Spring 1-2023

Man vs machine – Detecting deception in online reviews

Maria Petrescu
petrescm@erau.edu

Scholarly Commons Citation
Petrescu, M. (2023). Man vs machine – Detecting deception in online reviews. *Journal of Business Research, 154*(). https://doi.org/10.1016/j.jbusres.2022.113346

Man vs. machine – detecting deception in online reviews

ABSTRACT

This study focused on three main research objectives: analyzing the methods used to identify deceptive online consumer reviews, evaluating insights provided by multi-method automated approaches based on individual and aggregated review data, and formulating a review interpretation framework for identifying deception. The theoretical framework is based on two critical deception-related models, information manipulation theory and self-presentation theory. The findings confirm the interchangeable characteristics of the various automated text analysis methods in drawing insights about review characteristics and underline their significant complementary aspects. An integrative multi-method model that approaches the data at the individual and aggregate level provides more complex insights regarding the quantity and quality of review information, sentiment, cues about its relevance and contextual information, perceptual aspects, and cognitive material.

Man vs. machine – detecting deception in online reviews

1. Introduction

Online reviews are essential in consumer evaluations of products and their purchase decisions and are one of the primary sources of consumption information for modern society (Chakraborty & Bhat, 2018; De Langhe et al., 2016; Gössling, Hall & Andersson, 2018). Consumer comments regarding their purchases and service experiences can affect product sales and stock value as long as there is an equilibrium regarding the level of expected and perceived review credibility and consumer suspicion of deception (Moon, Kim, & Iacobucci, 2021; Riquelme, Román, & Iacobucci, 2016; Zhuang, Cui, & Peng, 2018).

Researchers have noted that online reviews influence 93% of individuals in their market decisions and emphasize the importance of peer-generated digital content in consumer decisions (Gentina, Chen, & Yang, 2021; Lee, Qiu, & Whinston, 2018; Schoenmueller, Netzer, & Stahl, 2020). Considering the widespread use of online reviews and their impact on consumer behavior, the level of deception and manipulation techniques have also increased (Cardoso, Silva, & Almeida, 2018; Hajek & Sahut, 2022; Malbon, 2013; Steward et al., 2020). Many marketers find themselves in a prisoner's dilemma, in which engaging in manipulation is the most rational choice in different competitive situations (Gössling, Hall & Andersson, 2018; Hajek & Sahut, 2022; Hu et al., 2012).

There are numerous instances in which we encounter deception in consumer reviews, including businesses incentivizing consumers to write about their brand and competing brands, as well as using modern digital entrepreneurs and online reputation management companies to manage this process (Choi et al., 2017; Ivanova & Scholz, 2017; Luca & Zervas, 2016; Petrescu et al., 2022; Sahut, Iandoli, & Teulon, 2021). Deceptive communication takes different forms, including automatically filtering out negative reviews, misleading aggregation algorithms, artificially written fake reviews, and incentivized consumer comments, which makes it difficult for consumers to evaluate this type of information (Dellarocas, 2006; Hu et al., 2011, 2012; Moon, Kim, & Iacobucci, 2021; Munzel, 2016; Plotkina, Munzel, & Pallud, 2020).

However, despite a significant number of research studies in business and data science on fake review detection, there is still no consensus on the efficacy of automated text classification methods and the best approaches to be employed by marketers and consumers, especially regarding the use of real-life vs. artificial reviews, valence, and procedure of analysis (Cardoso, Silva, & Almeida, 2018; Hajek & Sahut, 2022). Researchers note that semantic meaning, context and sentiment information are essential in evaluating deceptive reviews and review and reviewer characteristics (Hajek & Sahut, 2022; Heydari et al., 2015). Nevertheless, there is also a need for comprehensive theoretical models that focus on textual characteristics as indicators of review authenticity vs. deception (Banerjee & Chua, 2017; Petrescu et al., 2022). This is especially important considering that humans have a much lower accuracy of deception detection in online reviews than automated tools, even when primed with information on cues of fake online reviews (Plotkina, Munzel, & Pallud, 2020).

This study focuses on critical aspects related to deception identification in online reviews with three objectives: to analyze the current state of algorithms and tools used in the identification of fake consumer reviews, to evaluate a combination of insights provided by a multi-method automated approach based on individual and aggregated data, and to formulate a theoretical consumer review interpretation framework for identifying deception by stakeholders with the use of automated software analysis, based on our multi-method analysis and previous multidisciplinary theories.

The paper starts with analyzing the current research and practices on detecting fake reviews, considering consumers' options for manual detection and automated software analysis. Following this analysis and considering a theoretical framework based on self-presentation theory - SPT (Banerjee & Chua, 2017; DePaulo et al., 2003) and information manipulation theory – IMT (McCornack, 1992; McCornack et al., 2014), we perform a multi-method analysis of computational linguistics that draws insights about consumer review credibility and tests the application of an updated theoretical framework on online deceptive consumer messages. Finally, we formulate an integrative review interpretation framework for detecting fake reviews in online digital communication that considers all stakeholders of the communication ecosystem and can represent a theoretical and practical base for consumers, marketers, and researchers and further algorithm development.

The paper extends existing knowledge on deception use and identification in online consumer reviews and develops a theoretical framework on online deception, based on a multidisciplinary approach (De Bakker et al., 2019; Whetten, Felin, & King, 2009) involving marketing, communication, sociology, and the management of information systems theoretical elements. As recommended by research, its contributions represent an improved understanding of the online review phenomenon not only for theory, but also a relevant topic for business practitioners (Corley & Gioia, 2011). The mixed-methods approach, including both qualitative and quantitative methods analyzing digital data, allows us to better explain this phenomenon, as well as better answer questions related to "what, "how" and "why" (Crane, Henriques, & Husted, 2018; Gehman et al., 2018; Gioia, Corley, & Hamilton, 2012; Whetten, 1989).

2. Study 1: The state of deception in online consumer reviews

Marketing researchers and practitioners have tried to identify and quantify deception in online consumer reviews through various computer-assisted methods, considering consumer difficulties in detecting the various cues in the emotional and cognitive states that accompany it (DePaulo et al., 2003; Hartwig and Bond, 2011; Hu et al., 2011, 2012; Peng et al., 2016; Toma and Hancock, 2012). Specialists are also working on theoretical approaches, algorithms, and models to assess the degree of deception based on lexical and semantic characteristics, including review length, complexity, readability, subjectivity, valence, and sentiment (Banerjee & Chua, 2017; Mukherjee et al., 2012; Ott et al., 2012).

To provide an integrative review of the current state of deception detection studies and formulate a comprehensive theoretical framework, we first start with a bibliometric analysis of top business studies focused on analyzing deceptive consumer reviews. The preliminary analysis included 120 articles from Web of Science that focus on deception detection in online reviews, published in ranked marketing, communications, and the management of information systems journals, as included in Appendix 1. A bibliometric analysis helps identify key topics emerging in a field of research (Ferreira, 2018; Ferreira et al., 2016; Kapoor, Dwivedi, & Williams, 2014).

We performed network analysis and created a bibliometric of keyword co-occurrences focused on methods used to detect deception using VOSviewer. The network visualization of keyword co-occurrence is presented in Figure 1 and Table 1. The links and their strength attributes indicate the number of links of an item with other items and the total strength of an item's links (Waltman, Van Eck, & Noyons, 2010; Van Eck & Waltman, 2010).

***(Insert Figure 1** here)*

The results of the bibliometric analysis emphasize different approaches to deception detection discussed by researchers in the context of online reviews, as shown in Figure 1 and the quantitative details in Table 1. The results focus on AI-based deception detection tools, including machine learning, deep learning algorithms, and neural network analysis. We see different text analysis and classification methods and mention of cues, credibility, outlier detection, and feature extraction.

***(Insert Table 1** here)*

The bibliometric analysis results exhibit various technology-assisted deception detection tools based on distinct methods and different theoretical philosophies stemming from source-based identification, linguistic analysis, semantic anomaly identification, and opinion spam detection. The diversity of methods employed in these analyses creates an opportunity for this study to theoretically integrate the bases of deception detection in online reviews based on self-presentation theory (Banerjee & Chua, 2017; DePaulo et al., 2003) and information manipulation theory (McCornack, 1992; McCornack et al., 2014). In the following sections, we continue evaluating the critical methods of deception analysis based on a multi-method computational linguistic and semantic analysis. Our purpose is to formulate a comprehensive review interpretation framework that incorporates multiple methods, elements, and levels of analysis that both consumers and computerized algorithms can use as a base to identify fake reviews.

3. Unstructured review analysis

As reflected in the clusters identified in Study 1 (Table 1), there are numerous automated text analysis and sentiment mining methods and techniques at the disposal of marketers and consumers, including lexicon and statistical learning-based (Chatterjee et al., 2021). Previous studies have classified automated deception detection methods in online reviews in a few major categories, reflected in Table 1 too: machine learning, neural network-based, and pattern-mining approaches (Hajek & Sahut, 2022; Plotkina, Munzel, & Pallud, 2020). Opinion mining and sentiment analysis are some of the fundamental analyses performed in this context.

Different supervised machine learning techniques used in marketing can include logistic regression, naive Bayes analysis, and k-nearest neighbor. In contrast, unsupervised machine learning techniques can include the unsupervised topic-sentiment joint probabilistic model, statistics-based unsupervised clustering algorithm, and lexicon-based unsupervised model (Wu et al., 2020). These methods are subject to discussions related to their usefulness and role in deception detection in online reviews. Moreover, in this context, another debated topic is related to the use of real-life vs. artificial reviews in the assessment of deception and evaluation of methodological approaches, especially in the data training process for machine learning-based approaches (Cardoso, Silva, & Almeida, 2018; Hajek & Sahut, 2022; Wu et al., 2020).

Finally, the findings in the bibliometric analysis reflect the diversity of AI-based technologies involved in business in general and in review analysis, from broad discussions on AI to narrower uses of deep learning. Figure 2 integrates the differences and relationships among the most used terms of AI-related technology in the studies analyzed. As the broadest term discussed, artificial intelligence incorporates technologies that mimic human intelligence and helps predict and optimize tasks such as speech and image recognition and decision-making (Kavlakoglu, 2020). Machine learning (ML) is a part of AI that incorporates the classical ML component, requiring human experts for feature extraction and algorithm formulation, and the deep learning component, automating feature extraction and able to use unstructured data too. In this structure, a neural network represents a complex deep learning procedure that mimics the human brain, with more than three node “hidden” layers or depths incorporated in its algorithm (Kavlakoglu, 2020).

***(Insert Figure 2** here)*

Nevertheless, while in the case of classical, non-deep machine learning we are able to account for human intervention and determine the key algorithms employed, for deep learning and neural networks we do not necessarily require a labeled dataset and algorithm input, making their “black box” much more complex and difficult to analyze (Kavlakoglu, 2020). Considering the theoretical framework of this paper and the findings in the bibliometric analysis, our key research questions focus on drawing a combination of insights using standard, classical machine learning-based, widely available automated methods of text analysis that allow the identification of algorithms and determination of hierarchy of features. This can represent a step forward in transferring the problem of consumer review deception identification to more complex neural network models.

RQ1: Is there a significant difference between the performance of standard automated text analysis methods?

RQ2: Is there a potential for complementarity between standard automated text analysis methods?

RQ3: How can multiple analysis methods be combined to formulate a review interpretation framework?

4. Theoretical framework

Deception is a deliberate act performed by manipulating information to create or maintain a belief that the communicator knows to be false (DePaulo et al., 2003; Munzel, 2015; Peng et al., 2016; Xiao & Benbasat, 2011). In marketing communications, consumers have specific expectations regarding the characteristics and quality of the message and its credibility, which can be exploited through deceptive content (McCornack, 1992; McCornack et al., 1996, 2014).

