Review article

Don Bosco Institute of
Technology Delhi Journal of Research

Year 2024, Volume-1, Issue-1 (January - June)

www.acspublisher.com

Sentiment Analysis with ChatGPT Across Domains

Divya Verma

Assistant Professor, Sri Guru Tegh Bahadur Institute of Management & Information Technology, GGSIPU

ARTICLE INFO

Key Words: ChatGPT, Sentiment Analysis, Topic Model, Public Opinion, SWOT Analysis, Scientific Citations, App Reviews, Large Language Models, Multilingual, Artificial Intelligence.

doi: 10.48165/ dbitdjr.2024.1.01.04

ABSTRACT

This paper explores ChatGPT's applications in sentiment analysis across various domains, including business, Reddit data, and scientific citations. It demonstrates ChatGPT's utility in understanding customer needs and public opinion, as well as its ability to identify nuanced sentiment and biases in scholarly research evaluation. Analysis of app reviews on the Google Play Store shows predominantly positive sentiments, with models like Random Forest and SVM achieving high effectiveness. The research evaluates ChatGPT and other large language models such as Gemini and LLaMA for multilingual sentiment analysis, revealing their proficiency alongside biases and inconsistencies across languages. The study emphasizes the importance of standardized evaluation methodologies and the need for data and algorithm improvements to enhance ChatGPT's performance and applicability in sentiment analysis.

Introduction

The relentless flow of digital information and communication in today's era necessitates a thorough understanding and harnessing of human sentiments across a variety of applications. Sentiment analysis, also known as opinion mining, sentiment classification, or emotion AI, is a field that intersects Natural Language Processing (NLP), Machine Learning (ML), and computational linguistics. It encompasses techniques for the automated identification and analysis of emotions, opinions, and attitudes expressed in textual data. The primary goal of sentiment analysis is to determine the sentiment or emotional tone of a given piece of text, whether positive, negative, or neutral. This process involves deciphering the complex web of emotions, opinions, and attitudes within text data, providing insight into the latent sentiments of individuals, groups, or societies. This technology has found widespread use across various industries, including marketing, customer feedback analysis, social media monitoring, and product reviews. By offering businesses and organizations insight into public opinion, sentiment analysis supports informed decision-making processes. The advent of Large Language Models (LLMs) has revolutionized sentiment analysis capabilities. These potent NLP systems undergo training on extensive datasets to comprehend and produce human-like language. Models such as ChatGPT 3.5, ChatGPT 4, and others boast neural networks with hundreds of millions to billions of parameters, excelling at capturing intricate linguistic patterns and dependencies. Through

*Corresponding author.
E-mail address: divyaverma.sgtbimit@gmail.com (Divya Verma)
Received 22.08.2024; Accepted 10.09.2024

techniques like Reinforcement Learning from Human Feedback (RLHF), LLMs gain contextual understanding, empowering them to generate coherent and contextually appropriate responses across domains such as chatbots, virtual assistants, content generation, language translation, and sentiment analysis. In this research paper, we examine the capabilities and limitations of sentiment analysis models across diverse languages and scenarios, with a focus on state-of-the-art

LLMs such as ChatGPT 3.5, ChatGPT 4, Gemini Pro, and LLaMA2 7b. By testing these models on various sentiment analysis tasks in multiple languages, we aim to provide insights into the generalizability and robustness of these models in cross-cultural contexts. This comprehensive approach allows us to uncover potential variations or biases in LLM performance, contributing to the refinement and enhancement of sentiment analysis applications across different linguistic landscapes. In summary, our research investigates the performance and limitations of sentiment analysis models under real-world and cross-cultural conditions. The findings of this study are expected to contribute to the advancement of sentiment analysis applications in diverse settings, addressing linguistic and contextual variations that pose significant challenges. Our paper provides a comparative evaluation of LLMs across multiple languages, a verification of evaluations against a human benchmark, and an examination of safety filters and inherent biases in the models

