International Journal of Management, Finance and Accounting

Service Quality Analysis of Mhealth Services Using Text Mining Method : Alodokter and Halodoc

Nadya Paramita[1]*
Siska Noviarisanti[1]

*Corresponding author: nadyasetiawan@student.telkomuniversity.ac.id
[1]School of Economics and Business, Telkom University, Bandung, West Java, Indonesia

Abstract

The digital transformation of health services causes an increasing number of digital health service providers in Indonesia. The user shares their experiences and reviews each other on the online platform. This study aims to understand user perceptions of m-health services in Indonesia based on m-health service quality with a big data approach. Research using text mining is derived from the results of the reviews of the application Alodokter and Halodoc. User-generated content was gathered from the platform Google Play Store in the period April to December 2020. Based on the sentiment analysis, Alodokter performs well with 73% positive and 27% negative, while Halodoc also dominated with 86% positive and 14% negative. User reviews are categorized based on three dimensions of health service quality with a multiclass classification. It is possible to identify the word networks that often appear in user reviews through text network analysis. The dimension that reviews chiefly on Alodokter and Halodoc is perceived outcome quality. The result of this study could help or use as guidance to be a reference for evaluations to improve Indonesia's quality of m-health services.

Keywords: Alodokter, Halodoc, mhealth Service Quality, Text Mining.

Submitted on 25 June 2021; Accepted on 16 August 2021; Published on 31 August 2021.

1. Introduction

The digital transformation has disrupted all industries, including the healthcare industry. The emergence of technology can help health services become easily accessible, fast, and practical. Indonesia has achieved the fifth position globally as the most startups (Start-Up Ranking, 2020). Indonesia has around 2,074 startups per 2019. Of course, it will continue to grow until 2020 (Kata Data, 2019). In addition, the emergence of the COVID-19 pandemic is one of the challenges for the health industry to provide digital services. Reporting from the results of a press conference at Graha BNPB on April 23, 2020, the government stated that there are still indications of coronavirus transmission in the community so that the COVID-19 cases continue to increase (Watiknas, 2020).

The government issued regulations related to preventing Covid-19 in PP Number 21 of 2020, an appeal for the public to carry out social distancing to reduce the spread of the virus. The Indonesian Medical Council (KKI) issued Regulation No. 74 of 2020 concerning Clinical Authority and Medical Practice through Telemedicine during the Covid-19 Pandemic in Indonesia supports the continuity of the health service process by carrying out social distancing. Telemedicine can be developing into a solution in reducing the spread of Covid-19 (Watiknas, 2020). Even though m-health creates a positive change, there are concerns regarding service quality, as the deficiency in the reliability of platform services, competency-owned provider of services, privacy, and security of user data (Rodhiani, Nurcahyo, & Dachyar, 2020). Of course, this dramatically affects user satisfaction, sustainable use, and the quality of life of m-health service providers. Therefore, understanding the user's perceptions becomes a critical dimension to determine the service quality of m-health. There is an empirical concept to evaluate service quality using SERVQUAL developed by Parasuraman et al. (1988) to understand user's perceptions. However, the SERVQUAL evaluation concept has not explicitly been using in mapping the dynamics of service quality in the health sector. Then (Akter, D'Ambra, & Ray, 2010) developed a specific evaluation concept of m-health services, namely m-health service quality, which consists of 3 dimensions: perceived platform quality, perceived interaction quality, and perceived outcome quality. Considering the importance of service quality, researching and understanding the dimension of the concept of service quality takes much time (Parasuraman et al., 1988). Using text mining can reduce the time to analyze service quality by analyzing big data of online reviews, especially m-health's users.

User-generated content (UGC) can be online reviews, social media, or blogs. One of the most popular UGC platforms for app reviews is the Google Play Store. Online user reviews available on the Google Play Store platform can be used as a channel to extract user feedback (al-Ramahi & Noteboom, 2020). Based on the results of research conducted by DailySocial in an article entitled "The Penetration of Active and Healthy Urban Lifestyle, The Understanding of Wellness Market in Jakarta 2019 " states that Indonesia has the most popular m-health services with 45.3% of respondents using these applications, namely the Halodoc application and 32.2% of respondents who have used the Alodokter application. The Halodoc application is trendy for the variety of health services it provides. This application has features related to health information, teleconsultation with trusted doctors, check-up services at health laboratories in collaboration with Halodoc, and an online drug purchase feature. Alodokter's features provide an opportunity for users to make online bookings for doctor's appointments with online payments. Halodoc and Alodokter can be the object of research as a reference for evaluating Indonesia's quality of health services. This research is expecting to give recommendations that can help improve the quality of health services. In addition, m-health providers can collaborate with the government to increase health services by evaluating service quality dimensions and implementing government programs to reduce the spread of Covid-19 while still providing quality health services to the public.

