Social Sciences & Humanities Open 6 (2022) 100309

Contents lists available at ScienceDirect

Social Sciences & Humanities Open

journal homepage: www.sciencedirect.com/journal/social-sciences-and-humanities-open

Regular Article

Sentiment analysis researches story narrated by topic modeling approach

Saeed Rouhani *, Fatemeh Mozaffari

Department of IT Management, Faculty of Management, University of Tehran, Tehran, Iran

ARTICLE INFO

Keywords:
Sentiment analysis
Social network
Topic modeling
LDA

ABSTRACT

The wild growth of user-generated content like websites, social media, and mobile apps, conducts individuals to create enormous masses of opinions and reviews about products, services, and every day events. Sentiment analysis (SA) embraces a powerful tool for businesses and researchers to explore and study community attitudes, interpretations and, insightful consequences for decision support. This paper brings forward a comprehensive study about main research topics, research trends, and comparisons of research topics" in the field of "sentiment analysis" through "social media" using topic modeling, in specific LDA. The findings of this paper prove that "machine learning" methods are among the most important topics the studies worked on in recent years. Also, various social media platforms such as "Twitter, Facebook, YouTube, and blog" are the SA infrastructures. Among the applications, transportation, spam detection, and decision making are important in terms of the normalized frequency. Finally, findings verify the concept "service improvement by sentiment analysis" indicates the important topic which concentrates quality improvement of firm's service through analysis of customer reviews and it permits researchers and practitioners and also managers have better visions about the hot era of "sentiment analysis".

1. Introduction

Internet and digitalization, driven by the advances in information technology, have transformed our daily lives and shifted the concentration of various areas towards the digital environment in the past two decades (Vanhala et al., 2020). Widespread use of the Internet and social media platforms has provided a large amount of textual data (Greco & Polli, 2020). Specifically, with the pervasiveness of social media platforms, people nowadays share opinions on social networks such as Facebook, Twitter, TripAdvisor, and so on (Ali et al., 2019). Such a fast increase of available text data containing various user-generated content such as opinions, critics, reviews, forum debates has made sentiment analysis a domain of interest (El-Bèze et al., 2010). With the advent of the Big Data era, which contains structured, semi-structured, and unstructured data analytics, unstructured data processing technology of text mining has gained applications in different fields of research. Text analytics methods such as Topic Modeling and Sentiment Analysis are among the most popular methods that researchers employ to study themes, sentiments, viewpoints, etc. which can be conducted by Machine Learning (ML) algorithms (Keikhosrokiani & Pourya Asl, 2022). Moreover, sentiment analysis is considered as a Natural Language Processing (NLP) technique that is used to make human language accessible to computers (Eisenstein, 2019).

So that by automatically mining sentiments and opinions from online emotional reviews and posts, user concerned topics can be discovered (Wang et al., 2019). In other words, the growing importance of sentiment analysis coinciding with the increasing use of social media in different ways such as reviews, micro-blogs, and forum discussions resulted in applying sentiment analysis by using text mining techniques in most business domains because of the high influence of opinions on people's behaviors (Alamsyah et al., 2018). Therefore, as text mining has gained much more attention in recent years, automating information extraction from unstructured data become an important topic in researches (Ali et al., 2019).

The most challenging issue of using text mining in web content is the unstructured or semi-structured nature of it, which requires NLP techniques to deal with (Hemmatian & Sohrabi, 2019). Although existing sentiment analysis techniques can efficiently capture opinions from textual data, the extraction and analysis of relevant and valuable information from social networks have other challenges. Amongst these challenges is dealing with the informal data and imperfect and indirect language such as slang and idioms (Kumar & Garg, 2019), which has led to propose various NLP and ML based frameworks and models and a considerable body of literature in this domain. In addition to methods,

* Corresponding author. Faculty of Management, Jalal ale ahmad highway, Chamran, Tehran, P.O. Box: 1651849661, Iran.
E-mail address: SRouhani@ut.ac.ir (S. Rouhani).

https://doi.org/10.1016/j.ssaho.2022.100309
Received 11 January 2022; Received in revised form 1 May 2022; Accepted 5 June 2022
Available online 16 June 2022

classifying the literature based on the applications indicates that the related researches have been performed in different aspects of the applications; for example, Yue et al. (2019) identified three main aspects of such applications as commercial, political, and public security perspectives. They confirm that despite the plenty of works being conducted to make sentiment analysis viable, it is still in its early stage, and we are facing a multi-faceted problem when considering this research domain (Yue et al., 2019).

Some surveys and researches have been conducted to review this extensive literature in the field of sentiment analysis. A survey on sentiment analysis applications, major tasks, and common challenges was conducted by Ravi and Ravi (2015). Mäntylä et al. (2018) used computer-assisted literature review on sentiment analysis, where they used text mining and qualitative coding in order to analyze 6996 papers from Scopus (Mäntylä et al., 2018). In 2019, Kumar and Garg performed a systematic literature review on context-based sentiment analysis in social multimedia, in which they analyzed 37 studies (Kumar & Garg, 2019). Sentiment analysis and its methods over social media are also reviewed in some researches, such as Singh et al. (2020) and Yue et al. (2019).

Although the growing importance of sentiment analysis in social media has made this field a hot one, there are few surveys or systematic reviews in this multi-faceted problem that can clarify various aspects or topics, trends, and concepts of this field and its applications in the literature. Moreover, the fast development of sentiment analysis in social media makes the surveys outdated to some extent (Yue et al., 2019).

It can be highlighted from the mentioned gap that there is a need for more literature review in this field. Therefore, due to the large number of published documents, using text mining for literature review is considered in this research. In other words, Latent Dirichlet Allocation (LDA), a probabilistic topic modeling introduced by D. M. Blei et al. (2003), is used to discover latent topics from a large volume of data. So, this research has been designed to answer the following questions in the field of sentiment analysis in social media:

RQ1: What are the main research topics in the field of "sentiment analysis" through "social media"?
RQ2: What are the research trends about "sentiment analysis" in "social media" in terms of journal paper publication from 2011 to 2020?
RQ3: How can the research concentration in the field of "sentiment analysis" through "social media" be compared to each other?

In the current research, the above questions will be answered using topic modeling, in specific LDA, and then the topics and concepts related to them are entitled. By using the comparison of normalized frequency of each topic's first word during various years, the trends and hot topics can be found. Moreover, word cloud will be applied to compare the importance of different topics through time.

The scientific contribution of current research includes extracting the main research topics in the field of sentiment analysis in social media, elaborating on research trends, scientometric of published research in this area, and the comparison of the research concentration and their importance. Identifying the scientific fields and the trends in this domain may help the researchers to find the new and hot topics and concentrate on them. Moreover, from the practical perspective, customers' and users' analytics through sentiment analysis utilizing API can be useful for the business owners and those who work in the social network area. Besides, in the field of machine learning, techniques can be customized in order to be used in this domain.

The paper is organized as follows; Section 2 describes the background; Section 3 illustrates the research method. Identify search strategy and configure filtering criteria, data exporting and storing, data pre-processing, LDA model configuration, data analysis, and evaluation are described in this section. In the following, Section 4 is reporting findings for research questions. As a consequence, Section 5 discusses the contributions of the research. Section 6 sums up the findings and implications in the form of a conclusion.

2. Background

2.1. Sentiment analysis

Sentiment analysis (SA), also called opinion mining, review mining, appraisal extraction, or attitude analysis (Ravi & Ravi, 2015), is a research field whose purpose is to analyze people's sentiments toward various topics, events, individuals, issues, products, services, organizations, and their attributes (B. Liu, 2012). Since the early 2000s, sentiment analysis and opinion mining have been studied, and various techniques have been proposed to analyze emotions and opinions from social media (Ali et al., 2019). As mentioned by M. Hu and Liu (2006), in sentiment analysis, by using a collection of opinionated documents, the orientation or polarity (positive, negative, and objective) of an opinion toward a particular aspect of an entity at a given time are determined (M. Hu & Liu, 2006). SA is considered a branch of machine learning, data mining, NLP, and computational linguistics and some elements borrowed from sociology and psychology (Yue et al., 2019). Propelling by the growth of social media, special attention was paid to this field, while the history of NLP starts in the 1950s (Yue et al., 2019).

Koltsova and Koltcov (2013) divided the sentiment analysis literature into two streams (Koltsova & Koltcov, 2013). The first one is to extract the opinions on the preset issues which are presented in the texts (Pang & Lee, 2009; Thelwall et al., 2012), and the second stream is about revealing the agenda, which is a set of issues in a sample of texts by topic modeling (Daud et al., 2010; Steyvers & Griffiths, 2007), or using text clustering (Andrews & Fox, 2007; Carpineto et al., 2009). Sentiment analysis techniques can be classified into three groups: machine learning-based, lexicon-based, and hybrid methods (Abo et al., 2019). SA can have various applications, from observing public mood to measuring customer satisfaction about a product or movie sales prediction (Ravi & Ravi, 2015). Therefore, SA has been used in many areas such as healthcare (Abirami & Askarunisa, 2017; Rodrigues et al., 2016), transportation (Ali et al., 2017; Cao et al., 2013; Kwon et al., 2021), politics (Diakopoulos & Shamma, 2010; Jaidka et al., 2019), e-commerce (Chen et al., 2017; Ng & Law, 2020), financial (Rouhani & Abedin, 2019; Smailović et al., 2013), and environmental issues (Qiao & Williams, 2022).

