine Learning, Discourse Analysis, Functional Theory, Text Classification.**

I. Introduction

Social networking sites (SNS) have progressed to a new level in engagement and communication. Within political arena, outlets such as Facebook and Twitter are legitimate and paramount spaces for political participation [1]–[5]. The 2008 U.S. presidential campaign broke a new ground in campaigning strategies and public relations. President Barack Obama was the first U.S. candidate who utilized social media outlets as complementary tools for campaigning online. In addition to his offline campaign, President Obama leveraged various social media platforms such as Twitter, YouTube, Facebook, Digg, and many others to reach out to larger communities online [6].

President Obama's successful campaign in 2008 opened the door for leveraging SNS in the U.S. for political activities and set a trend that goes beyond U.S. to other parts of the globe. Historically, Millennials (or youth) in the U.S. have low political engagement; however, according to the Pew Research Center (2016), the emergence of social media has attracted most of them to learn about elections [7]. Moreover, according to the Pew Research Center (2019), 72% of Americans have adopted some type of social media with little variation across demographic characteristics like race, income, gender, education and ages 18-49; with Facebook being the most popular platform (69% of U.S. adults) and Twitter nearly 22% [8].

The type of social media platform is a major trigger for poltical participation. According to Williams and Gulati [9], Facebook is popular for political campaigns as it provides unique features such as *newsfeed* and *wall* that show opinions and discussions within user's network and interest. Twitter, the microblogging platform with unique features such as *hashtags* and *mentions*, is one of fastest growing SNS with 330 million monthly active users[1] and 20% of its daily users are Americans[2].

Extant literature presented several studies on elections and political engagements on SNS. Williams and Gulati [9] compared several SNS for political campaigning and concluded that Facebook is the leading platform as it provides additional features for mobilizing voters. In [10] Williams and Gulati concluded that the success of a political campaign is associated with online activities on SNS. The authors noted that the volume of Facebook supporters might be a cogent indicator for campaigning success. Enli and Skogerbø [11] studied Norwegian elections and found SNS enabled "personalized campaigning" and changed the way of how candidates and voters are communicating and engaging. Andersen and Medaglia [12] conducted a survey to study the role of SNS in connecting citizens with politicians in Denmark. They found nearly 57% of respondents utilized Facebook to connect and communicate with their favorite politician.

Numerous studies proposed methods to analyze SNS content to investigate various elements such as election outcomes, campaigning strategies, and candidates/voters behaviors. Hong and Nadler [13] studied the 2012 U.S. presidential elections to understand the relationship between candidates' online and offline activities. They found strong association between candidates' Twitter activities and the number of mentions they receive on Twitter and traditional media. Tumasjan et al. [14] analyzed over 100,000 tweets related to 6 parties in the 2009 German federal elections to investigate the utility of Twitter for political deliberation and its impact on the public. They found the number of tweets (mentions) is a valid indicator of campaign success and it reflects the traditional election polls. They also found the sentiment of Twitter messages matches offline public reaction.

There is substantial change in the use of SNS in the recent U.S. elections as compared to 2008 elections. During 2020 elections, all U.S. candidates have leveraged SNS as an

[1]https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/

[2]https://blog.hootsuite.com/twitter-statistics/

IEEE/ACM ASONAM 2020, December 7-10, 2020

TABLE I
DATASETS DISTRIBUTION

Dataset	Trump	Sanders	Biden
Facebook Posts	480	531	459
Twitter Posts	1,300	1,192	867
Comments	5,511,705	310,992	251,840

essential channel for communication and campaigning. The focus of our paper is to study the 2020 U.S. elections on Facebook and Twitter to investigate the type of candidates' messages from *Functional Theory* [15]–[17] perspective and public reaction on their messages.

Functional Theory is defined as the means that political candidates utilize to distinguish themselves and convince voters when campaining; which could be achieved through three functions: acclaiming, attacking, and defending [15]–[17]. These functions occur on two broader campaign topics: *Policy,* and *Character*. Policy topics addresses subtopics such as past deeds, general goals, and future plans; and Character topics accosts subtopics such as leadership ability, personal qualities, and values/principales [18].