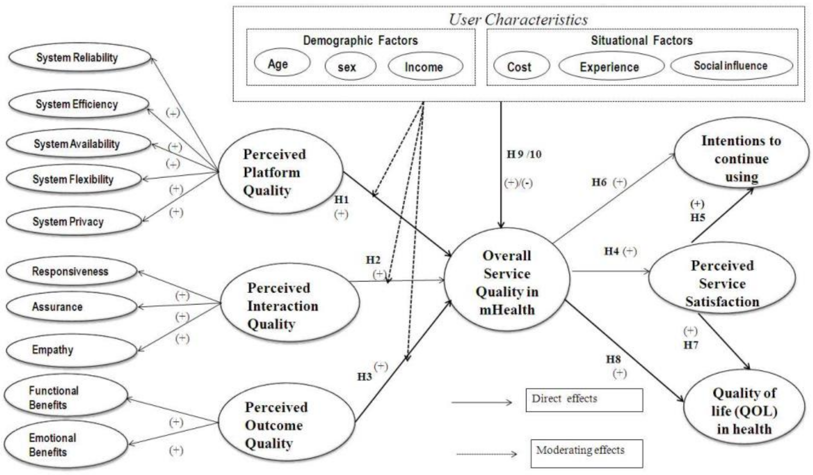
2. Literature review

In Indonesia, institutional providers of services of health and startup also begin to bloom the adoption of digital health as a media interaction with the patient without the need for face-to-face. The concept of digital health has the potential for development, seeing that 73.7% of Indonesia's population, namely 196.71 million people, are active internet users. In addition, 95.4% of active internet users access through a mobile phone (APJII, 2020). The penetration rate of mobile phones in Indonesia reaches 100% in terms of accessibility and coverage. Owners of mobile phones reached 96% of the total population of Indonesia case is an opportunity great for providers to develop applications mHealth (Hootsuite, 2020).

Industry mHealth in Indonesia continues to evolve rapidly, with more than 60% (MTPConnect & Business, 2020). Application of digital health in Indonesia there are 30 662 applications in the Google Play Store platform, with 200 applications new per day (42 Matters, 2020). The population under 25 years was highly literate and reached 110 million people (MTPConnect & Business, 2020). This opportunity has resulted in the emergence of many startups developing digital health applications, reaching 150 startups. Startup application developer m-health formed an association with the name Indonesia HealthTech Association. Community of the startup that moves on the application of mHealth The jointly develop services improve the quality of health facilities in Indonesia (Muhammad, 2019). The m-health application industry in Indonesia is a very competitive industrial market. As the number of players in the mhealth service industry increases, the quality standard of m-health services needs to be improving as a material for further health service development.

There is an empirical concept to evaluate service quality using SERVQUAL developed by Parasuraman et al. (1988) to understand user's perceptions. As technology develops, the limitations of this concept appear in evaluating service quality with digital platforms to trigger the emergence of the concept of service quality evaluation that adapts to technological transformation. The dynamics of online customers' changing needs have some standards that become claims against the provider shop online, with quality products, service, and quality. Seeing these dynamics, the concept of measurement using the ES-QUAL dimension emerged by (Parasuraman, Zeithaml, & Malhotra, 2005) related to evaluating the service quality of the online shopping platform. The ES-QUAL measurement concept is applied by (Santouridis, Trivellas, & Tsimonis, 2012) in evaluating the quality of service on e-commerce sites in Greek cities in their research. However, the ES-QUAL evaluation concept does not precisely map the dynamics of service quality in the health sector. The ES-QUAL evaluation concept triggers health service evaluation such as HEALTHQUAL by (Lee & Kim, 2017), Health Care Service Quality, which combines SERVQUAL dimensions with HEALTHQUAL by (Globenko & Sianova, 2012). Then (Akter, D'Ambra, & Ray, 2010) developed a specific evaluation concept of m-health services, namely m-health service quality.

Figure 1: M-Health Service Quality Conceptual Model



Source: Akter, D'Ambra, & Ray, 2010.

According to (Akter, D'Ambra, & Ray, 2010), service quality modelling for health services consists of three significant variables. The first variable, perceived platform quality, includes quality of the mobile network, ease of access, availability, privacy, and security. Second, perceived interaction quality between doctors and users is related to competence, credibility, manners, insight, customization, and assurance. Third, perceived outcome quality, which includes matters related to functional and emotional benefits for users. Dimensions platform quality as the construction quality of the service that is perceived users in the service of mHealth regarding the level of quality is technically in the communication process. This dimension measures the overall service delivery system, which refers to the dimensions of system reliability, system efficiency, system availability, system flexibility, and privacy of user medical information. When users using the services of mHealth will receive a quality that is available on the platform services that include ease of use, ease of access, availability, speed of response, network coverage, network stability. So the quality of the platform affects the overall quality of m-health services.

The Interaction Quality dimension is the construction of m-health services that involves intensive interaction between users and doctors in consultations or referrals. Interaction quality is the quality of activity during the period in which consumers interact directly with a service. This dimension measures the condition that the user interacts with the doctor using the mhealth platform. Then the user feels the quality in terms of knowledge and competence that the provider supplies. In addition, it is also against the solution's accuracy and user's attention that the provider of service gives.

Outcome quality is a dimension that sees benefit or experience of anything that the user perceives after use and a sense of service. It refers to the characteristics of the results offered by the system based on the aspect of accuracy, precision time, and completeness of the service. This dimension is critical for mhealth services to evaluate the quality of outcomes based on functional and emotional benefits. Dimension is measured under conditions when the user feels the benefit in the form of complete and accurate information to the complaint Medical (functional) and support for mental in addressing complaints medical.

