ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

COMPARATIVE ANALYSIS OF ONLINE REVIEW PLATFORMS: IMPLICATION IN ELECTRONIC SERVICE QUALITY FOR VIDEO PLAYERS AND EDITOR APPS (YOUTUBE AND TIKTOK)

NURVITA TRIANASARI [1], MAYA ARIYANTI [2] and LIDYA SRI MULYANI [3]*

[1, 2] Doctoral in Management, Telkom University, Bandung, Indonesia.
Email: [1]nurvitatrianasari@telkomuniversity.ac.id, [2]ariyanti@telkomuniversity.ac.id
[3] Magister Student, Telkom University, Bandung, Indonesia.
*Corresponding Author Email: lidyasrimulyani@gmail.com

Abstract

This study uses sentiment analysis, topic modeling, pearson correlation, and linear regression to identify the critical improvements that video-sharing platforms, particularly YouTube and TikTok, should make in the future to better handle user reviews. Given that video sharing provides viewers with an extremely immersive and engaging experience, its growing significance in relation to user feedback is becoming progressively apparent for companies. We initially preprocess 21,365 user reviews using Naïve Bayes classification, and we divide them into seven groups according to the dimensions of e-service quality: efficiency, responsiveness, system availability, compensation, fulfillment, contact, and privacy. The accuracy percentages for TikTok and YouTube came out as 80.33% and 75.56%, accordingly, indicating excellent performance in evaluating the quality of both platforms. Sentiment analysis revealed a higher prevalence of negative sentiment on TikTok and YouTube. Then, topic modeling based on Latent Dirichlet Allocation (LDA) evaluates the sentiment of the topic as well as the model of the topics discussed. The purpose of this research is to help both companies and individuals map public opinion toward a certain topic by analyzing the sentiment of the text and creating a topic model. We also measure the relationship between service quality using user sentiment and ratings, then predict future user ratings using predictive analysis methods such as linear regression. Regarding the classification of dimensions, we effectively draw attention to three dimensions— efficiency, fulfillment, and system availability —that are deemed important in the comparison between YouTube and TikTok. Then, the positive and significant relationship between service quality using user sentiment and ratings.

Keywords: Sentiment Analysis, Topic Modelling, Regression Linear, Correlation, Service Quality, Media Sharing Network.

INTRODUCTION

The twenty-first century is becoming proficient in numerous facets of technology, such as social media sharing networks (Novendri et al., 2020). Social media platforms known as "media-sharing networks" let users exchange visual content, like pictures and videos. Due to online video sharing providing audiences with an extremely immersive and engaging experience, its importance in relation to user feedback is growing for companies.

Videos are becoming one of the most popular social media and communication platforms (Sihag et al., 2023). Two of the most widely used social media sites that facilitate online video sharing are TikTok and YouTube.

Both of them are now engaged in intense competition with one another (Baegaegang & Cheol-soo, 2023).

ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

Now various news is starting to emerge regarding the popularity of TikTok which is starting to catch up with YouTube (Cepeda-López et al., 2019; Cuofanno, 2023; Hern, 2022; Ivanenko, 2023; Perez, 2023). First, regarding the decline in YouTube advertising revenue which was reported to have decreased by Evelyn Mutchell as Google Insider Intelligence analyst for the first time in Q4 2019, largely due to continued competition in streaming and short video applications. Second, based on the number of active users.

YouTube was launched in Indonesia in 2005 as a video sharing platform which has 139 million active YouTube users in Indonesia. When compared with TikTok, which was launched in 2017 and already has 109.9 million active users, this shows TikTok's significant growth. Third, based on time spent. Time spent Tiktok has a duration of 29 hours. This is longer compared to YouTube which lasts 26 hours 40 minutes (Kemp, 2023). Fourth, TikTok is ranked first in the video player and editor apps category on Google Play and YouTube is ranked sixth, and TikTok's rating is 4.4 per 5 and YouTube's rating is 4.2 per 5 (Google Play, 2023). Apart from that, based on examples of crawling review data, YouTube and TikTok users still feel dissatisfied with the services provided by YouTube and TikTok.

Based on the explanation above, to increase YouTube's popularity so that it is not overtaken by TikTok, YouTube needs to know more deeply about its customers through service quality analysis. Vice versa, for TikTok, to become the most popular application and beat YouTube, TikTok needs to know more deeply about its customers through quality service analysis.

This is because good service quality will produce customer satisfaction, thereby increasing customer loyalty. Customer satisfaction is also the basis for increasing customer loyalty amidst current competition on similar platforms. The opportunity to collect user perceptions through online user reviews is considered a faster methodology than conducting traditional methodologies with direct sampling (Sari et al., 2018). Service quality research has been carried out in various sectors and industries, including e-commerce, food and beverage, online auction companies, online airline ticket and hotel booking services, and airports. However, this research has not discussed the video sharing industry platforms that are currently in demand by the public.

Most companies conduct consumer review analysis using conventional methods for collecting data using sampling and/or questionnaires. This approach has limitations in terms of taking longer time and higher operating costs. The author proposes a different approach to analyze the level of service quality using user-generated content data from online review sites (Google Play Store) or also called User Generated Content (UCG), because of the unstructured nature of the data, we use sentiment analysis methods to classify dimensions service quality according to user perception and topic modeling. Using this method, we can determine customer perceptions regarding the quality of services provided (Sari et al., 2018) and evaluate user experience (Kiliç & Çadirci, 2022). Apart from that, researchers also carried out predictive analysis of the correlation between e-service quality and online customer ratings, where the e-service quality dimensions contribute significantly to the rating or value of electronic service quality (Ayo et al., 2016).

ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

This research aims to identify sentiments of YouTube and TikTok users based on e-service quality dimensions (E-Servqual), identify ratings of YouTube and TikTok service users, identify what topics are formed in each e-service quality dimension (E-Servqual) to measure the quality of YouTube and TikTok services, analyze the correlation between e-service quality (E-Servqual) and online customer ratings and determine the influence between service quality using sentiment on the ratings given by users. The results of this research contribute to enriching knowledge in the field of marketing management, developing knowledge in the field of big data and data analytics, and expanding understanding of customer satisfaction with services on video-sharing platforms. Apart from that, the YouTube Indonesia and TikTok Indonesia teams were able to determine the best strategy to improve the performance of their application services. Improving service performance will certainly have an impact on increasing customer satisfaction and customer loyalty in using the application platform.

LITERATURE REVIEW

Electronic Service Quality

In this industry, service quality is used by companies to differentiate themselves from competing companies. This results in increased competition, resulting in the competitive advantage that all companies want to achieve. Companies need to understand the relationships related to service quality, for example the relationship between service quality and customer satisfaction and the relationship between service quality, customer satisfaction and customer loyalty (Mushavhanamadi & Ratlhagane, 2018). Service quality is the ability of a service or product to carry out its functions and performance, which can meet customer needs and desires (Naini et al., 2022). According to Kotler & Keller (2022) service quality or service quality (ServQual) is the totality of features and characteristics of a product or service that influence its ability to satisfy stated or implied needs.

There are five general service quality dimensions: tangible evidence, reliability, responsiveness, assurance, and empathy (Kotler & Keller, 2022). The concept model for understanding and improving service quality in electronic form is called e-Servqual (Sari et al., 2018). ServQual theory developed into electronic service quality called ES-QUAL (Parasuraman et al., 2005). Electronic Service Quality (ES-QUAL) is all customer interactions on a website that includes the extent to which the website facilitates efficient and effective shopping, purchasing, and delivery. Electronic Service Quality (ES-QUAL) consists of two measurement scales (Parasuraman et al., 2005; Zahra, 2021). First, the e-core service quality scale (E-S-Qual) which consists of 4 dimensions: Efficiency, Fulfillment, System Availability, and Privacy. Second, the e-recovery service quality scale (E-Recs-QUAL) which consists of three dimensions: responsivenes, Compensation, and Contact. These seven criteria are used to evaluate the quality of YouTube and TikTok as video player and editor apps. Table I above, provides additional information about the definitions of each dimension.

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

Table 1: Electronic Service Quality

Dimension	Definition
Efficiency	Ease and speed in accessing the site by users.
Fulfillment	the extent to which site services can meet user needs.
System Availability	The service site's related to the function of the services provided.
Privacy	The service site's ability to protect user security and information.
Responsiveness	The site's ability to respond to user requests, questions and complaints.
Compensation	The service site's ability to provide compensation to customers who experience problems based on terms and conditions.
Contact	The service site's availability to provide services through good communication with customers.

Source: Result of data processing (2023)

To create customer value, companies must strive to provide superior service on all of these e-core (ES-Qual) and e-recovery (E-Recs-QUAL) dimensions. There are seven reasons why quality is necessary for a company as a company reputation. First, reducing the costs required to produce quality products or services but still oriented towards customer satisfaction. Second, increasing the market through minimizing costs is achieved because the company can reduce prices while still prioritizing quality. Third, accountable products and services require companies to always be responsible for meeting customer needs and expectations in increasing market competition by producing quality products or services. Fourth, if a quality product or service can be offered and introduced to the international market, the company will give a good impression of that quality. Fifth, the appearance of products and services will be known, where the company that produces the product will be known and trusted by customers. The things above will increase customer satisfaction and perceived quality (Naini et al., 2022).

