

Sentiment analysis of online licensing service quality in the energy and mineral resources sector of the Republic of Indonesia

Ahmad Hizqil, Yova Ruldeviani

Department of Technology Information, Faculty of Computer Science, Universitas Indonesia, Jakarta, Indonesia

Article Info

Article history:

Received Nov 19, 2023

Revised Jan 3, 2023

Accepted Jan 25, 2024

Keywords:

Online licensing service

Public satisfaction

Sentiment analysis

User feedback analysis

ABSTRACT

The Ministry of Energy and Mineral Resources of the Republic of Indonesia regularly assessed public satisfaction with its online licensing services. User rated their satisfaction at 3.42 on a scale of 4, below the organization's average of 3.53. Evaluating public service performance is crucial for quality improvement. Previous research relied solely on survey data to assess public satisfaction. This study goes further by analyzing user feedback in text form from an online licensing application to identify negative aspects of the service that need enhancement. The dataset spanned September 2019 to February 2023, with 24,112 entries. The choice of classification methods on the highest accuracy values among decision tree, random forest, naive bayes, stochastic gradient descent, logistic regression (LR), and k-nearest neighbor. The text data was converted into numerical form using CountVectorizer and term frequency-inverse document frequency (TF-IDF) techniques, along with unigrams and bigrams for dividing sentences into word segments. LR bigram CountVectorizer ranked highest with 89% for average precision, F1-score, and recall, compared to the other five classification methods. The sentiment analysis polarity level was 36.2% negative. Negative sentiment revealed expectations from the public to the ministry to improve the top three aspects: system, mechanism, and procedure; infrastructure and facilities; and service specification product types.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Ahmad Hizqil

Department of Technology Information, Faculty of Computer Science, Universitas Indonesia

Jakarta, 10440, Indonesia

Email: ahmad.hizqil@ac.id

1. INTRODUCTION

Since 2019, the Ministry of Energy and Mineral Resources (MEMR) has initiated digitalizing business licensing services in Indonesia's energy and mineral resource sectors. The MEMR provides risk-based business licensing services, which consists of several subsectors such as oil and natural gas, electricity, minerals and coal, new energy, renewable energy, and energy conservation through an integrated electronic business licensing system. Driven by the limitation of face-to-face services due to the impact of the COVID-19 pandemic in March 2020, the digitalization process for licensing services at the MEMR of Indonesia was further enhanced, experiencing a 700% increase in user uptake. The energy and mineral resources (EMR) sector was chosen because of the investment realization from 2019 to 2022 amounting to USD 28.45 billion, making it one of the largest investment-contributing sectors in Indonesia, and is considered stable amidst a general investment slowdown due to the pandemic and the economic growth following the pandemic.

As a public service provider, the online licensing service must conduct a community satisfaction survey periodically at least once a year. This survey is carried out periodically to obtain the community satisfaction index as a performance evaluation of the services provided to enhance the quality of public services. In the survey implementation, service recipients are asked to rate each service's importance and performance level based on nine indicators established by the Ministerial Regulation of the Ministry of State Apparatus Utilization and Bureaucratic Reform. From 2019 to 2022, the online licensing services of the MEMR achieved a community satisfaction score of 3.42, below the score of 3.53 for the directorate owning the service.

The satisfaction surveys conducted by the ministry also collect feedback in the form of appreciation, criticism, suggestions, and inputs about the services in text form filled out by users of the online-based licensing services every time they obtain or apply for benefits. However, this feedback data has not been optimally processed due to the lack of regulations on processing survey data in text form. In previous research, service quality was measured based on indicators and procedures listed in the Ministry of State Apparatus Utilization and Bureaucratic Reform number 14 of 2017 [1]. That research revealed the level of service satisfaction based on predetermined data entries and indicators, but it could not identify the complaints frequently experienced by users.

In this research, the main objectives are to determine the sentiment polarity of users of online licensing services in the EMR sector and to identify which aspects significantly impact service satisfaction levels. Sentiment analysis will be employed to evaluate the public satisfaction survey data. This method will classify the text data provided by users into positive or negative sentiment categories. The negative sentiments will be further analyzed to pinpoint the key dimensions of public satisfaction. The outcomes of this research aim to offer actionable recommendations to the MEMR for targeted improvements in public service delivery based on the issues most frequently raised by users.

2. METHOD

This subtitle delineates the foundational concepts driving this research and articulates the prevailing perspectives on implementing and evaluating information systems. It looks into different theories about using IT and explores various methods and ideas that support this study. Included are detailed descriptions of the concepts, models, research tools, methodologies for data collection, and the data processing techniques employed in this study.

