Purposed Framework for Understating Tourism Behavior for Discovering Business Opportunities

*Nitin Sharma
**Dr. K.P. Yadav

Abstract

The aim of this paper is to present and identify the main motivations for users to book personal accommodation found online through Airbnb, a sharing platform that links suppliers of living space with those requiring short-term accommodation. Since its first appearance in 2008, Airbnb has had a fast growth that raises questions about its current and future impacts on the traditional accommodation sector.

This research was guided primarily by concepts associated with disruptive innovation and the diffusion of innovations. Disruptive innovation describes a process through which new products, which underperform in comparison with existing products' key attributes, encroach upon a market by introducing an alternative package of benefits generally centred around being cheaper, simpler, smaller and/or more convenient. This framework provides a natural lens through which to view the rise of Airbnb, as traditional accommodations seemingly outperform Airbnb in many key areas, but Airbnb tends to be cheaper and offer some additional alternative benefits. The diffusion of innovations is a broad field examining topics related to how and why innovations spread and are adopted. Most relevant for this study, the diffusion of innovations literature has shown that an innovation's "relative advantage" and its "compatibility" with adopters directly influence its adoption, and communication channels play a fundamental role in the spread of innovations.

Keywords : Big Data, Tourism Management, Tourist Behavior, Literature Review

1. Introduction

There is no doubt that big data are now rapidly expanding in all science and engineering domains. While the potential of these massive data is undoubtedly significant, fully making sense of them requires new ways of thinking and novel learning techniques to address the various challenges. First, we review the machine learning techniques and highlight some promising learning methods in recent studies, such as representation learning, deep learning, distributed and parallel learning, transfer learning, active learning, and kernel-based learning. Next, we focus on the analysis and discussions about the challenges and possible solutions of machine learning for big data. Following that, we investigate the close connections of machine learning with signal processing techniques for big data processing. Eventually, we define several open issues and studies trends.

It is obvious that we are living in a data deluge era, evidenced by the phenomenon that enormous amount of data have been being continually generated at unprecedented and ever increasing scales. Large-scale data sets are collected and studied in numerous domains, from engineering sciences to social networks, commerce, bimolecular research, and security [1]. Particularly, digital data, generated from a variety of digital devices, are growing at astonishing rates. According to [2], in 2011,

digital information has grown nine times in volume in just 5 years and its amount in the world will reach 35 trillion gigabytes by 2020 [3]. Therefore, the term "Big Data" was coined to capture the profound meaning of this data explosion trend.

To clarify what the big data refers to, several good surveys have been presented recently and each of them views the big data from different perspectives, including challenges and opportunities [4], background and research status [5], and analytics platforms [6]. Among these surveys, a comprehensive overview of the big data from three different angles, i.e., innovation, competition, and productivity, was presented by the McKinsey Global Institute (MGI) [7]. Besides describing the fundamental techniques and technologies of big data, a number of more recent studies have investigated big data under particular context. For example, [8, 9] gave a brief review of the features of big data from Internet of Things (IoT). Some authors also analyzed the new characteristics of big data in wireless networks.

Over the past decade, machine learning techniques have been widely adopted in a number of massive and complex data-intensive fields such as medicine, astronomy, biology, and so on, for these techniques provide possible solutions to mine the information hidden in the data. Nevertheless, as the time for big data is coming, the collection of data sets is so large and complex that it is difficult to deal with using traditional learning methods since the established process of learning from conventional datasets was not designed to and will not work well with high volumes of data. For instance, most traditional machine learning algorithms are designed for data that would be completely loaded into memory which does not hold any more in the context of big data. Therefore, although learning from these numerous data is expected to bring significant science and engineering advances along with improvements in quality of our life, it brings tremendous challenges at the same time.

2. Theoretical Background

Social media sites have become a very important part of our lives nowadays and somewhere it affects our ability to decide. Big data generated on the social media sites has fabricated countless opportunities for bringing more sagacity to decision makers. Few studies on big data analytics have demonstrated its support for tactical decision making. Beside, a formal method for analyzing the big data generated on social media for decision support is yet to be developed, especially in tourism sector. The aim of this study is to design and valuate a 'big data analytics' method to support strategic decision making tourism destination management by using design science research approach. Everyday many types of data are constantly growing within social media (Such as Twitter, Facebook, Instagram, Line & we Chat etc.) because of their large user base, and frequent uploads of digital photos and videos.