Theory and Methods

This study explores the application of ChatGPT in improving customer sentiment analysis for businesses. The research used secondary data sources, conducting comprehensive searches in digital media and scientific databases using relevant keywords. The focus was on academic literature, including journal articles and essays published after 2015[1]. The study revolves around the exploration of public sentiment and response to ChatGPT, a prominent AI-based conversational agent. The public's opinion plays a crucial role in guiding ChatGPT's development and its application in various fields[2]. Sentiment analysis was applied to determine public sentiment towards ChatGPT across various platforms such as Reddit. Techniques such as topic model and Latent Dirichlet Allocation (LDA) were utilized to identify key topics and trends within the data, offering insights into public perception and concerns[2][4]. Knowledge mapping helped in visualizing data and understanding correlations and patterns related to ChatGPT usage and opinions. By integrating probability-based models, the research enhanced traditional methods, allowing for nuanced examination of data[2].For sentiment analysis of citations in scientific articles using ChatGPT, researchers can use automated sentiment classification to classify citations as positive, neutral, or negative, handling large volumes efficiently. Contextual analysis refines interpretation by categorizing citations based on their location in an article. ChatGPT's nuanced interpretation assesses language patterns and tone for accurate endorsements, acknowledgments, and criticisms. Temporal analysis tracks shifts in perception over time, while knowledge mapping reveals patterns across fields, highlighting potential biases and conflicts of interest. Data visualization supports ChatGPT's analysis by representing citation sentiments to identify trends and concerns[3][1].A literature review regarding ChatGPT combines both alarmist and enthusiastic perspectives sourced from both scholarly articles and news outlets. Recent literature on ChatGPT can be divided into these categories: • Evaluating ChatGPT's Performance • Distinguishing ChatGPT-Generated Work from Human-Generated Work: Methods and Insights • Exploring the Impacts of ChatGPT • User Opinion Analysis of ChatGPT on Social Media[4][2]. The findings reveal major distinctions between humans and ChatGPT: 1) ChatGPT's responses remain closely tied to the given question, whereas humans often digress and explore related topics. 2) ChatGPT provides objective answers, while humans tend to express subjective opinions. 3) ChatGPT produces safer, balanced, and neutral texts, whereas humans offer more specific information, drawing from diverse sources and including citations. 4) ChatGPT's responses maintain a formal tone, while humans employ colloquial language, abbreviations, slang, humour, and examples. 5) ChatGPT exhibits less emotional expression, relying on logical flow and conjunctions, whereas humans employ punctuation, grammar features, and brackets to convey emotions and explanations[4]. We interrogate a large twitter corpus (n = 4,251,662) of all publicly available English-language tweets containing the ChatGPT keyword. Our first research aim utilizes a prominent peaks model (upper-quartile significance threshold of prominence>20,000). Our second research aim utilized sentiment analysis to identify, week-on-week, highestfrequency negative, and positive keywords and emojis[5].This study conducts sentiment analysis on ChatGPT app reviews from the Google Play Store using three classification models: Random Forest, Support Vector Machine (SVM), and Naïve Bayes. Data were collected through scraping techniques from the app reviews, yielding a dataset of 2,652 reviews from July 28, 2023, to January 28, 2024. The data were preprocess and label for classification. Text preprocessing involved cleaning the reviews, removing stop words, and stemming. Term weighting was conducted using the Term Frequency-Inverse Document Frequency (TF-IDF) method to convert text into numerical values representing importance and significance within the context. The classification models were implemented to categorize reviews as positive, negative, or neutral. Random Forest leverages decision trees and ensemble learning for classification, while SVM uses hyperplane separation for categorization. Naïve Bayes applies probabilistic model based on Bayes' theorem .The models' performance was evaluated using metrics such as accuracy, precision, recall, and F1-score. This methodology aims to assess the sentiment of app reviews and compare the effectiveness of different classification models in sentiment analysis[7]. Methodology To evaluate the performance of ChatGPT, Gemini, and LLaMA on multilingual sentiment analysis, the study conducts a comprehensive comparative analysis across 20 carefully curated scenarios that encapsulate a wide range of emotions from joy to frustration. Each scenario is translated into 10 different languages to assess the language models' capabilities in sentiment analysis across diverse linguistic landscapes[8] Data Collection and Preparation: The study uses 20 distinct scenarios that span various emotional spectrums and translates them into 10 different languages. The languages cover a wide range of linguistic families to test the models' performance in multilingual sentiment analysis .Models Evaluation: The study focuses on three large language models (LLMs): ChatGPT (versions 3.5 and 4), Gemini (versions Pro and Ultra), and LLaMA (versions 1 and 2). Each model is evaluated based on its ability to accurately determine sentiment in each of the 20 scenarios across the 10 languages. Analysis and Metrics: Sentiment analysis is performed using each model, and the results are compared based on several key metrics, including accuracy, precision, recall, and F1-score. The models' performance is analysed for each language and scenario, allowing for a comprehensive assessment of their multilingual capabilities .Testing Framework: The testing framework involves running each scenario through the models and capturing their responses. The responses are then classified into positive, negative, or neutral sentiments. The models' performances are assessed based on their ability to accurately classify the sentiments and their consistency across languages[8] Comparison and Interpretation: The study compares the results from ChatGPT, Gemini, and LLaMA to identify strengths and weaknesses in their multilingual sentiment analysis performance. Differences in the models' responses across languages and scenarios are analysed to understand their respective capabilities and limitations. Insights and Implications: Based on the findings, the study provides insights into the strengths and weaknesses of each model in multilingual sentiment analysis. This analysis informs future development and refinement of LLMs in the context of multilingual sentiment analysis, highlighting areas for improvement and potential applications in various domains[8]