2.1. Service quality and sentiment analysis

Service quality refers to the caliber of integrated interaction, information sources, and internet-based services aimed at enhancing trust in an organization [2]. It can be measured by the gap between the service users' expectations and the actual services experienced. Better service quality is indicated by a narrower gap between expectations and reality, while a wider gap signifies poorer quality [3]–[5]. Evaluating service quality is expected to bridge this gap by providing feasible solutions [6].

Based on user feedback, sentiment analysis is one method to measure the gap between expectations and perceived reality. It offers insights into public opinions, evaluations, attitudes, and emotions regarding specific objects such as products, services, organizations, individuals, issues, events, and topics [7]. Public opinion can generate either positive or negative sentiments towards a product or service [8]. Improvements can be made to the object in question to enhance public satisfaction with the services [9].

2.2. Dataset and annotations

In analyzing a text, researchers collect a dataset. A dataset is a collection of data consisting of rows and columns that describe examples of data observed using machine learning. Datasets are necessary for creating predictive models during training. The use of datasets in machine learning for classification is divided into two parts: training and testing. Generally, the proportion of training data is more prominent, about more than 70%, than the total data [10].

The collection process of user feedback is conducted at the relevant agency, with the customer satisfaction survey data originating from online licensing service applications of the MEMR, spanning from September 21, 2019, to February 8, 2023, in the form of an Excel file. The survey data entries are filled out by contact persons from the business entities using the service after obtaining the licensing services. The survey data entries include the type of company, the importance and performance values of each dimension, and feedback on the services received in the form of text.

The dataset is created by assigning labels to each data entry. Data labeling is applied to each row of data and then labeled as positive or negative [8]. This labeling is done manually by annotators. The number of annotators is more than one and an odd number to ensure objective results. During the labeling process, if there

are differing labels, the majority label is chosen. The total dataset collected is 75% of the entire available data. Five people carried out annotation and assessed positive and negative ratings.

2.3. Data pre-processing

The pre-processing stage is performed by selecting raw data and transforming it into more structured data. Pre-processing reduces irrelevant and redundant data to optimize computational datasets during the machine learning model training phase to make it faster [11]. The platform used in this process is Python, with Conda as the package and environment management systems. Data pre-processing includes cleansing, transformation, tokenization, data stemming, and stopword removal and carried out in this stage include [12]–[17]:

- Cleansing: the process removes special characters such as hashtags, URL links, ASCII characters, numbers, and HTML attributes.
- Transformation: this stage begins with converting all words to lowercase.
- Tokenization: each word in a sentence is separated one by one.
- Data stemming: removing prefixes and suffixes to obtain the root word of each word.
- Stopword removal: deleting words that are commonly used but do not significantly influence the sentence. Examples of widely found stopwords include 'dan' (and), 'atau' (or), 'dari' (from), 'itu' (that).

An example of pre-processing on one of the service satisfaction data provided by the business entity is shown in Table 1.

Table 1. Pre-processing survey sentence

Process	Result
Raw Data	Dear Mr. Director, the process of waiting 7 working days is too long, please consider speeding up the response process regarding this permit to a maximum of 2 working days. :)
Cleansing	Dear Mr Director the process of waiting working days is too long please consider speeding up the response process regarding this permit to a maximum of working days
Transformation	dear mr director the process of waiting working days is too long please consider speeding up the response process regarding this permit to a maximum of working days
Tokenizing	dear working speeding permit mr days up to director is the a the too response maximum process long process of of please regarding working waiting consider this days
Stemming	dear mr director the process of wait work day is too long please consider speed up the response process regard this permit to a maximum of work day
Stopword Removal	process wait work day too long speed up response process regard permit maximum work day

2.4. Feature extraction

Feature extraction simplifies the original data by selecting key pieces of information to create a set of characteristics, known as a feature vector. This process involves using N-grams, which break down sentences into groups of words. The 'N' in N-grams refers to the quantity of words grouped in a segment. [18]. Applying N-grams can improve the accuracy of the classification process [18]–[21]. There are three kinds of N-Gram [22]: a unigram consists of one syllable, a bigram consists of two syllables, and a trigram consists of three. The result of the extraction process uses N-grams on a sentence to obtain the most relevant information, as shown in Table 2. Next, the extracted results are converted into a numerical representation to be processed through electronic media. CountVectorizer and term frequency-inverse document frequency (TF-IDF) transform the text into a numerical representation, then processed in machine learning. CountVectorizer prioritizes the frequency of words in a document, while TF-IDF combines the frequency of words and the inverse document frequency to create an informative text vector representation [17], [23].

