applied sciences

MDPI

Article

Comparative Study on Perceived Trust of Topic Modeling Based on Affective Level of Educational Text

Youngjae Im [1], Jaehyun Park [2,*], Minyeong Kim [2] and Kijung Park [2]

[1] Division of Design Engineering, Dong-eui University, Busan 47340, Korea; ergoim@deu.ac.kr
[2] Department of Industrial and Management Engineering, Incheon National University, Incheon 22012, Korea; cococha423@inu.ac.kr (M.K.); kjpark@inu.ac.kr (K.P.)
* Correspondence: jaehpark@inu.ac.kr; Tel.: +82-32-835-8867

Received: 9 August 2019; Accepted: 23 October 2019; Published: 28 October 2019



Abstract: Latent dirichlet allocation (LDA) is a representative topic model to extract keywords related to latent topics embedded in a document set. Despite its effectiveness in finding underlying topics in documents, the traditional algorithm of LDA does not have a process to reflect sentimental meanings in text for topic extraction. Focusing on this issue, this study aims to investigate the usability of both LDA and sentiment analysis (SA) algorithms based on the affective level of text. This study defines the affective level of a given set of paragraphs and attempts to analyze the perceived trust of the methodologies in regards to usability. In our experiments, the text of the college scholastic ability test was selected as the set of evaluation paragraphs, and the affective level of the paragraphs was manipulated into three levels (low, medium, and high) as an independent variable. The LDA algorithm was used to extract the keywords of the paragraph, while SA was used to identify the positive or negative mood of the extracted subject word. In addition, the perceived trust score of the algorithm was evaluated by the subjects, and this study verifies whether there is a difference in the score according to the affective levels of the paragraphs. The results show that paragraphs with low affect lead to the high perceived trust of LDA from the participants. However, the perceived trust of SA does not show a statistically significant difference between the affect levels. The findings from this study indicate that LDA is more effective to find topics in text that mainly contains objective information.

Keywords: latent dirichlet allocation (LDA); sentiment analysis (SA); topic modeling; affective level

1. Introduction

The amount of data processed in technical and social systems has exponentially increased with the advent of the fourth industrial revolution and the era of knowledge information processing. Massive text information and public opinions commonly recorded and shared through various social media services have led to the necessity of new technologies and methodologies to find meaningful information hidden in a large set of available unstructured text data. As a response, text mining has attained attention as a technique for extracting meaningful information from unstructured or semi-structured text, such as documents, emails, and hypertext markup language (HTML).

Especially, topic modeling is one of the popular text mining methods that enable us to extract highly interpretable topics in a document set. The latent dirichlet allocation (LDA) algorithm is a representative topic modeling approach [1] where a set of documents is grouped into latent topics with a distinct Dirichlet distribution and each topic is described as a Dirichlet distribution of occurring terms in the document set. The LDA algorithm has been applied to various domains, such as topic

extraction for the abstracts of a research paper set [2], analysis of news articles to interpret relevant social situations [3,4], and identification of consumer characteristics and market trends from social network service (SNS) data [5]. However, the traditional LDA algorithm uses the frequency of terms in text documents as a basis to extract their latent topics.

On the other hand, studies have also developed alternative models on the basis of existing topic modeling algorithms. Sentiment analysis (SA), which is referred to as opinion mining, has been used to identify user attitude, affect, subjectivity, or emotion in a user-generated text [6]. SA distinguishes the affective state, which indicates the positive or negative mood/sentiment of a word or sentence. Thus, SA can be used to collect and analyze vast amounts of data in real-time. Moreover, it can improve perceived trust (the level of trust determined by users based on the extracted information) and minimize errors due to time differences in the investigation process.

As mentioned above, since text mining algorithms such as LDA and SA are commercialized, information summary analysis and services are being provided in various fields. However, the usability evaluation of the results obtained through the algorithm is not yet actively discussed. The main purpose of this study is to investigate what evaluation criteria are more reliable or satisfactory for users who use both LDA and SA algorithms. Therefore, this study considers a perceived trust of the algorithms from users to reflect the interpretability of the algorithm (i.e., how well the user understands the results of the algorithm). In other words, it is considering not only the frequency of each term but also the affective characteristics of the term in extracting topics from text documents. In addition, this study evaluates the perceived trust characteristics of SA which was applied to the keywords extracted by LDA, and LDA which forms the main application algorithm for topic modeling.