When it comes to factors that can help identify deceptive reviews, previous studies have mentioned a lack of details, emotional exaggeration, and variability in valence, as well as review length (Luca & Zervas, 2016; Moon, Kim, & Iacobucci, 2021; Ott et al., 2013). Reviews can be differentiated based on numerous factors, such as comprehensibility, specificity, exaggeration, and negligence, as well as on syntactic elements like structure and format, writing style, and readability (Wu et al., 2020). Information Manipulation Theory (McCornack, 1992; McCornack et al., 1992, 2014) and Self-Presentation Theory (Banerjee & Chua, 2017; DePaulo et al., 2003) provide the most comprehensive and integrative models of deception identification in consumption communication used in both manual and automated consumer review analysis and are used as the bases for the theoretical framework explored and tested in our analysis.

4.1. *Information Manipulation Theory*

Information Manipulation Theory (IMT) states that consumers manipulate information simultaneously along different dimensions, which can be identified based on quantity (amount of information), quality (details), relation (relevance), and communication (style) manner (McCornack, 1992; McCornack & Levine, 1990; McCornack et al., 1996). This theory emphasizes that most of the everyday deceptive discourse includes numerous deceptive elements, including adjusting the amount of relevant information shared, incorporating false information, using irrelevant information, and employing a vague manner of communication (McCornack, 1992; McCornack & Levine, 1990).

Authors found that alterations of amount, veracity, relevance, and the clarity of information impact perceived message deceptiveness (McCornack, 1992; McCornack & Levine, 1990). In this context, studies also found that consumers have a minimal capacity for detecting deception because of a significant truth bias (Levine, Park, & McCornack, 1999; Plotkina, Munzel, & Pallud, 2020).

An updated version of IMT (McCornack et al., 2014) presents a propositional, testable theory of deceptive discourse production, enriches this theoretical framework and brings it up for modern digital communication by integrating elements of linguistics, cognitive neuroscience, speech production, and artificial intelligence. This new version focuses on individual intentional states, cognitive load, and information manipulation, including the intentional nature of deception, which is additionally completed by self-presentation theory.

4.2.Self-Presentation Theory

Self-presentation focuses on how individuals control the way they present themselves and try shaping others' opinions, based on controlled information about themselves, other individuals, and events. In the self-presentation perspective, the authors focus on the assumption that cues to deception are generally weak and authentic messages differ from deceptive ones as a function of perceptual aspects, such as sensory information, contextual details, such as information related to location and time, as well as cognitive information (Banerjee & Chua, 2017; DePaulo et al., 2003).

According to this theoretical framework, deceptive communicators are less forthcoming and provide fewer details and information. Moreover, consumers may also be likely to build an appearance of naturalness to signal low effort to their audience in a strategic manner in their interpersonal communication and self-presentation (Leary, 1995; Smith et al., 2022). Consumers often engage in self-presentational tactics with the purpose of manipulating and misrepresenting themselves to achieve positive outcomes (DeAndrea et al., 2012; Smith et al., 2022; Toma, Hancock, & Ellison, 2008).

Nevertheless, deception cues are not always identifiable and are more visible when consumers are motivated to succeed (Kashy & DePaulo, 1996; DePaulo et al., 2003). In the context of consumer reviews, it is crucial to consider that individuals are more inclined to be deceptive when they have a goal of appearing likable or competent, in which case the content of the deceptive message fluctuates as a function of the self-presentation goal (Feldman, Forrest, & Happ, 2002). Digital self-presentation can include negative, oppositional relationships relevant to consumers and the potential for individuals to place themselves concerning products and services (Schau & Gilly, 2003).

5. Study 2: Methodological comparison and integration

In the second study of this paper, we focus on comparing multiple methods of automated deception detection commonly used in marketing research and practice. Unlike previous studies, we are not making assumptions about the level of deceptions in the reviews we are using, but the purpose of the analysis is to decide which methods provide the most insights and how these approaches can be combined in an integrative model that can then be applied in manual and automated analysis. The analysis in this study is based on real-world reviews from Amazon for well-known brands, including various products in the industries of electronics and cosmetics, a type of modern digital data useful in qualitative and quantitative approaches and allowing for a better understanding of complex business problems (Crane, Henriques, & Husted, 2018). In the overall analysis, we employed 18,113 consumer reviews from 2021-2022 of three types, verified purchases, unverified, and the newly created program Vine reviewers, where Amazon invites to post opinions about products received for free. The reviews were downloaded at the beginning of 2022 and incorporated 20 different products with a number of reviews ranging from 26 to 5000, each product including the three types of reviews of interest, verified, unverified, and Vine incentivized.

Based on the findings from the bibliometric analysis in Study 1 and considering our theoretical framework based on information manipulation theory and self-presentation theory, we selected for this multi-method analysis three widely used methods of analysis, as well as readily available software options that can provide insights regarding quantitative and qualitative characteristics of text content, as well as context-related insights (Hajek & Sahut, 2022; Moon, Kim, & Iacobucci, 2021; Plotkina, Munzel, & Pallud, 2020). The first method includes a computational linguistic analysis that provides quantitative and qualitative information and sentiment insights. The second method is focused on a semantic analysis that can reveal additional contextual information and content themes. Finally, we complement the insights with a simple quantitative readability analysis. This combination of methods contributes to more insightful findings regarding the problem studied and provides more qualitative rigor to the data analysis process (Crane, Henriques, & Husted, 2018; Gioia, Corley, & Hamilton, 2012; Gioia et al., 2022).

5.1.Computational linguistic analysis

In the first step, we perform a computational linguistic analysis for each consumer review downloaded from Amazon. Linguistic Inquiry and Word Count (LIWC2015) has been used often in business and marketing research to evaluate the authenticity and emotionality/sentiment of text content in different circumstances, including social media and consumer reviews (Moon, Kim, & Iacobucci, 2021; Ott et al., 2011; Newman, Pennebaker, Berry, & Richards, 2003; Pennebaker, Boyd, Jordan, & Blackburn, 2015). This software is based on predefined dictionaries, established psychometrically tested scales, and algorithms from the Pennebaker Lab (Humphreys & Wang, 2018; Pennebaker et al., 2015). It includes over 90 indices and summary language variables such as authenticity, analytical thinking, clout, and emotional tone (Newman et al., 2003; Pennebaker et al., 2014). The LIWC2015 analytical thinking index analyzes formal communication in text based on function words and grammar words (Pennebaker et al., 2014; Plotkina et al., 2020). Clout is related to the social aspect of communication and externally focused messages, while tone reflects the overall positivity and the emotional style of the message (Pennebaker et al., 2014, 2001; Tausczik & Pennebaker, 2010). To measure the characteristics of our reviews, we employ previously used variables, including the four major indices provided by LIWC2015, authenticity index, analytics, clout, and tone, as well as consumer time orientation and focus on power in communication, found to be essential in deceptive reviews (Li et al., 2014; Moon, Kim, & Iacobucci, 2021; Ott et al., 2013; Petrescu et al., 2022; Plotkina, Munzel, & Pallud, 2020).

To evaluate the potential for this method to provide insights into the content of reviews without making assumptions, we separated the reviews into the three main categories downloaded, verified (17270), unverified (460), and Vine incentivized (378), and performed a MANOVA analysis using SPSS to assess the differences between these three categories. The results of the multivariate analysis are significant, as presented in Table 2, except for variable power. They show differences in review characteristics between the three types of reviews analyzed.

***(Insert Table 2** here)*

The multiple comparison results of the analysis included in Table 3 emphasize significant differences between review categories, starting with word count, showing longer text for reviewers incentivized through the Vine program, a more positive tone, but lower analytical characteristics. It is also interesting to note that for some indices, such as authenticity and clout, the analysis does not find significant differences between the unverified and incentivized reviews through Vine. The analysis also finds differences in time-anchoring in consumer communication, confirming findings of previous studies (Moon, Kim, & Iacobucci, 2021; Plotkina, Munzel, & Pallud, 2020).

***(Insert Table 3** here)*

To explore our research questions, we continue our analysis of the downloaded reviews with further semantic and sentiment analyses. We are also performing an analysis based on an aggregated level of reviews to investigate the possibility of drawing additional insights from the text.

5.2.Semantic analysis

Leximancer 4.5 is a semantic analysis tool designed to further understand conceptual themes and concepts as a series of associations and links (Krishen, Berezan, & Raab, 2019; Petrescu & Kachen, 2019; Petrescu et al., 2018). This is an unsupervised method of analysis based on deep learning, extracting a transparent three-level network model of meaning from qualitative data (Smith & Humphreys, 2006). Key themes and concepts are reflected in the conceptual map in Figure 3, based on a comparative semantic analysis of the three types of reviews.

***(Insert Figure 3** here)*

Also, the prominence scores for the key themes and concepts are included in Table 4. Prominence scores represent absolute measures of correlation between concept categories and attributes, with a value greater than 1 indicating a purposeful relationship (Smith, 2007; Smith & Humphreys, 2006). The themes of the three main types of comments and their relationships reflect a similar conclusion as the one in the LIWC analysis, showing the unverified and incentivized reviews clustering. The results in Table 4 emphasize the differences among the three categories of reviews analyze and underline the relationship between incentivization and positive content. The value highlighted in Table 4 show a high level of absolute correlation between positive, emotional content incorporating keywords such as easy, love and nice and the category of Vine incentivized reviews. These values show not only the much more positive content among these two types of reviews compared to the other two categories, but also the high level of the correlation effect among positive words and the incentivized category. While verified reviews are more informatively and negatively focused on the functionality of the product reviewed, tech support, time, problems, and money, the incentivized category presents significant relationships only for positive sentiment. The unverified reviews show a combination of significantly correlated variables in Table 4, including both positive sentiment and informative aspects.

***(Insert Table 4** here)*

Moreover, considering the essential themes that are common or distinct among these types of reviews, we can identify similar concepts evaluated in the computational linguistic analysis, such as time and emotional language, as well as informative or analytical clues related to price and product functioning, much more specific for verified reviews. To complete the insights drawn from these qualitative and quantitative analyses, in the next step, we are focusing on the aggregate level of reviews and aspects previously evaluated by the literature, such as sentiment and readability (Moon, Kim, & Iacobucci, 2021; Petrescu et al., 2022; Plotkina, Munzel, & Pallud, 2020).

5.3.Readability analysis

One of the debates regarding the use of readily available deception detection tools on websites like Yelp and Amazon is represented by the fact that the effectiveness of their algorithms is unknown, as they are proprietary (Moon, Kim, & Iacobucci, 2021). On a similar note, the effective yet complex neural network analysis also presents the same issue of the black box behind the algorithm and the need for large datasets for calibration and training (Hajek & Sahut, 2022). In this last step, we are integrating an approach readily available for consumers and marketers that can be used at individual and aggregated levels to evaluate text readability and sentiment. Textalyzer is a simple tool that calculates keyword density and text readability measures (Textalyzer, 2022), as reflected in the first part of Table 5.

***(Insert Table 5** here)*

The indicators in the table include different readability metrics, including the Coleman–Liau index, the Gunning fog index, the SMOG index, and the Automated Readability Index, which approximate the U.S. grade level thought necessary to comprehend the text. The second part of the table reflects the aggregated indicators from the LIWC analysis and the potential to integrate both linguistic and complementary readability measures as evaluators of review content.

6. Findings and discussion

Our analyses in Studies 1 and 2 focused on evaluating the current state of methods and tools used to assess fake consumer reviews and identify potential combinations of insights that a multi-method automated approach based on individual and aggregated data can provide. The bibliometric analysis of the current research and practices on detecting fake reviews considered consumers' options for manual detection and automated software analysis.

The bibliometric analysis reflects numerous classical machine learning-based automated tools of text processing and deception detection, based on distinct analysis methods and theoretical paradigms, as the main clusters emphasized in Table 1 also show. Some critical analysis methods focus on linguistic patterns and sentiment information, content-related statistics and metrics, behavioral and context-related information and content themes in the text. The insights drawn from the numerous studies analyzed also confirm some of the problems emphasized by marketing research, including the lack of integration and aggregation in the level of insights drawn, little knowledge and transparency about algorithms, and difficulty in sourcing real-life and artificial calibrating and training datasets (Hajek & Sahut, 2022; Moon, Kim, & Iacobucci, 2021; Plotkina, Munzel, & Pallud, 2020).

