**International Journal of Electrical and Computer Engineering (IJECE)**
Vol. 9, No. 1, February 2019, pp. 426~438
ISSN: 2088-8708, DOI: 10.11591/ijece.v9i1.pp426-438 

# Topic discovery of online course reviews using LDA with leveraging reviews helpfulness

**Fetty Fitriyanti Lubis, Yusep Rosmansyah, Suhono H. Supangkat**
School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia

**Article Info**

***Article history:***

Received Mar 1, 2018
Revised Jul 14, 2018
Accepted Sep 9, 2018

***Keywords:***

Learner reviews
Learning analytics
MOOCs
Sentiment analysis
Topic modeling

**ABSTRACT**

Despite the popularity of the Massive Open Online Courses, small-scale research has been done to understand the factors that influence the teaching-learning process through the massive online platform. Using topic modeling approach, our results show terms with prior knowledge to understand e.g.: Chuck as the instructor name. So, we proposed the topic modeling approach on helpful subjective reviews. The results show five influential factors: "learn easy excellent class program", "python learn class easy lot", "Program learn easy python time game", and "learn class python time game". Also, research results showed that the proposed method improved the perplexity score on the LDA model.

*Copyright © 2019 Institute of Advanced Engineering and Science.*
*All rights reserved.*

***Corresponding Author:***

Fetty Fitriyanti Lubis,
School of Electrical Engineering and Informatics,
Institut Teknologi Bandung,
Jl. Ganesha No. 10, Bandung, Indonesia.
Email: fettyfitriyanti@students.itb.ac.id

## 1. INTRODUCTION

Reviews play an essential role in the field of e-commerce and tourism through Amazon and Trip Advisor. Reviews present information and help users to make transaction decisions; hence, they increase business value for both companies [1], [2]. Both parties use summarizing methods to get useful information from the reviews. This helpful information is called aspects of the product or item. An aspect is the nature of the object that is commented on by reviewers [3].

In MOOCs, reviews are an accessible medium for learners to share opinions and experiences related to the course (such as the instructor, material, test, and assignments through the Class Central site). As in the field of e-commerce and tourism, reviews were used to analyze the user behavior through the aspects that the user criticizes. This paper's goal is to understand learner experiences through reviews.

Aspects are the properties of an object that users comment on [3], [4]. The list of issues is extracted by processing the reviews. However, a semantic lexicon is required for a specific domain to handle the reviews. A semantic lexicon approach to a particular field requires human intervention, takes time, and increases costs. Additionally, we did not know the aspect contained in the discussion. For example, some general issues of a hotel will be frequently found in reviews such as service, cleanliness, and price, since people are commonly interested in these domains. However, it will be difficult to make a list of aspects related to online courses, as they have never been interested, until recently, in the MOOCs. It is easy to see that online courses' aspects must also be subjective to reviewers and very few in number.

In this paper, we propose a method to find semantic classes automatically from student reviews related to the course aspects. We introduce helpful review sentences, in a way based on Blei's research on topic modeling [5]. However, Blei used consumer review data with known issues [6], whereas in this research, the elements of the course review data are unknown [7].

We developed a primary method for discovering aspects of learner experiences with a topic modeling approach. It is based on reviews that are voted helpful by a reader. The reviews contained specific information judged by readers to be meaningful to them. The rest of this paper has the following structure: section 2 reviews related works, section 3 describes our method, section 4 presents the results of our experiments and section 5 presents our conclusions and discussion of future work.

## 2. RELATED WORKS

Extraction of product aspects has been carried out by supervised, semi-supervised and unsupervised learning methods [3]. The supervised method requires annotating corpora for statistical classification training [3]. However, the data used in this study are not annotated for aspects. Most MOOCs also let their readers give reviews without structuring the review section with an issue. Therefore, this study uses an unsupervised clustering method. The goal is understanding the groups formed. Topic modeling is an unsupervised classification method for performing this process. The LDA method of modeling is a prevalent method of deciding the is-sues of a document [5].

Many studies on course review topic exploration and discussions on the MOOCs platform have used LDA [6], [8], [9]. However, LDA was used for different purposes in these studies. Ezen-can et al. used LDA to investigate qualitative discussion groups [6]. Atapattu et al. applied LDA to find groups of MOOCs topics of discussion related to the lectures [10]. However, LDA was unable to provide proper labeling due to a lack of source references. Thus, Atapattu et al. proposed automatic labeling by generating candidates for local labels from lecture courses [8]. Peng et al. Used LDA to detect a series of potential topics in the course review data set. Peng et al. combined LDA with the features of "like" behavior to improve the accuracy of topic detection and word coherence on each topic [11].

Ezen-can et al. proposed large-scale automatic dis-course analysis and mining to support student learning [6]. Their system uses a cluster approach to similar group reviews, then compares the clusters formed by groups annotated manually by MOOCs researchers. Their results suggest that unsupervised modeling frameworks for synchronous conversations with asynchronous discussions can offer insights from similar posts on a large scale and the topics covered by learners.

Atapattu et al. researched a visualization dashboard to find and classify emerging discussion topics [10]. The visualization aimed to explore the correlations between the issues discussed and other variables such as comments, posts, ideas, and interventions from the instructor. The output of this study showed the graph relations between the topic and the threads in lecture-related discussions.

Peng et al. used LDA to detect the interests of learners in the review-review course with real-life datasets. Peng et al. incorporated the LDA method with behavioral features known as LDA-like to obtain higher accuracy values when processing topics and keyword topics within each topic [9].

In contrast to the approach of the extraction conducted in various research works above, our study goal is to extract learner experiences by mining the course reviews from the Class-Central site. We used review data without knowing the aspects of the data and without experts in MOOCs who can help to label the data manually. Therefore, we proposed incorporating the LDA method with helpful review features.

