Data Technologies and Applications

Understanding customer behavior by mapping complaints to personality based on social media textual data

Andry Alamsyah, Fadiah Nadhila and Nabila Kalvina Izumi
Telkom University, Bandung, Indonesia

Received 7 February 2024
Revised 14 June 2024
Accepted 17 August 2024

Abstract

Purpose – Technology serves as a key catalyst in shaping society and the economy, significantly altering customer dynamics. Through a deep understanding of these evolving behaviors, a service can be tailored to address each customer's unique needs and personality. We introduce a strategy to integrate customer complaints with their personality traits, enabling responses that resonate with the customer's unique personality.

Design/methodology/approach – We propose a strategy to incorporate customer complaints with their personality traits, enabling responses that reflect the customer's unique personality. Our approach is twofold: firstly, we employ the customer complaints ontology (CCOntology) framework enforced with multi-class classification based on a machine learning algorithm, to classify complaints. Secondly, we leverage the personality measurement platform (PMP), powered by the big five personality model to predict customer's personalities. We develop the framework for the Indonesian language by extracting tweets containing customer complaints directed towards Indonesia's three biggest e-commerce services.

Findings – By mapping customer complaints and their personality type, we can identify specific personality traits associated with customer dissatisfaction. Thus, personalizing how we offer the solution based on specific characteristics.

Originality/value – The research enriches the state-of-the-art personalizing service research based on captured customer behavior. Thus, our research fills the research gap in considering customer personalities. We provide comprehensive insights by aligning customer feedback with corresponding personality traits extracted from social media data. The result is a highly customized response mechanism attuned to individual customer preferences and requirements.

Keywords Big five personality, Personality measurement platform, CCOntology, Machine learning, Personalized customer service, Indonesia language, Tweet

Paper type Research paper

1. Introduction

The advanced industry era exemplifies the reciprocal influence of industrial experiences and technological development (Castelo-Branco *et al.*, 2022). Industrial development positively affects financial status, enhancing economic growth (Panagariya, 2019). Nations can harness the power of the digital revolution to enhance their financial and economic performance (Gielens and Steenkamp, 2019). Information technology has notably emerged as the predominant technology for improving business efficiency, paving the way for the development of e-commerce (Onyancha, 2015; Baršauskas *et al.*, 2008). Additionally, the proliferation of mobile technologies, unlimited internet access, and the emergence of cloud computing have significantly transformed the digital economy era (Dudhat and Agarwal, 2023). Social media platforms have become essential for marketing and customer engagement (Zhao *et al.*, 2020). With the ubiquity of mobile devices, customers have unlimited access to online platforms (Hanaysha, 2022), facilitating the spread of information among internet users (Dudhat and Agarwal, 2023; Greenberg, 2010). Companies increasingly rely on social media to address customer complaints (Strauss and Hill, 2001; Prada *et al.*, 2022).

Data Technologies and Applications

2514-9288
DOI 10.1108/DTA-02-2024-0162

By utilizing large-scale data and implementing a data analytics approach, businesses could provide personalized experiences, products, and services (Mudambi and Schuff, 2010). Personalization has transformed the digital experience, offering tailored solutions that cater to individuals' unique needs and preferences. However, the effective implementation of personalization faces challenges in accurately understanding individual needs, preferences, and context due to the complexities of human behavior (Huang *et al.*, 2019). This effort is crucial in marketing, product recommendations, and user experience (Jiang and Liu, 2019; Liu *et al.*, 2019). Adapting to evolving preferences requires significant time and effort (McKinsey). Personalization was hard to achieve in the past due to limitations in the data source, knowledge extraction methodologies, and computation capabilities, but the big data phenomenon offers promising solutions (Bender, 2020). By providing valuable insights into individual behavior and preferences through statistical analysis and leveraging machine learning algorithms, they enable more accurate recommendations and increased personalization (Lopes *et al.*, 2016). As a result, research focuses on personalizing the customer experience in a business service context. It plays a crucial role in transforming various sectors, including government services, communication, politics, finance, and healthcare.

Customer behavior is crucial for businesses in the digital era. It refers to customers' actions and decisions when buying goods or services, influenced by psychological and emotional factors (Liu and Park, 2015). To increase sales, companies must conduct market research and develop effective marketing strategies (Zhao *et al.*, 2020). Positive customer experiences lead to positive reviews and word-of-mouth recommendations, while negative experiences result in negative feedback (Sun and Zhao, 2022; Li and Yang, 2012). Understanding consumer behavior is critical for marketing success. Customers have different personality traits when lodging complaints, which impact their interaction with the company. Companies must acknowledge their customers' personality characteristics and tailor the interaction methods accordingly (Mccoll-Kennedy and Smith, 2006; Tronvoll, 2011; Verhagen *et al.*, 2013).

The big five personality theory (BFT) is a central psychological framework characterized by five primary dimensions (Langford *et al.*, 2017). It is extensively utilized alongside other models such as the Myers-Briggs Type Indicator (MBTI), the Hexaco personality inventory, Eysenck's Personality Theory, Cattell's 16 Personality Factors (16PF), and Motive-based models. We choose BFT over these alternatives due to its robust empirical support and widespread acceptance within the psychological community, making it a reliable tool for predicting a range of life outcomes and behaviors. Its universalistic approach is supported by cross-cultural research, indicating that the five dimensions can effectively describe personalities across different cultures, thus enhancing its utility in diverse contexts. Moreover, the simplicity and comprehensiveness of BFT make it highly accessible for both academic research and practical applications in fields such as our study about customer opinion. Extensive research has demonstrated BFT's reliability and validity, making it a robust framework (Wright and Jackson, 2023; Dhelim *et al.*, 2021; Beck and Jackson, 2020). One approach is to measure one's personality based on textual expression on social media (Dudija *et al.*, 2022). We use a personality measurement platform (PMP) to classify people's personalities based on their posts on social media (Alamsyah *et al.*, 2021). Nevertheless, it is essential to acknowledge that BFT has limitations; it may occasionally reduce the complexity of human personality to overly broad categories and could reflect cultural biases, factors that must be considered when applying this model.

Most customers prefer filing complaints via social media as it provides an alternative avenue to voice their concerns (Agostino and Sidorova, 2017). Social media is accessible, allows for fast responses, and enables sharing experiences and receiving feedback (Strauss and Hill, 2001; Sashi *et al.*, 2019). Companies may utilize a CCOntology framework to analyze and categorize customer complaints (Jarrar, 2009). This framework breaks down complaints into its elements and helps identify patterns and trends in customer complaints (Jarrar *et al.*, 2003). This approach facilitates businesses in addressing recurring issues. While existing research predominantly centers on the service recovery paradox – suggesting that effectively addressed complaints can enhance customer loyalty and satisfaction (Kunathikornkit *et al.*, 2023; Umashankar *et al.*, 2017) – and the effects of personalized service responses on customer satisfaction and loyalty (Kunathikornkit *et al.*, 2023; Umashankar *et al.*, 2017), there remains a notable gap in understanding the role of specific personality traits in shaping customer complaint behaviors and their responses to service recovery. This oversight in the literature highlights a lack of detailed exploration into how individual personality differences influence the expression and resolution of complaints. Consequently, our research question seeks to bridge this gap by investigating the correlation between customer personality traits and their complaint behaviors. It aims to devise tailored, practical solutions for service recovery or develop further into applicable solutions.

Given the motivation and extensive phenomenon, our research question is how to construct a framework to understand customer complaints by learning about their personality. The specific research objective is to categorize customer complaints and analyze their personality traits, thereby providing companies with a comprehensive understanding of their customer's needs and preferences. Establishing a system to capture and analyze customer feedback allows businesses to anticipate the network effects of complaints (Fan and Niu, 2016). By leveraging CCOntology and the big five personality model, firms can better grasp customer issues, identify feedback patterns, and improve products and support (Eggert *et al.*, 2020; Shams *et al.*, 2020; Alamsyah and Adityawarman, 2017). Analyzing feedback data enables companies to identify themes, develop targeted solutions, and enhance satisfaction. As a case study, we collect customer complaint data on *Twitter* towards the top 3 e-commerce services in Indonesia using the informal Indonesian language. We compare three machine learning algorithms: Support Vector Machine, Naïve Bayes, and Maximum Entropy performance, to classify customer complaints based on CCOntology frameworks, followed by utilization of PMP to predict customer personality. The result shows that we can analyze the personality proportion of each complaint dimension. By understanding this, business organizations can determine the best approach to respond to complaints.

For the case study, we choose Indonesia, which has a population of over 276 million, is one of the most populous countries in the world, and 73,7% of the population is connected to the internet. Indonesia's digital transformation has led to an explosion of e-commerce, reshaping the way daily transactions take place. Online marketplaces and digital platforms have become go-to destinations for shopping, challenging the traditional offline retail landscape. Once bustling centers of commerce, shopping malls are now facing the threat of reduced foot traffic as more people turn to the convenience and variety offered by e-commerce platforms. We picked the 3 biggest and most popular Indonesia e-commerce sites: Tokopedia, Shoope, and Bukalapak. Indonesians are very open and transparent about their feeling toward services; often, they write their problem publicly on social media, although the solution process with the e-commerce company is still ongoing. This behavior indicates the perfect opportunity to collect massively available information on social media.

2. Theoretical background

2.1 Customer complaint ontology (CCOntology)

Customer complaints are an integral part of any business operation, and companies need to understand the nature and underlying causes of these complaints (Joseph and George, 2007). One way to achieve this is by using ontologies, formal representations of concepts, and their interrelationships within a given domain. In the context of customer complaints, ontologies can help businesses organize and categorize complaints, identify their underlying causes, and generate solutions that address the root of the problem.