Table 2. N-Gram extraction result

Unigram	Bigram
Process, wait, work, day, too, long, speed, up, response, process, regard, permit, maximum, work, day	Process wait, wait work, work day, day too, too long, long speed, speed up, up response, response process, process regard, regard permit, permit maximum, maximum work, work day

2.5. Classification

Generally, there are two methods for sentiment classification: the machine learning approach and the lexicon approach. The lexicon approach utilizes a dictionary of words with positive, neutral, or negative

connotations. Meanwhile, the machine learning approach develops learning algorithms and classification models from a dataset [24]. Machine learning methods are generally categorized into supervised and unsupervised learning. Supervised learning requires substantial pre-labeled training data, whereas unsupervised learning is applied when labeling training data is impractical [25]. Standard classification methods include decision tree (DT), random forest (RF), naive bayes (NB), stochastic gradient descent (SGD), logistic regression (LR), and k-nearest neighbor (k-NN) [11], [26].

2.6. Evaluation

The classification process is followed by measuring the performance of the machine learning classification and comparing the performance and effectiveness of each evaluated method. The evaluation uses the confusion matrix, as shown in Table 3, as a decision-making method that depicts the algorithm's confusion level [27], [28]. The confusion matrix is represented using a table containing the data correctly and incorrectly classified [29] and commonly used to calculate the accuracy of data mining. The values of true positives (TP) and true negatives (TN) provide information that the classification has been done correctly. In contrast, the importance of false positives (FP) and false negatives (FN) indicates that the data classification is incorrect [30]. Based on the confusion matrix in Table 1, performance measure of precision, recall, and F1-score [12] can be calculated in (1)-(3):

$$Precision = \left(\frac{TP}{TP+FP} \right) \quad (1)$$

$$Recall = \left(\frac{TP}{TP+FN} \right) \quad (2)$$

$$F1 - score = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall} \right) \quad (3)$$

Table 3. Confusion matrix

Confusion matrix	Actual sentiment	
	Positive	Negative
Predicted positive	TP	FP
Predicted negative	FN	TN

2.7. Negative sentiment dimension

The next step is to categorize each sentiment analysis according to each dimension. Each word is taken from the data collection and grouped according to a previously defined dictionary. Fundamentally, there are five dimensions of general service quality encompassing tangible evidence, assurance, empathy, responsiveness, and reliability [31]. In addition to previous research, various studies have been developed in different contexts and industries to obtain service quality measurements in various industries [32], including e-GovQual for e-government service quality [33] and democratic e-governance website evaluation model [34].

The factors that influence the quality of online governmental services are divided into functional aspects directly related to the scope of services provided by the system and non-functional aspects related to the ease of use, operation, availability, and security of the website [35], [36]. Service mechanisms and procedures include standardized service procedures for service providers and recipients, including complaints [37]. Procedure information in this study covers general news and standard service information available on websites regarding licensing procedures [9], [38]–[40]. Facilities and infrastructure in this study refer to elements that can be activated or mobilized in the system, such as computer technology and mechanical equipment. Infrastructure, on the other hand, refers to the foundational elements of the system, such as physical facilities and infrastructure.

The annotators compile a data dictionary, grouping words by each dimension derived from the survey responses. This reference tool is then employed to track the frequency of phrases categorized under negative sentiments, breaking down the counts according to respective service dimensions. The approach to defining service quality dimensions encompasses several steps: initially, each aspect of service quality is extracted from the analysis of sentences containing the categorized terms. This classification focuses exclusively on negative sentiments. The frequency of each grouped term is quantified and subsequently scrutinized concerning the most pertinent factor. If the context of the complaint is not clear from the word alone, the surrounding words—preceding or following—are considered to achieve a more precise interpretation. Ultimately, these calculated frequencies aid in pinpointing the factors influencing customer satisfaction with the service.

3. RESULTS AND DISCUSSION

This section presents a comprehensive analysis of the findings, thoroughly examining their implications in the context of existing literature. The results are discussed in detail, emphasizing how they align with or differ from prior studies in the field. This analysis not only contributes to a deeper understanding of the subject but also proposes and inspire further exploration and advancement in this area.

3.1. Dataset

Online licensing service has a total of 4,600 companies registered, which include 168 partnerships (commanditaire vennootschap) and 4,432 limited liability companies. The forms of registered companies encompass Regional State-Owned Enterprises, Private State-Owned Enterprises, and Permanent Establishments. A total of 24,112 entries were collected from registered companies distributed across all provinces in Indonesia, with the highest concentration of company domiciles in DKI Jakarta. The distribution of registered companies that filled out the satisfaction survey questionnaire can be seen in Table 4.

Based on data collected from registered companies, 80% of the collected data was labeled with negative and positive tags for training. The labeling was done by five annotators, with one person acting as a decider in the event of differing sentiments. The annotators are translators at the MEMR. The data collected is free text and may contain words, numbers, special characters, emojis, and is not limited by the number of letters or words. The results of the dataset labeling process will be featured in Table 5.

