JISKA (Jurnal Informatika Sunan Kalijaga)
Vol. 7, No. 3, SEPTEMBER, 2022, Pp. 177 – 191
ISSN: 2527 – 5836 (print) | 2528 – 0074 (online)

Summarizing Online Customer Review using Topic Modeling and Sentiment Analysis

Muhammad Rifqi Maarif
Teknik Mesin, Fakultas Teknik, Universitas Tidar, Magelang
e-mail : rifqi@untidar.ac.id.
Artikel ini diajukan 31 Juli 2022, direvisi 5 September 2022, diterima 5 September 2022, dan dipublikasikan 25 September 2022.

Abstract

With the massive implementation of social media in various forms in various business domains, business or product owners have the opportunity to be able to take advantage of user review data that is available free of charge to evaluate the products they issue. User reviews on social media platforms, marketplaces, and e-commerce are User Generated Content (UGC) which is very useful for product owners to find out the extent of user preferences for their products. However, to be able to comprehensively read the data, the right technology is needed considering that the data is in the form of text in very large quantities. Reading one by one and then drawing conclusions is certainly not the right approach because it will take quite a lot of time. So, in this study, the researcher will use a text analysis-based approach, especially topic modeling and sentiment analysis to summarize user reviews in the comments or reviews column on the e-commerce platform. The case study used in this study is user reviews in the comments column on the Amazon site for the Lenovo K8 Note smartphone product. From the experiments carried out, the approach used can summarize the reviews written by quite many users in one summary that can be easily understood.

Keywords: User Generated Content, Online Customer Review, Text Mining, Topic Modeling, Sentiment Analysis

Abstrak

Dengan masifnya implementasi media sosial dalam berbagai bentuk di berbagai domain bisnis, pemilik bisnis ataupun produk memiliki kesempatan untuk dapat memanfaatkan data ulasan pengguna yang tersedia secara gratis untuk mengevaluasi produk yang mereka keluarkan. Ulasan pengguna di platform media sosial, *marketplace* maupun *e-commerce* merupakan *User Generated Content* (UGC) yang sangat bermanfaat bagi pemilik produk untuk mengetahui sejauh mana preferensi pengguna terhadap produk mereka. Namun, untuk dapat secara komprehensif membaca data tersebut diperlukan teknologi yang tepat mengingat data tersebut berbentuk tekstual dengan jumlah yang sangat banyak. Membaca satu per satu untuk kemudian disimpulkan tentunya bukan pendekatan yang tepat karena akan memakan cukup banyak waktu. Sehingga, dalam penelitian ini, peneliti akan menggunakan pendekatan berbasis analisis teks khususnya pemodelan topik dan analisis sentimen untuk merangkum ulasan pengguna di kolom komentar atau *review* yang terdapat di platform *e-commerce*. Studi kasus yang digunakan dalam penelitian ini adalah ulasan pengguna pada kolom komentar di situs Amazon untuk produk *smartphone* Lenovo K8 Note. Dari percobaan yang dilakukan, pendekatan yang digunakan mampu merangkum ulasan-ulasan yang dituliskan oleh pengguna yang berjumlah cukup banyak dalam satu ringkasan yang dapat dengan mudah dipahami.

Kata Kunci: *User Generated Content*, Review Pengguna Online, Penambangan Teks, Pemodelan Topik, Analisis Sentimen

1. INTRODUCTION

The ultimate goal of product development is to deliver a product that meets the customer's expectations. Hence, considering customer reviews of the product they bought and used is a key to successful product development. Integrating customer perception or opinion could be done in various ways and stages, from an early stage of product development until product release (Cui & Wu, 2017; Zhan et al., 2019). In the early stage of product development, the customer could act as a source person for identifying and validating product requirements from their perspectives. In the middle stage of the product development cycle, the customer can also give their opinion on the concept of the product before it goes to production. In another way, the product owner can use the customers as active informants by constantly monitoring their opinion on online platforms like social media or the review section of online marketplace or e-commerce. Nowadays, by using online media, the customer is often shared their experience opinion, suggestion, and even complaint about the product they have bought and used. Those kinds of customer reviews could be potentially a source of ideas for a product owner to develop their product to the next version or even create new products (Hidayanti et al., 2018; Wang et al., 2020). With the advanced penetration of the internet, social media platforms including review sections on the online marketplace and e-commerce have changed the way customers interact and share their thoughts about the product they used. Specifically, Bashir et al. (2017) underlined that in correlation with product development, social media can be used to (1) identify user requirements and measure their review of the product, (2) recognize vocal customers that could be converted into a lead, (3) analyzing brand engagement within a certain group of society and (4) identifying the ideas for further development of a certain product.

User reviews of certain products on online platforms like social media or e-commerce are expressed in a very expressive and emotional way that makes product manufacturing companies have a more comprehensive review of products that have been marketed to the public (Elwalda et al., 2016). Thus, user comments and opinions on social media can be used as a basis for product evaluation as well as market analysis. In another study, online social media can also be used to identify potential customers by analyzing personal characteristics and their environments like friendship networks, demographic information, and other personal preference (Ramanathan et al., 2017). Furthermore, the use of social media data to obtain the voice of customers, apart from being more economical because we can get the review data for free by using the scraping technique, is also more accurate because it is sourced from a very large number of uploads from users who have very diverse characteristics (Elena, 2016).

Recently, research on social media analysis for summarizing customer reviews has been carried out with several approaches/techniques, one of which is text mining (Mahr et al., 2019). Text mining is a sub-type of data learning for textual data, so it is very suitable for analyzing customer uploaded data on social media, which is mostly in the form of text, in addition to images and videos. Techniques in text mining that are quite widely used are sentiment analysis and topic modeling. Sentiment analysis is a supervised computational technique by utilizes machine learning algorithms such as Support Vector Machine (SVM), Artificial Neural Network (ANN), Deep Learning, and so on to extract opinion polarity from text data (Baharudin et al., 2010). The polarity of opinion is the tendency of sentiment contained in a sentence from the simplest (positive, negative, and neutral) to the more complex (happy, enjoy, sad, angry, and so on). Several studies that analyze social media for the voice of customers in product development use sentiment analysis to analyze customer perceptions on social media towards a certain product brand (Hu et al., 2020). Another study was conducted by Greco & Polli (2020) to read people's perceptions of the release of a new product. Sentiment analysis is also used to evaluate customer reviews in general for products that have been sold and used (Mirtalaie et al., 2018; Ng et al., 2021).

