How do Politicians Garner Engagement on Social Media?

Julia Chock
University of Minnesota Duluth
chock023@d.umn.edu

Cameron Dahl
University of Minnesota Duluth
dahl1360@d.umn.edu

Abstract

We analyze 36,642 tweets from seven Minnesota U.S. Congress members (2019-2025) to identify audience engagement drivers, measured as replies, reposts, and likes. We tagged each tweet for calls to vote, policy discussion, and attack language, and measured sentiment, tweet length, hashtag counts, mention counts, and posting hour. Ordinary least-squares regression explained 15.2 % of variance in log engagement. The encouragement to vote increased participation significantly, while policy-focused posts reduced it by 31 %. A negative-binomial model confirmed these effects and showed that each additional hashtag, mention, unit increase in tweet length, and later posting hour significantly depressed engagement. Mixed-effects modeling showed substantial baseline differences between politicians, and positive sentiment was associated with a slight decrease in engagement.

1 Introduction

Social media platforms have transformed how elected officials communicate with constituents, enabling direct, accessible engagement that can shape public opinion and mobilize voter behavior [2, 4]. On Twitter (now X), politicians utilize a variety of strategies, from policy announcements, calls to vote, and critical messaging of opponents, yet the relative effectiveness of these strategies in driving audience interaction remains underexplored. Prior work has shown that more emotional and user engaging framing does better than neutral informational posts in garnering likes, retweets, and replies [1, 3, 5].

In this paper, we examine 36,642 tweets from seven Minnesota U.S. House and Senate members spanning January 2019 through April 2025. We classify each tweet by content (vote encouragement, policy discussion, attack language), quantify sentiment, tweet length, hashtag and mention counts, and posting hour. Using ordinary least-squares, negative-binomial, and mixed-effects models, we assess how these factors, independently and in combination, influence engagement metrics. Our findings offer strategic guidance for political communicators and contribute to a deeper understanding of how to civically engage and online audience.

2 Related Works

When reviewing related works for this study, we found a consistent conclusion with tweets featuring political content. Polarizing content is associated with a higher level of engagement. Rafail et al. (2024) found that all types of polarization increase engagement based on a sample of 134,442 tweets posted by 527 members of Congress during the 2022 midterm elections. Prior research defines polarizing rhetoric as content that is dividing, promotes political tribalism, or uses negative or insulting language (Rafail et al., 2024). In addition, the authors discuss the potential toxic feedback loop that can happen as politicians are rewarded with engagement and visibility to publish more polarizing tweets. The more frequently politicians publish more polarized tweets, the more engagement they receive from the public, thus encouraging them to become more polarized. Similarly, both Lazarus and Thornton (2020) and Ballard et al. (2022) have found that tweets containing negative sentiment are associated with better engagement.

In other contexts, such as the COVID-19 pandemic, McKay et al. (2022) found that tweets written with compassion and empathy were also associated with higher tweet engagement. This suggests that the relationship between sentiment and engagement may depend heavily on the context surrounding the tweet, such as an election versus a global pandemic.

Building on this prior work, we investigate how various features of tweets, including sentiment, policy discussions, and attack language, affect their level of engagement. While prior works have focused primarily on a national level, this study is aimed specifically at politicians within the State of Minnesota, and seeks to examine a significant period of time to look for temporal changes.

3 Methodology

3.1 Data Collection

Tweets were collected in JSON format using a third-party service called Octoparse from 2019-2025 in chronological order. The format of the JSON schema is shown in Figure 1. Retweets were not included in the analysis. The politicians included in this study are Amy Klobuchar, Angie Craig, Tina Smith, Dean Phillips, Betty McCollum, Ilhan Omar, and Tom Emmer. If the politician had multiple Twitter Accounts, their official campaign account was used.

After collecting tweets from all seven politicians, we wrote a Python script to perform sentiment analysis. Before analyzing the tweets, we preprocessed each tweet to ensure that the model was given the most accurate data. Several preprocessing steps were performed, including:

(1) mentions prefixed by the @ symbol were replaced with @USER (ex. @X was changed to @USER)
(2) converting emojis into text for better comprehension (ex. `:happy_face:`)
(3) break apart contractions (can't becomes can not)
(4) a.m/p.m expressions were fixed
(5) URLs were replaced with @URL as it is unnecessary for the analysis, and
(6) non-standard characters were removed.

