The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1463-5771.htm

Text classification using deep learning techniques: a bibliometric analysis and future research directions

Gaurav Sarin
Delhi School of Business, Delhi, India
Pradeep Kumar
Indian Institute of Management Lucknow, Lucknow, India, and
M. Mukund
VDX.tv, New Delhi, India

Received 14 July 2022
Revised 2 February 2023
18 March 2023
31 March 2023
22 April 2023
27 June 2023
Accepted 8 July 2023

Abstract

Purpose – Text classification is a widely accepted and adopted technique in organizations to mine and analyze unstructured and semi-structured data. With advancement of technological computing, deep learning has become more popular among academicians and professionals to perform mining and analytical operations. In this work, the authors study the research carried out in field of text classification using deep learning techniques to identify gaps and opportunities for doing research.

Design/methodology/approach – The authors adopted bibliometric-based approach in conjunction with visualization techniques to uncover new insights and findings. The authors collected data of two decades from Scopus global database to perform this study. The authors discuss business applications of deep learning techniques for text classification.

Findings – The study provides overview of various publication sources in field of text classification and deep learning together. The study also presents list of prominent authors and their countries working in this field. The authors also presented list of most cited articles based on citations and country of research. Various visualization techniques such as word cloud, network diagram and thematic map were used to identify collaboration network.

Originality/value – The study performed in this paper helped to understand research gaps that is original contribution to body of literature. To best of the authors' knowledge, in-depth study in the field of text classification and deep learning has not been performed in detail. The study provides high value to scholars and professionals by providing them opportunities of research in this area.

Keywords Data mining, Text analytics, Classification, Deep learning, Bibliometric analysis

Paper type Research paper

1. Introduction

Text classification (TC) is the process of dividing a specific text into organized groups from an unstructured data set by assigning labels to various text units. It is one of the classical problems in the natural language processing (NLP) domain. Based on the content, text classifiers use NLP to analyze the text automatically and subsequently determine a suitable set of predefined labels. TC has become a crucial part of businesses due to its ability to get remarkable insight from unstructured data and gradually assist businesses in making rational and comprehensive decisions about their future strategies for a product or a service. Some of the most prominent use cases of TC include sentiment analysis, which is a handy tool for understanding the polarity of a text. It is helpful to businesses in familiarizing themselves with the perception of customers about their brand or a specific product. Language detection is another application of TC that helps in detecting the language of a given text. Businesses that have a user base across the globe use language detection to help them with their requests.

Benchmarking: An International Journal

1463-5771
DOI 10.1108/BIJ-07-2022-0454

Another well-known application of TC is topic detection which helps in identifying the theme of a text. E-commerce websites with the facility for users to connect with the brand online use topic detection to understand if a query or feedback is related to one of their products or customer support. By assigning categories to each document, all these systems seek to arrange unstructured textual data. It is instrumental when thorough human coding is not possible owing to data volume or when categorical information is needed immediately (Hartmann *et al.*, 2019).

There lies an evident gap in the existing research on TC. This gap acts as a boon for the business community from a multi-domain application point of view, as only a handful of studies have discussed it. Machine learning tools for automatic TC have recently evolved by overtaking the knowledge engineering methodology. These technologies have a vast array of applications in the business industry, such as digital advertising, healthcare, education, e-commerce, etc. This article attempts to fill the gap by reviewing the various deep learning (DL) techniques used in TC, especially those which focus on their applications in business.

As per Figure 1, Statista's report on in-depth: Artificial Intelligence (AI), 2021: NLP is a branch of Artificial Intelligence that facilitates the interpretation and manipulation of human language using computer algorithms. It started to gain widespread recognition as early as the 1950s, as at that time, linguistic experts began exploring the usage of machines to automate the translation of languages. Today, multiple DL methods have developed to imitate the function of neurons in the human brain and act as a significant driver of advancement in NLP. DL and NLP have progressively retained from examples and their experiences, an attribute which is extremely useful when applied to sectors such as healthcare, consumer goods, defense, etc. In their study, Zhao *et al.* (2020) present a model to classify financial news that achieved an accuracy of 96.45% using 65,000 Chinese financial news. Other typical applications of translation products such as Google Translate, spell check features on Grammarly, MS Word, IVR used by banks, and AI assistants like ALexa, Siri, and Cortana that are extensively used by masses in their day-to-day lives. Some of the primary applications of NLP are stated as follows:

(1) Speech-to-text (Speech-Recognition): This technique involves the conversion of spoken language into text, which can be processed by other applications.

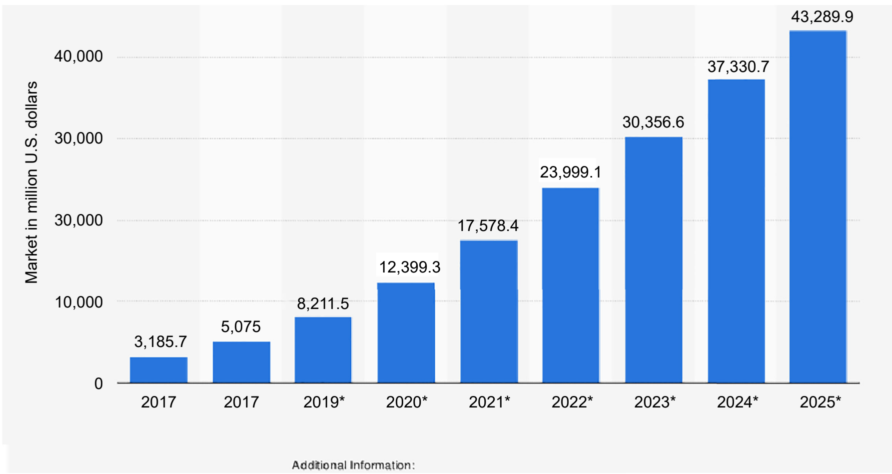


Figure 1.
NLP worldwide market

Source(s): Statista (2021)

(2) Text-to-speech (Speech-Synthesis): This involves the conversion of any text to related speech.

(3) Text processing: This technique involves making sense of the text and extracting information that can be used to derive meaningful insights.

(4) Natural language generation (NLG): This technique differs from text processing as it does not need much human intervention. In this, insights are directly inferred from large dataset.

(5) Chatbots: Bots driven by NLP have the capability to understand wide variety of human communication. Even though they cannot grasp the subtleties or nuances of human language, they can still be programmed while responding to specific questions.

(6) Machine translation: This includes the automatic translation of text from one language to another. This technique has advanced from being solely reliant on predefined rules to using complex statistical models and, most recently, using neural networks to mimic human-like thinking.

This study contributes to the field of TC research by first employing a bibliometric analysis to identify significant publications, doing a year-by-year analysis, topic trends analysis, and creating a thematic map. We go into detail on the TC tasks and applications in addition to the bibliometric study, which offers descriptive data about the literature. For our research, bibliometric analysis has become an essential research method because it helps in systematically analyzing the past work done in the field and summarizing the key insights in a single paper using graphs. High-quality publications from all fields, including business management and economics, have published bibliometric studies that have had a significant impact (Mukherjee *et al.*, 2022).

This leads to our first research question stated as:

RQ1. What are the areas of research in the field of Text Classification?

Although bibliometric reviews often do not capture the attention of young scholars because they focus primarily on citations and do not review specific theories, methods, and constructs (Paul *et al.*, 2021), we find this method suitable for our research because it is catered to academicians and industry professionals who can benefit from the insights drawn. Furthermore, by rigorously making sense of vast amounts of unstructured data, bibliometric analysis is helpful for unravelling and charting the cumulative scientific knowledge and evolutionary nuances of well-established domains (Donthu *et al.*, 2021).

This leads to our second research question:

RQ2. How research in these areas help academicians and industry professionals to solve business problems?

The major contributions can be summarized as:

(1) We delve into an in-depth analysis of TC processes through an exhaustive literature review of research papers published in the last two decades, from 2002 to 2022. Moreover, this paper aims to provide a detailed descriptive analysis regarding multiple facets, such as some of the most cited papers, authors, countries of origin, etc.

(2) We discuss the business applications of the DL techniques for TC to help industry practitioners build a successful strategy that includes a solid understanding of the different use cases and how the DL techniques can help different businesses.

(3) We acquaint business leaders and decision-makers with a thorough overview of DL techniques utilized for TC across different business applications.

This paper is organized into six sections. Section 2 states the latest research in study. Section 2 is further subdivided into two subsections on Text Classification and Deep Learning. In Section 3, the authors describe the methods and data used in current research. Section 4 presents the results and discussions. Section 5 presents Implications both theoretical and managerial. Finally, we conclude in Section 6 with limitations.

2. Literature review

Text classification have been extensively studied by researcher and practitioners (Luo and Zhang, 2022; Gupta and Lehal, 2009). Exploring the various applications and understanding the opportunities with textual data is always of major research objectives among the researchers. IBM report says that around 85% of available data is present in unstructured form including text, image, video and other. With the advancement in computing power deep learning techniques are extensively used to analyze data of all forms (Zhang *et al.*, 2019a, b). Text classification with deep learning techniques are nowadays attracting researcher of this field a lot. In this section, we present the literature Review related to the field of text classification and deep learning using bibliometric analysis.

