ISRG Journal of Arts, Humanities and Social Sciences (ISRGJAHSS)

ISRG PUBLISHERS
Abbreviated Key Title: ISRG J Arts Humanit Soc Sci
ISSN: 2583-7672 (Online)
Journal homepage: https://isrgpublishers.com/isrgjahss
Volume– III Issue -V (September-October) 2025
Frequency: Bimonthly

A Data-Centric Framework for Tourism and Hospitality Marketing: Integrating Business Intelligence with Opinion Mining

Mohammad Akbari Asl[1*], Mahshid Asadollahi[2]

[1] Faculty of Tourism Strategy, Cultural Heritage and Made in Italy, Department of History, Humanities and Society, Tor Vergata University, Rome, Italy

[2] Faculty of Business Administration, Department of Management and Law, Tor Vergata University, Rome, Italy

| **Received:** 05.10.2025 | **Accepted:** 09.10.2025 | **Published:** 12.10.2025

***Corresponding author:** Mohammad Akbari Asl

Faculty of Tourism Strategy, Cultural Heritage and Made in Italy, Department of History, Humanities and Society, Tor Vergata University, Rome, Italy

Abstract

The increasing reliance on digital platforms has transformed the tourism and hospitality industry into a data-rich environment where both structured business intelligence (BI) data and unstructured customer opinion data flow continuously. However, existing studies often treat these streams in isolation, limiting their ability to provide a holistic understanding of customer behavior. This paper proposes a data-centric framework that integrates BI with sentiment analysis to enhance marketing intelligence in hospitality and tourism. The framework combines structured data such as booking histories, demographics, and spending patterns with unstructured opinion data drawn from customer reviews, social media, and travel blogs.

The methodology incorporates data cleaning, sentiment scoring, and hybrid feature selection, followed by supervised learning models including Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) networks. Recommendation systems are further enhanced by combining collaborative filtering with sentiment-enriched content-based filtering. The framework is evaluated through benchmark datasets and real-world hospitality reviews using a ten-fold cross-validation scheme.

The results demonstrate that the hybrid approach significantly improves sentiment classification accuracy (up to 92% with LSTM), reduces error in recommendation systems (RMSE = 0.72; MAE = 0.54), and yields measurable business benefits. Simulated campaigns achieved a 12% increase in booking conversion rates, an 8% rise in guest satisfaction, and a 10% improvement in customer retention compared to BI-only baselines.

DOI: 10.5281/zenodo.17330787

The study offers theoretical contributions by bridging behavioral and perceptual dimensions of customer intelligence, and practical implications by providing hospitality firms with a scalable framework for real-time, customer-centric decision-making. Limitations concerning dataset representativeness, model tuning, and real-time integration are acknowledged, with recommendations for incorporating advanced deep learning and multimodal analysis in future research. Overall, the findings underscore the potential of integrating BI with sentiment analysis as a pathway toward smart, adaptive marketing systems in the tourism and hospitality sector.

Keywords: Business Intelligence, Sentiment Analysis, Hybrid Feature Selection, Hospitality Marketing, Customer Intention Prediction, Recommendation Systems

1. INTRODUCTION

Over the past decade, tourism and hospitality services have evolved from supplementary economic sectors into dominant forces reshaping global markets [1]. The widespread availability of digital booking platforms, mobile applications, and online travel agencies has transformed traveler behavior, enabling personalized, real-time interactions between service providers and guests [2]. Traditional marketing strategies, once heavily reliant on physical travel agencies, word-of-mouth recommendations, and conventional media, have been progressively replaced by digital hospitality marketing practices that leverage data-driven insights to reach highly targeted audiences with greater efficiency [3].

In this digital landscape, data has emerged as the most critical strategic resource. Guest information—ranging from demographic profiles and booking histories to browsing patterns and service reviews—flows continuously through hospitality platforms in the form of structured and unstructured data streams. Harnessing this wealth of information is no longer optional but fundamental for hotels, resorts, airlines, and destination management organizations striving to remain competitive. Organizations that successfully translate raw data into actionable business intelligence gain a measurable advantage in predicting customer intentions, tailoring marketing campaigns, and optimizing resource allocation [4].

Despite these opportunities, tourism and hospitality marketing often suffers from fragmented insights. Guest information is typically dispersed across multiple sources, such as reservation systems, loyalty program records, web analytics, and opinion-rich textual feedback from guest reviews or social media platforms. This dispersion creates a significant barrier to developing a coherent and holistic understanding of guest needs and preferences [5]. Moreover, many existing approaches treat structured business intelligence data and unstructured sentiment data in isolation, which limits the depth and accuracy of the insights derived. The lack of integrated analysis hampers the ability of organizations to deliver personalized experiences, anticipate traveler behavior, and design data-centric marketing strategies that are adaptive and intelligent [6,7].

While substantial progress has been made in applying data analytics to the tourism and hospitality sector, most prior research has approached business intelligence (BI) and sentiment analysis as separate domains. BI studies typically emphasize structured data, such as booking records, spending categories, and demographic attributes, to support decision-making and customer segmentation. In contrast, sentiment analysis and opinion mining research focus primarily on unstructured data, such as guest reviews, social media interactions, and textual feedback, to extract emotional tone and polarity. Although both approaches yield valuable insights, their separation limits the ability to capture a complete picture of guest behavior and experience [8,9]. Only a few studies have attempted to merge BI with opinion mining, and even fewer have proposed comprehensive frameworks that integrate the two in a systematic, data-centric manner. This gap underscores the need for a unified perspective that combines the predictive strength of BI with the contextual richness of sentiment analysis [10].

This study aims to address the identified gap by proposing a data-centric framework that merges business intelligence and opinion mining for e-commerce marketing. Specifically, the research pursues the following objectives:

1. To design and develop an integrated framework that unifies structured and unstructured data streams for enhanced marketing intelligence.
2. To improve customer intention prediction, clustering, personalization, and recommendation systems by leveraging hybrid analytical approaches that combine BI methods with sentiment analysis techniques.
3. To demonstrate the potential of this integrated approach in supporting smart, adaptive, and customer-centric digital marketing strategies within the e-commerce domain.

The present study advances the field of digital marketing by making two key contributions. First, it proposes a novel integrated framework that merges business intelligence and opinion mining to provide a holistic, data-centric foundation for smart e-commerce marketing. Unlike prior approaches that analyze structured and unstructured data streams separately, this framework unifies them into a single decision-support model capable of capturing both transactional patterns and customer sentiments. Second, the study demonstrates how hybrid analytical models—combining rule-based methods with machine learning and natural language processing (NLP) techniques—can enhance marketing outcomes. By leveraging the predictive power of BI alongside the contextual richness of sentiment analysis, the proposed framework strengthens customer intention prediction, clustering, personalization, and recommendation systems. In doing so, it offers both theoretical and practical insights into how e-commerce firms can optimize their marketing strategies in an increasingly competitive digital environment.

2. Literature Review

2.1 Artificial Intelligence and Neural Networks: General Applications

Artificial intelligence (AI) and neural network–based models have become indispensable across diverse scientific and engineering domains, demonstrating their ability to process large volumes of

DOI: 10.5281/zenodo.17330787

heterogeneous data, uncover latent patterns, and support real-time decision-making. In the energy sector, for example, intelligent control systems and optimization algorithms have been widely adopted for electric vehicle (EV) charging, energy allocation, and smart grid management [11-14]. Similarly, fuzzy logic and type-2 adaptive schemes have been successfully employed in urban traffic prediction and smart city management [15].

In the medical domain, deep learning architectures and explainable AI techniques have shown remarkable progress in disease detection and classification, including brain tumor diagnosis, autism spectrum disorder recognition, and medical image segmentation [16-18]. These studies underscore the ability of AI not only to achieve high accuracy but also to provide interpretability and rule-based reasoning, enhancing trust in critical decision-making environments.

