Research Article

AI Sentiment Analysis for Destination Branding: A Case Study of Buriram, Thailand

Received: *August 30, 2023*
Revised: *September 25, 2023*
Accepted: *October 12, 2023*

Itsaree Khumtaveeporn
Graduate of BADT Program,
School of Management, Asian Institute of Technology

Dr.Kritsadarat Wattanasuwan
Associate Professor of Department of Marketing,
Thammasat Business School, Thammasat University

ABSTRACT

This study aims to develop a machine learning model that extracts destination attributes from travel reviews for destination branding. It intends to assist destination management organisations and tourism practitioners to enhance emerging tourist destinations facing branding challenges. Big data analysis via Latent Dirichlet allocation is used to reveal unique attributes and TextBlob-Flair assesses the sentiment. TripAdvisor's reviews on Buriram served as a case study. The analysis demonstrated the effectiveness of the methodology in identifying destination attributes. The findings highlighted three distinctive attributes of Buriram: archaeological sites, sports, and local market. However, only archaeological and sports can be effectively leveraged for destination branding. This method demonstrated its effectiveness in rapidly and accurately analysing large amounts of text to reveal patterns that may not be immediately apparent to human readers. Consequently, sentiment analysis proves invaluable in identifying distinct attributes that can be utilised for destination branding on a global scale. Moreover, the methodology employed in this study can be easily adapted and implemented in various destinations. This innovative approach enhances the analysis of online reviews, providing valuable insights that contribute to the existing literature on destination branding.

Keywords: Sentiment Analysis, Destination Branding, LDA, Topic Modeling, eWOM

บทความวิจัย

การวิเคราะห์ความรู้สึกนึกคิดด้วย AI เพื่อการสร้างแบรนด์จุดหมายปลายทาง : กรณีศึกษาจังหวัดบุรีรัมย์

วันที่ได้รับต้นฉบับบทความ : *30 สิงหาคม 2566*
วันที่แก้ไขปรับปรุงบทความ : *25 กันยายน 2566*
วันที่ตอบรับตีพิมพ์บทความ : *12 ตุลาคม 2566*

อิสรีย์ คุ้มทวีพร
มหาบัณฑิตสาขาวิชาวิทยาการวิเคราะห์ธุรกิจ
และการปรับตัวทางเทคโนโลยีดิจิตอล (BADT)
สำนักวิชาวิทยาการการจัดการ สถาบันเทคโนโลยีแห่งเอเชีย

ดร.กฤษฎารัตน์ วัฒนสุวรรณ
รองศาสตราจารย์ประจำภาควิชาการตลาด
คณะพาณิชยศาสตร์และการบัญชี มหาวิทยาลัยธรรมศาสตร์

บทคัดย่อ

การศึกษานี้มีเป้าหมายเพื่อพัฒนาโมเดลการเรียนรู้ของเครื่อง (Machine Learning) ที่สามารถนำข้อมูลจากรีวิวท่องเที่ยวออนไลน์มาใช้ในการสร้างแบรนด์จุดหมายปลายทาง วัตถุประสงค์ของแบบจำลองนี้คือเพื่อช่วยองค์กรจัดการสถานที่ท่องเที่ยวและผู้ปฏิบัติงานด้านการท่องเที่ยวในการสร้างแบรนด์ของจุดหมายปลายทางใหม่ที่ต้องเผชิญกับความท้าทายในการสร้างแบรนด์ เพื่อให้บรรลุเป้าหมายนี้ การศึกษานี้ใช้การจัดสรรดีรีเคลแฝง (Latent Dirichlet allocation) เพื่อวิเคราะห์รีวิวการเดินทางในการค้นหาคุณลักษณะเฉพาะ นอกจากนี้ยังใช้เครื่องมือประมวลผลภาษาธรรมชาติ โดยเฉพาะ TextBlob-Flair ในการวิเคราะห์อารมณ์ความรู้สึก (Sentiment Analysis) ของนักท่องเที่ยวโดยใช้รีวิวจังหวัดบุรีรัมย์ที่เป็นภาษาอังกฤษ จากเว็บไซต์ท่องเที่ยว TripAdvisor เป็นกรณีศึกษา การวิเคราะห์แสดงให้เห็นถึงประสิทธิภาพของวิธีในการระบุคุณลักษณะของจุดหมายปลายทาง โดยพบว่าบุรีรัมย์มีคุณลักษณะที่โดดเด่น 3 ประการ ได้แก่ สถานที่ทางประวัติศาสตร์ กีฬา และตลาดท้องถิ่น อย่างไรก็ตาม เฉพาะสถานที่ทางประวัติศาสตร์และกีฬาเท่านั้นที่สามารถใช้ในการสร้างแบรนด์จุดหมายปลายทางได้อย่างมีประสิทธิภาพ วิธีนี้แสดงให้เห็นถึงประสิทธิภาพทั้งด้านความรวดเร็วและแม่นยำในการวิเคราะห์ข้อมูลจำนวนมาก ซึ่งเผยให้เห็นแบบแผนที่อาจไม่ปรากฏแก่ผู้อ่านในทันที ผลการศึกษานี้พิสูจน์ว่า การวิเคราะห์ความรู้สึกมีค่าอย่างยิ่งในการระบุคุณลักษณะเฉพาะ ที่สามารถนำไปปรับใช้สำหรับการสร้างแบรนด์จุดหมายปลายทางในระดับโลก แนวทางใหม่นี้สนับสนุนการวิเคราะห์รีวิวออนไลน์ โดยนำเสนอข้อมูลเชิงลึกอันมีค่าที่สนับสนุนวรรณกรรมเกี่ยวกับการสร้างแบรนด์จุดหมายปลายทางที่มีอยู่

คำสำคัญ : การวิเคราะห์ความรู้สึก การสร้างแบรนด์จุดหมายปลายทาง การจัดสรรดีรีเคลแฝง แบบจำลองหัวข้อ การสื่อสารแบบปากต่อปากบนอินเทอร์เน็ต

1. INTRODUCTION

Recent studies have highlighted the importance of transitioning from traditional top-down branding strategies to decentralised and co-creative approaches for the success of tourist destinations (Oliveira & Panyik, 2015). This shift recognises the significant role that customer experience plays in shaping travellers' perceptions and decision-making processes. Furthermore, the increasing influence of online travel reviews (OTRs) in tourism research and destination marketing underscores their impact on tourists' decision-making (Guo et al., 2021). The abundance of user-generated content (UGC) on various travel platforms emphasises the need for destinations to embrace information and communication technologies, including big data utilisation (Chang et al., 2020; Madyatmadja et al., 2021).

The exponential growth of UGC has led to a demand for automated methods to manage the vast amount of data (Gour et al., 2021). Manual reading and analysis of these reviews have become impractical and can lead to subjective interpretations by humans, prompting the adoption of natural language processing (NLP). NLP encompasses various tasks, such as information retrieval, information extraction, text summarisation, question answering, topic modelling, and sentiment polarity analysis, providing a comprehensive solution (Chang et al., 2020). Despite the extensive research on sentiment analysis in different fields, limited attention has been given to its application in destination branding within the tourism and hospitality industry.

