Customer Perceptions of AI-Powered Chatbots and Their Influence on Engagement in Retail Banking

University of Surrey

Faculty of Arts and Social Sciences

By Sophia Nelson

Submitted in part fulfilment of the requirements for the degree of Master of Science in Business Analytics

Word Count: 13,985

Executive Summary

AI-powered chatbots are now a standard feature of retail banking, increasingly used to automate customer service interactions, reduce operational costs, and enhance digital service availability. Despite widespread deployment, banks continue to face uncertainty regarding how customers engage with chatbot-related services and which system attributes meaningfully drive sustained usage. In a sector characterised by high perceived financial and privacy risk, engagement with AI-powered interfaces depends not only on efficiency but also on customer trust and confidence in system performance.

This study examines customer engagement with AI-powered chatbots in retail banking using a multi-method quantitative approach that integrates survey-based regression analysis with machine learning-based sentiment analysis of customer reviews. The research is grounded in established information systems and consumer behaviour frameworks, including the Technology Acceptance Model (TAM), the Stimulus-Organism-Response (SOR) framework, Customer Engagement Theory (CEB), and Service-Dominant (SD) logic, which frames chatbots as value-enabling service actors within banking ecosystems. Together, these perspectives conceptualise chatbot attributes as service stimuli that shape internal customer evaluations, most notably trust, which subsequently influence engagement behaviours.

Primary data were collected via an online survey capturing customer perceptions of chatbot engagement, trust, satisfaction, empathy, personalisation, and functional service quality attributes such as speed, accuracy, and ease of use. Descriptive analysis indicated generally positive evaluations of functional performance, while relational attributes were evaluated more modestly. Correlation analysis revealed substantial multicollinearity among functional service quality measures, suggesting that customers cognitively evaluate these attributes as a unified dimension of system performance rather than as discrete features. To preserve statistical validity, functional attributes were therefore excluded from the regression model and analysed descriptively, while relational constructs were retained for inferential testing.

Multiple regression analysis indicated that exhibited trust strong association with customer engagement, approaching conventional levels of statistical significance. Satisfaction and personalisation exhibited positive but non-significant relationships with engagement, while

empathy showed a negative, non-significant association. Although statistical power was constrained by sample size, the relative strength and direction of effects were consistent with theoretical expectations and prior research in high-stakes financial service environments, reinforcing the centrality of trust in sustained engagement.

To complement, validate and contextualise the survey findings, sentiment analysis was conducted on a large dataset of U.S. banking customer reviews using machine learning techniques, including logistic regression, XGBoost, and transformer-based (BERT) models. The primary interpretability model selected was the logistic regression model, with the sentiment analysis revealing that negative customer experiences were predominantly driven by functional service failures, such as delayed responses, incorrect information, unresolved issues, and difficulties escalating to human agents. Relational concerns, including empathy and emotional support, were less prominent in customer narratives. These findings indicate that customer dissatisfaction is primarily rooted in breakdowns in functional reliability and trust rather than deficiencies in conversational warmth or personalisation.

Taken together, the results demonstrate that while functional performance represents a base expectation, trust operates as the critical differentiator shaping customer engagement with AI-powered chatbots in retail banking. Customers appear willing to engage with chatbots for routine tasks provided the system is reliable, secure, and accurate, but disengagement is likely when trust is undermined by service failures or opaque escalation processes. Relational attributes such as empathy and personalisation play a secondary role, potentially reflecting both current technological limitations and customer expectations within task-oriented, Western banking contexts.

This study makes several contributions. Theoretically, it extends technology acceptance and engagement research by empirically distinguishing engagement from adoption and demonstrating that engagement in risk-intensive financial contexts is primarily influenced by cognitive trust than by relational or affective attributes. Methodologically, it illustrates the value of combining regression analysis with machine learning-based sentiment analysis to triangulate perceptual and behavioural evidence, while also highlighting the analytical implications of perceptual overlap among functional service quality measures.

From a managerial perspective, the findings suggest that banks should prioritise investments in functional reliability, data security, transparency, and escalation design when deploying chatbot systems. Enhancements to empathy or personalisation are unlikely to offset deficiencies in accuracy, response quality, or trustworthiness. Continuous monitoring of customer sentiment and operational performance metrics is therefore essential to identify service breakdowns before they result in disengagement. More broadly, chatbot strategies should be embedded within integrated human-AI service models that allow seamless transition to human agents when complexity or risk increases.

The study is subject to limitations, including a cross-sectional survey design, a modest sample size, and the use of general banking reviews rather than chatbot specific interaction data. Future research could address these limitations through longitudinal designs, analysis of chatbot interaction logs, or behavioural approaches examining hybrid human-AI service models in real-world banking environments.

This research thus provides evidence that trust is the primary driver of customer engagement with AI-powered chatbots in retail banking. By combining perceptual survey data with large-scale sentiment analysis, the study offers a robust, analytics-driven understanding of customer engagement and proposes actionable guidance for banks seeking to move beyond chatbot adoption toward sustained, trust-based digital service relationships.

Declaration of Originality

I hereby declare that this thesis has been composed by myself and has not been presented or accepted in any previous application for a degree. The work, of which this is a record, has been carried out by myself unless otherwise stated and where the work is mine, it reflects personal views and values. All quotations have been distinguished by quotation marks and all sources of information have been acknowledged by means of references including those of the Internet. ***I agree that the University has the right to submit my work to the plagiarism detection sources for originality checks.***

Author’s signature: Sophia Nelson

Full Name: Sophia Nelson

Date: 05/01/2026

Table of Contents

1. Introduction

Artificial intelligence (AI) has rapidly transitioned from a speculative technology to a core component of organisational strategy across industries. In financial services, and retail banking in particular, AI-driven systems are increasingly embedded into everyday customer interactions. According to the Bank of England (2024), approximately 75% of financial services firms are already using some form of AI, with a further 10% planning adoption within the next three years. This widespread uptake reflects both competitive pressure and the growing expectation that banks deliver faster, more accessible, and more efficient digital services.

One of the most visible applications of AI in retail banking is the use of conversational agents, commonly referred to as chatbots. These computer programs are designed to simulate conversation with a human user (Oxford English Dictionary, n.d.). Banking chatbots are typically deployed as customer-facing interfaces within mobile applications, websites, or messaging platforms, where they handle routine enquiries such as balance checks, transaction queries, and account responses, and scalability at a relatively low cost. As customer service volumes continue to increase and cost pressures intensify, chatbots have become an attractive solution for automating front-line service delivery.

Despite their growing prevalence, customer responses to banking chatbots remain mixed. While chatbots are often praised for speed and convenience, and subsequently leading to ongoing engagement, they are also frequently criticised for their limited problem-solving capability and perceived lack of human understanding. These shortcomings are particularly salient in banking contexts, where interactions often involve sensitive personal data, financial risk, and stress-inducing situations. As a result, customers may tolerate chatbots solely for simple tasks but disengage when trust, reassurance, or complex support is required (Upadhyay and Kamble, 2023). This apprehension raises important questions about which customer perceptions of AI-powered chatbots strongly influence engagement, as gains in speed and convenience coexist with concerns around trust and human understanding.

From a business analytics perspective, understanding customer engagement with AI-powered services is critical because engagement directly translates into measurable business outcomes. Engagement extends beyond initial usage and reflects the extent to which customers are willing to interact with, rely on, and continue using a service over time. In digital banking, sustained engagement is closely linked to customer retention, reduced service costs, cross-selling opportunities, and long-term profitability. For banks, chatbot interactions generate large volumes of behavioural and perceptual data, yet without analytical insight, these data offer limited managerial value (Li, Fang and Chiang, 2023). Engagement is not driven by technology deployment alone; it is shaped by how customers perceive the system's functionality, reliability, and trustworthiness. Therefore, identifying which chatbot attributes most strongly influence engagement represents a practical analytics challenge, as banks must determine where to allocate resources, optimise chatbot design, and intervene to prevent disengagement based on data-driven evidence rather than intuition.

Existing research on chatbots in banking and customer service has largely concentrated on technology adoption, usability, or satisfaction outcomes, with many studies emphasising functional attributes such as response speed, availability, and ease of use as predictors of positive evaluations. More recent studies have incorporated relational constructs, including trust, empathy, and personalisation, reflecting advances in conversational AI. However, much of this literature relies on experimental designs, short-term evaluations, or single-method approaches, which limits understanding of how customer perceptions shape engagement in real-world banking contexts (Hentzen et al., 2021).

From a business analytics perspective, a key limitation is the lack of integrated empirical approaches that link customer perceptions to observable behavioural evidence at scale. Survey-based studies provide structured measurement of theoretical constructs but are subject to perceptual and self-report bias, while large-scale customer reviews offer rich behavioural insight but lack alignment with established engagement and trust frameworks. This disconnect is particularly problematic for AI-enabled services, where engagement is influenced by both cognitive evaluations of system performance (e.g. accuracy, reliability, speed) and affective responses such as trust, reassurance, or frustration. Addressing this gap represents a practical analytics challenge for banks, as it requires combining inferential statistical methods with data-driven text analytics to identify which chatbot attributes meaningfully influence customer engagement, rather than merely generating positive surface-level evaluations.

This study addresses these limitations by adopting a multi-method quantitative design that integrates survey-based regression analysis with machine learning-based sentiment analysis. The survey examines the relationship between customer perceptions of key chatbot attributes such as trust, satisfaction, empathy, personalisation with engagement using regression analysis. Careful model design was applied to address overlap among functional service quality perceptions.

The sentiment analysis component complements this approach by analysing large scale customer reviews of digital banking services to identify recurring themes and emotional patterns in customer experiences. By integrating structured survey data with unstructured textual evidence, the study mitigates the limitations of single method designs and strengthens the robustness and managerial relevance of the findings through triangulation.

The research is theoretically grounded in the Technology Acceptance Model (TAM), Stimulus-Organism-Response (SOR) framework, Customer Engagement Theory (CEB), and Service-Dominant (S-D) logic, which frames AI-powered chatbots as technology-enabled service actors participating in value co-creation (Davis, 1989; Mehrabian and Russell, 1974; Van Doorn et al., 2010; Vargo and Lusch, 2016). In the elevated risk context of retail banking, this study extends existing theory by shifting analytical focus from technology adoption to customer engagement, examining how customers' perceptions of functional and relational chatbot attributes jointly influence engagement in AI-mediated service interactions.

Building on this theoretical foundation, the study aims to evaluate customer engagement with AI-powered chatbots in the context of retail banking. To achieve this aim, the study examines the influence of customer perceptions of AI-powered chatbots on customer engagement, assesses perceptions of key functional and relational chatbot attributes, and analyses the relationship between these perceptions and customer engagement using regression analysis. In addition, the study examines large-scale customer sentiment towards digital banking services using machine learning-based text analytics and triangulates survey sentiment analysis findings to identify key factors of customer engagement and dissatisfaction.

The remainder of this dissertation is structured as follows. Chapter 2 critically reviews the literature on AI chatbots, customer engagement, and relevant theoretical frameworks. Chapter 3 outlines the research methodology and analytical techniques. Chapter 4 presents the empirical results. Chapter 5 discusses the findings in relation to theory and prior research and offers actionable business recommendations for managers. Finally, Chapter 6 provides further research suggestions based on the current study. Chapter 7 concludes by outlining the study's contributions, practical implications, limitations, and directions for future research.

2. Literature Review

Artificial intelligence (AI) has become increasingly integrated into customer service functions within retail banking, particularly through the deployment of AI-powered chatbots. These systems enable banks to manage routine enquiries, reduce operational costs, and provide continuous service availability. While such technologies deliver clear efficiency gains, customer evaluations of chatbot interactions vary considerably. These evaluations extend beyond functional performance and play a critical role in shaping how customers engage with digital banking services over time. Consequently, understanding the perceptual and relational factors that influence not only initial acceptance but also sustained engagement with AI-powered chatbots has emerged as a key research priority.

This literature review examines existing research on AI-powered chatbots in retail banking, with a specific focus on how customer perceptions influence engagement rather than adoption alone. While prior studies have predominantly explored technology acceptance, intention to use, or task performance, emerging research underscores the need to investigate post-adoption outcomes such as continued usage, relational quality, and customer engagement. To address this gap, the review integrates insights from the Technology Acceptance Model (TAM), the Stimulus–Organism–Response (SOR) framework, customer engagement theory (CEB), and service-dominant (S-D) logic, alongside research on trust in digital and AI-mediated services.

Drawing on S-D logic, the review conceptualises AI-powered chatbots as technology-enabled service actors that contribute to value co-creation through ongoing interactions between customers, firms, and digital systems. From this perspective, customer engagement reflects not only technology use but also relational and experiential outcomes that emerge through repeated service exchanges. This theoretical lens complements adoption-based models by emphasising interaction, resource integration, and the conditions under which value is co-created or constrained in AI-enabled banking services.

In addition to synthesising theoretical perspectives, the review evaluates empirical findings on AI-based customer service in retail banking and examines the methodological approaches commonly employed in this domain. Particular attention is given to the predominance of survey-based, cross-sectional designs and the growing use of text analytics and sentiment analysis as complementary data sources for understanding customer perceptions and engagement.

The chapter is organised into four sections. First, it outlines the key theoretical frameworks underpinning the study. Second, it reviews prior research on AI-powered chatbots and customer service in retail banking. Third, it discusses the research methods used to examine customer perceptions and engagement with AI technologies. Finally, it identifies key gaps, debates, and limitations in the existing literature and explains how the present study seeks to contribute to academic understanding and practical decision-making in AI-enabled retail banking.

2.1. Technology Acceptance Model

The Technology Acceptance Model (TAM) developed by Davis (1989), explains why some users come to accept and use technology. TAM incorporates perceived usefulness and ease of use as key features for explaining user behaviour when interacting with technologies (Davis and Venkatesh, 1996). Several researchers have applied this technology adoption theory to explore customer interactions with AI-powered chatbots and found that these constructs play an important role in encouraging customer usage (Dhanya and Ramya, 2025). However, TAM has been criticised for offering limited guidance on how practitioners can actively improve technology usage (Venkatesh and Bala, 2008); accordingly, this study seeks to address this limitation.

Further extending TAM, Venkatesh et al. (2003) proposed the Unified Theory of Acceptance of Use of Technology (UTAUT), which integrates multiple adoption models and explains technology use through performance expectancy, effort expectancy, social influence, and facilitating conditions. However, UTAUT was primarily developed for organisational contexts in which technology use may be mandatory or institutionally driven. As the present study examines voluntary customer engagement with AI-powered chatbots in retail banking, where usage is shaped by perceptual and relational factors such as trust and relationship quality, UTAUT was not adopted as the primary theoretical framework (Ramya, 2025).

2.2. Stimulus-Organism-Response Framework

The Stimulus-Organism-Response (SOR) framework, originally proposed by Mehrabian and Russell (1974), is beneficial for understanding how environmental stimuli influence internal psychological states and successive behavioural responses. Within digital service contexts, SOR has been widely applied to explain how technological features and service characteristics affect user cognitive and emotional perceptions, which subsequently shape behavioural engagement (Upadhyay and Kamble, 2023; Vafaei-Zadeh et al., 2024). In the context of AI-powered chatbots, chatbot attributes such as responsiveness, reliability, transparency, and personalisation can be conceptualised as external stimuli that triggers internal evaluations within the user, including trust (organism), which in turn influences engagement outcomes such as continued usage or interaction (Rohit et al., 2025).

While TAM explains technology acceptance primarily through perceived usefulness and ease of use, it offers limited insight into the psychological processes that translate perceptions into sustained engagement. The SOR framework complements TAM by explicitly accounting for the internal perceptual mechanisms through which chatbot attributes influence behavioural engagement outcomes. Accordingly, SOR provides as suitable theoretical framework for examining customer cognitive and emotional processes indirectly through perceptions of trust, satisfaction, personalisation, and empathy which function as antecedents to behavioural engagement.

2.3. Customer Engagement Theory

Marketing thought within services has evolved through three broad phases over time. Transactional marketing, dominant during the 1970s and 1980s, focused solely on discrete exchanges, with success measured through short-term financial metrics such as sales and market share. During the 1990s and early 2000s, the emphasis shifted toward relationship marketing, which prioritised customer retention, trust, and long-term loyalty, particularly in service contexts characterised by repeated interactions. Since the 2010s, service marketing has increasingly adopted a customer engagement perspective, reflecting the growing importance of interactive technologies and ongoing customer involvement, as evidenced by 37% increase in annual revenue when customers are fully engaged rather than actively disengaged in their retail bank (Islam and Rahman, 2016). Thus, this shift emphasises customers' active cognitive, emotional, and behavioural connections with firms as a source of sustained competitive advantage (Pansari and Kumar 2017; Van Doorn et al., 2010; Brodie et al., 2013).