2.2. Social media

Since the early of 20 century, the growth of the World Wide Web has led to a group of Internet-based applications called Social Media build on the foundation of Web 2.0 that allow the creation and exchange of user-generated content (Kaplan & Haenlein, 2010). Although there is no universally agreed-upon classification for social media platforms, a generic categorization includes social networking services, social bookmarking sites, blogs, content-sharing sites, and opinion sharing sites (Lee, 2018). Therefore many platforms and sites such as Facebook, Google+, LinkedIn, Twitter, Reddit, Digg, Instagram, Pinterest, Flickr, YouTube, Yelp, and TripAdvisor can be included in this categorization (Lee, 2018).

Since online social networks are a rich source of information consisting of data variety, which are structured, semi-structured, and unstructured, it has become the basis of various research and studies, which is unprecedented (Abulaish & Fazil, 2018). Social media analytics can be considered as collecting and analyzing data from social media platforms to address specific problems by decision-makers (Lee, 2018). For example, online customer reviews, which contain information such as purchase experience, customer complaints, user experience, satisfaction, and customer ratings regarding products and services (Jin et al., 2019), can be used to obtain consumers' innovative ideas and improve customer relationships (Lee, 2018).

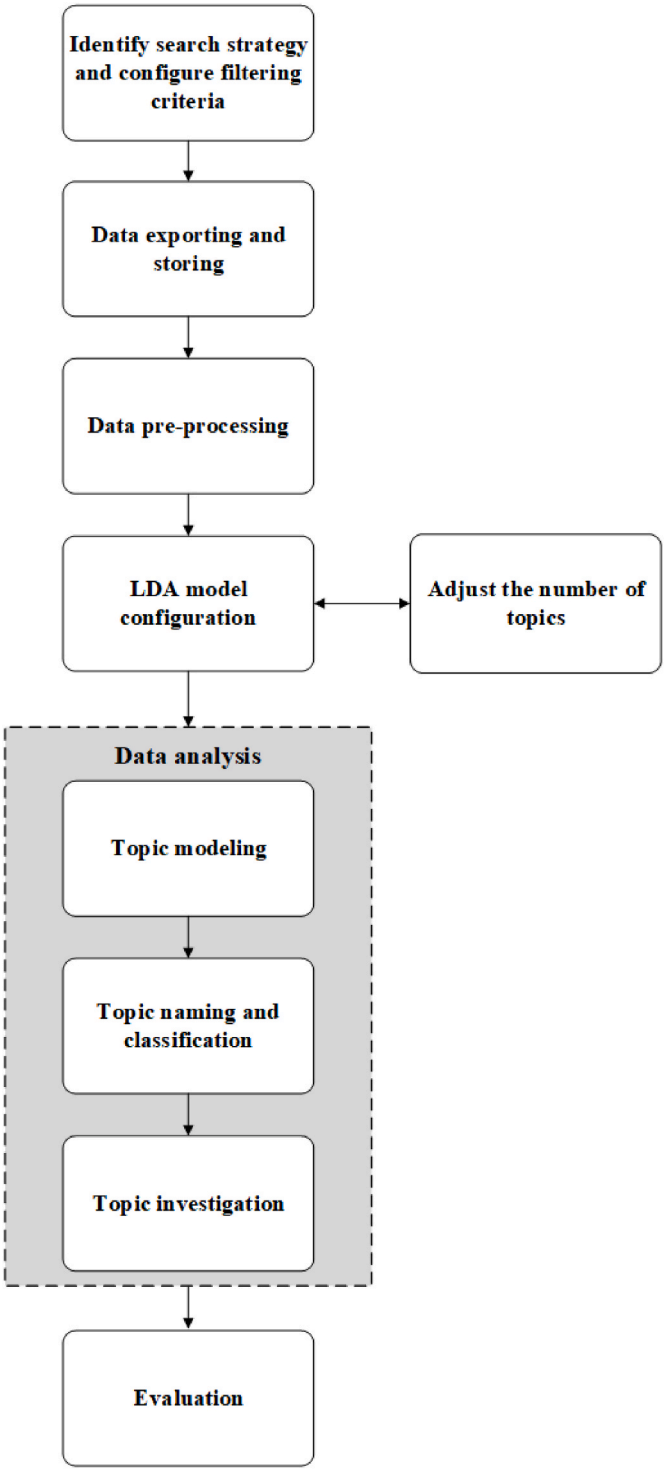


Fig. 1. Research method steps.

2.3. Topic modeling

One of the most relevant techniques used for modeling the evolution of event over time are topic models (Koller & Friedman, 2009), where the topic can be defined as the main subject matter of a text (Koltsova & Koltcov, 2013). A topic model can be considered a probabilistic model that relates documents and words through variables, called main topics (Dueñas-Fernández et al., 2014). Generally, topic discovery adopts clustering or machine-learning techniques to classify similar topics and track emerging issues (Kim et al., 2016). By using topic modeling, dominant themes of given texts from social media platforms, news articles, and purchase behavior can be detected (Akter et al., 2016). There are two main methods used for topic modeling (Rana et al., 2016): Probabilistic Latent Semantic Analysis (pLSA) (Hofmann, 2001) and Latent Dirichlet Allocation (LDA) (D. M. Blei et al., 2003), while there are several extensions to the standard LDA model such as Hierarchical Dirichlet Processes (Teh et al., 2006), Dynamic Topic Models (D. M. Blei & Lafferty, 2006), and Correlated Topic Models (D. Blei & J. Lafferty, 2006) amongst others.

Latent Dirichlet Allocation (LDA), which is a generative probabilistic model for collections of discrete data such as text corpora, is a "three-level hierarchical Bayesian model, in which each item of a collection is modeled as a finite mixture over an underlying set of topics" (D. M. Blei et al., 2003). This technique is a well-known unsupervised method in text mining in which latent topics of documents are derived from estimated probability distributions while disregards word order (Ali et al., 2019; Dueñas-Fernández et al., 2014). Hence, this model is based on the assumption that there are latent topics related to the document in the form of a multinomial distribution of words over a fixed vocabulary (Lau et al., 2014). In the literature, LDA has been used for topic modeling the texts related to many applications such as transportation, education, social networks, geo-tweets, e-commerce, marketing and brand management, etc. (Ali et al., 2019; García-Pablos et al., 2018; T.; Hu et al., 2019; X.; Liu, 2020; X.; Liu et al., 2017; Nimala & Jebakumar, 2019; Sommer et al., 2012; Wu et al., 2020). Moreover, LDA algorithm can be adapted for product aspect extraction (Ozyurt & Akcayol, 2021). LDA can also be used to conduct literature review over a large volume of documents in which each document consists of latent topics that correspond to various aspects of the document under review (Mäntylä et al., 2018).

3. Research method

In this section, the research method steps are elaborated to answer current research questions. Fig. 1 presents the stages of the research process.

3.1. Identify search strategy and configure filtering criteria

To conduct the literature review in the field of sentiment analysis in social media, Scopus, which is one of the most widely employed academic databases, was chosen. Then, the results became exclusive to 8 prominent publishers since these publishers are the most famous in the field of information systems. Therefore, this study was carried out over the publications of Science Direct, Springer, IEEE Explore, Emerald, Taylor & Francis, SAGE, Wiley Online Library, IGI, and ACM Digital library indexed in Scopus; Scopus, according to its developer Elsevier "is the largest abstract and citation database of peer-reviewed literature: scientific journals, books, and conference proceedings." Therefore, Scopus can provide good coverage of scientific literature. Moreover, only the journal or review articles were included. By applying these criteria, documents published from 2008 to the March of 2020 can be retrieved. In the search strategy, suitable keywords are defined to search the literature on "Sentiment analysis in social media."

The articles written in the English language were investigated since it is the language of academic writing, but this criterion can also be assumed as a research limitation. Finally, an abstract and title evaluation was carried out on the selected articles to choose the suitable and relevant ones. In other words, the stored documents have been monitored by abstracts and titles, and if the data used does not relate to social media, the article has been omitted. Fig. 2 briefly shows the search strategy, filtering criteria, and the proportion of the selected articles by each publisher.

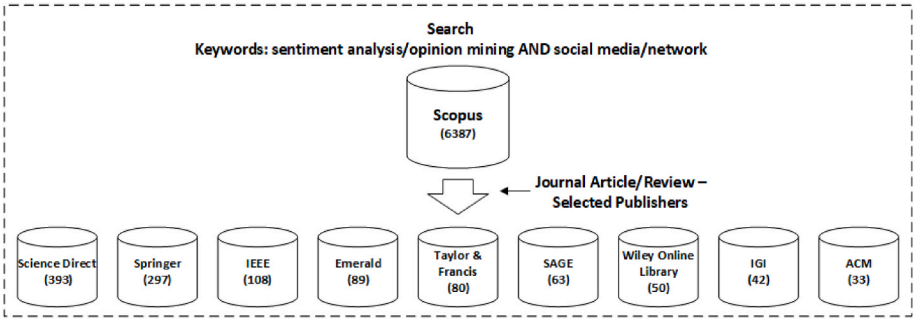


Fig. 2. Search strategy and proportion of articles by each publisher.

3.2. Data exporting and storing

As mentioned, Scopus abstract and citation database was used by searching ("sentiment analysis" OR "opinion mining") AND ("social media" OR "social network"), and the journal articles and reviews were selected. The results were exported and stored as an excel file containing citation information, abstracts, and keywords.

3.3. Data pre-processing

After data collection, the next step is data preprocessing (text preprocessing in our case). The R language was used to perform text preprocessing. Using R-package "tm" (Feinerer et al., 2015), lowercasing of all letters, removing punctuation and numbers, removing stop-words, which are common words in English that do not provide much insight to the sentence, have been applied to the abstracts of the selected documents from the previous stage. Then, stemming as another step of text preprocessing in which the variant word forms are mapped to their base form, as mentioned by J. Singh and Gupta (2016), was performed. Then, the words with less than three letters were removed. Next, a Term-Document Matrix was built to feed LDA. Term-Document matrix, also known as Document-Term matrix, describes each term's frequency in each document. The current research adopted the approach as Mäntylä et al. (2018) in which Term Frequency-Inverse Document Frequency (TF-IDF) computations was utilized in order to remove the words having less than median TF-IDF value from the document-term-matrix and improve the raw term frequency by considering the amount of information of each term represented by inverse document frequency.

