Factors Determination of Service Dissatisfaction at PT XYZ Based on Negative Google Maps Reviews Using Topic Modeling Method (Latent Dirichlet Allocation)

Fanji Hari Suhendro[1,*] and Yenny Maya Dora[1]

[1]*University of Widyatama, Bandung, Indonesia*
[*]*Corresponding author. Email: fanjiharisuhendro.fhs@gmail.com*

ABSTRACT

This research aims to analyze the main topics underlying negative reviews of PT XYZ's laboratory & clinic services. The negative review data was manually collected and processed using the Latent Dirichlet Allocation (LDA) method. The analysis was conducted on 1129 negative reviews, revealing a dominance of topic 3 with a percentage of 88.57%, which focuses on the interaction between dentists and patients. The research results indicate that the friendliness of doctors and the availability of medicine at clinics are significant factors influencing user dissatisfaction. Furthermore, other identified topics also highlight aspects such as patient experiences in obtaining healthcare services, medical procedures, and negative experiences with doctors. This research provides insights into areas requiring improvement to enhance the quality of healthcare services.

Keywords: *Topic modeling, user review, topic analysis.*

1. INTRODUCTION

The Covid-19 pandemic has transformed the landscape of the healthcare industry globally, including in Indonesia. Especially during the Covid-19 outbreak, there was a surge in demand for pharmaceutical products, indicated by a growth trend (Putri et al., 2020). The laboratory and clinic industry, which may have previously been considered a supporting sector, has now become a frontline in public health management. Post-pandemic, the need for laboratory and clinic services has not only increased in quantity but has also undergone changes in quality and public expectations. The demand for rapid, accurate, and reliable diagnostic tests has become increasingly urgent. The public is also increasingly aware of the importance of disease prevention and proactive health maintenance. Utilizing preventive care services can help promote healthier lifestyles, provide early disease detection, and reduce the need for inpatient care services (Tian et al., 2010).

To excel competitive environment, In facing increasingly fierce competition, enhancing service quality has become a crucial factor for the success of the laboratory and clinic industry. Service quality not only encompasses the accuracy of test results and the efficiency of processes but also the overall patient experience. This includes comfort, friendliness, and speed of service, as well as the ability to provide clear and easily understood information. The patient's perspective is increasingly seen as a meaningful and important indicator of healthcare service quality that must be considered as part of a comprehensive assessment of healthcare quality (Sandhyaduhita et al., 2016).

Nowadays, many service users seek information beforehand before utilizing a healthcare service, one of which is by looking through online reviews. With such changes in the customer journey, companies must begin to pay attention to the complaints felt by the public about the services provided to remain competitive. Online reviews have become a crucial source of information for companies to understand public perception and improve product quality (Pan et al., 2020). Thousands of user reviews are available on Google Maps, which is one of the sources that can be read by the public. However, analyzing large amounts of online review text data manually to generate knowledge is a challenging process. Topic modeling is a statistical method used to analyze extensive text data and discover hidden topics (Kumar et al., 2023). Latent Dirichlet allocation (LDA), a Bayesian probabilistic algorithm widely used in previous research, generates hidden topics (Aman et al., 2021). Hussain et al., 2023 used LDA to uncover the factors influencing the service quality of m-banking in 24 banks in Pakistan.

In this context, this research aims to analyze the dissatisfaction factors influencing the service quality of laboratories and clinics based on negative reviews on Google Maps that were manually collected from January 2023 to September 2024, in total 1129 reviews, and to identify the main dissatisfaction factors analyzed through topic modeling using LDA.

The contribution of this research is expected to determine the factors that cause complaints, as well as the main areas that need to be prioritized for service improvement at PT XYZ.

2. RESEARCH METHODS

In this research, LDA is used as the primary method to identify the dominant topics that emerge from Google Maps reviews. LDA is a topic modeling algorithm that allows for the discovery of hidden thematic structures within a collection of text documents. In this context, each user review is considered a "document," and the collection of reviews forms the “corpus” that is analyzed.

2.1. Negative Review Data Collection

This research is based on negative Google Maps reviews that were manually collected from January 2023 to September 2024, with a total of 1129 reviews.

2.2. Topic Modeling Process with LDA

Google Colab was used as a platform to run the LDA algorithm and conduct topic analysis. Google Colab provides a powerful and flexible computing environment for text data processing, topic modeling, and word cloud generation.

a. Data Pre-processing: Before applying LDA, the review data underwent pre-processing to clean the text from noise and prepare it for analysis. These pre-processing steps include tokenization to break down the text into individual tokens, such as words or phrases, stop word removal to eliminate meaningless words, normalization to replace certain words with more appropriate ones, and stemming to reduce words to their base form (Mustaqim et al., 2024; Puspasari et al., 2024).

b. Term-Document Matrix Formation: After pre-processing, the reviews were transformed into a term-document matrix, where each row represents a unique word in the corpus, each column represents a review, and each cell contains the frequency of the word in that review.

c. Topic Number Determination: Five main topics were selected, which are expected to be prioritized for service improvement at PT XYZ.

d. Word Frequency: Each topic is represented by 10 words.

e. LDA Application: The LDA algorithm is applied to the term-document matrix to identify the dominant topics in the reviews. LDA assumes that each review is a mixture of various topics, and each topic is a probability distribution over the words in the corpus.

f. LDA Topic Visualization: The LDA results are visualized in the form of word clouds, influenced by the frequency of occurrence and word weighting.