Customer engagement has been conceptualised in the literature through multiple theoretical lenses, including engagement as a behavioural manifestation, a psychological state, a dispositional tendency, and a dynamic process (Van Doorn et al., 2010; Brodie et al., 2011; Storbacka et al., 2016; Maslowska et al., 2016). Among these perspectives, the behavioural manifestation view has been particularly influential in applied and service marketing research, defining engagement as customers' observable behaviours that extend beyond transactions, such as repeated usage, interaction frequency, and voluntary participation with a firm, commonly referred to as customer engagement behaviours (CEB) (Van Doorn et al., 2010).

Brodie et al. (2011) conceptualises customer engagement as a multidimensional psychological state comprising cognitive, emotional and behavioural components. This was later extended to emphasise its cyclical and processual nature over time (Brodie et al., 2013). Subsequent reviews indicate that engagement is most frequently operationalised behaviourally (46%), particularly in cross-sectional and survey-based studies, due to its measurability and managerial relevance (Ng, Sweeney, Plewa, 2020). Consistent with this dominant empirical stream, the present study adopts a behavioural conceptualisation of customer engagement, focusing on customers' self-reported usage and continued interaction with AI-powered banking chatbots. While cognitive and emotional dimensions of engagement are acknowledged as theoretically important, they are not directly examined in this study, which concentrates on identifying perceptual drivers that influence customers' willingness to engage behaviourally with chatbot services in retail banking.

2.4. Service Marketing Perspective

Research on AI in customer-facing financial services has grown rapidly, prompting calls for systematic examination of how AI reshapes service interactions and customer relationships (Hentzen et al., 2021). Given the service-focused nature of financial services, several studies have drawn on S-D logic to conceptualise value co-creation processes in banking, viewing value as emerging through interactions among multiple actors within a service ecosystem (Payne et al., 2021). For customers this process could involve giving opinions or commenting and ultimately, resulting in engagement (Ordanini et al., 2011). From this perspective, AI-enabled technologies, such as chatbots, can be understood as operant resources that support resource integration and facilitate service exchange.

However, existing applications of S-D logic in AI-enabled financial services remain limited and are often conceptual rather than empirical, prompting calls for greater empirical scrutiny of S-D logic-derived concepts (Hentzen et al., 2021; Vargo & Lusch, 2017). While some studies have begun to empirically examine value co-creation, co-destruction and ecosystem dynamics in technology-enabled services (Castillo et al., 2020; Payne et al., 2021), many research avenues remain unexplored and connections among actors within service ecosystems (Lusch & Nambisan, 2015). Notably, the extent to which AI enhances or disrupts relational connections in the financial services sector remains insufficiently understood (Hentzen et al., 2021).

Responding to calls for greater empirical examination of S-D logic concepts in AI-enabled financial services, the present study contributes by examining customer engagement with AI-powered chatbots as a relational outcome of technology-mediated service interactions. While not explicitly testing S-D logic propositions, the study adopts its core premise that value emerges through service interactions and focuses on how customers' perceptions of AI-enabled service encounters shape behavioural engagement. By examining trust, satisfaction, personalisation, and empathy as drivers of engagement of banking chatbots, this research provides empirical insight into how AI functions as an operant resource that facilitates ongoing interaction and service exchange. In doing so, the study offers a customer-level perspective on value co-creation in AI-enabled banking services, addressing the lack of empirical work identified in prior S-D logic informed research.

2.5. Trust and Perceptions in Digital Service

Despite the growing importance of digital transformation research, there is no universally accepted definition of digital trust, as conceptualisations vary depending on the technological context and the nature of human-technology interaction (Saveljeva and Volkova, 2025). Early foundational work by McKnight, Choudhury and Kacmar (2002) established digital trust as a multidimensional construct that integrates both institutional/system-level mechanisms (e.g., structural assurances) and interpersonal trust beliefs (e.g., competence, integrity). More recent research in the context of AI systems emphasises the importance of both cognitive trust (rational evaluations of capability and reliability) and affective trust (confidence and comfort during interaction), indicating that human-centric factors such as transparency, interaction design are just as important in influencing trust formation in technologically mediated environments (Riley and Dixon, 2024; Glassberg et al., 2025). While both views recognise the role of user perceptions, they differ in the object of trust: confidence in technological functionality versus trust formed through sociotechnical relationships. In AI-powered customer service contexts, these dimensions jointly shape how users evaluate digital systems and their willingness to engage with them.

In retail banking, where financial risk, privacy concerns, and data sensitivity are particularly salient, trust becomes a critical determinant of customer engagement. While prior research has largely examined trust in relation to AI adoption or continued use, limited attention has been given to how trust in AI-powered banking chatbots translates into deeper engagement behaviours, such as sustained interaction and willingness to rely on chatbot services for more complex financial tasks (Mogaji et al., 2021; Nguyen and Le, 2024; Kumar et al., 2025). Privacy and data security concerns further complicate this relationship, as apprehensions about safeguarding sensitive financial information may undermine trust and constrain engagement AI-mediated services (Upadhyay and Kamble, 2023).

Accordingly, digital trust in this study is defined as a customer's willingness to depend on AI-powered banking chatbots, based on perceptions of the system's reliability, security, competence, and integrity (cognitive trust), as well as feelings of confidence and comfort during interactions (emotional trust). This conceptualisation aligns with the study's focus on understanding how customer perceptions of AI-enabled services shape behavioural customer engagement.

2.6. AI-Powered Chatbots in Retail Banking

Early chatbot system deployed in customer service, and later adopted within retail banking, were predominantly rule-based, relying on predefined scripts and keyword recognition rather than semantic understanding (Adamopoulou and Moussiades, 2020). In banking contexts, these systems were primarily designed to handle simple, informational interactions and were constrained by rigid rule sets, sensitivity to linguistic variation, and a lack of conversational context awareness. As a result, their ability to support flexible, adaptive, or relationship-oriented service interactions were limited (Kaakandikar et al., 2025).

Subsequent developments introduced limited flexibility through retrieval-based approaches; however, a more substantial shift occurred with advances in artificial intelligence, natural language processing (NLP), and machine learning. From around 2016 onwards, AI-powered and generative chatbots emerged, capable of learning from large language models (LLM) to understand and respond to natural human language by maintaining conversational context and human-like mannerisms across interactions, resulting in more adaptive and anthropomorphic customer service experiences (Nicolescu and Tudorache, 2022).

This technological evolution is particularly significant in retail banking, where chatbots increasingly operate as frontline service interfaces. The transition from scripted to AI-powered chatbots has expanded the potential for more interactive and personalised service delivery, while simultaneously intensifying the importance of trust, reliability, and customer engagement in financial environments with elevated risk (Nicolescu and Tudorache, 2022).

Empirical studies examining early chatbot systems consistently document limitations related to rigidity, poor language understanding, and an inability to manage complex or ambiguous customer requests due to inflexible architectural design (Adamopoulou and Moussiades, 2020; Nicolescu and Tudorache, 2022). This can lead to irrelevant responses, conversational breakdowns, and escalation to human agents, limiting continuity of interaction. While these studies do not explicitly measure customer engagement outcomes, engagement theory suggests that such interactional constraints are likely to undermine perceived value and relational quality, thereby limiting the conditions necessary for sustained customer engagement. Thus, customer willingness to interact repeatedly is weakened as continued chatbot interaction depends on sustained, meaningful, and reliable service interactions rather than one-off task completion (Brodie et al., 2013). This provides an important theoretical basis for examining whether more advanced AI-powered chatbots, through improved adaptability and contextual understanding, are better positioned to support engagement in retail banking contexts.

2.7. Customer Satisfaction, Service Quality, and Engagement in AI Interactions

Customer satisfaction has long been a central construct in service marketing literature and is commonly defined as a customer's evaluative judgment of a service experience, formed through a comparison between expected and perceived performance (Oliver and DeSarbo, 1988; Wirtz and Bateson, 1999). In AI-powered service environments, satisfaction reflects customers' assessments of chatbot performance, including attributes such as responsiveness, accuracy, reliability, and perceived service quality. Empirical studies suggest that positive evaluations of these functional attributes support continued usage of digital services and may contribute to favourable customer attitudes toward AI-enabled interactions (Zainol et al., 2023; Afshaq et al., 2020). Within retail banking, where chatbots increasingly serve as frontline service interfaces, satisfaction remains an important indicator of whether basic performance expectations are met.

However, an emerging body of literature argues that satisfaction alone is insufficient to explain sustained customer behaviour in digital and AI-mediated service contexts. Customers may report being satisfied with a service while remaining passive users who interact only when necessary and exhibit limited emotional or behavioural involvement. As a result, scholars increasingly distinguish between satisfaction as an evaluative judgement of prior service experiences and customer engagement as a broader relational construct that captures ongoing involvement, emotional connection, and behavioural investment over time (Hari, Iyer and Sampat, 2021; Pansari and Kumar, 2017; Moravec et al., 2025).

From this perspective, satisfaction may support engagement but does not fully account for why customers choose to actively interact with, rely on, or remain committed to AI-enabled services. This distinction reinforces the relevance of customer engagement as the primary outcome of the present study, while positioning satisfaction as one of several perceptual factors associated with engagement in retail banking chatbot interactions.

2.8. Research Approaches in Prior Studies

2.8.1. Survey / Quantitative Studies

Survey-based quantitative methods are the dominant research approach in studies examining customer perceptions of AI technologies in banking and digital financial services (Dhanya and Ramya, 2023; Thasleena and Santhi, 2023). Recent reviews indicate that nearly half of existing studies employ quantitative designs with surveys constituting the primary data collection method across multiple years. Specifically, surveys account for most quantitative studies, reflecting widespread adoption for measuring perceptual constructs such as perceived usefulness, ease of use, risk, trust, and customer engagement (Gouveia and Santos, 2025).

However, the literature also highlights several limitations associated with the heavy reliance on survey-based methods. Excessive use of self-reported surveys raises concerns regarding common method bias, social desirability bias, and recall bias (Fadnes et al., 2009). While geographically concentrated sampling in many studies limits cultural diversity and constrains the generalisability of findings (Gouveia and Santos, 2025).

Despite these limitations, the extensive use of quantitative surveys provides a robust methodological foundation for the present study. By focusing on customer perceptions of AI-powered chatbots in retail banking and complementing survey findings with sentiment analysis of naturally occurring customer feedback, this study enhances analytical robustness through a multi-method approach, while remaining transparent about its cross-sectional scope.

2.8.2. Text Analytics of Customer Reviews

Text analytics refers to the extraction of meaningful patterns and insights from unstructured textual data, such as online reviews, social media posts, and customer feedback (Anandarajan et al., 2019). One of the most widely used applications of text analytics is sentiment analysis, which employs Natural Language Processing (NLP) techniques to classify the emotional tone of text as either positive, neutral, or negative. In digital service environments, including retail banking, sentiment analysis has become increasingly important because customers' emotional expressions and evaluations strongly influence perceptions of trust, satisfaction, and engagement (Gallagher et al., 2019).

Recent studies have applied text analytics to examine customer sentiment toward digital banking services. The analysis of customer reviews for digital banking applications using topic modelling and sentiment analysis revealed that user discussion frequently centred on user experience, customer service quality, and technical errors, with user experience identified as the most underserved dimension. These findings illustrate the value of text-based approaches for uncovering customer pain points and informing service improvement strategies (Suryadi and Padlan, 2024).

However, much of existing literature adopts a single-method approach when analysing customer data. Prior research highlights that text analytics applied to social media posts, online reviews, and other consumer interfaces provides a rich and scalable source of insight into consumers' AI-related attitudes and sentiments (Heinonen and Medberg, 2018), while potentially mitigating common survey biases such as social desirability. Advances in text mining, artificial intelligence, and machine learning further enable efficient data collection and large-scale analysis of unstructured text (Berger et al., 2020).

To address these methodological limitations, the present study integrates sentiment analysis with survey-based quantitative data, while recognising the constraints of each approach. The sentiment dataset provides naturally expressed, real-world customer feedback on digital banking experiences, but sentiment analysis is inherently sensitive to data quality, linguistic variation, and model selection, and may be affected by informal writing style, sarcasm, contextual ambiguity, irony, and domain-specific terminology (Wankhade et al., 2022). To complement these limitations, the survey captures structured perceptions of AI-powered chatbots, including trust, satisfaction, personalisation, and empathy, using validated measurement scales. Although the datasets are analysed separately, their joint interpretation enhances the robustness and credibility of the findings by combining behavioural evidence with perceptual measures, thereby mitigating the limitations associated with reliance on a single methodological approach.

2.9. Gaps, Debates, and Challenges

Although technological acceptance models such as TAM and UTAUT are widely applied, they have been increasingly critiqued in the context of artificial intelligence-enabled services. While these frameworks offer parsimonious explanations of perceived usefulness, ease of use, and behavioural intention, recent research suggests that they were not originally designed to capture the autonomous, adaptive, and opaque characteristics of AI systems (Sebetci Özel, 2024; Pramanik and Jana, 2025). In AI-mediated service environments—particularly within high-risk domains such as retail banking—customer evaluations extend beyond functional efficiency to include concerns related to trust, transparency, control, and perceived risk, constructs that remain weakly theorised within traditional acceptance frameworks (Hentzen et al., 2021).

Although prior research has widely integrated and extended TAM and UTAUT to examine user responses to AI technologies (Bhattacherjee, 2001; Dwivedi et al., 2021; Su et al., 2013), such extensions largely retain and intention-centric perspective, prioritising adoption and usage intentions rather than post-adoption relational behaviours such as sustained engagement, depth of interaction or relationship continuity. This limitation is explicitly highlighted in systematic reviews of AI in financial services, which note that few studies have examined the role of AI in fostering customer engagement, despite growing interest in AI-enabled customer interfaces (Hentzen et al., 2021). These reviews call for further research investigating how AI influences customer engagement behaviours and their consequences including relational and attitudinal outcomes (Hentzen et al., 2021). Consequently, while TAM and UTAUT-based models are effective in explaining initial acceptance, they provide only partial insight into how customers evaluate and engage with AI-powered banking services.

Addressing this gap, the present study integrates TAM with CEB and positions trust as a central perceptual construct linking chatbot characteristics to customer engagement outcomes. By shifting the analytical focus from intention to use toward sustained engagement, this study extends prior acceptance-based research and offers a more relational and service-oriented explanation of customer behaviour in AI-powered retail banking.

Methodologically, a substantial proportion of customer engagement research relies on cross-sectional survey designs that capture customer perceptions at a single point in time. Islam and Rahman (2016) identify this dominance of survey-based methods in their systematic review, noting that engagement research has historically prioritised measurement efficiency over methodological diversity. More recently, Ashrafuzzaman (2025) similarly observes that studies examining digital and AI-enabled services frequently adopt cross-sectional designs, limiting insight into how customer perceptions and engagement evolve as users gain experience with technology-mediated service interactions.

Subsequent reviews reinforce this critique, arguing that cross-sectional designs provide only a partial understanding of customer engagement, which is inherently dynamic and context dependent. Rosado-Pinto and Loureiro (2020) emphasise that engagement is increasingly recognised as complex and multi-layered, yet much empirical work continues to rely on static research designs that are unable to capture changes in engagement across interaction episodes and service touchpoints. Likewise, ElHamd et al. (2021) highlight that engagement is often treated as a single outcome variable rather than an ongoing process, constraining understanding of how engagement is formed, sustained, or weakened over time.

In AI-enabled service contexts, these methodological limitations are particularly pronounced. AI-mediated interactions are continuous, adaptive, and embedded within broader digital service ecosystems, suggesting the need for research approaches that extend beyond single-method, self-reported data. Reflecting this, Islam and Rahman (2016) explicitly call for greater methodological pluralism in customer engagement research, including the integration of behavioural data, digital traces, and advanced analytical techniques.

While the present study also adopts a cross-sectional survey design and therefore does not resolve the longitudinal limitations identified in prior research (Ashrafuzzaman, 2025), its methodological contribution lies in the integration of multiple data sources. Survey-based perceptual data are complemented by sentiment analysis of naturally occurring customer reviews related to digital banking services. Although each dataset is analysed independently to preserve methodological rigour, sentiment analysis provides behavioural and contextual insight that supports and enriches interpretation of the survey findings. This triangulated approach directly responds to calls for greater methodological diversity in customer engagement research (Islam and Rahman, 2016; Rosado-Pinto and Loureiro, 2020).

Customer engagement has emerged as a central construct within service and marketing research; however systematic reviews and meta-analyses consistently highlight a lack of clarity and consistency in how engagement is defined and operationalised across studies. Engagement has been examined through multiple lenses, including behavioural, cognitive, emotional, dispositional, and process-based perspectives, yet many empirical studies fail to clearly specify their conceptual stance, limiting comparability across findings (Barari et al., 2020). Prior reviews emphasise the need for greater conceptual precision, particularly in technology-enabled service contexts where engagement behaviours may differ from traditional services (Ng, Sweeney and Plewa, 2020). Within this literature, behavioural engagement is identified as one of the most established and empirically tractable perspectives, focusing on observable customer actions such as continued usage and interaction (Van Doorn et al., 2010; Ng, Sweeney and Plewa, 2020). This behavioural conceptualisation is especially suited to applied, cross-sectional research designs where engagement is examined as an outcome of customer perceptions, providing a clear and appropriate foundation for the present study.