The initial efforts in analyzing campaign discourse *functions* relied on human analysis [16], [19], [20], which is time consuming and does not scale well when analyzing big data. In order to address these limitations, machine learning and data mining techniques can be employed to analyze and classify text automatically. In this paper we propose a model to detect the type of campaign topics on Facebook and Twitter: *Policy* vs. *Character*. Additionally, we analyze commentators comments on candidates posts to understand their linguistic and psychological characteristics.

To do so, we conducted an exploratory analysis for the official Facebook and Twitter pages of the top three U.S. presidential candidates: Donald Trump (Trump), Bernie Sanders (Sanders), and Joe Biden (Biden). We collected over 4,800 candidates posts, and over 6 million comments on those posts from November 1^{st}, 2019 to February 20^{th}, 2020.

The article proceeds in four parts. The next section describes the methodology. The results and analysis of findings are presented in Section III. Section IV concludes the paper and highlight directions for future work.

II. METHODOLOGY

A. Dataset

Our dataset is comprised of nearly 4,825 Facebook posts and Twitter tweets by the top three U.S. presidential candidates: Donald Trump, Bernie Sanders, and Joe Biden; and 6,074,537 comments on those posts from the public (commentators). The dataset spans a period of time from November 1st, 2019 to February 20th, 2020. We developed a Python script to collect the data from the official Facebook and Twitter pages. Table I presents the distribution of dataset.

B. Preprocessing

Preprocessing step is applied on the raw data by removing none textual posts and comments, as well as text that is not in English language. Further, spam and bot-generated comments is also removed; those comments typically follow similar and repeated pattern such as: *VOTE FOR SANDERS VOTE FOR SANDERS VOTE FOR SANDERS*, and *MAGA* (MAKE AMERICA GREAT AGAIN) *MAGA*. Additionally, stopwords were removed, and all words are converted into lowercase to standardize them when applying the analysis step.

C. Post-Type Detection

To detect whether candidates posted *Policy, Character,* or *Mixed/Other* posts on Facebook and Twitter, we build machine learning models to detect post type as follows:

- process text by first tokenizing words using NLTK (The Natural Language Toolkit) [21] and then apply stemming and lemmatization utilizing Gensim [22]
- generate two groups of features by utilizing *Word2Vec* [23] which is a representation of text in vector space, and TF-IDF (Term Frequency - Inverse Document Frequency) [24]–[26]
- experiment both three-way classification (three categories: *Policy, Character, Mixed/Other*) and two-way classification (two categories: *Policy, Character*)

We experimented various classification models namely: neural networks classification models both Convolutional Neural Network (CNN) as well as Multi-layer Perceptron (MLP) [27], Linear Support Vector Regression (L-SVR), and Multinomial Naive Bayes [28].

D. Analysis of Commentators

Once we identify post type by each candidate, we examined how public (commentators) reacted to those posts and what are their characteristics. We analyzed commentators comments utilizing Linguistic Inquiry and Word Count(LIWC) [29] to measure the linguistic and psychological indices of commentators as follows:

- measure commentators emotions such as: *positive emotion, negative emotion, sadness, anxiety*, and *anger*
- identify commentators characteristics such as *tone, analytical thinking, clout*, and *authenticity*

In LIWC, each score is measured from 0 to 100 based on the percentage of all words in the document (the total number of comments on each post-type). The *analytical thinking* measures how formal, logical, and hierarchical the writing is, and it is important in identifying how well-educated a commentator is [30]. *Clout* index measures the level of expertise and confidence in commentator writing [31]. *Authenticity* measures how authentic and honest the writing is; higher scores suggest honest writing and lower scores suggest deceptive writing [32].