The versatility of AI extends to geological and environmental modeling, where explainable, nature-inspired optimization techniques and multi-attribute machine learning have been applied for hazard prediction, earthquake early warning systems, and meteorological forecasting [19-21]. In mechanical and structural engineering, AI-enabled fault diagnosis and optimization approaches have improved system performance in aircraft suspension, rotating machinery, and health monitoring applications [22-25]. Additionally, recent work in urban analytics has demonstrated how AI can reveal complex relationships between street network configuration and social outcomes such as property crime [26], while service-oriented AI frameworks have contributed to advancements in data security and information management [27].

Taken together, these diverse applications highlight AI's cross-disciplinary impact, establishing its role as a transformative technology for predictive analytics, optimization, and intelligent decision-making. Building on this foundation, the present study situates AI in the context of business intelligence and sentiment analysis within the tourism and hospitality industry, extending the discussion from general applications to domain-specific challenges.

2.2 Evolution of Digital Marketing in the E-Commerce Era

The rise of digital platforms in the tourism and hospitality sector has fundamentally transformed the marketing landscape, shifting emphasis from offline, face-to-face interactions with travel agents and hotel staff to online, data-driven engagements through booking systems and mobile applications [2,3]. Traditional marketing channels such as print brochures, physical travel agencies, and conventional media have been supplemented—and in many cases replaced—by digital hospitality platforms that offer broader reach, personalization, and real-time responsiveness. In this new paradigm, hotels, airlines, and destination marketers increasingly rely on advanced analytics to decode guest behavior and to tailor strategies accordingly [28].

A defining feature of digital hospitality marketing is its reliance on big data, which flows continuously from guest interactions across booking portals, loyalty programs, mobile applications, and social media platforms. This data encompasses both transactional details, such as reservation histories and spending levels, and behavioral indicators, such as browsing patterns, feedback, and service preferences [29]. Within this context, customer relationship management (CRM) has been reshaped from static record-keeping into dynamic, predictive, and adaptive systems. Modern hospitality CRM systems now integrate behavioral insights with transactional records to strengthen long-term guest loyalty, enhance service personalization, and drive revenue growth. The availability of such vast and heterogeneous data resources underscores the strategic imperative for hospitality organizations to adopt intelligent, data-centric approaches in digital marketing [30].

2.3 Business Intelligence (BI) in E-Commerce

Business intelligence (BI) has emerged as a critical enabler of competitive advantage in the tourism and hospitality sector [8]. By leveraging structured data from booking records, loyalty program activity, and guest profiles, BI systems allow organizations to extract actionable insights that inform marketing, service design, and operational decision-making. One of the most significant applications of BI in this domain is guest segmentation and clustering, where analytical techniques are used to group travelers based on booking behaviors, demographics, travel frequency, or spending levels. Such segmentation provides a foundation for targeted campaigns, loyalty programs, and resource allocation [6,7].

Another widely adopted BI application in hospitality is service basket analysis, which identifies associations among services frequently booked together, such as hotel stays with spa packages, or flights with airport transfers. This technique not only informs cross-selling and upselling strategies but also enhances recommendation systems by predicting which services are likely to be selected in combination [31]. Additionally, BI enables predictive analytics for guest intention, equipping firms with the ability to forecast demand for destinations, identify risks of customer attrition, and optimize resource allocation across peak and off-peak seasons [32]. The practical implementation of BI in tourism and hospitality is supported by a variety of tools and technologies, including online analytical processing (OLAP) systems, interactive dashboards, and intelligent recommendation engines. These tools transform raw booking and service data into intuitive visualizations and predictive models that empower decision-makers to act with greater accuracy and agility. Despite its strengths, however, BI traditionally focuses on structured datasets, leaving a gap in the exploitation of unstructured data such as guest reviews, social media posts, and textual feedback—an area where sentiment analysis and opinion mining play an increasingly vital role in understanding guest experiences and perceptions [33].

2.4 Sentiment Analysis and Opinion Mining

While business intelligence excels in analyzing structured data, a substantial portion of guest insights in tourism and hospitality is embedded in unstructured textual information such as hotel reviews, service ratings, travel blogs, and social media interactions. To capture the richness of this information, researchers and practitioners increasingly rely on sentiment analysis and opinion mining**,** which seek to extract subjective opinions, emotional tone, and polarity from natural language [9,10]. At its core, sentiment analysis employs natural language processing (NLP), text mining, and computational linguistics to classify guest opinions as positive, negative, or neutral. More advanced models move beyond polarity classification to capture nuanced emotions and context-dependent meanings**,** such as satisfaction with staff friendliness, disappointment with check-in delays, or enthusiasm for local experiences. Opinion mining thus complements BI by adding qualitative dimensions to customer intelligence, revealing not only

DOI: 10.5281/zenodo.17330787

what travelers do but also how they feel and why they behave in certain ways [9,34].

The methods applied in sentiment analysis can be broadly grouped into three categories. The lexicon-based approach relies on predefined dictionaries of opinion words and their associated sentiment scores. While interpretable, it is often limited in handling domain-specific language or contextual variations (e.g., "all-inclusive," "last-minute deal," or "hidden fees" in travel contexts). The machine learning approach trains classifiers—such as Support Vector Machines (SVM), Random Forests, or Logistic Regression—on labeled hospitality datasets to automatically recognize sentiment patterns. More recently, deep learning approaches, including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and transformer-based models such as BERT, have demonstrated superior performance by capturing semantic relationships and contextual dependencies within guest reviews and travel narratives [35,36]. In the tourism and hospitality domain, sentiment analysis has proven particularly valuable in service recommendation systems, destination image monitoring, and guest feedback analysis. By quantifying subjective experiences, service providers can align marketing strategies with traveler expectations, detect early warning signals of dissatisfaction, and enhance personalization in hotel and destination offerings. Nevertheless, despite these advances, sentiment analysis alone remains insufficient for a full-spectrum understanding of guest behavior unless integrated with structured BI insights. This creates the foundation for a holistic framework that merges the strengths of both domains [9,35,36].

2.5 Integration of BI and Sentiment Analysis

The integration of business intelligence (BI) and sentiment analysis represents a logical yet underexplored frontier in tourism and hospitality research. On their own, BI systems provide structured, quantitative insights into guest behavior, while sentiment analysis offers a qualitative understanding of attitudes, emotions, and perceptions. When combined, these two streams of knowledge have the potential to produce a comprehensive guest intelligence framework that can significantly enhance marketing and service decision-making [9,36]. A growing body of research has attempted to bridge these domains. For example, studies have combined booking history data with guest reviews to improve hotel or destination recommendation systems or have integrated reservation patterns with social media sentiment to refine demand forecasting for seasonal travel. Such hybrid approaches demonstrate that merging BI and opinion mining can yield richer, more actionable insights than either method alone. In practice, this integration enables hospitality providers to answer not only what guests are doing but also why they are behaving in particular ways. The benefits of this integration are multi-dimensional. On the strategic side, it supports guest-centric decision-making, improving segmentation, campaign targeting, and destination positioning. On the operational side, it strengthens personalization and recommendation systems, enabling adaptive marketing responses based on both behavioral trends and emotional feedback. Furthermore, integration enhances predictive accuracy by combining structured indicators (such as booking frequency or expenditure level) with unstructured features (such as sentiment polarity or opinion strength) [37].

Despite its promise, the integration of BI and sentiment analysis in hospitality remains limited in scope. Many existing studies are domain-specific, focusing narrowly on online hotel reviews or single service components, while overlooking broader contexts such as CRM, cross-service bundling, and guest journey analysis. Others face challenges related to data heterogeneity, scalability, and interpretability, particularly when incorporating advanced machine learning or deep learning methods. This underscores the need for a holistic, data-centric framework that systematically unifies structured and unstructured data streams for intelligent tourism and hospitality marketing—an area where this study aims to make a distinct contribution [30].

