DECODING GOOD BOTS IN ONLINE COMMUNITIES: HOW TEXTUAL SENTIMENT AND PARALANGUAGE DRIVE POSITIVE SOCIAL INTERACTIONS

Yu Chen
Business School
Shandong University
Weihai, China
chrainy@mail.sdu.edu.cn

Ruijie Sun
School of Software and Microelectronics
Peking University
Beijing, China
rjsun25@stu.pku.edu.cn

Feng Liu[1]
Business School
Shandong University
Weihai, China
liufeng@sdu.edu.cn

ABSTRACT

Social bots have emerged as sophisticated agents in online communities, serving as a practical lens through which the societal implications of artificial general intelligence can be explored. These bots can participate in various positive social interactions, including the provision of entertainment services and community moderation, but some bots have negative effects, such as spreading misinformation. Therefore, to maintain the harmony of online communities, it is crucial to distinguish between "good" and "bad" bots and to identify the characteristics that can be used to make this distinction. Using the elaboration likelihood model, this study integrates text and sentiment analysis with machine learning algorithms to explore what makes bots good from a user perspective. Drawing on data from Reddit, we analyze how the textual sentiment and textual paralanguage embedded in bot-generated comments drive positive social interactions. Our results indicate that textual sentiment, processed via the central route, has a stronger impact on user evaluations than textual paralanguage, which operates through the peripheral route. Overall, our study offers important insights into bot design and the factors that promote positive social interactions within online communities, thus supporting platform governance.

Keywords: Social bot; Textual sentiment; Textual paralanguage; Social interaction; Machine learning

1. Introduction

The use of artificial intelligence (AI) on social media platforms has transformed online interactions, inviting the participation of entities other than human users (Safadi et al., 2024). This includes social bots, which are AI-driven algorithms that automatically generate content and emulate human behavior on social media (Ferrara et al., 2016). These bots offer an empirical lens through which to examine the societal implications of artificial general intelligence because they can be programmed to participate in a diverse range of activities, including posting content on social media (Cheng et al., 2023; Salge and Karahanna, 2018; Salge et al., 2022), moderating discussions (He et al., 2025), and assisting with editorial tasks (Halfaker and Riedl, 2012; Geiger, 2014). The rapid proliferation of social bots has increased the urgency to understand their impact on social interactions within online communities.

Research to date has shown that social bots can have both positive and negative impacts on online communities. For example, social bots can enhance user experiences (Tsai et al., 2021), participate in software development (Hukal

Cite: Chen, Y., Sun, R., & Liu, F. (2026, May). Decoding good bots in online communities: How textual sentiment and paralanguage drive positive social interactions. *Journal of Electronic Commerce Research, 27*(2).

[1] Corresponding author.

et al., 2019), and foster consumer engagement in e-commerce (Cheng et al., 2021), contributing to the growth and vibrancy of online platforms. However, social bots are often exploited to disseminate spam, misinformation, propaganda, and malware and to serve as vectors for phishing and other scams (Benjamin and Raghu, 2023). On social media platforms, malicious social bots can amplify low-credibility information and inflammatory content, thereby undermining user trust and the reputation of the platform (Shao et al., 2018; Stella et al, 2018). Platform operators are thus motivated to ensure that bots generate high-quality content because a deterioration in the user experience can undermine user engagement and potentially result in user attrition and financial losses.

Bot-generated content is the primary mechanism through which social bots influence online communities. Recent advances in large language models have allowed social bots to appear more anthropomorphic by producing contextually relevant and human-like content, which increasingly influences user perceptions and behaviors (Ferrara et al., 2016; Pentina et al., 2023). Notably, emotional expressions by social bots constitute a form of anthropomorphism that can strongly stimulate user engagement and interaction because emotions spread rapidly within online communities (Han et al., 2023; Kramer et al., 2014; Park et al., 2024). Emotional expressions can be delivered by social bots either as textual sentiment conveyed through words or as nonverbal textual elements such as emojis and emoticons (e.g., 😁, ✌️, 🤖), with the latter serving as textual paralanguage that specifically enhances emotional communication in online contexts (Luangrath et al., 2017, 2023). However, some bots deliberately amplify negative emotions such as fear and alarmism to manipulate users and exacerbate social conflict, thus distorting public discourse in online communities (Stella et al., 2018). Users in online communities have become increasingly aware of the influence of this negative bot-generated content on online interactions and frequently post comments warning others about its potential harms (Benjamin and Raghu, 2023; He et al., 2025). For example, Reddit users frequently express their approval or disapproval of a bot's comments by replying with phrases such as "good bot" or "bad bot." These evaluative responses can be interpreted as emotionally charged responses affected by the user's perception of social interaction quality, offering a useful lens for examining how bots are socially appraised (Sykora et al., 2022).

Drawing on the elaboration likelihood model (ELM), we investigate how the textual sentiment and paralanguage embedded in bot-generated comments influence social interactions. The ELM suggests that persuasive communication operates through two distinct processing routes: the central route, which involves deep cognitive elaboration and high user involvement, and the peripheral route, which is more automatic, relying on heuristics and requiring less cognitive effort (Petty and Cacioppo, 1986). In line with this framework, we conceptualize textual sentiment in bot-generated comments as a central-route cue that prompts deeper cognitive elaboration. In online communities, textual sentiment typically contains semantically rich and contextually grounded emotional language, which encourages careful reflection and sustained interpretation, thus increasing the likelihood of central route processing (Blanco et al., 2010; Han et al., 2023). In contrast, textual paralanguage embedded in bot-generated comments tends to activate the peripheral route because it is generally processed in a more intuitive and heuristic manner (Maiberger et al., 2024). The textual paralanguage (e.g., emojis and typographic symbols) primarily relies on visual and symbolic cues that are immediately perceptible, enabling users to make rapid affective responses with less conscious effort (Luangrath et al., 2023).

Despite growing interest in social bots, little research has systematically examined how the textual sentiment and paralanguage embedded in bot-generated content jointly influence social interactions, a gap this study seeks to address. Most prior research has focused on interpersonal contexts, examining how textual sentiment in customer service interactions (Altman et al., 2021), online drug reviews (Ball et al., 2025), and social media posts (Stieglitz and Dang-Xuan, 2013) affect user engagement, cognitive processing, and information diffusion. Similarly, studies on textual paralanguage have highlighted its role as a persuasive cue in shaping online attitudes, including in marketing communications and blog interactions (Orazi et al., 2023; Rodríguez-Hidalgo et al., 2017). However, these advances have mainly focused on human-human interactions (Maiberger et al., 2024; Oh et al., 2023), leaving open the question of how textual sentiment and paralanguage operate in human-bot interactions. To address these issues, the present study aims to examine the following research questions (RQs):

RQ1: *How does the textual sentiment embedded in bot-generated content engage central-route processing to affect human users' evaluations of social bots in online communities?*

RQ2: *How does the textual paralanguage embedded in bot-generated content engage peripheral-route processing to affect human users' evaluations of social bots in online communities?*

RQ3: *Which exerts a stronger influence on human users' evaluations of social bots in online communities: the textual sentiment or the textual paralanguage embedded in bot-generated content?*

To answer these RQs, we collected bot-generated content from a large online discussion and community platform and analyzed whether "good bot" responses could be predicted based on textual sentiment and paralanguage. We first conducted text and sentiment analysis to extract textual sentiment and paralanguage from bot-generated content.

Subsequently, we employed a set of machine learning algorithms to evaluate the predictive power of textual sentiment and paralanguage in determining whether bot-generated content would receive "good bot" responses. To systematically assess the contribution of textual sentiment and textual paralanguage to prediction performance, we sequentially introduced the two feature sets into each machine learning model. Of the eight algorithms evaluated, the Light Gradient Boosting Machine (LightGBM) model demonstrated the highest prediction accuracy. Finally, we interpreted the model output by visualizing the permutation importance of each feature, thereby quantifying their relative contribution to prediction performance.

Our research offers three notable contributions. First, this study extends the application of the ELM by offering a clearer understanding of how the textual sentiment and textual paralanguage embedded in bot-generated content shape user evaluations in online communities. While prior research has primarily applied the ELM to human-generated content (Chou et al., 2022; Shao et al., 2023), our study broadens its scope by examining how both central and peripheral cues influence users' perceptions of social bots, thus enriching the understanding of human-bot interactions. Second, we innovatively integrate machine learning with the ELM framework to systematically compare the persuasive effects of textual sentiment and paralanguage in bot-generated content. By distinguishing between the two persuasive routes, our approach offers a novel perspective on the mechanisms through which bot-generated content affects user evaluations. Our findings show that textual sentiment, which operates through the central route, exerts a greater impact on user evaluations than textual paralanguage via the peripheral route. Third, this study offers practical insights for platform operators seeking to foster constructive human-bot interactions. By identifying key factors associated with favorable user responses, this study provides potential guidelines for the design and moderation of bots to promote positive social interactions within online communities.

2. Literature Review and Theoretical Framework

2.1. Social Bots in Online Communities

Social bots are computer algorithms that automatically generate content and interact with human users on social media (Ferrara et al., 2016; Shao et al., 2018), emulating human-like behavior and potentially influencing social interactions within online communities (Stella et al., 2018; Chen et al., 2021). Due to their situatedness, autonomy, and flexibility, agentic information system artifacts such as bots can participate in various digital environments, act without human supervision, and respond to environmental stimuli (Safadi et al., 2024). These characteristics have been reinforced by advances in generative AI, enabling bots to moderate inappropriate content (He et al., 2025) and harassment (Nguyen et al., 2024). Bots can also swiftly and accurately analyze larger volumes of information than human moderators (Salge and Karahanna, 2018), whose processing range and speed are often limited (Safadi et al., 2024). In addition, because bots are not driven by emotions, their behavior tends to be more stable and reliable (Salge et al., 2022), and they can remain continuously active, unlike human users, who are constrained by time (Delkhosh et al., 2023).

Despite these advantages, social bots are also susceptible to manipulation and can negatively influence online platforms. For example, bots can amplify unverified content through trending topics and hashtags, accelerating the dissemination of fake news (Bessi and Ferrara, 2016; Shao et al., 2018; Song et al., 2025). In addition, platforms such as X (formerly Twitter), whose policies to curb the spread of misinformation primarily focus on regulating human user behavior, often struggle to effectively suppress bot-generated content (Hwang et al., 2024). Therefore, it is important to accurately identify malicious bot-generated content and understand what types of content foster user engagement and positive responses, so as to design bots that promote harmonious online interactions.

