Sentiment and Interest Detection in Social Media using GPT-based Large Language Models

Md Abdullah Al Asad
mdabdullah-.al-asad@students.cau.edu
Clark Atlanta University
Atlanta, Georgia, USA

Hasan Md Imran
Clark Atlanta University
Atlanta, Georgia, USA
hasanmd.imran@students.cau.edu

Md Alamin
Michigan State University
East Lansing, Michigan, USA

Tareque Abu Abdullah
Clark Atlanta University
Atlanta, Georgia, USA
arequeabu.abdullah@students.cau.edu

Suriya Islam Chowdhury
Clark Atlanta University
Atlanta, Georgia, USA
suriyaislam.chowdhur@students.cau.edu

ABSTRACT

In the ever-expanding realm of social media, deciphering sentiments and identifying relevant topics from various textual content is challenging. This paper comprehensively investigates sentiment analysis and interest/topic detection using cutting-edge language models, ChatGPT3.5 and gpt4All. The methodology encompasses data collection, meticulous text pre-processing, innovative prompt design, and the exploration of zero-shot, one-shot, and few-shot learning techniques. We unveil the nuances of model performance in sentiment analysis and interest/topic detection through a detailed comparative analysis. Our findings highlight the power of ChatGPT3.5 in achieving a substantial accuracy enhancement in sentiment analysis compared to gpt4All. Moreover, we delve into the intricacies of interest detection, demonstrating the complexities of linguistic structures and model biases. We offer a holistic view of social entities' preferences by categorizing topics into distinct domains. A web portal developed using Google's Flutter SDK facilitates the visualization of user-friendly sentiment and interest outcomes. This research contributes to the understanding of sentiment analysis and interest detection and underscores the evolving capabilities of AI and NLP in navigating the dynamic landscape of social media content.

CCS CONCEPTS

• **Computing methodologies** → **Information extraction**; • **Information systems** → **Language models**.

KEYWORDS

GPT3.5, Gpt4ALL, social media, sentiment analysis, topic modeling, interest detection

MLNLP 2023, December 27–29, 2023, Sanya, China

ACM ISBN 979-8-4007-0924-1/23/12
https://doi.org/10.1145/3639479.3639523

ACM Reference Format:
Md Abdullah Al Asad, Hasan Md Imran, Md Alamin, Tareque Abu Abdullah, and Suriya Islam Chowdhury. 2023. Sentiment and Interest Detection in Social Media using GPT-based Large Language Models. In *2023 6th International Conference on Machine Learning and Natural Language Processing (MLNLP 2023), December 27–29, 2023, Sanya, China.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3639479.3639523

1 INTRODUCTION

Sentiment analysis has long been a focal point of investigation in natural language processing (NLP). This area systematically explores people's opinions, emotions, and sentiments through computational approaches (Liu, 2015[11]; Poria et al., 2020[13]). Since its inception((Turney, 2002[14]; Hu and Liu, 2004 [9])), sentiment analysis has attracted considerable attention from both academia and industry due to its diverse range of applications, such as evaluating product reviews and gleaning insights from social media posts (Barbieri et al., 2020 [3]; Zhang et al., 2022 [20]). Furthermore, comprehending human subjective feelings through sentiment analysis represents a significant stride toward developing artificial general intelligence (Bubeck et al., 2023[5]).

Interest detection has become increasingly popular [6] as it analyzes social media posts to determine the themes and topics that attract users' attention. Revealing the interests and sentiments of social media users can be applied in various ways, including targeted marketing, political campaigning, and public opinion assessment. However, analyzing social media posts for sentiment and interest detection takes time and effort. Social media posts can be ambiguous or sarcastic or contain multiple sentiments or interests. Additionally, social media users may only sometimes be truthful or accurate in their posts. Recent years have witnessed the emergence of large language models (LLMs) like GPT-3 (Brown et al., 2020[4]), PaLM (Chowd- Hery et al., 2022[7]), and GPT-4 (OpenAI, 2023[12]), which have showcased impressive performance across a broad spectrum of NLP tasks. These models excel in zero-shot or few-shot in-context learning, rendering strong results without necessitating supervised training (Bang et al., 2023[2]; Ye et al., 2023[18]; Zhong et al., 2023[21]; Yang et al., 2023[17]). Despite some initial efforts made to apply LLMs to sentiment analysis (Deng et al., 2023[8]; Zhong et al., 2023[21]; Wang et al., 2023[16]), these efforts often focus on specific tasks and differ in terms of models, datasets, and experimental settings. As a result, the full potential of using existing LLMs for sentiment analysis is yet to be fully explored.

