

Analysis on social influence and public awareness of Republic of Korea on creative education using opinion mining techniques

Ji-Hoon Seo¹, Kil-Hong Joo², Nam Hun Park³

Dept. of Computer Science Engineering Incheon University, Korea
Dept. of Computer Education, Gyeongin National University of Education, Korea
Dept. of Computer Science, Anyang University, Korea
ssez@inu.ac.kr, nmhnpark@anyang.ac.kr, khjoo@ginue.ac.kr

Abstract. Creative education has become a global issue in recent years, and on that basis, the fostering of creative individuals and software education are developing immensely. The case of South Korea is setting a foundation in their direction of policies via the creative economy by adopting the free semester policy and STEAM education in line with the trend of the global market. However, despite the country's incessant efforts for the past five years, scales and solutions through which their performance levels can be verified are not being suggested. Therefore, this thesis had conducted an opinion mining analysis with South Korean online media as target by using semi-structured data for obtaining a quantitative scale and verification, while excluding the existing methods of analysis via checklists and questionnaires. The directivity of educational policies which South Korea is not obliged but has the option to pursue has been suggested through the analyzation of social impacts and public awareness on creative education currently being enforced via the opinion mining analysis.

Keywords: Creative education, big data, pattern algorithm, Opinion mining

1 Introduction

The method currently used for analyzing the big data using the big data of the education sector is suggesting six alternatives of analyzation, and as regards social issues, agendas, such as academic achievement, school violence, school lunch, students' physical fitness and et cetera were included. Moreover, in the case of South Korea, the significance of creative education is being emphasized on the basis of the creative economy and various improvements and new solutions are being suggested. Such instances are being accelerated for the utilizations of the big data in the education sector, whereas big data analysis on the effects which the creative education is having on people's awareness or its ripple effects on the society are not being accomplished [1, 2, 3]. Therefore, this study aims to suggest the directivity of the medium and long-term development of the South Korean education policy by analyzing the scales of the ripple effect being proliferated in the society and public

awareness on policies on creative education at the present time through the analyzation of texts of the public falling under the category big data by online media by using the opinion mining method.

2 Related Work

The social changes caused by the rapid development of science and technology is leading the world to enter into a creative society (Restrict, 2007) where previously existing information and knowledge are used in creative ways to recreate high value information that is beyond the information knowledge society. In order to foster creative individuals that are to lead a society where the creation of culture, technology and information which are based on such creativity, the role of education and paradigm of education are also in need of the courses that are oriented toward openness, diversification, liberalization and individualization (Chohui Yoon, 2013).

2.1 Opinion Mining

Opinion mining method is a technology that distinguishes the preference levels of positive, negative and neutral of standard and semi-structured texts of social media and et cetera. Opinion mining is used for predicting the size of the marketing market on specific products and services and is being applied to customer response analyses, viral analyses and et cetera. Accumulation of various resources indicators, expression and vocabularies that represent the preference levels by experts is required for accurate opinion mining [4].

2.2 Text Mining

Text mining method is a technology with its purpose on processing and extracting useful information on the basis of the processing technologies of natural language, which is semi-structure data. Result values that goes beyond what you can obtain through simple information search, such as locating the category of the text, can be obtained by extracting meaningful information from massive text clusters via the text mining technology and deciphering the connections that lie in between them [5-6].

3 Proposed Method

The opinion mining being suggested in this thesis, as a way of extracting the trends and reputation of the clients spread over the internet with semi-structured data as basis, makes selections on the sentiment data and classifications by gathering the semi-structured data that contain the keywords of 'creative education' with five online mediums spread over the internet, news, blogs, social media, community and blogs, as its target.

3.1 Data collection

The data collection suggested in this thesis, in general, is based on the semi-structured data comprised of opinions on creative education written by general clients on the web. These data fall under the category of historic data, and data mining is executed through fetching the data collected from the web by year.

3.2 Design of Korean grammar sentiment dictionary

The overall system configuration of the development of the sentiment dictionary suggested in this thesis is classified into the three models which are comprised of the storage server for preprocessing and data collection & storage, NLP learning model for the analyzation of morphemes and natural language processing and the phase in which the server is converted into a server for constructing a sentiment dictionary for the extraction of opinions.