Big data has allowed companies to track study shopping patterns, recommendations, and the purchasing behaviour that are known to influence sales and it has provided an exceptional insight into customers decision making process. Big data has given more power to the agencies and the merchants involved in tourism that they can find ingenious ways to use a variety of data resources to connect with a particular visitor or the potential visitor at every stage of a trip and they can use these big data sources to enhance an timely understand the fastest growing visitors demographics.

We are confident that both consumers and tourism product providers will be benefited by using big

data. It is clear that big data can provide better, targeted, and profitable products to consumers (Pries & Dunnigan, 2015). Personalized marketing and targeted product designs are very powerful opportunities for both consumers and tourism product providers. For instance, Big data gets the information about the interest of the consumer from photos posted on social networks by the consumer or what he searches for. This provide a very useful information to the product providers about the type of product the user is looking for. It is up to the researchers how efficiently they are going to examine the digital footprint left by activities on internet and make the use of these data or whether these data will create a new research epitome that tangles new methodologies and will develop our theoretical understanding of tourism. Till date the online data source is mainly used in applied research, and by means of that advantage was taken of the large and often free-of-charge volumes of data that brings the understanding of the activities of the tourism /travel industry and its patrons.

Business Intelligence and Big data are just about to unfold their full potential for the tourism domain. Big data can be a very useful source of information for the tourism sector seems to be still ignored by the researchers. Same situation seems to happen with Business intelligence also. Big data is to be considered an incredible opportunity it can provide answers about people's behaviours and views (Rodolfo Baggio 2016). In the last years, we need to recognised customer-focused approach in order to increase the tourism. Big data have also started to be a source for Business Intelligence Activities. The tourism scholars have well understood the capability of BD to provide insights that are useful or should be used for enriching the BI practices for destinations and operators. This note presents the results of an analysis of the recent literature on BD and BI and the application of the related technique to the fields of travel and tourism.

Smart tourism is become more and more important for the local government to raise the economy. Nowadays, many researchers draw much attention to the smart tourism. By the use of new technologies we can extract useful information from social media contents and then it help people to make more efficient and economic travel decisions. We use urban tourism check algorithm in which can find out the visitors who come from other country using social media check-in-data. We choose the Shanghai as a research target to analyze the data and the tourism place are Renmin Square, Yu Garden, The Bund and Chenhuang Temple in 2016 from Sino Weibo. We divide the tourists in two parts, one is local tourists and the other is outside tourists (Kai Yang, Wanggen Wan 2016).In order to make sure the accuracy of algorithm, we use the PLSA algorithm to analyse the Weibo text content. Through this analysis of outsider tourists we can provide a reference to the government which can use the tourism to rise in urban economy.

3. Problem Statement & Methodology

The aim of this research is to present and identify the main motivations for users to book personal accommodation found online through Airbnb, a sharing platform that links suppliers of living space with those requiring short-term accommodation. Since its first appearance in 2008, Airbnb has had a fast growth that raises questions about its current and future impacts on the traditional accommodation sector.

Trying to understand better what the motives of tourists when choosing Airbnb are is important for Airbnb owners and other stakeholders, such as tourism firms or destination organizations, because from them they can learn more about what customers need and want. Moreover, figuring out buyers'

preferences regarding Airbnb attributes helps to realise the advantages and disadvantages it has over other types of accommodation and vice-versa, the advantages and disadvantages other types of accommodation have over Airbnb. The substitution effect that Airbnb has on other types of accommodations could lead to the answer of the question regarding the effect it has over the traditional accommodation sector. It can also help the traditional accommodation sector to overcome Airbnb's threat. The study helps to figure out the new accommodation trend and the concept of sharing economy.

