

## **UNPACKING PUBLIC PERCEPTIONS OF QRIS WITH TWITTER DATA: A VADER AND LDA METHODOLOGY**

**Dzakiya Ishmatul Ulya, Anang Kunaefi, Dwi Rolliawati, Bayu Adhi Nugroho**

Departement of Information Systems, UIN Sunan Ampel Surabaya, Surabaya, Indonesia  
e-mail: [dishmatul@gmail.com](mailto:dishmatul@gmail.com), [akunaefi@uinsby.ac.id](mailto:akunaefi@uinsby.ac.id), [dwi\\_roll@uinsby.ac.id](mailto:dwi_roll@uinsby.ac.id), [bayu@uinsby.ac.id](mailto:bayu@uinsby.ac.id)

Received: 4 July 2023 – Revised: 1 November 2023 – Accepted: 10 November 2023

### **ABSTRACT**

*QRIS, a mobile payment transaction system standardized by Bank Indonesia, has become the subject of extensive public discourse on Twitter. Employing VADER for sentiment analysis and LDA for topic modeling, this study aims to capture the nuanced perspectives of the Indonesian public toward QRIS. Our methodology includes real human validation for tweets that have been initially labeled by VADER. Our unique contributions lie in employing a mixed-methods approach for comprehensive sentiment and topic analysis, as well as making our dataset publicly available for future research. We achieve a sentiment labeling accuracy of 81.66%, uncovering that 67% of the sentiment towards QRIS is positive, 28.2% negative, and 4.17% neutral. Positive tweets mostly cover six dominant topics with a  $C_v$  value of 0.488037, whereas negative sentiments are concentrated around three dominant topics with a  $C_v$  value of 0.383938. These findings not only affirm the generally positive public response towards QRIS but also highlight areas requiring attention for its continued success. Our study translates these insights into actionable recommendations, aiming to provide a multidimensional understanding that stakeholders can leverage for system enhancement. This study serves as a foundation for future works in sentiment analysis and public opinion mining related to financial technologies, particularly in the Indonesian context.*

**Keywords:** mobile payment, natural language processing, opinion mining, *Qris*, sentiment analysis, topic modeling.

### **I. INTRODUCTION**

**M**OBILE payment technology has rapidly spread worldwide. Digital wallets and digital currencies are predicted to replace debit and credit cards [1]. Contactless and mobile payments have become alternatives for individuals concerned about the health risks of using cash during the COVID-19 pandemic. Over 27% of small and medium-sized businesses in the United States reported an increase in mobile payment services during the pandemic [2].

In Indonesia, several mobile payment methods in the form of digital wallets are widely used and popular among various segments of society, such as GoPay, Dana, Ovo, and ShopeePay. The advantages of using these digital wallets lie in their convenience, comfort, and security. However, sellers or merchants must provide QR codes from each Payment System Service Provider (PJSP) separately during transactions. The emergence of various new PJSPs has made the QR code system fragmented and ineffective for both merchants and customers.

As the authority regulating the National Payment Gateway (GPN), Bank Indonesia has designed a system to integrate various types of PJSPs due to the ineffectiveness of using different QR codes for transactions. Bank Indonesia has created a standardized QR code called QRIS (Quick Response Code Indonesian Standard) to facilitate and streamline mobile payment transactions in Indonesia. The standardized QR code was developed in collaboration with the Indonesian Payment Systems Association (ASPI) to promote transaction efficiency, accelerate financial inclusion, support SMEs, and drive economic growth. QRIS is a standardized QR code that can be used for all types of digital transactions [3]. With the concept of integrating QR codes among PJSPs, QRIS was developed to fulfill the stages in implementing the Indonesian Payment System (SPI) vision by 2025. Governor of Bank Indonesia Perry Warjiyo stated that Indonesia has confirmed the commitment of four ASEAN countries

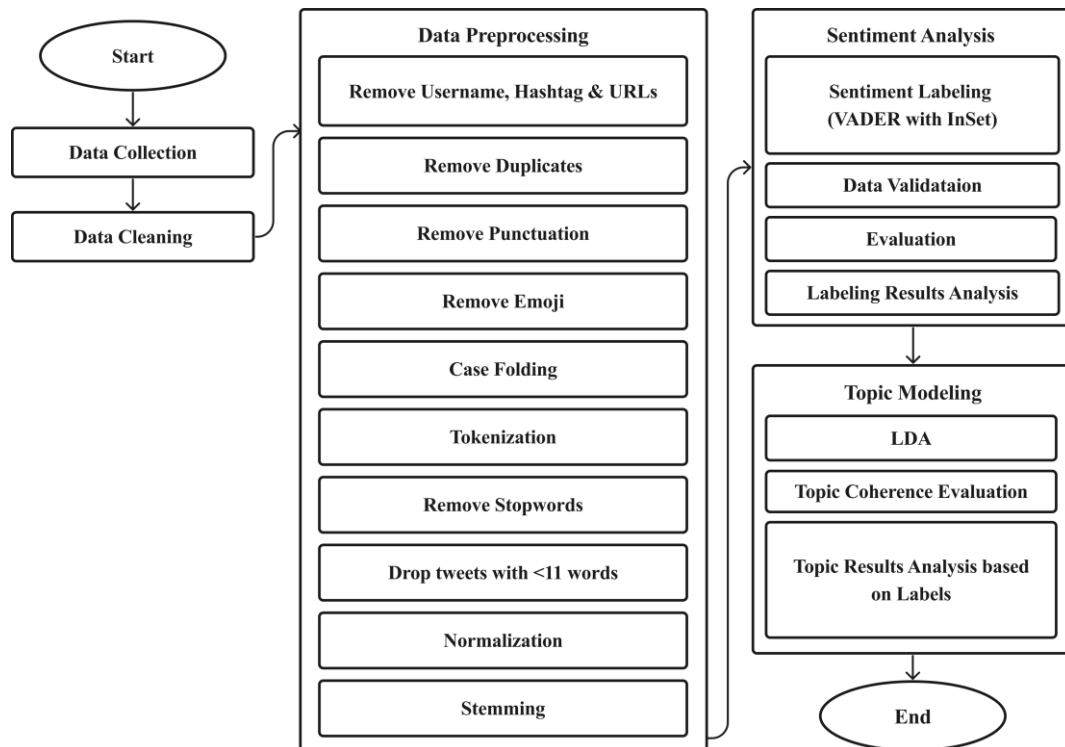


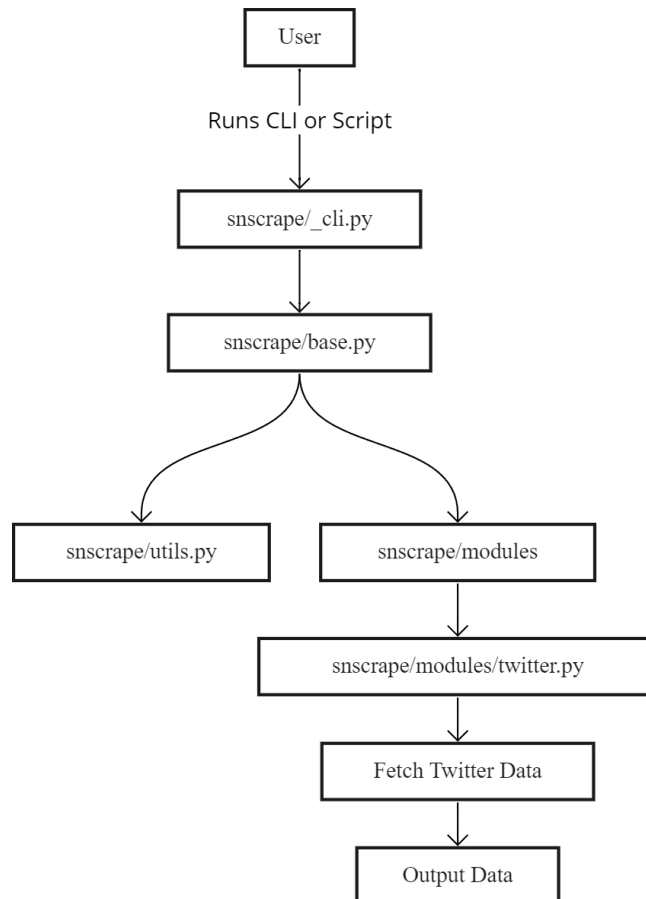
Figure 1. Research Methodology

to connect their payment systems with QRIS, including neighboring countries such as Singapore, Thailand, the Philippines, and Malaysia [4].