On the other hands, the Tourism Authority of Thailand (TAT) has played a pivotal role in promoting the tourism sector of Thailand. The agency has introduced various initiatives aimed at showcasing and marketing less-recognised destinations such as Buriram, Phetchabun, Nan, Loei, Samut Songkhram, and Nakhon Si Thammarat. These endeavours are designed to diversify Thailand's tourism offerings and alleviate overcrowding in popular destinations. Simultaneously, by lending support to these lesser-known destinations, TAT seeks to stimulate local economic development by creating employment opportunities and generating revenue.

In light of this gap and to align with the TAT's efforts, the authors propose a sentiment analysis of online travel reviews (OTRs) to discover tourists' unique experiences that can be leveraged for destination branding. Co-creation, a significant aspect of the tourism industry focused on selling experiential offerings (Mohammadi et al., 2021), reinforces the need for this analysis. Furthermore, a study indicated that destination management should prioritize the promotion of the destination through a comprehensive and emotionally resonant communication strategy (Költringer & Dickinger, 2015). By harnessing the power of OTRs, the authors aim to enhance the competitive advantage of tourist destinations and contribute to their development. Additionally, the authors envision the developed tools to address overcrowding issues in popular destinations. Through the utilisation of machine learning, this study seeks to analyse tourists' sentiments and provide insightful analysis and effective implementation strategies for lesser-known destination branding to promote and distinguish itself from its competitors. To achieve this goal, a topic modelling approach using Latent Dirichlet allocation (LDA) and sentiment analysis was employed to address three research questions (RQs):

1. What are the destination's attractions' identifiable attributes based on travellers' experiences? Python NLP techniques were employed to cluster recurring words or terms extracted from the OTRs on travel platforms. These word clusters represent the distinguishable attributes of a destination's attractions derived from the shared experiences of travellers.

2. What is the sentiment associated with each attribute of the attractions? Sentiment analysis was utilised to discover the sentiment of each sentence on the OTRs. The results indicate both the satisfaction and dissatisfaction the travellers experienced toward each attribute of the attraction.

3. How can the destination be effectively branded? The attributes that received a high level of satisfaction will be utilised for destination branding.

2. LITERATURE REVIEW

Developing a brand for a destination is a complex and challenging task due to the intangible nature of destination products, which cannot be returned if tourists are dissatisfied. This section highlights the significance of destination branding concepts, the pivotal role of UGC in shaping the perception and experience of destinations, and the sentiment analysis within the tourism and hospitality industry context. In this literature review, the authors will provide brief references to these three research fields and then determine the gap in the literature.

2.1 Destination Branding

Branding originated in the 1940s, but the focus on destination branding gained prominence in the late 1990s (Oppermann, 2000). While "identification" and "differentiation" are fundamental branding concepts (Aaker, 1991), destinations also need to consider the concept of "experience," and ideally, a "unique experience" of travellers (Ritchie and Ritchie, 1998). In 2005, the definition of destination branding was expanded from that proposed by Ritchie & Ritchie (1998) to include additional themes such as expectations, image, consolidation, reinforcement, recognition, consistency, brand messages, and emotional responses. Blain et al. (2005) proposed a new definition stating that destination branding involves marketing activities that: (1) create a recognisable and distinctive identity for the destination through a name, symbol, logo, or graphic; (2) consistently communicate the expectation of a unique and memorable travel experience associated with the destination; (3) strengthen the emotional connection between visitors and the destination; and (4) reduce consumer search costs and perceived risk, ultimately shaping consumer destination choice. These activities collectively shape a positive destination image influencing consumer choice (Blain et al., 2005). Moreover, the concept of destination branding is often used interchangeably with other related terms such as (re)positioning (Gilmore, 2002), image-building (Cai, 2002; Curtis, 2001), and image reconstruction (Hall, 2002) of a destination.

To establish a strong destination brand, destination management organisations (DMOs) need to consider the perceptions of stakeholders and the broader macro-environment, including the political, social, and economic issues affecting the destination (Almeyda-Ibáñez & George, 2017). Additionally, other studies suggest that DMOs should also consider psychological factors, particularly personal factors such as age, values, and motivations, as well as stimulus factors such as sources of information and past experiences, when targeting potential travellers (Baloglu & McCleary, 1999). By understanding these factors, DMOs can shape their destination brand strategy to align with the needs and preferences of their target audience.

Moreover, some researchers also indicated that destination brands can be viewed as co-created products. Scholars such as Aitken & Campelo (2011) and Oliveira & Panyik (2015) have highlighted the concept of co-creation in the context of destination branding. This perspective emphasises the active involvement and participation of various stakeholders, including tourists, in shaping and enhancing the destination brand. Giannopoulos et al., (2021) also proposed a framework for destination brand co-creation, emphasising the involvement of multiple stakeholders in creating value process. The framework highlights the idea that brand development is not solely determined by customers but also relies on the perspectives and contributions of various stakeholders. This implies that stakeholders at different levels can shape the brand's value and perception.

Additionally, Dellarocas (2003) found that UGC plays a crucial role in shaping destination image. With the rise of social media and UGC, tourists now could act as both customers and promoters for destinations (Wise & Farzin, 2018). They can share their experiences and create content on social media platforms, influencing potential travellers. This indicated that tourists' opinions and actions significantly contribute to destination branding, as they influence how destinations communicate with the public.

2.2 Electronic Word of Mouth (eWOM) and Tourist Co-creation Experience

With the rapid development of technology, web 2.0 applications have played a significant role in the tourism industry by enabling two-way communication and allowing internet users to generate a vast amount of UGC related to hotels, destinations, and services (Sigala, 2008; Ye et al., 2011). This UGC, when produced through online channels, is commonly referred to as electronic word of mouth (eWOM) (Filieri et al., 2015). It has been observed that these methods have gained popularity and are recognised as an official marketing strategy in the tourism industry (Kucukusta et al., 2015) since potential consumers perceive it as unbiased, free from commercial influence (Mir & Rehman, 2013). They also generally place more trust in electronic word of mouth (eWOM) than communication initiated by companies (Mukhopadhyay et al., 2023). It can be inferred that eWOM is a free marketing tool. Through their reviews, recommendations, and interactions on online platforms, tourists become active contributors to the co-creation process, influencing the overall tourism experience. Additionally, research suggests that co-creation and value creation enhance customer loyalty (Javed & Awan, 2022). Thus, tourist destinations that embrace co-creation in their branding strategies are likely to foster greater tourist loyalty.

In fact, both consumers and online sellers extensively utilise eWOM. From the supplier's perspective, eWOM carries significant implications for product development, brand establishment, and quality assurance (Dellarocas, 2003). In other words, reviews can serve as an indicator to evaluate the performance of service providers (Ganzaroli et al., 2017; Nieto et al., 2014) and shape their reputation (Baka, 2016), which in turn influences potential customers' purchase decisions and the overall popularity of the products. Despite being freely available and generated across various social media platforms, eWOM can be a game-changer for products and services if the insights, particularly the sentiment expressed by the authors, can be extracted, and utilised for further development. By understanding and analysing these sentiments, businesses can better cater to the preferences and needs of tourists based on their past purchases and experiences and use them to shape the products, services, and experiences that enhance customer satisfaction (Manosso & Ruiz, 2021). Therefore, eWOM plays a significant role in facilitating the co-creation experience between service providers and tourists.

