Generative AI and Discovery of Preferences for Single-Use Plastics Regulations

Catharina Hollauer,[1] Jorge Garcelán,[2] Nikhita Ragam,[1] Tia Vaish,[1] Omar Isaac Asensio[1,3*]
Georgia Institute of Technology,[1] University Charles III of Madrid,[2] Harvard Business School[3]
* Correspondence to: asensio@gatech.edu

Abstract

Given the heightened global awareness and attention to the negative externalities of plastics use, many state and local governments are considering legislation that will limit single-use plastics for consumers and retailers under extended producer responsibility laws. Considering the growing momentum of these single-use plastics regulations globally, there is a need for reliable and cost-effective measures of the public response to this rulemaking for inference and prediction. Automated computational approaches such as generative AI could enable real-time discovery of consumer preferences for regulations but have yet to see broad adoption in this domain due to concerns about evaluation costs and reliability across large-scale social data. In this study, we leveraged the zero and few-shot learning capabilities of GPT-4 to classify public sentiment towards regulations with increasing complexity in expert prompting. With a zero-shot approach, we achieved a 92% F1 score (s.d. 1%) and 91% accuracy (s.d. 1%), which resulted in three orders of magnitude lower research evaluation cost at 0.138 pennies per observation. We then use this model to analyze 5,132 tweets related to the policy process of the California SB-54 bill, which mandates user fees and limits plastic packaging. The policy study reveals a 12.4% increase in opposing public sentiment immediately after the bill was enacted with no significant changes earlier in the policy process. These findings shed light on the dynamics of public engagement with lower cost models for research evaluation. We find that public opposition to single-use plastics regulations becomes evident in social data only when a bill is effectively enacted.

Introduction

Policies targeting single-use plastics in the US and EU have increased in recent years due to concerns about plastic pollution and environmental impacts (Jebe 2022; Bachmann et al. 2023; Milbrandt et al. 2022; Xanthos and Walker 2017; Jambeck 2015). Climate change and plastics pollution policy are invariably linked, as currently 3.4% of global greenhouse gas (GHG) emissions are related to fossil fuel production and conversion in the plastics lifecycle. These emissions are projected to more than double by 2060, reaching an estimated 4.3 billion tons globally (OECD 2023). Reducing or recycling plastics reduces emissions by lowering the demand for primary plastics raw materials and increasing post-consumer recovery. In recent years, social movements around ocean microplastics and climate change have led to a wave of many US states enacting laws regulating single-use plastics under extended producer responsibility laws. These include plastics used in primary packaging, carryout bags, food service containers, plastic straws and other packaging materials, which have leaked into the environment and aquatic ecosystems due to inadequate disposal and management (Elliott, Gillie, and Thomson 2020; Xanthos and Walker 2017; Barnosky, Delmas, and Huysentruyt 2019). In response, social media has become a stage for supporters and opponents of these regulations to express their opinions, and even engage in the policy process (Mavrodieva et al. 2019; Anderson 2017; Pathak, Henry, and Volkova 2017).

Investigations of citizen engagement and participation on social media is a broad area of research in computational social science (Salganik, 2019; Loader, Vromen, and Xenos 2014; Siyam, Alqaryouti, and Abdallah 2020; Asensio et al. 2020; Tan, Cui, and Xi 2021). Platforms such as Twitter, now called X, and competing social platforms have facilitated the study of how individuals or organizations engage with their communities and participate in online discussions and make it easy for users to increase political and social participation. We examine whether AI-augmented analysis of this user data could be used to make broader inferences about policy preferences, based on the content of user shared information. However, reliably detecting large-scale support or opposition to single-use plastics regulations has been a difficult task, as current methods depend heavily on slow and costly government surveys or opinion polls for policy feedback. Additionally, prevailing survey-based approaches require human-intensive data curation and analysis, which can also be subject to measurement challenges related to hindsight, recency and other biases that can limit real-time analyses.

Recent advancements in generative AI models, which are capable of augmenting or even replacing human crowdsourcing in specific contexts, could allow us to understand the dynamics of public response to climate regulations