Table 4. Province distributions of the registered company

No	Province	Number of companies
1	DKI Jakarta	1600
2	Jawa Barat	413
3	Jawa Timur	365

33	Gorontalo	6
34	Sulawesi Barat	3
	Total	4600

Table 5. Dataset results

Resume	Total
Annotators	5
Raw data	24412
Data training	19530
Data testing	4882
Negative sentiment	8837
Positive sentiment	15575

3.2. Data pre-processing

The pre-processing stage is necessary to remove words that appear frequently but do not carry positive or negative sentiments. Researchers perform pre-processing to maximize accuracy and accelerate computation time during the data classification. Figure 1 illustrates data pre-processing activities where Figure 1(a) shows a word cloud before pre-processing and Figure 1(b) shows a word cloud after pre-processing.

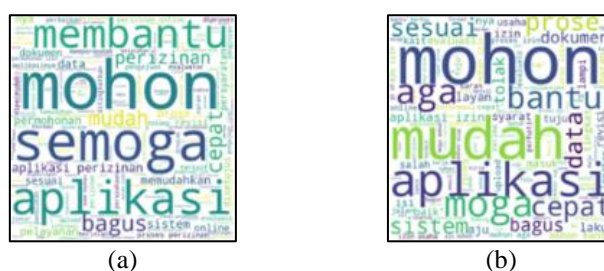


Figure 1. Data pre-processing (a) before and (b) after

3.3. Classification

This study employs six classification methods for both testing and training data. These methods include DT, RF, NB, SGD, LR, and k-NN. The classification methods are combined with converting text data to numeric through CountVectorizer and TF-IDF. The results of the tests are divided into two sentiments, positive and negative. All six methods are applied in both the training and testing phases. This research utilizes the N-gram process to enhance the accuracy of classification methods, limiting the use of N-grams to unigrams and bigrams. After segmenting sentences using N-grams and then classifying them, Table 6 presents the results of this process.

Table 6. Classification results

String extraction	Classifier	Confusion matrix	CountVectorizer		TF-IDF		
			Negative	Positive	Negative	Positive	
Unigram	k-NN	Negative	543	1240	187	1596	
		Positive	106	2952	54	3004	
	LR	Negative	1393	390	1422	361	
		Positive	244	2814	262	2796	
	NB	Negative	1392	391	1314	469	
		Positive	318	2740	256	2802	
	RF	Negative	1403	380	1482	301	
		Positive	285	2773	324	2734	
	DT	Negative	1286	497	1291	492	
		Positive	452	2606	448	2610	
	SGD	Negative	1423	360	1475	308	
		Positive	276	2782	287	2771	
	Bigram	k-NN	Negative	358	1425	59	1724
			Positive	88	2970	8	3040
LR		Negative	1451	332	1402	381	
		Positive	210	2848	239	2819	
NB		Negative	1406	377	1173	610	
		Positive	306	2752	146	2912	
RF		Negative	1337	446	1452	331	
		Positive	268	2790	267	2791	
DT		Negative	1299	484	1324	459	
		Positive	424	2634	446	2612	
SGD		Negative	1370	413	1469	314	
		Positive	252	2806	243	2815	

3.4. Evaluation

Precision, recall, and F1-score for each classification are calculated to correct the classification values mentioned [12]. Table 7 displays the results of these evaluations. These metrics help in understanding the performance of each classifier by providing details on the accuracy of the optimistic predictions (precision), the ability to find all the positive instances (recall), and a balance between precision and recall (F1-score), especially when dealing with imbalanced datasets. The comparison of each method using the confusion matrix in Table 7 indicates that bigram LR achieved the highest F1-score (89%). The F1-score is based on an imbalanced dataset between positive and negative classes. Based on these results, LR is chosen as the predictive model for the next steps. The sentiment analysis results using bigram CountVectorizer LR show that positive sentiments account for a more significant percentage, precisely 63.2%, compared to negative sentiments, which constitute 36.8%.

Table 7. Evaluation results

String algorithm	Classifier	CountVectorizer			Tf-Idf		
		Precision	Recall	F1-score	Precision	Recall	F1-score
Unigram	k-NN	0.75	0.72	0.68	0.70	0.66	0.56
	LR	0.87	0.87	0.87	0.87	0.87	0.87
	NB	0.85	0.85	0.85	0.85	0.85	0.85
	RF	0.86	0.86	0.86	0.87	0.87	0.87
	DT	0.8	0.8	0.8	0.8	0.81	0.81
	SGD	0.87	0.87	0.87	0.88	0.88	0.88
Bigram	k-NN	0.72	0.69	0.62	0.69	0.64	0.51
	LR	0.89	0.89	0.89	0.87	0.87	0.87
	NB	0.86	0.86	0.86	0.85	0.84	0.84
	RF	0.85	0.85	0.85	0.88	0.88	0.88
	DT	0.81	0.81	0.81	0.81	0.81	0.81
	SGD	0.86	0.86	0.86	0.88	0.88	0.88