Another text mining approach that is widely used for customer review analysis on social media is topic modeling (Ko et al., 2018). Unlike supervised sentiment analysis, topic modeling is a machine learning approach that is unsupervised. The results of topic modeling are keywords that represent the topics/themes contained in a text. Thus, in the voice of the customer, topic modeling is widely used to explore the most frequent and influential topics discussed by customers (Jeong et al., 2019). Some examples of the voice of customer research that utilize topic modeling include identifying product features that get the most reviews by customers (Irawan et al., 2020). Further implementation of text mining is to combine sentiment analysis and topic modeling into a more integrative approach. This approach uses sentiment analysis to evaluate customer opinions for a specific product feature. In this study, topic modeling was used to identify product features that were reviewed in the text of customer uploads on social media (Chehal et al., 2021).

Intensive monitoring of social media to analyze public opinion as users of industrial products will provide great benefits to find out the weaknesses of products that have been widely used by the public quickly and accurately, Hence, this research aims to produce a systematic and measurable summary by integrating text mining techniques for user review analysis of social media data.

2. METHODS

The proposed approach in this article is based on two theoretical backgrounds of text mining named topic modeling and sentiment analysis. First, the topic modeling approach was used for identifying the topics discussed by the user through their reviews on Amazon. The identified topics were further processed to select only topics which represent the user review statements. Second, the review value of users on each topic is measured by using a sentiment analysis approach. In this section, we will first describe the computational text mining approach by theoretically explaining topic modeling and sentiment analysis followed by an explanation of the step-by-step approach of gathering and processing user reviews for determining user review aspects and values.

2.1 Text Mining

Text mining is a transformation process of extracting structured data from unstructured text. Once extracted, those data can be further analyzed with various data analytic approaches to produce a meaningful pattern and facts. The process of text analysis employs a range of Natural Language Processing (NLP) methods. Text mining was commonly used alongside data mining techniques in which Naive Bayes, Support Vector Machine (SVM), and Artificial Neural Network (ANN) based methods including well-known deep learning were the most popular data mining techniques used in text mining. Employing advanced data mining techniques like deep learning in text processing works enables the computer to recognize any kind of insights hidden beneath the huge amount of textual documents (Suresh & Harshni, 2017).

The text mining process consists of several consecutive steps for deducting structured text data from the unstructured. The first phase of text mining is text preprocessing. In this phase, the collection text document so-called corpus would be cleaned and transformed into a standardized format. Text preprocessing is considered the phase of text mining that has the biggest portion of overall NLP activities (Vijayarani et al., 2018). This phase involves several technical works such as language identification, Part of Speech (POS) tagging, tokenization, parsing, stop word and non-standard character removal, and many more. The kind of preprocessing work was currently developing and even employees advance machine learning techniques. After the text is preprocessed, the next task is to apply text mining algorithms such as text summarization, topic modeling, text classification, sentiment analysis, named entity recognition, and many more (Maheswari; & Sathiaseelan, 2017). This research employed topic modeling and sentiment analysis consecutively for summarizing user review aspects from user reviews on Amazon.

2.2 Topic Modeling

Topic modeling is a common and widely adopted approach in text mining and machine learning for discovering semantic structures hidden within a collection of textual documents. In natural language processing, a topic model was defined as a statistical-based model for revealing a set of abstract representations of document collections. Those abstracts represent then further named as topics. In the topic model, topics are represented as a cluster of statistically similar words. Hence, a topic model discovers the topics of a set of documents based on the statistics of the words in each part of the document and estimates what the topics might be. There is some approach to topic modeling implementation (Alghamdi & Alfalqi, 2015). In our experiment, Latent Dirichlet Allocation (LDA) was used for implementing topic modeling. The use of LDA is based on the literature reviews that stated that LDA is a topic modeling algorithm that has the highest performance, specifically when operating over a large number of textual documents (Barde & Bainwad, 2017).

LDA is a probabilistic approach proposed by (Blei et al., 2003). This approach consists of three different phases which are performed sequentially. The basic idea is LDA represents a document as a set of topics. Then, each topic was represented as probability distributions containing keywords. The architecture of LDA also shows the consecutive steps involved depicted in figure 1. LDA operates on a corpus **D** which consists of **M** number documents, each document on D denoted as ***w*** with length ***N***. In figure 1, **α** is defined as Dirichlet's parameter before the per-document topic distribution. Consecutively, **β** is a parameter of the Dirichlet that need to be defined to determine the per-topic word distribution. The next parameter is **θ** which denotes the topic distribution for the document. The sum of all **θ** is 1.0. Finally, LDA defined ***z*** as the topic for document ***w***.

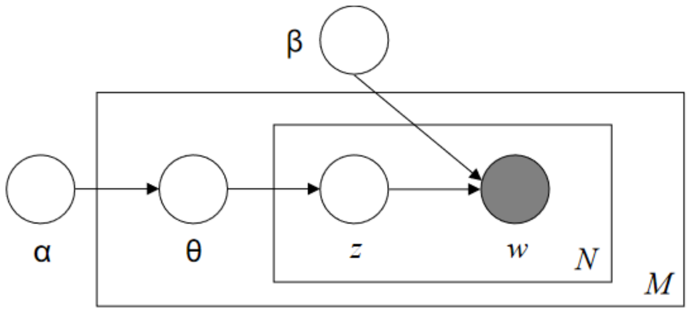

Figure 1 LDA Architecture (Blei et al., 2003)

When executing the LDA algorithm, **α** defined needs to be predefined manually as the expected number of generated topics. Then, assuming there is no prior knowledge on how many topics should be generated we can use the Elbow Method to determine the best number of topics in LDA-based topic modeling execution (Syakur et al., 2018).