at lower cost, offering insights from citizen data at a large-scale and in near-real time (Boussioux et al. 2023; Chung et. al 2022). Large language models (LLMs) like OpenAI's GPT-4 (OpenAI 2023) have demonstrated zero shot learning capabilities with the potential to scale, adapt, and perform a wide range of natural language processing tasks with reduced need for extensive and costly expert human-annotated data (Christiano et al. 2017; Ouyang et al. 2022, Zhao 2023, Dillion 2023). However, zero-shot classification of policy preferences with attributes or contextual descriptions from the public discourse on plastics regulations, often requires nuance understanding and social cues that have been hard to generalize. For example, a user writes: "LA plastic bag ban is like an anorexic putting on make-up ~ a pretense of a modicum of control while ignoring the pink elephant in the room". To fully understand how this statement reflects the user's stance on the regulation, whether in favor or opposition, one must have additional knowledge about the social context and idioms. This usually requires human expertise on the social context which goes beyond literal or keyword-based discovery. In this study, we investigate whether large pre-trained generative AI models can be used to measure public stated support or opposition towards single-use plastics regulations and policies. We investigate GPT-4 zero and few-shot learning capabilities with varying degrees of complexity in expert prompting to analyze its domain performance in climate change and environmental public discourse tailored to classification of single-use plastics policies.

Data and Methodology

We used the Twitter Academic API to retrieve relevant tweets related to single-use plastics regulations. To maintain a broad search, we did not restrict the tweets by geo-tagged locations (Malik et al. 2015) but included English language tweets, published from January 1st 2010 until October 31st 2023. Because the term 'plastics' can be used in a variety of contexts other than policy or regulations concerning single-use plastics, we implemented a Boolean search for the expression "plastic bag" with either the words "ban," "tax," "levy," or "fee" and collected a set of 1,457,420 tweets. We then removed duplicate tweets which resulted in a subset of 1,049,062 tweets. To build decision rules for classification, we developed guidelines that characterize behavioral intent of the user to construct different categories of tweets being favorable, opposing or neutral to single-use plastics regulation. In the next section we describe the human annotation experiments that served to build the ground truth for classification.

Human Experiments

A total of five human annotation experiments were conducted between September 2022 and April 2023 under Approved Institutional Review Board (IRB) Protocol Number H22242. Six annotators were divided into two groups and in each experiment, a training session was provided with guidelines for classifying the randomly selected tweets on single-use plastics policies. This was followed by an annotation session in which annotators classified a total of 400 tweets. Out of these 50 are ground truths, which were used to determine interrater agreement amongst annotators. Tweets were classified in one of three different labels: (1) favor, (2) oppose, or (3) neutral.

A "favor" label is defined as a tweet that advocates or shows support for plastic regulations, both from a first person or third person perspective or indicates favorable outcomes as a result of plastic regulations. For example, a user advocates: "Plastic bags are choking our marine life. Tell Environment Ministers to #banthebag now! https://t.co/RcVz8AJNBD via." A less common occurrence is in the form of a double negative. For example, a user shares: "Bette Midler Blasts Plastic Bag Ban Collapse in Calif. http://t.co/eVSyAdnmTT via @BreitbartNews," which is classified as favor.

A "neutral" label is defined as a tweet that provides information about the occurrence of a plastic policy or what it entails without additional affirming or dissenting commentary towards plastic regulations. A user shares: "NEW DETAILS: Proposed Plastic Bag Ban Bill http://t.co/CdjP2FSE." This tweet contains a news update of what a plastic bag ban entails with no opinionated reactions. Neutral tweets can also appear in the form of a question. For example, a user writes: "Have you heard about the plastic bag ban in Bali? https://t.co/jbVwKyPJQL."

Finally, a tweet is labeled "oppose" if it has commentary that advocates against plastic regulations from a first person or third person perspective or indicates adverse outcomes. One user shares: "Texas retailers sue city of Austin over plastic bag ban: http://t.co/AcpKVZryVA." Describing legal opposition from a third party against plastic regulations warrants this classification. Not as common, oppose tweets can be in favor of a plastic bag ban or plastic regulations but criticize its current execution. For example, a user writes:

> @user @user That's the point, the tax is not substantial enough to deter plastic bag usage, but calculated to seem like only a minor inconvenience\n\nFor every 1 million bags sold they make $100K, on top of the CA sales tax, which are as high as 10%...\n\nBut u a yang fan, so ur dumb, its OK

After screening tweets to learn about user patterns of communication regarding plastics regulation, a typology was developed with decision nodes for annotators with representative descriptions (Appendix B). A decision tree was

developed and made available at each experiment to serve as a guide to the annotators for classifying tweets. The most commonly occurring tweets express an opinion on plastic regulations through affirming or dissenting verbiage and the rarest node are tweets structured as a double negative. We provided the distribution of representative tweets by decision node in Appendix C. Through our internal experiments, every tweet sampled could be classified in exactly one of the existing nodes. The ground-truth decision tree provided to the annotators to classify tweets, followed a structured and hierarchical process to help the annotators to make decisions regarding the label for a specific tweet. In each node, the annotator must decide if the tweet still fits on that tree path leading to a specific label. The nodes account for edge cases, such as whether a user is in favor of banning plastic bags but takes issue with how current policies are being executed.