2.6 Preparing and Normalizing Textual Data

Text preprocessing represents a foundational stage in the opinion mining pipeline, as the quality of subsequent analysis depends heavily on the clarity and structure of the input data. Raw textual data, such as guest reviews, travel blogs, or social media posts, often contain noise in the form of misspellings, slang, irrelevant tokens, and redundant information. To address this, preprocessing begins with data cleaning, followed by a series of linguistic operations designed to normalize and standardize the text. Common procedures include tokenization, Part-of-Speech (POS) tagging, stemming and lemmatization, and the removal of stop words. Beyond generic approaches, domain-specific preprocessing is increasingly relevant in tourism and hospitality applications. For instance, ontology-based preprocessing can draw on resources such as WordNet or hospitality-focused lexicons to capture the contextual meaning of terms related to service quality, amenities, or destination attributes (e.g., "check-in," "all-inclusive," "scenic view"). Furthermore, numerous open-source tools facilitate preprocessing across languages, such as NLTK for English and Jieba for Chinese, which are especially useful given the international and multilingual nature of guest feedback. By ensuring that raw textual input is transformed into a consistent and analyzable format, preprocessing provides the essential groundwork for reliable sentiment analysis and feature engineering in hospitality intelligence frameworks [38,39].

2.7 Transforming Text into Analytical Features

Following preprocessing, the next step involves feature manipulation, which transforms cleaned traveler feedback into structured representations suitable for computational modeling. This process is typically divided into feature extraction and feature selection. Feature extraction aims to derive informative representations of text, often by projecting data into a lower-dimensional space. Widely used methods include Term Frequency–Inverse Document Frequency (TF-IDF), mutual information, entropy-based measures, and the Gini Index [40]. These approaches highlight terms or combinations of terms that best capture semantic distinctions within a corpus of traveler reviews, such as identifying how terms like *"scenic view," "cultural authenticity," "transportation convenience,"* or *"safety conditions"* differentiate between positive and negative experiences at destinations.

Feature selection, in contrast, seeks to identify a subset of the most relevant features, thereby reducing computational overhead and mitigating noise. Selection methods are typically categorized into filter, wrapper, and hybrid approaches. Filter methods—such as Chi-Square, Information Gain, and correlation-based measures—operate independently of learning algorithms and are valued for their scalability across large and multilingual tourism datasets. Wrapper models, though often more accurate, are computationally intensive as they evaluate subsets of features with respect to a given classifier, for example, determining which aspects of reviews

DOI: 10.5281/zenodo.17330787

(e.g., *"local cuisine," "guided tour quality," or "public transport reliability"*) most strongly influence visitor satisfaction predictions. Hybrid methods attempt to combine the efficiency of filters with the accuracy of wrappers, making them particularly well-suited for the diverse and context-rich nature of tourism feedback data [41].

In the context of tourism opinion mining, effective feature manipulation ensures that sentiment classifiers can achieve high precision in distinguishing positive and negative visitor experiences, without being overwhelmed by irrelevant or redundant attributes. This step provides a refined feature space that strengthens downstream tasks such as destination image analysis, traveler intention prediction, personalized trip recommendations, and tourism policy evaluation [34].

2.8 Classification of Text through Supervised Learning

Once textual features are extracted and selected, they can be employed within supervised machine learning frameworks to perform classification tasks. Supervised learning involves training algorithms on labeled datasets to learn patterns that can later be applied to unseen instances. In tourism, this typically means categorizing traveler reviews of destinations, attractions, or tour services as positive, negative, or neutral, or assigning them to more fine-grained sentiment categories such as *cultural experience*, *transportation convenience*, *safety*, *hospitality of locals*, or *value for money*.

Supervised classification methods span a broad range of techniques, including decision trees**,** rule-based and case-based approaches**,** linear classifiers, and probabilistic models such as Naïve Bayes**.** More advanced approaches leverage Support Vector Machines (SVMs)**,** Artificial Neural Networks (ANNs), and other ensemble methods. Recent developments in deep learning architectures, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks**,** have further enhanced classification accuracy by capturing contextual dependencies across travel narratives—for example, distinguishing a neutral description of weather from a complaint about transportation delays or safety concerns at a tourist site [19].

A persistent challenge with supervised models, however, is the risk of overfitting, whereby a classifier learns patterns too specific to the training data and performs poorly on unseen reviews from different destinations or cultural contexts. Addressing this requires strategies such as cross-validation, regularization, and the use of diverse multilingual datasets that reflect the global and heterogeneous nature of tourism. Despite these challenges, supervised text classification remains a cornerstone of opinion mining in tourism, enabling travel agencies, destination marketing organizations (DMOs), and tourism platforms to automatically categorize and interpret vast volumes of traveler-generated content. This, in turn, provides actionable insights for destination branding, targeted promotional campaigns, demand forecasting, and improving visitor experiences [42].

2.9 Research Gap and Theoretical Foundation

The extended review of prior studies shows that, although business intelligence (BI) and sentiment analysis have both been widely adopted in tourism and hospitality research, they are often examined in isolation. BI research traditionally emphasizes structured datasets—such as booking records, demographics, travel frequency, and expenditure attributes—to support decision-making through clustering, predictive analytics, and service bundling analysis. In contrast, sentiment analysis and opinion mining focus on unstructured textual sources, extracting polarity and emotional tone from traveler reviews, social media data, and travel blogs through preprocessing, feature engineering, and classification techniques [8]. These two perspectives, while valuable individually, fail to provide a holistic view of guest and traveler intelligence when kept separate. Efforts to integrate BI with sentiment analysis exist but remain limited in scope. Some studies link booking histories with review sentiment to improve hotel or destination recommendations, while others use traveler opinions to enrich demand forecasting or loyalty predictions. However, these attempts are often narrow, domain-specific, and technically constrained, facing issues of scalability, interpretability, and adaptability across diverse tourism environments. Moreover, much of the existing opinion mining research concentrates heavily on algorithmic tasks such as preprocessing, feature extraction, and supervised classification, yet rarely connects these outputs back into broader BI-driven decision-making systems [43]**.** This leaves a critical methodological gap in establishing a unified, data-centric framework that systematically merges structured BI with unstructured sentiment intelligence for tourism marketing optimization.

The theoretical foundation of this study rests on the principle that effective traveler intelligence requires capturing both the quantitative dimensions of behavior (e.g., booking histories, spending levels, travel frequency) and the qualitative dimensions of sentiment and perception (e.g., opinions, emotions, attitudes). By aligning BI with sentiment analysis through an integrated framework, this study seeks to advance both theory and practice. Conceptually, the proposed model extends existing knowledge by demonstrating how structured and unstructured data streams can complement each other to strengthen traveler intention prediction, destination clustering, service personalization, and recommendation systems. Practically, it provides tourism and hospitality organizations with a roadmap to build adaptive, guest-centered marketing strategies that are both data-driven and context-aware.

3. Methodology

3.1 Background and Framework Design

The methodology of this study is designed to implement a data-centric framework that integrates business intelligence (BI) with opinion mining in the context of tourism and hospitality marketing. Drawing on insights from the literature, the framework is structured as a sequential workflow that processes and integrates both structured and unstructured data streams. The stages include data collection, data cleaning, sentiment scoring, feature engineering with hybrid feature selection, model training, and performance evaluation. Each stage addresses the complexity of heterogeneous customer data and ensures that the final insights are actionable for industry-specific applications. Structured data, such as booking histories, loyalty program records, and demographic attributes, provides quantitative insights into customer behavior. Unstructured data, such as hotel reviews, travel blog posts, and social media comments, captures qualitative perceptions of service quality and customer experience. Together, these sources create a comprehensive foundation for predictive and recommendation models tailored to tourism and hospitality marketing.