Based on the explanation, it can be concluded that service quality is the ability of a service or product in its performance to meet customer needs and desires. Electronic Service Quality (ES-QUAL) consists of two measurement scales. First, the e-core service quality scale (E-S-Qual) which consists of four dimensions, namely efficiency, fulfillment, system availability and privacy. Second, the e-recovery service quality scale (E-Recs-QUAL) which consists of three dimensions, namely responsiveness, compensation, and contact. Service quality is something that differentiates one company from another which can be a competitive advantage for a company, so companies need to understand the relationship related to service quality, for example the relationship between service quality and customer satisfaction and the relationship between service quality, customer satisfaction, and customer loyalty

Sentiment Analysis

Sentiment analysis studies opinions that express or imply positive or negative sentiments. The term opinion is described as a broad concept that includes sentiment, evaluation, judgment, or attitude, and related information such as the target of the opinion and the person holding the opinion, and using the term sentiment to mean only the underlying positive or negative feelings (Feraco et al., 2017). In this research, the Naïve Bayes algorithm was used. Naïve Bayes is one of the most widely used and affordable classifiers for research. Apart from that, using Naïve

The Seybold REPORT

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

Bayes is faster and doesn't take up as much memory as other classification methods, namely Support Vector Machine (SVM). Support Vector Machine (SVM) is an algorithm commonly used to classify text (Zahra, 2021).

Topic Modelling

Text documents consist of words, topics mentioned in many documents can be expressed by combinations of closely related words. Each document consists of several topics. Topic Modeling is a technique used to infer hidden topics in text documents. Topic modeling represents each document as a complex combination of several topics and each topic as a complex combination of several words, also used as a text mining tool to classify documents based on topic inference results (Alamsyah et al., 2018).

Topic modeling is applied in classifying documents with similar content, finding possible topics in a collection of texts, identifying relationships between terms, and grouping trending topics. In essence, it collects data from various sources and analyzes it to produce more descriptive information or trends.

The results of this modeling are interesting because they help determine which areas (topics) receive attention and which do not, so it is necessary to carry out further research regarding areas (topics) that do not receive attention (Shaalan & Emran, 2022).

Topic modeling is one of the most powerful techniques in text mining for data mining, discovering latent data, and finding relationships between data and text documents.

There are various methods for topic modeling; Latent Dirichlet Allocation (LDA) is one of the most popular because LDA can be used to summarize, cluster, connect or process very large data because LDA produces a list of weighted topics for each document.

The topics that emerge from the data processing will then be tested for topic coherence, namely the relationship between the probability descriptions of words found with each other in compiling a topic (Putra & Kusumawardani, 2017).

Pearson Correlation

Pearson's correlation coefficient is a statistical measure that quantifies the strength and direction of the linear link between two variables. This correlation coefficient was employed to enhance the comprehension of the linear association between the two variables.

Correlation can exhibit a positive relationship, indicating that both variables tend to vary in the same direction, or a negative relationship, indicating that as one variable grows, the other tends to decrease. Alternatively, correlation can exhibit a neutral or negative value, indicating the absence of a relationship between the variables (Brownlee, 2019). The formula for the Pearson correlation, as in (1).

$$\textit{Pearson's correlation coefficient} = \frac{\mathrm{cov}(x,\ y)}{\mathrm{stdev}\ (x)\ \mathrm{X}\ \mathrm{stdev}\ (y)} \tag{1}$$

The interpretation of the correlation coefficient has a value ranging from +1 to -1. A value of 0 indicates that the two variables have no correlation.

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

Pearson correlation is a statistical hypothesis test based on the null hypothesis that there is no relationship between the samples.

The p-value is used in hypothesis testing and can be interpreted as follows:

- p-value ≤ alpha: significant results, reject null hypothesis; relationship (H_1) exists.
- p-value > alpha: The results are not significant; the null hypothesis cannot be rejected; there is no relationship (H_0).

Linear Regression

Subsequently, statistical models or forecasting techniques, such as regression analysis, were employed to conduct a predictive analysis (Savirani et al., 2021).

The basic linear regression model is employed to represent the linear association between the independent variables, which is postulated to have an impact on the dependent variable measured on a metric scale (Sekaran & Bougie, 2016).

Hence, the articulation of hypotheses for evaluating the appropriateness of the model, also known as model adequacy, is presented in equations (2) and (3).

$$H_0: a = \beta = 0 \text{ (the model is not suitable)} \quad (2)$$

$$H_1: \text{at least one sign} \neq \text{(the model suitable)} \quad (3)$$

If H_0 accepted, then a simple linear regression model like 4 below.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i \quad (4)$$

Hypothesis and Conceptual Framework

The conceptual framework underlying this research can be seen in Figure 1 below.

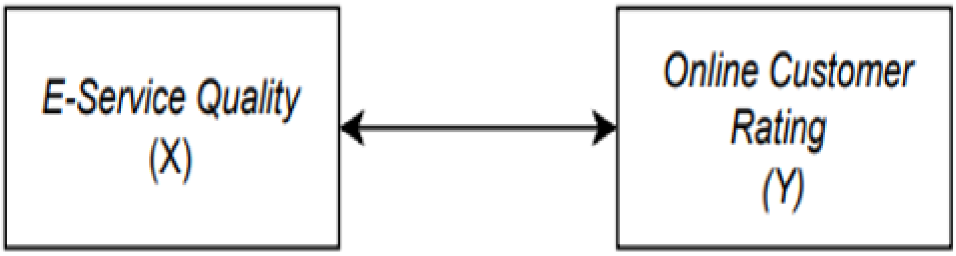


Figure 1: Conceptual Framework

Figure 1 displays the rationale for this research. This research is based on a problem formulation, namely that the emergence of TikTok in Indonesia has raised various issues regarding TikTok being more popular than YouTube and its popularity will shift YouTube.

Apart from that, the quality of service provided by YouTube and TikTok to users is still not optimal. Good quality will produce customer satisfaction, thereby increasing customer retention and loyalty in using the application (Suciptawati et al., 2019).

The Seybold REPORT
ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

Customer complaints in reviews and ratings describe the user's perception of the product or service that the company has provided to him (Sari et al., 2018). Google Play Store is present as an application download platform, including YouTube and TikTok which can be downloaded here. On this platform, users can write reviews and provide ratings on applications. The reviews and ratings provided are open-source. This research collects data from the Google Play Store.

Next, the researchers carried out data processing and data analysis using big data methods, namely sentiment analysis with the Naïve Bayes algorithm based on research (Alamsyah et al., 2018) and using topic modeling with the Latent Dirichlet Allocation (LDA) algorithm.

Topic modeling was carried out to see the main topics conveyed by users regarding the YouTube and TikTok applications. In the data processing process, the review data will be classified using sentiment analysis, topic modeling with E-Servqual dimensions. The E-Servqual classification used is ES-Qual and E-Recsqual, namely Efficiency, Fulfillment, System Availability, Privacy. Responsiveness, Compensation, and Contact.

Researchers also carried out predictive analysis by measuring the correlation between electronic service quality as a variable (X) and online customer ratings as a variable (Y), using logistic regression.

The results of this research will show the sentiments and mainstream topics conveyed by YouTube and TikTok users based on the E-Servqual dimensions, as well as determine the correlation between the quality of electronic services and online customer ratings. These results can be used as a reference in evaluating the quality of service on YouTube and TikTok.

In the company's efforts to produce high customer ratings, the company must provide and ensure the quality of service provided to customers operates well. Nowadays, online reviews have become an important factor in influencing the evaluation of service quality and play an important role for both customers and service providers.

Thus, managers must be sensitive to the dimensions of service quality that drive reviews and ratings to manage their online reputation. Although all dimensions of E-Servqual are important, there are differences in how these dimensions influence perceived customer satisfaction (Gunasekar et al., 2021).

Research by Brochado et al., (2019) examined online customer ratings and the quality of services provided in air transportation (airplanes). The current results show that customers classified their trip value as very good or very bad.

Passengers who provided low ratings wrote about negative experiences with airport operations and in-flight services, which can arise from factors over which the airline cannot control (e.g. delays and third-party ground operators). Therefore, companies need to know the number of negative incidents related to these themes, identify the root causes, and implement plans to minimize the negative impact on the consumer experience.

H1: There is a correlation between Service Quality and User Ratings

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

METHODOLOGY

The research workflow underlying this research can be seen in Figure 2 below.

