Review

A Review on Current Trends and Applications of Social Media Research in Sri Lanka

Isuru Udayangani Hewapathirana [ID]

Software Engineering Teaching Unit, University of Kelaniya, Kelaniya, Sri Lanka
Email: ihewapathirana@kln.ac.lk

Abstract: Standard research on social media and its applications has been widely disseminated in developed nations. But in Sri Lanka, research in this area has been released far less frequently. However, social media usage in the country is evolving regardless of age, sex, education level, or other limitations. This study aims to fill the gap by conducting a comprehensive review of social media-based research conducted in Sri Lanka between 2012 and 2022. A systematic search of reputable databases, including IEEE Xplore, ScienceDirect, Emerald Insight, Google Scholar, and Springer Link, identified 57 relevant papers for analysis. The review highlights the diversity of application areas where social media research has been employed in Sri Lanka, including disaster management, public health, marketing, education, and more. Additionally, the analysis highlights the methodological approaches employed in social media analytics and the specific social media platforms utilized by researchers in Sri Lanka. The results of the current study serve as a timely resource, enabling policymakers and decision-makers to identify the potential avenues of social media research in Sri Lanka. By understanding the existing trends and implications, stakeholders can harness the power of social media data to make informed policy decisions, develop effective marketing strategies, enhance public health initiatives, and revolutionize educational practices.

Keywords: social media, social media analytics, social media applications, literature review, Sri Lanka

1. Introduction

Sri Lanka, a developing country with a population of approximately 21 million, is experiencing a rapid expansion of social media's influence on its society. Noteworthy incidents, such as the spread of rumors on social media leading to panic buying during the COVID-19 epidemic and the dissemination of misinformation triggering ethnic conflict, highlight the importance of understanding social media usage patterns in the current information age. While innovative researchers and business experts worldwide have been exploring new approaches for gathering, combining, and analyzing vast amounts of social media data [1-2], it remains unclear if similar studies have been conducted in the Sri Lankan setting. Therefore, this study aims to address this gap by offering a comprehensive review of social media research conducted in Sri Lanka, providing unique insights on the most popular sectors in which social media research has been conducted, prevalent data analytic techniques utilized in such studies, and social media platforms that have generated the most discussion among the research community. This pioneering attempt represents an initial comprehensive analysis centered on the investigation of social media usage and applications in Sri Lanka. By offering

insights into the current state of social media research in the country, the current study provides substantial value to decision-makers interested in utilizing social media data for making informed decisions. Additionally, it serves as a foundational resource and offers insights into the difficulties encountered in conducting social media research in Sri Lanka, benefiting researchers in the field.

A literature review organizes, combines, and evaluates previously published research to educate the reader on how the present body of research is developing in the context of resolving a specific issue [3]. After reading a review article, readers can better spot discrepancies in earlier findings and potential explanations (such as moderators, mediators, measurements, and techniques) and quickly understand current gaps and potential future research objectives [4]. Several literature reviews exist on social media applications, analytics, and their effects concerning various fields. Examples include systematic literature reviews that discuss techniques, software tools and platforms for social media analytics [1-2, 5-6] and review articles that focus on domains such as education [7-8], health [9], politics [10], disaster management [11] and business and innovation [12]. Although there are several literature reviews on the data and usage of social media in different application fields, we could not locate any that discussed the general trends and applications of research on social media usage concerning a specific country.

The statement of the problem that motivated this research is, "In what ways has research on social media usage progressed in Sri Lanka over the past ten years?" This broad query is broken down into the primary research objective and secondary research questions. The primary research objective of this study is to provide a comprehensive, consolidated review of the research conducted in Sri Lanka on social media usage and related applications. Below are the secondary research questions (RQs) that correspond to it:

RQ1: What are the most popular sectors in which social media research was conducted?

RQ2: What data analytic techniques were primarily used in social media research?

RQ3: What social media platforms have generated the most discussion among the research community during the past ten years?

RQ4: What are the main challenges in researching social media usage in Sri Lanka?

To achieve the research objective and address the research questions, this paper undertook a comprehensive review of 57 research publications. The subsequent sections of this article are structured as follows: Section 2 outlines the methodology employed for conducting the literature survey. Section 3 presents the findings of the literature review, and Section 4 provides a detailed analysis of the reviewed articles together with a discussion on the key insights derived from the reviewed literature. Finally, Section 5 concludes the study by summarizing the main conclusions drawn and outlining potential avenues for future research.

2. Study methodology

This study followed the six-stage systematic review process described in [4]. The stages include i) outlining the study objectives and formulating research questions to be investigated, ii) defining the criteria for articles to be included and excluded, iii) Sampling relevant research to be included in the review process from among the research articles, iv) obtaining the identified relevant articles and determining whether the articles match the selection criterion, v) extracting the relevant information from the obtained articles and finally, vi) reporting the findings of the article reviews that have been conducted. The research objectives were discussed in Section 1. Consequently, the preceding sections will elaborate upon the remaining stages of the study methodology.

2.1 Inclusion and exclusion criteria

To ensure the selection of relevant research articles, specific inclusion and exclusion criteria were established. The studies included in the review needed to focus on social media usage and its applications in Sri Lanka. Only peer-reviewed articles published between the years 2010-2022 were considered eligible for inclusion. Additionally, articles had to be written in English and provide sufficient information pertaining to social media usage in Sri Lanka. Studies that did not meet these criteria were excluded from the review.

2.2 *Search strategy*

A systematic search strategy was employed to identify relevant research articles. Multiple academic databases, including Google Scholar, IEEE Xplore, ACM Digital Library, and relevant journals, were systematically searched using appropriate keywords and combinations. The search terms used were "social media", "social media analytics", and "Sri Lanka". The search strategy was refined iteratively to ensure comprehensive coverage of the available literature on the topic. The search retrieved 115 articles which were preceded by screening for selection.

2.3 *Study selection process*

The study selection process involved two stages: screening titles and abstracts, followed by a full-text review. The titles and abstracts of the identified articles were assessed to evaluate their relevance based on the predefined inclusion and exclusion criteria. Following the screening, the selected articles were distributed as follows: 23 articles from IEEE Xplore, 6 articles from the ACM digital library, 3 articles from ScienceDirect, 3 articles from SpringerLink, 17 articles from Google Scholar, and 5 articles from Emerald Insight. It is noted that many studies investigated people's perspectives concerning various social media applications through survey data collection. Even though these studies didn't use any methods for mining social media data, they were nevertheless helpful in recognizing social media-related issues unique to various business domains. They were, therefore, not excluded from the review. The selected articles then underwent a thorough full-text review to determine their suitability for inclusion in the review.

2.4 *Information extraction from articles*

A structured data extraction process was employed to collect relevant information from the selected articles. A predefined data extraction form was utilized to capture details such as author names, publication year, research objectives, methodology, key findings, and conclusions. This systematic approach ensured that essential data points were extracted consistently across all selected articles.

2.5 *Quality assessment*

The quality assessment of the included papers in this study was conducted by evaluating their methodological rigor, validity, and reliability. While a quantitative score was not calculated, a qualitative assessment was performed to determine the overall quality and suitability of the papers for inclusion in the review. For papers analyzing social media data, the appropriateness of the data pre-processing techniques was examined, including steps such as data cleaning, filtering, and transformation. The transparency and reproducibility of the analytical methods used for data analysis, such as text analysis, sentiment analysis, network analysis, or machine learning algorithms, were also evaluated. For survey papers, specific aspects such as the selection of study groups, the comparability of groups, and the assessment of outcomes or exposures were examined, focusing on factors like sample size, data collection methods, and statistical analysis techniques.

2.6 *Data synthesis and analysis*

To gain deeper insights into the research conducted on social media usage in Sri Lanka, a content analysis was performed on the selected articles. This involved systematically examining the articles to identify common themes, trends, and patterns that emerged from the research conducted in the field. During the content analysis, the articles were carefully reviewed and categorized based on their application domains, methodologies employed, social media platforms studied, and other relevant factors. This process allowed for a comprehensive understanding of the research landscape in Sri Lanka concerning social media usage. The information extracted is provided in the following sections.

3. Literature review

Social media research started to gain attraction mainly after 2005 [1], and the literature has extensively examined

several of its uses since 2010.

3.1 *Areas of social media research*

The articles reviewed span various application areas where social media research is conducted. The studies illustrate how social media platforms can be employed for decision-making in these areas.

3.1.1 *Social issues*

Adolescence is seen as a transition period between childhood and maturity because of the numerous psychological, physiological, and cognitive changes that occur during this time. Social media impacts adolescents' academic and social development positively and negatively. The effects of social media on social behavior, personal conduct, interactions, education, communication, attitudes, skills, and capacities of Sri Lankan adolescent users are investigated in [13]. The study population consisted of General Certificate of Education (GCE) Advanced Level students aged between 17 to 20 years in three educational zones: Bandarawela, Rathnapura and Balangoda. Convenience sampling was used to select the sample. Out of 300 collected questionnaires, 261 were selected for analysis. Among the respondents, the majority were from the Arts stream (41.76%) and the minority was from the Technology (5.36%) stream. Out of them, almost 69% were average social media users and only 12% were high social media users. The most popular social media was Youtube (99.23%) followed by Facebook (92.72%). It was found that 46.36% of the respondents use social media for entertainment purposes and only 27.97% were for education. The sign test results showed that social media showed no adverse effect on adolescents' studies (median = 3.0) and did not cause any mental confusion or health problem (median = 3.0). According to this study, slow internet connection and understanding of English were the greatest barriers to social media usage.