The presence of QRIS has helped many parties in facilitating transactions, leading to many Indonesian individuals sharing their experiences of using QRIS on social media. Social media platforms such as Twitter, Instagram, or Facebook have become popular communication channels in the past decade for individuals to express their feelings, thoughts, and ideas. There are no strict limitations on social media, allowing anyone to write about almost anything, covering all aspects and topics of life, ranging from daily life, politics, health, education, services, and more. In 2022, Indonesia ranked among the top five countries with the highest number of active Twitter users, accounting for 4.23% or 18.45 million users out of a total of 436 million Twitter users worldwide [5]. Supported by a 31.3% increase in the number of users compared to the previous year [6]. Based on the significant number of active Twitter users in Indonesia, the dataset from the Twitter platform can be utilized in this study.

The large number of Twitter users in Indonesia makes it highly interesting and beneficial to analyze the sentiment of the community towards specific topics. Sentiment analysis, also known as opinion mining, is a field that examines the analysis of opinions, feelings, judgments, emotions, and attitudes of individuals towards entities such as issues, services, products, individuals, organizations, events, and topics [7]. Generally, sentiment analysis can be performed using supervised machine learning approaches or lexicon-based rule-based approaches [8]. There are several lexicon-based sentiment analysis methods, and one commonly used method is VADER (Valence Aware Dictionary for sEntiment Reasoning) [9]–[15]. When analyzing sentiment using a dataset of tweets, VADER achieved the highest accuracy level with a score of 0.96, surpassing 11 other methods [16].

The sentiment analysis encompasses various fields of study, one of which is topic-based sentiment analysis. Topic-based sentiment analysis provides a foundation for managing online public opinions and plays a crucial role in identifying popular subjects and sentiment changes on the internet [17]. Topic Modeling is an unsupervised learning method that assumes each document is a mixture of topics, and each topic represents a probability distribution over words [7]. Latent Dirichlet Allocation (LDA) is considered a flexible and adaptive method for topic modeling [18]. Previous research has combined lexicon-based methods like VADER with LDA topic modeling to analyze text [11], [13], [14], [19], [20]. In this study, LDA-based Topic Modeling is employed to extract and identify the topics discussed by the public when expressing their opinions, evaluations, and sharing their experiences using QRIS, based on tweets labeled using VADER.



*Figure 2 How snsrape Works*

This study employs a comprehensive methodology that includes data collection from Twitter, data cleaning, and preprocessing. Sentiment analysis is performed using VADER, and topic modeling utilizes LDA. Validation is conducted manually with participants familiar with QRIS and technology. These methods aim to extract public opinion on QRIS, so that it can be evaluated and improved to enhance user experience.

Given this context, this research serves multiple purposes and makes several key contributions. Firstly, it can be utilized to determine public opinions on QRIS and evaluate its implementation effectiveness. This offers actionable recommendations for Bank Indonesia to enhance the payment system using QRIS. Secondly, we will publicly release our unique dataset of QRIS-related tweets, making it a valuable resource for future research. Lastly, our study provides specific and timely findings on QRIS, particularly focusing on the most recent period.

## II. RESEARCH METHOD

In order to make it easier to understand the preparation process and research steps in detail, a flow chart was made which can be seen in Figure 1.

### A. Data Collection

The tweet data was obtained through a scraping process using the snsrape library [21] in the Python programming language. The workflow of the snsrape library for scraping Twitter is as follows (see Figure 2 for a flow chart).

1. The user initiates the process by running a CLI command or a Python script, which calls the `_cli.py` module.
2. The `_cli.py` module then interacts with the `base.py` module, which serves as the foundation for the scraping process.
3. The `base.py` module utilizes utility functions from `utils.py` for various tasks like handling requests and parsing data.

4. The base.py module also calls the appropriate social media module from the modules directory, in this case, twitter.py.
5. The twitter.py module is responsible for fetching Twitter data based on the specified criteria.
6. Finally, the fetched Twitter data is outputted for further analysis.

The tweets acquired specifically contain the keyword "QRIS" and fall within the date range from the initial release of QRIS on August 17, 2019, to August 17, 2022. This targeted approach ensures that the dataset is relevant to our study of public sentiment towards QRIS during this specific period.

#### *B. Data Cleaning*

In the data cleaning stage, tweets that may be spam are cleaned. Spam tweet cleaning is done by removing all tweets from users who have tweets containing words from the list of words that are not relevant to the research.

#### *C. Data Preprocessing*

##### 1) Remove Username, Hashtag, URLs

Username, Hashtag, URLs removal needs to be done on tweets in the form of replies, tweets that mention usernames or hashtags, and tweets that contain URL links or image links. This stage is proven to improve the accuracy of sentiment labeling [22].

##### 2) Remove Duplicates

Removal of duplicate tweets is done after performing URL cleanup. URL links in tweets are usually image links, news links, and so on. Since the image listed in the tweet is a unique URL link, if the spam tweet contains a unique link, it will be detected in the form of the same tweet after URL cleanup.

##### 3) Remove Punctuation

Punctuation removal can also affect accuracy [23]. At this stage is assisted by the Python `string.punctuation` module which makes it easy to remove punctuation marks in each tweet.

##### 4) Remove Emoji

Emojis can influence in the process of labeling and modeling topics [22]. The emoji section was removed from tweets because there was no library available that could convert emoji into text that represented in Indonesian.

##### 5) Case Folding

The use of upper- or lower-case letters is considered to have no difference. Case folding is a stage in data pre-processing that converts all letters in tweets into lowercases to make them equal, so that 'GOOD' equals 'good'.

##### 6) Normalization

Indonesia, an archipelagic country, is culturally diverse, encompassing various races and ethnicities. In everyday communication, Indonesians frequently employ a mix of languages, including regional dialects and slang. Consequently, addressing out-of-vocabulary (OOV) words becomes crucial due to the extensive usage of conversational terms and phrases on social media platforms. One approach to tackle this issue involves normalizing OOV words based on the Colloquial Lexicon [24]. This lexicon, offers comprehensive coverage for normalizing not only slang but also commonly utilized words on social platforms, including elongated terms and misspellings.

##### 7) Tokenization

Tokenization is a method of dividing sentences into individual words. In addition, in the tokenization process, it can also be counted the number of times the word appears in a tweet, which is often referred to as a token.

##### 8) Remove Stopwords

Opinions in Indonesian contain many unimportant words (stop words) and slang words [25]. At this stage, the removal of unimportant words is done in order to increase efficiency in the classification of sentiment labeling [23].