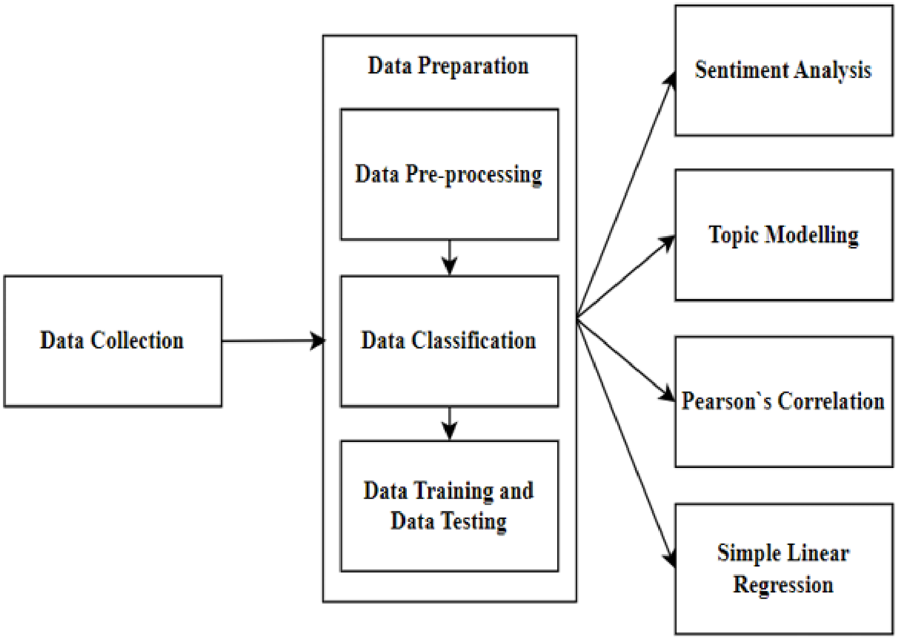


Figure 2: Research Workflow

As shown in Figure 2, the research workflow consists of six stages. The first stage begins with data collection, followed by the second stage, namely data preparation which consists of data pre-processing, data classification, creating training data and test data. The third stage carries out sentiment analysis using the Naïve Bayes machine learning model with Python scripts. Python is a well-known programming language in data science and provides extensive support for related functionality (Zahra, 2021). The fourth stage carries out topic modeling using Latent Dirichlet Allocation (LDA). The fifth stage measures the Pearson correlation between the independent variable in the form of service quality based on sentiment (X) and rating as the dependent variable (Y). The sixth stage is linear regression measurement.

This research uses a dataset from user reviews of the YouTube and TikTok applications on Google Play in Indonesian. The data collection period lasted from January 2023 to May 2023. We obtained 8,540 YouTube review data and 12,816 TikTok review data. The collected data undergoes cleaning during the preprocessing phase, which includes four steps, as illustrated in Figure 2. The user review data collected by the Python script contains some unnecessary formatting. Some properties should be removed and renamed for sentiment analysis and topic modeling. We removed properties like "userName", "score", "at", and "content". The final properties are “review” and “sentiment”. We renamed the “content” property to “review”. We introduce an additional property “sentiment”, and determine its value manually based on positive, neutral, and negative. The basic and final properties are shown in Tables 2 and 3, respectively.

The Seybold REPORT

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

Table 2: Basic Structure Data

Properties	Description	Example
userName	Name of the users review	Bahtiar FirmanSyah
score	The rating given by the customer	1
at	Date and the time of the review	1/1/2023
content	The review given by the customer	ini error apa gimana, kalau video di landscape terus kita keluar maka aplikasi jadi landscape juga. padahal dulu nggk lah.

Source: Result of data processing (2023)

Table 3: Final Structure Data

Properties	Description	Example
review	User Review	ini error apa gimana, kalau video di landscape terus kita keluar maka aplikasi jadi landscape juga. padahal dulu nggk lah.
sentiment	Classification for specific texts	Positif

Source: Result of data processing (2023)

RESULTS AND DISCUSSION

This research uses secondary data as a data source originating from the Google Play Store. Data was collected using data scrapping techniques via the Google Collaboratory website. The data collected contains ratings and reviews from users regarding their experience in the quality of services provided by YouTube and TikTok for the period 1 January 2023 to 30 May 2023. The results of scrapping data on the Google Collaboratory regarding two video player and editor apps platforms are: as follows:

Table 4: Scraping Review Results on Google Play Store

No.	Nama Aplikasi	Jumlah Data
1.	YouTube	8.540
2.	TikTok	12.816
Total Data Ulasan		21.358

Source: Result of data processing (2023)

Table 4 above shows that during the data collection period the researcher obtained data on 21,358 reviews. This number is still raw data and still needs to be processed. Next, researchers will carry out data preprocessing, namely cleaning noise data, in the form of spam reviews and irrelevant reviews. The main goal of this stage is to prepare higher quality data when entering the data processing stage. Then, data classification will be carried out based on dimensions, sentiment analysis and topic modeling, while rating data is used to analyze whether there is a

ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

correlation between service quality using user sentiment and user ratings. Next, data analysis and conclusions will be drawn.

Data Pre-processing

Pre-processing is the stage for preparing raw data into data that is ready to be processed. The first stage of pre-processing is data cleansing. Data cleaning is carried out to delete reviews in the form of spam, delete symbols, ambiguous data and data that is not relevant to research, such as discussions about other application services. The pre-processing results can be seen in table 5 below.

Table 5: Pre-processing Result

No.	Nama Aplikasi	Before Pre-processing	After Pre-processing
1.	YouTube	8.540	8.305
2.	TikTok	12.816	12.303
Total		**21.356**	**20.608**

Source: Result of data processing (2023)

Data preprocessing was carried out using Google Colaboratory in Python as seen in Figure 3 below:

Figure 3: Pre-processing Data

Source: Result of data processing (2023)

Figure 3 shows the stages of data pre-processing with several stages, namely:

1) Tokenization: dividing sentences into words, phrases, or symbols.
2) Filtering: to remove unnecessary stopwords or text attributes such as \`dan\`, \`atau\`, \`dari\`, \`untuk\`.
3) Stemming: changing words into basic words, such as \`berjualan\` becomes \`jual\`, \`membeli\` menjadi \`beli\`.

The Seybold REPORT

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

Data Labelling

Data labeling is determining data based on data characteristics by reading reviews given by users on the Google Play Store and categorizing these reviews into one of the E-Servqual dimensions according to the characteristics and indicators of the review. Determination of labels in each review is carried out based on indicators and operational variables, where each dimension has a definition based on (Chatterjee, 2019; Parasuraman et al., 2005; Zahra, 2021) which is the criteria for determining keywords. Each review will only be classified into one sentiment and one dimension. This aims to prevent repetition of review data.

This process was carried out manually by researchers together with other researchers by applying keywords for positive, negative and neutral labels obtained from previous research (Zahra, 2021) and peer debriefing to reduce researcher subjectivity. The labeling process can be seen in figure 4 below.

	A	B	C	D	E	F
1	Tiktok Review Negatif	Tiktok Review Positif	Negative Sentiment	Positive Sentiment	Youtube Review Positif	Youtube Review Negative
2	buka kira aplikasiku kep	buka kira aplikasiku kep	aneh	asik	baru setengah didownlo	baru setengah didownloac
3	cari video terjemahan c	cari video terjemahan c	risih	bagus	sakit hati subscribe iklar	sakit hati subscribe iklan w
4	tiktok zeh mala tambah	tiktok zeh mala tambah	berat	berfaidah	ngupdate lihat videovid	ngupdate lihat videovideo
5	tiktok fitur emosi bagia	tiktok fitur emosi bagiar	babi	berguna	komentar mengetik hab	komentar mengetik habis
6	jadi aplikasi bagus seha	jadi aplikasi bagus seha	bacot	bermanfaat	bintangnya youtube dic	bintangnya youtube didow
7	aplikasi tiktoknya tamb	aplikasi tiktoknya tamb	bangsat	bermutu	aplikasi youtube jaringa	aplikasi youtube jaringan
8	oky tiktok bagus cuman	oky tiktok bagus cuman	bug	informatif	bug update fitur subscri	bug update fitur subscripti
9	overal aplikasi bagus cu	overal aplikasi bagus cu	banned	inspirasi	bug aneh membuatku n	bug aneh membuatku mar
10	sinyla bagus tapi buka r	sinyla bagus tapi buka r	batal	keren	bug dibagian video didc	bug dibagian video didowr
11	parah tiktok kesini jelek	parah tiktok kesini jelek	takedown	menghibur	bug dimana sat menulis	bug dimana sat menulis kc
12	akun suka ngelag gelag	akun suka ngelag gelag j	block	menginspirasi	bug search search dowr	bug search search downlo
13	live lag lag jaringan bag	live lag lag jaringan bag	blokir	puas	bug tonton youtube ikl	bug tonton youtube iklan i
14	buka tiktok knapa vide	buka tiktok knapa video	blur	senang	bug serunya tonton tib	bug serunya tonton tiba ti
15	keluar aplikasi pencet p	keluar aplikasi pencet p	boros	sukses	bug otomatis keluar shc	bug otomatis keluar short
16	bagus fyp mpang rusak	bagus fyp mpang rusak	tambah	wajib	bug youtube short vide	bug youtube short video h
17	pedahal sinyal bagus te	pedahal sinyal bagus te	buram	wawasan	bug tonton video terus	bug tonton video terus tor
18	pelu tingkatkan kinerja	pelu tingkatkan kinerjar	buruk	menyukai	bug pause sepertiuntuk	bug pause sepertiuntuk m

Figure 4: Labeled Data Process

Source: Result of data processing (2023)

Figure 4 determining words for positive and negative keywords is based on literature studies from previous research. After determining keywords for positive and negative, the researcher carried out conditional formatting for each word with positive keywords colored green, negative keywords red and green-red for neutral. This data labeling process is carried out so that machine learning can be helped to study the data in determining data categories into e-servqual dimensions.

Then, proceed with classifying the data. Data that has been labeled will be divided into training data and testing data with a ratio of 80:20, of which 80% of the data obtained is for training data and 20% for testing data. YouTube has 6,644 training data and 1,661 testing data, while TikTok has 9,842 training data and 2,461 testing data.

