Text classification on user feedback: a systematic literatures review

Zuleaizal Sidek[1,2], Sharifah Sakinah Syed Ahmad[1], Yogan Jaya Kumar[1], Noor Hasimah Ibrahim Teo[3]
[1]Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia
[2]Institut Tun Perak, Melaka, Malaysia
[3]School of Computing, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA (Melaka Branch), Melaka, Malaysia

Article Info

Article history:

Received Feb 6, 2024
Revised Feb 19, 2024
Accepted Feb 22, 2024

Keywords:

Customer feedback
Customer reviews
Literature review
Text classification
Textual data

ABSTRACT

User feedback in text classification serves multiple purposes, including refining models, enhancing datasets, adapting to user preferences, identifying emerging topics, and evaluating system performance, all of which contribute to the creation of more effective and user-centric classification systems. Many text classification techniques, including data mining, machine learning, and deep learning approaches, have been employed in previous literature, each making significant contributions to the field. This paper aims to contribute by guiding researchers seeking commonly used classification techniques and evaluation metrics in text processing. Additionally, it identifies the classification technique that generates higher accuracy and works as a basis for researchers to synthesize studies within their respective fields. Preferred reporting items for systematic reviews and meta-analysis (PRISMA) methodology is adapted to systematically review 28 current literatures on text classification on user feedback. The results obtained are guided by four research questions; paper distribution year, dataset source and size; evaluation metric and model accuracy. The review has shown that support vector machines (SVM) are frequently employed and consistently achieve high levels of accuracy as high as 97.17% with various datasets used. The future direction of this work could explore models that integrate sentiment analysis and natural language understanding to more accurately capture nuanced user opinions and preferences.

Corresponding Author:

Zuleaizal Sidek
Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM)
76100, Durian Tunggal, Melaka, Malaysia
Email: p032220004@student.utem.edu.my

1. INTRODUCTION

The number of people actively participating virtually on the internet, particularly social media, is constantly rising, especially during the pandemic [1]. This has resulted in the rise of e-commerce, which refers to the provision of services via the internet and allows for instant communication and trading, thus cutting down on the amount of time needed to make a purchase or send an item. The internet has a vast quantity of text-based content, posing significant challenges for manual classification [2]. The complexity of handling various types of data, such as emails, patient reports derived from medical records, academic papers with specific technological needs, and news articles categorized by different topics, has increased significantly due to the large volume and numerous obstacles associated with identifying them as either malicious or safe. There exists a need for an automated technique that can effectively classify textual content in a manner that is both straightforward and compatible with machines [3].

Customer feedback texts play a crucial role in facilitating a direct channel for consumers to articulate their thoughts, sentiments, and experiences [4]. It has been proposed that users of social media platforms are not only engaged in product reviews, but also play a variety of other roles, including those of online activists, trolls, social critics, information seekers, and socialites. As a result, doing an appropriate analysis of their feelings is of the utmost importance [5]. This feedback encompasses a wide range of opinions, varying from highly positive recommendations to harsh criticisms, and serves as a valuable resource for businesses, offering a wealth of knowledge. Through the process of detecting reoccurring issues and client preferences, organizations can have the opportunity to modify their products, thus improving their overall attractiveness and functionality. The customer feedback may consist of complaints, reports, or inquiries about the services rendered by the company. Complaints should not be regarded as an inconvenience that needs to be evaded, dismissed, or concealed, as they have the potential to produce novel ideas for assessing and enhancing both products and services [6]. If a client’s complaint isn’t resolved to their satisfaction, they may decide to stop being a customer immediately. Therefore, businesses should motivate clients to provide feedback on services, particularly in the event of a negative experience.

Customer reviews often utilize a five-point rating scale to gauge customer satisfaction, along with textual input for additional comments or feedback. The additional remarks consist of text messages that possess significant utility, demanding the implementation of an intelligent analysis mechanism. This mechanism should be capable of analyzing these messages through the utilization of several techniques, including natural language processing, ontology [7], and classification approaches [8]. Putting the importance of text classification based on consumer feedback into perspective requires thinking about the massive amount of data created on a worldwide scale. A recent study estimates yearly global data production at an astonishing 64.2 zettabytes [9]. A substantial portion of this data is comprised of textual customer feedback, spanning multiple languages, cultures, and industries. This exponential growth in data volume underscores the urgent need for sophisticated techniques to extract actionable insights.

The problem of text classification refers to the task of assigning text documents to predetermined classes or groups depending on the content of the processing texts, utilizing data mining tools and approaches. Text classification applications have proven to be highly effective across several domains, serving diverse objectives [10], [11]. In the context of text classification, numerous researchers have contributed significant improvements in their studies. The subject matter encompasses the classification of textual datasets across domains using machine learning (ML) and deep learning (DL), as well as the utilization of statistical techniques.

Improve retrieval performance and speed up the process of extracting useful information from vast amounts of text data with the help of text classification technology, which is able to automate and simplify complex text information [12]. There are two main categories of textual representation: factual and opinion-based. Facts can be interpreted as objective representations of items, events, and their attributes, whereas opinions are subjective statements that reflect individuals’ sentiments towards these entities or occurrences [13]. Concerning understanding consumer feedback or reviews, sentiment analysis is becoming a viable technique for businesses [14]. In particular, customer reviews that include people’s attitudes toward products are being analysed with sentiment analysis algorithms [15]. Reviews, comments, surveys, online resources, and social media can all be leveraged for better customer relationship management through sentiment analysis.