3.5. Negative sentiment dimension

The researchers established a dictionary for each predetermined service quality dimension. The five annotators carry out the dictionary based on the dimensions of service quality, considering all the survey data available. Using the service quality dimension dictionary and the negative sentiment data, the researchers conducted data crawling and comparing the data dictionary and the survey responses previously filled out by respondents. Among the nine predetermined dimensions, the system, mechanism, and procedure dimensions dominate with a total of 4,487 negative sentiments from the overall negative sentiments expressed by service

users. Table 8 shows the mapping results between service quality dimensions, dictionary, and total negative sentiment.

Table 8. Dictionary of service quality dimension

Service quality dimension	Dictionary	Total negative sentiment	Highest word frequency
System, mechanism, and procedure	Guide, manual, mentor, socialization, register, example, reject, track, monitor, workflow, standard, government, draft, list, issue, evaluation, note, review, approve, validation, revision, tutorial, meet, process	4487	Process (1385)
Infrastructure and facilities	Browser, upload, show, file, chrome, firefox, download, access, integration, login, submit, feature, zip, limit, coordinate, error, character, connection, network, server, scroll, website, response, notification, chat, online, offline, synchronize, capacity	3789	Upload (1045)
Product specification of service type	Services, mining license (IUP) (<i>izin usaha pertambangan</i>) (IUP), minerba one data Indonesia (MODI), minerba one map Indonesia (MOMI), trade, transportation, import, online single submission (OSS), storage, warehouse, geothermal, fuel, new, extension, amendment, business license for supplying electricity (IUPTL) (<i>izin usaha penyediaan tenaga listrik</i>) (IUPTL), recommendation, mining business permit area (WIUP) (<i>wilayah izin usaha pertambangan</i>) (WIUP), ownership, Capital Investment Coordinating Board (BKPM) (<i>Badan Koordinasi Penanaman Modal</i>) (BKPM)	3197	IUP (853)
Requirements	Regulation, requirement, criteria, provision, condition, prerequisite, parameter, reference, demand, compliance, document, field	2462	Document (1022)
Completion time	Duration, length, period, speed, service-level agreement (SLA), time	1085	Time (730)
Handling of complaints, suggestions, and feedback	Response, action, acceptance, record, feedback, analysis, operator, follow, solution, supervision, report, management, RPIT, helpdesk, impact, contact, satisfaction	236	Response (91)
Executor's behavior	Attitude, professionalism, friendly, communication, discipline, cooperation, reliability, responsibility, responsive	186	Responsive (82)
Cost/tariff	Fee, price, charge, bill, invoice, payment, levy, surcharge	138	Payment (82)
Executor's competence	Ability, expertise, implementation, executor, performance, effectiveness	45	Executor (23)

The analysis identified three dimensions with the highest total negative sentiment based on word frequency. This prioritizes them for service improvement. The word frequency is linked with a word before or after to pinpoint more specific issues. In the dimensions of system, mechanism, and procedure, it was found that there is a need for a process evaluation of permits to ensure compliance with current rules and standard operating procedures. Additionally, there are processes related to integrated permits that businesses do not understand well. This issue necessitates socialization to business entities using the service about the service process through the application, both for internal ministry permits and those integrated with other ministries, providing easily understandable guidelines for users. Furthermore, an internal audit by the MEMR is expected to simplify the submission process and speed up the permit evaluation process.

The limitation in file format and size during the file upload process contributed the most to negative sentiment in the dimension of facilities and infrastructure, followed by difficulties filling out map coordinates and project locations. To address these issues, the ministry could implement a feature for validating file inputs, enhance the system's capacity to handle larger files according to user needs, and allow users more time to upload bigger files. Moreover, integration with the EMR single map service through Geoportal could simplify the process of filling out map coordinates, accompanied by a comprehensive user guide.

Finally, in the dimension of product service specifications, it was found that the IUP permit type received the highest frequency of negative sentiment, followed by OSS permits and WIUP permits, which were frequently mentioned in complaints. An internal audit could be conducted for these three types of services to optimize the application and business process. These findings suggest a need for targeted improvements in specific areas to enhance service efficiency and reduce negative sentiment among users. Furthermore, providing comprehensive integrated guidelines and strengthening the function of the contact center would be very helpful.