Following the bibliometric analysis, we formulated our three essential research questions based on self-presentation theory (Banerjee & Chua, 2017; DePaulo et al., 2003) and information manipulation theory (McCornack, 1992; McCornack et al., 2014) and performed a multi-method analysis to draws insights about consumer reviews in Study 2. Even though we made no assumptions regarding the characteristics of the three types of reviews analyzed and the purpose of our analysis is to evaluate the level of insights drawn, the findings of the computational linguistic analysis in LIWC show significant statistical differences in the levels of authenticity, analytical style, sentiment, and social orientation among review categories, emphasizing the different position of unverified and incentivized reviews vs. verified ones. The findings in Table 4 reflect a significant relationship between incentivized Amazon reviews through the Vine tool and the likelihood of consumer reviews including positive sentiment evaluations, including emotionally charged keywords such as easy, love, and nice.

Considering our research objectives and calls from previous literature (Hajek & Sahut, 2022; Moon, Kim, & Iacobucci, 2021; Plotkina, Munzel, & Pallud, 2020), we complemented this method with a semantic analysis in Leximancer. This provided an opportunity to explore RQ2 and evaluate the potential for complementarity between automated review analysis methods. The semantic analysis provided a few key findings. First, some of the themes extracted reflect information about the context of the reviews, related to the type of product analyzed (electronic vs. cosmetics), and consumer preoccupations with their most relevant characteristics, which were not emphasized in the linguistic analysis. Second, the themes and keywords reflected in Figure 3 corroborate some of the metrics provided by LIWC, such as an orientation toward time-related aspects, accentuating new concepts, including price and functionality-related topics. Finally, this analysis also provides visual confirmation of the differences and similarities between the three types of reviews analyzed and visualizes the relationships among them, confirming the findings of the linguistic analysis. As formulated in Figure 4, the semantic analysis comes not only as a confirmation but also complements the insights extracted with the previous method. It shows that, even though these methods are different, their results come to support one another and can be used in a complementary fashion, as we initially questioned in RQ1 and RQ2.

(Insert Figure 4 *here)*

Our third analysis focused on readability insights and comparing quantitative findings with the indices obtained in the computational linguistic analysis. The different readability indices provide additional information about the relevance and style of reviews and complement the LIWC results. Moreover, the aggregated-level comparison among the three groups of the text supports the individual-level and the aggregated linguistic insights.

Following our inquiries in RQ3, in Figure 4, we formulate an integrative review interpretation framework for evaluating reviews based on a multi-method approach, reflecting the critical theoretical elements of information manipulation theory and self-presentation theory. From a theoretical point of view, this framework integrates the quantity and quality of information, insights about its relevance and contextual information, and metrics about the clarity of the content (McCornack, 1992; McCornack et al., 1996, 2014). Moreover, this integrative model provides insights into perceptual aspects, contextual details, and cognitive material (Banerjee & Chua, 2017; DePaulo et al., 2003). The integrative framework in Figure 4 can represent a theoretical and practical model for consumers, marketers, researchers, and other stakeholders and is discussed further in the conclusions section.

7. Conclusions and future research

Previous literature and the bibliometric study showed many studies on deception detection and online review analysis but difficulty in managing the diversity of automated text classification methods and selecting the best approaches (Cardoso, Silva, & Almeida, 2018; Hajek & Sahut, 2022). Theory and practice have shown that all methods are useful and insightful, providing linguistic, semantic, and statistical information in evaluating deceptive reviews (Hajek & Sahut, 2022; Heydari et al., 2015). Nevertheless, considering the need for comprehensive theoretical and practical models that can help researchers, practitioners, and consumers, this paper analyzed deceptive consumer evaluation methods' current state and formulated a review interpretation framework provided in Figure 4.

7.1.Research implications

This study has three main research objectives: analyzing the methods used to identify deceptive consumer reviews, evaluating combinations of insights provided by multi-method automated analyses based on individual and aggregated data, and formulating a review interpretation framework for identifying deception. The essential research contribution of the paper is the development of a new theoretical framework reflecting online consumer deception that integrates information manipulation theory (McCornack, 1992; McCornack et al., 1992) and self-presentation theory (Banerjee & Chua, 2017; DePaulo et al., 2003) based on a combination of approaches and text analysis methods.

The newly formulated framework extends existing knowledge on deception in online consumer reviews and develops a new theoretical view based on a multidisciplinary approach (De Bakker et al., 2019; Whetten, Felin, & King, 2009) involving marketing, communication, sociology, and the management of information systems theoretical elements. The mixed-methods approach, including both qualitative and quantitative methods analyzing digital data, allows us to better explain this phenomenon, as well as better answer questions related to "what, "how" and "why" (Crane, Henriques, & Husted, 2018; Gehman et al., 2018; Gioia, Corley, & Hamilton, 2012; Whetten, 1989), and also explores the best methods that can be used by future research in answering these questions.

The findings of this study confirm the interchangeable characteristics of the various text analysis methods in drawing insights about review characteristics but emphasize the much more beneficial complementary aspects of distinct approaches. An integrative multi-method model that approaches the data at both the individual and aggregate level provides more complex insights regarding the quantity and quality of review information, cues about its relevance and contextual information, perceptual aspects, and cognitive material.

Researchers can use these findings in the analysis of deceptive communication as complementary to more complex machine learning and neural network models, as well as for improving quantitative and qualitative text exploration methods. A combination of linguistic, semantic, statistical, and neural network analysis can provide more insight while maintaining the desired level of transparency and visibility of algorithms and the theoretical base of the analysis.

7.2.Managerial implications

On the practitioner side, this paper provides managers with an overview of the main approaches available in analyzing review quality, deception levels, and sentiment. Also, the proposed framework emphasizes the effectiveness of a multi-method approach in evaluating review content, analyzing cues, and formulating communication strategies based on review sentiment, relevance, and style. The methods discussed can also be combined with more complex analyses such as neural networks and used as effective ways to manage difficulties in accessing calibration and training data or assessing algorithm elements. The linguistic analysis reflects the positive effects of review incentivization on consumer usage of keywords and content that reflect a positive sentiment.

For consumers and other stakeholders, this study is critical in offering easy solutions to analyze individual and aggregated review content and cues and review characteristics that can be used to identify deceptive content manually. While different review analysis tools offered by brands like Amazon and Yelp are proprietary and do not reveal their key variables and algorithms, this analysis provides easy ways to evaluate online comments through manual and automated methods. Nevertheless, the findings are also helpful for policymakers in evaluating deceptive marketing communications and formulating policies and regulations regarding online reviews.

7.3.Limitations and future research

This paper focused on evaluating the most widely used automated text analysis methods with the limitation of using some of the most accessible approaches and software options for marketing researchers and practitioners based on classical machine learning, including LIWC, Leximancer, and Textalyzer, based on assessing linguistic and semantic characteristics of reviews. As the purpose of the paper was to evaluate the interchangeable and complementary characteristics of these methods and the usefulness of the metrics and insights offered, we did not use more complex machine learning and neural network analyses. Another reason for not employing these methods was the current debate and difficulties related to using artificial vs. real-world reviews and the need for training datasets (Hajek & Sahut, 2022). In this context, we recommend exploring the theoretical model we formulated in a neural network context.

The literature on consumer reviews could also benefit from additional cross-cultural studies incorporating cultural characteristics and communication styles in the analysis of reviews and using multi-language approaches. Moreover, multidisciplinary projects that integrate insights from experts in linguistics, marketing, communications, sociology, psychology, and data science can provide further advancement in our field. Also, since there is a potential that some frequent verified reviewers included in this manuscript adopted a deceptive self-presentation behavior, such as writing informative and long reviews, to increase their chance to join the Amazon Vine program, we also recommend testing our theoretical framework in another review context and in an experimental setting to better evaluate the “how” and “why” research questions on deceptive review communication.

References

Amazon (2022). About Amazon Vine. Available at https://www.amazon.com/vine/about

Banerjee, S., & Chua, A. Y. K. (2017). Authentic versus fictitious online reviews: A textual analysis across luxury, budget, and mid-range hotels. *Journal of Information Science, 43*(1), 122–134. https://doi.org/10.1177/0165551515625027

Bond, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin*, *134*(4), 477–492. https://doi.org/10.1037/0033-2909.134.4.477

Cardoso, E. F., Silva, R. M., & Almeida, T. A. (2018). Towards automatic filtering of fake reviews. *Neurocomputing*, *309*, 106–116. https://doi.org/10.1016/j.neucom.2018.04.074

Chakraborty, U., & Bhat, S. (2018). The effects of credible online reviews on brand equity dimensions and its consequence on consumer behavior. *Journal of Promotion Management*, *24*(1), 57–82. https://doi.org/10.1080/10496491.2017.1346541

Chatterjee, S., Goyal, D., Prakash, A., & Sharma, J. (2021). Exploring healthcare/health-product ecommerce satisfaction: A text mining and machine learning application. *Journal of Business Research*, *131*(October 2020), 815–825. https://doi.org/10.1016/j.jbusres.2020.10.043

Chen, C.-D., & Huang, L.-T. (2011). Online deception investigation: Content analysis and cross-cultural comparison. *International Journal of Business and Information*, *6*(1), 91–111.

Choi, S., Mattila, A. S., Van Hoof, H. B., & Quadri-Felitti, D. (2017). The role of power and incentives in inducing fake reviews in the tourism industry. *Journal of Travel Research*, *56*(8), 975–987. https://doi.org/10.1177/0047287516677168

Corley, K. G., & Gioia, D. A. (2011). Building theory about theory building: what constitutes a theoretical contribution? *Academy of Management Review*, *36*(1), 12–32.

Crane, A., Henriques, I., & Husted, B. W. (2018). Quants and poets: Advancing methods and methodologies in business and society research. *Business and Society*, *57*(1), 3–25. https://doi.org/10.1177/0007650317718129

DeAndrea, D. C., TomTong, S., Liang, Y. J., Levine, T. R., & Walther, J. B. (2012). When do people misrepresent themselves to others? The effects of social desirability, ground truth, and accountability on deceptive self-presentations. *Journal of Communication*, *62*(3), 400–417.

De Bakker, F., Crane, A., Henriques, I., & Husted, B. W. (2019). Publishing Interdisciplinary Research in Business & Society. *Business and Society*, *58*(3), 443–452. https://doi.org/10.1177/0007650319826188

De Langhe, B., Fernbach, P. M., & Lichtenstein, D. R. (2016). Navigating by the stars: Investigating the actual and perceived validity of online user ratings. *Journal of Consumer Research*, *42*(6), 817–833. https://doi.org/10.1093/jcr/ucv047

Dellarocas, C. (2006). Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management Science*, *52*(10), 1577–1593. https://doi.org/10.1287/mnsc.1060.0567

DePaulo, B. M., Kirkendol, S. E., Kashy, D. A., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, *70*(5), 979–995. https://doi.org/10.1037/0022-3514.70.5.979

DePaulo, B. M., Malone, B. E., Lindsay, J. J., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*(1), 74–118. https://doi.org/10.1037/0033-2909.129.1.74

DePaulo, B. (1992). A self-presentational view of social phenomenon. *Psychological Bulletin*, *111*(2), 203–243.

Feldman, R. S., Forrest, J. A., & Happ, B. R. (2002). Self-presentation and verbal deception: Do self-presenters lie more? *Basic and Applied Social Psychology*, *24*(2), 163–170. https://doi.org/10.1207/S15324834BASP2402_8

Gehman, J., Glaser, V. L., Eisenhardt, K. M., Gioia, D., Langley, A., & Corley, K. G. (2018). Finding theory–method fit: A comparison of three qualitative approaches to theory building. *Journal of Management Inquiry*, *27*(3), 284–300. https://doi.org/10.1177/1056492617706029

Gentina, E., Chen, R., & Yang, Z. (2021). Development of theory of mind on online social networks: Evidence from Facebook, Twitter, Instagram, and Snapchat. *Journal of Business Research*, *124*(March 2020), 652–666. https://doi.org/10.1016/j.jbusres.2020.03.001

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia Methodology. *Organizational Research Methods*, *16*(1), 15–31. https://doi.org/10.1177/1094428112452151

Gioia, D., Corley, K., Eisenhardt, K., Feldman, M., Langley, A., Lê, J., Golden-Biddle, K., Locke, K., Mees-Buss, J., Piekkari, R., Ravasi, D., Rerup, C., Schmid, T., Silverman, D., & Welch, C. (2022). A curated debate: on using “templates” in qualitative research.

