



COVID-19 vaccine hesitancy: a social media analysis using deep learning

Serge Nyawa¹ · Dieudonné Tchuenté¹ · Samuel Fosso-Wamba¹

Accepted: 17 May 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Hesitant attitudes have been a significant issue since the development of the first vaccines—the WHO sees them as one of the most critical global health threats. The increasing use of social media to spread questionable information about vaccination strongly impacts the population's decision to get vaccinated. Developing text classification methods that can identify hesitant messages on social media could be useful for health campaigns in their efforts to address negative influences from social media platforms and provide reliable information to support their strategies against hesitant-vaccination sentiments. This study aims to evaluate the performance of different machine learning models and deep learning methods in identifying vaccine-hesitant tweets that are being published during the COVID-19 pandemic. Our concluding remarks are that Long Short-Term Memory and Recurrent Neural Network models have outperformed traditional machine learning models on detecting vaccine-hesitant messages in social media, with an accuracy rate of 86% against 83%.

Keywords Deep learning · Neural network · LSTM · Text classification · Vaccine hesitancy · COVID-19 · Twitter

1 Introduction

The Coronavirus disease, which was discovered in 2019, is an infectious disease caused by the Severe Acute Respiratory Syndrome coronavirus 2 (Henceforth, SARS-CoV-2). This highly disruptive pandemic (Henceforth, COVID-19) has caused major upheavals in virtually all industries across the globe, with severe implications in sectors such as manufacturing. Here, this global health disaster has particularly hit production networks, and the demand and supply chains underpinning manufacturing operations (Alam et al., 2021; Kapoor et al., 2021; Kumar

✉ Serge Nyawa
s.nyawa@tbs-education.fr

Dieudonné Tchuenté
d.tchuenté@tbs-education.fr

Samuel Fosso-Wamba
s.fosso-wamba@tbs-education.fr

¹ Department of Information, Operations and Management Sciences, TBS Business School, 1 Place Alphonse Jourdain, 31068 Toulouse, France

et al., 2021a, 2021b; Pujawan & Bah, 2021; Xiong et al., 2021). This pandemic has unfolded quickly to become a humanitarian crisis threatening billions of individuals globally, and demanding swift disaster relief efforts and humanitarian operations (Anparasan & Lejeune, 2019; Ivanov & Dolgui, 2020; Queiroz et al., 2020). Digital information and analytics are instrumental in the study and mitigation of such infectious diseases as well as in disaster situations (Dubey et al., 2019a, 2019b; DuHadway et al., 2019; Fast et al., 2018; Griffith et al., 2019; Gupta et al., 2021; Qayyum et al., 2021; Singh et al., 2019; Wamba et al., 2019). Social media, which is prevalent everywhere including among the younger generations, is used in real-time analysis for speed up trend predictions in many areas (e.g. Mishra & Singh, 2018; Moorhead et al., 2013). Social media analytics in particular can serve as a resource for efficient disease surveillance, and the diffusion of preventative actions to slow the spread of disease outbreaks (Anparasan & Lejeune, 2019; Kumar et al., 2021a, 2021b). For instance, Kumar et al. (2021a, 2021b) developed a dynamic transmission model to investigate the impact of social media (Twitter) on the number of infections and deaths due to influenza and COVID-19. Their findings indicate that social media is an integral part of the humanitarian logistics for pandemics emergency preparedness and contributes to the literature by informing best practices in the response to similar disasters.

In December 2019, the SARS-CoV-2 virus was first reported in Wuhan City by Chinese public health authorities. It rapidly spread through China and then to other Asian countries (Thailand, Japan, South Korea, ...). In Europe, the first cases were recorded in France on January 24, 2020, then in Germany and Italy respectively on January 28, 2020 and January 29, 2020. COVID-19 became a global pandemic on March 11, 2020.

Even though severe physical distancing and public health measures turned the epidemic curve downwards and decreased the risk of health system collapse while increasing the opportunity of developing treatments and vaccines, the economic impact of COVID-19 has been disastrous. Countries recorded declines in their economic activities (the pandemic allegedly pushed between 88 and 115 million people into extreme poverty in 2020 (See, e.g., Mahler, Laknerr, Castaneda and Wu (2020)), and millions of workers were deprived of their jobs. Globally, the estimated number of lost jobs amounted to hundreds of millions of jobs; only in the USA, more than 40 million became unemployed and found themselves filing unemployment insurance claims (See, e.g., Researcher (2020)). In parallel, firms faced far-reaching supply chain disruptions (According to Researcher (2020), losing more than 300 billion dollars, with long-lasting scars. This situation has been both socially and economically unbearable, forcing governments to cut short nationwide lockdowns (as these typically ended after only a few months) to reduce the economic impact. However, the virus was still circulating, and no pharmaceutical treatments nor vaccines were available.