4. CONCLUSION

This research presents an analysis of user sentiment from an online licensing service in the EMR sector, starting with the creation of a training dataset to build a predictive model. The process of developing the predictive model began with data transformation, which included steps like transformation, tokenization, stemming, and removing stopwords, followed by data classification to determine the most suitable predictive model for the case study. The bigram LR model emerged as the most accurate, with the best precision and

recall rates. An analysis of data from 4,600 companies within the sector yielded 24,412 responses from a public satisfaction survey, which were categorized into nine dimensions. From these responses, 36.8% expressed negative sentiments toward the service received, with the highest number of negative sentiments directed at the system, mechanism, and procedure, while the competency of implementers was the least criticized aspect. This research suggests that future studies could explore other variables, such as the correlation between service satisfaction levels and investment levels, and expand the scope to include more areas relevant to the Ministry of State Apparatus Utilization and Bureaucratic Reform, especially concerning public satisfaction surveys and their impact on online services.

ACKNOWLEDGEMENTS

I extend sincere gratitude to the University of Indonesia for their unwavering support and invaluable resources. Additionally, I am deeply thankful to the Center for Data and Information Technology MEMR for their essential collaboration and insights that were crucial in conducting the case study. The expertise and data provided by these esteemed institutions have significantly enriched this research, making it academically robust and practically relevant to the field of information technology and government services.




REFERENCES

- [1] M. A. Nasution and S. Y. Regif, "Study of SMS gateway service application at the one stop agency of Medan city investment and service for licensing customer satisfaction," *IOP Conference Series: Materials Science and Engineering*, vol. 648, no. 1, pp. 1–7, 2019, doi: 10.1088/1757-899X/648/1/012020.
- [2] R. Copeland and N. Crespi, "Analyzing consumerization - Should enterprise business context determine session policy?," in *2012 16th International Conference on Intelligence in Next Generation Networks*, 2012, pp. 187–193, doi: 10.1109/ICIN.2012.6376024.
- [3] S. M. Hizam, W. Ahmed, H. Akter, and I. Sentosa, "Understanding the public rail quality of service towards commuters' loyalty behavior in Greater Kuala Lumpur," *Transportation Research Procedia*, vol. 55, pp. 370–377, 2021, doi: 10.1016/j.trpro.2021.06.043.
- [4] A. G. Awan and M. Azhar, "Consumer Behaviour Towards Islamic Banking in Pakistan," *European Journal of Accounting Auditing and Finance Research*, vol. 2, no. 9, pp. 42–65, 2014.
- [5] P. Kotler, K. Keller, and M. Brady, *Marketing Management—European Edition*. Harlow, England: Pearson Prentice Hall Publishing, 2009.
- [6] V. H. Y. Chan, D. K. W. Chiu, and K. K. W. Ho, "Mediating effects on the relationship between perceived service quality and public library app loyalty during the COVID-19 era," *Journal of Retailing and Consumer Services*, vol. 67, pp. 1–16, 2022, doi: 10.1016/j.jretconser.2022.102960.
- [7] L. Zhang and B. Liu, "Sentiment Analysis and Opinion Mining," in *Encyclopedia of Machine Learning and Data Mining*, Boston, MA: Springer, 2017, pp. 1152–1161, doi: 10.1007/978-1-4899-7687-1_907.
- [8] B. Liu, *Sentiment Analysis and Opinion Mining*. California, USA: Morgan & Claypool Publishers, 2012.