III. RESULTS AND FINDINGS

A. Posts Annotations

Three expert coders annotated candidates posts into one of three categories: Policy, Character, Mixed/Other. Next, we

TABLE II
THREE-WAY POST-TYPE CLASSIFICATION

Classifier	Class Label	P	R	F-measure
CNN	Policy	0.68	0.68	0.68
	Character	0.64	0.65	0.64
	Mixed/Other	0.23	0.22	0.22
	Average	**0.52**	**0.52**	**0.52**
MLP	Policy	0.55	0.57	0.56
	Character	0.45	0.57	0.50
	Mixed/Other	0.21	0.04	0.07
	Average	0.40	0.39	0.39
Multinomial NB	Policy	0.51	0.61	0.56
	Character	0.52	0.61	0.56
	Mixed/Other	0.47	0.05	0.09
	Average	0.50	0.42	0.46

applied Fleiss' Kappa measure to measure the reliability of annotations among multiple coders [33] and the Kappa factor is 0.80 which can be interpreted as "Substantial Agreement" according to [34]. The total number of posts by all three candidates is 4,825 distributed among the three categories as follows: 39.7% *Character* posts, 43.8% *Policy* posts, and 16.5% *Mixed/Other* posts.

B. Post-Type Classification Results

To detect post type, we experimented different classifiers as well as features and performed both three-way classification with three categories: *Policy, Character, Mixed/Other* and two-way classification with two categories: *Policy, Character*. To evaluate the goodness of classifiers we utilized *precision (P), recall (R)* and *F-measure* as evaluation metrics and we performed ten-fold cross-validation. We found *Word2Vec* features outperformed *TF-IDF* features across all classifiers.

Table II presents the accuracies of top performing classifiers for identifying post-type using three-way classification models CNN, MLP, and Multinomial Naive Bayes. In this table, CNN classifier scored the highest average F-measure score of 52% which is 56% higher than random guess of 33.3%. In this table we observe that most classifiers did not perform well due to ambiguous and noisy posts in *Mixed/Other* category. Thus, we experimented two-way classification using only two categories: Policy, Character.

Table III illustrates the accuracies for identifying post-type using two-way classification {*Policy, Character*} using models: CNN, MLP, and Linear-SVR. We observe that classifiers performed well with CNN scoring the highest average F-measure accuracy of 78% for detecting {*Policy, Character*} posts and it is comparable to annotators scores of 80%.

C. Analysis of Commentators

Table IV shows the emotions and reactions of commentators on each candidate and post-type. On *Policy* posts, Trump's commentators scored highest scores across all indicies except *Sadness* index where Sanders commentators scored higher. On *Character* posts, Trump's commentators expressed more *Negative Emotion, Sadness, Anger,* and *Anxiety* compared to other candidates. Regardless of post type, we observed that Trump's commentators notably scored high on *Anger* index, nearly double for *Character* posts.

Table V illustrates the results of commentators characteristics. On *Policy* posts, Trump's commentators scored higher

TABLE III
TWO-WAY POST-TYPE CLASSIFICATION

Classifier	Class Label	P	R	F-measure
CNN	Policy	0.79	0.80	0.80
	Character	0.77	0.77	0.77
	Average	**0.78**	**0.78**	**0.78**
MLP	Policy	0.70	0.42	0.52
	Character	0.58	0.77	0.66
	Average	0.64	0.60	0.62
Linear-SVR	Policy	0.56	0.34	0.42
	Character	0.50	0.69	0.58
	Average	0.53	0.51	0.52

TABLE IV
COMMENTATORS EMOTIONS PER CANDIDATE AND POST TYPE

Post Type	Index	Trump	Sanders	Biden
Policy	Pos Emotion	**4.01**	3.67	3.32
	Neg Emotion	**2.97**	2.44	2.46
	Sadness	0.40	**0.41**	0.40
	Anger	**1.47**	1.01	1.00
	Anxiety	**0.34**	0.23	0.24
Character	Pos Emotion	3.63	**3.75**	3.31
	Neg Emotion	**2.99**	2.30	2.61
	Sadness	**0.44**	0.37	0.40
	Anger	**1.73**	0.90	0.98
	Anxiety	**0.35**	0.32	0.29

