

Sentiment Analysis of University Libraries during the Covid-19 Pandemic in Indonesia

Abdurrahman Rahim Thaha

Universitas Terbuka, South Tangerang, Indonesia

Abstract

The purpose of this study was to analyze the sentiments of university library visitors during the pandemic. Data collection techniques using web scraping techniques using a data scraper application. The data is taken from Google reviews as many as 261 reviews from ten universities in Indonesia. Data analysis technique using Vader method for sentiment analysis and Ekman method for classification of emotional sentiment. The results of the sentiment analysis in this study show that visitor satisfaction is quite high at the university library during the pandemic. Positive sentiment in the university library is 80.8% and the classification of emotional sentiment is dominated by joy as much as 71.6% compared to other sentiments such as negative, fear, and sadness. These results indicate that visits to the university library can still be carried out well even though there are many limitations in activities in the library because of the pandemic.

Keywords: sentiment analysis, university library, review, Indonesia.

1. Introduction

When the Covid-19 pandemic spreads worldwide, conventional educational institutions and their subordinate institutions, such as libraries, face challenges in carrying out their functions for the academic community. During the pandemic, the university library remains consistent in carrying out its role as an information-providing institution and responding to various problems during the pandemic to continue serving the academic community (Walsh and Rana 2020). The role of the library at the university in supporting the academic community in disseminating and providing information resources for the academic community has been widely recognized by students, lecturers, and professionals alike.

The growing popularity of digital media makes libraries use various online platforms to provide long-term promotions as institutions that are ready to help and support and get feedback from the academic community (Bhati and Kumar 2020). The pandemic presents unique information on how university libraries carry out their functions with the restrictions during the pandemic.

Research related to libraries during the pandemic has been widely studied regarding promotions in various digital media. In contrast, this study examines how the academic community's opinion responds to the state of university libraries in Indonesia during the pandemic from the beginning of the pandemic until the beginning of 2022. This study identifies the ten most prominent universities in Indonesia based on QS World data. It then crawls Google data reviews of each of the university libraries. to gain insight into the current state of the library during the pandemic.

2. Literature Review

2.1. Sentiment analysis

Sentiment analysis analyzes people's opinions, sentiments, judgments, attitudes, and emotions towards entities and their attributes expressed in written texts (Liu 2015). Sentiment analysis research, in general, is to classify all document opinions, both those expressing positive and negative sentiments. It is thus known as document-level sentiment classification. The results of sentiment analysis will be valuable knowledge for marketing, products, and

* Corresponding author.

E-mail address: abdurrahman@ecampus.ut.ac.id

other activities (Pang and Lee 2008). Positive sentiment words like beautiful, wonderful, and fantastic express some desired state or quality.

On the other hand, negative sentiment words like bad, terrible, and wrong express some undesirable state or quality. In addition to individual terms, there are phrases and idioms, such as high cost. Sentiment words and phrases are essential for sentiment analysis. Such a list of words and phrases is called a sentiment lexicon.

2.2. Text mining and text processing

Text mining extracts valuable information from data sources in a text by identifying and exploring exciting patterns. Textbooks consisting of an unstructured collection of documents will be processed and converted into a structured numerical representation (Makhabel et al., 2017).

According to Murnawan (2017), text preprocessing is a text mining process to process unstructured text into more structured text. Text preprocessing consists of four stages: case folding, tokenizing, filtering, and stemming. (1) Case folding is the stage to change all text in the document into all lowercase letters (lowercase). Only letters 'a' to 'z' are used in the case folding stage, while characters other than letters are omitted and considered delimiters. (2) Tokenizing is the stage of cutting the character set based on each word that composes it. Tokenizing breaks a set of characters in text into word units, distinguishing certain characters that can be treated as word separators or not. (3) Filtering is the stage of taking essential words from tokenizing results. Words like from, which, at, and for are examples of high-frequency words and can be found in almost every document (referred to as stopwords). The removal of this stopword can reduce the size of the runtime and reduce the noise level. (4) Stemming groups other words with the same root and meaning but different forms because they have different affixes.

2.3. University library

Higher education is the primary driver of any country's economic growth and is becoming an increasingly competitive market. Students are customers of higher education, so their satisfaction plays an essential factor in the success of educational institutions (Thomas and Galambos 2004). The more satisfied students are with the quality of services provided by an institution, the greater their chances of attracting prospective students in the long term.

