ASPECT-BASED SENTIMENT ANALYSIS OF TOURISTS' REVISIT INTENTION FROM TOURISTS' ONLINE REVIEWS ON THEME PARKS

NUR SHAMILLA BINTI SELAMAT

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Master of Philosophy

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

MARCH 2021

DEDICATION

To the person that matters the most –

my mother and my father.

ACKNOWLEDGEMENT

Praise be to **Allah The Almighty**. I am truly and forever indebted. All thanks to **Him** for taking me out from the shade of darkness to the shade of light.

Additionally, I wish to express my sincere appreciation to my research supervisor, **Dr Alex Sim Tze Hiang** for sharing his wisdom and providing me with guidance, support and motivation.

I wish to also express my sincere appreciation and gratitude to my research co-supervisor, **Associate Professor Dr Anazida binti Zainal** for sharing her wisdom and providing me with guidance, support and motivation. I am forever indebted to her assistance.

I am also indebted to **Universiti Teknologi Malaysia (UTM)** and **Jabatan Perkhidmatan Awam (JPA)** for funding my years of study.

ABSTRACT

The investigation of tourists' revisit intention (TRI) is crucial as it provides tourism destinations' practitioners (TDP) with insights to revolutionise their customer offerings and enhance their business processes while operating at optimum costs. User-generated content (UGC) is beneficial in offering TDP with an understanding of TRI by utilising aspect-based sentiment analysis. Nevertheless, studies primarily detect aspects based on UGC and utilise nouns for feature building and extraction. Additionally, previous studies that utilise text mining and UGC in the investigation of TRI in the tourism destinations domain are scarce as they are typically the result of surveys. Therefore, to address the gap in the studies and the tourism destinations domain, this study embarks on pre-determined aspects with sentiments detection from UGC. The first objective of this study is to identify TRI through a systematic literature review (SLR), establishing a TRI aspect that are supported by works of literature. Secondly, this study detects the TRI aspects from UGC by using Binary Relevance (BR), with an underlying Multinomial Naïve Bayes (MNB) classifier. This study also uses Latent Semantic Analysis (LSA) and Latent Dirichlet Allocation (LDA) to extract the topical meaning of the TRI aspects from UGC, establishing TRI meta-model from the perspective of text mining. Thirdly, this study compares the TRI meta-models from the perspective of text mining against the perspective of SLR by using the Jaccard similarity index coefficient. This study discovers that the text mining meta-model is reasonably similar to the meta-model from the works of literature, with an 83.98% similarity. This signifies that the traditional approach of using surveys in the investigation of TRI in the tourism destinations domain may be augmented using this study's proposed model.

ABSTRAK

Penyelidikan niat lawatan semula pelancong (TRI) adalah penting bagi melengkapkan pengusaha destinasi pelancongan (TDP) dengan pengetahuan untuk meningkatkan tawaran kepada pelanggan dan menambaik perniagaan di samping beroperasi pada kos yang optimum. Kandungan yang dihasilkan pengguna (UGC) adalah bermanfaat dalam memberi TDP pemahaman tentang TRI melalui penggunaan analisa sentimen berasaskan aspek. Namun, kajian terdahulu kebanyakannya mengesan aspek berdasarkan terma yang wujud dalam UGC dan menggunakan kata nama bagi pembangunan dan pengekstrakan ciri. Tambahan pula, kajian yang menggunakan perlombongan teks dan UGC dalam konteks penyelidikan TRI bagi destinasi pelancongan adalah kurang kerana kajian kebiasannya adalah menerusi kaji selidik. Oleh itu, bagi menangani jurang dalam kajian dan konteks destinasi pelancongan, kajian ini melaksanakan tugas pengesanan aspek yang dikenalpasti sebelumnya bersama sentimen dari UGC. Objektif pertama kajian ini adalah untuk mengenalpasti TRI melalui kajian literatur yang sistematik (SLR), bagi mewujudkan aspek TRI yang disokong oleh literatur. Kedua, kajian ini mengesan aspek TRI dari UGC menggunakan Binary Relevance (BR), dengan pengelas Multinomial Naïve Bayes (MNB). Kajian ini turut menggunakan Latent Semantic Analysis (LSA) dan Latent Dirichlet Allocation (LDA) untuk mengekstrak topik aspek TRI dari UGC untuk mewujudkan meta-model TRI dari perspektif perlombongan teks. Ketiga, kajian ini membandingkan meta-model TRI dari perspektif perlombongan teks dengan perspektif SLR menggunakan koefisien indeks kesamaan Jaccard. Kajian ini mendapati bahawa meta-model perlombongan teks agak sama dengan meta-model dari karya-karya sastera sebanyak 83.98%. Ini menandakan bahawa pendekatan tradisional menggunakan kaji selidik dalam penyiasatan TRI ke destinasi boleh diganti dengan model yang dicadangkan kajian ini.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABSA	-	Aspect-Based Sentiment Analysis
API	-	Application Programming Interface
BR	-	Binary Relevance
DT	-	Determiner
HTTP	-	HyperText Transfer Protocol
IN	-	Preposition
JJ	-	Adjective
JJR	-	Adjective, Comparative
JJS	-	Adjective, Superlative
JPA	-	Jabatan Perkhidmatan Awam
JSON	-	JavaScript Object Notation
LDA	-	Latent Dirichlet Allocation
LP	-	Label Powerset
LSA	-	Latent Semantic Analysis
LSI	-	Latent Semantic Indexing
MNB	-	Multinomial Naïve Bayes
NN	-	Noun, Singular
NNS	-	Noun, Plural
NNP	-	Proper Noun, Plural
OVA	-	One-Versus-All
OVO	-	One-Versus-One
PACIS 2017	-	Twenty-First Pacific Asia Conference on Information Systems 2017
PICOC	-	Population, Intervention, Comparison, Outcomes, and Context
POS	-	Part-Of-Speech
RAkEL	-	Random k-Labelsets
RB	-	Adverb
RBR	-	Adverb, Comparative

RBS	-	Adverb, Superlative
SVD	-	Singular Value Decomposition
TDP	-	Tourism Destinations' Practitioners
TF-IDF	-	Term Frequency-Inverse Document Frequency
TRI	-	Tourists' Revisit Intention
UGC	-	User-Generated Content
UNESCO	-	United Nations Educational, Scientific and Cultural Organisation
UTM	-	Universiti Teknologi Malaysia
VB	-	Verb, Base Form
VBD	-	Verb, Past Tense
VBG	-	Verb, Present Participle
VBN	-	Verb, Past Participle
VBP	-	Verb, Singular Present
VBZ	-	Verb, Third Person Singular Present

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1   Introduction

This study investigates tourists' revisit intention in the tourism destinations domain using user-generated content, or online reviews as a data source and the text mining approach aspect-based sentiment analysis. This contributes to the limited pool of studies that investigate tourists' revisit intention in the tourism destinations domain that mainly used surveys or content analysis. The utilisation of user-generated content as a data source is advantageous as it does not suffer from the expensive cost, effort and time in carrying out surveys. Apart from that, this study detects pre-determined tourists' revisit intention aspects and employs varying feature extraction approaches and Part-Of-Speech tags when detecting the aspects from user-generated content. This contributes to the pool of studies that in majority extract aspects based on the terms that exist in user-generated content and use nouns for feature extraction. This study aspires to detect aspects that are investigated in works of literature from user-generated content. Not only that, this study perceives that different tourists' revisit intention aspects convey different nature of subjects, which calls for the utilisation of varying feature extraction approaches and Part-Of-Speech tags.

This introductory chapter presents a synopsis of the thesis, beginning with the background and the problem statement of the study. The research questions, as well as the research objectives, are subsequently described, followed by a description of the

scope of the study and an introduction to the overall research design of the study. This chapter concludes with a synopsis of how the rest of the thesis unfolds.

1.2   Background of the Study

The tourism and hospitality industry propels the growth of a country's economy, whereby it contributes to job creation and poverty diminution, promotes environmental sustainability, peace and understanding amongst communities, across nations. In 2015, the travel and tourism sector gave rise to 107,833,000 jobs in the service industries such as hotels, travel agents, transportation, as well as restaurants and leisure industries (World Travel and Tourism Council, 2016). Malaysia's economy also benefitted from the tourism sector as it drives the nation's service trades, current account and foreign currency assets (Ooi, 2019). In 2017, the tourism sector contributes 14.9% to the nation's economy, with a 6.1% and a value of RM82.6 billion direct contribution to the nation's gross domestic product (Bernama, 2018). While Malaysia collected RM84.1 billion, the nation's highest overall tourism receipts in 2018, the Malaysian Tourism Board reported a gradual decline of 3.4% from 2017 in tourist arrivals (Ooi, 2019), which calls for the need for the government and tourism destinations' practitioners (TDP) to step up in their efforts.

In recovering and boosting tourist arrivals to Malaysia, Tussyadiah, Park and Fesenmaier (2011) suggest composing marketing narratives embedded with characters that are relatable to the target audience of a tourism destination. This will increase the target audience's knowledge of the destination, as well as their intention to visit or revisit the destination. Tussyadiah, Park and Fesenmaier (2011) add that employing customisation in advertising, such as incorporating different genres of narratives, or different characters in a narrative able to stimulate empathy among tourists of diverse characteristics. To compose relatable and customisable marketing strategies for

tourism destinations, TDP must recognise tourists' intention, specifically tourists' revisit intention (TRI) to a destination, especially in the face of continuous competitive challenges and increasing demand from consumers. Understanding TRI also empowers TDP with the appropriate information for them to be able to revolutionise their customer offerings and enhance their business processes while operating at optimum costs to encourage tourists to visit and revisit their tourism destinations.

This study focuses on the investigation of TRI to a destination, as also investigated in other studies (Chen *et al.*, 2011; Craggs and Schofield, 2011; Kozak and Duman, 2012; Lee *et al.*, 2012; Moutinho, Albayrak and Caber, 2012; Osti, Disegna and Brida, 2012; Eusébio and Vieira, 2013; Prayag and Jankee, 2013; Giraldi and Cesareo, 2014). TRI is the focus of the study as a consequence of the dwindling tourist arrivals to Malaysia in 2018 (Ooi, 2019). This study perceives that understanding TRI is crucial to encourage tourists with visitation experience to a destination for a repeat visit, apart from to be an instrument of positive word-of-mouth for the destination, hence, attracting new tourists to the destination, with minimal marketing costs from TDP.

In the investigation of TRI to a destination, it is discovered that studies on TRI are primarily the result of surveys. Studies that apply text mining approaches are limited. This study perceives text mining as a catalyst in the investigation of TRI. Especially as the ubiquity of the Internet and social media has given rise to the democratisation of user-generated content (UGC). In the context of this study, UGC refers to online reviews on travel social media platforms such as TripAdvisor, Expedia and Yelp. UGC are a wealth of information that offer insights on TRI to a destination. They document the voice of tourists, presenting TDP with TRI that are articulated in their own language and organised based on their own judgment of importance and priority. UGC are also extensively used by studies in decision analysis (Tirunillai and Tellis, 2014; Xiang *et al.*, 2015; Cheng, Shih and Lin, 2016; Culotta and Cutler, 2016; Farhadloo, Patterson and Rolland, 2016; Liu, Bi and Fan, 2017; Gao *et al.*, 2018; Siering, Deokar and Janze, 2018). Therefore, this study perceives UGC as a gold mine

of a data source that is able to offer TDP with an understanding of TRI through the utilisation of text mining.

Nevertheless, the utilisation of text mining on UGC in the tourism destinations domain, especially in the investigation of TRI is under-researched. Studies that utilised UGC as a data source in consumer behaviour research in the tourism and hospitality industry are primarily on the domains hotels and restaurants (Hu, Chen and Chou, 2017; Moro, Rita and Coelho, 2017; Hlee *et al.*, 2018; Li *et al.*, 2018; Mao, Yang and Wang, 2018). Studies that utilised UGC in the tourism destinations domain typically used content analysis (Larsen, 2014; Wong and Qi, 2017; Prakash *et al.*, 2018; Vo Thanh and Kirova, 2018). The drawback of using content analysis is it requires extensive effort to manually code, categorize and sort the UGC into themes. This calls for the necessity for this study to investigate TRI, utilising UGC as a data source and text mining approach that is other than content analysis in addressing this gap. Text mining, specifically aspect-based sentiment analysis automates the detection of themes from UGC without manual intervention of coding, categorisation and sorting, unlike content analysis.

To investigate TRI using UGC and text mining, this study initially looks towards AYLIEN aspect-based sentiment analysis (AYLIEN ABSA). AYLIEN ABSA is a sentiment analysis tool and extension in RapidMiner – a data science software platform for data pre-processing, machine learning and text mining, among others. AYLIEN ABSA is able to detect aspects for four industry domains namely airlines, cars, hotels and restaurants. However, the employment of AYLIEN ABSA on UGC for a tourism destination, specifically Berjaya Times Square Theme Park using all four domains is unsuccessful, whereby AYLIEN ABSA is unable to detect any aspects nor sentiments from UGC in the tourism destinations domain. This signifies that UGC in the tourism destinations domain discuss aspects that are different from the domains airlines and cars, as well as the domains hotels and restaurants that are in a similar industry as the tourism destinations domain, which is the tourism and hospitality industry. This is concurred by Pham and Le (2016).