Furthermore, it is noteworthy to emphasise that the recognition of destination brands as co-created products involving various stakeholders has only emerged relatively recently in academic discourse (Aitken & Campelo, 2011; Giannopoulos et al., 2021; Költringer & Dickinger, 2015; Oliveira & Panyik, 2015). Amongst these studies, only Költringer & Dickinger (2015) have utilised web content mining techniques to extract facets of destination brand identity and image from online sources. Surprisingly, despite the compelling evidence that user-generated content is the most extensive and diverse source of online information (Költringer & Dickinger, 2015), none of the mentioned studies have incorporated sentiment analysis to assess purely tourist-generated content for its potential value in destination branding efforts.

2.3 Sentiment Analysis

Sentiment analysis also referred to as opinion mining falls within the realm of NLP. Sentiment analysis is a specific type of content analysis focusing on texts' emotional polarity (Barbierato et al., 2021). In this approach, texts are classified into three distinct categories based on their emotional tone: positive, neutral, or negative (Barbierato et al., 2021; Puh & Babac, 2022). This classification is determined by assigning a scaled score to each text, indicating the prevailing sentiment expressed within it. This technique enables identifying and interpreting specific aspects of sentiment or emotion expressed in unstructured text, providing insights into human psychology. Furthermore, studies suggest that advanced machine learning methods, particularly those based on deep learning algorithms, can achieve high accuracy in determining the polarity of online reviews (Chang et al., 2020). According to Schuckert et al. (2015), the average accuracy of tourism and hospitality research is approximately 70%. By leveraging these techniques, researchers can better understand customer sentiments and preferences, contributing to improved decision-making and effective marketing strategies within the tourism industry.

In the destination branding context, tourists' sentiment and emotional attachment to a destination can provide a significant competitive advantage over rival destinations (Yağmur & Aksu, 2022). When tourists develop a strong emotional connection to a destination, it fosters a deeper bond and enhances their overall experience. This emotional commitment leads to increased loyalty, positive word-of-mouth promotion, and a higher probability of repeat visits. By leveraging sentiment analysis, destinations can identify areas that evoke strong positive emotions and capitalise on them for branding purposes.

In addition, various techniques can conduct sentiment analysis and have been studied in various tourism and hospitality aspects. For examples, Puh & Babac (2022) have explored machine and deep learning models for predicting sentiment and rating from tourist reviews. Barbierato et al. (2021) have analysed the key elements for the success of the wine tour by using the AFINN lexicon. Taecharungroj & Mathayomchan (2019) have applied LDA and Naïve Bayes modelling to analyse the sentiment polarity of the Phuket tourists. Hemalatha & Ramathmika (2019) have conducted sentiment analysis on restaurant reviews by applying Python and several algorithms, including Naïve Bayes, logistic regressions, and Linear Support Vector Machine (linear SVC). Prameswari et al. (2017) utilised the combination of sentiment analysis and text summarisation approaches to analyse hotel reviews in Labuan Bajo and Bali. Aljedaani et al. (2022) conducted sentiment analysis on the US airline industry using TextBlob and various deep learning models with the aim of enhancing sentiment accuracy. Mishra et al. (2021) conducted a study on people's sentiments towards healthcare tourism and hospitality during the COVID-19 outbreak. They utilised Recurrent Neural Network (RNN) and Support Vector Classifier (SVC) and found both algorithms to have approximately 81% accuracy. However, RNN proved more powerful due to its short-term and long-term memory components, making it suitable for sequential data analysis. In contrast, SVC does not have memory for sequential data.

2.4 Gaps in the Literature

The literature suggests that while tourists' actions and opinions, particularly eWOM, play a significant role in marketing, limited research explores the relationship between eWOM and destination branding. Additionally, while sentiment analysis has been widely employed in the tourism industry, there is a gap in research that specifically analyses the sentiment derived from eWOM for the purpose of destination branding, despite the potential for eWOM to facilitate the co-creation experience.

3. THE CASE STUDY DESTINATION

This study used Buriram, a northeastern Thailand province renowned for its ancient Khmer civilisation and volcanic landscape (TAT Buriram, 2019), as a case study. Located approximately 410 kilometres from Bangkok, Buriram is characterised by significant Khmer monuments, such as the well-preserved Angkorian monastery and Phanom Rung sanctuary (TAT Buriram, 2019). The province also boasts notable Khmer pottery kiln sites and a distinct local craft called Na Pho Mudmee silk. Buriram's cultural heritage is evidenced by archaeological discoveries, including prehistoric human habitation, ancient ruins from the Dvaravati period, and over 60 sandstone sanctuaries (TAT Buriram, 2019). Moreover, Buriram has gained recognition as a sports destination, featuring a prominent football stadium, an international motorsport racetrack, and the Buriram United Football Club (MICE Intelligence Center, 2019). Convenient access to Buriram is available through both air and land transportation.

According to Nasa & Hassan (2016), an analysis of Buriram's tourist sites and activities reveals a diverse range of attractions, including natural, historical, cultural, and recreational destinations. This makes Buriram suitable for educational tourism, particularly emphasising its historical resources. However, Chaigasem & Leruksa (2020) suggest that Buriram also has the potential to be a sport destination. Despite the implementation of the 6th National Sports Development Plan, which brought in a significant number of tourists and generated substantial revenue for local businesses in 2016, Buriram still lags behind Bangkok, the capital city of Thailand, in terms of tourist numbers (MICE Intelligence Center, 2019; Hedrick-Wong & Choong, 2016). This indicates that the current branding efforts in Buriram may not be effectively attracting a larger tourist market. Furthermore, there is a lack of promotion for tourists to discover Buriram's historical and natural attractions (Chaigasem & Leruksa, 2020). To address this discrepancy, it is crucial to evaluate Buriram from tourists' perspective and explore their demands. This evaluation will enable the implementation of effective marketing strategies to successfully brand Buriram as a desirable destination.

4. METHODOLOGY

The research methodology comprises three steps: data collection, data preprocessing, and data analysis, as depicted in Figure 1.

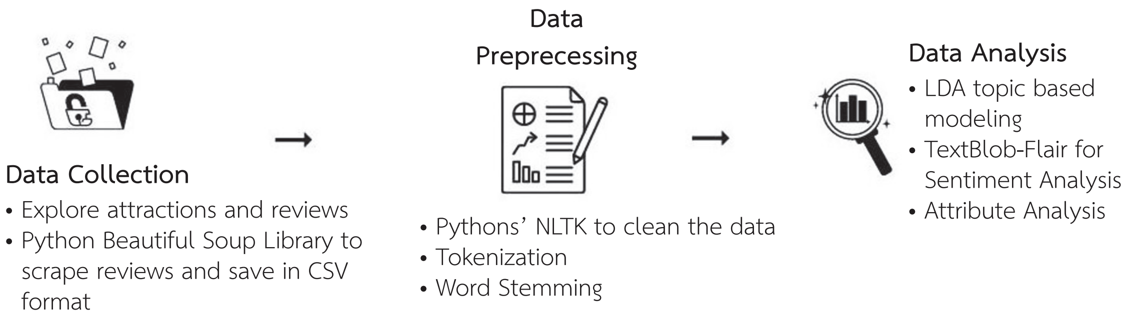

Figure 1: The overview of research methodology

4.1 Data Collection

The authors opted to collect the OTRs specifically from TripAdvisor. TripAdvisor is a travel platform that has gained immense popularity among both users and researchers (Taecharungroj & Mathayomchan, 2019). It is widely recognised as the leading review platform for marketing purposes (Barbierato et al., 2021). According to TripAdvisor's official website, it boasts being the world's largest travel guidance platform, with approximately 59 million global reviews and opinions in 2020, despite the impact of the COVID-19 pandemic. Moreover, TripAdvisor has implemented an advanced content moderation system, enabling the platform to remove or reject reviews that violate its community standards or contain false information. Hence, the reviews available on TripAdvisor can be deemed as reliable and high-quality data (Xiang et al., 2017), suitable for evaluating destination performance. Additionally, previous studies have reported high sentiment values associated with reviews on TripAdvisor (Mariani & Borghi, 2020).