3. RESULTS AND DISCUSSIONS

3.1. Result

Data was collected from all Google Maps reviews across all 338 laboratory and clinic outlets of PT XYZ. Negative reviews was collected from January 2023 to September 2024. The technique used in the review data collection process was manual extraction from each Google Maps outlet location, resulting in a total of 1129 negative reviews.

The collected reviews were processed using the LDA method with the help of Google Colab, as shown in Figure 1, which produced the percentage of each topic, as presented in Table 1.

Topic Modeling with sklearn

```
from sklearn.decomposition import LatentDirichletAllocation
lda_model=LatentDirichletAllocation(n_components=5,learning_method='online',random_state=42,max_iter=1)
# n_components is the number of topics
```

```
vect =TfidfVectorizer(max_features=10000)
vect_text=vect.fit_transform(df['textdata_tokens_stemmed2'])
idf=vect.idf_
lda_top=lda_model.fit_transform(vect_text)
```

```
sum=0
for i in lda_top[0]:
  sum=sum+i
print(sum)
```

```
0.9999999999999999
```

```
# composition of doc 0 for eg
print("Document 0: ")
for i,topic in enumerate(lda_top[0]):
  print("Topic ",i,": ",topic*100,"%")
```

```
Document 0:
Topic  0 :  2.8352601196010627 %
Topic  1 :  2.8473281047198142 %
Topic  2 :  88.56817618629314 %
Topic  3 :  2.8420732800802857 %
Topic  4 :  2.907162309305691 %
```

Figure 1 Code lines for the LDA process .

Table 1. Percentage of each topic

Topics	Persentage
1	2,84%
2	2,85%
3	88,57%
4	2,84%
5	2,91%
Total	100%

From this distribution, it is evident that topic 3 dominates the negative review documents with a percentage of 88.57%. This indicates that the overall reviews are very strongly related to the theme represented by topic 3. The 10 words for each topic can be seen in Table 2.

Table 2. Keywords of each topic

Topics	Keywords
1	service, teeth, hours, queue, bpjs, doctor, register, bad, medicine, patient
2	system, service, doctor, patient, call, bpjs, already, refer, recover, sick
3	doctor, hours, service, friendly, medicine, patient, teeth, bpjs, clinic, sick
4	doctor, injection, hours, rope, wait, release, service, college, recommend, there
5	doctor, patient, use, medicine, bpjs, I, rude, bad, correct, help

The generated keywords can help to identify the underlying themes of each resulting topic.

Additionally, word clouds for each topic can be derived from the LDA process results. The images below illustrate the word clouds for each topic, which can serve as a basis for further discussion.

Figure 1 Word cloud of topic 1 - 5.

3.2. Discussions

Each topic generated by LDA has unique characteristics reflected in its associated keywords. The description of each topic will help us understand the themes underlying the negative reviews and identify areas needing improvement. To better understand the implications of the topic distribution generated by LDA, here is an interpretation that can be developed from each topic:

a. Topic 1: Words like “service”, “teeth”, “hours”, “queue”, “bpjs”, “doctor”, “register”, “bad”, “medicine”, and “patient” indicate this topic relates to patients' experiences in obtaining BPJS health services, particularly dental services, focusing on aspects such as waiting times, registration processes, and service quality.

b. Topic 2: Words like “system”, “service”, “doctor”, “patient”, “call”, “bpjs”, “already”, “refer”, “recover”, and “sick” indicate this topic relates to the BPJS health service system, including patient referral and recovery processes.

c. Topic 3: Words like “doctor”, “hours”, “service”, “friendly”, “medicine”, “patient”, “teeth”, “bpjs”, “clinic”, and “sick” indicate this topic focuses on the interaction between dentists and BPJS patients, emphasizing service hours, doctor friendliness, and medicine availability at the clinic.

d. Topic 4: Words like “doctor”, “injection”, “hours”, “rope”, “wait”, “release”, “service”, “college”, “recommend”, and “there” indicate this topic relates to specific medical procedures (injections), waiting times, and service recommendations.

e. Topic 5: Words like “doctor”, “patient”, “use”, “medicine”, “bpjs”, “I”, “rude”, “bad”, “correct”, and “help” indicate this topic focuses on negative experiences of BPJS patients with doctors, including medicine usage and unpleasant interactions.

The LDA analysis results show that most people who have used PT XYZ's services are dissatisfied with Topic 3, which dominates the negative review documents with a percentage of 88.57%, indicating that the themes related to this topic significantly influence negative user sentiment. This topic, marked by keywords such as “doctor”, “hours”, “service”, “friendly”, “medicine”, “patient”, “teeth”, “bpjs”, “clinic”, and “sick”, highlights the interaction between dentists and BPJS patients, with an emphasis on service hours, doctor friendliness, and medicine availability at the clinic, focusing on doctor friendliness and medicine availability at the clinic. This indicates that the aspects of friendliness and medicine availability at dental clinics are significant factors influencing public satisfaction. Additionally, other topics provide important insights into public experiences. The characteristics of healthcare service quality have been found to have a significant influence on patient satisfaction in various studies. Medical service evaluations are always based on service quality, basic care, and timely responses from healthcare providers. Patient satisfaction is influenced by the quality of medical services and the value of services provided to patients (Akthar et al., 2023).