Customer engagement research increasingly recognises that engagement is shaped by engagement-facilitating technologies that enable real-time, interactive, and relational exchanges between firms and customers. Early technology-enabled platforms such as social media and online brand communities illustrate how two-way interactivity enhances engagement through co-creation, commenting, and sharing (Kumar et al., 2016; Van Doorn et al., 2010; Trunfio and Rossi, 2021). More recently, scholars argue that the advent of advanced engagement-facilitating technologies including virtual and augmented reality, service robots and other artificial intelligence-based technologies will fundamentally transform customer engagement processes and outcomes at various micro by altering how individuals interact with firms and at the macro level through changes in organisational structures and service systems (Rosado-Pinto and Loureiro (2020).

Although customer engagement research has increasingly acknowledged the role of artificial intelligence, recent reviews indicate that AI-based customer engagement remains a nascent and underdeveloped research domain. While prior studies have begun to examine AI-enabled interactions, scholars highlight paucity of research investigating, which attributes of AI technologies shape customer engagement with specific AI applications and call for greater attention to the interplay between AI and customer engagement in increasingly automated service environments (Hollebeek et al., 2024). Responding to this gap, the present study focuses on AI-powered chatbots as a distinct form of engagement facilitating technology to advance the understanding of customer engagement with AI-technologies by addressing recent calls for technology-focused, perception-driven research in increasingly automated service environments.

Recent systematic reviews highlight that emotional and relational perceptions of AI play a central role in shaping customer responses yet research in this area remains limited (Iryna Pentina et al., 2023). While prior research has identified several perception-based characteristics relevant to AI-human interactions such as intimacy, warmth, anthropomorphism, trust, empathy, and personalisation, these constructs are often studied in isolation and primarily linked to adoption-related outcomes, including acceptance and initial satisfaction (Gur and Yossi Maaravi, 2025; Alabed et al., 2022). As a result, there is limited understanding of how these perceived characteristics jointly influence sustained customer engagement with AI technologies.

The literature further indicates that negative perceptions such as privacy concerns and technological or AI-related anxiety, are frequently cited as barriers to AI adoption, yet their direct empirical effects on user behaviour and engagement remain underexplored (Dhiman et al., 2023; Puertas et al., 2024, Payne et al., 2021). Similarly, trust-related and emotional variables such as trust in AI, perceived risk, and perceived empathy are often examined as mediators rather than central drivers of engagement outcomes, leaving their broader relational effects insufficiently understood (Ahmed and Aziz, 2024; Chi and Hoang Vu, 2022; Gouveia, J. and Santos, S. (2025). Recent systematic reviews therefore call for research that examines whether these perception-based characteristics operate consistently across different AI technologies, contexts, and industries, and how they shape outcomes such as customer engagement, loyalty, and willingness to pay (Chaturvedi et al., 2023).

Responding to these calls, the present study focuses on AI-powered chatbots and examines how customers' perceived trust, satisfaction, personalisation, and empathy influence behavioural engagement. By integrating multiple perception-based characteristics within a single engagement framework, this study addresses the fragmentation identified in prior research and advances understanding of how customers engage with technologies beyond initial adoption.

There is increasing emphasis on the need for more context-aware research on artificial intelligence across the customer journey, highlighting the influence of social norms, cultural orientations, and demographic characteristics on customer responses to AI-enabled interactions (Gouveia, J. and Santos, S, 2025). Although these contextual factors are widely acknowledged as relevant, they remain under-integrated in empirical research, particularly in studies of customer engagement with AI-powered chatbots.

Existing research offers limited insight into how social norms, shared values, and cultural orientations such as individualism versus collectivism, shape customer perceptions of engagement with AI technologies (Nam & Kannan, 2020). Similarly, demographic characteristics including age, digital familiarity, and technological readiness are frequently discussed as relevant boundary conditions but are rarely examined systematically due to data and design constraints (Chen et al., 2021).

While the present study does not formally model these contextual factors, they are acknowledged and discussed to contextualise the findings, highlighting the need for future research to incorporate cultural and demographic variables within larger-scale and cross-contextual engagement models. Addressing these contextual influences remains an important avenue for advancing understanding of customer engagement with AI-powered chatbots in diverse service environments.

2.10. Summary and Link to the Present Study

This literature review has examined the theoretical, empirical, and methodological foundations relevant to understanding customer engagement with AI-powered chatbots in retail banking. The review highlighted how technology adoption models such as the TAM provide useful insight into initial technology use but offer limited explanatory power for understanding post-adoption engagement and relational outcomes. To address this limitation, the review drew on CEB theory and the SOR framework to explain how customer perceptions of AI-enabled service interactions shape behavioural engagement outcomes over time.

The review further demonstrated the relevance of service marketing perspectives, particularly S-D logic, in conceptualising AI-powered chatbots as operant resources that facilitate value co-creation through technology-mediated service encounters. Within this context, trust emerged as a central perceptual construct, encompassing both cognitive and emotional dimensions, alongside satisfaction, personalisation, and empathy as key drivers shaping customers' willingness to engage behaviourally with chatbot services. While prior research acknowledges the operational efficiency of chatbots, it also highlights persistent concerns related to reliability, privacy, and the quality of human-AI interaction, particularly in high-risk financial service environments.

Methodologically, the review identified a strong reliance on cross-sectional survey-based research, with limited integration of behavioural or naturally occurring customer data. In response to calls for greater methodological pluralism, the present study adopts a triangulated research design that combines survey-based measures of customer perceptions with sentiment analysis of customer reviews. Although cross-sectional in nature, this approach enhances analytical robustness by integrating structured perceptual data with real-world expressions of customer experience.

By integrating adoption, engagement, and service marketing perspectives, this study advances existing research by shifting the analytical focus from technology acceptance toward behavioural customer engagement with AI-powered chatbots. The study contributes theoretically by empirically examining how perception-based characteristics jointly influence engagement, and practically by offering insights to support the design and deployment of trustworthy, engaging, and customer-centred AI-powered chatbot services in retail banking.

3. Methodology

3.1. Research Philosophy and Design

This study adopted a quantitative research design to examine customer perceptions of AI-powered chatbots in retail banking and their relationship with customer engagement, operationalised as self-reported chatbot usage and continued interaction. Primary data were collected through a structured online survey using Likert-scale items to measure key constructs, including engagement, trust, satisfaction, functional service quality attributes (e.g. ease of use, response speed, accuracy, and convenience), empathy, and personalisation. Survey data were analysed using descriptive statistics and multiple regression analysis to assess the direction and relative strength of relationships between perceived chatbot attributes and engagement outcomes.

In addition to the survey analysis, the study incorporated machine learning-based sentiment analysis using secondary data drawn from online customer reviews of digital banking services. Textual data were processed using automated sentiment classification based on pre-trained natural language processing models, producing numerical sentiment scores, probability outputs, and term frequencies. Although the source data were unstructured text, the analytical approach was computational and quantitative rather than interpretive or qualitative.

The combination of survey-based regression analysis and sentiment analysis enabled methodological triangulation by linking self-reported perceptual data with behavioural evidence derived from naturally occurring customer feedback. This approach strengthened the robustness of the findings by allowing patterns observed in survey responses to be interpreted alongside real-world expressions of customer sentiment within digital banking environments.

Philosophically, the study is grounded in post-positivist paradigm, which assumes that an objective reality exists in relation to customer engagement and perceptions of AI-powered chatbots, while recognising that such phenomena can only be imperfectly observed through empirical measurement. The use of Likert-scale survey instruments and probabilistic machine learning models reflects this position, as both generate estimates subject to measurement error, model assumptions, and perceptual bias. Overall, the quantitative methods employed align with a post-positivist approach to knowledge generation through empirical observation, statistical modelling, and inference rather than absolute certainty.

3.2. Data Sources, Sampling Methods and Data Preprocessing

This study utilised two quantitative data sources: primary survey data and secondary textual data for sentiment analysis. Primary data were collected through an online survey hosted on Microsoft Forms, designed to capture customer perceptions of AI-powered chatbots in the retail banking sector. The survey consisted of close-ended questions measured on a five-point Likert scale and included constructs such as engagement, trust, satisfaction, ease of use, helpfulness, convenience, response speed, response accuracy, empathy, and personalisation, alongside demographic variables including country of residence, age, and gender.

Secondary data were obtained from an open-access dataset of U.S. banking customer reviews sourced from Hugging Face, a widely recognised platform for machine learning and artificial intelligence research. The dataset was anonymised and accompanied by documentation outlining its structure and intended use, supporting transparency and reproducibility. No web scraping techniques were employed. The raw customer review data for sentiment analysis is provided as a separate data file submitted alongside the dissertation.

For the survey component, a non-random convenience sampling approach was adopted. Participants were recruited online, and participation was voluntary. A total of 30 valid responses were collected and included in the analysis. While this sample size limits statistical power, generalisability, and causal inference, multiple regression analysis was employed in an exploratory manner to examine the direction and relative strength of relationships between key perceptual variables and engagement, rather than generate population-level estimates. Completed survey responses are provided as a separate supplementary file submitted alongside the dissertation.

The sentiment analysis dataset comprised a large corpus of naturally occurring customer reviews related to digital banking services. As the data were publicly available and did not involve direct participant recruitment, no sampling intervention was applied. Instead, the dataset reflects authentic customer feedback generated through real-world service interactions.

Survey data were exported from Microsoft Forms and processed using Microsoft Excel and Power Query. Likert-scale responses were converted into numerical values to enable statistical analysis. Negatively worded items, particularly within empathy and response accuracy measures, were reverse-coded to ensure directional consistency. Composite scores were created by averaging related items within each construct to reduce measurement error and capture latent perceptions. The usage construct was renamed as engagement to maintain theoretical consistency with the customer engagement framework adopted in this study. Data cleaning involved checking for missing values, ensuring consistent formatting, and validating variable labels.

For the sentiment analysis, textual data were cleaned prior to modelling. Preprocessing steps included lowercasing text for consistency, removing missing entries and stop words, and standardising text inputs to improve model performance. The processed text was transformed into numerical representations suitable for machine learning-based sentiment classification. Model outputs provided sentiment polarity labels (positive, neutral, negative) alongside confidence scores, which were used to examine overall sentiment distributions and identify frequently occurring terms associated with negative and positive digital banking experiences. The full programming codes used for sentiment data preprocessing, model development, and evaluation are provided as a separate ZIP file submitted alongside the dissertation.

3.3. Variables and Measurements

The dependent variable in this study was customer engagement, conceptualised as the extent to which customers actively and willingly interact with banking chatbots and intend to continue using them. Consistent with the dominant empirical stream in service marketing, engagement is treated as a behavioural outcome, reflecting observable usage-related actions rather than internal psychological states.

Customer engagement was operationalised using self-reported behavioural indicators. Although objective usage data can offer greater precision, such data are often inaccessible in banking contexts due to privacy and proprietary constraints; consequently, self-reported measures are widely used as valid proxies for behavioural engagement in cross-sectional research (Gupta and Khan, 2024; Torkzdeh, 2021). A composite engagement score was derived from five Likert-scale items measuring chatbot usage frequency, ease of initiating interaction, habitual reliance, familiarity with chatbot features, future usage intention.

Independent variables comprised relational and functional constructs associated with AI chatbot interactions. All constructs were measured using multi-item Likert scales adapted from prior research in digital banking and technology acceptance. For each construct, item responses were averaged to create composite scores, reducing random measurement error and capturing underlying latent perceptions rather than single-item evaluations.

Relational constructs included:

- **Trust**, measured through perceptions of reliability, honesty, data security, and integrity
- **Satisfaction**, capturing overall contentment and expectation fulfilment
- **Empathy**, assessing perceived emotional understanding and sensitivity, with negatively worded items reverse-coded
- **Personalisation**, measuring perceived adaptation and tailoring of chatbot interactions

Functional service attributes such as response speed, accuracy, ease of use, helpfulness, and convenience were initially measured separately. However, these variables exhibited very high intercorrelations, indicating that respondents perceived them as a single underlying dimension of functional performance. To avoid multicollinearity and inflated standard errors, these variables were excluded from the regression model and interpreted descriptively as a unified functional construct.

Demographic variables were used to describe the sample and explore descriptive differences but were not included as control variables due to limited sample size and variance.

For the sentiment analysis component, the primary analytical outputs were sentiment polarity classifications (positive, neutral, negative), accompanied by probability scores indicating classification confidence. In addition, term frequency analysis was used to identify recurring functional and service-related issues associated with positive and negative customer experiences. These outputs were used to contextualise and enrich interpretation of survey findings rather than as direct inputs to the regression model.

3.4. Analytical, Statistical and Machine Learning Techniques

Survey data were first analysed using descriptive statistics to summarise overall patterns in customer perceptions across all constructs. Mean values, standard deviations, and response distributions were computed. Country of origin was used as a descriptive grouping variable to explore broad cross-country patterns; however, no inferential statistical comparisons were conducted due to the small sample size.

Multiple linear regression analysis was then employed to examine the relationship between perceived chatbot attributes and behavioural customer engagement. Regression was selected due to its ability to estimate the direction and relative strength of multiple predictors simultaneously, supporting hypothesis testing and theoretical comparison with prior technology acceptance and service marketing. Given the modest sample size, regression analysis was used in an exploratory and theory-informed manner, with emphasis on pattern identification rather than causal inference.

Prior to regression analysis, correlation analysis was conducted to assess interrelationships among predictor variables. Functional service attributes demonstrated substantial multicollinearity, indicating that respondents evaluated these attributes holistically rather than as distinct dimensions. To preserve statistical validity and interpretability and to avoid inflated standard errors these variables were excluded from the regression model and analysed descriptively as a unified functional construct.

Given the limited sample size, inferential findings were interpreted cautiously, with emphasis placed on coefficient direction, relative magnitude, near-significance, and theoretical consistency, rather than statistical significance alone. Model diagnostics included R^2, adjusted R^2, F-statistics, and residual analysis to assess overall model fit and regression assumptions.

For the sentiment analysis component, automated machine learning techniques were applied to classify sentiment polarity in customer reviews. Logistic regression served as the primary analytical model due to its transparency and interpretability, while more advanced models—including XGBoost and a pretrained BERT-based transformer—were employed as performance comparators. Although the BERT model demonstrated superior contextual understanding and classification accuracy, its limited interpretability rendered it unsuitable as the primary explanatory model for this study.

The sentiment analysis output was used to contextualise and enrich interpretation of the survey findings rather than as standalone evidence. Integrating regression analysis with machine learning-based sentiment analysis strengthened the study by combining multiple data sources, enabling perceptual drivers of engagement to be examined alongside real-world expressions of customer satisfaction and frustration within digital banking environments.

3.5. Validity, Reliability, and Ethical Considerations

Construct validity was supported using measurement items adapted from established research in technology acceptance, service quality, and AI-mediated customer service. Multi-item Likert scales were employed to capture latent constructs more accurately and reduce the limitations associated with single-item measures.

Internal validity is inherently limited by the cross-sectional research design, which restricts causal inference. While regression analysis was used to examine associations between key constructs, the findings are interpreted as correlational rather than causal, particularly given the modest sample size and reliance on self-reported survey data.

Ecological validity was strengthened through the inclusion of sentiment analysis based on naturally occurring customer reviews. These data reflect real-world digital banking interactions rather than hypothetical or experimentally induced scenarios, enhancing the realism of the findings.

Reliability was supported by using composite measures derived from multiple items for each construct, reducing random measurement error. Internal consistency was assessed using standard reliability indicators, with detailed results reported in the subsequent analysis chapter. Sentiment analysis reliability was further supported through consistent model-based classification and the use of probability scores to indicate confidence in sentiment assignments.

Ethical considerations were addressed by ensuring voluntary participation, informed consent, and anonymisation of all survey responses. Secondary data used for sentiment analysis were obtained from publicly available, anonymised sources in accordance with platform guidelines. All data were stored securely and used exclusively for academic research purposes.

4. Results Analysis

4.1. Survey Descriptive Statistics & Data Summary

The survey comprised 30 respondents. Table 1 summarises the demographic characteristics of participants, providing an overview of age, gender, education, employment status, country of origin, and language.

In terms of age distribution, the sample was skewed towards younger respondents. Participants aged 18-24 and 25-34 collectively accounted for over half of the sample (56%), while smaller proportions were represented in the 35-44 and 45-54 age groups (17% each). Only a small number of respondents were aged 55-64 (7%). This age profile reflects a sample that is likely to be familiar with digital banking services and AI-driven technologies.

Gender distribution showed a higher proportion of female respondents (57%) compared with male respondents (37%), with a small percentage preferring not to disclose their gender (7%). A range of ethnic backgrounds was represented, with White (33%), Asian (30%), Mixed (13%), and Black (13%) respondents included, suggesting some ethnic diversity within the sample.

Regarding educational attainment, most respondents held at least a bachelor’s degree, with 33% holding a master’s and 23% a doctorate. Nearly half were employed full-time (47%), 27% were students, and the remainder reported part-time work, unemployment or self-employment.