TABLE V
COMMENTATORS CHARACTERISTICS PER CANDIDATE AND POST TYPE

Post Type	Index	Trump	Sanders	Biden
Policy	Tone	46.77	**48.06**	41.96
	Analytical Thinking	**66.05**	62.90	64.57
	Clout	**83.03**	75.72	79.22
	Authenticity	9.28	**15.26**	10.85
Character	Tone	36.56	**47.35**	38.96
	Analytical Thinking	63.51	**68.12**	65.46
	Clout	**83.09**	72.78	81.69
	Authenticity	8.09	**19.80**	8.12

scores on *Analytical Thinking,* and *Clout* indicies, which suggests formal and logical writing with more confidence and expertise. On *Authenticity* index, for both post types Trump's commentators scored lowest scores which suggests deceptive writing. For both post types, Sanders commentators scored highest scores on *Authenticity* index which suggest honest writing. Moreover, on *Character* posts Sanders commentators scored highest scores on nearly all indicies except *Clout* where Trump's commentators are higher.

IV. Conclusion and Future Work

This paper presents a new way in analyzing political discourse functions and interactions on SNS. Based on machine learning techniques, the proposed models presented promising accuracies in detecting $\{$*Policy, Character*$\}$ posts with average F-measure of 78% using Convolutional Neural Network (CNN) and it is comparable to human accuracy of 80%. Additionally, we also analyzed commentators comments and presented the results of various linguistic and psychological indices and how they differ across candidates as well as post type. In our future work, we intend to extend this study by analyzing text from other sources such as news and polls. We also intend to study the evolution of US political campaigns over decades from the functional theory point of view.

References

[1] D. L. Cogburn and F. K. Espinoza-Vasquez, "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*, vol. 10, no. 1-2, pp. 189–213, 2011.

[2] L. Rainie, A. Smith, K. L. Schlozman, H. Brady, and S. Verba, "Social media and political engagement," *Pew Internet & American Life Project*, vol. 19, pp. 2–13, 2012.

[3] H. Gil de Zúñiga, N. Jung, and S. Valenzuela, "Social media use for news and individuals' social capital, civic engagement and political participation," *Journal of computer-mediated communication*, vol. 17, no. 3, pp. 319–336, 2012.

[4] M. A. Xenos, A. Vromen, and B. D. Loader, "The great equalizer? patterns of social media use and youth political engagement in three advanced democracies," in *The Networked Young Citizen*. Routledge, 2014, pp. 33–54.

[5] S. Boulianne, "Social media use and participation: A meta-analysis of current research," *Information, communication & society*, vol. 18, no. 5, pp. 524–538, 2015.

[6] R. Effing, J. Van Hillegersberg, and T. Huibers, "Social media and political participation: are facebook, twitter and youtube democratizing our political systems?" in *International conference on electronic participation*. Springer, 2011, pp. 25–35.

[7] "Among millennials engaged in primaries, dems more likely to learn about the election from social media," https://www.pewresearch.org/fact-tank/2016/02/09/among-millennials-engaged-in-primaries-dems-more-likely-to-learn-about-the-election-from-social-media/, 2016, accessed: 2020-04-08.

[8] "Social media fact sheet," https://www.pewresearch.org/internet/fact-sheet/social-media/, 2019, accessed: 2020-04-08.

[9] C. B. Williams and G. J. Gulati, "Facebook grows up: An empirical assessment of its role in the 2008 congressional elections," *Proceedings from Midwest Political Science Association, Chicago*, vol. 32, p. 53, 2009.

[10] C. Williams and G. Gulati, "What is a social network worth? facebook and vote share in the 2008 presidential primaries." American Political Science Association, 2008.