After the OTR platform was selected, the authors studied several online travel websites, including TripAdvisor, Lonely Planet, Expedia, and Trip.com, that featured travel articles about the study area to choose the attractions. Various categories of attractions, such as religious sites, local markets, community malls, historical sites, museums, sports, entertainment parks, and spas and wellness, were identified. However, entertainment parks, spas, and wellness were excluded from the research due to their exclusive operation by private organisations. Hence, a total of 586 online English reviews related to 17 attractions were extracted from TripAdvisor using web scraping libraries, Selenium and Beautiful Soup, in October 2022. The extracted reviews were then stored in a CSV file.

4.2 Data Preprocessing

In machine learning, data preprocessing is a crucial step in preparing and training machine learning models to derive insights from the data. It involves cleaning and transforming raw data into a format that machine learning algorithms can easily process. In this research, the Natural Language Toolkit (NLTK) Python library is utilised for preprocessing and analysing the opinions extracted from the text dataset. NLTK offers various functionalities, including data cleaning, classification, tokenising, stemming, parsing, tagging, and semantic reasoning. Additionally, NLTK provides a set of stopwords, which are words that do not significantly contribute to the meaning of sentences and can be excluded from the dataset. This feature makes NLTK suitable for data preprocessing in this research.

4.2.1 Data Cleaning

The data cleaning process in this research involves several steps. Firstly, a punctuation eraser is used to remove special characters, punctuations, and symbols from the text data. This step helps to eliminate unnecessary noise from the dataset. Next, a Python script is created to remove stopwords, which are insignificant terms or commonly used words in NLP. Additionally, the word "Buriram" is also removed from the dataset, as it is specific to the context and not relevant for analysis. This pruning of stopwords helps to reduce the dataset size, resulting in faster training of machine learning models and improved performance and accuracy by focusing only on meaningful words or tokens. Furthermore, a case converter is applied to convert all the text to lowercase. This step ensures consistency in the text data, as the machine might interpret words written in different cases differently. Converting all characters to lowercase helps to avoid any inconsistencies and enhances accuracy in subsequent analysis and modelling tasks.

4.2.2 Tokenisation

Tokenisation is a crucial technique used to process raw data in machine learning. In this step, the authors employed tokenisation to transform the reviews, which were initially in the form of paragraphs, into machine-readable format. The tokenisation process involved breaking down the reviews into sentence tokens and then further splitting them into word tokens. These word tokens typically represent individual words, but they can also include numbers or symbols as necessary. Additionally, parts-of-speech tagging was applied to the tokens, assigning them specific tags based on their part of speech or morphemes. This process helps to provide contextual information about the tokens and allows for more accurate analysis. For example, the verb "visits" would be stored as "visit+s," while the noun "temples" would be stored as "temple+s." This tagging enables the identification and differentiation of various grammatical elements within the text, contributing to a more comprehensive data analysis.

4.2.3 Word Stemming

Word stemming is an essential step in preparing text data for effective machine learning modelling. It involves reducing words to their base or root forms, thereby removing any variations or inflections. This process helps to eliminate repetition and enhance the consistency of the data. In this research, the authors employed stemming techniques from the Python NLTK library to normalise the text data. Various stemming algorithms are available in NLTK, such as the Porter Stemmer, Snowball Stemmer, Lancaster Stemmer, and Regexp Stemmer.

To choose the stemmer for this study, the authors conducted tests on different stemmers to assess their ability in reducing tokens such as "companions," "companionship," and "companionships" to their root form. The findings indicated that only the Lancaster Stemmer was successful in normalising these words to the root form "companion." This process streamlined the data, eliminating redundancy and leading to more precise and efficient analysis. Thus, the Lancaster Stemmer was selected as the preferred stemmer algorithm for word stemming in this research.

4.3 Data Analytics

Data analytics is a process that involves examining data to identify patterns, extract insights, and develop knowledge. It encompasses various techniques, including data mining and data science, which aim to derive meaningful information from the data. Data analytics involves a range of methods, from basic approaches such as data visualisation and summary statistics to more advanced tools like machine learning models. In the context of this study, the authors employed two specific techniques, namely topic modelling and sentiment analysis, to analyse the collected data and gain valuable insights.

4.3.1 Topic Modelling

To answer RQ1, the acquired data needs to be extracted and analysed to gain the "insight" information. Topic modelling is a text mining approach that the authors adapted to discover a set of topics (attributes) from the OTRs. To illustrate, it is a technique used in machine learning and NLP that uses probabilities to identify recurring themes or topics in a collection of documents. It is a form of text mining that is commonly used to discover underlying semantic structures in textual data (Blei et al., 2003; Blei, 2012).

In this research, LDA, which is a generative probabilistic model, was utilized to detect the attributes in a collection of corpus of OTRs because it was proof to be the most effective model in managing big data (Blei et al., 2003). To set parameter values, the amounts of topics were selected based on the use of k- Means and silhouette method because it can interpret and validate crisp cluster data; it also graphically represents how well each topic lies within its cluster (Rousseeuw, 1987). Then the LDA algorithm was performed and generated the output containing each document's mapping to the assigned topic, and the probability of each document belonging to a specific topic were produced. Later, a list of 10 terms/words, under each topic, ranked by the frequency of appearance (weight) was also produced as another dataset. To answer RQ1, one researcher initially assigned names to the attributes based on the predominant underlying terms identified through the analysis. A second researcher then verified and confirmed these names (Guo et al., 2017; Taecharungroj & Mathayomchan, 2019). Figure 2 represents the overview of the topic modelling process by using LDA.