2.3 Sentiment Analysis

Sentiment analysis is one of the major research fields and widely adopted techniques of NLP. It is a technique that mainly focused on recognizing the polarity of people's opinions based on their textual expressions. Sentiment analysis applies to a wide range of domains including product development which involves analyzing user perspectives. Greco & Polli (2020) performed an experiment by using sentiment analysis for measuring the popularity and user engagement of certain brands. From a technical perspective, sentiment analysis is considered a classification problem over textual data with its sentiment acting as a class. In basic form, the sentiment was categorized as positive and negative which showed the polarity of the opinions of the speaker/writer of the text data (Ahmad et al., 2017). The sentiment of a certain text written by the speaker basically could be identified in two approaches lexicon-based or text classification based.

The lexicon-based approach employed a dictionary that is already defined and contains a sentiment label with its corresponding sentiment score. One of the famous and widely used predefined dictionaries is SentiWordNet (Baccianella et al., 2010). In this approach, a sentiment of a text is identified by calculating the polarity orientation score of words or phrases which construct the document one by one. The overall sentiment value for the text is then computed by summing up those polarity scores. The polarity score could be either categorical or continuous values. When the polarity score is a categorical value, then the overall text sentiment is determined by the category label with the most appearance. When the polarity score preserves in continuous form, the overall polarity is calculated by summing up all of the scores. On the continuous polarity score, positive sentiment represents positive values, and negative sentiment represents negative values. Even though this approach yields an accurate result, constructing a dictionary for every language was considered a costly activity.

Different from the former approach, the text classification-based approach employs machine learning techniques instead of relying on a dictionary. The text classification-based sentiment analysis is a supervised machine learning task. Thus, the classifier model was built from text data labeled by their corresponding sentiments to predict the sentiment orientation (Ahmad et al., 2017). Instead of resulting sentiment in a numerical or ordinal value. The supervised approach of sentiment analysis assumes a sentiment as a discrete class label (positive or negative). Given the training data, typical machine learning algorithm like Support Vector Machine (SVM), Naive Bayes Classifier (NBC), Logistic Regression (LR), K-Nearest Neighbor (KNN), and Artificial Neural Network (ANN)-based techniques including deep learning learns from the data and build a classification model with their representations. Then, when new text is coming this classification model was employed to predict its sentiment class.

Nowadays, with the advanced research of text analytics in the industry, there are plenty of ready-to-use cloud-based analytic services to perform sentiment analysis. Those kinds of services provide a pre-trained model which builds from a large amount of textual data. Hence, those services offer high-performance sentiment analysis with no programming works on the user side. One of the most popular analytic services to perform sentiment analysis is Cloud NLP provided by Google. In this article, we used the Cloud NLP service provided by Google to perform sentiment analysis. Google NLP service use BERT (Bidirectional Encoder Representations from Transformers) algorithm for performing sentiment analysis task. BERT is a deep learning-based algorithm which particularly designed for text analytics and NLP (Sun et al., 2019). Currently, BERT gives a state-of-the-art performance in text classification among the other machine learning or deep learning techniques.

2.4 Experimental Setup

In this section, we describe the step-by-step approach in our experiment to discover the user review aspect and its corresponding sentiment expressed by the user through their comments on Google Playstore. After the data was gathered and preprocessed, we first employed topic modeling techniques to identify the key topics talked about by the users. Those key topics were then further processed and eventually formed user review aspects. Then, after the aspects were constructed, the sentiment profile of each topic was estimated by using the sentiment analysis technique. A supervised machine learning-based approach was implemented to build the classifier model for sentiment classification. Figure 2 below illustrates the consecutive process of our proposed approach.

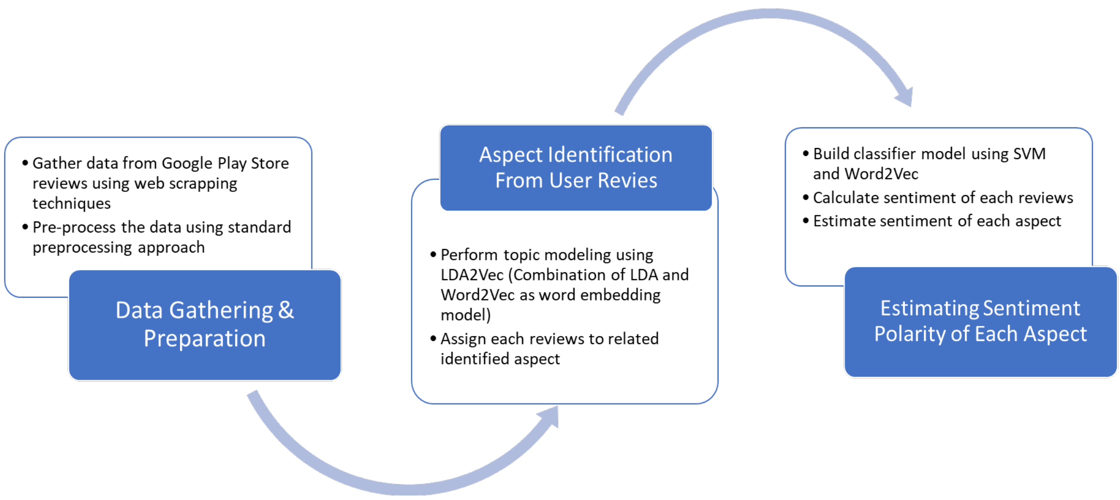

Figure 2 Proposed Framework from Identifying User Review Aspect and Its Sentiment

2.4.1 Data Preparation

The first phase of our approach is data collection. To pile up the summary of customer reviews, the data acquisition would be concentrated to acquire user reviews of one specific product named the Lenovo K80 smartphone which is available on Amazon. In this experiment, we collect data from user comments from the product review section on the e-commerce platform Amazon. Figure 3 shows the review section on a product page on Amazon. In that section, user can freely express their opinion and experience of the corresponding product they bought.