Examining users' evaluations of bot-generated content in online communities can provide valuable insights into the processes through which bots are socially appraised. These evaluations not only reflect judgments of specific bot-generated content but also extend to the bots themselves as social actors (Gao et al., 2025). Users may perceive bots producing high-quality, socially appropriate content to be competent and trustworthy actors, which increases their willingness to interact with these bots and to engage in related community discussions. In contrast, exposure to low-quality or misleading bot-generated content can foster algorithm aversion (Liu et al., 2023; Shao et al, 2018), discouraging user participation and weakening trust in the platform's governance (Shao et al., 2018). These evaluations made by individual users can also spread across online communities and affect how other users perceive particular bots (Gao et al., 2025), contributing to the formation of broader social perceptions.

Within the context of human-bot interactions, bot-generated content involving textual sentiment and textual paralanguage plays an increasingly significant role in shaping user perceptions. On the one hand, the anthropomorphic characteristics of social bots allow users to perceive human-like emotions in bot-generated content, which can evoke empathy and encourage more favorable user perceptions of the interaction (Chen et al., 2022; Lee et al., 2017). In addition, because emotional content is forwarded more frequently and more rapidly than neutral content (Stieglitz and Dang-Xuan, 2013), social bots tend to strategically disseminate various emotions to attract attention, increase their influence, and—in some cases—even manipulate public sentiment (Smith et al., 2022). For example, during the 2016 U.S. presidential election, social bots played a significant role in enhancing the positive reputation of the political candidates by generating a large number of tweets infused with positive sentiment (Heredia et al., 2018).

On the other hand, textual paralanguage, which refers to written manifestations of nonverbal cues such as symbols, images, or punctuation, is widely used in computer-mediated communication to substitute for face-to-face elements such as tone and body language (Luangrath et al., 2017). For example, it has been reported that the use of textual paralanguage (e.g., emojis) in human-bot interactions can enhance the perception of a bot's warmth (Yu and Zhao, 2024). Prior research also suggests that, when chatbots use interjections (e.g., "hmm" and "wow"), they can improve users' attitudes, increase satisfaction, and promote favorable behavioral intentions (Sheehan et al., 2024). Moreover, bots that use emojis in online communities are perceived as more socially attractive and more trustworthy than those relying solely on verbal messages (Beattie et al., 2020). Therefore, examining how textual paralanguage can be strategically incorporated into bot-generated content is essential for improving the quality of human-bot interactions and ensuring the effective governance of online communities.

Scholars often distinguish between "good" and "bad" bots based on their contrasting contributions to online communities. Massanari (2016) found that good bots on Reddit have some common characteristics, including that they are polite, useful, informative, and unobtrusive, while they interact with users over a longer period of time. Good bots can also conduct data analysis and increase user engagement, while bad bots spread misinformation and distribute advertising spam (Latah, 2020; Tsvetkova et al., 2024). However, although prior literature has discussed the positive and negative roles of social bots in online communities, limited attention has been given to understanding the factors that affect user evaluations of these bots. Our study fills this gap by exploring how textual sentiment and paralanguage influence the evaluation of social bots as "good", providing a clearer understanding of the mechanisms that drive positive human-bot interactions.

2.2. Elaboration Likelihood Model

User evaluations of social bots are influenced by textual sentiment and paralanguage, yet the mechanisms through which these features shape human cognition and judgment remain underexplored. This study uses the ELM as a theoretical lens through which to explore how textual sentiment and paralanguage shape users' evaluation of bots. The ELM is a dual process theory that posits that changing human perceptions, attitudes, and behaviors mainly depends on how persuasive information is processed, arguing that persuasion can act via a central or peripheral route based on the elaboration likelihood (Bhattacherjee and Sanford, 2006; Petty et al., 1983). The elaboration likelihood is determined by an individual's motivation and ability to process information, and it determines the effects of the central and peripheral routes on attitude formation (Petty and Cacioppo, 1986; Shao et al., 2023). Specifically, via the central route, individuals with high motivation and the ability to process information engage in extensive cognitive evaluation, which is referred to as a high elaboration likelihood, whereas, following the peripheral route, less cognitive elaboration is required and behavior is more intuitive. Therefore, the ELM has become a foundational framework for understanding how persuasive messages affect attitudes and behavioral responses (Sussman and Siegal, 2003).

The ELM has been widely applied to online community contexts to explain how individuals process persuasive information via the central and peripheral routes, leading to changes in attitude and subsequent decision making (Bao and Wang, 2021; Chou et al., 2022). Prior research suggests that central-route processing is triggered when users focus on substantive content, such as product descriptions and argument quality, while peripheral-route responses occur when judgments rely on superficial cues such as the page layout, color scheme, and background music (Blanco et al., 2010; Cyr et al., 2018). In user-generated content within online communities, argument quality indicators such as clarity, specificity, and readability require users to engage in effortful cognitive processing, whereas sentiment indicators and star ratings serve as heuristic, peripheral cues (Chou et al., 2022). However, these insights have primarily been derived from interactions among humans, leaving the dual-route dynamics of human-bot interactions underexamined. Bots differ from human communicators in terms of perceived agency, consistency, and sociality, which may alter how users engage with persuasive content.

Applying the ELM to bot-generated content, we distinguish between textual sentiment and textual paralanguage as features that may trigger particular cognitive processing routes. The ELM is particularly appropriate for our context because social bots are often designed to influence user perceptions and behaviors through persuasive messages that invite both thoughtful interpretation and intuitive responses (Cai et al., 2023; Dehnert and Mongeau, 2022). Specifically, bot-generated content often includes textual sentiment and paralanguage (Benjamin and Raghu, 2023; Maiberger et al., 2024), aligning with the dual-route structure of the ELM. The textual sentiment embedded in bot-generated content may prompt human users to engage in more effortful cognitive processing when it appears nuanced or strategically framed, thereby aligning with the central route (Chen et al., 2025). In contrast, textual paralanguage (e.g., emojis) functions as a peripheral cue that influences users' perceptions of warmth, credibility, or social presence, particularly when users rely on heuristics due to limited motivation or cognitive capacity (Beattie et al., 2020). By integrating the ELM with the context of social bots, this study addresses a key theoretical gap, with prior ELM research seldom considering how human perceptions of automated agents are affected by dual-route processing, particularly when message features are produced by artificial actors rather than humans. Understanding these mechanisms is critical for explaining why certain bot-generated content fosters favorable user evaluations and social interactions in online communities.

2.3. Conceptual Framework: Central and Peripheral Routes in Bot Evaluation

In online communities, text is the dominant form of user expression and a crucial medium through which social bots interact with human users. To systematically investigate how social bots are evaluated by users, it is essential to extract meaningful insights from text to measure, monitor, and understand the factors that define a good bot in the context of an online community. To achieve this, we analyze the factors affecting the evaluation of social bots from the perspective of textual sentiment and paralanguage using machine learning methods. Building on the ELM framework, we conceptualize textual sentiment as a central-route cue and textual paralanguage as a peripheral-route cue in the evaluation of social bots (see Figure 1).

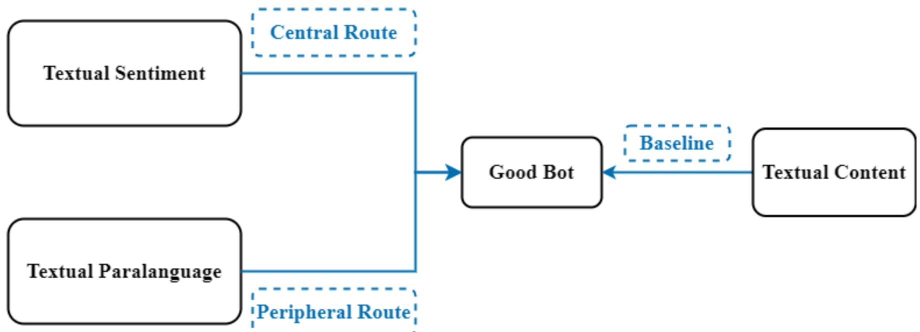

Figure 1. Conceptual Framework

Textual sentiment is more likely to activate the central route in users' evaluations of social bots in online communities. This route involves the careful and thoughtful processing of the persuasive content of a message when individuals are highly motivated and have sufficient cognitive resources (Petty et al., 1983). In online communities, emotional content in text often carries rich semantic nuances and contextually grounded meaning, which requires users to engage in deeper reflection and more effortful cognitive processing (Blanco et al., 2010; Han et al., 2023). Specifically, emotional language in bot-generated content can prime human users with corresponding emotions and elicit empathic responses, especially when the bot is perceived as socially present (Chen et al., 2025; Park et al., 2024). This empathic engagement increases the users' motivation to understand the bot's message, aligning with the central route of persuasion. Recent studies have explored the impact of textual sentiment on user activity in online communities (e.g., Altman et al., 2021; Ball et al., 2025; Zheng et al., 2024). For example, Ball et al. (2025) found that more strongly negative sentiment in drug reviews was correlated with a higher risk of a serious drug recall, and this relationship was especially true for female users. Although some related studies have proven that textual sentiment has an impact on the evaluation of the primary quality of a product (Zheng et al., 2024) and the quality of restaurant service (Mejia et al., 2021), little research has examined the relationship between textual sentiment and the evaluation of social bots. Therefore, to fill this gap, we explore the role of textual sentiment (e.g., anger, anticipation, disgust, fear, and joy) in evaluating whether a bot is good or bad via the central route.

Given its intuitive and heuristic nature under typical viewing conditions, textual paralanguage is positioned closer to the peripheral end of the elaboration continuum. The peripheral route involves the more superficial processing of a persuasive message (Petty et al., 1983). When individuals are less motivated or unable to carefully analyze a message due to factors such as low involvement, distraction, or a lack of knowledge, they rely on peripheral cues to make judgments (Petty and Cacioppo, 1986; Shao et al., 2023). Textual paralanguage consists mainly of visual elements such as emojis and symbols. These nonverbal communication cues are processed more quickly than verbal content, allowing users to form immediate affective impressions and process heuristic cues with minimal cognitive effort (Maiberger et al., 2024). The emotional meaning of textual paralanguage is often culturally encoded and easily recognizable, enabling rapid interpretation without requiring deep cognitive elaboration. For example, heart emojis (e.g., ❤️, 💓, and 💕) are often linked to positive sentiment, as in expressions like "LF 💓 HYL," thus offering valuable cues for valence evaluation. Likewise, the use of capital letters (e.g., "LF LOVE HYL") can intensify emotional expression or emphasize the key points of a message. Due to the crucial role of textual paralanguage in online communities (Maiberger et al., 2024; Orazi et al., 2023), we evaluated this as a peripheral route to explore the factors that characterize a good bot. By integrating sentiment and paralanguage into the dual-route structure of the ELM, our framework specifies how distinct textual features embedded in bot-generated content guide user evaluations through different processing mechanisms. This perspective addresses a critical gap in prior literature, which has largely focused on human-to-human persuasion, and advances our understanding of how users form judgments about social bots in online communities.