Overall, the LDA analysis results indicate that the interaction between doctors and patients, especially doctor friendliness and medicine availability, is a dominant factor influencing negative user sentiment. In addition, system issues, medical procedure efficiency, and negative experiences with doctors are also public concerns. These findings underscore the importance of improving service quality in these aspects to enhance public satisfaction.

2. CONCLUSION

The results of this study yielded dissatisfaction factors that influence the service quality of laboratories and clinics based on negative Google Maps reviews. By using the Latent Dirichlet Allocation (LDA) topic modeling, this study successfully identified the main themes underlying user negative reviews.

The research results indicate that most people who provided negative reviews were dissatisfied, with the 3rd topic, which relates to the interaction between dentists and BPJS patients, dominating the negative reviews (88.57%). This topic highlights the importance of service hours, doctor friendliness, and medicine availability at the clinic. Other identified topics also provide insights into aspects such as patient experiences in obtaining BPJS health services, the BPJS health service system, medical procedures, and negative patient experiences with doctors

The practical implications of this research are the need for improved service quality, especially in terms of doctor-patient interaction (doctor friendliness), medicine availability, service efficiency and accessibility, health service system improvements, medical procedure efficiency, and handling negative patient experiences.

This study has limitations in terms of the data used, which is only negative reviews manually collected from Google Maps. Future research could consider data from other sources, such as social media or surveys, to obtain a more comprehensive picture. Additionally, this study focuses on laboratory and clinic services specifically at PT XYZ, so the findings may not be specific to certain types of healthcare services.

Suggestions for future research include analyzing the sentiment of more varied data to understand the intensity of emotions associated with each theme, identifying trends over time to understand changes in user perception, comparing analysis results with operational data, and expanding the research scope to other types of healthcare services, such as hospital or mental health services.

References

Akthar, N., Nayak, S., & P, Y. P. (2023). Determinants of patient satisfaction in Asia : Evidence from systematic review of literature. *Clinical Epidemiology and Global Health*, *23*(April), 101393. https://doi.org/10.1016/j.cegh.2023.101393

Aman, J. J. C., Smith-Colin, J., & Zhang, W. (2021). Listen to E-scooter riders: Mining rider satisfaction factors from app store reviews. *Transportation Research Part D: Transport and Environment*, *95*, 102856. https://doi.org/https://doi.org/10.1016/j.trd.2021.102856

Hussain, A., Hannan, A., & Shafiq, M. (2023). Exploring mobile banking service quality dimensions in Pakistan: a text mining approach. *International Journal of Bank Marketing*, *41*(3), 601–618. https://doi.org/10.1108/IJBM-08-2022-0379

Kumar, A., Chakraborty, S., & Bala, P. K. (2023). Text mining approach to explore determinants of grocery mobile app satisfaction using online customer reviews. *Journal of Retailing and Consumer Services*, *73*, 103363. https://doi.org/https://doi.org/10.1016/j.jretconser.2023.103363

Mustaqim, I. Z., Puspasari, H. M., Utami, A. T., Syalevi, R., & Ruldeviyani, Y. (2024). Assessing Public Satisfaction of Public Service Application Using Supervised Machine Learning. *IAES International Journal of Artificial Intelligence (IJ-AI)*, *13*(2), 1606–1616.

Pan, X., Wang, H., You, W., Zhang, M., & Yang, Y. (2020). Assessing the reliability of electronic products using customer knowledge discovery. *Reliability Engineering and System Safety*, *199*(35), 106925. https://doi.org/10.1016/j.ress.2020.106925

Puspasari, H. M., Mustaqim, I. Z., Utami, A. T., & Syalevi, R. (2024). Evaluation of Indonesia ' s police public service platforms through sentiment and thematic analysis. *IAES International Journal of Artificial Intelligence*, *13*(2), 1596–1607. https://doi.org/10.11591/ijai.v13.i2.pp1596-1607

Putri, M. A., Septyanani, R., & Santoso, A. P. A. (2020). Dampak Covid-19 Pada Perekonomian Indonesia. *Prosiding HUBISINTEK*, *1*, 198.

Sandhyaduhita, P., Fajrina, H. R., Pinem, A. A., Hidayanto, A. N., Handayani, P., & Junus, K. (2016). *Hospital Service Quality from Patients Perspective : A Case of Indonesia*. *December 2017*. https://doi.org/10.4018/IJEHMC.2016100103

Tian, W.-H., Chen, C.-S., & Liu, T.-C. (2010). The demand for preventive care services and its relationship with inpatient services. *Health Policy*, *94*(2), 164–174. https://www.sciencedirect.com/science/article/abs/pii/S0168851009002450