Taking inspiration from the work of Pham and Le (2016) that extracts pre-determined aspects from UGC using aspect-based sentiment analysis in the hotels domain, and the work of Farhadloo, Patterson and Rolland (2016) that utilises aspect-based sentiment analysis in the investigation of consumers' satisfaction, this study aspires to detect pre-determined TRI aspects from UGC in the tourism destinations domain using aspect-based sentiment analysis. In applying aspect-based sentiment analysis, this study perceives that the seed-based aspect extraction approach is advantageous for this study. Seed-based approach extracts aspects based on the grammatical relation between the seed and opinion words. Nevertheless, the works of literature (Wang, Lu and Zhai, 2010; Zhu *et al.*, 2011; Mukherjee and Liu, 2012; Colhon, Bădică and Şendre, 2014) extracted aspects based on the terms that exist in UGC, not pre-determined aspects such as in the work of Pham and Le (2016). This highlighted the gap in the seed-based aspect extraction studies that are lacking in the detection of pre-determined aspects from UGC. Additionally, both the application of seed-based aspect extraction and lexicon-based aspect-based classification approaches on UGC in the tourism destinations domain is limited, which this study would like to contribute to.

Apart from that, in the aspect extraction task of aspect-based sentiment analysis, studies mainly Part-Of-Speech (POS) tagged terms and used nouns (Zhu *et al.*, 2011; Marrese-Taylor, Velasquez and Bravo-Marquez, 2013; Farhadloo, Patterson and Rolland, 2016) for feature extraction. This study would like to address this gap to enable the utilisation of different POS tags for feature extraction. This study perceives that different TRI aspects convey different nature of subjects, which calls for the utilisation of varying feature extraction approaches and POS tags to conform to the different subjects of the aspects.

In summary, this study aspires to investigate TRI using aspect-based sentiment analysis of UGC. This is to address the gap in the TRI studies that primarily used surveys and in the application of text mining in consumer behaviour research in the tourism and hospitality industry that is focusing on the domains hotels and restaurants. This study also seeks to detect pre-determined aspects that are genuinely relevant and

significant in the tourism destinations domain from UGC to contribute to the pool of studies that in majority extracted aspects based on the terms that exist in UGC. Additionally, this study aims to employ varying feature extraction approaches or POS tags when detecting aspects from UGC, to address the gap in the studies that mainly used nouns for feature extraction. The investigation of TRI using aspect-based sentiment analysis of UGC is beneficial for TDP as it supplies them with insightful information for the betterment of their tourism destinations that will encourage tourists for repeat visitation and glowing recommendation of their tourism destinations. The utilisation of UGC as a data source is also advantageous as it does not suffer from the expensive cost, effort and time in carrying out surveys.

1.3   Problem Statement

Through a systematic literature review (SLR), 91% of studies on TRI in the tourism destinations domain is discovered to be the result of surveys (see Chapter 2, Section 2.3 for the methodology of TRI investigation in the tourism destinations domain). Surveys are more often than not limited by the number of respondents, and expensive in terms of time, money and effort. Surveys also limit the number of aspects to be investigated to avoid the survey questionnaires from being lengthy, as to not exhaust the survey respondents.

Apart from surveys, the 2% of studies on TRI in the tourism destinations domain that utilised UGC and text mining used content analysis. Content analysis is also expensive in terms of time and effort to manually code, categorize and sort UGC into themes or aspects. The utilisation of aspect-based sentiment analysis of UGC, on the other hand, do not suffer the restrictions and limitations of carrying out surveys and content analysis.

As highlighted by the problem of AYLIEN ABSA that is unable to detect any aspects nor sentiments from UGC in the tourism destinations domain, albeit it is able to detect aspects for the hotels and restaurants domain, this study perceives that different aspects are of relevance and significance in the tourism destinations domain. Therefore, this study would like to contribute to the pool of studies in the tourism destinations domain, as studies are primarily focusing on the hotels and restaurants domains including studies in the seed-based aspect extraction and lexicon-based aspect-based classification approaches in aspect-based sentiment analysis.

Additionally, studies that used seed-based aspect extraction approaches are only detecting aspects based on the terms that exist in UGC(Wang, Lu and Zhai, 2010; Zhu *et al.*, 2011; Mukherjee and Liu, 2012; Colhon, Bădică and Şendre, 2014). They do not detect pre-determined aspects nor aspects that are already being tested in the works of literature, which calls the necessity for this study to address this gap. Furthermore, studies mainly POS tagged terms and used nouns when extracting features as part of the aspect extraction task. Hence, this study would like to contribute by enabling the utilisation of varying feature extraction approaches or POS tags when extracting features.

1.4   Purpose of the Study

This study seeks to investigate TRI in the tourism destinations domain by using UGC as a data source and the approach aspect-based sentiment analysis. Using the approach aspect-based sentiment analysis, this study aims to detect pre-determined TRI aspects that are supported by works of literature and their sentiments from UGC in the tourism destinations domain. Thus, beforehand, this study aspires to investigate TRI aspects that are genuinely relevant in the tourism destinations domain through an SLR.

7

Additionally, in the aspect extraction phase of the aspect-based sentiment analysis, this study aims to utilise multiple feature extraction approaches and POS tags for feature extraction. Ultimately, from the results of the proposed model of this study, this study also aims to compare with the results from the works of literature that are primarily based on surveys.

1.5   Research Questions

This study seeks to answer the following notable questions:

(a) What are the TRI aspects that are identifiable through a systematic literature review?

(b) How to detect and classify the TRI aspects into positive and negative sentiments from UGC?

(c) How does the result of investigating TRI to a destination using the proposed model of the study compare to the result from the works of literature?

1.6   Research Objectives

The objectives of this study include:

(a) To identify the TRI aspects through a systematic literature review.

(b) To detect and classify the TRI aspects into positive and negative sentiments from UGC by using a multi-label classification problem transformation approach, Binary Relevance (BR) with Multinomial Naïve Bayes (MNB) classifiers and

multiple feature extraction approaches for aspect extraction and aspect-based classification, and to identify hidden meaning or dimensions in TRI aspects that are detected by using topic modelling approaches, Latent Semantic Analysis (LSA) and Latent Dirichlet Allocation (LDA).

(c) To compare the result of investigating TRI to a destination using the proposed model of the study with the result from the works of literature by using the Jaccard similarity index coefficient.

1.7   Overall Research Design

This study seeks to investigate TRI in the tourism destinations domain using UGC as a data source and the text mining approach aspect-based sentiment analysis. It comprised of three main phases: (1) Aspect identification phase, (2) Aspect extraction and classification phase, and (3) Meta-model comparison phase. The purpose of the aspect identification phase is to identify TRI aspects that will be utilised in the detection of aspects and their sentiments in the aspect extraction and classification phase. The aspect extraction and classification phase is the primary phase of this study. It detects TRI aspects from UGC by implementing aspect-based sentiment analysis. The meta-model comparison phase, subsequently, compares the result based on the proposed model of the study with the result from the works of literature. Chapter 3 offers a more illustrious description in explaining each of the phases in this study and the overall research design choices.

1.8 Scope of the Study

The constraints and limitations of this study include:

(a) This study concentrates on the investigation of TRI to a destination. According to the theoretical lens of planned behaviour, intentions are significant predictors of actual behaviours (Ajzen, 1985). By investigating TRI to a destination, this study perceives that TDP are able to infer the aspects that influence tourists to actually revisit a destination. Thus, they are able to operate their tourism destinations in a manner that conforms to the aspects that influence tourists to revisit a destination, boosting tourist arrivals to the tourism destinations.

(b) This study also focuses on the tourism destinations theme parks with roller coaster-type attractions. Roller coaster-type attractions offer the utmost consumers' satisfaction in theme parks (Kim and Kim, 2016). Nevertheless, from an operational perspective, roller coaster-type attractions are costly as they require hefty installation areas, as well as expensive installation, maintenance and repair costs (Berkowitz and Stanton, 2013). Hence, this study believes that it is crucial to understand TRI to theme parks with roller coaster-type attractions, to enhance the manner in which TDP operate their theme parks that will be cost-effective and proliferate tourists' satisfaction, as well as their revisit intention to the theme parks.

1.9 Thesis Organisation

This section outlines the organisation of the thesis. An overview of each of the chapters in the thesis, whereby there are seven chapters in total, are as follows:

Chapter 1 – Introduction. This introductory chapter presents a synopsis of the thesis. It commences with an overview of the study and the motivations for the study. Due to the gap in the tourism destinations domain that is lacking in studies that apply aspect-based sentiment analysis, this study embarks on the task of detecting and classifying TRI aspects into positive and negative sentiments from UGC, to contribute to the studies in the tourism destinations domain that primarily used surveys or content analysis. This is deemed to be of assistance to TDP as it supplies them with insights on TRI for the betterment of their tourism destinations that will encourage tourists for repeat visitation and glowing recommendation of their tourism destinations. This introductory chapter also describes the research objectives and the research questions, as well as the scope of the study. It concludes with an overview of the overall research design of the study.

Chapter 2 – Literature Review. This chapter presents an overview of the importance of investigating TRI to a destination and using UGC as a data source in consumer behaviour research in the tourism and hospitality industry. This chapter reveals the gap in consumer behaviour research in the tourism destinations domain that primarily uses surveys. The text mining approach that is utilised is mainly content analysis. This highlights the necessity for the investigation of TRI from UGC using aspect-based sentiment analysis, to address the gap in the tourism destinations domain. This chapter also highlights the gap in the aspect extraction approach that primarily detect aspects based on the terms that exist in UGC, not pre-determined aspects that are established prior to aspect extraction, as well as the usage of nouns for feature extraction. This chapter also presents an overview of multi-label classification, MNB classification and topic modelling approaches, specifically LDA and LSA in order to implement the approaches in detecting TRI aspects from UGC in the aspect extraction and classification phase of this study.

Chapter 3 – Research Methodology. This chapter explains the research design of this study that comprises of three phases, including the aspect identification phase, followed by the aspect extraction and classification phase that detects TRI aspects from UGC. This chapter also describes the meta-model comparison phase that

11

compares the result based on the proposed model of the study from the aspect extraction and classification phase with the result from the works of literature from the aspect identification phase. This chapter also discusses the tools and techniques to be used, such as the use of BR with the underlying classifier MNB, and unigram-POS and skip bi-grams, as well as LSA and LDA in the aspect extraction and classification phase, and the Jaccard similarity index coefficient in the meta-model comparison phase.

Chapter 4 – Tourists' Revisit Intention Aspects Identification using a Systematic Literature Review. This chapter presents the SLR of TRI aspects from the works of literature from 2011 to 2020. To detect TRI aspects and their sentiments in tourists' online reviews, serving as the foundation for the rest of the phases of this study, this chapter identifies 12 aspects, namely (1) Image, (2) Accommodation, (3) Entertainment, (4) Food and cleanliness, (5) Hospitality, (6) Knowledge, (7) Perceived risk, (8) Novelty value, (9) Value for money, (10) Emotion, (11) Satisfaction, and (12) Revisit intention.

Chapter 5 – Tourists' Revisit Intention Aspects Detection and Classification using Aspect-Based Sentiment Analysis. This chapter presents the data analysis and the results of the aspect extraction and classification phase of this study. It documents the result of aspect-based sentiment analysis using BR with MNB classifiers and multiple feature extraction approaches, to detect pre-determined TRI aspects namely image, accommodation, entertainment, food and cleanliness, hospitality, knowledge, perceived risk, novelty value, value for money, emotion, satisfaction and revisit intention from UGC in the tourism destinations domain. It also presents the results of extracting dimensions of TRI aspects from UGC by using LSA and LDA.

Chapter 6 – Tourists' Revisit Intention from the Perspectives of Text Mining against Systematic Literature Review. This chapter presents the data analysis and results of the meta-model comparison phase. It documents the comparison of the result from the aspect extraction and classification phase with the result from the aspect

identification phase by using the Jaccard similarity index coefficient. This chapter also discusses the dimensions of aspects that are exclusive to a model and the dimensions that are shared by the two models, as well as the implication of such outcomes.

Chapter 7 – Conclusion. This chapter summarises how each of the research questions is addressed in the thesis. It also presents the contributions and implications of the study, as well as suggestions for future research.

REFERENCES

Afzaal, M. *et al.* (2016) 'Fuzzy Aspect Based Opinion Classification System for Mining Tourist Reviews', *Advances in Fuzzy Systems*, 2016. doi: 10.1155/2016/6965725.

Aguiló, E. and Rossello, J. (2012) 'Research note: The new all-inclusive board formula in mature destinations - From motivation to satisfaction', *Tourism Economics*, 18(5), pp. 1117–1123. doi: 10.5367/te.2012.0151.

Ahmadi, Z. and Kramer, S. (2018) 'A label compression method for online multi-label classification', *Pattern Recognition Letters*. North-Holland, 111, pp. 64–71. doi: 10.1016/J.PATREC.2018.04.015.

Ajzen, I. (1985) 'From Intentions to Actions: A Theory of Planned Behavior', in Kuhl, J. and Beckmann, J. (eds) *Action Control: From Cognition to Behavior*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 11–39. doi: 10.1007/978-3-642-69746-3_2.

