Machine learning-based model for customer emotion detection in hotel booking services

Abstract

Purpose – The purpose of this paper is to expand and analyze deeply customer emotions, concretize the levels of positive or negative emotions with the aim of using machine learning methods, and build a model to identify customer emotions.

Design/methodology/approach – The study proposed a customer emotion detection model and data mining method based on the collected dataset, including 80,593 online reviews on agoda.com and booking.com from 2009 to 2022.

Findings – By discerning specific emotions expressed in customers' comments, emotion detection, which refers to the process of identifying users' emotional states, assumes a crucial role in evaluating the brand value of a product. The research capitalizes on the vast and diverse data sources available on hotel booking websites, which, despite their richness, remain largely unexplored and unanalyzed. The outcomes of the model, pertaining to the detection and classification of customer emotions based on ratings and reviews into four distinct emotional states, offer a means to address the challenge of determining customer satisfaction regarding their actual service experiences. These findings hold substantial value for businesses operating in this domain, as they facilitate the evaluation and formulation of improvement strategies within their business models. The experimental study reveals that the proposed model attains an exact match ratio, precision, and recall rates of up to 81%, 90%, and 90%, respectively.

Research limitations/implications – the study has yet to mine real-time data. Prediction results may be influenced because the amount of data collected from the web is insufficient and preprocessing is not completely suppressed. Furthermore, the model in the study was not tested using all algorithms and multi-label classifiers. Future research should build databases to mine data in real-time and collect more data and enhance the current model.

Practical implications – Our results suggest that the emotion detection models can be applied to the real world to quickly analyze customer feedback. The proposed models enable the identification of customers' emotions, the discovery of customer demand, the enhancement of service, and the general customer experience. The established models can be used by many service sectors to learn more about customer satisfaction with the offered goods and services from customer reviews.

Social implications – The research paper helps businesses in the hospitality area analyze customer emotions in each specific aspect to ensure customer satisfaction. In addition, managers can come up with appropriate strategies to bring better products and services to society and people. Since then, developing the country's hotel tourism industry, promoting the national economy to develop in a sustainable way.

Originality/value – This study developed a customer emotions detection model for detecting and classifying customer ratings and reviews as 4 specific emotions: happy, angry, depressed, and hopeful based on online booking hotel websites agoda.com and booking.com that contains 80,593 reviews in Vietnamese. The research results help businesses check and evaluate the quality of their services, thereby offering appropriate improvement strategies to increase customers' satisfaction and demand more effectively.

Keywords: classification models, emotion detection, hotel booking services, machine learning, natural language processing

Paper type: Research paper

1 Introduction

Today, emotions are becoming especially crucial in shaping customers' experiences (Tuerlan *et al.*, 2021). Research on customer emotion has been done in the fields of hospitality and tourism management in an effort to understand how it affects the customer experience. Customers are influenced by marketing materials at the pre-consumption stage (Bastiaansen *et al.*, 2019), their emotional reactions to the creative process play a crucial part in determining their final decision over where to go (Walters *et al.*, 2012). During the consumption stage, emotions are key in forming memories of the associated experiences, and these emotions can vary throughout the experience (Servidio and Ruffolo, 2016). Mattila and Enz (2002) extended the understanding of how a customer's attitude and emotions impact how they evaluate the service provider and their whole service experience. They made hypotheses about the impact of emotions on customers. From this, the authors come to the conclusion that consumers' assessments of the customer experience are closely correlated with the emotions they expressed throughout the transaction and their post-experience moods. Therefore, businesses must pay close attention to customer emotions as they expand their businesses and train their workforce.

Kim and Fesenmaier (2015) suggested that emotions are crucial in shaping the overall experience of customers. Customers' emotions were monitored in real-time during their journey, and it was found that they changed as they visited websites that catered to various interests and experiences. Positive consumption emotions increase customers' contentment at the post-consumption stage, word-of-mouth (WOM) intention, and loyalty (Yüksel and Yüksel, 2007). Electronic word of mouth (eWOM) has emerged as a result of the advent and expansion of the Internet, which is regarded as one of the most significant informal media among customers, businesses, and the general population (Huete-Alcocer, 2017). Electronic Word of Mouth (eWOM) has occupied an important and essential position in the choice decision of customers (Xu, 2020). With the strong development of electronic commerce (e-commerce), feedback and reviews from customers increasingly become more crucial, directly affecting the customer's buying behavior, and online hotel booking service is also no exception (D'Acunto *et al.,* 2020; Xue *et al.*, 2021).

Future travelers will differ greatly from those of today as they become gradually more customized, time-driven, and technologically demanding. One possibility is that travelers would assign a number of planning duties to their own avatars, freeing them time to devote to other activities, such as taking in the sights while on vacation (Stylos, 2020). The critical roles of hotel service quality and related customer satisfaction are also underlined as significant antecedents in guests' decision-making processes and in enhancing customer loyalty. (Priporas *et al.,* 2017). The analysis of feedback and reviews is one of the concerns of many businesses (Ali and Anwar, 2021). They can significantly advance our knowledge of tourist decision-making and broaden our perspectives on tourism marketing (Stylos *et al.,* 2021). If a company doesn't care about customer satisfaction, then it can't expect their interest in its products (Demir *et al.*, 2020). However, with plenty of user feedback, it can be challenging for companies and organizations to pay attention, and categorizing and evaluating it requires a lot of time and work (Pantano *et al., 2017;* Nguyen *et al.*, 2021).

Thai *et al.* (2021) stated that customers' online reviews must be automatically gathered and used by businesses in the Big Data boom in order to track client shopping patterns and gauge their happiness with the caliber of goods and services. However, to be able to analyze the data source of these comments is still a tough problem for businesses. Chatterjee (2020) has shown that online hotel reviews (OHRs) are authentic, fast, and plentiful, but they can be challenging to assess for relevant insights. Many previous ED studies have taken advantage of large data sources from many fields such as comments on social networks, reviews on online food sites, etc. Despite being a rich source of data, online feedback reviews on hotel booking websites are rarely exploited and researched in depth. On the other hand, users can show emotions through many methods such as text, images, video, and audio… Therefore, classical text analysis methods, such as sentiment analysis are not enough to accurately explain the meaning conveyed in each message (Nguyen and Ho, 2021). Customers' emotions such as joy, fear, surprise, guilt, etc., play a vital role in understanding how real-life experiences affect customers. Accurately detecting and determining an individual's emotional or mental state is crucial (Nandwani and Verma, 2021).

Zad et al (2021) believe that one of the fastest-growing branches of Natural Language Processing (NLP), text-based emotion detection (TBED), involves grouping syntactic or semantic components of a corpus into a set of emotional categories indicated by a psychological model. Therefore, research by Haryadi and Kusuma (2019) has shown that using machine learning techniques, we can use computers to learn emotions from text. In machine learning, computers are trained to analyze examples and predict emotions rather than using a set of preprogrammed rules to solve issues.

This research expands and analyzes deeply into customer emotions, concretizes the levels of positive or negative emotions with the aim of using machine learning methods and building a model to identify customer emotions based on the collected data including 80,593 reviews in Vietnamese on hotel services websites agoda.com and booking.com. According to the study of (Nguyen and Tran, 2021), foreign online travel agencies (OTAs), especially those with headquarters in the United States, dominate the Vietnamese market. Two of the leading OTAs, Agoda and Booking, control more than 80% of the Vietnamese market for online bookings. Additionally, hotels in Vietnam are gaining market share, accounting for 30–40% of all visitors to each establishment. Therefore, there is an immense volume of tourist reviews that are valuable for analysis. Data collected on two e-commerce platforms were recorded from 2009 to 2022. In particular, data from 2017 and above accounted for the majority of 92.39%. Therefore, the data still maintains its novelty.