In order to achieve the objective of this study, the subjective or affective level was defined considering the degree of positivity and negativity in the designated text. In particular, the Stanford natural language processing algorithm, based on machine learning, was utilized to calculate the positive or negative level for each sentence automatically. Through this research process, the user can grasp the change of perceived trust on the result of existing topic modeling algorithms such as LDA and SA. This study also shows the necessity of a topic modeling algorithm that can reflect the affective level of each topic.

2. Literature Review

2.1. Topic Modeling

Topic modeling is a popular methodology used in text mining. It is an information technology approach that extracts "topical" information from a text document. As one of the most representative topic modeling techniques, LDA is a model for determining potentially meaningful topics [1]. LDA assumes that a set of words is grouped as per a specific topic or topics, and it calculates the probability that the words will be included in each topic and subsequently extracts them as a set of words likely to be included.

Since proposing LDA in a study in 2003, Blei proposed supervised LDA (sLDA) in a later study and compared it with the existing LDA [7]. Subsequent LDA studies have focused on analyzing and obtaining information from social media; these include an analysis of user responses to public events and topics of interest for Twitter users [8]. Related research has led to the proposal of new algorithms based on LDA [9], while other studies have focused on certain variants of the Bayesian inference algorithm [10,11]. Meanwhile, two Rao–Blackwellized online inference algorithms have also been proposed in this regard [12,13].

On the other hand, other streams of research have focused on the procedures and methods for using LDA algorithms. Song et al. [14] have presented several topics and keyword ranking methods to aid users to understand and use extracted topics using LDA in text analysis, while Anandkumar et al. [15] have suggested efficient learning procedures for a variety of subject models, including LDA. Some studies have suggested ways to efficiently improve ad-hoc search using LDA [16].

2.2. Sentiment Analysis

Opinion mining or sentiment analysis (SA) is a popular text mining technique that is mainly used to identify user sensibility, affect, and subjective opinions in texts. In previous studies, topic modeling has been conducted on reviews of articles and articles in SNS, and information and social flows were analyzed [17–19]. In this context, one study has conducted a thorough analysis of text mining of reviews to assess the impact of product reviews on economic performance indicators such as sales volume [20]. Further, Esuli and Sebastiani [21] have analyzed the objectivity, subjectivity, and effect of user opinions via analyzing the opinions and sensibilities of Twitter users. In addition, some studies have focused on the past, present, and future trends of SA while others have proposed a new probability model that can overcome the problems of SA and capture a mixture of a subject and affect at the same time [22,23]. Some studies have also proposed a sensible classifier that can determine the affirmative, negative, and neutral document properties by means of the corpus collection method [24,25].

The idea of user experience (UX) includes the concepts of usability and affective engineering and consists of all interactions between the user and the product [26,27]. Currently, products include not only physical and visible products but also invisible services and algorithms. Therefore, it is necessary to include a UX element in the evaluation of a text mining service or algorithm. According to previous research [28,29], the elements of UX can be classified as usability, affect, and user value (UV). These studies have also derived a quantification model that integrates these key elements into a single index. In particular, a total of 22 hierarchical dimensions, such as UX, and all elements and sub-elements were evaluated. As these studies on UX and affect are expected to be useful in the design of future products or services, they should be considered in this study.

Based on our review of the literature, the topic modeling algorithm and SA are widely applied to identify the main topics and customer opinions from texts provided through various sources. In the meantime, the efficiency aspect of keyword extraction has been extensively dealt with by the developers, but there is a lack of consideration on how users experience the algorithm and accept the results. Also, very few studies compare the characteristics of subject word extraction as per evaluation of the perceived trust of individual algorithms; this is the aspect of text mining that we address in this study.

3. Methodology

3.1. Participants

A total of 21 individuals participated in our study. All subjects were Korean and the mean age of the subjects was 33 years (standard deviation: ±9.05). All participants possessed basic English skills and had no problem understanding the English texts presented in our questionnaire. The participants expressed their agreement in taking part in the study after understanding the experimental content, precautions, and explanations on the use of personal information.

3.2. Contents of the Questionnaire

Our experiment was conducted on participants who received a questionnaire in English, which they subsequently read and evaluated. Figure 1 depicts the format of our questionnaire. The questionnaire consisted of 37 text sections, each of which contained a set of subject terms related to a topic. The sources used in the questionnaires are texts used in actual university entrance examinations in Korea [30]. The college scholastic ability test (CSAT) is a primary test to evaluate the study achievement and used by most universities for an admission decision. There are relevant studies [31–33] to identify significant topics in CSAT. Therefore, the topics presented in this questionnaire represent the titles provided in the college entrance examination and can be regarded as important clues in determining the correct answers.