3.2 Sentiment Analysis

Sentiment analysis was performed using the

`finiteautomata/bertweet-base-sentiment-analysis`

pre-trained model via Hugging Face, a BERTweet model trained specifically on English tweets designed to account for the abnormalities of language and platform-specific features present in tweets.

```
{
    "PostID": "1549823632288976903",
    "Time": "2022-07-20 18:28:42+00:00",
    "TweetURL": "https://x.com/Ilhan/status
        /1549823632288976903",
    "Text": "Honored to preside over the
        House floor today. \n\nGrateful to
        represent the 5th District in
        Congress every day! https://t.co/
        EHhK6UJO3A",
    "Username": "Ilhan",
    "Replies": "174",
    "Reposts": "167",
    "Likes": "1968",
    "Views": "0",
    "UserID": "1082334352711790593",
    "UserUrl": "https://x.com/Ilhan",
    "Location": "Minneapolis, MN and
        Washington, DC"
},
```


Figure 1: Example of a tweet authored by Illhan Omar (D-MN) used for analysis.

The model classifies tweets into three categories of negative, positive, and neutral. Tweets were preprocessed as described in Section 3.1 to align with BERTweet's specifications. The format of the output is shown in Figure 2.


```
{
    "politician": "Angie Craig (D-MN)",
    "text": "The United States has a duty to
        stand with our allies who share our
        values of freedom &amp; democracy. We
         must continue to support Ukraine in
        its fight for freedom.\n\u00a0\nLet\
        u2019s bring the Ukraine funding
        request to a vote.\n\u00a0\nhttps://t
        .co/PgF4rYTW7I",
    "predicted_sentiment": "Positive",
    "confidence_scores": {
        "negative": 0.0025500375777482986,
        "neutral": 0.25694161653518677,
        "positive": 0.7405083775520325
    },
    "date_published": "2023-10-25
        20:07:22+00:00"
}
```


Figure 2: Example of the output from the model.

In addition to the predicted sentiment, the output contains all confidence scores to verify results.

3.3 Data Analysis

All data processing and statistical modeling were implemented in Python using the `pandas`, `numpy`, `matplotlib` and `statsmodels` libraries. After loading and merging the raw tweet JSON (Figure 1) with the BERTweet sentiment output (Figure 2), we performed the following steps:

(1) **Engagement calculation.** We defined total engagement as the sum of replies, reposts, and likes.
(2) **Timestamp features.** The Time string was parsed into a `datetime` object. We extracted

$$year\_month_i = \text{the first day of the month of tweet } i,$$
$$hour\_of\_day_i = \text{the hour component (0–23) of the posting time.}$$

(3) **Content flags and categorization.** We applied case-insensitive regular expressions to flag each tweet for $\{policy,\ attack,\ encourage\ vote\}$. Each tweet was then assigned to one of eight mutually exclusive categories, belonging to a single group or an intersection of the tags, based on which flags matched.
(4) **Additional text-based features.** We computed tweet length, number of hastags, and number of mentions.
(5) **Continuous sentiment score.** From the BERTweet output, we defined

$$sentiment\_score_i = positive\_conf_i - negative\_conf_i,$$

yielding a continuous measure in $[-1, +1]$.

Because raw engagement was highly biased to the right (mean 1,589, median = 115), we applied a log transformation:

$$y_i = \log(engagement_i + 1),$$

resulting in an approximately normal distribution (Figure 3).

We then specified three complementary models:

1. Ordinary Least Squares (OLS). We regressed y_i on our predictors:

$$y_i = \beta_0 + \beta_1\, sentiment\_score_i + \sum_c \beta_c\, \mathbf{1}\{content\_cat_i = c\} + \beta_L\, tweet\_length_i + \beta_H\, n$$

where $\varepsilon_i \sim N(0, \sigma^2)$. We report coefficients β, robust standard errors, R^2, and 95 % confidence intervals.

2. Negative-Binomial GLM. For raw counts we used

$$\mathbb{E}\big[engagement_i\big] = \exp\big(\beta_0 + \beta_1\, sentiment\_score_i + \cdots + \beta_t\, hour\_of\_day_i\big),$$

with variance function $\mathrm{Var}(y_i) = \mu_i + \alpha\, \mu_i^2$. Exponentiated coefficients $\exp(\beta)$ yield incident-rate ratios (IRRs).

3. Mixed-Effects Linear Model. To account for the differences in popularity across the seven politicians, we added a random intercept $u_{j[i]}$ for the author of tweet i:

$$y_i = \beta_0 + \beta_1\, sentiment\_score_i + \cdots + \beta_t\, hour\_of\_day_i + u_{j[i]} + \varepsilon_i, \quad u_j \sim N(0, \tau^2),$$

These models confirmed improved normality after log transforming, strong over dispersion justifying the negative-binomial, and significant improvement in fit when including random politician intercepts (χ^2-tests, $p < .001$). All tests were two-tailed at $\alpha = 0.05$. Estimated coefficients, IRRs, and random intercepts are displayed in Figures 6-8.