3.4. LDA model configuration

The most important parameter of LDA is the number of topics that should be inferred from the corpus. Too many topics would lead to several topics representing one cohesive subject, while too few topics can result in an incomplete analysis (Dahal et al., 2019). To tackle this issue, Griffiths and Steyvers (2004) strategy to find the optimal number of topics by using Gibbs sampling and log-likelihood measure was used and the number of topics was selected recursively. Log-likelihood is the posterior likelihood of the corpus conditional on the topic assignments. So, $P(\mathbf{w}|T)$ should be computed in which $\mathbf{w}$ represents the words in the corpus, and T is the number of topics. However, summing over all possible assignments of words to topics, $\mathbf{z}$, would be complicated, so $P(\mathbf{w}|T)$ can be approximated by taking the harmonic mean of a set of values of $P(\mathbf{w}|\mathbf{z}, T)$ when $\mathbf{z}$ is sampled by Gibbs sampling algorithm from the posterior $P(\mathbf{z}|\mathbf{w}, T)$ (Griffiths & Steyvers, 2004).

As mentioned in Section 2, the LDA model is based on the latent correlation between words and themes in the documents and the assumption of bag-of-words, so the Document-Term Matrix was used as the input to the model. Moreover, topic distribution over documents and word distribution over topics are considered to have prior probability of Dirichlet, so the hyperparameter of the model are α and β which are prior probability distribution for topics over document and words over topic, respectively. Based on Griffiths and Steyvers (2004), α and β were assumed equal to 50/T and 0.1, respectively, where T is the number topics. By determining α and β, in order to find the optimal number of topics, harmonic average of the samples generated by Gibbs Sampling calculated to derive P(W|T) where W is the words of the dataset (Griffiths & Steyvers, 2004). Hence, by changing the number of topics from 10 to 200, the log-likelihood of each model can be computed and compared to each other by using R-package topicmodels (Grün et al., 2020). Log-likelihood is the posterior likelihood of the corpus conditional on the topic assignments (Ponweiser, 2012).

3.5. Topic modeling

Data analysis consists of three steps: Topic modeling, Topic Naming, and Topic investigation. After finding the optimal number of topics, the LDA function was applied, which is implemented in R-package topicmodels, to the available text to explore the latent topics. As a result, the subset terms for each topic were derived. These terms were arranged based on the probability of their occurrence in a given topic. Therefore, a hierarchical relationship exists between each topic and its comprising terms. Each topic can be represented by its most probable terms, derived from the posterior distribution over the assignments of words to topics.

3.6. Topic Naming

Latent topics were extracted as the clusters in which the terms were ordered by their probability of occurrence. Through labeling these clusters, the topics can be explored, and then by merging the topics, the concepts related to each group of topics were discovered. In other words, based on each topic's terms, the semantic or representative label for that topic was inferred by domain knowledge. Then, the labels merged hierarchically to form the main topics. So, each category or class, which was named as a main topic, consists of two or more labels that are related to each other. Next, the topics were combined to shape the concepts inferred from a set of topics.

3.7. Topic investigation

Topics aggregated to form the concepts, and these concepts compiled to discover the trend of researches in recent years. Therefore, the tool which was used in this research to investigate the research trends are the classes that were derived and named in the previous steps. To investigate the change of each topic's importance over different years, the normalized frequency of the most probable term in each cluster as the representative of the importance of that cluster was used, and the topics compared to each other in this regard. Several R-packages were used at this stage, such as wordcloud (Fellows et al., 2018). Comparing different clusters in terms of coherence and normalized point-wise mutual information are among other analyses presented in this research.

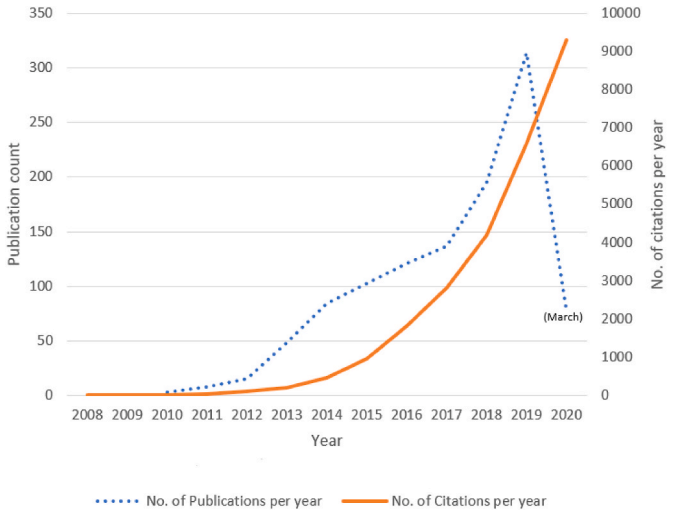


Fig. 3. Annual publications and citations based on Scopus database.

3.8. Evaluation

In the final stage, two measures were used to investigate the coherency and interpretability of the topics. These measures are introduced by Bouma (2009), Lau et al. (2014), and Mimno et al. (2011). Topic coherence measure and Normalized Point-wise Mutual Information were used to find out about fluctuations of the topics' quality and compare the topics' interpretability. As mentioned by several studies that use LDA, a good topic model should lead to human-interpretable topics that are coherent but distinct from each other (Dahal et al., 2019). However, some topics may be incoherent in terms of the most probable words for each topic (Mäntylä et al., 2018). Topic coherence score is an indicator of topic quality (Bakharia et al., 2016). In this regard, the topics are evaluated by two measures. The first one is the topic coherence measure introduced by Mimno et al. (2011), which can be obtained by formula (1) (Mimno et al., 2011).

$$C\left(t; V^{(t)}\right) = \sum_{m=2}^{M} \sum_{l=1}^{m-1} \log \frac{D\left(v_m^{(t)}, v_l^{(t)}\right) + 1}{D\left(v_l^{(t)}\right)}. \quad (1)$$

where D(v) is the frequency of the documents with at least one word of type v, and D(v, v') is the frequency of documents containing one or more words of type v and at least one word of type v'. $V^{(t)}$ is the vector comprised of a specific number of the most probable words in topic t (Mimno et al., 2011). Since the computed values are negative, the closer this value to zero, the higher is the coherence (Mimno et al., 2011). So, this metric indicates the semantic similarity between the most probable words of a given topic.

The other metric introduced in the literature is Normalized Point-wise Mutual Information, which can be derived by formula (2) (Bouma, 2009; Lau et al., 2014).

$$NPMI\left(t; V^{(t)}\right) = \sum_{m=1}^{M-1} \sum_{l=m+1}^{M} \frac{\log \frac{P\left(v_m^{(t)}, v_l^{(t)}\right)}{P\left(v_m^{(t)}\right) P\left(v_l^{(t)}\right)}}{-\log P\left(v_m^{(t)}, v_l^{(t)}\right)}. \quad (2)$$

where v represents the words. Therefore, in formula (2), the words' probabilities and their joint probabilities are used. NPMI metric measures the interpretability of the topics. The R-package text2vec (Selivanov & Wang, 2016) was used to apply NPMI to the dataset.

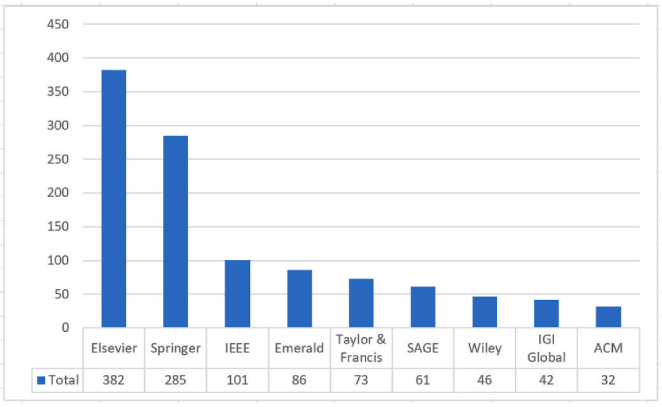


Fig. 4. Total number of published papers about sentiment analysis in social media per publisher.

Table 1
Dataset used for literature review.

Number of abstracts	1108
Total number of words	138271
Total number of characters (without space)	816237
Total number of characters (with space)	953400

4. Results

This section presents the results based on the research method and data analysis described in Section 3. As Fig. 3 shows, the number of publications and citations in sentiment analysis in social media that met our filtering criteria has increased during the years 2008–2020, while Fig. 4 presents the categorized number of papers per publisher. Since the search step was conducted by the March of 2020, the number of publications in 2020 can be estimated several times the current number. It can be seen that sentiment analysis in social media is a hot research area, and a considerable amount of research has been conducted in this area. Therefore, as depicted in Fig. 3, a sharp increase in the number of publications and citations from 2008 to 2020 has been achieved. Boosting in user generated contents, increase of infiltration coefficient of mobile social media in developing countries and also business application of social media can be assumed as researcher motivations to enter this hot research era.

Based on the search results, our corpus consists of 1108 abstracts of the papers about sentiment analysis in social media published by the selected publishers indexed in the Scopus. Table 1 provides some information about the collected dataset that includes 1108 papers.

The number of unique terms in the Term-Document Matrix was 5629, which was reduced to 1625 after the pre-processing phase and removing the terms with less than the median TF-IDF weights.