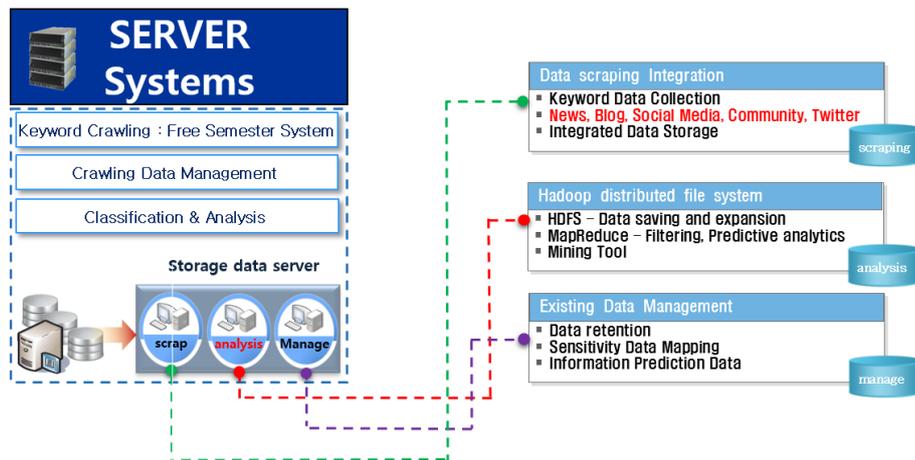


Figure 1. Opinion analysis model

As for the sentiment vocabularies, a sentiment dictionary appropriate for the Korean grammar was constructed on the basis of the priority of words, appearance frequency and weighted value, and, words, such as happy, perfect, great and et cetera, were categorized as positive vocabularies. On the contrary, negative vocabularies, such as bad, sad and et cetera, were included in the negative category and vocabularies and phrases that are consisted of expressions neither positive nor negative were included in the neutrality category, whereas other vocabularies were represented as others.

4 Performance Evaluation

Seen from a comprehensive perspective, as far as public awareness is considered and as a result of the analysis made on the opinion mining suggested in this thesis, it can be understood that creative education is receiving responses that are highly positive. In this thesis, information of online communities, blogs and Twitter were defined as public awareness and news articles that are exposed to the media were defined as social influence. The following [Fig.] illustrates the result of all opinions of creative education.

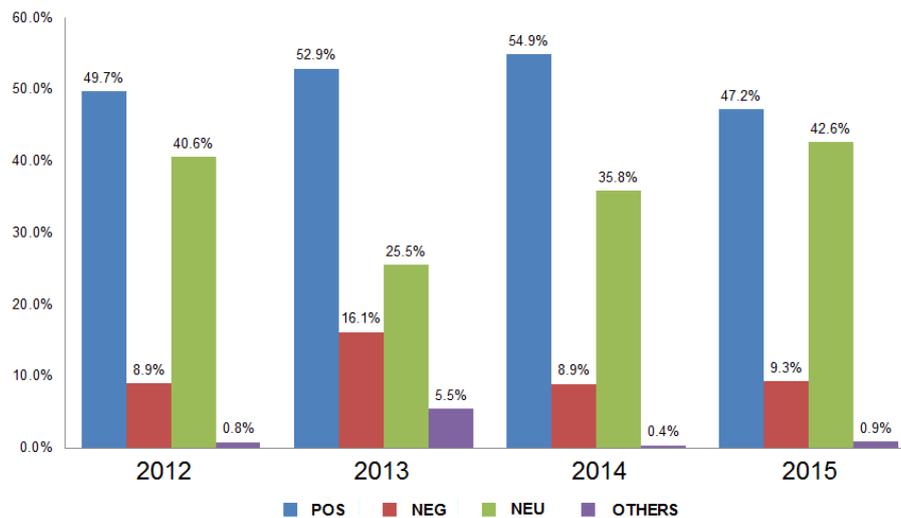


Figure 2. Opinion analysis of creative education

It can be understood that an average of 51.2% has evaluated creative education as positive whereas 10.7% stood on the other side. 36.2% were neutral about their evaluation and 1.9% stood for other opinions. On the assumption that neutral vocabularies are inclusive of positives vocabularies, it can be understood that 87.4% has had positive evaluations.

5 Conclusion

This thesis, with big data extraction from various online media as basis, has made an analysis on public awareness of South Korea on creative education and the social influence on the media via opinion mining. Although the South Korean public and media are showing positive responses toward the issue of creative education, which had become a huge issue around the world prior their adoption that took place in year 2013 after the new government came into office, the issue has entered into a period of lesser demand and thus negative views are also increasing since year 2014. The act of

adopting an advanced system into the South Korean education system not prepared for the very sake of meeting the expectations of the global world may cause negative repercussions difficult to revert. Therefore, the establishment of a customized education system more compatible and improvements on legal policies can be practiced in South Korea are required in lieu of the adoption of a new education system.

References

1. Narayanan, R., Liu, B., Choudhary, A.: "Sentiment analysis of conditional sentences", Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, Singapore, pp.180-189, 2009.
2. Hoon, S.J.: "Design of Opinion Sensitivity Dictionary Model for Big Data Management", 2015.
3. Ming, H., Wenying, N., Xu, L.: "An improved decision tree classification algorithm based on ID3 and the application in score analysis", Chinese Control and Decision Conference (CCDC), pp1876-1879, (2009).
4. Stonebraker, M.: "SQL Databases v. NoSQL Databases", Communications of the ACM, 53(4), pp.10-11, 2010.
5. Manyuka, J.: "Big Data: The Next Frontier for Innovation, Competition, and Productivity", McKinsey Global Institute, America, 2011.
6. Bo, P., Lee, L., Vaithyanathan, S.: "Thumbs up?: sentiment classification using machine learning techniques", Proceedings of the ACL-02 Conference on Empirical methods in Natural Language Processing, vol.10, pp.79-86, 2002.