2.1 Text classification

TC has been extensively studied in multiple studies. The specific set of classes selected by TC algorithms should be assigned to the text (Genkin *et al.*, 2007). Text classification research encompasses everything ranging from creating the best features to selecting the best machine learning classifiers (Shuang *et al.*, 2019). This problem closely relates to another problem with a pervasive range of applications, i.e., object classification. Due to the complex background surrounding the text and the variability of text features, this task of object classification from real scene photos is the most difficult (Khan and Mollah, 2019). Due to expansion in corpus sizes along with the number of fields growing as new horizons get explored (sub-fields within them), TC has become increasingly arduous over the past several years. The sub-fields continue to grow in number across a myriad of disciplines, including biology (e.g., material science, and health sciences, which were lesser-known areas of research only a couple of years ago). However, only a handful of studies provide a comprehensive overview of TC and DL that can be referred to by industry practitioners, which becomes essential to highlight given its increasing role in business applications. DL is one of the most popular ML techniques in these applications. Using unlabeled data sets has proven to be very effective for pre-training of language models to learn universal language representations (Moirangthem and Lee, 2021). Using unlabeled data to improve the performance of the classifier has recently emerged as an exciting topic in data mining (Zhang *et al.*, 2015). Organizations today produce and consume this unlabeled, unorganized and non-sequential data in copious amounts and on regular bases, and therefore, through our review, we have studied papers that categorically talk about DL techniques and their applications in TC. In the recent decade, online pages containing dubious and hazardous material, such as fraud, phishing, violence, extremism, pornography, and so on, have been multiplying. Qi and Davison (2009) discussed a similar problem. In their research, they suggest that the web page classification problem can be classified into multiple sub-problems, i.e., functional classification (concerned about the role of web page), subject classification (concerned about the subject/topic), sentiment classification (focuses on the opinion) and other types of classification. Blogging has become another popular medium of communication where we often see new abbreviations, slang, etc., added daily. To implement TC applications such as opinion mining or sentiment classification, it is required to keep track of these newly emerging terms (Dalal and Zaveri, 2011). TC is one of the most important techniques for organizing online information as it has been widely used in news categorization, opinion mining, and spam filtering. Therefore, TC has become an essential tool in the toolbox for industry practitioners who want to comprehend the complex and unstructured availability of data. For example, M. Ghiassi *et al.* (2013) discern that the complexity of language found on Twitter has always hindered the ability of researchers to determine the sentiment from the text. Many online pages covering a wide range of topics are becoming a hindrance in the process of delivering optimal topic-relevant results while utilizing information retrieval and extraction algorithms (Hashemi, 2020). In such cases, short text classification can help organize a jumble of data as well as provide more effective search tactics and search results for information retrieval (Yang *et al.*, 2021). One of the most prominent and mandatory phases in knowledge discovery is TC. Their study emphasizes techniques based on supervised learning methods as compared to semi-supervised and unsupervised learning.

2.2 Deep learning

Studies suggest that although different techniques perform in different ways depending on the textual data and their size, it was observed that K-NN (supervised learning method), with TF-IDF term weighting representation scheme, performs well in several TC algorithms. TF-IDF, or the term frequency-inverse document frequency, is a popular measure of significance for a word in a document. Almost all the papers we reviewed discuss and apply this measure in their research. However, despite being widely used, it has its own limitations. Yao *et al.* (2018) discuss some of its limitations in their work which are: (a) the same word has the same weight in each category due to TF-IDF's inability to incorporate class-based weights in its calculation, and (b) TF-IDF is unable to deal with synonyms and words with multiple meanings. The need of the industry reaches far beyond this rudimentary task, which can be already daunting depending on the domain, as some TC tasks may include multiple classes. In their work, they explain that the machine learning constructs such as Artificial Neural Networks (ANN) have demonstrated applicability in a myriad of domains while creating new performance records. Through their work, they elucidate the applications of NLP using DL and call out areas such as event extraction, relationship extraction, TC, text generation (poetry generation, joke and pun generation, story generation), text generation with generative adversarial networks (GANs), text generation with variational autoencoders (VAEs), summarization, question answering and, machine translation. Furthermore, they discuss that deep neural architectures have engendered high-performance models in natural language tasks. They also suggest that with pre-training and transfer learning playing crucial roles, DL techniques will become the standard norm in computational linguistics. Yao *et al.* (2019b) present a novel initialization to add label co-occurrence into NN models to investigate label correlations for NN models efficaciously. Wang *et al.* (2018) propose an attention-based framework to measure the compatibility of word embeddings between sequences of text and labels. Word embeddings can be pre-trained word structures generated offline by scanning huge volumes of textual data, or they can be initialized randomly and allowed to grow together with the remaining variables (Moreo *et al.*, 2021). Colace *et al.* (2014), in their research, exemplifies that a complex vector of features, subjected to weighted pairs of words, is effective in overpowering the drawbacks of basic structures when there is a small number of labeled samples. Liu *et al.* (2022) propose a short new TC method that combines contextual relevance and a multi-stage attention model on CNN. To improve the level of precision of the DL models, researchers often introduce disturbances in the input text known as adversarial examples. In the NLP domain, generating semantically and syntactically similar adversarial texts in the discrete input space of word symbols has proven to be more difficult as compared to images (Xu and Du, 2020a, b). Adversarial models have recently gained substantial attention. Grießhaber *et al.* (2020) investigate the use of domain-adversarial learning as a regularizer to avoid overfitting while training domain invariant features for DNN in low-resource and zero-resource scenarios. Deep neural models require a large amount of data to learn complex characteristics, and the quality of the data has a significant impact on the models' performance (Chen and Dai, 2021). When it comes to DNN, Lee *et al.* (2018) highlights two approaches to construct document-level representations, i.e., (1) context-based learning and (2) composition-based learning. Text classification models (CNN, SVM and RNN) may cluster vast construction text as per the predetermined topic, which subsequently increases information extraction and handling effectiveness, but a thorough understanding of text information still cannot be acquired (Tian *et al.*, 2021).

In the contemporary era, the profound availability of data at varying levels of complexity from multiple domains has led researchers to write about the growing field. This copious and complex information needs proficient classification algorithms that can be used to assign texts to one or more labels (Zulqarnain *et al.*, 2021). The semantic organization of available documents into their corresponding categories is a major goal of TC systems (Van Linh *et al.*, 2017). TC can filter English information ahead of time, classify the collected English information, and build a categorical library of an article, making it easier for the user to browse category material of interest and improve the user experience (Liu and Wang, 2020). One of the challenges has been explored by Yang *et al.* (2011), where they build text classifiers using keywords and unlabeled documents to classify text streams and use classifier ensemble algorithms. Furthermore, to deal with such complexity, an effective subfield within ML is ANN which enables the creation of learning models with a multi-level decision-making architecture. They are based on a supervised learning process and can be classified into three layers: input, hidden, and output. It is important to accentuate that DL and Neural Networks (NNs) are often perceived to be synonymous with each other. This leads to an ambiguity in understanding the difference between the two. A NN consisting of input, hidden, and output layers is the most basic type of its kind. Therefore, a DL technique is visualized to be a NN built of more than three layers mentioned above. When it comes to TC problems, researchers also use an ensemble of algorithms instead of a standalone algorithm to achieve higher accuracy. The ensemble techniques have become an effective classification method for numerous areas of application. For example, researchers look forward to combining ConvNets and RNNs into a hybrid structure to get the best of both worlds and complement the strengths of different neural architectures (Liu *et al.*, 2020). In topic classification, several researchers have achieved improvements in the accuracy of classification using the ensemble technique. The topic classification of news stories enhances their denomination to offer suitable information and improve the browsing experience of users (Zheng and Zhen, 2019). In their paper, Liang and Yi (2021) work to overcome the time-consuming and error-prone challenge of painstakingly extracting crucial information from policy texts. They suggest a two-stage-based, three-way technique to classify these texts into predefined categories automatically. Using ensemble learning algorithms in the first stage, they construct an ensemble CNN model to ensure the generalization ability and stability of text classification results. Subsequently, to improve the classification results in the second stage, they further utilize traditional ML methods as the secondary classifier. The results of their study show that the latter classification framework obtains a better performance. Hybrid approaches seek to exploit the strengths of the individual components, obtaining enhanced performance through their combination. Similarly, Jang *et al.* (2020) propose a hybrid model of attention-based Bi-LSTM and CNN that takes advantage of LSTM and CNN with additional attention mechanisms. They use the IMDB dataset for testing their model, and the results show that the hybrid model can attain higher classification accuracy along with F1 scores as compared to standalone versions of MLP, CNN or LSTM. Huang *et al.* (2021) investigate the semantic relationship between each document and extreme labels. Hajiabadi *et al.* (2020) call out the difference between two ensemble methods: dependent and independent ensemble methods. The output of every base classifier influences the learning of the following base classifiers in a dependent ensemble classifier. Conversely, each base classifier is trained individually in the independent ensemble techniques, and the outputs are then concatenated for the final decision. One of the objectives we fulfil through our study is elaborating upon the applications of TC in real life, especially in business. Text categorization is intrinsically tied to text representation approaches as an NLP problem. Sentiment analysis, which we discuss in detail in the following sections as one of the TC tasks, has been extensively written about. We also came across stance detection that has been described in different ways in diverse application setups. A widely accepted definition of stance detection is the "Automatic classification of the producer of a piece of text's stance towards a target into one of these three classes: Favour, Against, Neither" (Küçük and Can, 2020). The Internet's rapid expansion makes it challenging to identify web pages that contain relevant and valuable information by filtering away extraneous content.

Based on the literature review performed by us, we identify the following research gaps:

(1) Although the field is exploited yet, very little or no study has been performed to establish the maturity of NLP research based on the meta-information of the articles.

(2) We also identified that the research in this field is heavily done by various domain areas, but people have yet to focus on the prominent themes for management research.

(3) Scattered application research in multiple business areas has been performed by many researchers; however, there is a lack of literature that consolidates all the applications from the management area in one place.

This paper attempts to fill the mentioned research gaps by using the bibliometric analysis research methodology.

3. Methods and data

Bibliometric methods originate from research in the library and information sciences that involve large volume of bibliographic materials (Broadus, 1987). Kessler (1963) elucidated that scientific works exhibit intellectual convergence based on their common sources and patterns of referencing. Apart of this many other well excepted measures include co-authorship and co-occurrence. Co-authorship reveals the authorship pattern and connectivity among the collaborating authors, co-occurrence of keywords.

