Understanding Public Opinion through Social Media Posts: Summarization, Stance Annotation, and Demographic Inference

Anonymous authors
Paper under double-blind review

Abstract

Social media posts are a promising source of data for insight into the opinions held by members of the public (or at least users of a social media platform), since they can provide near real-time and lower-cost insights than more traditional methods like surveys and focus groups. Additionally, social media data may reveal the opinions of those who would not necessarily agree to participate in surveys or focus groups. However, there are challenges to using social media data for insights into public opinion: (a) the sheer volume far exceeds what a person can read and digest, and (b) they don't include demographic information, which is central to survey research. However, advances in AI can help address these challenges. We describe how three tools, embedded in a Social Media (SM) Browser, leverage language models to support the use of social media data in public opinion research. The three tools are: summarization (generating textual summaries of posts), stance annotation (e.g., whether a post expresses support or opposition for a proposition or topic), and inferring the demographic characteristics of the user who created each post (e.g., gender, age, education—not directly available within posts or in users' profiles). These tools can help researchers develop insights about the topics being discussed, the opinions held about those topics, and what kind of users hold those topics, despite the volume of posts and the paucity of information about users.

1 Introduction

Social media platforms offer a promising source of data for understanding public thinking (Murphy et al., 2014; Mneimneh et al., 2021; Jensen et al, 2021). Compared to traditional methods, such as surveys or focus groups, social media posts can provide more immediate and cost-effective insights, and because of their sheer volume, they may also reveal a broader spectrum of perspectives—including more nuanced, outlier, or minority opinions (especially from those who are unlikely to participate in conventional research formats).

Early promising findings suggested that analyses of social media posts—such as sentiment (are the words positive or negative?) of posts containing specific keywords—could align with survey results (e.g., Daas et al., 2015; O'Connor et al., 2010). While subsequent systematic efforts did not replicate these patterns across broader timeframes and with other statistical methods, valuable insights were gained about when and how survey data and social media may align (Schober et al., 2016). For example, stance—whether a person is for or against something—may be a more informative measure than sentiment, when using social media data.

Still, leveraging social media data for public opinion research presents some challenges. First, the volume of content vastly exceeds what researchers can manually review and retain, necessitating technological assistance. Second, social media posts lack demographic information about the users, so even if one can discern what opinions are posted in the corpus, it is impossible to know who holds those opinions, e.g., people with more education. Recent advances in AI, including summarization (Yang et al., 2024) and classification of text (Abdurahman et al., 2025; Gilardi et al., 2023; Törnberg, 2024), however, have made it possible to address these issues in ways that may not be feasible otherwise.

Here we describe three features of the Social Media (SM) Browser, an interactive tool designed to help its users gain insights about the discourse in a large social media corpus, by summarizing posts, classifying their stance, and inferring demographics of the users who created them.

2 Language Models for Public Opinion Research

2.1 Generating Textual Summaries of Posts

While social media is a possible source of qualitative information about public opinion, its massive volume prevents people from reading all posts in a corpus. Large language models (LLMs) can help by summarizing large sets of posts. Recent advances in language models—driven by faster and more efficient architectures, greater computational power, and access to very large-scale training sets (Naveed et al., 2024)—have made it possible for LLMs to generate coherent, accurate summaries. We implemented a feature in the SM Browser that generates paragraph-long abstractive summaries to help users grasp the main themes across large sets of posts. We fine-tuned LLaMa2-13B to generate narrative summaries based on samples of 50 to 200 posts drawn from a corpus of 3.5 million posts. To support evaluation and transparency, the model was also trained to identify the posts on which it based each sentence in the summary and which posts were omitted from the summary, both of which can be inspected by a user of the tool to assess the summary's fidelity to the source posts. An evaluation confirmed that the summaries are of high quality and fit for the intended purpose.

2.2 Labeling Stance of Posts

Whether someone is for or against something is central to public opinion research—and there is no shortage of opinions on social media—but is not feasible for a person to read thousands (or millions) of posts and categorize which express support, opposition, or neutrality toward a given issue. Moreover, social media posts can be ambiguous—e.g., ungrammatical, elliptical, abbreviated, out of context. Stance-detection classifies the position a user expresses in a post. While earlier work in stance detection relied on supervised deep learning models, i.e., which relied on human-annotated posts to train the models (Küçük & Can, 2022), LLMs can annotate stance, with zero or few shot training. The stance annotation feature first encourages users to identify posts that are semantically related to the topic of interest ("relevant") using a search feature based on SBERT. By annotating only relevant posts maximizes the proportion for which a for a user-provided stance applies. Then, the user specifies up to four stance categories, e.g., "agree," "neutral," "disagree" and a fine-tuned LLaMa2 model is deployed to classify each post with one of these stance labels. This process begins to transform the corpus into something analogous to a survey data set in which closed responses are accompanied by rich textual content, i.e., the post (See Conrad et al., 2023). A graphical time series tool allows SM Browser users to visualize how the prevalence of stances might change over time.