While the use of social media is rapidly growing, some individuals have turned it into a forum for disseminating racist ideas among communities. Due to these racial sentiments, there have been instances where riots have started or grown in Sri Lanka. The work of [14] proposes a method to identify racist social media comments that are in the Sinhala language using Microsoft Azure Machine Learning Studio to create a Text Analytics Model with a Two-Class Support Vector Machine. The Sinhala language data corpus included both randomly selected racist and non-racist remarks. After pre-processing the data set, n-gram features were retrieved. According to the findings, the two-class support vector machine classifier performed with an accuracy and precision of 70.8% and 100%, respectively.

The primary goal in [15] was to investigate whether or not hate speech can be detected automatically online. A local English dataset was created using reader feedback on stories published on the Colombo Telegraph website. 1,500 comments were collected, 1,000 of which were annotated to show whether or not they included hate. The study employs a machine learning approach combined with a lexicon-based method for hate speech identification. The Google bad words list was used as the hate lexicon to create the lexicon-based technique. After reading every comment in the dataset individually, a count of the number of hate words was extracted. This number was utilized to create the feature "Hate word count", which was employed in machine learning models to analyze every piece of data in the dataset. Four supervised learning algorithms and one unsupervised learning algorithm were combined to create five machine-learning models. The K-Means clustering method was employed for unsupervised learning models, and for supervised learning models, the Naive Bayes, Support Vector Machine, Logistic Regression, Decision Tree, and Naive Bayes algorithms. Features were extracted using CountVectorizer, Term Frequency-Inverse Document Frequency (TF-IDF) Vectorizer, and Bag of Features methods. The study of the findings revealed that, of all models, the Naive Bayes Model with Tf-Idf Features performs the best, with an F-score value of 0.719. Out of all five models K-Means clustering model was the model which had the worst performance in almost all the scenarios.

Similarly, [16] also proposes a system for detecting hate speech in the text that uses a combination of Sinhala, Romanized Sinhala, and English terms. An open-source data collection of 2,500 comments were utilized for training their classification algorithm, which classified social media remarks as hateful or non-hateful. An open-source stop word corpus and stem word corpus in Sinhala was used to pre-process the data set. The two corpora were manually transformed into a Romanized Sinhala stop word corpus and a Romanized Sinhala stem word corpus to identify stop words and stem words in Romanized Sinhala. An open-source English word corpus was used to recognize all English words, and a library was used to get stop word corpus and stem English words. The bag of words and term frequency-

inverse document frequency were compared for feature engineering. As classification algorithms, linear support vector classifiers, random forest classifiers, Stochastic Gradient Descent (SGD) classifiers, logistic regression, XGBoost classifiers, and multinomial Naive Bayes classifiers were evaluated. The maximum accuracy was 74.2% achieved by the SGD classifier using the term frequency-inverse document frequency.

Using a survey conducted on 420 undergraduates of the University of Sri Jayawardanapura in Sri Lanka [17], investigates the factors associated with crimes in social media that target youths. The techniques incorporated to analyze the results include descriptive analysis, binomial test, index construction, chi-square test and logistic regression. The results showed that several factors such as relationship with family, income level, digital literacy, awareness, security management, privacy concentration, number of Facebook friends, hours spent online, and purpose of online surfing have a significant impact on young people becoming victims of crime on social media.

On April 21, 2019, Sri Lanka was hit by a series of synchronized bomb explosions targeting three major churches and international hotels, killing over 300 people and injuring hundreds more, terrifying all Sri Lankans. After 30 years of civil conflict, Sri Lanka has become a victim of another form of intimidation: religious extremism, which has progressed to a hazardous level of religious extremism-based terrorism. Religious extremism is not a problem that exists alone in Sri Lanka. According to the Global Shapers Survey done by the World Economic Forum (with 31,495 participants from 180 countries) [18], religious issues are the fifth of the top 10 difficulties the world is experiencing today. In the age of modern technology, [19] discuss how religious extremist groups have used modern communication tactics such as the Internet and social media to directly contact their global audience to harass users, recruit new members, and inspire violence. This motivated the authors to investigate religious extremism and its prevalence on social media, as well as people who dangerously disseminate extreme views via social media. The study in [19] proposes a model that can take both Sinhala and Tamil tweets as inputs and apply text analytics to classify them as religious-extreme or not. Their unique approach identifies religious extremism by utilizing a context-based algorithm that addresses the pragmatic meaning of each tweet. The created, on the other hand, mostly applies to the Sri Lankan context because it analyzes both Sinhala and Tamil language posts on social media and generates successful results in both languages.

The primary goal of [20] is to develop an automated way to detect Sinhala language social media remarks posted to insult a person. A text analytics model with five rules is explicitly designed to extract features of social media text in the Sinhala language. After acquiring these features, the system used the K-NN (K-nearest Neighbor), SVM (Support Vector Machine), and Naive Bayes algorithms to learn the pattern of the features. According to the results, the SVM algorithm with RBF (Radial Basis Function) kernel showed the best performance detecting cyberbullying comments with an F1-score of 91%. The data for the study was collected from Twitter by accessing Twitter Rest API using the Tweepy Python library.

Individuals use social media to share knowledge and attitudes, as well as to comment on the opinions of others. This information can be used to determine a person's personality characteristics. An ontology-based technique was presented in [21] for detecting a person's personality by examining long-term Facebook status patterns. The research uses Python's Natural Language Toolkit (NLTK) library to perform the natural language processing tasks and the Web Ontology Language (OWL) ontology developing language in Protégé software for designing the ontology. The authors claim the developed method to be helpful for employee profiling in areas such as human resource management, psychology, medical and business intelligence. A similar study in [22] investigated the findings and posts of psychological stress using data from Facebook users who utilize and share their profiles. The data is collected using the Facebook Graph Application Programming Interface (API), then machine learning and natural language processing (NLP) techniques are applied, and an intelligent custom AI virtual application is created to quantify stress levels based on many criteria.

The study in [23] detects cyberbullying with high accuracy using natural language processing techniques such as TF-IDF, sentiment analysis, and a supervised machine learning algorithm. They identify themes or categories linked with cyberbullying, such as racism, sexual, physical meaning, and swearing.

The work of [24] calculates many social network statistics such as transitivity, clustering coefficient, modularity, and number of weakly linked components to understand cross-ideology exposure on two Facebook page network datasets generated using the NodeXL social media data collecting function. The authors' goal of the research is to prove that enclaves of exposure are socially mediated technology environments that exhibit exposure divergence. To accomplish this, they analyze Facebook activism connected to the controversial documentary Sri Lanka's Killing Fields,

broadcast by UK broadcaster Channel 4 as the empirical framework for the study.

3.1.2 *Politics*

Sri Lanka is a semi-presidential representative democratic republic with a multi-party system in which the President of Sri Lanka serves as both head of state and head of government. The President wields executive power on the advice of the Prime Minister and the Cabinet of Ministers. Parliament has legislative authority [25]. People voted for and elected a new government in 2015 in response to several issues from the previous government, including corruption in state institutions, security threats from extremist groups, growing inequality, high development debts, and a lack of media freedom. A study was conducted in [26] to determine whether there is any significant improvement in the present governance compared to the previous reign from the general public's perspective as expressed through their social media interactions. Posts in the famous Sri Lankan newspaper, Lankadeepa, Facebook page were analyzed to identify trends in Sri Lankan politics from 2011 to 2018. Facebook posts were categorized as political and non-political, and the political posts were further divided into several subcategories: functioning government, statutory boards, concession, consumer price, development work, etc. The collected news posts included data on past and present governments concerning 2018, and those posts were analyzed based on the Facebook reactions achieved for each post. According to the results, users are less engaged in politics and think it's funny. Additionally, it was found that user reactions to politics have been increasing monotonically throughout the period. In contrast, the trends for the current and previous governments have been dropping and rising monotonically, respectively.

According to [27], Twitter is widely used in Sri Lanka as a platform for news dissemination. They study Twitter trending topics to learn how Sinhala Twitter data influences news dissemination in the media. The results of the study are a set of features that provide a tweet "news value", which can be used by journalists or the media to locate social media information that is newsworthy.

3.1.3 *Education*

According to [28], more than 3.5 million of Sri Lanka's active internet users use Facebook, and the number of Instagram accounts has almost topped 60,000. Around 11.5% of Sri Lanka's total population is currently on Facebook, with 1,400,000 men and 720,000 women making up the majority (33% of the total base). Additionally, apps like Viber, Whatsapp, and Facebook Messenger have transformed Sri Lankans' way of life by facilitating more straightforward communication.

Sharing research data can benefit significantly from the use of social media. Effective social media use can lead to the creation of a public profile to highlight individual research interests and to connect with a large audience. Using Twitter in class is another way to keep students interested and learn more about their thoughts while listening to lectures. Teachers can develop their skills or alter the learning process using such an approach. The teaching and learning resources such as online courses, course materials, textbooks, videos, examination questions, mock tests, and software that are openly accessible on the internet are known as open educational resources (OER). Social media allow the creation and exchange of user-generated content using OER. A cross-sectional study was conducted in [29] to understand the utilization patterns of OER and social media among the medical students at the Faculty of Medicine, University of Kelaniya, Sri Lanka. Using stratified random sampling, students from the first, second, third and fourth years were selected for the survey. The study revealed that most students used OER (96%) and social media (90%) for academic purposes. Wiki sites were the most popular OER (84%), while Facebook was the preferred social media (80%). However, e-journals were considered one of the most reliable and valuable OER sources in medicine; only around 6% of students used e-journals provided by the faculty. It was apparent that more than 60% of the students obtained information without considering the resource's credibility. A few students opted for educational and government websites for information seeking. Social media was used by 79% of the students in educational and group activities. However, more than one-third of students (35.8%) could not find academic information from social media due to information overload, and 31.1% mentioned that SM distracted their education.