The identification of intent plays a vital role in dialogue systems since it enables the system to comprehend the user’s desired actions [16]. Therefore, this paper presents a study on text classification methods used for customer feedback texts, focusing on the methods and techniques used in the classification process. This paper will address the following research questions:

RQ1: what is the distribution year, author, domain application, classification types related to the text classification and also sources and sizes of the dataset used?

RQ2: what are the sources and sizes of datasets used in the reviewed studies?

RQ3: what are the most common performance evaluation metrics used?

RQ4: what are the techniques/approaches gives higher accuracy?

The subsequent sections of this paper are structured as follows. Section 2 presents the methodology employed in conducting a systematic review. Section 3 provides an analysis of the relevant literature, findings and discussions. The investigation is ultimately ended in section 4 with the conclusion and future work.

2. METHOD

The approach of conducting a systematic literature review (SLR) has been chosen as the research method for this research. The SLR is a protocol-based method that was developed in response to the research questions to locate, select, and evaluate the literature that is the most pertinent [17]. There are a few key principles such as coverage and focus were taken into consideration when deciding whether to use SLR as a method of review. The SLR procedure, which is suggested in the literature [18], [19], is followed by this study to give a more objective and comprehensive literature review. As can be seen in Figure 1, the SLR process is broken up into three key phases: planning, performing the review, and reporting the results of the review.

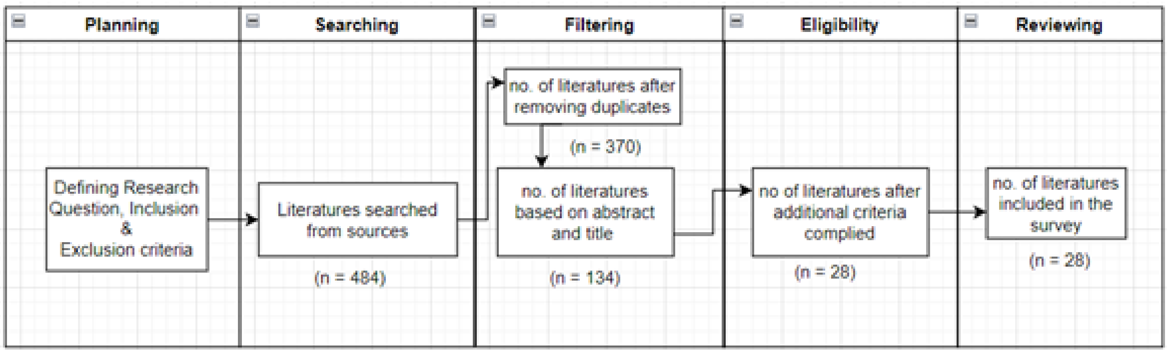

Figure 1. Strategic literature reviews process

This research employed the preferred reporting items for systematic reviews and meta-analysis (PRISMA) technique to perform a systematic review (SR) for this work. The PRISMA is a standardized collection of essential elements that serve as a foundation for the construction and organization of SRs and other forms of meta-analyses. Therefore, the reviewing process is comprised of five stages, namely planning, searching, filtering, eligibility and reviewing. It is grounded in empirical data and provides a framework for ensuring the quality and comprehensiveness of these research efforts. The details of these procedures, as they pertain to our research investigation, are going to be covered in the parts that follow.

2.1. Planning phase

2.1.1. Defining of research question

This SR is an organized approach that includes the extent of the research that was evaluated by classifying and analyzing the various publications that are already connected to this topic. Determining the research questions that will be asked to correctly describe the percentage of previously published works is the first step. When we read works that are related to one another, we can gain many insights that can later assist scholars in developing new concepts. Table 1 contains a description of the research questions that were utilized in our SR.

Table 1. Research question description

Research question	Description
RQ1: what is the distribution year, author, domain application, classification types related to the text classification and also sources and sizes of the dataset used?	The response to RQ1 facilitates the identification of the time, domain of application, classification type, and authors of the conducted research studies.
RQ2: what are the sources and sizes of datasets used in the reviewed studies?	The response to RQ2 enables researchers to gain an understanding of the biases or limitations inherent in the datasets.
RQ3: what are the most common performance evaluation metrics used?	The response to RQ3 outlines the metrics employed for assessing the efficacy of the developed techniques as the key findings of the reviewed studies.
RQ4: what are the techniques/approaches gives higher accuracy?	The response to RQ4 is to identify suitable tools and techniques employed by the latest applications.

This research employed the PRISMA technique to perform a SR for this work. The PRISMA is a standardized collection of essential elements that serve as a foundation for the construction and organization of SRs and other forms of meta-analyses. Therefore, the reviewing process is comprised of five stages, namely planning, searching, filtering, eligibility and reviewing. It is grounded in empirical data and provides a framework for ensuring the quality and comprehensiveness of these research efforts.