In all human experiments, conducted in-person or virtually, annotators were provided with a review from prior sessions along with the decision tree and corresponding examples as a refresher to allow for the continuous learning and improvement on tweet classification. After providing directed examples in each annotation session, we held a series of short debriefings to ensure common understanding amongst both groups. Each annotator group received a unique set of tweets to work within their respective groups on the classification. After each experiment, the interrater agreement was calculated.

The interrater agreement is assessed using Cohen's Kappa coefficient, a statistical measure designed to compute the reliability of agreement among annotators. The Cohen's Kappa coefficient quantifies the agreement between raters while considering the possibility of agreement occurring by chance alone. We interpret a value of 0 as complete disagreement beyond what would be expected by chance, while a value of 1 indicates perfect or complete agreement. The formula for Cohen's Kappa is provided below:

$$k = \frac{(P_0 - P_E)}{(1 - P_E)}$$

where P_0 is the actual observed agreement and P_E the chance agreement. In the case of our study, we have:

$$P_E = P(favor) + P(oppose) + P(neutral)$$

where $P(favor)$ is the probability that the annotators would randomly both say "favor" to a given tweet, while $P(oppose)$ is the probability to "oppose" and $P(neutral)$ to "neutral". As the experiments advanced, it became evident that annotators exhibited a learning curve concerning their agreement on tweets classification as observed in Fig 1.

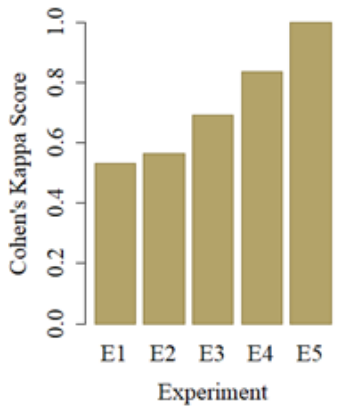


Fig 1. Interrater Agreement per Experiment

It is possible to observe the interrater agreement score increasing after each experiment. After five annotation sessions, the annotators achieved perfect agreement, which translated into a Cohen's Kappa score of 1. In this way we ensure that the human annotated data used for testing and validation were of high-quality.

ChatGPT-4 Prompt Engineering

The OpenAI GPT-3 and GPT-4 language models were employed to classify tweets about plastics reduction policies between three different labels: favor (1), oppose (2), or neutral (3). The data comprised of 50 ground truths tweets from the original sample of tweets that were collected, and each run was conducted over 10 replications. We tested both zero-shot and few-shot learning capabilities of GPT-4 and GPT-3, to allow for a fair comparison in performance between these two Generative AI models. While we utilized the same main prompt for all the runs, some additional prompts were included in some runs to allow for proper analysis of performance. Run 1 is considered the baseline model as it only utilizes the main prompt, we rely on the AI's pre-trained understanding of "favor", "oppose" and "neutral" with respect to plastics regulation. The chain-of-thought prompting included combinations of label definitions, context explanations and examples in order to be able to isolate the relative performance of the prompting. Examples of tweets were provided for the few-shot runs, while for the zero-shot runs no examples of tweets were provided. The main dialogue prompt for every run was: "I will input tweets about plastics reduction policies. Classify them in one word: favor (1), oppose (2), neutral (3)". Next, we describe the label definitions and context explanations.

Label definitions

For the "favor" label:

> "These tweets will be supportive of plastic regulations and policies. They might discuss favorable outcomes

because of such regulations, talk about someone else or an organization supporting these regulations, or show support for a potential plastic regulation."

And for the "oppose" label:

> "These tweets will show opposition to plastic regulations or policies. They might criticize a specific policy, discuss someone else or an organization being against these regulations, or point out negative outcomes due to such regulations."

And for the "neutral" label:

> "Neither favor nor oppose plastic regulations, this includes tweets in which people might just share information about regulation timelines or other general questions on these policies."

Context explanations

We included context explanations to define boundary conditions for each representative label to avoid commonly misclassified examples. These were developed following the ground-truth decision tree. For example, a user writes "Canada's ban on harmful single-use plastic starts taking effect today. We CAN eliminate #plasticpollution by 2030. #COP15." The context explanation provided in the prompt for this tweet classified as "favor" would be:

> "Although this is talking about when a plastic regulation begins, warranting a 3 (neutral) classification, this tweet is classified as 1 (favor) because of the adjective 'harmful' being used to describe single-use plastics. The use of this adjective indicates that the user believes plastics are not good for the environment, which means they are favorable towards plastic regulations."