We accessed bibliographic data used in this study from the Scopus database, the largest multi-disciplinary database of peer-reviewed literature in social science research (Bartol *et al.*, 2014). Scopus is widely recognized and frequently accessed for quantitative analyses (Durán-Sánchez *et al.*, 2019). The advanced query written to extract the relevant papers from the Scopus database is given as:

["text classification" AND ("deep learning" OR "neural network") (topic) OR "text classification" AND ("deep learning" OR "neural network") (abstract) OR "text classification" AND ("deep learning" OR "neural network") (author AND keywords)].

On executing the above query over the Scopus database returned 309 papers.

Query 2 with date as filter were executed to keep two decades of publication. To refine our search, we limited our search over two decades starting from 1/1/2002 and published on or before 31/12/2022. Query 2 is given below for the sake of reference.

["text classification" AND ("deep learning" OR "neural network") (topic) OR "text classification" AND ("deep learning" OR "neural network") (abstract) OR "text classification" AND ("deep learning" OR "neural network") (author AND keywords) AND PUBYEAR > 2001 AND PUBYEAR < 2023].

On executing query 2 we got 276 papers. To further refine the search, we limited our study for papers published in academic journals. Thus, the query 2 was refined to query 3 including only articles. Query 3 is given below for the reference. Query 3 resulted into 164 papers.

["text classification" AND ("deep learning" OR "neural network") (topic) OR "text classification" AND ("deep learning" OR "neural network") (abstract) OR "text classification" AND ("deep learning" OR "neural network") (author AND keywords) AND PUBYEAR > 2001 AND PUBYEAR < 2023 AND (LIMIT-TO (DOCTYPE, "ar"))].

Authors then focussed on journal entries in the Scopus database using the Query 4, which is stated as:

["text classification" AND ("deep learning" OR "neural network") (topic) OR "text classification" AND ("deep learning" OR "neural network") (abstract) OR "text classification" AND ("deep learning" OR "neural network") (author AND keywords) AND PUBYEAR > 2001 AND PUBYEAR < 2023 AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (SRCTYPE, "j"))].

Query 4 resulted in 162 articles.

Most of the researchers and practitioners are comfortable with English language. Hence we further filtered the data with respect to language as English. After adding the language as English our query 5 resulted into 156 articles. The query 5 is stated as:

["text classification" AND ("deep learning" OR "neural network") (topic) OR "text classification" AND ("deep learning" OR "neural network") (abstract) OR "text classification" AND ("deep learning" OR "neural network") (author AND keywords) AND PUBYEAR > 2001 AND PUBYEAR < 2023 AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (SRCTYPE, "j")) AND (LIMIT-TO (LANGUAGE, "English"))].

These 156 were manually analyzed by reading abstract only. On manual evaluation, we got 118 papers on which we performed the detailed bibliometric analysis.

Figure 2 outlines the steps in the bibliometric analysis process that the authors followed.

Figure 3 shows the different queries that authors ran on the Scopus global research database.

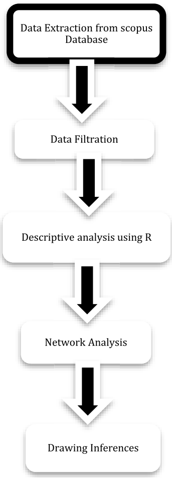
4. Results and discussion

This section talks in-depth about the results and discussion.

Figure 4 shows the number of top journal publications. The x-axis shows the number of documents. The y-axis shows the publication sources. The maximum number of documents come from IEEE Access with count of 59. The minimum number of documents come from PLOS One with count of 4.

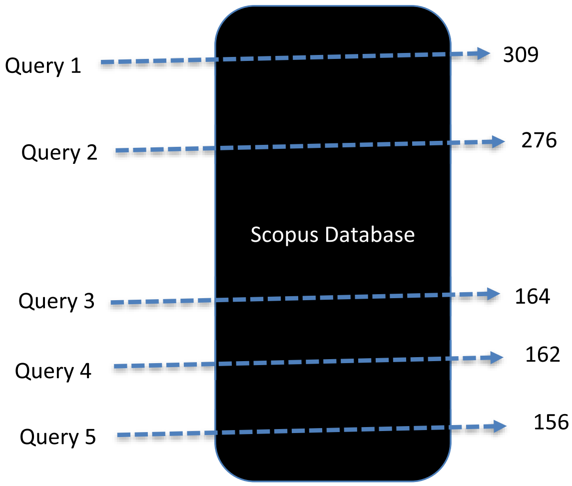
Figure 5 displays the top authors. The x-axis shows the number of documents. The y-axis shows the names of the authors. The maximum number of documents come from Zhang Y with a count of 9. The minimum number of documents come from Doherty J, Choi GS, Chen Z and Chen J with count of 3.

Figure 6 shows the year-wise topic trends. The x-axis shows the year. The y-axis shows the terms. The term "model" is trending the maximum during the period 2020–2023. The term "classification" is trending the maximum during the period 2019–2021. The term "information" is trending the maximum during the period 2019–2020. The term "social media" is trending the minimum during the period 2021–2023. The term "network" is



Source(s): Authors' own work

Figure 2.
Steps in bibliometric analysis



Source(s): Authors' own work

Figure 3.
Queries on Scopus database

trending the minimum during the period 2020–2021. The term "localization" is trending the minimum during the period 2017–2019.

Figure 7 displays the average citations per year country-wise. The x-axis shows the country of origin for the paper. The y-axis shows the citation count. The top two countries that have the high average citations per year are China and USA, while the bottom two countries that have low average citations per year are Saudi Arabia and Singapore.

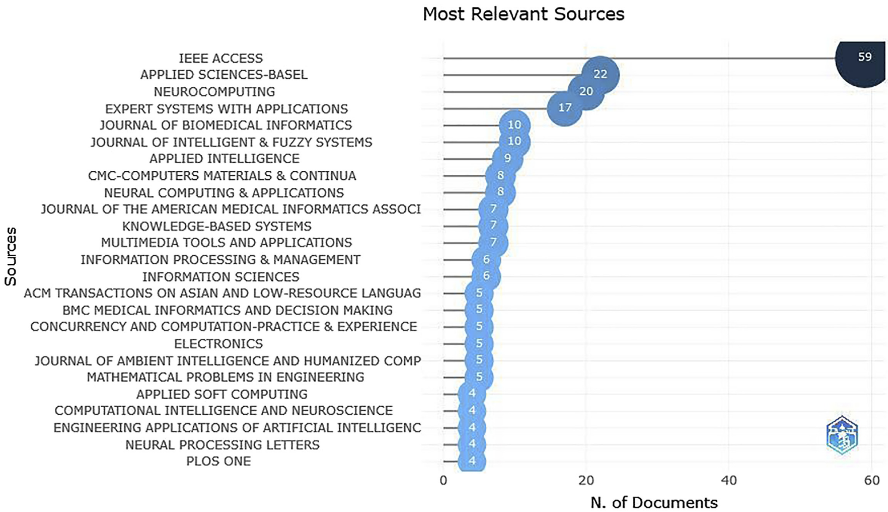


Figure 4.
Most relevant journals

Source(s): Authors' own work

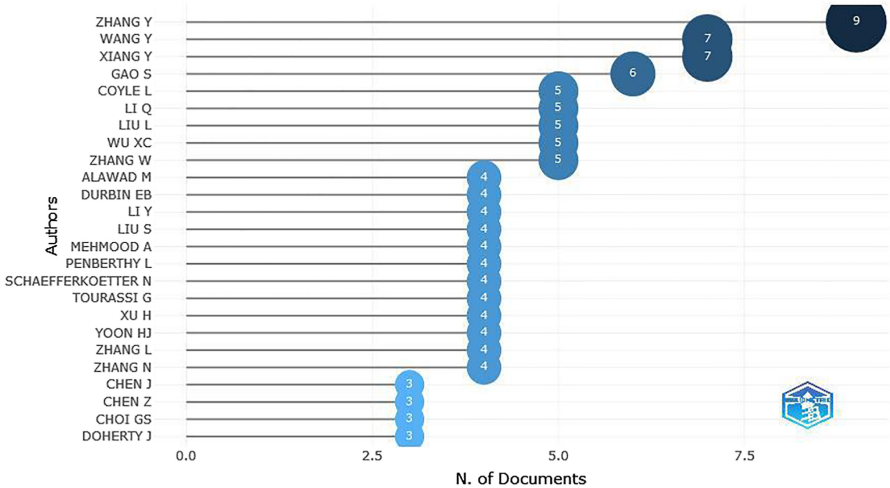
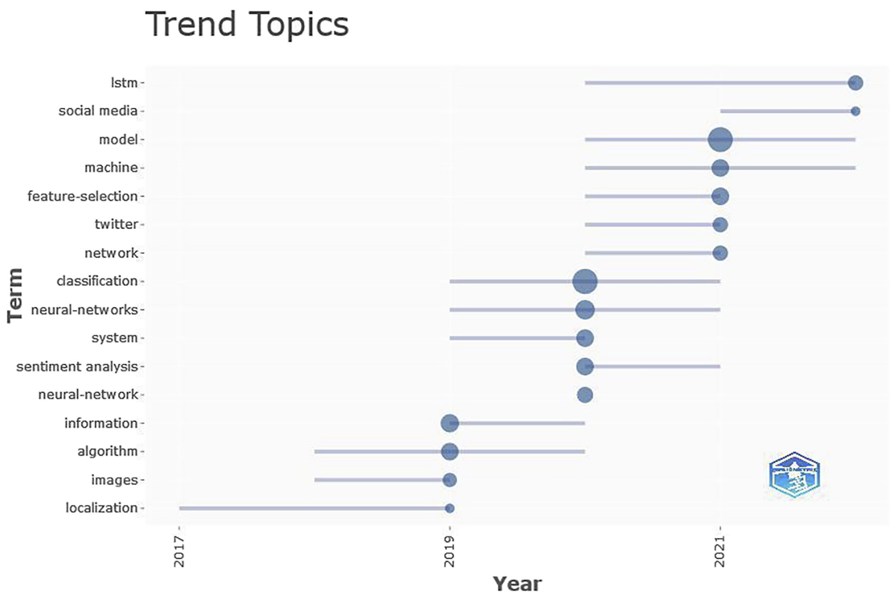