The sample was geographically concentrated, with most respondents based in England (83%). A smaller proportion reported being from outside the UK (10%). English was the primary language for most participants (70%), although a notable proportion did not report language information.

Table 1. Respondent Demographics (n = 30)[1]

Variable	Categories	Count	Percentage (%)
Age Group	18-24	7	23
	25-34	10	33
	35-44	5	17
	45-54	5	17
	55-64	2	7
	Prefer not to say	1	3
Gender	Male	11	37
	Female	17	57
	Prefer not to say	2	7
Ethnicity	White	10	33
	Mixed	4	13
	Asian	9	30
	Black	4	13
	Other	1	3
	Prefer not to say	2	7
Education Level	Secondary school	2	7
	Post-secondary	4	13
	Bachelor's degree	7	23
	Master's degree	10	33
	Doctoral	7	23
Employment Status	Full-time	14	47
	Part-time	4	13
	Self-employed	1	3
	Unemployed	2	7
	Student	8	27
	Prefer not to say	1	3
Marital Status	Married	6	20
	In a Relationship	9	30
	Single	10	33
	Divorced	2	7
	Prefer not to say	3	10
Country Origin	England	25	83
	Outside the UK	3	10
	Prefer not to say	1	3
Language	English	21	70
	Turkish	1	3
	Did not answer	8	27

In terms of banking usage, respondents reported using a range of retail banks. Monzo was the most frequently cited bank (27%), followed by Lloyds Bank (17%) and NatWest (13%). A substantial proportion of respondents selected “Other” banks (30%), indicating use of alternative or less commonly listed institutions. This distribution reflects a mix of both traditional high-street banks and digital-only banking providers within the sample.

[1] *Language is included as a background characteristic, as chatbot interactions rely on natural language processing and user comprehension may influence perceptions of service quality and engagement.*

Table 2. Banking Characteristics of Respondents (n = 30)

Category	Count	Percentage (%)
Natwest	4	13
Lloyds Bank	5	17
Monzo	8	27
Starling Bank	2	7
Other	9	30
Did not answer	6	20

Overall, the descriptive statistics indicate that the sample predominantly consists of younger, highly educated individuals with experience using digital banking platforms. While this profile limits the generalisability of the findings to broader populations, it provides an appropriate and relevant context for examining customer perceptions and behavioural engagement with AI-powered chatbots in retail banking.

As shown in Table 3, response speed recorded the highest mean score (M = 3.78, SD = 0.96), indicating that participants generally perceived the chatbot as delivering timely responses. In contrast, empathy (M = 2.87, SD = 1.10) and personalisation (M = 2.88, SD = 1.04) received the lowest mean scores, suggesting that respondents perceived these relational attributes less favourably. Overall, functional attributes tended to be rated more positively than relational attributes, highlighting a potential imbalance between chatbot operational performance and customers' perceptions of emotional understanding and personalised service.

Table 3. Descriptive Statistics of Survey Responses

Item	Mean	Max	Min	S.d.
Usage	3.17	5.00	1.00	1.22
Satisfaction	3.09	5.00	1.00	1.02
Trust	3.03	5.00	1.00	1.02
Ease of Use	3.33	5.00	1.00	1.07
Helpfulness	3.05	5.00	1.00	1.10
Convenience	3.23	5.00	1.00	1.10
Speed	3.78	5.00	1.00	0.96
Accuracy	3.05	5.00	1.00	1.03
Empathy	2.87	5.00	1.00	1.10
Personalisation	2.88	5.00	1.00	1.04

Figure 4.1 Average scores across items by country group
The graph shows that respondents that reside outside of the UK overall perceived chatbot interactions positively compared other country groups.

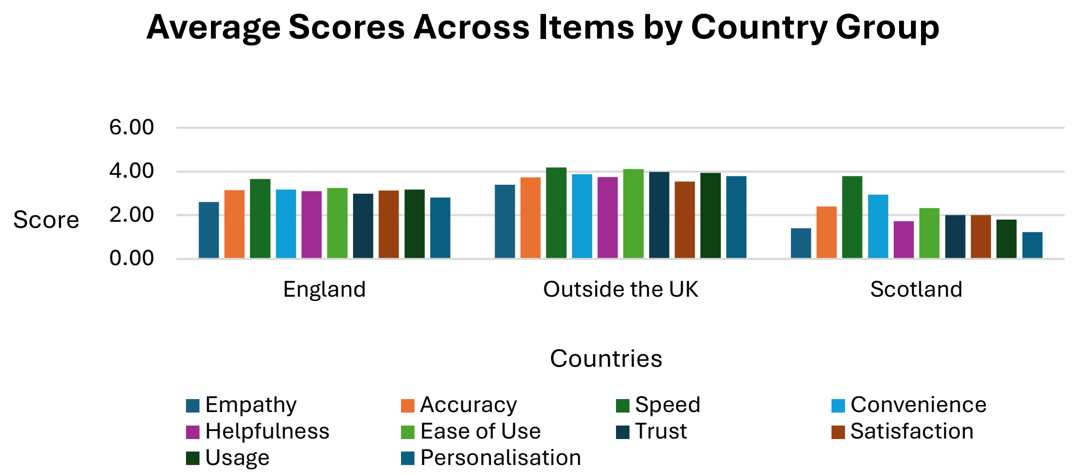


Figure 4.1 presents average construct scores by country group. Respondents residing outside the UK reported higher mean scores across all constructs compared with UK-based respondents, suggesting more positive overall perceptions of chatbot interactions among this group.

However, these differences should be interpreted with caution due to the very small number of non-UK participants (n = 3). Accordingly, the results are purely descriptive and are not intended to indicate statistically meaningful cross-country differences.

Figure 4.2 Distribution of Likert-Scale responses across survey items
Items including trust, empathy, personalisation and helpfulness received the highest scores for strong disagreement whereas chatbot speed received the highest score for strong agreement.

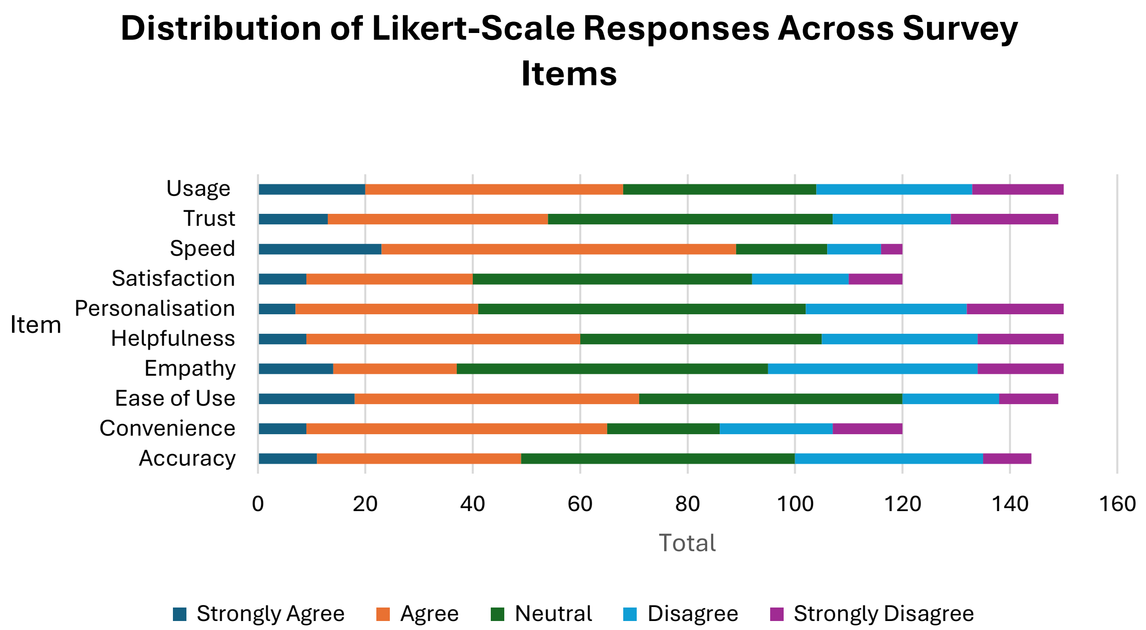


Figure 4.2 presents the distribution of Likert-scale responses across all survey items. Overall, responses clustered primarily around the neutral and agree categories, although notable differences emerged between functional and relational attributes.

Response speed exhibited the most positive distribution, with a high proportion of respondents indicating agreement or strong agreement and minimal disagreement, suggesting consistent perceptions of chatbot efficiency. In contrast, relational attributes, including empathy, personalisation, trust, and helpfulness, displayed greater dispersion across response categories. These constructs were characterised by higher neutral responses and relatively elevated levels of disengagement compared with functional attributes, indicating more mixed or uncertain evaluations.

Taken together, the distribution suggests that while respondents generally recognise the functional effectiveness of chatbots, perceptions of relational and affective service qualities are less settled and more variable across users.

4.2. Sentiment Analysis: Descriptive Summary

Secondary data were obtained from a publicly available dataset hosted on Hugging Face, comprising 19,271 customer reviews on 48 U.S. banks spanning the period 2017 to 2023. The dataset includes unstructured review text alongside associated metadata, enabling large-scale analysis of naturally occurring customer feedback related to digital banking services.

A small proportion of cases (0.47%) contained missing review text. As textual input is required for NLP-based modelling, these entries were removed, resulting 19,181 valid reviews for analysis. No duplicate entries or whitespace-only reviews were identified. Following standard text cleaning procedures, the mean review length was approximately 52 words (SD = 20), indicating that most reviews are relatively concise while still providing sufficient information to capture customer evaluations.

Figure 4.3 Distribution of Review Lengths
The distribution shows the typical review length and the approximate length of both short and long reviews within the dataset.

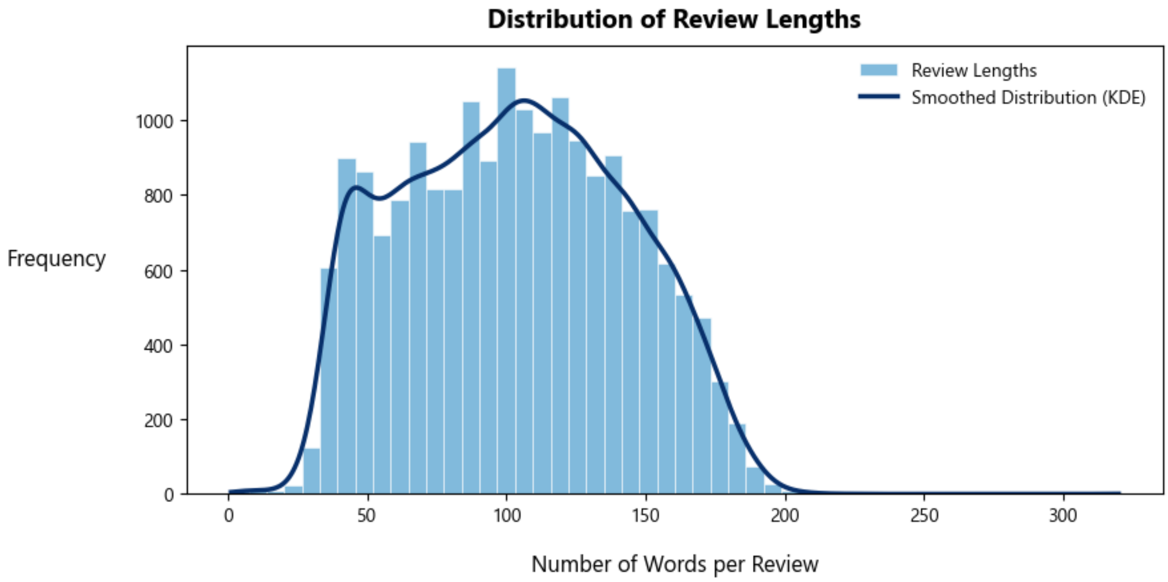


Figure 4.3 presents the distribution of review lengths across the dataset. The distribution was approximately symmetric (skewness = 0.08), with only a minor right tail, suggesting limited influence from extreme outliers and supporting the suitability of the dataset for supervised sentiment classification.

Text preprocessing was adapted to model requirements. For TF-IDF-based models, punctuation was removed to reduce feature sparsity and improve representation efficiency. In contrast, for transformer-based modelling (BERT), punctuation was retained, as tokenisation mechanisms incorporate punctuation as part of contextual meaning. All text was lowercased to improve token consistency and reduce vocabulary fragmentation in frequency-based approaches.

Exploratory analysis of model-generated sentiment labels indicated substantial class imbalance, with approximately 89% negative, 10% positive, and 2% neutral reviews. To mitigate biased learning, class imbalance was handled differently depending on model type. For logistic regression, imbalance was addressed using inverse-frequency weighting, which adjusts the loss contribution of minority classes without duplicating observations. For XGBoost, random oversampling was applied to increase representation of positive and neutral classes until class frequencies were balanced. For the BERT-based model, class weighting was applied during training to assign greater importance to underrepresented sentiment categories and reduce bias toward the dominant class. Across all models, data was partitioned into training and held-out evaluation sets, and model performance was assessed exclusively on unseen data.

4.3. Multiple Linear Regression Results

Correlation analysis (See Appendix A) revealed strong multicollinearity among several functional attributes. Specifically, Ease of Use, Accuracy, Helpfulness, Convenience, and Speed were highly correlated ($r > .75$), indicating that respondents perceived these items as a single underlying dimension of functional performance. Although helpfulness is conceptually associated with relational service support, its very high intercorrelation with other functional attributes suggests that respondents evaluated it primarily as an aspect of operational effectiveness rather than relational quality. To avoid redundancy and inflated standard errors, these variables were excluded from the regression model. Consequently, Satisfaction, Empathy, Personalisation, and Trust were retained as predictors due to their lower intercorrelations and conceptual distinctiveness.

The overall regression model was statistically significant ($F(4,25) = 5.54$, $p = .002$), indicating that the retained predictors jointly explained variation in customer engagement (see Appendix B). The model accounted for 47% of the variance in engagement (Adjusted $R^2 = .39$), suggesting moderate explanatory power.

As shown in Table 4, trust exhibited the largest positive coefficient ($\beta = .44$, $P = .053$), indicating a marginal positive association with engagement. Satisfaction ($\beta = .29$, $P = .131$) and personalisation ($\beta = .25$, $P = .209$) were positively related to engagement but did not reach statistical significance. Empathy displayed a negative but non-significant association with engagement ($\beta = -.34$, $P = .153$).

Table 4. Multiple Regression Predicting Engagement

Predictor	*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	1.135	0.551	2.061	0.050	0.001	2.269
Satisfaction	0.287	0.184	1.562	0.131	-0.092	0.666
Empathy	-0.339	0.220	-1.543	0.135	-0.793	0.114
Personalisation	0.252	0.195	1.290	0.209	-0.150	0.653
Trust	0.436	0.215	2.027	0.053	-0.007	0.878

4.4. Comparison of Sentiment Analysis Results Across Models

Model performance was evaluated using precision, recall, and F1-score for each sentiment class (Negative, Neutral, Positive), providing a detailed assessment of classification performance across sentiment categories. Performance results for Logistic Regression, XGBoost, and BERT are reported in Table 5.

Across all models, classification performance was highest for the negative sentiment class, reflecting its dominance in the dataset. In contrast, performance for the neutral class was substantially lower across models, indicating greater difficulty in distinguishing neutral sentiment from positive and negative classes.

Table 5. Classification Performance of Models by Sentiment Class

Class	*Metric*	*Logistic Regression*	*XGBoost*	*BERT*
Negative	*Precision*	0.96	0.97	0.98
	Recall	0.99	0.89	0.99
	F1-score	0.97	0.93	0.99
Neutral	*Precision*	0.28	0.06	0.28
	Recall	0.09	0.29	0.16
	F1-score	0.13	0.10	0.21
Positive	*Precision*	0.86	0.64	0.91
	Recall	0.74	0.74	0.89
	F1-score	0.79	0.69	0.90

To account for class imbalance, macro-averaged precision, recall, and F1-scores were computed, assigning equal weight to each sentiment category. As shown in Table 6, BERT achieved the strongest overall performance, based on macro F1, followed by logistic regression and XGBoost.

Table 6. Average Performance Across All Sentiment Classes

Model	*Macro Precision*	*Macro Recall*	*Macro F1*
Logistic Regression	0.70	0.60	0.63
XGBoost	0.56	0.64	0.57
BERT	0.73	0.68	0.70

Figure 4.4 Logistic Regression Confusion Matrix
The visualisation presents the confusion matrix for the logistic regression model, highlighting the model's classification performance across sentiment categories.