The library is one of the university's services to its students (Flemming 2012). Increased use of libraries by students is one indicator of library performance in universities, so the increase in library service users needs to be improved (Kiran and Diljit 2017). From a marketing perspective, the increase in the number of students using the library cannot be separated from the existence of loyal library users (Brunner, Stöcklin, and Opwis 2008; Kiran and Diljit 2017). Loyal library users will reuse library services, spread positive information, and recommend library services to other students (Helgesen and Nettet 2011; Kiran and Diljit 2017). In this context, understanding loyalty and user reviews in library services become essential to develop effective strategies to retain students who use library services.

3. Method

The text mining method used in this research is to mine text that is processed using numbers and adjusted to the pattern in the text data. Text data in the form of reviews of university libraries in Indonesia on google maps, processed and converted into structured and numerical representations. The result after numerical processing is further analysis, such as text analysis, opinion analysis, information search, and other summary information.

The data collection technique used is web scraping. Web scraping is retrieving structured or unstructured documents from the internet. The data taken is audience review data who visited the library and provided comments on the google maps review facility from April 2020 to mid-2022.

The data collection used comes from libraries on the ten best campuses in Indonesia based on QS World 2022. The campuses are (1) Gadjah Mada University (UGM), (2) University of Indonesia (UI), (3) Bandung Institute of Technology (ITB), (4) Airlangga University (Unair), (5) Bogor Agricultural University (IPB), (6) Sepuluh Nopember Institute of Technology (ITS), (7) Padjadjaran University (Unpad), (8) Bina Nusantara University (Binus), (9) Diponegoro University (Undip), and (10) Telkom University (Tel-U). The number of reviews obtained can be seen in Table 1.

The analytical technique used is a rule-based sentiment analysis called Vader, using the Orange application (Demšar et al. 2013). Data analysis aims to obtain sentiment class categories (positive, negative, neutral). Furthermore, the

classification of emotions is analyzed by calculating the score or emotional probability for each review. The classification of emotions previously used by researchers is the classification of emotions from Ekman et al. (1972). The basic human emotions are classified into six types: anger, disgust, fear, joy, sadness, dan surprise (Colnerić and Demsar 2018).

Table 1. Number of Reviews each University Library

Num.	University	Number of Reviews
1	UGM	73
2	UI	72
3	ITB	9
4	Unair	16
5	IPB	46
6	ITS	14
7	Unpad	5
8	Binus	3
9	Undip	7
10	Tel-U	16
Total		261

4. Results and Discussion

A total of 261 reviews were obtained from ten university libraries in Indonesia. The review data is extracted for other text mining processes to obtain patterns in word frequencies, terms, or tokens in a collection of text documents. These results can be seen in Table 2.

Table 2. Term Frequency

Num.	Term	Freq.	Num.	Term	Freq.	Num.	Term	Freq.
1	Place	135	21	Work	18	41	Public	11
2	Comfort	91	22	Quiet	17	42	Use	11
3	Librari	90	23	Realli	17	43	Visit	11
4	Studi	55	24	Servic	17	44	Want	11
5	Book	52	25	Collect	16	45	Assign	10
6	Good	50	26	Spaciou	16	46	Avail	10
7	Compleat	38	27	Ui	16	47	Campu	10
8	Room	35	28	Discuss	15	48	Come	10
9	Also	27	29	Pandem	15	49	Comput	10
10	Nice	26	30	Time	14	50	Friendli	10
11	Cool	23	31	Ugm	14	51	Quit	10
12	Clean	20	32	Atmospher	13	52	Best	9
13	Floor	20	33	Cozi	13	53	Fast	9
14	Student	20	34	Open	13	54	Find	9
15	Thesi	20	35	Access	12	55	Free	9
16	Facil	19	36	Borrow	12	56	Knowledg	9
17	Mani	19	37	Even	12	57	One	9
18	Build	18	38	Hour	12	58	Park	9
19	Read	18	39	Internet	12	59	Tabl	9
20	Wifi	18	40	Lot	12	60	Visitor	9

According to Calheiros (2015), the terms that appear can be classified into four categories referring to the sentiment dictionary: strong positive, ordinary positive, ordinary negative, and strong negative. From the list of terms, there are strong positive sentiments, namely best and knowledgeable, as well as ordinary positives, namely comfort, cool, cozy, good, nice, and spacious. Meanwhile, the terms normal negative and strong negative were not found in the list of words with a minimum of nine occurrences.