Figure 5.
Most relevant authors

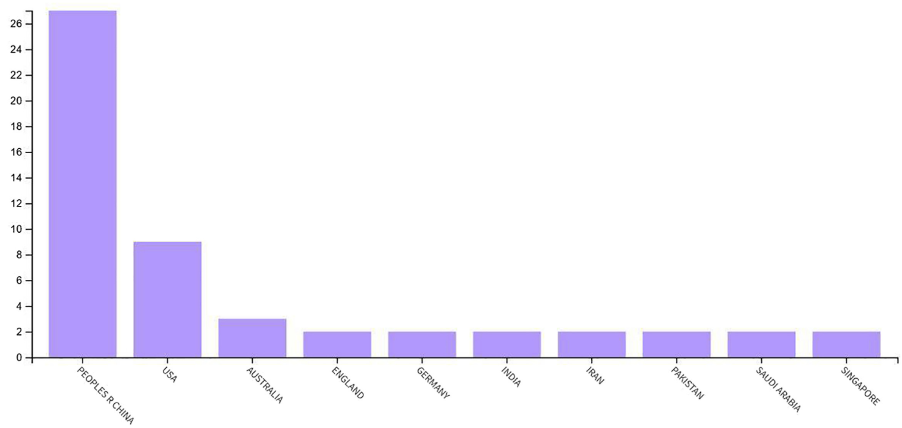
Source(s): Authors' own work

Figure 8 shows the top 20 globally cited articles. The x-axis shows the number of global citations. The y-axis shows the authors name with publication source. The maximum number of citations come from Lu *et al.* (2015), Knowledge Based Systems (ABDC, A category journal) having count of 395. The minimum number of citations come from Mitra (2007), Applied Soft Computing (ABDC, C category journal) having count of 64.



Source(s): Authors' own work

Figure 6.
Year-wise topic trends



Source(s): Authors' own work

Figure 7.
Average citations per year country wise

Figure 9 shows the top20 title wise word cloud. By looking at the word cloud, the authors inferred that the words – "model," "neural-networks," "information," "feature-selection" and "algorithm" appear the highest number of times due to their size and dimension.

Figure 10 shows the network visualization diagram created using the R bibliometrix package. There are number of interconnections between the terms – "text classification," "neural-network," "neural-networks," "model," "feature-selection" etc. These interconnections are between the nodes that are terms. Term bubble size indicates the number of times the term occurs in the articles. The edges indicate the terms linkages with other terms.

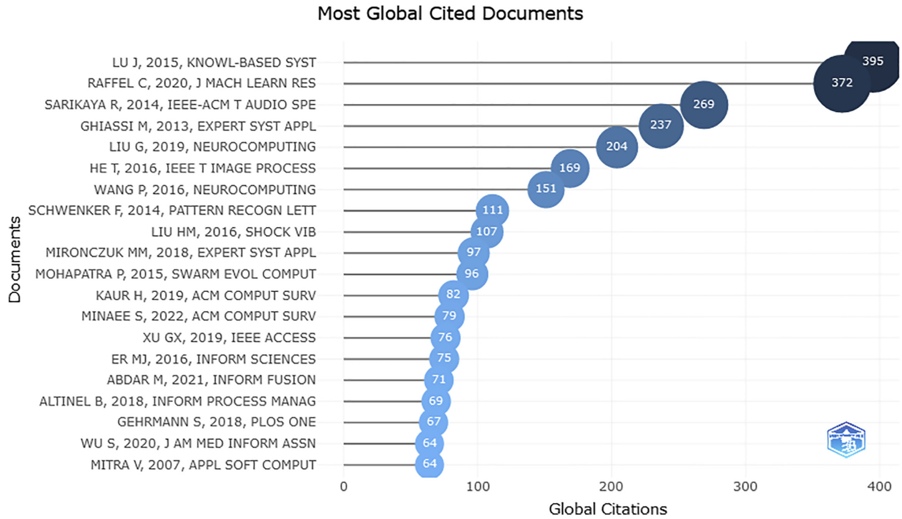


Figure 8.
Top20 globally cited articles

Source(s): Authors' own work

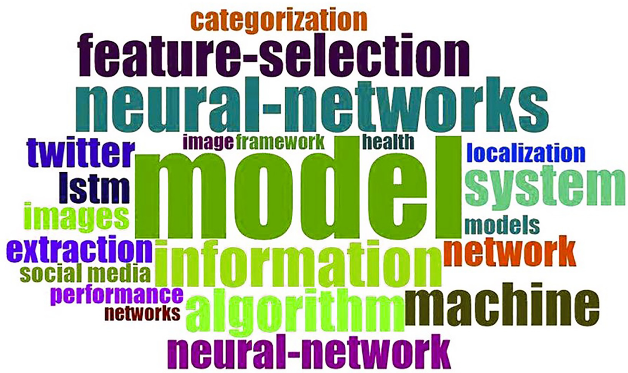


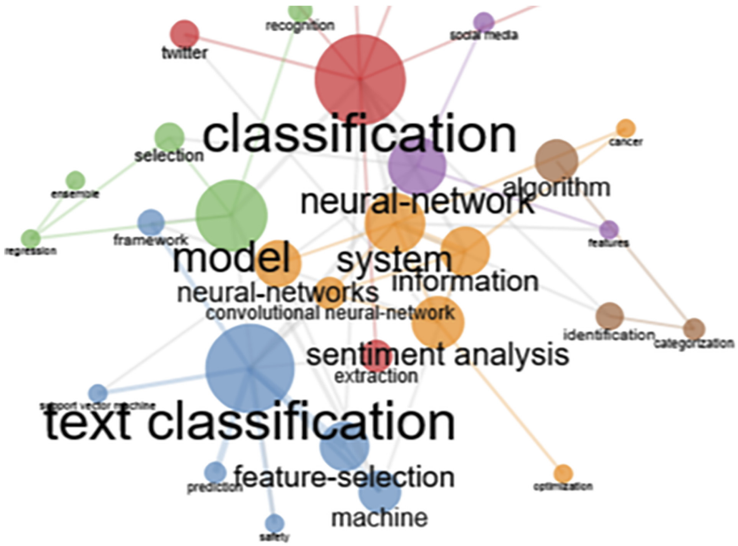
Figure 9.
Top20 word cloud title wise

Source(s): Authors' own work

Figure 11 shows the keyword-wise thematic map. The x-axis shows the degree of centrality in increasing order. The y-axis shows the development degree in increasing order. There are 4 quadrants in this map. The 1st quadrant shows the motor themes. The 2nd quadrant shows the basic themes. The 3rd quadrant shows the niche themes, and the 4th quadrant shows the emerging or declining themes.

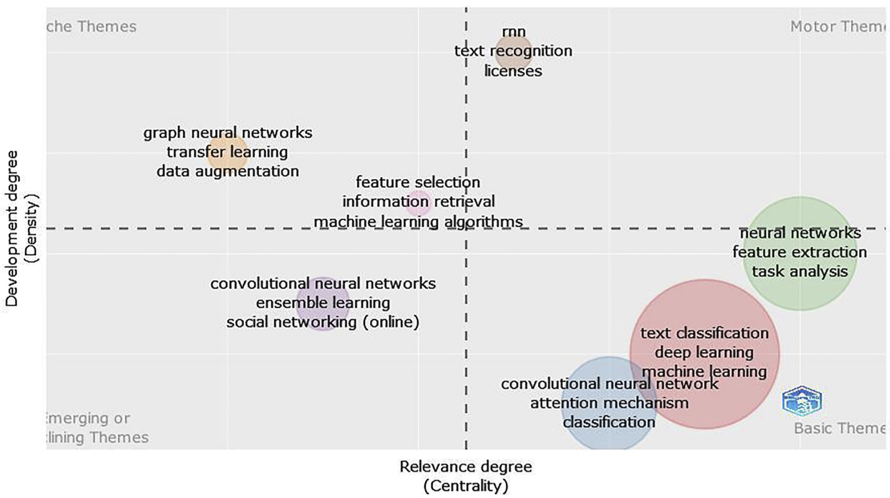
Figure 12 displays the country collaboration map. The x-axis shows the latitude, and the y-axis shows the longitude. This map is created using the R bibliometrix package. It depicts at the collaborations between different countries at a macro level.

Figure 13 shows the country-wise network map. There are number of interconnections between the countries that are nodes. These interconnections showcase the collaborations between the authors of different countries for papers on TC and DL at a micro level. The bigger the size of the bubble indicates the maximum number of contributions from that country. For example, China has the highest contribution to the TC and DL papers.