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

Classification Performance

After data classification, the data is divided into training data and testing data in a ratio of 80:20, where 80% of the testing data has been labeled and 20% of the testing data will be tested. Then, the data is input into a script in Google Collaboratory to produce performance. The confusion matrix is used to measure the performance of a model. This research uses the Naïve Bayes algorithm, the results can be seen in table 6 below.

Table 6: Confusion Matrix YouTube

Accuracy: 75.56%

	True Positive	True Neutral	True Negative	Class Precision
Pred. Positive	4044	376	110	80.22%
Pred. Neutral	615	1447	149	75.16%
Pred. Negative	191	488	784	60.16%
Class Recall	87.15%	50.15%	65.44%	

Source: Result of data processing (2023)

Table 6 above shows the model performance for YouTube with an accuracy value of 80.33%.

Table 7: Confusion Matrix TikTok

Accuracy: 80.33%

	True Positive	True Neutral	True Negative	Class Precision
Pred. Positive	3609	493	81	77%
Pred. Neutral	502	2409	493	87.13%
Pred. Negative	288	568	3886	69.42%
Class Recall	84.63%	81.94%	70.6%	

Source: Result of data processing (2023)

Table 7 above shows the model performance for TikTok with an accuracy value of 80.33%. According to Putri (2020), the performance of the model using naïve Bayes can be said to be good and valid if the accuracy value is above 75%, so it can be said that the accuracy of this YouTube TikTok model is good and valid because it has a value above 75%, namely 75.56% and 80.33%.

Data Classification Based on Sentiment

Sentiment analysis is carried out to classify user opinions or reviews into positive, negative and neutral based on the words contained in them (Christanto & Singgalen, 2022). Figure 5 below displays the results of the sentiment analysis that has been carried out, based on a dataset of YouTube application user reviews.

ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

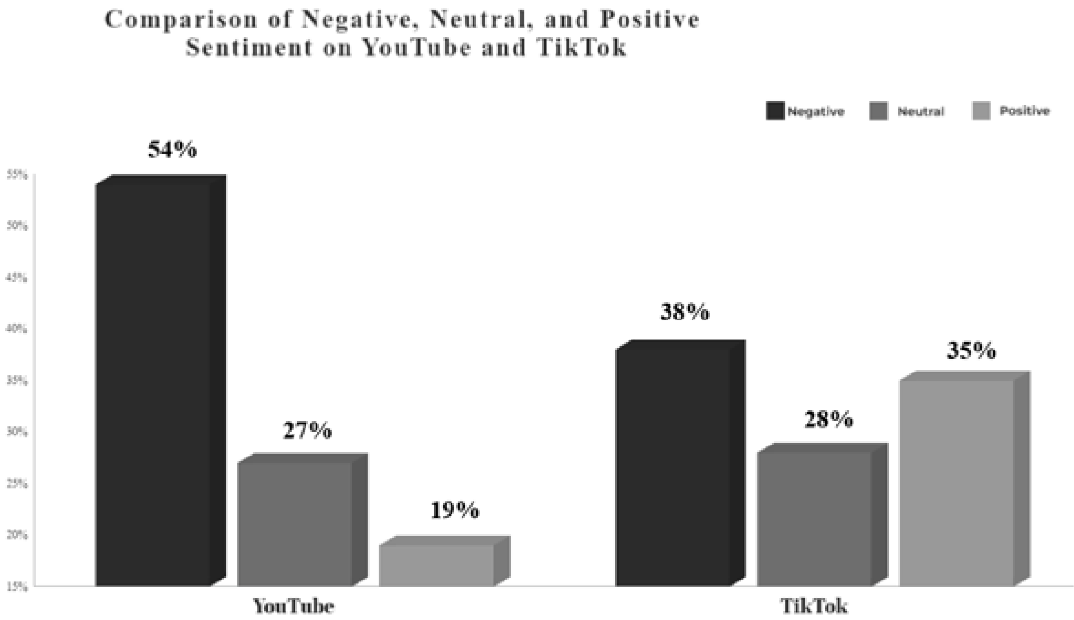


Figure 5: Comparison YouTube and TikTok Sentiment Classification

Source: Result of data processing (2023)

Figure 5 shows that 19% of the review data regarding the quality of YouTube services have positive sentiments, 27% have neutral sentiments, and 54% have negative sentiments. It can be seen that the largest percentage is owned by negative sentiment, which means that many users are dissatisfied with the quality of the YouTube services provided. Meanwhile, sentiment classification from the TikTok application user review dataset, a total of 12,303 review data regarding the quality of TikTok services had a negative sentiment of 37%, a neutral sentiment of 28% and a positive sentiment of 36%. It can be seen that the largest percentage is owned by negative sentiment, which means that many users are dissatisfied with the quality of the TikTok services provided.

Data Classification Based on Dimensions

Figure 6 below shows the percentage of dimensions appearing in YouTube user reviews and Figure 7 below shows the proportion of sentiment from each ESQual and ErescSQual dimension towards the quality of services provided by YouTube based on the reviews they wrote on the Google Play Store.

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

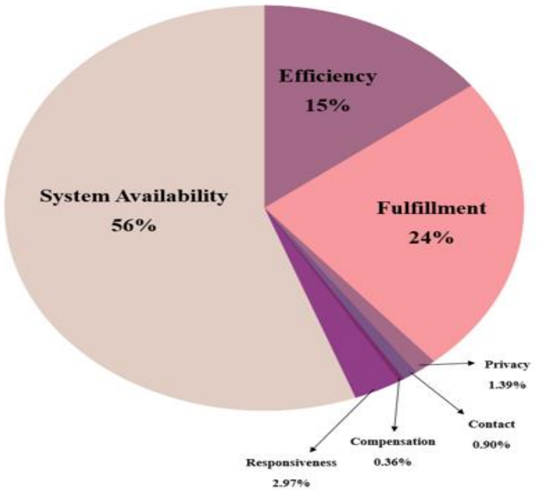


Figure 6: Proportion of 7 YouTube E-Servqual Dimensions

Source: Result of data processing (2023)

Figure 6 above shows that there are three dimensions that are most dominant or most talked about by YouTube users on the Google Play Store, namely system availability, fulfillment, and efficiency with a percentage order of 56%, 25%, and 15%. These dimensions are included in the ES-Qual scale, which is a scale that measures the company's core services. In fourth place there is responsiveness with a percentage of 2.97%. This responsiveness is included in the E-Resc-Qual scale, which is a scale that measures how well a company handles the problems faced by its users. The compensation, privacy and contact dimensions are in third place with a percentage of less than 2%. Furthermore, Figure 7 displays the proportion of user perceptions which describes users' feelings about the quality of services provided by YouTube based on the reviews they wrote on the Google Play Store.

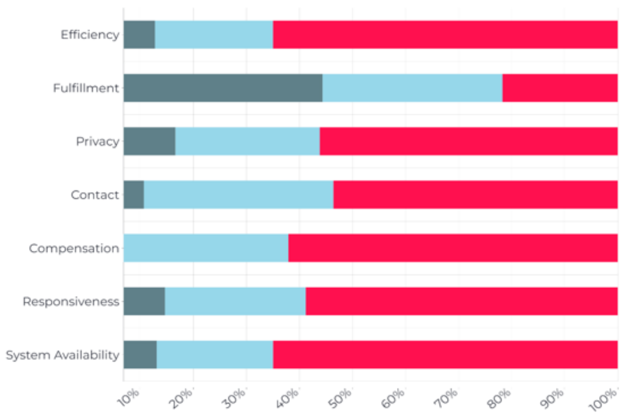


Figure 7: Proportion of Sentiment from 7 Dimensions of YouTube E-Servqual

Source: Result of data processing (2023)

ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

In Figure 7 above, red indicates negative sentiment, blue indicates neutral sentiment, and green indicates positive sentiment. When compared to the dimensions of service quality, the one that has the best performance is fulfillment.

This is indicated by the highest proportion of positive sentiment at 44%, the most neutral sentiment is fulfillment with a percentage of 33%, and the service quality dimension that has the most negative sentiment is efficiency with a proportion of 66%.

In sequence, the most negative sentiments are the dimensions of system availability, contact, compensation, responsiveness, privacy and fulfillment with proportion values of 64%, 64%, 59%, 56% and 21%. From these results, it is known that users' perceptions of 6 of the 7 dimensions of YouTube service quality are not good.

Figure 8 below shows the percentage of dimensions appearing in TikTok user reviews.

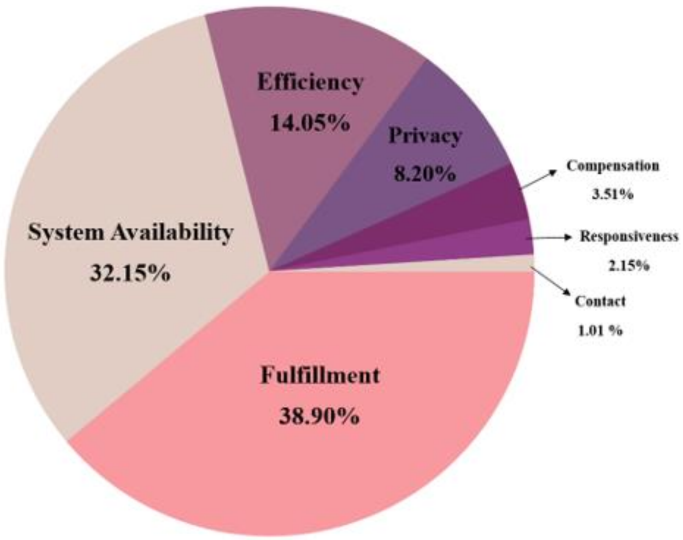


Figure 8: Proportion of Sentiment from 7 Dimensions of TikTok E-Servqual

Source: Result of data processing (2023)

Figure 8 above shows that there are three dimensions that are most dominant or most talked about by TikTok users on the Google Play Store, namely fulfillment, system availability, and efficiency with a percentage order of 38.90%, 32.15%, and 14.05%.

