Tomas Bata University in Zlín
Faculty of Management and Economics

Doctoral Thesis Summary

Customer Experience analysis based on big data analysis in grocery retail

Analýza zákaznické zkušenosti na základě analýzy velkých dat v maloobchodu

Author: **Taha Falatouri, Ph.D.**

Degree programme: P0413D050013 Economics and Management

Supervisor: prof. Ing. Felicita Chromjaková, PhD

Examiners: prof. Daniel Mertens, PhD.

doc. Ing. Milan Edl, Ph.D.

Zlín, February 2025

Published by **Tomas Bata University in Zlín** in the Edition **Doctoral Thesis.**

The publication was issued in the year 2025

Key words: *Service quality assessment, Text mining, Sentiment Analysis, LLMs*

Klíčové slova: *hodnocení kvality služeb, mapování textu, analýza sentimentu, velké jazykové modely*

ISBN 978-80-7678-334-8

ABSTRACT

This study examines the effectiveness of combining Large Language Models (LLMs) with traditional text mining methods to analyze Service Quality (SQ) dimensions and customer satisfaction in both physical and omnichannel retail settings. We investigate the usage of NLP techniques and LLMs to extract sentiment and SQ dimensions from user-generated content (UGC). Analyzing datasets of customer reviews (in English and Persian), we identify key factors influencing customer satisfaction and dissatisfaction in physical retail and omnichannel environments by applying unsupervised text mining to customer reviews from supermarkets. Our analysis reveals that among the general SQ dimensions, personal interaction, store policies, and product-related Dimensions positively impact customer satisfaction, while reliability concerns contribute to dissatisfaction. The importance of personal interaction is particularly pronounced in smaller stores and towns. Conversely, hypermarkets should focus on improving physical aspects and enhancing personal interaction to reduce negative feedback. Integrating LLMs with text mining provides a comprehensive approach to analyzing SQ dimensions across different retail formats, emphasizing the necessity for ongoing human oversight to ensure the accuracy and reliability of sentiment analysis and information extraction. Nonetheless, there are challenges, such as discrepancies between model predictions and human judgments and difficulties in accurately identifying specific dimensions from unstructured text.

ABSTRAKT

Tato studie zkoumá efektivitu kombinace velkých jazykových modelů (LLM – Large Language Models) a tradičních metod analýzy mapování textu s cílem analyzovat dimenze kvality služeb (SQ – Service Quality) a spokojenosti zákazníků v prostředí fyzického i online maloobchodu. Zkoumáme využití NLP technik a LLM k extrakci sentimentu a dimenzí SQ z hlediska uživatelsky generovaného obsahu (UGC – User Generated Content). Analyzujeme zákaznické recenze v angličtině a perštině a identifikujeme klíčové faktory ovlivňující spokojenost a nespokojenost zákazníků v prostředí fyzického maloobchodu a vícekanálového prodeje prostřednictvím nesupervizovaného mapování textu zákaznických recenzí ze supermarketů.

Naše analýza odhaluje, že mezi obecnými dimenzemi SQ mají osobní interakce, obchodní politika a produktové faktory pozitivní dopad na spokojenost zákazníků, zatímco obavy o spolehlivost služeb přispívají k nespokojenosti. Význam osobní interakce je obzvláště výrazný v menších obchodech a městech. Naopak hypermarkety by se měly zaměřit na zlepšení fyzických aspektů a kvality osobní interakce, aby snížily negativní zpětnou vazbu.

Integrace LLM s textovou analýzou umožňuje komplexní zkoumání dimenzí SQ napříč různými maloobchodními formáty, což podtrhuje nutnost neustálého lidského dohledu pro zajištění přesnosti a spolehlivosti analýzy sentimentu a extrakce informací. Přesto se objevují výzvy, například nesoulad mezi predikcemi modelu a lidskými úsudky či obtíže při přesné identifikaci konkrétních dimenzí v nestrukturovaném textu.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

SQ	Service Quality
UGC	User Generated Content
LLMs	Large Language Models
SQD	Service Quality Dimensions
WOM	word-of-mouth
NLP	Natural Language Processing
LDA	Latent Dirichlet allocation
SCM	Supply Chain Management
WOM	Word of mouth

1. INTRODUCTION

High service quality enhances customer satisfaction, strengthens brand loyalty, and reduces the likelihood of customers switching to competitors. However, assessing SQ in supermarkets presents unique challenges due to the complexity of interactions between customers and staff, store layout, product availability, and overall shopping experience (Dabholkar et al. 1996).

Service Quality (SQ) has been explored across various industries, including hotels (Nilashi et al. 2022), travel (Chandra Mahapatra and Bellamkonda 2023), and delivery services (Uzir et al. 2021), where it plays a crucial role in customer retention and competitive differentiation (Slack et al. 2020).

Unlike other retail sectors where online shopping has gained significant traction (Park and Kim 2022), grocery shopping remains largely conducted in physical stores (Daher 2021; Dominici et al. 2021). This reliance on in-store shopping highlights the importance of enhancing service experiences through improved customer service, efficient store management, and personalized engagement (Frasquet et al. 2021). Retailers must balance the expectations of those who seek efficiency and convenience with the traditional elements of customer service that emphasize personal interaction, product quality, and store ambiance.

Traditional methods of assessing SQ primarily rely on structured surveys, where customers rate predefined dimensions such as reliability, responsiveness, and product quality. While these methods have been widely used, they present limitations by constraining customer responses within rigid frameworks (Lu et al. 2023). Surveys often fail to capture spontaneous, nuanced feedback that reflects real customer experiences. Additionally, survey fatigue and response bias can reduce the accuracy of collected data, as respondents may provide socially desirable answers rather than their actual opinions. In contrast, user-generated content (UGC), such as online reviews, social media discussions, and forum posts, provides real-time, unbiased customer insights (Villeneuve and O'Brien 2020). Unlike structured surveys, UGC allows customers to express opinions freely, providing deeper insights into service quality perceptions. However, despite its value, UGC remains vastly underutilized in SQ assessment due to the challenges associated with processing and analyzing large volumes of unstructured textual data (Duan et al. 2013; Mejia et al. 2021). Businesses often struggle to extract meaningful patterns from UGC due to its diversity in language, sentiment expression, and contextual relevance.

Advances in artificial intelligence (AI) and large language models (LLMs), such as ChatGPT and BERT, have significantly improved natural language processing (NLP) capabilities, offering new possibilities for analyzing unstructured text data (Brown et al. 2020; Chu 2023). These models enable sentiment analysis, text classification, and keyword extraction, allowing businesses to gain actionable insights from customer feedback. The adoption of

LLMs has been widespread, with 92% of Fortune 500 companies integrating AI-powered tools into their operations (Open AI & GPT News 2023). Despite their widespread adoption, the effectiveness of LLMs in SQ assessment remains an underexplored area (Lu et al. 2023).

LLMs offer an efficient way to process large volumes of UGC, but they still face challenges in contextual understanding and causal relationship interpretation (Puyt and Madsen 2024). Sentiment analysis in SQ evaluation has shown that specific service attributes significantly impact customer perceptions. These results indicate that AI-driven sentiment analysis can help businesses detect patterns in customer feedback, identify service gaps, and adjust their strategies accordingly.

This study investigates how text mining and LLMs can improve SQ assessment in supermarket retail by capturing authentic customer feedback across different store formats, cultural contexts, and sales channels. This research addresses several key questions: To what extent can text mining and LLMs effectively assess SQ using UGC? How do SQ dimensions vary across different retail channels in supermarkets? How does store type influence customer perceptions of SQ? Does city size affect the perceived importance of SQ dimensions? How do different SQ dimensions impact customer satisfaction? This study seeks to bridge the gap between traditional SQ measurement methods and AI-driven approaches for assessing service quality in physical retail settings by answering these questions.

Retailers must adapt to changing consumer behaviors by leveraging advanced analytics to identify service gaps, enhance personalization, and optimize customer engagement. As omnichannel strategies continue to reshape the retail landscape, businesses must ensure that service quality remains prioritized across digital and physical touchpoints. Recent studies emphasize the growing importance of consistency in service quality across multiple retail channels, where customers expect seamless experiences whether they shop online, use mobile apps, or visit physical stores (Sousa and Voss 2006). Additionally, customer expectations evolve over time and are influenced by technological advancements, economic conditions, and cultural shifts. By integrating LLMs and text mining into SQ assessment, this study aims to provide a scalable, data-driven approach that enhances customer insights, improves service evaluation accuracy, and contributes to a more comprehensive understanding of SQ in supermarket retail environments.

To achieve these objectives, this research employs a multi-faceted methodology that includes text mining, sentiment analysis, and machine learning-based classification techniques. First, service quality issues are extracted from UGC using both traditional and LLMs models, enabling a more holistic analysis of customer perceptions. Second, customer feedback is analyzed across languages. Third, sentiment classification is applied to categorize service-related content and measure customer sentiment towards different aspects of SQ. Fourth, external factors such as store type and retail format are examined to determine their influence on customer expectations and SQ perceptions.

Despite the benefits of AI-driven text mining and LLMs, challenges remain in terms of data interpretability, algorithm bias, and ethical considerations. AI models require large datasets for training, and biases in these datasets can lead to skewed results, affecting the fairness of SQ assessments. Additionally, while AI can analyze vast amounts of text, it lacks human intuition and cultural sensitivity, which are often critical in understanding nuanced customer feedback. These challenges highlight the need for a balanced approach where AI complements traditional human-led analysis rather than completely replacing it. By combining machine learning techniques with expert evaluation, businesses can achieve a more accurate and meaningful interpretation of customer feedback, ultimately leading to better decision-making in service quality management.

2. CURRENT STATE OF THE ISSUES DEALT WITH

2.1 Retail Industry

Retail is a key global economic sector, generating over $25 trillion annually. Retail consistently holds the top two spots in the Fortune 500, underscoring its global influence and highlighting the industry's dominance. Retail’s dynamic nature, shaped by technological and behavioral trends, attracts research interest. Businesses expand online and offline, making efficiency a critical focus (Mou et al. 2018). Supermarkets, growing at 4.5% annually, represent the fastest-expanding retail segment, yet profits are increasingly fragmented, with nearly 50% of grocery profits being divided among multiple players (Kuijpers et al. 2018). To remain competitive, retailers must analyze consumer behavior, enhance service quality, and embrace digital trends (Galante et al. 2013).

The grocery sector, in particular, is shifting toward digital sales and marketing, accelerated by recent crises. Grocery shopping is an essential part of daily life, yet its dynamics vary across regions.