Results and Discussion

In exploring the application of ChatGPT for customer sentiment analysis, the research finds that ChatGPT's contextual understanding and automated processing capabilities enable businesses to efficiently analyse large volumes of customer feedback in real-time, identifying trends and issues to respond swiftly. This aligns with ChatGPT's role in citation sentiment analysis, where the model excels at capturing nuanced relationships between words and phrases, allowing for accurate assessment of sentiment in citations across various scientific articles. In both contexts, the transformer-based architecture of ChatGPT and its comprehensive training on diverse data empower the model to analyse sentiment efficiently, handling large volumes of data with speed and precision. This leads to informed decision-making in business settings, as well as deep insights into scholarly impact[1]. Challenges in Citation Analysis[3]

Challenges in citation analysis are prevalent due to the diversity, subjectivity, volume, evolution, and contextuality of citations. Sentiments within citations can vary greatly, from positive to neutral to negative, and interpreting these sentiments can be subjective, leading to potential discrepancies in analysis. Moreover, the large volume of citations within scholarly articles makes comprehensive sentiment analysis a challenging task .Another challenge arises from the evolving nature of scientific knowledge and perspectives, leading to shifts in how citations are interpreted over time. This can affect the perceived impact of older works as newer research emerges. Defending ChatGPT in Citation Analysis: Despite these challenges, leveraging ChatGPT's advanced language processing capabilities can provide efficient and accurate analysis of citation sentiments. ChatGPT's ability to understand contextual shifts and domainspecific expressions, combined with its transformer architecture, allows it to handle the complexities of citation analysis. By identifying and understanding both positive and negative sentiments in citations, ChatGPT aids in offering nuanced insights into a researcher's impact within their field. Furthermore, ChatGPT's speed and scalability in processing large volumes of citations enhance its effectiveness in scholarly analysis. Its comprehensive

Table 1: Examples of positive and negative citations with domain.