3. Methodology

Figure 2 presents the roadmap for our research, which consists of four key steps: data collection, feature extraction, model selection, and interpretation. First, we collected bot-generated content from a large online community on which human users and bots can discuss and share information on a wide range of topics. We then used text and sentiment analysis to extract textual content, sentiment, and paralanguage. Following this, we compared the accuracy of eight machine learning algorithms in predicting a good bot based on these cues and factors. Finally, we interpreted the results of the most accurate machine learning model.

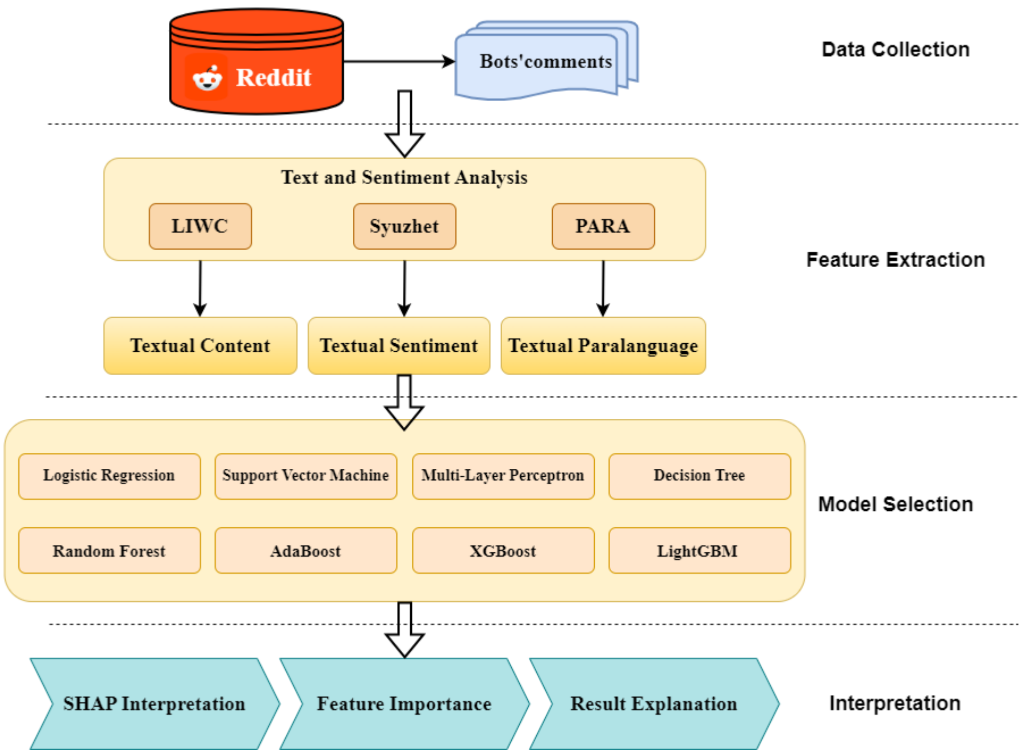

Figure 2. Roadmap for this Study

3.1. Data Collection

Our data sample was obtained from Reddit (https://www.reddit.com/), a leading global online community platform. Reddit offers a publicly accessible application programming interface (API) that allows for the retrieval of user accounts and communication data. We employed the Reddit API to gather bot-generated comments and users' evaluations of them. There are over 130,000 active communities on Reddit, which are referred to as subreddits, covering a wide array of topics, including global news, sports, politics, and science. Numerous social bots are active across various subreddits, participating in discussions by commenting on posts in a manner similar to human users. Human users, in turn, can respond to these comments, leading to interactions between humans and bots. According to Reddit community norms, users often evaluate a bot's comments by replying with phrases such as "good bot" or "bad bot", as illustrated in Figures 3a and 3b, respectively.

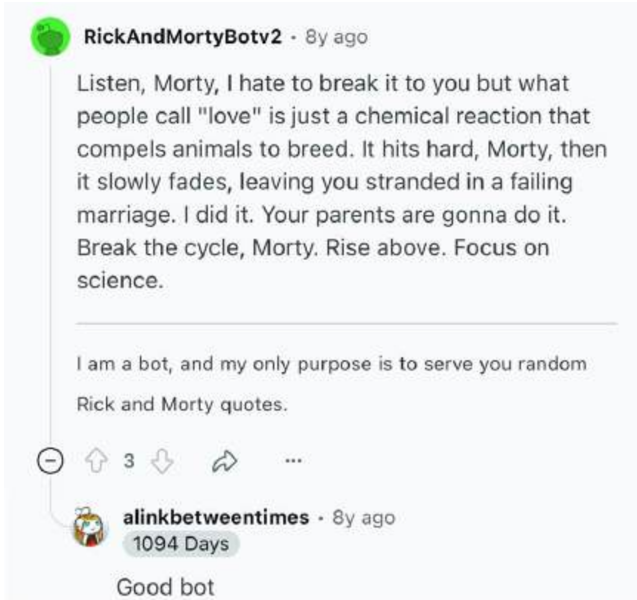

Figure 3a. Example of a Social Bot Receiving a "Good Bot" Response

Figure 3b. Example of a Social Bot Receiving a "Bad Bot" Response

We identified the most popular bots based on an online ranking that reflects how often users responded to their posts with "good bot" or "bad bot" comments[1]. This ranking was derived from the lower bound of the confidence interval for the Wilson score for the proportion of "good bot" votes (Wilson, 1927). Using this ranking list, we filtered out inactive accounts and human users misidentified as bots (Safadi et al., 2024). Our final dataset consisted of 15,195 comments from 1,863 social bots that received consistent "good bot" or "bad bot" replies across 3,356 subreddits between 2016 and 2024. Table 1 summarizes the dependent, central route, peripheral route, and control variables used in the present study, with the descriptive statistics for these variables provided in Appendix A.

[1] Additional information can be accessed at https://botrank.pastimes.eu.

Table 1. Description of the Variables

Variable	Description
Panel A. Dependent variable	
Good Bot	Binary variable: 1 if the bot comment received exclusively "good bot" responses; 0 if it received exclusively "bad bot" responses.
Panel B. Central route variables	
Text_ Anticipation	Percentage frequency of words that express anticipation
Text_ Joy	Percentage frequency of words that express joy
Text_Surprise	Percentage frequency of words that express surprise
Text_Trust	Percentage frequency of words that express trust
Text_Anger	Percentage frequency of words that express anger
Text_Disgust	Percentage frequency of words that express disgust
Text_Fear	Percentage frequency of words that express fear
Text_Sadness	Percentage frequency of words that express sadness
Panel C. Peripheral route variables	
Para_Stress	Number of stress elements
Para_Tempo	Number of tempo elements
Para_Rhythm	Number of rhythm elements
Para_Emphasis	Number of emphasis elements
Para_Pitch	Number of pitch elements
Para_Censorship	Number of censorship elements
Para_Spelling	Number of spelling errors
Para_Alternant	Number of alternant elements
Para_Differentiator	Number of differentiator elements
Para_Tactile_Emoticon	Number of tactile emoticon elements
Para_Alphahaptics	Number of alphahaptic elements
Para_Bodily_Emoji	Number of bodily emojis
Para_Bodily_Emoticon	Number of bodily emoticons
Para_Alphakinesics	Number of alphakinesic elements
Para_Nonbodily Emoji	Number of nonbodily emojis
Para_Nonbodily Emoticon	Number of nonbodily emoticons
Para_Formatting	Number of formatting elements
Para_Emoji_Count	Total number of emojis
Para_Emoji_Index	Total number of tactile emojis, bodily emojis, and nonbodily emojis
Para_Emoticon_Index	Total number of tactile emoticons, bodily emoticons, and nonbodily emoticons
Para_TPL_Index	Total number of all textual paralanguage nonverbal elements
Panel D. Control variables	
Liwc_Wc	Word count of the bot comment
60 LIWC Variables	Percentage frequency of words in 60 LIWC categories

3.2. Variables

The dependent variable in this study was whether a bot is perceived as good or not (*Good Bot*). The presence of "good bot" or "bad bot" replies to bot comments was used to assess this variable. The textual sentiment of bot comments was examined as a central route factor influencing the evaluations of a bot as good or bad because it required the user to engage in complex cognitive reflection and scrutiny. We conducted sentiment analysis on the bot comments using the Syuzhet package in R to obtain the frequency of use of four positive emotional terms (*Text_Anticipation*, *Text_Joy*, *Text_Surprise*, and *Text_Trust*) and four negative terms (*Text_Anger*, *Text_Disgust*, *Text_Fear*, and

Text_Sadness) (Jockers, 2016). This package cleaned the text mining data and then accurately identified the presence of emotional vocabulary within the text.

We believe that the textual paralanguage of bot comments is likely to affect user evaluations of bots via the peripheral route. Bot comments often contain nonverbal cues that signal a specific meaning. This textual paralanguage, which includes nonverbal auditory (e.g., "Hmmmm...."), tactile (e.g., "Would you like one? 🤗"), and visual elements (e.g., "👀"), serves as a complement or substitute for linguistic content in online text-based interactions (Luangrath et al., 2017). Figures 4a and 4b present examples of textual paralanguage in comments that received a "good bot" and "bad bot" response, respectively. PARA[2], an advanced automated classifier designed to detect nonverbal communication elements within text-based content (Luangrath et al., 2023), was used to extract 21 textual language features from bot comments, including 17 distinct auditory, tactile, and visual textual paralanguage elements and 4 aggregate variables (*Para_Emoji_Count*, *Para_Emoji_Index*, *Para_Emotion_Index*, and *Para_TPL_Index*).