[11] G. S. Enli and E. Skogerbø, "Personalized campaigns in party-centred politics: Twitter and facebook as arenas for political communication," *Information, communication & society*, vol. 16, no. 5, pp. 757–774, 2013.

[12] K. N. Andersen and R. Medaglia, "The use of facebook in national election campaigns: politics as usual?" in *International Conference on Electronic Participation*. Springer, 2009, pp. 101–111.

[13] S. Hong and D. Nadler, "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*, vol. 29, no. 4, pp. 455–461, 2012.

[14] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting elections with twitter: What 140 characters reveal about political sentiment," in *Fourth international AAAI conference on weblogs and social media*, 2010.

[15] W. L. Benoit and A. Harthcock, "Functions of the great debates: Acclaims, attacks, and defenses in the 1960 presidential debates," 1999.

[16] S. Cho and W. Benoit, "Primary presidential election campaign messages in 2004: A functional analysis of candidates' news releases," *Public Relations Review*, vol. 31, no. 2, pp. 175–183, 2005.

[17] ——, "2004 presidential campaign messages: A functional analysis of press releases from president bush and senator kerry," *Public Relations Review*, vol. 32, no. 1, pp. 47–52, 2006.

[18] W. L. Benoit, *Political election debates: Informing voters about policy and character*. Lexington Books, 2013.

[19] B. J. Chambers and S. L. Bichard, "Public opinion on youtube: A functional theory analysis of the frames employed in user comments following sarah palin's 2008 acceptance speech," *International Journal of E-Politics (IJEP)*, vol. 3, no. 2, pp. 1–15, 2012.

[20] J. M. Wilson-Kratzer and W. L. Benoit, "A functional analysis of press releases from senator barack obama and senator john mccain during the 2008 primary presidential election," *Public Relations Review*, vol. 36, no. 2, pp. 178–180, 2010.

[21] E. Loper and S. Bird, "Nltk: The natural language toolkit," *CoRR*, vol. cs.CL/0205028, 2002.

[22] R. Řehůřek and P. Sojka, "Software Framework for Topic Modelling with Large Corpora," in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. Valletta, Malta: ELRA, May 2010, pp. 45–50, http://is.muni.cz/publication/884893/en.

[23] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013, pp. 3111–3119.

[24] H. P. Luhn, "A statistical approach to mechanized encoding and searching of literary information," *IBM Journal of research and development*, vol. 1, no. 4, pp. 309–317, 1957.

[25] K. S. Jones, "A statistical interpretation of term specificity and its application in retrieval," *Journal of documentation*, 1972.

[26] G. Amati and C. J. Van Rijsbergen, "Probabilistic models of information retrieval based on measuring the divergence from randomness," *ACM Transactions on Information Systems (TOIS)*, vol. 20, no. 4, pp. 357–389, 2002.

[27] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[28] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[29] Y. R. Tausczik and J. W. Pennebaker, "The psychological meaning of words: Liwc and computerized text analysis methods," *Journal of language and social psychology*, vol. 29, no. 1, pp. 24–54, 2010.

[30] J. W. Pennebaker, C. K. Chung, J. Frazee, G. M. Lavergne, and D. I. Beaver, "When small words foretell academic success: The case of college admissions essays," *PloS one*, vol. 9, no. 12, p. e115844, 2014.

[31] E. Kacewicz, J. W. Pennebaker, M. Davis, M. Jeon, and A. C. Graesser, "Pronoun use reflects standings in social hierarchies," *Journal of Language and Social Psychology*, vol. 33, no. 2, pp. 125–143, 2014.

[32] M. L. Newman, J. W. Pennebaker, D. S. Berry, and J. M. Richards, "Lying words: Predicting deception from linguistic styles," *Personality and social psychology bulletin*, vol. 29, no. 5, pp. 665–675, 2003.

[33] J. L. Fleiss, "Measuring nominal scale agreement among many raters." *Psychological bulletin*, vol. 76, no. 5, p. 378, 1971.

[34] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *biometrics*, pp. 159–174, 1977.