2.1.2. Inclusion and exclusion criteria

We employ a set of inclusion criteria (IC) and exclusion criteria (EC) to identify whether papers are relevant to our search in Table 2. This allows us to refine the results of our search. Studies that do not address EC are disregarded, and an additional filtering technique is utilized to identify articles that meet our setting.

Table 2. Research question description

Inclusion criteria	Exclusion criteria
Literature is published within the timeframe from 2019 to 2023.	Duplicated literatures
Literature must satisfy at least one of the specified search terms	Research papers that are not written in English
Literature should be published in reputable academic journals, conferences, or academic magazines.	Literature that is not relevant to the topic.
The search is conducted leveraging the title, abstract, and complete text.	

The preliminary evaluation is comprised of the following three IC stages:

i) The following phase relies on the abstracts and entails eliminating important findings based on the information and keywords identified in the abstracts. Only the abstracts of articles approved by the IC were kept for further processing.
ii) Step based on the full text: we reject the results that neglect to address or describe the research terms in Table 3, such as publications that only represent minor elements of the search terms contained in their abstracts. This is the step that we take after the step based on the full text.
iii) Phase based on quality analysis: we assess the quality of the remaining results, deleting any of them that do not meet any of the following criteria:

C1: the publications present a thorough method for categorizing customer feedback.
C2: the publication details the technical aspects of how the suggested solution would be implemented.
C3: the publication cites several other studies.

2.2. Searching phase

There was a total of 484 papers found during the first phase of the search that had search phrases included in the title and had the potential to be the most relevant. The search queries that were utilized for this study are presented in Table 3. We considered the research presented at conferences, in journals, as part of theses, and in reports that were published between 2019 and 2023.

Table 3. Searching term

Search ID	Search term
S1	Text classification AND customer feedback
S2	Text classification AND customer reviews
S3	Intent classification AND customer reviews
S4	Sentiment analysis AND customer reviews
S5	Sentiment analysis AND customer feedback

The proposed approach comprises two main stages: i) analyzing search phrases from previous study questions to create a list of keywords, and ii) formulating queries using Boolean operators AND/OR to find and gather all relevant results. The first point generating a list of keywords by extracting search terms from previous research queries. The initial phase in conducting our SR was to identify the various data sets that would be used. As can be seen in Table 4, searches have been conducted using a variety of academic databases, digital libraries, and search engines, including academic and open-access options. The next step is to define the techniques that will be used to examine the academic and technical papers that these searches yielded to locate publications that are pertinent to our context.

Table 4. Literature database

Database	URL
Scopus digital library	www.scopus.com
IEEE Xplore digital library	https://ieeexplore.ieee.org/Xplore/home.jsp
ScienceDirect digital library	https://www.sciencedirect.com/
ResearchGate	https://researchgate.com

2.2.1. Filtering

Relevant literature was selected for this review by following the planning, searching, and filtering steps outlined in the PRISMA flow diagram. An illustration of the systematic search approach that will be utilized can be found in Figure 1. During the preliminary stage, we were successful in acquiring 484 records. We were left with 370 original results after we eliminated all the duplicate literature. Following that, criteria for filtering based on the title and abstract were applied to these 370, and therefore, to 134 after this round of elimination.

2.2.2. Eligibility

By including further criteria for eligibility outlined in the complete article, we were ultimately able to obtain a total of 28 relevant research studies. A thorough evaluation was conducted on a collection of 28 research publications to extract the findings that will be elaborated upon in the subsequent section. Out of the 28 publications that through the review process, 16 of them employed machine-learning approaches for text classification. Among these, 9 papers metrics DL methods, while 3 studies utilized statistical analysis. The remaining publications employed a combination of fuzzy logic and the Ernie approach. Out of a total of 28 published literature, many of the studies employ ML approaches for text classification.

3. RESULTS AND DISCUSSION

This section discusses the most important findings that emerged from our comprehensive literature reviews were based on primary research papers published between 2019 and 2023. This part focuses on the findings of the review related to the research questions that were previously established. It has a total of 28 papers that were systematically selected and relevant to the domain of text classification in the context of customer feedback and reviews. These discussions assist researchers in gaining information concerning recent literature on the topic. It provides information on the methods employed, the techniques used for evaluating text classification, and other specific tools employed in the research process.

RQ1: what is the distribution year, author, domain application, and classification types related to the text classification? The literature distribution result is shown in Figure 2. Based on the reviews, most of the literature was published in the year 2022, with a total count of 14 publications. In the year 2023, there are a total of eight publications. From 2019 to 2021, there were a total of six literatures, with one literature in 2019, three in 2020, and two in 2021 as depicted in Figure 2(a). It can be observed that in the year 2022, a significant number of academic researchers are engaged in the task of text classification, specifically focusing on customer feedback and reviews. Out of the 28 selected literature, more than half of them (16 sources) are utilizing ML as the primary methodology for researching the text classification of customer feedback and review. The remaining publications utilize DL techniques, as well as a combination of ML and statistical analysis methodologies as in Figure 2(b).