A quantitative research method was used in this research to gather multidimensional data. The data set is getting from the kaggle.com that was completed by tourists who already stayed with Airbnb. After the data was preprocessed and analyses, the results are presented and discussed accordingly. The research shows that some aspects have an essential importance in the decision-making process. The primary results presented here show that both the price and the authenticity of the experience are the most significant attributes in the choice of Airbnb. This analysis is supported by the concepts of sharing economy, collaborative consumption, peer-to-peer accommodation and consumer decision-making, which are underlined in this study.

Accuracy, Precision and Recall

Consider a classification task in which a machine learning system observes tumors and has to predict whether these tumors are benign or malignant. Accuracy, or the fraction of instances that were classified correctly, is an obvious measure of the program's performance. While accuracy does measure the program's performance, it does not make distinction between malignant tumors that were classified as being benign, and benign tumors that were classified as being malignant. In some applications, the costs incurred on all types of errors may be the same. In this problem, however, failing to identify malignant tumors is a more serious error than classifying benign tumors as being malignant by mistake.

We can measure each of the possible prediction outcomes to create different snapshots of the classifier's performance. When the system correctly classifies a tumor as being malignant, the prediction is called a true positive. When the system incorrectly classifies a benign tumor as being malignant, the prediction is a false positive. Similarly, a false negative is an incorrect prediction that the tumor is benign, and a true negative is a correct prediction that a tumor is benign. These four outcomes can be used to calculate several common measures of classification performance, like accuracy, precision, recall and so on.

Accuracy is calculated with the following formula –

 ACC = (TP + TN)/(TP + TN + FP + FN)

Where,

TP is the number of true positives

TN is the number of true negatives

FP is the number of false positives

FN is the number of false negatives.

Purposed Framework for Understating Tourism Behavior for Discovering Business Opportunities

Nitin Sharma & Dr. K.P. Yadav

Precision - Precision is the fraction of the tumors that were predicted to be malignant that are actually malignant. Precision is calculated with the following formula –

PREC = TP/(TP + FP)

Recall – Recall is the fraction of malignant tumors that the system identified. Recall is calculated with the following formula –

R = TP/(TP + FN)

In this example, precision measures the fraction of tumors that were predicted to be malignant that are actually malignant. Recall measures the fraction of truly malignant tumors that were detected. The precision and recall measures could reveal that a classifier with impressive accuracy actually fails to detect most of the malignant tumors. If most tumors are benign, even a classifier that never predicts malignancy could have high accuracy. A different classifier with lower accuracy and higher recall might be better suited to the task, since it will detect more of the malignant tumors. Many other performance measures for classification can also be used.

4. Implementation

Data visualization is a general term that describes any effort to help people understand the significance of data by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization software.

Today's data visualization tools go beyond the standard charts and graphs used in Microsoft Excel spreadsheets, displaying data in more sophisticated ways such as infographics, dials and gauges, geographic maps, sparklines, heat maps, and detailed bar, pie and fever charts. The images may include interactive capabilities, enabling users to manipulate them or drill into the data for querying and analysis. Indicators designed to alert users when data has been updated or predefined conditions occur can also be included.
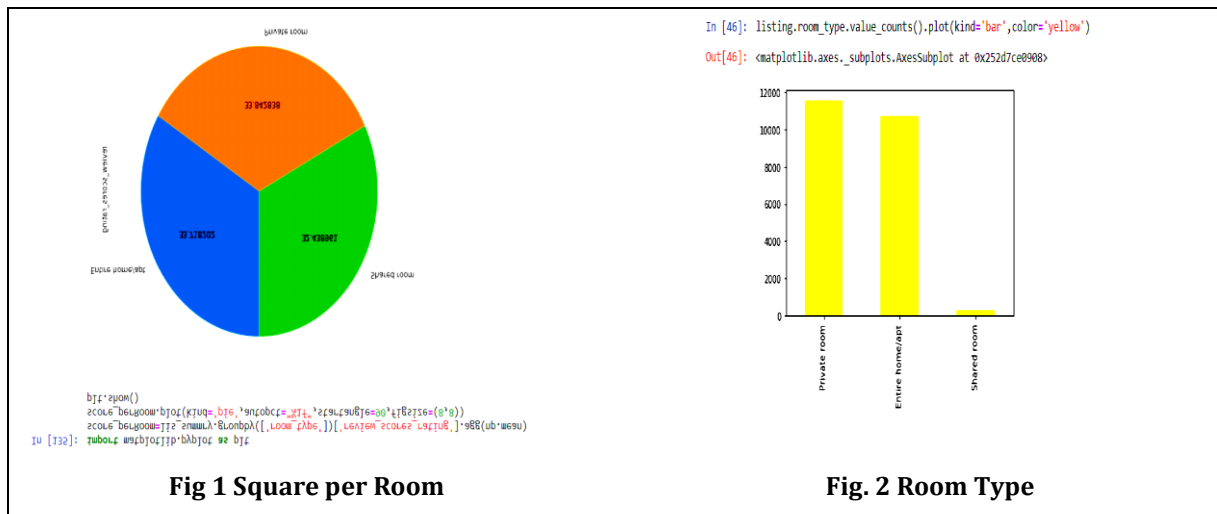
Fig 1 Square per Room   **Fig. 2 Room Type**