After being collected, the data will be preprocessed and labeled according to four types of emotions (happy, angry, depressed, hopeful). Next, machine learning methods are applied to build an emotion detection model from the above data. The purpose of this research is to identify and analyze actual customer experiences with products and services related to online hotel reservations. The study concentrated on four typical emotion types and the factors that matter to customers. Turn rich data sources from websites into knowledge that is actually helpful to businesses when developing products. The structure of this paper is divided into 5 sections. Section 1 introduces the topic, and its urgency, as well as summarizes previous studies, and highlights the new points of this study. The relevant theoretical background is presented in Section 2. In Section 3, the steps of data processing and model building will be clearly demonstrated. The results and significance of the study are detailed in Section 4. Finally, Section 5 summarizes and points out the contributions and limitations of the study.

2 Literature Review

2.1 Review of sentiment analysis research

Sentiment analysis (SA) is the process of gathering and analyzing people's opinions, feelings, attitudes, perceptions, etc., about various issues, products, and services. The rapid growth of Internet-based applications like websites, social networks, and blogs has led to a tremendous increase in the amount of thoughts and reviews that people are creating about products, services, and everyday activities (Birjali *et al.*, 2021). As a result, much research has been conducted regarding opinion mining using user comments and feedback via e-commerce platforms.

In the field of tourism, Nguyen and Ho (2023) have found the issues that customers care about and the set of keywords that have a high probability of attracting customers, assisting companies in the hotel industry to raise the standard of goods and services in accordance with customer preferences. The limitation of this study is that the model has not collected and classified customer comments in real time, leading to the inability to offer strategies to deal with timely issues.

Customers have provided feedback on the results over time based on various internet comments. In restaurants, hospitals, schools, and other public and private establishments, customers have expressed their opinions about goods and services. (Nguyen *et al.*, 2021) examined studies on opinion mining and recommended the application of a machine learning strategy to Vietnamese customer feedback. In this study, data from automatic programs, including 236,867 user reviews of online ordering services and eating establishment review websites in Vietnam, namely foody.vn and diadiemanuong.com, are analyzed using the knowledge mining method.

Sentiment analysis is now widely accepted by scholars, businesses, governments, and other groups (Birjali *et al.*, 2021). Sentiment analysis of review comments identifies polarity and classifies the emotions expressed in opinionated texts as positive, negative, or neutral (Xu and *et al.*, 2019). However, if you just stop here, it is still not specific to the level of customer emotions. For example, joy, fear, surprise, guilt, and other such emotions cannot be detected by sentiment analysis, which is essential to understand how real-life experiences affect customers. Therefore, the need to accurately detect and determine an individual's emotional or mental state (Nandwani and Verma, 2021).

2.2 Basic emotions model

Emotion models are the foundations of ED systems. The models make the assumption that emotions have different states, necessitating the need to discriminate between them. When engaging in any ED-related activity, it is crucial to establish the model of emotion that will be used (Acheampong *et al.*, 2020). The study's emotional model is based on Russell's (1980) model, which included 28 types of emotions as shown in Figure 1.

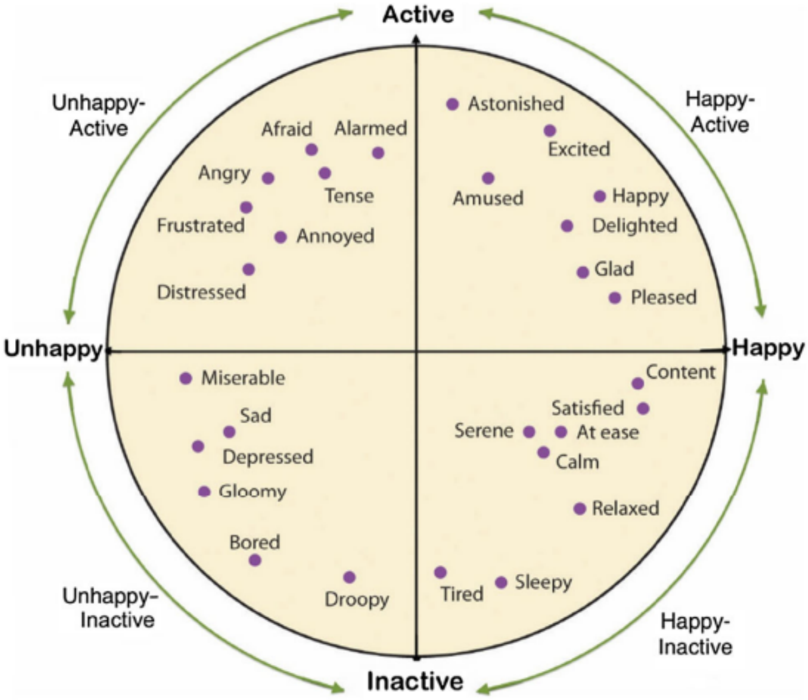

Figure 1. Circumplex model of effect including 28 affect words (Source: Russell, 1980)

In the work of Bagozzi *et al.* (1999), the authors emphasize the importance of emotions in the field of marketing as consumer decision-making processes and their behaviors are strongly influenced by emotions. The first point is how positive emotions can enhance consumers' attitudes and increase their likelihood of engaging with a particular product or brand. The positive emotions retrieved had a greater impact on decision-making in low-involvement situations compared to high-involvement contexts. The study demonstrated that viewers in a positive emotional state were more likely to recall the associated emotions when provided with the brand name as a retrieval cue, compared to viewers who were not in a positive emotional state when exposed to the ad. Conversely, negative emotions, such as fear or sadness, may lead to avoidance or rejection of a product. Emotions, as a mental state, activate individuals for action and provide the motivation to engage in behaviors, ultimately influencing human behavior. Furthermore, emotions have a significant impact on goal setting, specifically in three classes of goals: focal goals, subordinate goals, and superordinate goals. Focal goals represent what individuals strive for, while subordinate goals are the means to achieve them. Superordinate goals provide reasons for pursuing focal goals, with happiness being a central motive. Happiness, along with other emotions, plays a crucial role in customer satisfaction.

Experiences in travel are fundamentally emotional. It has an impact on different stages of the travel experience. However, emotional experiences fluctuate considerably over time (Hosany *et al.*, 2021; Ratnasari *et al.*, 2020). This study focuses mainly on 4 main types of emotions: happy, angry, depressed, and hopeful.

The needs of persons traveling differ from those of those going about their daily lives in the area of tourism. There are two major divisions in travel demand. both those that came before and those that came after the journey. Prior to the trip, the needs that drive the traveler's motivation will determine their decision. Travel motives might be thought of as these necessities. According to actual demand, travel engines are modified for specific locations (Lazoura *et al.*, 2021).

Along with the successful growth of upscale hotels and expanding transit decisions, travelers are embracing relaxation and taking pleasure in their travels. Today, some VIP clients are prepared to spend more and would rather go for a more pleasurable encounter. Travelers will be increasingly picky about how they organize their vacations and accommodation selections (Cheng and Jiang, 2021). Additionally, this implies that tourists will expect higher prices. They will then make more proposals with more ambitious demands.

2.3 Text-based emotion detection

Nandwani and Verma (2021) argued that emotion detection is a technique for recognizing various human emotion types, such as anger, joy, or depression. Emotions are an essential part of life for humans. The emotions from customer opinion transform into meaningful insight and can help managers in decision-making. A key component of communication is emotion, which can be expressed in a variety of ways. One of the fastest-growing areas of Natural Language Processing (NLP), Text-Based Emotion Detection (TBED), is the classification of syntactic or semantic units of a corpus into a specific set of emotion classes suggested by a psychological model (Zad *et al.* ,2021). Jayakrishnan *et al.* (2018) classified collected data from Malayalam (an Indian dialect) into one of the emotions of joy, sadness, fear, anger and surprise. The author pre-processed the data, extracted features, trained the data and evaluated it by using 500 Malayalam sentences for each emotion. Using the SVM classifier, then, they classified the texts into different emotional classes, their results showed the accuracy of 0.94, 0.92, 0.90, 0.93 and 0.90 for the emotions of joy, sadness, anger, fear, and surprise respectively. However, research studies are limited to semantic information in texts and do not consider the role of context in sentences.