2.1.1 Omnichannel Retail

The rise of new markets and suppliers has given consumers greater flexibility in shopping across multiple channels, allowing them to research in one channel and purchase in another (Marchet et al. 2018; Cai and Lo 2020). Retail has evolved from single-channel models to an omnichannel approach, where online and offline experiences merge to enhance convenience (Chopra 2018). Unlike traditional multichannel retail, where each channel operates independently, omnichannel retailing integrates all touchpoints into a seamless system, fostering customer satisfaction and brand loyalty (Shen et al. 2018). This strategy allows businesses to penetrate new markets while ensuring a cohesive, brand-driven customer experience (Manser Payne et al. 2017).

While omnichannel strategies are well-documented in retail (Saghiri et al. 2018), their adoption in grocery stores remains underexplored. Supermarkets have increasingly adopted omnichannel services, particularly in response to the COVID-19 crisis (Cocco and De-Juan-Vigaray 2022). However, they face unique challenges, including low profit margins on commodity items and complex logistics for perishable goods, requiring a strong focus on efficiency and agility. Additionally, serving a diverse customer base necessitates flexible solutions that cater to various shopping preferences. As supermarkets continue integrating omnichannel models, their success will depend on balancing customer experience and operational productivity within this evolving retail landscape.

2.2 Service Quality Assessment

Customer experience (CX) encompasses journey mapping, multichannel interactions, and adapting to mobile and digital landscapes. Measuring CX requires understanding the influence of various touchpoints and external factors such as economic conditions, crises, and technological advancements (Lemon and Verhoef 2016) SQ has evolved through multiple models addressing consumer expectations and perceptions. Customer value, defined as the trade-off between perceived benefits and sacrifices, is shaped by service quality, customer support, and personalization (Blocker et al. 2011). It is dynamic and influenced by changing expectations and external market forces.

(Gronroos 1990) distinguished between the "what" (service outcome) and the "how" (delivery process), emphasizing that both functional and interactive aspects define service quality (Prakash 2019; Albarq 2013). The intangible nature of services, human interactions, and environmental factors introduce variability, making it essential to systematically design, measure, and monitor service quality. Scholars argue that SQ is best assessed through disconfirmation theory, which compares actual performance with expectations while distinguishing SQ from customer satisfaction, which is more transaction-specific (Cronin and Taylor 1994).

The GAP Model (Parasuraman et al. 1985) identified ten SQ dimensions, later refined into the SERVQUAL model with five key components: Tangibles, Reliability, Responsiveness, Assurance, and Empathy (Parasuraman et al. 1988). While widely adopted, SERVQUAL has been criticized for its ambiguous definition of expectations, leading to alternative models like SERVPERF (Cronin and Taylor 1992), which focuses solely on performance outcomes. Retail-specific models such as RSQS (Dabholkar et al. 1996) and CALSUPER (Vázquez et al. 2001) were developed to assess SQ in supermarkets and physical retail. With the rise of e-commerce, frameworks like E-S-QUAL (Parasuraman et al. 2005) and eTailQ (Wolfinbarger and Gilly 2003) emerged to evaluate online service dimensions, including website design, fulfillment, and security.

The shift toward omnichannel retailing requires seamless integration of digital and physical touchpoints (Rigby 2011; Verhoef et al. 2015). Studies highlight the importance of transparency, consistency, and real-time data synchronization (Gao and Huang 2021). Omnichannel failures, including issues with logistics, pricing, and customer interactions, impact service quality (Rosenmayer et al. 2018).

Consumers expect integrated services, where loyalty programs, synchronized logistics, and cross-channel pricing contribute to a cohesive experience (Pantano and Viassone 2015). Online SQ research has expanded to include post-purchase support models like E-RecS-QUAL (Ulkhaq et al. 2017), which focuses on responsiveness, compensation, and customer support accessibility. As businesses prioritize CX in an increasingly competitive digital landscape, refining SQ models to capture omnichannel dynamics, customer engagement, and evolving consumer expectations remain critical (Lemon and Verhoef 2016). Traditional SQ frameworks must adapt to include seamless service delivery, real-time interactions, and cross-platform continuity for sustainable customer satisfaction.

2.3 User Generated Content and Text Mining

The rise of technology and social media has transformed consumer interactions, shifting engagement from face-to-face to digital platforms. Online research now guides purchasing decisions, with websites providing continuous brand interaction. After purchases, users share feedback, shaping others' perceptions and offering businesses valuable SQ insights (Falatouri et al. 2024a). UGC acts as digital word-of-mouth, capturing customer experiences and helping companies refine their services (Holmlund et al. 2020). The retail sector has been particularly affected, with the rise of online shopping contributing to store closures and discussions of a "retail apocalypse" (Helm et al. 2020).

Big data analytics enhances businesses' ability to process vast amounts of UGC, enabling real-time identification of trends and consumer needs (Holmlund et al. 2020). However, handling big data requires sophisticated tools to manage its complexity in volume, speed, and accuracy (Chen et al. 2012). Data analytics supports strategic, data-driven decisions, ensuring continuous service improvements when utilized effectively.

Businesses increasingly use text mining to process UGC efficiently, which extracts meaningful insights from unstructured text data (Sun and Reddy 2013). Unlike data mining, which deals with structured datasets, text mining applies to emails, blogs, and reviews. It supports decision-making, process automation, and trend analysis (Berger et al. 2022). Techniques include unsupervised methods like clustering and Latent Dirichlet Allocation (LDA) to detect latent customer needs without labeled data (Büschken and Allenby 2016), and supervised methods like Convolutional Neural Networks (CNNs) and BiLSTM networks for precise text classification (Wang et al. 2022). While unsupervised models struggle with semantic accuracy, supervised approaches require high-quality labeled data, making hybrid techniques a promising solution. By transforming raw text into actionable knowledge, text mining enhances business efficiency and supports strategic decision-making (Sun and Reddy 2013).

The rise of LLMs like GPT and BERT has transformed NLP, enabling applications such as text generation, sentiment analysis, and customer service automation (Devlin et al. 2019; Alec Radford et al. 2019). Unlike traditional models relying on predefined patterns, LLMs process vast datasets, capturing linguistic nuances for adaptive learning (Kojima et al. 2022; Chu 2023). Integrated with big data analytics, they extract insights from unstructured UGC, enhancing service quality (SQ) assessments beyond traditional surveys (Leippold 2023).

In healthcare, GPT-4o achieves 92.3% sensitivity in osteoarthritis detection, though medical oversight remains essential (Pagano et al. 2025). In education, GPT-4 approaches human benchmarks in automated essay scoring with structured prompt engineering (Tang et al. 2024). Businesses use LLMs to enhance customer service, personalize marketing, and optimize decision-making.

A key NLP application is sentiment analysis, which classifies textual emotions as positive, negative, or neutral (Eachempati et al. 2022; Nasseri et al. 2023). Machine learning and deep learning models now surpass traditional lexicon-based methods in accuracy (Falatouri et al. 2024b). Despite progress, sentiment analysis still struggles with understanding causal relationships. Causal prompts improve LLM performance, but challenges remain in fully grasping cause-effect dynamics (Lyu et al. 2024). Nevertheless, sentiment analysis continues evolving, helping businesses refine services and enhance customer satisfaction.

3. RESEARCH GAPS

The study of SQ in physical retail has advanced significantly through models like SERVQUAL and RSQS. However, critical gaps persist in adapting these frameworks to today's culturally diverse, technology-driven, and omnichannel retail environments. Traditional survey-based methods rely on static, predefined dimensions, limiting their ability to capture real-time, evolving customer expectations (Kaul 2007; Jain and Aggarwal 2018). Additionally, cultural and contextual variability in SQ perceptions remains underexplored despite evidence that regional norms influence feedback styles and satisfaction levels (Voss et al. 2004; Al-Deehani and Aldeehani 2017). Another major limitation is the underutilization of UGC as a data source for SQ assessment. While UGC provides rich, spontaneous customer insights, existing analytical methods struggle to process its unstructured nature at scale, limiting its practical application in retail service evaluation (Mejia et al. 2021; Korfiatis et al. 2019).

As omnichannel retailing becomes the norm, integrating online and offline experiences, its application in supermarkets remains understudied (Saghiri et al. 2018). Unlike other retail sectors, supermarkets operate with low-profit margins and complex logistics for perishable goods, requiring agility and operational efficiency (Fransoo et al. 2019). While omnichannel models aim to enhance convenience, supermarkets must also manage high transaction volumes and cater to a diverse customer base, necessitating highly adaptive SQ models that capture frequent, multi-touchpoint interactions.

A key gap in current research is the limited application of text mining and LLMs in analyzing UGC within supermarket retail. While text mining has been widely adopted in banking and hospitality, its role in grocery retail is underdeveloped, even though customer feedback in this sector is abundant and highly relevant. Traditional SQ models fail to capture the interplay between digital and physical retail channels, leaving supermarkets without a comprehensive understanding of omnichannel service experiences (Zhang et al. 2019).

The integration of LLMs and AI-driven text mining presents a transformative opportunity to address these research gaps. Unlike structured surveys, LLMs can process and analyze large-scale UGC, identifying patterns, sentiment, and emerging themes that would otherwise be missed. Their ability to track shifts in customer expectations over time allows for longitudinal SQ analysis, ensuring that service models remain relevant and adaptive. Furthermore, current studies often fail to link SQ insights from UGC to actual business outcomes, such as customer loyalty, repeat purchases, or behavior changes (Choi et al. 2019; Le et al. 2024). AI-driven analysis can bridge this gap by offering actionable insights that directly inform retail strategy and enhance customer satisfaction.

Addressing these challenges requires a new approach that leverages LLMs and advanced text mining techniques to modernize SQ assessment in omnichannel supermarket retail. Traditional methods, while foundational, are no longer sufficient for capturing the complexity of today's shopping behaviors. AI-powered solutions enable businesses to gain deeper, real-time insights, ensuring adaptive service strategies that align with evolving consumer demands. This shift will advance academic research and empower businesses with data-driven solutions to enhance customer experience, loyalty, and operational efficiency in an increasingly competitive retail landscape.

4. THEORETICAL FRAMEWORK

The assessment of SQ has traditionally relied on models such as SERVQUAL, GAP Theory, and Expectation-Confirmation Theory (Parasuraman et al. 1985). These models focus on measuring the gap between customer expectations and perceptions of service. However, conventional assessment methods often rely on static, questionnaire-based approaches, which limit their ability to capture dynamic customer expectations, real-time feedback, and cultural nuances (Kaul 2007; Jain and Aggarwal 2018). We propose integrating LLMs and text mining techniques into SQ assessment to overcome these limitations. By analyzing UGC including online reviews, social media posts, and feedback forms LLMs can extract actionable insights at scale, making SQ evaluation more accurate, contextual, and responsive to market changes (Mejia et al. 2021; Korfiatis et al. 2019).