- [9] C. A. Haryani, A. N. Hidayanto, N. F. A. Budi, and Herkules, "Sentiment Analysis of Online Auction Service Quality on Twitter Data: A case of E-Bay," in *2018 6th International Conference on Cyber and IT Service Management (CITSM)*, 2018, pp. 1–5, doi: 10.1109/CITSM.2018.8674365.
- [10] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," *Social Network Analysis and Mining*, vol. 11, no. 1, pp. 1–19, 2021, doi: 10.1007/s13278-021-00776-6.
- [11] B. Widoyono, I. Budi, P. K. Putra, and A. B. Santoso, "Sentiment analysis of learning from home during pandemic covid-19 in Indonesia," in *Proceedings of the International Conference on Economics, Business, Social, and Humanities (ICEBSH 2021)*, 2021, vol. 570, pp. 464–472, doi: 10.2991/assehr.k.210805.073.
- [12] M. S. Md Suhaimin, M. H. A. Hijazi, E. G. Moug, P. N. E. Nohuddin, S. Chua, and F. Coenen, "Social media sentiment analysis and opinion mining in public security: Taxonomy, trend analysis, issues and future directions," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 9, pp. 1–25, 2023, doi: 10.1016/j.jksuci.2023.101776.
- [13] Z. Li and W. Liu, "Automatic decision support for public opinion governance of urban public events," in *Recent Trends in Intelligent Computing, Communication and Devices*, Singapore: Springer, 2020, pp. 47–53, doi: 10.1007/978-981-13-9406-5_7.
- [14] T. Zhang and C. Cheng, "Temporal and Spatial Evolution and Influencing Factors of Public Sentiment in Natural Disasters—A Case Study of Typhoon Haiyan," *ISPRS International Journal of Geo-Information*, vol. 10, no. 5, pp. 1–19, 2021, doi: 10.3390/ijgi10050299.
- [15] R. Wadawadagi and V. Pagi, "Disaster Severity Analysis from Micro-Blog Texts Using Deep-NN," *Advances in Intelligent Systems and Computing*, vol. 1176, pp. 145–157, 2021, doi: 10.1007/978-981-15-5788-0_14.
- [16] K. M. Ridhwan and C. A. Hargreaves, "Leveraging Twitter data to understand public sentiment for the COVID-19 outbreak in Singapore," *International Journal of Information Management Data Insights*, vol. 1, no. 2, pp. 1–15, 2021, doi: 10.1016/j.ijmei.2021.100021.
- [17] B. A. Prasetyo and Subagyo, "Twitter User Sentiment Analysis for Indonesian Language Texts Against Home Fix Broadband Service Providers" (Analisis Sentimen Pengguna Twitter untuk Teks Berbahasa Indonesia Terhadap Penyedia Layanan Home Fix Broadband)," in *Seminar Nasional Teknik Industri Universitas Gadjah Mada*, 2021, pp. 18–23.
- [18] I. E. Tiffani, "Optimization of Naive Bayes Classifier By Implemented Unigram, Bigram, Trigram for Sentiment Analysis of Hotel Review," *Journal of Soft Computing Exploration*, vol. 1, no. 1, pp. 1–7, 2020, doi: 10.52465/joscev.v1i1.4.
- [19] W. Haitao, H. Jie, Z. Xiaohong, and L. Shufen, "A short text classification method based on n-gram and cmn," *Chinese Journal of Electronics*, vol. 29, no. 2, pp. 248–254, 2020, doi: 10.1049/cje.2020.01.001.
- [20] N. Y. Hassan, W. H. Goma, G. A. Khoriba, and M. H. Haggag, "Credibility detection in twitter using word n-gram analysis and supervised machine learning techniques," *International Food and Agribusiness Management Review*, vol. 23, no. 1, pp. 291–300, 2020, doi: 10.22434/IFAMR2019.0020.
- [21] J. Kruczek, P. Kruczek, and M. Kuta, "Are n-gram Categories Helpful in Text Classification?," in *Computational Science – ICCS 2020*, Cham: Springer, 2020, pp. 524–537, doi: 10.1007/978-3-030-50417-5_39.
- [22] Z. Drus and H. Khalid, "Sentiment analysis in social media and its application: Systematic literature review," *Procedia Computer Science*, vol. 161, pp. 707–714, 2019, doi: 10.1016/j.procs.2019.11.174.