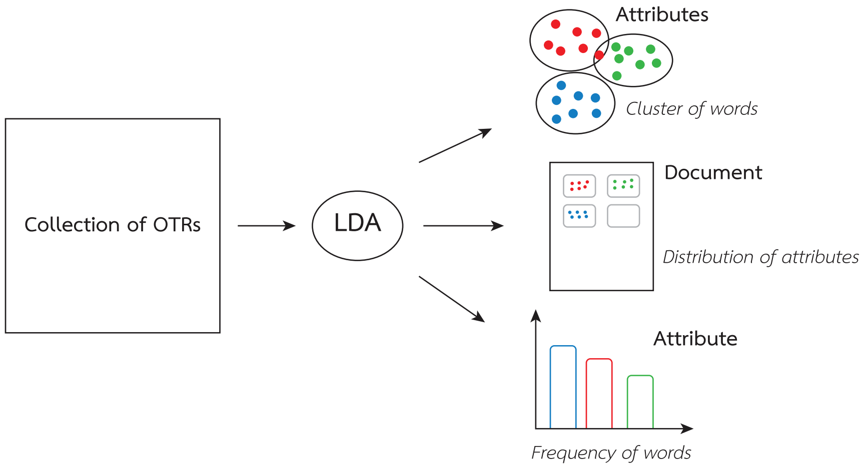

Figure 2: The LDA process

Note: Adapted from "Probabilistic topic models" by D. M. Blei, 2012, Communications of the ACM, 55(4), p. 78 (https://doi.org/10.1145/2133806.2133826)

4.3.2 Sentiment Analysis

Sentiment analysis can be performed using different techniques such as machine learning, deep learning, and lexicon-based methods. Several pre-trained sentiment analysis models are available that are both effective and freely accessible. TextBlob, VADER, and Flair are popular Python-based pre-trained models used for sentiment analysis. These models can accurately predict the polarity of sentiments, classifying them as positive, negative, and sometimes neutral.

To illustrate, TextBlob is a widely used library that processes textual data by utilising a sentiment lexicon containing predefined words (Dabade et al., 2021). However, a limitation of TextBlob is that it focuses primarily on individual words and does not consider the context within sentences. On the other hand, VADER, also known as Valence Aware Dictionary and Sentiment Reasoner, is a popular library specifically designed for sentiment analysis in social media (Dabade et al., 2021). It incorporates an algorithm that adjusts sentiment intensity based on factors such as slang, emojis, acronyms, and punctuation. Another approach is Flair, which is a contextualised representation framework built on PyTorch (Dabade et al., 2021). It breaks sentences into character sequences from larger strings and is pre-trained on a bidirectional language model, enabling it to learn embeddings at the character level.

To select the most effective sentiment analysis model for this research, each model needs to be tested on its accuracy. The following are the steps that the authors conducted to select the sentiment analysis model:

1. Find the test dataset.
2. Perform sentiment analysis using each model on the test dataset.
3. Compare the accuracy results of each model.

The authors obtained a test dataset from Kaggle.com (Dhar, 2022) that closely resembles the dataset obtained from OTRs in Buriram. This dataset comprises 20,490 online English reviews written by hotel customers on TripAdvisor, each with its corresponding sentiment analysis. The sentiment analysis models, namely VADER, TextBlob, and Flair, were applied to analyse the sentiment of each review in the test dataset. The polarity scores generated by the models were then compared to the labels of the test data to evaluate the accuracy of each model. The comparison of the results obtained from TextBlob, VADER, and Flair is presented in Table 1.

Table 1: Accuracy Results for TextBlob, VADER, and Flair

Pre-trained Model Library	Methodology	Output and its Meaning	Processing Time per Review	Overall Accuracy
TextBlob	Lexicon andrule-based	–1 : Very Negative 0 : Neutral 1 : Very Positive	0.2 seconds	84.5%
VADER	Lexicon andrule-based	–1 : Very Negative 0 : Neutral 1 : Very Positive	0.1 seconds	81.5%
Flair	Model-based	–1 : Very Negative 0 : Neutral 1 : Very Positive	~12.5 seconds (0.01 with pre-loaded model)	77.4%

Note: The output values of TextBlob and VADER are in the range between –1 to +1 where –1 is very negative, 0 is neutral, and +1 is very positive. While Flair directly outputs whether the review is positive or negative along with the confidence of the prediction.

To enhance the performance and accuracy of the sentiment analysis models, two modules were integrated and tested using the same test dataset. The accuracy of the integrated models is presented in Table 2. While Table 1 indicated that Flair achieves the lowest accuracy compared to the other models, when combined with TextBlob, which includes two sentiment analysis implementations (Pattern Analyzer and Naive Bayes Analyzer), the accuracy significantly improves to 88.2%, as displayed in Table 2. The effectiveness of this combination could be attributed to Flair's ability to disambiguate case-sensitive characters and handle tasks like part-of-speech tagging. It can identify negations, distinguish proper nouns from similar-sounding common nouns, and recognise other syntactic patterns in natural language. Therefore, in this research, the authors chose TextBlob-Flair as the sentiment classifier for conducting sentiment analysis to address RQ2, as it achieves an accuracy of approximately 88.2%, which surpasses the average accuracy reported in tourism and hospitality research by 18.2% (Schuckert et al., 2015).

Table 2: Accuracy Results for Integrated Sentiment Analysis Models

Sentiment Analysis Model	TextBlob (%)	VADER (%)	Flair (%)
TextBlob	84.5	87.5	88.2
VADER	87.5	81.5	86.9
Flair	88.2	86.9	77.4

4.3.3 Attribute Analysis

After conducting LDA and sentiment analysis on the acquired dataset, the authors follow a series of steps, as depicted in Figure 3.

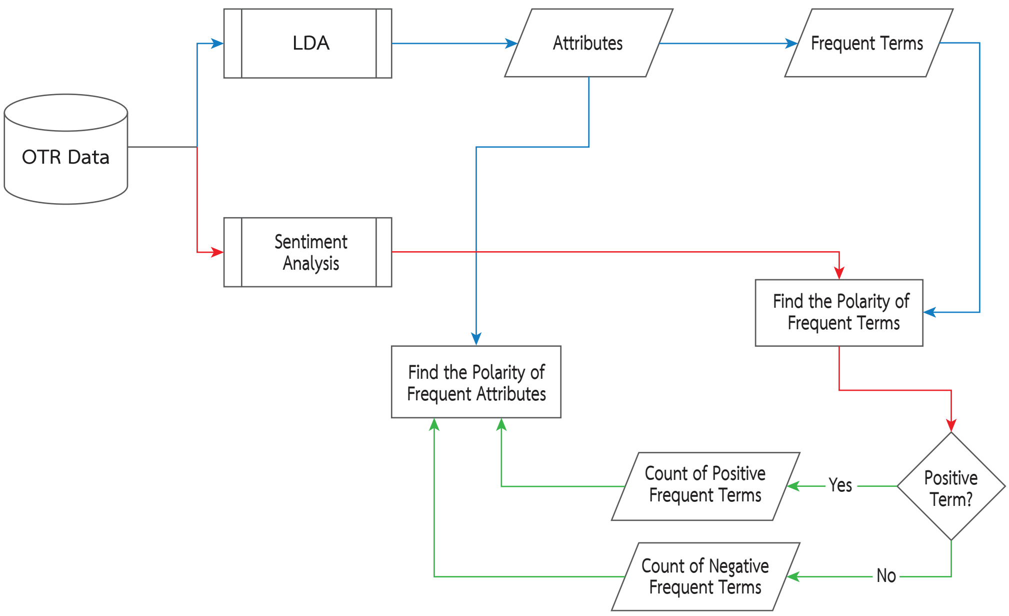

Figure 3: The Overview of the Attribute Analysis

Note: Frequently occurring terms and attributes in positive reviews are indicative of factors that positively influence tourists' experiences.