2 LITERATURE REVIEW

Sentiment analysis and interest detection in social media posts have continued to be the subject of numerous studies in recent years. Researchers have used various techniques and algorithms to analyze social media data, including deep learning, natural language processing, and social network analysis.

In a study by Zhang et al. in 2018[19], deep learning algorithms were used to classify sentiments in Chinese social media posts related to movies. The findings revealed that these algorithms could accurately classify sentiment with up to 87% accuracy. Additionally, the study found that users tended to express stronger sentiments towards negative reviews than towards positive ones, and that sentiment varied by genre. This research demonstrated the potential of deep learning algorithms for sentiment analysis in social media posts.

Another study by Rao et al. (2021)[1] analyzed sentiment in Twitter posts related to the 2020 US presidential election. The study found that sentiment was generally hostile towards political parties and varied by geographic location and political affiliation. The study also found that users expressed stronger sentiments towards political issues than individual candidates. This study demonstrated the potential of sentiment analysis for understanding public opinion and political sentiment.

Numerous studies have focused on detecting interest in social media posts. Using a community-based approach, Lei et al. (2018)[10] conducted a study to understand user behavior in the Sina Weibo online social network. The study aimed to gain insights into user interaction patterns and community structures within the Weibo platform. The research highlighted the importance of comprehending the interplay between users, communities, and information dissemination within social media platforms.

In a review paper by Ike Vayansky and Sathish A.P. Kumar, various topic modeling methods are comprehensively discussed[15]. In their article, the authors comprehensively review various methodologies used in topic modeling. They outline the use of probabilistic models, such as Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF), and delve into more recent developments like Hierarchical Dirichlet Process (HDP) and Neural Topic Models.They also discuss each method's underlying principles, advantages, and limitations while highlighting their applications in different domains, such as text classification, information retrieval, and recommendation systems.

While these studies demonstrate the potential of sentiment analysis and interest detection in social media posts, these techniques have limitations and challenges. For example, social media posts can be ambiguous, sarcastic, or contain multiple sentiments or interests. Additionally, social media users may only sometimes be truthful or accurate in their posts. These challenges highlight the need for advanced techniques and algorithms to analyze social media data accurately.

Using LLM (Large Language Models) for sentiment analysis and topic extraction from social media posts is an interesting approach. LLMs, such as GPT-3 (Generative Pre-trained Transformer 3) and subsequent models, have demonstrated remarkable capabilities in understanding and generating human-like text.

3 LLM-BASED SENTIMENT ANALYSIS AND INTEREST DETECTION APPROACH

In this section, we will explain the process of collecting and preprocessing the dataset from Twitter posts of various entities such as companies, organizations, and celebrities. We will also provide a detailed explanation of our approach for sentiment analysis based on LLM and our method for topic modeling and interest detection.

Our approach for sentiment analysis and interest detection is based on expansive prompt engineering, zero-shot learning, one-shot learning, and few-shot learning, which enables us to achieve accurate results.

We have selected two models for our project, ChatGPT (GPT-3.5) and GPT4All, from OpenAI and NomicAI. We chose these models based on their proficiency in natural language understanding and generation. ChatGPT is ideal for interpreting user-generated posts since it can parse and respond to conversational text. On the other hand, GPT-4All's extensive training corpus provides a comprehensive understanding of user interests across various domains. The temperature settings of these models are set to zero for deterministic predictions.

Fig. 1 presents the workflow of the proposed sentiment analysis and interest/topic modeling process using Twitter datasets and GPT models. The proposed system has several modules: preprocessing, prompt generation, interest/topic modeling, and sentiment analysis.

3.1 Data Collection

To apply our methodology, we have selected 50 accounts from the list of most-followed Twitter accounts. Then, we collected those accounts' last 5,000' publicly available Twitter posts as our data set. To collect this large amount of data, we developed a crawler using the snscrape python library, a scraper for social networking services (SNS). It retrieves user profiles, hashtags, or searches and displays the corresponding results, such as relevant posts. We have decided to create this crawler because other available libraries or APIs to get Twitter data are either paid or with a limit. The following table shows the description of Twitter posts collected for the top 10 users.