Several studies have explored the use of ontologies to manage customer complaints. Jarrar constructed a framework to understand online customer complaints using an Ontology Web Language named CCOntology (Jarrar, 2009). This framework approach may be used to improve the effectiveness and transparency of customer complaint management in e-business transactions. The proposed ontology captures relevant concepts and relationships in customer complaints and provides a standardized vocabulary to describe and exchange complaint-related information between different parties. This ontology is evaluated through a case study of a real-world e-business platform. It improved the efficiency and accuracy of the complaint resolution process while increasing customer satisfaction and trust. Using ontology-based reasoning is an intelligent approach to managing customer complaints (Lee *et al.*, 2015), which involves developing a CCOntology, a domain ontology, and a reasoning engine to analyze and classify complaints. Ontology-based methods can also assess the credibility of big social data by defining relevant concepts and relationships (Wongthongtham and Abu-Salih, 2018). The proposed system combines a sentiment analysis module that analyzes customer reviews to identify sentiments and an ontology-based recommendation module that matches customer interests with relevant products (Karthik and Ganapathy, 2021).

Based on the examples of ontology-based systems mentioned above, we conclude that ontology is a branch of philosophy that studies reality's existence and nature (Gustav and Špet, 2022). Ontologies are often used in artificial intelligence, knowledge management, customer complaints, the semantic web, and other fields to organize and structure data to support reasoning, decision-making, and communication between different systems and people (Ojino *et al.*, 2022). Ontology refers to a collection of problems, such as domains and sub-domains, in the form of entities specifically related (Rajapaksha *et al.*, 2017). The CCOntology framework becomes the basis to understand online customer complaints using an Ontology Web Language (OWL), thus can comprehensively integrate customer complaints in e-commerce, as depicted in Figure 1. However, not all the dimensions and sub-dimensions in the proposed framework have real-world applications in one particular instance, such as a business or a country. We identify only 7 characteristics based on sub-domains that consistently appear in Indonesian customer complaints; they are Conduct, Evidence, Privacy Issues, Pre-Purchase Phase, Purchase Phase, Post-Purchase Phase, and Selling Method.

2.2 Big five personality

A study examines the relationship between the big five personality traits and interpersonal functioning; the results exhibit that all Big Five traits had unique interpersonal profiles, with variations found across different facets and domains of interpersonal functioning (Du *et al.*, 2020). By grouping the words utilized in everyday conversations, this personality measurement methodology aims to provide a more nuanced understanding of an individual's unique disposition that distinguishes them from others (Lambiotte and Kosinski, 2014). Utilize the big five personality theory with a multitrait approach that observes one specific trait to depict a more dominant and accurate personality. In this study, the search for a customer's dominant personality trait is not solely based on their dominant nature but also on their principal characteristics (Hong *et al.*, 2008). The big five personality theory encompasses five major traits, each with a sub-trait: openness, conscientiousness, extraversion, agreeableness, and neuroticism (Geovanni *et al.*, 2021). Further elucidation of these traits and their respective sub-traits is presented in Table 1.

Source(s): Figure adapted from (Jarrar, 2009)

Figure 1. A taxonomy on customer complaint ontology (CCOntology) framework

Personality traits	Definition
Openness	Self-awareness and self-realization if something is not right, having an intellectual mind, being excited about finding out unpopular things, and always searching and looking forward to new experiences
Conscientiousness	The personality trait that has consistency in their efforts to orient themselves toward achieving grades and goals
Extraversion	Bound and embroiled to the outside world, searching for enthusiasm and positive emotions. Daily activities include going outside and socializing
Agreeableness	The possession of empathy and a propensity to participate in constructive conflict resolution, encompass differences in thoughts, feelings, or actions
Neuroticism	Psychological traits are caused by negative emotions, which result in unrealistic ideas and emotional overreacting

Source(s): Table summarized from Hong *et al.* (2008)

Table 1. Big five personality traits

2.3 Social media and personality

Web 2.0 pushes the evolution of social media, significantly impacting the ease of sharing knowledge and information (Douglas *et al.*, 2022). Many social media platforms accelerate user interaction and facilitate the emergence of user generated content (UGC), which drives content creation by anyone, including customers (Dennhardt, 2013). Among these platforms, Twitter stands out, becoming a primary medium for customers to voice their opinions on products and services (Ibrahim and Wang, 2019). Research indicates social media content mirrors its creator's personality (Du *et al.*, 2020; Amichai-Hamburger and Vinitzky, 2010). By analyzing vast amounts of textual data, insights into individuals' personalities can be gleaned from their language and writing styles on social media (Alamsyah *et al.*, 2020). For instance, extroverted individuals tend to be more active on social media and have larger online networks compared to their introverted counterparts (Amichai-Hamburger and Vinitzky, 2010).

Online reviews of products and services, primarily through UGC, have become ubiquitous (Haider *et al.*, 2018; Kusumawati *et al.*, 2019). Recognizing this, businesses increasingly harness customer feedback to gauge preferences and identify market opportunities (Sokolova and Lapalme, 2009; Lampropoulos *et al.*, 2022). Negative experiences can lead customers to express dissatisfaction, potentially influencing others' perceptions and diminishing their interest in a product or service (Mansur *et al.*, 2023). Highly open individuals often communicate their grievances directly and creatively (Marouf *et al.*, 2019). In contrast, those with high agreeableness typically convey complaints tactfully and positively, while those with low agreeableness might resort to more hostile expressions (Min and Han, 2014). People with pronounced neuroticism are inclined to post negative feedback (Ruben *et al.*, 2023). Interestingly, extroverted individuals are more prone to voice complaints related to information needs on social media (Golmohammadi *et al.*, 2021), while highly conscientious individuals offer constructive feedback and actively seek information (Zhang *et al.*, 2022).

2.4 Machine learning algorithm

Machine learning techniques have revolutionized the field of natural language processing (NLP), enabling machines to analyze and comprehend human language with increasing accuracy and efficiency (Sawicki *et al.*, 2023). The field of NLP has advanced significantly, with many recent approaches utilizing deep learning mechanisms for classification tasks. Recent studies have employed these techniques for various purposes, including personality detection in talent hiring (70), mental health detection (71), and extraction from video interviews (72). The recent NLP-based deep learning, such as transformers and larger language model (LLM), offer unparalleled accuracy but require higher advanced skill and computational cost, which is often hard to implement in practice. Conversely, research focused on classifying customer complaints predominantly employs non-deep learning NLP methodologies such as support vector machines (SVM) (Piccialli and Sciandrone, 2022), Naïve Bayes classifier (Li and Yang, 2023), and Maximum Entropy (Mazuelas *et al.*, 2020) for addressing issues in finance (76), telecommunications (77), and hospitality (78).

Overall, deep learning-based approaches require extensive datasets to identify accurate patterns. Although several pretrained models, such as IndoBERT for the Indonesian language (Hugging Face), are available, they are generally limited to major and formal languages and do not extend to local dialects or informal social media languages. Therefore, this research focuses on using a non-deep learning mechanism to classify complaints according to pre-determined classes. Some well-performed machine learning algorithms for classification tasks are used, including Support Vector Machine, which is based on finding a hyperplane in a high-dimensional space that separates different classes. Naïve Bayes classifier, which applies Bayes' theorem to calculate the probability of a given text belonging to a certain class, Maximum Entropy utilizes statistical modeling to make predictions based on the principle of maximum entropy. CCOntology is a multi-class classification problem, so we can employ these machine learning algorithms to classify customer complaints into different categories based on their linguistic features.

To find the best performance algorithm, we arrange a comparative analysis among those three algorithms based on the evaluation metrics, such as precision, recall, and F1 score. Precision in multi-class classification refers to the proportion of correctly predicted instances in a particular class out of all instances predicted as belonging to that class. Meanwhile, recall is the proportion of correctly predicted instances in a particular class out of all instances that actually belong to that class. Finally, the F1 score is a harmonic mean of precision and recall, which provides a balanced measure of the algorithms' performance.

2.5 Ontology-based knowledge

There are several approaches to identifying knowledge from large-scale data, particularly textual data. Among these, machine learning, as discussed in Subsection 2.4, stands out as the quickest and most accurate method for recognizing patterns, which is crucial for tasks such as complaint classification and personality identification. Numerous new algorithms have emerged as data analytics methodologies advance, enhancing pattern recognition for classifying textual data into distinct groups. However, older methodologies, such as ontology-based approaches, still hold significant value in terms of accuracy. Ontology-based methods involve domain experts mapping the scope of problems within classes based on their knowledge. A notable study proposes incorporating ontology and deep learning mechanisms to detect BFT-based personality (Biswas *et al.*, 2023), however, this study requires more complex multimodal data.

In general, machine learning excels in terms of speed and scalability. Nevertheless, ontology-based approaches offer unique, particularly intriguing advantages in this research context, especially the PMP (Alamsyah *et al.*, 2021). These advantages include:

(1) *Deeper contextual understanding:* Ontology-based methods are built by domain experts, in this case, psychologists, who provide a nuanced understanding of the context.

(2) *Expert validation*: The results are validated by experts, ensuring unparalleled accuracy and reliability.

(3) *Language specificity*: For Indonesian language data, the machine learning-based character measurement corpus may still need to be fully developed for specific personality identification cases. Ontology-based methods can better handle local dialects and informal language on social media.

(4) *Crowdsourced and expert-validated corpus*: The PMP invites the public to participate in enriching the corpus. Experts subsequently validate this crowdsourcing mechanism.

Given these benefits, ontology-based approaches present a compelling case for their continued use in research, especially when dealing with specific languages and nuanced contextual understanding. While machine learning offers unmatched speed and scalability, the depth and expert validation provided by ontology-based methods make them an invaluable tool for personality identification and similar tasks. This study uses both machine learning and ontology. This hybrid approach leverages the strengths of both techniques, enhancing classification accuracy and efficiency. Machine learning excels in handling large datasets and identifying patterns, while the ontological approach provides a structured method for personality detection.