## 3. EXPLORATORY ANALYSIS OF COURSE REVIEWS

In this section, we describe a series of early-stage explorations carried out in the course review data. The set of stages performed is as follows: acquisition of the course reviews, investigation of course review data, and study of the helpful reviews. The last two steps become the basis of the proposed method.

### 3.1. Data gathering

We collected the MOOCs student reviews data from the Class-Central website. Class-Central is a search engine with reviews for MOOCs and free online courses. More than six million learners have used the site to assist their enrollment decisions for online courses. The top 50 classes are quality courses with many reviews. Therefore, the first data collection focuses on the top 50 ranking course review data. Figure 1 shows a review with complete features. Detailed features complete with a review are as follows:

1. Rating (1-5 stars)
2. Review content
3. The number of people who have voted whether a review is helpful
4. The number of people who have stated that a review was necessary
5. Course status of learners
6. Learner id
7. Course id
8. Review id

16 out 17 people found the following review useful
2 years ago
Anonymous completed this course
★★★★★

I've taken and finished part 1 and part 2 of this course. There is a ton of information and skills learned.
The practice problems are not necessary for your final grade, but the quizzes and mini projects are. I
really recommend doing the practice problems as they help to pound in some fundamentals that are beyond synta

READ MORE ∨

Was this reviews help you? YES | NO

Figure 1. Full-featured course review

The next step is conducting data exploration. The goal is to find the features used and analyze the compatibility data with research problems. The target is filtering the elements and limiting the scope of data so that there is no bias.

### 3.2. Exploratory analysis

Reviews from the Class-Central website are publicly accessed, so other students can read them. The reviews also contain information that is usable to a recommenders system to provide personalized courses. Thus, learners will decide to enroll in the class more easily.

In this section, we evaluate the course's review features by analyzing the data visualization. Visu-alization helps to understand the characteristics. Visually analyzing the elements includes many reviews from the subject group, rating, sentence length, and frequency of word appearance in both positive and negative reviews. Figure 2 shows the subject review distribution. The subject with the most reviews is Programming. Moreover, an item with the least reviews is Theoretical Computer Science.

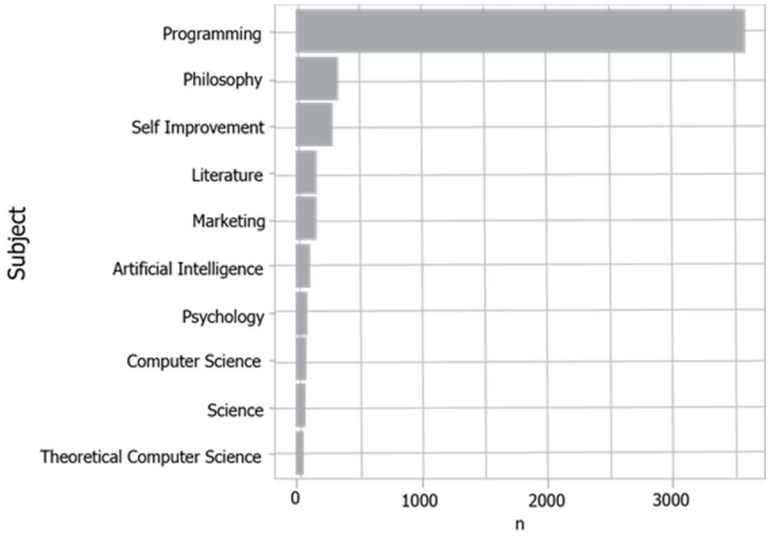



Figure 2. Distribution of reviews on each subject of the top 50 courses data on Class-Central

Figure 3 shows the Programming course's review rating distribution. Based on Figure 3, ratings of 4 and 5 are the dominant group. The positive association is assigned to ranks 4 and 5; the neutral team rates with 3, and the negative group includes ranks of 1 and 2. This means that courses with a rating range receive a positive response from learners.

Figure 4 shows the word cloud visualization from the two review categories, which are positive reviews and negative reviews based on the ratings obtained by each reviewer. The results of exploratory data analysis show that the programming subject is dominated by the number of reviews. Therefore, the next process is to collect the review data with a focus on the topic of programming to reduce the bias.

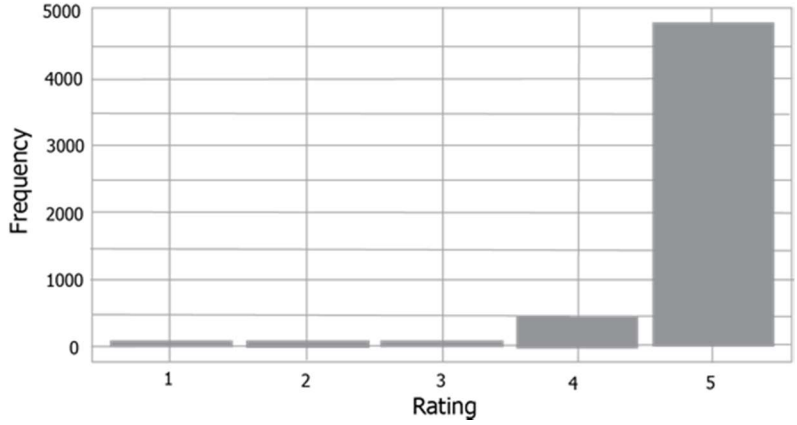



Figure 3. Ratings of the review data in the subject group programming




Figure 4. Common words in (a) positive reviews and (b) negative reviews from the review data in the Programming course category

### 3.3. Mining review helpfulness

User reviews play an important role in disseminating information, convey user confidence, and promote products in electronic commerce [10]. A large volume of discussions will lead to information overload for the reader. Providing helpful information can help to overcome the problem of information overload. Commerce sites, such as Amazon, use a community-based voting technique known as social navigation [11]. The method asks the reader to rate the usefulness of the product or service reviews and display the valuable information about the product or service provided by all the reviewers. However, many of the reviews were not getting votes from readers because those reviews were newly submitted.