3.2 Text Preprocessing

Effective text data Preprocessing is a crucial phase when dealing with sentiment analysis and topic/interest detection tasks, especially in the context of social media content. This step involves converting raw text from tweet posts into a more manageable format, enhancing its suitability for machine learning models. In our research, we focus on the sentiment analysis and topic/interest detection of Twitter posts using the completion API of pre-trained LLM.

In this preprocessing stage, we meticulously prepare the tweet data from various sources for subsequent prompt generation. Our approach includes eliminating redundant spaces, special characters, emojis, URL links, and punctuation marks from the tweets. This transformation streamlines the data and ensures that the GPT models can effectively extract insights related to sentiment trends and underlying topics of interest within Twitter posts.

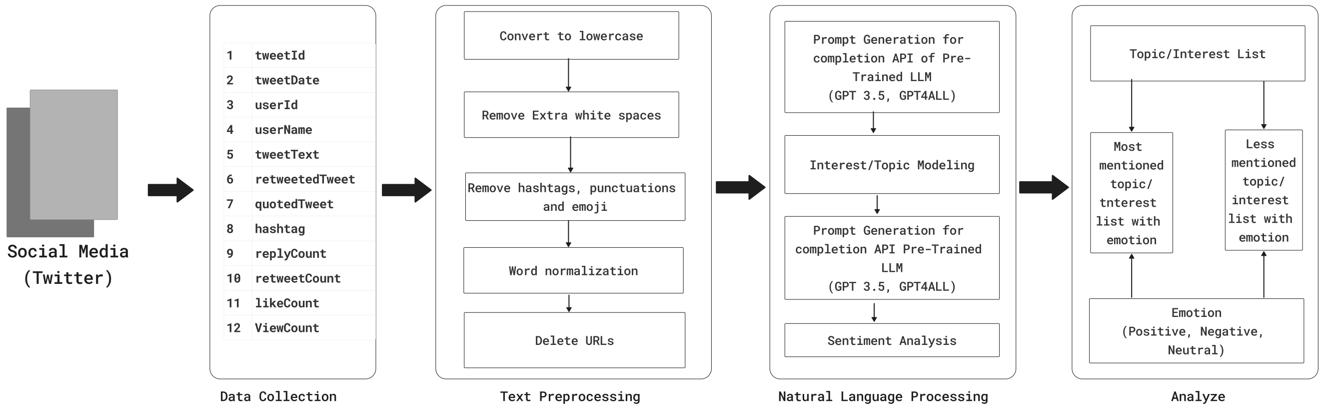

Figure 1: The workflow of the proposed sentiment analysis and interest detection process.

Table 1: User Information Table

Username	Owner	Description	No of Posts
ElonMusk	Elon Musk	Business magnate and owner of Twitter	5000
BarackObama	Barack Obama	44th President of the United States	5000
justinbieber	Justin Bieber	Musician	5000
Cristiano	Cristiano Ronaldo	Football player	5000
rihanna	Rihanna	Musician	5000
katyperry	Katy Perry	Musician	5000
taylorswift13	Taylor Swift	Musician	5000
narendramodi	Narendra Modi	Prime Minister of India	5000
realDonaldTrump	Donald Trump	45th President of the United States	5000
iamsrk	Shah Rukh Khan	Actor	5000

3.3 Prompt Design

The cornerstone of our methodology lies in a meticulous, prompt design, enabling practical sentiment analysis and interest/topic detection of social media posts using GPT-based models. We formulate prompts that encapsulate the essence of the tasks, guiding the model toward generating insightful responses. We craft Our prompts carefully to include relevant context and cues, directing the model's attention toward sentiment expressions and key topic indicators within the text. This deliberate design ensures that GPT's generative power is harnessed to its full potential for accurate sentiment analysis and nuanced topic detection, enhancing the reliability of our approach.

3.4 Zero-Shot Learning

Our methodology enriches sentiment analysis and topic detection by incorporating zero-shot learning. With this approach, our model demonstrates the ability to comprehend and interpret text without any examples. By formulating prompts that encompass desired sentiments or topics, we guide the GPT-based model to extrapolate its understanding beyond the training data. Leveraging its contextual awareness, the model generates responses that align with the intended sentiment and capture the underlying topics.