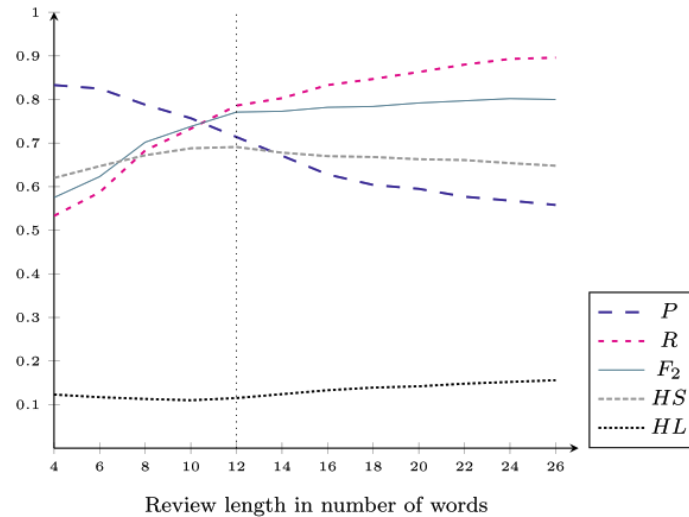


Figure 3. Dictionary Matching [27]

#### 9) Drop Tweets

The length of the text in topic modeling has a significant influence [26]. In this study, we restricted our analysis to tweets containing 12 words or more, in line with the methodology employed by [27] for analyzing App Store application reviews. This text length was found to yield good and balanced performance in dictionary matching, as evidenced by the generated precision and recall metrics. These findings are visually represented in Figure 3.

#### 10) Stemming

Stemming is the process of converting the root or word base of a word. This Indonesian stemming process is also carried out using the Sastrawi library [28].

#### D. Sentiment Analysis

Sentiment labeling classification involves utilizing the nltk VADER library [29], which employs the Lexicon InSet as a replacement for its default lexicon. Once the InSet Lexicon [30] is prepared for use, the compound score of preprocessed tweets will be calculated based on word weights from the InSet Lexicon.

Data validation (labeling validation) and evaluation were carried out by sampling data using random sampling techniques [31], [32]. Data validation is done manually with human participation. Participants must have native Indonesian language qualifications and have knowledge about technology, especially QRIS. Digital wallet users in Indonesia are dominated by people aged 18-24 years [33], so they are the main target of participants in this validation process.

#### E. Topic Modeling

Upon successful labeling of the tweets data, feature extraction is conducted for each label to enable processing by LDA. The conversion of tweets into a dictionary is necessary to assign a unique ID to each word. Subsequently, the conversion to the Bag of Words format is performed to facilitate topic distribution estimation by LDAs.

Following the conversion into the Bag of Words format, the LDA training model determines the range of topics, alphas, and passes to be utilized. Moreover, it is essential to specify the evaluation metrics for assessing the LDA model. In this case, the chosen metrics are  $C_V$  and  $C_{UMass}$ .

### III. RESULTS AND DISCUSSIONS

The study used data taken from Twitter. The data taken are tweets that have the keyword "QRIS" and tweets that are available from the release of QRIS on August 17, 2019 to August 17, 2022. Data collection used the snsrape library and collected 244,708 tweets.

Data cleaning was done by deleting all tweets from users who had tweets containing words from a list of words that were irrelevant to the study and could potentially be spam. The irrelevant words were

TABLE 1.  
EVALUATION RESULTS

Measure	Results
Accuracy	81.66%
Precision	82.32%
Recall	81.66%
F1 Score	81.69%

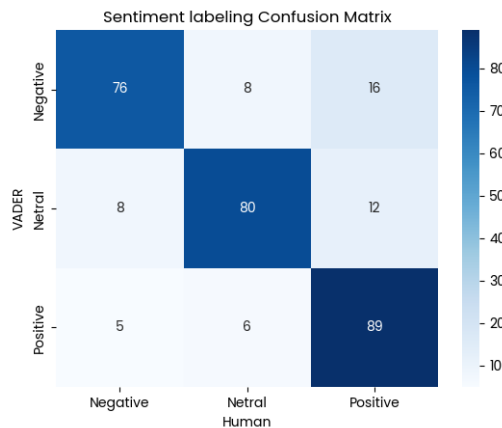


Figure 4. Confusion Matrix

collected manually. Examples of potentially spammy words such as: Netflix, iflix, Spotify, premium, etc. After data cleaning, a total of 73,764 tweets were obtained.

Data preprocessing is done to prepare the data to be ready for processing. The data preprocessing stage consists of Remove Username, Hashtag, URL, Remove Duplicate, Remove Punctuation, Remove Emoji, Case Folding, Normalization, Tokenization, Remove Stopwords, Drop Tweet, and Stemming. After preprocessing the data, sentiment labeling is done using VADER [29].

Sentiment labels using VADER [29] with InSet Lexicon [30] are derived from compound scores based on word weights on InSet Lexicon. There are three labels produced, namely positive, neutral, and negative.

Data validation was conducted to verify the labeling generated by VADER. To increase objectivity, each tweet was validated by three participants. A random sample of 300 tweets from a total of 21,536 labeled tweets was used for validation. The sample represented 14,555 positive tweets, 6,082 negative tweets, and 899 neutral tweets. With a confidence level of 95% and a margin of error of 5.62%, this sample accurately approximates the population, ensuring reliable results. In other words, if we perform sampling repeatedly, then 95% of the results will produce correct values or close to the true values [34].

Based on the results of labeling validation carried out manually, a confusion matrix is shown in Figure 4. VADER tends to be incorrect in labeling tweets that are actually positive. This is allegedly because people tend to make tweets related to QRIS using negative words that actually have the context of conveying the advantages of QRIS. Examples of tweets such as "*Eh sama dong, balik ke indo gih, skrg beli gorengan udah bisa pake QRIS loh jadi gausah ribet pake tunai dan ribet cari kembalian wkwk* (Hey, come back to Indo, now you can buy fried snacks using QRIS, so no need to bother with cash and change haha)" and "*Di sekitar kantor gue bnyk warteg dan warung2 kecil trs semua pake qris, ktnya gk terima cash krn ribet kalo pas lg rame kalo pas lg rame* (There are many food stalls and small shops near my office, and all of them accept QRIS. They don't take cash because it's inconvenient when it's crowded)". VADER, on the other hand, determines sentiment labeling based on the weight of the words contained in the tweets.

Accuracy, precision, recall, and F1 Score can be seen in Table 1. Sentiment labeling using VADER with the InSet lexicon on the QRIS tweets dataset yielded an accuracy of 81.66%, which falls within the range of good accuracy. This is supported by similar research on sentiment labeling related to handling COVID-19, which achieved an accuracy of 81% [35], sentiment labeling on video comments with an accuracy of 79.78% [9], and sentiment labeling on food delivery service with an accuracy of 71% [36].

It can be seen that the precision value is higher than accuracy, recall, and F1 Score. Precision measures how accurately VADER identifies positive or negative sentiment in data, and higher precision values indicate that VADER performs better at identifying true sentiment.