The reviews that were voted helpful by readers will be able to help other readers in making decisions that will impact the business of the product or service provider. The previous classification technique only used to label a review without weighting a vote value. Automated review classification helps readers and product or service providers to respond and act immediately [10].

The Amazon website displays helpful information from reviews based on reader votes with the format "x" from reader "y" finds helpful reviews based on the review's content. The regression technique used vote information to build the predictive model. The prediction model developed to determine the estimation value of helpful reviews [11].

The field of e-learning uses the same mechanism as e-commerce in determining the review's quality. The same problem also arises in the field of e-learning. However, in e-learning, the predictive

model's goal is to group the discussions as helpful or not helpful. Thus, the approach is different. This study used data collected from the Class-Central site using a Naive Bayes algorithm to classify reviews into the groups 'review helpful' and 'review unhelpful'.

## 4. LEARNERS' EXPERIENCES EXTRACTION

Provide a statement that what is expected, as stated in the "Introduction" chapter can ultimately result in "Results and Discussion" chapter, so there is compatibility. Moreover, it can also be added the prospect of the development of research results and application prospects of further studies into the next (based on result and discussion).

### 4.1. Overview

This study aims to extract the learners' experiences from the reviews dataset. That output is used to understand the learner focus during the course. Therefore, a learner will receive the appropriate recommendations. Term frequency is one of the primary and standard techniques used to understand the topic of a document. This method evolves into inverse document frequency (IDF), then term frequency-inverse document frequency (TF-IDF), so that the importance weight of every word in a document is obtained more objectively based on the context. Using one of these techniques, each word will have an influence. The word cloud model visualized the weight of each word. Figure 4 shows the word cloud from two different review groups, positive reviews and negative reviews. Based on Figure 4, many words are still general and less specifically describe learners' experiences.

LDA is used to visualize word detail from learners' reviews. LDA is a topic modeling technique to present topics found as a graphical model. Blei et al. [5] proposed an LDA model to apply to topic modeling on various domains in recent years [12], [13]. The use of LDA methods in the education field mostly focuses on the problem of analyzing and extracting semantic information from textual data. However, the research undertaken has not involved the characteristics of the specific users' behavior regarding the textual content, as suggested by Peng et al. [9]. Similar to Peng et al.'s suggestion, we use the helpful or unhelpful review categories obtained as a form of user ratings after reading the reviews submitted.

### 4.2. Review helpfulness extraction

The utility technique and the classification technique use data from a voting system to categorize reviews as helpful or unhelpful. The voting system displays the number of users who agree that the review is helpful or unhelpful. The equation is used to calculate the utility value (1) [14] by using the ratio of users who indicated that a review was helpful. Then, the rule in equation (2) is used to decide the review label. Meanwhile, a model classification is built to determine a user's review label data that have not received a vote. The classification algorithm is used in similar cases such as Support Vector Machine (SVM) [15], tree models [16], and linear regression [16].

The utility value is calculated using the value of variables x and y in equation (1). The "x" and "y" values represent helpful vote reviews displayed in the format "x of y people found these reviews helpful."

$$u = \frac{x}{y} \tag{1}$$

The utility value $(u)$ is used to decide the reviews' label using equation (2).

$$\begin{matrix} u \geq 1, helpful\ review \\ u < 1, unhelpful\ review \end{matrix} \tag{2}$$

For reviews without utility values, a classification technique is used to decide the review's label. The simplified algorithm that is often used to predict this case is a Support Vector Machine (SVM). However, in this study, we used the Naive Bayes algorithm. The Naive Bayes algorithm was selected because this algorithm has advantages that match the characteristics of the data used. These benefits include that the model performed well despite the training data being small. This algorithm has been proven to be effective with e-mail for spam filtering [17], for determining sentiment in social media data [18], and for security applications in computer networks.

Classification with the Naive Bayes algorithm aims to divide each review into the review categories helpful and unhelpful. Precision, recall, F-measure, and accuracy metrics in equations (3)-(6) are used to measure the model performance.

$$\text{Pr} = \frac{TP}{(TP + FP)} \quad (3)$$

$$Re = \frac{TP}{(TP + FN)} \quad (4)$$

$$F - measure = \frac{Re \times Pr}{Re + Pr} \quad (5)$$

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \quad (6)$$

The precision $(Pr)$ in equation (3) is the number of true positives $(TP)$ over the number of true positives $(TP)$ plus the number of false positives$(FP)$. Recall $(Re)$ in equation (4) is the number of true positives $(TP)$ over the number of true positives $(TP)$ plus the number of false negatives$(FN)$. Precision $(Pr)$ is a positive predictive value to decide the model confidence level. Meanwhile, recall is a measure of completeness of the results. Model accuracy uses precision and recall to explain accurately. F-measure or balanced F-score in equation (5) is the harmonic mean of precision and recall. The accuracy of equation (6) is the range of proximity to the actual value.

### 4.3. Text feature extraction

Text feature extraction was performed on reviews in this study. The purpose of text feature extraction is to understand the topic within the document. Text features were also extracted using topic modeling. Topic modeling is an unsupervised method of classification of documents, similar to the grouping methods in numerical data. The goal is to find a natural group despite searching without confidence. The technique for topic modeling used is LDA

LDA is a popular method with a generative process for defining topic models. This technique treats each document as a frequent topic and each topic as a word combination. Thus, documents may overlap in content and not be separated as discrete groups.