Social media platforms have produced thousands of new professions and revenue streams. Addiction is one of the harmful effects of social media or networks. Spending much time on social media can make it difficult to concentrate on one task. People become less motivated as a result, particularly teenagers and students. Instead of obtaining the

skills and knowledge necessary for everyday life, they mostly rely on technology and the internet [28]. The research conducted in [30] evaluated the computer literacy and learning styles of Sri Lankan medical students for the purpose of preparing e-Learning materials to deliver to medical students via the faculty's Moodle platform of the University of Colombo, which has just begun its online development and implementation. An anonymous, self-administered questionnaire was distributed among second-year medical students of the University of Colombo medical faculty. The only task required of the students was to indicate the responses from which they had selected one. The "Likert Scale" was employed when applicable to reduce the number of mixed responses. The study offered proof of the usefulness and acceptance of e-learning in medical education, particularly when integrated with conventional teacher-led activities in a blended learning environment. Although [30] focuses on an e-learning platform specifically for medical students, [31] presents a revolutionary e-learning platform that uses social media tactics, particularly those of "Facebook", to create an active and engaging learning environment for university students in any domain. Personal knowledge management, learning management, and collaborative learning are three significant parts of the platform. Newsfeed, Classmates Profile, Repository, Knowledgebase, Bookmark, Topic Map, Search Engine, Test Mark Prediction, and Slide Show Summary generator were some of the planned platform's features. Several platform functionalities were built using machine learning and natural language processing techniques. The evaluation results of the suggested system, "KnowMore", demonstrated student satisfaction through enhanced self-motivation, self-learning, and interaction.

In educational settings such as schools, colleges, universities, and other institutions for vocational training, English is studied as a second language in Sri Lanka. Significantly less emphasis has been placed on listening and speaking in formal classroom language learning and the traditional teaching approaches used in Sri Lanka's English as a Second Language curriculum. Because of this, most students following the national curriculum have a bad attitude about learning English. In [32], a sample of students of the National Apprentice & Industrial Training Authority (NAITA), Trincomalee, Sri Lanka were examined to determine how using social media can assist them to learn English. The study's findings show that most of the sampled students use social media and are interested in using it to learn English to advance their language skills. They assist the students in learning English because there are several helpful English learning pages on Facebook and YouTube, which support the growth of the student's linguistic abilities and vocabulary. Additionally, they can speak English with people from other cultures and languages worldwide.

WhatsApp, a new social media network, is becoming increasingly popular among students as a new mode of communication. Despite the availability of a supplementary learner management system (LMS) for all courses, educators in higher education may choose to conduct audio-visual sessions via WhatsApp following the outbreak of the Covid-19 pandemic. Yet, [33] claims that for developing countries like Sri Lanka, the fast transition to digital instruction may have a negative impact on many learners from disadvantaged groups due to Internet inaccessibility and unaffordable technology. It is thus advisable to discover students' priorities, problems, interests, and learning preferences before implementing unplanned and rushed teaching methods. A survey was conducted by [33] to explore the current degree of perceived engagement benefits among learners in informal academic WhatsApp groups. In a marketing sense, they examine the presence of customer social involvement among learners on WhatsApp groups established for academic reasons, viewing higher education as a service and students as customers receiving that service. The perceived benefits of participation in informal WhatsApp groups are high among learners, according to responses from a sample of 170 undergraduates enrolled in the BMS degree course at the Open University of Sri Lanka. Furthermore, the learners who join these informal WhatsApp groups expect significant participation benefits such as functional, social, and hedonic benefits.

3.1.4 *Tourism*

Through social media, tourists frequently discuss their experiences at specific destinations and other locations in general. Consequently, these social data are rich in information that can be cleaned and applied to decision-making. The paper by [34] suggests a social media analytics platform that can extract microblogs from publicly available sources, such as Twitter, TripAdvisor and Booking.com, related to an object or location, discover discussion pathways, subjects, and elements that appear there, and then assess sentiment and deeper emotions connected to each of those topics. The novelty of the research is that it combines five subtasks into one tool. These tasks are discussion pathway extraction, aspect extraction, aspect-based sentiment analysis, emotion detection, and visualizations.

There is a growing interest in using social media images as research data [35]. Flickr is one of the most popular

photo-sharing websites worldwide. Using crowdsourced images from the social media network Flickr, [36] analyze Sri Lanka's whale-watching tourists. To do this, research was done on the social tag analysis of photographs as well as the content of the images, their geographic distribution, temporal distribution, and demographics of the whale photographers. According to the current study, geo-tagged whale-watching photos' metadata can be used to identify how tourists behave when whale-watching, and geo-tagged images could also be used to estimate the distribution of whales in Sri Lanka.

Social media is also playing an essential part in advertising Sri Lankan tourism by promoting Sri Lanka as a destination and tourism item to international tourists worldwide. Currently, Sri Lanka's destination marketing groups and the Sri Lanka government are taking steps to promote Sri Lanka tourism around the world. The objective of [37] is to explore the influence of such social media marketing initiatives on establishing brand equity for Sri Lanka as a destination for international tourists. The study also investigated the social media marketing efforts of Sri Lanka as a destination brand and the brand equity of Sri Lanka as a destination. The study's findings show that all factors under social media marketing efforts, namely entertainment, customization, and e-word of mouth, positively contribute to the brand equity of Sri Lanka as a destination. In contrast, interaction and trendiness have a negative impact. The authors recommend that Sri Lankan destination marketers prioritize the interaction and trendiness of their social media marketing initiatives in order to improve social media marketing efforts further.

3.1.5 *Medical and health services*

The public health inspectors (PHIs) reporting and recording a single dengue case in Sri Lanka required a convoluted, time-consuming process that could take up to 10 days under the current paper-based data collection practices for dengue surveillance. Based on one-on-one in-depth interviews, [38] undertake a requirements assessment of PHIs in Colombo regarding their dengue-related responsibilities. They then created a new mobile-based system called Mo-Buzz to meet these needs while bolstering current systems. The Mo-Buzz system digitizes three primary functions of PHIs. It presents them on handheld mobile devices and Web interfaces: i) capturing, storing, and recording visual, textual, and geographical information from patient visits and house or area safety checks, ii) keeping track of dengue spread patterns in the Colombo region in real-time, and iii) providing dengue education to the public in an engaging format that will keep their attention and interest. Using the Mo-Buzz smartphone application, participants can report symptoms of dengue such as sickness, knowledge of mosquito exposure/visible bites, perceived mosquito density, and/ or share images of suspected Aedes mosquito breeding places. Health authorities get data that has been time-stamped and geotagged and are emailed to the health authorities, which prompts a public health response. Although it was reported that managing the old paper-based and the new tablet-based systems was complicated for the PHIs, the issue is expected to be sorted when the digital version entirely replaces the paper-based system.

The research conducted by [39] investigated how Facebook users in Sri Lanka reacted to the COVID-19 outbreak. 2398 posts from 50 Facebook sources were carefully gathered and categorized. Both profiles and pages chosen randomly were made up of these sources. Those in command of these sources represented various ideologies and occupations. It was believed that a single person was in control of profiles, while a group of like-minded individuals was in charge of pages. The gathered information was then categorized and then normalized before analysis. Such analysis was carried out to produce major insights in areas including behavioral patterns, significant subjects, and the co-occurrence of topics.

The studies in [40-41] look into the suitability of several tablet applications for children with Autism Spectrum Disorder (ASD) in Sri Lanka. They determined that when developing applications to improve the social skills of Sri Lankan children with ASD, several functionalities, including multi-player functionalities, must be considered. Some essential functions include incorporating social elements while teaching other skills such as academics, adaptability to accommodate relevant cultural contexts, and monitoring progress. Arachchi et al. [42] studied how to create eLearning tools that improve the interactions of young adults with intellectual disabilities in Sri Lanka with web search tools.

Xinli, a social media app that proposes to predict emotions using an aggregated modalities method is proposed in [43]. They further predict personalized activities based on the user's mental state using reinforcement learning and also track the improvement of emotional state with the impact of recommended activities and social support groups.

Pets are seen as family members, and owners build significant ties with their animals. Pet owners confront challenges when their pet becomes ill, and there is no emergency veterinary care available or when the owner is unclear

about problems threatening the pet's life. A pet-specific social media platform (Zilla) is created in [44] that allows all parties involved in the pet industry to collaborate and promote better services and pet owner participation. Pet owners, pet lovers, and pet groups can use Zilla, a reliable method to contact veterinarians.

3.1.6 *Disaster and crisis management*

Over the years, Sri Lanka has been subjected to various hazards that have resulted in disasters [45]. One of the essential tactics for lessening and minimizing the effects of disasters is the establishment of appropriate communication channels for the efficient dissemination of information linked to disasters. The study in [46] investigated the variables that influence the use of social media during disasters to communicate essential information. An online questionnaire was circulated, and 352 responses were selected for analysis. Descriptive statistics were used to explore the response patterns, and logistic regression was conducted to identify the relationships between the dependent and predictor variables. The survey revealed that people aged 18-24 are more likely to use social media during disasters. Also, the results showed that both males and females use social media during disasters, and the previous experience of disasters has also increased the use of social media for obtaining information.