Our proposed framework, illustrated in Figure 1, integrates external influences such as store type, business type, city type, and cultural context with theoretical SQ models to assess how service quality is perceived in different environments. The GAP Model measures discrepancies between expected and actual service performance, while Expectation-Confirmation Theory (ECT) evaluates whether post-service experiences align with initial customer expectations (Hossain and Quaddus 2012). The Disconfirmation Model assesses whether a service exceeds, meets, or fails to meet expectations, while SERVQUAL categorizes SQ into five key dimensions: tangibles, reliability, responsiveness, assurance, and empathy. This framework accounts for both objective service performance and subjective customer perceptions within specific business and cultural contexts.

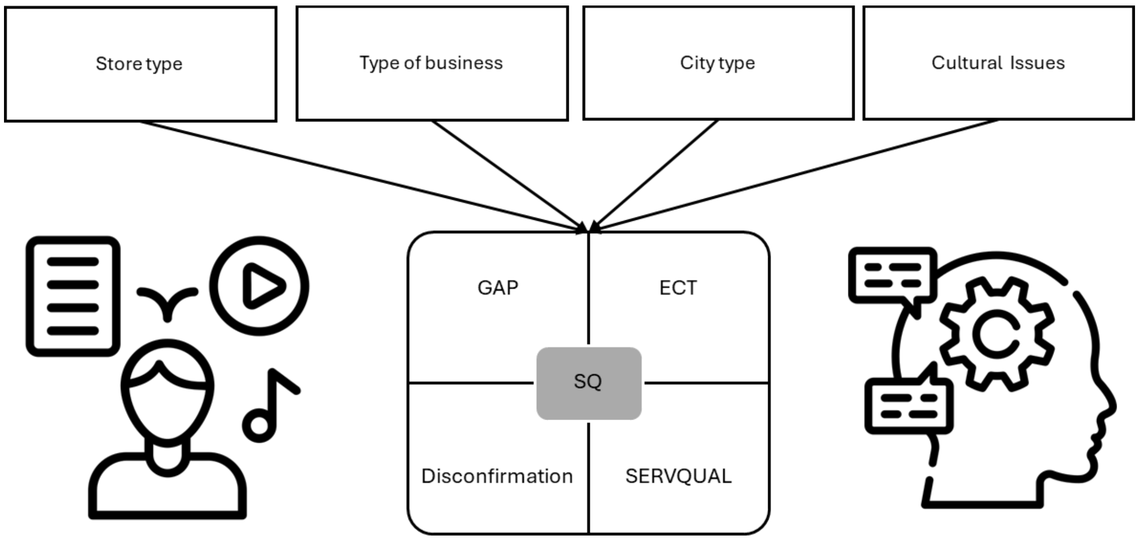

Figure 1 Theoretical Framework (Self reference)

Traditional sentiment analysis methods struggle with contextual nuances and multilingual data, often overrepresenting negative feedback, as dissatisfied customers are more likely to express their opinions (Martin-Domingo et al. 2019; Falatouri et al. 2024b). By leveraging machine learning-based sentiment analysis, this research demonstrates that LLMs can provide more balanced, comprehensive assessments of SQ across different channels.

Q1: Text mining and LLMs significantly improve the accuracy of service quality assessments?

H1a: Text mining and LLMs significantly improve the accuracy of SQ assessments in the supermarket industry.

H1b: Text mining and LLMs provide more comprehensive insights into customer expectations and preferences compared to traditional service quality assessment methods.

H1c: The use of LLMs allows for a higher degree of contextual understanding in service quality assessments, making it possible to detect complex patterns in customer feedback across different channels.

Previous research demonstrates the effectiveness of text mining and LLMs in various sectors, including restaurants, banking, and healthcare, but their use in supermarket retail remains underexplored (Mejia et al. 2021; Rula and D'Souza 2023). Given the unique challenges of supermarket service, such as inventory management, logistics, and high transaction volumes, exploring the role of AI-driven insights is critical for optimizing service quality, customer satisfaction, and operational efficiency.

Q2: How do the dimensions of service quality vary across different retail channels in the supermarket context?

H2a: Different retail channels (online vs. in-store) prioritize distinct SQ dimensions, with in-store experiences emphasizing tangibles and staff interaction, while online services focus on convenience and ease of us.

H2b: Omnichannel supermarkets demonstrate a higher integration of service quality dimensions, particularly with respect to channel integrity, than single-channel stores.

H2c: Customer expectations regarding specific service quality dimensions (e.g., reliability, responsiveness) differ significantly between online and physical retail channels.

Retail studies suggest that in-store service quality depends on store layout, cleanliness, and personal engagement (Dabholkar et al. 1996), while online experiences emphasize website usability, delivery speed, and seamless navigation (Verhoef et al. 2015). However, supermarket-specific research is limited, requiring further exploration into how SQ dimensions vary between channels and how businesses can optimize omnichannel strategies.

Q3: How does the importance of service quality dimensions vary across different store types?

H3a: Larger stores place a higher emphasis on service quality dimensions related to variety and convenience, whereas smaller stores prioritize personalized interactions and product quality

H3b: The impact of reliability and responsiveness on customer satisfaction is more significant in medium-sized supermarkets than in smaller specialty stores.

H3c: Smaller stores demonstrate higher customer satisfaction when emphasizing interpersonal engagement, while larger stores benefit more from well-organized and accessible environments.

Studies indicate that large-format stores (e.g., hypermarkets) succeed by offering efficiency and product variety, (Khare 2013; Tannady et al. 2018)while small local stores gain loyalty through personalized service and community engagement .Supermarkets must adapt service strategies to meet customer expectations across different store formats to ensure sustained competitiveness.

Q4: Is there a difference in the perceived importance of service quality dimensions across different city types?

H4a: In larger cities, service quality dimensions related to convenience and product variety have a higher perceived importance compared to smaller cities.

H4b: Customers in smaller cities place greater emphasis on personal interaction and familiarity in service quality than those in larger cities.

Retail environments differ based on geographic and demographic factors. Urban customers prioritize fast, efficient service, while rural consumers place greater importance on social connections and trust in service providers (Le et al. 2024). Understanding these regional preferences is essential for tailoring service offerings to different market segments.

Q5: How does the significance of different service quality dimensions (based on the frequency of review comments) vary in relation to customer satisfaction levels?

H5a: Service quality dimensions with a higher frequency of review comments correlate positively with higher customer satisfaction levels.

Analyzing online reviews provides valuable insights into which service attributes influence satisfaction the most. Higher engagement with specific service dimensions indicates customer priorities, allowing businesses to focus on key areas for improvement (Zhang et al. 2014). Identifying patterns in review frequency helps retailers refine their service strategies, track customer sentiment in real-time, and proactively address emerging service gaps.

5. PROCESSING METHODS

Service quality (SQ) evaluation has traditionally utilized methods like SERVQUAL (Dabholkar et al., 1996; Shokouhyar et al., 2021), surveys (Omar et al., 2021), post-service feedback (Dholakia et al., 2010), expert opinions (Korfiatis et al., 2019), and sentiment analysis (Dhaoui et al., 2017). However, these approaches face limitations such as response bias, narrow focus, and lack of generalizability (Basfirinci & Mitra, 2015). Our study leverages this data to apply text-mining techniques across multiple food retail companies, overcoming traditional biases and enhancing SQ assessment.

5.1 City Size Classification

In comparing store locations by city size, we categorized these into three groups: small cities with fewer than 5,000 residents, cities with populations between 5,000 and 40,000, and large cities with populations exceeding 40,000

5.2 Store Type Classification

Bonfrer et al. (2022) examined food retail types and identified key similarities and differences based on store size and services. Supermarkets generally offer medium to large spaces with varied products. Discounters, while similar in size and location, emphasize low prices, fewer products, limited customer service, and private-label offerings. Mass merchandisers, such as warehouse clubs (e.g., Costco) and supercenters (e.g., Walmart), cater to bulk and convenience needs. Convenience stores include both chain-operated outlets (e.g., 7-Eleven) and those adjacent to gas stations, focusing on quick service. Specialty stores, often smaller and independently owned, provide focused product lines. We further refined this classification for retail. defining five store types based on size and distinctive characteristics.

5.3 Data Acquisition and Data Anlytic Techniques

We collected datasets to demonstrate the effectiveness of Text Mining and LLMs for SQ assessment.

The first data set relied on publicly accessible review data. Specifically, the dataset comprises over 500,000 Google reviews related to Austria's five largest food retail chains, from which 163,055 textual reviews were extracted for this study. The dataset spans approximately 41 months, covering the period from January 2017 to May 2020. The extraction method was developed with reference to previous studies (Udokwu et al. 2020). Each review entry includes a set of structured features ID, author name, author, link, star rating, review text, and review date. The second datasets include customer reviews of mobile apps, comprising more than 8,000 reviews of an Iranian supermarket's mobile app (OKALA) and over 1,000 reviews of the supermarket’s customer club app (OK Club). These reviews were obtained from cafebazaar.ir, Iran's most popular site for rating Android apps. This dataset was collected in February 2022 and includes reviews from April 2019 to February 2022.

5.4 Data Preprocessing

In the initial cleaning phase, we removed comments that contained no text. Since we employed different methods, text mining relies heavily on thorough text preparation, while using advanced LLMs requires less preprocessing as the model can filter out irrelevant textual data. Additionally, keeping emojis and other non-textual elements might be beneficial in revealing the writer's sentiment.

For traditional timing models, To prepare and clean the text corpus, we implemented a Python script utilizing regular expressions. We began by leveraging the "stop words" library to eliminate uninformative words, we then tokenized the corpus, transforming sample sentences from the original text into lists of meaningful words, Next, we removed all punctuation, non-text elements, numbers, and emojis. We then isolated unique words within the corpus. Initially, the corpus contained unique words, but after excluding high-frequency, low-value words like "the" this was reduced to unique words. After filtering out these terms, we applied the Lancaster Stemmer to extract the root forms of words, further refining the corpus *Figure 2*.

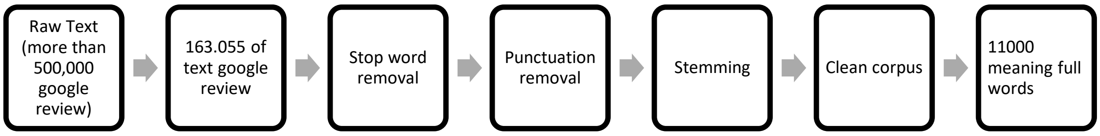

Figure 2 Text preprocessing pipeline and steps (Falatouri et al. 2024a)

To enhance tokenization, we employed the tidytext N-gram analysis package. Bigram words were then combined and treated as single units. Subsequently, we focused on selecting words with the greatest variation in per-topic-per-word probabilities across subjects. This approach allowed us to identify the key features of each topic and conduct a comprehensive comparison of all topics simultaneously.