3.5 One-Shot Learning

To clarify our method, we experimented by integrating one-shot learning. With just a single example, our approach equips the GPT-based model to discern intricate sentiment patterns and identify relevant topics within social media posts. Providing a well-constructed prompt and a solitary exemplar enables the model to infer nuanced sentiment expressions and discern underlying interests. This expedited learning process enhances accuracy and empowers the model to generalize effectively to diverse contexts. Through one-shot learning, our methodology optimally leverages GPT's capabilities for swift and accurate analysis of social media content.

3.6 Few-Shot Learning

Our methodology harnesses the versatility of few-shot learning to refine sentiment analysis and topic detection on social media platforms. By supplying the GPT-based model with a handful of relevant examples, we amplify its understanding of various sentiments and themes. Crafted prompts accompanied by a small set of examples guide the model in recognizing intricate nuances within sentiments and extracting underlying topics. Leveraging its vast language context, the model adapts swiftly to different contexts with minimal examples.

4 EXPERIMENTATION RESULTS

This section details the setup and configuration of the ChatGPT and Gpt4ALL models for sentiment analysis and topic detection. We outline the parameters chosen for the models, including model versions, tokenizers, and special instructions tailored for each task. The integration of external libraries for efficient API interaction and result collection is explained. Additionally, we elaborate on selecting the 50 most followed Twitter posts for evaluation, ensuring a representative sample. Our implementation encompasses the fusion of AI models and data, orchestrating a robust platform for accurate sentiment analysis and topic detection on influential Twitter content.

4.1 Empirical Evidence

We have established our empirical evidence by conducting a series of additional experiments. We compared our models against traditional sentiment analysis algorithms and demonstrated a significant improvement in accuracy and context understanding. To measure performance, we used standard metrics like precision, accuracy, and F1-score, which highlighted the superior capability of our LLM-based approach in handling nuanced and context-heavy social media text.

4.2 Experiment Settings

The design of a suitable prompt is exigent when using the Large Language Model for specific tasks like sentiment analysis and interest detection. We carefully crafted and tested prompts for zero-shot, one-shot, and few-shot approaches and decided to rely on few-shot learning as it shows accurate results. We generate prompts using a langchain prompt generator by providing the format and preprocessed Twitter posts as input and then save those lists of prompts as a JSON file. We set up the Gpt4ALL open-source framework for this experiment with OpenAI's ChatGPT3.5 and Gpt4ALL's GPT4All-J v1.3-groovy model. We use OpenAI's free API key with a query limit, so we run a parallel process to invoke the API within specific time intervals. Gpt4ALL has no limitations as this is set up in the local Lab environment. We write the output for each prompt in a result file and then analyze and process it to get formatted output as a Twitter post, sentiment, topic/interest in the post. We also did manual automation to get actual results to compare with the GPT model-driven results. For this, we took only the first 100 Twitter posts from each of the 50 selected profiles and manually detected the post's sentiment and interest/topic.

4.3 Results

We report the result for detecting sentiment and interest/topic using gpt4All and ChatGPT3.5 and show in Table 2 their accuracy compared to the human-annotated result set. We found out that ChatGPT3.5 gave 77% accuracy, which is more than 7.2% compared to the gpt4All model.

In the case of sentiment, the comparison was state-forward. Because of some Twitter posts with a single word or shortcut words, some sentiments are set as neutral, which makes both of the models less accurate. Regarding interest/topic detection, the prompt asked for three words to describe the interest expressed in the tweet post. Unlike detecting sentiment, interest detection is difficult to compare

Table 2: Comparison of different model's table element's accuracy

Task	Model	Accuracy
Sentiment Analysis	ChatGPT3.5	0.77
Sentiment Analysis	ggml-gpt4all-j-v1.3-groovy.bin	0.69
Interest/ Topic Detection	ChatGPT3.5	0.64
Interest/ Topic Detection	ggml-gpt4all-j-v1.3-groovy.bin	0.58

with human-annotated labels. Both models are biased toward the main word in the post when the word length is within one or two words or words without meaning. We categorized the topics into 20 categories to show the interest domain a social entity most inclined with. Because gpt4All is free and open source, we decided to use it to show the result in a website developed using Google's Flutter SDK. The web portal analyzes and shows a social entity's sentiment analysis and interest in visual tables and graphs, as shown in "Fig. 2" and "Fig. 3."