In another instance, a user writes: "Grocery shopping was an enjoyable errand before the plastic bag ban." Similarly, the context explanation clarifies that this tweet expresses opposition towards the plastic bag ban with an ironic tone despite the use of positive words or phrases. Neutral classifications also can require context explanations. For example: "New York State Plastic Bag Ban starts March 1, 2020 https://bit.ly/3Ocx716The." The context explanation provided was:

> "This is an example of a Tweet with classification neutral because it is informational and does not use affirming or opposing jargon. It is simply spreading information on when a plastic bag ban begins."

In the next section, we present the results on the chain-of-thought experiments that allowed us to evaluate the relative performance of the generative AI with varying levels of expert prompting.

Analysis and Discussion

We first explored the zero-shot learning capabilities of generative AI without including any tweet examples, label definitions or context explanations. The results for both GPT-3 and GPT-4 are presented in Table 1 and also in Appendix A. The baseline model produced an accuracy of 0.80 (0.01) and an F1 Score of 0.77 (0.02), which is remarkable as in dozen of previous domain-specific or general-purpose systems, the performance for sentiment classification does not achieve comparable accuracy, commonly staying within the range of 0.65 and 0.70 (Giachanou and Crestani 2016; Zimbra et al. 2018). In domains such as climate change and environmental regulations, there is a growing interest in understanding citizen engagement through social media (Pathak, Henry, and Volkova 2017).

In the zero-shot learning capability, the inclusion of label definitions in the prompt as in Run 2, observes a significant improvement in accuracy of 0.92 (0.01) and F1 score of 0.91 (0.01) in comparison to the baseline. In the few-shot learning capability, we experimented with variations of prompts, and the performance of the model remained consistently superior to that of the baseline. In all the experiments, GPT-4 achieved a higher performance than GPT-3, and thus, it is the model chosen to be leveraged for the policy case study.

Models	Experiment	Tweet examples	Label definitions	Context explanations	Accuracy	F1 Score	Precision	Recall
	Zero-shot							
GPT4	Run 1	No	No	No	0.80 (0.01)	0.77 (0.02)	0.82 (0.02)	0.76 (0.02)
GPT3	Run 1	No	No	No	0.72 (0.02)	0.71 (0.01)	0.73 (0.02)	0.71 (0.02)
GPT4	Run 2	No	Yes	No	0.92 (0.01)	0.91 (0.01)	0.91 (0.01)	0.93 (0.01)
GPT3	Run 2	No	Yes	No	0.74 (0.01)	0.71 (0.01)	0.73 (0.02)	0.71 (0.02)
	Few-shot							
GPT4	Run 3	Yes	No	No	0.87 (0.02)	0.86 (0.03)	0.87 (0.02)	0.86 (0.03)
GPT3	Run 3	Yes	No	No	0.71 (0.02)	0.68 (0.02)	0.70 (0.02)	0.68 (0.02)
GPT4	Run 4	Yes	No	No	0.89 (0.01)	0.88 (0.01)	0.89 (0.01)	0.90 (0.01)
GPT3	Run 4	Yes	No	No	0.74 (0.02)	0.72 (0.03)	0.74 (0.03)	0.71 (0.03)
GPT4	Run 5	Yes	Yes	Yes	0.84 (0.01)	0.84 (0.01)	0.85 (0.01)	0.86 (0.02)
GPT3	Run 5	Yes	Yes	Yes	0.73 (0.02)	0.71 (0.02)	0.72 (0.02)	0.70 (0.02)

Table 1: Prompt engineering results for few- and zero-shot learning for GPT3 (text-davinci-003) and GPT4 (gpt4-0314)

Policy Case Study

We leveraged GPT-4 model for sentiment classification to evaluate public's perception on California Senate Bill 54 (SB-54). SB-54 aims to prevent plastic pollution and encourage responsibility among producers. The goal is to have 100% of packaging in California recyclable or compostable by 2032, cut plastic packaging by 25%, and recycle 65% of all single-use plastic packaging (California Legislative Information, 2022). The policy process in California commenced in December of 2020 when the bill was introduced in the Senate. Over a two-year period, the bill was refined and reviewed and ultimately enacted in June of 2022. The policy process and dates are presented in Figure 2.