Figure 4a. Example of Textual Paralanguage in a Bot Comment Receiving a "Good Bot" Response

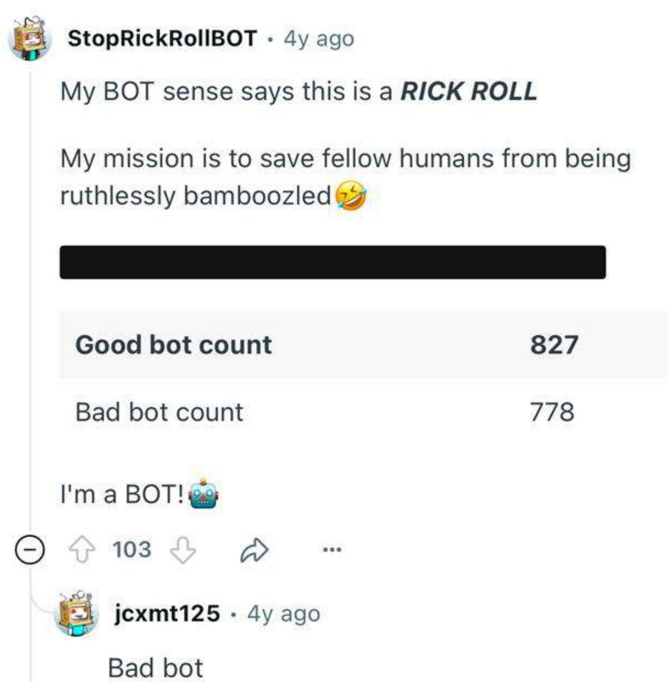

Figure 4b. Example of Textual Paralanguage in a Bot Comment Receiving a "Bad Bot" Response

[2] Further details about textual paralanguage are available at https://www.textualparalanguage.com.

To control for factors potentially affecting the user evaluation of bots, we included textual content features in the baseline model. In particular, we extracted 61 linguistic inquiry and word count (LIWC) features as baseline variables (e.g., *Liwc_Function* and *Liwc_Wc*) (Pennebaker et al., 2015). LIWC is extensively used in social science research to analyze text by establishing the correspondence between textual content and predefined linguistic variables across a wide range of contexts (Ludwig et al., 2013; McHaney et al., 2018). This approach can be used to calculate the percentage frequency of specific word categories within a given text to build a linguistic profile.

3.3. Machine Learning Algorithms

This study employed a machine learning approach to examine how textual sentiment and paralanguage influence the evaluation of bots in online communities. Compared to traditional statistical methods, machine learning is more adept at handling problems with ambiguous functional forms by directly identifying patterns in data (Shao et al., 2023), while also reducing the risk of overfitting and enhancing the interpretability and replicability of analytical results (Sun et al., 2024). In this study, we aimed to predict whether a bot was perceived as a good bot by users. To this end, we tested four individual machine learning algorithms (Logistic Regression [LR], Support Vector Machine [SVM], Multi-Layer Perceptron [MLP], and Decision Tree [DT]) and four bagging and boosting ensemble machine learning algorithms (Random Forest [RF], Adaptive Boosting [AdaBoost], eXtreme Gradient Boosting [XGBoost], and LightGBM). A brief explanation of these models is presented below.[3]

LR, a widely used classification method in machine learning, predicts the probability of categorical outcomes by modeling the relationship between features and the likelihood of a particular outcome using a logistic function (Harrell, 2015). SVM is a binary classification model designed to identify a hyperplane in feature space, maximizing the margin between the nearest data points from each class. This approach enhances the model's ability to classify binary outcomes (Kuhn and Johnson, 2013). MLP is an artificial neural network with interconnected input, hidden, and output layers that is capable of complex signal mapping for classification while being robust to noise and missing data (Han et al., 2022). DT is a supervised learning method for regression and classification composed of nodes, branches, and leaves (Dietterich, 2000).

The ensemble models were used to enhance the generalizability and robustness of the estimator (Chen and Guestrin, 2016). RF is a bagging algorithm that combines predictions from multiple decision trees trained on random subsets of data, enhancing model interpretability and reducing variance (Breiman, 2001). The selected boosting methods integrated weak models to construct a more powerful ensemble for more accurate forecasting. AdaBoost iteratively focuses on previously misclassified samples, refining each subsequent model to improve accuracy through weighted adjustments (Kuhn and Johnson, 2013). XGBoost, known for its computational efficiency, uses a second derivative to produce a more accurate loss function, with the regular terms preventing overfitting (Chen and Guestrin, 2016). Finally, LightGBM, a highly efficient gradient boosting model, optimizes computational speed and accuracy using techniques such as histogram algorithms, making it suitable for large-scale data (Ke et al., 2017).

We first divided the collected data into training and testing sets and used the synthetic minority oversampling technique (SMOTE) with the training set to address the significant class imbalance, where only 1,598 of the 15,195 bot comments received "bad bot" replies. The SMOTE generates synthetic samples for the minority class using a *k*-nearest-neighbor approach to produce a more balanced class distribution (Chawla et al., 2002). For each machine learning model, we performed a grid search to define the optimal range for the core hyperparameters. We then conducted five-fold cross-validation, randomly dividing the training set into five equal parts, using one for validation and the remaining four for training. This process was repeated across different hyperparameter combinations to identify the configuration that yielded the best performance, as measured using the classification accuracy (Sun et al., 2024). The out-of-sample predictive performance of the final models was then evaluated using the test set. Appendix B provides more details on the hyperparameter tuning process used in this study to predict cases of "good bot" classification.

To systematically compare the persuasive effects of the textual sentiment and paralanguage embedded in bot-generated content, this study integrated machine learning with the ELM framework. We evaluated the contribution of the peripheral route and central route variables by sequentially including them in the models. First, we tested a baseline model with only the control variables (i.e., the 61 LIWC variables). We then introduced the peripheral and central route variables to the baseline model, employing three feature sets to assess the overall performance of our framework and identify the unique contribution of each route to the prediction of "good bot" classifications.

[3] More information can be found at https://scikit-learn.org/stable/supervised_learning.

4. Results

4.1. Performance Evaluation Indicators

We employed accuracy, recall, and the F1 score as performance metrics to evaluate and compare the effectiveness of the tested machine learning models. Our study conceptualized the user evaluation of bot comments as a binary outcome, categorizing "good bot" and "bad bot" responses as positive and negative observations, respectively. The classification output of the models was classified into one of four groups: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). A TP occurs when the model correctly identifies a bot comment that received a "good bot" response, while a TN reflects the correct identification of a comment that received a "bad bot" response. Conversely, an FP arises when the model incorrectly predicts a "good bot" response for a comment that did not receive one, and an FN occurs when a comment that actually received a "good bot" response is predicted to have received a "bad bot" response.

Accuracy was assessed by calculating the proportion of correct predictions (i.e., TP+TN) relative to the total number of cases (Eq. 1). A higher accuracy reflected a model's ability to effectively distinguish between "good bot" and "bad bot" replies.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Recall represents the proportion of "good bots" that were correctly identified by the model relative to the total number of actual "good bot" cases (Eq. 2). A high recall indicated that a model was effective in recognizing "good bot" replies, demonstrating its sensitivity in detecting TPs.

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

Precision (Eq. 3) and recall are inversely related, with the F1 score (Eq. 4) balancing the two. A higher F1 score indicates a more effective model because it reflects a favorable balance between precision and recall, ultimately improving overall prediction quality.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$F1score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

4.2. Performance of the Machine Learning Models

Table 2 summarizes the predictive performance of the machine learning models, showing that LightGBM was the best-performing algorithm. Most of the ensemble machine learning models (i.e., RF, XGBoost, and LightGBM) outperformed the other methods, with the textual sentiment and paralinguistic variables consistently improving the performance compared to the baseline model, which relied solely on LIWC variables. This improvement was observed across all models but was most pronounced for the ensemble methods. In particular, LightGBM demonstrated the highest predictive accuracy (0.8206) when trained on the full set of variables. The recall score for LightGBM was 0.8690, indicating that it correctly identified 86.9% of actual "good bot" cases. Its F1 score of 0.8968 also highlighted its robustness in avoiding FPs and FNs.

Table 2. Predictive Performance of the Machine Learning Models

	LR			**SVM**			**DT**			**MLP**		
	Accuracy	Recall	F1	Accuracy	Recall	F1	Accuracy	Recall	F1	Accuracy	Recall	F1
Baseline	0.6596	0.6981	0.7863	0.6475	0.6649	0.7720	0.7203	0.7473	0.8275	0.7405	0.7749	0.8428
Baseline + Central route	0.5782	0.5864	0.7129	0.6897	0.7153	0.8046	0.7466	0.7812	0.8464	0.7631	0.8048	0.8586
Baseline + Peripheral route	0.5791	0.5864	0.7143	0.6931	0.7179	0.8076	0.7570	0.7886	0.8534	0.7978	0.8441	0.8822
All features	0.6054	0.6219	0.7388	0.7478	0.7878	0.8486	0.7644	0.7986	0.8588	0.7403	0.7746	0.8426
	RF			**AdaBoost**			**XGBoost**			**LightGBM**		
	Accuracy	Recall	F1	Accuracy	Recall	F1	Accuracy	Recall	F1	Accuracy	Recall	F1
Baseline	0.7671	0.8054	0.8612	0.6550	0.6751	0.7784	0.7682	0.8064	0.8619	0.7660	0.8037	0.8604
Baseline + Central route	0.7976	0.8435	0.8816	0.7187	0.7473	0.8260	0.7980	0.8424	0.8817	0.8016	0.8468	0.8841
Baseline + Peripheral route	0.8076	0.8514	0.8882	0.7280	0.7600	0.8337	0.7982	0.8397	0.8819	0.7964	0.8384	0.8808
All features	0.8190	0.8683	0.8960	0.7289	0.7629	0.8347	0.8147	0.8631	0.8931	**0.8206**	**0.8690**	**0.8968**

Note: The best performance is indicated in bold font

4.3. Model Interpretation

To interpret the output of LightGBM, the best-performing machine learning model, we employed SHapley Additive exPlanations (SHAP) to quantify the importance of the individual variables in predicting "good bot" evaluations

(Lundberg and Lee, 2017). The impact of each feature on the model's predictions was assessed by considering the output of all possible feature combinations. The significance of the top 25 most important variables and their influence on the evaluation of bots are illustrated in Figure 5. In this figure, the red bars indicate a variable that is positively associated with "good bot" evaluations, whereas the blue bars represent a negative relationship. Overall, *Liwc_Function, Liwc_Wc, Liwc_Prep, Text_Fear,* and *Liwc_Article* were the five most influential variables in predicting "good bot" responses. In addition, *Liwc_Wc, Liwc_Prep, Text_Fear, Liwc_Article,* and *Text_Trust* were the top five variables that most positively impacted "good bot" evaluations, while *Liwc_Function, Liwc_Cogproc, Para_TPL_Index, Text_Disgust,* and *Text_Surprise* were the top five variables with the strongest negative impact on "good bot" evaluations.