Words are the basic unit from a document in LDA A word is an element of the vocabulary arranged as a vocabulary vector $\{1, \dots, V\}$. The word is represented as the base of the vector unit, whose value of each part is 0 unless the corresponding word has a value of 1.

LDA is a generative probabilistic model of a corpus. LDA represents the document as a random combination of potential topics, and the distribution of words poses an issue. The document $\boldsymbol{w}$ used the following steps to find words as a topic member.

1. Determine document size, N referring to Poisson, $\xi$

$$N \sim Poisson(\xi) \quad (7)$$

2. Determine the topic distribution $\theta$ with Dirichlet distribution, $\alpha$:

$$\theta \sim Dir(\alpha) \quad (8)$$

3. For every $N$ words, $w_n$
   a. Determine the topic $z_n$ based on the multinomial distribution obtained from $\theta$ in step (2)

$$z_n \sim Multinomial(\theta) \quad (9)$$

   b. A word $w_n$ defined from a multinomial distribution probability obtained from the $\beta$ and on condition $z_n$

$$w_n \sim p(w_n | z_n, \beta) \quad (10)$$

Thus, mixture topics $\theta$, range of issues $z$, and a set of words $w$ are used in equation (11) to calculate the joint distribution probability.

$$p(\theta, z, w | \alpha, \beta) = p(\theta | \alpha) \prod_{i=1}^{N} p(z_n | \theta)\, p(w_n | z_n, \beta) \quad (11)$$

Thus, $w$ is experimentally observed, and the latent variables are $\theta$ and$z$. Then, we used Bayesian inference to estimate the posterior density from $\theta$ and $z$ using the following equation:

$$p(\theta, z|w, \alpha, \beta) = \frac{p(\theta, z, w|\alpha, \beta)}{p(w|\alpha, \beta)} \quad (12)$$

## 5. EXPERIMENTAL RESULTS

We conducted topic modeling for extracting learners' experiences from course reviews in MOOCs. Before that, we performed several steps such as collecting data, exploring data, processing data, and predicting helpful reviews. We proposed helpful LDA as an improvement of topic modeling. We compared our helpful LDA with the basic LDA to measure the model performance. We used perplexity as the metric of performance.

### 5.1. Data preparation

We explored the collected data. We limited the subject to Programming based on the exploration result and selected the features. The features are course title, review id, review content, and the class of review (helpful review, unhelpful review, and unlabeled review).

The stage performed after data exploration was data checking. At this stage, this checking process was carried out among others: first, we checked for duplicate reviews and removed all duplicates, and second, we checked if the language used was not English.

The next step is the first processing stage of the review content. In this stage, the abbreviated words were converted into long words (e.g., do not, will not), symbols were translated into text, words were transformed into their basic forms (lemmatization), and then stop words were removed.

### 5.2. Predicting reviews' helpfulness

At this stage, pre-processing data is used to construct a category review classification model [19]. The review classification was necessary because useful category reviews have information that is specific to the learner experience. However, in the data collected, only a few reviews were labeled by category.

The review classification model was built using the Naive Bayes algorithm because the algorithm performs well. However, based on our experiments, the Naive Bayes model's performance has not fulfilled our goals for accuracy in classifying reviews that help with low levels of classification errors. Therefore, we developed the Naive Bayes model by adding sentiment analysis [20]. The model showed good performance with a more moderate error rate [20].

### 5.3. Understanding the experiences of learners with LDA

We used topic modeling to discover and understand learner experiences through LDA. However, we need more analysis to confirm the topics' quality. Figure 5 shows two parts of the LDA modeling process. The first part is a text mining process, consisting of the term matrix and the document-term matrix. The second part is visualization. In the first part, reviews were transformed into two types of matrix data as required by the LDA model. Then, the LDA model built topics using the matrix data. Later, a mixture of words from each topic was visualized.

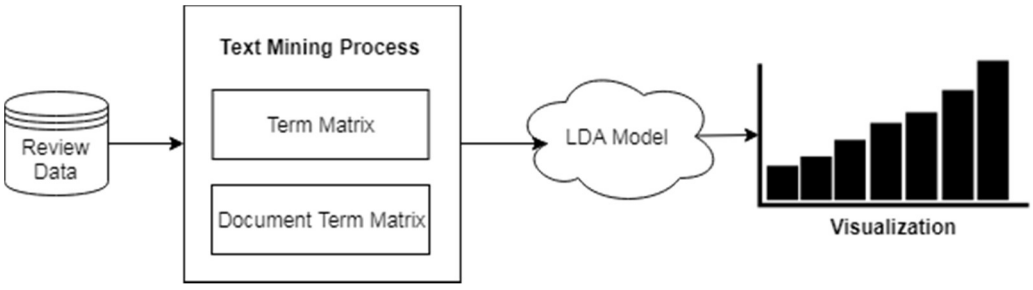



Figure 5. The topic modeling diagram with LDA

The first step of topic modeling is determined the number of topics, $k$. This experiment used $k = 2, 4, 10, 20, 50, \text{and } 100$. The goal was to find $k$ with the lowest perplexity score. The topics generated from LDA model was shown by topic visualization. The topic visualizations have shown the word topic probability distribution and the topic probability distribution because the goals are to discover the learner's experience from the reviews by understanding word topic usage with probability distribution only. Figure 6 indicates the word-topic distribution with $k = 2$, whereas Figure 7 shows the word-topic distribution with $k = 10$.