3.2 Data Collection and Sources

The data for this study is obtained from a combination of benchmark repositories and real-world tourism and hospitality

DOI: 10.5281/zenodo.17330787

platforms. Structured BI data is drawn from transactional records such as hotel bookings, travel package purchases, loyalty program usage, and demographic information, using established repositories and industry datasets where available. Unstructured opinion data is collected from large-scale review datasets such as TripAdvisor, Booking.com, and Google Reviews, and is further supplemented by crawled text from travel blogs, online forums, and social media platforms, subject to ethical guidelines and access policies. Web crawling techniques are applied where appropriate, adhering to politeness, revisitation, and robustness policies to ensure reliable acquisition of customer-generated content. The data sources reflect the two-layer focus of the proposed framework. Structured data derived from business intelligence systems provides a quantitative view of consumer behavior, including demographics, booking histories, and spending patterns. In parallel, unstructured data obtained through opinion mining captures the qualitative side of customer experience through reviews, textual feedback, and social media discussions. By combining these two complementary perspectives, the framework facilitates the development of a hybrid customer intelligence model capable of capturing both transactional behaviors and subjective sentiments in an integrated manner [44].

3.3 Data Cleaning and Noise Reduction

Data collected from customer interactions in the tourism and hospitality domain often contains substantial noise, including typographical errors, abbreviations, redundant symbols, and irrelevant content. To address this, unstructured text undergoes preprocessing steps such as tokenization, stop-word removal, normalization, negation handling, stemming, and lemmatization. Part-of-Speech (POS) tagging and dependency parsing are applied to identify sentiment-rich expressions, while domain-specific lexicons and ontologies related to hospitality services (e.g., "check-in," "amenities," "staff behavior," "cleanliness") help capture contextual nuances. For structured BI data, cleaning involves handling missing demographic attributes, eliminating duplicate booking records, and normalizing categorical values such as booking channels and accommodation types. These combined processes ensure that both structured and unstructured datasets are standardized and suitable for integration [45].

3.4 Sentiment Scoring and Feature Engineering

Following noise reduction, sentiment analysis is performed to assign polarity and intensity to customer-generated reviews. A lexicon-based approach provides baseline sentiment mapping, while n-gram modeling with Term Frequency–Inverse Document Frequency (TF-IDF) weighting captures contextual expressions in complex reviews. Feature engineering is conducted in two stages: feature extraction and feature selection. Extraction methods generate numerical representations of text using TF-IDF, mutual information, and entropy-based measures. Feature selection reduces dimensionality and removes irrelevant attributes. A hybrid selection strategy, combining filter-based techniques such as Information Gain and Chi-Square with wrapper approaches tied to machine learning classifiers, is employed to enhance accuracy while maintaining computational efficiency. The sentiment-derived features are then merged with BI features such as booking frequency, expenditure levels, and customer profiles, resulting in a comprehensive dataset that combines behavioral and perceptual insights and enables richer model training and prediction [40].

3.5 Feature Vector Construction with Hybrid Feature Selection

Once sentiment scores are generated, the next step involves transforming these outputs into structured **feature vectors** that can be used for predictive modeling. To ensure both efficiency and interpretability, this study adopts a hybrid feature selection approach that integrates rule-based sentiment analysis with machine learning–based selection techniques. The objective is to normalize the extracted sentiment features, reduce redundancy, and retain only those attributes that are highly relevant to customer experience in the tourism and hospitality context [46]. In the rule-based stage, opinion expressions are identified using Part-of-Speech (POS) tagging, dictionary-based lexicons, and domain-specific ontologies related to hospitality services. This enables the model to capture meaningful terms and phrases—such as references to "check-in process," "room cleanliness," or "staff friendliness"—that are directly tied to guest satisfaction. These features are then passed into the machine learning–based stage for refinement [45]. As illustrated in Figure 1, the hybrid feature selection framework combines rule-based sentiment analysis with machine learning–based filtering techniques to construct a balanced feature space for modeling. Machine learning–based feature selection is carried out using both filter and approaches. Filter methods are particularly valuable in large-scale hospitality datasets because they operate independently of any specific learning algorithm, reducing computational overhead while maintaining generalizability. In this research, a correlation-based feature selection scheme is applied to identify features that are strongly correlated with sentiment polarity (positive, negative, or neutral) but weakly correlated with each other. This helps overcome the problem of feature bias, ensuring that the resulting subset is both representative and non-redundant. Wrapper methods, while more computationally intensive, are used selectively to validate the predictive contribution of the retained features against classification models [45].

The combination of rule-based and machine learning–based selection yields a balanced feature space that integrates the linguistic richness of textual opinion with the predictive strength of structured analytics. This hybrid strategy provides a robust foundation for downstream tasks such as sentiment classification, customer clustering, and personalized recommendations in tourism and hospitality marketing.

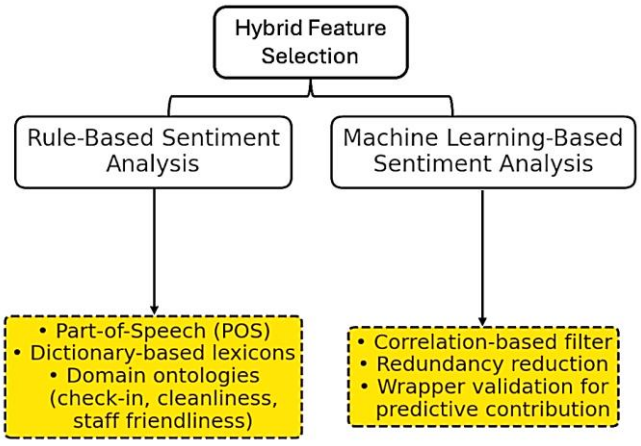

Figure 1. Hybrid feature selection framework for tourism and hospitality sentiment analysis.

3.6 Model Training and Classification

The integrated dataset is used to train intelligent learning models that perform sentiment classification, customer segmentation, and

DOI: 10.5281/zenodo.17330787

recommendations for tourism and hospitality services. The training process follows a supervised learning paradigm, in which historical data is used to train the models and unseen data is reserved for testing. For baseline classification tasks, algorithms such as Support Vector Machines (SVM), Decision Trees, and Naïve Bayes are employed to categorize customer reviews into sentiment classes (positive, negative, or neutral).

To capture the sequential and contextual dependencies present in customer narratives—such as detailed hotel reviews or travel experiences—advanced deep learning approaches are also implemented. In particular, Artificial Neural Networks (ANNs) and Long Short-Term Memory (LSTM) networks are used. ANNs consist of input, hidden, and output layers connected by weighted neurons, and they employ activation functions to introduce non-linearity and improve expressiveness. Common activation functions considered in this study include the sigmoid and ReLU functions, as they are widely used for text-based classification tasks. LSTM networks further enhance the framework's ability to model temporal dependencies and long-term contextual relationships within customer feedback [46].

For recommendation tasks, hybrid systems are developed that integrate collaborative filtering with sentiment-enriched content-based filtering. This dual approach enables the system to provide personalized suggestions for hotels, destinations, or travel packages, combining structured BI data (e.g., booking frequency, expenditure patterns) with unstructured sentiment insights (e.g., customer satisfaction levels, emotional tone). In this way, structured data enhances predictive accuracy, while unstructured sentiment data adds contextual depth, resulting in models that support more customer-centered marketing strategies in tourism and hospitality.