Akgunduz, Y. and Coşar, Y. (2018) 'Motivations of Event Tourism Participants and Behavioural Intentions', *Tourism and hospitality management*, 24(2), pp. 341–357. doi: 10.20867/thm.24.2.4.

Akhtar, N. *et al.* (2017) 'Aspect based Sentiment Oriented Summarization of Hotel Reviews', *Procedia Computer Science*. Elsevier B.V., 115, pp. 563–571. doi: 10.1016/j.procs.2017.09.115.

Al-Mosaiwi, M. and Johnstone, T. (2018) 'Linguistic markers of moderate and absolute natural language', *Personality and Individual Differences*. Elsevier, 134(May), pp. 119–124. doi: 10.1016/j.paid.2018.06.004.

De Albornoz, J. C. *et al.* (2011) 'A joint model of feature mining and sentiment analysis for product review rating', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6611 LNCS, pp. 55–66. doi: 10.1007/978-3-642-20161-5_8.

Albrecht, R. and Merkl, D. (1998) 'Knowledge Discovery in Literature Data Bases',

in Grothkopf, U. et al. (eds) *Library and Information Services in Astronomy III (LISA III) ASP Conference Series*. Spain: Astronomical Society of the Pacific, p. 320. Available at: http://adsabs.harvard.edu/abs/1998ASPC..153...93A (Accessed: 4 May 2017).

Almeida, A. M. G. *et al.* (2018) 'Applying multi-label techniques in emotion identification of short texts', *Neurocomputing*. Elsevier B.V., 320, pp. 35–46. doi: 10.1016/j.neucom.2018.08.053.

Altheneyan, A. S. and Menai, M. E. B. (2014) 'Naïve Bayes classifiers for authorship attribution of Arabic texts', *Journal of King Saud University - Computer and Information Sciences*. Elsevier, 26(4), pp. 473–484. doi: 10.1016/J.JKSUCI.2014.06.006.

Aly, M. (2005) *Survey on Multiclass Classification Methods*.

Anand, D. and Naorem, D. (2016) 'Semi-supervised Aspect Based Sentiment Analysis for Movies Using Review Filtering', *Procedia Computer Science*. Elsevier Masson SAS, 84, pp. 86–93. doi: 10.1016/j.procs.2016.04.070.

Arnaiz-González, Á. *et al.* (2018) 'Study of data transformation techniques for adapting single-label prototype selection algorithms to multi-label learning', *Expert Systems with Applications*. Pergamon, 109, pp. 114–130. doi: 10.1016/J.ESWA.2018.05.017.

Assaker, G. and Hallak, R. (2013) 'Moderating Effects of Tourists' Novelty-Seeking Tendencies on Destination Image, Visitor Satisfaction, and Short- and Long-Term Revisit Intentions', *Journal of Travel Research*, 52(5), pp. 600–613. doi: 10.1177/0047287513478497.

AYLIEN (2019) *API Endpoints | Text Analysis API Documentation*. Available at: https://docs.aylien.com/textapi/endpoints/#api-endpoints (Accessed: 9 May 2019).

Bajs, I. P. (2015) 'Tourist Perceived Value, Relationship to Satisfaction, and Behavioral Intentions: The Example of the Croatian Tourist Destination Dubrovnik', *Journal of Travel Research*. SAGE Publications Ltd, 54(1), pp. 122–134. doi: 10.1177/0047287513513158.

Barros, C. P. and Assaf, A. G. (2012) 'Analyzing Tourism Return Intention to an Urban Destination', *Journal of Hospitality & Tourism Research*, 36(2), pp. 216–231. doi: 10.1177/1096348010388658.

Beckman, E., Kumar, A. and Kim, Y.-K. (2013) 'The Impact of Brand Experience on

Downtown Success', *Journal of Travel Research*, 52(5), pp. 646–658. doi: 10.1177/0047287513478502.

Berjaya Times Square (2019a) *Awards : Berjaya Times Square Theme Park Kuala Lumpur*. Available at: https://www.berjayatimessquarethemeparkkl.com/awards (Accessed: 10 February 2019).

Berjaya Times Square (2019b) *Key Attractions : Berjaya Times Square Theme Park Kuala Lumpur*. Available at: https://www.berjayatimessquarethemeparkkl.com/key-attractions (Accessed: 10 February 2019).

Berkowitz, B. and Stanton, L. (2013) 'Roller coasters: Feeling loopy', *The Washington Post*. Available at: https://www.washingtonpost.com/wp-stat/graphics/health/why-roller-coasters-make-us-scream/index.html (Accessed: 4 June 2017).

Bernama (2018) 'Tourism contributes 14.9 pct to Malaysia's economy', *New Straits Times*. Available at: https://www.nst.com.my/news/nation/2018/09/411526/tourism-contributes-149-pct-malaysias-economy (Accessed: 17 July 2019).

Bielza, C., Li, G. and Larrañaga, P. (2011) 'Multi-dimensional classification with Bayesian networks', *International Journal of Approximate Reasoning*, 52(6), pp. 705–727. doi: 10.1016/j.ijar.2011.01.007.

Bilgihan, A. *et al.* (2016) 'Consumer perception of knowledge-sharing in travel-related Online Social Networks', *Tourism Management*. Elsevier Ltd, 52, pp. 287–296. doi: 10.1016/j.tourman.2015.07.002.

Bilici, E. and Saygın, Y. (2017) 'Why do people (not) like me?: Mining opinion influencing factors from reviews', *Expert Systems with Applications*. Elsevier Ltd, 68, pp. 185–195. doi: 10.1016/j.eswa.2016.10.001.

Bitext Innovations (2019) *NLP for CX Analytics Tools | Bitext. We help AI understand humans*. Available at: https://www.bitext.com/text-analytics-tools/ (Accessed: 9 May 2019).

Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003) 'Latent Dirichlet Allocation', *Journal of Machine Learning Research*, 3, pp. 993–1022. doi: 10.1162/jmlr.2003.3.4-5.993.

Boateng, H., Okoe, A. F. and Hinson, R. E. (2018) 'Dark tourism: Exploring tourist's experience at the Cape Coast Castle, Ghana', *Tourism Management Perspectives*. Elsevier, 27, pp. 104–110. doi: 10.1016/j.tmp.2018.05.004.

Bogaert, M. *et al.* (2019) 'Evaluating multi-label classifiers and recommender systems in the financial service sector', *European Journal of Operational Research*. North-

Holland, 279(2), pp. 620–634. doi: 10.1016/J.EJOR.2019.05.037.

Boussalis, C. and Coan, T. G. (2016) 'Text-mining the signals of climate change doubt', *Global Environmental Change*. Elsevier Ltd, 36, pp. 89–100. doi: 10.1016/j.gloenvcha.2015.12.001.

Boutell, M. R. *et al.* (2004) 'Learning multi-label scene classification', in *Pattern Recognition*, pp. 1757–1771. doi: 10.1016/j.patcog.2004.03.009.

Brida, J. G., Meleddu, M. and Pulina, M. (2012) 'Understanding Urban Tourism Attractiveness: The Case of the Archaeological Ötzi Museum in Bolzano', *Journal of Travel Research*. SAGE Publications Ltd, 51(6), pp. 730–741. doi: 10.1177/0047287512437858.

Briggs, F. *et al.* (2012) 'Acoustic classification of multiple simultaneous bird species: A multi-instance multi-label approach', *The Journal of the Acoustical Society of America*, 131(6), pp. 4640–4650. doi: 10.1121/1.4707424.

Britt, M. A. *et al.* (2004) 'Using intelligent feedback to improve sourcing and integration in students' essays', *International Journal of Artificial Intelligence in Education*, 14(3), pp. 359–374.

Callen, D. *et al.* (2017) *Roller Coaster Database*. Available at: http://www.rcdb.com/ (Accessed: 2 July 2017).

Cambria, E. *et al.* (2013) 'New Avenues in Opinion Mining and Sentiment Analysis', *IEEE Intelligent Systems*, 28(2), pp. 15–21. doi: 10.1109/MIS.2013.30.

Cao, Q., Duan, W. and Gan, Q. (2011) 'Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach', *Decision Support Systems*, 50, pp. 511–521. doi: 10.1016/j.dss.2010.11.009.

de Carvalho, A. C. P. L. F. and Freitas, A. A. (2009) 'A Tutorial on Multi-Label Classification Techniques', in *Foundations of Computational Intelligence*. Springer, Berlin, Heidelberg, pp. 177–195. doi: 10.1007/978-3-642-01536-6.

Castro, D. W. *et al.* (2017) 'Smoothed N-gram based Models for Tweet Language Identification: A Case Study of the Brazilian and European Portuguese National Varieties', *Applied Soft Computing*. Elsevier, 61, pp. 1160–1172. doi: 10.1016/J.ASOC.2017.05.065.

Charte, F. *et al.* (2015) 'QUINTA: A question tagging assistant to improve the answering ratio in electronic forums', in *IEEE EUROCON 2015 - International Conference on Computer as a Tool (EUROCON)*. IEEE, pp. 1–6. doi:

10.1109/EUROCON.2015.7313677.

Chen, C.-C. and Lin, Y.-H. (2012) 'Segmenting Mainland Chinese Tourists to Taiwan by Destination Familiarity: a Factor-cluster Approach', *International Journal of Tourism Research*, 14, pp. 339–352. doi: 10.1002/jtr.

Chen, C. M. *et al.* (2011) 'Tourist behavioural intentions in relation to service quality and customer satisfaction in Kinmen National Park, Taiwan', *International Journal of Tourism Research*, 13(5), pp. 416–432. doi: 10.1002/jtr.810.

Chen, W. and Hirschheim, R. (2004) 'A paradigmatic and methodological examination of information systems research from 1991 to 2001', *European Journal of Information Systems*, 14, pp. 197–235.

Chen, Xiao-jun *et al.* (2016) 'Complex video event detection via pairwise fusion of trajectory and multi-label hypergraphs', *Multimedia Tools and Applications*. Springer US, 75(22), pp. 15079–15100. doi: 10.1007/s11042-015-2514-8.

Chen, Y. *et al.* (2015) 'Why Can Package Tours Hurt Tourists? Evidence from China's Tourism Demand in Hong Kong', *Journal of Travel Research*. SAGE Publications Ltd, 55(4), pp. 427–439. doi: 10.1177/0047287515612597.

Cheng, J.-S., Shih, H.-Y. and Lin, H.-Y. (2016) 'Key success factors in resort hotels practicing moving experience', *Revista de Cercetare si Interventie Sociala*. Editura Lumen, 53, pp. 80–88.

Cheng, X. *et al.* (2019) 'An investigation on online reviews in sharing economy driven hospitality platforms: A viewpoint of trust', *Tourism Management*. Elsevier, 71, pp. 366–377. doi: 10.1016/j.tourman.2018.10.020.

Chi, C. G. (2012) 'An Examination of Destination Loyalty: Differences Between First-Time and Repeat Visitors', *Journal of Hospitality & Tourism Research*, 36(1), pp. 3–24. doi: 10.1177/1096348010382235.

Del Chiappa, G., Alarcón-Del-Amo, M.-D.-C. and Lorenzo-Romero, C. (2016) 'Internet and User-Generated Content Versus High Street Travel Agencies: A Latent Gold Segmentation in the Context of Italy', *Journal of Hospitality Marketing & Management*, 25(2), pp. 197–217. doi: 10.1080/19368623.2014.1001933.

Del Chiappa, G., Lorenzo-Romero, C. and Alarcón-del-Amo, M.-C. (2018) 'Profiling tourists based on their perceptions of the trustworthiness of different types of peer-to-peer applications', *Current Issues in Tourism*, 21(3), pp. 259–276. doi: 10.1080/13683500.2015.1077208.

Cini, F., Leone, L. and Passafaro, P. (2012) 'Promoting Ecotourism Among Young People: A Segmentation Strategy', *Environment and Behavior*, 44(1), pp. 87–106. doi: 10.1177/0013916510389313.

Cole, S. T. and Chancellor, H. C. (2009) 'Examining the festival attributes that impact visitor experience, satisfaction and re-visit intention', *Journal of Vacation Marketing*. SAGE PublicationsSage UK: London, England, 15(4), pp. 323–333. doi: 10.1177/1356766709335831.

Colhon, M., Bădică, C. and Şendre, A. (2014) 'Relating the opinion holder and the review accuracy in sentiment analysis of tourist reviews', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8793(October), pp. 246–257. doi: 10.1007/978-3-319-12096-6_22.

Collie, M. (2014) *Internet Leads Travel, Geton Marketing Inc.* Available at: https://www.geton.com/blog/internet-leads-travel/ (Accessed: 17 March 2019).

Cong, H. and Tong, L. H. (2008) 'Grouping of TRIZ Inventive Principles to facilitate automatic patent classification', *Expert Systems with Applications*. Pergamon, 34(1), pp. 788–795. doi: 10.1016/J.ESWA.2006.10.015.

Corbau, C. *et al.* (2019) 'Tourism analysis at Asinara Island (Italy): Carrying capacity and web evaluations in two pocket beaches', *Ocean and Coastal Management*. Elsevier, 169, pp. 27–36. doi: 10.1016/j.ocecoaman.2018.12.004.