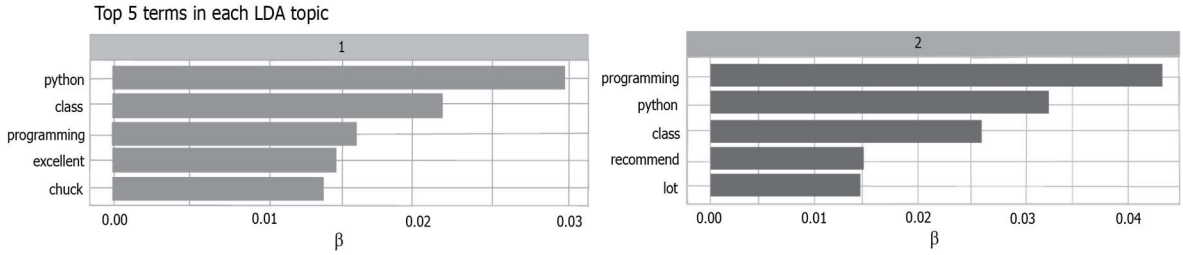



Figure 6. Visualization of word-topic probability from LDA model with number of topics $k = 2$

Based on the word topic distribution in Figure 6, we obtained an overview of each topic as follows:

1. Topic 1: python class programming excellent chuck (instructor)
2. Topic 2: programming python class recommend lot

Meanwhile, Figure 7 with ten topics provided a topic overview from learner experiences, as follows:

1. Python class fun programming learning
2. Programming python class experience fun
3. Python experience easy doctor fun
4. Programming learn python easy recommend
5. Programming class learn excellent lot
6. Lot python fun recommend learn
7. Programming learn excellent lot class
8. Python chuck (instructor) class learning doctor
9. Python programming excellent recommend fun
10. Class programming learn concepts chuck (instructor).

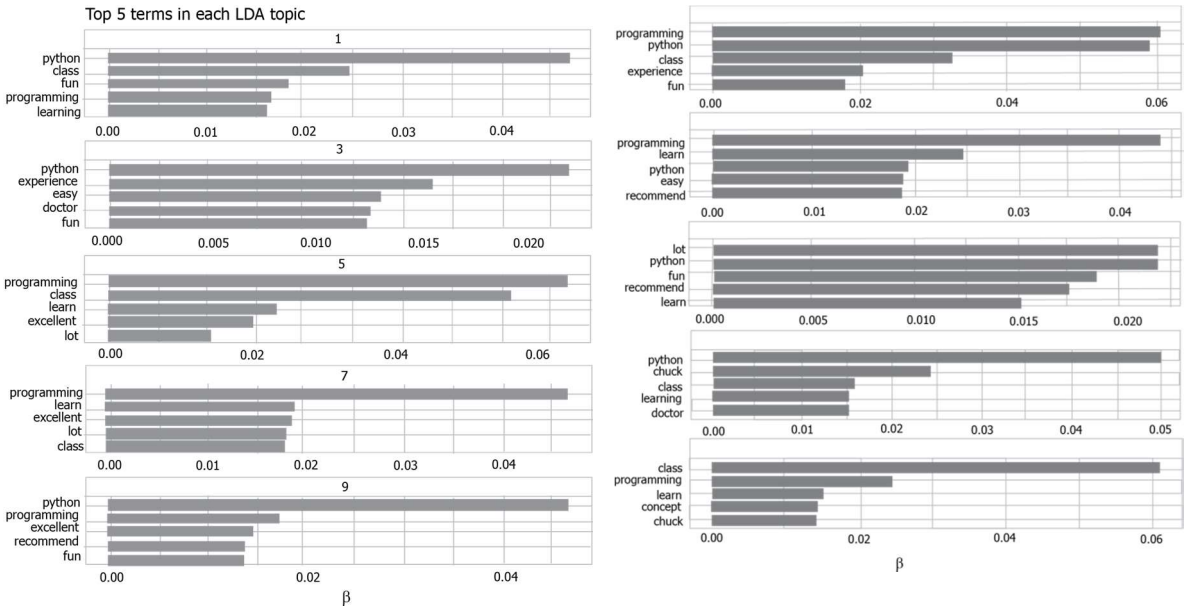



Figure 7. Visualization of word-topic probability from LDA model with number of topics $k = 10$

The visualization of two topics and display of ten topics showed that some topics need prior knowledge to decide the topic name. For example, the word "Chuck" in word-topic distribution with two topics means the instructor name. Thus, analyzing the LDA model needs prior knowledge related to the course. The visualization of two topics and ten topics has ambiguous interpretation. For example, on a

probabilistic version the word-topics, the word “chuck” was found in two topics; Chuck is the name of the course instructor. Thus, the LDA model has required prior knowledge related to the course to interpret this case. Meanwhile, the word-topic distribution with ten topics among the topics found a similar interpretation. Thus, the number of issues affected the topic clarity interpretation. Increasing the topic numbers also reduced the clarity. Topic modeling research from Fang et al. used perplexity as a qualitative evaluation criterion [21].

Perplexity is a statistical measurement that aims to measure the model’s ability to predict the sample. The perplexity score describes the unseen data generalization. A low perplexity score indicates the model’s generalization ability. The formula to calculate the perplexity in test documents is as follows:

$$perplexity(D_{test}) = exp - \frac{\sum_{d=1}^{|D_{test}|} log(p(o_d))}{\sum_{d=1}^{|D_{test}|} N_{o,d}} \quad (13)$$

$$p(o_d) = \prod_{i=1}^{N_{o,d}} \sum_{k=1}^{K} p(o_i|z_i = k)\, p(z_i = k|d) \quad (14)$$

where $o_d$ is a set of words that appear in the test document $d$, while $p(o_i|z_i = k)$ is the probability learned during the training process, and $p(z_i = k|d)$ was concluded from the Sampling Gibbs process against the test data based on the observed parameters of the training data. We performed a perplexity test on the model with the number of topics, $k = 2, 4, 10, 20, 50, \text{and } 100$. Figure 8 shows the perplexity of the learners’ review data by the number of topics, $k = 2, 4, 10, 20, 50, \text{and } 100$. Based on the plot in Figure 8, we see that the LDA model achieved the minimum perplexity score with 50 topics, while 100 topics obtained the second lowest position.