Domain	Sentiment	Example Sentences
Computer Science	Positive	"The algorithm proposed by Smith et al. (2022) significantly improves efficiency in data processing tasks." "Jones' software development methodology sets a new standard for code quality and reliability."
	Negative	"Contrary to the findings of Johnson et al. (2021), recent studies have identified vulnerabilities in their security model." "The limitations of Brown's computational model highlight the need for more robust error handling mechanisms."
Architecture	Positive	"The innovative design approach by Smith and team (2024) showcases a harmonious blend of functionality and aesthetics." "Jones' sustainable architecture principles demonstrate a commitment to environmental conservation."
	Negative	"Despite the acclaim, Johnson's architectural design lacks practicality and usability." "Brown's architectural concept has been criticised for its lack of structural integrity."
English Literature	Positive	"Smith's literary analysis sheds new light on the complexities of Shakespearean tragedies." "Jones' critical review of modern poetry enriches our understanding of contemporary literary trends."
	Negative	"Contrary to prevailing opinions, Johnson's literary interpretation overlooks crucial thematic elements." "The limitations of Brown's literary critique lie in its superficial examination of character motivations."
Physics	Positive	"The experimental findings of Smith et al. (2020) confirm the existence of quantum entanglement at macroscopic scales." "Jones' theoretical framework provides a novel perspective on the unified theory of particle physics."
	Negative	"Johnson's theoretical model fails to account for observed phenomena in quantum mechanics." "The conclusions drawn by Brown et al. (2018) contradict established principles of thermodynamics."
Chemistry	Positive	"The synthesis method developed by Smith and colleagues (2024) offers a sustainable solution for chemical waste reduction." "Jones' research on catalyst efficiency has wide-ranging applications in industrial processes."
	Negative	"Despite initial promise, Johnson's chemical reaction mechanism lacks reproducibility in experimental settings." "Brown's chemical analysis overlooks key factors impacting reaction kinetics."
Electrical Engineering	Positive	"The innovative circuit design by Smith et al. (2020) significantly improves power efficiency in electronic devices." "Jones' research on renewable energy sources contributes to advancements in sustainable technologies."
	Negative	"Johnson's electrical circuit design is prone to voltage fluctuations under varying load conditions." "The limitations of Brown's electrical system design result in inefficiencies and overheating."
Biology	Positive	"The discovery by Smith's research team (2020) offers new insights into genetic mechanisms underlying disease progression." "Jones' experimental methodology in cell biology sets a benchmark for precision and reproducibility."
	Negative	"Contrary to expectations, Johnson's experimental results show no significant impact on cellular function." "The limitations of Brown's biological model restrict its applicability to real-world scenarios."
Medicine	Positive	"The clinical trial led by Smith et al. (2020) demonstrates promising results in cancer treatment efficacy." "Jones' medical research on public health interventions has positive implications for population health."
	Negative	"Johnson's clinical study methodology lacks adequate control measures, raising questions about result validity." "Brown's medical diagnosis algorithm exhibits high false-positive rates, impacting diagnostic accuracy."

training on vast amounts of scientific literature enables it to recognize patterns and associations specific to different research fields, offering a more precise and informed evaluation of citation sentiments ChatGPT possesses the capability to comprehend and . Analytical Approaches for Positive and Negative effectively react to human language, thereby enabling Citations[3] Analysing positive and negative citations involves commercial enterprises to conduct large-scale and identifying key phrases and expressions indicative of specific instantaneous analysis of customer sentiment. Through the sentiments. For positive citations, keywords such as utilisation of ChatGPT's capacity to discern customer "groundbreaking" and "remarkable findings" often signify sentiment, enterprises can detect concerns, quantify levels of praise, while negative citations may include terms like customer contentment, and acquire profound understandings "limitations" and "inconsistencies." By leveraging both into customer predilections and requirements. It is crucial to keywords and contextual understanding, ChatGPT can integrate both human perspective and expertise in the effectively discern the sentiment expressed in citations, aiding business domain when analysing ChatGPT data and making researchers in assessing their impact and potential areas for well informed decisions.[1] improvement. Table 1 summarises the examples of positive and negative citations, showcasing the different sentiments conveyed in scholarly writing in many domains including computer science, architecture, physics chemistry, etc.[3]