In order to derive a set of subjects, LDA was used in this study, while SA was applied for additional analysis. LDA is an algorithm that summarizes the central topics in a given set of paragraphs, while SA is an algorithm that categorizes whether words or sentences are positive, negative, or neutral. In this study, we used the MALLET JAVA module to extract the keywords using LDA. In the MALLET module, the "subject" attribute consists of a set of words that occur frequently and together [34]. In this study, two topics were extracted per paragraph, and each topic included less than five keywords. In addition, SA was applied to the extracted keywords to confirm them as positive or negative words. Further, the Stanford CoreNLP toolkit was utilized for SA. Stanford CoreNLP provides a set of human language skills tools and functions. In this study, we used the result of which noun phrase expresses affect through the toolkit.

Here, we briefly describe the notation used in our study. For example, for a set ["point", "tank", "great", "subsequent", "watch"] derived from LDA in a given paragraph, the result obtained with application of SA method is shown in bold font ["point", "tank", "great", "subsequent", "watch"]. At this time, "tank" has a negative characteristic, and "great" has a positive characteristic. In the application of the questionnaire to the participants, we divided the algorithm into two parts to aid participant understanding. LDA was classified into "Expression Method A" and LDA with SA method was classified as "Expression Method B".

The Adventures of Tom Sawyer is regarded as one of the great American novels, but at one point its author despaired of finishing it. In his Autobiography, Mark Twain describes reaching a point in the story where he felt unable to go on: "My tank had run dry." He abandoned it for two years and turned his mind to other things. When he eventually picked up the manuscript again, he made the "great discovery" that the "tank" of his imagination had refilled itself in the meantime, and he was able to complete the story. This discovery was a turning point in Twain's writing career: he learned to watch out for the point in each subsequent book when his tank ran dry, and to take a break before finishing it.

Title: the way Mark Twain overcame his creative block

Expression Method	The following words summarize the subject of the text. Evaluate how trustworthy the algorithm for deriving the following words is.	1	2	3	4	5
A	● point; tank; great; subsequent; watch ● discovery; dry; story; finishing; break					
B	● point; **tank**; **great**; subsequent; watch ● discovery; dry; story; finishing; break					

* Expression Method A: LDA (Displayed as two sets of less than 5 subject words)

* Expression Method B: LDA (Displayed as two sets of less than 5 subject words) + Sentiment (Bold)

Figure 1. Format of the questionnaire provided to study participants.

3.3. Experimental Design

In this study, we defined "Affective Level" as an independent variable to classify the paragraph characteristics. The level of affect per paragraph was determined by how the paragraph reveals the attributes of a positive or negative mood. The levels of the independent variable were selected as high, medium, and low. SA calculates the degree of affective level as very negative, negative, neutral, positive, and very positive in each sentence. In this experiment, the sum of attribute scores (very negative (−2), negative (−1), neutral (0), positive (1), very positive (2)) for each sentence is calculated. An absolute value of the sum is used as the final affective level of the corresponding paragraph. If the total value of the sum is 0, 1, or 2 points, it is classified as "low", whereas 3 or 4 points corresponds to "medium", and 5 or more as "high".

Paragraphs corresponding to low affective levels included "How did Mark Twain overcome the clogged creativity?" (paragraph 2), "Changes in the function of classical music" (paragraph 7), and "The rising cause of the Himalayas in progress" (paragraph 31), which contained primarily objective information or facts. From a different perspective, paragraphs with high affective levels included "Difficulties in establishing causal relationships in social science" (paragraph 10), "The effect of naming on children's identity" (paragraph 21), and "The ethical problems associated with creation" (paragraph 33), which contain primarily subjective opinions rather than objective information. Meanwhile, paragraphs with moderate levels of affect are blended with objective information and subjective opinions, as in the case of those titled "The Impact of Situations on Color Preference" (paragraph 11) and "Mandeville's book that caused misunderstandings in the Middle Ages" (paragraph 27).

In this study, the dependent variable was defined as "perceived trust" for each algorithm. The perceived trust represents how well the keywords in Expression Method A (LDA) and B (LDA + SA) represent the topic of each paragraph. The evaluation method was based on participants' subjective scoring. The evaluation scale was selected from the five points of the Likert scale considering the burden and accuracy based on the participant's feeling. Points 1 to 5 on the scale corresponded to "very dissatisfied", "slightly dissatisfied", "average", "slightly satisfied", and "very satisfied", respectively.