Text mining is an interdisciplinary field related to information retrieval, data mining, machine learning, statistics, and computational linguistics. Text mining is implementing algorithms to convert unstructured text into structured data objects through quantitative methods when analyzing it. The purpose of text mining is to generate high-quality information from the text (Singh, 2016). The text mining application in this research consists of sentiment analysis, multiclass classification, and text network analysis. Sentiment analysis is a Natural Language Processing (NLP) and information extraction process classified in positive or negative comments from large-scale document analysis. Sentiment analysis is a derivative of the research branch in text mining, including opinion mining (Adawiyah & Nugraha, 2018). In this study, sentiment analysis uses to identify user opinions based on their positive and negative experiences regarding the quality of m-health services.

Multiclass classification is a classification method that targets more than two classes (Alamsyah & Rachmadiansyah, 2018). This study uses the multiclass classification method to map reviews of m-health service users by understanding the topic or subject of discussion based on three dimensions of m-health Service Quality. This mapping process aims to find out user reviews of the quality of m-health services in each dimension.

Text network analysis involves the process of converting words into concepts, making relationships between these concepts, and then analyzing them to produce a text network (Hunter, 2014). Text network analysis is used in this study to determine the dynamics of the quality of m-health services in the form of word networks. The word network uses to identify keywords that represent each dimension of Service Quality m-health. Keywords are often written in user reviews to be a reference for the factors that most influence the quality of m-health services.

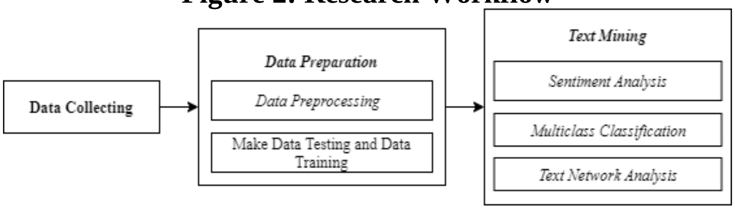
Technological developments also provide a medium for users to share experiences via the internet. Likewise, users of m-health services are active users who share experiences using m-health services through an online platform. Understanding m-health services quality is necessary to analyze users' feedback collected from electronic word-of-mouth (eWOM). Feedback content is created directly by users without any rules and published via the internet called user-generated content (al-Riendi & Noteboom, 2020).

User-generated content has characteristics based on the level of an individual's contribution. Users can contribute their content by only accepting or sharing content and activities without certain conditions. Content contributions can be in comments in framework services such as online letters for editors, user comments for online articles, reviews for applications, or comments on blogs. The content can be the research results and information such as Wikinews, uploading individual text, images, and audio such as blog posts, forum posts, and photo and video platforms. User-generated content has the characteristic that it must be published on the internet to allow discussions between users or communities (Naab & Sehl, 2017).

3. Methodology

The following are the steps taken in the process of analysing this research data:

Figure 2: Research Workflow



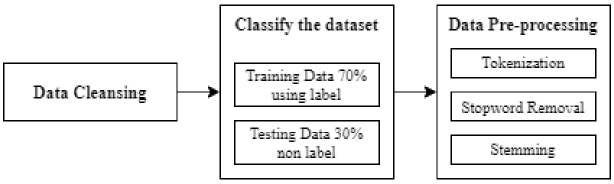
Source: Researcher

3.1 Data Collection

Data collection for this study used a scrapping technique on the Google Play Store platform with the Python language programming algorithm, especially in the Halodoc and Alodokter m-health applications. Google Play Store is a platform that provides features for users to provide feedback as reviews and ratings related to applications that have been downloaded and used. Data collected from the Google Play Store is User Generated Content in user reviews related to their experience using the Halodoc and Alodokter applications for nine months starting April 2020 to December 2020. User reviews collected are in Indonesian and English, which will translate into Indonesian. The data taken includes the user name, date, and review of the scrapping results as a .csv file will transforming into .xls.

3.2 Data Preparation

Figure 3: Data Preparation Diagram



Source: Researcher

After data collection is carried out in the data preparation stage, which consists of pre-processing data and making test data and training data. This stage is essential so that the data generated from the processing process has good quality. Before pre-processing the data, the dataset will be clean first after being collected and cleaned according to the desired data criteria. Furthermore, the dataset is prepared to become test data and training data. Test data and training data were collected from a dataset of user reviews related to health services consisting of Halodoc and Alodokter. Test data consists of data that does not use labelling and is used to test machine learning from training data.

Meanwhile, training data is labelled or grouped according to categories for learning materials for machine learning. The percentage of sharing training data and test data in order to be able to produce maximum data is the ratio of 70:30. Training data is 70%, and test data is 30% of the total data (Ting & Tsang, 2011).

The data review user also gives a label based on the three-dimensional m-health service quality: Perceived Quality Platform, Perceived Quality Interaction, and Perceived Quality Outcome. The m-health service quality dimension will use as a label on the training data. The training data results will become the reference for machine learning to predict the test data that do not have labels associated with three-dimensional m-health service quality. The result of learning will establish a model that will calculate the level of validation and performance. Dimensions m-health service quality uses to measure the quality of service of the application Alodokter and Halodoc. So it can be in identifying the dimensions that have good and which still need improvement in the user's perception. The labelling process for training data manually uses a reference based on indicators of the dimensions of mHealth service quality by Akter, D'Ambra, & Ray, 2010.