Figure 2: SB-54 dates

We collected a specific set of 5,132 tweets focused on the timeline of the SB-54, out of which, 2,412 tweets were related to the enactment period. We evaluated the opposition towards SB-54 across social media over time with a window of periods of 6 months before and after Jun 30, 2022. We classified the enactment related tweets in a zero-shot learning and "label definitions" approach as described in the previous section. To better understand the opposition to SB-54 and to draw inference on public perception on plastics regulations, we adopted a regression discontinuity design (RDD) for time to event outcomes (Thistlethwaite and Campbell 1960; Gelman and Imbens 2019). Since the choice of bandwidth can quantitively affect the estimates, we used the Imbens-Kalyanaraman method to automatically find the optimal bandwidth using the expected-squared-error-loss criterion and selected several values at or below the optimal bandwidth (Imbens and Kalyanaraman 2012). This event study approach is appropriate as treatment assignment is measured and determined by the enactment of SB-54 in

the state of California. We present the results of the opposition to SB-54 in a daily aggregation in Figure 3 and the event study RDD estimates in Table 2.

Following the SB-54 enactment, opposition observes a significant increase of 12.4%. This also suggests that engagement and participation during the policy process may be limited or not readily evident through tweets activity. We find that this approach allows for the monitoring of social movements and engagement in real-time which can provide valuable insights to policymakers.

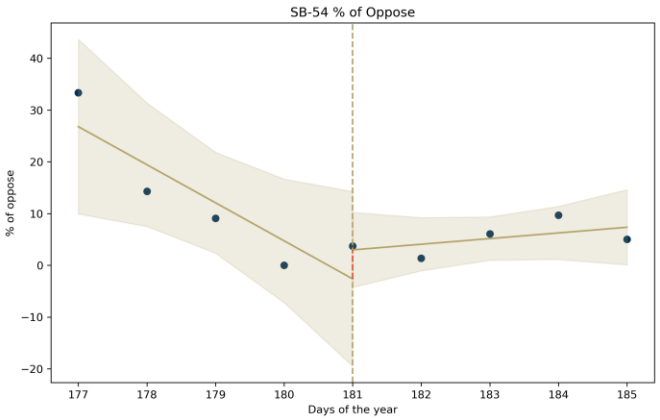


Figure 3: SB-54 Opposition Over Time in Days Surrounding the Enactment of the Bill

Bandwidth	Estimate	p-value
5.0	12.424 (2.290)	0.000***
2.5	12.546 (0.281)	0.000***

Note: The dependent variable is percentage of opposing tweets aggregated daily. Robust standard errors in parenthesis.
*Significant to * p<.05 ** p<.01 *** p<.001.*

Table 2: RDD Estimates for SB-54 Opposition Over Time in Days Surrounding the Enactment of the Bill

To evaluate potential reductions in research evaluation costs, we compared the relative cost of human expert annotation to an AI-driven approach for classifying user tweets related to the SB54 policy experiment. Assuming that each annotator is paid at least minimum wage ($7.25 is currently the federal minimum) and an individual expert can annotate 100 tweets per hour, in the case of human annotators, it would cost $362.50 for 5,000 tweets per annotator in the policy experiment; and $76,057 per annotator for tweets in the total dataset of 1,049,062 tweets. Annotation occurs in groups of 6 individuals, resulting in a total cost of $2,175 for the policy experiment and $456,341.97 for the total dataset. This does not account for coordination time and cost to administer the research. In contrast, the AI-driven approach would cost $6.90 for the policy case and $1,447.7 for the total dataset of tweets (currently approximately $.06 per 1000 prompt tokens for prediction in a zero-shot approach and an average of 23 tokens per tweet), making it significantly more cost-effective than relying on human experts alone.

Closing

This study presented an approach to real-time monitoring of public perception towards single-use plastics regulations and policies on social media. Through a variety of expert prompting strategies with minimal fine-tuning, we achieved high performance with GPT-4 in this domain. This study demonstrates that generative AI models with zero-shot learning may be readily deployed as an input for subsequent analysis in a variety of causal inference and prediction settings to track public responses to policy processes in near real-time, and at relatively lower cost. We demonstrate that high-performing and scalable human-in-the-loop AI systems can be deployed with a substantial cost reduction in research evaluation, of up to three orders of magnitude compared to prevailing annotation approaches. Such capabilities will make it possible to accommodate millions of users and ultimately help foster more responsive and effective governance with expanded citizen intelligence and public participation.

Ethical Statement

This research was conducted under approved Institutional Review Board (IRB) Protocol Number H22242.