Risk and crisis communication are essential for enhancing public preparedness, knowledge, and response to a disaster because it increases public awareness and empowers citizens to act appropriately in a catastrophe. The COVID-19 pandemic-affected communities are examined in [47] at different stages and the influence of social media, policymakers, and health-focused organizations during the crisis. This study used a quantitative approach to evaluate how Sri Lankan politicians and top health-focused organizations used Facebook and LinkedIn to reach out to their constituents during the COVID-19 pandemic. The study's findings supported the applicability of the Crisis and Emergency Risk Communication Model (CERC) to enhance leadership and health-related organizations' crisis communication.

3.1.7 *Industrial sector*

A rising issue for brands across the breadth of digital marketing is choosing the correct social media influencer (SMI). The study in [48] investigated how Sri Lankan cosmetic and personal care companies choose to use social media influencers for their influencer marketing efforts on Instagram. Semi-structured interviews were conducted with staff in charge of overseeing influencer marketing efforts for cosmetic and personal care firms in Sri Lanka. Additionally, the primary criteria that Sri Lankan cosmetic and personal care companies used to select social media influencers were identified.

The impact of factors such as social media visibility, engagement, and social media influence on customers' brand preference is investigated in [49]. A sample of 186 young males and females was selected using convenience sampling from the Anuradhapura district, and a questionnaire was distributed. The validity of the collected data for factor analysis was proved using Kaiser-Meyer-Olkin (KMO) test. The name of the value loaded most frequently for each component was used to identify the factors. Following the factor analysis, Cronbach's alpha was used to assess the reliability of the factor components. Finally, multiple regression analysis was performed after confirming all latter mentioned tests. The $R2$ score was 0.320, which indicated that social media influence, engagement, and visibility accounted for 32.00% of the variance in brand choice.

The online shopping concept is currently prevalent in Sri Lanka. In the study by [50], the authors use classification models to predict the type and color of clothing based on the ethnicities of the consumers. Online vendors from Sri Lanka will be able to clearly comprehend customers' expectations and purchasing habits based on their ethnicities. Additionally, retailers may increase their sales, and online customers would enjoy a pleasant and appealing experience.

The majority of Sri Lankan business organizations are accustomed to traditional business practices. Still, most Sri Lankan organizations are unfamiliar with social media networking (SMN). According to the study in [51], a clear relationship exists between SMN usage patterns and business success factors such as revenue, customer satisfaction level, and customer growth rate. Using a disproportionate stratified random sampling technique, the study samples 25 purely SMN-based organizations, 50 partially SMN-based organizations, and 25 non-SMN-based organizations. Ten business success factors were picked as a set of Key Performance Indicators (KPIs) to measure the success of a business organization. Furthermore, the industry type and organizational scale influence the relationship between SMN usage

patterns and KPIs. Further study findings revealed that trading-related organizations generated higher revenues in pure SMN-based small-scale organizations, and service-related organizations generated the most revenue in partially SMN-based large-scale organizations. Based on the findings, the authors recommend that businesses use SMN as a supporting service because they can gain more financial and other benefits than traditional businesses.

The content in user profiles can extract thinking patterns such as interests, likes, hobbies, and thoughts. This data has immense potential for delivering personalized advertisements. Most commercial advertisements and web applications provide results based on previously searched queries most of the time and do not employ social media-based advertisement classification mechanisms. Hence the advertisements proposed by these systems are not based on user's preferences. In [52], an effective method is introduced for delivering personalized advertisements to the right people at the right time and a mechanism for pushing advertisements based on user preferences. They present the AdSeeker advertisement engine, which employs semantic analysis, ontology mapping, and an advertisement classification mechanism to achieve this goal. The user preference-based advertisement pushing method was built using Protege software to create ontology domains. The proposed system was tested for advertisement relevancy. Essentially, a survey was launched to test the Adseeker engine's accuracy with the eBay advertisement engine by comparing relevant adverts on the suggested system home feed. The results showed that the AdSeeker home feed provides the user with around 90% of relevant advertisements, whereas the eBay engine provided only 23%.

When it comes to purchasing things, customers rely heavily on online reviews. People's subjective thoughts and feelings about a product can be found in online reviews. These reviews can be gathered and analyzed from multiple social media platforms to rank products. In Sri Lanka, however, individuals frequently utilize Singlish (Sinhala-Engish) to comment on and review things. In this case, traditional text analysis approaches are difficult to employ. In [53], data from social media platforms on various brands in the automobile sector were extracted, and a mechanism was proposed to rank the vehicle brands in Sri Lanka based on social media comments posted in Singlish. When rating products, the study considered people's moods toward a specific brand and the frequency of comments for that brand. Furthermore, several ranking strategies based on various aspects such as market value, country of origin and second-hand market, vehicle performance, product features that people pay the most attention to in the automobile industry, and an overall ranking taking all of these aspects into account were proposed. After the rating system was implemented, 100,000 social media comments were extracted and annotated. The primary model was created using a convolutional neural network. Of the many approaches used to predict sentiment as part of the primary model, the random forest method provided the highest accuracy of 96.7, making it a more robust combination. The proposed model identified which vehicle type or brand has the highest and lowest market demand, allowing automobile manufacturers to understand better where a specific product stands out compared to other brands and customize their tactics accordingly.

As banks begin to rely on social media to conduct market sales promotions, it is critical to understand what aspects contribute to a campaign's effective launch and execution. The aforesaid topic is addressed in [54] using a case study of twelve market sales promotions from the Sri Lankan banking sector. They used a qualitative study using Straussian grounded theory to identify the factors to be examined when formulating research questions. Their study showed that the type of promotion, target audience, timing, and platform for communication, mode of communication, resources have a direct positive impact on the success of a sales promotion on social media.

Due to their concern about the epidemic, many customers have resorted to internet purchasing in recent years. In their efforts to offer and sell items online, online retailers can profit from a deeper understanding of online customer behaviour. Following this path, [55] investigate the role of trust in mediating the relationship between customers and online purchase intentions (PI) in the context of social media marketing relating to the fashion industry in Sri Lanka. A survey is conducted based on the conceptual model with online PI as the dependent variable and social motivation (SM) as the independent variable. Consumer trust in retailers on social media (TRSM) is regarded as a mediator. According to sources, consumers in the Colombo district are better educated and have higher income levels. Hence, the study's sample consisted of fashion product consumers residing there. Data was collected by distributing a Google form questionnaire, and 120 responses were collected. The data were analyzed using the IBM SPSS application. Statistical techniques such as reliability checking using Cronbach's Alpha and relationship analysis using regression and correlation calculations were performed. According to the findings of [55], SM has a strong positive impact on consumer trust and online PI. Furthermore, the results demonstrated that customers' online PIs are highly influenced by their trust level and customers' online PI is primarily mediated by trust in fashion shops.

3.2 *Social media analytics methods and tools*

Social media is especially essential in computational social science research, which analyzes problems using quantitative tools such as computational statistics, machine learning, data mining and so-called big data analytics techniques. Social media analytics is acquiring and analyzing data from social media data utilizing these techniques to acquire insights into people's sentiments, intentions, and behaviors. Our study reviewed several insightful methodological papers about new algorithms or models for deriving insights from social media data. Below we summarize these papers based on different aspects of the data analytics process. The main focus is highlighting the tools and techniques used to acquire data and gain insights into social media content.

The research in [56] collected news posts from the 'Lankadeepa' Facebook page. Data is collected from Facebook using the Facebook Graph API Explorer version 2.11. Graph API is a low-level HTTP-based API that can query data programmatically. Since the API is HTTP-based, it can be used with any language with an HTTP library. It can be executed directly on a browser by browsing for the Facebook developer's URL. A GET request can be used to obtain necessary information regarding a node. The results obtained through a GET request are in JSON format, and generally, they will not be received through a single response. Hence, the responses are paginated by default. The Facebook Software Development Kit (SDK) for PHP, a library with powerful features that enable PHP developers to easily make requests to the Graph API with the integration of Facebook Login, has been used in this study. A single PHP script retrieves the required data and transforms it into JSON format. The JSON data obtained is parsed using OpenRefine, an open-source tool used to handle extensive, messy data, and is exported as an Excel sheet. Unescaped double quotes in JSON description fields were found while parsing the data and eliminated.

NodeXL provides a complete solution for collecting, analyzing, and visualizing social media network data using the Microsoft spreadsheet application. This application enables anyone, regardless of programming experience, to import data from Twitter, Flickr, and Wikipedia and provide network-based content analysis insights. Before 2019, NodeXL could also access the Facebook API. Unfortunately, when Facebook restricted access to public postings in September 2019, the NodeXL program eliminated the Facebook option. The study in [24] was able to utilize NodeXL to extract data from two Facebook sites, 'Sri Lanka's Killing Fields: A Channel 4 Investigation' and 'Sri Lankans Hate Channel 4'. Nodes in both networks represented individuals who liked or commented on posts, and the edges were created based on co-likes and/or co-comments. In the networks created, the Facebook users' first name was used to predict their ethnicity.

Owing to Facebook's constraints on retrieving public data via the Facebook API, [39] manually extracted publicly published content, anonymized it, and used it for their research. They analyze content posted by a representative selection of Sri Lankan Facebook users. Data collection comprises posts from 50 randomly selected profiles and pages spanning about 60 days. Although this approach of manual data extraction limited the number of sources that could be taken practically, the authors argue that it allowed for a considerably more extensive study that included not only image-based postings but also those that were indirectly connected to their research field.