3.4. Procedure

In our study, subjects first read a given paragraph and subsequently asked for an evaluation questionnaire containing the keywords. Subjects were individually selected for the evaluation, and the questionnaire was freely proceeded between the method of email and face to face in consideration of the subject's situation and intention. In addition, there was no restriction on the assessment time in order to not pressurize the participants. These notifications were indicated on the cover of the questionnaire in advance. In the analysis phase, data from the perceived trust evaluation of LDA and SA collected from a total of 21 subjects and 37 paragraphs were used. Consequently, it was possible to judge the average perceived trust per paragraph.

3.5. Data Analysis

In our study, analysis of variance (ANOVA) and post-analysis were conducted to analyze the perceived trust difference by affective level. Since the level of the independent variable according to the experimental design was three groups (low, medium, high), ANOVA analysis was used to test the difference in perceived trust among the groups. In addition, the number of cases in each group was the same and the normality assumption was achieved, and the Student–Newman–Keuls (SNK) method was used for post-analysis. The statistical program used for the analysis was the IBM SPSS Statistics package (Ver. 25.0).

4. Results

The main result of this study is the difference in response of the dependent variable (perceived trust of main word summarization algorithm) according to the independent variable after analysis of the difference of the perceived trust score between LDA and SA according to the paragraph affective level.

4.1. Perceived Trust of Latent Dirichlet Allocation (LDA)

4.1.1. Average of Perceived Trust

The results of LDA perceived trust evaluation according to paragraph affective levels are as follows: The mean score of the paragraph at the low level is 3.40 (standard deviation: ±0.41), the middle level is 2.97 (standard deviation: ±0.28), and the high level is 3.11 (standard deviation: ±0.35). In other words, if the affective level is high or low, the level of perceived trust usually indicates a level above normal. But, if the affective level is medium, the level of perceived trust shows a level below normal.

4.1.2. ANOVA and Post-Analysis

The results of ANOVA analysis and post-analysis using SPSS are as follows. First, our ANOVA analysis indicated that the influence of the affective level on the perceived trust of the LDA evaluation algorithm was statistically significant (p-value < 0.05), as shown in Table 1. In other words, there is a perceived trust difference between the users according to the affective level of the paragraph. In addition, SNK as the post-hoc analysis was conducted to verify the differences in affective levels, as shown in Table 2. As a result of this analysis, paragraphs with low affective levels were classified into one group, and those of medium and high affective levels were classified into one group. Figure 2 depicts the average LDA confidence score for each affective level in the paragraphs.

Table 1. Analysis of variance (ANOVA) results corresponding to the application of Latent Dirichlet Allocation (LDA).

		Sum of Squares (Type III)	DF	Mean Square	F	p-Value
intercept	theory	7688.995	1	7688.995	1387.875	0.000
	error	110.802	20	5.540		
affective level	theory	23.952	2	11.976	13.153	0.000
	error	36.421	40	0.911		
subject	theory	110.802	20	5.540	6.086	0.000
	error	36.746	40.366	0.910		
affective level×subject	theory	36.421	40	0.911	1.055	0.382
	error	616.491	714	0.863		

Table 2. Student–Newman–Keuls (SNK) analysis of LDA results.

Affective Level	N	Subset	
		1	2
Medium	231	2.9740	
High	294	3.1054	
Low	252		3.4048
p-value		0.110	1.000

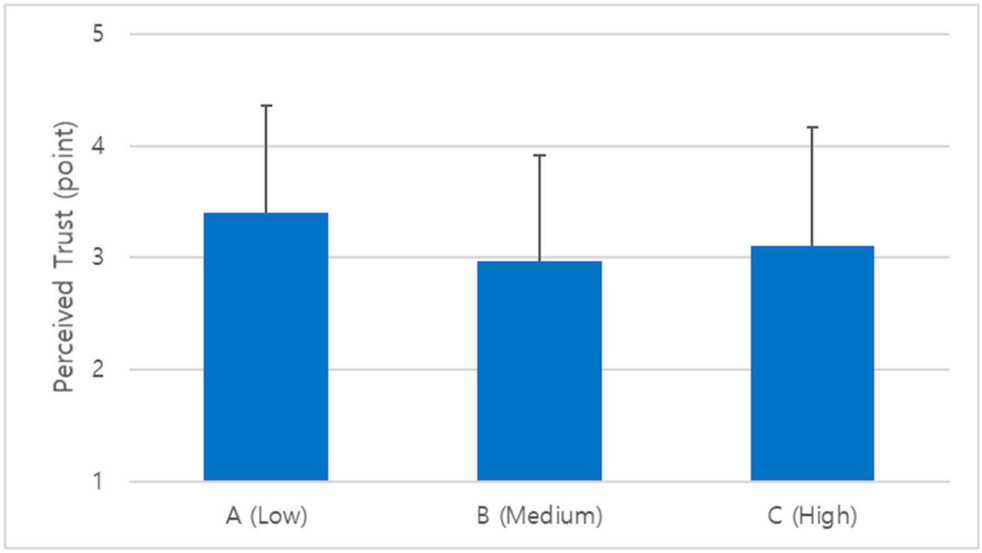


Figure 2. Perceived trust of LDA according to affective level.