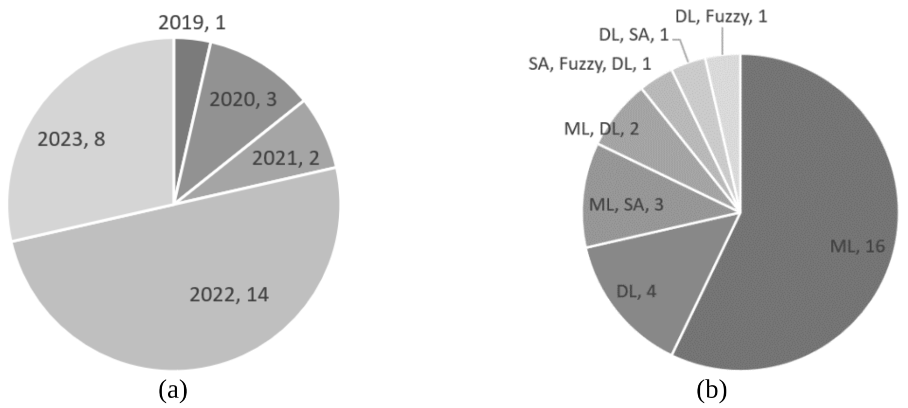

Figure 2. Literature distribution based on (a) year of publication and (b) classification type

RQ2: what are the sources and sizes of datasets used in the reviewed studies? The datasets were obtained from social media platforms such as Twitter, and other online platforms, as indicated in Table 5. Most of the literature utilizes datasets consisting of Amazon customer reviews and tweets from the social media platform Twitter. Three different studies have utilized TripAdvisor reviews, whereas two studies used ATIS as datasets for their research. The rest of the literature uses several textual datasets for their analyses.

Table 5. List of literature and dataset sources

Literature	Sources
[20]–[24]	Amazon
[20], [21], [25]–[27]	Twitter
[28], [29]	TripAdvisor
[30], [31]	ATIS
[22], [32]–[43]	Others

Due to the large number of individuals who use it and the diversity of intriguing topics on which they express their thoughts, Twitter has become the most popular social media network. In addition, scraping the public live tweets from Twitter is made simpler when the Twitter API is used. Since it is the largest online purchasing platform in the world, Amazon.com has become an increasingly popular source of customer feedback for researchers conducting sentiment analysis. There are several of the listed literature, such as [20]–[22], that believe that the quantity of datasets influences the accuracy of classification. When common approaches that involve ML, one crucial consideration is the size of the dataset. According to what is found in the research literature, the quality of model learning improves and the accuracy of the generalization phase increases as the dataset size is increased. Because the dimensions of datasets may vary from one research project to the next, we decided to include dataset size as one of the parameters in our systematic reviews. The dataset sizes are categorized into five intervals, commencing with datasets containing less than 5,000 reviews or cases, and completing with datasets including over 100,000 comments.

According to the data depicted in Table 6, it is evident that the dataset group with a size of less than 20,000 reviews or comments exhibits the majority of literature which is 42.86%. This observation implies that a dataset of this magnitude is enough to achieve effective training and generalization outcomes. Approximately 14% of the literature surveyed used datasets above a size of 100,000. The observed outcome can be explained by the fact that certain machine-learning methods require a smaller dataset. Therefore, it is feasible to conduct a comprehensive examination of the performance of ML and DL algorithms based on the varying sizes of datasets.

RQ3: what are the most common performance evaluation metrics used? The purpose of this part of the findings is to provide the answer to the third question, which focuses on an analysis of the evaluation metric. This research question also aims to limit down potential methods by providing a comparison of the performance evaluation statistic. Validating the effectiveness of a text classification model can be done with the use of a wide variety of performance evaluation criteria. These metrics can also be used to determine how accurate a learning model such as machine and DL was trained.

Table 6. List of literature and dataset size

Literature	Dataset size
[23], [25], [29], [31], [44], [45]	<5,000
[20], [22], [26], [30], [43], [46]	5,000-20,000
[24], [36], [38], [47]	20,000-50,000
[21]	50,000-100,000
[33], [34], [35], [40]	>100,000
[27], [28], [32], [37], [39], [41], [42]	Not available

When assessing the efficacy of a ML model, accuracy is a crucial metric to take into consideration. The model has successfully learned the relationships among the input samples if the output value is very accurate, and it is then ready to classify future values. The most common metric for evaluating classification performance is accuracy. However, it is important to note that accuracy may not provide an accurate representation of performance when dealing with imbalanced datasets [48]. The F1 measure is recommended for evaluating skewed datasets [49]. If DL models are employed and a decline in precision is observed, an alternative measure such as macro-F1 can be utilized [50]. According to the data presented in Table 7, it is observed that a significant proportion of the literature reviewed uses accuracy as the primary performance criterion for evaluating their respective models. The implementation of machine and DL approaches is common in the field. It is noticeable that few literatures used Perception score, mapping quality characteristics, and exploring the relationship between internet reviews and hotel sales.

RQ4: what are the techniques/approaches gives higher accuracy? The evaluation of a text classification model's performance heavily relies on the aspect of accuracy, which is considered to be of utmost importance. A value that exhibits a high level of accuracy signifies that the model has effectively acquired a comprehensive understanding of the connections between the input samples.