The study conducted by [57] aimed to assess the level of accuracy of news shared by social media users as status updates in the news feed. The categorization of news updates is determined by analyzing the semantics of each update, allowing for the identification of various categories such as political, weather, sports, economic, and others. The system then drags the URLs of online news articles that match the social media news update content from RSS feeds, web crawling, and custom Google searches. After dragging the URLs, their contents are extracted using web scraping and summarized using automatic extractive summarizing techniques. This summary is used to calculate a confidence score, which indicates the percentage of the relevance of that news article to the social media news post. This relevance score from all URLs is used to calculate the accuracy score for the social media news post so the user can get an idea of how far this news is mentioned in different online news sites and, thus, whether the particular post is actual or not.

Sinhala is the native language of Sri Lanka. In [58], a system is created and deployed for extracting Sinhala text from social media text images. Because Sinhala characters are distinct in shape, existing approaches for other languages needed to be adaptable. The authors concentrated on Sinhala's touching and overlapping character identification, touching and overlapping chatter segmentation, and character recognition. On the other hand, there are few approaches to analyzing social media text data in the Sinhala language. Jayasuriya et al. [59] proposed a new approach to lexicon library construction for sentiment classification of Sinhala comments on Youtube. The study also uses a machine learning approach in conjunction with a lexicon-based system to improve the classification accuracy of informal

language in social media comments.

Natural Language Processing studies how computers understand and manipulate human language. Automatic summarization, translation, named entity recognition, relationship extraction, sentiment analysis, speech recognition, and topic segmentation are some tasks that can be performed using Natural Language Processing (NLP) [60]. Semantic analysis is applied in [57] in their research to match social media news posts with online news articles to determine the meanings of web content. They used Python to extract data from the social media application programming interfaces. The Newspaper library in Python is used for web scraping, where the contents of web pages relevant to the project are extracted.

Most people use social networking platforms to express their feelings and thoughts, generating large amounts of data. Thus, effective methods are necessary to extract meaningful sentiment information from such user data from social networks [61]. It is more difficult to analyse the sentiments of less commonly used languages, such as Sinhala, due to a lack of necessary tools, such as WordNet, special lexicons, and coding libraries. To expand the subjective lexicon for sentiment classification, Levenshtein ratio analysis was used combined with the Sinhala-English dictionary [62]. Stop-word removal was also included in the method. The study's findings revealed that combining machine-learning and lexicon-based techniques produced more accurate sentiments than either machine-learning-based or lexicon-based methods alone.

In [19], a methodology for sentiment analysis of tweets that equally considers both Sinhala and Tamil languages is proposed. Extracted tweets are first subjected to general preprocessing steps such as removing URLs, hashtags, duplicated tweets, empty tweets and emoticon-only tweets, numbers, punctuation, and replacing all emoticons with their sentiment. In the second phase, Part of Speech (POS) tagging, Tokenization and stop word removal were committed using tools unique to each language. In the third phase, the tokenized and POS-tagged, stop-words removed tweets were translated to English using Google Translator, which achieved 76% accuracy for Sinhala tweets and 70% accuracy for Tamil tweets. The next step was feature extraction, which involved converting the preprocessed dataset to a numerical form. Count Vectors and Term Frequency-Inverse Document Frequency (TF-IDF) are the methodologies investigated here. The sentiment classification model was built using a hybrid technique in the final step. Specifically, the feature vectors were passed to SentiWordNet to extract sentiment values before being fed into the machine learning model for classification. For both languages, six machine-learning models were tested. Eventually, the model constructed with the Nave Bayes Classification algorithm recognized religious extremism tweets in Sinhala, while the Random Forest method performed best in Tamil.

In [20], a rule-based approach is used to extract features from Twitter social media tweets in the Sinhala language and then used to detect whether a particular tweet contains offensive or insulting content targeting a specific person. Each tweet was assessed for the following rules: (i) the percentage of bad words in the tweet is greater than 10%, (ii) the tweet contains a combination of a first-person pronoun, a bad word and a second-person pronoun, (iii) the tweet contains a combination of second person pronoun with a bad word, (iv) the tweet contains a combination of the third-person pronoun with a bad word, and (v) the tweet contains a combination of the first-person pronoun, bad word and a third-person pronoun. Based on the satisfaction of these rules, binary features were calculated for each tweet and saved in a Python pandas data frame to train the classification algorithm.

In Sri Lanka, people use English (Latin) letters to write and type Sinhala words. It can be found on various social media platforms and in text messages. This transliteration technique is known as "Singlish". The term "Singlish" is mainly associated with Singapore's mixed English language, while in Sri Lanka, "Singlish" refers to transliterated Sinhalese [63]. To analyze and process text in Singlish, it is necessary to convert Singlish text into English or Sinhala language. Such a transliteration model is presented in [63]. The authors utilized the N-gram approach with a text corpus containing comments from YouTube videos to create this model. The probability is used in the N-gram to consider the sequence of n words from a given sample text. There are three types of N-grams: unigram, bigram, and trigram. The N-gram tagger is implemented using Natural Language Toolkit (NLTK), a Python natural language processing library package that supports part of speech tagging. This package has various taggers that can be trained with a specified corpus. N-gram taggers were trained with a Singlish training corpus that contains a Sinhala term for each Singlish word to execute the Singlish transliterator. The proposed transliteration method outperformed existing methods by achieving 62% word accuracy and 77% letter-by-letter accuracy.

People use social media to express their opinions on various topics, such as politics, religion, the economy, and

sports, through images. The text in these images conveys critical information. Extracting these texts to analyze social media content and decision-making is beneficial. On the other hand, digital images with sentences include touching and overlapping characters. It is critical to segment these characters before employing algorithms to identify them. According to the study [58], the Convolutional Neural Network (CNN) based approach showed optimal results for character identification, while the Background thinning and Ridge Regression approach was best for character segmentation.

4. Analysis and findings

Figure 1 summarizes the distribution of the reviewed articles according to the application areas where social media research is utilized for decision-making. The majority of the articles (28%) focused on social media's influence on issues such as racism, addiction, hate speech, and cyberbullying. These studies shed light on the negative aspects and challenges associated with social media usage. Another significant theme identified in the literature (21%) was the examination of how social media can enhance brand awareness, customer attraction, and customer retention. These studies explored the positive impact of social media on businesses and marketing strategies. Approximately 18% of the articles discussed the role of social media in education and highlighted its potential to improve the quality of educational services. These studies explored the use of social media in facilitating learning, engaging students, and creating interactive educational environments. Additionally, 15% of the reviewed articles focused on the applications of social media in the healthcare industry, particularly in areas such as professional education, patient care, and public health programs. The reviewed articles also addressed social media applications in other domains. Approximately 5% of the articles discussed the utilization of social media in disaster and crisis management, highlighting its role in emergency communication and response. Similarly, 5% of the articles examined the impact of social media on politics, exploring its influence on political campaigns, public opinion, and mobilization. Tourism-related studies accounted for 8% of the reviewed articles, discussing the use of social media platforms for destination marketing and tourism promotion.
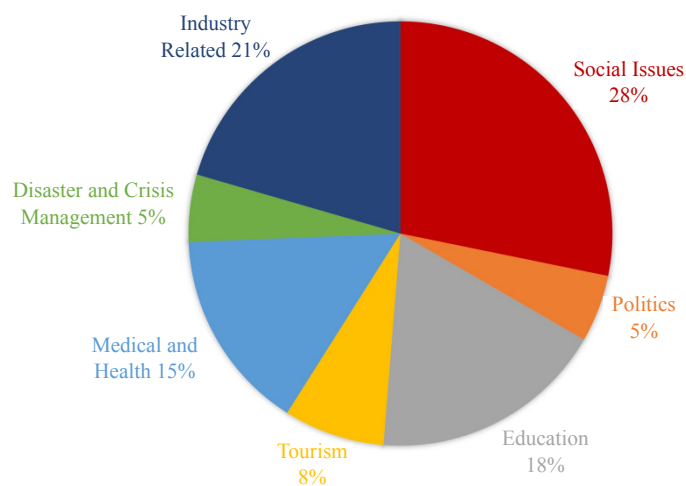

Figure 1. Distribution of papers based on the application area

Table 1 provides a comprehensive summary of the reviewed papers in Section 3.1, categorizing them according to their primary decision-making area. It outlines the key findings, main themes, and trends identified in each application area. This table serves as a valuable resource for understanding the diverse range of research conducted in Sri Lanka and its implications for decision-making in various domains.