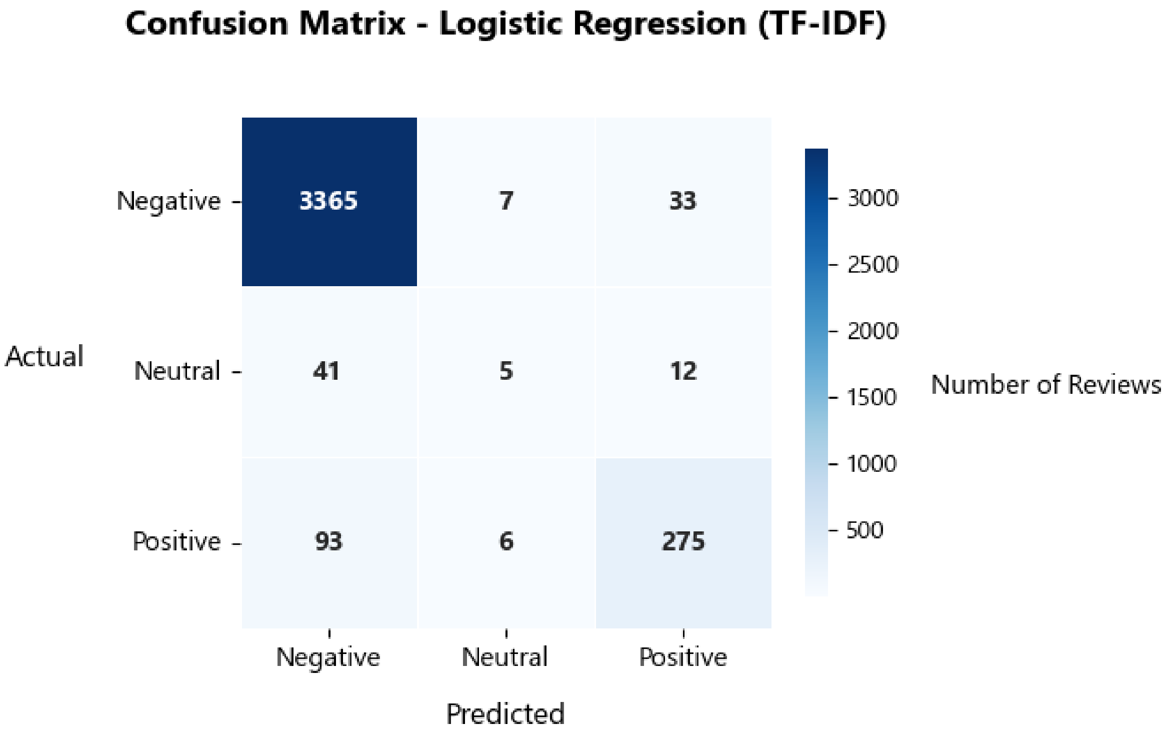


Figure 4.5 Logistic Regression Feature Importance
The plot displays the most influential features contributing to negative sentiment classification.

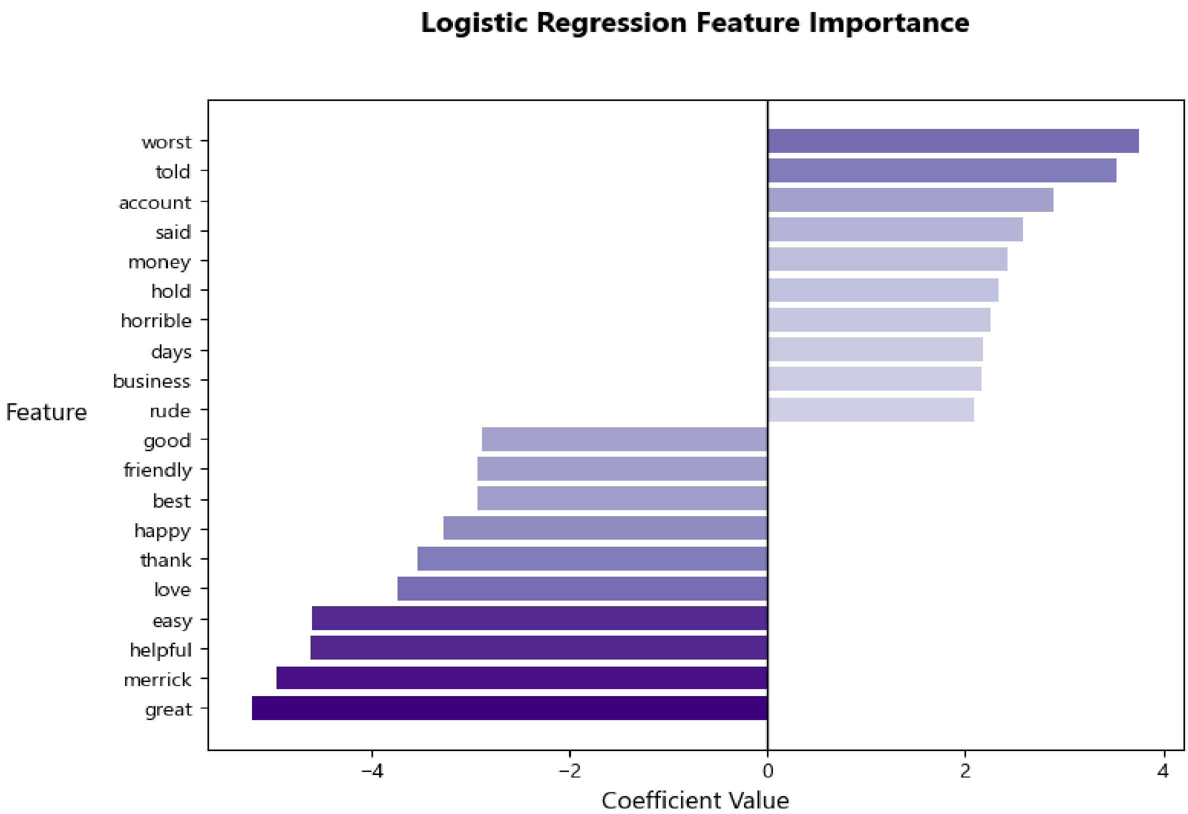
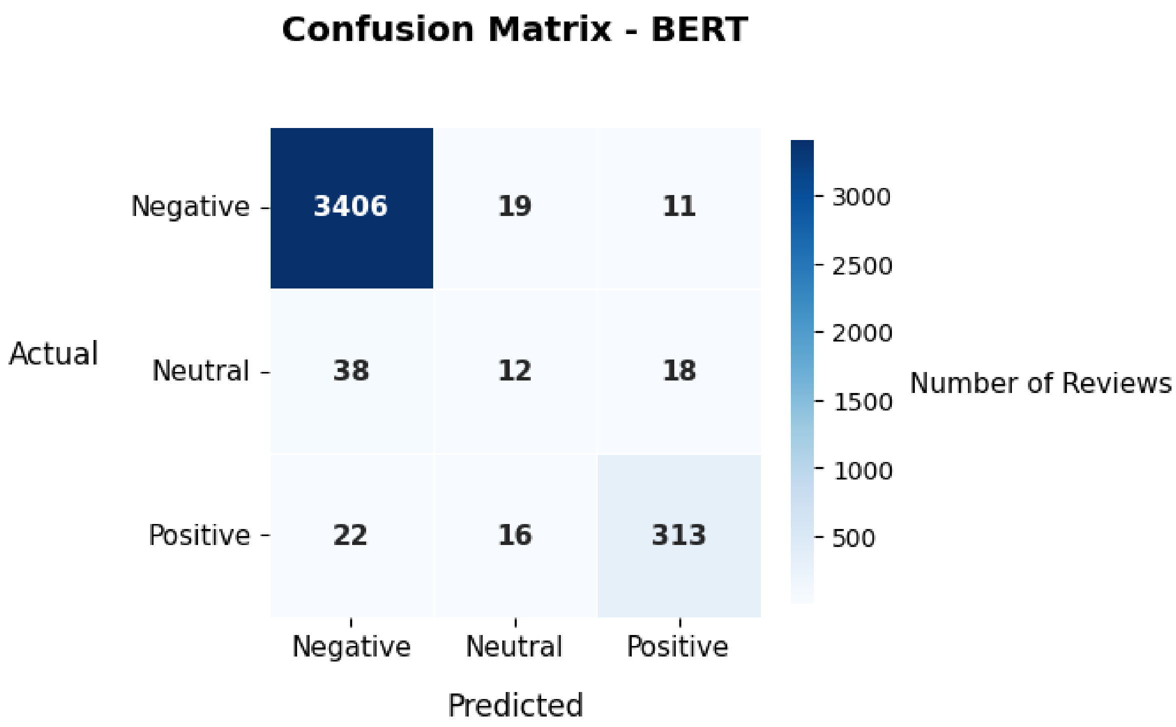


Figure 4.4 presents the confusion matrix for the logistic regression sentiment classification model using TF-IDF features. The model performed strongly in identifying negative sentiment, while performance for neutral sentiment was limited, with frequent misclassification into negative and positive classes. Positive sentiment was identified with moderate accuracy, although some overlap with negative sentiment remains. This pattern shows that while logistic regression is easy to interpret, its reliance on individual word features limits its ability to distinguish subtle differences in sentiment.

Moreover, coefficient magnitudes were examined, with particular attention given to the negative sentiment class due to its prevalence and relevance for identifying customer pain points. Words with larger positive coefficients were more strongly associated with negative sentiment (see Figure 4.5) indicating a higher likelihood that their presence contributed to negative sentiment classification.

Figure 4.6 Confusion Matrix for the BERT sentiment classifier
The figure displays the distribution of correct and incorrect predictions across Negative, Neutral and Positive sentiment classes.



The confusion matrix for the BERT model (Figure 4.6) illustrates classification performance across sentiment classes. The neutral class exhibited the highest rate of misclassification, likely due to semantic ambiguity and limited representation within the dataset. In contrast, negative reviews were classified with the highest accuracy, reflecting their dominance in the dataset and clearer linguistic signals.

XGBoost was evaluated as a benchmark model and achieved weaker average performance compared to logistic regression and BERT models. The corresponding confusion matrix is provided in Appendix C.

Although BERT demonstrated superior classification performance, its limited interpretability restricted feature-level analysis. Additionally, XGBoost was solely evaluated as a benchmark model and was not included in the detailed analysis due to its weaker average performance compared to the logistic regression and BERT models, as well as its limited interpretability in the context of the study. As a result, logistic regression was selected as the primary interpretability model due to its simplicity, transparency, and suitability for generating business-relevant insights.

4.5. Research Hypotheses

Based on the research aim and prior literature, nine hypotheses were proposed:

- H1: Higher perceived trust positively predicts customer engagement.
- H2-H4: Higher perceived ease of use, response speed, and response accuracy positively predict customer engagement.
- H5-H7: Higher perceived satisfaction, personalisation, and empathy positively predict customer engagement.
- H8-H9: Higher perceived helpfulness and convenience positively predict customer engagement.

Due to substantial multicollinearity among functional service attributes, only four hypotheses (H1, H5-H7) were tested in the regression model. The model was statistically significant and explained 47% of the variance in customer engagement.

Trust exhibited the largest positive coefficient and demonstrated a marginal association with engagement, providing indicative support for H1. Satisfaction, personalisation, and empathy did not show statistically significant relationship with engagement; therefore, H5-H7 were not supported.

Hypotheses relating to functional attributes (H2-H4, H8-H9) could not be tested inferentially due to multicollinearity. However, descriptive statistics indicated generally positive evaluations of these attributes, suggesting their relevance to chatbot interactions despite exclusion from the regression model.

4.6. Residuals and Assumption Diagnostics

Diagnostic tests indicated that the assumptions of multiple linear regression were not violated. Residual plots (Appendix D) showed random dispersion around zero, supporting the assumption of linearity. The variance of residuals remained relatively constant across predictor values, indicating homoscedasticity. No standardised residuals exceeded ±2, suggesting the absence of influential outliers. Overall, the diagnostic results support the validity of the regression results.

5. Discussion

5.1. Interpretation of Regression Findings

The regression analysis indicates that customers evaluate functional attributes such as speed, accuracy, ease of use as a unified dimension of performance rather than as independent drivers of engagement. Their exclusion from the regression model reflects this holistic evaluation rather than a lack of relevance.

The relative prominence of trust in the present findings suggests that trust represents a distinct evaluative dimension in customers' engagement with AI-powered chatbots. While functional service attributes were perceived holistically and excluded due to multicollinearity, trust remained conceptually and statistically distinct and exhibited the largest coefficient among the retained predictors. This pattern indicates that customer engagement in retail banking may be shaped not only by perceptions of operational performance but also by broader judgments of system reliability, security, and confidence. While the effect did not reach conventional levels of statistical significance, its relative magnitude indicates that trust explains engagement variance beyond that shared with functional service perceptions and the other retained predictors.

Satisfaction, personalisation and empathy demonstrated divergent relationships with customer engagement in the present study. While satisfaction and personalisation exhibited a positive but non-significant association with engagement, empathy displayed a negative, non-significant relationship. This asymmetric pattern suggests that not all relational attributes operate uniformly in AI-mediated service contexts, particularly within high-stakes and task-oriented financial service environments such as retail banking. As a result, these relational perceptions play a secondary role compared to trust in shaping engagement within retail banking chatbot interactions.

5.2. Interpretation of Sentiment Analysis Results

Despite the substantial class imbalance, the BERT model achieving a macro F1-score of 0.70 indicates that sentiment classification tasks achieved reasonable classification performance, particularly given the dominance of negative reviews. This result suggests that transfer learning using BERT may generalise more effectively than traditional machine learning approaches in imbalanced sentiment classification settings. However, for the purposes of this study, the logistic regression model was selected for further analysis due to its greater interpretability and suitability for exploratory and explanatory research.

The sentiment analysis findings indicate that negative customer evaluations were more frequently associated with functional issues, such as delayed responses, inaccurate information, and unresolved account-related problems, rather than explicit expressions of emotional dissatisfaction. This pattern indicates that customers primarily evaluate chatbot interactions in retail banking through a functional performance lens, with concerns related to system reliability and task completion appearing more salient than affective reactions.

These findings indicate that negative sentiment is primarily associated with functional service failures rather than resistance to chatbot technology. Accordingly, dissatisfaction observed in the data is more indicative of unmet functional expectations than resistance to AI-mediated service channels. Overall, the results highlight the importance of functional reliability in shaping customer sentiment, although conclusions should be interpreted cautiously considering the exploratory nature of the analysis.

5.3. Findings in Relation to the Literature

Overall, although the regression coefficients did not reach conventional levels of statistical significance, their direction and relative magnitude were broadly consistent with prior research on perception-driven engagement in AI-mediated service.

The prominence of trust observed in the present study aligns with prior literature positioning trust as a central antecedent of customer engagement in digital service contexts characterised by uncertainty and perceived risk (Van Doorn et al., 2010; Islam and Rahman, 2016, Yousafzai et al., 2003). In retail banking, where customers rely on automated systems to manage sensitive financial transactions, trust has consistently been identified as a critical condition for ongoing interaction with AI-enabled service interfaces (Rohit et al., 2025). Interpreted through the SOR framework, trust can be understood as an internal evaluative state through which customers process and interpret AI-mediated interactions, thereby shaping subsequent engagement behaviours (Alamoudi et al., 2025; Rohit et al., 2025). This conceptualisation is consistent with SOR-based studies in retail banking and FinTech that model trust as an organism-level variable, which mediates the relationship between system-related stimuli and user outcomes.

From a S-D logic perspective, trust is also fundamental for value co-creation in AI-enabled banking services. Concerns related to security, privacy, and system reliability may constrain customers' willingness to engage with digital service exchanges and share information through AI interfaces (Lusch and Nambisan, 2015; Laukkanen, 2016; Kristensson, 2019). Recent systematic reviews of AI across the customer journey further reinforce the salience of trust beyond initial adoption, highlighting its role in shaping post-adoption engagement and continued reliance on AI-powered service (Gouveia and Santos, 2025). The present findings contribute to this stream of research by empirically supporting the prominence of trust in engagement-focused evaluations of AI-powered chatbots within risk-sensitive banking environments.

Satisfaction demonstrated a positive but relatively weak association with customer engagement, suggesting that satisfaction alone may be insufficient to explain sustained engagement with AI-powered chatbots. This pattern aligns with customer engagement literature recognising satisfaction as a customer-focused construct that may function both as an antecedent and a consequence of engagement rather than a stable, unidirectional driver (Van Doorn et al., 2010; Brodie et al., 2011). In the context of retail banking, satisfaction may reflect confirmation, whereas engagement appears more closely associated with perceptions of trust and reliability (Hari, Iyer and Sampat, 2021).

The weaker influence of personalisation on customer engagement observed in this study is consistent with digital banking research. Although personalisation is widely recognised as a driver of customer satisfaction (Arora et al., 2023; Ashrafuzzaman 2025), prior studies indicate that it does not always exert a straightforward positive influence on trust formation, particularly when privacy concerns or transparency issues are salient (Lappeman et al., 2022; Nivedha et al., 2025). Given that trust is frequently positioned as a key antecedent of customer engagement in digital service environments (Van Doorn et al., 2010; Brodie et al., 2011), the limited role of personalisation in fostering trust provides a plausible explanation for its non-significant association with engagement in this study. Thus, in high-stakes banking settings, customers may prioritise secure, reliable, and consistent service delivery over personalised interactions (Rohit et al., 2025).