4.2. Perceived Trust of LDA+ Sentiment Analysis (SA)

4.2.1. Average of Perceived Trust

The results of the perceived trust evaluation as per LDA+SA according to the affective levels of the paragraph are as follows. The mean of the paragraph at the low level is 2.90 (standard deviation: ±0.42), the middle level is 2.73 (standard deviation: ±0.50), and the high level is 2.62 (standard deviation: ±0.22). Perceived trust for all affective level was below 3 points. In general, users' satisfaction with algorithm results was slightly dissatisfied.

4.2.2. ANOVA and Post-Analysis

The results of ANOVA analysis and post-analysis using SPSS are as follows. As per ANOVA analysis, the influence of the affective level on the perceived trust of the SA evaluation algorithm was not statistically significant (p-value = 0.101) as shown in Table 3. In other words, LDA+SA perceived trust score can be observed in the same group regardless of the affective level of the paragraph. Figure 3 shows the average LDA+SA perceived trust score based on the affective level of the paragraph.

Table 3. Results of analysis of variance (ANOVA) applied to LDA+ sentiment analysis (SA).

		Sum of Squares (Type III)	DF	Mean Square	F	*p*-Value
intercept	theory	4245.003	1	4245.003	447.374	0.000
	error	189.774	20	9.489		
affective level	theory	4.964	2	2.482	2.425	0.101
	error	40.945	40	1.024		
subject	theory	189.774	20	9.489	9.273	0.000
	error	41.242	40.305	1.023		
affective level×subject	theory	40.945	40	1.024	1.099	0.317
	error	469.425	504	0.931		

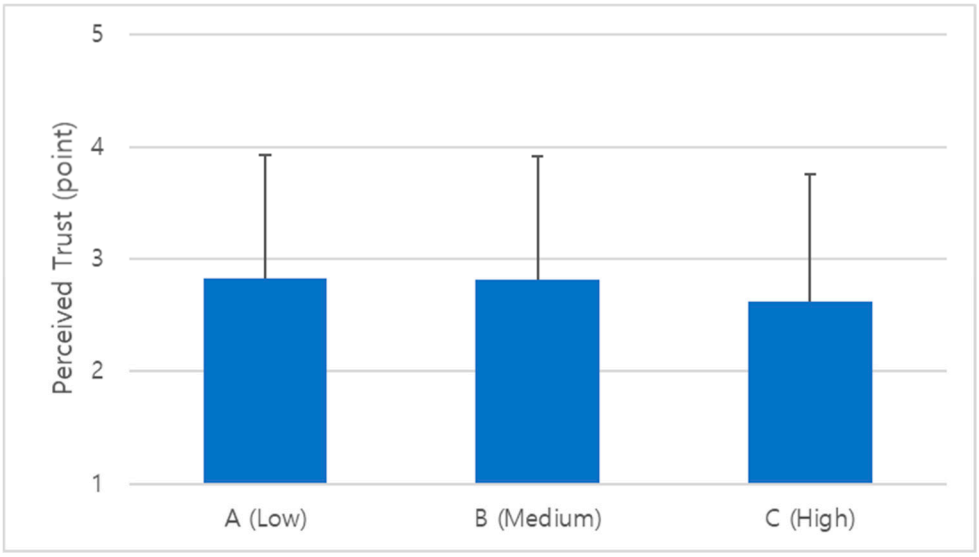


Figure 3. Perceived trust of LDA+SA according to affective level.

5. Discussion

The participants revealed a relatively high perceived trust score for LDA processing low affective text. This shows that the LDA algorithm is effective in extracting the topics of text with a low affective level that mainly consists of terms indicating objective information or facts. According to related research [35], we note that SA is used to summarize the characteristics of paragraphs that mainly include subjective opinions or affective expressions, such as highly affective paragraphs. In the study by Turney [36], SA was used to analyze review texts corresponding to specific writing interests such as automobiles, movies, and travel. In Turney's study, the sentiment orientation of the text was determined based on the amount of information regarding the words used with apparent affective vocabulary, such as "excellent" or "poor". As result, the more emotional vocabularies that are contained in text, the sentiment orientation has a greater value.