As described in Section 3.3, the number of topics was changed to find the optimal number of topics based on Griffiths and Steyvers (2004) approach. The result of model selection based on the number of topics is shown in Fig. 5.

There is a slight difference between the log-likelihood of models with topics changes from 50 to 100. Therefore, the mean of the coherence score was computed for various states with topics between 50 and 80. The results showed that the coherence score for 50 topics is higher (closer to zero) than the other number of topics. Moreover, since if fewer

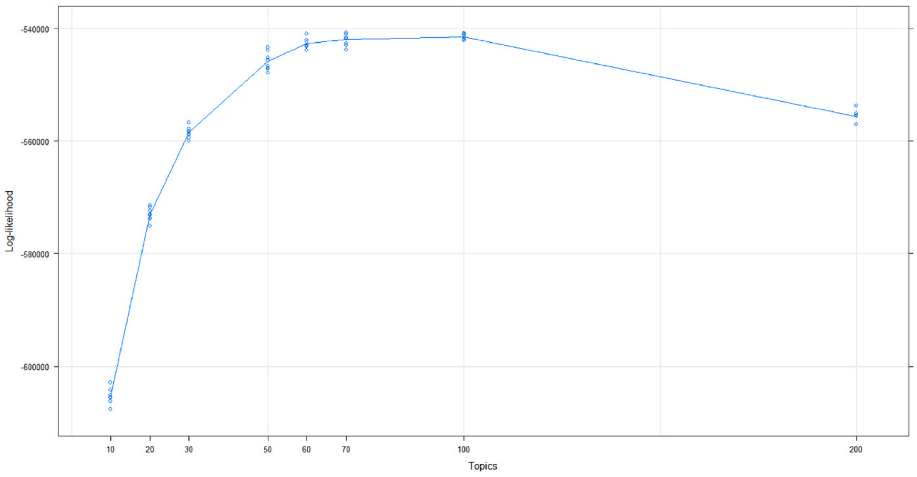


Fig. 5. Log-likelihood of each LDA model based on topic number sampled by Gibbs sampling.

Table 2
Final topics and their five representative terms.

	Topic 2	Topic 3	Topic 4	Topic 6	Topic 8	Topic 9	Topic 10
1	topic	classifi	web	data	make	network	social
2	model	machin	blog	big	decis	social	collect
3	sentiment	learn	search	process	use	communiti	interact
4	detect	classif	network	social	inform	group	individu
5	latent	use	semant	analyt	way	cluster	can
	Topic 11	**Topic 12**	**Topic 13**	**Topic 15**	**Topic 16**	**Topic 17**	**Topic 18**
1	featur	public	sentiment	text	user	emot	languag
2	detect	issue	messag	imag	content	affect	process
3	set	onlin	microblog	textual	network	express	text
4	use	media	analysi	visual	generat	state	natur
5	extract	monitor	express	sentiment	social	human	corpus
	Topic 19	**Topic 20**	**Topic 21**	**Topic 22**	**Topic 23**	**Topic 24**	**Topic 25**
1	method	post	review	tool	data	tweet	inform
2	base	comment	onlin	news	use	twitter	subject
3	propos	content	rate	develop	mobil	use	howev
4	use	facebook	product	applic	locat	relat	analysi
5	dataset	signific	tourism	softwar	citi	follow	question
	Topic 26	**Topic 27**	**Topic 28**	**Topic 29**	**Topic 30**	**Topic 31**	**Topic 33**
1	system	measur	media	servic	domain	framework	market
2	user	base	social	use	train	new	stock
3	recommend	use	platform	qualiti	label	propos	predict
4	rank	result	also	experi	supervis	base	financi
5	prefer	metric	use	analysi	annot	present	model
	Topic 34	**Topic 35**	**Topic 36**	**Topic 37**	**Topic 38**	**Topic 39**	**Topic 40**
1	polar	negat	model	relat	health	purpos	sentiment
2	sentiment	posit	learn	knowledg	onlin	methodolog	word
3	context	brand	network	can	support	approach	lexicon
4	score	sentiment	deep	base	use	research	specif
5	term	neutral	neural	associ	communiti	find	construct
	Topic 41	**Topic 42**	**Topic 43**	**Topic 44**	**Topic 45**	**Topic 46**	**Topic 47**
1	product	opinion	sentiment	aspect	communic	techniqu	polit
2	custom	mine	analysi	level	analys	analysi	predict
3	consum	extract	exist	base	use	paper	elect
4	busi	express	mani	sentenc	polici	also	campaign
5	compani	peopl	demonstr	entiti	examin	analyz	result

	Topic 48	Topic 49	Topic 50
1	larg	differ	event
2	comput	structur	time
3	analysi	text	real
4	scale	understand	propos
5	effici	concept	detect

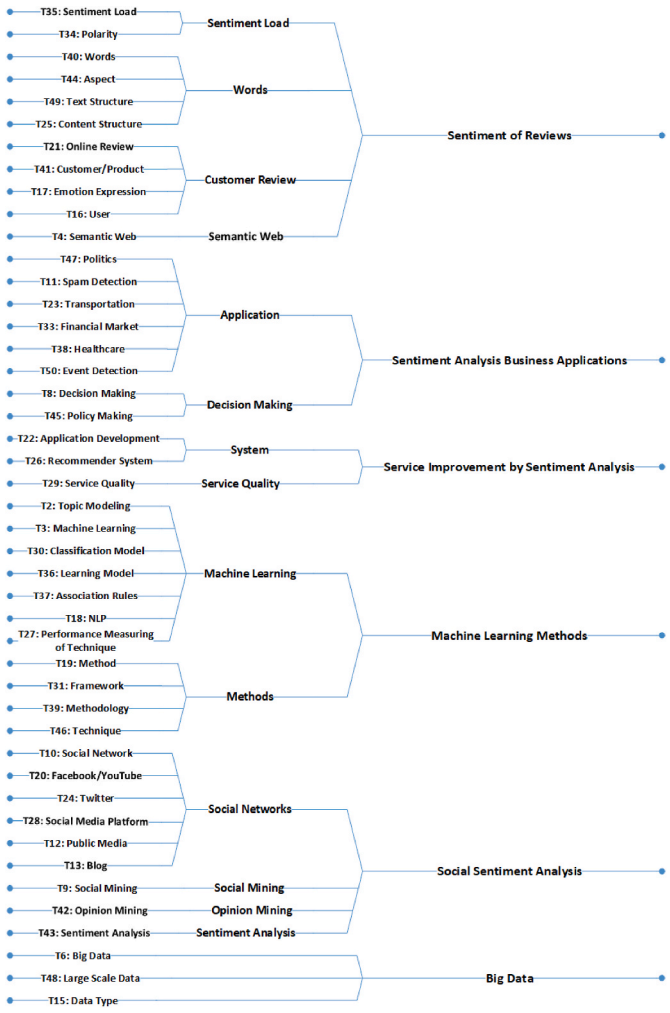


Fig. 6. Graph of labels, topics, and concepts extracted by topic modeling.

topics are chosen, more interpretability and labeling capability can be achieved, to facilitate the concept extraction and avoid divergence of topics, 50 was chosen as the LDA model's number of topics.

The selected LDA model with 50 topics was applied to the Term-Document Matrix using the LDA function implemented in R-package topicmodels (Grün et al., 2020; Hornik & Grün, 2011). Since LDA provides soft clustering of the terms, there are overlaps between the terms in various clusters. After evaluation of topics, five clusters were eliminated. The most probable terms of these clusters were the common research words such as "research," "studi," "result," so on. Therefore, finally, 45 topics were chosen. Table 2 presents the five most probable terms in each cluster. Each topic's terms are ordered by the probability that each term is assigned to a given cluster. In other words, a weight assigned to each term by each topic which indicates the probability of that term relates to the topic.

By investigating the most probable (representative) terms in each topic, each cluster can be labeled accordingly. Then, clusters that relate to each other are grouped to develop main topics, and finally, the main concepts of the dataset by merging main topics could be derived. Fig. 6 depicts the graph of labels, topics, and concepts of the corpus. The concepts were extracted in a 3-stage process:

1. By reviewing the semantic of each cluster's terms, a label wasassigned to each cluster (T1 to T50).
2. The main topics were derived by putting together the related labels.
3. Fundamental concepts of the domain under consideration were extracted by domain knowledge.

All the stages of this process (labeling, topics and concepts naming) were conducted by expert opinions and domain knowledge.

In the next step, to investigate the change of each topic's importance and each concept through the topics that form it, the first probable term in each topic was considered as the representative of that topic. Then, the normalized frequency of each term was used by dividing its frequency by the total number of terms in all the documents of the years from 2015 to 2019. Normalizing the terms' frequency can be an indicator of the real weight of these terms in each year. Fig. 7 provides the trends of main topics in each concept through various years.

As depicted in Fig. 7, the topic "words" has a significant frequency compared to other topics related to the concept of "Sentiment of Reviews." Some other topics in this concept show a slight increase, such as "aspect." Amongst "Sentiment Analysis Business Applications," "transportation" is above the others, followed by "Spam Detection." "Decision Making" also shows relatively an increasing change. In the plot related to "Machine Learning Methods," "learning model," and "method" are above the others and show increasing changes while the dynamic of "association rules" is decreasing. Topics included in the "Social Sentiment Analysis" concept show constant dynamics, while "sentiment analysis" and "social network" have more normalized frequencies than other topics through various years.

Table 3 provides the number of documents that support each concept according to the terms related to the topics that comprise the concept. "Social sentiment analysis" is the most supported concept in terms of the number of documents since it contains the search keywords of this study. Next, "machine learning methods" have gained the most support and attention by the researchers that indicates among various methods proposed, machine learning ones have become the most established to analyze data in social networks. "Sentiment of reviews" placed third, which indicates the importance of the semantics and the structure of reviews and their related sentiment load.