These dimensions are included in the ES-Qual scale, which is a scale that measures the company's core services.

In fourth place is privacy with a percentage of 8.20%. The dimensions of compensation, responsiveness and contact are in third place with a percentage of less than 2%.

These three dimensions are included in the E-Resc-Qual scale, which is a scale that measures how well a company handles the problems faced by its users.

ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

Figure 9 below shows the proportion of users' perceptions of the quality of service provided by TikTok based on the reviews they wrote on the Google Play Store.

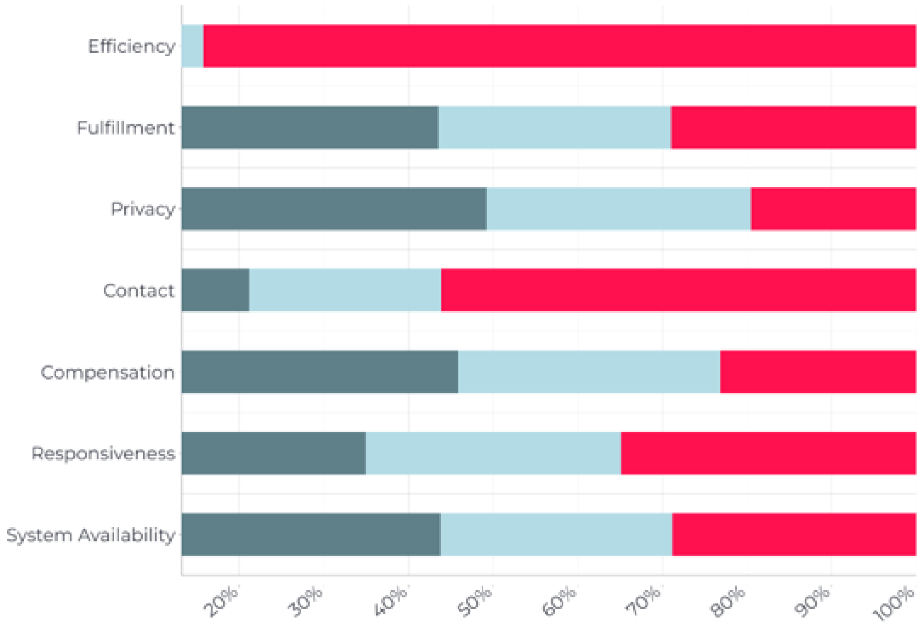


Figure 9: Proportion of Sentiment from 7 Dimensions of TikTok E-Servqual

Source: Result of data processing (2023)

In Figure 9, red indicates negative sentiment, blue indicates neutral sentiment, and green indicates positive sentiment. When compared to the dimensions of service quality, the one that performs best is privacy. This is characterized by the highest proportion of positive sentiment at 49%. Meanwhile, the service quality dimension that has the worst performance because it has the most negative sentiments is efficiency with a proportion of 92%, sequentially the most negative sentiments are the dimensions of contact, responsiveness, system availability, compensation and privacy with a proportion value of 56%, 38 %, 29%, 23%, and 20%. From these results, it is known that user perceptions regarding the efficiency dimension of TikTok's service quality are not good. TikTok's service quality performance still needs to be improved both in terms of fulfilling service promises (ESQual) and in terms of dealing with problems experienced by customers (ErescSQual).

Topic Modelling

The dimensions most frequently discussed in this review are efficiency, compliance, and system availability, as revealed by the classification per dimension and sentiment analysis above. As shown in Tables 8 to 13, we focus on these three aspects for topic modeling. Table 8 presents aspects of YouTube's topic modeling efficiency, revealing a tendency to overemphasize ads and intrusive updates. Table 9 presents aspects of TikTok's topic modeling efficiency, showing a tendency to overemphasize the number of effects and filters. In particular, TikTok effects and filters are seen as fun and humorous. Table 10 describes the fulfillment

The Seybold REPORT

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

dimensions of YouTube topic modeling, tending to highlight the increasing quality, highly useful nature of information available on YouTube.

Table 11 explains the dimensions of TikTok's topic modeling fulfillment; it tends to highlight improvements in quality. TikTok and TikTok Shop have gained significant popularity due to their perceived use and efficacy. Table 12 presents the system availability (SA) aspects of YouTube topic modeling, highlighting identified improvement needs related to bug fixes, downloads, and unexpected outages. Table 13 presents the system availability (SA) aspects of TikTok topic modeling, placing special emphasis on the identified improvement requirements, which are in line with those observed on YouTube. These terms include bug fixes and addressing unexpected outages.

Table 8: Topic Modeling Youtube Efficiency Dimension

Top Topic	Top word					Most Relevant Term Distribution
Topic 0	tolong	banyak	iklannya	aneh	skip	29.6%
Topic 1	update	tolong	kurangi	banyak	mengganggu	24.1%
Topic 2	update	tonton	jelek	skip	aneh	23.8%
Topic 3	iklannya	tonton	update	aneh	detik	22.5%

Table 9: Topic Modeling TikTok Efficiency Dimension

Top Topic	Top word					Most Relevant Term Distribution
Topic 0	mantap	keren	efek	enak	suka	27.5%
Topic 1	senang	best	filter	mantap	efeknya	26.8%
Topic 2	mantap	hiburan	filter	suka	seru	26.2%
Topic 3	wawasan	best	efek	positif	mantap	19.5%

Table 10: Topic Modeling YouTube Fulfillment Dimension

Top Topic	Top word					Most Relevant Term Distribution
Topic 0	perbaiki	tonton	short	tolong	suka	27.5%
Topic 1	boros	kuota	skip	malas	suka	26.8%
Topic 2	informasi	membantu	filter	mohon	seru	26.2%
Topic 3	kualitas	tolong	perbaiki	bermanfaat	muncul	19.5%

Table 11: Topic Modeling TikTok Fulfillment Dimension

Top Topic	Top word					Most Relevant Term Distribution
Topic 0	menarik	membantu	suka	senang	terima_kasih	27.5%
Topic 1	tolong	perbaiki	update	lelet	susah	26.8%
Topic 2	perbaiki	suka	patah	update	mohon	26.2%
Topic 3	kualitas	shop	suka	susah	senang	19.5%

Table 12: Topic Modeling YouTube System Availability Dimension

Top Topic	Top word					Most Relevant Term Distribution
Topic 0	bug	buka	perbaiki	terus	keluar	28.5%
Topic 1	tolong	download	offline	keluar	tiba	24.5%
Topic 2	download	force_close	kurang	kualitas	kembali	24%
Topic 3	perbaiki	lelet	tolong	kembali	keluar	23.1%

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

Table 13: Topic Modeling TikTok System Availability Dimension

Top Topic	Top word					Most Relevant Term Distribution
Topic 0	perbaiki	tolong	bug	tiba	keluar	30%
Topic 1	fyp	jadi	lelet	patah	bug	27.2%
Topic 2	perbaiki	macet	keluar	tiba	tolong	23.4%
Topic 3	ngelag	susah	lelet	bug	keluar	19.5%

Source: Result of data processing (2023)

Pearson Correlation

Pearson correlation test will be carried out using the independent variable in the form of service quality based on sentiment (X) and rating as the dependent variable (Y). The Pearson correlation test aims to measure the extent to which two numerical variables on an interval or ratio scale have a linear relationship with each other. Pearson correlation helps in understanding whether there is a statistically significant relationship between the two variables and how strong or weak the relationship is. Figure 10 below displays the results of the YouTube Pearson correlation test using the Python programming language in Google Collaboratory.

Creating Correlation of Sentiment & Rating

```
[5] # Calculate Pearson correlation coefficient and p-value
    pearson_corr, pearson_p_value = pearsonr(rating, sentiment)

    # Calculate Spearman rank correlation coefficient and p-value
    spearman_corr, spearman_p_value = spearmanr(rating, sentiment)

    # Print the correlation results
    print(f'Pearson Correlation Coefficient: {pearson_corr}')
    print(f'Pearson p-value: {pearson_p_value}')

    Pearson Correlation Coefficient: 0.9076242613888645
    Pearson p-value: 0.0
```

Figure 10: YouTube Pearson Correlation Results

Source: Result of data processing (2023)

In figure 10 above shows:

Correlation coefficient or r value (Pearson correlation) is used to measure the level of strength of the relationship between variables and the direction of the variable relationship. The correlation coefficient value ranges between negative one (-1) and positive (1). The closer it is to 1, the closer the correlation value is to perfect and the positive or negative value indicates the direction of the relationship.

ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

The following is an interpretation of the correlation figures.