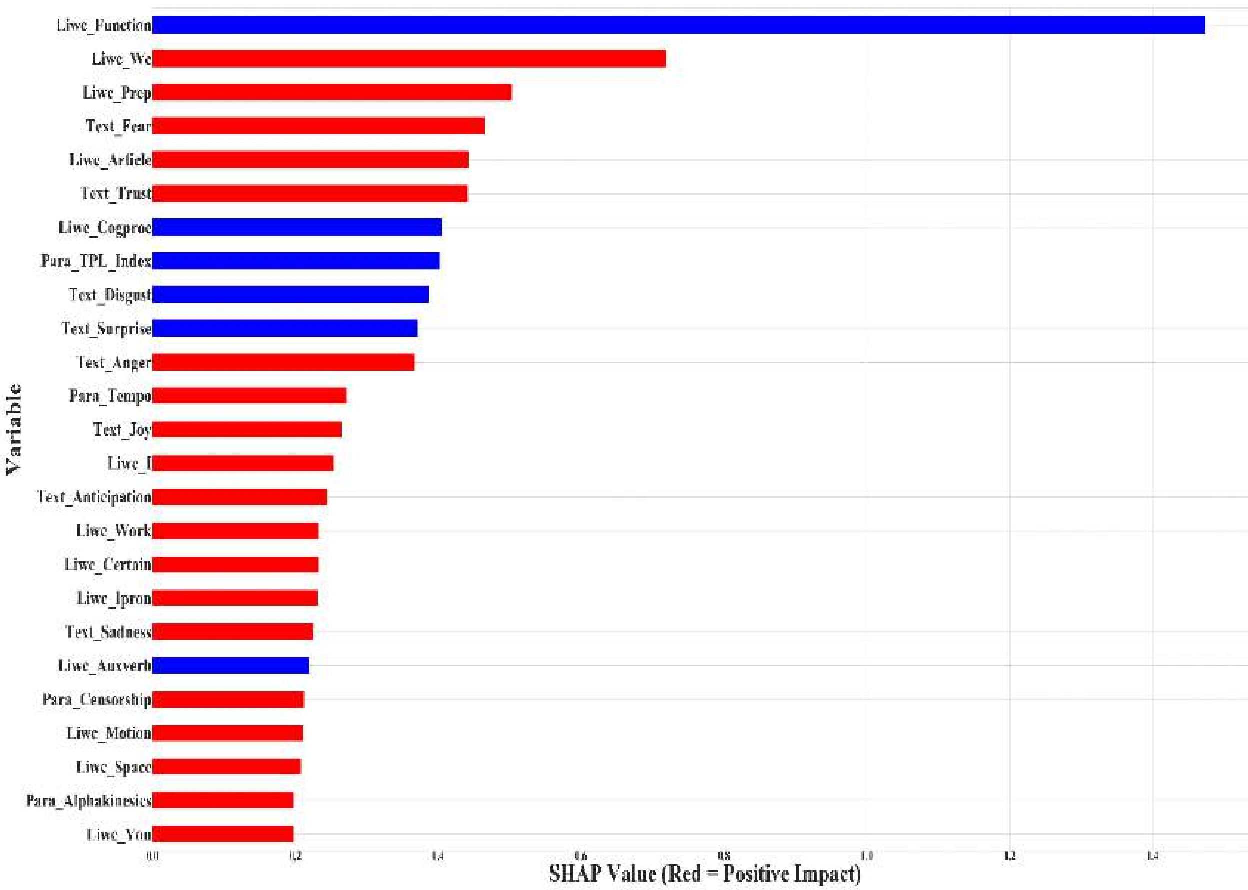

Figure 5. SHAP Values for the 25 Most Important Variables in Predicting a "Good Bot" Classification

Table 3 presents the effects and importance of the textual sentiment and paralanguage variables in predicting a "good bot" response. Panel A shows that *Text_Fear* had the highest predictive power and was positively associated with "good bot" evaluations, followed by *Text_Trust, Text_Disgust, Text_Surprise, Text_Anger, Text_Joy, Text_Anticipation,* and *Text_Sadness*. Notably, only *Text_Disgust* and *Text_Surprise* exhibited a negative association. Panel B summarizes the effects of textual paralanguage variables, with most of these variables having a positive association with "good bot" evaluations, except for *Para_TPL_Index* and *Para_Emoji_Count*, which exhibited a negative association. Notably, the cumulative SHAP value for the textual sentiment variables (2.7629) surpassed that of the textual paralanguage variables (2.1449), while the textual sentiment variables also exhibited greater importance than most of the textual paralanguage variables. These results highlight the greater importance of textual sentiment in accurately predicting a "good bot" evaluation.

Table 3. Effects of Textual Sentiment and Paralanguage Variables on “Good Bot” Classification

Variable	SHAP	Effect	Sign
Panel A. Effects of Textual Sentiment Variables on “Good Bot” Classifications			
Text_Fear	0.4647	0.1972	Positive
Text_Trust	0.4407	0.1133	Positive
Text_Disgust	0.3860	–0.0544	Negative
Text_Surprise	0.3708	–0.1569	Negative
Text_Anger	0.3669	0.0449	Positive
Text_Joy	0.2641	0.5935	Positive
Text_Anticipation	0.2448	0.5468	Positive
Text_Sadness	0.2249	0.4153	Positive
Panel B. Effects of Textual Paralanguage Variables on “Good Bot” Classifications			
Para_TPL_Index	0.4025	–0.2464	Negative
Para_Tempo	0.2713	0.8606	Positive
Para_Censorship	0.2116	0.4508	Positive
Para_Alphakinesics	0.1975	0.4537	Positive
Para_Bodily_Emoticon	0.1602	0.5298	Positive
Para_Alternant	0.1462	0.6729	Positive
Para_Emoticon_Index	0.1139	0.5745	Positive
Para_Stress	0.1135	0.6040	Positive
Para_Differentiator	0.0945	0.9177	Positive
Para_Emoji_Count	0.0766	–0.1481	Negative
Para_Emphasis	0.0637	0.1778	Positive
Para_Rhythm	0.0633	0.6755	Positive
Para_Emoji_Index	0.0541	0.2779	Positive
Para_Pitch	0.0537	0.6956	Positive
Para_Nonbodily_Emoticon	0.0316	0.5877	Positive
Para_Formatting	0.0212	0.5983	Positive
Para_Bodily_Emoji	0.0196	0.2536	Positive
Para_Nonbodily_Emoji	0.0179	0.3811	Positive
Para_Tactile_Emoticon	0.0154	0.5765	Positive
Para_Alphahaptics	0.0094	0.7787	Positive
Para_Spelling	0.0072	0.9765	Positive

5. Discussion and Conclusion

This study examined how the textual sentiment and paralanguage embedded in bot-generated content shape user evaluations in online communities, providing insights into the mechanisms that drive positive social interactions with bots. Drawing on the theoretical framework of the ELM, we conducted text and sentiment analysis on comments made by social bots that received consistent “good bot” or “bad bot” responses using eight machine learning models and identified LightGBM as the optimal algorithm for the identification of good bots. Our results indicate that textual sentiment, processed through the central route with deeper cognitive elaboration, has a considerably stronger influence on “good bot” evaluations than textual paralanguage, which operates via the peripheral route. Taken together, this study advances the understanding of the dynamics of positive human-bot interactions in online communities.

5.1. Key Findings

Our analysis revealed several important findings. First, the fear expressed in bot-generated content emerged as the most influential textual sentiment in predicting a “good bot” response, followed by trust, disgust, surprise, anger, joy, anticipation, and sadness. Specifically, of the positive emotions, words expressing trust, joy, and anticipation were all positively associated with “good bot” evaluations. This aligns with prior research suggesting that positive emotional expressions by AI agents can elicit emotional contagion, leading to more favorable perceptions (Han et al., 2023). Interestingly, certain negative emotions (i.e., fear, anger, and sadness) also showed a positive association with "good bot" evaluations. This may appear counterintuitive, but it can be understood in terms of empathic engagement. Emotional content that evokes sympathy, such as sadness, or signals moral outrage, such as anger, may lead users to perceive the bot as socially attuned or morally aligned, thus strengthening their relational trust and evaluative judgments (Chen et al., 2025; Park et al., 2024). In contrast, the negative association of disgust and surprise with "good bot" evaluations was reflective of a more complex dynamic. Disgust, often considered a visceral response to moral or physical impurity (Haidt, 2003), may signal aggression, condemnation, or divisiveness, potentially making bots appear to be malicious or socially inappropriate actors rather than helpful contributors. Surprise, though valence-ambiguous (Neta and Kim, 2023), may trigger user skepticism about bot-generated content due to increased uncertainty and perceived artificiality. Taken together, these findings answer RQ1 by demonstrating that different types of textual sentiment embedded in bot-generated comments influence user evaluations through the central route.

Secondly, we found that the five most important textual paralanguage features were the total number of all textual paralanguage nonverbal parts of speech, tempo (e.g., "A grrrreat person"), censorship (e.g., "OUT! OUT, DAMN YOU!"), alphakinesics (e.g., "*smile*"), and bodily emoticons (e.g., ✌️). In contrast, the number of alphahaptics (e.g., "1st Cakeday") and spelling (e.g., "C.O.L.O.U.R.S.") were less important in identifying good bots. Specifically, we observed that there was a negative relationship between the total number of all textual paralanguage nonverbal parts of speech and "good bot" evaluations. A possible explanation is that social bots excessively use nonverbal parts of speech to attract users' attention, which reduces the interaction between the users and bots, potentially leading to "bad bot" judgments. In contrast, tempo, censorship, alphakinesics, and bodily emoticons were positively associated with "good bot" evaluations. This could be because this textual paralanguage acts as an anthropomorphic feature of social bots, which makes it easier for users to feel that they are real when interacting with them (Nass and Moon, 2000), and users are more likely to be persuaded, respond positively to the bots, and willing to continue interacting (Feine et al., 2019). Overall, these findings address RQ2 by highlighting how different forms of textual paralanguage shape user perceptions via the peripheral route.

Finally, our findings indicate that the textual sentiment conveyed through the central route has a more influential impact on user evaluations of social bots than textual paralanguage via the peripheral route, as evidenced by the higher cumulative SHAP values for textual sentiment. This dominance of textual sentiment can be explained through the lens of verbal superiority; textual sentiment tends to offer more explicit, nuanced, and contextually grounded emotional information that facilitates users' deliberative evaluation under uncertainty (Blanco et al., 2010). In contrast, textual paralanguage often conveys affective meanings that are context-dependent or ambiguous, with inconsistent mappings to sentiment valence, reducing its predictive clarity compared to verbal sentiment (Luangrath et al., 2017; Maiberger et al., 2024). In summary, these findings answer RQ3 by demonstrating that textual sentiment processed via the central route exerts a stronger influence on user evaluations of social bots than does textual paralanguage via the peripheral route, thus highlighting the key affective cues that promote positive social interactions.