The steps to explore and draw conclusions regarding the attributes that can be leveraged for destination branding include:

1. Assessing term frequency in positive and negative reviews: The frequency of the top 10 terms in both positive and negative reviews is determined by counting the number of occurrences.
2. Analysing attribute frequency in positive and negative reviews: For each attribute, the sentiment percentage is calculated by dividing the total number of positive reviews by the total number of reviews. The same calculation is performed for negative reviews to determine the sentiment distribution.
3. Drawing conclusions about influential attributes: The results of the positive and negative sentiment percentages are compared to identify the most promising attribute that can be utilised for destination branding.

5. DATA ANALYSIS AND RESULTS

To answer RQ1, LDA was applied to analyse a large dataset and identify the attributes associated with the destination. This process is also known as topic modelling. And to address RQ2, TextBlob-Flair was employed to determine the sentiment expressed in the reviews.

5.1 LDA Analysis

A k-Means algorithm was implemented to cluster the data based on similarity. The determination of the parameter value, k, was crucial but challenging. To address this, the authors utilised the silhouette coefficient method, a well-known technique for identifying the optimal value of k (Kaoungku et al., 2018). The silhouette coefficient measures how well an object belongs to its own cluster compared to other clusters. For this research, the highest silhouette coefficient score for Buriram’s attractions is approximately 0.36, and the corresponding parameter value is 3. Therefore, the analysis revealed that there are three distinct topics or attributes of Buriram's tourist attractions represented in the collected OTRs.

Then, the authors employed LDA to extract a predetermined number of themes, underlying terms, and weight scores (WS) from the OTR dataset. To identify the attributes, one researcher interpreted the names of each attribute by examining the 10 most frequent and dominant terms within each topic and followed by confirmation by another researcher. The process of naming the topics was determined by establishing a logical link between the most commonly occurring words within each topic (Guo et al., 2017). For instance, in Table 3, the topic labelled as “Sport” was named based on the presence of highly weighted terms such as “circuit” with a weight of 99% and “race” with a weight of 99%, which both were prominently ranked at the top of the word list. Based on this analysis, three attributes emerged from the data: archeological sites, sport, and local market.

Table 3: Example of Identified Attributes

Archeological Sites		Sport		Local Market	
Term	WS	Term	WS	Term	WS
Site	1	Circuit	0.99	Market	0.98
Temple	1	Race	0.99	Bar	0.93
Khmer	1	Track	0.98	Night	0.89
Baht	1	Grandstand	0.96	Castle	0.65
Beautiful	1	Event	0.96	Live	0.55
Visit	1	Football	0.95	Street	0.46
Top	1	Facility	0.84	Restaurant	0.38
Walk	0.99	International	0.82	Shop	0.38
Volcano	0.99	Watch	0.75	Food	0.33
See	0.99	Stadium	0.67	Local	0.23

Note: Weight score is calculated by dividing the total number of specific terms under an attribute by the total number of the same terms for the entire set of reviews. A higher weight score indicates greater dominance of a term in representing the attribute.

5.2 Sentiment Analysis

To examine the sentiment conveyed by tourists towards the identified attributes, the authors employed TextBlob-Flair sentiment analysis. This tool has demonstrated an impressive sentiment accuracy of 88.2%, surpassing the average accuracy observed in the field of tourism and hospitality research by 18.2% (Schuckert et al., 2015). To ascertain the sentiment polarity of each attribute, the 10 underlying terms associated with it were subjected to multiple analyses using TextBlob-Flair. This approach aimed to ensure the highest possible prediction accuracy and overall reliability of the sentiment analysis. To address RQ2 and RQ3, sentiment percentages were computed to identify the most promising destination attribute in Buriram. The results of the sentiment analysis for each attribute are presented in Tables 4-6, which provide information on the frequency of the 10 most common terms associated with each attribute, the number of reviews containing these terms, and the distribution of positive, neutral, and negative reviews featuring these terms.

Table 4: Sentiment Analysis Results: Archeological sites

Term	Review Count	Positive	Neutral	Negative
Site	211	106	96	9
Temple	263	167	88	8
Khmer	118	61	56	1
Baht	38	8	28	2
Beautiful	103	101	2	0
Visit	326	229	83	14
Top	161	88	67	6
Walk	182	111	62	9
Volcano	95	37	53	5
See	215	144	58	13
Total	**1712**	**1052**	**593**	**67**

Note: Positive percentage = 61.45 %, Negative percentage = 3.91%

Table 5: Sentiment Analysis Results: Sport

Term	Review Count	Positive	Neutral	Negative
Circuit	68	35	31	2
Race	73	45	22	6
Track	59	33	23	3
Grandstand	32	20	9	3
Event	30	20	9	1
Football	17	7	10	0
Facility	7	6	1	0
International	34	17	14	3
Watch	23	15	7	1
Stadium	31	21	12	0
Total	**376**	**219**	**138**	**19**

Note: positive percentage = 58.25%, negative percentage = 5.05%

Table 6: Sentiment Analysis Results: Local Market

Term	Review Count	Positive	Neutral	Negative
Market	43	16	22	5
Bar	29	12	15	2
Night	40	26	13	1
Castle	29	12	17	0
Live	20	7	8	5
Street	14	7	5	2
Restaurant	39	21	15	3
Shop	67	32	33	2
Food	91	49	35	7
Local	61	23	32	6
Total	**433**	**205**	**195**	**33**

Note: positive percentage = 47.34%, negative percentage = 7.62%

Based on the sentiment percentages presented in Tables 4-6, the analysis reveals that the attributes of archaeological sites and sport receive high positive sentiment percentages and low negative sentiment percentages. On the other hand, the attribute of the local market exhibits a low positive sentiment percentage and a high negative sentiment percentage. This suggests that the local market attribute may not be the most favourable aspect to emphasis in Buriram's branding efforts. Instead, Buriram can position itself as a distinctive destination that combines modernisation with ancient history. By promoting its sports facilities and archeological sites, Buriram can attract a larger number of tourists interested in sports and historical experiences. Adopting this branding strategy will enable the TAT to develop marketing campaigns that showcase Buriram's unique offerings and appeal to the preferences of target visitors.

6. CONCLUSION, LIMITATIONS, AND FUTURE RESEARCH

The contribution of this study lies in bridging a gap in the existing literature concerning the relationship between destination branding, co-creation, and eWOM by using sentiment analysis. While previous researchers have explored the influence of eWOM on destination image (Dellarocas, 2003), they have largely overlooked the potential of sentiment analysis in identifying distinctive and memorable tourist experiences that can contribute to destination branding. This study addresses this gap by introducing an innovative approach that utilises sentiment analysis to discover unique experiences and attributes associated with a destination. The method has demonstrated its effectiveness in swiftly and accurately scrutinising significant volumes of text, revealing patterns that might not be immediately apparent to human readers. As a result, sentiment analysis proves invaluable for identifying distinctive features that can be utilised to enhance the global branding of destinations, with a particular focus on providing support to less well-known tourist destinations. Furthermore, the research methodology employed in this study can be readily adapted and implemented in various destination contexts.