To exploit customer opinions in the Vietnamese language, a study by Ho *et al.* (2020) focuses a higher approach on a special case of sentiment analysis. The results of sentiment analysis or prediction do not stop at two emotional aspects: positive or negative or in the form of ratings, but at a more detailed level of analysis such as emotions of sadness, joy, anger, disgust, fear, and surprise. The study has built a standard Vietnamese social media emotion database (Vietnamese Social Media Emotion Corpus - UIT-VSMEC) with data including 6,927 sentences annotated with emojis, contributing a significant part. into natural language processing (NLP) of emotion recognition studies in Vietnamese, which is a language with few sources to exploit. By experimental methods of evaluating and measuring machine learning models and convolutional neural networks (CNN), the research results have achieved the highest performance with a weighted F1 score of 59.74%.

2.4 Machine learning approach to emotion detection

One of the most challenging jobs in natural language processing (NLP) and one that involves natural language understanding (NLU) is emotion detection in the text, particularly implicit emotion detection (Alswaidan and Menai, 2020). Figure 2 shows that ML algorithms are used to categorize texts into different emotion groups in order to tackle the ED problem. Both supervised and unsupervised ML techniques are frequently used for detection (Acheampong *et al.*, 2020). These approaches for unsupervised machine learning (TUML), include support vector machines (SVM), Logistic Regression (LR), Decision Tree (DT), Random Forest (RF) and so on.

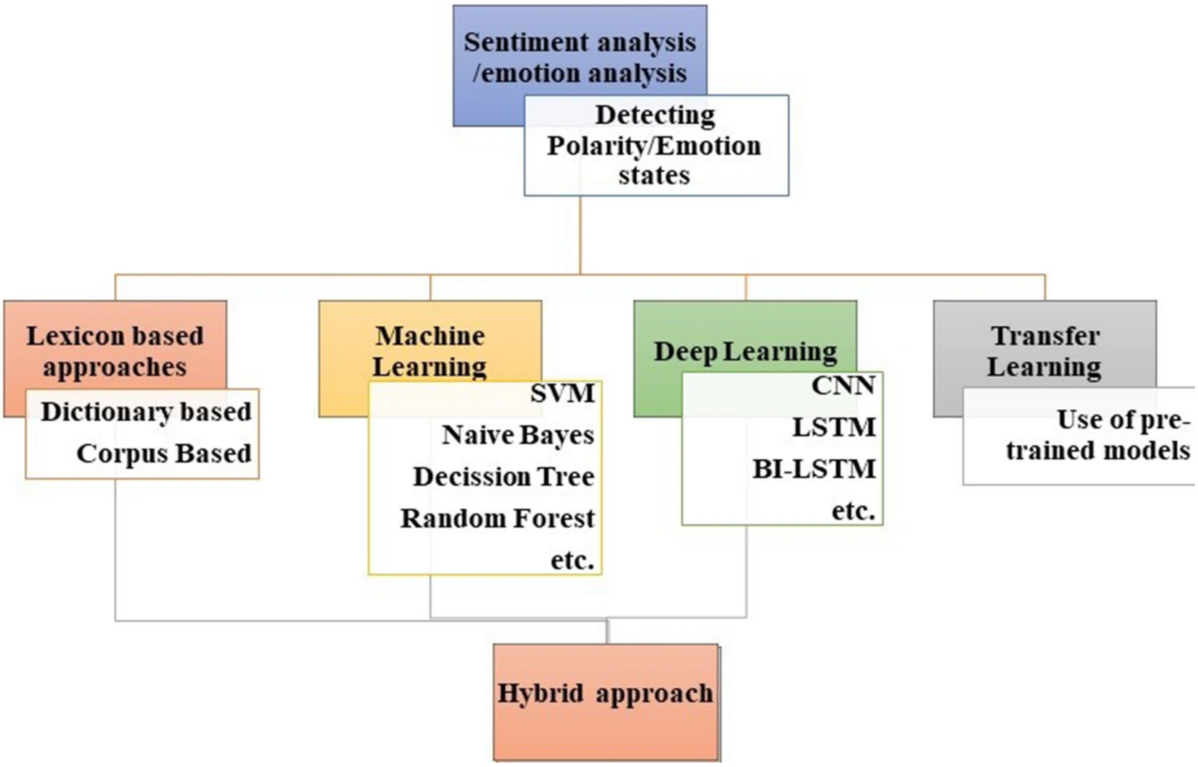

Figure 2. Techniques for sentiment analysis and emotion detection (Source: Nandwani and Verma, 2021)

For text categorization, a variety of classification techniques have been used, including logistic regression, decision trees, Bayesian classifiers, nearest neighbor classifiers and support vector machines (Hasan *et al.*, 2019). To classify emotion, we explored three different classifiers. We selected Logistic Regression as a probabilistic classifier, SVM as a decision boundary classifier, and decision tree as a rule-based classifier.

2.4.1 Support Vector Machine

Support Vector Machine (SVM) is one of the relatively new and promising methods for learning decomposition functions in pattern recognition (classification) tasks or for performing function estimation in regression problems (Ukil, 2007). However, this method is mainly used in solving classification problems. SVM is considered a powerful classification method (Guo and Chou, 2020).

SVM is essentially an optimization problem, the goal of this algorithm is to find an optimal hyperplane with the maximum margin that serves as the decision boundary to separate two different classes.

2.4.2 Logistics Regression

The logistic regression model is a continuation of the idea of linear regression into classification problems. Logistic regression needs a function that projects the predicted value on a probability space in the range [0,1] and at the same time creates nonlinearity for the regression equation to help it have a dividing boundary between two groups better. From the output of the linear function, we put on the Sigmoid function to find the probability distribution of the data. The Sigmoid function is only used in binary classification problems. For the classification problem of more than two labels, the Softmax function which is a generalized form of Sigmoid will be used.

2.4.3 Decision Tree

Decision Tree is a Supervised Machine Learning strategy to address classification and regression issues by repeatedly separating data based on a particular parameter. The decisions are in the leaves and the data is split into the nodes. The Decision tree has the following advantages: suitable for regression as well as classification problems, easy to interpret, easy to handle categorical and quantitative values, and capable of fully handling missing values in the most probable, high-performance valuable properties (Ray, 2019). The Decision tree is one of the power-enhancing methods commonly used in various fields, such as machine learning, image processing, and identification of patterns (Damanik *et al.*, 2019; Charbuty and Abdulazeez, 2021).

2.5 Multi-label classification approach by problem transformation methods

Multi-label classifier machine learning methods are divided into two main groups, including problem transformation and algorithm adaptation. In this study, the problem transformation method is used because of its simplicity when it converts the problem from multi-label classification to binary or multi-class single-label classification. This method has high flexibility (Gibaja and Ventura, 2014) when it can use any single-label algorithms such as Support Vector Machine, Logistic Regression, k-Nearest Neighbor, Decision Tree, Random Forest. These methods are divided into the following techniques in turn:

2.5.1 Binary Relevance (BR): The first technique Binary Relevance transforms the problem from multi-label classification into a single-label classification problem. The number of classifiers needed for the binary compatibility problem is equal to the number of labels. This classification algorithm will work in parallel, independent of each other to classify the binary corresponding to the label in the label space as Figure 3 shown. Therefore, this technique has the advantage that it is easy to perform, but also has the disadvantage of ignoring the correlation between labels (Gharroudi, 2017).