3. Research framework

This study proposes a framework model to enhance research concerning customer complaints, as depicted in Figure 2. Regarding customer traits, it is crucial to consider that specific individuals exhibit extroverted and assertive personality traits, whereas others demonstrate more introverted and reticent characteristics. These varying dispositions can influence how customers articulate their grievances, their choice of vocabulary, the intensity of emotional expression, and other aspects. Furthermore, complaints within organizations may stem from ambiguous policies and procedures. Employees or customers requiring role and responsibility clarification, or problem resolution guidance, encounter confusion and difficulties. This lack of clarity leads to errors, postponements, or disputes, adversely affecting organizational performance, employee and customer satisfaction, and loyalty.

Historically, companies collected complaints via conventional channels such as email or designated complaint numbers, resulting in protracted complaint management processes. Customers lodge grievances on social media platforms regarding product or service issues, expediting complaint submissions due to the platforms' accessibility, prompt response times, experience-sharing capabilities, and ease of providing feedback. The proposed framework addresses customer behavior and the challenges experienced therein. It consists of two distinct stages: personality and complaint analysis.

The first stage of our research framework involves personality analysis. It has been explained previously that customer personalities are diverse, and to facilitate our understanding of them, we utilize the big five personality (BFT) model. The framework broadly classifies personality traits into five categories: openness, conscientiousness, extraversion, agreeableness, and neuroticism. This analysis helps identify the root causes of customer complaints, as certain personality traits may indicate underlying problems beyond the product, leading customers to boycott the company and share negative reviews on social media. Each personality category is associated with specific characteristics and behaviors that we observe in customer complaint texts. Measuring these traits can provide valuable

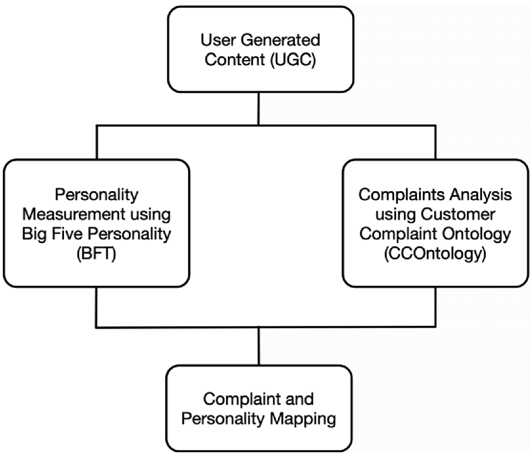


Figure 2. The research framework

Source(s): Figure by authors

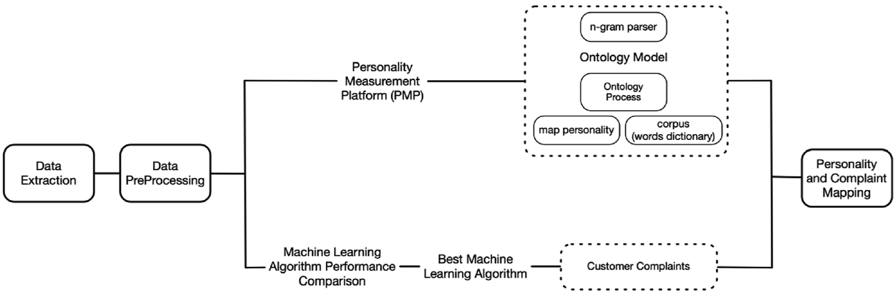
insights into customers' thought processes and behaviors. To carry out this analysis, we implement a personality measurement platform (PMP) on the informal Indonesian language, which will be further explained in Section 4.3.

The second stage involves analyzing customer complaints by examining their interactions on social media, specifically in e-commerce. The CCOntology framework facilitates a more effective complaint management process and enhances comprehension in categorizing and organizing information about products, customers, or other business process challenges. The selected machine learning algorithm is utilized to classify customer grievances into a hierarchical arrangement of categories and subcategories, as illustrated in Figure 1. This methodology aids in organizing and examining customer data, affording businesses a more profound understanding of their client's needs and preferences. A comprehensive elucidation of this analysis is furnished in the methodology section.

In the final stage, we map customer complaints based on CCOntology against their respective personalities based on the BFT. Previously, we have discussed the advantages of the two frameworks. We analyze how the nature of customer complaints in each complaint area. A detailed explanation is provided in Section 5.2. This approach enables companies to adapt to social media as it requires more proactiveness and interaction. It helps identify problem areas of the company, understands customer personalities effectively and efficiently, and anticipates social media effects.

4. Methodology

We outline the research methodology used to analyze customer complaint tweets directed to the official Twitter accounts of Indonesia's top three e-commerce companies: Tokopedia, Shopee, and Bukalapak. Their respective *Twitter* accounts are @TokopediaCare, @ShopeeCare, and @BukaBantuan. Our research framework is designed to accurately map customer complaints and their associated personality traits. To achieve this, we utilize the PMP ontology-based to analyze customer personality and the CCOntology framework to categorize customer complaints. We automatically classify customer complaints by comparing three machine learning algorithms (Support Vector Machine, Naïve Bayes, and Maximum Entropy). The best algorithm performer will be used as part of the proposed model. To fulfill the research objective, in the final step, we map customer complaints and their personality to identify dominant personality traits associated with specific complaint categories, providing a more comprehensive understanding of customer needs and preferences. The details of the research methodology are shown in Figure 3.



Source(s): Figure by authors

Figure 3.
Research methodology

4.1 Data extraction

We utilized the Twitter Application Programming Interface (API) to extract tweets related to customer complaints directed to e-commerce accounts. The case study includes the three biggest e-commerce services in Indonesia; they are @TokopediaCare, @ShopeeCare, and @BukaBantuan. The tweets are in Bahasa (Indonesian language), and the data was collected from August 2020 to August 2021. We only collect negative sentiment tweets to classify the problem based on CCOntology. To select the negative sentiment, we use Fasttext (FastText), a light implementation for sentiment analysis, which has Indonesian sentiment corpus. Further step is to analyze the personality behind the written tweets that represent customer characters using PMP. In total, we analyzed 30,000 tweets. The proportion tweets number of each e-commerce is listed in Table 2.

4.2 Data preprocessing

The primary objective of textual data preprocessing is to enhance the efficiency and accuracy of language models and ontologies by eliminating noise, inconsistencies, and irrelevant information. The preprocessing step ensures that the words in the tweets are easily readable and can be used for personality analysis (Hong *et al.*, 2008) and complaint analysis. The process typically includes several steps, such as tokenization, which involves splitting the text into individual words or tokens; lowercasing, to ensure uniformity and mitigate case sensitivity issues; removing stop words, which are common yet insignificant words (e.g. "and" "the" "is") that do not contribute to the text's meaning; stemming and lemmatization, to reduce words to their root form and improve model generalization; and removing special characters, punctuation, and numbers, to focus on the core textual content. Text preprocessing lays the groundwork for more effective and accurate language model training and analysis by performing these steps.

4.3 Personality measurement platform

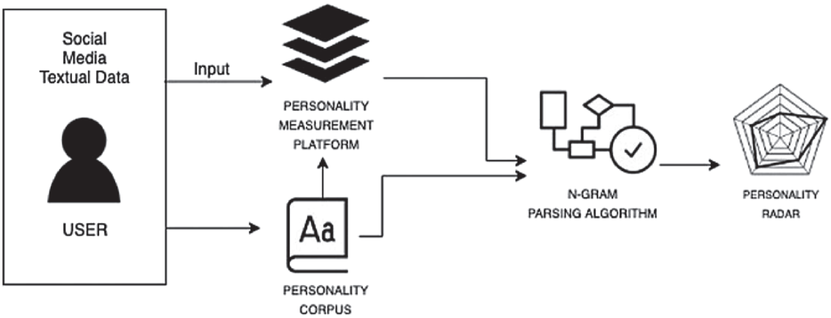
After preprocessing, the data is subjected to the BFT traits measurement approach. The refined tweets are input into the personality measurement platform (PMP), accessible at http://kepribadian.labscbd.id (Alamsyah *et al.*, 2021). PMP initially constructs an ontology using an ontology-based method, representing a formal and explicit specification of a domain's concepts, relationships, and attributes. This approach entails creating a standardized set of concepts and categories reflecting various aspects of the BFT, serving as the foundation for measurement. The n-gram language model parser assists in executing the application. Personality recognition involves identifying and analyzing individual personality traits based on behavior, speech, and other features. The PMP process diagram is shown in Figure 4.

Simultaneously, domain experts play a crucial role in enhancing the platform's accuracy by constructing a term-based library, known as a corpus. This corpus significantly bolsters text analysis by detecting language patterns and examining word sequences within a

Table 2. Number of tweets

E-commerce name	Account twitter	Number of tweets
Tokopedia	@tokopedia	10,207
Bukalapak	@bukalapak	4,506
Shopee	@ShopeeID	15,287
Total tweets	30,000	

Source(s): Table by authors



Source(s): Figure from Alamsyah *et al*. (2021)

Figure 4. PMP process diagram

specific text set. Currently, the personality measurement platform (PMP) operates exclusively in the Indonesian language, meticulously analyzing e-commerce customer personalities based on the gathered results. The e-commerce customer's personality is analyzed upon obtaining the results, with the outcomes discussed in Section 5. Applying personality measures analysis improves the understanding of customer behavior through their complaints. Text-based complaints undergo preprocessing and categorization using PMP to identify the dominant personality traits composed of several sub-traits. Table 3 presents sub-traits descriptions, which help to identify the corpus and categorized them into suitable BFT traits. Figure 5 shows a result example using a radar graph, visual representing personality measurement through PMP.