Although LDA discovers the underlying themes of the textual content and its hidden topics, word cloud also can provide a quick and basic view of the texts through the word frequencies. So the R-package wordcloud was used, which uses Time-Document Matrix as its input to show the importance of each word in terms of its frequency (Fellows et al., 2018). "wordcloud" function applied to the keywords and titles of the documents under review in three different years: 2011, 2015, and 2020. Abstracts were not considered in this analysis because they were investigated in topic modeling. Fig. 8 presents the results. These word clouds can be used to better understand the frequency and importance of each topic in visualization aspects. Some common words such as sentiment and opinion were removed from the dataset in order to detect the differences.

The results from the word clouds proves that while in 2011 words such as "review," "system," and "user" had the most frequencies, in 2015, words such as "media," "Twitter," and "text" became more frequent, and words "learn," "machine," and "network" are among the new words that can be seen with the most frequencies in 2020.

In the last part of this section, a concept which indicates the semantic validity of the clusters derived by LDA is investigated. Fig. 9 presents the topics ordered by topic coherence measure.

As mentioned in section 3.7, topic coherence score is an indicator of topic quality. This measure was used in this research for two reasons.

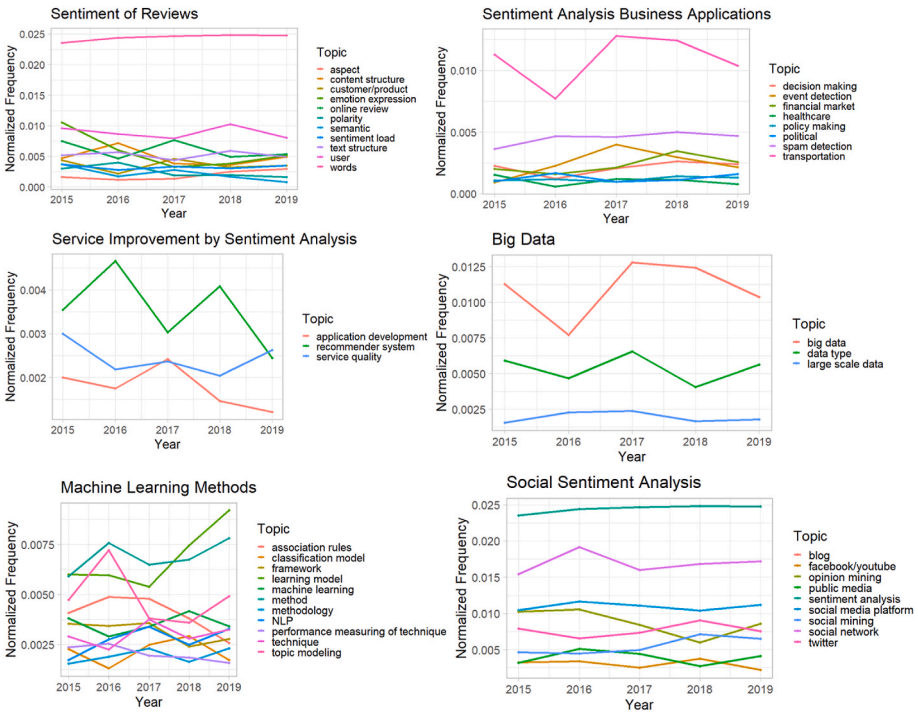


Fig. 7. Dynamics of the topics from 2015 to 2019.

Table 3
Number of documents that support each concept.

Concepts	Sentiment of Reviews		Sentiment Analysis Business Applications		Service Improvement by Sentiment Analysis	Machine Learning Methods		Social Sentiment Analysis		Big Data
	web	user	polit	spam	tool	methodolog	classifi	social	post	big
	emot	subject	transport	market	system	framework	associ	tweet	media	larg
	negat	product	health	event	servic	techniqu	languag	public	network	text
	word	structur	decis	polici		measur	method	microblog		
	review	polar				train	topic			
	aspect					model				
No. of Documents	949		532		466	1005		1039		541

First, to choose the number of topics after using log-likelihood, as shown in Fig. 5, the means of coherence scores for topics with a slight difference in log-likelihood were computed, and 50 was chosen as the number of topics. Second, the coherence score was used to choose and prioritize the topics. In this regard, as it can be seen in Fig. 9, topics 39, 3, and 30, which relate to "Methodology," "Machine Learning," and "Classification Model," respectively, are the three most important labels achieved by LDA. So, by using the coherence score, the main topics which have more quality can be recognized.

Fig. 10 shows the NPMI of the topics. The more the normalized pointwise mutual information, the more interpretable the topic is.

The proven fact is that most of the topics with high coherence based on coherence score also shows high NMPI metric. These clusters usually show the similar probability among the most probable terms, and have more interpretability, so that domain expert can judge it as an interpretable topic which belongs to unified concept. However, there are some other semantically coherent topics whose most probable term has relatively more probability than the other words. For example, while in topic_17, the term "emot" has much more probability than the other terms of the topic, there are interpretable relation between the most probable terms in this topic to infer "emotion expression" as the label. Therefore, better coherence does not essentially indicate the importance of the LDA model topic and vice versa. However, it should be noted that when the coherency or NPMI are higher, more terms associate with extracting the label from a topic. In contrast, when there are one or two terms with far higher probabilities, they would have a more prominent role in naming the topic. The topics and 10 of their represented terms ordered by each term's probability are presented in the Appendix.

5. Discussion

Based on the collected dataset, which contains studies from 2008 to the March of 2020, and analysis of the data by topic modeling, our findings are now presented to answer the research questions.

RQ1: What are the main research topics in the field of "sentiment analysis" through "social media"?

LDA model was applied to the dataset, and after adjusting the number of topics, generated 50 clusters, in other words, the latent topics. After applying LDA, one of the critical tasks is to give labels to these latent topics based on the domain knowledge, which represent the main research sentiment analysis areas. After assigning labels to 50 extracted topics and eliminating 5 of them because they contain

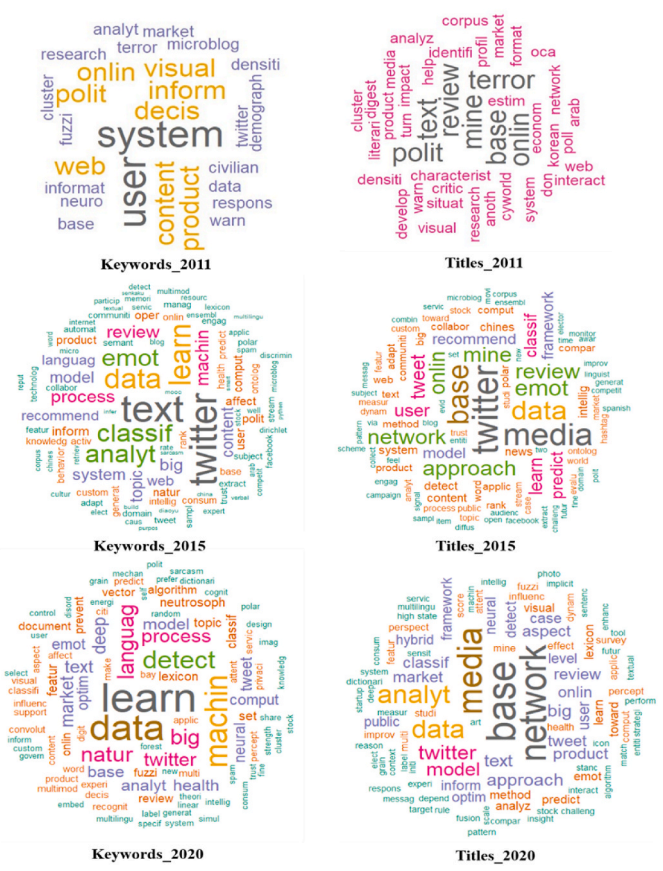


Fig. 8. Word clouds of the terms in the documents under review in the years 2011, 2015, and 2020.

common research words, these 45 labels were categorized into 15 main topics. Table 4 presents these 15 topics, which were derived by merging the relevant labels.

Each main topic includes several labels. For example, "Application" consists of recognized applications studied in the literature in the sentiment analysis through social media such as "political," "spam detection," "transportation," "financial market," "healthcare," and "event detection."

Then, another step was taken to identify the concepts of this domain by using the relations amongst extracted topics and also considering beta, a parameter obtained from the LDA model which indicates the probability of assigning each term to a given topic by using the posterior distribution over the assignments of words to topics. This step was conducted by inferring the concepts from two or more topics. The derived concepts are illustrated as follows.

Main concepts derived from 15 topics:

- Sentiment of Reviews
- Service Improvement by Sentiment Analysis
- Social Sentiment Analysis
- Sentiment Analysis Business Applications
- Machine Learning Methods
- Big Data

In a similar study which was conducted by Mäntylä et al. (2018) in the field of sentiment analysis, some similar key concepts such as "social," "reviews," "media," "applications," and "techniques" were identified. In the study of Mäntylä et al. (2018), the key concepts of sentiment analysis identified as "social," "online," "reviews," "media," and "product." In the current research, broader concepts have been inferred by merging the topics to make them more universal.

RQ2: What are the research trends about "sentiment analysis" in "social media" in terms of journal paper publication from 2011 to 2020?

As presented in Section 4, the normalized frequency of the most probable terms in each topic and each concept can be compared through years (Fig. 7). Moreover, obtaining word clouds of keywords and titles,

Table 4
Main research topics obtained by applying LDA and merging related labels.