Craggs, R. and Schofield, P. (2011) 'The Quays in Salford: An analysis of visitor perceptions, satisfaction and behavioural intention', *International Journal of Tourism Research*, 13(6), pp. 583–599. doi: 10.1002/jtr.831.

Crammer, K. *et al.* (2007) 'Automatic Code Assignment to Medical Text', in *Proceedings of the Workshop on BioNLP 2007: Biological, Translational, and Clinical Language Processing*, pp. 129–136. Available at: http://www.cdc.gov/nchs/about/otheract/ (Accessed: 27 March 2019).

Crotts, J. C., Dickson, D. R. and Ford, R. C. (2005) 'Aligning organizational processes with mission: The case of service excellence', *The Academy of Management Executives*, 19(3), pp. 54–68. doi: 10.5465/ame.2005.18733215.

Culotta, A. and Cutler, J. (2016) 'Mining Brand Perceptions from Twitter Social Networks', *Marketing Science*. INFORMS , 35(3), pp. 343–362. doi: 10.1287/mksc.2015.0968.

Deng, Q. and Li, M. (2014) 'A Model of Event-Destination Image Transfer', *Journal of Travel Research*. SAGE Publications Ltd, 53(1), pp. 69–82. doi: 10.1177/0047287513491331.

Denzin, N. K. and Lincoln, Y. S. (2013) *Collecting and Interpreting Qualitative Materials*. SAGE Publications.

Diplaris, S. *et al.* (2005) 'Protein Classification with Multiple Algorithms', in *Advances in Informatics. Panhellenic Conference on Informatics*. Springer, Berlin, Heidelberg, pp. 448–456. doi: 10.1007/11573036_42.

Dolnicar, S., Coltman, T. and Sharma, R. (2015) 'Do Satisfied Tourists Really Intend to Come Back? Three Concerns with Empirical Studies of the Link between Satisfaction and Behavioral Intention', *Journal of Travel Research*. SAGE Publications Ltd, 54(2), pp. 152–178. doi: 10.1177/0047287513513167.

Doong, H.-S., Wang, H.-C. and Law, R. (2012) 'An Examination of the Determinants of In-Flight Duty-Free Shopping: Hedonic and Utilitarian Motivations', *International Journal of Tourism Research*, 14(3), pp. 303–306. doi: 10.1002/jtr.

Doran, R. and Larsen, S. (2016) 'The Relative Importance of Social and Personal Norms in Explaining Intentions to Choose Eco-Friendly Travel Options', *International Journal of Tourism Research*. John Wiley and Sons Ltd, 18(2), pp. 159–166. doi: 10.1002/jtr.2042.

Dumais, S. T. (1994) 'Latent Semantic Indexing (LSI) and TREC-2', in *Proceedings of the Second Text Retrieval Conference (TREC-2)*, pp. 105–115. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.54.5141 (Accessed: 6 May 2017).

Duygulu, P. *et al.* (2002) 'Object Recognition as Machine Translation: Learning a Lexicon for a Fixed Image Vocabulary', in *Computer Vision*. Springer, Berlin, Heidelberg, pp. 97–112. doi: 10.1007/3-540-47979-1_7.

EffectiveSoft (2019) *Intellexer API*, *EffectiveSoft*. Available at: https://www.intellexer.com/intellexer_api.html (Accessed: 9 May 2019).

Engeset, M. G. and Elvekrok, I. (2015) 'Authentic Concepts: Effects on Tourist Satisfaction', *Journal of Travel Research*. SAGE Publications Ltd, 54(4), pp. 456–466. doi: 10.1177/0047287514522876.

Enrique Sucar, L. *et al.* (2014) 'Multi-label classification with Bayesian network-based chain classifiers', *Pattern Recognition Letters*. North-Holland, 41, pp. 14–22.

doi: 10.1016/J.PATREC.2013.11.007.

Eusébio, C. and Vieira, A. L. (2013) 'Destination Attributes' Evaluation, Satisfaction and Behavioural Intentions: a Structural Modelling Approach', *International Journal of Tourism Research*, 15, pp. 66–80. doi: 10.1002/jtr.

Evangelopoulos, N., Zhang, X. and Prybutok, V. R. (2012) 'Latent Semantic Analysis: Five methodological recommendations', *European Journal of Information Systems*, 21(1), pp. 70–86. doi: 10.1057/ejis.2010.61.

Fan, T. *et al.* (2019) 'The impact of disaster of a national airline on the Nation's Tourism: An empirical investigation', *Sustainability*, 11(5), pp. 1–20. doi: 10.3390/su11051233.

Fang, B. *et al.* (2016) 'Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics', *Tourism Management*. Pergamon, 52, pp. 498–506. doi: 10.1016/J.TOURMAN.2015.07.018.

Farhadloo, M., Patterson, R. A. and Rolland, E. (2016) 'Modeling customer satisfaction from unstructured data using a Bayesian approach', *Decision Support Systems*. North-Holland, 90, pp. 1–11. doi: 10.1016/J.DSS.2016.06.010.

Feldman, R. and Sanger, J. (2007) *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*, *Cambridge University Press*. doi: 10.1017/CBO9780511546914.

Filo, K. *et al.* (2011) 'Sport tourists' involvement with a destination: A stage-based examination', *Journal of Hospitality & Tourism Research*, 37(1), pp. 100–124. doi: 10.1177/1096348011425496.

Fu, X., Tanyatanaboon, M. and Lehto, X. Y. (2015) 'Conceptualizing transformative guest experience at retreat centers', *International Journal of Hospitality Management*. Elsevier Ltd, 49, pp. 83–92. doi: 10.1016/j.ijhm.2015.06.004.

Gable, G. G. (1991) *Consultant Engagement Success Factors: A Case Study and Survey of Factors Affecting Client Involvement in, and Satisfaction with, Consultant Engagement in Computer System Selection Projects, Carried Out for the Small Enterprise Computerisation Programme, in Singapore*. University of Bradford. Available at: https://eprints.qut.edu.au/126514/7/Consultant engagement success factors_Thesis.pdf (Accessed: 20 July 2019).

Galar, M. *et al.* (2011) 'An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes',

Pattern Recognition. Pergamon, 44(8), pp. 1761–1776. doi: 10.1016/J.PATCOG.2011.01.017.

Galliers, R. D. (1991) 'Choosing appropriate information systems research approaches: a revised taxonomy', *Contemporary approaches and emergent traditions*, (March), pp. 327–345.

Gallivan, M., Oppenheim, B. and Madhav, N. K. (2019) 'Using social media to estimate Zika's impact on tourism: #babymoon, 2014-2017', *PLoS ONE*, 14(2), pp. 1–10. doi: 10.1371/journal.pone.0212507.

Gálvez, R. H. and Gravano, A. (2017) 'Assessing the usefulness of online message board mining in automatic stock prediction systems', *Journal of Computational Science*, 19, pp. 43–56. doi: 10.1016/j.jocs.2017.01.001.

Gao, L. and Eldin, N. (2014) 'Employers' expectations: A probabilistic text mining model', *Procedia Engineering*, 85, pp. 175–182. doi: 10.1016/j.proeng.2014.10.542.

Gao, S. *et al.* (2018) 'Identifying competitors through comparative relation mining of online reviews in the restaurant industry', *International Journal of Hospitality Management*. Pergamon, 71, pp. 19–32. doi: 10.1016/J.IJHM.2017.09.004.

Garcia, J. *et al.* (2015) 'A study of traveller decision-making determinants: Prioritizing destination or travel mode?', *Tourism Economics*, 21(6), pp. 1149–1167. doi: 10.5367/te.2015.0517.

Gardiner, S., Grace, D. and King, C. (2015) 'Is the Australian domestic holiday a thing of the past? Understanding Baby Boomer, Generation X and Generation Y perceptions and attitude to domestic and international holidays', *Journal of Vacation Marketing*. SAGE Publications Ltd, 21(4), pp. 336–350. doi: 10.1177/1356766715580188.

Gholami, V. *et al.* (2015) 'Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers', *Journal of Hydrology*. Elsevier B.V., 529, pp. 1060–1069. doi: 10.1016/j.jhydrol.2015.09.028.

Ghorbanzade, D., Mehrani, H. and Rahehagh, A. (2019) 'The effect of experience quality on behavioral intentions of domestic tourists in visiting water parks', *Cogent Business and Management*. Cogent, 6(1), pp. 1–24. doi: 10.1080/23311975.2019.1580843.

Giraldi, A. and Cesareo, L. (2014) 'Destination image differences between first-time and return visitors: An exploratory study on the city of Rome', *Tourism and Hospitality Research*, 14(4), pp. 197–205. doi: 10.1177/1467358414543971.

Gonçalves, T. and Quaresma, P. (2003) 'A Preliminary Approach to the Multilabel Classification Problem of Portuguese Juridical Documents', in *Progress in Artificial Intelligence*. Springer, Berlin, Heidelberg, pp. 435–444. doi: 10.1007/978-3-540-24580-3_50.

Gretzel, U. *et al.* (2010) 'Narrating travel experiences: the role of new media', *Tourist Experiences: Contemporary Perspectives*, pp. 171–182.

Grigaliūnaitė, V. and Pilelienė, L. (2018) 'Assessment of Lithuanian rural homesteads' websites quality', *Research for Rural Development*, 2, pp. 231–237. doi: 10.22616/rrd.24.2018.077.

Guan, X. *et al.* (2017) 'A multi-view OVA model based on decision tree for multi-classification tasks', *Knowledge-Based Systems*. Elsevier, 138, pp. 208–219. doi: 10.1016/J.KNOSYS.2017.10.004.

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

Gursoy, D., Del Chiappa, G. and Zhang, Y. (2017) 'Preferences regarding external information sources: a conjoint analysis of visitors to Sardinia, Italy', *Journal of Travel & Tourism Marketing*, 34(6), pp. 806–820. doi: 10.1080/10548408.2016.1237405.

Hague, P. and Hague, N. (2019) *Customer Satisfaction Survey Questions: How to Measure Satisfaction, B2B International*. Available at: https://www.b2binternational.com/publications/customer-satisfaction-survey/ (Accessed: 5 March 2019).

Hallmann, K., Zehrer, A. and Müller, S. (2015) 'Perceived Destination Image: An Image Model for a Winter Sports Destination and Its Effect on Intention to Revisit', *Journal of Travel Research*. SAGE Publications Ltd, 54(1), pp. 94–106. doi: 10.1177/0047287513513161.

Hardan Suleiman, J. S. and Mohamed, B. (2012) 'Palestine's International Visitors Barriers: Case Studies of Bethlehem and Ramallah', *International Journal of Tourism Research*, 14(2), pp. 177–191. doi: 10.1002/jtr.

Hasan, M. K., Ismail, A. R. and Islam, M. F. (2017) 'Tourist risk perceptions and revisit intention: A critical review of literature', *Cogent Business and Management*. Cogent, 4, pp. 1–21. doi: 10.1080/23311975.2017.1412874.

Hearst, M. A. (1999) 'Untangling Text Data Mining', in *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*. Morristown, NJ, USA: Association for Computational Linguistics, pp. 3–10. doi: 10.3115/1034678.1034679.

Hernández, J. M., Kirilenko, A. P. and Stepchenkova, S. (2018) 'Network approach to tourist segmentation via user generated content', *Annals of Tourism Research*. Elsevier, 73, pp. 35–47. doi: 10.1016/j.annals.2018.09.002.

Herrera, F. *et al.* (2016) 'Multilabel classification', in *Multilabel Classification: Problem Analysis, Metrics and Techniques*. Springer International Publishing Switzerland, pp. 17–31. doi: 10.1007/978-3-319-41111-8.

Hirschmann, R. (2020a) *Gross value added tourism industries (GVATI) in Malaysia from 2011 to 2019*. Available at: https://www.statista.com/statistics/954663/gross-value-added-tourism-industries-malaysia/.

Hirschmann, R. (2020b) *Number of people employed in the tourism industry in Malaysia from 2011 to 2019*. Available at: https://www.statista.com/statistics/954719/employment-tourism-industries-malaysia/.

Hirschmann, R. (2020c) *Travel and tourism in Malaysia - Statistics & Facts*. Available at: https://www.statista.com/topics/5741/travel-and-tourism-in-malaysia/#:~:text=Tourism is the third biggest,percent to the total GDP.

Hirschmann, R. (2020d) *Year on year growth of tourist arrivals to Malaysia from 2010 to 2019*. Available at: https://www.statista.com/statistics/1004724/tourist-arrival-growth-malaysia/.

Hlee, S. *et al.* (2018) 'The moderating effect of restaurant type on hedonic versus utilitarian review evaluations', *International Journal of Hospitality Management*. Pergamon, 77, pp. 195–206. doi: 10.1016/J.IJHM.2018.06.030.

Hovold, J. (2005) 'Naive Bayes Spam Filtering Using Word-Position-Based Attributes', *Proceedings of the Second Conference on Email and Anti-Spam (CEAS)*. Available at: http://fileadmin.cs.lth.se/ai/psfiles/ceas-05b.pdf.

Hsu, C. H. C. and Huang, S. (Sam) (2012) 'An Extension of the Theory of Planned Behavior Model for Tourists', *Journal of Hospitality & Tourism Research*, 36(3), pp. 390–417. doi: 10.1177/1096348010390817.