Journal of Management Inquiry, *31*(3), 231–252. https://doi.org/10.1177/10564926221098955

Gössling, S., Hall, C. M., & Andersson, A. C. (2018). The manager's dilemma: a conceptualization of online review manipulation strategies. *Current Issues in Tourism*, *21*(5), 484–503. https://doi.org/10.1080/13683500.2015.1127337

Hajek, P., & Sahut, J. M. (2022). Mining behavioural and sentiment-dependent linguistic patterns from restaurant reviews for fake review detection. *Technological Forecasting and Social Change*, *177*(January), 121532. https://doi.org/10.1016/j.techfore.2022.121532

Heydari, A., Tavakoli, M. A., Salim, N., & Heydari, Z. (2015). Detection of review spam: A survey. *Expert Systems with Applications*, *42*(7), 3634–3642. https://doi.org/10.1016/j.eswa.2014.12.029

Hu, N., Bose, I., Gao, Y., & Liu, L. (2011). Manipulation in digital word-of-mouth: A reality check for book reviews. *Decision Support Systems*, *50*(3), 627–635. https://doi.org/10.1016/j.dss.2010.08.013

Hu, N., Bose, I., Koh, N. S., & Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems*, *52*(3), 674–684. https://doi.org/10.1016/j.dss.2011.11.002

Huang, A. H., Chen, K., Yen, D. C., & Tran, T. P. (2015). A study of factors that contribute to online review helpfulness. *Computers in Human Behavior*, *48*, 17–27. https://doi.org/10.1016/j.chb.2015.01.010

Ivanova, O., & Scholz, M. (2017). How can online marketplaces reduce rating manipulation? A new approach on dynamic aggregation of online ratings. *Decision Support Systems*, *104*, 64–78. https://doi.org/10.1016/j.dss.2017.10.003

Kashy, D. A., & DePaulo, B. M. (1996). Who lies? *Journal of Personality and Social Psychology*, *70*(5), 1037–1051. https://doi.org/10.1037/0022-3514.70.5.1037

Kavlakoglu, E. (2020). https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks

Kim, J., Naylor, G., Sivadas, E., & Sugumaran, V. (2016). The unrealized value of incentivized eWOM recommendations. *Marketing Letters*, *27*(3), 411–421. https://doi.org/10.1007/s11002-015-9360-3

Kim, J., Kim, J. E., & Marshall, R. (2016). Are two arguments always better than one? Persuasion knowledge moderating the effect of integrated marketing communications. *European Journal of Marketing*, *50*(7–8), 1399–1425. https://doi.org/10.1108/EJM-06-2014-0344

Krishen, A. S., Agarwal, S., & Kachroo, P. (2016). Is having accurate knowledge necessary for implementing safe practices? A consumer folk theories-of-mind perspective on the impact of price. *European Journal of Marketing*, *50*(5–6), 1073–1093. https://doi.org/10.1108/EJM-01-2015-0027

Leary, M. R. (1995). *Self-presentation: Impression management and interpersonal behavior*. Westview Press.

Lee, H. H., & Ma, Y. J. (2012). Consumer perceptions of online consumer product and service reviews: Focusing on information processing confidence and susceptibility to peer influence. *Journal of Research in Interactive Marketing*, *6*(2), 110–132. https://doi.org/10.1108/17505931211265426

Lee, J., Park, D. H., & Han, I. (2011). The different effects of online consumer reviews on consumers' purchase intentions depending on trust in online shopping malls: An advertising perspective. *Internet Research*, *21*(2), 187–206. https://doi.org/10.1108/10662241111123766

Lee, S. Y., Qiu, L., & Whinston, A. (2018). Sentiment manipulation in online platforms: an analysis of movie tweets. *Production and Operations Management*, *27*(3), 393–416. https://doi.org/10.1111/poms.12805

Levine, T. R., & McCornack, S. A. (2014). Theorizing about deception. *Journal of Language and Social Psychology*, *33*(4), 431–440. https://doi.org/10.1177/0261927X14536397

Libai, B., Biyalogorsky, E., & Gerstner, E. (2003). Setting referral fees in affiliate marketing. *Journal of Service Research*, *5*(4), 303–315. https://doi.org/10.1177/1094670503005004003

Lin, C. A., & Xu, X. (2017). Effectiveness of online consumer reviews: The influence of valence, reviewer ethnicity, social distance and source trustworthiness. *Internet Research*, *27*(2), 362–380. https://doi.org/10.1108/IntR-01-2016-0017

Liu, A. H., Chugh, R., & Noel Gould, A. (2016). Working smart to win back lost customers the role of coping choices and justice mechanisms. *European Journal of Marketing*, *50*(3–4), 397–420. https://doi.org/10.1108/EJM-10-2014-0642

Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science, 62*(12), 3412–3427. https://doi.org/10.1287/mnsc.2015.2304

Ma, Y. J., & Lee, H. H. (2014). Consumer responses toward online review manipulation. *Journal of Research in Interactive Marketing*, *8*(3), 224–244. https://doi.org/10.1108/JRIM-04-2013-0022

Malbon, J. (2013). Taking fake online consumer reviews seriously. *Journal of Consumer Policy*, *36*(2), 139–157. https://doi.org/10.1007/s10603-012-9216-7

McCornack, S. A., Morrison, K., Paik, J. E., Wisner, A. M., & Zhu, X. (2014). Information Manipulation Theory 2: A propositional theory of deceptive discourse production. *Journal of Language and Social Psychology*, *33*(4), 348–377. https://doi.org/10.1177/0261927X14534656

McCornack, S. A., Levine, T. R., Morrison, K., & Lapinski, M. (1996). Speaking of information manipulation: A critical rejoinder. *Communication Monographs*, *63*(March), 83-95.

McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs, 59*, 1-16.

McCornack, S. A., & Levine, T. R (1990). When lies are uncovered: Emotional and relational outcomes of discovered deception. *Communication Monographs, 57*, 119-138.

Moon, S., & Kamakura, W. A. (2017). A picture is worth a thousand words: Translating product reviews into a product positioning map. *International Journal of Research in Marketing*, *34*(1), 265–285. https://doi.org/10.1016/j.ijresmar.2016.05.007

Moon, S., Kim, M. Y., & Iacobucci, D. (2021). Content analysis of fake consumer reviews by survey-based text categorization. *International Journal of Research in Marketing*, *38*(2), 343–364. https://doi.org/10.1016/j.ijresmar.2020.08.001

Mukherjee, A., Liu, B., & Glance, N. (2012). Spotting fake reviewer groups in consumer reviews. *WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web*, 191–200. https://doi.org/10.1145/2187836.2187863

Munzel, A. (2016). Assisting consumers in detecting fake reviews: The role of identity information disclosure and consensus. *Journal of Retailing and Consumer Services*, *32*, 96–108. https://doi.org/10.1016/j.jretconser.2016.06.002

Munzel, A. (2015). Malicious practice of fake reviews: Experimental insight into the potential of contextual indicators in assisting consumers to detect deceptive opinion spam. *Recherche et Applications en Marketing (English Edition)*, *30*(4), 24–50. https://doi.org/10.1177/2051570715604155

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*. *29*(5), 665–675. https://doi.org/10.1177/0146167203251529

Ott, M., Cardie, C., & Hancock, J. (2012). Estimating the prevalence of deception in online review communities. *WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web*, 201–210. https://doi.org/10.1145/2187836.2187864

Peng, L., Cui, G., Zhuang, M., & Li, C. (2016). Consumer perceptions of online review deceptions: an empirical study in China. *Journal of Consumer Marketing*, *33*(4), 269–280. https://doi.org/10.1108/JCM-01-2015-1281

Petrescu, M., O'Leary, K., Goldring, D., & Ben Mrad, S. (2018). Incentivized reviews: Promising the moon for a few stars. *Journal of Retailing and Consumer Services, 41*, 288-295. 10.1016/j.jretconser.2017.04.005

Petrescu, M., Kitchen, P., Dobre, C., Ben Mrad, S., Milovan-Ciuta, A., Goldring, D., & Fiedler, A. (2022). Innocent until proven guilty: suspicion of deception in online reviews. *European Journal of Marketing*, *56*(4), 1184-1209. https://doi.org/10.1108/EJM-10-2019-0776

Plotkina, D., & Munzel, A. (2016). Delight the experts, but never dissatisfy your customers! A multi-category study on the effects of online review source on intention to buy a new product. *Journal of Retailing and Consumer Services*, *29*, 1–11. https://doi.org/10.1016/j.jretconser.2015.11.002

Plotkina, D., Munzel, A., & Pallud, J. (2020). Illusions of truth—Experimental insights into human and algorithmic detections of fake online reviews. *Journal of Business Research*, *109*, 511–523. https://doi.org/10.1016/j.jbusres.2018.12.009

Ren, J., Yeoh, W., Shan Ee, M., & Popovič, A. (2018). Online consumer reviews and sales: Examining the chicken-egg relationships. *Journal of the Association for Information Science and Technology*, *69*(3), 449–460. https://doi.org/10.1002/asi.23967

Rese, A., Schreiber, S., & Baier, D. (2014). Technology acceptance modeling of augmented reality at the point of sale: Can surveys be replaced by an analysis of online reviews? *Journal of Retailing and Consumer Services*, *21*(5), 869–876. https://doi.org/10.1016/j.jretconser.2014.02.011

Riquelme, I. P., & Román, S. (2014). The influence of consumers' cognitive and psychographic traits on perceived deception: A comparison between online and offline retailing contexts. *Journal of Business Ethics*, *119*(3), 405–422. https://doi.org/10.1007/s10551-013-1628-z

Román, S. (2010). Relational consequences of perceived deception in online shopping: The moderating roles of type of product, consumer's attitude toward the internet and consumer's demographics. *Journal of Business Ethics*, *95*(3), 373–391. https://doi.org/10.1007/s10551-010-0365-9

Román, S., & Cuestas, P. J. (2008). The perceptions of consumers regarding online retailers' ethics and their relationship with consumers' general internet expertise and word of mouth: A preliminary analysis. *Journal of Business Ethics*, *83*(4), 641–656. https://doi.org/10.1007/s10551-007-9645-4

Sahut, J. M., Iandoli, L., & Teulon, F. (2021). The age of digital entrepreneurship. *Small Business Economics*, *56*(3), 1159–1169. https://doi.org/10.1007/s11187-019-00260-8

Schau , H. J., & Gilly, M C. (2003). We are what we post self presentation in personal web space. *Journal of Consumer Research*, *30*(3), 385–404.

Schoenmueller, V., Netzer, O., & Stahl, F. (2020). The polarity of online reviews: prevalence, drivers and implications. *Journal of Marketing Research*, *57*(5), 853–877. https://doi.org/10.1177/0022243720941832

Smith, A. E., & Humphreys, M. S. (2006). Evaluation of unsupervised semantic text mapping. *Behavior Research Methods*, *38*(2), 262–279.

Smith, R. K., Yazdani, E., Wang, P., Soleymani, S., & Ton, L. A. N. (2022). The cost of looking natural: Why the no-makeup movement may fail to discourage cosmetic use. *Journal of the Academy of Marketing Science*, *50*(2), 324–337. https://doi.org/10.1007/s11747-021-00801-2

Steward, M. D., Burns, A. C., Morgan, F. N., & Roehm, M. L. (2020). Credible effects: the impact of disclosure of material connections within online product reviews. *Journal of*

Public Policy and Marketing, *39*(3), 353–368. https://doi.org/10.1177/0743915619864543

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Textalyzer (2022). Textalyzer: Keyword Density + Word Count Tool. https://seoscout.com/tools/text-analyzer

Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Personality and Social Psychology Bulletin*, 34(8), 1023–1036.