Furthermore, after obtaining the test and training data, the training data will go through the pre-processing data process before becoming machine learning input. Pre-processing data is the preparation of data for use when the classification process reviews Service Quality Health. Following are the stages in pre-processing data based on (Monali & Sandip, 2014):

a. Tokenization:

   This stage is solving the words in the review sentence, separating the words in the sentence with a space character. The primary purpose of this stage is to convert unstructured documents into structured documents because the data generated from text mining is unstructured.

b. Stopword Removal

   Removing words belonging to the stopword category or words with a high occurrence rate is deemed meaningless.

c. Stemming

   The stage is finding the root word in a word by removing the affixes. The purpose of this stage is to improve system performance.

3.3 Text Mining

This research uses text classification methods, namely sentiment analysis and multiclass classification, as well as knowing the word network using text network analysis. Sentiment analysis is used to identify user reviews of the quality of m-health services based on positive and negative sentiments regarding satisfaction in using m-health services. The multiclass classification method is used to map the user reviews of health services based on three dimensions of health service quality. It is making a classification model using a text classification algorithm, namely the Naïve Bayes algorithm. The Naïve Bayes algorithm is an algorithm that is simple, efficient, and does not require a large quantity of training data datasets (Widyawati, Irawan, & Ghina, 2020).

Before processing data using the Naïve Bayes algorithm method, it is necessary to determine the modelling validation first. K-Fold Cross-validation is a method for evaluating model performance. Confusion Matrix is a tool for evaluating classification and clustering models in estimating the truth of objects. The Confusion Matrix uses a matrix to predict the comparison of the training data class with the actual value information from the testing data (Gorunescu, 2011). The indicators that can specify in the Confusion Matrix are accuracy, precision, recall, f-measure and kappa.

The text network analysis method explores the quality of new health services based on topic findings so that new dimensions of health service quality can be identified from user reviews. The text network analysis phase begins with cleansing a text document broken down into a word list. The steps consisted of three steps, namely word frequency count, word co-occurrence analysis, and visualization. In this step, the word frequency count aims to determine the frequency of appearance of words in a document. The word frequency count step in this study was assisted by the Wordij software, which resulted in a list of words and their frequencies in the .csv document. The next step is word co-occurrence analysis, namely giving weight to the relationships between words in the document. The final step is to visualize the results of the relationship between words with the help of Gephi software. The visualization results can display keywords to explore indicators of new dimensions of health services quality.

Table 1: Example of Text Classification by Sentiment

Review	Sentiment
Pertama kali pakai halodoc nyaman banget di saat pandemi, gak perlu keluar rumah untuk ketemu dokter dan beli obat. Semua beres via app. Dokternya sgt profesional jg. Bravo halodoc sudah standby 24 jam (The first time I used Halodoc, it was very comfortable during a pandemic. There was no need to leave the house to see a doctor and buy medicine. Everything is done via the app. The doctor is very professional too. Bravo Halodoc is on standby 24 hours)	Positive
Worst service, antigen test didalam kategori hasil 1 jam, ternyata hasil nya ditunggu sampai jam 23.59. ketika ditanya ke CS ternyata tidak ada yg mengerti, pelaporan dalam 15-30 menit dijanjikan ada respon ternyata hanya tipuan. (Worst service, the antigen test is in the 1-hour result category. It turns out that the results have waited until 23.59 hours. When asked to customer service, it turned out that no one understood. Reporting within 15-30 minutes was promised a response, it turned out to be a hoax.)	Negative

Notes: The review using Indonesian.

Table 2: Example of Text Classification by Multiclass

Review	Indicator	Dimension
Sistem operasi mudah dan cepat responsenya (The operating system is easy, and the response is fast)	*System Reliability*	*Perceived Platform Quality*
Super mudah bisa pilih mau konsul dengan dokter manapun, bisa gratis (It is super easy to choose whether customers want to free consult with any doctor.)	*System Efficiency*	
Aplikasinya stabil, lancar dengan fitur dan tampilan sederhana (The application is stable, smooth with simple features and an interface.)	*System Availability*	
Konsultasi bisa dilakukan dari rumah, sekalipun malam (Online consultation from home, even at night)	*System Flexibility*	
Meminta data pribadi dan rekening. (Request personal and account data.)	*System Privacy*	
Dokter cepat tanggap dan sangat rinci dalam penjelasannya (The doctor is quick to respond and very detailed in his explanations)	*Responsive*	*Perceived Interaction Quality*
Dokter profesional dan obat cepat sampai (Professional doctor and fast-delivery medicine)	*Assurance*	
Konsultasi bisa menggunakan suara dan telfon video dengan dokternya langsung. (Consultations can use voice and video calls with the doctor directly)	*Empathy*	
Saya sangat senang berhasil transaksi pertama kalinya dengan metode pembayaran asuransi, di masa pandemi ini sangat terbantu sekali (I am pleased that the first transaction of insurance payment method was easy during this pandemic.)	*Functional Benefit*	*Perceived Outcome Quality*
Mempermudah orang sakit tidak perlu datang ke RS yang saat ini rawan tempat tertular virus COVID. (Make it easier for sick people not to have to come to hospitals that are currently prone to be infected with the COVID virus.)	*Emotional Benefit*	

Notes: The review using Indonesian.