Empathy exhibited a negative association with engagement, suggesting that perceived emotional responsiveness may not only be ineffective but potentially incongruent with customer expectations in banking interactions. Prior research comparing AI and human frontline service indicates that chatbots are evaluated more favourably in functional and competence-based tasks, whereas emotional and empathic interactions are better suited to human agents (Ruan and Mezei, 2022; Yang and Hu 2021; Markovitch et al., 2024). In high-stakes financial contexts, attempts by AI systems to display empathy may be perceived as inauthentic, distracting, or misaligned with customers' primary goals of accuracy, efficiency, and problem resolution.

Although cultural effects were not formally examined in this study, the findings can be contextualised within broader patterns observed in Western banking environments. Prior research suggests that customers in individualistic service contexts tend to prioritise functional efficiency, accuracy, and problem resolution over relational warmth or emotional expressiveness, particularly in high-stakes domains such as financial services (Picoto and Pinto, 2020; Vafaei-Zadeh et al., 2024). This interpretation is reinforced by the sentiment analysis findings based on U.S. banking customer reviews, where negative sentiment was predominantly associated with functional issues rather than emotional dissatisfaction. These observations are offered as contextual interpretation rather than empirical evidence of cultural differences.

5.4. Theoretical Implications

This study contributes to customer engagement and digital service literature by advancing understanding of how customers engage with AI-powered chatbots in high-risk service contexts such as retail banking. Drawing on the SOR framework, the findings suggest that engagement with banking chatbots is primarily shaped through cognitive and evaluative states, most notably trust, rather than through affective or relational cures such as empathy, satisfaction, personalisation. While functional attributes were not modelled inferentially, their strong intercorrelation and salience in descriptive and sentiment-based analyses indicate that customers evaluate chatbot performance holistically, with trust emerging as the key internal state linking service perceptions to behavioural engagement, particularly in digital banking (Mehrabian and Russell, 1974; Francos and Bruckstein, 2023).

This refines prior customer engagement research that has emphasised emotional connection, anthropomorphism, and relational warmth as central drivers of engagement in AI-mediated interactions (Brodie et al., 2011; Hollebeek et al., 2024). The present findings indicate that, in high-risk and task-oriented service environments, cognitive evaluations of system reliability, security, and competence take precedence over emotional evaluations when customers for engagement-related judgments. This supports a more context-sensitive view of customer engagement, suggesting that antecedents are not uniform across service domains but vary according to perceived risk and service criticality (Van Doorn et al., 2010).

The findings also contribute to CEB by clarifying the role of satisfaction. Although satisfaction has frequently been conceptualised as an antecedent of engagement, the weak association observed in this study aligns with theoretical perspectives that position satisfaction as both a consequence and reinforcing outcome of engagement rather than a primary driver (Van Doorn et al., 2010; Brodie et al., 2011). In trust-intensive contexts such as retail banking, engagement appears more closely associated with perceptions of reliability and security than with general contentment following service interactions (Hari, Iyer and Sampat, 2021).

From a S-D logic perspective, the results reinforce the importance of trust as a prerequisite for value co-creation in AI-enabled service ecosystems. Engagement with banking chatbots depends on customers' willingness to participate in digital service exchanges and share sensitive information, which may be constrained when trust in data security and system integrity is lacking (Lusch and Nambisan, 2015; Laukkanen, 2016; Kristensson, 2019). In this sense, trust functions as an enabling condition for relational value creation rather than as a by-product of interaction.

Finally, the findings highlight the context-dependent nature of AI service evaluations. The limited influence of empathy and personalisation may reflect both the task-oriented nature of retail banking and broader tendencies within Western service contexts to prioritise functional efficiency and problem resolution over relational warmth (Picoto and Pinto, 2010; Vafaei-Zadeh et al., 2024). Consistent with CEB, cultural orientation appears to shape how engagement is expressed rather than determine whether engagement occurs (Connell et al., 2022). Together, these insights advance theory by positioning customer engagement with AI-powered chatbots as the outcome of an interaction between service risk, technological characteristics, and cognitive trust formation.

5.5. Practical Implications

For practitioners, the findings suggest that chatbot success metrics should extend beyond deployment rates or interaction volume to include trust-related indicators such as error resolution success, escalation transparency, and perceived reliability, as these factors are more closely associated with sustained customer engagement in retail banking environments.

Another managerial implication is that privacy protection and data security should be treated as core service attributes, rather than as background technical requirements. Customers' willingness to engage with AI-powered chatbots is strongly shaped by perceptions of how securely personal and financial data are handled, stored, and governed. Banks should therefore design chatbot services that make security and privacy protections visible and intelligible to users. This includes clearly communicating what data is collected, how they are used, and when human oversight is involved, thereby supporting cognitive trust formation and reducing perceived vulnerability.

Transparency plays a critical role in this process. Explicit disclosure of chatbot capabilities and limitations such as clarifying the types of queries that can be handled autonomously and those that required human intervention, which may help align customer expectations with actual service delivery. Providing customers with perceived control mechanisms, such as the ability to limit data sharing, opt out of automated interactions, or seamlessly escalate sensitive issues to human agents, can further reinforce trust by signalling accountability and governance (Lappeman et al., 2022). From a service marketing perspective, such practices enhance perceived relationship quality by reducing uncertainty and reinforcing institutional responsibility.

While empathy and personalisation did not significantly predict engagement in this study, this does not imply that these attributes are irrelevant. Rather, the findings suggest these relational features are unlikely to compensate for deficiencies in functional performance or perceived security. Investments in highly personalised or emotionally expressive chatbot behaviours may therefore have limited impact on engagement unless foundational expectations regarding accuracy, reliability, and privacy are first satisfied. Banks should prioritise robust intent recognition, accurate information retrieval, and consistent service performance before layering more advanced relational features.

Insights from sentiment analysis further indicate that negative customer sentiment is predominantly associated with service failures such as delayed responses, unresolved issues, or inadequate escalation when deployed for routine, low-risk tasks, including balance inquiries or transactional support. More advanced AI capabilities such as context-aware routing, improved intent detection, and selective use of generative AI may enhance service continuity but should be implemented cautiously and framed as augmenting rather than replacing human judgment. This will help to maintain deterministic, transparent, and auditable chatbot responses that are likely to be more conducive to trust and ongoing engagement (Zhang, Liang and Wu, 2024).

Finally, banks operating across multiple markets should recognise that expectations surrounding chatbot interaction may vary by cultural and institutional context. While functional reliability appears paramount in Western banking environments, modest localisation such as culturally appropriate tone, explanation styles, or support pathways may enhance perceived relevance without undermining consistency. A modular chatbot architecture that preserves core security and performance standards while allowing limited contextual adaption may therefore offer a balanced approach to engagement across diverse customer segments (Naidoo and Chadha, 2025).

6. Future Research

While this study provides empirical insight into customer perceptions of AI-powered chatbots in retail banking, several avenues for future research remain, particularly in relation to methodological approaches, data sources, and research design.

First, future research should adopt cross-country comparative designs to formally examine cultural effects on chatbot engagement. Although descriptive findings in this study suggested higher ratings for empathy, trust, satisfaction, and personalisation among respondents outside the UK, small and uneven subgroup sizes prevented inferential testing. Larger, internationally balanced samples would enable multigroup analysis to examine cultural moderators such as individualism, uncertainty avoidance, and power distance. This would allow researchers to test whether relational attributes such as empathy and personalisation exert stronger effects in collectivist cultures, as suggested in prior literature.

Second, future research should adopt longitudinal or behavioural data approaches to examine how trust and engagement evolved over time. The present study adopts a cross-sectional design and therefore captures customer perceptions at a single point time. Longitudinal research could examine how trust develops, stabilises, or deteriorates following repeated interactions, service failures, or successful problem resolution. Such approaches are particularly relevant in banking contexts, where customer-AI interactions are continuous rather than episodic. Future studies could further examine trust recovery processes, including how errors, data concerns, or failed chatbot interactions influence engagement trajectories and long-term reliance on AI-enabled services.

Third, integrating neuromarketing methodologies may be beneficial for capturing trust formation processes that are difficult to observe using self-report measures alone. Neuromarketing techniques such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) enable the examination of neural responses associated with trust, risk perceptions, and decision-making ((Gouveia and Santos, 2025). These methods have been successfully applied to chatbot-mediated interactions (Yen and Chiang, 2021) and could provide valuable insight into how customers cognitively process AI-driven banking services, particularly in emerging contexts involving generative AI.

Fourth, although sentiment analysis of U.S. banking reviews provided valuable contextual support, future studies should incorporate chatbot-specific interaction data, such as conversational transcripts, post-interaction feedback, or customer support logs. Combining such chatbot-level interaction data with survey-based perception measures would strengthen methodological triangulation and enhance construct validity by jointly capturing behavioural engagement and underlying perceptual drivers.

Firth, future research should more explicitly examine causal relationships between AI service characteristics, trust formation, and customer engagement. While experimental designs have been widely applied and allow for causal inference, artificial laboratory-based experiments may influence participant behaviour (Hawthorne effect) and oversimplify complex service interactions (Gouveia and Santos, 2025). Thus, field-based experiments and quasi-experimental designs such as A/B testing of chatbot features, natural experiments following service updates, or staggered rollout designs could enable causal inference while preserving ecological validity (Yin et al., 2023). Such approaches would allow researchers to examine how changes in functional performance, transparency, or escalation mechanism causally influence trust and engagement in real-world banking environments.

Sixth, the recent hybrid human-AI service models should be examined in greater depth. Netnography offers a valuable methodology for capturing qualitative data on how customers experience chatbot-human handoffs in real-world banking contexts. Future research could use netnographic approaches to examine customer preferences for human versus AI support in hybrid service models, particularly relevant for exploring complex, sensitive, or high-stake service encounters. Given the sensitive nature of this research topic, netnography could be a vital and unobtrusive way to gather relevant naturalistic data from customers (Heinonen and Medberg, 2018). Combining these insights with other data sources such as survey's or adopting a longitudinal netnographic approach may provide deeper understanding for trust recovery, engagement retention, and the conditions under which seamless integration between AI and human agents become critical (Hentzen et al., 2022; Heinonen and Medberg, 2018).

Finally, future research could explore multi-criteria decision-making frameworks, such as the Fuzzy Analytic Hierarchy Process (Fuzzy AHP), to systematically evaluate the relative importance of engagement drivers across customer journey touchpoints. Fuzzy AHP accommodates uncertainty and subjective judgement, making it well suited for analysing complex constructs such as trust, satisfaction, and personalisation in AI-enabled financial services (Arora et al., 2023). This approach could be used to rank chatbot features and engagement drivers across different stages of the customer journey, offering structured insight into where AI-enabled service investments generate the greatest value.

7. Conclusion

This research examined customer perceptions of AI-powered chatbots in the retail banking industry and assessed whether these perceptions influence customer engagement. The primary objective was to identify which factors most strongly shape customer evaluations of banking chatbots and to understand how these factors relate to engagement outcomes, with the aim of informing both theory and managerial practice in AI-enabled financial services.

The present study makes three key contributions. First, it extends customer engagement research by demonstrating that in high-stakes financial service contexts, engagement with AI-powered chatbots is influenced primarily by trust rather than by satisfaction, empathy or personalisation. Second, it shows how combining survey-based regression with machine learning-based sentiment analysis can triangulate perceptual and behavioural evidence while addressing overlap among functional service quality measures. Third, it provides managerial insight by indicating reliability, transparency, and trust-building mechanisms are more important for sustaining engagement that conversational warmth or personalisation features.

The study adopted a multi-method quantitative design, combining survey-based perceptual data analysed using descriptive statistics and multiple regression with machine learning-based sentiment analysis of real-world customer reviews. This approach enabled methodological triangulation between structured self-reported perceptions and naturally occurring behavioural evidence, strengthening the robustness and practical relevance of the findings.

The descriptive results indicated potential differences in how chatbot attributes are evaluated across regions, with non-UK respondents reporting more favourable perceptions of relational attributes. Regression analysis showed that trust emerged as the most prominent predictor of customer engagement among the constructs examined. The findings underscore the central role of perceived security, reliability, and transparency in shaping customers' willingness to engage with AI-powered chatbots in risk-intensive financial contexts. Although functional service attributes such as speed, accuracy, and ease of use were excluded from regression analysis due to high multicollinearity, their consistently high descriptive scores suggest that functional performance is perceived as a baseline expectation rather than a differentiating factor. Within this context, trust appears to play a critical role in sustaining engagement once basic performance standards are assumed.

Conversely, satisfaction, personalisation, and empathy exhibited weaker and statistically non-significant relationships with customer engagement. While these attributes may enhance overall perceptions of chatbot interactions, they appear insufficient to independently sustain engagement within retail banking contexts. This finding aligns with prior research indicating that satisfaction often operates because of service experiences rather than as a primary behavioural determinant, particularly in digital environments characterised by high risk and task-oriented interactions.

Findings from sentiment analysis reinforced these conclusions. Negative customer sentiment was predominantly associated with functional service failures, including delayed responses, unresolved account issues, and inaccurate information, rather than with chatbot design or emotional shortcomings. As the sentiment data were derived from U.S. banking customers, these findings support the view that customers in Western banking contexts prioritise effective problem resolution and operational reliability over relational or emotional interaction cues. Importantly, these issues appear to reflect broader digital service delivery challenges rather than resistance to chatbot technology itself.

Taken together, these findings are consistent with established perspectives in customer engagement research and align with the SOR framework as well as technology acceptance theory. Service-related stimuli—particularly those associated with functional performance and trust—emerge as critical factors shaping customers' internal evaluations, which in turn influence engagement behaviours. Although cultural context may affect how relational attributes are perceived, the evidence suggests that in high-stakes environments such as retail banking, cognitive and institution-based trust—anchored in competence, security, and reliability—plays a more decisive role in sustaining engagement than affective or emotionally expressive service features.

Methodologically, this study advances business analytics and service research by illustrating the value of integrating survey-based inferential analysis with machine learning-driven sentiment analysis to examine customer engagement with AI-powered services. The findings also reveal an important modelling implication: perceptual measures of functional service quality often exhibit high multicollinearity because customers tend to evaluate these attributes holistically as an overall assessment of system performance. This observation underscores the need for alternative research designs—such as behavioural log data, field experiments, or longitudinal approaches—to isolate specific functional effects and strengthen causal inference in future studies.

From a practical perspective, the findings indicate that banks and AI service providers should prioritise trust-building mechanisms when implementing chatbots in retail banking. This involves ensuring accurate and reliable information delivery, transparent communication regarding data usage, robust security assurances, and seamless escalation to human agents when issues exceed chatbot capabilities. Insights from sentiment analysis further reveal that negative customer experiences are primarily driven by operational failures rather than interface design, underscoring the importance of continuous service monitoring, system integration, and effective human–AI handover processes. Although empathy and personalisation did not significantly predict engagement in this study, banks operating across diverse markets may benefit from culturally adaptive chatbot designs that allow limited relational flexibility while maintaining core functional consistency.

Several limitations should be acknowledged. The cross-sectional design captures customer perceptions at a single point in time, restricting insight into how trust and engagement evolve through repeated interactions or service recovery episodes. The relatively small survey sample also limits statistical power and generalisability. Furthermore, reliance on self-reported measures introduces potential perceptual bias, which future research could address through behavioural interaction data, longitudinal designs, or field-based experiments.

In conclusion, this study provides empirical evidence that while functional performance forms a necessary foundation for chatbot use in retail banking, trust is pivotal in sustaining customer engagement. By integrating perceptual survey data with sentiment-driven behavioural insights, the research offers a nuanced understanding of how customers evaluate AI-powered banking services and identifies trust as a critical lever for advancing theory and guiding effective digital banking strategies.

References

Abo ElHamd, E., Shamma, H., Saleh, M. and Elkhodary, E. (2021). Customer engagement value: process, limitations and future research. *Journal of Modelling in Management*. Available at: https://doi.org/10.1108/jm2-12-2020-0319.

Adamopoulou, E. and Moussiades, L. (2020) 'An overview of chatbot technology', *IFIP Advances in Information and Communication Technology*, 584(1), pp. 373–383. Available at: https://doi.org/10.1007/978-3-030-49186-4_31.

Alabed, A., Javornik, A. and Gregory-Smith, D. (2022) 'AI anthropomorphism and its effect on users' self-congruence and self–AI integration: A theoretical framework and research agenda', *Technological Forecasting and Social Change*, 182, p. 121786. Available at: https://doi.org/10.1016/j.techfore.2022.121786.

Ashrafuzzaman, M. *et al.* (2025) 'AI-powered personalization in digital banking: a review of customer behaviour analytics and engagement', *American Journal of Interdisciplinary Studies*, 6(1), pp. 40–71. Available at: https://doi.org/10.63125/z9s39s47.