Hu, Y.-H., Chen, Y.-L. and Chou, H.-L. (2017) 'Opinion mining from online hotel

reviews – A text summarization approach', *Information Processing and Management*. Pergamon, 53(2), pp. 436–449. doi: 10.1016/j.ipm.2016.12.002.

Huang, S. (Sam), Weiler, B. and Assaker, G. (2015) 'Effects of interpretive guiding outcomes on tourist satisfaction and behavioral intention', *Journal of Travel Research*. SAGE Publications Ltd, 54(3), pp. 344–358. doi: 10.1177/0047287513517426.

Huang, Y. *et al.* (2012) 'Impression of Liusanjie: Effect of mood on experience and satisfaction', *International Journal of Tourism Research*, 14(1), pp. 91–102. doi: 10.1002/jtr.829.

Huang, Y. C. *et al.* (2016) 'Exploring the Implications of Virtual Reality Technology in Tourism Marketing: An Integrated Research Framework', *International Journal of Tourism Research*. John Wiley and Sons Ltd, 18(2), pp. 116–128. doi: 10.1002/jtr.2038.

Hung, K. and Petrick, J. F. (2011) 'The role of self- and functional congruity in cruising intentions', *Journal of Travel Research*. SAGE Publications Ltd, 50(1), pp. 100–112. doi: 10.1177/0047287509355321.

Iowa State University of Science and Technology (2019) *Google Scholar, Scopus, Web of Science - Database Comparisons, Iowa State University of Science and Technology.* Available at: https://instr.iastate.libguides.com/c.php?g=901522&p=6492159.

Jiménez Barreto, J., Rubio, N. and Campo Martínez, S. (2018) 'The online destination brand experience: Development of a sensorial–cognitive–conative model', *International Journal of Tourism Research*, 21(2), pp. 245–258. doi: 10.1002/jtr.2258.

Jin, N. (Paul_, Lee, S. and Lee, H. (2015) 'The effect of experience quality on perceived value, satisfaction, image and behavioral intention of water park patrons: New versus repeat visitors', *International Journal of Tourism Research*. John Wiley and Sons Ltd, 17(1), pp. 82–95. doi: 10.1002/jtr.1968.

Joshi, A. (2016) 'Comparison Between Scopus & ISI Web of Science', *Journal Global Values ISSN*, VII(1), pp. 976–9447.

Juan, P. J. and Chen, H. M. (2012) 'Taiwanese Cruise Tourist Behavior during Different Phases of Experience', *International Journal of Tourism Research*, 14(5), pp. 485–494. doi: 10.1002/jtr.882.

Jun, J., Kim, J. and Tang, L. (Rebecca) (2017) 'Does Social Capital Matter on Social Media? An Examination Into Negative e-WOM Toward Competing Brands', *Journal*

of Hospitality Marketing & Management, 26(4), pp. 378–394. doi: 10.1080/19368623.2017.1251869.

Jung, T. *et al.* (2015) 'Influence of festival attribute qualities on Slow Food tourists' experience, satisfaction level and revisitintention: The case of the Mold Slow Food and Drink Festival', *Journal of Vacation Marketing*, 21(3), pp. 277–288. doi: 10.1177/1356766715571389.

Karakatsanis, I. *et al.* (2017) 'Data mining approach to monitoring the requirements of the job market: A case study', *Information Systems*, 65, pp. 1–6. doi: 10.1016/j.is.2016.10.009.

Katakis, I., Tsoumakas, G. and Vlahavas, I. P. (2008) 'Multilabel Text Classification for Automated Tag Suggestion', in *Proceedings of the ECML/PKDD 2008 Discovery Challenge*.

Khan, M. J. *et al.* (2017) 'Role of travel motivations, perceived risks and travel constraints on destination image and visit intention in medical tourism: Theoretical model', *Sultan Qaboos University Medical Journal*, 17(1), pp. 11–17. doi: 10.18295/squmj.2016.17.01.003.

Kim, C. and Kim, S. (2016) 'Measuring the operational efficiency of individual theme park attractions', *SpringerPlus*. Springer International Publishing, 5(1), pp. 834–843. doi: 10.1186/s40064-016-2530-9.

Kim, D. *et al.* (2019) 'Determinants and outcomes of volunteer satisfaction in mega sports events', *Sustainability*, 11(7), pp. 1–18. doi: 10.3390/su11071859.

Kim, K. *et al.* (2017) 'What makes tourists feel negatively about tourism destinations? Application of hybrid text mining methodology to smart destination management', *Technological Forecasting & Social Change*. Elsevier Inc., 123, pp. 362–369. doi: 10.1016/j.techfore.2017.01.001.

Kim, M.-S., Thapa, B. and Kim, H. (2018) 'International tourists' perceived sustainability of Jeju Island, South Korea', *Sustainability*, 10(1), pp. 73–86. doi: 10.3390/su10010073.

Kim, M. J. *et al.* (2013) 'Factors Affecting Online Tourism Group Buying and the Moderating Role of Loyalty', *Journal of Travel Research*. SAGE Publications Ltd, 53(3), pp. 380–394. doi: 10.1177/0047287513497837.

Kim, S.-B. and Kwon, K.-J. (2018) 'Examining the relationships of image and attitude on visit intention to Korea among Tanzanian college students: The moderating effect

of familiarity', *Sustainability*, 10(2), pp. 360–375. doi: 10.3390/su10020360.

Kim, S.-H., Holland, S. and Han, H.-S. (2013) 'A structural model for examining how destination image, perceived value, and service quality affect destination loyalty: A case study of Orlando', *International Journal of Tourism Research*, 15(4), pp. 313–328. doi: 10.1002/jtr.1877.

Kim, W. *et al.* (2017) 'Destination personality, destination image, and intent to recommend: The role of gender, age, cultural background, and prior experiences', *Sustainability*, 10(1), pp. 87–105. doi: 10.3390/su10010087.

Kirilenko, A. P., Stepchenkova, S. O. and Hernandez, J. M. (2019) 'Comparative clustering of destination attractions for different origin markets with network and spatial analyses of online reviews', *Tourism Management*. Pergamon, 72, pp. 400–410. doi: 10.1016/J.TOURMAN.2019.01.001.

Kitchenham, B. (2004) 'Procedures for performing systematic reviews', *Keele, UK, Keele University*, 33(TR/SE-0401), p. 28. doi: 10.1.1.122.3308.

Klimt, B. and Yang, Y. (2004) 'The Enron Corpus: A New Dataset for Email Classification Research', in JF., B. et al. (eds) *Machine Learning*. Springer, Berlin, Heidelberg, pp. 217–226. doi: 10.1007/978-3-540-30115-8_22.

Kozak, M. and Duman, T. (2012) 'Family Members and Vacation Satisfaction: Proposal of a Conceptual Framework', *International Journal of Tourism Research*, 14, pp. 192–204. doi: 10.1002/jtr.

Kwon, H., Kim, J. and Park, Y. (2017) 'Applying LSA text mining technique in envisioning social impacts of emerging technologies: The case of drone technology', *Technovation*. doi: 10.1016/j.technovation.2017.01.001.

Larsen, H. G. (2014) 'The emerging Shanghai city brand: A netnographic study of image perception among foreigners', *Journal of Destination Marketing & Management*. Elsevier, 3(1), pp. 18–28. doi: 10.1016/j.jdmm.2013.12.003.

Lee, C.-H. *et al.* (2018) 'Evaluating international tourists' perceptions on cultural distance and recreation demand', *Sustainability*, 10(12), pp. 1–14. doi: 10.3390/su10124360.

Lee, C.-K. *et al.* (2012) 'Thanatourism or peace tourism: Perceived value at a North Korean resort from an indigenous perspective', *International Journal of Tourism Research*, 14(1), pp. 71–90. doi: 10.1002/jtr.836.

Lee, H., Chung, N. and Nam, Y. (2018) 'Do online information sources really make

tourists visit more diverse places?: Based on the social networking analysis', *Information Processing and Management*. Elsevier. doi: 10.1016/j.ipm.2018.01.005.

Lee, J. (Jiyeon) and Kyle, G. T. (2014) 'Segmenting Festival Visitors Using Psychological Commitment', *Journal of Travel Research*. SAGE Publications Ltd, 53(5), pp. 656–669. doi: 10.1177/0047287513513168.

Lee, J. H., Jung, S. H. and Park, J. (2017) 'The role of entropy of review text sentiments on online WOM and movie box office sales', *Electronic Commerce Research and Applications*, 22, pp. 42–52. doi: 10.1016/j.elerap.2017.03.001.

Lee, R. and Lockshin, L. (2012) 'Reverse Country-of-Origin Effects of Product Perceptions on Destination Image', *Journal of Travel Research*. SAGE Publications Ltd, 51(4), pp. 502–511. doi: 10.1177/0047287511418371.

Leelapattana, W. *et al.* (2019) 'Understanding the impact of environmental education on tourists' future Visit Intentions to Leisure Farms in Mountain Regions', *Sustainability*, 11(6), pp. 1567–1580. doi: 10.3390/su11061567.

Lexalytics (2019) *Semantria Cloud API Text & Sentiment Analysis, Lexalytics*. Available at: https://www.lexalytics.com/semantria (Accessed: 9 May 2019).

Li, C. Y. and Tsai, B. K. (2013) 'Impact of extraversion and sensation seeking on international tourism choices', *Social Behaviour and Personality*, 41(2), pp. 327–334. doi: 10.2224/sbp.2013.41.2.327.

Li, H. *et al.* (2018) 'Making restaurant reviews useful and/or enjoyable? The impacts of temporal, explanatory, and sensory cues', *International Journal of Hospitality Management*. Pergamon. doi: 10.1016/J.IJHM.2018.11.002.

Li, J. *et al.* (2016) 'Up in Smoke? The Impact of Smog on Risk Perception and Satisfaction of International Tourists in Beijing', *International Journal of Tourism Research*. John Wiley and Sons Ltd, 18(4), pp. 373–386. doi: 10.1002/jtr.2055.

Li, M. and Cai, L. A. (2012) 'The Effects of Personal Values on Travel Motivation and Behavioral Intention', *Journal of Travel Research*. SAGE Publications Ltd, 51(4), pp. 473–487. doi: 10.1177/0047287511418366.

Lin, C.-H. (2014) 'Effects of Cuisine Experience, Psychological Well-Being, and Self-Health Perception on the Revisit Intention of Hot Springs Tourists', *Journal of Hospitality & Tourism Research*, 38(2), pp. 243–265. doi: 10.1177/1096348012451460.

Lin, C. H. (2014) 'Effects of Cuisine Experience, Psychological Well-Being, and Self-

Health Perception on the Revisit Intention of Hot Springs Tourists'. SAGE Publications Inc., 38(2), pp. 243–265. doi: 10.1177/1096348012451460.

Lin, L.-Z. and Hsu, T.-H. (2013) 'The analysis of risk perception with fuzzy means-end approach', *Qual Quant*, 47, pp. 713–734. doi: 10.1007/s11135-011-9540-z.

Littman, M. L., Dumais, S. T. and Landauer, T. K. (1998) 'Automatic Cross-Language Information Retrieval Using Latent Semantic Indexing', in Grefenstette, G. (ed.) *Cross-Language Information Retrieval*. 2nd edn. Boston, MA: Springer, Boston, MA, pp. 51–62. doi: 10.1007/978-1-4615-5661-9_5.

Liu, S. M. and Chen, J.-H. (2015) 'A multi-label classification based approach for sentiment classification', *Expert Systems with Applications*. Elsevier Ltd, 42(3), pp. 1083–1093. doi: 10.1016/j.eswa.2014.08.036.

Liu, X., Li, J. J. and Yang, Y. (2015) 'Travel arrangement as a moderator in image-satisfaction-behavior relations: An investigation of Chinese outbound travelers', *Journal of Vacation Marketing*, 21(3), pp. 225–236. doi: 10.1177/1356766714567797.

Liu, Y. *et al.* (2017) 'Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews', *Tourism Management*. Pergamon, 59, pp. 554–563. doi: 10.1016/J.TOURMAN.2016.08.012.

Liu, Y. *et al.* (2019) 'Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations', *Tourism Management*. Pergamon, 71, pp. 337–347. doi: 10.1016/J.TOURMAN.2018.10.004.

Liu, Y., Bi, J.-W. and Fan, Z.-P. (2017) 'Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory', *Information Fusion*. Elsevier, 36, pp. 149–161. doi: 10.1016/J.INFFUS.2016.11.012.

Liu, Y., Jiang, C. and Zhao, H. (2018) 'Using contextual features and multi-view ensemble learning in product defect identification from online discussion forums', *Decision Support Systems*. North-Holland, 105, pp. 1–12. doi: 10.1016/J.DSS.2017.10.009.

Lochbaum, K. E. and Streeter, L. A. (1989) 'Comparing and combining the effectiveness of latent semantic indexing and the ordinary vector space model for information retrieval', *Information Processing & Management*, 25(6), pp. 665–676. doi: 10.1016/0306-4573(89)90100-3.

Lončarić, D., Perišić Prodan, M. and Bagarić, L. (2018) 'The Relationship Between Tourism Experience Co-Creation, Life Satisfaction and Behavioural Intentions', in

Central European Business Review, pp. 1–14.