- [23] A. Patel and K. Meehan, "Fake News Detection on Reddit Utilising CountVectorizer and Term Frequency-Inverse Document Frequency with Logistic Regression, MultinomialNB and Support Vector Machine," in *2021 32nd Irish Signals and Systems Conference (ISSC)*, 2021, pp. 1–6, doi: 10.1109/ISSC52156.2021.9467842.
- [24] S. Smetanin and M. Komarov, "Sentiment Analysis of Product Reviews in Russian using Convolutional Neural Networks," in *2019 IEEE 21st Conference on Business Informatics (CBI)*, 2019, pp. 482–486, doi: 10.1109/CBI.2019.00062.
- [25] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014, doi: 10.1016/j.asej.2014.04.011.
- [26] M. Goswami and P. Sabata, "Evaluation of ML-Based Sentiment Analysis Techniques with Stochastic Gradient Descent and Logistic Regression," in *Trends in Wireless Communication and Information Security*, Singapore: Springer, 2021, pp. 153–163, doi: 10.1007/978-981-33-6393-9_17.
- [27] H. Christian, M. P. Agus, and D. Suhartono, "Single document automatic text summarization using term frequency-inverse document frequency (TF-IDF)," *ComTech: Computer, Mathematics and Engineering Applications*, vol. 7, no. 4, pp. 285–294, 2016, doi: 10.21512/comtech.v7i4.3746.
- [28] J. García-Balboa, M. Alba-Fernández, F. Ariza-López, and J. Rodríguez-Avi, "Analysis of Thematic Similarity Using Confusion Matrices," *International Journal of Geo-Information*, vol. 7, no. 6, pp. 1–10, 2018, doi: 10.3390/ijgi7060233.
- [29] M. F. Rahman, D. Alamsah, M. I. Darmawidjaja, and I. Nurma, "Classification for Diabetes Diagnosis Using the Method Bayesian Regularization Neural Network (RBNN)" (Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization Neural Network (RBNN)), *Jurnal Informatika*, vol. 11, no. 1, pp. 36–45, 2017, doi: 10.26555/jifo.v11i1.a5452.
- [30] M. F. Fibrianda and dan Adhitya Bhawiyuga, Comparative analysis of the accuracy of attack detection on computer networks with metode naïve bayes dan support vector machine (SVM)" (Analisis perbandingan akurasi deteksi serangan pada jaringan komputer dengan metode naïve bayes dan support vector machine (SVM)), *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 9, pp. 3112–3123, 2018.
- [31] A. Parasuraman, V. A. Zeithaml, and L. L. Berry, "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality.," *Journal of Retailing*, vol. 64, no. 1, pp. 12–40, 1988.
- [32] N. Paoprasert, B. Saputro, A. N. Hidayanto, and Z. Abidin, "Measuring service quality in the telecommunications industry from customer reviews using sentiment analysis: a case study in PT XL Axiata," *International Journal of Innovation and Learning*, vol. 30, no. 2, pp. 188–200, 2021, doi: 10.1504/IJIL.2021.10040271.
- [33] X. Papadomichelaki and G. Mentzas, "e-GovQual: A Multiple-Item Scale for Assessing e-Government Service Quality," *Government Information Quarterly*, vol. 29, no. 1, pp. 98–109, 2012.
- [34] T. (David) Lee, S. Lee-Geiller, and B.-K. Lee, "A validation of the modified democratic e-governance website evaluation model," *Government Information Quarterly*, vol. 38, no. 4, 2021, doi: 10.1016/j.giq.2021.101616.
- [35] L. Chung and J. C. S. Do Prado Leite, "On non-functional requirements in software engineering," in *Conceptual Modeling: Foundations and Applications*, Berlin, Heidelberg: Springer, 2009, pp. 363–379, doi: 10.1007/978-3-642-02463-4_19.
- [36] W. J. A. Al-Nidawi, S. K. J. Al-Wassiti, M. A. Maan, and M. Othman, "A review in E-government service quality measurement," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 10, no. 3, pp. 1257–1265, 2018, doi: 10.11591/ijeecs.v10.i3.pp1257-1265.
- [37] C. W. Tan, I. Benbasat, and R. T. Cenfetelli, "It-mediated customer service content and delivery in electronic governments: An empirical investigation of the antecedents of service quality," *MIS Quarterly: Management Information Systems*, vol. 37, no. 1, pp. 77–109, 2013, doi: 10.25300/MISQ/2013/37.1.04.
- [38] F. Sá, A. Rocha, and M. P. Cota, "Potential dimensions for a local e-Government services quality model," *Telematics and Informatics*, vol. 33, no. 2, pp. 270–276, 2016, doi: 10.1016/j.tele.2015.08.005.
- [39] N. Widiyanto, P. I. Sandhyaduhita, A. N. Hidayanto, and Q. Munajat, "Exploring information quality dimensions of government agency's information services through social media: a case of the Ministry of Education and Culture in Indonesia," *Electronic Government, an International Journal*, vol. 12, no. 3, p. 256, 2016, doi: 10.1504/EG.2016.078421.
- [40] M. S. Janita and F. J. Miranda, "Quality in e-Government services: A proposal of dimensions from the perspective of public sector employees," *Telematics and Informatics*, vol. 35, no. 2, pp. 457–469, 2018, doi: 10.1016/j.tele.2018.01.004.

BIOGRAPHIES OF AUTHORS



Ahmad Hizqil    is a Bachelor of Information Technology graduate from Islamic State University Syarif Hidayatullah Jakarta. He is an IT professional specializing as a computer system manager in the government sector. His experience includes integrating various government systems and developing government services in EMR sectors, where he has served as a project manager and system analyst. He holds certifications as an assessor of competency and a certified ICT project manager (CIPTM). He is pursuing his Master of Information Technology at the University of Indonesia. He can be contacted at email: ahmad.hizqil@ui.ac.id.



Yova Ruldeviyani    is a lecturer and researcher at the Faculty of Computer Science, Universitas Indonesia. She holds a bachelor's and a master's from Universitas Indonesia. Her research interests include databases, business intelligence, enterprise resource planning, e-government, and software development. Her expertise spans several areas, including IS/IT management and governance, software engineering, data science and analytics, and e-government. Among her notable publications are research on customer personal data protection in ride-sharing applications, knowledge management system design in Indonesian hospitals, and data governance structure design based on the data management body of knowledge (DMBOK) framework. She can be contacted at email: yova@cs.ui.ac.id.