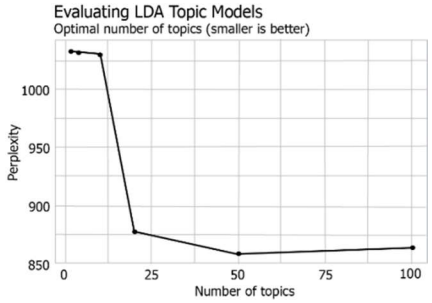



Figure 8. LDA model’s perplexity value

### 5.4. Understanding the learner experience with helpful LDA

The most helpful reviews refer to the research of Li et al. [22], manifested as a credible source perceived by the voter based on content, product or item information related to the rating. Based on the research of Li et al., we developed the LDA model by adding helpful reviews. The goal is finding more specific learners' experience topics through visualization of word-probabilistic topics and enhancing the model’s capabilities, as shown by the lower perplexity score compared with the LDA model. Figure 9 shows a helpful LDA flow diagram.

Based on the flow diagram in Figure 9 we proposed a review feature to filter the reviews to be more specific from the user side. Then, the filtered data based on the helpful review feature are filtered back in every sentence with a sentiment analysis. The goal is to obtain subjective sentences. The flow diagram of the study of sentiments on the helpful reviews is shown in Figure 10. Detailed analysis steps according to Figure 10 are as follows. First, we separated paragraphs into sentences. Second, sentiment analysis was performed with Lexicon Bing to categorize sentences as belonging to the positive, negative, or neutral categories. The third step was to filter out the phrases with positive and negative groups. This was done because the positive and negative classes contained subjective learner experience.

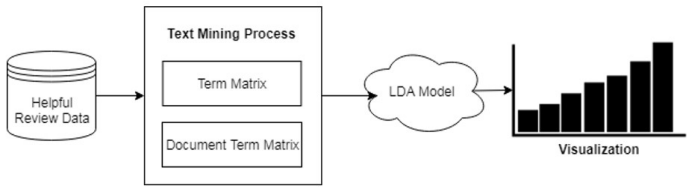



Figure 9. The topic modeling diagram with Helpful LDA

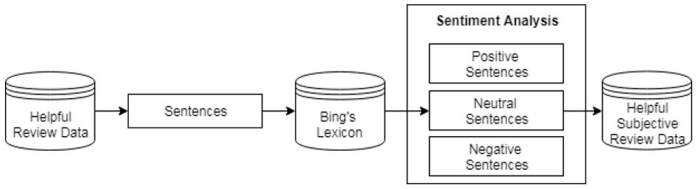



Figure 10. Flow diagram of sentiment analysis on helpful reviews

Figure 9 shows the process performed after filtering the data, such as tokenization, term matrix, document-term matrices, and LDA modeling. Then, the next stage is the visualization and performance measurement of the helpful LDA model with the number of topics, $k = 2, 4, 10, 20, 50, \text{and } 100$. The output of the word-term probability is shown in Figures 11 and 12 by the number of topics, $k = 2$ and $k = 10$.

Figure 11 shows the probability of word-topics with the number of topics $k = 2$ that provide learner experience information related to the course. Topic 1 concerns the class situation. Then, topic 2 addresses the activity recommendation to learn in the course. The word distribution with helpful LDA gives an overview of every clear topic as in the LDA model. Thus, interpretation of the experience of learners can be done without the need for early knowledge related to the course, such as the name of the instructor.

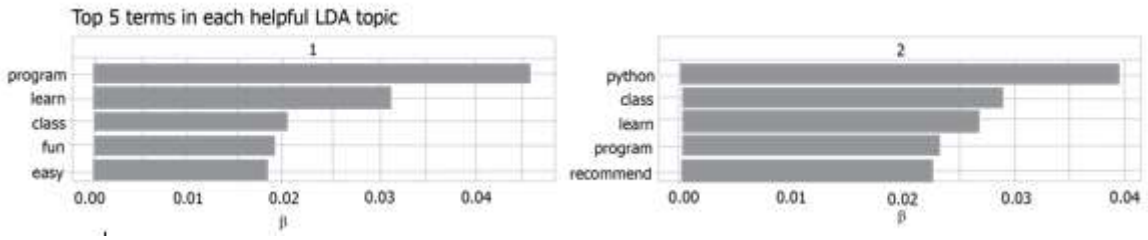



Figure 11. Visualization of word-topic probability helpful LDA model with number of topics $k = 2$

Figure 12 shows a visualization of word-topic probability with the number of topics $k = 10$. With the helpful LDA model, the topics are formed as follows:

1. Program learn easy excellent class
2. Python learn class easy lot
3. Program fun learn lot easy
4. Learn class python time game
5. Program excellent learn instructor recommend
6. Recommend python program fun teach
7. Fun program teach python learn
8. Program learn python class teacher
9. Class easy code python code excellent
10. Python fun teach video recommend

We observed the topic structure of this model take a slightly different form in the LDA model with the same number of topics, $k = 10$. Additionally, we identified from the topics three topics that required prior knowledge to interpret the title of a topic related to the instructor's name in the word-topic distribution visualization, such as "Chuck."

Figure 13 shows the helpful LDA model's perplexity value for topics number, $k = 2, 4, 10, 20, 50, \text{and } 100$. Based on that figure, we observed that topic $k = 100$ has the lowest perplexity value. Based on the perplexity value shows in Figure 14, helpful LDA has better performance than LDA. The model with the lowest perplexity is generally considered the "best". The "best" means getting even more specific topics now.