4. Results and Discussion

4.1 Text Classification

In the performance of the sentiment analysis classification model with the Naïve Bayes algorithm, the accuracy values for Halodoc and Alodokter are 87.95% and 89.33%. The resulting accuracy value in the sentiment analysis model of the two m-health applications is above 60%, so the modelling is suitable. The precision, recall, and f-measure values in the Halodoc and Alodokter sentiment analysis modelling have values above 60%, so that the performance of the modelling is good. Meanwhile, the kappa values generated from the sentiment analysis modelling on Halodoc and Alodokter were 0.58 and 0.69. According to Landis and Koch, 1977 in (Widyawati, Irawan, & Ghina, 2020), the value range of 0.61 - 0.80 interpreted that the classification model was correct. The sentiment analysis result is Halodoc shows a positive percentage of 86%, while negative perceptions are only 14%. Likewise, with the Alodokter application, positive perceptions reached 73%, while the percentage of negative perceptions was 27%.

In the performance of the multiclass classification model with the Naïve Bayes algorithm, the accuracy scores for Halodoc and Alodokter are 65.85% and 89.33%. The resulting accuracy value in the multiclass classification model for both health applications is above 60%, so that the modelling used is suitable. The precision, recall, and f-measure values in the Halodoc and Alodokter's multiclass classification modelling have a value above 60%, so that the performance of the modelling is good. Meanwhile, the kappa values generated from the sentiment analysis modelling on Halodoc and Alodokter were 0.470 and 0.423. Halodoc has the largest percentage of user perceptions based on data processing results, namely the Perceived Outcome Quality dimension of 40%. The perception of Halodoc users on the Perceived Platform Quality dimension is 32%, and the Perceived Interaction Quality dimension is 28%. While the results of identifying the perception of users of Alodokter, the dimension with the largest percentage is the Perceived Interaction Quality dimension of 44%, the dimensions of Perceived Platform Quality are 16%, and the dimensions of Perceived Outcome Quality are 39%.

Table 3: Alodokter and Halodoc Dimension Proportion

Dimension	Alodokter		Halodoc	
	Positive	Negative	Positive	Negative
Perceived Platform Quality	74%	26%	73%	27%
Perceived Interaction Quality	86%	14%	83%	17%
Perceived Outcome Quality	96%	4%	98%	2%

Based on the percentage of positive sentiment, the quality of the Perceived Platform Quality dimension has a dominant proportion in the Halodoc and Alodokter applications. The negative sentiment regarding the quality of this dimension at Halodoc stated that it was related to problems on the platform such as old loading, jammed screens, the location of drug purchases that were too far away, did not have proof of payment, problematic verification codes, failed payment systems, had not been chat but paid for and stock availability. Meanwhile, the Perceived Platform Quality dimension in the Alodokter application is that it is difficult to log in with a phone number, protect the user's telephone number, problems in entering user data, and problems in the chat feature.

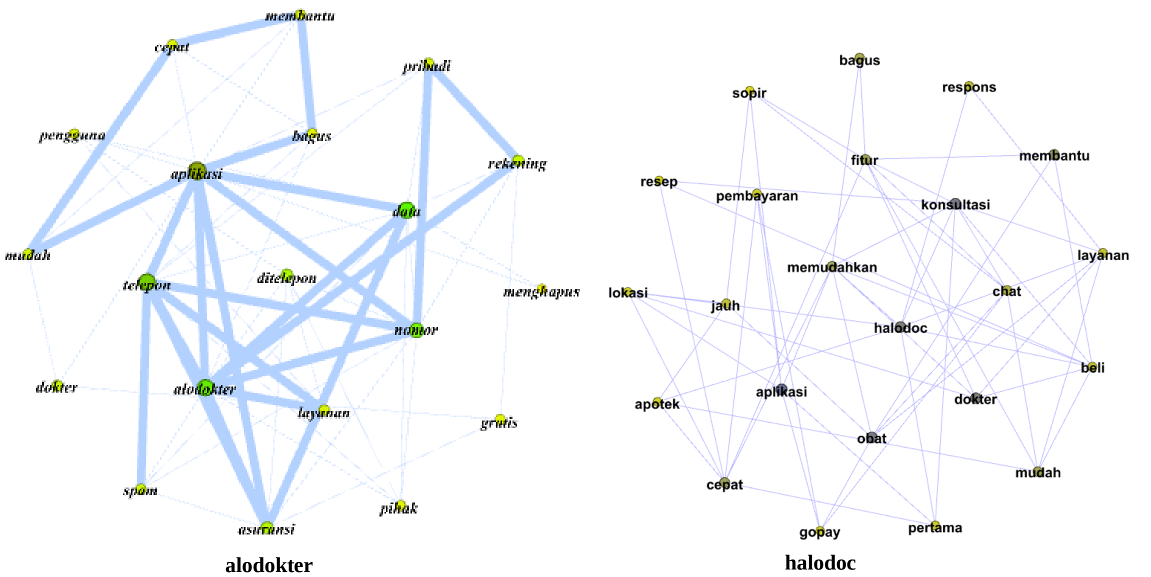
In the Perceived Interaction Quality dimension, the proportion of positive sentiments dominates the Halodoc and Alodokter applications. In the Halodoc application, negative sentiments regarding user service are slow to respond, and orders are not delivered, swab schedule mismatches, difficult reimbursements, and inappropriate prices. Meanwhile, many negative sentiments on this dimension are related to dissatisfaction with the doctor's response, insurance offers, and inappropriate prices. The proportion of positive sentiment in the Perceived Outcome Quality dimension still has a high percentage in the Halodoc and Alodokter applications. Meanwhile, the negative sentiment on the Halodoc application was due to sudden schedule changes, subjective consultations, and expired drugs. Whereas the Alodokter application is mainly related to testimonials, is not satisfied with the consultation, and cancels the consultation experience.