*A. Sentiment Analysis Results*

Based on the results of VADER [29] sentiment labeling on tweets data, it was found that Indonesian QRIS Tweets Sentiment Distribution

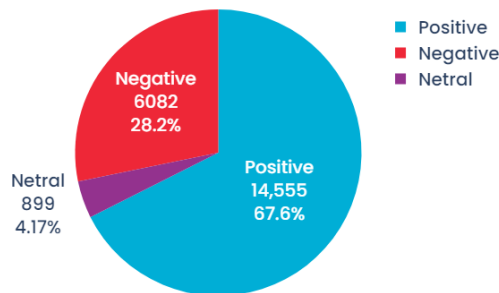


Figure 5. Sentiment Distribution



Figure 6. QRIS Trend Over Three Years

people tend to have positive sentiments. The distribution of Indonesian sentiment towards QRIS on Twitter can be seen in Figure 5.

QRIS trends since the release of QRIS on August 17, 2019 to August 17, 2022 continue to increase (Figure 6). This trend analysis is based on the top tweets with the most likes.

The first surge (December 2019 to April 2020) was dominated by tweets from various official accounts of Bank Jatim, Kompas, detik.com that informed, introduced, and educated the public about QRIS. Not only that, in this period there were also many discount and cashback promos from various merchants which were a form of introduction and efforts to increase QRIS adoption. This period also officially announced the first COVID-19 case in Indonesia and the implementation of Large-Scale Social Restrictions (PSBB), making this period the initial period of introduction and adoption of QRIS for the people of Indonesia.

During the second surge (December 2020 to April 2021), many Indonesians shared their experiences with QRIS. The implementation of Community Activity Restrictions (PPKM) during this period forced Indonesians to reduce the use of cash, leading to an increase in QRIS adoption. This shows the important role of QRIS in daily transactions during the pandemic, with many merchants' tweets promoting payment with QRIS.

During the third surge (August 2021 to April 2022), netizen discussion forums dominated the top tweets related to QRIS. Users expressed their frustration regarding the use of QRIS at various merchants due to the lack of education regarding QRIS among cashiers. QRIS only accepts payments from certain PJSPs, even though it is designed to accept payments from all PJSPs registered with Bank Indonesia [37], [38]. There is also a viral debate among netizens regarding the pronunciation of QRIS, either as "kyuris" or "keris", which indirectly increases public awareness and encourages participation in QRIS usage. Currently, top tweets continue to promote QRIS through various promotional offers from various merchants.

The fourth surge (May to August 2022) was characterized by viral threads and tweets on discussion forums, where individuals shared their experiences with QRIS adoption. Some praised its convenience







TABLE 2.  
LIST OF "TIDAK BISA"

Associated Words	Total
<i>Tidak bisa / Gak bisa / Ga bisa / Gabisa pakai</i>	119
<i>Tidak bisa / Gak bisa / Ga bisa / Gabisa QRIS</i>	86
<i>Tidak bisa / Gak bisa / Ga bisa / Gabisa transfer</i>	56
<i>Tidak bisa / Gak bisa / Ga bisa / Gabisa dipakai</i>	37
<i>Tidak bisa / Gak bisa / Ga bisa / Gabisa scan</i>	37

TABLE 3.  
LIST OF COHERENCE VALUE FOR POSITIVE LABELS TOPICS

Model ID	No. of Topics	Alpha	Passes	$C_V$	$C_{U\text{Mass}}$
48	6	1	100	0.488	-3.460
67	8	0.1	50	0.479	-4.333
71	8	1	50	0.478	-3.675
15	4	0.01	50	0.469	-3.017
22	4	1	20	0.466	-3.236

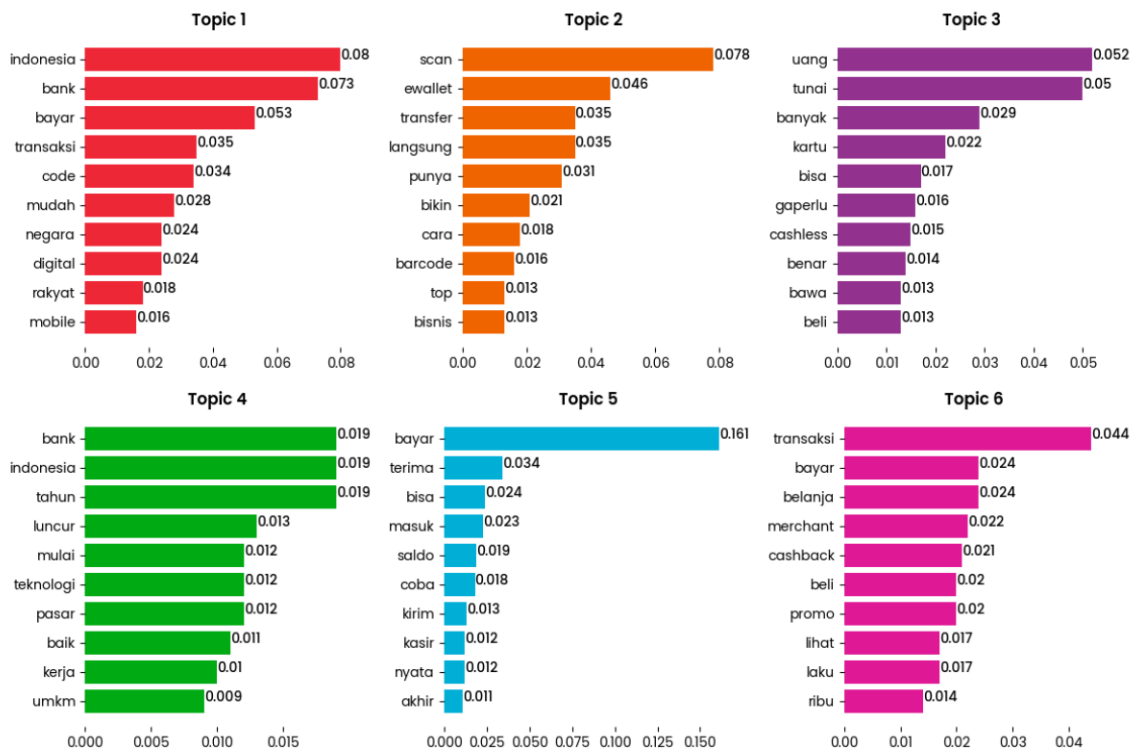


Figure 9. Positive Tweets Topic Modeling Results

Other words that often appear in the negative category are the words "*beli* (buy)", "*masuk* (credited)", "*tunai* (cash)", and "transfer" are also often mentioned by users who experience difficulties when making transactions using QRIS. In addition, terms such as "*gabisa*" or "*tidak bisa*" (meaning "unable") often appear in the context of technical problems or transaction failures. The term "*gabisa*" is most frequently used and is commonly followed by specific words, as detailed in Table 2.

### B. Topic Modeling Results

Topic modeling was done on each of the positive and negative label tweets. The LDA model running process is carried out iteratively with a range of 3-10 topics, alpha ranges of 0.01, 0.1, and 1, and passes 10, 20, 50, 100. This is done to find the model with the highest  $C_V$  and  $C_{U\text{Mass}}$ .

#### 1) Positive Labels

Table 3 shows top five models with highest  $C_V$  and  $C_{U\text{Mass}}$  for tweets with positive labels. Obtained models with six topics have  $C_V$  the highest with which  $C_{U\text{Mass}}$  is also relatively high. The study prioritizes how easy the topic is to interpret, so the model with the  $C_V$  highest is chosen for analysis.

There are six topics, each of which provides a different perspective on the use of QRIS by Indonesians. Figure 9 displays visualization of topic modeling results equipped with words and word weights on each topic.