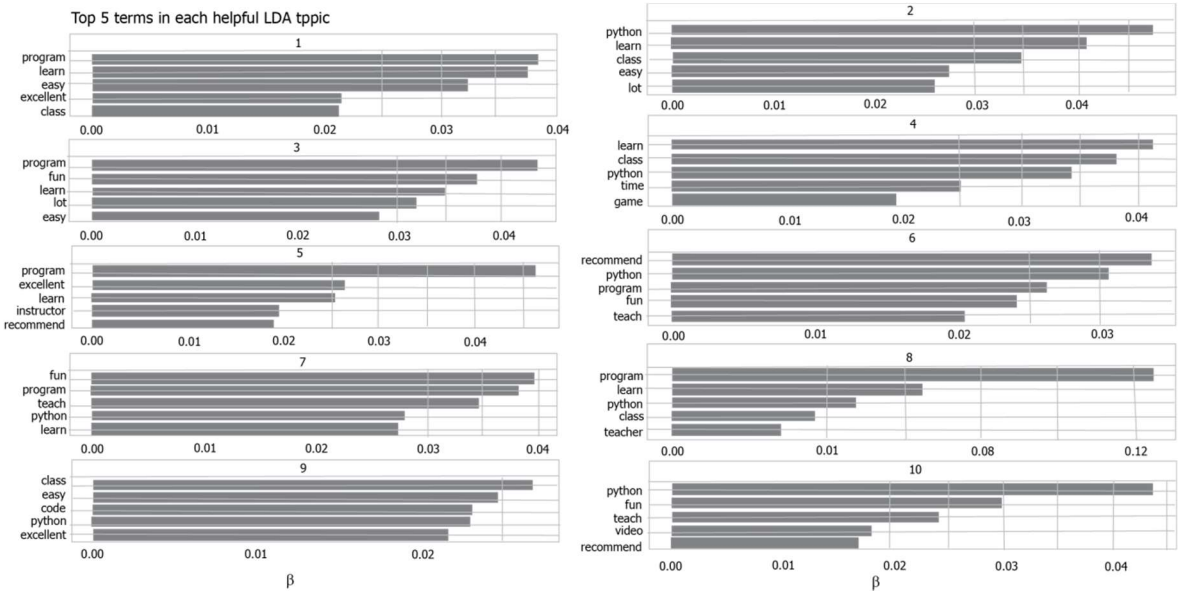



Figure 12. Visualization of word-topic probability helpful LDA model with number of topics $k = 10$

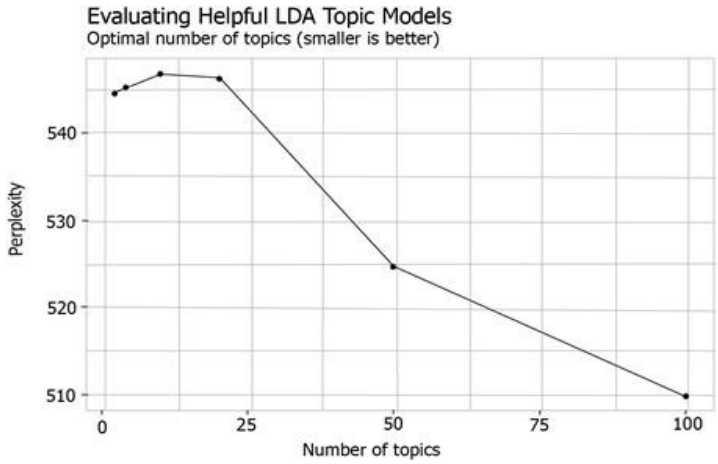



Figure 13. Helpful LDA model's perplexity value

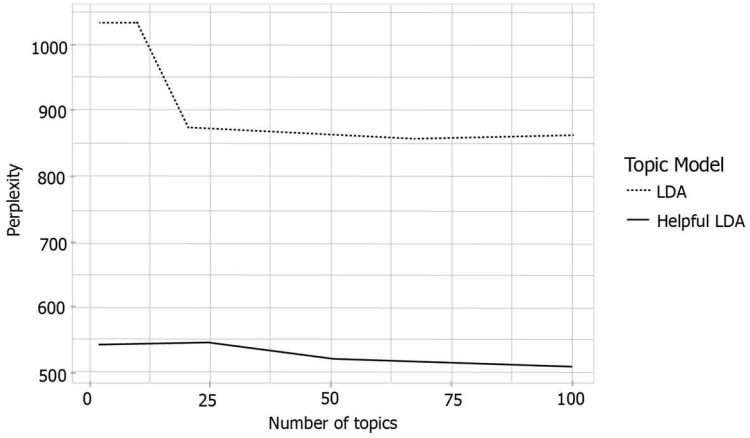



Figure 14. Perplexity value between LDA and helpful LDA

## 6. CONCLUSIONS

In this paper, we have developed topic modeling with LDA to understand the course experience of learners on the MOOCs platform. Our first focus is on understanding the factors that influence the teaching-learning process through topic modeling using the LDA method. However, based on research results obtained, the approach still need prior knowledge related to course. So, we developed LDA model by adding sentences filtering through helpful subjective reviews and sentiment analysis. The results show that the proposed method of reducing the prior knowledge.

We used perplexity to compare words distributions represented by the topics. The result shows that the proposed method can be decreasing the perplexity score compare to LDA method which is the lowest perplexity is considered the "best". In the future, we will add model performance metrics, such as inter-word coherence in topic and inter-topic coherence. Additionally, we will use the topics of user experience as the basis for building a recommendation system.

## ACKNOWLEDGEMENTS

We thank BlackBerry Innovation Center, Institut Teknologi Bandung (ITB), Indonesia, and Smart City & Community Innovation Center (SCCIC) laboratory, Institut Teknologi Bandung, for the support.