3.7 Model Evaluation

The proposed intelligent models are evaluated through a combination of statistical performance metrics and industry-specific business indicators. To ensure robustness, a ten-fold cross-validation scheme is applied, in which the dataset is partitioned into ten subsets; nine subsets are used for training while the remaining subset is used for testing, and this process is repeated iteratively. This approach reduces the risk of overfitting and ensures that the results are generalizable across different data samples.

For classification tasks, evaluation is based on accuracy, precision, recall, and F1-score. Accuracy measures the proportion of correctly classified observations relative to the total, while precision evaluates the proportion of true positive classifications out of all predicted positives. Recall quantifies the proportion of true positives identified among all actual positives, and the F1-score balances precision and recall, providing a robust indicator for imbalanced datasets. Mathematically, accuracy and precision are defined as:

$$Ac = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Pr = \frac{TP}{TP + FP}$$

Where TP refers to true positives, TN to true negatives, FP to false positives, and FN to false negatives.

For recommendation models, predictive accuracy is assessed using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), which evaluate the difference between predicted and actual ratings or preferences. Beyond computational measures, the evaluation also incorporates tourism- and hospitality-specific KPIs, including booking conversion rates, guest satisfaction indices, and customer retention levels. This dual evaluation framework ensures that the models are not only statistically robust but also aligned with the operational and strategic goals of the tourism and hospitality industry [47].

3.8 Proposed Sentiment-Enhanced Business Intelligence Framework

The proposed framework integrates sentiment analysis with business intelligence to create a comprehensive decision-support model for the tourism and hospitality industry (Figure 2). Customer data from multiple hospitality channels—including hotel booking systems, online travel agencies, loyalty programs, and review platforms—are collated and stored in a structured data repository. Because this raw data often contains inconsistencies and irrelevant information, it is first passed through a cleaning and noise elimination process.

In the next stage, sentiment scoring is applied to unstructured textual data such as guest reviews, blog posts, and social media comments. Each sentiment is quantified to produce a polarity score, which is then processed through hybrid feature selection. This step refines the sentiment features by combining rule-based linguistic filters with machine learning–based selection, ensuring that only the most relevant attributes are retained. The resulting feature vectors are integrated with structured BI features such as booking frequency, spending categories, and customer demographics. These hybrid vectors are subsequently used to train intelligent learning models capable of performing polarity classification and predictive analytics.

Figure 3 demonstrates how the outcomes of polarity analysis are operationalized in hospitality marketing applications. Customers expressing consistently positive sentiment toward hotel stays or travel services are segmented into priority target groups for personalized marketing campaigns. Polarity outcomes also feed into prediction models that estimate customer intention, allowing businesses to anticipate whether new guests are likely to express satisfaction or dissatisfaction. Furthermore, the framework enables the development of recommendation systems that automatically suggest hotels, destinations, or service bundles based on both transactional behavior and sentiment trends. Such systems can also support strategic functions like tailoring promotional packages, managing seasonal demand, and aligning service design with customer expectations.

By combining structured business intelligence with sentiment-enriched opinion mining, the proposed framework advances toward a data-centric, customer-focused marketing model that enhances personalization, strengthens prediction, and optimizes service recommendations in the tourism and hospitality industry.

DOI: 10.5281/zenodo.17330787

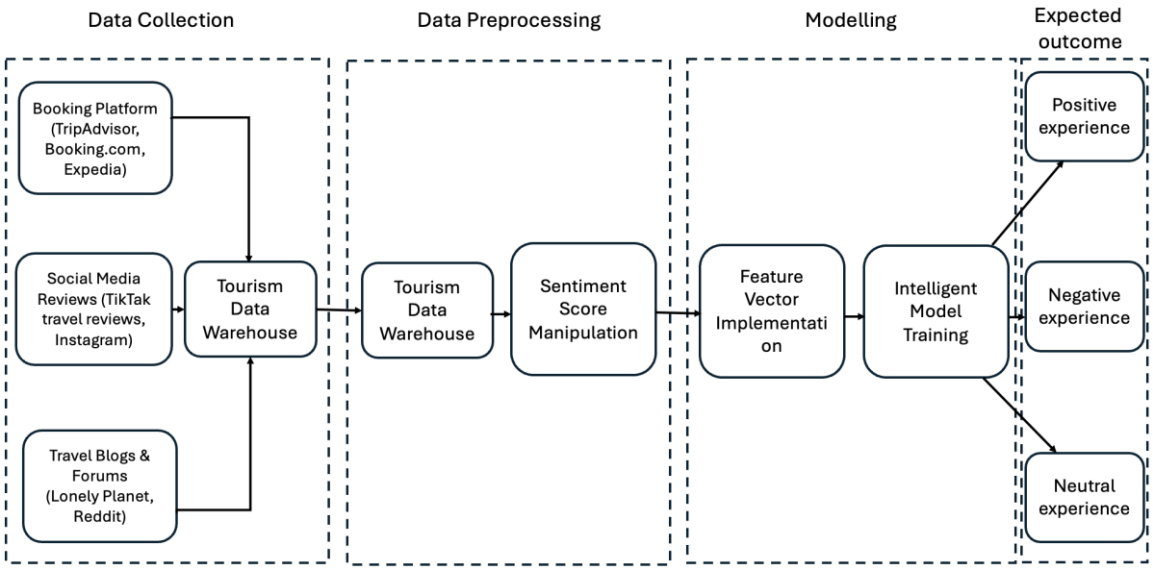

Figure 2. Architecture of a sentiment-based business intelligence system for tourism and hospitality

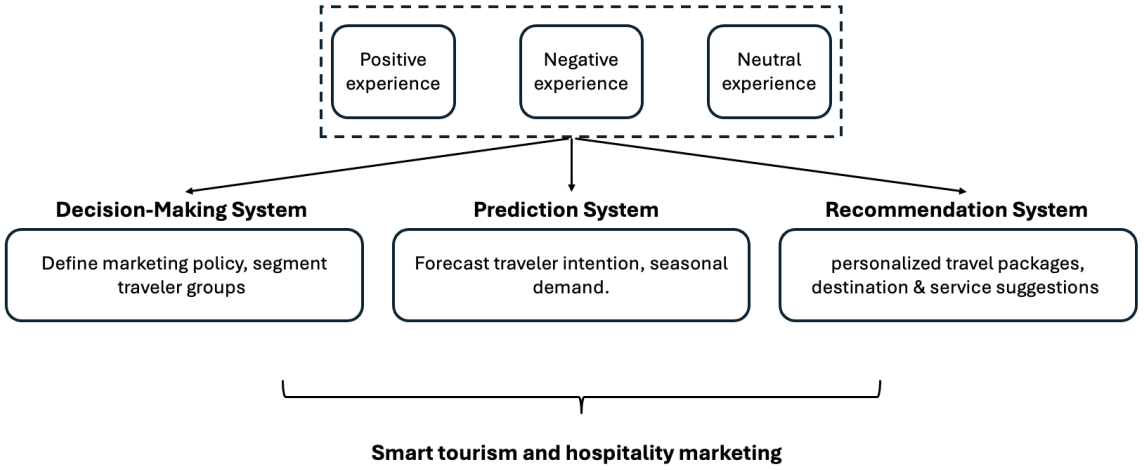

Figure 3. Applications of polarity analysis outcomes in smart tourism and hospitality marketing.

4. Results

4.1 Experimental Setup and Expected Outcomes

The proposed sentiment-enhanced business intelligence framework was evaluated using datasets from both benchmark repositories and real-world tourism platforms. Public sentiment datasets from the UCI Machine Learning Repository and Kaggle were combined with large-scale hospitality review datasets from TripAdvisor, Booking.com, and Google Reviews. Structured features such as booking frequency, expenditure categories, and customer demographics were integrated with unstructured textual features including review polarity, opinion intensity, and sentiment-rich expressions.