The first topic includes words such as “Indonesia”, “bank”, “bayar (pay)”, “transaksi (transaction)”, “kode (code)”, “digital”, and “mobile” related to digital payments in Indonesia, proving that QRIS is considered a digital solution for efficient and accessible financial transactions via mobile devices. Payment transactions here not only refer to buying and selling payment transactions but also to zakat, infak, and alms payments [41]. Words like “rakyat (people)”, “negara (country)”, “Indonesia”, and “mudah (easy)” reveal the importance of ease and accessibility in digital payments for all Indonesians. This topic can represent that QRIS can provide an efficient digital alternative to help Indonesians make transactions easily and effectively. One of the tweets that represents this first topic is “Mantap deh pada era digital sekarang, apa-apa udah pada serba digital, mau bayar ini itu jg pake cara digital. Tinggal tap-tap atau scan QRIS beres deh (It's great in the digital era now, everything is already digital, even payments for this and that are done digitally. Just tap or scan QRIS and it's done)”.

The second topic includes words such as “scan”, “e-wallet”, and “transfer”, which discuss the use of QR codes or QRIS itself, using e-wallets in the business sector in Indonesia. Keywords such as “punya (have)” and “bikin (create)” indicate that QRIS users tend to create their own e-wallet accounts to store their money in digital form and make transactions using QRIS. In addition, keywords such as “bisnis (business)” and “top” can be attributed to the fact that QRIS is gaining popularity among businesses and helps entrepreneurs increase efficiency in making payments or transactions with customers. This topic proves that QRIS can be an efficient and popular payment solution for businesses in Indonesia. The following is one of the tweets, a businessman shared his experience while using QRIS: “Ini benar. QRIS memudahkan transaksi bisnis kami. Pembeli bisa langsung transfer ke rekening tanpa ribet memasukkan nomor rekening, dari sisi penjual juga nyaman karena langsung ada pemberitahuan adanya dana masuk sekian ke rekening... :) (This is true. QRIS facilitates our business transactions. The buyer can directly transfer to the account without the hassle of entering the account number, from the seller's side it is also convenient because there is immediately a notification of funds entering the account... :))”.

The third topic includes words such as “uang (money)”, “tunai (cash)”, and “kartu (card)” which are related to transaction security in cashless transformation in Indonesia. These words indicate the need to reduce dependence on cash in transactions. Words such as “bawa (carry)”, “gaperlu (no need)”, and “cashless” describe the practicality of QRIS as an alternative to digital payments that is easier and safer than carrying cash. This topic reveals that QRIS can help reduce the use of cash in Indonesia and improve transaction security. The following is one of the tweets related to this topic: “Skrng udah hampir ga pernah bayar ini itu pake cash, pake debit card pun udah jarang banget. Qris the best sih emang, mempermudah hidup makhluk cashless macem hamba dan lebih aman juga (I've almost never paid this and that using cash, using debit cards is already very rare. Qris the best indeed, makes life easier for cashless creatures like servants and safer too)”.

The fourth topic includes words such as “bank”, “Indonesia”, “teknologi (technology)”, and “UMKM (MSMEs)”, which can be related to Bank Indonesia's role in improving the ability of MSMEs to transact and expand markets through QRIS technology. Words such as “pasar (market)” and “baik (good)” represent the positive impact of using QRIS for MSMEs in increasing the accessibility of products and services to their markets. This topic represents that QRIS can provide important technological support for MSMEs in Indonesia. Research in Pematangsiantar City has also proven that QRIS implementation has a positive influence on the progress of MSMEs [42].

The fifth topic includes words such as “bayar (pay)”, “terima (receive)”, and “saldo (balance)”, which is about the positive experience of QRIS users who can pay and receive balances quickly and easily. Users can easily check their balance. Words like “coba (try)”, “bisa (can)” and “nyata (real)” mean QRIS has proven successful in improving transaction efficiency. This topic reveals how QRIS can help users make transactions quickly and easily, and provide a positive experience for users. One of the tweets with this topic is “Sekarang beramal semakin mudah ya, bisa diam-diam dengan memindai QRIS pakai GoPay, OVO, Link aja, langsung terdeteksi nama penerima. Lebih cepat dan praktis (Now doing charity is getting easier, you can secretly scan QRIS using GoPay, OVO, Link, immediately detected the recipient's name. Faster and more practical)”.

The last topic includes words such as “transaksi (transaction)”, “bayar (pay)”, “cashback”, “promo”, “belanja (shopping)”, “beli (buy)”, and “lihat (see)” can be associated with benefits or benefits for users in using QRIS in the form of cashback and special promos in making payments when shopping. This proves that QRIS can provide additional incentives for users to use digital payments so as to increase QRIS usage in Indonesia. In addition, the word “merchant” represents how businesses can use

TABLE 4.  
LIST OF COHERENCE VALUE FOR NEGATIVE LABELS TOPIC

Model ID	No. of Topics	Alpha	Passes	$C_V$	$C_{U,Mass}$
7	3	0.1	50	0.383	-3.837
12	3	1	100	0.373	-3.324
11	3	1	50	0.366	-3.585
88	10	0.01	100	0.364	-5.603
39	6	0.01	50	0.363	-5.070

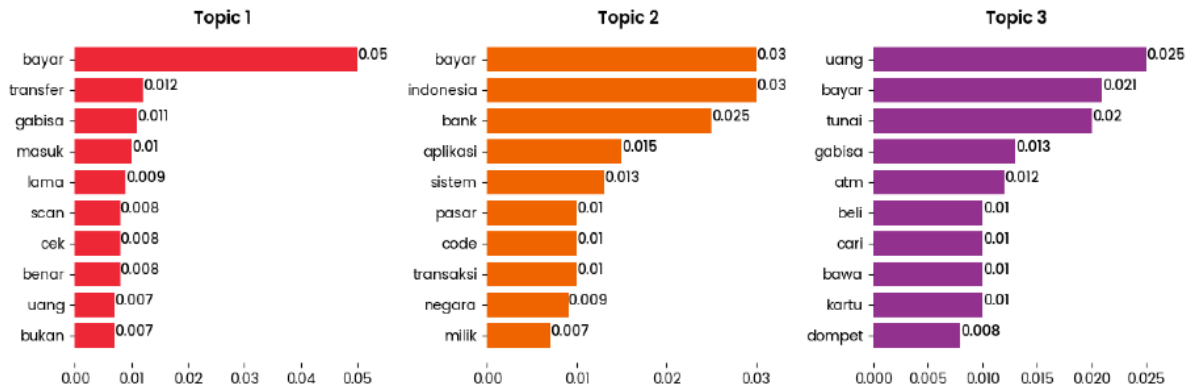


Figure 10. Negative Tweets Topic Modeling Results

QRIS to offer promos and cashback for their customers as an effective marketing strategy. This topic shows that QRIS can be an effective marketing tool for businesses in Indonesia and provide incentives for users to use QRIS. The following is one of the tweets that represents the sixth topic “saya ke Giant Blok M Plaza hari Minggu malam (21/3) bisa bayar pake QRIS. Besoknya dapat cashback 50% 🤗 (I went to Giant Blok M Plaza on Sunday night (21/3) to pay using QRIS. The next day I got 50% 🤗 cashback)”.

Overall, our findings underscore the multifaceted benefits of QRIS, indicating its positive impact on digital payments, business transactions, and transaction security. The sentiment analysis further validates these benefits, suggesting a predominantly favorable public opinion.

## 2) Negative Labels

The top five models for tweets with negative labels can be seen in Table 4. There is one model with three topics that have the highest  $C_V$  score.