Toy, D., Wright, L., & Olson, J. (2001). A conceptual framework for analyzing deception and debriefing effects in marketing research. *Psychology and Marketing*, *18*(7), 691–719. https://doi.org/10.1002/mar.1026

van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84, 523–538. https://doi.org/10.1007/s11192-009-0146-3

Waltman, Ludo & van Eck, Nees Jan & Noyons, Ed C.M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, *4*(4), 629-635.

Weathers, D., Swain, S. D., & Grover, V. (2015). Can online product reviews be more helpful? Examining characteristics of information content by product type. *Decision Support Systems*, *79*, 12–23. https://doi.org/10.1016/j.dss.2015.07.009

Whetten, D. A., Felin, T., & King, B. G. (2009). The practice of theory borrowing in organizational studies: Current issues and future directions. *Journal of Management*, *35*(3), 537–563. https://doi.org/10.1177/0149206308330556

Whetten, D. A. (1989). *What Constit utes a Theore tical Contribution ? 14*(4), 490–495.

Wu, Y., Ngai, E. W. T., Wu, P., & Wu, C. (2020). Fake online reviews: Literature review, synthesis, and directions for future research. *Decision Support Systems*, *132*(March), 113280. https://doi.org/10.1016/j.dss.2020.113280

Xiang, L., Zheng, X., Zhang, K. Z. K., & Lee, M. K. O. (2018). Understanding consumers' continuance intention to contribute online reviews. *Industrial Management and Data Systems*, *118*(1), 22–40. https://doi.org/10.1108/IMDS-09-2016-0395

Yang, C. S., Chen, C. H., & Chang, P. C. (2015). Harnessing consumer reviews for marketing intelligence: a domain-adapted sentiment classification approach. *Information Systems and E-Business Management*, *13*(3), 403–419. https://doi.org/10.1007/s10257-014-0266-z

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on Sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, *74*(2), 133–148. https://doi.org/10.1509/jmkg.74.2.133

Zhuang, M., Cui, G., & Peng, L. (2018). Manufactured opinions: The effect of manipulating online product reviews. *Journal of Business Research*, *87*(February 2017), 24–35. https://doi.org/10.1016/j.jbusres.2018.02.016

Zourrig, H., Zhang, M., El Hedhli, K., & Becheur, I. (2021). The influence of culture on consumer perceptions of deceptiveness. *Journal of Consumer Marketing*, *38*(5), 469–483. https://doi.org/10.1108/JCM-09-2020-4150

Fig. 1. Study 1: Keyword co-occurrence map.

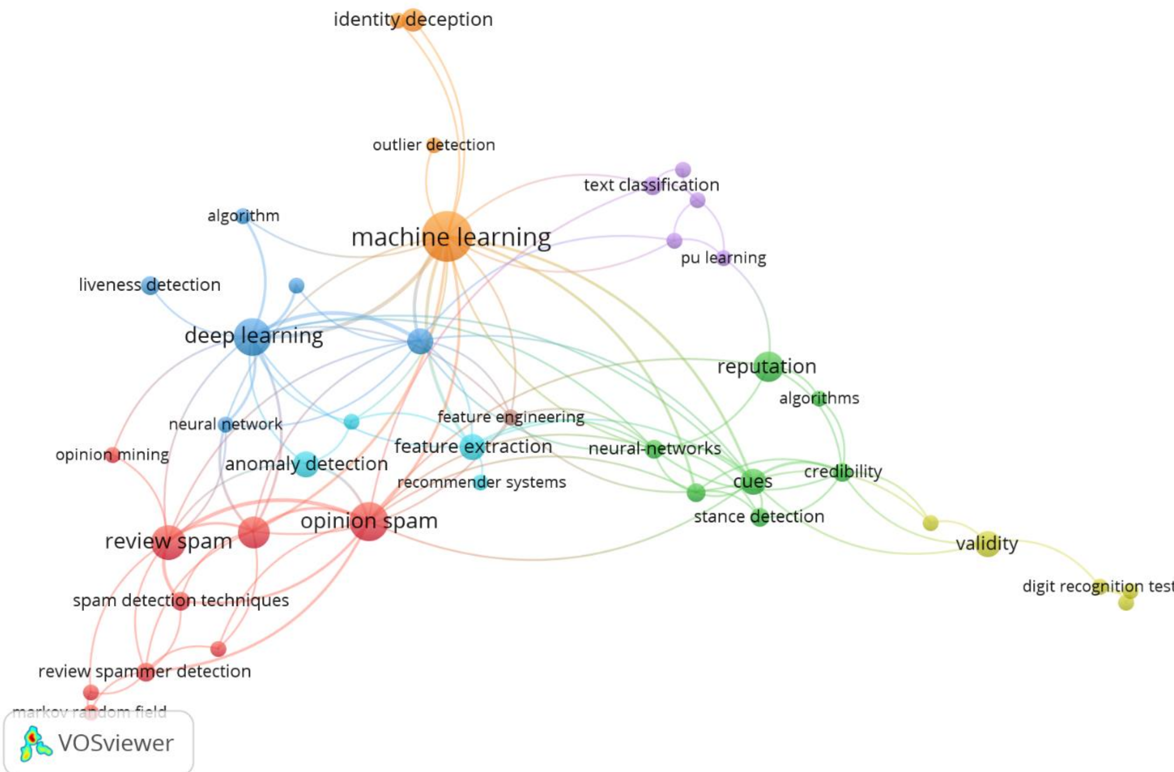

Fig. 2. Artificial intelligence technologies

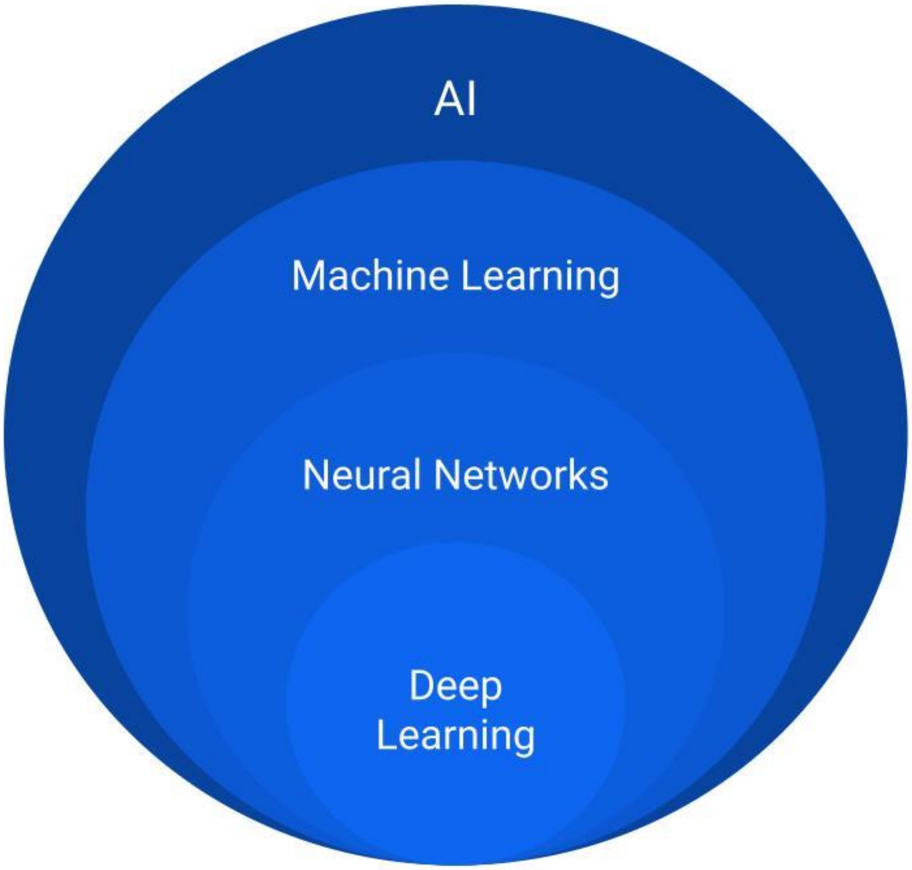

Fig. 3. Study 2: Conceptual map.

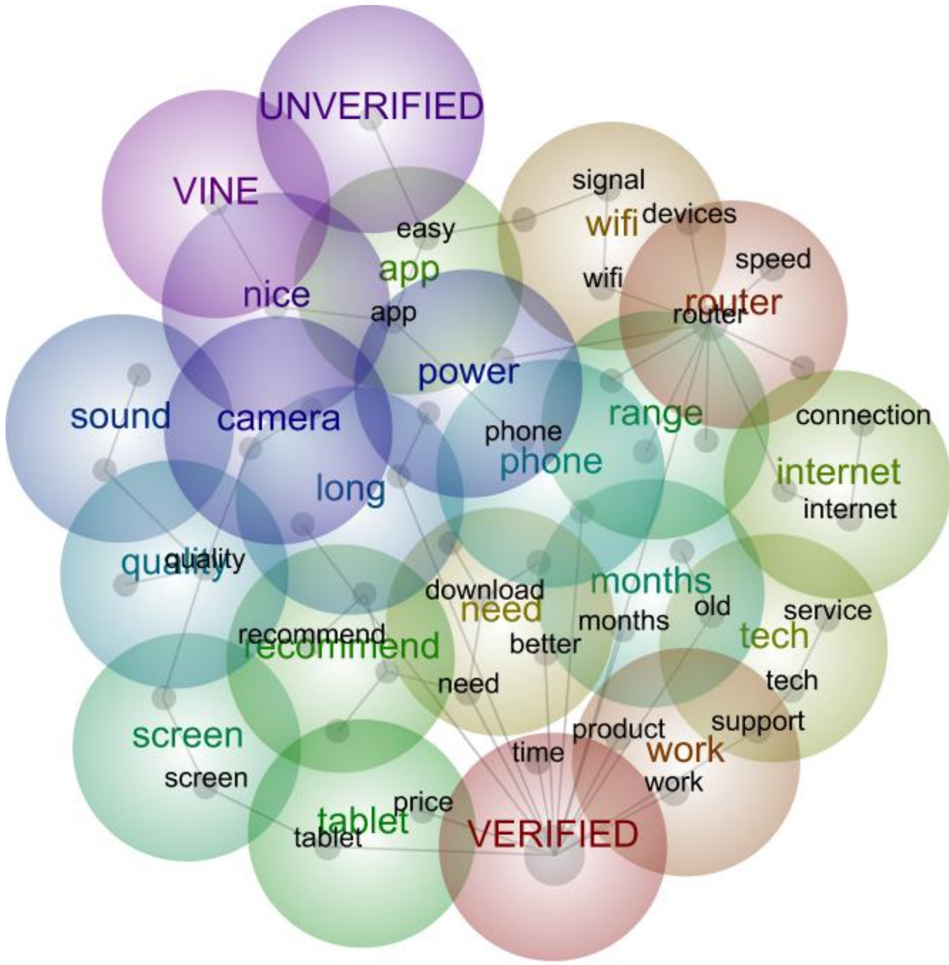

Fig. 4. Review interpretation model.