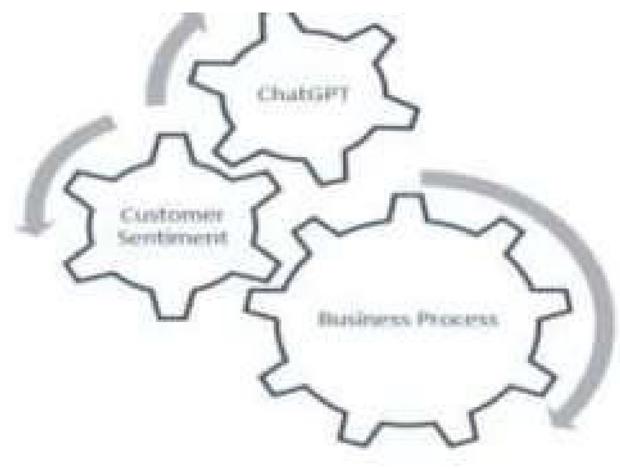

Figure 2. Correlation between ChatGPT, customer sentiment, and business process

The results and discussion integrate the findings from the study of user responses to ChatGPT with the evaluation of ChatGPT's performance, distinguishing ChatGPT-generated work from human-generated work, and exploring the impacts of ChatGPT across various domains[2][4]ChatGPT's Performance Research shows that ChatGPT performs well in a range of natural language processing tasks, including reasoning and dialogue. For instance, Kung et al. [6] demonstrated ChatGPT's ability to perform with high accuracy on the United States Medical Licensing Examinations (USMLE), suggesting potential applications in medical education and decision-making. This performance reflects ChatGPT's proficiency in various domains, including translation tasks, where it competes well with commercial products, especially for high-resource languages. [2][4]However, the limitations of ChatGPT are evident in specific tasks such as sequence tagging and machine translations, where it may generate lengthy summaries or incorrect translations. These nuances highlight the need for continuous improvements and cautious application of ChatGPT in critical tasks .User Responses and Public Perception[4]User opinion analysis from social media platforms such as Weibo and Bilibili indicates a range of sentiments toward ChatGPT. Most users express positive views regarding ChatGPT's capabilities, such as its ability to engage in conversations and generate creative content. However, concerns regarding its limitations and potential misuse, particularly in education, are prevalent[4][2] Studies on opinion mining from social media posts found that users were optimistic about ChatGPT's potential in various domains like software development, entertainment, and creativity. Yet, other studies point out that concerns persist about ChatGPT's impact on academic integrity and the potential decline in higher-order cognitive skills .Distinguishing ChatGPTGenerated Work from Human-Generated Work[4]

Methods for detecting AI-generated content have emerged, such as the ChatGPT Detector and Detect GPT, which effectively identify patterns in AI-generated responses. These approaches help maintain the integrity of academic and professional work by distinguishing AI-generated content from human-generated work. The challenge lies in the potential for ChatGPT to produce responses that closely mimic human writing. As a result, researchers suggest adjusting educational and professional evaluation methods to account for AI's influence, such as using oral exams or advanced proctoring techniques.[4] Impact of ChatGPT in Education and Beyond

The discussion extends to the impact of ChatGPT in education, where there is debate on its use for student assignments. Some argue for outright bans due to concerns about academic integrity, while others advocate for guidelines to integrate AI tools responsibly in teaching and learning.

The impact of ChatGPT on professional careers, particularly in knowledge-centric fields, remains a topic of debate. While ChatGPT offers opportunities for efficient work, questions arise about the reliability of its outputs and potential risks associated with over-reliance on AI[4]. The study examining early public discourse surrounding ChatGPT across 4 million tweets and structured interviews with marketing major students highlights the evolving perceptions of Generative Artificial Intelligence (GenAI) in marketing and education. The analysis revealed a complex interplay between the opportunities and challenges of using GenAI tools like ChatGPT, reflecting a nuanced understanding of its potential impact on various domains. [5][6]Negativity surrounding ChatGPT and GenAI center on concerns about credibility, ethical usage, and over-reliance on AI in tasks traditionally performed by humans. Users and students highlighted ethical issues such as implicit bias in AI responses, environmental concerns, and employment rights of data annotators and programmers. Additionally, job displacement and the risk of ethical violations were mentioned as potential downsides of using GenAI.[5]However, there was notable positivity in early public discourse and students' attitudes toward GenAI, particularly in the context of college education and professional domains. Users and students recognized the potential for task automation, creativity enhancement, and personalized marketing strategies enabled by GenAI tools. ChatGPT's application in coding, content creation, education, and personal productivity was viewed favorable, contributing to a generally positive sentiment.[6][5]Topical model further highlighted key themes such as the need for a strategic understanding of GenAI tools, balancing human intuition with GenAI efficiency, and emphasizing ethical training and transparency. Students' diverse perspectives also underscored the necessity for academic programs to adapt to the complexities of GenAI, preparing career-ready students for the marketing field's future demands.[6].Limitations and future directions While ChatGPT offers significant potential in enhancing various aspects of academic publishing, it is important to acknowledge its limitations and explore future directions for improvement and innovation. • Model limitations: ChatGPT's performance in sentiment analysis may be impacted by biases in the training data and the model's reliance on pre-existing knowledge. Future research should focus on mitigating these biases and improving the model's ability to handle diverse contexts and terminology. • Contextual understanding: One of the challenges faced by ChatGPT is its limited ability to grasp nuanced context, leading to potential misinterpretations in sentiment analysis. Future developments should aim to enhance ChatGPT's contextual understanding, especially in domain-specific language and academic discourse. • Ethical considerations: The integration of AI tools like ChatGPT in decision-making processes raises ethical considerations, including transparency, accountability, and potential biases