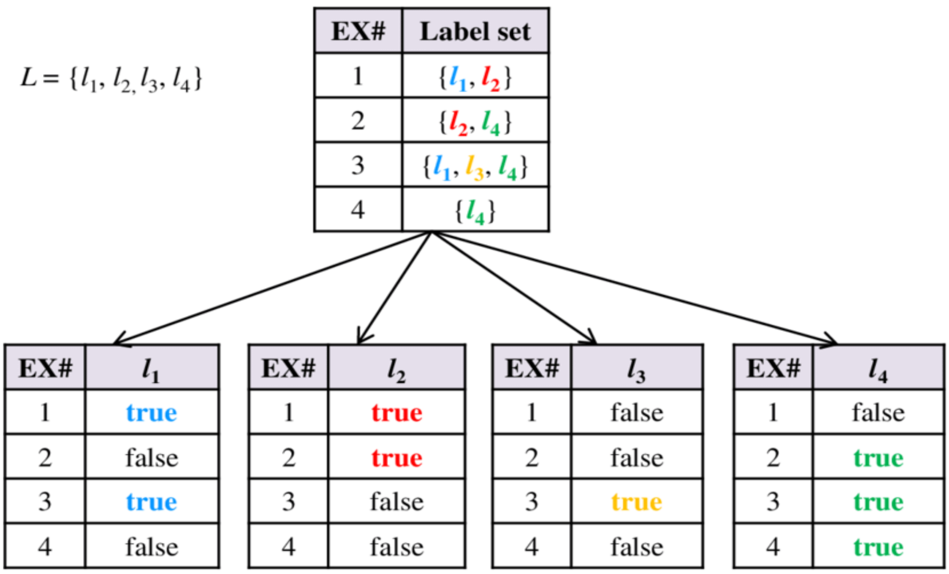

Figure 3. Binary Relevance illustration (Source: Vidulin, 2013)

2.5.2 Label Powerset (LP): The second technique is the label powerset, which transforms the problem from classifying multiple labels into a multi-class classification problem. This technique requires 2^N a classifier that operates with N being the number of labels present in the data set, and 2^N is the number of possible classes. As the number of classes increases, the number of distinct combinations of labels can increase exponentially, easily leading to combinatorial explosion and thus making computation infeasible. Figure 4 presents the principle of Label Powerset technique. In this study, the total number of labels is only 4, so this technique is perfectly suitable.

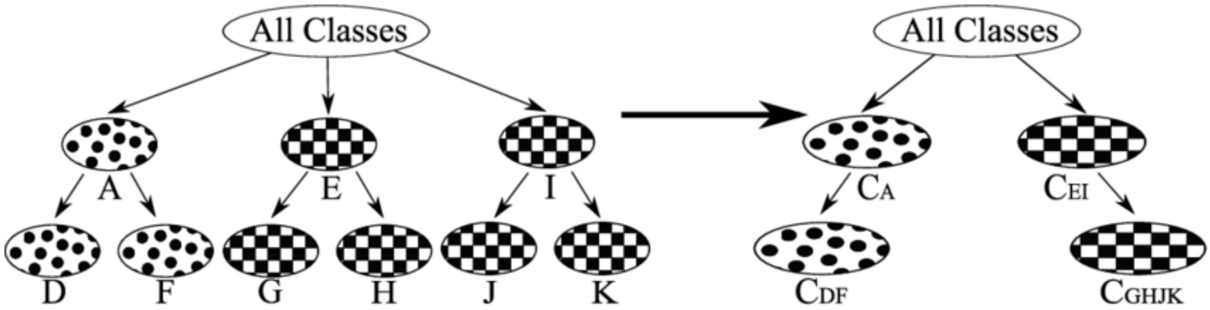

Figure 4. Label Powerset illustration (Source: Cerri *et al*., 2011)

2.5.3 Classifier Chain (CC): The last technique is classifier chaining, somewhat similar to binary compatibility, but instead of being trained in parallel, the classifiers are trained sequentially as shown in Figure 5. The total number of classifiers needed for this approach is equal to the number of classes, but training the classifiers is more involved. Because many sequences are executed sequentially, the correlation between labels will not be ignored but has the disadvantage of not executing in parallel (Gibaja and Ventura, 2014).

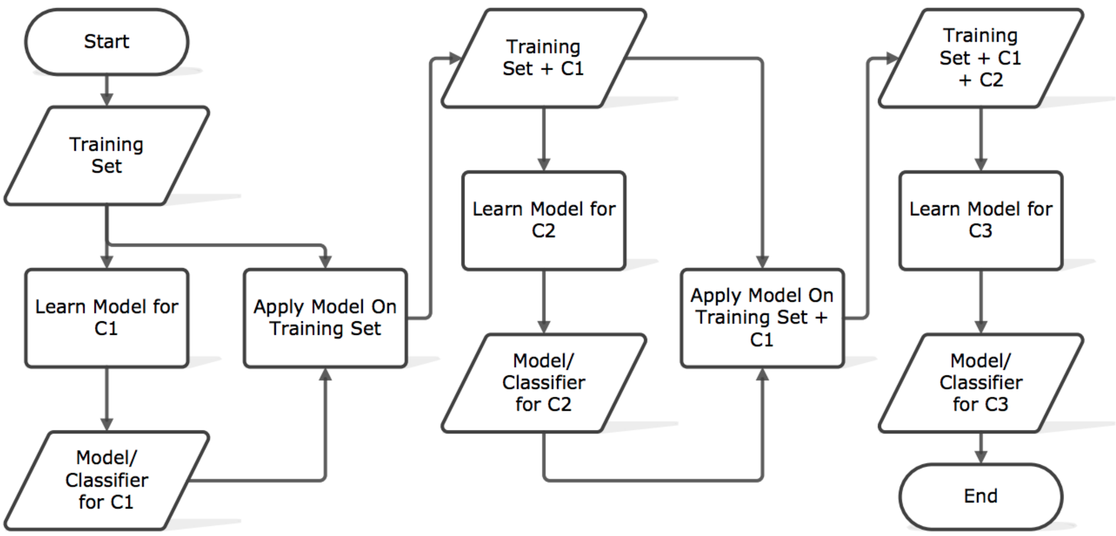

Figure 5. Classifier Chain illustration (Source: Kariuki, 2021)

3 Research Methodology

The study used a qualitative method. The research team examined and synthesized the results of related studies on the topic of emotion detection and found gaps in their research. In addition, data collection was collected in July 2022, with a data set of 80,593 reviews spanning from 2009 to the time of data collection. In particular, data from 2017 and above account for the majority of 92.39%. In addition, the research team also used a quantitative method when building a model to identify customer emotions based on the collected data. Based on the original purpose, the result of this model is to identify each customer's emotions after experiencing the hotel service. Then, compare the predicted results with the actual results, and consider the accuracy rate to evaluate the effectiveness of the model.

3.1 Overview model

First, data is collected by the research team from reviews through the hotel service websites for research purposes. Before implementing machine learning methods, it is necessary to process, sample and label them. The training data is used to train machine learning, and the test dataset is a type of data to test the accuracy of the model. After building the model, testing and selecting the best model, the model will be used for data mining. Finally, visualization is used to show the performance of the model. The steps are described in Figure 6.