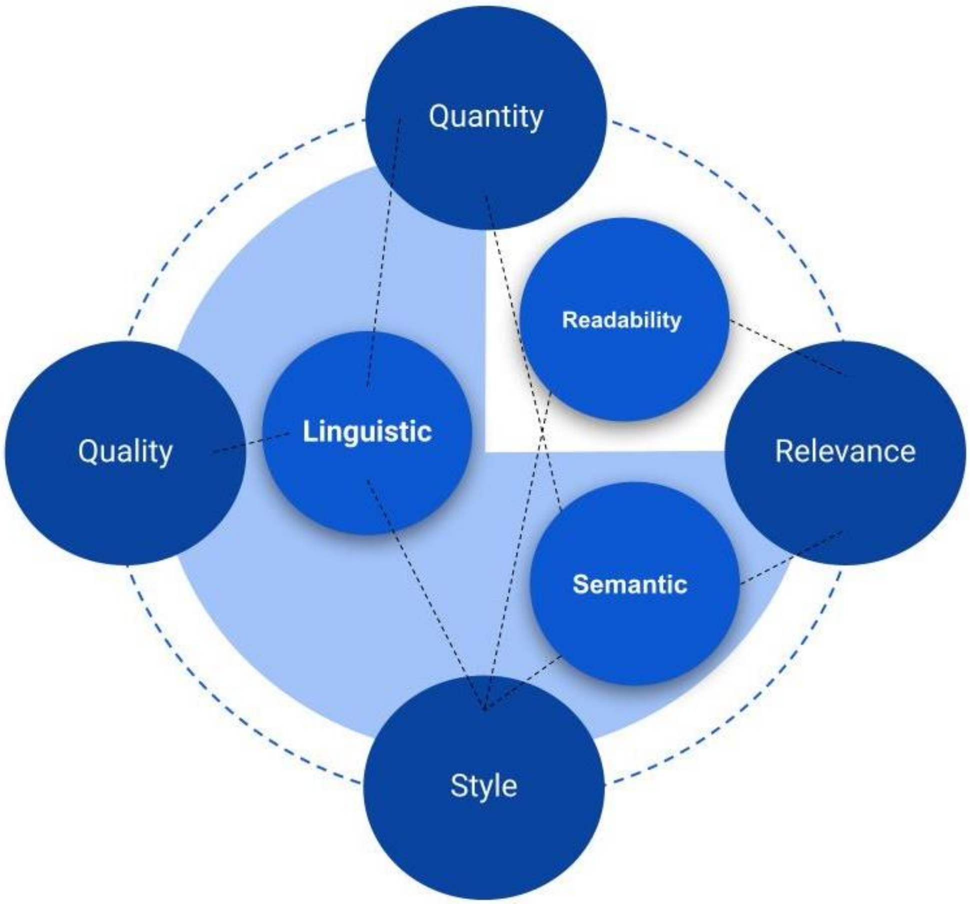

Table 1 Study 1: Keyword co-occurrence clusters.

keyword	cluster	Links	Link strength	Avg. citations
group spamming	Spam detection	3	3	10
markov random field		2	2	2
opinion mining		2	2	16
opinion spam		12	20	25
review spam		10	16	34
review spammer detection		7	8	44
sentiment analysis		10	15	17
spam detection techniques		4	7	52
spam review detection		3	3	8
algorithms	NLP	2	2	59
credibility		8	9	26
cues		12	15	21
natural language processing		6	7	12
reputation		5	5	16
stance detection		4	4	4
convolutional neural network	Neural network	2	3	3
deep learning		15	25	7
neural network		5	7	18
opinion spam detection		11	16	3
digit recognition test	Computation	3	3	27
forced-choice method		2	2	13
knowledge graph		3	3	22
validity		4	4	28
em algorithm	Semi-supervised	4	4	24
pu learning		3	3	17
semi-supervised learning		4	4	18
spammer group detection		2	2	10
text classification		4	4	13
anomaly detection	Supervised	4	5	13
feature extraction		8	10	5
recommender systems		1	1	5
identity deception	ML	2	2	6
machine learning		17	28	16
outlier detection		1	1	17
sybil detection		2	2	13

Table 2 Study 2: MANOVA analysis results.

DV	Sum of Squares	df	Mean²	F	Sig.	Partial Eta²	Noncent. Param.	Obs. Power[a]
Word count	1487698.86	2	743849.43	159.860	0.001	0.017	319.72	1.000
Analytic	10411.19	2	5205.60	5.954	0.003	0.001	11.91	0.881
Clout	12184.18	2	6092.09	9.569	0.001	0.001	19.14	0.981
Authentic	29974.36	2	14987.18	11.813	0.001	0.001	23.63	0.995
Tone	96480.41	2	48240.20	38.056	0.001	0.004	76.11	1.000
power	56.60	2	28.30	2.364	0.094	0.001	4.73	0.480
time	1307.78	2	653.89	18.772	0.001	0.002	37.54	1.000

The F tests the effect of type. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.
a. Computed using alpha = .05

Table 3 Study 2: Multiple comparison results.

DV			Mean Diff. (I-J)	Std. Error	Sig.[b]	95% Conf. Int. Dif.[b] Lower Bound	Upper Bound
Word count	Unverified	Verified	32.699*	3.223	**0.001**	26.382	39.015
		Vine	-20.354*	4.736	**0.001**	-29.636	-11.071
	Verified	Unverified	-32.699*	3.223	**0.001**	-39.015	-26.382
		Vine	-53.052*	3.547	**0.001**	-60.004	-46.100
	Vine	Unverified	20.354*	4.736	**0.001**	11.071	29.636
		Verified	53.052*	3.547	**0.001**	46.100	60.004
Analytic	Unverified	Verified	2.241	1.397	0.109	-0.497	4.979
		Vine	6.879*	2.053	**0.001**	2.855	10.902
	Verified	Unverified	-2.241	1.397	0.109	-4.979	0.497
		Vine	4.637*	1.537	**0.003**	1.624	7.651
	Vine	Unverified	-6.879*	2.053	**0.001**	-10.902	-2.855
		Verified	-4.637*	1.537	**0.003**	-7.651	-1.624
Clout	Unverified	Verified	4.250*	1.192	**0.001**	1.914	6.587
		Vine	0.816	1.752	0.641	-2.618	4.249
	Verified	Unverified	-4.250*	1.192	**0.001**	-6.587	-1.914
		Vine	-3.435*	1.312	**0.009**	-6.006	-0.863
	Vine	Unverified	-0.816	1.752	0.641	-4.249	2.618
		Verified	3.435*	1.312	**0.009**	0.863	6.006
Authentic	Unverified	Verified	-4.223*	1.683	**0.012**	-7.521	-0.924
		Vine	3.594	2.473	0.146	-1.253	8.441
	Verified	Unverified	4.223*	1.683	**0.012**	0.924	7.521
		Vine	7.817*	1.852	**0.001**	4.187	11.447
	Vine	Unverified	-3.594	2.473	0.146	-8.441	1.253
		Verified	-7.817*	1.852	**0.001**	-11.447	-4.187
Tone	Unverified	Verified	-6.139*	1.682	**0.001**	-9.436	-2.842
		Vine	-20.644*	2.472	**0.001**	-25.489	-15.800
	Verified	Unverified	6.139*	1.682	**0.001**	2.842	9.436
		Vine	-14.506*	1.851	**0.001**	-18.134	-10.877
	Vine	Unverified	20.644*	2.472	**0.001**	15.800	25.489
		Verified	14.506*	1.851	**0.001**	10.877	18.134
power	Unverified	Verified	-0.052	0.163	0.750	-0.372	0.268
		Vine	0.336	0.240	0.162	-0.135	0.807
	Verified	Unverified	0.052	0.163	0.750	-0.268	0.372
		Vine	.388*	0.180	**0.031**	0.036	0.741
	Vine	Unverified	-0.336	0.240	0.162	-0.807	0.135
		Verified	-.388*	0.180	**0.031**	-0.741	-0.036
time	Unverified	Verified	-0.213	0.279	0.446	-0.759	0.334
		Vine	1.658*	0.410	**0.001**	0.855	2.461
	Verified	Unverified	0.213	0.279	0.446	-0.334	0.759
		Vine	1.871*	0.307	**0.001**	1.269	2.472
	Vine	Unverified	-1.658*	0.410	**0.001**	-2.461	-0.855
		Verified	-1.871*	0.307	**0.001**	-2.472	-1.269

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference

Table 4 Study 2: Concept prominence.

	concept	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	Verified																	
2	Unverified	0.00																
3	Vine	0.00	0.00															
4	work	**1.01**	**1.05**	0.77														
5	support	**1.05**	0.52	0.09	1.32													
6	tech	**1.05**	0.46	0.10	1.31	19.10												
7	easy	0.99	0.90	**1.38**	1.09	0.82	0.87											
8	time	**1.00**	**1.26**	0.78	1.19	1.52	1.46	0.95										
9	product	**1.00**	**1.05**	**1.00**	1.05	1.73	1.68	0.93	1.19									
10	recommend	0.99	**1.07**	**1.10**	0.70	0.92	0.92	1.10	0.86	2.54								
11	better	0.98	**1.20**	**1.33**	1.13	0.64	0.66	0.81	1.14	1.22	0.81							
12	quality	0.92	**2.38**	**1.86**	0.65	0.27	0.26	1.05	0.90	1.13	1.82	1.92						
13	price	0.99	**1.07**	**1.15**	0.80	0.41	0.42	0.92	0.87	1.50	1.38	1.82	2.57					
14	service	**1.03**	0.92	0.26	1.13	6.04	6.06	0.59	1.46	2.14	1.02	1.22	0.43	0.58				
15	problem	**1.01**	**1.01**	0.72	1.20	3.00	2.82	0.56	1.63	1.11	1.01	0.71	0.53	0.68	2.35			
16	love	0.94	0.72	**2.70**	0.51	0.16	0.19	1.51	0.67	0.72	0.89	1.35	1.92	0.76	0.12	0.73		
17	money	**1.01**	**1.48**	0.38	1.05	1.39	1.38	0.60	2.80	1.78	1.11	1.72	1.02	0.83	1.69	1.32	0.45	
18	nice	0.77	**1.17**	**6.72**	0.88	0.38	0.42	1.68	0.95	1.02	1.14	1.43	1.62	1.43	0.52	0.58	1.94	0.86

Table 5 Study 2: Review metrics.

Metrics	Unverified	Verified	Vine
Average Words/Sentence	14.7	12.3	14.4
Average Syllables/Word	1.4	1.4	1.4
Lexical Density	41%	44%	41%
Lexical Diversity	11%	11%	10%
Reading Ease	74.20%	76.90%	77%
Grade Level	6.5	5.6	6.1
Gunning Fog	9.1	7.8	8.7
Coleman Liau Index	8.8	9	8.6
Smog Index	6.8	6.1	6.5
Automated Reading Index	5.6	4.6	5.3
Word count	88.117	55.419	108.471
Analytic	61.447	59.206	54.568
Clout	39.593	35.343	38.777
Authentic	46.008	50.231	42.414
Tone	60.863	67.001	81.507
Power	1.903	1.955	1.567
Time	5.419	5.632	3.761