5.2. Theoretical Implications

This study provides several theoretical insights for the literature on social bots and online persuasion. First, our study extends the application of the ELM by providing a more refined understanding of how the textual sentiment and paralanguage embedded in bot-generated content shape user evaluations in online communities. While prior studies on online persuasion based on the ELM have primarily examined human-generated content, such as advertisements (Shao et al., 2023) and customer reviews (Chou et al., 2022), limited research has employed this framework to investigate how users process persuasive messages in human-bot interactions. To address this gap, we conducted textual sentiment and paralanguage analysis to explore what leads social bots to be perceived as "good bots," with textual sentiment conceptualized as the central route and textual paralanguage as the peripheral route. Our findings extend the boundary conditions of the ELM by showing that bot-generated persuasive content can trigger the central and peripheral routes, suggesting that the model can be applied to non-human communicators.

Second, we integrated a machine learning approach with the dual-route framework of the ELM to systematically examine how textual sentiment and paralanguage shape user evaluations of social bots. Machine learning models can offer significant advantages in handling large, complex datasets by uncovering patterns and making predictions with minimal human intervention, thus enhancing decision-making processes and improving overall efficiency. Specifically, we evaluated the prediction performance of several widely used machine learning methods, with the assessment metrics collectively indicating that LightGBM was best suited to identifying good bots. Moreover, the application of SHAP-based interpretability techniques enabled us to quantify and compare the relative contribution of textual sentiment through the central route and textual paralanguage via the peripheral route. This integration enriches our understanding of how textual sentiment and paralanguage operate through the ELM's dual routes to identify good bots and foster positive social interactions.

Third, our study contributes to the social bot literature by advancing the understanding of how good bots promote positive social interactions within online communities. Past studies on social bots have explored the effects of malicious bot content (Benjamin and Raghu, 2023; He et al., 2025; Hwang et al., 2024), while our research focuses on bot-generated content that influences users' evaluations of good bots. Our findings show that textual sentiment and paralanguage play an important role in interpreting good bots, with textual sentiment demonstrating greater significance. Thus, our study provides useful information for bot design by offering guidelines for the implementation of machine learning in the evaluation of social bots and the factors that promote positive social interactions within online communities.

5.3. Practical Implications

Maintaining the health of online communities and ensuring member retention are critical objectives with significant practical and scholarly value (Kuem et al., 2020). In line with this, our study provides guidelines for the design of social bots. To maintain harmony within online communities and foster positive social interactions, it is essential to know what factors are associated with good bots. In terms of textual sentiment, increasing the levels of fear, trust, anger, joy, anticipation, and sadness and reducing the levels of disgust and surprise conveyed by social bots' comments are more likely to see them perceived as good bots by users. In terms of textual paralanguage, a higher total number of all textual paralanguage nonverbal parts of speech and a higher total number of all emojis in the comments of social bots can lead to classification as a bad bot. In contrast, if the comments contain specific textual paralanguage such as emphasis (e.g., "Thank you!!!!") and stress (e.g., "I AM A BOT."), social bots are more likely to be seen as good bots.

5.4. Limitations and Future Directions

Our research has several limitations that future research could address. First, while we investigate the role of various forms of textual sentiment and paralanguage in the evaluation of good bots, additional content-related factors such as timing and topic relevance may influence user evaluations. Future studies could incorporate these factors to provide a more comprehensive understanding of how users assess bot credibility and intention. Second, although our findings suggest that textual sentiment is more likely to exert influence through the central route whereas textual paralanguage may operate through the peripheral route within the ELM framework, future work could employ controlled experiments that manipulate bot-generated messages to directly validate whether these cues indeed trigger distinct processing routes, thereby offering stronger causal evidence for the proposed mechanisms. Third, our analysis concentrated on textual cues, yet visual and multimodal signals (e.g., profile images, formatting styles, and embedded media) are increasingly integral to online interactions. Exploring how these cues interact with textual content may help uncover more complex mechanisms of persuasion in human-bot interactions. Finally, while Reddit provides a rich and diverse dataset, its platform-specific norms and community structures may limit the generalizability of our findings. In particular, social bots may be evaluated differently on platforms such as YouTube, TikTok, or Instagram. Cross-platform studies would enhance the ecological validity and identify context-dependent variation in human-bot interactions. Therefore, it is strongly recommended to explore other factors that potentially influence the evaluation of social bots using data from various social media platforms.

Acknowledgment

This research was supported by the National Natural Science Foundation of China (Grant No. 72404175), the Social Science Planning Research Project of Shandong Province (Grant No. 25CLJJ14), and the Graduate Science Research Fund of the Business School, Shandong University. We gratefully acknowledge the insightful comments and suggestions from the editors and anonymous reviewers, which have substantially improved this article. We also thank the members of the Star-lights Research Team (https://www.fengliu.net/) for their valuable research assistance on earlier versions of this manuscript.

REFERENCES

Altman, D., Yom-Tov, G. B., Olivares, M., Ashtar, S., & Rafaeli, A. (2021). Do customer emotions affect agent speed? An empirical study of emotional load in online customer contact centers. *Manufacturing & Service Operations Management*, *23*(4), 854–875. https://doi.org/10.1287/msom.2020.0897

Ball, G. P., Bavafa, H., Blanco, C. C., Park, H., & Wowak, K. D. (2025). Gender and serious drug recalls: A textual sentiment analysis of drug reviews on WebMD. *Production and Operations Management*, *34*(4), 698–710. https://doi.org/10.1177/10591478241256644

Bao, Z., & Wang, D. (2021). Examining consumer participation on brand microblogs in China: Perspectives from elaboration likelihood model, commitment–trust theory and social presence. *Journal of Research in Interactive Marketing*, *15*(1), 10–29. https://doi.org/10.1108/jrim-02-2019-0027

Beattie, A., Edwards, A. P., & Edwards, C. (2020). A bot and a smile: Interpersonal impressions of chatbots and humans using emoji in computer-mediated communication. *Communication Studies*, *71*(3), 409–427. https://doi.org/10.1080/10510974.2020.1725082

Benjamin, V., & Raghu, T. S. (2023). Augmenting social bot detection with crowd-generated labels. *Information Systems Research*, *34*(2), 487–507. https://doi.org/10.1287/isre.2022.1136

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 US Presidential election online discussion. *First Monday*, *21*(11). https://doi.org/10.5210/fm.v21i11.7090

Bhattacherjee, A., & Sanford, C. (2006). Influence processes for information technology acceptance: An elaboration likelihood model. *MIS Quarterly*, *30*(4), 805–825. https://doi.org/10.2307/25148755

Blanco, C. F., Sarasa, R. G., & Sanclemente, C. O. (2010). Effects of visual and textual information in online product presentations: Looking for the best combination in website design. *European Journal of Information Systems*, *19*(6), 668–686. https://doi.org/10.1057/ejis.2010.42

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Cai, M., Luo, H., Meng, X., Cui, Y., & Wang, W. (2023). Network distribution and sentiment interaction: Information diffusion mechanisms between social bots and human users on social media. *Information Processing & Management*, *60*(2), 103197. https://doi.org/10.1016/j.ipm.2022.103197

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Chen, Q., Gong, Y., Lu, Y., & Tang, J. (2022). Classifying and measuring the service quality of AI chatbot in frontline service. *Journal of Business Research*, *145*, 552–568. https://doi.org/10.1016/j.jbusres.2022.02.088

Chen, Q., Yin, C., & Gong, Y. (2025). Would an AI chatbot persuade you: An empirical answer from the elaboration likelihood model. *Information Technology & People*, *38*(2), 937–962. https://doi.org/10.1108/ITP-10-2021-0764

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Chen, W., Pacheco, D., Yang, K. C., & Menczer, F. (2021). Neutral bots probe political bias on social media. *Nature Communications*, *12*(1), 5580. https://doi.org/10.1038/s41467-021-25738-6

Cheng, X., Bao, Y., Zarifis, A., Gong, W., & Mou, J. (2021). Exploring consumers’ response to text-based chatbots in e-commerce: The moderating role of task complexity and chatbot disclosure. *Internet Research*, *32*(2), 496–517. https://doi.org/10.1108/intr-08-2020-0460

Cheng, X., Cohen, J., & Mou, J. (2023). AI-enabled technology innovation in e-commerce. *Journal of Electronic Commerce Research*, *24*(1), 1–6.

Chou, Y. C., Chuang, H. H. C., & Liang, T. P. (2022). Elaboration likelihood model, endogenous quality indicators, and online review helpfulness. *Decision Support Systems*, *153*, 113683. https://doi.org/10.1016/j.dss.2021.113683

Cyr, D., Head, M., Lim, E., & Stibe, A. (2018). Using the elaboration likelihood model to examine online persuasion through website design. *Information & Management*, *55*(7), 807–821. https://doi.org/10.1016/j.im.2018.03.009

Dehnert, M., & Mongeau, P. A. (2022). Persuasion in the age of artificial intelligence (AI): Theories and complications of AI-based persuasion. *Human Communication Research*, *48*(3), 386–403. https://doi.org/10.1093/hcr/hqac006

Delkhosh, F., Gopal, R. D., Patterson, R. A., & Yaraghi, N. (2023). Impact of bot involvement in an incentivized blockchain-based online social media platform. *Journal of Management Information Systems*, *40*(3), 778–806. https://doi.org/10.1080/07421222.2023.2229124

Dietterich, T. G. (2000). An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine Learning*, *40*(2), 139–157. https://doi.org/10.1023/A:1007607513941

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies*, *132*, 138–161. https://doi.org/10.1016/j.ijhcs.2019.07.009

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, *59*(7), 96–104. https://doi.org/10.1145/2818717

Gao, Y., Zhang, M. M., & Lysyakov, M. (2025). Does social bot help socialize? Evidence from a microblogging platform. *Information Systems Research.* Advance online publication. https://doi.org/10.1287/isre.2024.1089

Geiger, R. S. (2014). Bots, bespoke, code and the materiality of software platforms. *Information, Communication & Society*, *17*(3), 342–356. https://doi.org/10.1080/1369118x.2013.873069

Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences* (pp. 852–870). Oxford University Press.

Halfaker, A., & Riedl, J. (2012). Bots and cyborgs: Wikipedia's immune system. *Computer, 45*(3), 79–82. https://doi.org/10.1109/mc.2012.82

Han, E., Yin, D., & Zhang, H. (2023). Bots with feelings: Should AI agents express positive emotion in customer service? *Information Systems Research, 34*(3), 1296–1311. https://doi.org/10.1287/isre.2022.1179

Han, J., Pei, J., & Tong, H. (2022). *Data mining: Concepts and techniques* (4th ed.). Elsevier

Harrell, F. E. (2015). *Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis*. Springer.