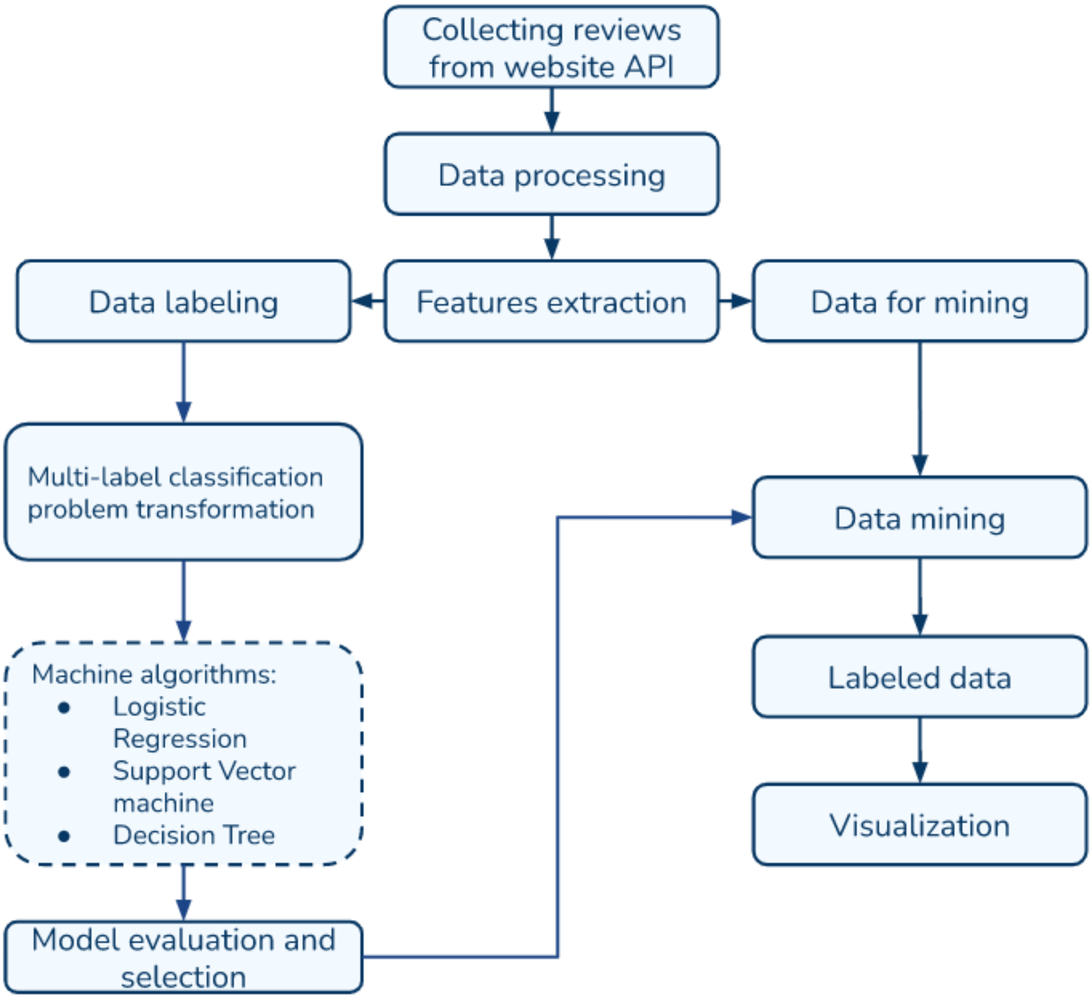

Figure 6. Overview of research model (Sources: Authors)

3.2 Implementation process

The research is carried out through 5 phases including data preparation, data preprocessing, data analysis, data mining, and result visualization.

- *Stage 1: Data preparation*

The experimental corpus in the study was collected from the websites of companies providing online hotel booking services including booking.com and agoda.com. According to the study of Nguyen and Tran (2021), the Vietnamese market is being dominated by foreign online travel agents (OTAs), primarily those based in the United States. Agoda and Booking, two of the top OTAs, hold over 80% of the market share for online reservations in Vietnam. Furthermore, the market share for hotels in Vietnam is on the rise, representing 30-40% of the total number of guests per hotel. Therefore, there is an immense volume of tourist reviews that are valuable for analysis. In the data collection step, the research team used the *aiohttp* and t*hebeautifulsoup* libraries from the Python programming language to access the website's API to collect customer reviews from the reviews section. The response data from the API is JSON and then converted and aggregated into an Excel file with the extension .xlsx with important extracted attributes such as hotel_id, comments, rating, and review_date. After implementation, 80,593 customer reviews were collected including 50,918 reviews from Agoda and 29,675 from Booking used to build the model.

- *Stage 2: Data preprocessing*

Collected data is raw one that has not been processed so it is not completely ready to use for models (Nguyen and Ho, 2021). The research team conducted preprocessing of the data so that the model could work best. First, duplicate and lost data that were not statistically significant were removed. In the next step, special characters such as icons and punctuation will be removed, and sentences will be converted to lowercase or uppercase because they greatly affect the performance of the model. For the Vietnamese language, words with punctuation marks will affect their meaning, so comments that are not added by the user will be re-added using the Pyvi library. Words continue to be standardized to their normal form, for example, the word “không” (no) can be written by users in many different ways such as “hok”, “khum”, “hông”, “khônggg”... do reduce model performance. Also, some words look the same but they use different Unicode codes, causing the computer to see them as two different words. The next step is word encoding, because Vietnamese features compound words, the order of words is very important, so the research team created a word code that is paired together as “friendly”. As a final step, stop words that do not provide analytical meaning are discarded. The preprocessing process is illustrated in the steps in Table 1.

Table 1. Illustrate the data preprocessing process (Source: Authors)

Process	Output
Raw data	Nhan vien ksan than thien va de thuong, toi thich lam!!!
Remove, change special characters	Nhan vien ksan than thien va de thuong toi thich lam
Add accents	nhân viên ksan thân thiện và dễ thương tôi thích lắm
Complements	nhân viên khách sạn thân thiện và dễ thương tôi thích lắm
Tokenization	“nhân_viên”, “khách_sạn”, “thân_thiện”, “và”, “dễ_thương", “tôi”, “thích”, “lắm”
Remove stop word	“nhân_viên”, “khách_sạn”, “thân_thiện”, “dễ_thương”, “thích”, “lắm”

- *Stage 3: Data Analysis*

To prepare the data set for the machine learning model, the comments are manually labeled, then label classification for the remaining data. At the same time, to ensure the accuracy of the model, the comments after being labeled continue to be cross-checked. The dataset used to train the model includes 40,176 comments. It can be seen that the comment data labeled hope and anger is less than happy and depressed, as shown in Figure 7.

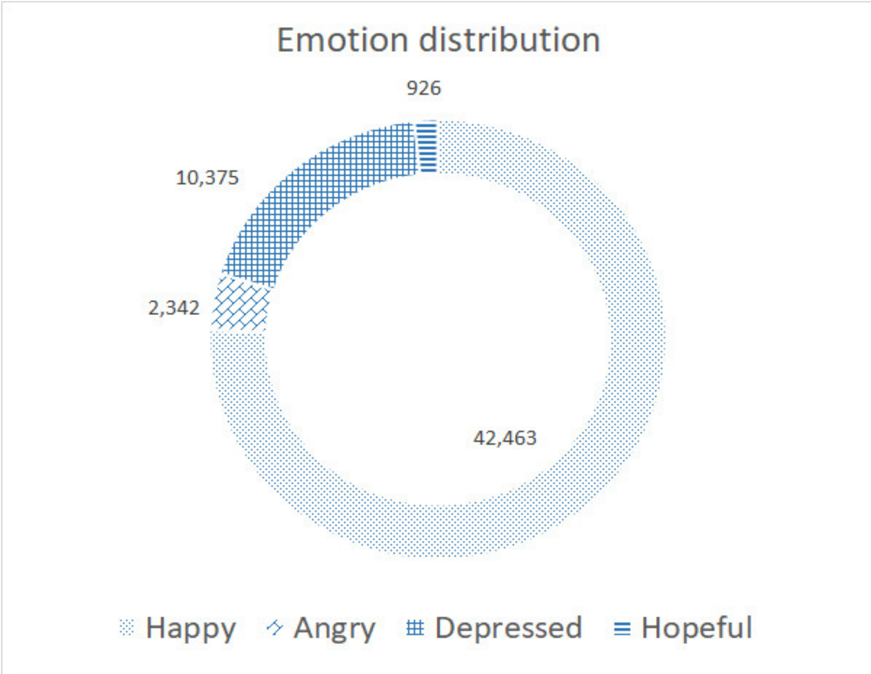

Figure 7. Emotion Distribution (Source: Authors)

- *Stage 4: Data mining*

After finishing the training model, the mined data is used to classify the remaining data for the information collection process.