**Table 1.** Summary of reviewed articles by application area and decision-making scenario

Application Area	Decision Making Scenario	Identified Key Themes and Trends
Social Issues	Impact of social media on adolescents' academic and social development [13], the detection and prevention of harmful content such as racism and hate speech ([15-16, 64]), the factors associated with crimes targeting youths on social media [17], the prevalence of religious extremism [19], and the issues of cyberbullying and psychological stress ([20-23]), cross-ideology exposure on social networks [24]	A complex interplay exists between social media and societal dynamics in Sri Lanka, emphasizing the need for effective measures to ensure responsible and safe social media usage.
Politics	Changes in political perceptions [26], sinhala news dissemination [27]	There is a decrease in public engagement with politics on social media, a humorous approach towards political content, and an increase in user reactions over time. Additionally, the studies highlighted the influence of Sinhala Twitter data on news dissemination, providing insights into the "news value" of tweets.
Education	Utilization of OER [29], e-learning patterns [30], social media in enhancing English proficiency [32], benefits of academic Whatsapp groups [33]	Social media are utilized for academic purposes, accessing open educational resources, and facilitating communication. However, concerns about addiction and information credibility exist. E-learning platforms integrated with conventional teaching methods show promise, while social media can enhance English language learning and learner engagement. Challenges related to internet accessibility need to be addressed for equitable access to digital education.
Tourism	Tourist opinion identification [34], social tag analysis [36], influence of social media marketing [37]	Social media marketing initiatives are influential in establishing brand equity for Sri Lanka as a tourism destination, with factors like entertainment and customization contributing positively, while interaction and trendiness have a negative impact.
Medical and Health Services	Dengue surveillence [38], reaction to the COVID-19 outbreak [39], social skills related to ASD [40], emotion detection [43], veterinary-based [44]	There is an increasing reliance on technology and social media for various aspects of the health domain such as disease surveillance, support individuals with specific needs, and facilitate collaboration to improve animal health.
Disaster and Crisis Management	Information dissemination [46], risk and crisis communication [47]	The use of social media platforms, particularly among young adults, has been identified as a significant channel for disseminating information during disasters.
Industrial Sector	Marketing [48], brand preference [49], social media networking [51], personalized advertisements [52], automobile sector [53], banking [54], fashion industry [50, 55]	Businesses in the cosmetic and personal care industry are utilizing social media influencers to enhance their marketing efforts. Factors such as social media visibility, engagement, and trust influence customers' brand preference and online purchase intentions.

Regarding social media platform usage from Section 3.2, Facebook and Twitter emerged as the most widely studied platforms, each accounting for 28% of the articles. However, it was observed that research focusing on Facebook postings was predominantly published on or before 2019, potentially due to the restrictions introduced by Facebook on API-based data retrieval in September 2019. In contrast, Twitter has no such restrictions and has a wealth of resources that can be used to acquire and analyze Twitter data [65]. Also, a single Twitter post can contain many attributes, such as hashtags, URLs, emotions and mentions. This makes the data retrieval process simple for social media analytics based on the necessary criteria. Moreover, Twitter is more popular for gauging people's perceptions of different events, as statements on Twitter are shortened and frequently updated [66]. YouTube, with its 2 billion monthly active users and extensive engagement, ranked as the second most popular social media platform in the reviewed articles. Figure 2 illustrates the distribution of articles based on social media platform usage, highlighting the prominence of Facebook, Twitter, and YouTube. Notably, YouTube research in Sri Lanka represented a smaller proportion compared to Facebook and Twitter, despite its significant user base and potential insights available through comments and engagement.
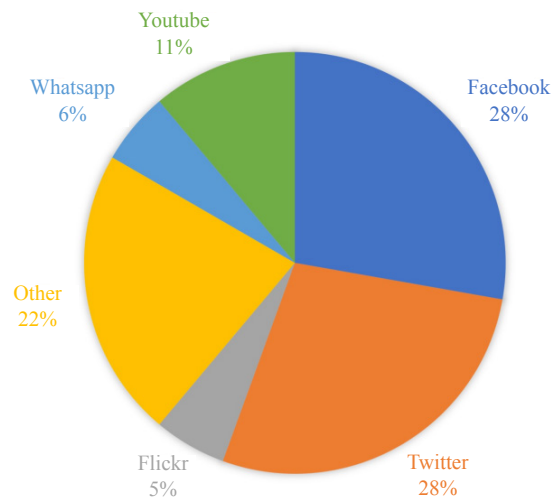
Figure 2. Distribution of articles based on social media platforms used

As depicted in Figure 3, the analysis of the reviewed articles in Section 3 also revealed a range of approaches employed in social media research in Sri Lanka. Survey methods were prevalent, accounting for 30% of the studies. These surveys employed various sampling techniques, questionnaire designs, and data processing methodologies, often utilizing software such as IBM SPSS for data analysis. Text analytics played a significant role, featured in 30% of the articles, enabling the extraction of insights and patterns from unstructured text data. Other approaches included Sinhala text processing (10%), sentiment analysis (10%), topic modeling (4%), and social network analysis (2%), each providing unique perspectives on social media research. Additionally, 12% of the publications proposed approaches for developing social media systems, contributing to the advancement of social media platforms and tools.
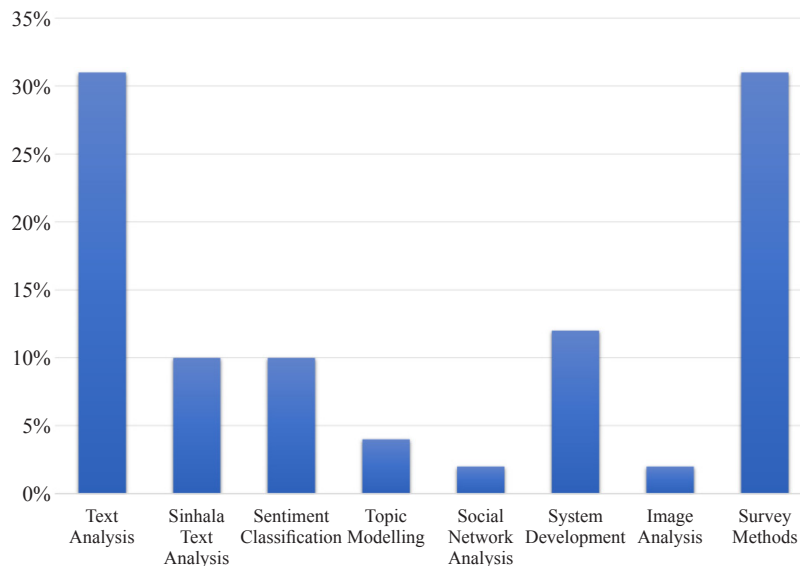
Figure 3. Methods used in papers to research social media

Overall, the comprehensive analysis of the reviewed articles showcases the breadth and depth of social media

research in Sri Lanka, covering various application domains, social media platforms, and methodological approaches. These findings contribute to a richer understanding of social media's impact and potential in the Sri Lankan context, offering insights for researchers, practitioners, and policymakers seeking to leverage social media effectively.

5. Conclusion and future research

The study's objective is to provide information on research conducted on social media usage in Sri Lanka over the last decade. Based on the studies reviewed, this research contributes to several major areas. The study begins by identifying the primary application areas of social media research. According to the articles analyzed, social media analytics are employed in various fields, including local government services, the industrial sector, tourism, politics, social issues, medical and health-related services, disaster and crisis management, online media, and supply chain activities. The studies show how social media platforms can be used to make decisions in these industries. According to peer-reviewed articles, investigating social issues has received more attention than other application domains. The study also helps in the identification of platforms that are utilized for social media analytics. The majority of the research was carried out on Facebook and Twitter. YouTube, WhatsApp, Flickr, and Blogs were among the other platforms. Finally, the research identifies various approaches and technologies utilized in social media analytics. Text analysis, Sinhala text processing, sentiment classification, topic modeling, image analysis, and social network analysis are some of the analytics that has been presented. According to peer-reviewed studies, text analysis is the most employed than other types of techniques. Many surveys were also conducted to better understand the many aspects of social media usage and its impact.

After reviewing all articles for social media research, some common challenges were identified. Most survey researchers have recognized obtaining a representative sample using probabilistic sampling methods as challenging. Because non-probability sampling produces significant sampling errors and biased results, survey findings could not be extended to the greater Sri Lankan population. Many problems were observed for those researchers who conducted social media scraping and analytics regarding data collecting, pre-processing, and analysis. Most researchers faced numerous obstacles when analyzing Sinhala text data, owing to a scarcity of Sinhala resources and the inclusion of Singlish terminology in the text.

The applications of social media usage in those examined publications indicate the benefits and importance of this social media research in providing insight into various applications. Because of its benefits, research on social media usage must be conducted reasonably and manageably for the outcomes to be beneficial. Looking ahead, there are several promising avenues for future research in the field of social media analytics in Sri Lanka. Firstly, there is a need for studies that delve deeper into the specific impact and implications of social media usage in areas such as politics, health, and tourism. Understanding how social media platforms like Twitter, YouTube, and Instagram are utilized in these sectors can provide valuable insights for policymakers, marketers, and public health professionals. Additionally, given the challenges associated with analyzing Sinhala text data and the inclusion of Singlish terminology, future research could focus on developing robust language processing techniques and resources specifically tailored to the Sri Lankan context. This would enhance the accuracy and effectiveness of sentiment analysis, topic modeling, and other text analytics approaches. Furthermore, as social media platforms continue to evolve, there is a need to explore emerging platforms and their implications for social media analytics. Platforms like TikTok, Snapchat, or emerging local platforms may offer unique opportunities and challenges that warrant investigation.

In summary, this review significantly contributes to the existing body of knowledge by providing a comprehensive overview and analysis of social media usage research in Sri Lanka. It fills a critical gap in the literature by focusing specifically on the Sri Lankan context, which is essential for understanding the unique dynamics and challenges of social media in the country. By synthesizing and consolidating the existing research, this study offers new insights into the patterns, trends, and applications of social media in Sri Lanka, enabling a deeper understanding of its impact on society, the economy, and decision-making processes. The review also highlights the significance of social media research in Sri Lanka, emphasizing its relevance for decision-makers, practitioners, and researchers who can benefit from the study's findings and recommendations. Overall, this review provides a valuable resource that enhances our understanding of social media usage in Sri Lanka and opens avenues for further research and exploration in this field.