However, in this study, we observed no statistically significant difference in the perceived trust of LDA+SA with respect to the affective level. A plausible reason for this result is that there are differences in the text attributes used across studies. The results of keywords extraction through LDA+SA are basically affected by the attributes of the text itself to be analyzed. Therefore, the vocabulary represented by the attributes of text (papers, magazines, newspaper articles, etc.) may vary, even if the text indicates the same subject.

This study used the university entrance examination texts as a case study to investigate the effect of LDA+SA. The main topics of the texts cover the contents, such as how to improve creativity, efforts to improve security systems, lack of organized efforts by disaster response organizations, and the role of sports as a means of sustainable development. Although there are some text sections that express subjective opinions, the exam texts mostly include the descriptions of specific methods, problems, and objective information. However, in other related research [17–19], mainly reviews of products and articles in SNS were used, and the relevant paragraphs revealed a clearer subjectivity and included a relatively large number of expressions of positive and negative affect. In fact, Myung et al. [37] collected review text from an online shopping mall as experimental data and then analyzed it based on the polarity information of the vocabulary, indicating the characteristics of each product. The polarity information for the product is expressed as "uncomfortable" or "ease of use".

The implications of this study are as follows. There are limits to the applicability of SA itself, which can impact the result. SA used in this study is applied at the word level, and the performance of the algorithm itself may be somewhat restricted. For example, after analyzing that the text is composed of three positive words and one negative word, it is judged to be "positive text" consisting of two

positive words as a whole. In other words, simply counting the number of positive or negative words is not sufficient to interpret the overall meaning of the text. In addition, there are some parts of the collected text data that are not related to the affect. Thus, it is possible to propose a process extracting only portions to be subjected to SA after data collection is completed. For example, statements that only address facts, such as "buying a new laptop today", can be categorized as objective texts. These statements could be initially excluded from the analysis. Consequently, the major contribution of this study is to lay a foundation for a transition of prevailing technical viewpoints in the integration of LDA and SA to a user viewpoint. Indeed, there are various studies addressed the integration of LDA and SA through an improved semantic algorithm for LDA [38–40]. However, existing studies lack important discussion for how users perceive information delivered from LDA and SA. The findings from this study may provide a plausible explanation for the necessity of a topic modeling algorithm that provides more trustworthy outputs, depending on the extent to which affective level is associated with the text.

In addition, there are limitations to applying LDA as well. There may be limitations due to the Dirichlet distribution modeling the variability among the ratios of keywords. Analyzing the keywords on the basis of occurrence of the words, rather than grasping them in the overall context, may reduce the user's confidence in the interpretation of the results. For example, even if a paper on sports is a subject that is relatively more relevant to health than international finance, it is not possible to model accurate subject associations if health-related words appear frequently in the paragraph on international finance. Therefore, in order to overcome these limitations, we can consider the correlation topic model (CTM) instead of LDA in future studies. CTM is superior to LDA in the predictability of modeling and is a more realistic approach to visualizing and navigating unstructured data sets [41]. In fact, the LDA model predicts keywords based on potential topics implied by the observations, but the CTM can predict items on additional topics that may be conditionally related. Moreover, the documents used in this study were English paragraphs, but the subjects were all Koreans. Although we recruited the subjects with appropriate English ability, we did not fully address the difference in language ability among individuals.

6. Conclusions

The purpose of this study was to assess the usability of topic modeling algorithm as the user-centric aspects. Also, the affective level of text in a paragraph, as well as the frequency of existing words, were considered for the method of extracting subject words from text. The algorithms that summarize the characteristics of paragraph are employed for the case where LDA is used singly and the case where LDA and SA are applied together. The paragraphs provided to the users were composed of assignment tests. Subjects were asked to evaluate the perceived trust of a set of keywords derived using algorithms. At this time, the affective level of the paragraph was classified using the Stanford NLP to analyze the difference of the perceived trust evaluation according to the affective level of the paragraph. In analyzing this perceived trust, we also interpreted the results not only from the technical characteristics of the algorithm but also from the ergonomic viewpoint. As a result, the effect of affective level of text on the perceived trust of LDA algorithm was found to be statistically significant, and we found through post-analysis that the perceived trust of the paragraphs with low affective level was higher than those of the mid- and high-affective-level ones. In the case of LDA combined with SA, the effect of the affective level of text on the perceived trust was not statistically significant.