Appendix

Authors	Title	Journal	Year
Adewole, Kayode Sakariyah; Anuar, Nor Badrul; Kamsin, Amirrudin; Varathan, Kasturi Dewi; Razak, Syed Abdul	Malicious accounts: Dark of the social networks	Journal of Network and Computer Applications	2017
Agarwal, Rohit; Jalal, Anand Singh	Presentation attack detection system for fake Iris: a review	Multimedia Tools and Applications	2021
Akhter, Muhammad Pervez; Zheng, Jiangbin; Afzal, Farkhanda; Lin, Hui; Riaz, Saleem; Mehmood, Atif	Supervised ensemble learning methods towards automatically filtering Urdu fake news within social media	PEER Journal of Computer Science	2021
Akram, Abubakker Usman; Khan, Hikmat Ullah; Iqbal, Saqib; Iqbal, Tassawar; Munir, Ehsan Ullah; Shafi, Muhammad	Finding Rotten Eggs: A Review Spam Detection Model using Diverse Feature Sets	KSII Transactions on Internet and Information Systems	2018
Alarifi, Abdulrahman; Alsaleh, Mansour; Al-Salman, AbdulMalik	Twitter turing test: Identifying social machines	Information Sciences	2016
ALDayel, Abeer; Magdy, Walid	Stance detection on social media: State of the art and trends	Information Processing & Management	2021
Alharbi, Ahmed; Dong, Hai; Yi, Xun; Tari, Zahir; Khalil, Ibrahim	Social Media Identity Deception Detection: A Survey	Acm Computing Surveys	2021
Alotaibi, Aziz; Mahmood, Ausif	Deep face liveness detection based on nonlinear diffusion using convolution neural network	Signal Image and Video Processing	2017
Alsubari, Saleh Nagi; Deshmukh, Sachin N.; Al-Adhaileh, Mosleh Hmoud; Alsaade, Fawaz Waselalla; Aldhyani, Theyazn H. H.	Development of Integrated Neural Network Model for Identification of Fake Reviews in E-Commerce Using Multidomain Datasets	Applied Bionics and Biomechanics	2021
Ananthakrishnan, Uttara M.; Li, Beibei; Smith, Michael D.	A Tangled Web: Should Online Review Portals Display Fraudulent Reviews?	Information Systems Research	2020
Asghar, Muhammad Zubair; Ullah, Asmat; Ahmad, Shakeel; Khan, Aurangzeb	Opinion spam detection framework using hybrid classification scheme	Soft Computing	2020
Asr, Fatemeh Torabi; Taboada, Maite	Big Data and quality data for fake news and misinformation detection	Big Data & Society	2019
Barbado, Rodrigo; Araque, Oscar; Iglesias, Carlos A.	A framework for fake review detection in online consumer electronics retailers	Information Processing & Management	2019
Basit, Abdul; Zafar, Maham; Liu, Xuan; Javed, Abdul Rehman; Jalil, Zunera; Kifayat, Kashif	A comprehensive survey of AI-enabled phishing attacks detection techniques	Telecommunication Systems	2021
Ben Sassi, Imen; Ben Yahia, Sadok	Malicious accounts detection from online social networks: a systematic review of literature	International Journal of General Systems	2021
Bindu, P. V.; Mishra, Rahul; Thilagam, P. Santhi	Discovering spammer communities in twitter	Journal of Intelligent Information Systems	2018

Budhi, Gregorius Satia; Chiong, Raymond; Wang, Zuli	Resampling imbalanced data to detect fake reviews using machine learning classifiers and textual-based features	Multimedia Tools and Applications	2021
Budhi, Gregorius Satia; Chiong, Raymond; Wang, Zuli; Dhakal, Sandeep	Using a hybrid content-based and behaviour-based featuring approach in a parallel environment to detect fake reviews	Electronic Commerce Research and Applications	2021
Cagnina, Leticia C.; Rosso, Paolo	Detecting Deceptive Opinions: Intra and Cross-Domain Classification Using an Efficient Representation	International Journal of Uncertainty Fuzziness and Knowledge-Based Systems	2017
Cao, Jiuxin; Xia, Rongqing; Guo, Yifang; Ma, Zhuo	Collusion-aware detection of review spammers in location based social networks	World Wide Web-Internet and Web Information Systems	2019
Cardoso, Emerson F.; Silva, Renato M.; Almeida, Tiago A.	Towards automatic filtering of fake reviews	Neurocomputing	2018
Chen, Lirong; Li, Wenli; Chen, Hao; Geng, Shidao	Detection of Fake Reviews: Analysis of Sellers' Manipulation Behavior	Sustainability	2019
Cheng, Li-Chen; Hu, Hsiao-Wei; Wu, Chia-Chi	Spammer Group Detection Using Machine Learning Technology for Observation of New Spammer Behavioral Features	Journal of Global Information Management	2021
Cresci, Stefano; Di Pietro, Roberto; Petrocchi, Marinella; Spognardi, Angelo; Tesconi, Maurizio	Fame for sale: Efficient detection of fake Twitter followers	Decision Support Systems	2015
Dewang, Rupesh Kumar; Singh, Anil Kumar	State-of-art approaches for review spammer detection: a survey	Journal of Intelligent Information Systems	2018
Dlamini, M. T.; Venter, H. S.; Eloff, J. H. P.; Eloff, M. M.	Digital deception in cybersecurity: an information behaviour lens	Information Research-An International Electronic Journal	2020
Dong, Lu-yu; Ji, Shu-juan; Zhang, Chun-jin; Zhang, Qi; Chiu, DicksonK. W.; Qiu, Li-Qing; Li, Da	An unsupervised topic-sentiment joint probabilistic model for detecting deceptive reviews	Expert Systems with Applications	2018
Dong, Manqing; Yao, Lina; Wang, Xianzhi; Benatallah, Boualem; Huang, Chaoran; Ning, Xiaodong	Opinion fraud detection via neural autoencoder decision forest	Pattern Recognition Letters	2020
D'Ulizia, Arianna; Caschera, Maria Chiara; Ferri, Fernando; Grifoni, Patrizia	Fake news detection: a survey of evaluation datasets	Peerj Computer Science	2021
Echizen, Isao; Babaguchi, Noboru; Yamagishi, Junichi; Nitta, Naoko; Nakashima, Yuta; Nakamura, Kazuaki; Kono, Kazuhiro; Fang, Fuming; Myojin, Seiko; Kuang, Zhenzhong; Nguyen, Huy H.; Tieu, Ngoc-Dung T.	Generation and Detection of Media Clones	IEICE Transactions on Information and Systems	2021
Fang, Youli; Wang, Hong; Zhao, Lili; Yu, Fengping; Wang, Caiyu	Dynamic knowledge graph based fake-review detection	Applied Intelligence	2020

Fernandez, Facundo M.; Green, Michael D.; Newton, Paul N.	Prevalence and detection of counterfeit pharmaceuticals: A mini review	Industrial & Engineering Chemistry Research	2008
Fornaciari, Tommaso; Cagnina, Leticia; Rosso, Paolo; Poesio, Massimo	Fake opinion detection: how similar are crowdsourced datasets to real data?	Language Resources and Evaluation	2020
Galan-Garcia, Patxi; de la Puerta, Jose Gaviria; Gomez, Carlos Laorden; Santos, Igor; Bringas, Pablo Garcia	Supervised machine learning for the detection of troll profiles in twitter social network: application to a real case of cyberbullying	Logic Journal of the IGPL	2016
Guo, Bin; Ding, Yasan; Sun, Yueheng; Ma, Shuai; Li, Ke; Yu, Zhiwen	The mass, fake news, and cognition security	Frontiers of Computer Science	2021
Guo, Bin; Ding, Yasan; Yao, Lina; Liang, Yunji; Yu, Zhiwen	The Future of False Information Detection on Social Media: New Perspectives and Trends	ACM Computing Surveys	2020
Hajek, Petr; Barushka, Aliaksandr; Munk, Michal	Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining	Neural Computing & Applications	2020
He, Daojing; Pan, Menghan; Hong, Kai; Cheng, Yao; Chan, Sammy; Liu, Xiaowen; Guizani, Nadra	Fake Review Detection Based on PU Learning and Behavior Density	IEEE Network	2020
Heydari, Atefeh; Tavakoli, Mohammad Ali; Salim, Naomie; Heydari, Zahra	Detection of review spam: A survey	Expert Systems with Applications	2015
Heydari, Atefeh; Tavakoli, Mohammadali; Salim, Naomie	Detection of fake opinions using time series	Expert Systems with Applications	2016
Hooi, Bryan; Shin, Kijung; Song, Hyun Ah; Beutel, Alex; Shah, Neil; Faloutsos, Christos	Graph-Based Fraud Detection in the Face of Camouflage	Acm Transactions on Knowledge Discovery from Data	2017
Hunt, Kate Mathews	Gaming the system: Fake online reviews v. consumer law	Computer Law & Security Review	2015
Hussain, Naveed; Mirza, Hamid Turab; Rasool, Ghulam; Hussain, Ibrar; Kaleem, Mohammad	Spam Review Detection Techniques: A Systematic Literature Review	Applied Sciences-Basel	2019
Hyman, Michael R.; Sierra, Jeremy J.	Adjusting self-reported attitudinal data for mischievous respondents	International Journal of Market Research	2012
Impedovo, Donato; Dentamaro, Vincenzo; Abbattista, Giacomo; Gattulli, Vincenzo; Pirlo, Giuseppe	A comparative study of shallow learning and deep transfer learning techniques for accurate fingerprints vitality detection	Pattern Recognition Letters	2021
Jacob, Minu Susan; Rajendran, P. Selvi	Fuzzy artificial bee colony-based CNN-LSTM and semantic feature for fake product review classification	Concurrency and Computation-Practice & Experience	2021
Jain, Praphula Kumar; Pamula, Rajendra; Ansari, Sarfraj	A Supervised Machine Learning Approach for the Credibility Assessment of User-Generated Content	Wireless Personal Communications	2021

Ji, Shu-Juan; Zhang, Qi; Li, Jinpeng; Chiu, Dickson K. W.; Xu, Shaohua; Yi, Lei; Gong, Maoguo	A burst-based unsupervised method for detecting review spammer groups	Information Sciences	2020
Jiang, Cuixia; Zhu, Jun; Xu, Qifa	Dissecting click farming on the Taobao platform in China via PU learning and weighted logistic regression	Electronic Commerce Research	2021
Kaliyar, Rohit Kumar; Goswami, Anurag; Narang, Pratik; Sinha, Soumendu	FNDNet - A deep convolutional neural network for fake news detection	Cognitive Systems Research	2020
Kauffmann, Erick; Peral, Jesus; Gil, David; Ferrandez, Antonio; Sellers, Ricardo; Mora, Higinio	A framework for big data analytics in commercial social networks: A case study on sentiment analysis and fake review detection for marketing decision-making	Industrial Marketing Management	2020
Khalid, Waqar; Ullah, Zahid; Ahmed, Naveed; Cao, Yue; Khalid, Muhammad; Arshad, Muhammad; Ahmad, Farhan; Cruickshank, Haitham	A taxonomy on misbehaving nodes in delay tolerant networks	Computers & Security	2018
Khurshid, Faisal; Zhu, Yan; Xu, Zhuang; Ahmad, Mushtaq; Ahmad, Muqeet	Enactment of Ensemble Learning for Review Spam Detection on Selected Features	International Journal of Computational Intelligence Systems	2019
Kumar, Naveen; Venugopal, Deepak; Qiu, Liangfei; Kumar, Subodha	Detecting Anomalous Online Reviewers: An Unsupervised Approach Using Mixture Models	Journal of Management Information Systems	2019
Kumar, Naveen; Venugopal, Deepak; Qiu, Liangfei; Kumar, Subodha	Detecting Review Manipulation on Online Platforms with Hierarchical Supervised Learning	Journal of Management Information Systems	2018
Lappas, Theodoros; Sabnis, Gaurav; Valkanas, Georgios	The Impact of Fake Reviews on Online Visibility: A Vulnerability Assessment of the Hotel Industry	Information Systems Research	2016
Li, Neng; Du, Suguo; Zheng, Haizhong; Xue, Minhui; Zhu, Haojin	Fake Reviews Tell No Tales? Dissecting Click Farming in Content-Generated Social Networks	China Communications	2018
Li, Wentao; Gao, Min; Li, Hua; Zeng, Jun; Xiong, Qingyu; Hirokawa, Sachio	Shilling Attack Detection in Recommender Systems via Selecting Patterns Analysis	IEICE Transactions on Information and Systems	2016
Li, Yuejun; Wang, Fangxin; Zhang, Shuwu; Niu, Xiaofei	Detection of Fake Reviews Using Group Model	Mobile Networks & Applications	2021
Ligthart, Alexander; Catal, Cagatay; Tekinerdogan, Bedir	Analyzing the effectiveness of semi-supervised learning approaches for opinion spam classification	Applied Soft Computing	2021
Liong, Sze-Teng; Gan, Y. S.; Zheng, Danna; Li, Shu-Meng; Xu, Hao-Xuan; Zhang, Han-Zhe; Lyu, Ran-Ke; Liu, Kun-Hong	Evaluation of the Spatio-Temporal Features and GAN for Micro-Expression Recognition System	Journal of Signal Processing Systems for Signal Image and Video Technology	2020
Liu, Meiling; Shang, Yue; Yue, Qi; Zhou, Jiyun	Detecting Fake Reviews Using Multidimensional	IEEE Access	2021