Table 7. List of literature and its evaluation metric

Literature	Performance metric
[21], [25], [41], [47]	Accuracy
[28], [32]	Accuracy, precision, recall
[30], [44]	Accuracy, F1-score
[20], [23], [24], [26], [29], [31], [36], [37]	Accuracy, precision, recall, F1-score
[22], [27]	Precision, recall, F1-score
[38]	F1-score
[34]	Perception score
[46]	Mapping quality characteristic
[40]	Relationship (online reviews–hotel sales)
[43]	Classify user component
[33]	Confusion matrix

Consequently, it is deemed proficient at classifying forthcoming values. According to the data presented in Table 8, it can be observed that the random forest technique exhibits the highest level of accuracy among the ML algorithms, achieving a remarkable accuracy rate of 99.2% when applied to the TripAdvisor dataset. The following technique involved is the utilization of bidirectional encoder representations from transformers (BERT) and support vector machine (SVM) techniques, which achieved the highest accuracies of 98.8% and 97.17% respectively. It was observed that the SVM is the predominant ML model employed for text classification, as evidenced by its appearance in 9 research papers. The SVM achieved a peak accuracy of 97.17% when applied to the e-commerce dataset. The SVM is a powerful classifier capable of achieving high performance in various circumstances, particularly in multilabel problems. It exhibits the ability to handle both linear and nonlinear classification tasks [51]. The model's long-short term memory (LSTM) is also widely recognized and utilized in the field of text classification as it produces accuracy as high as 97%. The data presented in Table 8 also demonstrates that most of the ML algorithms achieved accuracy levels exceeding 90%.

Table 8. List of top ten accuracy score

Literature	Dataset	Technique	Accuracy
[29]	Tripadvisor	Random forest	99.20%
[30]	ATIS and Snips	BERT	98.8%
[36]	Mobile banking	SVM	97.17%
[26]	Twitter	LSTM	97%
[23]	E-commerce	LSTM	96.92%
[44]	E-commerce	SVM	96.57%
[33]	Web news	SVM	95.04%
[31]	ATIS	SVM	95%
[24]	Amazon	Random forest	94%
[47]	Amazon	CNN	92.30%

4. CONCLUSION

This paper provides a comprehensive review of the distribution of publications on text classification across several domains. This research presents the contributions by providing a critical analysis of the text classification methodologies used in recent studies, along with the evaluation metrics used to assess the model's performance. A systematic literature review was implemented using the PRISMA as methodology, which consists of a selection procedure, identification, screening, and eligibility. From the original pool of 484 retrieved materials, 28 literatures were selected after a thorough assessment based on four research questions. The academic publications chosen for this study's systematic reviews were published between 2019 and 2023 by four respectable publishers. Consists in detail: the type of datasets used, the size of datasets, the classification type, the evaluation metrics and the best performance in terms of accuracy. SVM is the main text classification method, delivering greater accuracy. For text classification, numerous datasets from multiple areas were examined. The most commonly used datasets in literature published in the last five years were Amazon customer reviews and tweets from Twitter. The popularity of these datasets is due to the ease of crawling and extracting these data. Based on our comprehensive review of the articles, it is evident that the adoption of machine learning and DL methodologies holds significant potential for advancing the field of text classification. By using these techniques, organizations may effectively streamline sentiment analysis processes, resulting in notable reductions in both cost and time spending. As for future work, the utilization of machine learning and DL techniques could improve efficacy and minimize dependency on human judgment in classifying consumer feedback and reviews.

ACKNOWLEDGEMENTS

The authors express their sincere appreciation to the reviewers for their valuable feedback, which helped enhance the quality of this article. The funding for this research was provided by 'Skim Zamalah UTeM' of Universiti Teknikal Malaysia Melaka (UTeM), Malaysia.

REFERENCES

[1] N. Krishnadoss and L. K. Ramasamy, "A study on high dimensional big data using predictive data analytics model," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 30, no. 1, pp. 174–182, Apr. 2023, doi: 10.11591/ijeecs.v30.i1.pp174-182.

[2] A. K. Al-Mashhadany, M. B. Mohammed, and M. M. Alrawi, "Inheritance issues' features extraction using Arabic text analyzer (IFAA)," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 28, no. 1, pp. 611–624, Oct. 2022, doi: 10.11591/ijeecs.v28.i1.pp611-624.

[3] R. A. Albeer, H. F. Al-Shahad, H. J. Aleqabie, and N. D. Al-Shakarchy, "Automatic summarization of YouTube video transcription text using term frequency-inverse document frequency," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 26, no. 3, pp. 1512–1519, Jun. 2022, doi: 10.11591/ijeecs.v26.i3.pp1512-1519.

[4] A. Ibrahim, D. R. Indah, and D. I. Meytri, "The implementation of social customer relationship management for tourism information system," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 24, no. 3, pp. 1578–1588, Dec. 2021, doi: 10.11591/ijeecs.v24.i3.pp1578-1588.

[5] M. Mkono and J. Tribe, "Beyond reviewing: uncovering the multiple roles of tourism social media users," *Journal of Travel Research*, vol. 56, no. 3, pp. 287–298, Aug. 2017, doi: 10.1177/0047287516636236.