Figure 3 The Customer Reviews Section on Amazon

The next step in this phase is to preprocess the data. The data preprocessing consists of several common steps including the removal of stopwords, punctuation, and non-standard character followed by a steaming process and lemmatization. The detailed processes of the preprocessing steps are outlined below:

1) **Case folding**. Case folding is a process in text preprocessing that is done to standardize the form of characters in the data. The case folding process is the process of converting all letters to lowercase. In this process, the characters 'A'-'Z' contained in the data are converted into characters 'a'-'z'.
2) **Stop-word removal**. Stop words are common words that usually appear in large numbers and are considered meaningless. Examples of stop words for English include "of", "the" and so on. Stopword Removal is a filtering process, selecting important words from the token beyond the stopwords to represent documents.
3) **Punctuation removal**. Punctuation removal is a process where the system will remove punctuation marks or symbols that exist in the dataset. These punctuation marks or symbols are removed because they do not affect the results of text analysis.
4) **Stemming**. Stemming is the process of forming basic words. The terms obtained from the stopword disposal stage will be carried out by the steaming process. Stemming is used to reduce the form of terms to avoid mismatches that can reduce recall, where terms that are different but have the same basic meaning are reduced to a single form.
5) **Lemmatization**. Lemmatization is a process that aims to normalize text/words based on the basic form which is the lemma form. The distinction between stemming and lemmatization is while stemming changes a word into a root word without knowing the context of the word like cutting off the ends of words, lemmatization changes a word into a root word by knowing the context of the word (Balakrishnan & Ethel, 2014).
6) **Part of speech tagging**. Part-of-speech (POS) tagging or in short it can be written as tagging is the process of assigning POS tags or syntactic classes to each word in the corpus (Gimpel et al., 2011). Part of speech is a classification of words that are categorized through their

role and function in sentences of a language. When studying English, we will come across the terms noun, adjective, pronoun, and so on. These terms are part of the part of speech. Part of speech has an important role to form a sentence so that it is coherent and follows the grammar of the sentence. In this experiment, we used POS Tagging to extract nouns from each user review in the text corpus.

Afterward, once a set of sentences represents the single user reviews of the target been prepared, keywords (or key phrases) are extracted from the sentence to construct the structure of the reviews. Eventually, each user review was represented as a list of keywords and their frequency of appearance in their related review.

2.4.2 Identifying User Review Aspect

In this step, the user review aspects are defined based on topic modeling which is computed using the LDA algorithm. LDA-based topic modeling calculation requires three parameters for its operation. Those three parameters are corpus, word dictionary, and the number of topics. User reviews gathered from Google Playstore would be the corpus parameter. Then dictionary parameter was produced automatically from the corpus. Dictionary is simply defined as a set of words that appears on the corpus regardless of the frequency of its appearance. From the constructed dictionary, LDA builds a bag-of-words model for representing the documents. A major problem of the bag-of-words representation comes from the inability to capture the semantic relation of a word within the word vector. Whereas, revealing the semantically related words and using the correlation value on the word vector would greatly improve the representation model (Xun et al., 2016). Hence, in this experiment word embedding was used for empowering the document representations.

In our experiments, we implemented LDA-based topic modeling by using LDA2vec to integrate word embedding representation through the Word2Vec model. LDA2Vec generates a context vector from a document vector built around the pivot word vector. The pivot word vector itself is a result of LDA expansion by the Word2Vec model over the document vectors (Bojanowski et al., 2017). The context vector is then used for estimating the context between words within vectors. That operation enables LDA to generate more human-interpretable LDA topics The set topics from user reviews generated by LDA2vec topic modeling become the major subjects by which users are expressing their experience of using the mobile learning applications. Those set of topics is then further processed for determining the appropriate terms and phrases which closely related to user satisfaction or dissatisfaction expressions. The final operation of this phase was to generate the user review aspects in which each aspect was represented by a set of keywords (terms and/or phrases).

2.4.3 Computing User Review Sentiment on Each Aspect

In this second step, a user review of each aspect that was identified in the previous stage was calculated. Sentiment analysis was employed as a basis for user review calculation. The first step in this phase was assigning every user review in the corpus into identified aspects. These steps are done by simply matching the keywords found in the user reviews with the keywords on every topic. Then, the sentiment of every sentence was estimated by using a classifier model constructed with BERT which was provided by the Google NLP service. Prior works showed that BERT gained a high-quality performance in terms of sentiment analysis and work well even though the training data was limited. Hence, we consider this machine learning approach suitable for our work. The BERT method provided by the Google NLP service was improved by employing word embedding as feature representation. Word embedding would empower feature representation with contextual information of words in high-dimensional vectors. Thus, the sentiment classification of each user review was improved in terms of accuracy. The second step in this phase involved calculating the sentiment of each identified aspect. This calculation is simply done by taking the percentage of positive and negative sentiment.

3. RESULTS AND DISCUSSION

In this section, the results of our experiment have been outlined in line with the scenario on methods section chronologically. We will start with a discussion about the dataset and exploratory analysis of our text corpus. Afterward, the discussion goes to the extraction of the user review aspect using topic modeling. Lastly, we will discuss the result of sentiment analysis on each aspect of user review.

3.1 Dataset and Exploratory Analysis

We used approximately 14.520 textual review data gathered from amazon.com. The dataset contains a user-generated review of the Lenovo K8 smartphone product. Lenovo K8 smartphone is one of the most popular budget “high-end” smartphones. Hence, there are a lot of user reviews of this device on Amazon. We use the web scrapping technique to acquire the dataset. The tool used in this experiment for scrapping purposes is Scrappy, an Open Source and Python-based web scrapping library freely available for any use. Table 1 shows the excerpt of our dataset in a row form.