5.5 Text Mining Approach

5.5.1 LDA Approach

LDA effectively categorizes document content by associating topics with words through a Dirichlet-based framework (Zhou and Zhang 2016). It autonomously groups words into themes, aiding various fields, including retail for product strategy (Park et al. 2023) and supermarket customer satisfaction analysis (Falatouri et al., 2024a).

LDA was chosen for its ability to analyze diverse customer comments and handle synonymy, grouping related terms (e.g., "location," "area") using Gibbs sampling to enhance topic coherence. It models topics (α) and word associations (β), where each document links to multiple topics. The learning process involves human-guided topic selection, enabling the algorithm to autonomously analyze new documents. LDA constructs “topic-per-document” and “words-per-topic” models using Dirichlet distributions, allowing automated topic clustering and document relationship identification. Our LDA implementation procedure for analyzing customer reviews is summarized in Figure 3:

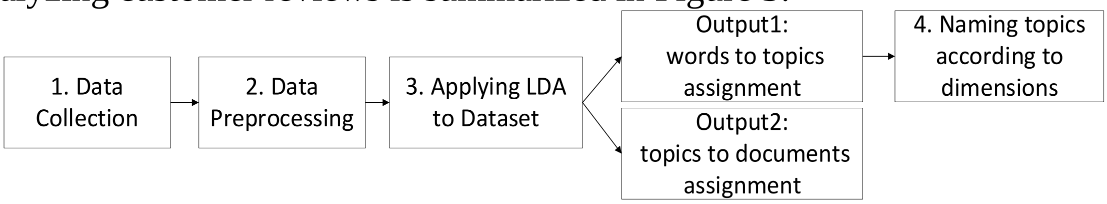

Figure 3 LDA Analysis Method (Falatouri et al. 2024a)

5.5.2 Sentiment Analysis Models

Three traditional text-processing models utilized to ensure robust analysis. The first model, TextBlob, is a Python library that employs probabilistic methods, including the Naive Bayes classifier, to analyze text and perform sentiment classification (Steven 2024).TextBlob relies on rule-based methods and probabilistic techniques, including the Naive Bayes classifier. It analyzes text by assigning polarity scores, ranging from -1 (negative sentiment) to +1 (positive sentiment), and subjectivity scores, which indicate the degree of opinion versus fact within the text. This dual-scoring system enables nuanced evaluations of sentiment while offering insights into the emotional tone of the content (Steven 2024). VADER (Valence Aware Dictionary and Sentiment Reasoner), is a widely used rule-based tool explicitly designed for sentiment analysis in textual data. It employs a lexicon-based approach, leveraging a dictionary of words, phrases, and emoticons, each of which is pre-assigned a polarity score. These scores indicate whether the sentiment associated with a term is positive, negative, or neutral, and they also include intensity measures to capture varying degrees of sentiment (Hutto and Gilbert 2014; Aljedaani et al. 2022).

Unlike earlier models, such as recurrent neural networks, which process text sequentially, Transformers process the entire input at once, enabling parallel computation and improving both efficiency and scalability (Naseem et al. 2020).

The core of the Transformer architecture lies in its self-attention mechanism, which calculates the importance of each word in relation to every other word in a sentence. This allows the model to understand context and meaning across long text spans, making it exceptionally effective for tasks involving complex dependencies, such as sentiment analysis, translation, and summarization (Vaswani et al. 2017).

5.5.3 LLMs Approach

LLMs are powerful tools for evaluating SQ through advanced NLP. The evaluation process consists of two key stages: (1) identifying and categorizing SQ-related content and (2) assessing its alignment with customer expectations. By automating these processes, LLMs enhance efficiency, consistency, and scalability.

In the first stage, LLMs extract and categorize content from customer reviews, online feedback, and UGC, classifying it into dimensions like reliability, responsiveness, assurance, and empathy (Devlin et al. 2019). They recognize both explicit and implicit SQ indicators, helping service providers pinpoint areas for improvement. Research has shown their effectiveness in healthcare and other domains, providing structured feedback while reducing manual effort (Alexander and Kent 2022). The second stage involves sentiment analysis, where LLMs detect emotional tones and nuanced sentiments in large datasets (Wankhade et al. 2022). Beyond basic sentiment classification, they capture mixed sentiments and subtle emotional cues, offering a more accurate representation of customer perceptions. Their ability to process vast amounts of feedback makes them valuable for large-scale businesses, enabling proactive service improvements.

To validate this approach, we used ChatGPT 3.5 (OpenAI) and Claude 3 (Anthropic), both known for advanced text analysis and reasoning capabilities (Dinesh & Nathan 2023; Konrad & Cai 2024). These models, with general-purpose pre-training and Python-compatible APIs, allow seamless integration into workflows, making SQ evaluation more data-driven and scalable.

6. EXPERIMENTAL ANYLISIS

6.1 Text Mining Analysis

We conducted LDA analysis on datasets from physical retail stores and omnichannel supermarkets to identify key dimensions. Using the TM library in R . Topic effectiveness was assessed via a human review-based approach, where 2% of reviews were manually evaluated to compare machine-generated and human-assigned topics.

To quantify agreement, we calculated Cohen's Kappa coefficient (Wongpakaran et al. 2013), accounting for chance agreement. The highest inter-rater reliability (IRR) was achieved with five clusters for retail stores and six for omnichannel supermarkets, yielding a Cohen's Kappa above 0.7, indicating substantial agreement.

Key topic characteristics were identified by analyzing "beta" values, representing per-topic word probabilities. Words with beta > 1/1000 were selected for further examination using:

$$log_2 \frac{beta_i}{average\ of\ other\ topics\ beta} \quad (6)$$

This approach enabled cluster comparison and linked key topics to Service Quality (SQ) dimensions.

Physical Store SQ Analysis, among SQ dimensions, policy was the most significant, appearing in 30.9% of reviews, followed by personal interaction (21.7%), reliability (20.1%), and physical aspects (14.9%).

Table 1 Share of SQ dimension in customer reviews Physical store (Falatouri et al. 2024a)

SQ dimension	Percentage of comments	Frequency of comments
System availability	41.88%	3768
Efficiency	27.73%	2495
Fulfillment & Product	18.73%	1685

Integration	5.18%	465
Contact	4.03%	362

Omnichannel Store SQ Analysis, for omnichannel stores, system availability was the most dominant dimension (41.88%), followed by efficiency (27.73%), fulfillment & product (18.73%), and integration (5.18%). Contact (4.03%) played a minor role, while compensation (2.45%) had minimal impact.

Table 2 Share of SQ dimension in customer reviews Omnichannel store(Self refrence)

SQ dimension	Percentage of comments	Frequency of comments
System availability	41.88%	3768
Efficiency	27.73%	2495
Fulfillment & Product	18.73%	1685
Integration	5.18%	465
Contact	4.03%	362
Compensation	2.45%	220

6.2 LLMs Analysis

Selecting the appropriate prompt for an LLM is crucial, as it significantly affects the quality of the model's output. To optimize the prompt's effectiveness, we experimented with various prompts tailored to specific use cases, such as identifying SQ-related content and assessing review sentiment. Our method drew examples from Zhang et al. (2019), followed the OpenAI guidelines (OpenAI 2023), and focused on creating clear and concise prompts.

Our prompts are made up of two main parts. The first part is the command or query that is used to interact with the model to generate a response. In particular scenarios, like comment categorization and sentiment analysis, the second part includes the data that the model must process to produce the desired results. This data is formatted as a JSON string. Below is an example of how such an interaction works:

Here are customer reviews for an Iranian Omnichannel retailer in the given json string. Based on the review answer following questions in English.

In one word what type of service quality dimension has been mentioned in the text?

What is the sentiment of the text based on these three classes: positive, neutral, or negative?

Print your result in a json string including 'ReviewId', 'Channel', 'Service_Quality_Dimension', 'Main_Concern', 'Sentiment', 'Satisfaction_Source'. Just print a json string nothing else.

[{"Id":75552883,"Review":"فوق العاده ست واقعا دارن زحمت میکشند به کوری چششششششم بعضیا م برای ی"},{"Id":76792157,"Review":"فروشگاهیه که توتموم"},{"Id":74217171,"Review":"ایراد دااااره"},{"Id":71481956,"Review":"گرونی پناه مردم هست"},{"Id":71271524,"Review":"کار نمی کونه"},{"Id":70530031,"Review":"ایرادات دارد باز"},{"Id":68768897,"Review":"متاسفانه نرمافزار مناسب ی نیست"},{"Id":63169771,"Review":"واقعاً بی کاربرده"},{"Id":60818080,"Review":"هیچ کارایی ندارد"}]

Our workflow involved importing customer reviews into Pandas, crafting custom prompts for ChatGPT and Claude 3, and batch-processing reviews via Anthropic and OpenAI APIs. A key prompt, "In one word, what type of service quality dimension has been mentioned?" yielded 171 unique outcomes from ChatGPT and 148 from Claude 3. ChatGPT's top 15 responses covered 60% of reviews, while Claude 3's top 10 covered 80%. ChatGPT left 7% of cases unclassified as "N/A," whereas Claude 3 assigned a dimension to nearly all reviews, even when irrelevant.

ChatGPT provided structured, customer-centric outputs, classifying physical aspects, reliability, and personal interaction with a neutral tone. In contrast, Claude 3 emphasized negative experiences, highlighting slow service, technical issues, and dissatisfaction. ChatGPT viewed policies broadly (pricing, promotions), while Claude 3 focused on cost-related concerns. Additionally, ChatGPT summarized online services functionally, while Claude 3 detailed technical issues like app errors.

For a systematic comparison, we analyzed 363 reviews and 87 non-SQ-related reviews. ChatGPT labeled 22% as "N/A", whereas Claude 3 assigned labels indiscriminately. Table 4 summarizes how both LLMs handled undefined SQ cases.

To compare LLM outputs with established Service Quality (SQ) models, we referenced the CALSUPER model (Vázquez et al., 2001), which includes physical aspects, reliability, personal interaction, and policies. Due to app-related feedback, we introduced "Online Service" and "Overall Service" dimensions. Applying Jaccard Similarity (Vijaymeena and Kavitha 2016), we found <2% overlap between LLM-generated dimensions and theoretical SQ models, highlighting discrepancies.

After manual abstraction, ChatGPT correctly classified 76% of reviews into theoretical SQ dimensions, while Claude 3 achieved 68% accuracy. Claude 3 frequently assigned "Overall Service" to short comments, reducing specificity. Applying a Kolmogorov-Smirnov (KS) test ($P = 0.93$) revealed no significant distribution differences between the two LLMs.

6.3 Sentiment Analysis

We reused the sub-sample of 363 Service Quality (SQ) reviews, classifying their sentiment as negative, neutral, or positive based on human raters and five automated methods: ChatGPT, Claude 3, TextBlob (Aljedaani et al., 2022), VADER (Hutto & Gilbert, 2014), and a transformer-based model (Naseem et al., 2020). Disagreements among human raters were resolved through discussion.