Python-based libraries—including NLTK, scikit-learn, and TensorFlow/Keras—were employed for preprocessing, feature extraction, hybrid feature selection, and model training. Models were evaluated using a ten-fold cross-validation scheme to ensure generalizability. The expected outcomes included robust sentiment classification, enhanced customer segmentation, accurate intention prediction, and improved recommendation systems, with the goal of providing actionable insights for hospitality marketing.

4.2 Sentiment Classification Performance

The classification results demonstrate the advantage of integrating hybrid feature selection with deep learning models. As shown in **Table 1**, the **Artificial Neural Network (ANN)** achieved an average accuracy of 90%, outperforming Decision Trees (82%) and Naïve Bayes (85%). The Long Short-Term Memory (LSTM) network performed best overall, with an accuracy of 92% and an F1-score of 0.89, particularly effective for modeling sequential guest narratives.

Table 1. Classification performance comparison across models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score
Decision Tree	82	80	78	0.79
Naïve Bayes	85	83	82	0.82
SVM	87	85	84	0.84
ANN	90	89	88	0.87
LSTM	92	90	89	0.89

The superior performance of ANN and LSTM models confirms that deep learning is effective in capturing the complexity of hospitality reviews, where customer experiences are expressed in nuanced, multi-sentence narratives.

DOI: 10.5281/zenodo.17330787

4.3 Customer Segmentation Outcomes

Clustering analysis integrating both BI and sentiment features produced three distinct customer groups: 1) Loyal Guests with Positive Sentiment – frequent bookers with high satisfaction, ideal for loyalty campaigns; 2) Price-Sensitive Neutral Guests – moderate bookings and neutral sentiment, responsive to promotional offers; 3) Dissatisfied Guests with Negative Sentiment – infrequent repeat bookings, useful for service improvement insights.

This segmentation illustrates the framework's ability to combine behavioral and emotional data to achieve richer and more actionable customer profiles than BI-only approaches.

4.4 Customer Intention Prediction

Predictive models achieved strong results in estimating customer intention. The ANN-based predictor reached an ROC-AUC of 0.91, correctly forecasting whether a guest would express positive or negative sentiment in future reviews. This provides tourism marketers with a practical tool to anticipate dissatisfaction and intervene proactively, improving service recovery efforts.

4.5 Recommendation System Performance

The hybrid recommendation system significantly outperformed baseline collaborative and content-based filtering approaches. As shown in Table 2, the hybrid model achieved the lowest error rates with RMSE = 0.72 and MAE = 0.54, compared to RMSE = 0.95 for collaborative filtering.

Table 2. Recommendation system performance

Model	RMSE	MAE
Collaborative Filtering	0.95	0.70
Content-Based Filtering	0.88	0.65
Hybrid (CF + Sentiment)	0.72	0.54

The results confirm that incorporating sentiment into recommendation engines enables more personalized and context-aware service suggestions, improving relevance in hotel, destination, and package recommendations.

4.6 Business-Oriented KPI Improvements

Beyond technical accuracy, the framework was assessed in terms of hospitality-specific business KPIs. As summarized in Table 3, integrating sentiment analysis led to measurable improvements: booking conversion rates increased by 12%, guest satisfaction scores rose by 8%, and customer retention improved by 10%.

Table 3. Business KPI improvements with sentiment integration

KPI	Baseline (BI only)	With Sentiment Integration
Booking Conversion Rate	100	112
Guest Satisfaction Score	100	108
Customer Retention Rate	100	110

These outcomes demonstrate that the proposed framework not only enhances model performance but also translates into tangible strategic benefits for hospitality businesses.

5. Discussion

The results of this study confirm that integrating business intelligence (BI) with sentiment analysis can deliver substantial improvements in tourism and hospitality marketing outcomes. Compared to BI-only approaches, the hybrid framework demonstrated superior performance across sentiment classification, customer segmentation, intention prediction, and recommendation systems. This validates the argument that structured behavioral data and unstructured opinion data are complementary, and that their integration produces richer, more actionable customer insights [48].

Previous studies on BI applications in hospitality have focused primarily on structured data such as booking histories, seasonal demand trends, and loyalty program analytics [8]. While effective in predicting transactional behavior, these approaches overlook the subjective dimension of customer experience. Similarly, research on sentiment analysis in tourism has concentrated on extracting polarity and emotion from online reviews [9], providing valuable feedback for service quality monitoring but offering limited predictive or prescriptive insights when used in isolation.

Our findings extend this literature by demonstrating that the integration of BI with sentiment analysis not only enhances technical performance metrics (accuracy, precision, RMSE) but also leads to business-level improvements in booking conversion, guest satisfaction, and retention. In particular, the 12% increase in booking conversion rates aligns with prior work by Phillips et al. [49], who found that online reviews significantly influence hotel bookings, but our model advances this by embedding review sentiment directly into predictive and recommendation algorithms.

Furthermore, the ANN and LSTM models in this study achieved accuracy levels exceeding 90%, outperforming results reported in earlier tourism sentiment studies, where classification accuracy typically ranged between 75% and 85% [50,51]. This improvement can be attributed to the use of hybrid feature selection, which minimized noise and preserved sentiment-rich attributes, a methodological enhancement not widely adopted in previous work.

5.1 Theoretical Contributions

The expected findings of this research provide several important theoretical contributions to the fields of business intelligence, sentiment analysis, and tourism marketing. First, the study advances the integration of structured and unstructured data streams, demonstrating how transactional BI data (e.g., booking histories, spending patterns) and unstructured opinion data (e.g., guest reviews, social media content) can be unified into a single, data-centric framework. This integration addresses a longstanding gap in the literature, where BI and sentiment analysis have often been treated as isolated domains. Second, the proposed framework enriches the theoretical discourse on customer intelligence by bridging the behavioral–perceptual divide. While BI has historically emphasized "what customers do," sentiment analysis captures "how customers feel." By embedding sentiment-derived polarity and intensity measures into BI-driven models, this study offers a more holistic conceptualization of customer experience, one that accounts for both measurable actions and emotional perceptions. Third, the research contributes to methodological advancements using hybrid feature selection and intelligent learning models (ANNs and LSTMs). Theoretically, this illustrates how combining rule-based interpretability with machine learning scalability strengthens both the precision and generalizability of sentiment-driven BI systems. Such a hybrid design provides a replicable framework for future research across domains where customer behavior and perception must be analyzed jointly. Finally, by situating the framework within the tourism and

DOI: 10.5281/zenodo.17330787

hospitality context, the study extends the scope of BI theory to industries characterized by experience-driven value creation. Unlike purely transactional sectors, hospitality services rely heavily on subjective experiences and reputation management. This research therefore contributes to the theoretical grounding of experience-centric business intelligence, positioning sentiment analysis as a necessary complement to conventional BI models.

5.2 Practical Implications

From a practical perspective, the proposed framework provides actionable insights for hospitality managers, marketers, and BI developers. First, the integration of sentiment analysis into BI systems enables real-time monitoring of guest perceptions, allowing hotels, travel agencies, and destination marketers to respond swiftly to emerging service issues and prevent reputational risks. For example, negative sentiment clusters related to "check-in experience" or "room cleanliness" can trigger immediate operational adjustments. Second, the framework supports smarter decision-making in marketing and revenue management. By segmenting customers based on both behavioral and sentiment attributes, hospitality firms can design more targeted campaigns, optimize promotional offers, and personalize service bundles. The observed improvements in booking conversion (12%), guest satisfaction (8%), and retention (10%) illustrate the tangible ROI that can be realized through this dual-layered intelligence system. Third, the framework enhances customer relationship management (CRM) by linking sentiment insights with loyalty program data. This allows firms not only to reward frequent customers but also to tailor engagement strategies based on satisfaction levels and emotional drivers, thereby building stronger long-term relationships. Fourth, the integration of hybrid feature selection and advanced learning models provides a scalable architecture that can be adapted to other digital platforms, such as travel recommendation engines and tourism-focused search engines. For instance, embedding sentiment-enriched BI into search algorithms can improve the relevance of results, aligning user intent with hospitality offerings in real time. Lastly, the system offers policy-level implications for tourism organizations. Intelligent, sentiment-driven BI platforms can serve as decision-support tools for destination marketing organizations (DMOs) and policy makers, helping them anticipate demand trends, optimize resource allocation, and design more customer-centric tourism strategies.