Acknowledgments

Conflict of interest

The authors declare no competing financial interest.

References

[1] Rathore AK, Kar AK, Ilavarasan PV. Social media analytics: Literature review and directions for future research. *Decision Analysis*. 2017; 14(4): 229-249.

[2] Madila SS, Dida MA, Kaijage S. A review of usage and applications of social media analytics. *Journal of Information Systems Engineering and Management*. 2021; 6(3): em0141.

[3] Bern DJ. Writing a review article for psychological bulletin. *Psychol Bull*. 1995; 118(2): 172-177.

[4] Palmatier RW, Houston MB, Hulland J. Review articles: Purpose, process, and structure. *Journal of the Academy of Marketing Science*. 2018; 46: 1-5. Available from: https://doi.org/10.1007/s11747-017-0563-4.

[5] Nanayakkara AC, Kumara BTGS, Rathnayaka RMKT. A survey of finding trends in data mining techniques for social media analysis. *Sri Lanka Journal of Social Sciences and Humanities*. 2021; 1(2): 37.

[6] Batrinca B, Treleaven PC. Social media analytics: A survey of techniques, tools and platforms. *AI Soc*. 2015; 30(1): 89-116.

[7] Keles B, McCrae N, Grealish A. A systematic review: The influence of social media on depression, anxiety and psychological distress in adolescents. *International Journal of Adolescence and Youth*. 2020; 25(1): 79-93. Available from: https://doi.org/10.1080/02673843.2019.1590851.

[8] Ahmed YA, Ahmad MN, Ahmad N, Zakaria NH. Social media for knowledge-sharing: A systematic literature review. *Telematics and Informatics*. 2019; 37: 72-112.

[9] Chen J, Wang Y. Social media use for health purposes: Systematic review. *Journal of Medical Internet Research*. 2021; 23(5): e17917. Available from: https://doi.org/10.2196/17917.

[10] Zhuravskaya E, Petrova M, Enikolopov R. Political effects of the internet and social media. *Annual Review of Economics*. 2020; 12: 415-438. Available from: https://doi.org/10.1146/annurev-economics-081919-050239.

[11] Zhang C, Fan C, Yao W, Hu X, Mostafavi A. Social media for intelligent public information and warning in disasters: An interdisciplinary review. *International Journal of Information Management*. 2019; 49: 190-207. Available from: https://doi.org/10.1016/j.ijinfomgt.2019.04.004.

[12] Bhimani H, Mention AL, Barlatier PJ. Social media and innovation: A systematic literature review and future research directions. *Technol Forecast Soc Change*. 2019; 144: 251-269.

[13] Wickramanayake L. Social media use by adolescent students of Sri Lanka: Impact on learning and behavior. *Global Knowledge, Memory and Communication*. 2022; 71(1-2): 70-85.

[14] Dias DS, Welikala MD, Dias NGJ. Identifying racist social media comments in sinhala language using text analytics models with machine learning. *2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer)*. Colombo, Sri Lanka: IEEE; 2018. p.1-6. Available from: https://doi.org/10.1109/ICTER.2018.8615492.

[15] Ruwandika NDT. Identification of hate speech in social media. *2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer)*. Colombo, Sri Lanka: IEEE; 2018. p.273-278. Available from: https://doi.org/10.1109/ICTER.2018.8615517.

[16] Shalinda JADU, Munasinghe L. Hate words detection among Sri Lankan social media text messages. *2022 International Research Conference on Smart Computing and Systems Engineering (SCSE)*. Colombo, Sri Lanka: IEEE; 2022. p.55-60. Available from: https://doi.org/10.1109/SCSE56529.2022.9905129.

[17] Nalaka S, Diunugala H. Factors associating with social media related crime victimization: Evidence from the undergraduates at a public university in Sri Lanka. *International Journal of Cyber Criminology*. 2020; 14(1): 174-

184.

[18] Loudenback T, Jackson A. *The 10 most critical problems in the world, according to millennials*. Business Insider; 2018. Available from: https://www.businessinsider.com/world-economic-forum-world-biggest-problems-concerning-millennials-2016-8#10-lack-of-economic-opportunity-and-unemployment-121-1 [Accessed 17th Feb 2023].

[19] Fernando A, Wijayasiriwardhane TK. Identifying religious extremism-based threats in Sri Lanka using bilingual social media intelligence. *2020 International Research Conference on Smart Computing and Systems Engineering (SCSE)*. Colombo, Sri Lanka: IEEE; 2020. p.103-110. Available from: https://doi.org/10.1109/SCSE49731.2020.9313010.

[20] Amali HMAI, Jayalal S. Classification of cyberbullying sinhala language comments on social media. *2020 Moratuwa Engineering Research Conference (MERCon)*. Moratuwa, Sri Lanka: IEEE; 2020. p.266-271. Available from: https://doi.org/10.1109/MERCon50084.2020.9185209.

[21] Sewwandi D, Perera K, Sandaruwan S, Lakchani O, Nugaliyadde A, Thelijjagoda S. Linguistic features based personality recognition using social media data. *2017 6th National Conference on Technology and Management (NCTM)*. Malabe, Sri Lanka: IEEE; 2017. p.63-68. Available from: https://doi.org/10.1109/NCTM.2017.7872829.

[22] Lanerolle TY, Perera WTH, Andrado YDS, Wickramasinghe WAPC, Bandara PS, Kishara J. Measuring psychological stress rate using social media posts engagement. *2022 14th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*. Ploiesti, Romania: IEEE; 2022. p.1-6. Available from: https://doi.org/10.1109/ECAI54874.2022.9847471.

[23] Perera A, Fernando P. Accurate cyberbullying detection and prevention on social media. *Procedia Comput Sci*. 2021; 181: 605-611. Available from: https://www.sciencedirect.com/science/article/pii/S1877050921002507.

[24] Rathnayake C, Suthers DD. 'Enclaves of exposure': A conceptual viewpoint to explore cross-ideology exposure on social network sites. *Soc Sci J*. 2019; 56(2): 145-155. Available from: https://www.sciencedirect.com/science/article/pii/S0362331918301241.

[25] Wikipedia. *Politics in Sri Lanka*. Wikipedia; 2022. Available from: https://en.wikipedia.org/wiki/Politics_of_Sri_Lanka [Accessed 27th Jan 2023].

[26] Sally MS, Wickramasinghe MIE. A trend analysis on Sri Lankan politics based on facebook user reactions. *Technology in Society*. 2020; 62: 101321. Available from: https://doi.org/10.1016/j.techsoc.2020.101321.

[27] Madhushika MD, Ahangama S, Rajapaksha DA. Analyzing the impact of social media on sinhala news dissemination in mass media. *2022 2nd International Conference on Advanced Research in Computing (ICARC)*. Belihuloya, Sri Lanka: IEEE; 2022. p.177-182. Available from: https://doi.org/10.1109/ICARC54489.2022.9754025.

[28] Wijekoon MWAS. Comparative analysis on social media and its positive and negative impact on Sri Lankan culture. *International Journal of Research in Economics and Social Sciences (IJRESS)*. 2017; 7(2): 180-185. Available from: http://euroasiapub.org.

[29] Hettige S, Dasanayaka E, Ediriweera DS. Student usage of open educational resources and social media at a Sri Lanka Medical School. *BMC Med Educ*. 2022; 22(1): 35. Available from: https://doi.org/10.1186/s12909-022-03106-2.

[30] Yapa YMMM, Dilan MMNS, Karunaratne WCD, Widisinghe CC. Computer literacy and attitudes towards e-learning among Sri Lankan medical students. *Sri Lanka Journal of Bio-Medical Informatics*. 2012; 3(3): 82-96.

[31] Nasome V, Malavige O, Costa M, Jayasinghe B, Karunasena A, Samarakoon U. KnowMore: Social media based student centric e-learning platform with machine learning approaches. *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*. Mumbai, India: IEEE; 2022. p.1-7. Available from: https://doi.org/10.1109/I2CT54291.2022.9824235.

[32] Halik AF, Rafeena HM. The role of social media in learning english as a second language: A study based on the students of National Apprentice & Industrial Training Authority (NAITA), Trincomalee, Sri Lanka. *International Journal of Scientific and Research Publications*. 2020; 10(7): 162-166. Available from: https://doi.org/10.29322/IJSRP.10.07.2020.p10321.

[33] Lakmali AAI, Abeysekera N, Silva DACS. Effectiveness of customer social participation for academic purposes: a case of informal WhatsApp groups. *Asian Association of Open Universities Journal*. 2021; 16(3): 326-343. Available from: https://doi.org/10.1108/AAOUJ-08-2021-0093.

[34] Abeysinghe S, Manchanayake I, Samarajeewa C, Rathnayaka P, Walpola MJ, Nawaratne R, et al. Enhancing decision making capacity in tourism domain using social media analytics. *2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer)*. Colombo, Sri Lanka: IEEE; 2019. p.369-375. Available from: https://doi.org/10.1109/ICTER.2018.8615462.

[35] Chen Y, Sherren K, Smit M, Lee KY. Using social media images as data in social science research. *New Media Soc*.

2023; 25(4): 849-871. Available from: http://journals.sagepub.com/doi/10.1177/14614448211038761.

[36] Bandara T, Bandara TP. Whale watching in Sri Lanka: Understanding the metadata of crowd-sourced photographs on FlickrTM social media platform. *Sri Lanka Journal of Aquatic Sciences*. 2019; 24(2): 41.