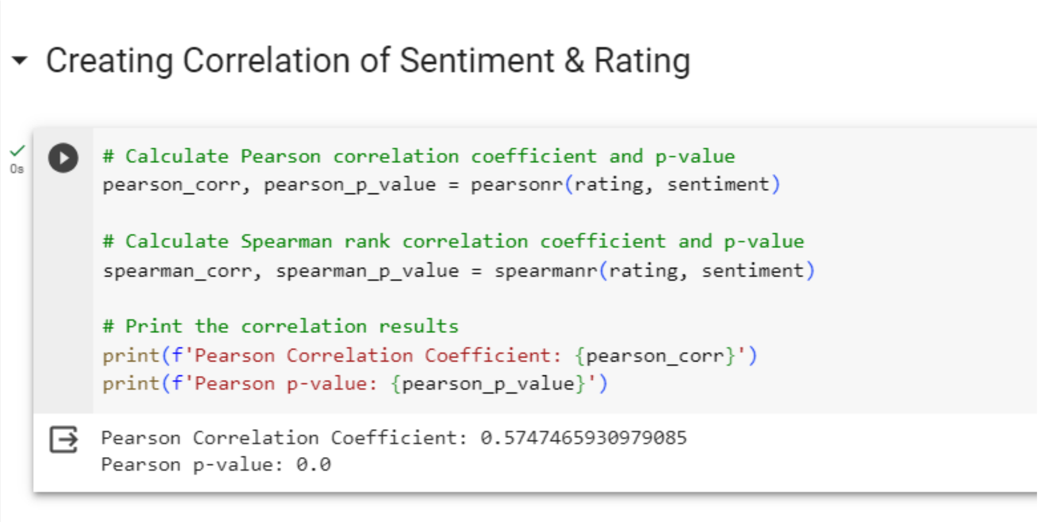
- 0.9 to close to 1 (positive or negative) indicates a very high degree of relationship.
- 0.7 to 0.8 (positive or negative) indicates a high degree of relationship.
- 0.5 to 0.6 (positive or negative) indicates a moderate degree of relationship.
- 0.3 to 0.4 (positive or negative) indicates a low degree of relationship.
- 0.1 to 0.2 (positive or negative) indicates a very low degree of relationship.
- 0.0 means the two variables do not have a linear relationship.

• P-value is a significance level used to measure the significant level of relationship between variables.

This test was carried out to see whether there was a correlation between the service quality variables using sentiment (x) and rating (y). This test uses a sample of 8,305 data.

• Pearson correlation results. The correlation value, or correlation coefficient, shows the strength of the relationship between service quality using sentiment (x) and rating (y) is positive 0.90. The magnitude of the correlation number shows that the correlation between service quality using sentiment (x) and rating (y) is in the strong relationship category, while a positive value indicates the direction of the relationship is in the same direction, meaning that the better the quality of service the user sentiment provided increases, the higher the rating given by users tend to move up at a perfect rate. Apart from that, a significance value of 0.00 was obtained, which is smaller than 0.05, meaning that there is a significant relationship between the service quality variable using sentiment (x) and rating (y).

Next, in figure 11 below, the results of the TikTok Pearson correlation test are displayed.


Creating Correlation of Sentiment & Rating

```
# Calculate Pearson correlation coefficient and p-value
pearson_corr, pearson_p_value = pearsonr(rating, sentiment)

# Calculate Spearman rank correlation coefficient and p-value
spearman_corr, spearman_p_value = spearmanr(rating, sentiment)

# Print the correlation results
print(f'Pearson Correlation Coefficient: {pearson_corr}')
print(f'Pearson p-value: {pearson_p_value}')
```

```
Pearson Correlation Coefficient: 0.5747465930979085
Pearson p-value: 0.0
```

Figure 11: TikTok Pearson Correlation Results

Source: Result of data processing (2023)

ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

In figure 11 it shows:

- Correlation coefficient or r value (Pearson correlation) is used to measure the level of strength of the relationship between variables and the direction of the variable relationship. The correlation coefficient value ranges between negative one (-1) and positive (1). The closer it is to 1, the closer the correlation value is to perfect and the positive or negative value indicates the direction of the relationship. The following is an interpretation of the correlation figures.
  - 0.9 to close to 1 (positive or negative) indicates a very high degree of relationship.
  - 0.7 to 0.8 (positive or negative) indicates a high degree of relationship.
  - 0.5 to 0.6 (positive or negative) indicates a moderate degree of relationship.
  - 0.3 to 0.4 (positive or negative) indicates a low degree of relationship.
  - 0.1 to 0.2 (positive or negative) indicates a very low degree of relationship.
  - 0.0 means the two variables do not have a linear relationship.
- P-value is a significance level used to measure the significant level of relationship between variables.

This test was carried out to see whether there was a correlation between service quality variables using sentiment (x) and rating (y). This test uses a sample of 12,303 data.

- Pearson correlation results. The correlation value, or correlation coefficient, shows the strength of the relationship between service quality using sentiment (x) and rating (y) is positive 0.55. The magnitude of the correlation number shows that the correlation between service quality using sentiment (x) and rating (y) is in the medium relationship category, while a positive value indicates the direction of the relationship is in the same direction, meaning that when the quality of service provided by users increases, the rating given by users tends to increase. rising, but not at a perfect rate. Apart from that, a significance value of 0.00 was obtained, which is smaller than 0.05, meaning that there is a significant relationship between the service quality variable using sentiment (x) and rating (y).

Simple Linear Regression

In Figure 12 below, the results of a linear regression between service quality using sentiment (x) and rating as a variable (y) are displayed. With the following simple linear regression equation model (5).

$$Y = a + bX \qquad (5)$$

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

Figure 12: YouTube Simple Linear Regression Results

Source: Result of data processing (2023)

Figure 12 above displays the test results using linear regression on YouTube with the systematic regression equation being Y = 2.84 + 1.503X, the explanation is as follows.

- Mean R-Squared Score is the extent to which the linear regression model fits the existing data. The R-squared (R2) value ranges from 0 to 1, the higher the value, the better the model is at explaining variations in the data.
- F-score is the significance level of the regression model.
- P-value is the level of significance used to measure the significant level of relationship between variables.
- Slope is used to measure how much influence changes in the independent variable have on the dependent variable.

This test was carried out to see whether this model could be used to predict the independent variable (service quality using sentiment) based on the dependent variable (rating). The linear regression results in Figure 4.83 above show that the Mean R-squared value is around 0.82, indicating that this model can explain around 82% of the variation in the data. This shows that the model has a good ability to explain the independent variable (service quality using sentiment) and the dependent variable (rating). The high F-score value of 3810.65 shows that your model can predict the "rating" value well. This shows that the model as a whole is a significant model. The P-value is 0.00 (zero), this indicates that "rating" has a significant impact on the dependent variable. This coefficient value is positive 1.5035 which indicates that every increase in service quality uses sentiment to contribute positively around 1.5035 units to the "rating" variable.

ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

It can be stated that the service quality variable using sentiment has a significant impact on the independent variable, and the relationship between the two is positive. The "rating" coefficient (slope) is around 2.8405, which shows a significant positive relationship between service quality using sentiment and rating. Intercept is the predicted value when all independent variables are zero. Thus, these results indicate that this linear regression model can be used to predict the dependent variable (service quality using sentiment) based on the "rating" variable with high statistical significance.

Furthermore, in Figure 13 below, the results of the linear regression between service quality using sentiment (x) and rating as a variable (y) are displayed. With the following simple splenic regression equation model (6).

$$Y = a + bX \qquad (6)$$

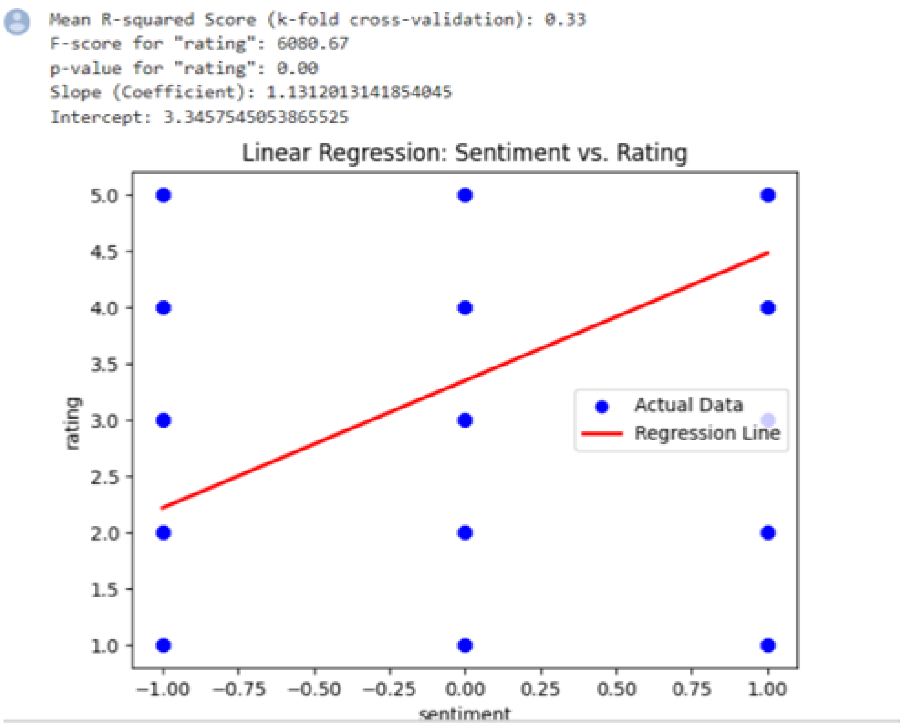


Figure 13: TikTok Simple Linear Regression Results

Source: Result of data processing (2023)

Figure 13 above displays the test results using linear regression on TikTok with the systematic regression equation being Y = 3,345 + 1,131X, the explanation is as follows.

