

**FROM POSTS TO VOTES: USING XLM-RoBERTa (XLM-R) AND
AGENT-BASED MODELING (ABM) TO ANALYZE AND PREDICT
ELECTION WINS IN THE PHILIPPINES THROUGH TWITTER/X**

A Thesis

Presented to the

Department of Information Systems

and Computer Science

Ateneo de Manila University

In Partial Fulfillment

of the Requirements for the Degree of

Bachelor of Science in Computer Science

by

Michael Gavin N. Del Castillo

Mary June Aubrey C. San Jose

Shaira Jane L. Sy

2026

ABSTRACT

Recent election periods have seen the rise of social media platforms as tools through which political ideologies are spread, and support for candidates are garnered. The most recent and popular example of which was the landslide victory of Rodrigo Roa Duterte during the 2016 Philippine presidential election, which many attributed to his social media campaigns. The same cannot be said for when former vice president Maria Leonor “Leni” Robredo, despite widespread support in online spaces, lost the 2022 Philippine presidential election to Ferdinand “Bongbong” Marcos Jr.

The main motivation for this research is to determine the effectiveness of social media to predict electoral wins. Through the Robustly Optimized BERT Pre-trained Approach (RoBERTA), its multilingual variant (XLM-RoBERTa), and Agent-based modeling (ABM), this paper aims to perform sentiment analysis on posts from X (formerly Twitter), a popular social media platform when it comes to public discussions on politics in the Philippines, to draw insights on how citizens perceived the winning and first runner-up electoral candidates during pre- and proper election periods and see if it can be used as an indicator of an election win to serve as a tool for predicting future elections.

TABLE OF CONTENTS

ABSTRACT	ii
LIST OF FIGURES	iv
LIST OF TABLES	v
CHAPTER	
I INTRODUCTION	1
1.1 Context of Study	1
1.2 Research Questions	3
1.3 Research Objectives	4
1.4 Scope and Limitations of the Study	5
1.5 Significance of the Study	6
BIBLIOGRAPHY	8

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

1.1 Context of Study

In recent years, social media has had a major role as a platform through which political ideologies are spread and political discussions occur. This is especially apparent when observing the flow of recent elections in certain countries. During election seasons in the Philippines, renewed posts circulate social media characterising voters as ‘ill-informed’ and not understanding how democracy works for selling their votes to corrupt politicians [2]. However, many may not realize that even the mere act of engaging with social media campaigns, even to criticize it, may actually be supporting the candidate.

When it comes to social media campaign regulations, the Philippines, along with Malaysia and Indonesia, are generally a *laissez faire*— not subjected to serious scrutiny or obligated to be transparent about campaign spending [5]. The 2016 Philippine presidential election is widely considered the first “social media election” for the Philippines, mainly due to how its winner, Rodrigo Roa Duterte, was able to utilize social media to establish a controversial image which mobilized his wide follower-base to rally in support of him, both online and offline [4].

Tapsell's [2020] analysis brings attention to an interview with Nic Gabunada, Duterte's 'campaign friend', where he says Facebook became vital to their campaign after realizing that 45% of Filipinos are on Facebook, via mobile phone, and that their goal for the campaign was to maximize awareness.

Cases, however, have occurred in which online support did not directly translate to election wins. The 2022 Philippine presidential elections saw the Angat Buhay campaign of former Vice President Leni Robredo garner the attention of young audiences on social media. Rallies in support of Robredo alongside Robredo's track record as a politician made her a popular choice for millions as a capable presidential candidate [3]. Despite massive online support, Robredo lost the elections, garnering some 14.8 million votes (as opposed to Ferdinand "Bongbong" Marcos, Jr, who gathered 31.1 million) [1]. Given the possible disparity between social media popularity and election votes, the aim of this research is to provide a data-driven analysis on the effectiveness of social media as an indicator of election wins by observing social media trends at the time of both 2022 and 2024 elections, as well as comparing and contrasting these elections in terms of said trends.

Through Natural Language Processing (NLP) Algorithms, this paper aims to analyze the conversations on the aforementioned presidential candidates that had transpired in online spaces during pre-election seasons, namely X (formerly

Twitter). Previous research endeavors have already shown the effectiveness of sentiment analysis in determining key themes behind social media posts, especially in the context of events such as elections. Thus, this research would like to push this idea further by not only contextualizing the data within a single setting. Rather, this paper aims to compare and contrast different election periods of the Philippines to create an Agent-based model (ABM) to, given the proper datasets, simulate and serve as a tool to predict future (Philippine) elections.