Main Research Topics		
Sentiment Load	Words	Customer Review
Semantic Web	Application	Decision Making
System	Service Quality	Machine Learning
Methods	Social Networks	Social Mining
Opinion Mining	Sentiment Analysis	Big Data

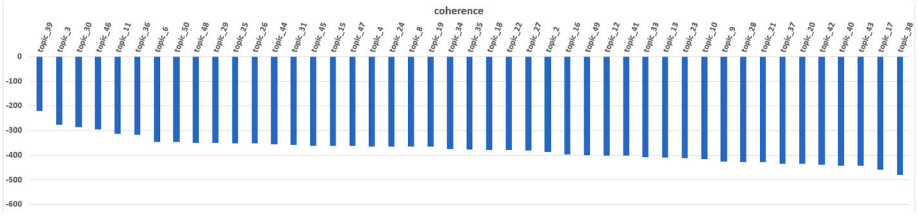


Fig. 9. The coherence score of the topics.

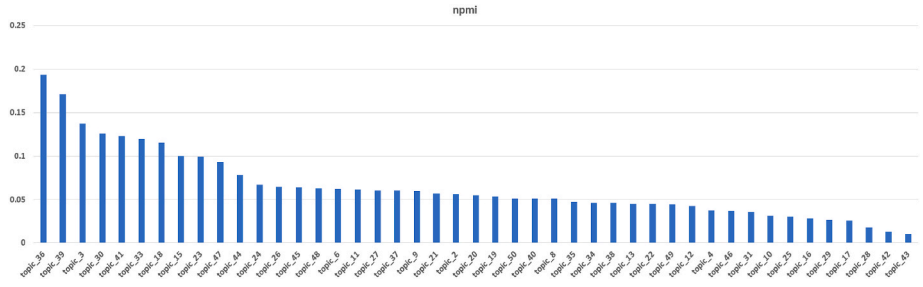


Fig. 10. The NPMI of the topics.

the frequency of the most important words in terms of overall occurrence can be compared in three different years: 2011, 2015, and 2020. The results of both comparisons indicate that topics and words such as "learn," "machine learning," "media," "network," "language," "detect," and "aspect" are increasing in recent years while the words such as "system," "user," "review," "product," and "online" can be recognized in the early years that the studies were conducted in this field. Among the social media, "Twitter" occurred more frequently in the literature. Kumar and Garg (2019) also mentioned the same. In other words, in the sentiment analysis context, more studies have been done on Twitter than on other social media (Kumar & Garg, 2019). Therefore, a shift from the sentiment analysis of organizational system level to the social networks can be observed from 2011 to 2020. In other words, in 2011, sentiment analysis had more concentration on organizational system software users, while in 2020, the analysis becomes applicable for the social networks.

RQ3: How can the research concentration in the field of "sentiment analysis" through "social media" be compared to each other?

As depicted in Table 3, the concepts can be compared in terms of their consisting topics' representative terms. By investigating the number of abstracts that include these terms and removing the repeated ones, the number of documents that support each concept in terms of containing its related topics' representative terms, which is among the most probable terms in each cluster, can be achieved. Therefore, it can be inferred that after "social sentiment analysis," "machine learning methods" support with more documents, that is 1005, followed by "sentiment of reviews," 949, "big data," 541, "sentiment analysis business applications," 532, and "service improvement by sentiment analysis," 466, respectively. In addition to "social sentiment analysis," which contains the fundamentals of the research domain, the results indicate that machine learning techniques are the most important tools for sentiment analysis; it encompasses all methods used to accomplish sentiment analysis by using machine learning algorithms. Moreover, the high number of documents that support the "sentiment of reviews" highlights the significance of structure of review and topics such as NLP and words. "Big data" identified as the fourth most supported concept since this domain relates to one of its important subsets. In other words, text as unstructured data is one of the aspects investigated in big data analytics. Next, "sentiment analysis business applications," as mentioned before, encompass a broad range of businesses. Lastly, documents support "service improvement by sentiment analysis," consider the methods to improve software, social network, services, and product.

6. Conclusion

With the fast-growing use of social media platforms, a large volume of user-generated content, mostly unstructured texts, sentiment analysis has become useful in various applications and domains. Mining and analysis of this large volume of unstructured data necessitates applying text mining and NLP techniques and encompass many challenges and at the same time different applications that have led to much work in this field and formed a large number of literature studies. Reviewing the fast-growing field of sentiment analysis through social media by the LDA model can reveal the most important and new topics and concepts latent in the literature. Moreover, finding out about the dynamics of topics in various years and the importance of each concept in terms of the documents support it were among the main objectives of this research. The method to fulfill the goals of the research was conducted through selecting a dataset consisting of abstracts of documents that were extracted from Scopus and limited by filtering criteria. Adjusting the number of clusters by Gibbs sampling and log-likelihood measure as used by Griffiths and Steyvers (2004), 50 recognizable clusters were identified by applying LDA to the pre-processed texts. The labels were then assigned to the clusters, and by merging them into 15 main topics, six significant concepts in this field could be found. Each of them represents one of the main areas that researches are conducted. Further analyses were done in order to investigate the trends and importance of concepts through the years. Interpretations of topic coherence were also provided. The findings of this study show that machine learning methods such as classification, NLP, and topic modeling are among the most important topics the studies worked on in recent years. Various social media platforms such as Twitter, Facebook, YouTube, and blog each dedicated a cluster to themselves. Sentiment of reviews, which is another important concept recognized in this research, consists of sentiment load, words, customer review, and semantic web while some topics such as "aspect" increased in recent years. Big data and Business applications are also other concepts based on the findings of this research. Among the applications, transportation, spam detection, and decision making are important in terms of the normalized frequency. Finally, the concept "service improvement by sentiment analysis" indicates the topics such as service quality and recommender systems, which are all the subset of sentiment analysis through social media.

It should be noted that this literature review has some limitations in terms of data source, filtering criteria, and method. Scopus was used as a source of collected literature, and only the journal articles and reviews were selected. Moreover, there are several clustering methods that can be used and compared in order to find the most coherent and interpretable topics.

The implications of this research fall into two categories. From the scientific point of view, identifying the scientific fields and the trends in this domain may help the researchers to find the new and hot topics and concentrate on them. Moreover, from the practical perspective, customers' and users' analytics through sentiment analysis utilizing API can be useful for the business owners and those who work in the social network area. Besides, in the field of machine learning, techniques can be customized in order to be used in this domain. Future research can extend our work in various ways, such as more coverage in the database and using other databases, other types of documents, and other clustering techniques, which means applying other techniques instead of LDA. Using sentiment analysis of each topic also can be considered an extension of current research. Comparing topics based on publisher or journal and their citations can be another future policy.

CRediT authorship contribution statement

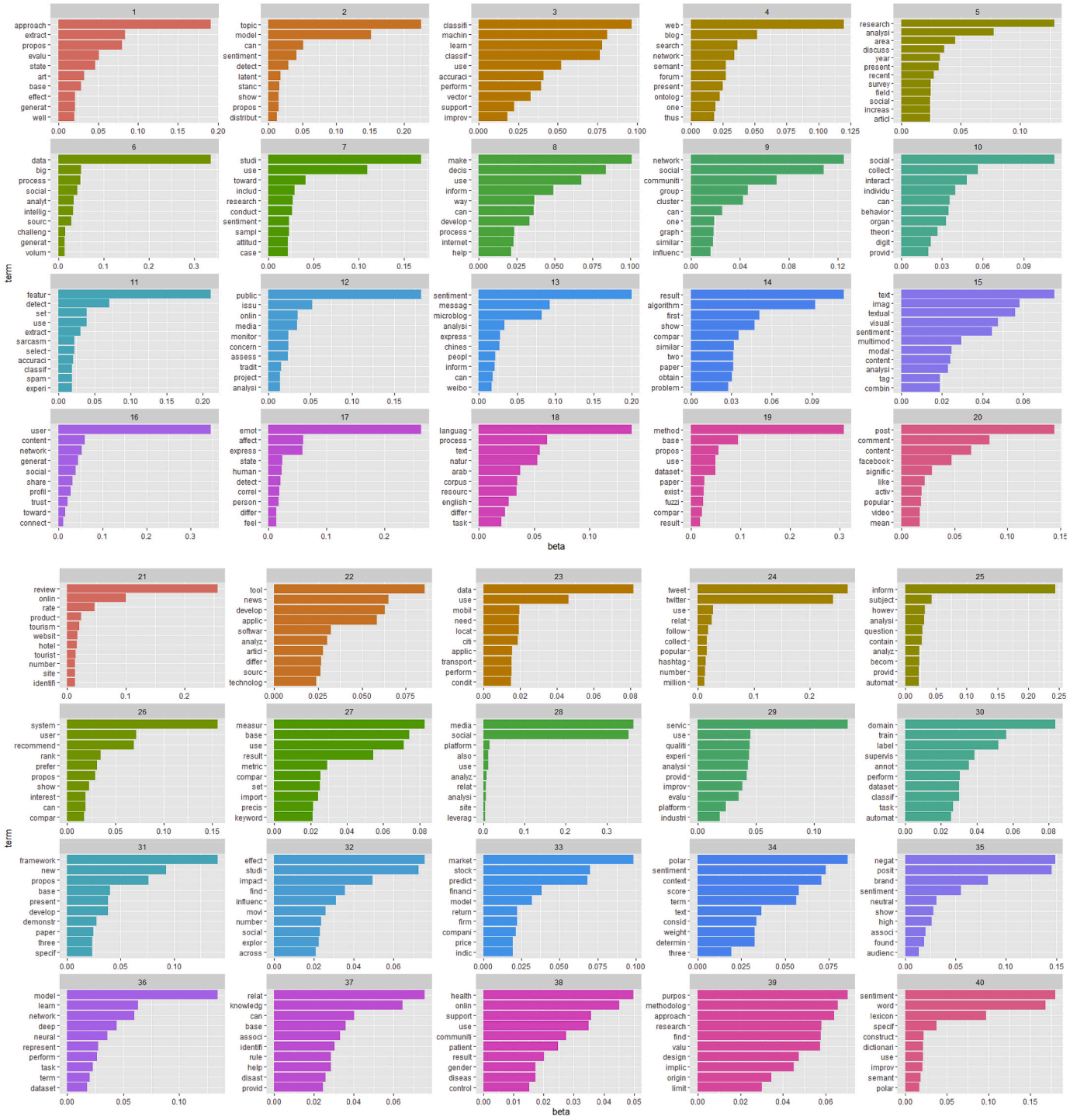
Saeed Rouhani: Conception and design of study, Formal analysis, Writing – original draft, revising the manuscript critically for important intellectual content, Approval of the version of the manuscript to be published. **Fatemeh Mozaffari:** acquisition of data, Formal analysis, Writing – original draft, revising the manuscript critically for important intellectual content, Approval of the version of the manuscript to be published.