Source(s): Authors' own work

Figure 10.
Network diagram



Source(s): Authors' own work

Figure 11.
Thematic map keyword wise

Our analysis of papers written in the field of Text Classification (TC) and Deep Learning (DL) in the past 20 years showed the variety of applications such as analysis of medical texts, financial news, Google Translate, spell check features on Grammarly, MS Word, IVR used by banks, and AI assistants such as ALexa, Siri, and Cortana. Past literature shows that the top five relevant journals in this field are namely – IEEE Access, Applied Sciences, Neurocomputing, Expert Systems with Application, and Journal of Biomedical Informatics. One of the key findings from our examination of the literature is that there has been a

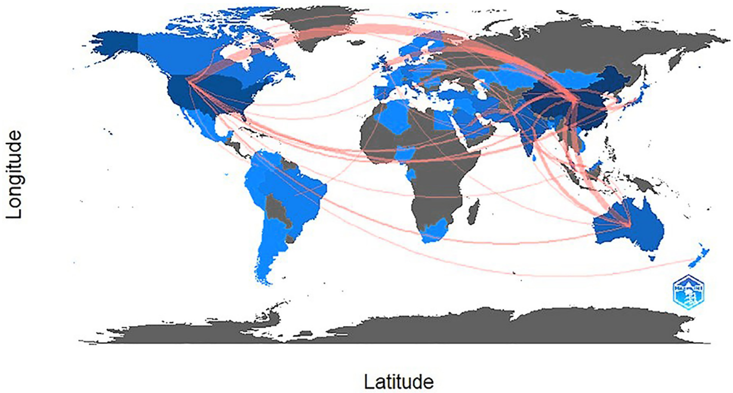


Figure 12.
Country collaboration map

Source(s): Author's own work

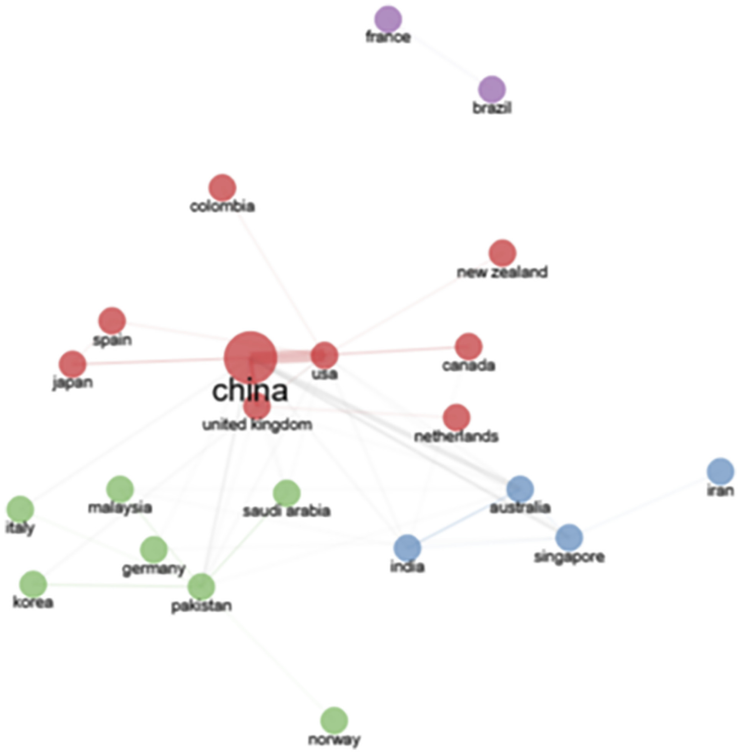


Figure 13.
Country wise network map

Source(s): Authors' own work

significant emphasis on hybrid approaches as they can incorporate the strengths of multiple techniques for accurate outcomes. For example, one of the analyses done on the IMDB dataset showed that hybrid models could attain higher classification accuracy along with F1 scores compared to standalone versions of MLP, CNN, or LSTM. This is not just consistent with our findings but also has a major implication on industry applications as hybrid models can help organizations that focus on building and optimizing language models for dialogue. The top five relevant authors are – Zhang Y, Wang Y, Xiang Y, Gao S, and Coyle L. Terms, namely – model, classification, neural networks, feature-selection, information, etc., are used mostly in research papers surveyed between 2017 and 2022. The top five countries based on the average citations per year are namely – the Peoples' Republic of China, the USA, Australia, England, and Germany. The top five research papers are namely – Lu *et al.* (2015), Knowledge-Based Systems (395 citations) | Raffel *et al.* (2020), Journal of Machine Learning Research (372 citations) | Sarikaya (2014), IEEE-ACM Transactions on Audio, Speech, and Language Processing (269 citations) | Ghiassi (2013), Expert Systems with Applications (237 citations) | Liu (2019), Neurocomputing (204 citations). The major themes that emerged out of the bibliometric analysis are namely – text classification, deep learning, machine learning, neural networks, feature extraction, and attention mechanism. The emerging themes are namely – convolutional neural networks, ensemble learning, and social networking. China is collaborating majorly on TC research with multiple countries, namely – the USA, Australia, the United Kingdom, and others. At the same time, the USA is collaborating with countries, namely – Canada, New Zealand, Colombia, Netherlands, and others. France and Brazil are collaborating with no other collaboration link. Asian countries are collaborating more than non-Asian countries. Since our research focuses on industry applications, we believe our inclusive approach, where we are not just doing a review of past work but also thoroughly discussing the techniques and business applications, provides a single repository to industry practitioners and increases the importance of this kind of research.

5. Implications

This section states both the theoretical and managerial implications of this paper.

5.1 Theoretical implications

We provide a few theoretical implications for improving future research on TC considering the findings. First, our findings help researchers comprehend the limitations and range of present research in this field. As a result, to encourage broader adoption of TC in managerial domains, academics may make use of our findings to draw attention to the less examined and unique challenges. Second, the Identification of eminent individuals as possible partners and driving factors for the advancement of this field's study may also be beneficial to researchers. Third, the crucial information about eminent and significant papers may be the cornerstones of this research field. This knowledge will be useful to upcoming researchers. These articles serve as a starting point for more in-depth investigations into the problems revealed by the network and citation studies. Finally, by conducting research based on mathematical modeling and empirical investigations, our study could be used as a platform for pushing methodological improvements in future studies.

According to (Mukherjee *et al.*, 2022), bibliometric research advances theory in five different ways: (a) by encouraging the objective discovery of knowledge clusters; (b) by clarifying the nomological networks to present the state of the field; (c) by mapping social patterns to understand social processes supporting knowledge development in the field; (d) by tracking evolutionary nuances to understand where the field is going; and € by recognizing crucial knowledge gaps to situate future research directions. The steps to conducting bibliometric analysis are as follows, according to (Donthu *et al.*, 2021): a. Define the objectives and parameters of the bibliometric study; b. Select the methods for bibliometric analysis; c. Gather the data for bibliometric analysis; and d. Execute the bibliometric analysis and present the results. Our research contributions are in-line with the above-stated studies.

5.2 Managerial implications

From an industry point of view, there are significant implications of TC that further needs to be explored and developed for accuracy. One of the major use cases lies in the online advertising industry. This industry is now moving towards a "cookie-less future," which means relatively more user privacy. Data collection giants such as Google and Apple are now planning to let go of third-party cookies in Chrome and access device identifiers on iOS. These steps have had a mixed response from consumers and advertisers. Regardless, more privacy to users means less flexibility to advertisers in targeting specific individuals whose user data was accessible through data management platforms. Moreover, in the last two decades, online advertising has seen some stern and far-reaching regulations being enforced by the advertising authorities in tandem with the government. Some of these are the Children's Online Privacy Protection Act of 1998 (COPPA), the General Data Protection Regulation (GDPR) of 2018, and even some state-level regulations implemented in the US, such as the California Online Privacy Protection Act (CalOPPA) adopted in 2013. What does this mean for brands allocating millions of dollars to online advertising? As the visibility into user data for targeting gets restricted, advertisers are likely going to be increasingly solicitous about brand measurement metrics of their online campaigns – Brand Safety, Brand Suitability, Invalid Traffic (IVT; fraudulent inventory), and Viewability. Out of these, TC, combined with Deep Learning (DL), has a wide array of applications in Brand Safety, Brand Suitability, and detecting Invalid Traffic coming from spoofed domains. Improved algorithms for Brand Safety and Brand Suitability will safeguard the reputation of a brand and, at the same time, maintain the contextual relevance of ads being served online. IVT, on the other hand, is an even bigger concern that continues to be a major reason for revenue leakage because of the ad server and the security system's inability to distinguish between a real and spoofed domain. These fake domains have different Uniform Resource Locators (URLs) from the actual domain but a similar, if not exact, User Interface (UI) that makes it easy to fool an average user; apart from brands, such malpractices can be damaging for users as well. Phishing attacks are used by fraudsters to obtain access to sensitive information, such as financial credentials, which leads to huge losses. Sophisticated TC algorithms will prove to be extremely beneficial in ensuring the safety of users and the prevention of revenue leakage in the online mode of advertising. Along the same lines, in the information age, where on a given topic there exists an abounding number of sources to provide information, it is crucial to provide a distinction between information and "misinformation", i.e., the information based on or directed towards misleading and questionable facts. This becomes necessary, especially to ensure social decorum and harmony. Islam *et al.* (2020) in their work explain five types of misinformation that exist on social media and can also be seen as interrelated: False information, Rumours, Spam, Fake news, and Disinformation. Fake news can lead to social distress, and it is our firm belief that an effective way to ensure the prevention of such instances is by using AI technologies on social media platforms. Facebook, in their quarterly published Community Standards Enforcement Report for Nov. 2020, claims that "AI now proactively detects 94.7% of hate speech we remove from Facebook, up from 80.5% a year ago and up from just 24% in 2017". Hate speech has been a popular subject, as seen by increasing media coverage as well as increased government attention to the issue (Fortuna and Nunes, 2018). To improve the classification accuracy of such content, more research in TC and DL techniques becomes essential and important.

6. Conclusion and limitations

TC has shown to be a valuable resource for anyone conducting research in the field of NLP. This is due to the versatility of its applications. When combined with DL approaches, this tool's processing powers, and precision become even more potent. Using a complete literature study, we examined the applications of TC, DL models, and business applications of TC in this work. We analyzed 118 research papers published in the fields of NLP, TC, DL, and TC tasks to accomplish two fundamental goals: 1) To showcase TC business applications and 2) To illustrate the use of common DL models and TC tasks, a comprehensive evaluation of previous research in this field is conducted. The greatest number of articles in this field were published between 2016 and 2019, according to our comprehensive evaluation of the past 20 years' worth of scholarly works. We created our own database containing all the papers we studied and filtered out those that covered TC, NLP, TC tasks, and DL for TC.