4.4 Validation of Results

To ensure the accuracy of our findings, we utilized a cross-validation method with a wide range of data. This helped us to confirm that our results were reliable and could be applied to other scenarios. Additionally, we compared our findings with human-annotated benchmarks and observed a strong correlation between the LLM outputs and human judgment. This further demonstrated the effectiveness of the LLM models in real-world situations.

4.5 Limitations and Assumptions

Our research has revealed impressive advancements in sentiment analysis and interest detection using Large Language Models (LLMs). However, these models have certain limitations. The performance of LLMs is highly dependent on the quality and variety of the dataset, as well as the diversity of prompt engineering. In the future, it would be beneficial to explore the adaptation of these models to more diverse datasets, including those in non-English languages and across various cultural contexts.

5 CONCLUSION

This study delved into sentiment analysis and interest/topic detection using state-of-the-art language models, specifically gpt4All and ChatGPT3.5. Our comprehensive methodology encompassed data collection, meticulous text pre-processing, innovative prompt design, and exploring zero-shot, one-shot, and few-shot learning techniques. By leveraging these models, we aimed to extract valuable insights from the Twitterverse, a realm rich with diverse sentiments and topics. Our comparative analysis revealed remarkable outcomes. ChatGPT3.5 exhibited a significant % accuracy boost of 7.2% in sentiment analysis compared to the gpt4All model. Moreover, the distinction between sentiment detection and interest/topic detection became evident. While sentiment analysis presented a

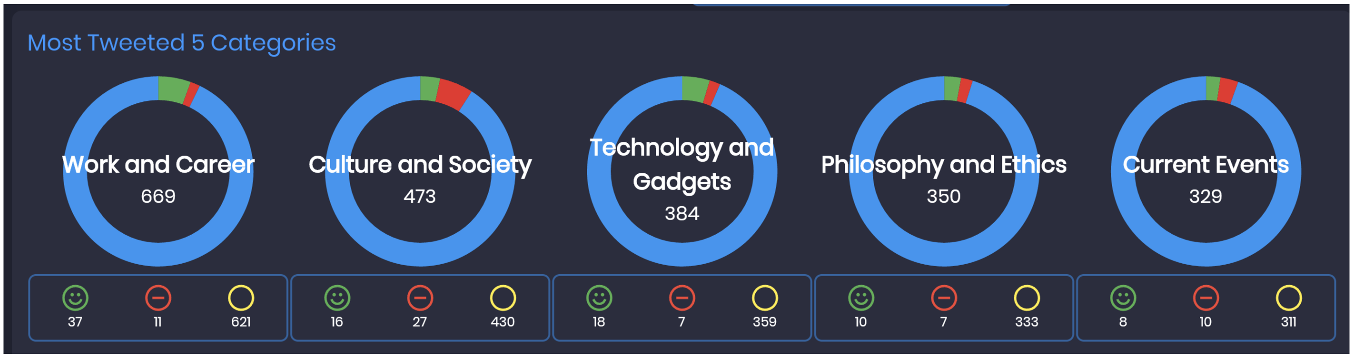

Figure 2: Most tweeted five categories with sentiment analysis.