Long, N. T. and Nguyen, T.-L. (2018) 'Sustainable development of rural tourism in An Giang Province, Vietnam', *Sustainability*, 10(4), pp. 953–973. doi: 10.3390/su10040953.

Loureiro, S. M. C. (2015) 'The Role of Website Quality on PAD, Attitude and Intentions to Visit and Recommend Island Destination', *International Journal of Tourism Research*. John Wiley and Sons Ltd, 17(6), pp. 545–554. doi: 10.1002/jtr.2022.

Lu, A. C. C., Gursoy, D. and Lu, C. Y. R. (2016) 'Antecedents and outcomes of consumers' confusion in the online tourism domain', *Annals of Tourism Research*. Pergamon, 57, pp. 76–93. doi: 10.1016/J.ANNALS.2015.11.020.

Lu, W. and Stepchenkova, S. (2012) 'Ecotourism experiences reported online: Classification of satisfaction attributes', *Tourism Management*. Pergamon, 33(3), pp. 702–712. doi: 10.1016/J.TOURMAN.2011.08.003.

Luaces, O. *et al.* (2012) 'Binary relevance efficacy for multilabel classification', *Progress in Artificial Intelligence*. Springer-Verlag, 1(4), pp. 303–313. doi: 10.1007/s13748-012-0030-x.

Lucas, M. (2000) 'Mining in Textual Mountains, an Interview with Marti Hearst', *Mappa Mundi Magazine*, pp. 1–3. Available at: http://mappa.mundi.net/trip-m/hearst/ (Accessed: 4 May 2017).

Luo, S. J. and Hsieh, L. Y. (2013) 'Reconstructing revisit intention scale in tourism', *Journal of Applied Science*, 13(18), pp. 3638–3648. doi: 10.3923/jas.2013.3638.3648.

Magnini, V. P., Crotts, J. C. and Zehrer, A. (2011) 'Understanding customer delight: An application of travel blog analysis', *Journal of Travel Research*. SAGE Publications Ltd, 50(5), pp. 535–545. doi: 10.1177/0047287510379162.

Maharani, W., Widyantoro, D. H. and Khodra, M. L. (2015) 'Aspect Extraction in Customer Reviews Using Syntactic Pattern', *Procedia Computer Science*. Elsevier Masson SAS, 59(Iccsci), pp. 244–253. doi: 10.1016/j.procs.2015.07.545.

Majeed, S. *et al.* (2018) 'Health resorts and multi-textured perceptions of international health tourists', *Sustainability*, 10(4), pp. 1063–1089. doi: 10.3390/su10041063.

Maletic, J. I. and Marcus, A. (2000) 'Using latent semantic analysis to identify similarities in source code to support program understanding', in *Proceedings of the 12th IEEE International Conference on the Tools with Artificial Intelligence*.

Vancouver, BC: IEEE Comput. Soc, pp. 46–53. doi: 10.1109/TAI.2000.889845.

Manning, C. D., Raghavan, P. and Schutze, H. (2009) *An Introduction To Informational Retrieval*, *Cambridge University Press*. doi: 10.1109/LPT.2009.2020494.

Mao, Z., Yang, Y. and Wang, M. (2018) 'Sleepless nights in hotels? Understanding factors that influence hotel sleep quality', *International Journal of Hospitality Management*. Pergamon, 74, pp. 189–201. doi: 10.1016/J.IJHM.2018.05.002.

Marine-Roig, E. and Ferrer-Rosell, B. (2018) 'Measuring the gap between projected and perceived destination images of Catalonia using compositional analysis', *Tourism Management*. Pergamon, 68, pp. 236–249. doi: 10.1016/J.TOURMAN.2018.03.020.

Marrese-Taylor, E., Velasquez, J. D. and Bravo-Marquez, F. (2013) 'OpinionZoom, a modular tool to explore tourism opinions on the Web', *Proceedings - 2013 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Workshops, WI-IATW 2013*, 3, pp. 261–264. doi: 10.1109/WI-IAT.2013.193.

Marrese-Taylor, E., Velásquez, J. D. and Bravo-Marquez, F. (2014) 'A novel deterministic approach for aspect-based opinion mining in tourism products reviews', *Expert Systems with Applications*, 41(17), pp. 7764–7775. doi: 10.1016/j.eswa.2014.05.045.

Marso and Gunawan (2018) 'Destination image and its consequences in the perspective of four-stage loyalty model (An empirical evidence from visitors of Tarakan City, Indonesia)', *Problems and Perspectives in Management*, 16(2), pp. 269–278. doi: 10.21511/ppm.16(2).2018.25.

Martín-Valdivia, M. *et al.* (2013) 'Sentiment polarity detection in Spanish reviews combining supervised and unsupervised approaches', *Expert Systems with Applications*, 40, pp. 3934–3942. doi: 10.1016/j.eswa.2012.12.084.

McCabe, S., Li, C. (Spring) and Chen, Z. (2016) 'Time for a Radical Reappraisal of Tourist Decision Making? Toward a New Conceptual Model', *Journal of Travel Research*. SAGE Publications Ltd, 55(1), pp. 3–15. doi: 10.1177/0047287515592973.

McKercher, B., Guillet, B. D. and Denizci Guillet, B. (2010) 'Are Tourists or Markets Destination Loyal?', *Journal of Travel Research*. SAGE Publications Ltd, 50(2), pp. 121–132. doi: 10.1177/0047287510362779.

McKercher, B. and Tse, T. S. (2012) 'Is Intention to Return a Valid Proxy for Actual

Repeat Visitation?' SAGE Publications Ltd, 51(6), pp. 671–686. doi: 10.1177/0047287512451140.

MeaningCloud (2019) *Sentiment Analysis API, MeaningCloud.* Available at: https://www.meaningcloud.com/developer/sentiment-analysis (Accessed: 9 May 2019).

Mehmood, S., Liang, C. and Gu, D. (2018) 'Heritage image and attitudes toward a heritage site: Do they really mediate the relationship between user-generated content and travel intentions toward a heritage site?', *Sustainability*, 10(12), pp. 4403–4429. doi: 10.3390/su10124403.

Microsoft Azure (2019) *Text Analytics API, Microsoft.* Available at: https://azure.microsoft.com/en-in/services/cognitive-services/text-analytics/ (Accessed: 9 May 2019).

Millán, A., Fanjul, M. L. and Moital, M. (2016) 'Segmenting the Business Traveler Based on Emotions, Satisfaction, and Behavioral Intention', *Psychology & Marketing.* Wiley-Liss Inc., 33(2), pp. 82–93. doi: 10.1002/mar.20856.

Mohammed Abubakar, A. (2016) 'Does eWOM influence destination trust and travel intention: A medical tourism perspective', *Economic Research.* Routledge, 29(1), pp. 598–611. doi: 10.1080/1331677X.2016.1189841.

Mohammed Badr, B. and Sameen Fatima, S. (2015) 'Using Skipgrams, Bigrams, and Part of Speech Features for Sentiment Classification of Twitter Messages', in *Proceedings of the 12th International Conference on Natural Language Processing.* NLPAI, pp. 268–275. Available at: http://www.sentiment140.com (Accessed: 10 March 2019).

Moro, S. *et al.* (2019) 'Can we trace back hotel online reviews' characteristics using gamification features?', *International Journal of Information Management.* Pergamon, 44, pp. 88–95. doi: 10.1016/J.IJINFOMGT.2018.09.015.

Moro, S., Cortez, P. and Rita, P. (2015) 'Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation', *Expert Systems With Applications*, 42, pp. 1314–1324. doi: 10.1016/j.eswa.2014.09.024.

Moro, S., Rita, P. and Coelho, J. (2017) 'Stripping customers' feedback on hotels through data mining: The case of Las Vegas Strip', *Tourism Management Perspectives.* Elsevier, 23, pp. 41–52. doi: 10.1016/J.TMP.2017.04.003.

Morris, C. (2010) *Boston Tech-Mafia Mondays: TripAdvisor - Pioneers in Online*

Travel, BostInno. Available at: https://www.americaninno.com/boston/boston-tech-mafia-mondays-tripadvisor-pioneers-in-online-travel/ (Accessed: 7 February 2019).

Moufakkir, O. (2013) 'Culture shock, what culture shock? Conceptualizing culture unrest in intercultural tourism and assessing its effect on tourists' perceptions and travel propensity', *Tourist Studies*, 13(3), pp. 322–340. doi: 10.1177/1468797613498166.

Moutinho, L., Albayrak, T. and Caber, M. (2012) 'How Far does Overall Service Quality of a Destination Affect Customers' Post-Purchase Behaviours?', *International Journal of Tourism Research*, 14(4), pp. 307–322. doi: 10.1002/jtr.856.

Mudinas, A., Zhang, D. and Levene, M. (2012) 'Combining lexicon and learning based approaches for concept-level sentiment analysis'. doi: 10.1145/2346676.2346681.

Mukherjee, A. and Liu, B. (2012) 'Aspect extraction through semi-supervised modeling', in *50th Annual Meeting of the Association for Computational Linguistics, ACL 2012 - Proceedings of the Conference*, pp. 339–348.

Mustelier-Puig, L. C., Anjum, A. and Ming, X. (2018) 'Interaction quality and satisfaction: An empirical study of international tourists when buying Shanghai tourist attraction services', *Cogent Business & Management*. Cogent, 5(1), pp. 1–20. doi: 10.1080/23311975.2018.1470890.

Nakayama, M. and Wan, Y. (2018) 'Is culture of origin associated with more expressions? An analysis of Yelp reviews on Japanese restaurants', *Tourism Management*. Pergamon, 66, pp. 329–338. doi: 10.1016/J.TOURMAN.2017.10.019.

Nakayama, M. and Wan, Y. (2019) 'The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews', *Information & Management*. North-Holland, 56(2), pp. 271–279. doi: 10.1016/J.IM.2018.09.004.

Nassirtoussi, A. K. *et al.* (2014) 'Text mining for market prediction: A systematic review'. doi: 10.1016/j.eswa.2014.06.009.

Nawijn, J. and Fricke, M.-C. (2015) 'Visitor Emotions and Behavioral Intentions: The Case of Concentration Camp Memorial Neuengamme', *International Journal of Tourism Research*, 17, pp. 221–228. doi: 10.1002/jtr.

Nghiêm-Phú, B. and Suter, J. R. (2018) 'Airport image: An exploratory study of McCarran International Airport', *Journal of Air Transport Management*. Elsevier Ltd, 67, pp. 72–84. doi: 10.1016/j.jairtraman.2017.11.011.

Niemczyk, A. (2014) 'The Application of Path Modelling in the Analysis of Consumer

Behaviour in the Cultural Tourism Market', *Economics & Sociology*, 7(1), pp. 204–216. doi: 10.14254/2071-789X.2014/7-1/18.

Nilashi, M. *et al.* (2018) 'Travelers decision making using online review in social network sites: A case on TripAdvisor', *Journal of Computational Science*. Elsevier, 28, pp. 168–179. doi: 10.1016/J.JOCS.2018.09.006.

De Nisco, A. *et al.* (2015) 'Tourism satisfaction effect on general country image, destination image, and post-visit intentions', *Journal of Vacation Marketing*. SAGE Publications Ltd, 21(4), pp. 305–317. doi: 10.1177/1356766715577502.

Oliver, R. L. (1977) 'Effect of expectation and disconfirmation on postexposure product evaluations: An alternative interpretation', *Journal of Applied Psychology*, 62(4), pp. 480–486. doi: 10.1037/0021-9010.62.4.480.

Ooi, T. C. (2019) 'Malaysia's economy received RM84b from tourists', *New Straits Times*. Available at: https://www.nst.com.my/business/2019/05/487762/malaysias-economy-received-rm84b-tourists (Accessed: 17 July 2019).

Osti, L., Disegna, M. and Brida, J. G. (2012) 'Repeat visits and intentions to revisit a sporting event and its nearby destinations', *Journal of Vacation Marketing*, 18(1), pp. 31–42. doi: 10.1177/1356766711428803.

Palau-Saumell, R. *et al.* (2012) 'Tourist behavior intentions and the moderator effect of knowledge of UNESCO world heritage sites: The case of La Sagrada Familia.', *Journal of Travel Research*. SAGE Publications Ltd, 52(3), pp. 364–376. doi: 10.1177/0047287512465959.

Pantano, E., Priporas, C.-V. and Stylos, N. (2017) '"You will like it!" using open data to predict tourists' response to a tourist attraction', *Tourism Management*. Pergamon, 60, pp. 430–438. doi: 10.1016/J.TOURMAN.2016.12.020.

Papadimitriou, D., Apostolopoulou, A. and Kaplanidou, K. (Kiki) (2015) 'Destination Personality, Affective Image, and Behavioral Intentions in Domestic Urban Tourism', *Journal of Travel Research*. SAGE Publications Ltd, 54(3), pp. 302–315. doi: 10.1177/0047287513516389.

Park, J.-Y. and Jang, S. (Shawn) (2014) 'Psychographics: Static or Dynamic?', *International Journal of Tourism Research*, 16, pp. 351–354. doi: 10.1002/jtr.

Park, S. and Kim, D.-Y. (2017) 'Assessing language discrepancies between travelers and online travel recommendation systems: Application of the Jaccard distance score to web data mining', *Technological Forecasting & Social Change*. doi:

10.1016/j.techfore.2017.03.031.