Table 1 Example of an Acquired Customer Review of Lenovo K8 Smartphone on Amazon

No.	Review
1	Superb product. A few of the features are awesome. Dual camera, front 13mp camera, back and front flash, dedicated music button, dedicated memory card slot, free transparent case, and split window for multitasking. These are some features I like about the product in my budget
2	Hello, The phone starts resetting itself randomly. This issue starts after 3 days of use, especially when you use it for a longer time like watching a youtube video for an hour or two. The issue starts to repeat after 1 or 2 days, then I tried a few options mentioned in Lenovo help APP which is pre-installed(I am feeling stupid for having done this), and after that, for 4 or 5 days it didn't show this issue. Now it is back with a bang and almost resets itself every night. Use the phone at night, then lock it and keep it aside and when you take the phone again in the morning, boom, it is dead already. And when you manually power it on there is enough juice left in the battery. Not sure what to do with this kind of issue, the phone is useless. When I approached amazon for a return, they are quoting policy and suggested that I should run pillar to post to get the problem rectified (not sure even if it is the rectifiable issue),
3	Great experience with this amazing product from Lenovo. it is equipped with almost every feature that a smartphone required. Deca core processor long-lasting battery and 64GB internal memory with 4GB RAM is just awesome. I would certainly recommend this phone for users having usage and game lovers
4	Over Heating Issue while light usage like just an internet connection, and always warm while connected to the internet. this update is After usage of 10 days. Don't prefer this one.
5	It's the jack of all trades but the king of none. Battery backup could have been better if they used some other processor. The battery drains quite fast. The camera is better than average. And I think there is no option to keep external media as your ringtone. Only custom build ringtones available to set as your ringtone. Kinda bums me out. Update after 3-day use: Battery backup is horrible, normal usage like WhatsApp and Instagram browsing consume more than 25% battery in an hour or so. Then there is a turbocharging issue, it starts with fast charging then after 10-20 minutes, depending upon mood, the rate decreases, it took 7 hours to charge it by 40% in total. Would appreciate it if Amazon could take this matter seriously and take it up with Lenovo and return the money of its customers for defective models. I would not trust my 14K bucks with Lenovo or Moto from this point on.

The acquired data from Amazon is raw and not ready for further text analysis. Thus, we perform a sequence of preprocessing tasks before feeding the data into the topic model algorithm. We

employ the removal of stopwords, and punctuation standard characters followed by a steaming process and lemmatization as a sequence of preprocessing tasks in our experiment.
After preprocessing task, we then performed POS Tagging over the preprocessed dataset. This task aims to extract nouns from each user review. We consider that the core meaning of every review written by the user is in the nouns in the sentences. Hence, we can achieve a more precise topic model by neglecting another sentence structure beyond the noun forms.

We perform a simple exploratory analysis by generating a word cloud for early investigation of the text dataset. Word cloud is a visualization approach to reveal some of the most frequent words in a text corpus. Through the word cloud, we can gain a raw insight into what the customers frequently talk about the topic. This insight can be helpful for confirmation when a topic model was constructed. Figure 4 shows the word cloud of extracted nouns from our text corpus. From the word cloud of user reviews about the Lenovo K8 smartphone product, we know that two product features, namely camera, and battery, are the most frequent object talked about by the user. Later, by using topic modeling, these two main keywords and potentially some other keywords will be revealed further using the topic modeling approach.

Figure 4 Word Cloud of Extracted Nouns from the Text Corpus

The next exploratory text analysis performed in this experiment is to reveal the most frequent pair of nouns found on the corpus. Table 2 outlines the most frequent pair of nouns found in our text corpus. In line with the word cloud depicted in figure 4, the terms battery and camera remain dominant. By analyzing the most frequent pair we can further investigate the interrelation of most frequent words even further. For instance, from table 4 we can gain an insight that the user reviews about "camera" are predominantly about the "camera quality" or "camera performance".

Table 2 Most Frequent Topic Keyword Pair Within a Text Corpus

Rank	Topic Pair	Frequency
1	quality - camera	1110
2	battery - backup	662
3	performance - camera	545
4	quality - battery	512
5	problem - battery	499
6	battery - performance	470
7	issue - battery	437
8	battery - time	424
9	battery - hours	421
10	camera - product	416
11	mode - camera	405
12	life - battery	344
13	camera - problem	337
14	issue - camera	328
15	day - battery	325

3.2 Identified User Review Aspects

The aspect of every user review is considered as a particular topic contained in the sentences written by the users. Hence, topic modeling is the proper approach to reveal the aspects of the reviews generated by the users. In this experiment, we employ the LDA algorithm to extract the topic model from our text corpus. The LDA algorithm's main requirement is first to determine the number of topics we want to be extracted. Since we do not have prior knowledge of the data or domain, we employed the Elbow statistical approach to determine the optimal number of topics potentially generated by LDA.

In the Elbow method, the number of topics is considered optimal if the average cosine similarity between all pairs of distribution vectors of topic-words generated in topic modeling operation is minimum. Those cosine similarity values are then represented as coherence scores. The greater the coherence value, the more optimal the topic is. We then conduct some LDA algorithm experiments with various topics ranging from 4 to 10 topics. To ease the analysis process, we avoid determining the number of topics more than 10. Figure 5 shows the experimental results of determining the coherence score. From the graph in Figure 5, we can outline that the highest coherence score was achieved from our text corpus by determining the topic number as 5. Thus for the later tasks, we will use five topic clusters.

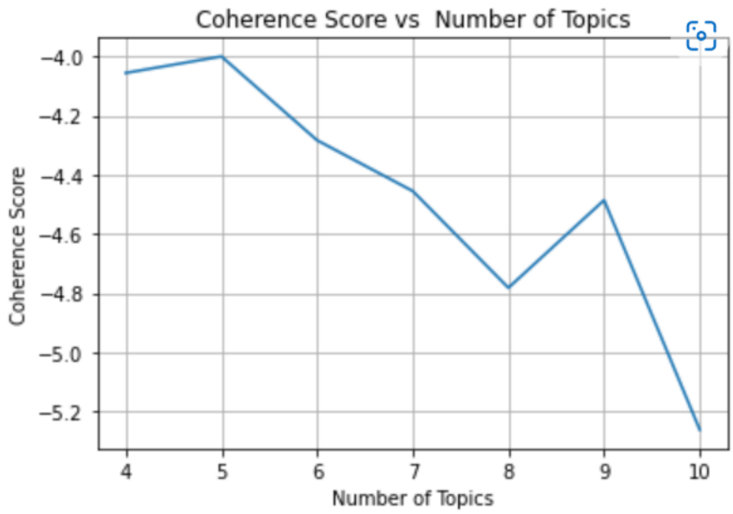

Figure 5 Coherence Score to Determine the Optimal Number of Topics

To investigate and validate whether the determined cluster number is optimum, we then visualize the cluster using an inter-topic distance map. To perform this task, we use Gensim. Gensim is an Open-Source library to perform topic modeling, including visualizing the topic clusters. We implement Gensim using Python language. Figure 6 shows the visualization of the topic cluster generated by the LDA algorithm with five clusters. The clusters are considered as good if they are well separated from each other. Hence, from Figure 6, we can see that the determined number of a cluster as five is considered an excellent or optimum number of clusters.