Vaccination remained the most successful public health intervention that could contain the COVID-19 virus and halt the rise in the mortality, morbidity, and disability rates of collateral infectious diseases.¹ According to the World Health Organization (Henceforth, WHO), over 3 million deaths and 75,000 cases of disability are prevented annually owing to vaccination. Therefore, vaccination is being given great consideration and hope as a powerful containment measure, especially given past successes.

Scientists ensured about the potential for vaccination to successfully render COVID-19 less lethal, following the release of the genetic sequence of the coronavirus in January 2020. This sparked the commitment and desire of scientists to develop a vaccine against this deadly disease. Technological innovations, impressive efforts, unprecedented research frenzy, and colossal resources helped scientists develop COVID-19 vaccines faster than ever. In January

¹ <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/covid-19-vaccines>.

2021, two vaccines with proven efficacy in reducing infectious risk by more than 90% were approved in Europe. Earlier in the preceding year, lab progress had already led to a precise characterization of the virus as the "severe acute respiratory syndrome coronavirus 2" (SARS-CoV-2). A total of 16 candidate vaccines underwent phase 3 trials and were widely credited for their potential to curb the spread, severity, and lethality of COVID-19. Such vaccine efficacy exceeded 80%, with the potential to cover herd immunity when the percentage of the vaccinated population would range between 75 and 90% (see, e.g., Chevallier et al., 2021).

From June 2020, countries started deploying their respective vaccine campaigns. Despite the availability of vaccine doses, the vaccination rate remained low among the population (some figures for October 2021: France (68%), Russia (34%) and South Africa (21%), the USA (58%), the UK (68%)). As a result, full-herd immunity remained seriously threatened. Moreover, this highlighted the existence of vaccine hesitancy.

Vaccine hesitancy, which refers to people's reluctance to get vaccinated on various grounds (doubts about vaccines, for instance), is considered by the WHO as one of the more important global health threats.² It negatively impacts both the vaccine demand and efforts to maintain a reasonable level of immunization coverage. Vaccine hesitancy is complex and context-specific and varies across time, place, and the types of vaccines. It is often at the centre stage of debates or partial information about a specific disease. People can express reluctance to be vaccinated for many reasons, including lack of confidence, complacency, non-perception of the vaccine's usefulness, and challenging vaccine accessibility (See, e.g., Dubé et al., 2021). While investigating the feelings of U.S. people about COVID-19 vaccines, Thelwall et al. (2021) found that vaccine hesitators surveyed in that country blamed such vaccines for being rushed (37%) or simply mistrusting them (12%) (Reinhardt & Rossmann, 2020). In the UK, vaccine hesitancy is rather explained by mistrust toward the country's healthcare system (Freeman et al., 2022). Across the world, other reasons why people are hesitant about COVID-19 vaccines include conspiracy theory beliefs, personal freedom, and disgust about blood/needles (Hornsey; Harris; Fielding, 2018), vaccine ineffectiveness due to virus mutations, disbelief in the severity of COVID-19, fear of side effects, mistrust of Bill Gates and the mainstream media, and a perception of being an experimental subject (ElonPoll, 2020). In developing countries, the weight of traditions and beliefs, alternative conceptions of health, and lack of knowledge can be added to the list.

As for the degree of vaccine hesitancy, it can vary depending on several factors. A survey conducted in the USA between October and November 2020 showed that approximately 40% of adults were hesitant about getting vaccinated (Reinhardt & Rossmann, 2020). The proportion was higher among Black adults (Funk; Tyson, 2020). In the UK, 16% of the surveyed population reported in June 2020 that they were hesitant about vaccination (Skinner, 2020). In October, it was very unlikely for 12% to get vaccinated while 17% were still weaving in uncertainty (Freeman et al., 2022). In Italy, concerns about vaccines propagated at the early stages of COVID-19 (Palamenghi et al., 2020). People also cast doubt on COVID-19 vaccination in China in 2020 (Wagner et al., 2020). A survey with 19 countries in June 2020 revealed that globally, the intensity of vaccine hesitancy could vary between 45% (Russia) and 10% (China) (Lazarus et al., 2021).