Arora, A. *et al.* (2023) 'Customer experiences in the era of artificial intelligence (AI) in context to FinTech: a fuzzy AHP approach', *Benchmarking: An International Journal*, 30(10). Available at: https://doi.org/10.1108/bij-10-2021-0621.

Ahmed, S. and Aziz, N.A. (2024) 'Impact of AI on customer experience in video streaming services: a focus on personalization and trust', *International Journal of Human-Computer Interaction*, pp. 1–20. Available at: https://doi.org/10.1080/10447318.2024.2400395.

Bank of England (2024) *Artificial intelligence in UK financial services – 2024*. Available at: https://www.bankofengland.co.uk/report/2024/artificial-intelligence-in-uk-financial-services-2024 (Accessed: 27 December 2025).

Barari, M. *et al.* (2020) 'A meta-analysis of customer engagement behaviour', *International Journal of Consumer Studies*, 45(4).

Berger, J. *et al.* (2020) 'Uniting the tribes: using text for marketing insight', *Journal of Marketing*, 84(1), pp. 1–25. Available at: https://doi.org/10.1177/0022242919873106.

Bhattacherjee, A. (2001) ‘Understanding information systems continuance: an expectation-confirmation model’, *MIS Quarterly*, 25(3), pp. 351–370. Available at: https://doi.org/10.2307/3250921.

Brodie, R.J. *et al.* (2011) ‘Customer engagement: conceptual domain, fundamental propositions, and implications for research’, *Journal of Service Research*, 14(3), pp. 252–271. Available at: https://doi.org/10.1177/1094670511411703.

Brodie, R.J. *et al.* (2013) ‘Consumer engagement in a virtual brand community: an exploratory analysis’, *Journal of Business Research*, 66(1), pp. 105–114.

Connell, C., Marciniak, R. and Carey, L.D. (2022) ‘The effect of cross-cultural dimensions on the manifestation of customer engagement behaviours’, *Journal of International Marketing*, 31(1), p. 1069031X2211306.

Chaturvedi, R. *et al.* (2023) ‘Social companionship with artificial intelligence: recent trends and future avenues’, *Technological Forecasting and Social Change*, 193, pp. 122634–122634. Available at: https://doi.org/10.1016/j.techfore.2023.122634.

Chen, J.-S., Le, T.-T.-Y. and Florence, D. (2021) ‘Usability and responsiveness of artificial intelligence chatbot on online customer experience in e-retailing’, *International Journal of Retail & Distribution Management*, 49(11). Available at: https://doi.org/10.1108/ijrdm-08-2020-0312.

Davis, F.D. (1989) ‘Perceived usefulness, perceived ease of use, and user acceptance of information technology’, *MIS Quarterly*, 13(3), pp. 319–340. Available at: https://doi.org/10.2307/249008.

Davis, F.D. and Venkatesh, V. (1996) ‘A critical assessment of potential measurement biases in the technology acceptance model: three experiments’, *Information Systems Research,* 7(3), pp. 219–240.

Dhanya, C. and Ramya, K. (2023) ‘Customer perception of chatbots in banking: a study on technology acceptance’, *IUP Journal of Management Research*, 22(3), pp. 22–39.

Dhanya, C. and Ramya, K. (2025) ‘Unlocking banking chatbot adoption: a unified approach through extended TAM and UTAUT model’, *Proquest.com*, 16(1). Available at:

https://doi.org/10.1831l/sdmimd/2025/48908.

Dhiman, N., Jamwal, M. and Kumar, A. (2023) 'Enhancing value in customer journey by considering the (ad)option of artificial intelligence tools', *Journal of Business Research*, 167, p. 114142. Available at: https://doi.org/10.1016/j.jbusres.2023.114142.

Dwivedi, Y.K. *et al.* (2021) 'Artificial Intelligence (AI): multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy', *International Journal of Information Management*, 57(101994), p. 101994. Available at: https://doi.org/10.1016/j.ijinfomgt.2019.08.002.

Francos, R.M. and Bruckstein, A.M. (2023) 'On the role and opportunities in teamwork design for advanced multi-robot search systems', *Frontiers in Robotics and AI*, 10. Available at: https://doi.org/10.3389/frobt.2023.1089062.

Castillo, D., Canhoto, A.I. and Said, E. (2020) 'The dark side of AI-powered service interactions: exploring the process of co-destruction from the customer perspective', *The Service Industries Journal*, 41(13-14), pp. 1–26. Available at: https://doi.org/10.1080/02642069.2020.1787993.

Chi, N.T.K. and Hoang Vu, N. (2022) 'Investigating the customer trust in artificial intelligence: the role of anthropomorphism, empathy response, and interaction', *CAAI Transactions on Intelligence Technology*, 8(1). Available at: https://doi.org/10.1049/cit2.12133.

Islam, J.U. and Rahman, Z. (2016). The transpiring journey of customer engagement research in marketing. *Management Decision*, 54(8), pp.2008–2034. Available at: https://doi.org/10.1108/md-01-2016-0028.

Kumar, V. and Pansari, A. (2016) 'Competitive advantage through engagement', *Journal of Marketing Research*, 53(4), pp. 497–514. Available at: https://doi.org/10.1509/jmr.15.0044.

Ordanini, A., Miceli, L., Pizzetti, M., & Parasuraman, A. (2011). Crowd-funding: transforming customers into investors through innovative service platforms. *Journal of Service Management*, *22*(4), 443–470. https://doi.org/10.1108/09564231111155079

Oxford English Dictionary (no date) *Chatbot*. Oxford University Press. Available at: https://doi.org/10.1093/OED/1140056711 (Accessed: 27 December 2025).

Fadnes, L.T., Taube, A. and Tylleskär, T. (2009) 'How to identify information bias due to self-reporting in epidemiological research', *The Internet Journal of Epidemiology*, 7(2). Available at: https://doi.org/10.5580/1818.

Gallagher, C., Furey, E. and Curran, K. (2019) 'The application of sentiment analysis and text analytics to customer experience reviews to understand what customers are really saying', *International Journal of Data Warehousing and Mining*, 15(4), pp. 21–47. Available at: https://doi.org/10.4018/ijdwm.2019100102.

Glassberg, I., Ilan, Y.B. and Zwilling, M. (2025) 'The key role of design and transparency in enhancing trust in AI-powered digital agents', *Journal of Innovation & Knowledge*, 10(5), pp. 100770–100770. Available at: https://doi.org/10.1016/j.jik.2025.100770.

Guo, Y. (2022) 'Digital trust and the reconstruction of trust in the digital society: an integrated model based on trust theory and expectation confirmation theory', *Digital Government: Research and Practice*, 3(4). Available at: https://doi.org/10.1145/3543860.

Gouveia, J. & Santos, S. (2025) Rethinking the customer journey: impact of AI for consumers and businesses—a systematic literature review and research agenda. *International journal of consumer studies*, 49 (6).

Gur, T. and Yossi Maaravi (2025) 'The algorithm of friendship: literature review and integrative model of relationships between humans and artificial intelligence (AI)', *Behaviour and Information Technology*, pp. 1–21. Available at: https://doi.org/10.1080/0144929x.2025.2502467.

Harrison, M.D., Choudhury, V. and Kacmar, C. (2002) 'Developing and validating trust measures for ecommerce: an integrative typology', *Information Systems Research*, 13(3), pp. 334–359. Available at: https://doi.org/10.2307/23015741.

Hari, H., Iyer, R. and Sampat, B. (2021) 'Customer brand engagement through chatbots on bank websites – examining the antecedents and consequences', *International Journal of Human–Computer Interaction*, 38(13), pp. 1–16. Available at: https://doi.org/10.1080/10447318.2021.1988487.

Hentzen, J.K. *et al.* (2021) ‘Artificial intelligence in customer-facing financial services: a systematic literature review and agenda for future research’, *International Journal of Bank Marketing*, 40(6). Available at: https://doi.org/10.1108/ijbm-09-2021-0417.

Hollebeek, L.D., Menidjel, C., Sarstedt, M., Jansson, J. and Urbonavicius, S. (2024). Engaging consumers through artificially intelligent technologies: Systematic review, conceptual model, and further research. *Psychology & Marketing*, 41(4), pp.880–898. Available at: https://doi.org/10.1002/mar.21957.

Gupta Y, Khan FM (2024) ‘Role of artificial intelligence in customer engagement: a systematic review and future research directions. *Journal of Modelling in Management*, Vol. 19 No. 5 pp. 1535–1565. Available at: https://doi.org/10.1108/JM2-01-2023-0016

Heinonen, K. and Medberg, G. (2018) ‘Netnography as a tool for understanding customers: implications for service research and practice’, *Journal of Services Marketing*, 32(6), pp. 657–679. Available at: https://doi.org/10.1108/jsm-08-2017-0294.

Julija Saveljeva and Volkova, T. (2025) ‘A survey on digital trust: towards a validated definition’, *Digital*, 5(2), pp. 14–14. Available at: https://doi.org/10.3390/digital5020014

Islam, J.U. and Rahman, Z. (2016) ‘The transpiring journey of customer engagement research in marketing’, *Management Decision*, 54(8), pp. 2008–2034. Available at: https://doi.org/10.1108/md-01-2016-0028.

Iryna Pentina *et al.* (2023) ‘Consumer–machine relationships in the age of artificial intelligence: Systematic literature review and research directions’, *Psychology & Marketing*, 40(8). Available at: https://doi.org/10.1002/mar.21853.

Laukkanen, T. (2016) ‘Consumer adoption versus rejection decisions in seemingly similar service innovations: The case of the Internet and mobile banking’, *Journal of Business Research*, 69(7), pp. 2432–2439. Available at: https://doi.org/10.1016/j.jbusres.2016.01.013.

Li, C.-Y., Fang, Y.-H. and Chiang, Y.-H. (2023) ‘Can AI chatbots help retain customers? An integrative perspective using affordance theory and service-domain logic’, *Technological Forecasting and Social Change*, 197, p. 122921. Available at: https://doi.org/10.1016/j.techfore.2023.122921.

Pentina, I. *et al.* (2023) ‘Consumer–machine relationships in the age of artificial intelligence: Systematic literature review and research directions’, *Psychology & Marketing*, 40(8). Available at: https://doi.org/10.1002/mar.21853.

Pramanik, P. and Jana, R.K. (2025). A consumer acceptance model in the artificial intelligence era. *Management Decision*. Available at: https://doi.org/10.1108/md-03-2024-0574.

Kaakandikar, R. *et al.* (2025) ‘Study of customer perception of AI-driven chatbots in banking services’, *Advances in Consumer Research*, 2(2), pp. 693–703.

Kumar, A. *et al.* (2025) ‘Anthropomorphic generative AI chatbots for enhancing customer engagement, experience and recommendation’, *Journal of Consumer Marketing*. Available at: https://doi.org/10.1108/jcm-06-2024-6922

Kristensson, P. (2019) ‘Future service technologies and value creation’, *Journal of Services Marketing*. Available at: https://doi.org/10.1108/jsm-01-2019-0031.

Lappeman, J. *et al.* (2022) ‘Trust and digital privacy: willingness to disclose personal information to banking chatbot services’, *Journal of Financial Services Marketing*, 28. Available at: https://doi.org/10.1057/s41264-022-00154-z.

Lusch, R.F. and Nambisan, S. (2015) ‘Service innovation: a service-dominant logic perspective’, *MIS Quarterly*, 39(1), pp. 155–176. Available at: https://doi.org/10.2307/26628345.

Maslowska, E., Malthouse, E.C. and Collinger, T. (2016) ‘The customer engagement ecosystem’, *Journal of Marketing Management*, 32(5-6), pp. 469–501. Available at: https://doi.org/10.1080/0267257x.2015.1134628

Manser Payne, E.H., Dahl, A.J. and Peltier, J. (2021) ‘Digital servitization value co-creation framework for AI services: a research agenda for digital transformation in financial service ecosystems’, *Journal of Research in Interactive Marketing*, 15(2). Available at: https://doi.org/10.1108/jrim-12-2020-0252.

Manser Payne, E.H., Peltier, J. and Barger, V.A. (2021) ‘Enhancing the value co-creation process: artificial intelligence and mobile banking service platforms’, *Journal of Research* *in Interactive Marketing*, 15(1), pp. 68–85. Available at: https://doi.org/10.1108/jrim-10-2020-0214.

Markovitch, D.G., Stough, R.A. and Huang, D. (2024) ‘Consumer reactions to chatbot versus human service: an investigation in the role of outcome valence and perceived empathy’, *Journal of retailing and consumer services*, 79, pp. 103847–103847. Available at: https://doi.org/10.1016/j.jretconser.2024.103847.

Mogaji, E. *et al.* (2021) ‘Emerging-market consumers’ interactions with banking chatbots’, *Telematics and Informatics*, 65(0736-5853), p. 101711. Available at: https://doi.org/10.1016/j.tele.2021.101711.

McKinsey & Company (2020) *Reimagining customer engagement for the AI bank of the future*. Available at: https://www.mckinsey.com (Accessed: 27 December 2025).

Mehrabian, A. and Russell, J.A., 1974. *An approach to environmental psychology*. the MIT Press.

Moravec, V. *et al.* (2025) ‘Human-machine in the vortex of digital synergy’, *Humanities and Social Sciences Communications*, 12(1). Available at: https://doi.org/10.1057/s41599-025-05014-4.

Torkzdeh, S. (2021) 'Customer engagement and participation behavior: the underlying mechanisms of the virtual servicescape and service outcomes,' *e-Service Journal*, 13(2). Available at: https://doi.org/10.2979/eservicej.13.2.01.

Mostafa, R.B. (2025) ‘AI and value co-creation in the banking sector: a bibliometric analysis and a systematic literature review’, *International Journal of Bank Marketing*, 43(6). Available at: https://doi.org/10.1108/ijbm-12-2024-0753.

Munira, M.S.K. (2025) ‘Digital transformation in banking: a systematic review of trends, technologies, and challenges’, *Strategic Data Management and Innovation*, 2(01), pp. 78–95. Available at: https://doi.org/10.71292/sdmi.v2i01.12.

Anandarajan, M. *et al.* (2019) *Practical text analytics: maximizing the value of text data*. Cham: Springer.

Naidoo, V. and Chadha, K.K. (2025) ‘Culturally responsive AI chatbots: from framework to field evidence', *Computers in Human Behaviour: Artificial Humans*, p. 100224. Available at: https://doi.org/10.1016/j.chbah.2025.100224.

Nam, H. and Kannan, P.K. (2020) 'Digital environment in global markets: cross-cultural implications for evolving customer journeys', *Journal of International Marketing*, 28(1), p. 1069031X1989876. Available at: https://doi.org/10.1177/1069031X19898767.

Ng, S.C., Sweeney, J.C. and Plewa, C. (2020) 'Customer engagement: a systematic review and future research priorities,' *Australasian Marketing Journal (AMJ)*, 28(4), pp. 235–252. Available at: https://doi.org/10.1016/j.ausmj.2020.05.004.

Nguyen, T.H. and Le, X.C. (2024) 'Artificial intelligence-based chatbots – a motivation underlying sustainable development in banking: standpoint of customer experience and behavioral outcomes', *Cogent Business & Management*, 12(1). Available at: https://doi.org/10.1080/23311975.2024.2443570.

Nicolescu, L. and Tudorache, M.T. (2022) 'Human-computer interaction in customer service: the experience with AI chatbots—a systematic literature review', *Electronics*, 11(10), p. 1579. Available at: https://doi.org/10.3390/electronics11101579.

Nivedha V. *et al.* (2025) 'AI in FinTech: redefining customer trust and personalization in digital finance', *Advances in Consumer Research*, 2(5), pp. 153–160. Available at: https://research.ebsco.com/linkprocessor/plink?id=375cc49f-0ddc-35c9-aced-0b87367dd1b7.

Oliver, R.L. and DeSarbo, W.S. (1988) 'Response determinants in satisfaction judgments', *Journal of Consumer Research*, 14(4), pp. 495–507. Available at: https://doi.org/10.2307/2489156.

Oyeniyi, L.D., Ugochukwu, C.E. and Mhlongo, N.Z. (2024) 'Implementing AI in banking customer service: A review of current trends and future applications', *International Journal of Science and Research Archive*, 11(2), pp. 1492–1509. Available at: https://doi.org/10.30574/ijsra.2024.11.2.0639.

Pansari, A. and Kumar, V. (2017) 'Customer engagement: the construct, antecedents, and consequences', *Journal of the Academy of Marketing Science*, 45(3), pp. 294–311. Available at: https://link.springer.com/article/10.1007/s11747-016-0485-6.

Picoto, W.N. and Pinto, I. (2020) ‘Cultural impact on mobile banking use – a multi-method approach’, *Journal of Business Research*, 124, pp. 620–628. Available at: https://doi.org/10.1016/j.jbusres.2020.10.024.

Puertas, S.M. *et al.* (2024) ‘Purchase intentions in a chatbot environment: an examination of the effects of customer experience’, *Oeconomia Copernicana*, 15(1), pp. 145–194. Available at: https://doi.org/10.24136/oc.2914.