After going through the preprocessing process on 261 reviews, sentiment analysis was carried out to obtain sentiment class categories (positive, negative, neutral) using the Vader method. The data categorization process is carried out

automatically by the Lexicon dictionary by calculating the sentiment score of each review. The results of class sentiment categorization from all reviews obtained data as shown in Figure 1.

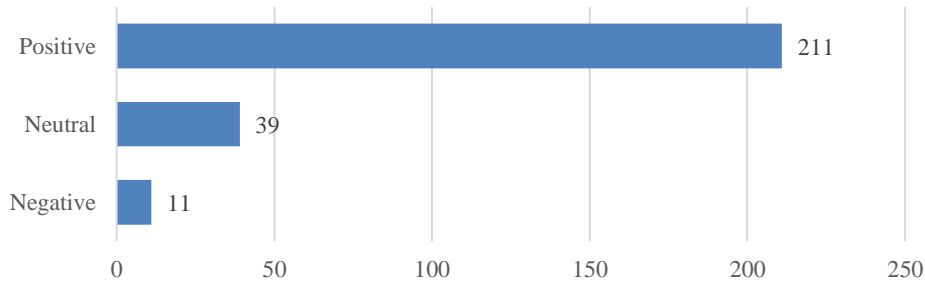


Figure 1. Class Sentiment Categorization Results

Based on Figure 1, the results of the sentiment class categorization show that the number of positive reviews has a higher frequency, namely 211 out of 261 reviews or 80.8% of the total reviews. At the same time, the number of neutral reviews has a frequency of 39 reviews, or 14.9% of the total reviews. In addition, negative reviews have the lowest frequency, with 11 reviews or 4.2% of the total reviews.

The results of the sentiment class show that the level of satisfaction of university library users is relatively high, by looking at positive reviews reaching 80.8% and negative reviews only around 4.2%. Furthermore, each review is classified by emotion in Figure 2.

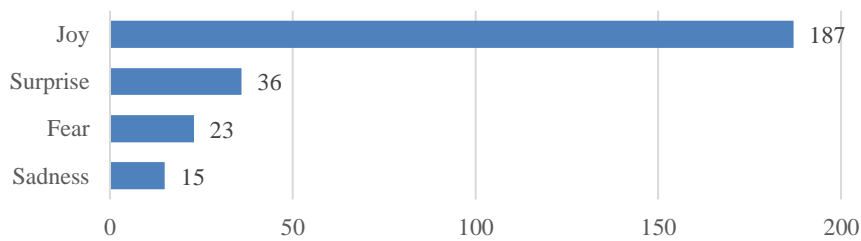


Figure 2. Emotion Classification Results

Based on Figure 2, the classification of emotional sentiments for all reviews shows that emotional sentiment is dominated by happy emotions, with a total of 187 reviews or 71.6% of the overall reviews. These results indicate that more than 50% of university library users are happy and enjoy the university library. The second-order surprise sentiment was surprise with a total of 36 reviews, or 13.8% of all reviews, the emotion of fear was 23 reviews or 8.8% of all reviews, and the emotion of sadness was 15 reviews or 5.7% of all reviews. More in-depth exploration can be done by processing the frequency of terms in each emotion classification (Thaha and Aziz 2020). Exploration will return the frequency terms with the most frequent and relevant occurrences. The results of the term frequency process for each emotion classification are outlined in the word cloud form in Table 4.

Figure 3 shows the most frequently occurring terms in each emotion classification. The classification of joy emotions shows the happy feelings of library visitors with the words library, place, comfort, book, complete, good, study, and nice. In this classification, library visitors are satisfied with the facilities and activities in the library despite the limitations of the pandemic.



Figure 3. Word Cloud each Emotion Classification

The surprise emotions category tends to be filled with positive words like good, nice, comfort, and best, as well as the addition of the word pandemic, which describes the library's condition during the pandemic. This classification illustrates that library visitors tend to be surprised by the changes that occurred during the pandemic in a positive sense. The fear emotion category shows the caution of visitors to the library situation during a pandemic by appearing words such as comfort, place, clean, build, and public. Classification of sadness emotions related to place, unfortunate, and good. The emotion category here describes visitors' sadness due to limited activities in different libraries before the pandemic. This insight can serve as an illustration for stakeholders regarding the reviews that appear in the four categories.

