Appendix A

Classifying opinions toward candidates using GPT-4o

This Pew Research Center analysis examines a total of 208,561 posts on Facebook, Instagram TikTok, X (Twitter) and YouTube posted by [news influencers](#) during the weeks of July 15-21, July 29-Aug. 4, Aug. 19-25, and Oct. 28-Nov. 4, before the 2024 U.S. presidential election, as well as the week of Nov. 5-11 covering Election Day and the days following. To understand the opinions expressed toward the presidential candidates in these posts, researchers first checked for direct mentions of candidate names in the text and/or transcript of the posts, then passed the text and/or transcript of each post containing a candidate name to OpenAI's GPT-4o model for opinion classification. The model was configured using the following instruction prompt:

> You are an AI assistant designed to help extract information from social media posts.
>
> It is currently November 2024.
> The United States just had a presidential election.
> The Democratic candidate was initially the incumbent Joe Biden, but he withdrew this past summer and was replaced by the incumbent Vice President, Kamala Harris.
> The Democratic ticket leading into the election consisted of Kamala Harris for president, and Tim Walz (currently governor of Minnesota) for vice president.
> The Republican ticket leading into the election consisted of Donald Trump for president, and JD Vance (currently junior senator from Ohio) for vice president.
> The outcome of the election was that Donald Trump won both the electoral college and the popular vote. Republicans also won narrow majorities in the House of Representatives and in the Senate.
>
> You will be provided with the text of a social media post, written either shortly before, or shortly after the 2024 Presidential Election. Please respond with the following information:
>
> What stance or opinion does the post express toward {candidate}?
> Respond with 'SUPPORTIVE', 'CRITICAL', or 'NO OPINION'.
> Opinions about a candidate can be nuanced and multidimensional--the primary concern here is simply whether the opinion expressed is generally positive (even if there is also some criticism), or generally negative (even if there are some positive points conceded). Neutral or factually presented coverage should be considered 'NO OPINION', as should posts where a clear opinion is not expressed. Note that on its own, a statement predicting the election's outcome, or commenting on how a given candidate performed is not necessarily an expressed opinion, although it can be a contributing factor.

Model validation

To validate the opinion labels generated by the GPT-4o model, we used a sample of 500 posts. These posts were independently annotated as indicating support for the candidate, opposition to the candidate, or no strong opinion toward the candidate by four researchers. Ground truth consensus labels were then established from the set of multiple annotations per post using a Dawid-Skene aggregation model.

We compared the predicted labels generated by the model to these consensus labels in two ways: first, as the share of labeled posts where the model categorized the post with the exact same opinion label as the human annotators' consensus label; and second, as the share of labeled posts where the model categorized the post with the same directionality of opinion as the human annotators' consensus label. For instance, under this metric, if the consensus label was that a post was supportive toward Donald Trump, the model is graded as correct if it predicts either "supportive" or "no opinion" and incorrect if it predicts "critical." Both sets of metrics are shown in the table below:

Opinion labeling accuracy metrics

	% agreement: SUPPORTIVE vs. NO OPINION vs. CRITICAL	% agreement: SUPPORTIVE/NO OPINION vs. CRITICAL/NO OPINION
Opinion expressed toward Joe Biden	77%	99%
Opinion expressed toward Kamala Harris	77%	99%
Opinion expressed toward Donald Trump	72%	99%

Source: Pew Research Center analysis of 208,561 posts published July 15-21, July 29-Aug. 4, Aug. 19-25, Oct. 28-Nov. 4, and Nov. 5-11, 2024, by 500 influencers with over 100,000 followers who regularly posted about news on Facebook, Instagram, TikTok, X or YouTube in early 2024. Refer to methodology for details.

PEW-KNIGHT INITIATIVE