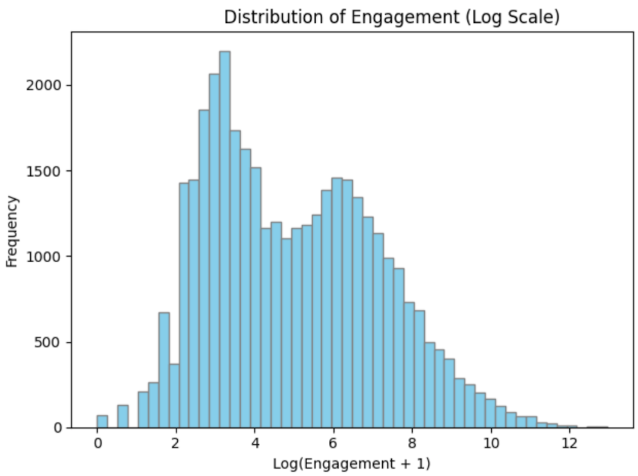

Figure 3: Distribution of log transformed engagement counts.

4 Results

Once raw engagement counts were log transformed, we produced a distribution more closely approximating normality (Figure 3).

Monthly aggregates (Figure 4) showed two contrasting trends. First, average sentiment scores remained stable (mean = 0.3337, SD = ±0.05). Second, average monthly engagement rose from approximately 5,000 in early 2019 to over 80,000 by April 2025.

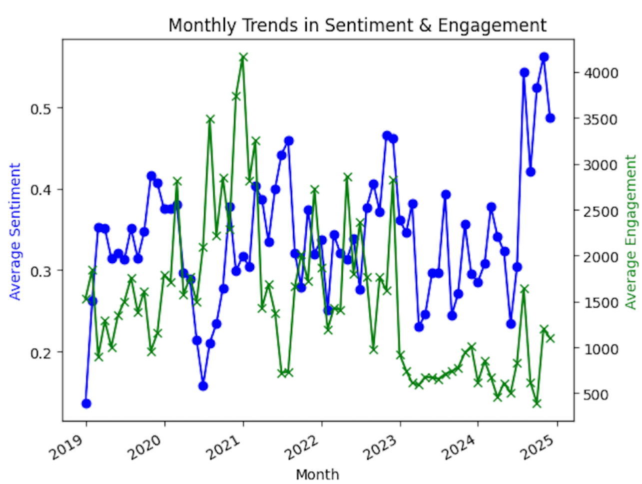

Figure 4: Monthly trends in average sentiment and engagement.

When examining tweet categories, those explicitly urging voting garnered the highest median log-engagement (attack+vote: 6.8; vote only: 6.2), while attack-only tweets had a median of 5.0. Policy-only tweets were weakest (4.0), and mixed frames (policy+attack 4.1; policy+vote 4.4) outperformed pure policy but remained below pure voting frames (Figure 5).

We applied ordinary least squares (OLS) regression to model $\log(\text{engagement} + 1)$ as a function of sentiment, content category,

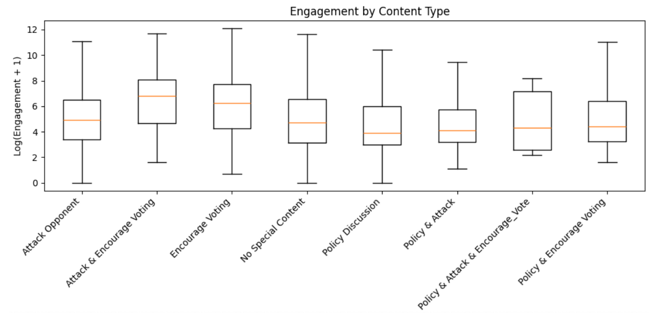

Figure 5: Median log engagement by tweet category.

tweet length, hashtags, mentions, and posting hour, explaining $R^2 = 0.152$ ($p < 0.001$). Key coefficients (95% CI):

- **Sentiment**: $\beta = -0.474$ (SE = 0.016, $p < 0.001$), indicating a one-unit increase in positivity reduces log-engagement by 0.474 (38% fewer interactions).
- **Encourage Vote**: $\beta = 1.014$ (SE = 0.056, $p < 0.001$; IRR $= e^{1.014} = 2.76$).
- **Policy only**: $\beta = -0.371$ (SE = 0.051, $p < 0.001$; IRR = 0.69).
- **Hashtags**: $\beta = -0.821$ (SE = 0.020, $p < 0.001$), each extra tag cuts engagement by 56% (IRR = 0.44).
- **Mentions**: $\beta = -0.213$ (SE = 0.009, $p < 0.001$), each mention cuts engagement by 19.2% (IRR = 0.81).
- **Posting hour**: $\beta = -0.048$ (SE = 0.002, $p < 0.001$), each hour later reduces engagement by 4.7% (IRR = 0.953).
- **Tweet length**: $\beta = -0.0029$ (SE 0.000, $p < 0.001$), i.e., 0.29% fewer interactions per additional character (IRR per char 0.997).