Hesitancy to vaccination is not a recent fact. People have been distrustful of vaccines since the first human vaccination programs. In 1982, a TV documentary called "DTP Vaccine Roulette" in the USA claimed that the DTP vaccine was causing severe brain damage, seizures, and mental retardation. This generated detrimental effects on parental acceptance

² <http://www.who.int/emergencies/ten-threats-to-global-health-in-2019>.

of this vaccine. In 1990, a vaccination roll-out against hepatitis B in France raised doubts about the reliability of the vaccine, thus constraining the campaign. The vaccine was accused of causing multiple sclerosis. Between 2002 and 2005, unfounded rumors about the polio vaccine coupled with distrust of the Nigerian government led to the vaccine boycott. The direct consequence was that the incidence of polio cases increased fivefold. Therefore, anti-vaccine movements impeded the vaccination campaign against influenza A(H1N1)pdm09. In Denmark and Ireland, between 2015 and 2016, Human Papillomavirus vaccination programs underwent a drastic slowdown (vaccine uptake rates dropped from above 85% to less than 40%) because of misinformation about the vaccine's side effects.

Vaccination communication and outreach have improved significantly, and social networks are currently playing a critical role in the dissemination of both information and fake news, which strongly impacts the population's decision to get vaccinated. For example, rumours that the COVID-19 pandemic is a trick to sell vaccines are circulating faster on social media platforms than the virus itself. Quyen et al. (2021) argue that in 2019, some 31 million Facebook users and about 17 million YouTube users followed anti-vaxxers on these respective platforms. Continuous exposure to anti-vaccination messages increases the likelihood of being hesitant about the vaccine, since people shape their opinions after having been in contact with online information on the topic, and the majority of individuals do not consider the credibility of the source of information (Germani & Biller-Andorno, 2021).

The spread of warnings and misinformation via social networks leads to increased hesitation attitudes among the population and undermines any potential sales success of COVID-19 vaccines. This is a major concern for governments, medical and social sciences researchers, who fear huge losses of funds and time to develop vaccines. Therefore, it becomes necessary to use the huge number of rich data generated from social media to detect concerns about vaccination among the population and control changes in vaccination behaviours. To do this effectively, digital technologies to monitor pandemic situations are required (Bag et al., 2021).

Information digitalization is being intensively used to support the public-health response to COVID-19 worldwide. Relying on mobile phone messaging, social media datasets, connected devices, low-cost computing resources, and progress in machine learning includes population surveillance, case identification, contact tracing, and evaluation of interventions based on mobility data and communication with the public (Budd et al., 2020). Digital interventions such as regular webinars, dissemination of digital newsletters and toolkits, text messages, email-based communication, and smartphone apps are being increasingly used to promote the uptake of vaccinations in all age groups, including in low- and middle-income countries (Odone et al., 2021). In light of the COVID-19 vaccine roll-out, the use of real-time monitoring (activities that employ digital technologies to accelerate the sharing, analysis, and use of data to improve campaign quality) to support vaccination campaigns is more important than ever.³ Health institutions implement social listening to identify and understand posts about vaccines and vaccination on social media to learn what topics are being discussed, what information is being shared, and whether it is accurate. This paper contributes to this growing literature consisting on using digital technologies to improve vaccination and protect society.

The primary interest of this paper resides in the development of text classification methods that can identify hesitant messages on social media. It is worth mentioning the difference between identifying global opinion about vaccination (positive, negative, neutral) or anti-vax messages (for which there is growing literature) and identifying vaccine hesitancy messages.

³ The Use of Real Time Monitoring Approaches and Tools for Immunization Campaigns: Good Practices and Lessons Learned New York: United Nations Children's Fund (UNICEF), 2021.

The two groups of concepts do not refer to the same reality (Dubé et al., 2021). The concept of vaccine hesitancy entails a shift from a dichotomous perspective (for or against a vaccine) to a behavioral approach regarding a spectrum of potential attitudes (active demand for vaccines, total rejection of all vaccines). Vaccine-hesitant individuals are a heterogeneous group along this continuum. A vaccine-hesitant person can delay or be reluctant about the vaccine; he has legitimate doubts and concerns about the vaccine (Fornell et al., 2015), while an anti-vaxxer is opposed to vaccination. Thus, in this paper, we specifically address vaccine hesitancy in accordance with the etymological definition of this concept by classifying social media contents into hesitant or non-hesitant, using machine learning and deep learning methods.