By achieving our first primary objective, we were able to offer key business applications of TC that will assist industry practitioners in developing a successful strategy. This involves a thorough comprehension of the many use cases applicable to their organizations. In this section, we discuss its applications, which include Opinion Spam Detection, Intelligent Document Analysis, Sentiment Analysis, Chatbots, Spam Classifiers, Social Media Analytics, Fraud Detection, Brand Safety Detection, Search Engine Optimization, and CRM Automation. The examination of TC-related literature was our second objective. We were especially interested in identifying TC tasks with business implications. We encountered TC tasks such as Sentiment Analysis (SA), Machine Translation (MT), Summarization, Question Answering (QA), Topic Labelling (TL), Information Retrieval (IR), Paraphrase Identification (PI), and Syntactic Parsing (SP), among others. Each of these TC responsibilities can be matched to an enterprise application. DL has demonstrated efficacy in the development of algorithms for various TC problems. In addition, we define DL and describe DL models that employ NLP and TC-related research extensively.

One of the limitations of our study is that being a thorough survey of research done in the domain of TC and NLP areas in the last two decades, our paper needs to explicitly discuss a single technique for TC or its applications in detail. The objective of our research is to elucidate industry practitioners and NLP researchers about the future avenues of research. Therefore, our objective and research methodology do not allow us to emphasize standalone applications categorically. However, in our follow-up research, we plan to narrow down to standalone applications where we not only expound on their business use cases but also propose new methodologies and novel algorithms to solve these problems.

References

Bartol, T., Budimir, G., Dekleva-Smrekar, D., Pusnik, M. and Juznic, P. (2014), "Assessment of research fields in Scopus and Web of Science in the view of national research evaluation in Slovenia", *Scientometrics*, Vol. 98, pp. 1491-1504.

Broadus, R.N. (1987), "Toward a definition of "bibliometrics"", *Scientometrics*, Vol. 12, pp. 373-379.

Chen, C. and Dai, J. (2021), "Mitigating backdoor attacks in LSTM-based text classification systems by backdoor keyword identification", *Neurocomputing*, Vol. 452, pp. 253-262.

Colace, F., De Santo, M., Greco, L. and Napoletano, P. (2014), "Text classification using a few labeled examples", *Computers in Human Behavior*, Vol. 30, pp. 689-697.

Dalal, M.K. and Zaveri, M.A. (2011), "Automatic text classification: a technical review", *International Journal of Computer Applications*, Vol. 28 No. 2, pp. 37-40.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. and Lim, W.M. (2021), "How to conduct a bibliometric analysis: an overview and guidelines", *Journal of Business Research*, Vol. 133, pp. 285-296.

Durán-Sánchez, A., Del Río, M.D.L.C., Álvarez-García, J. and García-Vélez, D.F. (2019), "Mapping of scientific coverage on education for entrepreneurship in higher education", *Journal of Enterprising Communities: People and Places in the Global Economy*, Vol. 13 Nos 1/2, pp. 84-104.

Fortuna, P. and Nunes, S. (2018), "A survey on automatic detection of hate speech in text", *ACM Computing Surveys (CSUR)*, Vol. 51 No. 4, pp. 1-30.

Genkin, A., Lewis, D.D. and Madigan, D. (2007), "Large-scale Bayesian logistic regression for text categorization", *Technometrics*, Vol. 49 No. 3, pp. 291-304.

Ghiassi, M., Skinner, J. and Zimbra, D. (2013), "Twitter brand sentiment analysis: a hybrid system using n-gram analysis and dynamic artificial neural network", *Expert Systems with Applications*, Vol. 40 No. 16, pp. 6266-6282.

Grießhaber, D., Vu, N.T. and Maucher, J. (2020), "Low-resource text classification using domain-adversarial learning", *Computer Speech and Language*, Vol. 62, 101056.

Gupta, V. and Lehal, G.S. (2009), "A survey of text mining techniques and applications", *Journal of Emerging Technologies in Web Intelligence*, Vol. 1 No. 1, pp. 60-76.

Hajiabadi, H., Molla-Aliod, D., Monsefi, R. and Yazdi, H.S. (2020), "Combination of loss functions for deep text classification", *International Journal of Machine Learning and Cybernetics*, Vol. 11 No. 4, pp. 751-761.

Hashemi, M. (2020), "Web page classification: a survey of perspectives, gaps, and future directions", *Multimedia Tools and Applications*, Vol. 79 No. 17, pp. 11921-11945.

Hartmann, J., Huppertz, J., Schamp, C. and Heitmann, M. (2019), "Comparing automated text classification methods", *International Journal of Research in Marketing*, Vol. 36 No. 1, pp. 20-38.

Huang, X., Chen, B., Xiao, L., Yu, J. and Jing, L. (2021), "Label-aware document representation via hybrid attention for extreme multi-label text classification", *Neural Processing Letters*, pp. 1-17.

Islam, M.R., Liu, S., Wang, X. and Xu, G. (2020), "Deep learning for misinformation detection on online social networks: a survey and new perspectives", *Social Network Analysis and Mining*, Vol. 10 No. 1, pp. 1-20.

Jang, B., Kim, M., Harerimana, G., Kang, S.U. and Kim, J.W. (2020), "Bi-LSTM model to increase accuracy in text classification: combining Word2vec CNN and attention mechanism", *Applied Sciences*, Vol. 10 No. 17, p. 5841.

Kessler, M.M. (1963), "Bibliographic coupling between scientific papers", *American Documentation*, Vol. 14 No. 1, pp. 10-25.

Khan, T. and Mollah, A.F. (2019), "AUTNT-A component level dataset for text non-text classification and benchmarking with novel script invariant feature descriptors and D-CNN", *Multimedia Tools and Applications*, Vol. 78 No. 22, pp. 32159-32186.

Küçük, D. and Can, F. (2020), "Stance detection: a survey", *ACM Computing Surveys (CSUR)*, Vol. 53 No. 1, pp. 1-37.

Lee, K., Han, S. and Myaeng, S.H. (2018), "A discourse-aware neural network-based text model for document-level text classification", *Journal of Information Science*, Vol. 44 No. 6, pp. 715-735.

Liang, D. and Yi, B. (2021), "Two-stage three-way enhanced technique for ensemble learning in inclusive policy text classification", *Information Sciences*, Vol. 547, pp. 271-288.

Liu, C. and Wang, X. (2020), "Quality-related English text classification based on recurrent neural network", *Journal of Visual Communication and Image Representation*, Vol. 71, 102724.

Liu, J., Xia, C., Yan, H., Xie, Z. and Sun, J. (2019), "Hierarchical comprehensive context modeling for Chinese text classification", *IEEE Access*, Vol. 7, pp. 154546-154559.

Liu, Y., Li, P. and Hu, X. (2022), "Combining context-relevant features with multi-stage attention network for short text classification", *Computer Speech and Language*, Vol. 71, 101268.

Liu, Z., Lu, C., Huang, H., Lyu, S. and Tao, Z. (2020), "Hierarchical multi-granularity attention-based hybrid neural network for text classification", *IEEE Access*, Vol. 8, pp. 149362-149371.

Lu, J., Behbood, V., Hao, P., Zuo, H., Xue, S. and Zhang, G. (2015), "Transfer learning using computational intelligence: a survey", *Knowledge-Based Systems*, Vol. 80, pp. 14-23.

Luo, W. and Zhang, L. (2022), "Question text classification method of tourism based on deep learning model", *Wireless Communications and Mobile Computing*, Vol. 2022, pp. 1-9.

Mitra, V., Wang, C.J. and Banerjee, S. (2007), "Text classification: a least square support vector machine approach", *Applied Soft Computing*, Vol. 7 No. 3, pp. 908-914.

Moirangthem, D.S. and Lee, M. (2021), "Hierarchical and lateral multiple timescales gated recurrent units with pre-trained encoder for long text classification", *Expert Systems with Applications*, Vol. 165, 113898.

Moreo, A., Esuli, A. and Sebastiani, F. (2021), "Word-class embeddings for multiclass text classification", *Data Mining and Knowledge Discovery*, Vol. 35 No. 3, pp. 911-963.

Mukherjee, D., Lim, W.M., Kumar, S. and Donthu, N. (2022), "Guidelines for advancing theory and practice through bibliometric research", *Journal of Business Research*, Vol. 148, pp. 101-115.

Paul, J., Merchant, A., Dwivedi, Y.K. and Rose, G. (2021), "Writing an impactful review article: what do we know and what do we need to know?", *Journal of Business Research*, Vol. 133, pp. 337-340.

Qi, X. and Davison, B.D. (2009), "Web page classification: features and algorithms", *ACM Computing Surveys (CSUR)*, Vol. 41 No. 2, pp. 1-31.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. and Liu, P.J. (2020), "Exploring the limits of transfer learning with a unified text-to-text transformer", *Journal of Machine Learning Research*, Vol. 21 No. 140, pp. 1-67.

Sarikaya, R., Hinton, G.E. and Deoras, A. (2014), "Application of deep belief networks for natural language understanding", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Vol. 22 No. 4, pp. 778-784.

Shuang, K., Guo, H., Zhang, Z., Loo, J. and Su, S. (2019), "A word-building method based on neural network for text classification", *Journal of Experimental and Theoretical Artificial Intelligence*, Vol. 31 No. 3, pp. 455-474.

Tian, D., Li, M., Shi, J., Shen, Y. and Han, S. (2021), "On-site text classification and knowledge mining for large-scale projects construction by integrated intelligent approach", *Advanced Engineering Informatics*, Vol. 49, 101355.