PMP relies on the principle of ontology text mapping to personality traits, where increased textual data enhances the accuracy of the assessment. Evaluating personality based on a single or few tweets is insufficient; however, a thorough analysis of a large corpus of tweets can provide a nuanced understanding of an individual's character. By aggregating and analyzing a large collection of tweets, we can accurately map them to specific personality traits, thus offering a comprehensive and insightful view of a person's or group's true nature.

4.4 Machine learning classifier

Parallel to the PMP process, customer complaint tweets are classified according to CCOntology. We extracted from Section 2.1 seven categories suitable for Indonesian customer characteristics: Conduct, Evidence, Privacy Issues, Pre-Purchase Phase, Purchase Phase, Post-Purchase Phase, and Selling Method. We first analyze the structure used to organize and categorize customer complaints based on the problem and language of the text sub-problem to categorize them according to this framework. Analyzing customer complaints via text, such as complaints about unusable transactions, may be classified based on the Pre-Purchase Phase category. Then, the sub-problem for the Personal Selling category.

The details process contains several steps. The first step is to construct training data to feed machine learning to automate multi-class classification tweets containing customer complaints into respective CCOntology, the seven categories mentioned above. This step is done by manually labeling or annotating tweets on each customer complaint category, supervised by language and language experts. We sample on average 500 tweets in each category. The second step is to choose several candidates of machine learning algorithms that perform well in Indonesian language categorization. They are Support Vector Machine, Naïve Bayes, and Maximum Entropy. In the third step, we compare the performance of the three machine learning algorithms for the NLP Indonesian language to classify the rest of the

Personality traits	Sub-trait	Description
Openness	Fantasy	The tendency to have a vivid imagination and daydream
	Aesthetic	Show an appreciation for art, beauty, and nature
	Feelings	Being in touch with one's own feelings and emotions
	Actions	Willingness to try new activities and experiences
	Idea	Curiosity and openness to unconventional ideas and beliefs
	Values	Readiness to re-examine traditional values and beliefs
Conscientiousness	Competence	A sense of capability and effectiveness
	Order	A preference for organization, structure, and planning
	Dutifulness	A sense of responsibility and adherence to rules and obligations
	Achievement Striving	The drive to set and meet high standards and goals
	Self-Discipline	The ability to begin tasks and follow through to completion despite distractions
	Deliberation	The tendency to think carefully before acting
Extraversion	Warmth	Affection and friendliness towards others
	Gregariousness	The enjoyment of being with groups and socializing
	Assertiveness	The tendency to speak up, take charge, and express opinions
	Activity-Level	A high energy level, leading a fast-paced life
	Excitement Seeking	A desire for stimulation and a taste for adventure
	Positive Emotions	Experiencing joy, enthusiasm, and other positive feelings
Agreeableness	Trust	Believing that others are honest and well-intentioned
	Straightforwardness	Being frank, sincere, and candid with others
	Compliance	A tendency to avoid conflict and get along with others
	Altruism	A selfless concern for the well-being of others
	Modesty	A lack of arrogance or self-importance
	Tender Mindedness	Being empathetic and sympathetic to the feelings of others
Neuroticism	Anxiety	A tendency to worry and be nervous
	Hostility	A tendency to feel anger and resentment
	Depression	A tendency to feel sad, discouraged, and hopeless
	Self-Consciousness	Feeling uncomfortable in social situations due to perceived judgments from others
	Impulsiveness	Difficulty controlling cravings and urges
	Vulnerability	A susceptibility to stress and the feeling of being overwhelmed

Source(s): Table adapted from Alamsyah *et al.* (2021)

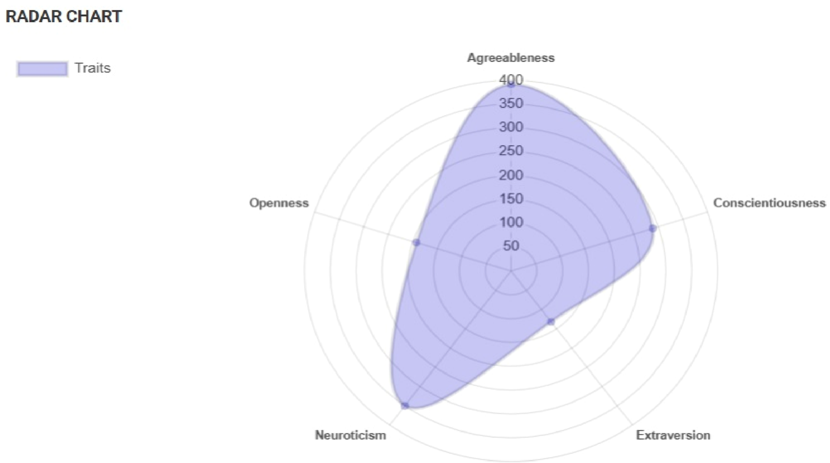
Table 3.
The big five personalities traits and sub-traits description on PMP

tweets, all 30,000 tweets. We applied 10 cross-validations to measure the robustness of algorithm performance. Multi-class classification problem performance is measured by Micro Precision (MicroP), Macro Precision (MacroP), Micro Recall (MicroR), Macro Recall (MacroR), Micro F1 Score (MicroF1), Macro F1 Score (MacroF1), and overall accuracy (Sokolova and Lapalme, 2009).

Micro metrics (MicroP, MicroR, MicroF1) give equal weight to each instance, making them suitable for imbalanced datasets where each instance is equally important. Macro metrics (MacroP, MacroR, MacroF1) treat all classes equally, providing an average performance across all classes, which can be helpful when the class distribution is balanced or when you want to give equal importance to each class. Accuracy is a straightforward measure of overall correctness but may not be informative for imbalanced datasets. Table 4 shows a detailed multi-class classification metrics definition and formulas.

4.5 Personality and complaint mapping

The final stage of our methodology is to map complaints and personality traits. We employ a descriptive approach to identify the predominant personality corresponding to each issue.



Source(s): Figure by authors

Figure 5. An example of visual personality measurement on PMP

Table 4. Multi-class classification metric

Definition	Formula
MicroP is micro-averaged precision that measures the overall precision of all classes (c_i) by calculating the sum of True Positives (TP) and dividing it by the combined sum of each class's TP and False Positives (FP)	$\text{MicroP} = \frac{\sum_{c_i \in C} \text{TPs}(c_i)}{\sum_{c_i \in C} \text{TPs}(c_i) + \text{FPs}(c_i)}$
MacroP is macro-averaged precision that calculates the average precision of all classes (c_i) by determining each class's precision and then averaging the values	$\text{MacroP} = \frac{\sum_{c_i \in C} P(D, c_i)}{\lvert C \rvert}$
MicroR is micro-averaged recall that calculates the average recall of all classes (c_i) by summing True Positives (TP) and False Negatives (FN) across all classes, then dividing the total TP by the combined sum of TP and FN	$\text{MicroR} = \frac{\sum_{c_i \in C} \text{TPs}(c_i)}{\sum_{c_i \in C} \text{TPs}(c_i) + \text{FNs}(c_i)}$
MacroR is macro-averaged recall that calculates the overall average recall by determining the recall value for each class (c) and then averaging all the recall values	$\text{MacroR} = \frac{\sum_{c_i \in C} R(D, c_i)}{\lvert C \rvert}$
MicroF1 is micro-averaged F1-score that aggregates F1-scores of all classes. Calculate it by multiplying microP and microR, then multiplying the result by two and dividing it by the sum of microP and microR	$\text{MicroF1} = 2 \cdot \frac{\text{MicroP} \cdot \text{MicroR}}{\text{MicroP} + \text{MicroR}}$
MacroF1 is macro-averaged F1-score that calculates the average of all F1-score by determining each label's F1-Score and then averaging them	$\text{MacroF1} = \frac{1}{N} \sum_{i=0}^{N} \text{F1}$

Source(s): Table summarized from Sokolova and Lapalme (2009)

The selected traits represent the dominant personality in general terms. The categorization of customer complaints is structured according to the CCOntology framework, which includes the categories of Conduct, Evidence, Privacy Problem, Pre-Purchase Phase, Purchase Phase, Post-Purchase Phase, and Sales Method. We have established a correlation between these customer complaint categories and personality traits. Each category is analyzed to determine the proportion of personality traits, based on the word count converted into percentages, to identify the predominant personality traits associated with each type of customer complaint. A thorough examination of these mappings and the analysis results is presented in Section 5, Table 7.

5. Result and analysis

5.1 Customer complaints classification

We have compared three machine learning algorithms for multi-class classification of the informal Indonesian language: Support Vector Machine (SVM), Naive Bayes, and Maximum Entropy. Among these, Naive Bayes outperformed the others across several evaluation metrics, including Micro Precision (MicroP), Macro Precision (MacroP), Micro Recall (MicroR), Macro Recall (MacroR), Micro F1 Score (MicroF1), Macro F1 Score (MacroF1), and overall accuracy, as demonstrated in Table 5. Consequently, we have chosen Naive Bayes for customer complaints classification based on the informal Indonesian language, categorizing them into seven common categories derived from CCOntology. The superior performance of Naive Bayes highlights its effectiveness in handling the nuances of complaints expressed in the informal Indonesian language, making it the most suitable algorithm for this task.

Naïve Bayes classifier automatically classifies complaints into problems and sub-problems based on CCOntology categories illustrated in Figure 1 to detect customer complaints from the text. For example, when a customer types the Indonesian translated text into *"admin, i am having problems with my voucher claim.",* This kind of text is detected to be included in the Pre-Purchase Phase problem since it discusses how one cannot claim the voucher offers for the intended purchase. Some examples of issues, complaints tweets (in the original informal Indonesian language), translations and contexts, and explanations on the CCOntology category are shown in Table 6.