5. Conclusion

Sentiment analysis using text mining can be done to get a certain pattern from 261 university library review data collected. From the review pattern, it is known that positive sentiment is the most frequent review, with a total of 211 times. Neutral sentiment is the second review with a frequency of 39 times and negative sentiment with a frequency of 11 times.

The results of sentiment analysis in this study indicate the level of satisfaction of visitors to the university library is quite good. The emotional sentiment classification of all reviews by emotion was joy with a total of 187 reviews or 71.6% of the overall reviews followed by the emotion classification of surprise, fear, and sadness. These results indicate that visits to the university library can still be carried out well even though there are many limitations in activities in the library because of pandemic.

References

- Bhati, Pankaj, and Inder Kumar. 2020. "Role of Library Professionals in a Pandemic Situation like COVID-19." *International Journal of Library and Information Studies* 10, no. 2: 33–48.
- Brunner, Thomas A, Markus Stöcklin, and Klaus Opwis. 2008. "Satisfaction, Image and Loyalty: New versus Experienced Customers." *European Journal of Marketing*.
- Calheiros, Ana Catarina dos Santos. 2015. "Sentiment Analysis in Hospitality Using Text Mining: The Case of a Portuguese Eco-Hotel."
- Colnerić, Niko, and Janez Demsar. 2018. "Emotion Recognition on Twitter: Comparative Study and Training a Unison Model." *IEEE Transactions on Affective Computing*.
- Demšar, Janez, Tomaž Curk, Aleš Erjavec, Črt Gorup, Tomaž Hočevar, Mitar Milutinović, Martin Možina, Matija Polajnar, Marko Toplak, and Anže Starič. 2013. "Orange: Data Mining Toolbox in Python." *The Journal of Machine Learning Research* 14, no. 1: 2349–53.
- Ekman, Paul, Wallace V Friesen, and Phoebe Ellsworth. 1972. *Emotion in the Human Face: Guide-Lines for Research and an Integration of Findings: Guidelines for Research and an Integration of Findings*. Pergamon.
- Flemming, Christina. 2012. "A Joint University Library: Vision and Reality." *Library Management*.
- Helgesen, Øyvind, and Erik Nettet. 2011. "Does LibQUAL+ TM Account for Student Loyalty to a University College Library?" *Quality Assurance in Education*.
- Kiran, K, and Singh Diljit. 2017. "Antecedents of Customer Loyalty: Does Service Quality Suffice?" *Malaysian Journal of Library & Information Science* 16, no. 2: 95–113.
- Liu, Bing. 2015. *Opinions, Sentiment, and Emotion in Text*. Cambridge University Press.
- Makhabel, B., Mishra, P., Danneman, N., & Heimann, R. 2017. *R: Mining Spatial, Text, Web, and Social Media Data*. Birmingham: Packt Publishing Ltd.
- Murnawan, Murnawan. 2017. "Pemanfaatan Analisis Sentimen Untuk Pemeringkatan Popularitas Tujuan Wisata." *Jurnal Penelitian Pos Dan Informatika* 7, no. 2: 109.
- Pang, Bo, and Lillian Lee. 2008. "Opinion Mining and Sentiment Analysis." *Foundations and Trends® in Information Retrieval* 2, no. 1–2: 1–135.

- Thaha, Abdurrahman Rahim, and Fauzan Aziz. 2020. “Penambangan Teks Pada Tujuan Wisata Di Bandung Raya (Studi Kasus: Tangkuban Perahu Dan Kawah Putih).” *Jurnal Sekretaris Dan Administrasi Bisnis* 4, no. 2: 146–56.
- Thomas, Emily H, and Nora Galambos. 2004. “What Satisfies Students? Mining Student-Opinion Data with Regression and Decision Tree Analysis.” *Research in Higher Education* 45, no. 3: 251–69.
- Walsh, Benjamin, and Harjinder Rana. 2020. “Continuity of Academic Library Services during the Pandemic the University of Toronto Libraries’ Response.” *Journal of Scholarly Publishing* 51, no. 4: 237–45.