He, Q., Hong, Y., & Raghu, T. S. (2025). Platform governance with algorithm-based content moderation: An empirical study on Reddit. *Information Systems Research, 36*(2), 1078–1095. https://doi.org/10.1287/isre.2021.0036

Heredia, B., Prusa, J. D., & Khoshgoftaar, T. M. (2018). The impact of malicious accounts on political tweet sentiment. In *2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC)* (pp. 197–202). IEEE. https://doi.org/10.1109/CIC.2018.00035

Hukal, P., Berente, N., Germonprez, M., & Schecter, A. (2019). Bots coordinating work in open source software projects. *Computer, 52*(9), 52–60. https://doi.org/10.1109/MC.2018.2885970

Hwang, E. H., & Lee, S. (2024). A nudge to credible information as a countermeasure to misinformation: Evidence from Twitter. *Information Systems Research, 36*(1), 621-636. https://doi.org/10.1287/isre.2021.0491

Jockers, M. (2016). *Syuzhet: Extracts sentiment and sentiment derived plot arcs from text* (R package version 1.0.7). [Computer software] https://cran.r-project.org/web/packages/syuzhet/index.html

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems 30* (pp. 3146–3154).

Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences, 111*(24), 8788–8790. https://doi.org/10.1073/pnas.1320040111

Kuem, J., Khansa, L., & Kim, S. S. (2020). Prominence and engagement: Different mechanisms regulating continuance and contribution in online communities. *Journal of Management Information Systems, 37*(1), 162–190. https://doi.org/10.1080/07421222.2019.1705510

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer. https://doi.org/10.1007/978-1-4614-6849-3

Latah, M. (2020). Detection of malicious social bots: A survey and a refined taxonomy. *Expert Systems with Applications, 151*, 113383. https://doi.org/10.1016/j.eswa.2020.113383

Lee, D., Oh, K. J., & Choi, H. J. (2017). The chatbot feels you-a counseling service using emotional response generation. In *2017 IEEE International Conference on Big Data and Smart Computing (BigComp)* (pp. 437–440). IEEE. https://doi.org/10.1109/BIGCOMP.2017.7881752

Liu, M., Tang, X., Xia, S., Zhang, S., Zhu, Y., & Meng, Q. (2023). Algorithm aversion: Evidence from ridesharing drivers. *Management Science*. Advance online publication. https://doi.org/10.1287/mnsc.2022.02475

Luangrath, A. W., Peck, J., & Barger, V. A. (2017). Textual paralanguage and its implications for marketing communications. *Journal of Consumer Psychology, 27*(1), 98–107. https://doi.org/10.1016/j.jcps.2016.05.002

Luangrath, A. W., Xu, Y., & Wang, T. (2023). Paralanguage classifier (PARA): An algorithm for automatic coding of paralinguistic nonverbal parts of speech in text. *Journal of Marketing Research*, *60*(2), 388–408. https://doi.org/10.1177/00222437221116058

Ludwig, S., De Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing, 77*(1), 87–103. https://doi.org/10.1509/jm.11.0560

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st Conference on Neural Information Processing System* (pp. 4765–4774).

Maiberger, T., Schindler, D., & Koschate-Fischer, N. (2024). Let's face it: When and how facial emojis increase the persuasiveness of electronic word of mouth. *Journal of the Academy of Marketing Science, 52*(1), 119–139. https://doi.org/10.1007/s11747-023-00932-8

Massanari, A. L. (2016). Contested play: The culture and politics of Reddit bots. In *Socialbots and Their Friends* (pp. 126–143). Routledge. https://doi.org/10.4324/9781315637228-13

McHaney, R., Tako, A., & Robinson, S. (2018). Using LIWC to choose simulation approaches: A feasibility study. *Decision Support Systems, 111*, 1–12. https://doi.org/10.1016/j.dss.2018.04.002

Mejia, J., Mankad, S., & Gopal, A. (2021). Service quality using text mining: Measurement and consequences. *Manufacturing & Service Operations Management, 23*(6), 1354–1372. https://doi.org/10.1287/msom.2020.0883

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Neta, M., & Kim, M. J. (2023). Surprise as an emotion: A response to Ortony. *Perspectives on Psychological Science, 18*(4), 854–862. https://doi.org/10.1177/17456916221132789

Nguyen, A., Rai, A., & Maruping, L. (2024). Understanding the unintended effects of human-machine moderation in addressing harassment within online communities. *Journal of Management Information Systems, 41*(2), 341–366. https://doi.org/10.1080/07421222.2024.2340831

Oh, H., Goh, K. Y., & Phan, T. Q. (2023). Are you what you tweet? The impact of sentiment on digital news consumption and social media sharing. *Information Systems Research, 34*(1), 111–136. https://doi.org/10.1287/isre.2022.1112

Orazi, D. C., Ranjan, B., & Cheng, Y. (2023). Non-face emojis in digital marketing: Effects, contingencies, and strategic recommendations. *Journal of the Academy of Marketing Science, 51*(3), 570–597. https://doi.org/10.1007/s11747-022-00917-z

Park, G., Chung, J., & Lee, S. (2024). Scope and limits of AI fundraisers: Moderated serial multiple mediation model between artificial emotions and willingness to donate via humanness and empathy. *Technological Forecasting and Social Change, 201*, 123211. https://doi.org/10.1016/j.techfore.2024.123211

Pennebaker J. W., Booth R. J., Boyd R. L., Francis M. E. (2015). *Linguistic inquiry and word count: LIWC 2015 operator's manual*. Pennebaker Conglomerates.

Pentina, I., Hancock, T., & Xie, T. (2023). Exploring relationship development with social chatbots: A mixed-method study of replika. *Computers in Human Behavior, 140*, 107600. https://doi.org/10.1016/j.chb.2022.107600

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer-Verlag.

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research, 10*(2), 135–146. https://doi.org/10.1086/208954

Rodríguez-Hidalgo, C., Tan, E. S., & Verlegh, P. W. (2017). Expressing emotions in blogs: The role of textual paralinguistic cues in online venting and social sharing posts. *Computers in Human Behavior, 73*, 638–649. https://doi.org/10.1016/j.chb.2017.04.007

Safadi, H., Lalor, J. P., & Berente, N. (2024). The effect of bots on human interaction in online communities. *MIS Quarterly, 48*(3), 1279-1296. https://doi.org/10.25300/MISQ/2023/17901

Salge, C. A. D. L., & Karahanna, E. (2018). Protesting corruption on Twitter: Is it a bot or is it a person? *Academy of Management Discoveries, 4*(1), 32–49. https://doi.org/10.5465/amd.2015.0121

Salge, C. A. D. L., Karahanna, E., & Thatcher, J. B. (2022). Algorithmic processes of social alertness and social transmission: How bots disseminate information on Twitter. *MIS Quarterly, 46*(1), 229–259. https://doi.org/10.25300/MISQ/2021/15598

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K. C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications, 9*(1), 1–9. https://doi.org/10.1038/s41467-018-06930-7

Shao, Z., Zhang, L., Pan, Z., & Benitez, J. (2023). Uncovering the dual influence processes for click-through intention in the mobile social platform: An elaboration likelihood model perspective. *Information & Management, 60*(5), 103799. https://doi.org/10.1016/j.im.2023.103799

Sheehan, B., Jin, H. S., Martin, B., & Kim, H. J. (2024). Wow! Interjections improve chatbot performance: The mediating role of anthropomorphism and perceived listening. *Communication Research, 51*(7), 891–917. https://doi.org/10.1177/00936502241273259

Smith, C. E., Alam, I., Tan, C., Keegan, B. C., & Blanchard, A. L. (2022). The impact of governance bots on sense of virtual community: Development and validation of the gov-bots scale. *Proceedings of the ACM on Human-Computer Interaction, 6*(CSCW2), 1–30. https://doi.org/10.1145/3555563

Song, M., Zhu, S., Duan, Y., Xing, X., & Mou, J. (2025). Escaping the clutches of fake news? Exploring the mechanisms of online opinion climate on social media users' immunity. *Journal of Electronic Commerce Research, 26*(1), 34–49.

Stella, M., Ferrara, E., & De Domenico, M. (2018). Bots increase exposure to negative and inflammatory content in online social systems. *Proceedings of the National Academy of Sciences, 115*(49), 12435–12440. https://doi.org/10.1073/pnas.1803470115

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems, 29*(4), 217–248. https://doi.org/10.2753/MIS0742-1222290408

Sun, R., Liu, F., Li, Y., Wang, R., & Luo, J. (2024). Machine learning for predicting corporate violations: How do CEO characteristics matter? *Journal of Business Ethics, 195*(1), 151–166. https://doi.org/10.1007/s10551-024-05685-0

Sussman, S. W., & Siegal, W. S. (2003). Informational influence in organizations: An integrated approach to knowledge adoption. *Information Systems Research*, *14*(1), 47–65. https://doi.org/10.1287/isre.14.1.47.14767

Sykora, M., Elayan, S., Hodgkinson, I. R., Jackson, T. W., & West, A. (2022). The power of emotions: Leveraging user generated content for customer experience management. *Journal of Business Research*, *144*, 997–1006. https://doi.org/10.1016/j.jbusres.2022.02.048

Tsai, W. H. S., Liu, Y., & Chuan, C. H. (2021). How chatbots' social presence communication enhances consumer engagement: The mediating role of parasocial interaction and dialogue. *Journal of Research in Interactive Marketing*, *15*(3), 460–482. https://doi.org/10.1108/JRIM-12-2019-0200

Tsvetkova, M., Yasseri, T., Pescetelli, N., & Werner, T. (2024). A new sociology of humans and machines. *Nature Human Behaviour*, *8*(10), 1864–1876. https://doi.org/10.1038/s41562-024-02001-8

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, *22*(158), 209–212. https://doi.org/10.1080/01621459.1927.10502953

Yu, S., & Zhao, L. (2024). Emojifying chatbot interactions: An exploration of emoji utilization in human-chatbot communications. *Telematics and Informatics*, *86*, 102071. https://doi.org/10.1016/j.tele.2023.102071

Zheng, J., Yin, G., Tan, Y., & Ding, J. (2024). Does help help? An empirical analysis of social desirability bias in ratings. *Information Systems Research*, *35*(3), 1052–1073. https://doi.org/10.1287/isre.2020.0406