[37] Hilal MIM. Impact of social media marketing efforts on Destination's brand equity: A study among International Tourists arriving to Sri Lanka. *2019 3rd International Conference on Computing and Communications Technologies (ICCCT)*. Chennai, India: IEEE; 2019. p.59-62. Available from: https://doi.org/10.1109/ICCCT2.2019.8824976.

[38] Lwin MO, Vijaykumar S, Rathnayake VS, Lim G, Panchapakesan C, Foo S, et al. A social media mHealth solution to address the needs of dengue prevention and management in Sri Lanka. *J Med Internet Res*. 2016; 18(7): e149. Available from: https://doi.org/10.2196/jmir.4657.

[39] Lenadora D, Gamage G, Haputhanthri D, Meedeniya D, Perera I. *Exploratory analysis of a social media network in Sri Lanka during the COVID-19 Virus Outbreak*. arXiv:2006.07855 [cs.SI]. 2020. Available from: https://doi.org/10.48550/arXiv.2006.07855.

[40] Soysa AI, Al Mahmud A. Assessing tablet applications focused on social interactions: What functionalities do Sri Lankan practitioners want for children with ASD? *Proceedings of the 30th Australian Conference on Computer-Human Interaction*. New York, NY, USA: Association for Computing Machinery; 2018. p.32-41. Available from: https://doi.org/10.1145/3292147.3292164.

[41] Soysa AI, Al Mahmud A. Beyond digital displays: Design considerations for tablet applications targeting children with ASD in Sri Lanka. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery; 2018. p.1-6. Available from: https://doi.org/10.1145/3170427.3188524.

[42] Arachchi TK, Sitbon L, Zhang J, Gamage R, Hewagamage P. Enhancing internet search abilities for people with intellectual disabilities in Sri Lanka. *ACM Trans Access Comput*. 2021; 14(2): 1-36. Available from: https://doi.org/10.1145/3460202.

[43] Kalansooriya S, Kaluarachchi A, Weerawickrama C, Nanayakkara D, Kasthurirathna D, Adeepa D. "xīnlǐ" The social media app to replenish mental health with the aid of an egocentric network. *2022 IEEE 10th Region 10 Humanitarian Technology Conference (R10-HTC)*. Hyderabad, India: IEEE; 2022. p.348-354. Available from: https://doi.org/10.1109/R10-HTC54060.2022.9929642.

[44] Kasun R, Mahesh LGH, Yapa YADI, Suwendra SMSD, Kodagoda N, Suriyawansa K. Zilla: An animal based social media platform. *2019 International Conference on Advancements in Computing (ICAC)*. Malabe, Sri Lanka: IEEE; 2019. p.267-272. Available from: https://doi.org/10.1109/ICAC49085.2019.9103368.

[45] Ministry of Health. *Epidemiology Unit*. Global Health Data Exchange; 2022. Available from: http://www.epid.gov.lk/web/index.php?option=com_content&view=article&id=149&Item [Accessed 27th Jan 2023].

[46] Jayathilaka HADGS, Siriwardana CSA, Amaratunga D, Haigh RP, Dias N. Investigating the variables that influence the use of social media for disaster risk communication in Sri Lanka. In: Dissanayake R, Mendis P, Weerasekera K, De Silva S, Fernando S. (eds.) *ICSBE 2020*. Singapore: Springer Singapore; 2022. p.259-276.

[47] Nagahawatta R, Kaluarachchi C, Warren M, Sedera D. Strategic use of social media in covid-19 pandemic. pandemic management by Sri Lankan Leaders and Health Organisations Using the CERC Model. *Australasian Conference on Information Systems Nagahawatta*. Melbourne, Australia; 2022. p.1-11. Available from: https://aisel.aisnet.org/acis2022/99.

[48] Lakmal KGP, Hettiarachchi HAH, Anuranga BKHD. How brands opt social media influencers for influencer marketing on instagram: A study on Sri Lankan beauty & personal care brands. *Sri Lanka Journal of Marketing*. 2019; 5(2): 135-158.

[49] Kumaradeepan V. Nexus between social media and brand preference of smart mobile phones: An empirical study in Sri Lanka. *Journal of Asian Finance*. 2021; 8(8): 241-249.

[50] Ginige TNDS, Mahima KTY. Ethnicity based consumer buying behavior analysis and prediction on online clothing platforms in Sri Lanka. *2021 the 5th International Conference on Information System and Data Mining*. New York, NY, USA: Association for Computing Machinery; 2021. p.121-127. Available from: https://doi.org/10.1145/3471287.3471291.

[51] Thelijjagoda S, Hennayake TM. The Impact of Social Media Networking (SMN) towards business environment in Sri Lanka. *2015 Fifteenth International Conference on Advances in ICT for Emerging Regions (ICTer)*. Colombo, Sri Lanka: IEEE; 2015. p.207-213. Available from: https://doi.org/10.1109/ICTER.2015.7377690.

[52] De Silva H, Jayasinghe P, Perera A, Pramudith S, Kasthurirathna D. Social media based personalized advertisement engine. *2017 11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*. Malabe, Sri Lanka: IEEE; 2017. p.1-6. Available from: https://doi.org/10.1109/SKIMA.2017.8294102.

[53] Warnakulasooriya A, Sandanayake T, Gamps W, Radw R, Kasn S. Automobile product ranking based on the

singlish comments in social media platforms. *2022 7th International Conference on Information Technology Research (ICITR)*. Moratuwa, Sri Lanka: IEEE; 2022. p.1-6. Available from: https://doi.org/10.1109/ICITR57877.2022.9992795.

[54] Bandara N, Bandara HMND. Critical factors for market sales promotion on social media in banking sector. *2017 Moratuwa Engineering Research Conference (MERCon)*. Moratuwa, Sri Lanka: IEEE; 2017. p.491-496. Available from: https://doi.org/10.1109/MERCon.2017.7980533.

[55] Karunasingha A, Abeysekera N. The mediating effect of trust on consumer behavior in social media marketing environments. *South Asian Journal of Marketing*. 2022; 3(2): 135-149. Available from: https://doi.org/10.1108/SAJM-10-2021-0126.

[56] Sally MS, Wickramasinghe MIE. A trend analysis on Sri Lankan politics based on facebook user reactions. *Technol Soc*. 2020; 62: 101321. Available from: https://www.sciencedirect.com/science/article/pii/S0160791X19306256.

[57] Chandrathlake R, Ranathunga L, Wijethunge S, Wijerathne P, Ishara D. A semantic similarity measure based news posts validation on social media. *2018 3rd International Conference on Information Technology Research (ICITR)*. Moratuwa, Sri Lanka: IEEE; 2018. p.1-6. Available from: https://doi.org/10.1109/ICITR.2018.8736136.

[58] Samarajeewa S, Ranathunga L. An approach for resolving double character segmentation in sinhala social media text images. *2020 From Innovation to Impact (FITI)*. Colombo, Sri Lanka: IEEE; 2020. p.1-6. Available from: https://doi.org/10.1109/FITI52050.2020.9424892.

[59] Jayasuriya P, Munasinghe R, Thelijjagoda S. Sentiment classification of sinhala content in social media: An ensemble approach. *2021 IEEE 16th International Conference on Industrial and Information Systems (ICIIS)*. Colombo, Sri Lanka: IEEE; 2021. p.140-145. Available from: https://doi.org/10.1109/SCSE49731.2020.9313023.

[60] Mengoni P, Santucci V. Special issue "Recent trends in natural language processing and its applications." *Applied Sciences*. 2023; 13(12): 7284. Available from: https://www.mdpi.com/2076-3417/13/12/7284.

[61] Meena G, Mohbey KK, Indian A. Categorizing sentiment polarities in social networks data using convolutional neural network. *SN Comput Sci*. 2022; 3(2): 116. Available from: https://link.springer.com/10.1007/s42979-021-00993-y.

[62] Jayasuriya P, Ekanayake S, Munasinghe R, Kumarasinghe B, Weerasinghe I, Thelijjagoda S. Sentiment classification of Sinhala content in social media. *2020 International Research Conference on Smart Computing and Systems Engineering (SCSE)*. Colombo, Sri Lanka: IEEE; 2020. p.136-141. Available from: https://doi.org/10.1109/SCSE49731.2020.9313023.

[63] Liwera WMP, Ranathunga L. Combination of trigram and rule-based model for singlish to sinhala transliteration by focusing social media text. *2020 From Innovation to Impact (FITI)*. Colombo, Sri Lanka: IEEE; 2020. p.1-5. Available from: https://doi.org/10.1109/FITI52050.2020.9424880.

[64] Dias DS, Welikala MD, Dias NGJ. Identifying racist social media comments in sinhala language using text analytics models with machine learning. *2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer)*. Colombo, Sri Lanka: IEEE; 2019. p.1-6. Available from: https://doi.org/10.1109/ICTER.2018.8615492.

[65] Ahmed W. *Using Twitter as a data source an overview of social media research tools*. The London School of Economics and Political Science; 2021. Available from: https://blogs.lse.ac.uk/impactofsocialsciences/2021/05/18/using-twitter-as-a-data-source-an-overview-of-social-media-research-tools-2021/ [Accessed 22th Feb 2023].

[66] Meena G, Mohbey KK, Kumar S, Lokesh K. A hybrid deep learning approach for detecting sentiment polarities and knowledge graph representation on monkeypox tweets. *Decision Analytics Journal*. 2023; 7: 100243. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2772662223000838.