[6] J. Joung, K. Jung, S. Ko, and K. Kim, "Customer complaints analysis using text mining and outcome-driven innovation method for market-oriented product development," *Sustainability (Switzerland)*, vol. 11, no. 1, p. 40, Dec. 2019, doi: 10.3390/su11010040.

[7] P. Zhou and N. El-Gohary, "Ontology-based multilabel text classification of construction regulatory documents," *Journal of Computing in Civil Engineering*, vol. 30, no. 4, Jul. 2016, doi: 10.1061/(asce)cp.1943-5487.0000530.

[8] V. Korde, "Text classification and classifiers:a survey," *International Journal of Artificial Intelligence & Applications*, vol. 3, no. 2, pp. 85–99, Mar. 2012, doi: 10.5121/ijaia.2012.3208.

[9] P. Taylor, "Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025," *Statista*, 2023. https://www.statista.com/statistics/871513/worldwide-data-created/ (accessed Sep. 16, 2023).

[10] V. Gupta and G. S. Lehal, "A survey of text mining techniques and applications," *Journal of Emerging Technologies in Web Intelligence*, vol. 1, no. 1, pp. 60–76, Aug. 2009, doi: 10.4304/jetwi.1.1.60-76.

[11] M. A. Tayal, V. Bajaj, A. Gore, P. Yadav, and V. Chouhan, "Automatic domain classification of text using machine learning," in *2023 International Conference on Communication, Circuits, and Systems, IC3S 2023*, May 2023, pp. 1–5, doi: 10.1109/IC3S57698.2023.10169470.

[12] R. Risnantoyo, A. Nugroho, and K. Mandara, "Sentiment analysis on corona virus pandemic using machine learning algorithm," *Journal of Informatics and Telecommunication Engineering*, vol. 4, no. 1, pp. 86–96, Jul. 2020, doi: 10.31289/jite.v4i1.3798.

[13] V. S. Anoop and S. Asharaf, "Aspect-oriented sentiment analysis: a topic modeling-powered approach," *Journal of Intelligent Systems*, vol. 29, no. 1, pp. 1166–1178, Dec. 2020, doi: 10.1515/jisys-2018-0299.

[14] L. M. Rojas-Barahona, "Deep learning for sentiment analysis," *Language and Linguistics Compass*, vol. 10, no. 12, pp. 701–719, Dec. 2016, doi: 10.1111/lnc3.12228.

[15] B. Liu, "Sentiment analysis and opinion mining," *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1–167, May 2012, doi: 10.2200/S00416ED1V01Y201204HLT016.

[16] A. T. Al-Tuama and D. A. Nasrawi, "Intent classification using machine learning algorithms and augmented data," in *2022 International Conference on Data Science and Intelligent Computing, ICDSIC 2022*, Nov. 2022, pp. 234–239, doi: 10.1109/ICDSIC56987.2022.10075794.

[17] H. Zhang, M. A. Babar, and P. Tell, "Identifying relevant studies in software engineering," *Information and Software Technology*, vol. 53, no. 6, pp. 625–637, Jun. 2011, doi: 10.1016/j.infsof.2010.12.010.

[18] H. Snyder, "Literature review as a research methodology: an overview and guidelines," *Journal of Business Research*, vol. 104, pp. 333–339, Nov. 2019, doi: 10.1016/j.jbusres.2019.07.039.

[19] O. Kaiwartya *et al.*, "Internet of vehicles: motivation, layered architecture, network model, challenges, and future aspects," *IEEE Access*, vol. 4, pp. 5356–5373, 2016, doi: 10.1109/ACCESS.2016.2603219.

[20] R. Akila and S. Revathi, "Opinion mining of social media reviews using fuzzy inference and intent prediction using deep learning," *Advances in Engineering Software*, vol. 174, p. 103293, Dec. 2022, doi: 10.1016/j.advengsoft.2022.103293.

[21] A. M. Alayba and V. Palade, "Leveraging Arabic sentiment classification using an enhanced CNN-LSTM approach and effective Arabic text preparation," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 10, pp. 9710–9722, Nov. 2022, doi: 10.1016/j.jksuci.2021.12.004.

[22] W. Huang, M. Lin, and Y. Wang, "Sentiment analysis of Chinese E-commerce product reviews using ERNIE word embedding and attention mechanism," *Applied Sciences (Switzerland)*, vol. 12, no. 14, p. 7182, Jul. 2022, doi: 10.3390/app12147182.

[23] A. Sharma and M. O. Shafiq, "A comprehensive artificial intelligence based user intention assessment model from online reviews and social media," *Applied Artificial Intelligence*, vol. 36, no. 1, Jan. 2022, doi: 10.1080/08839514.2021.2014193.

[24] Prathyakshini, K. P. Salian, B. R. Puneeth, T. Nargis, and S. Salian, "A novel approach for classification of online product reviews using various machine learning techniques," in *6th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2022 - Proceedings*, Dec. 2022, pp. 878–884, doi: 10.1109/ICECA55336.2022.10009447.

[25] Z. A. Diekson, M. R. B. Prakoso, M. S. Q. Putra, M. S. A. F. Syaputra, S. Achmad, and R. Sutoyo, "Sentiment analysis for customer review: case study of Traveloka," *Procedia Computer Science*, vol. 216, pp. 682–690, 2022, doi: 10.1016/j.procs.2022.12.184.