- *Stage 5: Visualization of results*

The overview of the review presents model results with each matching algorithm that uses word cloud dashboards to clearly represent each aspect of all types of research emotions.

4 Results and Discussion

4.1 Context

The Vietnamese language has few resources for natural language processing (Nguyen *et al.*, 2020). The data of this study contains only Vietnamese, so we had several examples of Vietnamese language compared to English translation shown in Table 2.

Table 2. Several examples for Vietnamese language compared to English translation (Source: Authors)

Vietnamese	English meaning
Vui vẻ	Happy
Giận dữ	Angry
Chán nản	Depressed
Hy vọng	Hopeful
Đẹp	Beautiful
Tốt	Good
Xuất sắc	Excellent
Nhân viên	Staff
Nhiệt tình	Enthusiastic
Thân thiện	Friendly
Tàm tạm	So - so
Tệ	Terrible
Thất vọng	Disappointed
Giá	Price
Không	Not
Vị trí	Location
Hôi	Stinky
Lễ tân	Front desk
Hơi	Quite
Thêm	Add
Cải thiện	Improve

4.2 The result of model

After being trained, the assessment sample will be labeled illustratively in Table 3:

Table 3. Label sample example (Source: Authors)

Comment	**Label**
Khách sạn 5 sao gì mà hôi và tệ, nhân viên bất lịch sự. Thật thất vọng!!! Tôi sẽ không đến lần sau. **English:** What a smelly and bad 5-star hotel with impolite staff. Very disappointing!!! I won't come next time.	Angry
Đồ ăn ngon, nhân viên phục vụ cũng thân thiện. Mong lần sau đến khách sạn có cải tiến bữa ăn sáng cho đa dạng hơn. **English:** The food is delicious, and the staff is also friendly. Hope the next time I come to the hotel, the breakfast will be improved with a wider range of selections.	Happy, hopeful
Kỳ nghỉ khá chán khi phòng bên cạnh cách âm không tốt, làm mình không tận hưởng hoàn toàn nhưng bù lại, nhân viên rất dễ thương và giảm giá buổi sáng cho mình. **English:** The trip was quite boring when the next room was not soundproof, making me not fully enjoy it, but in return, the staff was very nice and gave me a discount in the morning.	Depressed, happy

The indicators used to measure include the equations (1), (2), (3), (4):

Let T be a multi-label dataset consisting multi-label examples (x_i, Y_i), $1 \le i \le n$, $(x_i \in \chi, Y_i \in \gamma = \{0,1\}^k)$, with a label set $\mathcal{L}$, $|\mathcal{L}| = k$. Let h be a multi-label classifier and $Z_i = h(x_i) = \{0,1\}^k$ be the set of label memberships predicted by h for the example x_i.

Exact match ratio (MR): is an index representing the percentage of comments that are completely correct in 4 labels. Comments that are only partially correct (3 or 2 labels) will not be counted. The formula for calculating the MR index is shown in equation (1):

$$MR = \frac{1}{n}\sum_{i=1}^{n} I(Y_i = Z_i) \qquad (1)$$

Precision (P): is the ratio of the total number of correctly predicted model TP scores of all labels to the total number of TP + FP predicted model scores of all labels. The formula for calculating the P index is shown in equation (2):

$$P = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i \cap Z_i|}{|Z_i|} \quad (2)$$

Recall (R): is the ratio of the correctly predicted model TP scores of all labels to the total number of initially labeled Positive scores of all labels (TP + FN). The formula for calculating the R index is shown in equation (3):

$$R = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i \cap Z_i|}{|Y_i|} \quad (3)$$

F1-score: is the harmonic mean, used to evaluate the overall model more if both Recall and Precision are important. The formula for calculating the F1-score is shown in equation (4):

$$F1 - score = \frac{2 \times precision \times recall}{precision + recall} \quad (4)$$

Table 4. Evaluation results for each model (Source: Authors)

Algorithms	Support Vector Machine			Decision Tree			Logistic Regression		
	BR	**CC**	**LP**	**BR**	**CC**	**LP**	**BR**	**CC**	**LP**
MR	0.80	**0.81**	0.79	0.70	0.71	0.70	0.78	0.78	0.80
Precision	**0.90**	**0.90**	**0.90**	0.84	0.84	0.84	**0.90**	**0.90**	**0.90**
Recall	**0.90**	**0.90**	0.89	0.84	0.84	0.84	0.88	0.89	**0.90**
F1 - score	**0.90**	**0.90**	**0.90**	0.84	0.84	0.84	0.89	0.89	**0.90**
Training time (seconds)	9.3	22.3	**5.3**	1362.1	335.9	299.1	55.7	77.6	155.6

Table 4 described the model evaluation results. It summarizes the performance of our models. The numbers in bold are the best results in the model. It is easy to see, the MR index reached the maximum value of 0.81 in the SVM algorithm. The maximum value of the Precision index is 0.90; Recall 0.90 and F1-score 0.90 are both in the SVM algorithm. Besides, the indexes of LR algorithm are also quite equal to SVM algorithm, especially when approaching multi-label by LP. The learning time of the SVM algorithm is also the shortest compared to the other two algorithms, the lowest at 5.3 when approaching by LP. The index of DT with the lowest results and the highest learning time shows that DT does not perform as well as the other two algorithms.

4.3 Visualization of results

The Word Cloud chart represents each of the emotions happy, angry, depressed, and hopeful for comparison, which is described in figures 8, 9, 10 and 11. In the images below, we can easily see which words are mentioned the most in customer comments and the larger the word, the more mentioned. In the ranking of WordCloud_Happy, the words “đẹp”, “xuất_sắc”, “tốt” appear the most in customer reviews. In the WordCloud_Angry chart, the words “không”, “nhân_viên” and “tệ” are mentioned the most.

In the WordCloud_Hopeful chart of Figure 11, the most mentioned words are the words “thêm”, “tốt” và “cải_thiện” showing the expectation of change from the hotel with many mentioned aspects such as “nhân_viên”, “nhiệt_tình” và “thân_thiện”. And similarly, as Figure 8, in the WordCloud_Depressed chart, the words “tàm_tạm”, “được” reveal misleading feelings about the most mentioned “nhân_viên” aspect.

WordCloud charts of four types of emotions all mention the word “nhân_viên” a lot, showing that customers care a lot about this factor. However, each type of emotion will have different meanings. For example, in happy emotions, “nhân_viên” is a factor of satisfaction. The word “nhân_viên” in the emotion of hope expresses the expectation of better improvement. In anger, “nhân_viên” is a negative factor that makes the customer experience bad. This meaning is similar to the feeling of depression but to a lesser extent by expressing the word "hơi".

The word “không” appears in common in the three emotions of anger, hope, and depression. However, the terms and accompanying aspects are different. Figure 9 presents WordCloud_Angry, the word “không” represents anger when customer needs are not met as expected with the words “tệ”, “thất_vọng”. Meanwhile, the word "không" in Wordcloud_Depressed comes with the words “tạm”, “hơi” to express disappointment as shown in Figure 10, unlike thinking in terms of “giá”, “vị_trí” và “hôi”. Similarly, for Wordcloud_Hope, it is the desire to improve from customers with keywords accompanied by “thêm”, “để”, “cải_thiện” issues such as “nhân_viên”, “lễ_tân” và “giá”.