In other words, for destination management organisations and tourism practitioners, the insights gained from sentiment analysis can play a crucial role in crafting destination branding strategies. By utilising sentiment analysis to identify the most positively perceived attributes and experiences associated with a destination, practitioners can direct their efforts towards highlighting and promoting these aspects. This focused approach can assist in constructing a more compelling and authentic destination image, which is particularly vital for emerging or lesser-known tourist destinations aiming to attract and engage a global audience. Moreover, the methodology presented in this study provides a practical tool for practitioners to evaluate and refine branding strategies based on tourists' sentiments, ultimately enhancing their destination's competitiveness in the tourism market.

Furthermore, by leveraging unsupervised LDA modelling of OTRs, the study identifies destination attributes with minimal bias and provides qualitative insights into the underlying terms associated with these attributes. In addition, the silhouette method, employed in this research, ensured an optimal number of attributes for the destination, while also extracting distinctive characteristics of each attribute. The integration of Textblob and Flair also evidently indicate the robustness in providing high accuracy for sentiment analysis of approximately 88.2%, which surpasses the average accuracy reported in tourism and hospitality research by 18.2% (Schuckert et al., 2015). This contribution not only expands the understanding of the role of eWOM in destination branding but also offers a practical methodology for researchers and practitioners to develop effective destination branding strategies based on the sentiment analysis of tourists' experiences on a global scale.

There are, however, limitations to this study. First, the methodology employed in this study aims to assist DMOs in creating effective branding strategies for lesser-known or emerging tourist destinations. Nonetheless, these destinations typically have a smaller number of reviews available, which can limit the applicability and generalizability of the findings. Moreover, the analysis in this study was conducted solely on English reviews, further restricting the pool of available data for destination branding purposes. Thus, it is crucial to recognise the constraint imposed by the limited number of reviews for these specific destinations. Second, it is essential for researchers to consider that unsupervised techniques yield dynamic attributes that are not firmly grounded in established theories. As a result, these attributes may change over time as new data is incorporated, indicating the need for ongoing evaluation and refinement. Additionally, it is worth noting that this research solely focused on capturing sentiment from the perspective of tourists. However, according to Giannopoulos et al. (2021), brand development is influenced by the hearts and minds of various stakeholders, extending beyond customers alone. Therefore, future research should consider incorporating sentiment analysis from other stakeholders to gain a comprehensive understanding of brand value. Exploring sentiments expressed in other languages and even through the use of emojis would also be valuable avenues for future investigation.

REFERENCES

Aaker, D. A. (1991), *Managing Brand Equity*. New York, NY: The Free Press.

Aitken, R. & Campelo, A. (2011). The four Rs of place branding. *Journal of Marketing Management, 27*(9/10), 913–933. doi: 10.1080/0267257X.2011.560718

Aljedaani, W., Rustam, F., Mkaouer, M. W., Ghallab, A., Rupapara, V., Washington, P. B., Lee, E., & Ashraf, I. (2022). Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry. *Knowledge-Based Systems*, 255. doi: 10.1016/j.knosys.2022.109780

Almeyda-Ibáñez, M., & George, B. P. (2017). The evolution of destination branding: A review of branding literature in tourism. *Journal of Tourism, Heritage & Services Marketing, 3*(1), 9–17. doi: 10.5281/zenodo.401370

Baka, V. (2016). The becoming of user-generated reviews: Looking at the past to understand the future of managing reputation in the travel sector. *Tourism Management, 53*, 148–162. doi: 10.1016/j.tourman.2015.09.004

Baloglu, S., & McCleary, K. W. (1999). A model of destination image formation. *Annals of Tourism Research, 26*(4), 868–897. doi: 10.1016/S0160-7383(99)00030-4

Barbierato, E., Bernetti, I. & Capecchi, I. (2022). analysing TripAdvisor reviews of wine tours: an approach based on text mining and sentiment analysis. *International Journal of Wine Business Research, 34*(2), 212–236. https://doi.org/10.1108/IJWBR-04-2021-0025

Blain, C., & Levy, S. E. (2005). Destination Branding: Insights and Practices from Destination Management Organisations. *Journal of Travel Research, 43*(4), 328–338. doi: 10.1177/0047287505274646

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77–84. doi: 10.1145/2133806.2133826

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*(1), 993–1022. doi: 10.5555/944919.944937

Cai, L. (2002). Cooperative branding for rural destinations. *Annals of Tourism Research, 29*(3), 720–742. doi: 10.1016/S0160-7383(01)00080-9

Chaigasem, T., & Leruksa, C. (2020). Ready to be a host?: The potential development guidelines of sport tourism in Buriram Province, Thailand. *African Journal of Hospitality, Tourism and Leisure, 9*(1), https://www.ajhtl.com/uploads/7/1/6/3/7163688/article_14_vol_9_1 2020_thailand.pdf

Chang, Y. -C., Ku, C. -H., & Chen, C. -H. (2020). Using deep learning and visual analytics to explore hotel reviews and responses. *Tourism Management, 80*, 104–129. doi: 10.1016/j.tourman.2020.104129

Curtis, J. (2001). Branding a state: The evolution of brand Oregon. *Journal of Vacation Marketing, 7*(1), 75–81. doi: 10.1177/135676670100700107

Dabade, M. S., Sale, M. D., Dhokate, D. D., & Kambare, S. M. (2021). Sentiment analysis of Twitter data by using deep learning and machine learning. *Turkish Journal of Computer and Mathematics Education, 12*(6), 962–970. doi: 10.17762/turcomat.v12i6.2375

Dellarocas, C. (2003). The digitisation of word-of-mouth: promise and challenges of online feedback mechanisms. *Management Science, 49*(10), 1407–1424.

Dhar, S. (2022). TripAdvisor Hotel Reviews 20k Dataset. https://www.kaggle.com/datasets/shoumikdhar/tripadvisor-hotel-reviews-20k-dataset. Accessed the 20th of November 2022, at 14:44.

Filieri, R., Alguezaui, S., & McLeay, F. (2015). Why do travellers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth. *Tourism Management, 51*, 174–185. doi: 10.1016/j.tourman.2015.05.007

Ganzaroli, A., Noni, I. D., & Baalen, P. V. (2017). Vicious advice: analysing the impact of TripAdvisor on the quality of restaurants as part of the cultural heritage of Venice. *Tourism Management, 61*, 501–510. doi: 10.1016/j.tourman.2017.03.019

Giannopoulos, A., Piha, L. & Skourtis, G. (2021). Destination branding and co-creation: a service ecosystem perspective. *Journal of Product & Brand Management, 30*(1), 148–166. doi: 10.1108/JPBM-08-2019-2504

Gilmore, F. (2002). Branding for success. Morgan, N., Pritchard, A., and Pride, R. (Eds.), *Destination branding: Creating a unique destination proposition*. Butterworth Heinemann. 57–65.

Gour, A., Aggarwal, S. & Erdem, M. (2021). Reading between the lines: analysing online reviews by using a multi-method Web-analytics approach. *International Journal of Contemporary Hospitality Management, 33*(2), 490–512.

