



Analysis of Satisfaction Factors by language groups on tourist destinations using Text Mining

DaHee Kim^{1,2}, KangWoo Lee², JiWon Lim², Soon-Goo Hong³

¹Department of Computer Engineering, Dong-A Univ. 37, Nakdong-daero 550beon-gil, Saha-gu, Busan, Republic of Korea.

²Smart Governance Research Center, Dong-a Univ. 225 Gudeok-ro, Seo-gu, Busan, Republic of Korea.

³Department of Management Information System, Dong-a Univ. 225 Gudeok-ro, Seo-gu, Busan, Republic of Korea.

Abstract

Social media influences the decision about tourist destinations. On the other hand, the tourism industry can provide an effective tourism service by analyzing social shares and identifying which aspect of a tourist destination is satisfied or dissatisfied. By analyzing the recommendations of tourists from different language groups, this study intends to understand their satisfaction/dissatisfaction with tourist attractions. The reviews on tourist attractions in Busan, South Korea, written in Korean, English, Japanese, and Chinese, were collected from TripAdvisor. The reviews were split into sentences, and then sentences containing recommendation-related words were extracted. Using BERTopic and word co-occurrence analysis, we investigated what topics occur and what words are associated with the recommendation-related words in the different language groups. Characteristically, the topics associated with 'food' and 'price' were found in Japanese reviews, 'photo' and 'shopping' in Chinese reviews, and topics associated with '(Buddhist) temple' and 'spa' in English. For Korean reviews, recommendations mainly concern which 'season' and 'time' are the best for traveling in Busan. Similar patterns were also found in word co-occurrence analysis. It is identified that things recommended are different depending on different language groups. It seems that tourists in Korean, Japanese, and Chinese groups share similar cultural backgrounds, such as Buddhism, and their recommendations mainly concern what to eat and buy. In contrast, tourists in the English language group seem fascinated with unique Korean cultures - temple-stay and huge-sized public spas. Our study suggests that customized tourism strategies would be necessary to enhance tourists' satisfaction from different nations.

Keywords: topic modeling; word co-occurrence; BERTopic; text mining; tourist destination satisfaction