4.2 Text Network Analysis

To strengthen understanding regarding the quality of m-health services illustrated through user reviews through the Google Play Store platform, it uses text network analysis. The text network analysis technique is carried out based on each dimension to determine the word network formed from each dimension. So can see the words that users often utter. According to user reviews, the results of text network analysis can become a reference for other m-health providers to improve service quality. The data from the user review by calculating the frequency first using the Wordij software, and the word pairs that appear in the user review are obtained. The relationship between the word pairs is visualizing using Gephi in the form of a text network. Text networks consist of nodes and edges. Nodes that have the most significant size are the words that appear most frequently in user reviews. Then the edges are the relationships between words that are connected with lines. The thicker the relationship lines, the stronger the relationship between words. The text network layout used to visualize the word network is Fruchterman Reingold. The following is the result of text network analysis based on each dimension of m-health service quality for the Halodoc and Alodokter applications:

Perceived Platform Quality

Figure 4: Perceived Platform Quality Text Network

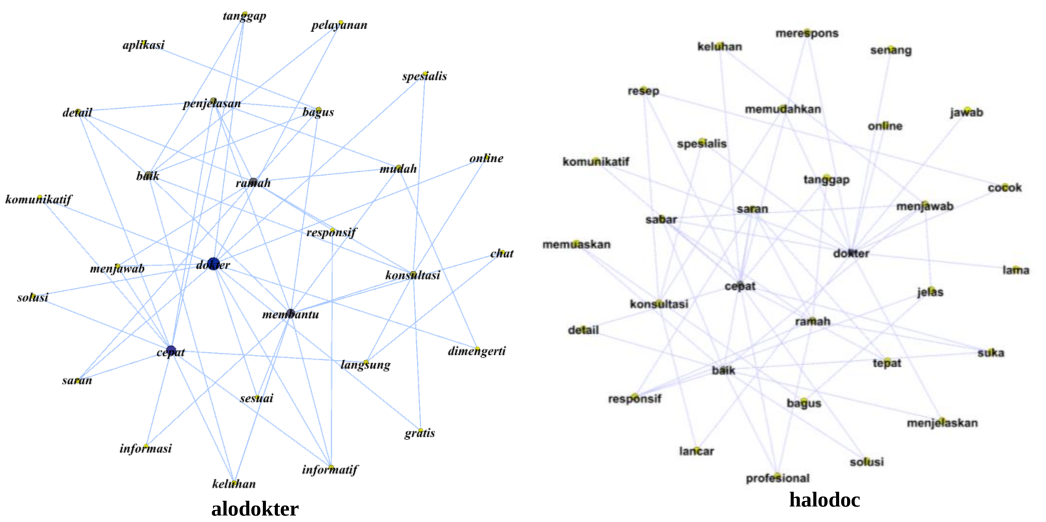


Source: Researcher

Text networks formed from the relationship between words related to the discussion of the dimensions of perceived platform quality in the Halodoc and Alodokter applications are dominated by discussions containing positive perceptions. It can be concluded that users are satisfied with the features available on the Halodoc and Alodokter platforms. It is proven by the emergence of several words that contain positive sentiments in discussing user reviews of Halodoc and Alodokter. The positive words that appear in the Halodoc application are "fast", "easy", "good", "response", and "help". Alodokter users also give positive perceptions through the appearance of positive words, namely "fast", "easy", "safe", "helpful", and "good". The largest nodes and the green ones are the words that have the greatest frequency of occurrence. In the Halodoc application, the green nodes, namely. Meanwhile, the Alodokter application consists of “phone”, “data”, “number” and “alodokter”. In addition, the text network also found words that contained negative sentiments or user complaints regarding the perceived platform quality dimension.