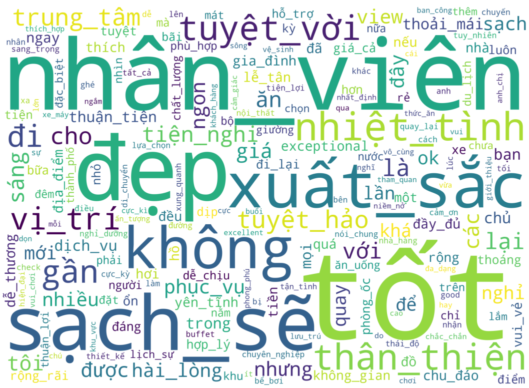

Figure 8. Wordcloud_Happy

Figure 9. Wordcloud_Angry

Figure 10. Wordcloud_Depressed

Figure 11. Wordcloud_Hopeful

5 Conclusion and Future Work

5.1 Conclusion

This study proposed an emotion detection model that concretizes customer emotions based on reviews posted on agoda.com and booking.com websites. According to experimental research results, the multi-label classification model has the highest absolute accuracy, depth, and coverage of up to 81%, 90%, and 90%, respectively. In addition, the results are directly visualized, serving the decision-making needs of enterprises. Our research was conducted on a dataset based on Vietnamese customer comments and reviews on hotel booking websites. This is a field that is considered quite specific because the emotional state of each individual in the comments will be rich and diverse. The research paper also contributes a dataset including reviews and comments in Vietnamese. We believe that each culture will have a different way of assessing emotions based on the language of each country. Vietnamese was categorized as a language with significant linguistic and/or differences from English according to the FSI language difficulty ranking at the US State Department. Furthermore, although the authors of the research paper GoEmotions: A Dataset of Fine-Grained Emotions used a strict process to train and validate the human annotators but annotators were all native English speakers from India so they may still contain potentially problematic content. Therefore, we found that our research results can be a new point that can be a reference for further research.

5.2 Theoretical Implications

Currently, there have been many research papers on the hotel industry, but only about point-view analysis (Sentiment Analysis - SA). These studies mainly focused on emotion-oriented analysis under three categories: positive, negative, and neutral. These three categories are useful in different applications, but will not provide valued details about any of the emotions highlighted in a text. This research paper delves into detecting specific emotions and splits out the positive and negative aspects according to human emotions. The four emotions the study focused on were happy, angry, depressed, and hopeful. The research contributes to the development of previous emotion detection research papers applied to the hospitality sector. Create a set of four main emotions that are common in the field. In particular, the paper used a little-tapped emotion that was hopeful to be incorporated into the emotion set of the paper when conducted in this field of hospitality. At the same time, it creates a premise for the development of a set of emotions in future research papers on the same topic.

5.3 Practical Implications

The research paper by Demszky *et al.* (2020) focuses on the performance according to each emotion and many different sets of emotions, while our research paper focuses on machine learning algorithms without division to evaluate the overall performance for each model. We look forward to running the dataset on multiple models and seeing their performance. This is also one of our goals. The visualization of the results will serve as a basis for the practical implications of the study. The results of the research have also aided the hospitality industry in Vietnam by providing an overview of the four emotions that visitors leave with after receiving hotel services. Businesses may get to know the customers they serve, learn what aspects of their services make them happy, angry, or depressed, and use this information to make improvements according to the aspects in which customers place their hopes. The outcomes serve as a consultative resource for businesses in the post-COVID-19 era as Vietnam's tourism slowly resumes after a brief stop.

5.4 Limitations and Future Research

This study has yet to mine real-time data. Prediction results may be affected because the amount of data collected from the web is not large enough and the preprocessing is not 100% suppressed. In addition, the model in the study has not been tested by all algorithms and multi-label classifiers.

In the future, the research can be extended by building model to mine data in real-time and collect more data. Additionally, the models can develop more spam filtering functions, sarcasm detection, etc. Besides, the research team plans to build deep learning and ensemble learning models with the goal of improving and enhancing the current model. By integrating the emotion detection model with product and service recommendations grounded on customer emotions, a more profound business significance can be achieved.

REFERENCES

Acheampong, F. A., Wenyu, C., and Nunoo-Mensah, H. (2020). Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports*, *2*(7), e12189. https://doi.org/10.1002/ENG2.12189

Ali, B. J., and Anwar, G. (2021). An empirical study of employees' motivation and its influence job satisfaction. Ali, BJ, and Anwar, G.(2021). An Empirical Study of Employees' Motivation and its Influence Job Satisfaction. International Journal of Engineering, Business and Management, 5(2), 21-30.

Alswaidan, N., and Menai, M. E. B. (2020). A survey of state-of-the-art approaches for emotion recognition in text. *Knowledge and Information Systems*, *62*(8), 2937–2987. https://doi.org/10.1007/S10115-020-01449-0/METRICS

Bagozzi, R. P., Gopinath, M., and Nyer, P. U. (1999). The Role of Emotions in Marketing. *Http://Dx.Doi.Org/10.1177/0092070399272005*, *27*(2), 184–206. https://doi.org/10.1177/0092070399272005

Bastiaansen, M., Lub, X. D., Mitas, O., Jung, T. H., Ascenção, M. P., Han, D. I., Moilanen, T., Smit, B., and Strijbosch, W. (2019). Emotions as core building blocks of an experience. *International Journal of Contemporary Hospitality Management*, *31*(2), 651–668. https://doi.org/10.1108/IJCHM-11-2017-0761/FULL/XML

Birjali, M., Kasri, M., and Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, *226*, 107134. https://doi.org/10.1016/J.KNOSYS.2021.107134

Cerri, R., de Carvalho, A. C. P., and Freitas, A. A. (2011). Adapting non-hierarchical multilabel classification methods for hierarchical multilabel classification. *Intelligent Data Analysis*, *15*(6), 861-887.

Charbuty, B., and Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. Journal of Applied Science and Technology Trends, 2(01), 20-28.

Chatterjee, S. (2020). Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach. International Journal of Hospitality Management, 85, 102356.

Cheng, L., and Jiang, Z. (2021). *What drives consumers to keep the top-tier elite membership of premium hotels : Economic Needs or Spiritual Needs?* https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-448556

D'Acunto, D., Tuan, A., Dalli, D., Viglia, G., & Okumus, F. (2020). Do consumers care about CSR in their online reviews? An empirical analysis. *International Journal of Hospitality Management*, *85*, 102342.

Damanik, I. S., Windarto, A. P., Wanto, A., Andani, S. R., and Saputra, W. (2019, August). Decision tree optimization in C4. 5 algorithm using genetic algorithm. In Journal of Physics: Conference Series (Vol. 1255, No. 1, p. 012012). IOP Publishing.

Demir, A., Maroof, L., Khan, N. U. S., and Ali, B. J. (2020). The role of E-service quality in shaping online meeting platforms: a case study from higher education sector. Journal of Applied Research in Higher Education.

Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., and Ravi, S. (2020). GoEmotions: A Dataset of Fine-Grained Emotions. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 4040–4054. https://doi.org/10.18653/v1/2020.acl-main.372

Ghanbari-Adivi, F., and Mosleh, M. (2019). Text emotion detection in social networks using a novel ensemble classifier based on Parzen Tree Estimator (TPE). Neural Computing and Applications, 31(12), 8971-8983.

Ghanbari-Adivi, F., and Mosleh, M. (2019). Text emotion detection in social networks using a novel ensemble classifier based on Parzen Tree Estimator (TPE). *Neural Computing and Applications*, 31(12), 8971-8983.

Gharroudi, O. (2017). *Ensemble multi-label learning in supervised and semi-supervised settings* (Doctoral dissertation, Université de Lyon).

Gibaja, E., & Ventura, S. (2014). Multi-label learning: a review of the state of the art and ongoing research. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *4*(6), 411-444.

Guo, C. Y., and Chou, Y. C. (2020). A novel machine learning strategy for model selections-Stepwise Support Vector Machine (StepSVM). Plos one, 15(8), e0238384.