5.2 Personality and complaint mapping

In personality and complaints mapping, we explore each problem we found in the data with their respective personality. The framework enables the determination of the proportion of customer personality traits for each complaint area, as demonstrated in Table 7, which shows the distribution of customer personality in each complaints category. The personality measurement produces a radar graph, shown in Figure 5, which contains each BFT score. The aggregate score portrays the personalities of the people who make the complaints. We analyze the relationship between customer complaints and personality traits by identifying the two most frequent complaint characteristics in each area of concern. As a result, it becomes apparent which customer personality traits frequently emerge across different complaint areas. For example, Table 7 no 1 shows that rudeness or untruthfulness as part of a conduct problem is commonly voiced by neuroticsm and conscientiousness personalities.

E-commerce customer complaints in the Conduct category predominantly stem from traits of neuroticism and conscientiousness. Neurotic customers frequently articulate their dissatisfaction with strong emotional expressions (Lampropoulos *et al.*, 2022). In contrast, conscientious individuals tend to deliver meticulously detailed information and accurately report issues (Mansur *et al.*, 2023). By recognizing these traits, customer service departments can significantly improve their responsiveness. For neurotic customers, adopting strategies such as active listening, demonstrating empathy, and providing practical solutions can effectively address their concerns and enhance their satisfaction. Meanwhile, acknowledging the precision and thoroughness of conscientious customers can lead to more efficient and

Table 5. Machine learning algorithms performance comparison

Algorithm	MicroP	MacroP	MicroR	MacroR	MicroF1	MacroF1	Accuracy
Support vector machine	0.85	0.83	0.84	0.82	0.845	0.825	0.85
Naïve Bayes	0.92	0.90	0.91	0.89	0.915	0.895	0.92
Maximum entropy	0.89	0.87	0.88	0.86	0.885	0.865	0.89

Source(s): Table by authors

Problem	Complaint in original tweet	Translation and context	Explanation and CCOntology category
Conduct	*"dm tidak menyelesaikan masalah, parah banget cara penyelesaian kalian."*	*"(contacting via) dm (direct message) does not solve the problem, and the way (they) solve the problem is really bad."*	Complaints about the effectiveness of e-commerce service to handle the situation. (Rudeness)
	"akun tiba tiba di batasi dengan dalih melanggar snk padahal saya sama sekali tidak melanggar. di mintai bukti jawabannya hanya rahasia internal."	*"my account access suddenly got limited without any good reason, when (i) asked about the explanation, the answer was internal discretion."*	Lack of process/information clarity. (Untruthfulness)
Evidence	*"pembeli mengunggah foto yang diklaim sebagai barang kami. produk tersebut berbeda dari yg kami kirim."*	*"buyer uploaded the photo for the claiming process, but (the photo shows) a different product than the seller delivered."*	Chronological explanation. (Declaration)
	"iya sampe kapan bos, dokumen yang kalian minta udah dikirim semua"	*"until when (we have to wait), all the requested documents have been delivered."*	Customers have already sent the document to solve their problem. (Attachment)
Pre-purchase phase	*"min kok gak bisa pake voucher gratis ongkir sih, padahal udah klaim voucher."*	*"(admin) why can't i use the free voucher for the delivery fee, though i have already claimed the voucher."*	Cannot use the voucher after claiming. (Offer Problem)
	"halo. kenapa ya kalau saya klik deskripsi, langsung frozen ngga bisa balik ke halaman pencarian."	*"hello. why is the page frozen each time i click the (product) description? I cannot go back to the search page."*	Customer cannot see the product description. (Information Problem)
	"susah banget sekarang kalo mau update stok produk. harus bgt ya diperiksa sampe berjam-jam? kan yg mau beli juga susah."	*"(it is) getting hard to update (item) stock. (why) does it take hours to verify (the item)?. Buyers will have to wait for this process."*	Seller complains that the product verification is taking too long. (Personal Selling)
Purchase phase	*"mohon bantuannya transaksi saya dibatalkan karena kadaluarsa waktu pembayaran. tapi saya sudah transfer."*	*"please help, my transaction has been canceled because of the expiration time session, but i have transferred (money)."*	The transaction is canceled due to the expiration time session. (Billing or Payment Problem)
	"kendala tidak di pick up oleh kurir min"	*"the problem is the courier did not pick up the (item)."*	Complaint that the courier did not pick up the item. (Delivery Problem)
	"baca dm saya dong. aneh banget gak ada beli tapi ada tagihan paylater."	*"please read my dm(direct message). So strange, i did not buy anything, but i got the invoice for paylater (program)."*	Complaint that they are being charged, even though they did not make any purchase. (Unexpected Charges)

(*continued*)

Table 6. Example of defined problem class explanation

Problem	Complaint in original tweet	Translation and context	Explanation and CCOntology category
Post purchase phase	*"aku melakukan pembelian barang, yg pada akhirnya aku cancel. tetapi seller tetap mengirimkan barang yg sudah di cancel sehari sebelumnya. Gimana atas solusi ini ?."*	*"i made a purchase, which at the end i canceled, but the seller kept sending the item, although i had canceled the order a day before. how to solve this (problem)?."*	Canceling the order, but the seller keeps sending the item. (Cancellation or Withdrawal)
	"barang sudah sampai tapi dana belum masuk saldo"	*"the item has been received (by the buyer), but the (money) has not been added yet to (my) saldo."*	The buyer has received the item, but the seller still has not received the money. (Deposit Withheld)
	"kok orderan saya ga masuk ke penjual"	*"(why) my order has not been received by seller."*	Complaint about the order that they made was not received by the seller. (Documentation Problem)
	"cashback ga masuk"	*"(i) have not received the cashback yet."*	Have not received the cashback. (Guarantee Problem)
	"aku pesen tapi yang dateng ga sesuai pesanan"	*"i ordered the item, but the one that arrived is different than (i) ordered."*	Receive different products than they purchase. (Product Problem)
	"pelapak selalu lama balas chat"	*"the seller is taking too long to answer my chat (question)."*	Seller response is slow. (Service Problem)
	"jadi biaya kirim ke pelapak buat perbaiki aku yang tanggung?"	*"so do i need to pay for the delivery cost to send the items back for repairing?"*	Questioning about repairing expedition cost. (Repair Problem)

Table 6.

Source(s): Table by authors

effective resolutions, fostering a positive customer experience. By tailoring their approach to these distinct personality traits, customer service teams can resolve issues more effectively and build stronger, more trusting relationships with their customers.

In the Evidence category, high levels of agreeableness result in complaints that are less hostile and aggressive (Marouf *et al.*, 2019). Customers exhibiting low conscientiousness are more likely to report issues such as damaged or incorrect items, often due to the company's oversight. Addressing these complaints efficiently and thoroughly can lead to higher customer satisfaction. Recognizing that customers' personality traits are multifaceted and may span multiple categories is crucial for e-commerce entities. By understanding and accommodating these complexities, companies can tailor their customer service approaches more effectively. Providing exceptional service that acknowledges these personality nuances can not only resolve issues more successfully but also build long-term relationships and uphold a positive market reputation. By valuing and responding to the diverse characteristics of their customer base, e-commerce companies can significantly enhance customer loyalty and trust.

In the Privacy Problem complaint category, neurotic customers often perceive situations as severe or threatening, making it difficult for them to calm down (Lampropoulos *et al.*, 2022). Additionally, these customers may exhibit openness traits (Min and Han, 2014), indicating a willingness to embrace new ideas. E-commerce companies must handle privacy issues with utmost care, as inadequate responses can lead to significant negative

No	Problem	Problem description	Personality	Words detected	Percentage
1	Conduct	Rudeness, Untruthfulness	Agreeableness	142	15.3%
			Conscientiousness	249	26.9%
			Extraversion	103	11.1%
			Neuroticism	277	29.9%
			Openness	153	16.5%
2	Evidence	Declaration, Attachment	Agreeableness	66	31.5%
			Conscientiousness	60	28.7%
			Extraversion	30	14.3%
			Neuroticism	36	17.2%
			Openness	17	8.1%
3	Privacy Problem	Hacked Accounts, Personal Data Disrupted, and personal electronic money jacked	Agreeableness	63	38.8%
			Conscientiousness	64	17.6%
			Extraversion	52	14.3%
			Neuroticism	102	28.1%
			Openness	81	22.3%
4	Pre-Purchase Phase	Offer Problem, Information Problem, Personal Selling	Agreeableness	190	22.5%
			Conscientiousness	173	20.5%
			Extraversion	134	15.9%
			Neuroticism	211	25.08%
			Openness	133	15.81%
5	Purchase Phase	Billing or payment Problems, Delivery Problems, Unexpected Charges	Agreeableness	139	18.7%
			Conscientiousness	162	21.8%
			Extraversion	79	10.6%
			Neuroticism	276	37.1%
			Openness	87	11.7%
6	Post Purchase Phase	Cancellation or Withdrawal, Deposit Withheld, Documentation Problem	Agreeableness	838	34%
			Conscientiousness	361	14%
			Extraversion	212	8.6%
			Neuroticism	836	33%
			Openness	214	8.6%
7	Sales Methods	Sellers' complaints about selling features or terms	Agreeableness	33	20.7%
			Conscientiousness	51	32%
			Extraversion	35	22%
			Neuroticism	16	10%
			Openness	24	15.09%

Source(s): Table by authors

Table 7. Mapping customer complaints and personality

consequences, including customer backlash and reputational damage. Poor management of personal information can erode customer trust, severely impacting loyalty and revenue. By effectively addressing privacy concerns with sensitivity and transparency, companies can mitigate risks, preserve their reputation, and foster stronger customer relationships. Prioritizing robust privacy protection measures and clear communication reassures customers, reinforcing their confidence and loyalty in the long term.