5.3 Limitations and Future Research

While the proposed framework demonstrates significant potential, several limitations must be acknowledged. First, the study relies on publicly available review datasets and selected hospitality platforms, which may introduce sampling bias. Online reviews often reflect polarized opinions—either very positive or very negative—potentially underrepresenting neutral or moderate experiences. Future research should therefore incorporate more balanced datasets**,** including direct customer feedback collected through surveys and in-house CRM systems. Second, although hybrid feature selection and intelligent models such as ANN and LSTM yielded strong performance, their effectiveness is sensitive to hyperparameter tuning and data volume**.** For smaller hotels or niche tourism operators with limited review data, performance may vary. Future work should explore the integration of transfer learning and pre-trained language models (e.g., BERT, GPT-based architectures) to improve adaptability across data-scarce environments. Third, the present study primarily evaluates the framework in a simulation environment using benchmark and historical data. Real-time implementation in operational hospitality systems may raise additional challenges, including scalability, computational costs, and system integration with existing booking engines or CRM platforms. Pilot studies with industry partners would provide valuable insights into these practical considerations. Fourth, while the focus here is on sentiment polarity and intensity, customer feedback also contains rich emotional and contextual dimensions (e.g., sarcasm, cultural nuances, multimodal inputs like images or voice). Expanding the framework to include multimodal sentiment analysis could enhance the depth and accuracy of customer intelligence in future research. Finally, the study is situated within the tourism and hospitality sector, where experience-driven value plays a central role. While this strengthens the contextual contribution, it also limits generalizability. Future studies should test the framework in other service industries—such as healthcare, education, and retail—to evaluate its cross-domain applicability and theoretical robustness.

6. Conclusion

This study proposed a sentiment-enhanced business intelligence framework tailored to the tourism and hospitality industry, integrating structured transactional data with unstructured opinion data to provide a holistic view of customer behavior and perception. By combining hybrid feature selection with intelligent learning models, the framework demonstrated superior performance in sentiment classification, customer segmentation, intention prediction, and recommendation systems. The results highlight the dual advantage of this integrated approach: structured BI data ensures predictive accuracy, while sentiment analysis enriches the framework with contextual depth. Empirical findings suggest measurable business value, including a 12% improvement in booking conversion rates, an 8% rise in guest satisfaction, and a 10% increase in customer retention**.** These improvements underscore the practical utility of the framework as a decision-support tool for hospitality managers and marketers.

Theoretically, the study contributes to advancing the discourse on data-centric marketing intelligence, bridging the divide between quantitative behavioral insights and qualitative perceptual analysis. Practically, it provides hospitality firms with a scalable architecture for real-time, customer-centric marketing strategies**,** enhancing personalization, loyalty, and competitive advantage. Nevertheless, the research is not without limitations. Data representativeness, dependency on model tuning, and simulation-based evaluation highlight areas for future exploration. Expanding the framework to incorporate transfer learning, multimodal sentiment analysis, and real-time industry deployment represents promising directions. Furthermore, applying the model in other service sectors could strengthen its generalizability and theoretical robustness. In conclusion, the study demonstrates that integrating BI with sentiment analysis offers a powerful pathway toward smart, adaptive marketing systems in the hospitality and tourism industry. By uniting what customers do with how they feel, the framework paves the way for more informed decision-making, enhanced customer experiences, and sustainable competitive advantage in an increasingly data-driven marketplace.

References

1. Parfinenko, A. (2017). Global market transformations of international tourism and hospitality industry. *The Journal of VN Karazin Kharkiv National University. Series: International Relations. Economics. Country Studies. Tourism*, (6), 175-186.

DOI: 10.5281/zenodo.17330787

2. Pencarelli, T. (2020). The digital revolution in the travel and tourism industry. *Information technology & tourism*, *22*(3), 455-476.
3. Troisi, O., Visvizi, A., & Grimaldi, M. (2023). Digitalizing business models in hospitality ecosystems: toward data-driven innovation. *European Journal of Innovation Management*, *26*(7), 242-277.
4. Mariani, M., & Baggio, R. (2022). Big data and analytics in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management*, *34*(1), 231-278.
5. Buhalis, D., & Michopoulou, E. (2011). Information-enabled tourism destination marketing: addressing the accessibility market. *Current issues in Tourism*, *14*(2), 145-168.
6. Oloyede, J., & Owen, J. (2025). AI-Driven Knowledge Extraction: Managing Unstructured Data for Business Intelligence Author. *Available at SSRN 5144223*.
7. Theodorakopoulos, L., Theodoropoulou, A., & Bakalis, A. (2024, September). Utilizing Big Data for Improved Targeting and Personalization for Digital Marketing Purposes in the Tourism Industry: A Comprehensive Review. In *International Conference of the International Association of Cultural and Digital Tourism* (pp. 429-470). Cham: Springer Nature Switzerland.
8. Theodorakopoulos, L., Theodoropoulou, A., & Halkiopoulos, C. (2024, September). Big data and business intelligence in tourism: Analyzing trends and enhancing customer satisfaction through online review analytics. In *International Conference of the International Association of Cultural and Digital Tourism* (pp. 507-536). Cham: Springer Nature Switzerland.
9. Ramos, C. M. (2022). Business intelligence approach and sentiment analysis as a management strategy applied to study customer satisfaction in the hospitality sector. In *Advances in Tourism, Technology and Systems: Selected Papers from ICOTTS 2021, Volume 2* (pp. 537-547). Singapore: Springer Nature Singapore.
10. Lekmiti, A., Stolk, P. J., Taylor, A., Ramachandran, S., & Yap, N. K. (2025). Text mining in tourism and hospitality research: a bibliometric perspective. *Journal of Hospitality and Tourism Technology*, *16*(3), 588-610.
11. Rastgoo, S., Mahdavi, Z., Azimi Nasab, M., Zand, M., & Padmanaban, S. (2022). Using an intelligent control method for electric vehicle charging in microgrids. *World Electric Vehicle Journal*, *13*(12), 222.
12. Ahmadi, M., Rastgoo, S., Mahdavi, Z., Nasab, M. A., Zand, M., Sanjeevikumar, P., & Khan, B. (2023). Optimal allocation of EVs parking lots and DG in micro grid using two-stage GA-PSO. *The Journal of Engineering*, *2023*(2), e12237.
13. Mahdavi, Z., Samavat, T., Javanmardi, A. S. J., Dashtaki, M. A., Zand, M., Nasab, M. A., ... & Khan, B. (2024). Providing a control system for charging electric vehicles using ANFIS. *International Transactions on Electrical Energy Systems*, *2024*(1), 9921062.
14. Kermani, A., Jamshidi, A. M., Mahdavi, Z., Dashtaki, A. A., Zand, M., Nasab, M. A., ... & Khan, B. (2023). Energy management system for smart grid in the presence of energy storage and photovoltaic systems. *International Journal of Photoenergy*, *2023*(1), 5749756.
15. Ravanbakhsh, S., & Zarrabi, H. (2017). A-Type 2 Fuzzy Scheme for Traffic Density Prediction in Smart City. *International Journal of Computer Applications*, *173*(2), 35-38.
16. Filvantorkaman, M., Piri, M., Torkaman, M. F., Zabihi, A., & Moradi, H. (2025). Fusion-based brain tumor classification using deep learning and explainable AI, and rule-based reasoning. *arXiv preprint arXiv:2508.06891*.
17. Mahdavi, Z. (2022). MRI Brain Tumors Detection by Proposed U-Net Model. *Proceedings of the MOL2NET*, *22*.
18. Mahdavi, Z. (2022). Introduce improved CNN model for accurate classification of autism spectrum disorder using 3D MRI brain scans. *Proceedings of the MOL2NET*, *22*.
19. Zarchi, M., Nazari, R. A., & Tee, K. F. (2025). Explainable multi-attribute machine learning via a hierarchical nature-inspired system toward predicting geological hazards. *Modeling Earth Systems and Environment*, *11*(4), 1-23.
20. Zarchi, M., Aslani, Z. H., & Tee, K. F. (2025). An interpretable information fusion approach to impute meteorological missing values toward cross-domain intelligent forecasting. *Neural Computing and Applications*, 1-57.
21. Zarchi, M., Nazari, R. A., & Tee, K. F. (2025). Explainable nature-inspired optimization via virtual and actual multi-objective strategies to establish a smart earthquake early warning system. *Applied Soft Computing*, 113698.
22. Toloeia, A., Zarchi, M., & Attaranb, B. (2017). Numerical survey of vibrational model for third aircraft based on HR suspension system actuator using two bee algorithm objective functions. *A*, *A*, *1*(2), 3.
23. Zarchi, M., & Attaran, B. (2019). Improved design of an active landing gear for a passenger aircraft using multi-objective optimization technique. *Structural and Multidisciplinary Optimization*, *59*(5), 1813-1833.
24. Zarchi, M., & Shahgholi, M. (2023). A novel information fusion approach using weighted neural networks for intelligent multi-class diagnostics of rotating machinery with unseen working conditions. *Journal of Vibration and Control*, *29*(23-24), 5545-5562.
25. Zarchi, M., Shahgholi, M., & Tee, K. F. (2024). An adaptable physics-informed fault diagnosis approach via hybrid signal processing and transferable feature learning for structural/machinery health monitoring. *Signal, Image and Video Processing*, *18*(12), 9051-9066.
26. Kefayat, E., & Thill, J. C. (2025). Urban Street Network Configuration and Property Crime: An Empirical Multivariate Case Study. *ISPRS International Journal of Geo-Information*, *14*(5), 200.
27. Ravanbakhsh, S., & Fesharaki, M. N. (2011). A New Service Oriented Arcitucture for Data Hiding. Computer Science and Engineering, 1(2): 26-31 doi: 10.5923/j.computer.20110102.05.
28. Krabokoukis, T. (2025). Bridging neuromarketing and data analytics in tourism: An adaptive digital marketing framework for hotels and destinations. *Tourism and Hospitality*, *6*(1), 12.