Visualization of topic modeling results can be seen in Figure 10. There are several keywords that appear many times that show the concerns and problems experienced by Indonesians in using QRIS in their daily lives.

The first topic that came up was the issue of payments and transfers using QRIS. Many users complain about difficulties when making payments or transfers using QRIS. The words “*gabis* (can’t)”, “pay (bayar)”, “transfer”, “scan”, “login”, and “lama (slow)” represent some of the problems that are often faced are payments or transfers that cannot be made, payments or transfers that take a long time, and difficulties when scanning QRIS and checking whether the money has really entered the merchant account balance. In addition, there are also users who feel that the money they pay does not go into the merchant account balance. The following is an example of tweets complaining from one of the customers who uses QRIS to one of the PJSP “Min saya scan qris dana bisnis.. uang saya sudah terdebit tapi ngak masuk mercant toko dana.. gimana nih min.. bayar belanjaan tapi uangnya ngak masuk ke akun DANA yg punya toko.. solusinya bagaimana (I have scanned the QRIS of Dana Bisnis, but my money has been debited without being credited to the merchant's DANA. What should I do? I have paid for my purchases, but the money has not been transferred to the DANA account of the store. What is the solution for this?)”.

The second topic includes several keywords such as “bayar (pay)”, “Indonesia”, “bank”, “aplikasi (application)”, “sistem (system)”, “pasar (market)”, “code”, “transaksi (transaction)”, “negara (country)”, and “milik (own)”. Many tweets complained about QRIS accessibility issues, such as unresponsive PJSP applications, inadequate QRIS access in the market, and other technical problems that caused difficulties in making transactions. One of the tweets representing this topic was “jangan

pake qris dari bank juga kalo sinyal lagi jelek. setiap transaksi gagal ke penerima uang, biasanya transaksi berhasil kalo dicek dari mutasi rekening si yang bayar. saldo udah kepotong tapi transaksi ga berhasil (It is not advisable to use the QRIS from the bank if the signal is poor. In such cases, every transaction fails to reach the recipient, although the payment can be verified through the payer's account statement. The balance has already been deducted, but the transaction itself is unsuccessful)".

The third topic includes several keywords such as "uang (money)", "bayar (pay)", "tunai (cash)", "gabisa (can't)", "ATM", "beli (buy)", "cari (search)", "bawa (carry)", "kartu (card)", and "dompet (wallet)". Many tweets complain about the difficulty of making payments due to the uneven availability of QRIS, Indonesians tend to have difficulty finding merchants that accept payments with QRIS, so there is still dependence on cash and still have to have cards, wallets, or looking for ATMs to make transactions. Here's an example of tweets that represent this topic "Gara2 ATM pada trouble, terpaksa harus jauh2 ke kampus buat ngambil duit aja. Bener sih tinggal di Padang ini masih blm bisa buat belanja cashless. Di Grapari aja kmren harus bayar cash dan gak terima bayar pakai kartu, qris, ataupun transfer (Due to the ATM malfunction, I had to go all the way to campus just to withdraw some cash. It's true that living in Padang still doesn't fully support cashless transactions. Even at Grapari, I had to pay in cash yesterday and they didn't accept payment through cards, QRIS, or transfers.)".

Regarding payment and transfer issues in the first topic and QRIS accessibility issues in the second topic, this does not entirely depend on the QRIS system alone but can also come from the PJSP system used or the internet connection of both parties from both buyers and merchants [43] When compared to payment cards, it is indeed a limitation of QRIS that requires you to be connected to the internet in order to make transactions. It is undeniable that the fact that infrastructure in Indonesia in terms of internet connection is still uneven in certain areas is one of the obstacles to QRIS adoption and QRIS transaction constraints [38]. However, one of the backgrounds for QRIS development is based on Bank Indonesia data at the end of 2021, the ratio between the amount of electronic money in circulation is more than double that of circulating debit, namely 575 million and [44]. It is expected that Bank Indonesia can develop digital payment methods that can be made even when not connected to the internet, such as Near Field Communication (NFC) technology [45].

Bank Indonesia launched QRIS with two types, static and dynamic. Static QRIS is usually in the form of images, stickers, or posters containing QRIS QR Codes that are unique to each merchant and are fixed. The way this static QRIS works is by translating letters and numbers into dots. Static QRIS cannot be changed at all because each shape and position of the dot, represents a certain letter or number. Different from static QRIS, the links linked to dynamic QRIS can change. For example, for cashless payments at a store, the dynamic QRIS used will be different every time you make a transaction so that it is safer. Another advantage of dynamic QRIS is that it can track data. The tracking data referred to here includes the frequency with which QRIS is scanned, when the code is used, where the code is used, and what type of device is used to scan a particular QRIS code [46].

The most cost-effective and commonly used type of QRIS among MSMEs is the static QRIS. It is not surprising that there have been several security vulnerabilities and fraudulent schemes that harm merchants or MSMEs. Several viral tweets discuss various QRIS fraud schemes experienced by merchants and users, such as the following examples "UMKM yg pake QRIS hati2 ya. Yg jual ikan di pasar tadi kena tipu 500rb karena yg beli tunjukin bukti transfer/QRIS palsu dan yg jual ga cek dulu (MSMEs that use QRIS be careful. Those who sold fish at the market earlier were tricked by 500 thousand because those who bought showed fake transfer/QRIS proof and those who were selling didn't check first)", "Modus "kelebihan 0 waktu transfer bank dgn screen capture slip transfer palsu" dimodifikasi ke QRIS. Penipu memang laknat (The "excess 0 bank transfer time with fake transfer slip screen capture" fraud mode was modified to QRIS. Fraudsters are indeed cursed)", and "oh ini stiker QRIS nya ditempel dengan milik penipu ya 🤪 memang perlu edukasi customer kalau nama merchant harus di double check saat pembayaran (misal transaksi merchant A, saat scan QR kok bayarnya jadi ke X)) (Oh, this QRIS sticker is attached with the fraudster's own sticker. It's necessary to educate customers to double-check the merchant's name during payment (e.g., transaction with merchant A, but when scanning the QR code, the payment goes to X))". Although some tweets explain QRIS fraud schemes in 2022, as of now (12/04/2023), similar fraud schemes involving the placement of fake QRIS codes, disguised as donations at various mosque locations, still occur [47]. It is the responsibility of Bank Indonesia as the developer and registered PJSPs to improve the security of static QRIS and to be more proactive in responding to new types of fraud without having to wait for them to go viral before

taking action. On the other hand, Bank Indonesia should continue to expand educational efforts on transaction security for users with low financial literacy and digital literacy levels.

The current solution that can be implemented by SMEs to minimize fraud and ensure transaction success is to always be connected to the internet and consistently verify if the funds have been credited to their account for each transaction. It would be even better for SMEs with additional resources to upgrade to dynamic QRIS, which enhances the security of payments. As QRIS users on the payer's side, it is also crucial to be cautious when conducting transactions and be aware of cases involving fraudulent individuals who attach fake QRIS stickers (not the genuine QRIS intended for the merchant).

Regarding the limited availability of QRIS among merchants or traditional markets mentioned in the third topic, it is because the dataset of tweets used in this study consists of tweets from the past three years, which coincides with the government's efforts to digitally transform the traditional market sector, MSMEs, and merchants across Indonesia. In fact, the process of digitalization is still ongoing. According to reports from [idxchannel.com](https://www.idxchannel.com) (24/03/2023), it is rumored that all traditional markets in Depok are now able to accept payments through QRIS [48].