A key limitation of traditional sentiment analysis models was their inability to process Persian text, requiring translation before analysis. The LLMs and traditional methods were then compared using Cohen's Kappa to measure agreement with human ratings (Wongpakaran et al., 2013). Table 5 summarizes sentiment classifications and interrater agreement scores.

The results show that LLMs outperformed traditional models, with ChatGPT achieving the highest agreement (Kappa = 0.64), followed by Claude 3 (0.57) and the transformer model (0.51). Traditional models struggled, with TextBlob and VADER scoring 0.28, indicating weak alignment with human assessments. Notably, ChatGPT slightly overestimated negativity while Claude 3 showed a stronger bias toward negative sentiment, underclassifying neutral and positive reviews. Traditional models misclassified a large proportion of reviews, with TextBlob and VADER overestimating neutrality and positivity.

An interesting finding was that translation did not significantly impact performance, as sentiment classification results remained consistent across datasets, suggesting the robustness of the employed methods. These results confirm that LLMs significantly outperform traditional sentiment analysis models, particularly in handling nuanced sentiment across different languages and datasets.

Table 3 Interrater agreement for human raters and text mining methods (Falatouri et al. 2024b)

	Human raters	**ChatGPT**	**Claude 3**	**TextBlob**	**Transformer**	**Vader**
Negative	209	236	251	88	249	106
Neutral	77	53	46	137	-	137
Positive	77	74	66	138	114	120
Kappa index – dataset 2	-	.64	.57	.28	.51	.28

7. DISCUSSION

This study confirms the effectiveness of text mining and Large Language Models (LLMs) in enhancing Service Quality (SQ) assessment. The acceptance of H1a, H1b, and H1c demonstrates that LLMs significantly improve the accuracy of SQ analysis, provide richer insights into customer expectations, and enable a higher degree of contextual understanding. These findings align with (Tang et al. 2024), who emphasized the role of LLMs in extracting structured insights from unstructured text while highlighting the need for advanced prompt engineering to refine results. Similarly, Wan et al. (2024) confirmed that LLMs outperform traditional text mining techniques, particularly in automating label taxonomy generation and reducing manual effort.

The results also provide strong evidence for H2a, H2b, and H2c, indicating that retail channels exhibit distinct service quality dimensions, with in-store experiences emphasizing tangible factors, online channels prioritizing convenience, and omnichannel supermarkets achieving greater integration of SQ dimensions. These findings align with Parasuraman et al. (2005) and Lemon & Verhoef (2016), who argued that service quality expectations differ significantly between physical and digital retail formats. Furthermore, Piotrowicz and Cuthbertson (2014) highlighted the necessity of integrated service experiences in omnichannel environments, reinforcing the current study's findings that omnichannel supermarkets have more highlighted dimensions than single-channel stores in SQ integration.

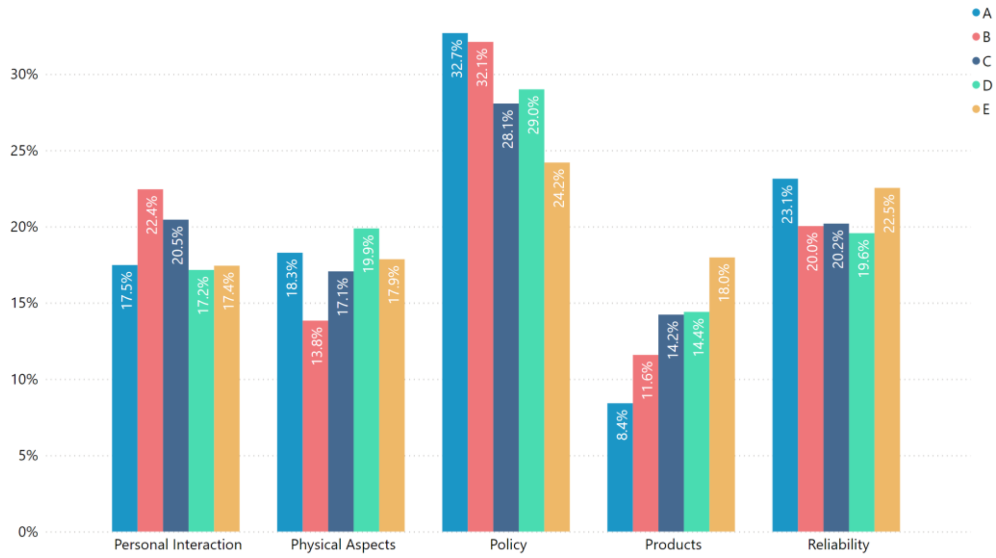

Figure 4 SQ dimension importance in different store types (Falatouri et al. 2024a)

In contrast, H3a, H3b, and H3c were rejected, suggesting that store size does not strictly dictate SQ dimensions or customer satisfaction drivers. Larger stores do not necessarily gain an advantage from increased variety or accessibility, nor do smaller stores always benefit from interpersonal engagement. Instead, store type and operational model are more decisive in shaping SQ expectations. These findings contrast with earlier studies, such as Vázquez et al. (2001) and Kaul (2007), which suggested that store size directly influences customer perceptions of service quality. However, the results align more closely with Hossain et al. (2020), who argued that SQ expectations are increasingly defined by a retailer's specialization rather than its physical scale. The findings also support Xu and Jackson (2019), who demonstrated that omnichannel supermarkets focus on seamless integration, whereas single-channel stores emphasize individual SQ attributes Figure 4.

Findings related to H4a and H4b confirm that city size significantly impacts SQ expectations. Customers in large cities prioritize efficiency, convenience, and product variety, while those in smaller cities emphasize personal interaction and familiarity. These results align with Parasuraman et al. (1988) and Blocker et al. (2011), who found that urban customers demand standardized, high-speed service, whereas small-town customers prioritize relational engagement. Similarly, Pantano & Viassone (2015) highlighted that smaller retailers in rural areas rely on relationship-driven strategies to maintain competitiveness, which is consistent with the present study's findings that personal interactions are more valued in smaller cities.

The rejection of H5a further underscores the complexity of SQ assessment. The assumption that frequently mentioned SQ dimensions correlate with customer satisfaction was disproven, as high-frequency dimensions often indicate pain points rather than key satisfaction drivers. The findings confirm that negative experiences are more likely to generate customer feedback, leading to overrepresenting complaints in review data. Similarly, Pantano & Priporas (2016) emphasized that technology-dependent dimensions, such as system availability, amplify frustration in omnichannel contexts, further reinforcing the study's finding that reliability and system failures generate the highest review volume but are primarily associated with dissatisfaction.

These findings collectively highlight the transformative role of LLMs and text mining in service quality assessment, emphasizing the importance of context-aware evaluations that account for retail format, regional differences, and evolving consumer expectations. While LLMs enhance automation and efficiency, human oversight remains crucial for ensuring theoretical alignment and minimizing biases (Shahbazi & Byun, 2020). The results suggest that a hybrid approach, integrating LLM-driven analysis with domain expertise, may yield the most accurate and actionable SQ insights.

8. CONCLUSION

Service quality (SQ) dimensions vary across retail contexts, shaped by store type, city size, and cultural norms. This study applied LDA and Large Language Models (LLMs) such as ChatGPT and Claude 3 to analyze user-generated content (UGC) from physical and omnichannel retail environments, demonstrating their efficiency in processing multilingual data and extracting meaningful SQ insights. Cultural habits strongly influence customer expectations, with personal interaction being a key driver of satisfaction in smaller towns, while urban consumers prioritize efficiency, reliability, and store layout due to greater competition and alternative options. In large retail settings such as hypermarkets and gas stations, functional elements like store organization and service speed play a more significant role, whereas smaller retailers thrive on familiarity and personalized engagement. Payment preferences also reflect cultural values, necessitating a balance between traditional and digital payment options to accommodate diverse consumer habits. Retailers must adapt their strategies accordingly small stores should invest in interpersonal service training to enhance staff-customer interactions, while larger retailers should optimize store layouts, navigation, and signage to streamline customer experiences. Omnichannel businesses must ensure seamless integration between digital and physical shopping experiences, synchronizing pricing, promotions, and inventory across platforms while maintaining system reliability to prevent service disruptions. Customers increasingly expect fast and transparent service, with real-time assistance, smooth fulfillment, and reliable inventory tracking being crucial to satisfaction. Additionally, strong customer support mechanisms, including chat, email, and phone support, are essential in mitigating dissatisfaction, while fair compensation policies, such as automated refunds and efficient complaint resolution, enhance trust and loyalty. AI-driven personalization through recommendation engines and automated chatbot support can further improve customer engagement and satisfaction. Applying LLMs to SQ assessment, this study found that these models significantly improved text classification and sentiment analysis, outperforming traditional NLP approaches like TextBlob and VADER. ChatGPT's single-label classification approach allowed for faster processing, whereas Claude 3's multi-label strategy captured the complexity of reviews that addressed multiple concerns. While LLMs demonstrated strong potential in handling unstructured data, challenges such as misclassification and occasional inaccuracies underline the need for human supervision and refined prompting strategies to ensure reliable results. These findings highlight the necessity of adapting service quality strategies to local market conditions while leveraging advanced AI tools for more effective and scalable service assessments. Future research should explore the integration of LLMs with domain-specific lexicons and hybrid analytical frameworks to further improve SQ dimension extraction and interpretation. Additionally, cross-cultural comparisons could provide further validation of the findings, ensuring the applicability of LLM-enhanced SQ assessment methods across diverse retail environments.

8.1 Contribution for Science

This research advances service quality assessment methodologies by demonstrating the effectiveness of Large Language Models (LLMs) in processing unstructured textual data. Traditional text mining approaches often struggle with language nuances, contextual interpretation, and scalability, especially in multilingual datasets. By leveraging LLMs like ChatGPT and Claude 3, this study addresses these limitations. LLMs significantly reduce noise in large datasets, improving analysis accuracy. Unlike LDA-based topic models, LLMs capture more complex and interconnected SQ dimensions, While LLMs outperform rule-based models, human validation remains necessary to correct misclassifications and improve domain-specific performance.

The integration of sentiment analysis further strengthens the methodological rigor of this study. Unlike traditional models like VADER or TextBlob, which struggle with domain-specific expressions, LLMs demonstrate greater adaptability and accuracy across different languages and contexts.

8.2 Contribution to Practice

Optimizing customer satisfaction requires tailored strategies based on store type, retail format, and customer expectations. Findings indicate that contact quality and system availability are major satisfaction drivers, while compensation policies strongly influence long-term loyalty.

Small stores in close-knit communities should focus on personalized engagement through staff training in empathetic communication and loyalty programs.