Perceived Interaction Quality

Figure 5: Perceived Interaction Quality Text Network

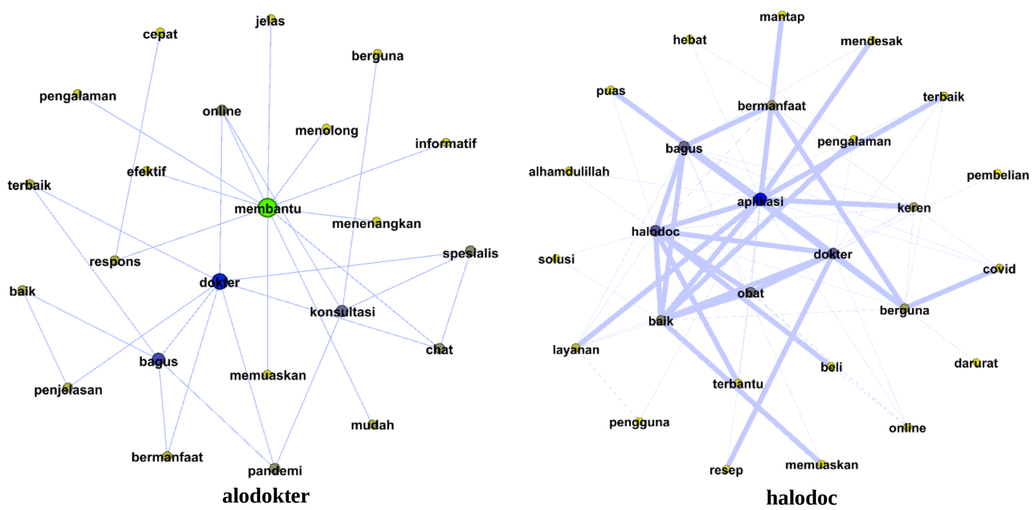


Source: Researcher

Text networks formed from the relationship between words related to the discussion of the dimensions of perceived interaction quality in the Halodoc and Alodokter applications are dominated by discussions that contain positive perceptions. Especially Alodokter, in this dimension, the words that are often shown contain negative sentiments. It can be concluded that the Halodoc and Alodokter applications users are satisfied with the features available on the Halodoc and Alodokter platforms. It is proven by the emergence of several words that contain positive sentiments in discussing user reviews of Halodoc and Alodokter. The positive words that appear in the Halodoc application are “specialist”, “fast”, “friendly”, “fluent”, “communicative”, “professional”, “responsive”, “clear”, “patient”, and “responsive”. The Alodokter application also has the appearance of positive words, namely “fast”, “friendly”, “appropriate”, “responsive”, “communicative”, "informative", "detail", "specialist", "answer", "understandable" and "good ". The largest nodes and the green ones are the words that have the greatest frequency of occurrence. The Halodoc and Alodokter applications do not have green nodes, so that in this dimension, the frequency of occurrence of words is balanced. Apart from that, words that contain negative sentiments were also found on the dimension of perceived interaction quality in the Halodoc application.

Perceived Outcome Quality

Figure 6: Perceived Outcome Quality Text Network



Source: Researcher

Text networks formed from the relationship between words related to the discussion of the dimensions of perceived platform quality in the Halodoc and Alodokter applications are dominated by discussions containing positive perceptions. Halodoc and Alodokter"s applications in this dimension, the words often appear, do not contain negative sentiments. It can be concluded that users of the Halodoc and Alodokter applications are satisfied with the features available on the Halodoc and Alodokter platforms. It is proven by the emergence of several words that contain positive sentiments in discussing user reviews of Halodoc and Alodokter. The positive words that appear in the Halodoc application are "useful", "good", "satisfying", "helpful", "useful", "satisfied", "great", "cool" and "steady". The Alodokter application also has the occurrence of positive words, namely "effective", "satisfying", "useful", "helpful", "informative", "best", "useful" and "calming".The largest nodes and the green ones are the words that have the greatest frequency of occurrence. Halodoc does not have green nodes, so in this dimension, the frequency of occurrence of words is balanced. Meanwhile, the Alodokter application consists of the word “helping”. In addition, the text network also found words that contained negative sentiments or user complaints regarding the perceived platform quality dimension.

Based on the research results related to the analysis of the service quality of m-health applications in Indonesia using data from user reviews of the Halodoc and Alodokter applications, it can help understand user requests for application-based health facilities. In this study, an analysis of service quality was carried out based on the dimensions of m-health service quality, divided into three terms of perceived quality of the platform, interaction, and outcome. The results of this study can be used as a reference for developing the quality of m-health applications in Indonesia. Through understanding user perceptions, it is possible to create appropriate application-based health services. This study aims to be a reference in improving the quality of health services in Indonesia. This research is useful in identifying problems experienced by users based on negative sentiment mapping. So that it can be used as a focus for handling or subsequent repairs. The following is a summary of the problems with m-health application services that are most often the topic of user complaints:

Table 3: Alodokter and Halodoc Dimension Proportion

Application	Description	Keyword
Halodoc	The location of the pharmacy selection to buy drugs is far from the user’s destination	“jauh”,”lokasi”,”apotek”
	The response from user services and doctors to user complaints feels that it still takes a long time.	“lama”,”keluhan”
Alodokter	Distracted users get calls offering insurance to ask for the account number	“telepon”,”spam”, “penipuan”,”asuransi”, “rekening”
	User cannot delete the account	“menghapus”, “akun”
	Users feel they have less privacy, and their phone number data is leaked.	“data”,”bocor”, “privasi”,”nomor”

5. Conclusion

This study uses text classification methods, namely sentiment analysis and multiclass classification, and knowing word networks using text network analysis. Sentiment analysis identifies user reviews of the quality of m-health services based on positive and negative sentiments. The multiclass classification method uses to determine the quality of health services based on the review classification of m-health service users, which refers to three dimensions of health service quality. Meanwhile, text network analysis explores the word that describes the quality of m-health services found from user reviews of m-health applications.