Neuroticism significantly influences customer complaints in the Pre-Purchase Phase category, driving heightened vigilance in detecting and reporting potential issues. During this stage, customers often scrutinize detailed and informative reviews to guide their decisions. Additionally, agreeable individuals, known for their patience, understanding, and empathy (Marouf *et al.*, 2019), play a vital role in this phase. Effective handling of their complaints is crucial, as it can significantly impact their purchasing decisions. E-commerce companies must prioritize addressing complaints with these personality traits in mind. Failure to do so can diminish the likelihood of purchase, potentially leading to lost sales. By proactively managing and resolving complaints with empathy and attention to detail, companies can positively influence customer decisions, enhance satisfaction, and boost overall sales. Tailoring complaint handling strategies to align with customers' personality traits is essential for fostering trust, encouraging positive purchasing behavior, and maintaining a competitive edge in the market.

The Purchase Phase, a pivotal stage in a customer's buying journey, encompasses the finalization of transactions after selecting a product or service. During this critical phase, conscientious customers offer detailed descriptions of any issues and actively collaborate with companies to resolve them. In contrast, neurotic individuals may experience heightened emotions, becoming upset or angry and finding it challenging to manage their reactions. Effectively addressing complaints during the Purchase Phase is crucial, as it profoundly impacts the customer's overall experience and influences future purchasing decisions. By promptly and empathetically resolving issues, companies can enhance customer satisfaction, foster loyalty, and encourage repeat business. Prioritizing attentive and tailored complaint handling during this phase not only mitigates potential dissatisfaction but also reinforces the customer's confidence in the brand, ultimately contributing to long-term success and a positive market reputation.

The Post-Purchase Phase is a crucial period for e-commerce businesses, as customers assess their products, services, and overall experiences. During this phase, agreeable customers tend to express complaints politely, whereas neurotic individuals often perceive situations as severe and promptly report any issues. Effectively addressing complaints in this phase is vital, as negative experiences can deter customers from making future purchases and lead them to share their dissatisfaction publicly. The negative experience can significantly damage the company's reputation and hinder future sales. By prioritizing exceptional customer service and resolving issues with empathy and efficiency, businesses can turn potential setbacks into opportunities for building stronger relationships. Ensuring a positive post-purchase experience fosters customer loyalty, encourages repeat business, and enhances the company's reputation, leading to sustained growth and success in the competitive e-commerce landscape.

In the Sales Method category, conscientiousness scores are generally lower, often requiring customers to provide detailed information about their problems, necessitating follow-up. However, extraversion is highly evident in this phase, as customers are more inclined to voice their complaints publicly on platforms such as social media or through public reviews (Mansur *et al.*, 2023). Effectively addressing issues during the sales phase is crucial for persuading customers to complete their purchases. Prompt and thorough resolution of complaints enhances customer satisfaction and mitigates the risk of negative public feedback, which can deter potential buyers. By actively engaging with customers and swiftly resolving their concerns, businesses can build trust, enhance their reputation, and ultimately drive sales. Ensuring a seamless and responsive sales process is essential for converting potential customers into loyal buyers and maintaining a competitive edge in the market.

The CCOntology framework categorizes customer complaints into seven distinct areas, frequently highlighting the traits of neuroticism and conscientiousness across these categories. Neurotic customers, characterized by emotional instability and psychological stress, often express their dissatisfaction through negative perceptions and harsh language, as observed in e-commerce platforms like Tokopedia, Shopee, and Bukalapak. Conversely, conscientious customers are goal-driven and collaborative problem solvers, providing detailed feedback to aid in issue resolution.

In testing the model's accuracy, 30 respondents were selected to post complaints reflective of their inherent personalities and behavioral patterns. With the help of psychology experts, the model accurately identified the correct personalities or behavioral tendencies in 25 out of the 30 participants, achieving an impressive 83% success rate. This high level of accuracy underscores the model's exceptional performance in assessing and categorizing individual personalities based on expressed complaints. These results demonstrate the model's potential efficacy in applications requiring a nuanced understanding of human personality and behavior. The ability to accurately identify personality traits from customer complaints makes this model a valuable tool for enhancing customer service, enabling personalized marketing, and improving overall user experience. By leveraging these insights, businesses can foster stronger customer relationships, tailor their approaches to individual needs, and maintain a competitive edge in the market. The CCOntology framework, supported by this highly accurate model, offers a promising solution for understanding and addressing the complex dynamics of customer behavior.

6. Discussion

(1) *Accuracy of the proposed model*:

The foremost discussion point is the impressive accuracy of our proposed model, which boasts an 83% success rate in predicting individual personalities, as detailed in the previous section. This high accuracy underscores the model's robustness and potential for practical application. By leveraging open data from social media – a platform where Indonesian digital natives frequently express their concerns publicly – we tap into a rich vein of unfiltered feedback. While our outcomes approximate actual behavior, the reliance on open data offers a unique advantage in social control and feedback for digital services. Nevertheless, broader testing is essential to generalize these promising results further. The introduction of newer and more advanced techniques, such as NLP based on deep learning, transformers, and large language models (LLM) could potentially increase the accuracy of a similar framework. However, these techniques are more complex to implement in a practical sense, require higher expertise in computer science, and higher costs in computational resources.

(2) *Model generality:*

The second critical discussion point pertains to the model's generality. Our model has undergone meticulous adjustments to enhance its sensitivity and responsiveness to the Indonesian language, capturing its linguistic nuances and idiomatic expressions. While this fine-tuning benefits applications targeting the Indonesian-speaking population, it inherently limits the model's applicability across different linguistic and cultural contexts. As a result, the model's current state does not support broad generalization of its effectiveness and accuracy to other languages and cultures, highlighting an area for future development.

(3) *Ethical considerations:*

The third discussion addresses the ethical considerations of using customer data without explicit written consent. We posit that data shared publicly on social media is intended for public attention and scrutiny, a practice deeply ingrained in Indonesian culture. Public airing of grievances is preferred over private negotiations with customer service, often seen as a faster route to issue resolution. This cultural tendency reflects a diminished trust in online companies when collective issues are not openly shared and addressed. Nonetheless, the ethical implications of this approach necessitate careful consideration and ongoing dialogue.

(4) *Significance of the study:*

The fourth discussion centers on the study's significance in terms of effectiveness and efficiency. This study is pivotal for its potential to revolutionize the strategies and communications used in public discourse. Tailoring responses to customer complaints based on personality traits can lead to more productive issue resolution. This prompts an inquiry into whether customizing customer service in this manner can enhance problem-solving efficacy. Moreover, the study explores the potential for such a tailored approach to aid businesses in the optimal allocation of resources, identifying complaints that may require heightened scrutiny and intervention.

(5) *Managerial implications:*

Implementing this approach has profound managerial implications, particularly in enhancing customer relationship management through personalization rather than intrusion. This strategy can provide a competitive advantage, supporting long-term business performance by fulfilling customer satisfaction and fostering loyalty. Management must closely monitor several potential issues, including the impact on customer behavior, the long-term reliability and validity of the methods, potential legal issues related to data protection laws, and the effect on customer service employees. Addressing these considerations can ensure the sustainable integration of this innovative approach into business strategies.

7. Conclusion

Using digital footprints on social media, we propose a methodology to map customer complaints against their personality. By integrating the big five personality model via an ontology-based platform (PMP) and CCOntology via Naive Bayes machine learning to categorize complaint areas, the procedure shows how the underlying personality can explain individual customers' behavior, reactions, and complaints. Future research can help companies recognize patterns and trends in customer complaints, enabling them to address recurring issues and proactively provide personalized customer service.

In today's data-rich age, companies can harness extensive review data strategically to gain insights into diverse customers and offer tailored treatment. Technological advancements and the utilization of big data approaches have provided us with exceptional capabilities to address customer complaints. Through the analysis of personality traits and customer complaints, we can acquire a deeper understanding of issues spanning various domains and industries. This approach enables more precise customer personalization, empowering companies to identify areas for improvement and comprehend customer characteristics. Ultimately, this fosters customer loyalty and attracts new buyers to the business.

The research limitation lies in the dependency on personality models' accuracy, which relies on the limited Indonesian corpus. Building language models presents a significant challenge, as influenced by social and cultural aspects. The progress of the large language model (LLM) research may significantly increase the speed and accuracy of model creation and customer complaints classification. We suggest incorporating the LLM for future research to mitigate the limitations and model generalization.

References

Agostino, D. and Sidorova, Y. (2017), "How social media reshapes action on distant customers: some empirical evidence", *Accounting, Auditing and Accountability Journal*, Vol. 30 No. 4, pp. 777-794, doi: 10.1108/aaaj-07-2015-2136.

Alamsyah, A. and Adityawarman, F. (2017), "Hybrid sentiment and network analysis of social opinion polarization", *2017 5th International Conference on Information and Communication Technology (ICoIC7)*, pp. 1-6.

Alamsyah, A., Widiyanesti, S., Putra, R.D. and Sari, P.K. (2020), “Personality measurement design for ontology based platform using social media text”, *Advances in Science, Technology and Engineering Systems Journal*, Vol. 5 No. 3, pp. 100-107, doi: 10.25046/aj050313.

Alamsyah, A., Dudija, N. and Widiyanesti, S. (2021), “New approach of measuring human personality traits using ontology-based model from social media data”, *OR Informatie*, Vol. 12 No. 10, p. 413, doi: 10.3390/info12100413.