Purposed Framework for Understating Tourism Behavior for Discovering Business Opportunities

Nitin Sharma & Dr. K.P. Yadav

Fig 3 Neighbourhood Group   Fig. 4 Scatter Plot
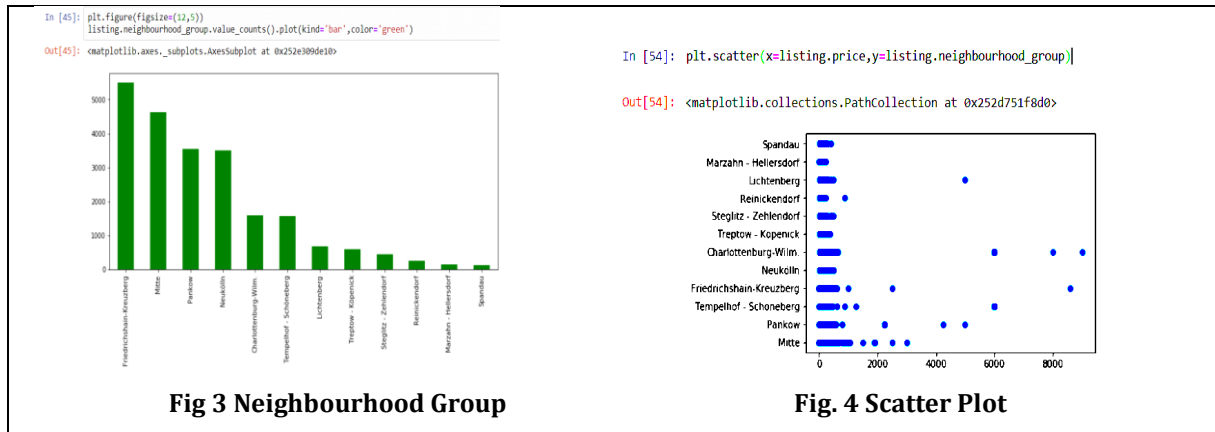
5. Conclusion

When the general idea for this research project was born several years ago, most discussion of it had to be preceded with an introductory description of Airbnb. Today, Airbnb is arguably the most talked-about subject in the entire tourism sector. In fact, a large portion of the research studies and media stories that have been referenced in this thesis were only published within the last year or so, long after this project started. Such mounting interest highlights the important role this study can play in providing new insights on the consumer side of Airbnb.