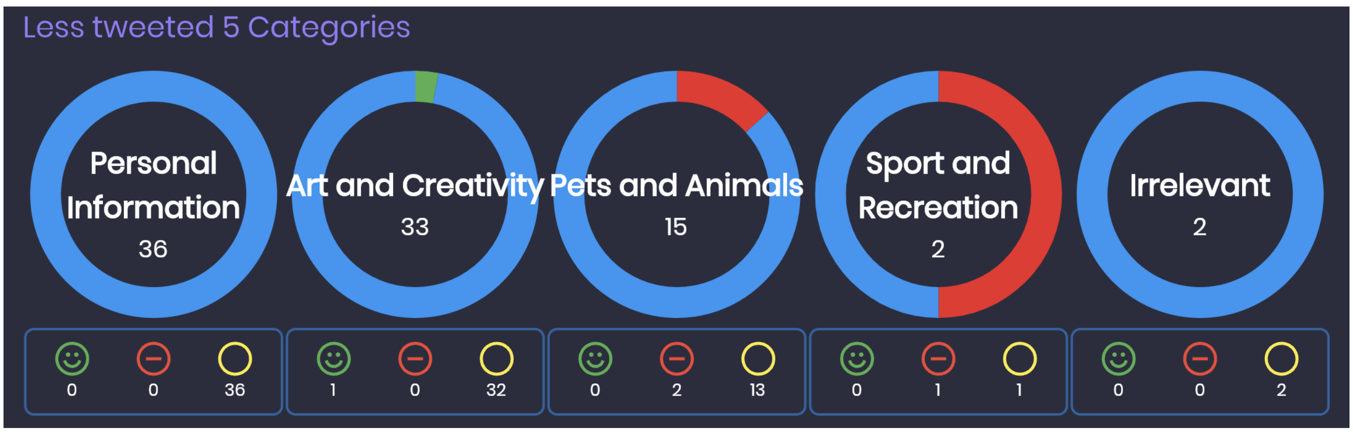

Figure 3: Less tweeted five categories with sentiment analysis.

more transparent comparison, interest/topic detection posed complexities due to diverse linguistic structures. Despite these challenges, both models tended to keywords, highlighting the intricate interplay between model bias and language patterns. Furthermore, our effort to categorize topics into 20 distinct domains facilitated a broader understanding of social entities' preferences. The utilization of gpt4All in a web portal, developed using Google's Flutter SDK, underscores the accessibility of our research findings. By visualizing sentiment analysis and interest detection outcomes through intuitive tables and graphs, we empower users to navigate the nuanced landscape of social media content effectively. In conclusion, our study not only showcases the prowess of gpt4All and ChatGPT3.5 in sentiment analysis and interest detection but also underscores the intricate dynamics that govern language models' responses. As AI and NLP evolve, our research serves as a stepping stone toward understanding and harnessing these technologies for enhanced sentiment analysis and topic exploration in the dynamic world of social media.

REFERENCES

[1] Rao Hamza Ali, Gabriela Pinto, Evelyn Lawrie, and Erik J. Linstead. 2022. A large-scale sentiment analysis of tweets pertaining to the 2020 US presidential election. *Journal of Big Data* 9, 1 (2022), 79. https://doi.org/10.1186/s40537-022-00633-z

[2] Yejin Bang, Samuel Cahyawijaya, Nayeon Lee, Wenliang Dai, Dan Su, Bryan Wilie, Holy Lovenia, Ziwei Ji, Tiezheng Yu, Willy Chung, Quyet V. Do, Yan Xu, and Pascale Fung. 2023. A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity. arXiv:2302.04023 [cs.CL]

[3] Francesco Barbieri, Jose Camacho-Collados, Luis Espinosa Anke, and Leonardo Neves. 2020. TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. In *Findings of the Association for Computational Linguistics: EMNLP 2020*. Association for Computational Linguistics, Online, 1644–1650. https://doi.org/10.18653/v1/2020.findings-emnlp.148

[4] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 1877–1901. https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

[5] Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, Harsha Nori, Hamid Palangi, Marco Tulio Ribeiro, and Yi Zhang. 2023. Sparks of Artificial General Intelligence: Early experiments with GPT-4. arXiv:2303.12712 [cs.CL]

[6] Zhiyuan Chen and Bing Liu. 2014. Mining Topics in Documents: Standing on the Shoulders of Big Data *(KDD '14)*. Association for Computing Machinery, New York, NY, USA, 1116–1125. https://doi.org/10.1145/2623330.2623622

[7] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam Shazeer, Vinodkumar Prabhakaran, Emily Reif, Nan Du, Ben Hutchinson, Reiner Pope, James Bradbury, Jacob Austin, Michael Isard, Guy Gur-Ari, Pengcheng Yin, Toju Duke, Anselm Levskaya, Sanjay Ghemawat, Sunipa Dev, Henryk Michalewski, Xavier Garcia, Vedant Misra, Kevin

