Summer 7-5-2020

Research on Online Word-of-mouth Sentiment Analysis and Attribute Extraction Based on Deep Learning

Jiamei Yang
School of Economics and Management, Nanjing University of Science and Technology, Nanjing, 210094, China

Li Li
School of Economics and Management, Nanjing University of Science and Technology, Nanjing, 210094, China; Jiangsu Collaborative Innovation Center of Social Safety Science and Technology, Nanjing, 210094, China, lily691111@njust.edu.cn

Yulan Lin
School of Economics and Management, Nanjing University of Science and Technology, Nanjing, 210094, China

Research on Online Word-of-mouth Sentiment Analysis and Attribute Extraction Based on Deep Learning

Jiamei Yang[1], *Li Li*[1,2*], *Yulan Lin*[1]

[1] School of Economics and Management, Nanjing University of Science and Technology, Nanjing, 210094, China

[2] Jiangsu Collaborative Innovation Center of Social Safety Science and Technology, Nanjing, 210094, China

Abstract: Online word-of-mouth content mining is of great significance to product, service improvement and demand prediction of online marketing enterprises. However, most studies have focused on the identification of the sentiment tendency of online word-of-mouth, and lack of text content mining for online word-of-mouth, especially negative word-of-mouth. This paper introduces deep learning into online word-of-mouth sentiment tendency analysis and negative word-of-mouth word attribute feature extraction, and builds an online word-of-mouth sentiment tendency analysis and attribute extraction model based on LSTM deep learning algorithm. The model was trained and tested through online word-of-mouth data of a fashion apparel e-commerce company. The results show that the LSTM model has a good effect on sentiment analysis and negative word-of-mouth attribute feature extraction. Through comparative experiments, it is shown that the model has a better effect than the traditional machine learning methods (SVM, Naive Bayes) in the analysis of sentiment tendency.

Keywords: online word-of-mouth, LSTM, attribute extraction, sentiment analysis, deep learning

RESEARCH QUESTION

With the popularity of online shopping, information exchange based on the information interaction platform of e-commerce websites has gradually increased. These interaction information include user generated content such as user experience, perception and evaluation information of a certain product or service, forming an online reputation. The user sentiment and the descriptions of products, services and other attributes contained in online word-of-mouth are of great significance to the product, service improvement and demand prediction of online marketing companies. However, most studies focused on the identification of the sentiment tendency of online word-of-mouth[1][2][3], or only extract the subject of online word-of-mouth[4][5], but lack of text content mining for online word-of-mouth, especially negative word-of-mouth. Some studies have pointed out that in the decision-making process of consumers, the impact of negative reviews is far greater than positive reviews[6]. Word-of-mouth text information contains specific issues that users are more interested in, especially negative word-of-mouth. Moreover, the negative word-of-mouth contains the user's feedback information, which is the user's experience about the quality of the product and service, and such information is the focus of the user's purchase process. Therefore, it is necessary to dig the text information in negative word-of-mouth.

With the rapid increase in the number of online word-of-mouth in major e-commerce platforms, the manual filtering and word-of-mouth content extraction implemented in traditional management is becoming increasingly difficult to achieve. Deep learning model has strong feature learning ability and can overcome many difficulties in artificial feature extraction[7]. In particular, LSTM (long short term memory) model can effectively fit serialized data. Therefore, we can consider designing a comprehensive model that can simultaneously implement sentiment analysis and attribute extraction to fully mine online word-of-mouth content, and provide automated and intelligent methods and tools for product design, service improvement, and demand forecasting.

* Corresponding author. Email: lily691111@njust.edu.cn (Li Li)

MAIN FINDINGS

This paper divides the manual online word-of-mouth of a fashion apparel e-commerce platform into a training set and a test set. First, the natural language processing method is used to preprocess the word-of-mouth content of the training set and the test set, and the initial text is converted into a text corpus. Then use the Word2vec model to realize the word vector construction of the word segmentation. On the training set, the LSTM deep learning algorithm is trained on the online word-of-mouth sentiment analysis and negative word-of-mouth attribute feature extraction. Furthermore, the effectiveness of this model in extracting sentiment tendency features and negative word-of-mouth attribute features is verified on the test set. Finally, the effectiveness of this model compared to traditional machine learning algorithms in extracting sentiment tendency is further verified through comparative experiments. The results show that the LSTM model has a better effect on the extraction of sentiment tendency features and negative word-of-mouth attribute extraction. At the same time, compared with the SVM model and the Naive Bayes model, the LSTM model is superior to the other two models in accuracy, recall, and F value.

RESEARCH CONTRIBUTION

This paper delves into the text content of online word-of-mouth. Based on the feature that LSTM model can fit serialized data effectively, an online word-of-mouth sentiment identification and negative word-of-mouth attribute extraction model based on deep learning is constructed. Online word-of-mouth, a text that has a word order relationship, is modeled, and at the same time, sentiment analysis and attribute extraction are realized, so that online word-of-mouth content can be fully mined. This provides automated, intelligent methods and tools for product design, service improvement, and demand forecasting.

ACKNOWLEDGEMENT

This research was supported by the National Natural Science Foundation of China under Grant 71771122 and the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant SJCX19_0044.

REFERENCES

[1] Yan Z, Xing M, Zhang D, et al. (2015). EXPRS: an extended pagerank method for product feature extraction from online consumer reviews. Information & Management, 52(7): 850-858.

[2] Tripathy A, Agrawal A, Rath S K. (2016). Classification of Sentiment Reviews using N-gram Machine Learning Approach. Expert Systems with Applications, 57: 117-126.

[3] Tang D Y, Wei F R, Qin B, et al. (2016). Sentiment Embeddings with Applications to Sentiment Analysis. IEEE Transactions on Knowledge & Data Engineering, 28 (2): 496-509.

[4] Cruz I, Gelbukh A, Sidorov G. (2014). Implicit Aspect Indicator Extraction for Aspect Based Opinion Mining. International Journal of Computational Linguistics and Applications, 5(2): 135-152.

[5] Poria S, Cambria E, Gelbukh A. (2016). Aspect Extraction for Opinion Mining with a Deep Convolutional Neural Network. Knowledge-Based Systems, 108: 42-49.

[6] Jin Liyin. (2007). The Impact of Internet Word-of-mouth Information on Consumers' Purchase Decisions: An Experimental Study. Economic Management, (22): 36-42.(in Chinese)

[7] Lecun Y, Bengio Y, Hinton G. (2015). Deep learning. Nature, 521(7553): 436-444.