[26] W. Aljedaani *et al.*, "Sentiment analysis on Twitter data integrating TextBlob and deep learning models: the case of US airline industry," *Knowledge-Based Systems*, vol. 255, p. 109780, Nov. 2022, doi: 10.1016/j.knosys.2022.109780.

[27] M. Tummawat and C. Jareanpon, "The sentiment classification of hotel reviews and hotel description using feature-based technique for customer relationship management," *ICIC Express Letters*, vol. 17, no. 3, pp. 279–287, 2023, doi: 10.24507/icicel.17.03.279.

[28] L. Gunawan, M. S. Anggreainy, L. Wihan, Santy, G. Y. Lesmana, and S. Yusuf, "Support vector machine based emotional analysis of restaurant reviews," *Procedia Computer Science*, vol. 216, pp. 479–484, 2022, doi: 10.1016/j.procs.2022.12.160.

[29] C. Huda, Y. Heryadi, Lukas, and W. Budiharto, "Aspect-based sentiment analysis in tourism industry for tourism recommender system," in *2022 5th International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2022*, Dec. 2022, pp. 407–412, doi: 10.1109/ISRITI56927.2022.10053083.
[30] Y. Chen and Z. Luo, "Pre-trained joint model for intent classification and slot filling with semantic feature fusion," *Sensors*, vol. 23, no. 5, p. 2848, Mar. 2023, doi: 10.3390/s23052848.
[31] M. Menda and G. S. Keerthi, "Intent classification in conversational system using machine learning techniques," *International Journal of Computer Applications*, vol. 183, no. 51, pp. 6–11, Feb. 2022, doi: 10.5120/ijca2022921913.
[32] A. Ghazzawi and B. Alharbi, "Analysis of customer complaints data using data mining techniques," *Procedia Computer Science*, vol. 163, pp. 62–69, 2019, doi: 10.1016/j.procs.2019.12.087.
[33] A. Mulahuwaish, K. Gyorick, K. Z. Ghafoor, H. S. Maghdid, and D. B. Rawat, "Efficient classification model of web news documents using machine learning algorithms for accurate information," *Computers and Security*, vol. 98, p. 102006, Nov. 2020, doi: 10.1016/j.cose.2020.102006.
[34] S. Dhar and I. Bose, "Walking on air or hopping mad? Understanding the impact of emotions, sentiments and reactions on ratings in online customer reviews of mobile apps," *Decision Support Systems*, vol. 162, p. 113769, Nov. 2022, doi: 10.1016/j.dss.2022.113769.
[35] S. Park, J. Cho, K. Park, and H. Shin, "Customer sentiment analysis with more sensibility," *Engineering Applications of Artificial Intelligence*, vol. 104, p. 104356, Sep. 2021, doi: 10.1016/j.engappai.2021.104356.
[36] N. A. Rahman, S. D. Idrus, and N. L. Adam, "Classification of customer feedbacks using sentiment analysis towards mobile banking applications," *IAES International Journal of Artificial Intelligence*, vol. 11, no. 4, pp. 1579–1587, Dec. 2022, doi: 10.11591/ijai.v11.i4.pp1579-1587.
[37] A. Londhe and P. V. R. D. P. Rao, "Dynamic classification of sentiments from restaurant reviews using novel fuzzy-encoded LSTM," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 10, no. 9, pp. 112–124, Sep. 2022, doi: 10.17762/ijritcc.v10i9.5714.
[38] Á. Aldunate, S. Maldonado, C. Vairetti, and G. Armelini, "Understanding customer satisfaction via deep learning and natural language processing," *Expert Systems with Applications*, vol. 209, p. 118309, Dec. 2022, doi: 10.1016/j.eswa.2022.118309.
[39] A. Patel, P. Oza, and S. Agrawal, "Sentiment analysis of customer feedback and reviews for airline services using language representation model," *Procedia Computer Science*, vol. 218, pp. 2459–2467, 2022, doi: 10.1016/j.procs.2023.01.221.
[40] Y. Sim, S. K. Lee, and I. Sutherland, "The impact of latent topic valence of online reviews on purchase intention for the accommodation industry," *Tourism Management Perspectives*, vol. 40, p. 100903, Oct. 2021, doi: 10.1016/j.tmp.2021.100903.
[41] V. Desai, S. Wadhwa, A. A, and B. Bajaj, "Text-based intent analysis using deep learning," *International Journal of Innovative Science and Research Technology*, vol. 5, no. 7, pp. 267–274, Jul. 2020, doi: 10.38124/ijisrt20jul342.
[42] V. R. Azhaguramyaa, J. Janet, G. R. Madhavan, S. Balakrishnan, and K. Arunkumar, "Sentiment analysis on book reviews using machine learning techniques," in *8th International Conference on Advanced Computing and Communication Systems, ICACCS 2022*, Mar. 2022, pp. 1530–1534, doi: 10.1109/ICACCS54159.2022.9785311.
[43] Y. Son and W. Kim, "Development of methodology for classification of user experience (UX) in online customer review," *Journal of Retailing and Consumer Services*, vol. 71, p. 103210, Mar. 2023, doi: 10.1016/j.jretconser.2022.103210.
[44] P. Vijayaragavan, R. Ponnusamy, and M. Aramudhan, "An optimal support vector machine based classification model for sentimental analysis of online product reviews," *Future Generation Computer Systems*, vol. 111, pp. 234–240, Oct. 2020, doi: 10.1016/j.future.2020.04.046.
[45] P. Yu, X. Zhou, Z. Li, and Y. Yan, "Inactivation and change of tetracycline-resistant Escherichia coli in secondary effluent by visible light-driven photocatalytic process using Ag/AgBr/g-C3N4," *Science of The Total Environment*, vol. 705, p. 135639, Feb. 2020, doi: 10.1016/j.scitotenv.2019.135639.
[46] H. Liu, S. Wu, C. Zhong, and Y. Liu, "The effects of customer online reviews on sales performance: the role of mobile phone's quality characteristics," *Electronic Commerce Research and Applications*, vol. 57, 2023, doi: 10.1016/j.elerap.2022.101229.
[47] L. Arora, P. Srivastava, and S. A. Kumar, "Evaluation of product reviews using deep learning classifier models," in *2022 13th International Conference on Computing Communication and Networking Technologies, ICCCNT 2022*, Oct. 2022, pp. 1–5, doi: 10.1109/ICCCNT54827.2022.9984463.
[48] B. Espejo-Garcia, J. Martinez-Guanter, M. Pérez-Ruiz, F. J. Lopez-Pellicer, and F. J. Zarazaga-Soria, "Machine learning for automatic rule classification of agricultural regulations: a case study in Spain," *Computers and Electronics in Agriculture*, vol. 150, pp. 343–352, Jul. 2018, doi: 10.1016/j.compag.2018.05.007.
[49] A. Ligthart, C. Catal, and B. Tekinerdogan, "Analyzing the effectiveness of semi-supervised learning approaches for opinion spam classification," *Applied Soft Computing*, vol. 101, p. 107023, Mar. 2021, doi: 10.1016/j.asoc.2020.107023.
[50] D. Song, A. Vold, K. Madan, and F. Schilder, "Multi-label legal document classification: A deep learning-based approach with label-attention and domain-specific pre-training," *Information Systems*, vol. 106, p. 101718, May 2022, doi: 10.1016/j.is.2021.101718.
[51] N. A. P. Rostam and N. H. A. H. Malim, "Text categorisation in Quran and Hadith: overcoming the interrelation challenges using machine learning and term weighting," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 6, pp. 658–667, Jul. 2021, doi: 10.1016/j.jksuci.2019.03.007.