- Men R-Squared Score is the extent to which the linear regression model fits the existing data. The R-squared (R2) value ranges from 0 to 1, the higher the value, the better the model is at explaining variations in the data.
- F-score is the significance level of the regression model.

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

- P-value is a significance level used to measure the significant level of relationship between variables.
- Slope is used to measure how much influence changes in the independent variable have on the dependent variable.

This test was carried out to see whether this model could be used to predict the independent variable (service quality using sentiment) based on the dependent variable (rating). The linear regression results in Figure 4.84 above show that the Mean R-squared value is around 0.33, indicating that this model can explain around 33% of the variation in the data.

This indicates that the model has a moderate level of fit to the data. The high F-score value is 6080.67. This shows that the model as a whole is a significant model. In this case, the overall regression model is useful for explaining variations in the data.

The P-value is 0.00 (zero), this indicates that "rating" has a significant impact on the dependent variable. This coefficient value is positive (1.1312), which indicates that there is a positive relationship between "rating" and the dependent variable. In other words, as the “rating” increases, the dependent variable also tends to increase at a rate of about 1.1312 units.

It can be stated that the "rating" variable has a significant impact on the dependent variable, and the relationship between the two is positive. The overall regression model was also significant in explaining the variation in the data, although only about 33% of the variation could be explained by this model.

The coefficient of “rating” (slope) is approximately 1.1312, which indicates a significant positive relationship between “rating” and the dependent variable. Intercept is the predicted value when all independent variables are zero. Thus, these results indicate that this linear regression model can be used to predict the dependent variable (service quality) based on the "rating" variable with high statistical significance.

However, it is important to remember that this model can only explain some of the variation in the data, and there are other factors that may influence the dependent variable that are not included in this model.

Based on the results, it can be seen that:

1) This research uses data obtained from crawling user review data on the Google Play Store using Google Collaboratory using the Python programming language for the time period from 1 January 2023 to 30 May 2023. The data obtained was 8,540 YouTube review data and 12,816 TikTok review data. After data collection, researchers carried out data pre-processing. The amount of data after pre-processing for YouTube is 8,305 data and for TikTok is 12,303 data. From the results of the classification analysis of the two applications' datasets, it is shown that the YouTube application discusses system availability (56%) and fulfillment (24%) the most, while TikTok discusses system availability (38.90%) and fulfillment (32.15%) the most. This shows that from the review data, TikTok and YouTube mostly discuss system availability and fulfillment.

The Seybold REPORT

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

2) The results of sentiment analysis on the dataset for the two applications show that user reviews of the YouTube application in Figure 4.3 are dominated by negative sentiment at 54%, neutral sentiment at 27% and positive sentiment at 19%. The results of this sentiment analysis can support the results of research conducted (Musleh et al., 2023) where there was 63% negative sentiment from the YouTube review data used in the research, many YouTube users felt dissatisfied with YouTube. The efficiency dimension is the dimension that has the most negative sentiment, namely 66% compared to other dimensions. In sequence, the most negative sentiments are the dimensions of system availability, efficiency, compensation, responsiveness, privacy and fulfillment with proportion values of 64%, 64%, 59%, 56% and 21%. Meanwhile, user reviews of the TikTok application in figure 4.4 are dominated by negative sentiment at 37%, negative sentiment at 36% and neutral sentiment at 28%. The efficiency dimension is the dimension that has the most negative sentiment, namely 92% of the efficiency dimension reviews have negative sentiment. In sequence, the most negative sentiments are the dimensions of contact, responsiveness, system availability, compensation and privacy with proportion values of 56%, 38%, 29%, 23% and 20%.

3) To strengthen the statement from the results of sentiment analysis, the researcher carried out topic modeling. Topic modeling is carried out to find out what topics are formed in each e-servqual dimension. The results of YouTube's topic modeling based on the efficiency dimension show that YouTube often asks for updates and has lots of advertisements, thus disturbing user comfort, which is represented by the words 'update' and 'advertisement' which always appear in every topic circle on this efficiency dimension. Meanwhile, TikTok shows that TikTok is a cool and useful application, but unfortunately it suddenly asks for updates which are represented by the words 'like', 'suddenly' and 'update' which always appear in every topic circle on this efficiency dimension. In the fulfillment dimension, YouTube shows that after the YouTube update is more exciting, but there are more and more advertisements so that the quota is wasted, which is represented by the words 'exciting', 'ads' and 'wasteful_quota'. Meanwhile, TikTok shows that TikTok is an easy-to-use application, the content displayed by TikTok is fun, but since it was updated, the video is broken, slow and difficult to open. This is represented by the words 'easy' and 'patah_patah'. In the privacy dimension, YouTube shows that to use YouTube, verification, copyright and monetization are required so that the channel is not blocked, which is represented by the words 'verification', 'copyright', 'monetization' and 'block'. Meanwhile, TikTok shows that the TikTok application has suddenly been subject to copyright infringement and has been blocked, even though it has fulfilled the requirements represented by the words 'violation', 'copyright' and 'block'. In the system availability dimension, YouTube shows well but often has bugs, lags and force closes, which makes users disappointed. This is represented by the words 'nge_bug', 'nge_lag' and 'force_close'. Meanwhile, TikTok shows that the TikTok application when watching a video suddenly exits on its own, often lags, bugs and has serious errors, so users are lazy to open TikTok. This is represented by the words 'out', 'nge_lag', 'bug', 'error' and 'lazy'. In the responsiveness dimension, it displays please improve the response of the live chat service,

The Seybold REPORT
ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

reporting a problem and even being disappointed because they were not given a solution. This is represented by the words 'response', 'chat' and 'solution'. Meanwhile, TikTok shows that users are complaining because they cannot open TikTok, which requires an OTP code that is not sent to enter TikTok. They have complained that the response is not friendly, which makes them disappointed. This is represented by the words 'complain', 'response', and 'friendly'. In the compensation dimension, it displays please improve the response of the live chat service, reporting a problem and even being disappointed because they were not given a solution. This is represented by the words 'response', 'chat' and 'solution'. Meanwhile, TikTok shows that users have complained to TikTok because they need an OTP code that was not sent to enter TikTok. They have complained that the unfriendly response has made them disappointed. This is represented by the words 'complain', 'response', and 'friendly'. In the contact dimension, it displays please improve YouTube premium monetization, which is difficult to appear and annoying, which makes you disappointed. This is represented by the words 'monetization', 'premium' and 'difficult'. Meanwhile, TikTok shows that users are complaining to TikTok because TikTok is bad and makes them disappointed because they block accounts, photos or live, there is no email notification from TikTok. This is represented by the words 'notification', 'email', and 'disappointed'.

4) Comparison between the results of sentiment analysis and topic modeling shows that the quality of service provided by the YouTube and TikTok applications based on user review data on the Google Play Store is considered to be still not good, so improvements are needed regarding service quality, especially in the system availability dimension from the results of sentiment analysis and complaints from the topic modeling results in the form of bugs, lag and force close. Apart from that, to increase the efficiency dimension, YouTube and TikTok need to pay more attention to the quality of the services provided because based on the results of topic modeling, in this dimension user complaints focus on the application frequently requesting updates and advertisements that annoy users. To increase the privacy dimension, YouTube and TikTok need to pay more attention to copyright, verification and blocking.

5) After knowing what dimensions need to be improved, this can effectively increase user satisfaction so as to maintain user loyalty. This statement is supported by (Zahra, 2021) by conducting this research it can help find important information to improve service quality more accurately. Apart from that, paying attention to what topics frequently appear can improve user perception because each user can see information related to a service.

6) Based on the correlation results between service quality using sentiment (x) and rating (y) for YouTube, a Pearson coefficient of 0.9076 with a p-value of 0.0 was produced and TikTok had a Pearson coefficient of 0.5546 with a p-value of 0.0. These results show that there is a strong positive correlation between service quality using sentiments written by users (x) and ratings given by users (y) on both YouTube and TikTok platforms. The higher correlation on YouTube (0.9076) compared to TikTok (0.5747) indicates that the relationship between sentiment and ratings on YouTube is stronger. This means that the sentiments expressed by users on YouTube have a greater influence on the ratings they give

ISSN 1533-9211
DOI: 10.5281/zenodo.10300129

compared to TikTok users. The very low p-value (0.0) on both platforms indicates that this relationship is highly statistically significant, meaning these results did not occur by chance. So, both on TikTok and on YouTube, the feelings expressed by users have a significant influence on the ratings they give. However, this influence is stronger on YouTube than on TikTok.

7) The results of the linear regression analysis show that the YouTube regression model overall has better performance than the TikTok regression model. This is indicated by a higher R-squared score, a much higher F-score, and a lower intercept. A higher R-squared score for the YouTube model indicates that the model better explains the variability of the data.