Van Linh, N., Anh, N.K., Than, K. and Dang, C.N. (2017), "An effective and interpretable method for document classification", *Knowledge and Information Systems*, Vol. 50 No. 3, pp. 763-793.

Wang, S., Huang, M. and Deng, Z. (2018), "Densely connected CNN with multi-scale feature attention for text classification", *IJCAI*, pp. 4468-4474.

Xu, J. and Du, Q. (2020a), "Adversarial attacks on text classification models using layer-wise relevance propagation", *International Journal of Intelligent Systems*, Vol. 35 No. 9, pp. 1397-1415.

Xu, J. and Du, Q. (2020b), "TextTricker: loss-based and gradient-based adversarial attacks on text classification models", *Engineering Applications of Artificial Intelligence*, Vol. 92, 103641.

Yang, B., Zhang, Y. and Li, X. (2011), "Classifying text streams by keywords using classifier ensemble", *Data and Knowledge Engineering*, Vol. 70 No. 9, pp. 775-793.

Yang, K., Gao, Y., Liang, L., Bian, S., Chen, L. and Zheng, B. (2021), "CrowdTC: crowd-powered learning for text classification", *ACM Transactions on Knowledge Discovery from Data (TKDD)*, Vol. 16 No. 1, pp. 1-23.

Yao, H., Zhang, B., Zhang, P. and Li, M. (2018), "A novel kernel for text classification based on semantic and statistical information", *Computing and Informatics*, Vol. 37 No. 4, pp. 992-1010.

Yao, L., Mao, C. and Luo, Y. (2019b), "Clinical text classification with rule-based features and knowledge-guided convolutional neural networks", *BMC Medical Informatics and Decision Making*, Vol. 19 No. 3, pp. 31-39.

Zhang, W., Tang, X. and Yoshida, T. (2015), "TESC: an approach to TExt classification using Semi-Supervised Clustering", *Knowledge-Based Systems*, Vol. 75, pp. 152-160.

Zhang, S., Chen, Y., Huang, X. and Cai, Y. (2019a), "Text classification of public Feedbacks using convolutional neural network based on differential evolution algorithm", *International Journal of Computers Communications and Control*, Vol. 14 No. 1, pp. 124-134.

Zhang, S., Yao, L., Sun, A. and Tay, Y. (2019b), "Deep learning based recommender system: a survey and new perspectives", *ACM Computing Surveys (CSUR)*, Vol. 52 No. 1, pp. 1-38.

Zhao, W., Zhang, G., Yuan, G., Liu, J., Shan, H. and Zhang, S. (2020), "The study on the text classification for financial news based on partial information", *IEEE Access*, Vol. 8, pp. 100426-100437.

Zheng, J. and Zheng, L. (2019), "A hybrid bidirectional recurrent convolutional neural network attention-based model for text classification", *IEEE Access*, Vol. 7, pp. 106673-106685.

Zulqarnain, M., Ghazali, R., Hassim, Y.M.M. and Aamir, M. (2021), "An enhanced gated recurrent unit with auto-encoder for solving text classification problems", *Arabian Journal for Science and Engineering*, Vol. 46 No. 9, pp. 8953-8967.

Further reading

Akhter, M.P., Jiangbin, Z., Naqvi, I.R., Abdelmajeed, M., Mehmood, A. and Sadiq, M.T. (2020), "Document-level text classification using single-layer multisize filters convolutional neural network", *IEEE Access*, Vol. 8, pp. 42689-42707.

Akhter, M.P., Jiangbin, Z., Naqvi, I.R., Abdelmajeed, M. and Fayyaz, M. (2022), "Exploring deep learning approaches for Urdu text classification in product manufacturing", *Enterprise Information Systems*, Vol. 16 No. 2, pp. 223-248.

Al-Garadi, M.A., Yang, Y.C., Cai, H., Ruan, Y., O'Connor, K., Graciela, G.H., Perrone, J. and Sarker, A. (2021), "Text classification models for the automatic detection of nonmedical prescription medication use from social media", *BMC Medical Informatics and Decision Making*, Vol. 21 No. 1, pp. 1-13.

Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Hasan, M., Van Essen, B.C., Awwal, A.A. and Asari, V.K. (2019), "A state-of-the-art survey on deep learning theory and architectures", *Electronics*, Vol. 8 No. 3, p. 292.

Alsaleh, D. and Larabi-Marie-Sainte, S. (2021), "Arabic text classification using convolutional neural network and genetic algorithms", *IEEE Access*, Vol. 9, pp. 91670-91685.

Altınel, B. and Ganiz, M.C. (2018), "Semantic text classification: a survey of past and recent advances", *Information Processing and Management*, Vol. 54 No. 6, pp. 1129-1153.

Alvesson, M. and Sandberg, J. (2011), "Generating research questions through problematization", *Academy of Management Review*, Vol. 36 No. 2, pp. 247-271.

Burkhardt, S. and Kramer, S. (2018), "Online multi-label dependency topic models for text classification", *Machine Learning*, Vol. 107 No. 5, pp. 859-886.

Cai, L., Song, Y., Liu, T. and Zhang, K. (2020), "A hybrid BERT model that incorporates label semantics via adjustive attention for multi-label text classification", *IEEE Access*, Vol. 8, pp. 152183-152192.

Chen, H., Wu, L., Chen, J., Lu, W. and Ding, J. (2022), "A comparative study of automated legal text classification using random forests and deep learning", *Information Processing and Management*, Vol. 59 No. 2, 102798.

Dong, Y., Liu, P., Zhu, Z., Wang, Q. and Zhang, Q. (2019), "A fusion model-based label embedding and self-interaction attention for text classification", *IEEE Access*, Vol. 8, pp. 30548-30559.

El Rifai, H., Al Qadi, L. and Elnagar, A. (2021), "Arabic text classification: the need for multi-labeling systems", *Neural Computing and Applications*, Vol. 34 No. 2, pp. 1135-1159.

El-Alfy, E.S.M. and Mohammed, S.A. (2020), "A review of machine learning for big data analytics: bibliometric approach", *Technology Analysis and Strategic Management*, Vol. 32 No. 8, pp. 984-1005.

El-Halees, A.M. (2015), "Arabic text classification using maximum entropy", *IUG Journal of Natural Studies*, Vol. 15 No. 1.

Er, M.J., Zhang, Y., Wang, N. and Pratama, M. (2016), "Attention pooling-based convolutional neural network for sentence modeling", *Information Sciences*, Vol. 373, pp. 388-403.

Fang, W., Luo, H., Xu, S., Love, P.E., Lu, Z. and Ye, C. (2020), "Automated text classification of near-misses from safety reports: an improved deep learning approach", *Advanced Engineering Informatics*, Vol. 44, 101060.

Gandotra, E. and Gupta, D. (2021), "Improving spoofed website detection using machine learning", *Cybernetics and Systems*, Vol. 52 No. 2, pp. 169-190.

Gao, S., Alawad, M., Young, M.T., Gounley, J., Schaefferkoetter, N., Yoon, H.J., Wu, X.C., Durbin, E.B., Doherty, J., Stroup, A., Coyle, L. and Tourassi, G. (2021), "Limitations of transformers on clinical text classification", *IEEE Journal of Biomedical and Health Informatics*, Vol. 25 No. 9, pp. 3596-3607.

Gargiulo, F., Silvestri, S., Ciampi, M. and De Pietro, G. (2019), "Deep neural network for hierarchical extreme multi-label text classification", *Applied Soft Computing*, Vol. 79, pp. 125-138.

Ghiassi, M., Olschimke, M., Moon, B. and Arnaudo, P. (2012), "Automated text classification using a dynamic artificial neural network model", *Expert Systems with Applications*, Vol. 39 No. 12, pp. 10967-10976.

Gupta, V. and Gupta, V. (2012), "Algorithm for Punjabi text classification", *International Journal of Computer Applications*, Vol. 37 No. 11, pp. 30-35.

Hao, M., Xu, B., Liang, J.Y., Zhang, B.W. and Yin, X.C. (2020), "Chinese short text classification with mutual-attention convolutional neural networks", *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, Vol. 19 No. 5, pp. 1-13.

Huan, J.L., Sekh, A.A., Quek, C. and Prasad, D.K. (2021), "Emotionally charged text classification with deep learning and sentiment semantic", *Neural Computing and Applications*, pp. 1-11.

Ibrahim, M.A., Khan, M.U.G., Mehmood, F., Asim, M.N. and Mahmood, W. (2021), "GHS-NET a generic hybridized shallow neural network for multi-label biomedical text classification", *Journal of Biomedical Informatics*, Vol. 116, 103699.

Irfan, R., King, C.K., Grages, D., Ewen, S., Khan, S.U., Madani, S.A., Kolodziej, J., Wang, L., Chen, D., Rayes, A., Tziritas, N. and Li, H. (2015), "A survey on text mining in social networks", *The Knowledge Engineering Review*, Vol. 30 No. 2, pp. 157-170.

Jing, W., Song, X., Di, D. and Song, H. (2021), "geoGAT: graph model based on attention mechanism for geographic text classification", *Transactions on Asian and Low-Resource Language Information Processing*, Vol. 20 No. 5, pp. 1-18.

Kadhim, A.I. (2019), "Survey on supervised machine learning techniques for automatic text classification", *Artificial Intelligence Review*, Vol. 52 No. 1, pp. 273-292.

Kardakis, S., Perikos, I., Grivokostopoulou, F. and Hatzilygeroudis, I. (2021), "Examining attention mechanisms in deep learning models for sentiment analysis", *Applied Sciences*, Vol. 11 No. 9, p. 3883.

Khan, A., Baharudin, B., Lee, L.H. and Khan, K. (2010), "A review of machine learning algorithms for text-documents classification", *Journal of Advances in Information Technology*, Vol. 1 No. 1, pp. 4-20.