Large urban stores must prioritize efficiency using queue management and real-time inventory tracking.

Omnichannel retailers should integrate online and offline channels, ensuring consistent pricing, promotions, and inventory synchronization.

System availability is a persistent challenge investing in redundant server infrastructure and 24/7 IT support is essential for maintaining customer trust.

Regional payment preferences also play a role in shaping customer experiences.

8.3 Limitations of Research

While this study provides novel insights, certain limitations remain:The rapid digitization of retail post-pandemic continues to reshape consumer behavior, requiring ongoing refinements in SQ evaluation models.

AI-driven service quality assessment faces ethical concerns, including biases in training data and misclassifications in multilingual analysis, Merging the captured data with personal data could influence the result radically.

REFERENCES

Albarq, Abbas N. (2013): Applying a SERVQUAL Model to Measure the Impact of Service Quality on Customer Loyalty among Local Saudi Banks in Riyadh. In *AJIBM* 03 (08), pp. 700–707. DOI: 10.4236/ajibm.2013.38079.

Al-Deehani, Mohammad T.; Aldeehani, M. T. (2017): Perceptions and predictions of service quality-customer satisfaction of conventional and Islamic banks in the GCC region. In *Journal of Research in Business, Economics and Management* 9 (1), pp. 1641–1656.

Alec Radford; Jeffrey Wu; Rewon Child; David Luan; Dario Amodei; Ilya Sutskever (2019): Language Models are Unsupervised Multitask Learners, pp. 1–9.

Alexander, Bethan; Kent, Anthony (2022): Change in technology-enabled omnichannel customer experiences in-store. In *Journal of Retailing and Consumer Services* 65, p. 102338. DOI: 10.1016/j.jretconser.2020.102338.

Aljedaani, Wajdi; Rustam, Furqan; Mkaouer, Mohamed Wiem; Ghallab, Abdullatif; Rupapara, Vaibhav; Washington, Patrick Bernard et al. (2022): Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry. In *Knowledge-Based Systems* 255, p. 109780. DOI: 10.1016/j.knosys.2022.109780.

Berger, Jonah; Packard, Grant; Boghrati, Reihane; Hsu, Ming; Humphreys, Ashlee; Luangrath, Andrea et al. (2022): Marketing insights from text analysis. In *Mark Lett* 33 (3), pp. 365–377. DOI: 10.1007/s11002-022-09635-6.

Blocker, Christopher P.; Flint, Daniel J.; Myers, Matthew B.; Slater, Stanley F. (2011): Proactive customer orientation and its role for creating customer value in global markets. In *JAMS* 39 (2), pp. 216–233. DOI: 10.1007/s11747-010-0202-9.

Brown, Tom B.; Mann, Benjamin; Ryder, Nick; Subbiah, Melanie; Kaplan, Jared; Dhariwal, Prafulla et al. (Eds.) (2020): Language Models are Few-Shot Learners. 34th Conference on Neural Information Processing Systems. Vancouver, Canada.

Büschken, Joachim; Allenby, Greg M. (2016): Sentence-Based Text Analysis for Customer Reviews. In *Marketing Science* 35 (6), pp. 953–975. DOI: 10.1287/mksc.2016.0993.

Cai, Ya-Jun; Lo, Chris K. Y. (2020): Omni-channel management in the new retailing era: A systematic review and future research agenda. In *International Journal of Production Economics* 229, p. 107729.

Chandra Mahapatra, Subas; Bellamkonda, Raja Shekhar (2023): Higher expectations of passengers do really sense: Development and validation a multiple scale-FliQual for air transport service quality. In *Journal of Retailing and Consumer Services* 70, p. 103162. DOI: 10.1016/j.jretconser.2022.103162.

Chen, Hsinchun; Chiang, Roger H. L.; Storey, Veda C. (2012): Business intelligence and analytics: From big data to big impact. In *MIS quarterly*, pp. 1165–1188.

Choi, Laee; Lawry, Charles A.; Kim, MiRan (2019): Contextualizing customer organizational citizenship behaviors: The changing nature of value cocreation and customer satisfaction across service settings. In *Psychology and Marketing* 36 (5), pp. 455–472. DOI: 10.1002/mar.21190.

Chopra, Sunil (2018): The Evolution of Omni-Channel Retailing and its Impact on Supply Chains. In *Transportation Research Procedia* 30, pp. 4–13. DOI: 10.1016/j.trpro.2018.09.002.

Chu, Mi-Na (2023): Assessing the Benefits of ChatGPT for Business: An Empirical Study on Organizational Performance. In *IEEE Access* 11, pp. 76427–76436. DOI: 10.1109/ACCESS.2023.3297447.

Cocco, Helen; De-Juan-Vigaray, María D. (2022): A typology of omnichannel retailer activities during the COVID-19 pandemic. In *IJRDM* 50 (8/9), pp. 1062–1094. DOI: 10.1108/IJRDM-10-2021-0506.

Cronin, J. Joseph; Taylor, Steven A. (1992): Measuring Service Quality: A Reexamination and Extension. In *Journal of Marketing* 56 (3), pp. 55–68. DOI: 10.1177/002224299205600304.

Cronin, J. Joseph; Taylor, Steven A. (1994): Servperf versus Servqual: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality. In *Journal of Marketing* 58 (1), pp. 125–131. DOI: 10.1177/002224299405800110.

Dabholkar, Pratibha A.; Thorpe, Dayle I.; Rentz, Joseph O. (1996): A measure of service quality for retail stores: Scale development and validation. In *JAMS* 24 (1), pp. 3–16. DOI: 10.1007/BF02893933.

Daher, Amal (2021): Old habits die hard. insights into the reasons why people don’t switch to online retail grocery shopping despite COVID-19 in Austria and Peru in 2020./Author Amal Daher, Andrea Zevallos. Universität Linz.

Devlin, Jacob; Chang, Ming-Wei; Lee, Kenton; Toutanova, Kristina (2019): BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (1), pp. 4171–4186. DOI: 10.18653/v1/N19-1423.

Dominici, Andrea; Boncinelli, Fabio; Gerini, Francesca; Marone, Enrico (2021): Determinants of online food purchasing: The impact of socio-demographic and situational factors. In *Journal of Retailing and Consumer Services* 60, p. 102473. DOI: 10.1016/j.jretconser.2021.102473.

Duan, Wenjing; Cao, Qing; Yu, Yang; Levy, Stuart (Eds.) (2013): Mining online user-generated content: using sentiment analysis technique to study hotel service quality. 46th Hawaii International Conference on System Sciences (pp. 3119-3128).

Eachempati, Prajwal; Srivastava, Praveen Ranjan; Kumar, Ajay; Muñoz de Prat, Javier; Delen, Dursun (2022): Can customer sentiment impact firm value? An integrated text mining approach. In *Technological Forecasting and Social Change* 174, p. 121265. DOI: 10.1016/j.techfore.2021.121265.

Falatouri, Taha; Brandtner, Patrick; Nasseri, Mehran; Darbanian, Farzaneh (2024a): Service quality dimensions in Austrian food retailing – a text mining approach for physical retail stores. In *The International Review of Retail, Distribution and Consumer Research*, pp. 1–36. DOI: 10.1080/09593969.2024.2371456.

Falatouri, Taha; Hrušecká, Denisa; Fischer, Thomas (2024b): Harnessing the Power of LLMs for Service Quality Assessment From User-Generated Content. In *IEEE Access* 12, pp. 99755–99767. DOI: 10.1109/ACCESS.2024.3429290.

Fransoo, Jan C.; Zhao, Lei; Martinelli, José Larco (2019): Omnichannel and Traditional Retail: Platforms to Seamlessly Connect Retail, Service, and Delivery. In Santiago Gallino, Antonio Moreno (Eds.): Operations in an Omnichannel World, vol. 8. Cham: Springer International Publishing (Springer Series in Supply Chain Management), pp. 341–353.

Frasquet, Marta; Ieva, Marco; Ziliani, Cristina (2021): Online channel adoption in supermarket retailing. In *Journal of Retailing and Consumer Services* 59, p. 102374. DOI: 10.1016/j.jretconser.2020.102374.

Galante, N.; López, E.; Monroe, S. (2013): The Future of Online Grocery in Europe. McKinsey & Company: Available at McKinsey: www. mckinsey. com [Accessed 1 November 2017].

Gao, Mengjia; Huang, Lin (2021): Quality of channel integration and customer loyalty in omnichannel retailing: The mediating role of customer engagement and relationship program receptiveness. In *Journal of Retailing and Consumer Services* 63, p. 102688.

Gronroos, Christian (1990): Service Management: A Management Focus for Service Competition. In *International journal of service industry management* 1 (1), pp. 6–14. DOI: 10.1108/09564239010139125.

Helm, Sabrina; Kim, Soo Hyun; van Riper, Silvia (2020): Navigating the 'retail apocalypse': a framework of consumer evaluations of the new retail landscape. In *Journal of Retailing and Consumer Services* 54.

Holmlund, Maria; van Vaerenbergh, Yves; Ciuchita, Robert; Ravald, Annika; Sarantopoulos, Panagiotis; Ordenes, Francisco Villarroel; Zaki, Mohamed (2020): Customer experience management in the age of big data analytics: A strategic framework. In *Journal of Business Research* 116, pp. 356–365. DOI: 10.1016/j.jbusres.2020.01.022.

Hossain, Mohammad Alamgir; Quaddus, Mohammed (2012): Expectation–Confirmation Theory in Information System Research: A Review and Analysis. In Yogesh K. Dwivedi, Michael R. Wade, Scott L. Schneberger (Eds.): Information Systems Theory, vol. 28. New York, NY: Springer New York (Integrated Series in Information Systems), pp. 441–469.

Hutto, C.; Gilbert, Eric (Eds.) (2014): VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. ICWSM (8).

Jain, Prachi; Aggarwal, Vijita Singh (2018): Developing a service quality scale in context of organized grocery retail of India. In *MD* 56 (9), pp. 1969–1990. DOI: 10.1108/MD-08-2017-0790.

Kaul, Subhashini (2007): Measuring Retail Service Quality: Examining Applicability of International Research Perspectives in India. In *Vikalpa: The Journal for Decision Makers* 32 (1), pp. 15–26. DOI: 10.1177/0256090920070102.

Khare, Arpita (2013): Retail service quality in small retail sector: the Indian experience. In *Facilities* 31 (5/6), pp. 208–222. DOI: 10.1108/02632771311307089.

Kojima, Takeshi; Gu, Shixiang Shane; Reid, Machel; Matsuo, Yutaka; Iwasawa, Yusuke (2022): Large Language Models are Zero-Shot Reasoners. In *Advances in neural information processing systems* 35, pp. 22199–22213.