A cluster of topics contains a set of keywords that construct the topic model itself. Hence, after generating the topic clusters, we then reveal the set of keywords of every cluster and perform qualitative analysis to summarize those keywords into a specified label. The specified label determines the most closely related keywords within a cluster. Table 3 outline the keywords which dominantly appear on each cluster and their approximate labels. The labels outlined in table 3 were determined manually by looking at the semantic tendencies of the keywords contained in each cluster.

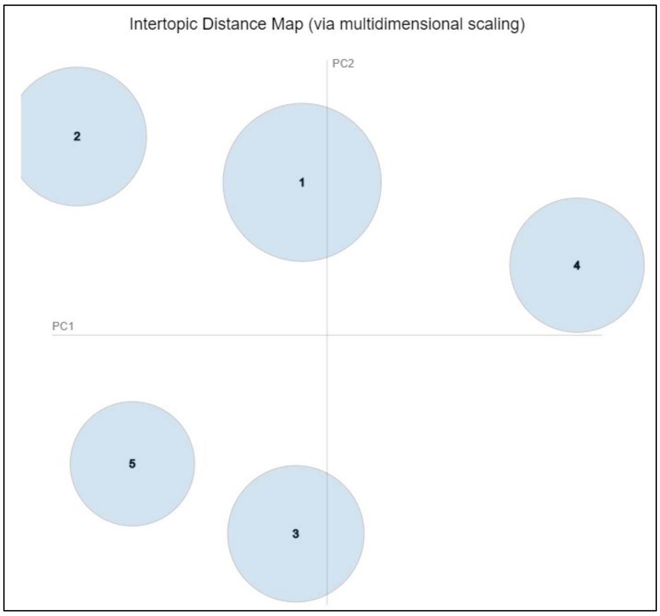

Figure 6 Topic Clusters Generated from LDA Operation

Table 3 The Keyword Generated by LDA and Its Corresponding Label

Topic Number	Keywords	Topic Label
1	camera, product, quality, performance, price, glass, range, awesome, display	camera and display quality
2	price, money, feature, screen, waste, value, option, phone	price money value
3	battery, camera, backup, quality, hour, day, life, mode	battery performance
4	the problem, issue, battery, month, heating, time, charger, heat, use	
5	product, service, amazon, call, device, return, buy, day	after-sales service

We can see in table 3 that from five topic clusters, we can derive four major topics. Topic 3 and topic 4 which shared similar semantics then be integrated as one. The four major topics derived from table 3 are "camera and display quality", "price money value", "battery performance", and "after-sales service". Those four topics are summarized as the major user concern about the product they bought on Amazon. In this case, the product is Lenovo K8 Note Smartphone. In the next section, the sentiment polarity of each major topic would be revealed.

3.3 Sentiment Analysis of User Review Aspects

The next step of online customer review summarization is to infer the sentiments of online customer reviews. Before going further on the sentiments of each topic, we will take a look at the overall sentiment distribution of the entire text corpus. Figure 7 below shows the sentiment distributions of our customer review dataset. From Figure 7, we can see that the negative sentiments slightly dominated the customer opinion about the Lenovo K8 Note smartphone on Amazon.

For better insight, we then reveal the dominant keywords in both negative and positive sentiments. Figure 8 and Figure 9 shows the predominant keywords in negative and positive review subsequently. The interesting parts of the visualizations in Figure 8 and Figure 9 are both negative and positive sentiments have similar dominant words namely "battery", "camera", and "product". Slightly different from the positive sentiment, in the negative sentiment keyword "battery" is appear much bigger. Hence, we can conclude that the customers of the Lenovo K8 Note smartphone are mostly unhappy with the smartphone's battery performance.