In the initial phase, the study began with a dataset of 2652 review data collected from July 28, 2023, to January 28, 2024. [7]

Fig. 2: Dataset Collection

	content	score	cleaned_text	label	content_len	punct
2651	Menjadi yang pertama itu sangatlah baik	5	menjadi yang pertama itu sangatlah baik	1	34	0.0
2650	Next geoglee in blockchain .... Bano cryto	5	next geoglee in blockchain bano cryto ...	1	46	10.9
2649	Nice app	5	nice app	1	7	0.0
2648	Bagus sekali	5	bagus sekali	1	17	0.0
2647	Menurut saya aplikasi ini sudah oke untuk pemb...	5	menurut saya aplikasi ini sudah oke untuk pemb...	1	60	0.0

Fig. 3: Gambar 3 Data Cleansing and Case Folding

After successfully collecting the dataset, the next step is to clean the data, such as removing emojis, numbers, and punctuation marks and changing uppercase letters to lowercase.

Fig. 4: Tokenizing a Dataset [7]

The final stage is removing words that have no effect and the removal of affixes in words.

Fig. 5: Stopword Removal and Stemming [7]

Researchers conducted sentiment analysis using Google Colab and Python programming language. The study was conducted on 2652 data, with 2326 data labelled as positive (87.71%) and 326 as negative (12.29%). Researchers divided the data into training data as much as 70% (1,856 reviews) and testing data as much as 30% (796 reviews). [7]

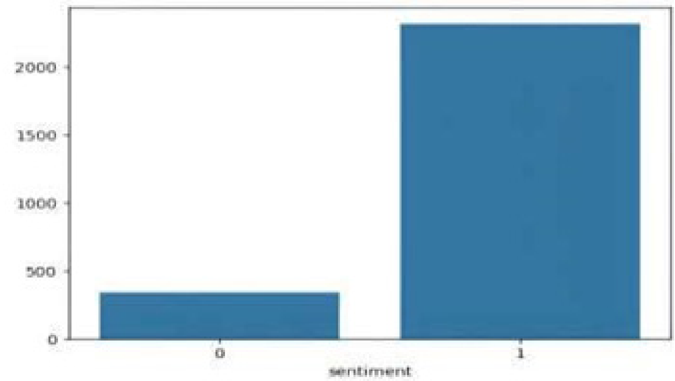

Fig: 6. Sentiment chart before in SMOTE [7]

An imbalance in the amount of data between positive reviews and negative reviews can result in an imbalance of data that can lead to errors in classifying minority classes that tend to be majority classes. Therefore, researchers use oversampling to balance data by adding data in minority classes. One of the oversampling methods used is the Synthetic Minority Oversampling Technique (SMOTE), which deals with unbalanced data problems or overfitting problems (Utami, 2022).