### *C. Recommendations for Stakeholders in the QRIS Ecosystem*

The data generated from our sentiment analysis and topic modeling offer a multi-dimensional view of public perceptions and experiences with QRIS. These insights have practical implications, suggesting specific areas for improvement for both primary and secondary participants in the QRIS ecosystem for Bank Indonesia & ASPI.

#### 1) Security Enhancements:

We have identified several cases of fraud and security vulnerabilities, particularly with static QRIS. Focus should be on enhancing security measures, possibly by encouraging the adoption of dynamic QRIS among MSMEs.

#### 2) Educating Users:

Our findings suggest that more extensive educational efforts are crucial for fraud prevention and encouraging responsible QRIS use, especially among users with low financial and digital literacy levels.

#### 3) Enhancing User Experience with NFC:

The positive sentiment reveals that users find QRIS convenient. However, there's room for improvement. Consider introducing NFC technology for more seamless transactions. For PJSPs:

#### 4) Enhancing Payment and Transfer Systems:

Our topic modeling for negative labels revealed issues related to payments and transfers. These providers should prioritize improving the speed and reliability of these processes.

#### 5) Accessibility and Technical Issues:

Tweets indicated that QRIS apps are sometimes unresponsive and not adequately accessible in certain markets. Addressing these issues can lead to increased adoption rates.

#### 6) Promotions and Incentives:

Our analysis showed that promotions and cashbacks effectively encourage QRIS adoption. More such strategic promotions could attract a broader user base.

By implementing these recommendations, Bank Indonesia, ASPI, and PJSPs in the QRIS ecosystem can address current pain points and enhance user experience, thereby contributing positively to Indonesia's financial ecosystem.

## **IV. CONCLUSION**

The primary problem this research addresses is the lack of comprehensive sentiment and topic analysis on QRIS, a widely-used mobile payment system in Indonesia. By employing VADER for sentiment analysis and LDA for topic modeling, this study aims to fill this gap and provide a nuanced understanding of public sentiment and issues surrounding QRIS.

Sentiment analysis on QRIS using VADER reveals a largely positive sentiment from Indonesian society, comprising 67.6% positive, 28.2% negative, and 4.17% neutral sentiment. Our methodology

includes real human validation for tweets that have been initially labeled by VADER, enhancing the robustness of our findings. VADER and InSet achieve an accuracy of 81.66% in sentiment labeling. However, the labeling of positive tweets can sometimes be inaccurate due to the use of negative words that actually convey the advantages of QRIS.

Topic modeling on QRIS using LDA, based on VADER's sentiment labeling, identifies six topics for positive tweets with a  $C_V$  value of 0.488 and  $C_{U\text{ Mass}}$  value of -3.460. These topics range from digital payments to QRIS usage in the business sector and transaction security. Conversely, negative tweets focus on three main topics with a  $C_V$  value of 0.383 and  $C_{U\text{ Mass}}$  value of -3.837, which include criticisms of transaction fees, technical issues, and data privacy concerns.

By analyzing these sentiments and topics, our research offers actionable insights that can guide improvements in QRIS implementation. Specifically, the findings can help stakeholders understand areas requiring attention, such as transaction fees and technical issues, while also confirming the positive impact QRIS has had on digital payments in Indonesia.

## APPENDIX

The datasets used for implementing the experiments described in this paper is available on GitHub at the following link: <https://github.com/dz4kiya/qris-vader-lda/>.

## REFERENCES

- [1] A. Pal, T. Herath, and H. R. Rao, "Why do people use mobile payment technologies and why would they continue? An examination and implications from India," *Res. Policy*, vol. 50, no. 6, p. 104228, 2021.
- [2] O. Balch, "Digital payments: how COVID-19 sped up adoption." Accessed: Jan. 14, 2023. [Online]. Available: <https://www.raconteur.net/finance/payments/digital-payments-covid-19/>
- [3] O. B. Saputri, "Preferensi konsumen dalam menggunakan quick response code indonesia standard (qris) sebagai alat pembayaran digital," *Kinerja*, vol. 17, no. 2, pp. 237–247, 2020.
- [4] Tim TvOne, "QRIS Bisa Digunakan di Thailand, Transaksi jadi Lebih Mudah." Accessed: Jan. 14, 2023. [Online]. Available: <https://www.tvonenews.com/ekonomi/63678-qris-bisa-digunakan-di-thailand-transaksi-jadi-lebih-mudah>
- [5] Statista, "Leading countries based on number of X (formerly Twitter) users as of January 2023." Accessed: Jan. 15, 2023. [Online]. Available: <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>
- [6] M. A. Rizaty, "Pengguna Twitter di Indonesia Capai 18,45 Juta pada 2022." Accessed: Jan. 15, 2023. [Online]. Available: <https://dataindonesia.id/digital/detail/pengguna-twitter-di-indonesia-capai-1845-juta-pada-2022>
- [7] L. Bing, "Sentiment analysis and opinion mining (synthesis lectures on human language technologies)," *Univ. Illinois Chicago, IL, USA*, 2012.
- [8] J. Albrecht, S. Ramachandran, and C. Winkler, *Blueprints for Text Analytics Using Python*. "O'Reilly Media, Inc.," 2020.
- [9] V. D. Chaithra, "Hybrid approach: naive bayes and sentiment VADER for analyzing sentiment of mobile unboxing video comments," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 5, pp. 4452–4459, 2019.
- [10] A. P. Ekaputri and S. Akbar, "Financial News Sentiment Analysis using Modified VADER for Stock Price Prediction," in *2022 9th International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*, IEEE, 2022, pp. 1–6.
- [11] B. Mathayomchan, V. Taecharungroj, and W. Wattanacharoensil, "Evolution of COVID-19 tweets about Southeast Asian Countries: Topic modelling and sentiment analyses," *Place Brand. Public Dipl.*, vol. 19, no. 3, pp. 317–334, 2023.
- [12] B. S. Rintyarna *et al.*, "Mapping acceptance of Indonesian organic food consumption under Covid-19 pandemic using Sentiment Analysis of Twitter dataset," *J. Theor. Appl. Inf. Technol.*, vol. 99, no. 5, pp. 1009–1019, 2021.
- [13] S. Shurrah, Y. Shannak, A. Almsnhanah, H. Khazaleh, and H. Najadat, "Attitudes evaluation toward covid-19 pandemic: An application of twitter sentiment analysis and latent dirichlet allocation," in *2021 12th International Conference on Information and Communication Systems (ICICS)*, IEEE, 2021, pp. 265–272.
- [14] I. Wedel, M. Palk, and S. Voß, "A bilingual comparison of sentiment and topics for a product event on Twitter," *Inf. Syst. Front.*, pp. 1–12, 2021.
- [15] H. Yin, X. Song, S. Yang, and J. Li, "Sentiment analysis and topic modeling for COVID-19 vaccine discussions," *World Wide Web*, vol. 25, no. 3, pp. 1067–1083, 2022.
- [16] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the international AAAI conference on web and social media*, 2014, pp. 216–225.
- [17] T. Ali, B. Omar, and K. Soulaïmane, "Analyzing tourism reviews using an LDA topic-based sentiment analysis approach," *MethodsX*, vol. 9, p. 101894, 2022.
- [18] I. Vayansky and S. A. P. Kumar, "A review of topic modeling methods," *Inf. Syst.*, vol. 94, p. 101582, 2020.
- [19] K. M. Ridhwan and C. A. Hargreaves, "Leveraging Twitter data to understand public sentiment for the COVID-19 outbreak in Singapore," *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 2, p. 100021, 2021.
- [20] Y. Yiran and S. Srivastava, "Aspect-based Sentiment Analysis on mobile phone reviews with LDA," in *Proceedings of the 2019 4th International Conference on Machine Learning Technologies*, 2019, pp. 101–105.
- [21] JustAnotherArchivist, "snsrape." Accessed: Oct. 08, 2023. [Online]. Available: <https://github.com/JustAnotherArchivist/snsrape>
- [22] H. M. Keerthi Kumar and B. S. Harish, "Classification of short text using various preprocessing techniques: An empirical evaluation," in *Recent Findings in Intelligent Computing Techniques: Proceedings of the 5th ICACNI 2017, Volume 3*, Springer, 2018, pp. 19–30.
- [23] W. Etaïwi and G. Naymat, "The impact of applying different preprocessing steps on review spam detection," *Procedia Comput. Sci.*, vol. 113, pp. 273–279, 2017.
- [24] N. A. Salsabila, Y. A. Winatmoko, A. A. Septiandri, and A. Jamal, "Colloquial indonesian lexicon," in *2018 International Conference on Asian Language Processing (IALP)*, IEEE, 2018, pp. 226–229.
- [25] Y. Fauziah, B. Yuwono, and A. S. Aribowo, "Lexicon Based Sentiment Analysis in Indonesia Languages: A Systematic".
- [26] S. Sbalchiero and M. Eder, "Topic modeling, long texts and the best number of topics. Some Problems and solutions," *Qual. Quant.*, vol. 54, pp. 1095–1108, 2020.