Kim, H.J., Kim, J., Kim, J. and Lim, P. (2018), "Towards perfect text classification with Wikipedia-based semantic Naïve Bayes learning", *Neurocomputing*, Vol. 315, pp. 128-134.

Kushwaha, A.K., Kar, A.K. and Dwivedi, Y.K. (2021), "Applications of big data in emerging management disciplines: a literature review using text mining", *International Journal of Information Management Data Insights*, Vol. 1 No. 2, 100017.

Kwon, D., Kim, H., Kim, J., Suh, S.C., Kim, I. and Kim, K.J. (2019), "A survey of deep learning-based network anomaly detection", *Cluster Computing*, Vol. 22 No. 1, pp. 949-961.

Lewis, D.D., Yang, Y., Russell-Rose, T. and Li, F. (2004), "Rcv1: a new benchmark collection for text categorization research", *Journal of Machine Learning Research*, Vol. 5 Apr, pp. 361-397.

Li, Q., Li, P., Mao, K. and Lo, E.Y.M. (2020), "Improving convolutional neural network for text classification by recursive data pruning", *Neurocomputing*, Vol. 414, pp. 143-152.

Li, X., Cui, M., Li, J., Bai, R., Lu, Z. and Aickelin, U. (2021), "A hybrid medical text classification framework: integrating attentive rule construction and neural network", *Neurocomputing*, Vol. 443, pp. 345-355.

Liang, Y., Li, H., Guo, B., Yu, Z., Zheng, X., Samtani, S. and Zeng, D.D. (2021), "Fusion of heterogeneous attention mechanisms in multi-view convolutional neural network for text classification", *Information Sciences*, Vol. 548, pp. 295-312.

Liu, G. and Guo, J. (2019), "Bidirectional LSTM with attention mechanism and convolutional layer for text classification", *Neurocomputing*, Vol. 337, pp. 325-338.

Martarelli, N.J. and Nagano, M.S. (2022), "How to undertake reviews of large collections of articles and establish main contributions: an ontology-based literature review approach", *International Journal of Information Management Data Insights*, Vol. 2 No. 2, 100091.

Maw, M., Balakrishnan, V., Rana, O. and Ravana, S.D. (2020), "Trends and patterns of text classification techniques: a systematic mapping study", *Malaysian Journal of Computer Science*, Vol. 33 No. 2, pp. 102-117.

Mohammad, A.S., Jaradat, Z., Mahmoud, A.A. and Jararweh, Y. (2017), "Paraphrase identification and semantic text similarity analysis in Arabic news tweets using lexical, syntactic, and semantic features", *Information Processing and Management*, Vol. 53 No. 3, pp. 640-652.

Oleynik, M., Kugic, A., Kasáč, Z. and Kreuzthaler, M. (2019), "Evaluating shallow and deep learning strategies for the 2018 n2c2 shared task on clinical text classification", *Journal of the American Medical Informatics Association*, Vol. 26 No. 11, pp. 1247-1254.

Omayio, E.O., Sreedevi, I. and Panda, J. (2021), "Language-based text categorization: a survey. Digital techniques for heritage presentation and preservation", *Digital Techniques for Heritage Presentation and Preservation*, pp. 11-36.

Pan, C., Huang, J., Gong, J. and Yuan, X. (2019), "Few-shot transfer learning for text classification with lightweight word embedding based models", *IEEE Access*, Vol. 7, pp. 53296-53304.

Queiroz, M.M., Ivanov, D., Dolgui, A. and Fosso Wamba, S. (2022), "Impacts of epidemic outbreaks on supply chains: mapping a research agenda amid the COVID-19 pandemic through a structured literature review", *Annals of Operations Research*, Vol. 319 No. 1, pp. 1159-1196.

Ranjan, N., Mundada, K., Phaltane, K. and Ahmad, S. (2016), "A survey on techniques in NLP", *International Journal of Computer Applications*, Vol. 134 No. 8, pp. 6-9.

Ren, J., Wu, W., Liu, G., Chen, Z. and Wang, R. (2021), "Bidirectional gated temporal convolution with attention for text classification", *Neurocomputing*, Vol. 455, pp. 265-273.

Rogers, D., Preece, A., Innes, M. and Spasić, I. (2021), "Real-time text classification of user-generated content on social media: systematic review", *IEEE Transactions on Computational Social Systems*.

Sahingoz, O.K., Buber, E., Demir, O. and Diri, B. (2019), "Machine learning based phishing detection from URLs", *Expert Systems with Applications*, Vol. 117, pp. 345-357.

Salloum, S.A., AlHamad, A.Q., Al-Emran, M. and Shaalan, K. (2018), "A survey of Arabic text mining", *Intelligent Natural Language Processing: Trends and Applications*, Springer, Cham, pp. 417-431.

Schwenker, F. and Trentin, E. (2014), "Pattern classification and clustering: a review of partially supervised learning approaches", *Pattern Recognition Letters*, Vol. 37, pp. 4-14.

Shin, H.S., Kwon, H.Y. and Ryu, S.J. (2020), "A new text classification model based on contrastive word embedding for detecting cybersecurity intelligence in twitter", *Electronics*, Vol. 9 No. 9, p. 1527.

Stein, R.A., Jaques, P.A. and Valiati, J.F. (2019), "An analysis of hierarchical text classification using word embeddings", *Information Sciences*, Vol. 471, pp. 216-232.

Sun, S., Luo, C. and Chen, J. (2017), "A review of natural language processing techniques for opinion mining systems", *Information Fusion*, Vol. 36, pp. 10-25.

Tan, Z., Chen, J., Kang, Q., Zhou, M., Abusorrah, A. and Sedraoui, K. (2021), "Dynamic embedding projection-gated convolutional neural networks for text classification", *IEEE Transactions on Neural Networks and Learning Systems*.

Tan, C., Ren, Y. and Wang, C. (2022), "An adaptive convolution with label embedding for text classification", *Applied Intelligence*, pp. 1-9.

Van Phan, T. and Nakagawa, M. (2016), "Combination of global and local contexts for text/non-text classification in heterogeneous online handwritten documents", *Pattern Recognition*, Vol. 51, pp. 112-124.

Vanhala, M., Lu, C., Peltonen, J., Sundqvist, S., Nummenmaa, J. and 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*, Vol. 106, pp. 46-59.

Wagner, G., Lukyanenko, R. and Paré, G. (2022), "Artificial intelligence and the conduct of literature reviews", *Journal of Information Technology*, Vol. 37 No. 2, pp. 209-226.

Wang, P., Xu, B., Xu, J., Tian, G., Liu, C.L. and Hao, H. (2016), "Semantic expansion using word embedding clustering and convolutional neural network for improving short text classification", *Neurocomputing*, Vol. 174, pp. 806-814.

Wang, J., Wang, Z., Zhang, D. and Yan, J. (2017), "Combining knowledge with deep convolutional neural networks for short text classification", *IJCAI*, Vol. 350.

Wang, J., Li, Y., Shan, J., Bao, J., Zong, C. and Zhao, L. (2019), "Large-scale text classification using scope-based convolutional neural network: a deep learning approach", *IEEE Access*, Vol. 7, pp. 171548-171558.

Wang, C., Ju, S., Liu, Y., Chen, R. and Huang, X. (2020), "Adaptive feature extractor of global representation and local semantics for text classification", *IEEE Access*, Vol. 8, pp. 202687-202695.

Wang, B., Hu, X., Li, P. and Philip, S.Y. (2021), "Cognitive structure learning model for hierarchical multi-label text classification", *Knowledge-Based Systems*, Vol. 218, 106876.

Whetten, D.A. (1989), "What constitutes a theoretical contribution?", *Academy of Management Review*, Vol. 14 No. 4, pp. 490-495.

Woźniak, M., Grana, M. and Corchado, E. (2014), "A survey of multiple classifier systems as hybrid systems", *Information Fusion*, Vol. 16, pp. 3-17.

Xia, R., Zong, C. and Li, S. (2011), "Ensemble of feature sets and classification algorithms for sentiment classification", *Information Sciences*, Vol. 181 No. 6, pp. 1138-1152.

Xia, R., Xu, F., Yu, J., Qi, Y. and Cambria, E. (2016), "Polarity shift detection, elimination, and ensemble: a three-stage model for document-level sentiment analysis", *Information Processing and Management*, Vol. 52 No. 1, pp. 36-45.

Xing, Z., Pei, J. and Keogh, E. (2010), "A brief survey on sequence classification", *ACM SIGKDD Explorations Newsletter*, Vol. 12 No. 1, pp. 40-48.

Xu, S. and Xiang, Y. (2021), "Frog-GNN: multi-perspective aggregation based graph neural network for few-shot text classification", *Expert Systems with Applications*, Vol. 176, 114795.

Yang, Z. and Liu, G. (2019), "Hierarchical sequence-to-sequence model for multi-label text classification", *IEEE Access*, Vol. 7, pp. 153012-153020.

Yao, J., Wang, K. and Yan, J. (2019a), "Incorporating label co-occurrence into neural network-based models for multi-label text classification", *IEEE Access*, Vol. 7, pp. 183580-183588.

Zeng, K., Pan, Z., Xu, Y. and Qu, Y. (2020), "An ensemble learning strategy for eligibility criteria text classification for clinical trial recruitment: algorithm development and validation", *JMIR Medical Informatics*, Vol. 8 No. 7, e17832.

Zhang, J.Z., Srivastava, P.R., Sharma, D. and Eachempati, P. (2021), "Big data analytics and machine learning: a retrospective overview and bibliometric analysis", *Expert Systems with Applications*, Vol. 184, 115561.

Corresponding author
Gaurav Sarin can be contacted at: gaurav.sarin@iiml.org