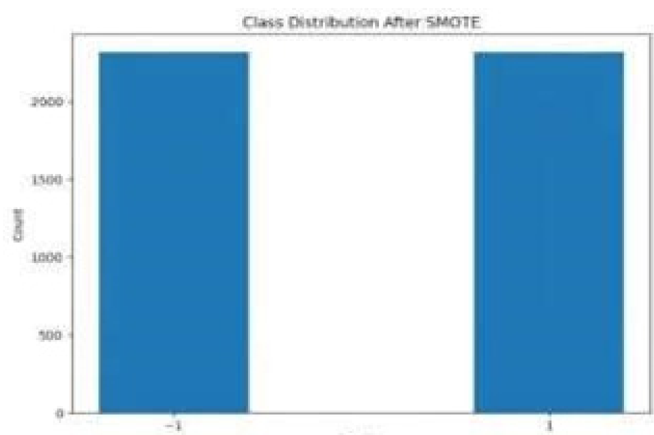

Fig. 7: Sentiment Chart after in SMOTE

Figure 7 is a form of the dataset that has been in SMOTE by obtaining a balanced amount of data, namely 2326 positive and 2326 negative data. After that, researchers enter the data into each classification algorithm and get the following results.

Random Forest

Researchers used the Scikit-Learn library to apply Random Forest classification to data. The analysis showed that out of 796 cases predicted to be positive, 696 (or 87.43%) were completely positive (True Positive), indicating that the model had high accuracy in identifying positive cases. In addition, of the 796 instances predicted negative, 31 (or 3.89%) were negative (True Negative), illustrating the model's ability to identify negative instances correctly. On the other hand, 63 cases (7.91%) were incorrectly predicted as positive when, in fact, they were negative (False Positive), indicating an error in classifying these cases. In comparison, only 6 cases (0.75%) were incorrectly predicted as negative when they were positive (False Negative), indicating that the model may tend to ignore some positive cases (Oktavia, Ramadahan, & Minarto, 2023).

Table 2: Results with Random Forest

	Precision	Recall
Negative	0.84	0.33
Positive	0.92	0.99
	Accuracy: 0.91	
	F1-Score: 0.90	

Table 2[7] shows the results of classification using the random forest algorithm. 91% accuracy indicates how well the model classifies all data correctly. The 90% F1score is a combined measure of precision and recall, with a precision class negative of 84% and a precision class positive of 92%, describing the model's accuracy in classifying each class. Meanwhile, recall class negative reached 33% and recall class positive reached 99%, indicating the model's ability to identify negative and positive sentiments specifically. These results show that the recall value of the positive class is much higher than that of the recall class negative, indicating that the model is superior in recognising and classifying positive sentiment (Dey, Chakraborty, Biswas, Bose, & Tiwari, 2016).[7]