Korfiatis, Nikolaos; Stamolampros, Panagiotis; Kourouthanassis, Panos; Sagiadinos, Vasileios (2019): Measuring service quality from unstructured data: A topic modeling application on airline passengers' online reviews. In *Expert Systems with Applications* 116, pp. 472–486. DOI: 10.1016/j.eswa.2018.09.037.

Kuijpers, Dymfke; Simmons, Virginia; van Wamelen, Jasper (2018): Reviving grocery retail: Six imperatives. In *December, available at: https://www. mckinsey. com/industries/retail/our-insights/reviving-groceryretail-six-imperatives (accessed 28 October 2019)*.

Le, Hoang Tran Phuoc Mai; Kim, Dongyoup; Park, Jungkun (2024): The way to generate customer citizenship behavior with customer experience. In *Journal of Retailing and Consumer Services* 76, p. 103608. DOI: 10.1016/j.jretconser.2023.103608.

Leippold, Markus (2023): Thus spoke GPT-3: Interviewing a large-language model on climate finance. In *Finance Research Letters* 53, p. 103617. DOI: 10.1016/j.frl.2022.103617.

Lemon, Katherine N.; Verhoef, Peter C. (2016): Understanding Customer Experience Throughout the Customer Journey. In *Journal of Marketing* 80 (6), pp. 69–96. DOI: 10.1509/jm.15.0420.

Lu, Lin; Xu, Pei; Wang, Yen-Yao; Wang, Yu (2023): Measuring service quality with text analytics: Considering both importance and performance of consumer opinions on social and non-social online platforms. In *Journal of Business Research* 169, p. 114298. DOI: 10.1016/j.jbusres.2023.114298.

Lyu, Zhiheng; Jin, Zhijing; Gonzalez Adauto, Fernando; Mihalcea, Rada; Schölkopf, Bernhard; Sachan, Mrinmaya (2024): Do LLMs Think Fast and Slow? A Causal Study on Sentiment Analysis. In Yaser Al-Onaizan, Mohit Bansal, Yun-Nung Chen (Eds.): Findings of the Association for Computational Linguistics: EMNLP 2024. Findings of the Association for Computational Linguistics: EMNLP 2024. Miami, Florida, USA. Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 9353–9372.

Manser Payne, Elizabeth; Peltier, James W.; Barger, Victor A. (2017): Omni-channel marketing, integrated marketing communications and consumer engagement. In *JRIM* 11 (2), pp. 185–197. DOI: 10.1108/JRIM-08-2016-0091.

Marchet, Gino; Melacini, Marco; Perotti, Sara; Rasini, Monica; Tappia, Elena (2018): Business logistics models in omni-channel: a classification framework and empirical analysis. In *International Journal of Physical Distribution & Logistics Management*.

Martin-Domingo, Luis; Martín, Juan Carlos; Mandsberg, Glen (2019): Social media as a resource for sentiment analysis of Airport Service Quality (ASQ). In *Journal of Air Transport Management* 78, pp. 106–115. DOI: 10.1016/j.jairtraman.2019.01.004.

Mejia, Jorge; Mankad, Shawn; Gopal, Anandasivam (2021): Service Quality Using Text Mining: Measurement and Consequences. In *M&SOM* 23 (6), pp. 1354–1372. DOI: 10.1287/msom.2020.0883.

Mou, Shandong; Robb, David J.; DeHoratius, Nicole (2018): Retail store operations: Literature review and research directions. In *European Journal of Operational Research* 265 (2), pp. 399–422.

Naseem, Usman; Razzak, Imran; Musial, Katarzyna; Imran, Muhammad (2020): Transformer based Deep Intelligent Contextual Embedding for Twitter sentiment analysis. In *Future Generation Computer Systems* 113, pp. 58–69. DOI: 10.1016/j.future.2020.06.050.

Nasseri, Mehran; Brandtner, Patrick; Zimmermann, Robert; Falatouri, Taha; Darbanian, Farzaneh; Obinwanne, Tobechi (2023): Applications of Large Language Models (LLMs) in Business Analytics – Exemplary Use Cases in Data Preparation Tasks. In Helmut Degen, Stavroula Ntoa, Abbas Moallem (Eds.): HCI International 2023 – Late Breaking Papers, vol. 14059. Cham: Springer Nature Switzerland (Lecture Notes in Computer Science), pp. 182–198.

Nilashi, Mehrbakhsh; Abumalloh, Rabab Ali; Minaei-Bidgoli, Behrouz; Abdu Zogaan, Waleed; Alhargan, Ashwaq; Mohd, Saidatulakmal et al. (2022): Revealing travellers' satisfaction during COVID-19 outbreak: Moderating role of service quality. In *Journal of Retailing and Consumer Services* 64, p. 102783. DOI: 10.1016/j.jretconser.2021.102783.

Open AI & GPT News (2023): 92% of Fortune 500, 100M using ChatGPT Weekly. With assistance of Open AI & GPT News. Available online at https://www.linkedin.com/pulse/92-fortune-500-100m-using-chatgpt-weekly-open-ai-gpt-news-pietc/, updated on 11/8/2023.

OpenAI (2023): Best practices for prompt engineering with OpenAI API. Available online at https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-openai-api, updated on 5/31/2023, checked on 5/31/2023.

Pagano, Stefano; Strumolo, Luigi; Michalk, Katrin; Schiegl, Julia; Pulido, Loreto C.; Reinhard, Jan et al. (2025): Evaluating ChatGPT, Gemini and other Large Language Models (LLMs) in orthopaedic diagnostics: A prospective clinical study. In *Computational and Structural Biotechnology Journal* 28, pp. 9–15. DOI: 10.1016/j.csbj.2024.12.013.

Pantano, Eleonora; Viassone, Milena (2015): Engaging consumers on new integrated multichannel retail settings: Challenges for retailers. In *Journal of Retailing and Consumer Services* 25, pp. 106–114. DOI: 10.1016/j.jretconser.2015.04.003.

Parasuraman, Ananthanarayanan; Zeithaml, Valarie A.; Berry, L. (1988): SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. In *1988* 64 (1), pp. 12–40.

Parasuraman, Ananthanarayanan; Zeithaml, Valarie A.; Berry, Leonard L. (1985): A Conceptual Model of Service Quality and Its Implications for Future Research. In *Journal of Marketing* 49 (4), pp. 41–50. DOI: 10.1177/002224298504900403.

Parasuraman, Ananthanarayanan; Zeithaml, Valarie A.; Malhotra, Arvind (2005): ES-QUAL: A multiple-item scale for assessing electronic service quality. In *Journal of service research* 7 (3), pp. 213–233.

Park, Jeongeun; Yang, Donguk; Kim, Ha Young (2023): Text mining-based four-step framework for smart speaker product improvement and sales planning. In *Journal of Retailing and Consumer Services* 71, p. 103186. DOI: 10.1016/j.jretconser.2022.103186.

Park, Joonyong; Kim, Renee B. (2022): Importance of offline service quality in building loyalty of OC service brand. In *Journal of Retailing and Consumer Services* 65, p. 102493. DOI: 10.1016/j.jretconser.2021.102493.

Piotrowicz, Wojciech; Cuthbertson, Richard (2014): Introduction to the Special Issue Information Technology in Retail: Toward Omnichannel Retailing. In *International Journal of Electronic Commerce* 18 (4), pp. 5–16. DOI: 10.2753/JEC1086-4415180400.

Prakash, Gyan (2019): Understanding service quality: insights from the literature. In *JAMR* 16 (1), pp. 64–90. DOI: 10.1108/JAMR-01-2018-0008.

Puyt, Richard W.; Madsen, Dag Øivind (2024): Evaluating ChatGPT-4's historical accuracy: a case study on the origins of SWOT analysis. In *Front. Artif. Intell.* 7, p. 1402047. DOI: 10.3389/frai.2024.1402047.

Rigby, Darrell (2011): The future of shopping. In *Harvard business review* 89 (12), pp. 65–76.

Rosenmayer, Anneliese; McQuilken, Lisa; Robertson, Nichola; Ogden, Steve (2018): Omni-channel service failures and recoveries: refined typologies using Facebook complaints. In *Journal of Services Marketing*.

Rula, Anisa; D'Souza, Jennifer (2023): Procedural Text Mining with Large Language Models. In Brent Venable, Daniel Garijo, Brian Jalaian (Eds.): Proceedings of the 12th Knowledge Capture Conference 2023. K-CAP '23: Knowledge Capture Conference 2023. Pensacola FL USA, 05 12 2023 07 12 2023. New York, NY, USA: ACM, pp. 9–16.

Saghiri, Soroosh Sam; Bernon, Michael; Bourlakis, Michael; Wilding, Richard (2018): Omni-channel logistics special issue. In *International Journal of Physical Distribution & Logistics Management* 48 (4), pp. 362–364. DOI: 10.1108/IJPDLM-05-2018-361.

Shen, Xiao-Liang; Li, Yang-Jun; Sun, Yongqiang; Wang, Nan (2018): Channel integration quality, perceived fluency and omnichannel service usage: The moderating roles of internal and external usage experience. In *Decision Support Systems* 109, pp. 61–73.

Slack, Neale; Singh, Gurmeet; Sharma, Shavneet (2020): The effect of supermarket service quality dimensions and customer satisfaction on customer loyalty and disloyalty dimensions. In *International Journal of Quality and Service Sciences* 12 (3), pp. 297–318. DOI: 10.1108/IJQSS-10-2019-0114.

Sousa, Rui; Voss, Christopher A. (2006): Service Quality in Multichannel Services Employing Virtual Channels. In *Journal of service research* 8 (4), pp. 356–371. DOI: 10.1177/1094670506286324.

Steven, Loria (2024): textblob Documentation. Release 0.18.0.post0. Available online at https://readthedocs.org/projects/textblob/downloads/pdf/dev/, updated on 6/12/2024.

Sun, Jimeng; Reddy, Chandan K. (2013): Big data analytics for healthcare. In Rayid Ghani, Ted E. Senator, Paul Bradley, Rajesh Parekh, Jingrui He, Robert L. Grossman et al. (Eds.): Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. KDD' 13: The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Chicago Illinois USA, 11 08 2013 14 08 2013. New York, NY, USA: ACM, p. 1525.

Tang, Xiaoyi; Chen, Hongwei; Lin, Daoyu; Li, Kexin (2024): Harnessing LLMs for multi-dimensional writing assessment: Reliability and alignment with human judgments. In *Heliyon* 10 (14), e34262. DOI: 10.1016/j.heliyon.2024.e34262.

Tannady, Hendy; Nurprihatin, Filscha; Hartono, Hendy (2018): SERVICE QUALITY ANALYSIS OF TWO OF THE LARGEST RETAIL CHAINS WITH MINIMART CONCEPT IN INDONESIA. In *Business: Theory and Practice* 19 (0), pp. 177–185. DOI: 10.3846/btp.2018.18.