- [27] N. Jha and A. Mahmoud, "Mining non-functional requirements from app store reviews," *Empir. Softw. Eng.*, vol. 24, pp. 3659–3695, 2019.
- [28] Sastrawi, "Sastrawi." Accessed: Jan. 17, 2023. [Online]. Available: <https://github.com/sastrawi/sastrawi>
- [29] C. J. Hutto, "VADER-Sentiment-Analysis." Accessed: Oct. 08, 2023. [Online]. Available: <https://github.com/cjhutto/vaderSentiment>
- [30] F. Koto and G. Y. Rahmangtyas, "Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs," in *2017 International Conference on Asian Language Processing (IALP)*, IEEE, 2017, pp. 391–394.
- [31] S. Anastasia and I. Budi, "Twitter sentiment analysis of online transportation service providers," in *2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, IEEE, 2016, pp. 359–365.
- [32] D. S. Putri and T. Ridwan, "Analisis Sentimen Ulasan Aplikasi Pospay dengan Algoritma Support Vector Machine," *J. Ilm. Inform.*, vol. 11, no. 01, pp. 32–40, 2023.
- [33] D. Mardiansyah, "Pengguna Aplikasi Dompot Digital Capai 87%." Accessed: Mar. 06, 2023. [Online]. Available: <https://keuangan.kontan.co.id/news/pengguna-aplikasi-dompot-digital-capai-87>
- [34] A. S. Singh and M. B. Masuku, "Sampling techniques & determination of sample size in applied statistics research: An overview," *Int. J. Econ. Commer. Manag.*, vol. 2, no. 11, pp. 1–22, 2014.
- [35] U. Ependi, S. Aliya, and A. Wibowo, "Sentiment Analysis of Covid-19 Handling in Indonesia Based on Lexicon Weighting," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 12, no. 1, pp. 76–82, 2023.
- [36] A. P. Yulivia, "Sentiment Analysis Layanan Pesan Antar Makanan pada Twitter Menggunakan Inset Lexicon (Studi Kasus: GoFood dan GrabFood)," 2023.
- [37] R. A. Nasution, "Analisis Persepsi Pedagang Pada Penggunaan Qris Sebagai Alat Transaksi Umkm Di Kota Medan." Universitas Islam Negeri Sumatera Utara Medan, 2020.
- [38] G. J. Tobing, L. Abubakar, and T. Handayani, "Analisis peraturan penggunaan QRIS sebagai kanal pembayaran pada praktik UMKM dalam rangka mendorong perkembangan ekonomi digital," *Acta Com. J. Huk. Kenotariatan*, vol. 6, no. 03, pp. 491–509, 2021.
- [39] J. Liu, R. J. Kauffman, and D. Ma, "Competition, cooperation, and regulation: Understanding the evolution of the mobile payments technology ecosystem," *Electron. Commer. Res. Appl.*, vol. 14, no. 5, pp. 372–391, 2015.
- [40] Bank Indonesia, "QR Code Indonesian Standard (QRIS)." Accessed: Feb. 02, 2023. [Online]. Available: <https://www.bi.go.id/QRIS/default.aspx>
- [41] M. A. Prayitno and W. Fadly, "Pelatihan Pemanfaatan dan Pendampingan Pembuatan QRIS (QR Code Indonesian Standard) Sebagai Media Digitalisasi ZIS di Desa Glinggang Kabupaten Ponorogo," *Bubungan Tinggi J. Pengabd. Masy.*, vol. 4, no. 2, p. 543, 2022.
- [42] R. A. Hutagalung, P. Nainggolan, and P. D. Panjaitan, "Analisis Perbandingan Keberhasilan UMKM Sebelum Dan Saat Menggunakan Quick Response Indonesia Standard (QRIS) Di Kota Pematangsiantar," *J. Ekuilnoma*, vol. 3, no. 2, pp. 94–103, 2021.
- [43] S. A. Natalina, A. Zunaidi, and R. Rahmah, "Quick Response Code Indonesia Standard (Qris) Sebagai Strategi Survive Usaha Mikro Kecil Dan Menengah (UMKM) Di Masa Pandemi Di Kota Kediri," *Istithmar J. Stud. Ekon. Syariah*, vol. 5, no. 2, 2021.
- [44] Bank Indonesia, "Statistik Sistem Pembayaran dan Infrastruktur Pasar Keuangan (SPIP)." Accessed: Mar. 31, 2023. [Online]. Available: <https://www.bi.go.id/id/statistik/ekonomi-keuangan/spip/Default.aspx>
- [45] I. S. Igboanusi, K. P. Dirgantoro, J.-M. Lee, and D.-S. Kim, "Blockchain side implementation of pure wallet (pw): An offline transaction architecture," *ICT Express*, vol. 7, no. 3, pp. 327–334, 2021.
- [46] PT Aplikasi Karya Anak Bangsa, "Perbedaan QRIS Statis dan Dinamis yang Kelihatan Sama." Accessed: May 31, 2023. [Online]. Available: <https://gobiz.co.id/pusat-pengetahuan/perbedaan-qr-is-statis-dan-dinamis/>
- [47] W. Noviansah, "Duit dari QRIS Palsu Masuk Rekening Iman Mahlil, Bukan 'Restorasi Masjid.'" Accessed: Apr. 22, 2023. [Online]. Available: <https://news.detik.com/berita/d-6668548/duit-dari-qr-is-palsu-masuk-rekening-iman-mahlil-bukan-restorasi-masjid>
- [48] M. R. Sandi, "Permudah Transaksi, Seluruh Pasar Tradisional di Depok Sudah Terfasilitasi QRIS." Accessed: Apr. 30, 2023. [Online]. Available: <https://www.idxchannel.com/economics/permudah-transaksi-seluruh-pasar-tradisional-di-depok-sudah-terfasilitasi-qr-is>