## REFERENCSES

[1] Ye Q, Law R, Gu B. "The Impact of Online User Reviews on Hotel Room Sales," *Int J Hosp Manag,* 28: 180–182, 2009.
[2] Fauzi MA, Nur F, Afirianto T. "Improving Sentiment Analysis of Short Informal Indonesian Product Reviews using Synonym Based Feature Expansion," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control),* 16: 1345–1350, 2018.
[3] Suleman K, Vechtomova O. "Discovering Aspects of Online Consumer Reviews," *J Inf Sci* 2016; 42: 492–506.
[4] Candra S, Putrama IK. "Applied Healthcare Knowledge Management for Hospital in Clinical Aspect," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control),* 16: 1760–1770, 2018.
[5] Blei DM, Ng AY, Jordan MI. "Latent Dirichlet Allocation," *J Mach Learn Res*; 3: 993–1022, 2003.
[6] Ezen-can A, Boyer KE, Kellogg S, et al. "Unsupervised Modeling for Understanding MOOC Discussion Forums : A Learning Analytics Approach Categories and Subject Descriptors," In: *LAK '15 Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. ACM Press, pp. 146–150, 2015.
[7] Liu H, Chen H, Lin M, et al. "Community Detection Based on Topic Distance in Social Tagging Networks," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control),* 12: 4038–4049, 2014.
[8] Atapattu T, Falkner K, Tarmazdi H. "Topic-wise Classification of MOOC Discussions : A Visual Analytics Approach.," In: *Proceedings of the 9th International Conference on Educational Data Mining*, pp. 276–281, 2016.
[9] Peng X, Liu S, Liu Z, et al. "Mining Learners' Topic Interests in Course Reviews based on Like-LDA Model," *Int J Innov Comput Inf Control*, 12, pp. 2099–2110, 2016.
[10] Ngo-Ye, L. T, Sinha AP. "Analyzing Online Review Helpfulness Using a Regressional ReliefF-Enhanced Text Mining Method," *ACMTransactions Manag Inf Syst,* 3: 10:1--10:20, 2012.
[11] Gilbert E, Karahalios K. "Understanding Deja Reviewers," In: *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10*. Savannah, Georgia, USA: ACM, pp. 225--228, 2010.
[12] Song Y, Pan S, Liu S, et al. "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: ACM, pp. 1757–1760.
[13] Pyo S, Kim E, Kim M, et al. "LDA-Based Unified Topic Modeling for Similar TV User Grouping and TV Program Recommendation," *IEEE Trans Cybern*, 45 1476–1490, 2015.
[14] Zhang Z, Varadarajan B. "Utility Scoring of Product Reviews," In: *Proceedings of the 15th ACM International Conference on Information and Knowledge Management*. Arlington, Virginia, USA: ACM, pp. 51–57.
[15] Zhang Y, Zhang D. "Automatically Predicting the Helpfulness of Online Reviews," In: *Information Reuse and Integration (IRI)*. Redwood City, CA, USA, pp. 662–668, 2014.
[16] Hu Y, Chen K. "Predicting Hotel Review Helpfulness: The Impact of Review Visibility, and Interaction between Hotel Stars and Review Ratings," *Int J Inf Manage*, 36: 929–944, 2016.
[17] Androutsopoulos I, Koutsias J, Chandrinos K V, et al. "An Evaluation of Naive Bayesian Anti-Spam Filtering," In: *Proceedings of the workshop on Machine Learning in the New Information Age*. Barcelona, Spain, pp. 9–17, 2000.
[18] Ding W, Yu S, Wang Q, et al. "A Novel Naive Bayesian Text Classifier," In: *Information Processing (ISIP), 2008 International Symposiums on*. Moscow, Russia, pp. 78–82, 2008.
[19] Shah S, Kumar K, Saravanaguru RK. "Sentimental Analysis of Twitter Data Using Classifier Algorithms," *Int J Electr Comput Eng*; 6, Epub ahead of print, DOI: 10.11591/ijece.v6i1.8982, 2016.
[20] Lubis FF, Rosmansyah Y, Supangkat SH. "Improving Course Review Helpfulness Prediction Through Sentiment Analysis," In: *2017 The International Conference on ICT for Smart Society (ICISS)*. Tangerang: IEEE, 2017.
[21] Fang Y, Si L, Somasundaram N, et al. "Mining Contrastive Opinions on Political Texts using Cross-Perspective Topic Model Categories and Subject Descriptors," In: *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*. ACM Press, pp. 63--72.
[22] Li M, Huang L, Tan C-H, et al. "Helpfulness of Online Product Reviews as Seen by Consumers: Source and Content Features," *Int J Electron Commer,* 17: 101–136, 2016.

BIBLIOGRAPHY OF AUTHORS



Received a master's degree from the School of Electrical Engineering and Informatics Engineering, Institut Teknologi Bandung, Indonesia, in 2008. She is working toward a doctoral degree in the School of Electrical Engineering and Informatics Engineering at Institut Teknologi Bandung, Indonesia. She is also a research assistant in Smart City and Community Innovation Center (SCCIC) laboratory at the Institut Teknologi Bandung, Indonesia. Her primary areas of interest include recommenders systems for technology-enhanced learning, learning analytics, and educational data mining.



An Associate Professor in School of Electrical Engineering and Informatics Engineering at Institut Teknologi Bandung, West Java, Indonesia. He received a master's and doctoral degree from the University of Surrey, United Kingdom. His research interests include learning technology, learning analytics, mobile learning, computer-supported collaborative learning, and ubiquitous learning environments.



A Professor in the School of Electrical Engineering and Informatics Engineering at the Institut Teknologi Bandung, West Java, Indonesia. He received a master's degree from Meisei University, Japan, and a doctoral degree from the Uni-versity of Electro-Communications, Tokyo, Japan, in 1998. His research interests include smart cities, smart education, smart health, smart energy, and smart mobility.