Acknowledgments

We gratefully acknowledge funding awards from Microsoft and National Science Foundation Award #2028998 and Award #1945332. We thank our EFRI collaborators Carsten Sievers, Christopher Jones, Sankar Nair and Fani Boukouvala. For expert annotation support, we also thank Elisavet Anglou, Yuchen Chang, Natechanok Yutthasaksunthorn, Arvind Ganesan, Erin Phillips, Van Son Nguyen. For institutional support, we thank Debra Spar and colleagues at the Institute for the Study of Business in Global Society (BiGS) at the Harvard Business School. For valuable exploratory research assistance, we thank Sarah Canastra, Abby Bower, and Matteo Zullo. This research was also supported in part through research cyberinfrastructure resources and services provided by the Partnership for an Advanced Computing

Environment (PACE) at the Georgia Institute of Technology, Atlanta, Georgia, USA.

Author Contributions

Conceptualization: O.I.A.; **Data Curation:** N.R., T.V.; **Formal Analysis:** C.H., J.G.; **Funding acquisition:** O.I.A.; **Investigation:** C.H., J.G., N.R., T.V., O.I.A.; **Methodology:** C.H., J.G., N.R., T.V., O.I.A.; **Project administration:** O.I.A.; **Resources:** O.I.A.; **Software:** C.H., J.G., T.V.; **Supervision:** C.H.; **Validation:** C.H., J.G., N.R., T.V.; **Visualization:** J.G., N.R., T.V.; **Writing original draft:** C.H., J.G., N.R., T.V., O.I.A.; **Writing review & editing:** C.H., O.I.A.

Appendix A: GPT-3 and GPT-4 Evaluation

To allow for comparison among the most recent transformers models, we evaluated both GPT3 and GPT4 and the results are shown in Figure B1 below. A description of the prompt used in the experimental runs is provided in Table 1.

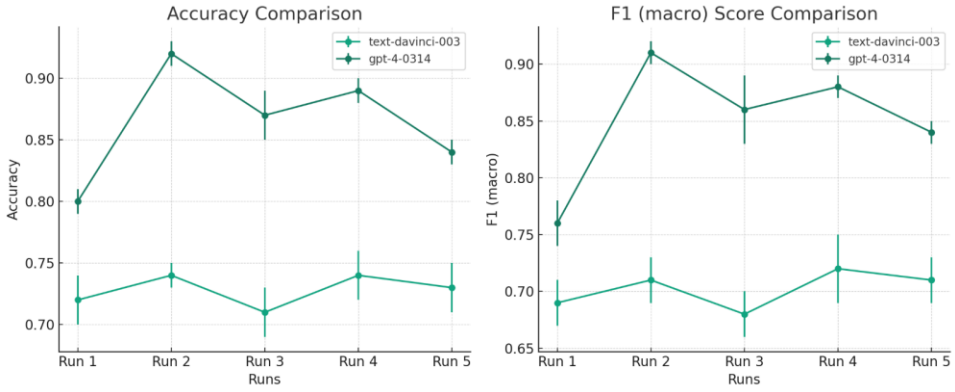
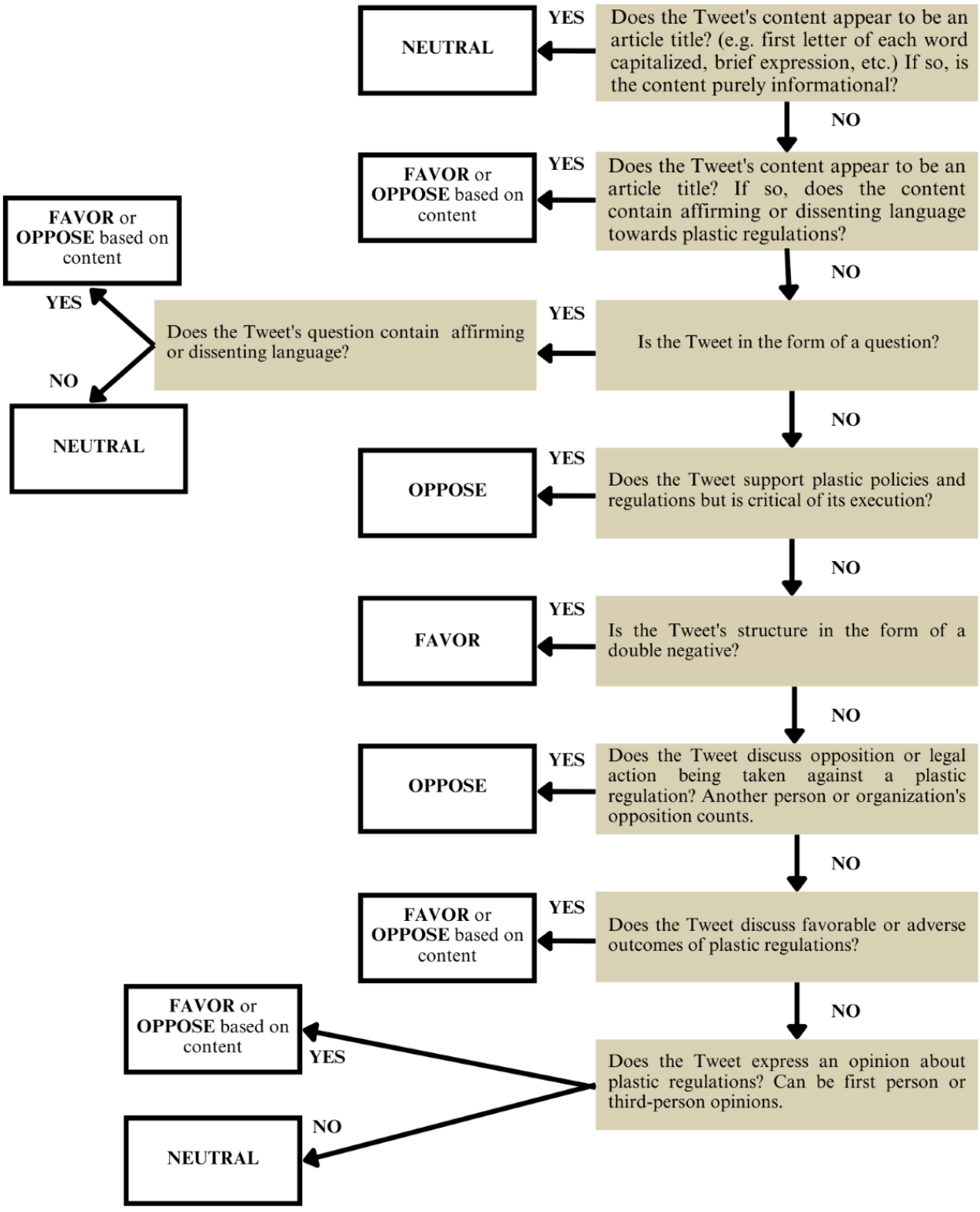