Amichai-Hamburger, Y. and Vinitzky, G. (2010), “Social network use and personality”, *Computers in Human Behavior*, Vol. 26 No. 6, pp. 1289-1295, doi: 10.1016/j.chb.2010.03.018.

Baršauskas, P., Šarapovas, T. and Cvilikas, A. (2008), “The evaluation of e-commerce impact on business efficiency”, *Baltic Journal of Management*, Vol. 3 No. 1, pp. 71-91, doi: 10.1108/17465260810844275.

Beck, E.D. and Jackson, J.J. (2020), “A mega-analysis of personality prediction: robustness and boundary conditions”, *Journal of Personality and Social Psychology*, Vol. 122 No. 3, pp. 523-553, doi: 10.1037/pspp0000386.

Bender, P.M. (2020), “Limits of personalization of default rules – towards a normative theory”, *Law and Economics eJournal*. doi: 10.2139/ssrn.3544029.

Biswas, K., Shivakumara, P., Pal, U. and Lu, T. (2023), “A new ontology-based multimodal classification system for social media images of personality traits”, *SIViP*, Vol. 17 No. 2, pp. 543-551, doi: 10.1007/s11760-022-02259-3.

Castelo-Branco, I., Oliveira, T., Coelho, P.S., Portugal, J. and Filipe, I. (2022), “Measuring the fourth industrial revolution through the Industry 4.0 lens: the relevance of resources, capabilities, and the value chain”, *Computers in Industry*, Vol. 138, p. 103639, doi: 10.1016/j.compind.2022.103639.

Dennhardt, S. (2013), *User-Generated Content and its Impact on Branding: How Users and Communities Create and Manage Brands in Social Media*, Springer Science & Business Media, New York.

Dhelim, S., Chen, L.L., Aung, N., Zhang, W. and Ning, H. (2021), “Big-five, MPTI, Eysenck or HEXACO: the ideal personality model for personality-aware recommendation systems”, *ArXiv, abs/2106.03060*.

Douglas, J., Cumming, Sofia, J., Zaheer, K. and Martin, M. (2022), “E-commerce policy and international business”, *Management International Review*, Vol. 63 No. 1, pp. 3-25, doi: 10.1007/s11575-022-00489-8.

Du, T.V., Yardley, A.E. and Thomas, K.M. (2020), “Mapping big five personality traits within and across domains of interpersonal functioning”, *Assessment*, Vol. 28 No. 5, pp. 1358-1375, doi: 10.1177/1073191120913952.

Dudhat, A. and Agarwal, V. (2023), “Indonesia’s digital economy’s development”, *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, Vol. 4 No. 2, pp. 109-118, doi: 10.34306/itsdi.v4i2.580.

Dudija, N., Natalia, L., Alamsyah, A. and Romadhony, A. (2022), “Identification of extraversion and neuroticism personality dimensions using IndoBERT’s deep learning model”, *2022 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, pp. 155-159.

Eggert, A., Frow, P., Payne, A. and Steinhoff, L. (2020), “Understanding and managing customer value propositions: introduction to the special issue”, *Industrial Marketing Management*, Vol. 87, pp. 242-243, doi: 10.1016/j.indmarman.2020.01.007.

Fan, Y. and Niu, R.H. (2016), “To tweet or not to tweet? Exploring the effectiveness of service recovery strategies using social media”, *International Journal of Operations and Production Management*, Vol. 36 No. 9, pp. 1014-1036, doi: 10.1108/ijopm-10-2013-0461.

FastText (n.d.), “Crawl vectors”, *FastText*, available at: https://fasttext.cc/docs/en/crawl-vectors.html (accessed 14 June 2024).

Geovanni, F.Y.M., Alamsyah, A. and Dudija, N. (2021), "Identifying personality of the new job applicants using the ontology model on twitter data", *2021 2nd International Conference on ICT for Rural Development (IC-ICTRuDev)*, pp. 1-5.

Gielens, K. and Steenkamp, J. (2019), "Branding in the era of digital (dis)intermediation", *International Journal of Research in Marketing*, Vol. 36 No. 3, pp. 367-384, doi: 10.1016/j.ijresmar.2019.01.005.

Golmohammadi, A., Havakhor, T., Gauri, D.K. and Comprix, J.J. (2021), "Complaint publicization in social media", *Journal of Marketing*, Vol. 85 No. 6, pp. 1-23, doi: 10.1177/00222429211002183.

Greenberg, P. (2010), "The impact of CRM 2.0 on customer insight", *Journal of Business and Industrial Marketing*, Vol. 25 No. 6, pp. 410-419, doi: 10.1108/08858621011066008.

Gustav, G. and Špet. (2022), "The ontology of the social as a theory of social forms", in *The Centrality of Sociality* (Vol. 39, pp. 149-173). Emerald. doi: 10.1108/s0278-120420220000039008.

Haider, S., Afzal, M.T., Asif, M., Maurer, H.A., Ahmad, A. and Abuarqoub, A. (2018), "Impact analysis of adverbs for sentiment classification on Twitter product reviews", *Concurrency and Computation: Practice and Experience*, Vol. 33 No. 4, doi: 10.1002/cpe.4956.

Hanaysha, J.R. (2022), "Impact of social media marketing features on consumer's purchase decision in the fast-food industry: brand trust as a mediator", *International Journal of Infosis Management Data Insights*, Vol. 2, p. 100102, doi: 10.1016/j.jjimei.2022.100102.

Hong, R.Y., Paunonen, S.V. and Slade, H.P. (2008), "Big Five personality factors and the prediction of behavior: a multitrait–multimethod approach", *Personality and Individual Differences*, Vol. 45 No. 2, pp. 160-166, doi: 10.1016/j.paid.2008.03.015.

Huang, Y., Chai, Y., Liu, Y. and Shen, J. (2019), "Architecture of next-generation e-commerce platform", *Tsinghua Science and Technology*, Vol. 24 No. 1, pp. 18-29, doi: 10.26599/tst.2018.9010067.

Hugging Face (n.d.), "bert-base-Indonesian-522M", *Hugging Face*, available at: https://huggingface.co/cahya/bert-base-indonesian-522M (accessed 14 June 2024).

Ibrahim, N.F. and Wang, X. (2019), "A text analytics approach for online retailing service improvement: evidence from Twitter", *Decision Support Systems*, Vol. 121, pp. 37-50, doi: 10.1016/j.dss.2019.03.002.

Jarrar, M. (2009), "Towards effectiveness and transparency in E-business transactions: an ontology for customer complaint management", in *Semantic Web for Business: Cases and Applications*, (pp. 127-149,). IGI Global. doi: 10.4018/978-1-60566-066-0.ch007.

Jarrar, M., Verlinden, R. and Meersman, R. (2003), *Ontology-Based Customer Complaint Management*, eBusiness & eCommerce, New York.

Jiang, M. and Liu, J. (2019), "Research on internet financial personalized services based on big data analysis", *Prosiding International Conference on Financial Management, Education and Social Science (FMESS)*.

Joseph, G. and George, A. (2007), "A framework to integrate the enterprise domain ontology and organizational change application domain", *Accounting Research Journal*, Vol. 15 No. 2, pp. 3-23, doi: 10.1108/18347640710837326.

Karthik, R.V. and Ganapathy, S. (2021), "A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce", *Applied Soft Computing*, Vol. 108, p. 107396, doi: 10.1016/j.asoc.2021.107396.

Kunathikornkit, S., Piriyakul, I. and Piriyakul, R. (2023), "One-to-one marketing management via customer complaint", *Social Network Analysis and Mining*, Vol. 13, pp. 1-17, doi: 10.1007/s13278-023-01082-z.

Kusumawati, R.D., D'arofah, A. and Pramana, P.A. (2019), "Comparison performance of naive Bayes classifier and support vector machine algorithm for twitter's classification of Tokopedia services", *Journal of Physics: Conference Series*, Vol. 1320 No. 1, p. 012016, doi: 10.1088/1742-6596/1320/1/012016.

Lambiotte, R. and Kosinski, M. (2014), "Tracking the digital footprints of personality", *Proceedings of the IEEE*, Vol. 102, pp. 1934-1939, doi: 10.1109/jproc.2014.2359054.

Lampropoulos, G., Anastasiadis, T., Siakas, K. and Siakas, E. (2022), "The impact of personality traits on social media use and engagement: an overview", *International Journal on Social and Education Sciences (IJonSES)*, Vol. 4 No. 1, pp. 34-51, doi: 10.46328/ijonses.264.

Langford, P.H., Dougall, C.B. and Parkes, L.P. (2017), "Measuring leader behaviour: evidence for a "big five" model of leadership", *The Leadership and Organization Development Journal*, Vol. 38 No. 1, pp. 126-144, doi: 10.1108/lodj-05-2015-0103.

Lee, C., Wang, Y. and Trappey, A.J. (2015), "Ontology-based reasoning for the intelligent handling of customer complaints", *Computers and Industrial Engineering*, Vol. 84, pp. 144-155, doi: 10.1016/j.cie.2014.11.019.

Li, Y. and Yang, Y. (2012), "The study on the impaction of customer service needs on customer behaviors in E-commerce circumstances", *2012 International Joint Conference on Service Sciences*, pp. 181-184.

Li, L. and Yang, L. (2023), "A classification and extraction method of attribute hybrid big data based on Naive Bayes algorithm", *Journal of Computational Methods in Science and Engineering*, Vol. 23 No. 4, pp. 1955-1970, doi: 10.3233/jcm-226802.

Liu, Z. and Park, S. (2015), "What makes a useful online review? Implication for travel product websites", *Tourism Management*, Vol. 47, pp. 140-151, doi: 10.1016/j.tourman.2014.09.020.