Airbnb has already shaken up the tourism accommodation industry, with little indication that its growth will slow significantly anytime soon. The speed with which it has impacted the tourism sector is both quite rare and remarkable. Nonetheless, there continues to be little understanding of why so many tourists are choosing this innovative service, and what the implications of its emergence will be for the traditional accommodation industry. This study sheds some important light on these questions by showing the strongest motivations tend to involve cost and other practical considerations, whereas the experiential motivations are mostly secondary. This finding raises questions regarding Airbnb's marketing and regarding the common belief among hoteliers that Airbnb is not competing with hotels because it appeals to a separate market. This latter notion is further contradicted by the finding that the majority of respondents had used Airbnb as a hotel substitute. The study also importantly identified and profiled five distinct motivation-based segments of Airbnb users, and this exercise led to myriad practical implications that were discussed. The concept of disruptive innovation was used to compare Airbnb with hotels along several traditional accommodation attributes, which highlighted some of Airbnb's strengths and weaknesses, while also offering conceptual implications regarding the notion of disruptive innovation. Finally, the various findings from this study were drawn upon to anticipate future changes within the tourism accommodation industry. Although it is a future filled with question marks, it is hoped that this study can provide all of the relevant stakeholders with a better understanding of the consumer motivations that serve as the foundation for these changes.

Airbnb is having such a significant impact on the tourism accommodation industry that it is important to discover the incentives from the consumer side for using this innovative service. There

Purposed Framework for Understating Tourism Behavior for Discovering Business Opportunities

Nitin Sharma & Dr. K.P. Yadav

still exists the question of why people tend to choose Airbnb rather than a traditional accommodation and what will be the future impact of it for the traditional accommodation industry

*Research Scholar
**Professor (Research Supervisor)
Computer Science & Engineering
Sunrise University, Alwar (Raj.)

References

1. Yanxia Rong, S. K. Lee, and H. A. Choi, "Detecting stations cheating on backoff rules in 802.11 networks using sequential analysis," In Proc. IEEE INFOCOM '06, 15-jun-2014.

2. Cenker Demir and Cristina Comaniciu, "An Auction based AODV Protocol for Mobile Ad Hoc Networks with Selfish Nodes", IEEE International Conference on 24-28- June- 2011.

3. Zahara Safaei , Mohammad Hossein Anisi an Fatemeh Torgheh, " A Reputation-Based Mechanisms to enforce Cooperation in MANETs", Software, Telecommunications and Computer Networks, (SoftCOM 2008) 16th International Conference on 25-27 Sept.- 2012.

4. C.-K. Toh, D. Kim, S. Oh, and H. Yoo, "The controversy of selfish nodes in ad hoc networks," in Proceedings of the Twelveth international conference on Advanced communication technology, pp.20-26, 20-June- 2010.

5. F. Kargl, A. Klenk, S. Schlot, and M. Webber, "Advanced detection of selfish or malicious nodes in ad hoc networks," springer - verlag berlin Heidelberg, vol. 05, pp. 152-165, 15-Jan. - 2009.

6. G. Corradi, J. Janssen, and R. Manca, "Numerical Treatment of Homogeneous Semi-Markov Processes in Transient Case – a Straightforward Approach," Methodology and Computing in Applied Probability, vol. 6, pp. 233–246, Feb. -2014.

7. J.-K. Lee and J. C. Hou, "Modeling Steady-state and Transient Behaviors of User Mobility: Formulation, Analysis, and Application," in Proc. of ACM MobiHoc 06, vol.08, pp. 85–96, May- 2012.

8. S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating Routing Misbehavior in Mobile Ad hoc Networks," in Proc. of ACM MobiCom', pp. 255–265, 2010.

9. D. Chen, S. Garg, and K. S. Trivedi, "Network Survivability Performance Evaluation: A Quantitative Approach with Applications in Wireless Ad-hoc Networks," in Proc. of the ACM International Workshop on Modeling, Analysis, and Simulation of Wireless and Mobile Systems, pp. 61–68, Sep.- 2002.

10. Bahmani B, Moseley B, Vattani A, Kumar R, Vassilvitskii S. Scalable k-means++. PVLDB 2012, 5:622–633.

11. Jin C, Liu R, Chen Z, Hendrix W, Agrawal A, Choudhary A. A scalable hierarchical clustering algorithm using Spark. In: Proceedings of IEEE First Inter-national Conference on Big Data Computing Service and Applications (Big Data Service), San Francisco Bay, CA, USA, 2015, 418–426. doi: 10.1109/BigDataService.2015.67.

Purposed Framework for Understating Tourism Behavior for Discovering Business Opportunities

Nitin Sharma & Dr. K.P. Yadav