Pavesi, A., Gartner, W. and Denizci-Guillet, B. (2015) 'The Effects of a Negative Travel Experience on Tourists' Decisional Behavior', *International Journal of Tourism Research*. doi: 10.1002/jtr.

Pekar, V. and Ou, S. (2008) 'Discovery of subjective evaluations of product features in hotel reviews', *Journal of Vacation Marketing*, 14(2), pp. 145–155. doi: 10.1177/1356766707087522.

Peranginangin, J. (2019) 'Antecedents of revisited intention: Evidence from Indonesia zoo tourism', *Contaduría y Administración*, 64(1), pp. 1–14. doi: 10.22201/fca.24488410e.2018.1518.

Perikos, I. *et al.* (2017) 'A system for aspect-based opinion mining of hotel reviews', *WEBIST 2017 - Proceedings of the 13th International Conference on Web Information Systems and Technologies*, (Webist), pp. 388–394. doi: 10.5220/0006377103880394.

Perkins, J. (2010) *API Documentation for text-processing.com*. Available at: https://text-processing.com/docs/ (Accessed: 9 May 2019).

Pham, D. H. and Le, A. C. (2016) 'A neural network based model for determining overall aspect weights in opinion mining and sentiment analysis', *Indian Journal of Science and Technology*, 9(18), pp. 1–6. doi: 10.17485/ijst/2016/v9i18/93164.

Phillips, W. J. *et al.* (2013) 'Tourist Word of Mouth and Revisit Intentions to Rural Tourism Destinations: A Case of North Dakota, USA', *International Journal of Tourism Research*, 15(1), pp. 93–104. doi: 10.1002/jtr.879.

Pike, S. and Bianchi, C. (2016) 'Destination Brand Equity for Australia: Testing a Model of CBBE in Short-Haul and Long-Haul Markets', *Journal of Hospitality & Tourism Research*, 40(1), pp. 114–134. doi: 10.1177/1096348013491604.

Pillay, S. R. and Solorio, T. (2010) 'Authorship attribution of web forum posts', in *2010 eCrime Researchers Summit*. IEEE, pp. 1–7. doi: 10.1109/ecrime.2010.5706693.

Plank, A. (2016) 'The hidden risk in user-generated content: An investigation of ski tourers' revealed risk-taking behavior on an online outdoor sports platform', *Tourism Management*. Elsevier Ltd, 55, pp. 289–296. doi: 10.1016/j.tourman.2016.02.013.

Plog, S. (2001) 'Why destination areas rise and fall in popularity: an update of a Cornell Quarterly classic', *The Cornell Hotel and Restaurant Administration Quarterly*. No longer published by Elsevier, 42(3), pp. 13–24. doi: 10.1016/S0010-8804(01)81020-X.

Plog, S. C. (2002) 'The Power of Psychographics and the Concept of Venturesomeness', *Journal of Travel Research*, 40(3), pp. 244–251. doi: 10.1177/004728750204000302.

Polo Peña, A. I., Frías Jamilena, D. M. and Rodríguez Molina, M. Á. (2012) 'Market Orientation as a Strategy for the Rural Tourism Sector: Its Effect on Tourist Behavior and the Performance of Enterprises', *Journal of Travel Research*. SAGE Publications Ltd, 52(2), pp. 225–239. doi: 10.1177/0047287512459108.

Pontiki, M. *et al.* (2014) 'SemEval-2014 Task 4: Aspect Based Sentiment Analysis', in *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pp. 27–35. doi: 10.3115/v1/s14-2004.

Pontiki, M. *et al.* (2015) 'SemEval-2015 Task 12: Aspect Based Sentiment Analysis', in *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*. Association for Computational Linguistics, pp. 486–495. doi: 10.18653/v1/s15-2082.

Pontiki, M. *et al.* (2019) 'SemEval-2016 Task 5 : Aspect Based Sentiment Analysis To cite this version : HAL Id : hal-02407165 SemEval-2016 Task 5 : Aspect Based Sentiment Analysis', in *10th International Workshop on Se- mantic Evaluation (SemEval 2016)*.

Pournarakis, D. E., Sotiropoulos, D. N. and Giaglis, G. M. (2017) 'A computational model for mining consumer perceptions in social media', *Decision Support Systems*. North-Holland, 93, pp. 98–110. doi: 10.1016/J.DSS.2016.09.018.

Prakash, S. L. *et al.* (2018) 'Reasons for visitor dissatisfaction with wildlife tourism experiences at highly visited national parks in Sri Lanka', *Journal of Outdoor Recreation and Tourism*. Elsevier Ltd. doi: 10.1016/j.jort.2018.07.004.

Prayag, G. (2012) 'Paradise for who? Segmenting visitors' satisfaction with cognitive image and predicting behavioural loyalty', *International Journal of Tourism Research*, 14(1), pp. 1–15. doi: 10.1002/jtr.837.

Prayag, G. and Jankee, M. (2013) 'Perceived risk, satisfaction and future behaviour of windsurfers: A segmentation perspective', *International Journal of Tourism Policy*, 5(1), pp. 19–33. doi: 10.1504/IJTP.2013.054054.

Prebensen, N. K. and Lee, Y.-S. (2013) 'Why visit an eco-friendly destination? Perspectives of four European nationalities', *Journal of Vacation Marketing*, 19(2), pp. 105–116. doi: 10.1177/1356766712457671.

Preko, A., Doe, F. and Dadzie, S. A. (2018) 'The future of youth tourism in Ghana: Motives, satisfaction and behavioural intentions', *Journal of Tourism Futures*. doi: 10.1108/JTF-12-2016-0059.

Priyantina, R. A. and Sarno, R. (2019) 'Sentiment analysis of hotel reviews using Latent Dirichlet Allocation, semantic similarity and LSTM', *International Journal of Intelligent Engineering and Systems*, 12(4), pp. 142–155. doi: 10.22266/ijies2019.0831.14.

Puri, C. *et al.* (2017) 'Opinion Ensembling for Improving Economic Growth through Tourism', *Procedia Computer Science*. Elsevier, 122, pp. 237–244. doi: 10.1016/J.PROCS.2017.11.366.

Qi, J. *et al.* (2016) 'Mining customer requirements from online reviews: A product improvement perspective', *Information & Management*. Elsevier B.V., 53(8), pp. 951–963. doi: 10.1016/j.im.2016.06.002.

Qiao, G. *et al.* (2019) 'Identifying causes for the decline in international arrivals to China-perspective of sustainable inbound tourism development', *Sustainability*, 11(6), pp. 1723–1741. doi: 10.3390/su11061723.

Quadri-Felitti, D. L. and Fiore, A. M. (2013) 'Destination loyalty: Effects of wine tourists' experiences, memories, and satisfaction on intentions', *Tourism and Hospitality Research*. SAGE Publications Inc., 13(1), pp. 47–62. doi: 10.1177/1467358413510017.

Rakuten RapidAPI (2018) *Top 10 Sentiment Analysis APIs: Microsoft, Aylien and more, Rakuten RapidAPI Blog*. Available at: https://blog.api.rakuten.net/top-sentiment-analysis-api/ (Accessed: 9 May 2019).

Rehder, B. *et al.* (1998) 'Using Latent Semantic Analysis to assess knowledge: Some technical consierations', *Discourse Processes*, 25(2–3), pp. 337–354.

Ren, G. and Hong, T. (2017) 'Investigating online destination images using a topic-based sentiment analysis approach', *Sustainability (Switzerland)*, 9(10). doi: 10.3390/su9101765.

Salleh, N., Mendes, E. and Grundy, J. C. (2011) 'Empirical Studies of Pair Programming for CS / SE Teaching in Higher Education : A Systematic Literature Review', 37(4), pp. 509–525.

Sebastiani, F. (2002) 'Machine learning in automated text categorization', *ACM Computing Surveys*. ACM, 34(1), pp. 1–47. doi: 10.1145/505282.505283.

Shan, J. *et al.* (2019) 'Co-learning binary classifiers for LP-based multi-label classification', *Cognitive Systems Research*. Elsevier, 55, pp. 146–152. doi: 10.1016/J.COGSYS.2019.01.009.

Shen, S. (2014) 'Intention to revisit traditional folk events: A case study of Qinhuai Lantern Festival, China', *International Journal of Tourism Research*. John Wiley and Sons Ltd, 16(5), pp. 513–520. doi: 10.1002/jtr.1949.

Sheridan, M. F., Lee, S. and Roehl, W. (2013) 'Effects of hotel discounting practice on visitors perceptions and visitation intentions', *Tourism Economics*, 19(3), pp. 599–611. doi: 10.5367/te.2013.0220.

Sidorova, A. *et al.* (2008) 'Uncovering the Intellectual Core of the Information Systems Discipline', *MIS Quarterly*, 32(3), pp. 467–482. Available at: http://aisel.aisnet.org/misq/vol32/iss3/3 (Accessed: 6 May 2017).

Siering, M., Deokar, A. V. and Janze, C. (2018) 'Disentangling consumer recommendations: Explaining and predicting airline recommendations based on online reviews', *Decision Support Systems*. North-Holland, 107, pp. 52–63. doi: 10.1016/J.DSS.2018.01.002.

Simeon, M. I. and Martone, A. (2016) 'Travel Communities, Innovative Tools to Support Decisions for Local Tourism Development', *Procedia - Social and Behavioral Sciences*, 223, pp. 681–686. doi: 10.1016/j.sbspro.2016.05.237.

Simpson, P., Cruz-Milán, O. and Gressel, J. (2014) 'Perceived Crime and Violence: Effects on Winter Migrants', *Journal of Travel Research*. SAGE Publications Ltd, 53(5), pp. 597–609. doi: 10.1177/0047287513513162.

Simpson, P. M., Siguaw, J. A. and Sheng, X. (2016) 'Tourists' Life Satisfaction at Home and Away: A Tale of Two Cities ', *Journal of Travel Research* . SAGE Publications Ltd, 55(2), pp. 161–175. doi: 10.1177/0047287514541004.

Sivic, J. and Zisserman, A. (2003) 'Video Google: A Text Retrieval Approach to Object Matching in Videos', in *Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003)*, pp. 1–8. doi: http://dx.doi.org/10.1177/1933719112442493.

Snoek, C. G. M. *et al.* (2006) 'The Challenge Problem for Automated Detection of 101 Semantic Concepts in Multimedia', in *Proceedings of the 14th ACM International Conference on Multimedia*. New York, NY, USA: ACM (MM '06), pp. 421–430. doi: 10.1145/1180639.1180727.

Sobol-Shikler, T. and Robinson, P. (2010) 'Classification of Complex Information: Inference of Co-Occurring Affective States from Their Expressions in Speech', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(7), pp. 1284–1297. doi: 10.1109/TPAMI.2009.107.

Song, H. and Kim, H. (2018) 'Value-based profiles of visitors to a world heritage site: The case of Suwon Hwaseong fortress (in South Korea)', *Sustainability*, 11(1), pp. 132–151. doi: 10.3390/su11010132.

Sotiriadis, M. (2017) 'Experiential dimensions and their influence on behavioral intentions within the context of nature-based tourism', *Tourism and Hospitality Management*, 23(1), pp. 35–50. doi: 10.20867/thm.23.1.7.

Sparks, B. A., So, K. K. F. and Bradley, G. L. (2016) 'Responding to negative online reviews: The effects of hotel responses on customer inferences of trust and concern', *Tourism Management*. Pergamon, 53, pp. 74–85. doi: 10.1016/J.TOURMAN.2015.09.011.

Sriram, B. *et al.* (2010) 'Short text classification in twitter to improve information filtering', in *Proceeding of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2010*. New York, New York, USA: ACM Press, p. 841. doi: 10.1145/1835449.1835643.

Su, L. and Huang, Y. (2019) 'How does perceived destination social responsibility impact revisit intentions: The mediating roles of destination preference and relationship quality', *Sustainability*, 11, pp. 133–155. doi: 10.3390/su11010133.

Sunway Lagoon Malaysia (2019a) *About Us - Sunway Lagoon Theme Park*. Available at: https://sunwaylagoon.com/about-us/ (Accessed: 10 February 2019).

Sunway Lagoon Malaysia (2019b) *Awards and Accolades - Sunway Lagoon Theme Park*. Available at: https://sunwaylagoon.com/awards-accolades/#1527481474357-3cdce297-2d2a (Accessed: 10 February 2019).

Tang, L. (Rebecca) and Jang, S. (Shawn) (2012) 'Investigating the routes of communication on destination websites', *Journal of Travel Research*. SAGE Publications Ltd, 51(1), pp. 94–108. doi: 10.1177/0047287510394196.

Thammaboosadee, S., Muangon, A. and Haruechaiyasak, C. (2014) 'A Lexiconizing Framework of Feature-based Opinion Mining in Tourism Industry', in *2014 4th International Conference on Digital Information and Communication Technology and Its Applications, DICTAP 2014*. doi: 10.1109/DICTAP.2014.6821677.

The LEGO Group (2014) *About LEGOLAND® Malaysia Resort | LEGOLAND® Malaysia Resort.* Available at: https://www.legoland.com.my/en/explore-legoland/about-legoland-malaysia-resort/ (Accessed: 10 February 2019).