Liu, Z., Long, C., Lu, X., Hu, Z., Zhang, J. and Wang, Y. (2019), "Which Channel to ask my question?: personalized customer service request stream routing using deep reinforcement learning", *IEEE Access*, Vol. 7, pp. 107744-107756, doi: 10.1109/access.2019.2932047.

Lopes, C.L., Cabral, B. and Bernardino, J. (2016), "Personalization using big data analytics platforms", *Proceedings of the Ninth International Conference on Computer Science & Software Engineering*, pp. 131-132, doi: 10.1145/2948992.2949000.

Mansur, Z., Omar, N. and Tiun, S. (2023), "Twitter hate speech detection: a systematic review of methods, Taxonomy analysis, challenges, and opportunities", *IEEE Access*, Vol. 11, pp. 16226-16249, doi: 10.1109/access.2023.3239375.

Marouf, A.A., Hasan, M.K. and Mahmud, H. (2019), "Identifying neuroticism from user generated content of social media based on psycholinguistic cues", *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, pp. 1-5.

Mazuelas, S., Shen, Y. and Pérez, A. (2020), "Generalized maximum entropy for supervised classification", *IEEE Transactions on Information Theory*, Vol. 68 No. 4, pp. 2530-2550, doi: 10.1109/tit.2022.3143764.

Mccoll-Kennedy, J.R. and Smith, A.K. (2006), "Chapter 10 customer emotions in service failure and recovery encounters", *Research on Emotion in Organizations*, Vol. 2, pp. 237-268, doi: 10.1016/S1746-9791(06)02010-4/FULL/XML.

Min, W.Z. and Han, Y. (2014), "Emotions as strategic information: examining the direct and ripple effect of emotions in negotiations", *2014 International Conference on Management Science & Engineering 21th Annual Conference Proceedings*, pp. 958-963.

Mudambi, S.M. and Schuff, D. (2010), "What makes a helpful online review? A study of customer reviews on Amazon.com", *MIS Quarterly*, Vol. 34 No. 1, pp. 185-200, doi: 10.2307/20721420.

Ojino, R.O., Mich, L. and Mvungi, N.H. (2022), "Hotel room personalization via ontology and rule-based reasoning", *International Journal of Web Information Systems*, Vol. 18 Nos 5/6, pp. 369-387, doi: 10.1108/ijwis-02-2022-0045.

Onyancha, O.B. (2015), "An informetrics view of the relationship between internet ethics, computer ethics and cyberethics", *Library Hi Tech*, Vol. 33 No. 3, pp. 387-408, doi: 10.1108/lht-04-2015-0033.

Panagariya, A.S. (2019), "Digital revolution, financial infrastructure and entrepreneurship: the case of India", *Emerging Markets: Finance eJournal*.

Picciali, V. and Sciandrone, M. (2022), "Nonlinear optimization and support vector machines", *Annals of Operations Research*, Vol. 314 No. 1, pp. 15-47, doi: 10.1007/s10479-022-04655-x.

Prada, M., Saraiva, M., Cruz, S., Xavier, S. and Rodrigues, D.L. (2022), "Using emoji in response to customer reservation requests and service reviews", *Human Behavior and Emerging Technologies*, Vol. 2022, pp. 1-10, doi: 10.1155/2022/1433055.

Rajapaksha, D.S., Perera, K., Sandaruwan, S.G., Lakchani, O., Nugaliyadde, A. and Thelijjagoda, S. (2017), "Linguistic features based personality recognition using social media data", *2017 6th National Conference on Technology and Management (NCTM)*, pp. 63-68.

Ruben, C., Arslan., Stefan, C., Schmukle., Julia, M. and Rohrer. (2023), "Emotional (in)stability: neuroticism is associated with increased variability in negative emotion after all", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 120 No. 23, p. e2212154120, doi: 10.1073/pnas.2212154120.

Sashi, C.M., Brynildsen, G. and Bilgihan, A. (2019), "Social media, customer engagement and advocacy", *International Journal of Contemporary Hospitality Management*, Vol. 31 No. 3, pp. 1247-1272, doi: 10.1108/IJCHM-02-2018-0108.

Sawicki, J., Ganzha, M. and Paprzycki, M. (2023), "The state of the art of natural language processing—a systematic automated review of NLP literature using NLP techniques", *Data Intelligence*, Vol. 5 No. 3, pp. 707-749, doi: 10.1162/dint_a_00213.

Shams, G., Rehman, M.A., Samad, S. and Rather, R.A. (2020), "The impact of the magnitude of service failure and complaint handling on satisfaction and brand credibility in the banking industry", *Journal of Financial Services Marketing*, Vol. 25 Nos 1-2, pp. 25-34, doi: 10.1057/s41264-020-00070-0.

Sokolova, M. and Lapalme, G. (2009), "A systematic analysis of performance measures for classification tasks", *Information Processing and Management*, Vol. 45 No. 4, pp. 427-437, doi: 10.1016/j.ipm.2009.03.002.

Strauss, J.P. and Hill, D.J. (2001), "Consumer complaints by E-mail: an exploratory investigation of corporate responses and customer reactions", *Journal of Interactive Marketing*, Vol. 15 No. 1, pp. 63-73, doi: 10.1002/1520-6653(200124)15:1<63::aid-dir1004>3.0.co;2-c.

Sun, M. and Zhao, J. (2022), "Behavioral patterns beyond posting negative reviews online: an empirical view", *Journal of Theoritical Applied Electronic Commerce Research*, Vol. 17 No. 3, pp. 949-983, doi: 10.3390/jtaer17030049.

Tronvoll, B. (2011), "Negative emotions and their effect on customer complaint behaviour", *Journal of Service Management*, Vol. 22 No. 1, pp. 111-134, doi: 10.1108/09564231111106947.

Umashankar, N., Ward, M.K. and Dahl, D.W. (2017), "The benefit of becoming friends: complaining after service failures leads customers with strong ties to increase loyalty", *Journal of Marketing*, Vol. 81 No. 6, pp. 79-98, doi: 10.1509/jm.16.0125.

Verhagen, T., Nauta, A. and Feldberg, F. (2013), "Negative online word-of-mouth: behavioral indicator or emotional release?", *Computers in Human Behavior*, Vol. 29 No. 4, pp. 1430-1440, doi: 10.1016/j.chb.2013.01.043.

Wongthongtham, P. and Abu-Salih, B. (2018), "Ontology-based approach for identifying the credibility domain in social Big Data", *Journal of Organizational Computing and Electronic Commerce*, Vol. 28 No. 4, pp. 354-377, doi: 10.1080/10919392.2018.1517481.

Wright, A.J. and Jackson, J.J. (2023), "Do changes in personality predict life outcomes?", *Journal of Personality and Social Psychology*, Vol. 125 No. 6, pp. 1495-1518, doi: 10.1037/pspp0000472.

Zhang, X., Wang, X., Tian, F., Xu, D. and Fan, L. (2022), "Anticipating the antecedents of feedback-seeking behavior in digital environments: a socio-technical system perspective", *Internet Research*, Vol. 33 No. 1, pp. 388-409, doi: 10.1108/intr-09-2021-0672.

Zhao, Z., Wang, J., Sun, H., Liu, Y., Fan, Z. and Xuan, F. (2020), "What factors influence online product sales? Online reviews, review system curation, online promotional marketing and seller guarantees analysis", *IEEE Access*, Vol. 8, pp. 3920-3931, doi: 10.1109/access.2019.2963047.

Further reading

Junior, V.G.D.R., Kulkarni, A. and Antonio, A.D.A. (2023), "Classifying customer complaints of a large fixed broadband service provider using machine learning", *2023 International Wireless Communications and Mobile Computing (IWCMC)*, (pp. 1334-1339), IEEE, doi: 10.1109/IWCMC58020.2023.10182843.

Ball, A.D., Coelho, P.S. and Vilares, M.J. (2006), "Service personalization and loyalty", *Journal of Services Marketing*, Vol. 20 No. 6, pp. 391-403, doi: 10.1108/08876040610691284.

Churchi, M. (2023), "Prediction of personality traits and suitable job through an intelligent interview agent using machine learning", *Journal of Computer Sciences and Applications*, Vol. 15 No. 1, pp. 21-31, doi: 10.37624/jcsa/15.1.2023.21-31.

Coelho, P.S. and Henseler, J. (2012), "Creating customer loyalty through service customization", *European Journal of Marketing*, Vol. 46 Nos 3/4, pp. 331-356, doi: 10.1108/03090561211202503.

Emily, J. and Getzen. (2022), "Behavioral analytics of consumer complaints", in *Advances in Business Information Systems and Analytics Book Series*, pp. 42-67, doi: 10.4018/978-1-6684-4246-3.ch003.

Indira, D., R., Athilakshmi. and Amulya (2023), *Personality Prediction Using Deep Learning*. IEEE, New York, pp. 1-5, doi: 10.1109/ICAECT57570.2023.10117573.

Isha, G., Manasvi, J. and Prashant, J. (2023), "Smart-hire personality prediction using ML", *2023 International Conference on Disruptive Technologies (ICDT)*, (pp. 381-385), IEEE. doi: 10.1109/ICDT57929.2023.10151367.

McKinsey (n.d.), "The value of getting personalization right—or wrong—is multiplying", available at: https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-value-of-getting-personalization-right-or-wrong-is-multiplying (accessed 7 June 2023).

Raksmey, S., Pei-Chun, L. and Shu-Yi, L. (2023), "Understanding customers' insights using attribution theory: a text mining and rule-based machine learning two-step multifaceted method", *Applied Sciences*, Vol. 13 No. 5, p. 3073, doi: 10.3390/app13053073.

Corresponding author
Andry Alamsyah can be contacted at: andrya@telkomuniversity.ac.id