APPENDIXES

Appendix A. Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	50%	Max.
Good Bot	0.8948	0.3068	0.0000	1.0000	1.0000
Text_Anticipation	0.0047	0.0150	0.0000	0.0000	0.2500
Text_Joy	0.0049	0.0156	0.0000	0.0000	0.2500
Text_Surprise	0.0031	0.0133	0.0000	0.0000	0.2500
Text_Trust	0.0057	0.0152	0.0000	0.0000	0.2500
Text_Anger	0.0020	0.0083	0.0000	0.0000	0.2500
Text_Disgust	0.0015	0.0072	0.0000	0.0000	0.2500
Text_Fear	0.0023	0.0087	0.0000	0.0000	0.2500
Text_Sadness	0.0021	0.0076	0.0000	0.0000	0.2500
Para_Stress	0.1243	0.6010	0.0000	0.0000	36.0000
Para_Tempo	0.1364	0.5018	0.0000	0.0000	15.0000
Para_Rhythm	0.0571	0.4981	0.0000	0.0000	22.0000
Para_Emphasis	0.0536	1.1427	0.0000	0.0000	119.0000
Para_Pitch	0.0249	0.2809	0.0000	0.0000	16.0000
Para_Censorship	0.3813	5.3252	0.0000	0.0000	119.0000
Para_Spelling	0.0028	0.0550	0.0000	0.0000	2.0000
Para_Alternant	0.0814	0.3555	0.0000	0.0000	17.0000
Para_Differentiator	0.0296	0.1879	0.0000	0.0000	4.0000
Para_Tactile_Emoticon	0.0088	0.2142	0.0000	0.0000	15.0000
Para_Alphahaptics	0.0079	0.1222	0.0000	0.0000	9.0000
Para_Bodily_Emoji	0.0287	0.5947	0.0000	0.0000	44.0000
Para_Bodily_Emoticon	1.5748	6.4832	0.0000	0.0000	124.0000
Para_Alphakinesics	0.6101	3.4865	0.0000	0.0000	202.0000
Para_Nonbodily_Emoji	0.1705	1.2378	0.0000	0.0000	61.0000
Para_Nonbodily_Emoticon	0.3799	2.8274	0.0000	0.0000	86.0000
Para_Formatting	0.0129	0.1613	0.0000	0.0000	8.0000
Para_Emoji_Count	0.2285	2.5275	0.0000	0.0000	189.0000
Para_Emoji_Index	0.2013	1.6495	0.0000	0.0000	105.0000
Para_Emoticon_Index	1.9635	8.1128	0.0000	0.0000	172.0000
Para_TPL_Index	3.6886	11.3060	0.0000	1.0000	228.0000
Liwc_Wc	47.5964	90.9984	1.0000	24.0000	1593.0000
Liwc_Function	14.2291	35.0591	0.0000	5.0000	799.0000
Liwc_Pronoun	2.7708	6.9756	0.0000	1.0000	171.0000
Liwc_Ppron	1.5649	4.6766	0.0000	0.0000	94.0000
Liwc_I	0.3033	1.0642	0.0000	0.0000	26.0000
Liwc_We	0.0907	0.5965	0.0000	0.0000	31.0000
Liwc_You	0.6823	1.6188	0.0000	0.0000	60.0000
Liwc_Shehe	0.3409	3.3637	0.0000	0.0000	85.0000
Liwc_They	0.1478	0.7229	0.0000	0.0000	23.0000
Liwc_Ipron	1.2034	3.0306	0.0000	0.0000	78.0000
Liwc_Article	2.4331	6.1691	0.0000	1.0000	132.0000
Liwc_Prep	4.2743	11.6050	0.0000	1.0000	237.0000
Liwc_Auxverb	2.5749	5.6858	0.0000	1.0000	166.0000
Liwc_Adverb	0.9265	2.7510	0.0000	0.0000	115.0000

Liwc_Conj	1.5944	4.9679	0.0000	0.0000	92.0000
Liwc_Negate	0.4047	1.2613	0.0000	0.0000	28.0000
Liwc_Social	2.7109	7.3043	0.0000	1.0000	176.0000
Liwc_Family	0.0476	0.3736	0.0000	0.0000	12.0000
Liwc_Friend	0.0729	0.4461	0.0000	0.0000	36.0000
Liwc_Female	0.1991	2.3426	0.0000	0.0000	47.0000
Liwc_Male	0.2536	1.7442	0.0000	0.0000	70.0000
Liwc_Cogproc	3.1918	7.9222	0.0000	1.0000	210.0000
Liwc_Insight	0.6973	1.6595	0.0000	0.0000	29.0000
Liwc_Cause	0.5541	1.7551	0.0000	0.0000	42.0000
Liwc_Discrep	0.4055	1.3069	0.0000	0.0000	47.0000
Liwc_Tentat	0.8278	2.1713	0.0000	0.0000	46.0000
Liwc_Certain	0.3425	1.1023	0.0000	0.0000	56.0000
Liwc_Differ	0.8600	2.5016	0.0000	0.0000	50.0000
Liwc_Percept	0.6376	2.0182	0.0000	0.0000	41.0000
Liwc_See	0.3446	1.0323	0.0000	0.0000	29.0000
Liwc_Hear	0.1441	0.7825	0.0000	0.0000	24.0000
Liwc_Feel	0.1201	0.7080	0.0000	0.0000	19.0000
Liwc_Bio	0.5124	2.6540	0.0000	0.0000	48.0000
Liwc_Body	0.2349	2.1003	0.0000	0.0000	43.0000
Liwc_Health	0.0815	0.5413	0.0000	0.0000	17.0000
Liwc_Sexual	0.0821	0.6996	0.0000	0.0000	28.0000
Liwc_Ingest	0.1304	0.9172	0.0000	0.0000	33.0000
Liwc_Drives	1.7514	5.0488	0.0000	0.0000	104.0000
Liwc_Affiliation	0.4368	1.5625	0.0000	0.0000	49.0000
Liwc_Achiev	0.4292	1.6014	0.0000	0.0000	36.0000
Liwc_Power	0.6303	2.1357	0.0000	0.0000	40.0000
Liwc_Reward	0.3764	1.2419	0.0000	0.0000	31.0000
Liwc_Risk	0.1418	0.6198	0.0000	0.0000	25.0000
Liwc_Focuspast	1.0422	4.0804	0.0000	0.0000	95.0000
Liwc_Focuspresent	2.9912	6.3506	0.0000	1.0000	139.0000
Liwc_Focusfuture	0.2819	1.0582	0.0000	0.0000	25.0000
Liwc_Relativ	3.5489	11.2382	0.0000	1.0000	242.0000
Liwc_Motion	0.4939	1.8282	0.0000	0.0000	38.0000
Liwc_Space	1.9299	6.6316	0.0000	0.0000	117.0000
Liwc_Time	1.0808	3.7082	0.0000	0.0000	114.0000
Liwc_Work	0.6282	2.2378	0.0000	0.0000	89.0000
Liwc_Leisure	0.3002	1.3294	0.0000	0.0000	43.0000
Liwc_Home	0.0696	0.5247	0.0000	0.0000	19.0000
Liwc_Money	0.2295	1.2445	0.0000	0.0000	83.0000
Liwc_Relig	0.0321	0.2753	0.0000	0.0000	15.0000
Liwc_Death	0.0367	0.3075	0.0000	0.0000	19.0000
Liwc_Informal	0.2064	0.9156	0.0000	0.0000	45.0000
Liwc_Swear	0.0724	0.6157	0.0000	0.0000	29.0000
Liwc_Netspeak	0.0915	0.4893	0.0000	0.0000	28.0000
Liwc_Assent	0.0221	0.1804	0.0000	0.0000	12.0000
Liwc_Nonflu	0.0141	0.1288	0.0000	0.0000	3.0000

Appendix B. Hyperparameters for Machine Learning Models in Predicting a "Good Bot" Classification

	Specification	The optimal parameters of Baseline (LIWC)	The optimal parameters of LIWC+Textual Sentiment	The optimal parameters of LIWC+Textual Paralinguistic	The optimal parameters of All Features
LR	• C ∈ {0.1, 1, 10, 100} • max_iter ∈ {10, 200, 500, 1000}	• C = 1 • max_iter = 500	• C = 10 • max_iter = 500	• C = 0.1 • max_iter = 500	• C = 100 • max_iter = 500
SVM	• C ∈ {0.1, 1, 10, 100} • gamma ∈ {scale, auto}	• C = 100 • gamma = scale	• C = 100 • gamma = scale	• C = 100 • gamma = scale	• C = 100 • gamma = scale
DT	• max_depth ∈ {None, 10, 20, 30} • min_samples_split ∈ {5, 10, 50, 100}	• max_depth = None • min_samples_split = 5	• max_depth = None • min_samples_split = 5	• max_depth = None • min_samples_split = 5	• max_depth = None • min_samples_split = 5
MLP	• hidden_layer_sizes ∈ {50, 100, 150, 200} • learning_rate ∈ {0.01, 0.1, 0.3, 0.5}	• hidden_layer_sizes = 200 • learning_rate = 0.01	• hidden_layer_sizes = 200 • learning_rate = 0.01	• hidden_layer_sizes = 200 • learning_rate = 0.01	• hidden_layer_sizes = 200 • learning_rate = 0.01
RF	• n_estimators ∈ {100, 300, 500, 1000} • max_depth ∈ {None, 10, 20, 30} • min_samples_split ∈ {5, 10, 50, 100}	• n_estimators = 300 • max_depth = None • min_samples_split = 3	• n_estimators = 1000 • max_depth = None • min_samples_split = 5	• n_estimators = 1000 • max_depth = None • min_samples_split = 5	• n_estimators = 500 • max_depth = None • min_samples_split = 5
AdaBoost	• n_estimators ∈ {100, 300, 500, 1000} • learning_rate ∈ {0.01, 0.1, 0.3, 0.5}	• n_estimators = 1000 • learning_rate = 0.5	• n_estimators = 1000 • learning_rate = 0.5	• n_estimators = 1000 • learning_rate = 0.5	• n_estimators = 1000 • learning_rate = 0.5
LightGBM	• n_estimators ∈ {100, 300, 500, 1000} • learning_rate ∈ {0.01, 0.1, 0.3, 0.5}	• n_estimators = 1000 • learning_rate = 0.1	• n_estimators = 1000 • learning_rate = 0.1	• n_estimators = 1000 • learning_rate = 0.1	• n_estimators = 1000 • learning_rate = 0.1
XGBoost	• n_estimators ∈ {100, 300, 500, 1000} • learning_rate ∈ {0.01, 0.1, 0.3, 0.5}	• n_estimators = 1000 • learning_rate = 0.5	• n_estimators = 1000 • learning_rate = 0.5	• n_estimators = 1000 • learning_rate = 0.1	• n_estimators = 1000 • learning_rate = 0.5