The LEGO Group (2016) *LEGOLAND® Malaysia Resort Ranked Number 1 In Tripadvisor Award.* Available at: https://www.legoland.com.my/en/about-us/press/ranked-number-1-in-tripadvisor-award/ (Accessed: 10 February 2019).

The LEGO Group (2017) *LEGOLAND® Malaysia Resort Launches The World's First LEGO® Virtual Reality Roller Coaster.*

Thuraisingham, B. M. (1999) *Data mining : technologies, techniques, tools, and trends.* CRC Press. Available at: https://www.crcpress.com/Data-Mining-Technologies-Techniques-Tools-and-Trends/Thuraisingham/p/book/9780849318153 (Accessed: 4 May 2017).

Tirunillai, S. and Tellis, G. (2014) 'Extracting Dimensions of Consumer Satisfaction with Quality From Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation', *SSRN Electronic Journal*, pp. 463–479. Available at: http://ssrn.com/abstract=2408855Electroniccopyavailableat:http://ssrn.com/abstract=2408855 (Accessed: 11 March 2019).

Tourism Malaysia (2020) *Malaysia Tourism Performance January - December 2019.* Available at: http://mytourismdata.tourism.gov.my/wp-content/uploads/2020/03/Msia-Tourism-Performance-Fast-Facts-Q4-2019-final1.pdf.

TripAdvisor LLC (2013) *24 Insights to shape your TripAdvisor strategy.* Available at: https://mk0tainsights9mcv7wv.kinstacdn.com/wp-content/uploads/2018/01/24-Insights-to-Shape-TripAdvisor-Strategy-Infographic-en-us.pdf (Accessed: 18 March 2019).

TripAdvisor LLC (2018a) *Jumpshot for TripAdvisor Sites.* Available at: https://tripadvisor.mediaroom.com/my-about-us (Accessed: 7 February 2019).

TripAdvisor LLC (2018b) *TripAdvisor Log Files.*

Tsai, C.-T. (Simon) (2016) 'Memorable Tourist Experiences and Place Attachment When Consuming Local Food', *International Journal of Tourism Research.* John Wiley and Sons Ltd. doi: 10.1002/jtr.2070.

Tsoumakas, G., Katakis, I. and Vlahavas, I. (2010) 'Mining Multi-label Data', in Maimon, O. and Rokach, L. (eds) *Data Mining and Knowledge Discovery Handbook.* Springer, Boston, MA, pp. 667–685. doi: 10.1007/978-0-387-09823-4_34.

Tsoumakas, G., Katakis, I. and Vlahavas, I. (2011) 'Random k-Labelsets for Multilabel Classification', *IEEE Transactions on Knowledge and Data Engineering*, 23(7), pp. 1079–1089. doi: 10.1109/TKDE.2010.164.

Turnbull, D. *et al.* (2008) 'Semantic Annotation and Retrieval of Music and Sound Effects', *IEEE Transactions on Audio, Speech, and Language Processing*, 16(2), pp. 467–476. doi: 10.1109/TASL.2007.913750.

Tussyadiah, I. P., Park, S. and Fesenmaier, D. R. (2011) 'Assessing the effectiveness of consumer narratives for destination marketing', *Journal of Hospitality & Tourism Research*, 35(1), pp. 64–78. doi: 10.1177/1096348010384594.

Twinword Inc. (2019) *Sentiment Analysis API, Twinword Inc.* Available at: https://www.twinword.com/api/sentiment-analysis.php (Accessed: 9 May 2019).

Ukpabi, D. C. and Karjaluoto, H. (2017) 'Consumers' acceptance of information and communications technology in tourism: A review', *Telematics and Informatics*. Pergamon, 34(5), pp. 618–644. doi: 10.1016/J.TELE.2016.12.002.

Umar, H., Munir, A. R. and Syam, N. (2019) 'Intention to return after disaster', *IOP Conference Series: Earth and Environmental Science*, 235(1), pp. 1–8. doi: 10.1088/1755-1315/235/1/012100.

van der Veen, R. and Song, H. (2014) 'Impact of the Perceived Image of Celebrity Endorsers on Tourists' Intentions to Visit', *Journal of Travel Research*. SAGE Publications Ltd, 53(2), pp. 211–224. doi: 10.1177/0047287513496473.

Viegas, F. *et al.* (2018) 'Exploiting efficient and effective lazy Semi-Bayesian strategies for text classification', *Neurocomputing*. Elsevier B.V., 307, pp. 153–171. doi: 10.1016/j.neucom.2018.04.033.

Vo Thanh, T. and Kirova, V. (2018) 'Wine tourism experience: A netnography study', *Journal of Business Research*, 83, pp. 30–37. doi: 10.1016/j.jbusres.2017.10.008.

Wang, H., Lu, Y. and Zhai, C. (2010) 'Latent aspect rating analysis on review text data: A rating regression approach', *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 783–792. doi: 10.1145/1835804.1835903.

Weaver, D. B. and Lawton, L. J. (2011) 'Visitor Loyalty at a Private South Carolina Protected Area', *Journal of Travel Research*. SAGE Publications Ltd, 50(3), pp. 335–346. doi: 10.1177/0047287510362920.

Wieczorkowska, A., Synak, P. and Raś, Z. W. (2006) 'Multi-Label Classification of

Emotions in Music', in *Intelligent Information Processing and Web Mining*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 307–315. doi: 10.1007/3-540-33521-8_30.

Wild, F. *et al.* (2005) 'Parameters Driving Effectiveness of Automated Essay Scoring', *Information Systems Journal*, 80(18), pp. 485–494. Available at: http://scholar.google.de/scholar?hl=de&lr=&q=factors+influencing+the+effectiveness+of+lsa&btnG=Suche&lr=#0.

Wolpert, D. H. and Macready, W. G. (1997) 'No free lunch theorems for optimization', *IEEE Transactions on Evolutionary Computation*, 1(1), pp. 67–82. doi: 10.1109/4235.585893.

Wong, C. U. I. and Qi, S. (2017) 'Tracking the evolution of a destination's image by text-mining online reviews - the case of Macau', *Tourism Management Perspectives*. Elsevier, 23, pp. 19–29. doi: 10.1016/j.tmp.2017.03.009.

Wong, J.-Y. and Lai, T.-C. (2015) 'Celebrity attachment and behavioral intentions: The mediating role of place attachment', *International Journal of Tourism Research*. John Wiley and Sons Ltd, 17(2), pp. 161–170. doi: 10.1002/jtr.1974.

Wong, J. Y., Lee, S. J. and Lee, W. H. (2015) '"Does it Really Affect Me?" Tourism Destination Narratives, Destination Image, and the Intention to Visit: Examining the Moderating Effect of Narrative Transportation', *International Journal of Tourism Research*. doi: 10.1002/jtr.2063.

Wongso, R. *et al.* (2017) 'News Article Text Classification in Indonesian Language', *Procedia Computer Science*. Elsevier, 116, pp. 137–143. doi: 10.1016/J.PROCS.2017.10.039.

World Travel and Tourism Council (2016) *Travel and Tourism: World Economic Impact 2016*. United Kingdom.

Wu, M. Y., Wall, G. and Pearce, P. L. (2014) 'Shopping experiences: International tourists in Beijing's Silk Market', *Tourism Management*. Elsevier Ltd, 41, pp. 96–106. doi: 10.1016/j.tourman.2013.09.010.

Wu, Y. and Ester, M. (2015) 'FLAME: A Probabilistic Model Combining Aspect Based Opinion Mining and Collaborative Filtering', in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. New York, NY, USA: Association for Computing Machinery (WSDM '15), pp. 199–208. doi: 10.1145/2684822.2685291.

Xiang, Z. *et al.* (2015) 'What can big data and text analytics tell us about hotel guest experience and satisfaction?', *International Journal of Hospitality Management*. Pergamon, 44, pp. 120–130. doi: 10.1016/J.IJHM.2014.10.013.

Xiang, Z. *et al.* (2017) 'A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism', *Tourism Management*. Pergamon, 58, pp. 51–65. doi: 10.1016/J.TOURMAN.2016.10.001.

Xianghua, F. *et al.* (2013) 'Multi-Aspect Sentiment Analysis for Chinese Online Social Reviews Based on Topic Modeling and HowNet Lexicon', *Knowledge-Based Systems*. NLD: Elsevier Science Publishers B. V., 37, pp. 186–195. doi: 10.1016/j.knosys.2012.08.003.

Xiao, S., Wei, C.-P. and Dong, M. (2016) 'Crowd intelligence: Analyzing online product reviews for preference measurement', *Information & Management*. North-Holland, 53(2), pp. 169–182. doi: 10.1016/J.IM.2015.09.010.

Xu, F. *et al.* (2018) 'Is Southern Xinjiang really unsafe?', *Sustainability*, 10(12), pp. 4639–4660. doi: 10.3390/su10124639.

Xu, X. *et al.* (2013) 'Aspect-level opinion mining of online customer reviews', *China Communications*, 10(3), pp. 25–41. doi: 10.1109/CC.2013.6488828.

Xu, X. and Li, Y. (2016) 'The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach', *International Journal of Hospitality Management*. Pergamon, 55, pp. 57–69. doi: 10.1016/j.ijhm.2016.03.003.

Xu, Z. *et al.* (2019) 'Exploring the role of emotion in the relationship between museum image and tourists' behavioral intention: The case of three museums in Xi'an', *Sustainability*, 11(3), pp. 559–579. doi: 10.3390/su11030559.

Yeh, T.-M., Chen, S.-H. and Chen, T.-F. (2019) 'The relationships among experiential marketing, service innovation, and customer satisfaction-A case study of tourism factories in Taiwan', *Sustainability*, 11(4), pp. 1041–1053. doi: 10.3390/su11041041.

Yen, C.-H. and Teng, H.-Y. (2015) 'Celebrity Involvement, Perceived Value, and Behavioral Intentions in Popular Media-Induced Tourism', *Journal of Hospitality & Tourism Research*. SAGE Publications Inc., 39(2), pp. 225–244. doi: 10.1177/1096348012471382.

Yoo, K. and Gretzel, U. (2018) 'What motivates consumers to write online travel reviews?', *Information Technology & Tourism*, 10(4), pp. 283–295. Available at: https://www.ingentaconnect.com/content/cog/itt/2008/00000010/00000004/art00003

(Accessed: 17 March 2019).

Zhang, J.-J., Fang, M. and Li, X. (2017) 'Clustered intrinsic label correlations for multi-label classification', *Expert Systems with Applications*. Pergamon, 81, pp. 134–146. doi: 10.1016/J.ESWA.2017.03.054.

Zhang, L. and Liu, B. (2014) 'Aspect and Entity Extraction for Opinion Mining', in, pp. 1–40. doi: 10.1007/978-3-642-40837-3_1.

Zhang, M. L. and Zhou, Z. H. (2007) 'ML-KNN: A lazy learning approach to multi-label learning', *Pattern Recognition*, 40(7), pp. 2038–2048. doi: 10.1016/j.patcog.2006.12.019.

Zhang, Y. and Wang, L. (2019) 'Influence of sustainable development by tourists' place emotion: Analysis of the multiply mediating effect of attitude', *Sustainability*, 11(5), pp. 1384–1399. doi: 10.3390/su11051384.

Zhang, Zili *et al.* (2019) 'Booking now or later: Do online peer reviews matter?', *International Journal of Hospitality Management*. Pergamon, 77, pp. 147–158. doi: 10.1016/J.IJHM.2018.06.024.

Zhang, Ziqiong *et al.* (2011) 'Sentiment classification of Internet restaurant reviews written in Cantonese', *Expert Systems with Applications*. Pergamon, 38(6), pp. 7674–7682. doi: 10.1016/J.ESWA.2010.12.147.

Zhao, Y., Xu, X. and Wang, M. (2019) 'Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews', *International Journal of Hospitality Management*. Pergamon, 76, pp. 111–121. doi: 10.1016/J.IJHM.2018.03.017.

Zhu, J. *et al.* (2011) 'Aspect-Based Opinion Polling from Customer Reviews', *IEEE Transactions on Affective Computing*, 2(1), pp. 37–49. doi: 10.1109/T-AFFC.2011.2.

Zhu, J. and Liu, W. (2020) 'A tale of two databases: the use of Web of Science and Scopus in academic papers', *Scientometrics*, 123(1), pp. 321–335. doi: 10.1007/s11192-020-03387-8.

Zhu, R. (2018) 'Analysis on the Impact of haze on Beijing Residents' Traveling Intention and Decision-Making', *IOP Conference Series: Materials Science and Engineering*, 394(5), pp. 1–6. doi: 10.1088/1757-899X/394/5/052054.

Zorn, P. *et al.* (1999) 'Finding Needles in the Haystack: Mining Meets the Web'. United States. Available at: https://www.questia.com/magazine/1P3-44243170/mining-meets-the-web.

LIST OF PUBLICATIONS

Indexed Conference Proceedings

1. **Selamat, N. S.**, Sim, A. T. H., & Hee, J. M. (2017). Transforming Tourism Destinations' Marketing Strategies by Understanding Tourists' Satisfaction. In *PACIS 2017 Proceedings*, 56. https://aisel.aisnet.org/pacis2017/56. **(Indexed by SCOPUS)**