Figure A1: Comparison of results for zero-shot learning for GPT3 (text-davinci-003) and GPT4 (gpt4-0314)

Appendix B: Ground-Truth Decision Tree



Appendix C: Ground-Truth Decision Tree Nodes

We conducted multiple expert annotator experiments, in which thousands of tweets were read and analyzed to develop a logic-based decision tree. Each ground truth was classified into 1 of 7 decision nodes. We validated this decision tree in subsequent experiments to confirm that our decision tree was comprehensive and generalizable to fresh samples of tweets.

Decision node	Description	Counts of examples
Article title, informational	Content has no affirming or dissenting language and appears to be from news source	11
Article title, affirming or dissenting language	Content appears to be from news source and conveys a stance on plastic regulations	5
Question	Generally classified as neutral but can be favor or oppose if it takes a stance on plastic regulations	3
Supportive but critical of execution	User is in favor of plastic regulations but does not agree with current implementation so classified as oppose	3
Double negative	Uses double negative format to convey support so classified as favor	1
Favorable or adverse outcomes	Classified as favor or oppose based on whether it discusses positive or negative outcomes of plastic regulations	8
Express opinion	If tweet expresses an opinion, classified as favor or oppose based on opinion otherwise neutral	19

Table C1: Decision nodes used for prompting experiments.

References

Anderson, A.A., 2017. Effects of social media use on climate change opinion, knowledge, and behavior. *In Oxford research encyclopedia of climate science*. doi.org/10.1093/acrefore/9780190228620.013.369.

Asensio, O.I., Alvarez, K., Dror, A., Wenzel, E., Hollauer, C. and Ha, S., 2020. Real-time data from mobile platforms to evaluate sustainable transportation infrastructure. *Nature Sustainability*, 3(6): 463-471. doi.org/10.1038/s41893-020-0533-6

Bachmann, M., Zibunas, C., Hartmann, J., Tulus, V., Suh, S., Guillén-Gosálbez, G. and Bardow, A., 2023. Towards circular plastics within planetary boundaries. *Nature Sustainability*, 6(5): 599-610. doi.org/10.1038/s41893-022-01054-9.

Barnosky, E., Delmas, M. A., and Huysentruyt, M., 2019. The circular economy: motivating recycling behavior for a more effective system. Available at SSRN 3466359.

Boussioux, L., N Lane, J., Zhang, M., Jacimovic, V., and Lakhani, K. R. 2023. The Crowdless Future? How Generative AI Is Shaping the Future of Human Crowdsourcing. The Crowdless Future.