	Representations With Fine-Grained Aspects Plan		
Liu, Rui; Liu, Runze; Pugliese, Andrea; Subrahmanian, V. S.	STARS: Defending against Sockpuppet-Based Targeted Attacks on Reviewing Systems	ACM Transactions on Intelligent Systems and Technology	2020
Liu, Yuanchao; Pang, Bo	A unified framework for detecting author spamicity by modeling review deviation	Expert Systems with Applications	2018
Majadi, Nazia; Trevathan, Jarrod; Gray, Heather; Estivill-Castro, Vladimir; Bergmann, Neil	Real-time detection of shill bidding in online auctions: A literature review	Computer Science Review	2017
Makkar, Aaisha; Kumar, Neeraj; Zomaya, Albert Y.; Dhiman, Shalini	SPAMI: A cognitive spam protector for advertisement malicious images	Information Sciences	2020
Manaskasemsak, Bundit; Tantisuwankul, Jirateep; Rungsawang, Arnon	Fake review and reviewer detection through behavioral graph partitioning integrating deep neural network	Neural Computing & Applications	2021
Maseri, Aimi Nadrah; Norman, Azah Anir; Eke, Christopher Ifeanyi; Ahmad, Atif; Molok, Nurul Nuha Abdul	Socio-Technical Mitigation Effort to Combat Cyber Propaganda: A Systematic Literature Mapping	IEEE Access	2020
Masood, Faiza; Ammad, Ghana; Almogren, Ahmad; Abbas, Assad; Khattak, Hasan Ali; Din, Ikram Ud; Guizani, Mohsen; Zuair, Mansour	Spammer Detection and Fake User Identification on Social Networks	IEEE Access	2019
Melchers, Klaus G.; Roulin, Nicolas; Buehl, Anne-Kathrin	A review of applicant faking in selection interviews	International Journal of Selection and Assessment	2020
Miller, David J.; Xiang, Zhen; Kesidis, George	Adversarial Learning Targeting Deep Neural Network Classification: A Comprehensive Review of Defenses Against Attacks	Proceedings of the IEEE	2020
Mohawesh, Rami; Tran, Son; Ollington, Robert; Xu, Shuxiang	Analysis of concept drift in fake reviews detection	Expert Systems with Applications	2021
Mohawesh, Rami; Xu, Shuxiang; Tran, Son N.; Ollington, Robert; Springer, Matthew; Jararweh, Yaser; Maqsood, Sumbal	Fake Reviews Detection: A Survey	IEEE Access	2021
Monaro, Merylin; Cannonito, Emanuela; Gamberini, Luciano; Sartori, Giuseppe	Spotting faked 5 stars ratings in E-Commerce using mouse dynamics	Computers in Human Behavior	2020
Moon, Sangkil; Kim, Moon-Yong; Iacobucci, Dawn	Content analysis of fake consumer reviews by survey-based text categorization*	International Journal of Research in Marketing	2021
Neisari, Ashraf; Rueda, Luis; Saad, Sherif	Spam review detection using self-organizing maps and convolutional neural networks	Computers & Security	2021
Oh, Yu Won; Park, Chong Hyun	Machine Cleaning of Online Opinion Spam: Developing a Machine-Learning Algorithm for Detecting Deceptive Comments	American Behavioral Scientist	2021

Pandey, Avinash Chandra; Tikkiwal, Vinay Anand	Stance detection using improved whale optimization algorithm	Complex & Intelligent Systems	2021
Paul, Himangshu; Nikolaev, Alexander	Fake review detection on online E-commerce platforms: a systematic literature review	Data Mining and Knowledge Discovery	2021
Plotkina, Daria; Munzel, Andreas; Pallud, Jessie	Illusions of truth-Experimental insights into human and algorithmic detections of fake online reviews	Journal of Business Research	2020
Rahman, Mahmudur; Carbunar, Bogdan; Ballesteros, Jaime; Chau, Duen Horng (Polo)	To Catch a Fake: Curbing Deceptive Yelp Ratings and Venues	Statistical Analysis and Data Mining	2015
Ramalingam, Devakunchari; Chinnaiah, Valliyammai	Fake profile detection techniques in large-scale online social networks: A comprehensive review	Computers & Electrical Engineering	2018
Rani, Neetu; Das, Prasenjit; Bhardwaj, Amit Kumar	Rumor, misinformation among web: A contemporary review of rumor detection techniques during different web waves	Concurrency and Computation-Practice & Experience	2021
Reyes-Menendez, Ana; Ramon Saura, Jose; Filipe, Ferrao	The importance of behavioral data to identify online fake reviews for tourism businesses: a systematic review	PEERJ Computer Science	2019
Rout, Jitendra Kumar; Dalmia, Anmol; Choo, Kim-Kwang Raymond; Bakshi, Sambit; Jena, Sanjay Kumar	Revisiting Semi-Supervised Learning for Online Deceptive Review Detection	IEEE Access	2017
Rout, Jitendra Kumar; Singh, Smriti; Jena, Sanjay Kumar; Bakshi, Sambit	Deceptive review detection using labeled and unlabeled	Multimedia Tools and Applications	2017
Ruan, Na; Deng, Ruoyu; Su, Chunhua	GADM: Manual fake review detection for O2O commercial platforms	Computers & Security	2020
Sadiq, Saima; Umer, Muhammad; Ullah, Saleem; Mirjalili, Seyedali; Rupapara, Vaibhav; Nappi, Michele	Discrepancy detection between actual user reviews and numeric ratings of Google App store using deep learning	Expert Systems with Applications	2021
Saquete, Estela; Tomas, David; Moreda, Paloma; Martinez-Barco, Patricio; Palomar, Manuel	Fighting post-truth using natural language processing: A review and open challenges	Expert Systems with Applications	2020
Sato, Koichi; Wang, Junbo; Cheng, Zixue	Credibility Evaluation of Twitter-Based Event Detection by a Mixing Analysis of Heterogeneous Data	IEEE Access	2019
Savage, David; Zhang, Xiuzhen; Yu, Xinghuo; Chou, Pauline; Wang, Qingmai	Detection of opinion spam based on anomalous rating deviation	Expert Systems with Applications	2015
Savyan, P. V.; Bhanu, S. Mary Saira	UbCadet: detection of compromised accounts in twitter based on user behavioural profiling	Multimedia Tools and Applications	2020
Shan, Guohou; Zhou, Lina; Zhang, Dongsong	From conflicts and confusion to doubts: Examining review	Decision Support Systems	2021

	inconsistency for fake review detection		
Shehnepoor, Saeedreza; Salehi, Mostafa; Farahbakhsh, Reza; Crespi, Noel	NetSpam: A Network-Based Spam Detection Framework for Reviews in Online Social Media	IEEE Transactions on Information Forensics and Security	2017
Sun, Chengai; Du, Qiaolin; Tian, Gang	Exploiting Product Related Review Features for Fake Review Detection	Mathematical Problems in Engineering	2016
Taneja, Harsh; Kaur, Supreet	An ensemble classification model for fake feedback detection using proposed labeled CloudArmor dataset	Computers & Electrical Engineering	2021
Tolosana, Ruben; Vera-Rodriguez, Ruben; Fierrez, Julian; Morales, Aythami; Ortega-Garcia, Javier	Deepfakes and beyond: A Survey of face manipulation and fake detection	Information Fusion	2020
Verdoliva, Luisa	Media Forensics and DeepFakes: An Overview	IEEE Journal of Selected Topics in Signal Processing	2020
Viviani, Marco; Pasi, Gabriella	Credibility in social media: opinions, news, and health information-a survey	Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery	2017
Wang, Guan; Xie, Sihong; Liu, Bing; Yu, Philip S.	Identify Online Store Review Spammers via Social Review Graph	Acm Transactions on Intelligent Systems and Technology	2012
Wang, Jianyu; Wu, Chunming	Camouflage is NOT easy: Uncovering adversarial fraudsters in large online app review platform	Measurement & Control	2020
Wang, Jingdong; Kan, Haitao; Meng, Fanqi; Mu, Qizi; Shi, Genhua; Xiao, Xixi	Fake Review Detection Based on Multiple Feature Fusion and Rolling Collaborative Training	IEEE Access	2020
Wang, Jitao; Sun, Guozi; Gu, Yu; Liu, Kun	ConGradetect: Blockchain-based detection of code and identity privacy vulnerabilities in crowdsourcing	Journal of Systems Architecture	2021
Wang, Ye; Liu, Bixin; Wu, Hongjia; Zhao, Shan; Cai, Zhiping; Li, Donghui; Fong, Cheang Chak	An Opinion Spam Detection Method Based on Multi-Filters Convolutional Neural Network	Cmc-Computers Materials & Continua	2020
Wang, Zhuo; Chen, Qian	Monitoring online reviews for reputation fraud campaigns	Knowledge-Based Systems	2020
Wang, Zhuo; Gu, Songmin; Xu, Xiaowei	GSLDA: LDA-based group spamming detection in product reviews	Applied Intelligence	2018
Wang, Zhuo; Hou, Tingting; Song, Dawei; Li, Zhun; Kong, Tianqi	Detecting Review Spammer Groups via Bipartite Graph Projection	Computer Journal	2016
Wang, Zhuo; Hu, Runlong; Chen, Qian; Gao, Pei; Xu, Xiaowei	ColluEagle: collusive review spammer detection using Markov random fields	Data Mining and Knowledge Discovery	2020

Xu, Yuanbo; Yang, Yongjian; Han, Jiayu; Wang, En; Ming, Jingci; Xiong, Hui	Slanderous user detection with modified recurrent neural networks in recommender system	Information Sciences	2019
Yang, Zhi; Xue, Jilong; Yang, Xiaoyong; Wang, Xiao; Dai, Yafei	Vote Trust: Leveraging Friend Invitation Graph to Defend against Social Network Sybils	IEEE Transactions on Dependable and Secure Computing	2016
Yao, Jianrong; Zheng, Yuan; Jiang, Hui	An Ensemble Model for Fake Online Review Detection Based on Data Resampling, Feature Pruning, and Parameter Optimization	IEEE Access	2021
Yuan, Shuhan; Wu, Xintao; Xiang, Yang	Task-specific word identification from short texts using a convolutional neural network	Intelligent Data Analysis	2018
Zhang, Dongsong; Zhou, Lina; Kehoe, Juan Luo; Kilic, Isil Yakut	What Online Reviewer Behaviors Really Matter? Effects of Verbal and Nonverbal Behaviors on Detection of Fake Online Reviews	Journal of Management Information Systems	2016
Zhang, Fuzhi; Hao, Xiaoyan; Chao, Jinbo; Yuan, Shuai	Label propagation-based approach for detecting review spammer groups on e-commerce websites	Knowledge-Based Systems	2020
Zhang, Lu; He, Gaofeng; Cao, Jie; Zhu, Haiting; Xu, Bingfeng	Spotting review spammer groups: A cosine pattern and network based method	Concurrency and Computation-Practice & Experience	2018
Zhang, Lu; Wu, Zhiang; Cao, Jie	Detecting Spammer Groups From Product Reviews: A Partially Supervised Learning Model	IEEE Access	2018
Zhong, Minjuan; Li, Zhenjin; Liu, Shengzong; Yang, Bo; Tan, Rui; Qu, Xilong	Fast Detection of Deceptive Reviews by Combining the Time Series and Machine Learning	Complexity	2021
Zhou, Tuoyu; Han, Huawen; Liu, Pu; Xiong, Jian; Tian, Fake; Li, Xiangkai	Microbial Fuels Cell-Based Biosensor for Toxicity Detection: A Review	Sensors	2017
Zhou, Xinyi; Zafarani, Reza	A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities	ACM Computing Surveys	2020
Zhu, Chengzhang; Zhao, Wentao; Li, Qian; Li, Pan; Da, Qiaobo	Network Embedding-Based Anomalous Density Searching for Multi-Group Collaborative Fraudsters Detection in Social Media	CMC-Computers Materials & Cont.	2019