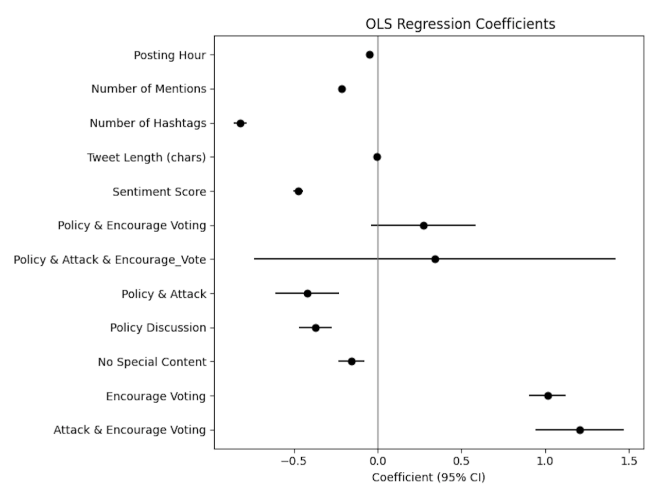

Figure 6: OLS regression coefficients (95% confidence intervals).

A negative-binomial generalized linear model on raw counts achieved pseudo-$R^2 = 0.526$. Voting tweets had IRR = 2.51 (+151%, $p < 0.001$), while policy tweets had IRR = 0.93 (−7%, $p < 0.01$). Hashtags (IRR = 0.44), mentions (IRR = 0.83), posting hour (IRR =

0.96), and sentiment (IRR = 0.58 per unit) all significantly depressed engagement (Figure 7).

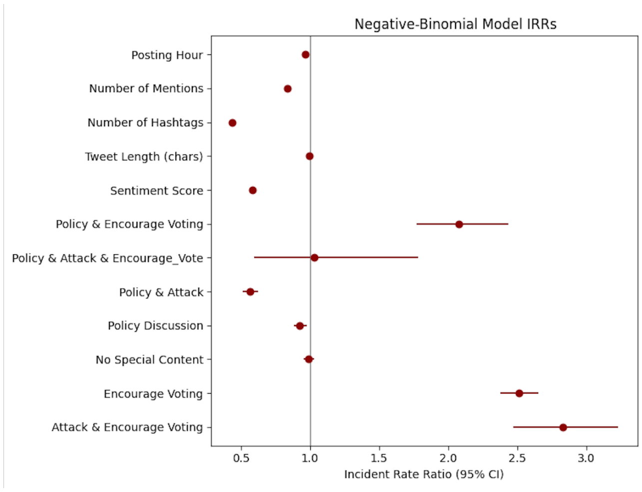

Figure 7: Negative-binomial model incident rate ratios (95% CIs).

To account for popularity differences, we also used a mixed-effects model with random intercepts (between-politician variance $\tau^2 = 2.048$). Fixed effects still retained strong vote-appeal boosts (attack+vote: $\beta = 0.554$, $p < 0.001$; vote only: $\beta = 0.566$, $p < 0.001$) and a sentiment penalty (= −0.445, $p < 0.001$), while policy-only was non-significant (= 0.008, $p = 0.793$). Tweet length became non-significant.

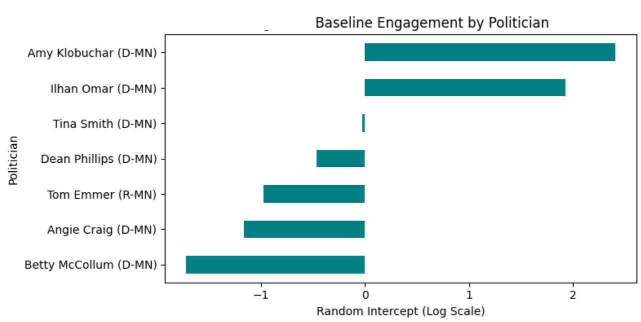

Figure 8: Mixed-effects random intercepts by politician (log scale) and fixed-effect coefficients (95% CIs).