Guo, X., Pesonen, J., & Komppula, R. (2021). Comparing online travel review platforms as destination image information agents. *Information Technology & Tourism, 23*, 159–187. doi: 10.1007/s40558-021-00201-w

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management, 59*, 467–483. doi: 10.1016/j.tourman.2016.09.009

Hall, D. (2002). Brand development, tourism and national identity: The re-imaging of former Yugoslavia. *Journal of Brand Management, 9*, 323–334. doi: 10.1057/palgrave.bm.2540081

Hedrick-Wong, Y. & Choong, D. (2016). Global Destination Cities Index by Mastercard. https://www.mastercard.com/news/media/1m1dffxp/mastercard-global-destination- cities-index-report-2016.pdf. Accessed the 8th of September 2022, at 12:44.

Hemalatha, S., & Ramathmika, R. (2019). Sentiment analysis of Yelp reviews by machine learning, *Paper presented at the 2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, Madurai, India, 700–704. doi: 10.1109/ICCS45141.2019.9065812

Javed, M. & Awan, T.M. (2023). The young tourist's co-creation nexus: market mavens and existential authenticity as driving forces of intentions to revisit and recommend. *Journal of Hospitality and Tourism Insights, 6*(2), 716–734. https://doi.org/10.1108/JHTI-12-2020-0240

Kaoungku, N., Suksut, K., Chanklan, R., Kerdprasop, K., & Kerdprasop, N. (2018). The Silhouette Width Criterion for Clustering and Association Mining to Select Image Features. *International Journal of Machine Learning and Computing, 8*(1), 69–73

Költringer, C., Dickinger, A. (2015). Analysing destination branding and image from online sources: A web content mining approach. *Journal of Business Research, 68*(9), 1836–1843. https://doi.org/10.1016/j.jbusres.2015.01.011

Kucukusta, D., Law, R., Besbes, A., & Legoherel, P. (2015). Re-examining perceived usefulness and ease of use in online booking: The case of Hong Kong online users. *International Journal of Contemporary Hospitality Management, 27*(2), 185–198. doi: 10.1108/IJCHM-09-2013-0413

Madyatmadja, E.D., Pristinella, D., Rahardja, N. & Ginting, R.B. (2021). Smart tourism services: a systematic literature review. *Paper presented at the 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI)*, 329–333

Manosso, F.C. & Domareski Ruiz, T.C. (2021). Using sentiment analysis in tourism research: A systematic, bibliometric, and integrative review. *Journal of Tourism, Heritage & Services Marketing, 7*(2), 16–27.

Mariani, M., & Borghi, M. (2020). Environmental discourse in hotel online reviews: A big data analysis. *Journal of Sustainable Tourism, 29*(12), 1–19. doi: 10.1080/09669582.2020.1858303

MICE Intelligence Center. (2019). Ready for sport city. https://intelligence.business eventsthailand.com/th /blog/ready-for-sport-city-blog. Accessed the 8th of September 2022, at 18:33.

Mir, I. A., & Rehman, K. U. (2013). Factors Affecting Consumer Attitudes and Intentions toward User – Generated Product Content on YouTube. *Management & Marketing Challenges for the Knowledge Society, 8*(4), 637–654.

Mishra, R. K., Urolagin, S., Jothi, J. A.A., Neogi, A. S., & Nawaz, N. (2021). Deep learning-based sentiment analysis and topic modelling on tourism during Covid-19 pandemic""", *Frontiers in Computer Science, 3*(14), 1–14. doi: 10.3389/fcomp.2021.775368

Mohammadi, F., Yazdani, H.R., Jami Pour, M. & Soltani, M. (2021). Co-creation in tourism: a systematic mapping study. *Tourism Review, 76*(2), 305–343. doi: /10.1108/TR-10-2019-0425

Mukhopadhyay, S., Pandey, R. & Rishi, B. (2023). Electronic word of mouth (eWOM) research – a comparative bibliometric analysis and future research insight", *Journal of Hospitality and Tourism Insights, 6*(2), 404–424. doi: 10.1108/JHTI-07-2021-0174

Nasa, M., & Hassan, F. B. (2016). Assessment of Tourism Resource Potential at Buriram Province, Thailand. *Asian Social Science, 12*(10), 27–34. doi: 10.5539/ass.v12n10p27

Nieto, J., Hernández-Maestro, R. M., & Muñoz-Gallego, P. A. (2014). Marketing decisions, customer reviews, and business performance: The use of the Top rural website by Spanish rural lodging establishments. *Tourism Management, 45*, 115–123. doi: 10.1016/j.tourman.2014.03.009

Oliveira, E. H., & Panyik, E. (2015). Content, context and co-creation: Digital challenges in destination branding with references to Portugal as a tourist destination. *Journal of Vacation Marketing, 21*(1), 53–74. doi: 10.1177/13567667145442

Oppermann, M. (2000). Tourism destination loyalty. *Journal of Travel Research, 39*(1), 78–84. doi: 10.1177/004728750003900110

Prameswari, P., Zulkarnain, Surjandari, I., & Laoh, E. (2017). Mining online reviews in Indonesia's priority tourist destinations using sentiment analysis and text summarization approach. *Paper presented at the 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST 2017)*, Taichung, Taiwan, 121–126. doi: 10.1109/ICAwST.2017.8256429

Puh, K. & Bagić Babac, M. (2022). Predicting sentiment and rating of tourist reviews using machine learning. *Journal of Hospitality and Tourism Insights, 6*(3), 1188–1204. doi: 10.1108/JHTI-02-2022-0078

Ritchie, J. R. B., & Ritchie, J. R. R. (1998), "The branding of tourism destinations", *Paper presented at the Annual congress of the international association of scientific experts in tourism (AIEST)*. Marrakech, Morocco.

Rousseeuw, P. J. (1987). Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Computational and Applied Mathematics, 20*, 53–65. doi: 10.1016/0377-0427(87)90125-7

Schuckert, M., Liu, X., & Law, R. (2015). Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing, 32*(5), 608–621, doi: 10.1080/10548408.2014.933154

Sigala, M. (2009). Web 2.0, social marketing strategies and distribution channels for city destinations: enhancing the participatory role of travellers and exploiting their collective intelligence. M. Gascó-Hernández & T. Torres-Coronas. (Ed.s), *Information Communication Technologies and City Marketing: Digital Opportunities for Cities Around the World*, IGI Global, 221–245.

Taecharungroj, V., & Mathayomchan, B. (2019). Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand. *Tourism Management, 75*, 550–568. doi: 10.1016/j.tourman.2019.06.020

TAT Buriram. (2019), Buriram. https://www.amazingthailandebook.com/files/book_125/75f8cccf31c9fc6c533fd2f707f452d81566376494268.pdf. Accessed the 8th of September 2022, at 14:44.

Wise, N., & Farzin, F. (2018). Authenticity & Tourism: Materialities, Perceptions, Experiences. J. M. Rickly & E. S. Vidon, (Eds.). *Emerald Publishing Limited*, Bingley, 33–52.

Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management, 58*, 51–65. doi: 10.1016/j.tourman.2016.10.001

Yağmur, Y., & Aksu, A. (2022). Investigation of destination image mediating effect on 'tourists' risk assessment, behavioural intentions and satisfaction. *Journal of Tourism, Heritage & Services Marketing, 8*(1), 27–37.

Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior, 27*(2), 634–639. doi: 10.1016/j.chb.2010.04.014