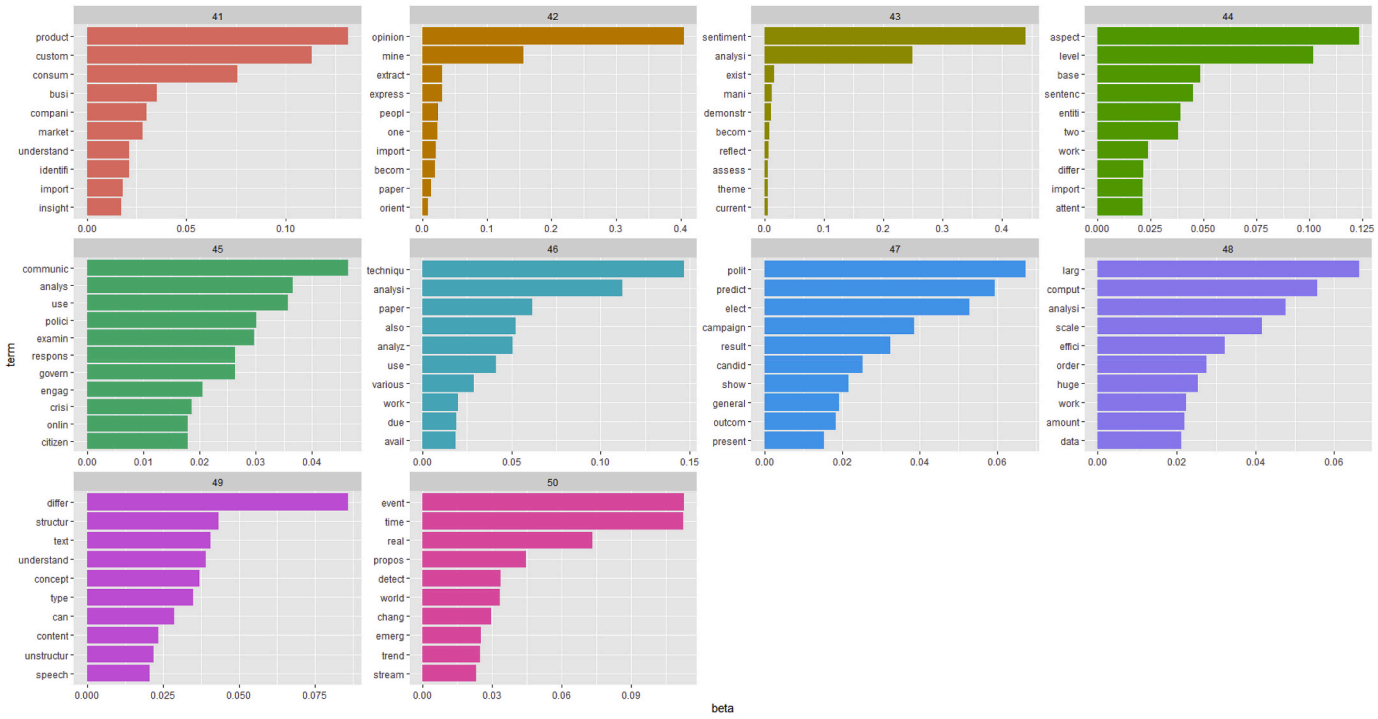
Acknowledgements

All persons who have made substantial contributions to the work reported in the manuscript (e.g., technical help, writing and editing assistance, general support), but who do not meet the criteria for authorship, are named in the Acknowledgements and have given us their written permission to be named. If we have not included an Acknowledgements, then that indicates that we have not received substantial contributions from non-authors.

Appendix

Topics and their representative terms are presented.


1
approach
extract
propos
evalu
state
art
base
effect
generat
well
2
topic
model
can
sentiment
detect
latent
stanc
show
propos
distribut
3
classifi
machin
learn
classif
use
accuraci
perform
vector
support
improv
4
web
blog
search
network
semant
forum
present
ontolog
one
thus
5
research
analysi
area
discuss
year
present
recent
survey
field
social
increas
artici
6
data
big
process
social
analyt
intellig
sourc
challeng
generat
volum
7
studi
use
toward
includ
research
conduct
sentiment
sampl
attitud
case
8
make
decis
use
inform
way
can
develop
process
internet
help
9
network
social
communiti
group
cluster
can
one
graph
similar
influenc
10
social
collect
interact
individu
can
behavior
organ
theori
digit
provid
11
featur
detect
set
use
extract
sarcasm
select
accuraci
classif
spam
experi
12
public
issu
onlin
media
monitor
concern
assess
tradit
project
analysi
13
sentiment
messag
microblog
analysi
express
chines
peopl
inform
can
weibo
14
result
algorithm
first
show
compar
similar
two
paper
obtain
problem
15
text
imag
textual
visual
sentiment
multimod
modal
content
analysi
tag
combin
16
user
content
network
generat
social
share
profil
trust
toward
connect
17
emot
affect
express
state
human
detect
correl
person
differ
feel
18
languag
process
text
natur
arab
corpus
resourc
english
differ
task
19
method
base
propos
use
dataset
paper
exist
fuzzi
compar
result
20
post
comment
content
facebook
signific
like
activ
popular
video
mean
term
beta
21
review
onlin
rate
product
tourism
websit
hotel
tourist
number
site
identifi
22
tool
news
develop
applic
softwar
analyz
articl
differ
sourc
technolog
23
data
use
mobil
need
locat
citi
applic
transport
perform
condit
24
tweet
twitter
use
relat
follow
collect
popular
hashtag
number
million
25
inform
subject
howev
analysi
question
contain
analyz
becom
provid
automat
26
system
user
recommend
rank
prefer
propos
show
interest
can
compar
27
measur
base
use
result
metric
compar
set
import
precis
keyword
28
media
social
platform
also
use
analyz
relat
analysi
site
leverag
29
servic
use
qualiti
experi
analysi
provid
improv
evalu
platform
industri
30
domain
train
label
supervis
annot
perform
dataset
classif
task
automat
31
framework
new
propos
base
present
develop
demonstr
paper
three
specif
32
effect
studi
impact
find
influenc
movi
number
social
explor
across
33
market
stock
predict
financi
model
return
firm
compani
price
indic
34
polar
sentiment
context
score
term
text
consid
weight
determin
three
35
negat
posit
brand
sentiment
neutral
show
high
associ
found
audienc
36
model
learn
network
deep
neural
represent
perform
task
term
dataset
37
relat
knowledg
can
base
associ
identifi
rule
help
disast
provid
38
health
onlin
support
use
communiti
patient
result
gender
diseas
control
39
purpos
methodolog
approach
research
find
valu
design
implic
origin
limit
40
sentiment
word
lexicon
specif
construct
dictionari
use
improv
semant
polar



. (*continued*).

References

Abirami, A. M., & Askarunisa, A. (2017). *Sentiment analysis model to emphasize the impact of online reviews in healthcare industry*. Online Information Review.

Abo, M. E. M., Raj, R. G., & Qazi, A. (2019). A review on Arabic sentiment analysis: State-of-the-Art, taxonomy and open research challenges. *IEEE Access, 7*, 162008–162024.

Abulaish, M., & Fazil, M. (2018). Modeling topic evolution in twitter: An embedding-based approach. *IEEE Access, 6*, 64847–64857.

Akter, S., Bhattacharyya, M., Wamba, S. F., & Aditya, S. (2016). How does social media analytics create value? *Journal of Organizational and End User Computing, 28*(3), 1–9.

Alamsyah, A., Rizkika, W., Nugroho, D. D. A., Renaldi, F., & Saadah, S. (2018). Dynamic large scale data on Twitter using sentiment analysis and topic modeling. In *2018 6th international conference on information and communication technology (ICoICT)* (pp. 254–258).

Ali, F., Kwak, D., Khan, P., El-Sappagh, S., Ali, A., Ullah, S., Kim, K. H., & Kwak, K.-S. (2019). Transportation sentiment analysis using word embedding and ontology-based topic modeling. *Knowledge-Based Systems, 174*, 27–42.

Ali, F., Kwak, D., Khan, P., Islam, S. M. R., Kim, K. H., & Kwak, K. S. (2017). Fuzzy ontology-based sentiment analysis of transportation and city feature reviews for safe traveling. *Transportation Research Part C: Emerging Technologies, 77*, 33–48.

Andrews, N. O., & Fox, E. A. (2007). *Recent developments in document clustering*. Department of Computer Science, Virginia Polytechnic Institute & State ...

Bakharia, A., Bruza, P., Watters, J., Narayan, B., & Sitbon, L. (2016). Interactive topic modeling for aiding qualitative content analysis. In *Proceedings of the 2016 ACM on conference on human information interaction and retrieval* (pp. 213–222).