Conclusion

The analysis of public discourse and sentiment surrounding ChatGPT across various online platforms reveals a complex and multifaceted view of this emerging technology. ChatGPT's capabilities, particularly in generating context-aware conversation and producing code, have been met with enthusiasm, as they hold significant potential for applications in software development, entertainment, and education. Many envision ChatGPT as a catalyst for advancing AIdriven solutions that simplify daily tasks and enhance productivity.[2][4]. However, there are notable concerns regarding the potential adverse effects of ChatGPT, especially in the education sector and job market. Critics question the ethical use of data by AI companies, as well as the risk of over-reliance on AI, which could lead to job displacement and affect the integrity of academic work. Issues such as the creation of ChatGPT's alter ego, "DAN," raise questions about the ethical boundaries of AI experimentation.[4]To address these challenges, future efforts should focus on establishing clear policies and guidelines for the ethical use of ChatGPT and other AI technologies, particularly in areas such as data collection and AI misuse. Policymakers should strive to transform potential threats into opportunities by creating job opportunities where AI influence is limited and encouraging the development of AI-driven tools in various domains[2] Despite its challenges, ChatGPT represents a significant technological advancement, and many users and institutions have begun to embrace its potential. To maximize its benefits while minimizing risks, collaboration between technology developers, regulators, educators, and other stakeholders is essential. By doing so, the broader public can reap the numerous benefits of ChatGPT and other GenAI technologies while ensuring responsible and ethical use in the future[2][4]. The integration of ChatGPT in business and marketing education offers significant advantages, as it can improve customer sentiment analysis and help students understand customer needs and preferences more effectively. By leveraging ChatGPT's natural language understanding and realtime data analysis capabilities, businesses can gain valuable insights for better decision-making[1][6]For marketing students, exposure to ChatGPT and its applications in sentiment analysis and customer engagement provides practical experience and prepares them for the complexities of the marketing field. By incorporating a strategic approach to GenAI applications in education, students can learn how to maximize the benefits of these tools while also understanding the ethical considerations and limitations involved[6]To fully harness ChatGPT's potential in business, companies should follow best practices such as collecting and organizing diverse customer data, training the model with appropriate datasets, and validating results using other approaches. Combining ChatGPT analysis with human judgment and business knowledge can enhance decision-making and improve overall outcomes[1] ChatGPT can play a crucial role in enhancing business and marketing education by providing students with hands-on experience in using AI for customer sentiment analysis. This exposure helps bridge gaps in marketing education, ensuring students are well-equipped for future challenges in the industry. By responsibly leveraging ChatGPT, businesses and marketing students can unlock new opportunities and achieve greater success in understanding and responding to customer needs[1][6]. For software applications like marketing search or chatbot development, our results can thus provide an additional feature to consider when choosing the kernel LLM, on top of computational cost, efficiency, accuracy in other tasks, and more. Our results can also support informed development and management of automated filters in the context of content moderation and hate speech detection , as well as for industrial applications for companion technologies tailored to different countries[8].ChatGPT holds significant promise for transforming various sectors, especially academic publishing, by enhancing sentiment analysis and streamlining decision-making for reviewers and editors. Its potential to improve transparency and efficiency in scholarly communication is notable. However, careful consideration is needed to address ethical concerns and bias mitigation to ensure responsible use. Collaborative efforts across disciplines will be essential to maximizing the benefits of ChatGPT while promoting continuous improvement and ethical AI practices. By doing so, ChatGPT can contribute to a more objective, efficient, and inclusive academic ecosystem. Based on an analysis of 2,652 reviews on ChatGPT performance from July 28, 2023, to January 28, 2024, 87.71% of users expressed satisfaction, compared to 12.29% negative reviews. This suggests a generally favrouable public perception of ChatGPT's performance on the Google Play Store .Using the f1score as an evaluation metric due to the dataset's imbalance, three classification algorithms were tested on a subset of 796 reviews. Random Forest and Support Vector Machine both achieved an f1-score of 90%, while Naïve Bayes received 87%. The results indicate that Random Forest and Support Vector Machine outperformed Naïve Bayes in sentiment analysis accuracy, reflecting positive user responses to ChatGPT. This high satisfaction suggests ChatGPT is effectively meeting users' expectations and solidifying its position as a leading tool in AI and conversational agents.

References

Frans Sudirjo1 , Karno Diantoro2 , Jassim Ahmad AlGasawneh3, Hizbul Khootimah Azzaakiyyah4 , Abu Muna Almaududi Ausat5.

Jinqiao Zhou 1 , Ziqi Liang 2, Yuhua Fang 3 , Zhanxi Zhou B. Shawar, E.

Walid Hariri Labged Laboratory, Computer Science department Badji Mokhtar Annaba University, Algeria hariri@labged.netP.

SHWEZINSUNAING E-mail: shwezinsunaing_s@cmu.ac.th DR. PIYACHAT UDOMWONG E-mail: piyachat.u@cmu.ac.

Reuben NgID1,2, Ting Yu Joanne ChowID1

Kelly La Venture Bemidji State University, kelly.laventure@bemidjistate.edu Hyun Sang An Minnesota State University, Moorhead, hyunsang.an@mnstate.edu Wooyang Kim "Minnesota State University, Moorhead", wooyang.kim@mnstate.edu

Gilbert Jeffson Sagala1 , Yusran Timur Samuel2

Alessio Buscemi, Daniele Proverbio Department of Industrial Engineering, University of Trento