From our results, we can draw the conclusion that it is possible to select and use algorithms that summarize subject words according to the affective level of the document. In terms of SA, there has been a focus on the classification of affirmative and negative vocabularies in text and then calculating the frequency of these polar vocabularies. In the future, in order to improve the effectiveness of SA, it is necessary to analyze not just the word level but the property of the text. Furthermore, it is expected that a corpus that accumulates affective expressions utilized by actual users will need to be constructed.

Author Contributions: Conceptualization, J.P.; methodology, J.P.; formal analysis, Y.I. and M.K.; writing—original draft preparation, Y.I. and M.K.; writing—review and editing, Y.I., J.P. and K.P.

Funding: This work was supported by the Incheon National University Research Grant in 2018 (Grant No.: 20180402).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
2. Gerrish, S.; Blei, D.M. A Language-based Approach to Measuring Scholarly Impact. *ICML* **2010**, *10*, 375–382.
3. Newman, D.; Block, S. Probabilistic Topic Decomposition of An Eighteenth-century American Newspaper. *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 753–767. [CrossRef]
4. Rajasundari, T.; Subathra, P.; Kumar, P.N. Performance analysis of topic modeling algorithms for news articles. *J. Adv. Res. Dyn. Control Syst.* **2017**, *11*, 175–183.
5. Hu, Y.; John, A.; Seligmann, D.D. Event Analytics via Social Media. In Proceedings of the Workshop on Social and Behavioural Networked Media Access, Scottsdale, AZ, USA, 1 December 2011.
6. Chen, H.; Zimbra, D. AI and opinion mining. *IEEE Intell. Syst.* **2010**, *25*, 74–80. [CrossRef]
7. Mcauliffe, J.D.; Blei, D.M. Supervised Topic Models. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–11December 2008.
8. Michelson, M.; Macskassy, S.A. Discovering Users' Topics of Interest on Twitter: A First Look. In Proceedings of the Fourth Workshop on Analytics for Noisy Unstructured Text Data, Toronto, ON, Canada, 26 October 2010.
9. Hoque, E.; Carenini, G. MultiConVis: A Visual Text Analytics System for Exploring a Collection of Online Conversations. In Proceedings of the 21st International Conference on Intelligent User Interfaces, Sonoma, CA, USA, 7–10 March 2016.
10. Burkhardt, S.; Kramer, S. Multi-label classification using stacked hierarchical Dirichlet processes with reduced sampling complexity. *Knowl. Inf. Syst.* **2019**, *59*, 93–115. [CrossRef]
11. Papanikolaou, Y.; Foulds, J.R.; Rubin, T.N.; Tsoumakas, G. Dense distributions from sparse samples: Improved Gibbs sampling parameter estimators for LDA. *J. Mach. Learn. Res.* **2017**, *18*, 1–58.
12. Teh, Y.W.; Newman, D.; Welling, M. A Collapsed Variational Bayesian Inference Algorithm for Latent Dirichlet Allocation. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 3–6 December 2007.
13. Canini, K.; Shi, L.; Griffiths, T. Online Inference of Topics with Latent Dirichlet Allocation. In *Artificial Intelligence and Statistics*; Springer: Clearwater Beach, FL, USA, 2009.
14. Song, Y.; Pan, S.; Liu, S.; Zhou, M.X.; Qian, W. Topic and keyword re-ranking for LDA-based topic modeling. In Proceedings of the 18th ACM Conference on Information and Knowledge Management, Hong Kong, China, 2–6 November 2009.
15. Anandkumar, A.; Foster, D.P.; Hsu, D.J.; Kakade, S.M.; Liu, Y.K. A spectral algorithm for latent dirichlet allocation. In Proceedings of the Advances in Neural Information Processing Systems, Stateline, NV, USA, 3–8 December 2012.
16. Wei, X.; Croft, W.B. LDA-based document models for ad-hoc retrieval. In Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, WA, USA, 6–11 August 2006.
17. Jin, S. Reliability Analysis of VOC Data for Opinion Mining. *J. Intell. Inf. Syst.* **2016**, *22*, 217–245.
18. González-Bailón, S.; Paltoglou, G. Signals of public opinion in online communication: A comparison of methods and data sources. *Ann. Am. Acad. Political Soc. Sci.* **2015**, *659*, 95–107. [CrossRef]
19. Hu, M.; Liu, B. Mining and summarizing customer reviews. In Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 22–25 August 2004.
20. Ghose, A.; Ipeirotis, P.G. Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Trans. Knowl. Data Eng.* **2011**, *23*, 1498–1512. [CrossRef]
21. Esuli, A.; Sebastiani, F. Determining Term Subjectivity and Term Orientation for Opinion Mining. In Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics, Trento, Italy, 5–6 April 2006.