Text content classification is a growing topic, and a myriad of methods are used to allocate texts to categories. Machine learning algorithms have been found to deliver higher predictive performances in text classification tasks (Hussain et al., 2021). However, only the best model must be implemented when deploying text classifiers. Thus, the choice of the best machine learning model is crucial. The empirical strategy often consists of implementing all the adapted machine learning algorithms and selecting only the one with the best predictive performances. This leads to a comparison map of different algorithms useful for choosing the best final model. This paper uses a similar approach.

We choose Machine Learning and Deep Learning (hereafter, ML and DL) perspectives, using well-established models such as Artificial Neural Networks, Ensemble Learning (random forest, Gradient boosting, AdaBoost), Support Vector Machine, K-nearest neighbors, Decision Tree, Logistic Regression, Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Bidirectional Encoder Representations from Transformers (BERT). As part of the study, we have set out the following research question:

RQ: How efficient can Machine Learning and Deep Learning Models contribute to identifying vaccine-hesitant messages on social networks?

The main contribution of this paper is to provide a comparison map of Machine Learning and Deep Learning models for the classification of vaccine-hesitant messages. For future studies, the findings of this study should provide useful information for developing appropriate models enabling the detection of vaccine-hesitant messages. This can help health campaigns to address negative influences among social network users and provide useful information for coping with hesitant-vaccination sentiments.

2 Related works

Studying people's perceptions on social media to understand their sentiment presents a powerful medium for researchers to identify the causes of vaccine hesitancy and therefore develop appropriate public health messages and interventions (Alamoodi et al., 2021; Karafillakis et al., 2021). Sentiment analysis involves categorizing subjective opinions to determine polarities (e.g. positive, negative, and neutral), emotions (e.g. anger, sadness, and happiness), or states of mind (e.g. interest vs. disinterest) toward target topics, themes, or aspects of interest (Hussain et al., 2021). A complementary approach termed stance detection (Majumdar et al., 2020) assigns a stance label (favorable, against, and none) to a post on a specific predetermined target, which itself may not be referred to or be the target of opinions expressed in the post. These recent years, many studies have been interested in performing sentiment or stance analysis (SA) on vaccine hesitancy for many diseases, including not only COVID-19

pandemic, but also many other diseases (e.g., HPV, Measles, Rubella, Influenza, Hepatitis B, Diphtheria, Chickenpox, Tetanus, Polio) through machine learning and deep learning methods as synthesized in Table 1.

Most of these studies classify social media's messages into positive, negative, and neutral opinions toward vaccination (Abd Rahim & Rafie, 2020; Bar-Lev et al., 2021; Hussain et al., 2021; Piedrahita-Valdés et al., 2021; Tavošchi et al., 2020; Yuan et al., 2019). Tavošchi et al. (2020) use 1,80,620 tweets collected in 13 months (September 2016 to August 2017) to monitor public opinion on vaccination in Italy. Using an SVM machine learning model, they classified 60% of tweets as neutral, 23% against vaccination, and 17% in favor of vaccination. Hussain et al. (2021) relied on 3,00,000 posts and tweets from Twitter or Facebook in 9 months (March 2020 to November 2020) to analyse public sentiment about COVID-19 in the UK and the USA. They used a deep learning BERT model and found that the overall averaged positive, negative, and neutral sentiments were at 58%, 22%, and 17% in the UK, compared to 56%, 24%, and 18% in the USA, respectively. Bar-Lev et al. (2021) collect 9,596 posts on Facebook groups and the Tapuz platform in Israel 5-year period (2013–2018). They used several machine learning methods (Logistic Regression, Random Forest, Neural Networks and Linear Regression) to assess how online content regarding vaccination affects vaccine hesitancy. They compared models that only use demographic variables with models adding aggregated sentiments from social media to demographic variables. They found that higher hesitancy was associated with more social media traffic, for most of the vaccinations, and social media traffic features improved the performances of most of the models. Piedrahita-Valdés et al. (2021) collected 1,499,227 tweets in an 8-year period (2011–2019) to evaluate the public perceptions regarding vaccination among several countries in the world. By using a Lexicon Analysis and a SVM model, they reached a classification accuracy of 85%, and classified 69.36% of the tweets as neutral, 21.78% as positive, and 8.86% as negative. They also performed a trend analysis. The percentage of neutral tweets showed a decreasing tendency, while the proportion of positive and negative tweets increased over time. Peaks in positive tweets were observed every April. The proportion of positive tweets was significantly higher in the middle of the week before decreasing during weekends. Negative tweets followed the opposite pattern. While Switzerland recorded more positive tweets (71.43%), most negative tweets were to be found in the Netherlands (15.53%), Canada (11.32%), Japan (10.74%), and the United States (10.49%). Yuan et al. (2019) collect 6,69,136 tweets in a 2-month period (February 2015 to March 2015) for investigating the communication patterns of anti- and pro-vaccine users and the role of bots in the USA. They compared five machine learning algorithms (Logistic Regression, SVM, kNN, Nearest Centroid and Naïve Bayes) for tweets classification. SVM provides the best accuracy. Additionally, they use a clustering algorithm (from social networks analysis) to identify groups in the retweet network, and a bot detection algorithm to identify potential bots among users. They found that pro- and antivaccine users retweet predominantly from their own opinion group. In addition, the bot analysis disclosed that 1.45% of the corpus users were identified as likely bots which produced 4.59% of all tweets within the dataset. They also found that bots display hyper-social tendencies by initiating retweets at higher frequencies with users within the same opinion group.