BIOGRAPHIES OF AUTHORS

Zuleaizal Sidek is the founder and managing director of Kencana Niaga, an IT start-up and independent provider of data analytics for interdisciplinary analysis. He is also a postgraduate student in Philosophy Doctor in Data Science at the Universiti Teknikal Malaysia (UTeM). He is currently working as Data Scientist at the Institut Tun Perak (Melaka state owned company). He has multiple experiences in IT for more than 10 years at the Universiti Teknologi MARA as an IT Manager. His first paper on e-government services: a value of E-service in local government: a fuzzy approach evaluation". His areas of interest are big data management and analytics, machine learning and blockchain technology. He can be contacted at email: zulsidek@gmail.com.

Assoc. Prof. Dr. Sharifah Sakinah Syed Ahmad is currently an associate professor in the Department of Intelligent Computing and Analytics (ICA), Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). She received her bachelor's and master's degrees in applied mathematics from the School of Mathematics at the University of Science, Malaysia. Following this, she received her Ph.D. from the University of Alberta, Canada in 2012 in intelligent systems. She can be contacted at email: sakinah@utem.edu.my.

Dr. Yogan Jaya Kumar is a Senior Lecturer at the Department of Intelligent Computing and Analytic in the Faculty of Information and Communication Technology (FTMK), Universiti Teknikal Malaysia Melaka (UTeM). He earned both his bachelor degree and master degree from Universiti Sains Malaysia (USM), in the field of Mathematical Science in year 2003 and 2005. He joined the Centre for Affiliate and Diploma Programme, MMU as assistant lecturer in 2006. He was then offered a lecturer position at UTeM in 2008. He went on to complete his Ph.D. studies a Universiti Teknologi Malaysia, in 2014 in the field of Computer Science under the supervision of Professor Dr. Naomie Salim. Currently, his research involves in the field of text mining, information extraction and AI applications. He has published a number of journal articles in applied soft computing, mainly in the field of text document summarization. He can be contacted at email: yogan@utem.edu.my.

Dr. Noor Hasimah Ibrahim Teo is a senior lecturer at the Universiti Teknologi Teknologi MARA Melaka, Malaysia. She received her Ph.D. in Computer Science from the University of Warwick, United Kingdom in the year 2019 and an M.Sc. in Computer Science from Universiti Teknologi MARA, Malaysia. She received Ernst Mach Grant ASEA-UNINET, OeAD Austria for a postdoctoral visit to the Technological University of Vienna in 2022. Her research interests are concentrated in the field of ontology, text processing, machine learning and question answering system. She can be contacted at email: shimateo@uitm.edu.my.