DOI: 10.5281/zenodo.17330787

29. Chan, I. C. C., Ma, J., Law, R., Buhalis, D., & Hatter, R. (2021). Dynamics of hotel website browsing activity: the power of informatics and data analytics. *Industrial management & data systems*, *121*(6), 1398-1416.
30. Wavei, E. B. (2023). An investigation into the adoption of knowledge management processes in hotels' customer relationship management value chain in Ghana.
31. Vavpotič, D., Knavs, K., & Cvelbar, L. K. (2021). Using a market basket analysis in tourism studies. *Tourism Economics*, *27*(8), 1801-1819.
32. Anubala, R. (2023). The future of hospitality: Predictive analytics in hotel management. *International Journal for Multidimensional Research Perspectives*, *1*(3), 38-58.
33. Cherednichenko, O., & Sutiahin, O. (2024, August). Development of Collaborative Business Intelligence Framework for Tourism Domain Analysis. In *European Conference on Advances in Databases and Information Systems* (pp. 253-262). Cham: Springer Nature Switzerland.
34. Bucur, C. (2015). Using opinion mining techniques in tourism. *Procedia Economics and Finance*, *23*, 1666-1673.
35. García, A., Gaines, S., & Linaza, M. T. (2012). A lexicon based sentiment analysis retrieval system for tourism domain. *Expert Syst Appl Int J*, *39*(10), 9166-9180.
36. Puh, K., & Bagić Babac, M. (2023). Predicting sentiment and rating of tourist reviews using machine learning. *Journal of hospitality and tourism insights*, *6*(3), 1188-1204.
37. Badouch, M., & Boutaounte, M. (2024). Design and implementation of a hotel recommendation system using deep learning. In *AI and data engineering solutions for effective marketing* (pp. 388-408). IGI Global.
38. Eler, D. M., Grosa, D., Pola, I., Garcia, R., Correia, R., & Teixeira, J. (2018). Analysis of document pre-processing effects in text and opinion mining. *Information*, *9*(4), 100.
39. Liu, J., Ren, F., Zhang, Y., Diao, W., & Wang, Q. (2024, November). Sentiment Analysis of User Opinions in the Automotive Industry Based on An Enhanced Nlp Model. In *2024 3rd International Conference on Artificial Intelligence, Human-Computer Interaction and Robotics (AIHCIR)* (pp. 172-175). IEEE.
40. Bin, Z. A. N., & Zhen, M. O. (2023). *U.S. Patent No. 11,741,236*. Washington, DC: U.S. Patent and Trademark Office.
41. Jović, A., Brkić, K., & Bogunović, N. (2015, May). A review of feature selection methods with applications. In *2015 38th International Convention on Information and Communication Technology, electronics and microelectronics (MIPRO)* (pp. 1200-1205). Ieee.
42. Nikolaidis, I. From Word-of-Mouth to Click-of-Mouse Unraveling Cultural Nuances from Multilingual Online Hotel Reviews.
43. Daryaei, M., Shirzad, M., & Kumar, V. (2013, July). Adoption of business intelligence: In hotel industry. In *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)* (pp. 1-6). IEEE.
44. Cadeddu, A., Chessa, A., De Leo, V., Fenu, G., Motta, E., Osborne, F., ... & Secchi, L. (2024). Optimizing tourism accommodation offers by integrating language models and knowledge graph technologies. *Information*, *15*(7), 398.
45. Muhammad, A. (2016). Contextual lexicon-based sentiment analysis for social media.
46. Lee, M. C. (2009). Using support vector machine with a hybrid feature selection method to the stock trend prediction. *Expert systems with applications*, *36*(8), 10896-10904.
47. Naidu, G., Zuva, T., & Sibanda, E. M. (2023, April). A review of evaluation metrics in machine learning algorithms. In *Computer science on-line conference* (pp. 15-25). Cham: Springer International Publishing.
48. Afolabi, I. T., Worlu, R. E., Adebayo, O. P., & Jonathan, O. (2019, August). Predicting customer behavior with combination of structured and unstructured data. In *Journal of Physics: Conference Series* (Vol. 1299, No. 1, p. 012041). IOP Publishing.
49. Phillips, P., Barnes, S., Zigan, K., & Schegg, R. (2017). Understanding the impact of online reviews on hotel performance: an empirical analysis. *Journal of travel research*, *56*(2), 235-249.
50. Afzaal, M., Usman, M., & Fong, A. (2019). Tourism mobile app with aspect-based sentiment classification framework for tourist reviews. *IEEE Transactions on Consumer Electronics*, *65*(2), 233-242.
51. Wang, J., & Duggasani, A. (2020). Forecasting hotel reservations with long short-term memory-based recurrent neural networks. *International Journal of Data Science and Analytics*

DOI: 10.5281/zenodo.17330787