22. Mei, Q.; Ling, X.; Wondra, M.; Su, H.; Zhai, C. Topic sentiment mixture: Modeling facets and opinions in weblogs. In Proceedings of the 16th international conference on World Wide Web, Banff, AB, Canada, 8–12 May 2007.
23. Cambria, E.; Schuller, B.; Xia, Y.; Havasi, C. New avenues in opinion mining and sentiment analysis. *IEEE Intell. Syst.* **2013**, *28*, 15–21. [CrossRef]
24. Alnawas, A.; Arici, N. The Corpus Based Approach to Sentiment Analysis in Modern Standard Arabic and Arabic Dialects: A Literature Review. *J. Polytech.* **2018**, *21*, 461–470. [CrossRef]
25. Pak, A.; Paroubek, P. Twitter as a corpus for sentiment analysis and opinion mining. In Proceedings of the Seventh Conference on International Language Resources and Evaluation, Valletta, Malta, 19–21 May 2010.
26. Kim, H.K.; Han, S.H.; Park, J.; Park, J. Identifying affect elements based on a conceptual model of affect: A case study on a smartphone. *Int. J. Ind. Ergon.* **2016**, *53*, 193–204. [CrossRef]
27. Kim, H.K.; Park, J.; Park, K.; Choe, M. Analyzing thumb interaction on mobile touchpad devices. *Int. J. Ind. Ergon.* **2018**, *67*, 201–209. [CrossRef]
28. Park, J.; Han, S.H.; Kim, H.K.; Cho, Y.; Park, W. Developing elements of user experience for mobile phones and services: Survey, interview, and observation approaches. *Hum. Factors Ergon. Manuf. Serv. Ind.* **2013**, *23*, 279–293. [CrossRef]
29. Park, J.; Han, S.H.; Kim, H.K.; Oh, S.; Moon, H. Modeling user experience: A case study on a mobile device. *Int. J. Ind. Ergon.* **2013**, *43*, 187–196. [CrossRef]
30. Available online: http://www.ebsi.co.kr/ebs/xip/xipc/previousPaperList.ebs (accessed on 9 August 2019).
31. Kim, N.B. A corpus-based lexical analysis of the foreign language (English) test for the college scholastic ability test (CSAT). *Engl. Lang. Lit. Teach.* **2008**, *14*, 201–221.
32. Kang, M.K.; Kim, Y.M. The internal analysis of the validation on item-types of Foreign (English) Language Domain of the current 2005 CSAT for designing the level-differentiated English tests of the 2014 CSAT. *J. Korea Engl. Educ. Soc.* **2013**, *12*, 1–35.
33. Park, H.J.; Jang, K.Y.; Lee, Y.H.; Kim, W.J.; Kang, P.S. Prediction of Correct Answer Rate and Identification of Significant Factors for CSAT English Test Based on Data Mining Techniques. *Kips Tr. Softw. Data Eng.* **2015**, *4*, 509–520. [CrossRef]
34. Hu, Z.; Fang, S.; Liang, T. Empirical study of constructing a knowledge organization system of patent documents using topic modeling. *Scientometrics* **2014**, *100*, 787–799. [CrossRef]
35. Rao, Y.; Lei, J.; Wenyin, L.; Li, Q.; Chen, M. Building emotional dictionary for sentiment analysis of online news. *World Wide Web* **2014**, *17*, 723–742. [CrossRef]
36. Turney, P.D. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proceeding of the 40th Annual Meeting on Association for Computational Linguistics, Philadelphia, PA, USA, 7–12 July 2002.
37. Myung, J.; Lee, D.; Lee, S. A Korean Product Review Analysis System Using a Semi-Automatically Constructed Semantic Dictionary. *J. Kiise Softw. Appl.* **2008**, *35*, 392–403.
38. Poria, S.; Chaturvedi, I.; Cambria, E.; Bisio, F. Sentic LDA: Improving on LDA with semantic similarity for aspect-based sentiment analysis. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 4465–4473.
39. Shams, M.; Shakery, A.; Faili, H. A non-parametric LDA-based induction method for sentiment analysis. In Proceedings of the 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP), Shiraz, Iran, 2–3 May 2012; pp. 216–221.
40. Yuan, B.; Wu, G. A Hybrid HDP-ME-LDA Model for Sentiment Analysis. In *2nd International Conference on Automation, Mechanical Control and Computational Engineering (AMCCE)*; Atlantis Press: Paris, France, 2017.
41. Blei, D.; Lafferty, J. Correlated Topic Models. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 5–8 December 2005.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).