5 Discussion

Our results yield four clear recommendations for increasing engagement on social media.

First, explicit calls to vote, especially when combined with critical language, drive engagement by more than 2× compared to neutral or informational posts. Considering appeals to user participation in messaging and considering pairing them with targeted critiques of opponents can maximize reach during election periods. Exclusively discussing information or policy fails to connect with audiences.

Second, we find a strong negativity bias. Each one-point increase in sentiment predicts a 38 % drop in engagement. Neutral or slightly critical wording outperforms purely positive tones. Those aiming to spark discussion should therefore try to balance optimistic content with measured critiques rather than rely exclusively on upbeat messaging to draw in users.

Third, structural choices have large impacts. Every additional hashtag reduces engagement by 56 %, each mention by 17 %, and delaying a post by one hour cuts engagement by 4 %. To improve visibility, limiting hashtags and mentions to the most essential tags (e.g. one or two per tweet) and scheduling posts during known peak-traffic windows, typically midday and early evening, will be most successful.

Fourth, baseline popularity varies up to 36× across politicians. Any analysis of content effectiveness must control for a user's follower base and historical engagement levels, mixed-effects models are an effective tool for separating content effects from celebrity effects.

Future research should test these findings on other platforms and could include dynamic follower growth data, sponsored post effects, and examine how algorithmic ranking changes these effects.

6 Limitations

While sentiment analysis was conducted on preprocessed tweets, feature tagging utilized the raw text, which may introduce minor inconsistencies.

Our text-based tagging using regular expressions may misclassify tweets that employ sarcasm, idioms, slang, etc. and cannot capture nuance used to convey tone. If more time and resources allowed, manually categorizing tweets may provide more insightful results. Similarly, sentiment models struggle with sarcasm, double negatives, and regional language variations, and may not recognize cultural references that arose after their last training update, which in this case was 2021. These factors could add noise to sentiment scores, affecting the true effect sizes.

Engagement was measured as the sum of replies, reposts, and likes, omitting quote-tweets, impressions, link clicks, and views. These other metrics may also provide insight into how politicians reach their audiences, and fill the gaps in this paper. Future work should integrate more extensive engagement metrics to further conclusions.

Their was also other causality that may not have been considered like current events or sponsored posts, which may drive both content choices and engagement. Although mixed-effects models control for differences in baseline follower counts, they do not account for changes in follower counts or promotional strategies over time.

Finally, our sample is limited to Minnesota's federal legislators on Twitter (X), and results may not generalize to other regions or social-media platforms with different audience demographics or algorithmic ranking mechanisms. Platform changes have been apparent for Twitter during 2019–2025, which may also influence engagement trends independently of tweet features.

Data and Code Availability

All scripts, raw results, and processed data for this analysis are hosted at https://github.umn.edu/DAHL1360/SentimentAnalysis.git.

References

[1] Andrew O. Ballard, Ryan DeTamble, Spencer Dorsey, Michael Heseltine, and Marcus Johnson. 2023. Dynamics of Polarizing Rhetoric in Congressional Tweets. *Legislative Studies Quarterly* 48, 1 (2023), 105–144. doi:10.1111/lsq.12374 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/lsq.12374

[2] Jeffrey Lazarus and Judd R. Thornton. 2021. Bully Pulpit? Twitter Users' Engagement With President Trump's Tweets. *Social Science Computer Review* 39, 5 (2021), 961–980. doi:10.1177/0894439320946578 arXiv:https://doi.org/10.1177/0894439320946578

[3] Melissa MacKay, Andrea Cimino, Samira Yousefinaghani, Jennifer E. McWhirter, Rozita Dara, and Andrew Papadopoulos. 2022. Canadian COVID-19 Crisis Communication on Twitter: Mixed Methods Research Examining Tweets from Government, Politicians, and Public Health for Crisis Communication Guiding Principles and Tweet Engagement. *International journal of environmental research and public health* 19, 11 (2022), 6954–.

[4] Juan Manuel Pérez, Juan Carlos Giudici, and Franco Luque. 2021. pysentimiento: A Python Toolkit for Sentiment Analysis and SocialNLP tasks. arXiv:2106.09462 [cs.CL]

[5] Patrick Rafail, Whitney E. O'Connell, and Emma Sager. 2024. Polarizing Feedback Loops on Twitter: Congressional Tweets during the 2022 Midterm Elections. *Socius* 10 (2024), 23780231241228924. doi:10.1177/23780231241228924 arXiv:https://doi.org/10.1177/23780231241228924