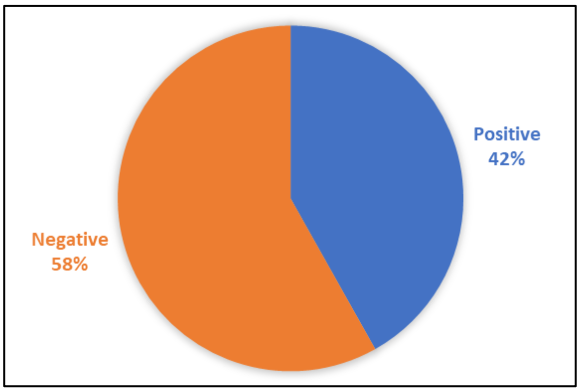

Figure 7 Sentiment Analysis Results Over the Entire Corpus

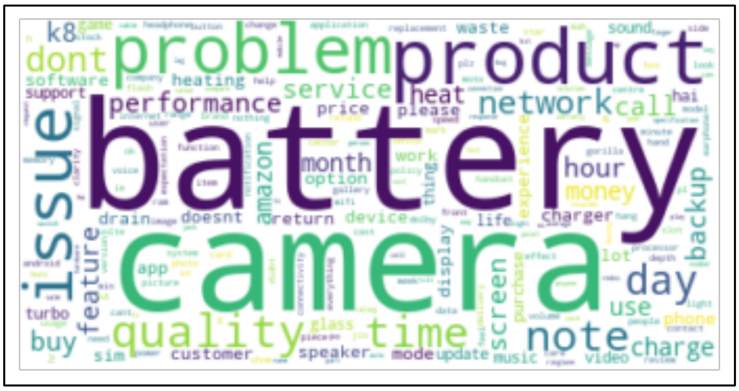

Figure 8 Word Cloud of Nouns Found in a Negative Review

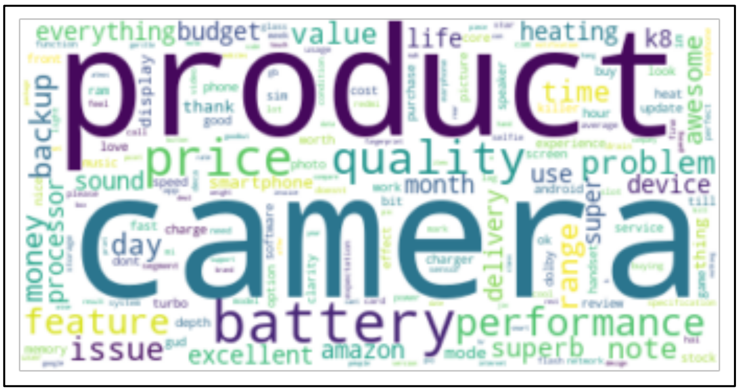

Figure 9 Word Cloud of Nouns Found in a Positive Review

The last step of this experiment is to infer the sentiment of each topic review. This step aims to provide a deeper insight into what customer likes and do not like about the Lenovo K8 Note smartphone product. Figure 10 shows the sentiment distribution of each major topic. From Figure 10, we can see that in three major aspects which are "Camera and Display", "Battery Quality" and "After Sales Service", the negative sentiments are dominant. Only one aspect named "Price and Money Value" users give mostly positive sentiments. Hence, we can conclude that even though the Lenovo K8 Note smartphone product is not meet user satisfaction in terms of its product features (camera, display, and battery) and after-sales service, the user of this product feels that this product is worth the price. According to the customer reviews on Amazon, Lenovo K8 Smartphone is considered good as a low-budget high-end smartphone.

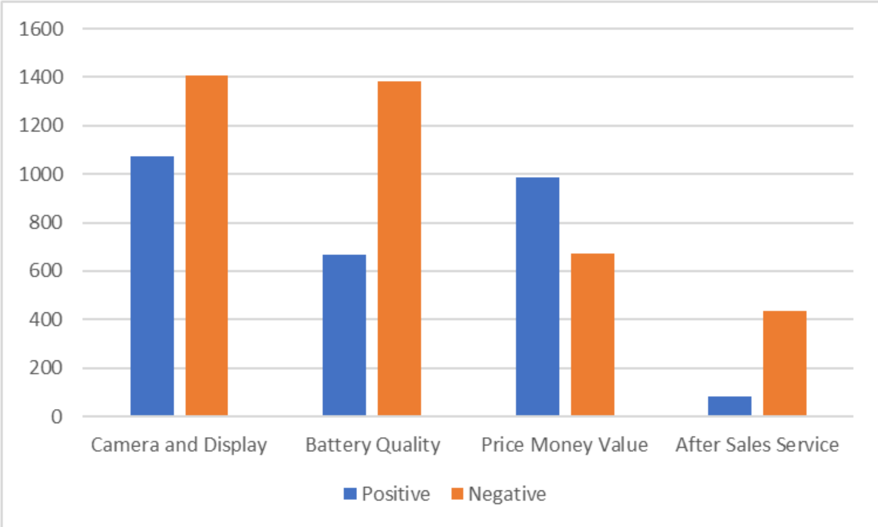

Figure 10 Sentiment Analysis Results in Each Aspect

4. CONCLUSIONS

From the conducted research we can conclude that the use of text mining specifically topic modeling and sentiment analysis was very beneficial in summarizing a large amount of text user reviews gathered from the internet. The results of running topic modeling and sentiment analysis give the product owner clear evidence quantitatively of what aspects of their product are liked by its customer and what are not. This research also gives an insight that combining two techniques of text analysis namely topic modeling and sentiment analysis gives a comprehensive result in terms of summarizing user reviews. This combination allows us to go further with user sentiment polarity on each aspect of the product instead of user sentiment in more general terms. Nevertheless, this research still has a drawback since the label of the topic was manually determined from the keywords. Therefore, this research can be further developed in terms of the automatic identification of proper labels for a cluster that contains a set of keywords.

REFERENCES

Ahmad, M., Aftab, S., Muhammad, S. S., & Ahmad, S. (2017). Machine Learning Techniques for Sentiment Analysis: A Review. *International Journal of Multidisciplinary Sciences and Engineering*, *8*(3).

Alghamdi, R., & Alfalqi, K. (2015). A Survey of Topic Modeling in Text Mining. *International Journal of Advanced Computer Science and Applications*, *6*(1). https://doi.org/10.14569/IJACSA.2015.060121

Baccianella, S., Esuli, A., & Sebastiani, F. (2010). SENTIWORDNET 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010*, 2200–2204.

Baharudin, B., Lee, L. H., & Khan, K. (2010). A Review of Machine Learning Algorithms for Text-Documents Classification. *Journal of Advances in Information Technology*, *1*(1). https://doi.org/10.4304/jait.1.1.4-20

Balakrishnan, V., & Ethel, L.-Y. (2014). Stemming and Lemmatization: A Comparison of Retrieval Performances. *Lecture Notes on Software Engineering*, *2*(3), 262–267. https://doi.org/10.7763/LNSE.2014.V2.134

Barde, B. V., & Bainwad, A. M. (2017). An overview of topic modeling methods and tools. *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*, 745–750. https://doi.org/10.1109/ICCONS.2017.8250563

Bashir, N., Papamichail, K. N., & Malik, K. (2017). Use of Social Media Applications for Supporting New Product Development Processes in Multinational Corporations. *Technological Forecasting and Social Change*, *120*, 176–183. https://doi.org/10.1016/j.techfore.2017.02.028

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*(4–5), 993–1022. https://doi.org/10.1016/b978-0-12-411519-4.00006-9