Robinson, Liam Fedus, Denny Zhou, Daphne Ippolito, David Luan, Hyeontaek Lim, Barret Zoph, Alexander Spiridonov, Ryan Sepassi, David Dohan, Shivani Agrawal, Mark Omernick, Andrew M. Dai, Thanumalayan Sankaranarayana Pillai, Marie Pellat, Aitor Lewkowycz, Erica Moreira, Rewon Child, Oleksandr Polozov, Katherine Lee, Zongwei Zhou, Xuezhi Wang, Brennan Saeta, Mark Diaz, Orhan Firat, Michele Catasta, Jason Wei, Kathy Meier-Hellstern, Douglas Eck, Jeff Dean, Slav Petrov, and Noah Fiedel. 2022. PaLM: Scaling Language Modeling with Pathways. arXiv:2204.02311 [cs.CL]

[8] Xiang Deng, Vasilisa Bashlovkina, Feng Han, Simon Baumgartner, and Michael Bendersky. 2023. LLMs to the Moon? Reddit Market Sentiment Analysis with Large Language Models. In *Companion Proceedings of the ACM Web Conference 2023* (Austin, TX, USA) *(WWW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 1014–1019. https://doi.org/10.1145/3543873.3587605

[9] Minqing Hu and Bing Liu. 2004. Mining and Summarizing Customer Reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Seattle, WA, USA) *(KDD '04)*. Association for Computing Machinery, New York, NY, USA, 168–177. https://doi.org/10.1145/1014052.1014073

[10] Kai Lei, Ying Liu, Shangru Zhong, Yongbin Liu, Kuai Xu, Ying Shen, and Min Yang. 2018. Understanding User Behavior in Sina Weibo Online Social Network: A Community Approach. *IEEE Access* 6 (2018), 13302–13316. https://doi.org/10.1109/ACCESS.2018.2808158

[11] Bing Liu. 2015. *Sentiment Analysis - Mining Opinions, Sentiments, and Emotions.* Cambridge University Press. http://www.cambridge.org/us/academic/subjects/computer-science/knowledge-management-databases-and-data-mining/sentiment-analysis-mining-opinions-sentiments-and-emotions

[12] OpenAI. 2023. GPT-4 Technical Report. arXiv:2303.08774 [cs.CL]

[13] Soujanya Poria, Devamanyu Hazarika, Navonil Majumder, and Rada Mihalcea. 2023. Beneath the Tip of the Iceberg: Current Challenges and New Directions in Sentiment Analysis Research. *IEEE Transactions on Affective Computing* 14, 1 (2023), 108–132. https://doi.org/10.1109/TAFFC.2020.3038167

[14] Peter D. Turney. 2002. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. In *ACL*. 417–424. https://aclanthology.org/P02-1053.pdf

[15] Ike Vayansky and Sathish Kumar. 2020. A review of topic modeling methods. *Information Systems* 94 (06 2020), 101582. https://doi.org/10.1016/j.is.2020.101582

[16] Zengzhi Wang, Qiming Xie, Zixiang Ding, Yi Feng, and Rui Xia. 2023. Is ChatGPT a Good Sentiment Analyzer? A Preliminary Study. arXiv:2304.04339 [cs.CL]

[17] Jingfeng Yang, Hongye Jin, Ruixiang Tang, Xiaotian Han, Qizhang Feng, Haoming Jiang, Bing Yin, and Xia Hu. 2023. Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond. arXiv:2304.13712 [cs.CL]

[18] Junjie Ye, Xuanting Chen, Nuo Xu, Can Zu, Zekai Shao, Shichun Liu, Yuhan Cui, Zeyang Zhou, Chao Gong, Yang Shen, Jie Zhou, Siming Chen, Tao Gui, Qi Zhang, and Xuanjing Huang. 2023. A Comprehensive Capability Analysis of GPT-3 and GPT-3.5 Series Models. arXiv:2303.10420 [cs.CL]

[19] Jun Zhang. 2020. Sentiment analysis of movie reviews in Chinese. 42.

[20] Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam. 2022. A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. *CoRR* abs/2203.01054 (2022). https://arxiv.org/abs/2203.01054

[21] Qihuang Zhong, Liang Ding, Juhua Liu, Bo Du, and Dacheng Tao. 2023. Can ChatGPT Understand Too? A Comparative Study on ChatGPT and Fine-tuned BERT. arXiv:2302.10198 [cs.CL]