Blei, D., & Lafferty, J. (2006). Correlated topic models. *Advances in Neural Information Processing Systems, 18*, 147.

Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. *Proceedings of the 23rd International Conference on Machine Learning*, 113–120.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*(Jan), 993–1022.

Bouma, G. (2009). Normalized (pointwise) mutual information in collocation extraction. *Proceedings of GSCL*, 31–40.

Cao, J., Zeng, K., Wang, H., Cheng, J., Qiao, F., Wen, D., & Gao, Y. (2013). Web-based traffic sentiment analysis: Methods and applications. *IEEE Transactions on Intelligent Transportation Systems, 15*(2), 844–853.

Carpineto, C., Osiński, S., Romano, G., & Weiss, D. (2009). A survey of web clustering engines. *ACM Computing Surveys, 41*(3), 1–38.

Chen, K., Luo, P., & Wang, H. (2017). An influence framework on product word-of-mouth (WoM) measurement. *Information & Management, 54*(2), 228–240.

Dahal, B., Kumar, S. A. P., & Li, Z. (2019). Topic modeling and sentiment analysis of global climate change tweets. *Social Network Analysis and Mining, 9*(1), 24.

Daud, A., Li, J., Zhou, L., & Muhammad, F. (2010). Knowledge discovery through directed probabilistic topic models: A survey. *Frontiers of Computer Science in China, 4*(2), 280–301.

Diakopoulos, N. A., & Shamma, D. A. (2010). Characterizing debate performance via aggregated twitter sentiment. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1195–1198.

Dueñas-Fernández, R., Velásquez, J. D., & L'Huillier, G. (2014). Detecting trends on the web: A multidisciplinary approach. *Information Fusion, 20*, 129–135.

Eisenstein, J. (2019). *Introduction to natural language processing*. MIT press.

El-Bèze, M., Jackiewicz, A., & Hunston, S. (2010). Opinions, sentiments et jugements d'évaluation. *Traitement Automatique des Langues, 51*(3), 7–17.

Feinerer, I., Hornik, K., & Feinerer, M. I. (2015). Package 'tm. *Corpus, 10*(1).

Fellows, I., Fellows, M. I., Rcpp, L., & Rcpp, L. (2018). *Package 'wordcloud*.

García-Pablos, A., Cuadros, M., & Rigau, G. (2018). W2VLDA: Almost unsupervised system for aspect based sentiment analysis. *Expert Systems with Applications, 91*, 127–137.

Greco, F., & Polli, A. (2020). Emotional text mining: Customer profiling in brand management. *International Journal of Information Management, 51*, Article 101934.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences, 101*(suppl 1), 5228–5235.

Grün, B., Hornik, K., & Grün, M. B. (2020). *Package 'topicmodels*.

Hemmatian, F., & Sohrabi, M. K. (2019). A survey on classification techniques for opinion mining and sentiment analysis. *Artificial Intelligence Review*, 1–51.

Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning, 42*(1–2), 177–196.

Hornik, K., & Grün, B. (2011). topicmodels: An R package for fitting topic models. *Journal of Statistical Software, 40*(13), 1–30.

Hu, M., & Liu, B. (2006). Opinion extraction and summarization on the web. *AAAI, 7*, 1621–1624.

Hu, T., She, B., Duan, L., Yue, H., & Clunis, J. (2019). A systematic spatial and temporal sentiment analysis on geo-tweets. *IEEE Access, 8*, 8658–8667.

Jaidka, K., Ahmed, S., Skoric, M., & Hilbert, M. (2019). Predicting elections from social media: A three-country, three-method comparative study. *Asian Journal of Communication, 29*(3), 252–273.

Jin, J., Liu, Y., Ji, P., & Kwong, C. K. (2019). Review on recent advances in information mining from big consumer opinion data for product design. *Journal of Computing and Information Science in Engineering, 19*(1).

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! the challenges and opportunities of Social Media. *Business Horizons, 53*(1), 59–68.

Keikhosrokiani, P., & Pourya Asl, M. (2022). *Handbook of research on opinion mining and text analytics on literary works and social media*. IGI Global.

Kim, E. H.-J., Jeong, Y. K., Kim, Y., Kang, K. Y., & Song, M. (2016). Topic-based content and sentiment analysis of Ebola virus on Twitter and in the news. *Journal of Information Science, 42*(6), 763–781.

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques*. MIT press.

Koltsova, O., & Koltcov, S. (2013). Mapping the public agenda with topic modeling: The case of the Russian LiveJournal. *Policy & Internet, 5*(2), 207–227.

Kumar, A., & Garg, G. (2019). Systematic literature review on context-based sentiment analysis in social multimedia. *Multimedia Tools and Applications*, 1–32.
Kwon, H.-J., Ban, H.-J., Jun, J.-K., & Kim, H.-S. (2021). Topic modeling and sentiment analysis of online review for airlines. *Information, 12*(2), 78.
Lau, J. H., Newman, D., & Baldwin, T. (2014). Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality. In *Proceedings of the 14th conference of the European chapter of the association for computational linguistics* (pp. 530–539).
Lee, I. (2018). Social media analytics for enterprises: Typology, methods, and processes. *Business Horizons, 61*(2), 199–210.
Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies, 5*(1), 1–167.
Liu, X. (2020). Analyzing the impact of user-generated content on B2B Firms' stock performance: Big data analysis with machine learning methods. *Industrial Marketing Management, 86*, 30–39.
Liu, X., Burns, A. C., & Hou, Y. (2017). An investigation of brand-related user-generated content on Twitter. *Journal of Advertising, 46*(2), 236–247.
Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—a review of research topics, venues, and top cited papers. *Computer Science Review, 27*, 16–32.
Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. In *Proceedings of the 2011 conference on empirical methods in Natural Language processing* (pp. 262–272).
Ng, C. Y., & Law, K. M. Y. (2020). Investigating consumer preferences on product designs by analyzing opinions from social networks using evidential reasoning. *Computers & Industrial Engineering, 139*, Article 106180.
Nimala, K., & Jebakumar, R. (2019). Sentiment topic emotion model on students feedback for educational benefits and practices. *Behaviour & Information Technology*, 1–9.
Ozyurt, B., & Akcayol, M. A. (2021). A new topic modeling based approach for aspect extraction in aspect based sentiment analysis: SS-LDA. *Expert Systems with Applications, 168*, Article 114231.
Pang, B., & Lee, L. (2009). Opinion mining and sentiment analysis. *Computational Linguistics, 35*(2), 311–312.
Ponweiser, M. (2012). *Latent dirichlet allocation in R*.
Qiao, F., & Williams, J. (2022). Topic modelling and sentiment analysis of global warming tweets: Evidence from big data analysis. *Journal of Organizational and End User Computing, 34*(3), 1–18.
Rana, T. A., Cheah, Y.-N., & Letchmunan, S. (2016). Topic modeling in sentiment analysis: A systematic review. *Journal of ICT Research and Applications, 10*(1), 76–93.
Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems, 89*, 14–46.
Rodrigues, R. G., das Dores, R. M., Camilo-Junior, C. G., & Rosa, T. C. (2016). SentiHealth-cancer: A sentiment analysis tool to help detecting mood of patients in online social networks. *International Journal of Medical Informatics, 85*(1), 80–95.
Rouhani, S., & Abedin, E. (2019). Crypto-currencies narrated on tweets: A sentiment analysis approach. *International Journal of Ethics and Systems, 36*(1), 58–72.
Selivanov, D., & Wang, Q. (2016). text2vec: Modern text mining framework for r. *Computer Software Manual*. R Package Version 0.4. 0). Retrieved from Https://CRAN. R-Project. Org/Package= Text2vec.
Singh, J., & Gupta, V. (2016). Text stemming: Approaches, applications, and challenges. *ACM Computing Surveys, 49*(3), 1–46.
Singh, N. K., Tomar, D. S., & Sangaiah, A. K. (2020). Sentiment analysis: A review and comparative analysis over social media. *Journal of Ambient Intelligence and Humanized Computing, 11*(1), 97–117.
Smailović, J., Grčar, M., Lavrač, N., & Žnidaršič, M. (2013). Predictive sentiment analysis of tweets: A stock market application. In *International workshop on human-computer interaction and knowledge discovery in complex, unstructured, big data* (pp. 77–88).
Sommer, S., Schieber, A., Heinrich, K., & Hilbert, A. (2012). What is the conversation about?: A topic-model-based approach for analyzing customer sentiments in twitter. *International Journal of Intelligent Information Technologies, 8*(1), 10–25.
Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. *Handbook of Latent Semantic Analysis, 427*(7), 424–440.
Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2006). Hierarchical dirichlet processes. *Journal of the American Statistical Association, 101*(476), 1566–1581.
Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology, 63*(1), 163–173.
Vanhala, M., Lu, C., Peltonen, J., Sundqvist, S., Nummenmaa, J., & Järvelin, K. (2020). The usage of large data sets in online consumer behaviour: A bibliometric and computational text-mining–driven analysis of previous research. *Journal of Business Research, 106*, 46–59.
Wang, R., Zhou, D., Jiang, M., Si, J., & Yang, Y. (2019). A survey on opinion mining: From stance to product aspect. *IEEE Access, 7*, 41101–41124.
Wu, Q., Sun, Y., Yan, H., & Wu, X. (2020). ECG signal classification with binarized convolutional neural network. *Computers in Biology and Medicine, 121*. https://doi.org/10.1016/j.compbiomed.2020.103800
Yue, L., Chen, W., Li, X., Zuo, W., & Yin, M. (2019). A survey of sentiment analysis in social media. *Knowledge and Information Systems*, 1–47.