California Legislative Information, 2022. Senate Bill No. 54: SB-54 Solid waste: reporting, packaging, and plastic food service ware. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202120220SB54. Accessed: 2023-

28-07.

Christiano, P.F., Leike, J., Brown, T., Martic, M., Legg, S. and Amodei, D., 2017. Deep reinforcement learning from human preferences. *Advances in neural information processing systems*, 30.

Chung, J. J. Y., Kim, W., Yoo, K. M., Lee, H., Adar, E., and Chang, M., 2022. TaleBrush: Sketching stories with generative pretrained language models. *In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1-19).

Dillion, D., Tandon, N., Gu, Y. and Gray, K., 2023. Can AI language models replace human participants?. *Trends in Cognitive Sciences*. doi.org/10.1016/j.tics.2023.04.008.

Elliott, T., Gillie, H. and Thomson, A., 2020. European Union's plastic strategy and an impact assessment of the proposed directive on tackling single-use plastics items. *In Plastic waste and recycling*: 601-633. Academic Press.

Gelman, A., and Imbens, G., 2019. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3), 447-456.

Giachanou, A. and Crestani, F., 2016. Like it or not: A survey of twitter sentiment analysis methods. *ACM Computing Surveys (CSUR)*, 49(2): 1-41. doi.org/10.1145/2938640.

Imbens, G., and Kalyanaraman, K., 2012. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3), 933-959.

Jambeck, J.R., Geyer, R., Wilcox, C., Siegler, T.R., Perryman, M., Andrady, A., Narayan, R. and Law, K.L., 2015. Plastic waste inputs from land into the ocean. *Science*, 347(6223): 768-771. Doi.org/10.1126/science.1260352

Jebe, R., 2022. The US Plastics Problem: The Road to Circularity. *Env't L. Rep*., 52: 10018.

Loader, B.D., Vromen, A. and Xenos, M.A., 2014. The networked young citizen: social media, political participation and civic engagement. Information, *Communication & Society*, 17(2): 143-150. doi.org/10.1080/1369118X.2013.871571.

Malik, M.; Lamba, H.; Nakos, C.; and Pfeffer, J. 2015. Population bias in geotagged tweets. In proceedings of the international AAAI conference on web and social media. 9(4): 18-27. doi.org/10.1609/icwsm.v9i4.14688.

Mavrodieva, A.V., Rachman, O.K., Harahap, V.B. and Shaw, R., 2019. Role of social media as a soft power tool in raising public awareness and engagement in addressing climate change. *Climate*, 7(10): 122. doi.org/10.3390/cli7100122.

Milbrandt, A., Coney, K., Badgett, A. and Beckham, G.T., 2022. Quantification and evaluation of plastic waste in the United States. *Resources, Conservation and Recycling*, 183: 106363. doi.org/10.1016/j.resconrec.2022.106363.

Pathak, N., Henry, M.J. and Volkova, S., 2017. Under-standing social media's take on climate change through large-scale analysis of targeted opinions and emotions. In AAAI 2017 Spring Symposium on Artificial Intelligence for the Social Good, Technical Report SS-17-01.

OECD, 2023. Policy Highlights: Climate Change and Plastics Pollution, Synergies between two crucial environmental challenges. https://www.oecd.org/environment/plastics/Policy-Highlights-Climate-change-and-plastics-pollution-Synergies-between-two-crucial-environmental-challenges.pdf . Accessed: 2023-28-07.

OpenAI, 2023. Gpt-4 technical report. https://cdn.openai.com/papers/gpt-4.pdf. Accessed:2023-28-07.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A. and Schulman, J., 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35: 27730-27744.

Salganik, M.J., 2019. Bit by bit: Social research in the digital age. Princeton University Press.

Siyam, N., Alqaryouti, O. and Abdallah, S., 2020. Mining government tweets to identify and predict citizens engagement. *Technology in Society*, 60: 101211. doi.org/10.1016/j.techsoc.2019.101211.

Tan, W., Cui, D. and Xi, B., 2021. Moving policy and regulation forward for single-use plastic alternatives. *Frontiers of Environmental Science & Engineering*, 15: 1-4. doi.org/10.1007/s11783-021-1423-5.

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, 51(6), 309.

Xanthos, D. and Walker, T.R., 2017. International policies to reduce plastic marine pollution from single-use plastics (plastic bags and microbeads): A review. *Marine pollution bulletin*, 118(1-2): 17-26. doi.org/10.1016/j.marpolbul.2017.02.048.

Zhao, Y., 2023. The State-of-art Applications of NLP: Evidence from ChatGPT. *Highlights in Science, Engineering and Technology*, 49: 237-243. doi.org/10.54097/hset.v49i.8512