The researcher observed that the three text analysis methods were suitable for analysing the results of user reviews of m-health applications, namely Halodoc and Alodokter. The result of processing such data can be used for reference in improving service quality healthcare applications in Indonesia. The following is a conclusion based on the results of the research as well as the research discussion :

1. User reviews on the Google Play Store about m-health service applications, namely Halodoc and Alodokter, illustrate the perception of positive and negative sentiments. The sentiment analysis found that the Halodoc application had a positive sentiment perception of 86% and a negative sentiment of 14%. Meanwhile, the Alodokter application also dominates by perceptions with a positive sentiment of 73% and a negative sentiment of 27%. Based on the sentiment percentage, the Halodoc and Alodokter applications have the highest positive sentiment compared to negative sentiment. So it can be concluded that the Halodoc and Alodokter applications have good quality health services according to user experience.
2. User reviews regarding their experiences using the Halodoc and Alodokter healthcare applications are classified based on three dimensions of health service quality. The multiclass classification results show that Halodoc and Alodokter users tell more about the benefits and complaints after using the application, namely the perceived outcome quality dimension. Positive sentiments dominate the proportion of sentiments in each dimension in the Halodoc and Alodokter applications. So it can be concluded that the Halodoc and Alodoc application service quality performance is reasonable based on the m-health service quality dimension.

3. Analysis of text networks on user reviews of the Halodoc and Alodokter applications, the intensity of the appearance of positive words was more than the appearance of negative words in each dimension of health service quality. Some words with negative sentiments are a problem that can be a reference for improving and developing m-health applications. In the Halodoc application, the user's negative perception is that choosing a pharmacy to buy drugs is far from the user's destination and the responses from user services and doctors in responding to user complaints feel that it still takes a long time. While the negative perception of users of the Alodokter application is that users are disturbed by getting calls that offer insurance to ask for account numbers, users cannot delete accounts, and users feel they have low privacy and their phone number data is leaking.

This study aims to better understand the implementation of big data analytics on service management, especially in the health sector. This research is still limited to the scope of user reviews of the two most popular health applications in Indonesia, namely Halodoc and Alodokter. User review data was collected only from one UGC platform channel (user-generated content), the Google Play Store. The sample is taken at a limited range from the beginning pandemic to the end of 2020. So other health applications that have just released applications during that period still have limited UGCs than Halodoc and Alodokter. In this study, the variables used in text mining analysis only use user review content. Hopefully that further research can use variables such as ratings to maximize the measurement of service quality. Future research can also add methods other than sentiment analysis, multiclass classification, and text network analysis.

References

Akter, S., D'Ambra, J., & Ray, P. (2010). User Perceived Service Quality of mHealth Services in Developing Countries. *18th European Conference on Information Systems*, 1-12. DOI: http: //aisel.aisnet.org/ecis2010/134-riendi

al-Ramahi, M., & Noteboom, C. (2020). Mining User-generated Content of Mobile Patient Portal: Dimensions of User Experience. *ACM Transactions on Social Computing*, 3(3), 1-24.

Barutçu, S. (2019). mHealth apps design using quality function deployment. *International Journal of Health Care Quality Assurance*, *32*(4), 698-708.

DailySocial, R. (2019). The Penetration of Active and Healthy Urban Lifestyle, The Understanding. *White Paper*.

Fitzsimmons, JA, & Fitzsimmons, MJ (2011). *Service Management: Operations, Strategy, Information Technology, Seventh Edition.* New York: McGraw-Hill.

Globenko, A., & Sianova, Z. (2012). Service quality in healthcare:quality improvement initiatives through the prism of patients’ and providers’ perspectives. Unpublished Master Thesis: Umeå School of Business and Economics.

Gorunescu, F. (2011). *Data Mining Concept, Models and Techniques.* Berlin: Springer.

Kata Data. (2019, Maret 21). Indonesia Memiliki 2.000 Startup. Retrieved from Databoks: https://databoks.katadata.co.id/datapublish/2019/03/22/indonesia-memiliki-2000-startup

Lee, D., & Kim, K. K. (2017). Assessing healthcare service quality: a comparative study of patient treatment types. International Journal of Quality Innovation, 3(1), 1-15. doi:10.1186/s40887-016-0010-5

Monali, P., & Sandip, K. (2014). A Concise Survey on Text Data Mining. *International Journal of Advanced Research in Computer and Communication Engineering, 3*(9), 8040-8043.

Naab, T. K., & Sehl, A. (2017). Studies of user-generated content: A systematic review. Journalism, 18(10), 126-1273. doi:10.1177/1464884916673557

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. doi:10.1177/1094670504271156

Rodhiani, A. N., Nurcahyo, R., & Dachyar, M. (2020). Analysis of Service Quality in m-Health. International Journal of Advanced Science and Technology, 29(7), 3511-3517.

Santouridis, I., Trivellas, P., & Tsimonis, G. (2012). Using E-S-QUAL to measure internet service quality of e-commerce web sites in Greece. International Journal of Quality and Service Sciences, 4(1), 86-98. doi:10.1108/17566691211219751

Startup Ranking. (2020). Startup each Country. Retrieved from Startup Ranking: https://www.startupranking.com/countries

Sunjaya, AP (2019). Potential, Application and Development of Digital Health in Indonesia. *Journal of the Indonesian Medical Association*, *69*(4), 167-169.

Ting, S., & Tsang, A. H. (2011). Is Naive Bayes a good classifier for document classification? *International Journal of Software Engineering and Its Applications, 5*(3), 37-46.

Watiknas. (2020). Solusi Telemedicine Di Tengah Pandemi. Jakarta: Watiknas.

Widyawati, R. S., Irawan, H., & Ghina, A. (2020). Content Analysis of Tourist Opinion Based On Tourism Quality (TOURQUAL) By Text Mining Online Reviews - The Case of Borobudur. *International Conference on Sustainable Management And Innovation 2020*.