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146. https://doi.org/10.1162/tacl_a_00051

Chehal, D., Gupta, P., & Gulati, P. (2021). Implementation and comparison of topic modeling techniques based on user reviews in e-commerce recommendations. *Journal of Ambient Intelligence and Humanized Computing*, *12*(5), 5055–5070. https://doi.org/10.1007/s12652-020-01956-6

Cui, A. S., & Wu, F. (2017). The Impact of Customer Involvement on New Product Development: Contingent and Substitutive Effects. *Journal of Product Innovation Management*, *34*(1), 60–80. https://doi.org/10.1111/jpim.12326

Elena, C. A. (2016). Social Media – A Strategy in Developing Customer Relationship Management. *Procedia Economics and Finance*, *39*, 785–790. https://doi.org/10.1016/S2212-5671(16)30266-0

Elwalda, A., Lü, K., & Ali, M. (2016). Perceived derived attributes of online customer reviews. *Computers in Human Behavior*, *56*, 306–319. https://doi.org/10.1016/j.chb.2015.11.051

Gimpel, K., Schneider, N., O'Connor, B., Das, D., Mills, D., Eisenstein, J., Heilman, M., Yogatama, D., Flanigan, J., & Smith, N. A. (2011). Part-of-speech tagging for twitter: Annotation, features, and experiments. *ACL-HLT 2011 - Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, *2*, 42–47.

Greco, F., & Polli, A. (2020). Emotional Text Mining: Customer profiling in brand management. *International Journal of Information Management*, *51*, 101934. https://doi.org/10.1016/j.ijinfomgt.2019.04.007

Hidayanti, I., Herman, L. E., & Farida, N. (2018). Engaging Customers through Social Media to Improve Industrial Product Development: The Role of Customer Co-Creation Value. *Journal of Relationship Marketing*, *17*(1), 17–28. https://doi.org/10.1080/15332667.2018.1440137

Hu, S., Kumar, A., Al-Turjman, F., Gupta, S., Seth, S., & Shubham. (2020). Reviewer Credibility and Sentiment Analysis Based User Profile Modelling for Online Product Recommendation. *IEEE Access*, *8*, 26172–26189. https://doi.org/10.1109/ACCESS.2020.2971087

Irawan, M. I., Wijayanto, R., Shahab, M. L., Hidayat, N., & Rukmi, A. M. (2020). Implementation of social media mining for decision making in product planning based on topic modeling and sentiment analysis. *Journal of Physics: Conference Series*, *1490*(1), 012068. https://doi.org/10.1088/1742-6596/1490/1/012068

Jeong, B., Yoon, J., & Lee, J.-M. (2019). Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. *International Journal of Information Management*, *48*, 280–290. https://doi.org/10.1016/j.ijinfomgt.2017.09.009

Ko, N., Jeong, B., Choi, S., & Yoon, J. (2018). Identifying Product Opportunities Using Social Media Mining: Application of Topic Modeling and Chance Discovery Theory. *IEEE Access*, *6*, 1680–1693. https://doi.org/10.1109/ACCESS.2017.2780046

Maheswari;, M. U., & Sathiaseelan, D. J. G. R. (2017). Text Mining: Survey on Techniques and Applications. *International Journal of Science and Research (IJSR)*, *6*(6), 1660–1664.

Mahr, D., Stead, S., & Odekerken-Schröder, G. (2019). Making sense of customer service experiences: a text mining review. *Journal of Services Marketing*, *33*(1), 88–103. https://doi.org/10.1108/JSM-10-2018-0295

Mirtalaie, M. A., Hussain, O. K., Chang, E., & Hussain, F. K. (2018). Sentiment Analysis of Specific Product's Features Using Product Tree for Application in New Product Development. In *Lecture Notes on Data Engineering and Communications Technologies* (Vol. 8, pp. 82–95). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-319-65636-6_8

Ng, C. Y., Law, K. M. Y., & Ip, A. W. H. (2021). Assessing Public Opinions of Products Through Sentiment Analysis. *Journal of Organizational and End User Computing*, *33*(4), 125–141. https://doi.org/10.4018/JOEUC.20210701.oa6

Ramanathan, U., Subramanian, N., & Parrott, G. (2017). Role of social media in retail network operations and marketing to enhance customer satisfaction. *International Journal of Operations and Production Management*, *37*(1), 105–123. https://doi.org/10.1108/IJOPM-03-2015-0153

Sun, F., Liu, J., Wu, J., Pei, C., Lin, X., Ou, W., & Jiang, P. (2019). Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer. *International Conference on Information and Knowledge Management, Proceedings*, *11*, 1441–1450. https://doi.org/10.1145/3357384.3357895

Suresh, R., & Harshni, S. R. (2017). Data mining and text mining — A survey. *2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)*, *2018-Janua*, 412–420. https://doi.org/10.1109/ICCPEIC.2017.8290404

Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering*, *336*(1), 012017. https://doi.org/10.1088/1757-899X/336/1/012017

Vijayarani, S., Ilamathi, J., & Nithya. (2018). Preprocessing Techniques for Text Mining - An Overview. *International Journal of Computer Science & Communication Networks*, *5*(1), 7–16. https://doi.org/10.1016/j.procs.2013.05.286

Wang, L., Jin, J. L., Zhou, K. Z., Li, C. B., & Yin, E. (2020). Does customer participation hurt new product development performance? Customer role, product newness, and conflict. *Journal of Business Research*, *109*, 246–259. https://doi.org/10.1016/j.jbusres.2019.12.013

Xun, G., Gopalakrishnan, V., Ma, F., Li, Y., Gao, J., & Zhang, A. (2016). Topic Discovery for Short Texts Using Word Embeddings. *2016 IEEE 16th International Conference on Data Mining (ICDM)*, 1299–1304. https://doi.org/10.1109/ICDM.2016.0176

Zhan, Y., Tan, K. H., & Huo, B. (2019). Bridging customer knowledge to innovative product development: a data mining approach. *International Journal of Production Research*, *57*(20), 6335–6350. https://doi.org/10.1080/00207543.2019.1566662