Haryadi, D., and Kusuma, G. P. (2019). Emotion detection in text using nested long short-term memory. International Journal of Advanced Computer Science and Applications, 10(6).

Hasan, M., Rundensteiner, E., and Agu, E. (2019). Automatic emotion detection in text streams by analyzing Twitter data. *International Journal of Data Science and Analytics*, *7*(1), 35–51. https://doi.org/10.1007/S41060-018-0096-Z/METRICS

Ho, V. A., Nguyen, D. H. C., Nguyen, D. H., Pham, L. T. V., Nguyen, D. V., Nguyen, K. V., and Nguyen, N. L. T. (2019, October). Emotion recognition for vietnamese social media text. In International Conference of the Pacific Association for Computational Linguistics (pp. 319-333). Springer, Singapore).

Hosany, S., Martin, D., and Woodside, A. G. (2021). Emotions in Tourism: Theoretical Designs, Measurements, Analytics, and Interpretations. *Journal of Travel Research*, *60*(7), 1391–1407. https://doi.org/10.1177/0047287520937079

Huete-Alcocer, N. (2017). A literature review of word of mouth and electronic word of mouth: Implications for consumer behavior. *Frontiers in Physiology*, *8*(JUL), 1256. https://doi.org/10.3389/FPSYG.2017.01256/BIBTEX

Jayakrishnan, R., Gopal, G. N., and Santhikrishna, M. S. (2018, January). Multi-class emotion detection and annotation in Malayalam novels. In 2018 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-5). IEEE.

Kariuki, C. (2021). *Multi-label classification with Scikit-MultiLearn*. Section.io. https://www.section.io/engineering-education/multi-label-classification-with-scikit-multilearn/.

Kim, J., & Fesenmaier, D. R. (2015). Measuring emotions in real time: Implications for tourism experience design. *Journal of Travel Research*, 54(4), 419-429.

Kusal, S., Patil, S., Kotecha, K., Aluvalu, R., and Varadarajan, V. (2021). AI Based Emotion Detection for Textual Big Data: Techniques and Contribution. Big Data and Cognitive Computing, 5(3), 43.

Lazoura, J., Athanasios, L., Harris, G., & Dimitrios, M. (2021). Examining Tourist Needs Through Maslow's Hierarchy of Needs and the Role of Hotel Animation. *JOURNAL OF TOURISM RESEARCH* V. 26 June 2021, 229.

Mattila, A. S., and Enz, C. A. (2002). The Role of Emotions in Service Encounters. *Http://Dx.Doi.Org/10.1177/1094670502004004004*, *4*(4), 268–277. https://doi.org/10.1177/1094670502004004004

Nandwani, P., and Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1), 81.

Nandwani, P., and Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, *11*(1), 1–19. https://doi.org/10.1007/S13278-021-00776-6/TABLES/1

Nguyen, B., Nguyen, V.-H., and Ho, T. (2021). Sentiment Analysis of Customer Feedbacks in Online Food Ordering Services. *Business Systems Research: International Journal of the Society for Advancing Innovation and Research in Economy*, *12*(2), 46–59. https://doi.org/10.2478/bsrj-2021-0018

Nguyen, K., Nguyen, V., Nguyen, A., and Nguyen, N. (2020, December). A Vietnamese dataset for evaluating machine reading comprehension. In Proceedings of the 28th International Conference on Computational Linguistics (pp. 2595-2605).

Nguyen, P. H., and Tran, T. T. (2021). Need to build a River tourism information system for tourism students in Ho Chi Minh City. *AIP Conference Proceedings*, *2406*(1). https://doi.org/10.1063/5.0066628/604158

Nguyen, V. H., and Ho, T. (2021). Analyzing customer experience in hotel services using topic modeling. *Journal of Information Processing Systems*, *17*(3), 586-598.

Nguyen, V. H., and Ho, T. (2023). Analysing online customer experience in hotel sector using dynamic topic modelling and net promoter score. *Journal of Hospitality and Tourism Technology*, 14(2), 258-277.

Pantano, E., Priporas, C. V., & Stylos, N. (2017). 'You will like it!'using open data to predict tourists' response to a tourist attraction. *Tourism Management*, *60*, 430-438.

Phung, T. K., Te, N. A., and Ha, T. T. T. (2021, January). A machine learning approach for opinion mining online customer reviews. In *2021 21st ACIS International Winter Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD-Winter)* (pp. 243-246). IEEE.

Pokhun, L., and Chuttur, M. Y. . (2020). Emotions in texts. Bulletin of Social Informatics Theory and Application, 4(2), 59–69. https://doi.org/10.31763/businta.v4i2.256

Priporas, C. V., Stylos, N., Vedanthachari, L. N., & Santiwatana, P. (2017). Service quality, satisfaction, and customer loyalty in Airbnb accommodation in Thailand. *International Journal of Tourism Research*, *19*(6), 693-704.

Ratnasari, R. T., Gunawan, S., Mawardi, I., and Kirana, K. C. (2020). Emotional experience on behavioral intention for halal tourism. *Journal of Islamic Marketing*, *12*(4), 864–881. https://doi.org/10.1108/JIMA-12-2019-0256/FULL/XML

Ray, S. (2019, February). A quick review of machine learning algorithms. In 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon) (pp. 35-39). IEEE.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/H0077714

Servidio, R., and Ruffolo, I. (2016). Exploring the relationship between emotions and memorable tourism experiences through narratives. *Tourism Management Perspectives*, *20*, 151–160. https://doi.org/10.1016/J.TMP.2016.07.010

Stylos, N. (2020). Technological evolution and tourist decision-making: a perspective article. *Tourism Review*, *75*(1), 273-278.

Stylos, N., Zwiegelaar, J., & Buhalis, D. (2021). Big data empowered agility for dynamic, volatile, and time-sensitive service industries: the case of tourism sector. *International Journal of Contemporary Hospitality Management*.

Tuerlan, T., Li, S., and Scott, N. (2021). Customer emotion research in hospitality and tourism: conceptualization, measurements, antecedents and consequences. *International Journal of Contemporary Hospitality Management*, *33*(8), 2741–2772. https://doi.org/10.1108/IJCHM-11-2020-1257/FULL/PDF

Ukil, A. (2007). Support Vector Machine Intelligent Systems and Signal Processing in Power Engineering (pp. 161-226). Berlin, Heidelberg: Springer Berlin Heidelberg.

Vidulin, V. (2013). Searching for Credible Relations in Machine Learning. *Informatica*, *37*(3).

Walters, G., Sparks, B., & Herington, C. (2012). The impact of consumption vision and emotion on the tourism consumer's decision behavior. *Journal of Hospitality & Tourism Research*, 36(3), 366-389.

Xu, G., Meng, Y., Qiu, X., Yu, Z., and Wu, X. (2019). Sentiment analysis of comment texts based on BiLSTM. *IEEE Access*, 7, 51522-51532.

Xu, X. (2020). Examining consumer emotion and behavior in online reviews of hotels when expecting managerial response. International Journal of Hospitality Management, 89, 102559.

Xue, J., He, Y., Liu, M., Tang, Y., and Xu, H. (2021). Incentives for Corporate Environmental Information Disclosure in China: Public Media Pressure, Local Government Supervision and Interactive Effects. Sustainability, 13(18), 10016.

Yüksel, A., and Yüksel, F. (2007). Shopping risk perceptions: Effects on tourists' emotions, satisfaction and expressed loyalty intentions. *Tourism Management*, *28*(3), 703–713. https://doi.org/10.1016/J.TOURMAN.2006.04.025

Zad, S., Heidari, M., James Jr, H., and Uzuner, O. (2021, May). Emotion detection of textual data: An interdisciplinary survey. In 2021 IEEE World AI IoT Congress (AIIoT) (pp. 0255-0261). IEEE.