Riley, B.K. and Dixon, A. (2024) ‘Emotional and cognitive trust in artificial intelligence: a framework for identifying research opportunities’, *Current Opinion in Psychology*, 58, pp. 101833–101833. Available at: https://doi.org/10.1016/j.copsyc.2024.101833.

Rohit, K. *et al.* (2025) ‘Smart banking chatbots and consumer engagement: the role of trust and privacy in AI-driven banking’, *Journal of Strategic Marketing*, pp. 1–18. Available at: https://doi.org/10.1080/0965254x.2025.2481140.

Rosado-Pinto, F. and Loureiro, S.M.C. (2020). The growing complexity of customer engagement: a systematic review. *EuroMed Journal of Business*, 15(2), pp.167–203. Available at: https://doi.org/10.1108/emjb-10-2019-0126.

Ruan, Y. and Mezei, J. (2022) ‘When do AI chatbots lead to higher customer satisfaction than human frontline employees in online shopping assistance? considering product attribute type’, *Journal of Retailing and Consumer Services*, 68(103059), p. 103059. Available at: https://doi.org/10.1016/j.jretconser.2022.103059.

Sebetci Özel (2025). Unpacking the drivers of AI technology acceptance. *Journal of Computer Information Systems*, pp.1–12. Available at: https://doi.org/10.1080/08874417.2025.2564439.

Statista (no date) *Bank customer retention in the UK.* Available at: https://www.statista.com/study/27479/bank-customer-retention-in-the-uk-statista-dossier (Accessed: 8 December 2025).

Storbacka, K. *et al.* (2016) ‘Actor engagement as a microfoundation for value co-creation’, *Journal of Business Research*, 69(8), pp. 3008–3017. Available at: https://doi.org/10.1016/j.jbusres.2016.02.034.

Suryadi, D. and Padlan, K.O. (2024) 'Leveraging topic modeling and sentiment analysis to improve digital bank applications', *International Conference on Data Analytics for Business and Industry (ICDABI)*, pp. 86–90. Available at: https://doi.org/10.1109/icdabi63787.2024.10799999.

Su, S., Tsai, C. and Hsu, W. (2013) 'Extending the TAM model to explore the factors affecting intention to use telecare systems', *Journal of Computers*, 8(2). Available at: https://doi.org/10.4304/jcp.8.2.525-532.

Thasleena K, F. and Santhi, P. (2025) 'Artificial intelligence driven e-services of new generation banks: customer perception, attitude and response analysis', *International Research Journal of Multidisciplinary Scope*, 06(03), pp. 876–888. Available at: https://doi.org/10.47857/irjms.2025.v06i03.04268.

Trunfio, M. and Rossi, S. (2021) 'Conceptualising and measuring social media engagement: a systematic literature review', *Italian Journal of Marketing*, 2021(3), pp. 267–292. Available at: https://doi.org/10.1007/s43039-021-00035-8.

Upadhyay, N. and Kamble, A. (2023) 'Why can't we help but love mobile banking chatbots? perspective of stimulus-organism-response', *Journal of Financial Services Marketing*, 29(3), pp. 855–872. Available at: https://doi.org/10.1057/s41264-023-00237-5.

Vafaei-Zadeh, A. *et al.* (2024) 'Investigating factors influencing AI customer service adoption: an integrated model of stimulus–organism–response (SOR) and task-technology fit (TTF) theory', *Asia Pacific Journal of Marketing and Logistics*. Available at: https://doi.org/10.1108/apjml-05-2024-0570.

Van Doorn, J. *et al.* (2010) 'Customer engagement behavior: theoretical foundations and research directions', *Journal of Service Research*, 13(3), pp. 253–266. Available at: https://doi.org/10.1177/1094670510375599.

Vargo, S.L. and Lusch, R.F. (2016) 'Institutions and axioms: An extension and update of service-dominant logic', *Journal of the Academy of Marketing Science*, 44(1), pp. 5–23. Available at: https://doi.org/10.1007/s11747-015-0456-3.

Venkatesh, V. and Bala, H. (2008) 'Technology acceptance model 3 and a research agenda on interventions', *Decision Sciences*, 39(2), pp. 273–315.

Venkatesh, V. *et al.* (2003) ‘User acceptance of information technology: toward a unified view’, *MIS Quarterly*, 27(3), pp. 425–478. Available at: https://doi.org/10.2307/30036540.

Wankhade, M., Rao, A.C.S. and Kulkarni, C. (2022) ‘A survey on sentiment analysis methods, applications, and challenges’, *Artificial Intelligence Review*, 55(55). Available at: https://doi.org/10.1007/s10462-022-10144-1.

Williams, M.D., Rana, N.P. and Dwivedi, Y.K. (2015). The unified theory of acceptance and use of technology (UTAUT): A literature review. *Journal of Enterprise Information Management*, 28(3), pp.443–488. Available at: https://doi.org/10.1108/jeim-09-2014-0088.

Wirtz, J. and Bateson, J.E.G. (1999) ‘Consumer satisfaction with services’, *Journal of Business Research*, 44(1), pp. 55–66. Available at: https://doi.org/10.1016/s0148-2963(97)00178-1.

Yang, C. and Hu, J. (2021) ‘When do consumers prefer AI-enabled customer service? the interaction effect of brand personality and service provision type on brand attitudes and purchase intentions’, *Journal of Brand Management*, 29(2). Available at: https://doi.org/10.1057/s41262-021-00261-7.

Yousafzai, S.Y., Pallister, J.G. and Foxall, G.R. (2003) 'A proposed model of e-trust for electronic banking,' *Technovation*, 23(11), pp. 847–860. Available at: https://doi.org/10.1016/s0166-4972(03)00130-5.

Yen, C. and Chiang, M.-C. (2020) ‘Trust me, if you can: a study on the factors that influence consumers’ purchase intention triggered by chatbots based on brain image evidence and self-reported assessments’, *Behaviour & Information Technology*, 40(11), pp. 1–18. Available at: https://doi.org/10.1080/0144929x.2020.1743362.

Yin, D., Li, M. and Qiu, H. (2023) ‘Do customers exhibit engagement behaviors in AI environments? The role of psychological benefits and technology readiness’, *Tourism Management*, 97, p. 104745. Available at: https://doi.org/10.1016/j.tourman.2023.104745.

Zainol, S. *et al.* (2023) ‘Understanding customer satisfaction of chatbots service and system quality in banking services’, *Journal of Information Technology Management*, 15(Special Issue), pp. 142–152. Available at: https://doi.org/10.22059/jitm.2022.89417.

Zhang, R.W., Liang, X. and Wu, S.-H. (2024) 'When chatbots fail: exploring user coping following a chatbots-induced service failure', *Information technology & people*, 37(8), pp. 175–195. Available at: https://doi.org/10.1108/itp-08-2023-0745.

Alamoudi, H. *et al.* (2025) 'Exploring trust and outcome expectancy in FinTech digital payments: insights from the stimulus-organism-response model', *International Journal of Bank Marketing*. Available at: https://doi.org/10.1108/ijbm-04-2024-0252.

Appendices

Appendix A: Correlation Matrix

Table A1. Correlation Matrix

	1	2	3	4	5	6	7	8	9	10
1. Engagement	1.00									
2. Satisfaction	0.60	1.00								
3. Trust	0.57	0.66	1.00							
4. Ease of Use	0.59	0.79	0.80	1.00						
5. Helpfulness	0.53	0.73	0.62	0.83	1.00					
6. Convenience	0.45	0.57	0.49	0.68	0.81	1.00				
7. Speed	0.50	0.69	0.59	0.76	0.69	0.72	1.00			
8. Accuracy	0.43	0.73	0.59	0.81	0.84	0.66	0.76	1.00		
9. Empathy	0.16	0.29	0.52	0.52	0.46	0.35	0.30	0.47	1.00	
10. Personalisation	0.31	0.34	0.36	0.48	0.65	0.67	0.40	0.43	0.63	1.00

Appendix B: Multiple Regression Output

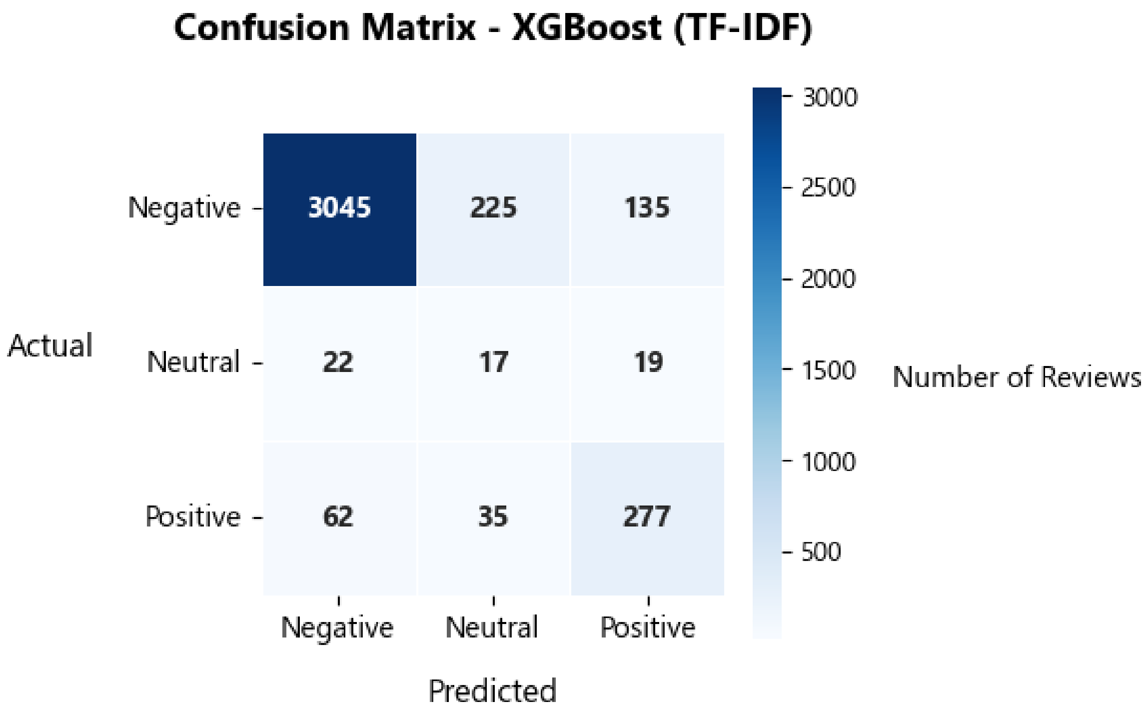
Table B1. Regression Statistics

Multiple R	0.686
R Square	0.470
Adjusted R Square	0.385
Standard Error	0.554
Observations	30

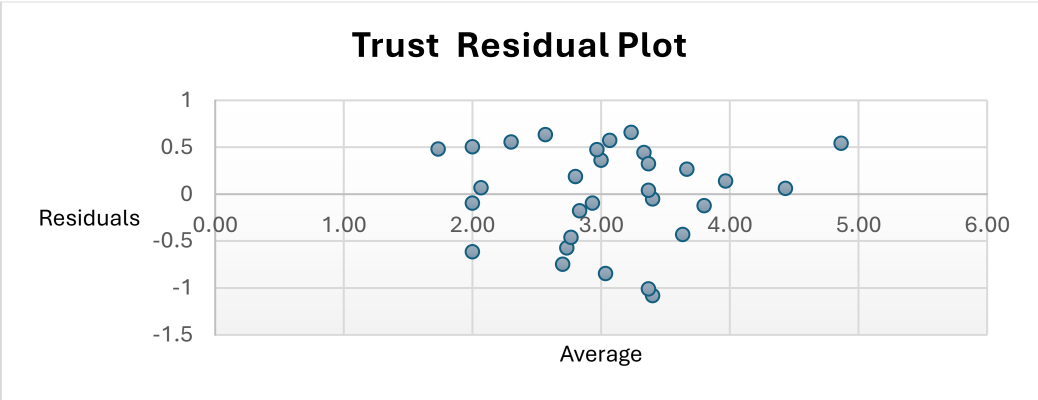
Table B2. ANOVA

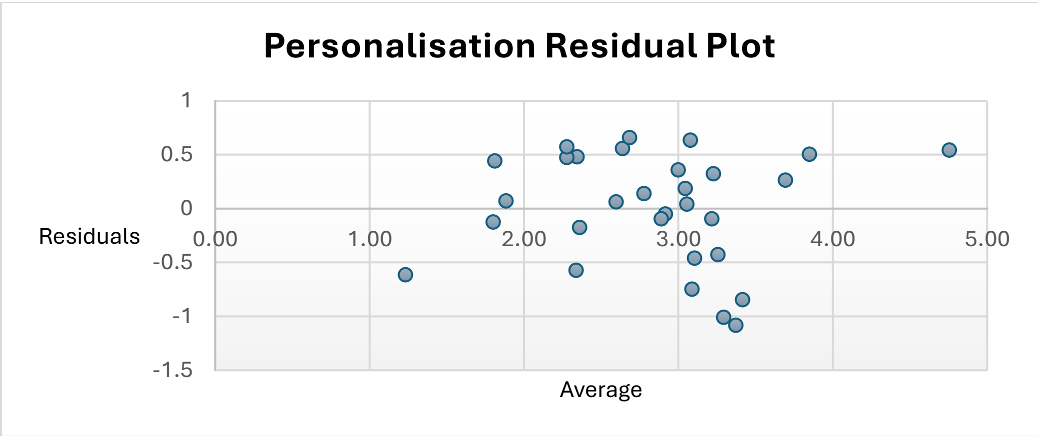
	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	6.809	1.702	5.542	0.002
Residual	25	7.678	0.307	–	–
Total	29	14.487	–	–	–

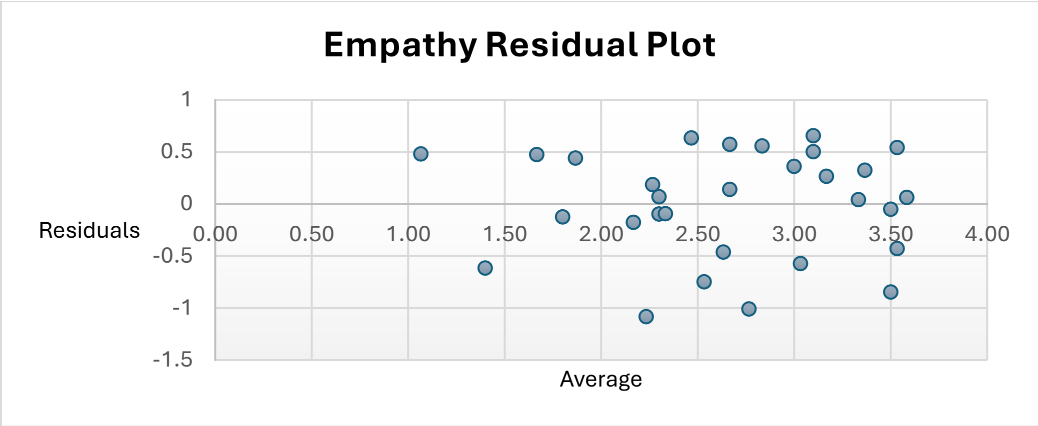
Appendix C: Confusion Matrix for XGBoost Model

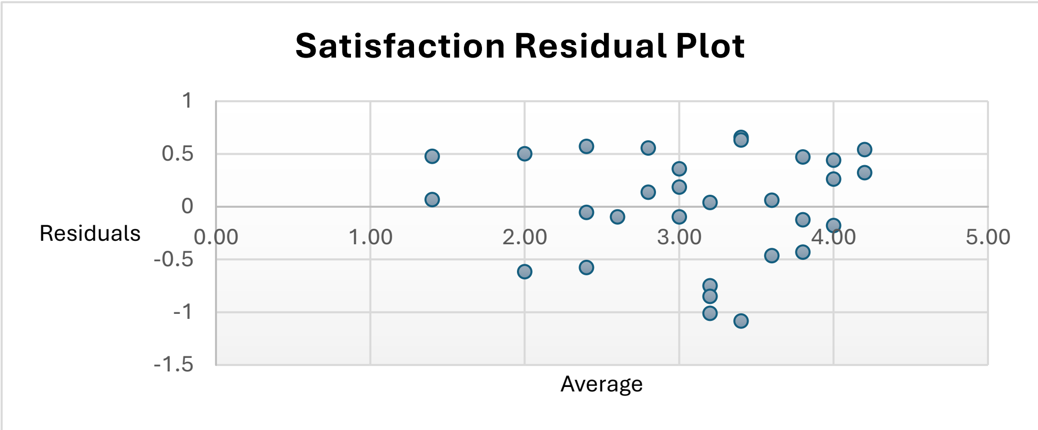


Appendix D: Residual and Assumption Diagnostics






Empathy Residual Plot
1
0.5
0
-0.5
-1
-1.5
Residuals
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
Average


Satisfaction Residual Plot
1
0.5
0
-0.5
-1
-1.5
Residuals
0.00
1.00
2.00
3.00
4.00
5.00
Average