CONCLUSION

A comparison between the results of sentiment analysis and topic modeling shows that the quality of service provided by the YouTube and TikTok applications based on user review data on the Google Play Store is considered to be still not good, so improvements are needed regarding service quality, especially in the efficiency dimension from the results of sentiment analysis and complaints from The results of topic modeling are updates, advertisements, and slowdowns. Next, Pearson correlation measurements were carried out to determine the extent to which service quality variables using sentiment (x) are related to each other's ratings (y). The results show that there is a strong correlation between the feelings expressed by users and the ratings given by users on both platforms TikTok and YouTube. This suggests that the positive or negative feelings expressed by users can influence how they rate content or experiences on the platform. The YouTube regression model overall performs better than the TikTok regression model. This is indicated by a higher R-squared score, a much higher F-score, and a lower intercept. A higher R-squared score for the YouTube model indicates that the model better explains the variability of the data.

The researcher realizes that there are limitations in this research, therefore it is hoped that future researchers will be more exploratory. For example, using different research objects and other data sources, collecting data for a longer duration and using the Indobert method or other methods in accordance with technological developments. YouTube and TikTok can consider the results of this research to evaluate service quality, especially e-servqual, to increase user satisfaction in order to maintain customer loyalty and improve user perception. In conclusion, it has been explained that the main priority for service improvements can be made in the efficiency dimension, where there are many complaints regarding sudden updates and advertisements, so that they interfere when users watch the application.

Therefore, the improvement effort that can be recommended is to evaluate the management sector to determine an application update schedule that is not close together and reduce the duration of advertisements. In second place there is the system availability dimension, YouTube and TikTok can make regular system improvements to avoid bugs, lags and slowness when used by users. In the fulfillment dimension, YouTube and TikTok can make improvements in providing maximum feature performance and privacy dimensions, improving security, monetization and blocking.

The Seybold REPORT

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

References

1) Alamsyah, A., Rizkika, W., Nugroho, D. D. A., Renaldi, F., & Saadah, S. (2018). Dynamic large scale data on Twitter using sentiment analysis and topic modeling case study: Uber. *2018 6th International Conference on Information and Communication Technology, ICoICT 2018*, 254–258. https://doi.org/10.1109/ICoICT.2018.8528776

2) Ayo, C., Oni, A. A., Adewoye, O. J., & Eweoya, I. O. (2016). E-banking users' behaviour: e-service quality, attitude, and customer satisfaction. *International Journal of Bank Marketing*. https://www.emerald.com/insight/content/doi/10.1108/IJBM-12-2014-0175/full/html?fullSc=1&mbSc=1&fullSc=1

3) Baegaegang, & Cheol-soo, K. (2023). A Study on the Competitive Relationship among Short-form Content - Focusing on YouTube and TikTok's User Gratifications Niche Analysis. *Journal of The Korean Society of Illustration Research*. https://www.earticle.net/Article/A429006

4) Brownlee, J. (2019). *Statistical Methods for Machine Learning Discover how to Transform Data into Knowledge with Python* (A. Koshy & A. Cheremskoy (eds.)).

5) Cepeda-López, F., Gamboa-Estrada, F., León, C., & Rincón-Castro, H. (2019). The evolution of world trade from 1995 to 2014: A network approach. *Journal of International Trade and Economic Development*, *28*(4), 452–485. https://doi.org/10.1080/09638199.2018.1549588

6) Chatterjee, S. (2019). Explaining customer ratings and recommendations by combining qualitative and quantitative user generated contents. In *Decision Support Systems* (Vol. 119, pp. 14–22). https://doi.org/10.1016/j.dss.2019.02.008

7) Christanto, H. J., & Singgalen, Y. A. (2022). Sentiment Analysis of Customer Feedback Reviews Towards Hotel's Products and Services in Labuan Bajo. *Journal of Information Systems and Informatics*, *4*(4), 805–822. https://doi.org/10.51519/journalisi.v4i4.294

8) Cuofanno, G. (2023). *YouTube Vs. TikTok*. FourWeekMBA. https://fourweekmba.com/youtube-vs-tiktok/

9) Feraco, A., Das, D., Cambria, E., & Bandyopadhyay, S. (2017). *A Practical Guide to Sentiment Analysis*. Springer International Publishing. https://www.google.co.id/books/edition/A_Practical_Guide_to_Sentiment_Analysis/8QunDgAAQBAJ?hl=id&gbpv=0

10) Google Play. (2023). *Aplikasi di Google Play*. Https://Play.Google.Com/. https://play.google.com/store/apps/details?id=com.google.android.youtube&hl=id

11) Gunasekar, S., Kumar, D. S., Purani, K., Sudhakar, S., Dixit, S. K., & Menon, D. (2021). Understanding service quality attributes that drive user ratings: A text mining approach. *Journal of Vacation Marketing*, *27*(4), 400–419. https://doi.org/10.1177/13567667211003246

12) Hern, A. (2022). *YouTube's battle with TikTok takes its toll as revenues dip _ YouTube _ The Guardian*. Guardian News & Media Limited. https://www.theguardian.com/technology/2022/oct/26/youtube-battle-with-tiktok-takes-its-toll-as-revenues-dip-alphabet

13) Ivanenko, N. (2023). TikTok has a serious competitor: YouTube Shorts have exceeded 50 billion daily views. *Mezha*. https://mezha.media/en/2023/02/03/tiktok-has-a-serious-competitor-youtube-shorts-have-exceeded-50-billion-daily-views/

14) Kemp, S. (2023, February 9). *Digital 2023: Indonesia — DataReportal – Global Digital Insights*. Kepios. https://datareportal.com/reports/digital-2023-indonesia

The Seybold REPORT

ISSN 1533-9211

DOI: 10.5281/zenodo.10300129

15) Kiliç, S., & Çadirci, T. O. (2022). An evaluation of airport service experience: An identification of service improvement opportunities based on topic modeling and sentiment analysis. *Research in Transportation Business and Management*, *43*(November 2021). https://doi.org/10.1016/j.rtbm.2021.100744

16) Naini, N. F., Sugeng Santoso, Andriani, T. S., Claudia, U. G., & Nurfadillah. (2022). The Effect of Product Quality, Service Quality, Customer Satisfaction on Customer Loyalty. *Journal of Consumer Sciences*, *7*(1), 34–50. https://doi.org/10.29244/jcs.7.1.34-50

17) Novendri, R., Callista, A. S., Pratama, D. N., & Puspita, C. E. (2020). Sentiment Analysis of YouTube Movie Trailer Comments Using Naïve Bayes. *Bulletin of Computer Science and Electrical Engineering*, *1*(1), 26–32. https://doi.org/10.25008/bcsee.v1i1.5

18) Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). E-S-QUAL a multiple-item scale for assessing electronic service quality. *Journal of Service Research*, *7*(3), 213–233. https://doi.org/10.1177/1094670504271156

19) Perez, S. (2023). TikTok is crushing YouTube in annual study of kids' and teens' app usage. *Techcrunch*. https://techcrunch.com/2023/02/07/tiktok-is-crushing-youtube-in-annual-study-of-kids-and-teens-app-usage/

20) Putra, I. M. K. B., & Kusumawardani, R. P. (2017). Analisis Topik Informasi Publik Media Sosial Di Surabaya Menggunakan Pemodelan Latent Dirichlet Allocation (Lda) Topic Analysis of Public Information in Social Media in Surabaya Based on Latent Dirichlet Allocation (Lda) Topic Modelling. *Jurnal Teknik Its*, *6*(2), 2–7.

21) Sari, P. K., Alamsyah, A., & Wibowo, S. (2018). Measuring e-Commerce service quality from online customer review using sentiment analysis. *Journal of Physics: Conference Series*, *971*(1). https://doi.org/10.1088/1742-6596/971/1/012053

22) Savirani, A., Kurnia, N., Eddyono, S., Kusumasari, B., Rajiyem, Widhyarto, D. S., Madya, S. H., Monggilo, Z. M. Z., Djindan, M., Putri, T. E., Larasati, Z. W., Prasetyo, W., Santoso, A. D., & Nityasari, A. (2021). *Big Data Untuk Ilmu Sosial Antara Metode Riset Dan Realitas Sosial* (Savirani,). UGM Press. https://www.google.co.id/books/edition/BIG_DATA_UNTUK_ILMU_SOSIAL/yHxJEAAAQBAJ?hl=id&gbpv=0

23) Sekaran, U., & Bougie, R. (2016). Research Methods for Business. In *John Wiley & Sons Ltd. All* (Sevent Edi). John Wiley & Sons Ltd. All. https://doi.org/10.1007/978-94-007-0753-5_102084

24) Shaalan, K., & Emran, M. Al. (2022). *Recent Innovations in Artificial Intelligence and Smart Applications* (K. Shaalan & M. Al Emran (eds.)). Springer International Publishing. google.co.id/books/edition/Recent_Innovations_in_Artificial_Intelli/0tqREAAAQBAJ?hl=id&gbpv=1

25) Sihag, M., Li, Z. S., Dash, A., Arony, N. N., Devathasan, K., Ernst, N., Albu, A. B., & Damian, D. (2023). A Data-Driven Approach for Finding Requirements Relevant Feedback from TikTok and YouTube. *2023 IEEE 31st International Requirements Engineering Conference (RE)*, 111–122. https://doi.org/10.1109/RE57278.2023.00020

26) Suciptawati, N. L. P., Paramita, N. L. P. S. P., & Aristayasa, I. P. (2019). Customer satisfaction analysis based on service quality: Case of local credit provider in Bali. *Journal of Physics: Conference Series*, *1321*(2). https://doi.org/10.1088/1742-6596/1321/2/022055

27) Zahra, S. (2021). *Digital Wallet Service Quality Analysis using Multiclass Classification and Sentiment Analysis*.