Tong, Zhou; Zhang, Haiyi: A text mining research based on LDA topic modelling. In : International conference on computer science, engineering and information technology, pp. 201–210.

Udokwu, Chibuzor; Darbanian, Farzaneh; Falatouri, Taha Nejad; Brandtner, Patrick (2020): Evaluating Technique for Capturing Customer Satisfaction Data in Retail Supply Chain. In : 2020 The 4th International Conference on E-commerce, E-Business and E-Government. ICEEG 2020: 2020 The 4th International Conference on E-commerce, E-Business and E-Government. Arenthon France, 17 06 2020 19 06 2020. New York, NY, USA: ACM, pp. 89–95.

Ulkhaq, M. Mujiya; Rabbani, Murad; Wibowo, Agus T.; Rachmania, Bedietra A. (2017): Assessing electronic service quality using ES-QUAL and E-RecS-QUAL scales. In *ITMSOC Transactions on Innovation & Business Engineering* 2 (1), pp. 20–26.

Uzir, Md. Uzir Hossain; Al Halbusi, Hussam; Thurasamy, Ramayah; Thiam Hock, Rodney Lim; Aljaberi, Musheer A.; Hasan, Najmul; Hamid, Mahmud (2021): The effects of service quality, perceived value and trust in home delivery service personnel on customer satisfaction: Evidence from a developing country. In *Journal of Retailing and Consumer Services* 63, p. 102721. DOI: 10.1016/j.jretconser.2021.102721.

Vaswani, Ashish; Shazeer, Noam; Parmar, Niki; Uszkoreit, Jakob; Jones, Llion; Gomez, Aidan N. et al. (2017): Attention Is All You Need. In *Advances in neural information processing systems*, pp. 5998–6008. DOI: 10.48550/arXiv.1706.03762.

Vázquez, Rodolfo; Rodríguez-Del Bosque, Ignacio A.; Ma Díaz, Ana; Ruiz, Agustín V. (2001): Service quality in supermarket retailing: identifying critical service experiences. In *Journal of Retailing and Consumer Services* 8 (1), pp. 1–14. DOI: 10.1016/S0969-6989(99)00018-1.

Verhoef, Peter C.; Kannan, P. K.; Inman, J. Jeffrey (2015): From Multi-Channel Retailing to Omni-Channel Retailing. In *Journal of Retailing* 91 (2), pp. 174–181. DOI: 10.1016/j.jretai.2015.02.005.

Vijaymeena, M.k; Kavitha, K (2016): A Survey on Similarity Measures in Text Mining. In *Machine Learning and Applications: An International Journal* 3 (1), pp. 19–28. DOI: 10.5121/mlaij.2016.3103.

Villeneuve, Hannah; O'Brien, William (2020): Listen to the guests: Text-mining Airbnb reviews to explore indoor environmental quality. In *Building and Environment* 169, p. 106555.

Voss, Christopher A.; Roth, Aleda V.; Rosenzweig, Eve D.; Blackmon, Kate; Chase, Richard B. (2004): A Tale of Two Countries' Conservatism, Service Quality, and Feedback on Customer Satisfaction. In *Journal of service research* 6 (3), pp. 212–230. DOI: 10.1177/1094670503260120.

Wan, Mengting; Safavi, Tara; Jauhar, Sujay Kumar; Kim, Yujin; Counts, Scott; Neville, Jennifer et al. (2024): TnT-LLM: Text Mining at Scale with Large Language Models. In Ricardo Baeza-Yates, Francesco Bonchi (Eds.): Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. KDD '24: The 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Barcelona Spain, 25 08 2024 29 08 2024. New York, NY, USA: ACM, pp. 5836–5847.

Wang, Yue; Luo, Linkai; Liu, Hai (2022): Bridging the Semantic Gap Between Customer Needs and Design Specifications Using User-Generated Content. In *IEEE Trans. Eng. Manage.* 69 (4), pp. 1622–1634. DOI: 10.1109/TEM.2020.3021698.

Wankhade, Mayur; Rao, Annavarapu Chandra Sekhara; Kulkarni, Chaitanya (2022): A survey on sentiment analysis methods, applications, and challenges. In *Artif Intell Rev* 55 (7), pp. 5731–5780. DOI: 10.1007/s10462-022-10144-1.

Wolfinbarger, Mary; Gilly, Mary C. (2003): eTailQ: dimensionalizing, measuring and predicting etail quality. In *Journal of Retailing* 79 (3), pp. 183–198. DOI: 10.1016/S0022-4359(03)00034-4.

Wongpakaran, Nahathai; Wongpakaran, Tinakon; Wedding, Danny; Gwet, Kilem L. (2013): A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. In *BMC medical research methodology* 13, p. 61. DOI: 10.1186/1471-2288-13-61.

Xu, Xun; Jackson, Jonathan E. (2019): Examining customer channel selection intention in the omni-channel retail environment. In *International Journal of Production Economics* 208, pp. 434–445. DOI: 10.1016/j.ijpe.2018.12.009.

Zhang, Kem Z.K.; Zhao, Sesia J.; Cheung, Christy M.K.; Lee, Matthew K.O. (2014): Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. In *Decision Support Systems* 67, pp. 78–89. DOI: 10.1016/j.dss.2014.08.005.

Zhang, Min; He, Xueping; Qin, Fang; Fu, Wenbiao; He, Zhen (2019): Service quality measurement for omni-channel retail: scale development and validation. In *Total Quality Management & Business Excellence* 30 (sup1), S210-S226. DOI: 10.1080/14783363.2019.1665846.

Zhou, Tong; Zhang, Haiyi (Eds.) (2016): A text mining research based on LDA topic modelling. International conference on computer science, engineering and information technology.

LIST OF PUBLICATIONS BY THE AUTHOR

1. FALATOURI, Taha; HRUŠECKÁ, Denisa; FISCHER, Thomas. Harnessing the Power of LLMs for Service Quality Assessment from User-Generated Content. IEEE Access, 2024. doi.org/10.1109/ACCESS.2024.3429290
2. Falatouri, T., Brandtner, P., Nasseri, M., & Darbanian, F. (2024). Service quality dimensions in Austrian food retailing–a text mining approach for physical retail stores. The International Review of Retail, Distribution and Consumer Research, 1-36.http://dx.doi.org/10.1080/09593969.2024.2371456
3. Darbanian, F., Brandtner, P., Falatouri, T., & Nasseri, M. (2024). Data Analytics in Supply Chain Management: A State-of-the-Art Literature Review. Operations and Supply Chain Management: An International Journal, 17(1), 1-31. http://doi.org/10.31387/oscm0560411
4. Nejad Falatouri Moghaddam, Taha. "Conceptual model to analyze the influence of customer participation and value co-creation in social media on brand performance." 15th Annual International Bata Conference for Ph. D. Students and Young Researchers (DOKBAT). Tomas Bata Univ Zlin, 2019. https://doi.org/10.7441/dokbat.2019.075
5. Ladislav Burita, Taha Falatouri, @Analysis of Data from the Social Media." CERes Journal, Volume 5, Issue 1, 2019. Faculty of Management Science and Informatics, University of Zilina
6. Falatouri Taha. "SUPPLY CHAIN EXCLUSIVITY IN OMNICHANNEL RETAIL." 12th International Scientific Conference Karviná Ph.D. Conference on Business and Economics,28-34

CURRICULUM VITAE

TAHA **FALATOURI**

Experience

2023 - Current

Head Of Research Field - Advanced Industrial Analytics, UNIVERSITY OF APPLIED SCIENCE UPPER AUSTRIA

- Teaching “business informatics” course and “Statistics” course to master students
- interpret business questions into data questions
- Contributed in preparing and publishing scientific journal and conference papers
- Supervising research projects

2019 – 2023

Research Associate, UNIVERSITY OF APPLIED SCIENCE UPPER AUSTRIA

- Implemented various forecasting techniques of stores using SPSS Modeler and Python
- Streamlined feature selection for the model to predict the daily sale of the stores
- Performed sentiment analysis to +1M customer text reviews of +2000 store branches
- Conducted market basket analysis on cashier data of +60 grocery stores using SPSS Modeler
- Conducted in-depth qualitative research, reviewed articles, and extracted key insights for data

2018 – 2019

Head Of Sales Analytics, OFOQ KOOROSH SUPER MARKET CHAIN

- Managed team to Devised store performance KPIs across company in collaboration with cross-functional teams
- Created store clustering using R to analyses stores performance on two level of region based and KPI based
- Prepared and presented monthly analytical reports for 8 regions and 2000 stores.

2012 – 2015

Strategic Expert, GOLRANG INDUSTRIAL GROUP

- Identified global industry trends and new market opportunities.
- Led strategic planning, business development, and deal evaluation.
- Developed product categorization for 3000+ SKUs in a Master Data project.

- Prepared analytical reports for 9+ subsidiaries, presented to C-level executives.
- Analyzed markets and business cases to identify growth opportunities.
- Collaborated with internal teams and external partners to achieve business goals.
- Managed and trained analysts, ensuring cross-functional communication.
- Led the launch of Iran's largest cinema complex as Executive Director.

Education

2019 – current

PHD. Industrial Engineering, TOMAS BATA UNIVERSITY, ZLIN, CZECH REPUBLIC

2009 – 2011

MSC. IT MANAGEMENT, TARBIAT MODARES UNIVERSITY, TEHRAN, IRAN

2004 – 2008

BSC. Industrial Management, ALLAMEH TABATABA'I UNIVERSITY, TEHRAN, IRAN

PROJECTS

- PREVAIL - Predictive Analytics and Data-Driven Intelligence in Value Networks 2023 - 2028
- NEXT - Patern recognition and Foresight 2019 - 2023.
- Predictive Maintenance Vaillant Group, Austria 2021-2022 .
- Strategic planning Arian Chimia Tech 2018,Iran.
- Strategic development planning on cellulose industries Marinasun Group,Iran 2017.
- IGA/FaMe/2020/002 The Impact of digital transformation on customer behavior and firm's sustainable performance, Czech Republic 2020-2021.
- Maintenance schedule planning Alps, Czech Republic 2019

SKILLS

- Programming: SQL, Python, R
- Reporting: Google Analytics, PowerBI
- Statistical analysis: SPSS, SAS
- Project management: MS Project

Taha Falatouri

Customer Experience analysis based on big data analysis in grocery retail

Analýza zákaznické zkušenosti na základě analýzy velkých dat v maloobchodě s potravinami

Doctoral Thesis Summary

Published by: Tomas Bata University in Zlín,
nám. T. G. Masaryka 5555, 760 01 Zlín.

Edition: published electronically
1st edition

Typesetting by: Taha Falatouri
This publication has not undergone any proofreading or editorial review.

Publication year: 2025

ISBN 978-80-7678-334-8