2.3 Inferring Demographics of Users

Most social media posts contain no information about the user who created it and user profiles do not provide much if any user description (only a small percent of Twitter users provide a location in their profile, and beyond this, little else is available). WIthout demographic information about users researchers cannot say much about how opinions and attitudes may differ between groups. A solution is to prompt an LLM to infer attributes of the users who create each post, e.g., age, gender, and education, based just on the posted textand,. This could open the door to asking more nuanced questions of the sort one might ask of survey data. To enable SM Browser users to do this, we make available models trained to infer each of seven demographic characteristics. To do this we constructed a data set of all the Twitter and Reddit posts created by about 500 users, linked to their self-reported demographics characteristics and trained an open source LLM, Gemma-3-12B-IT, to predict these characteristics of each post's author. SM Browser users can plot change in opinion by group (age, gender, etc.) over time. Model predictions are good for all seven demographic characteristics and very good for several (Li, et al, 2025).

3 Discussion

These AI-enabled features fundamentally change how researchers can approach social media data for public opinion analysis by giving them a means to overcome the challenges inherent in this kind of data. These features allow them to extract themes without reading all posts and examine how opinions (e.g., favor a proposition) vary across (inferred) demographic groups over time.

4 Limitations

LLMs are a powerful new tool for public opinion researchers. But, there are limitations and practical considerations. First, LLMs can hallucinate or create "compelling misinformation" (Spitale et al., 2023), and so they require constant and intensive validation to ensure that results make sense (although we found no evidence of hallucination in the summaries of posts). Second, inferred information—such as user stance or demographics—should be interpreted cautiously, with an understanding that these are predictions that open the door further to quantitative analysis, not ground truth. Finally, the lack of transparency and reproducibility are cause for concern. Most LLMs operate as black boxes, with outputs that may vary across models—or even across time—as underlying algorithms are updated. This means that recommendations for researchers must also evolve; for example, including model justification, model name, and query date on research projects (Abdurahman et al., 2025).

5 References

Abdurahman, S., Salkhordeh Ziabari, A., Moore, A. K., Bartels, D. M., & Dehghani, M. (2025). A Primer for Evaluating Large Language Models in Social-Science Research. Advances in Methods and Practices in Psychological Science, 8(2), 25152459251325174. https://doi.org/10.1177/25152459251325174

Conrad, F., Li. M., Gagnon-Bartsch, J., Schober, M.F., Ferg, R., Dolgin, R., & Beatty, P. (May, 2023). Uncovering Alignment between Survey Responses and Social Media by Detecting the Stance Expressed in Posts. Paper presented at the American Association of Public Opinion Research 78th Annual Conference, Philadelphia, PA.

Daas, P. J. H., Puts, M. J., Buelens, B., & Hurk, P. A. M. van den. (2015). Big Data as a Source for Official Statistics. Journal of Official Statistics, 31(2), 249–262. https://doi.org/10.1515/jos-2015-0016

Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT Outperforms Crowd-Workers for Text-Annotation Tasks. http://arxiv.org/abs/2303.15056

Jensen, B., Singh, L., Davis-Kean, P., Abraham, K., Beatty, P., Bode, et al., (2021). Analysis and visualization considerations for quantitative social science research using social media data.

Küçük, D., & Can, F. (2022). A Tutorial on Stance Detection. Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, 1626–1628. https://doi.org/10.1145/3488560.3501391

Li, M. , Conrad, F.G., Schober, M.F., Gagnon-Bartsch, J., Dolgin, R., Chen, P., Zhou, E., & Beatty, Paul. (May, 2025). Who posted that? Automatically inferring characteristics of social media users. Paper presented at the American Association of Public Opinion Research 80th Annual Conference, St. Louis, MO.

Mneimneh, M., Pasek, J., Singh, L., Best, R., Bode, L., Bruch, E., Budak, C., Davis-Kean, P., Donato, K., Ellison, N., et al. 2021. Data acquisition, sampling, and data preparation considerations for quantitative social science research using social media data.

Murphy, J. Link, M.W., Childs, J. H., Tesfaye, C.L., Dean, E., Stern, M., Pasek, J., Cohen, J., Callegaro, M., Harwood, P., et al. 2014. Social media in public opinion research: report of the AAPOR task force on emerging technologies in public opinion research. American Association for Public Opinion Research.

Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N., & Mian, A. (2024). A Comprehensive Overview of Large Language Models (arXiv:2307.06435). arXiv. https://doi.org/10.48550/arXiv.2307.06435

O'Connor, B., Balasubramanyan, R., Routledge, B., & Smith, N. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. Proceedings of the International AAAI Conference on Web and Social Media, 4(1), Article 1. https://doi.org/10.1609/icwsm.v4i1.14031

Schober, M. F., Pasek, J., Guggenheim, L., Lampe, C., & Conrad, F. G. (2016). Social Media Analyses for Social Measurement. Public Opinion Quarterly, 80(1), 180–211. https://doi.org/10.1093/poq/nfv048

Spitale, G., Biller-Andorno, N., & Germani, F. (2023). AI model GPT-3 (dis)informs us better than humans. Science Advances, 9(26), eadh1850. https://doi.org/10.1126/sciadv.adh1850

Törnberg, P. (2024). Large Language Models Outperform Expert Coders and Supervised Classifiers at Annotating Political Social Media Messages. Social Science Computer Review, 08944393241286471. https://doi.org/10.1177/08944393241286471

Yang, J., Jin, H., Tang, R., Han, X., Feng, Q., Jiang, H., Zhong, S., Yin, B., & Hu, X. (2024). Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond. ACM Trans. Knowl. Discov. Data, 18(6), 160:1-160:32. https://doi.org/10.1145/3649506