This paper intends to determine whether or not social media support directly translates to election success, or if other factors were present which had contributed to the losses of Robredo and worked to raise Duterte and Marcos Jr's popularity.

1.2 Research Questions

1. Can Agent-Based Modelling effectively predict election results by way of simulating interactions between users on social media that are centered around electoral candidates?
 - (a) The XLM-RoBERTA model can be used to perform sentiment analysis on social media posts in Filipino and English. Can XLM-RoBERTA extract the necessary sentiments needed to define a social media user

as an agent?

- (b) By using Agent-Based Modelling on social media data on the 2019 Philippine senatorial elections and 2022 Philippine presidential elections, can this model simulate the recent 2025 Philippine senatorial elections and accurately determine the winners of that election?
- (c) Can the Agent-Based Model not only capture an accurate simulation of previous presidential elections, but also predict future elections as well, e.g. 2028 Philippine presidential elections?

1.3 Research Objectives

1. Agent-Based Modelling can be used for creating simulations of various social phenomena. For this paper, the researchers intend to test its ability to simulate not only candidate-user interactions but also user-user interactions in social media during Philippine election seasons.
 - (a) Through XLM-R, perform sentiment analysis on social media posts about both the 2019 and 2022 Philippine elections and create a list of parameters with which Agents and interactions in the simulation are defined.
 - (b) Through our baseline model, perform an Agent-Based Simulation to

evaluate the efficacy of the model to predict winning candidates of the 2025 Philippine senatorial elections, using information diffusion to represent the amount of social interactions done in the simulation with respect to a given candidate.

- (c) After verifying the accuracy of the Agent-Based model, perform a simulation to predict potential winning candidates of the upcoming 2028 Philippine presidential elections, given a hypothetical candidate's on-line presence and sentiments by users surrounding them.

1.4 Scope and Limitations of the Study

The scope of the study is to create a virtual environment that simulates the Philippine social media during an election season, with both users and candidates as agents. Thus, concerns of fake news content within the environment can be a topic of discussion in drawing insights from the results, but will not be thoroughly tackled in the course of this study's experiments and results.

Data collection will be limited to three Philippine election seasons for training and evaluation of the model; namely, the 2019 and 2022 elections for training purposes, and the 2025 elections for validation. Scraping of social media posts will be restricted within X (formerly Twitter) due to not only its ease of data scraping through external APIs, but also the short character limit of 280 for

each post, which makes performing sentiment analysis more manageable than on other platforms. The ranges of dates of tweets collected depend on the election campaign period of each Philippine election season.

As for the number of candidates, all presidential candidates for the 2022 elections will be the main agents; for senatorial elections, on the other hand, only tweets about the top 20 candidates will be collected. This is so that the model is still able to capture the contribution of those outside the winning 12 candidates to online public discourse on which candidates to vote for.

1.5 Significance of the Study

This paper aims to contribute to the field of social computing by analyzing the behavior of users and political candidates in online spaces during election periods, especially in the Philippine context. Using sentiment analysis and other NLP techniques, the research aims to gather statistically large amounts of social media data to represent the interaction between candidates and the general public, and between users as well, who might be potential voters. Then it utilizes models that process such data and draw insights from the sentiments expressed in social media posts by candidates and their audiences, particularly during presidential and midterm elections.

The analytics and findings of this study will benefit the following: Philip-

pine political campaign teams in strategizing social media posts to build towards this particular sentiment to gather shifts in voting preference, the general public in understanding the information diffusion of Philippine social media during elections, and political analysts in grasping a wider view of social media networks during an election season.

BIBLIOGRAPHY

- [1] ABS-CBN. Halalan results 2022. <https://halalanresults.abs-cbn.com/>, 2022.
- [2] CALIMBAHIN, C. A. A. *Current Electoral Processes in Southeast Asia. Regional Learnings*. Research Institute on Contemporary Southeast Asia, 2023, ch. The 2022 Philippine Elections: Unity, Continuity, and Impunity, pp. 11–18.
- [3] JOHNSON, H., AND HEAD, J. Leni robredo: The woman leading the philippines' 'pink revolution'. *BBC* (7 May 2022).
- [4] SINPENG, A., GUEORGUIEV, D., AND ARUGAY, A. A. Strong fans, weak campaigns: Social media and duterte in the 2016 philippine election. *Journal of East Asian Studies* 20, 3 (2020), 353–374.
- [5] TAPSELL, R. Social media and elections in southeast asia: The emergence of subversive, underground campaigning. *Asian Studies Review* 45 (09 Nov 2020), 117–134.