Rather than classifying messages into positive, negative, and neutral opinion toward vaccination, other studies mostly focus on detecting negative content, antivaccine content or misinformation content by using only two classes (e.g. negative and non-negative; antivaccine and others; true information and false information) (Rodríguez-González et al., 2020; To et al., 2021; Wang et al., 2020; Zhou et al., 2015). To et al. (2021) collected 1,651,687 tweets in an 8-month period (January 2020 to August 2020) to evaluate the performance of different machine learning (SVM, Naïve Bayes) and deep learning (BERT, Bi-LSTM) models

Table 1 Synthesis of existing studies

Paper	Disease(s)	Goal	SA	TM	Type SA	Method SA	Method TM	Other analysis	Data source	Period	Size
Tavoschi et al. (2020)	General (mainly measles)	Monitoring the public opinion on vaccination in Italy	x		Positive Negative Neutral	SVM			Twitter	13 months (sept 2016–Aug 2017)	180,620 tweets
Zhou et al. (2015)	HPV	Examining if social connection information from tweets about human papilloma virus (HPV) vaccines could be used to train classifiers that identify antivaccine opinions	x		Antivaccine Other	SVM		Social network analysis	Twitter	7 months (Oct. 2013–March 2014)	42,533 tweets
Hussain et al. (2021)	COVID-19	Analyzing public sentiments on social media in the UK and the USA	x		Positive Negative Neutral	BERT			Twitter Facebook	9 months (March 2020–Nov 20, 2020)	300,000 posts or tweets
To et al. (2021)	COVID-19	Evaluating the performance of different natural language processing models to identify anti-vaccination tweets	x		Antivaccine Other	BERT Bi-LSTM SVM Naive Bayes			Twitter	8 months (Jan 2020–Aug 2020)	1,651,687 tweets
Ma et al. (2021)	COVID-19	Comparing two different topic models to identify topics related to vaccine hesitancy in the USA		x			Top2Vec LDA		Twitter	3 months (Jan 2021–March 2021)	3,403,166 tweets

Table 1 (continued)

Paper	Disease(s)	Goal	SA	TM	Type SA	Method SA	Method TM	Other analysis	Data source	Period	Size
Rodríguez-González et al. (2020)	General (e.g., HPV, measles, Influenza, Hepatitis, Chickenpox, Varicella)	Identifying sentiment in tweets by using different machine-learning techniques and methods, and dealing with the unbalanced data problem in Spain	x		Negative Non-negative	C5.0 Logit Boost Bayesian GLM Neural Networks Random Forest SVM			Twitter	3 years (2015–2018)	1,028,742 tweets
Bar-Lev et al. (2021)	General (e.g., Hepatitis B, Diphtheria, Tetanus, Whooping cough, Polio, Pneumococcal, Rotavirus, MMR)	Using machine-learning strategies to assess how online content regarding vaccination affects vaccine hesitancy in Israel	x		Positive Negative Neutral	Logistic regression Random Forest Neural Networks Linear Regression			Facebook Tapuz	5 years (2013–2018)	9,596 posts on Facebook groups or Tapuz platform
Piedrahita-Valdés et al. (2021)	General	Evaluating public perceptions regarding vaccination and comparison among several countries	x		Positive Negative Neutral	Lexicon Analysis SVM		Trend analysis	Twitter	8 years (2011–2019)	1,499,227 tweets
Yuan et al. (2019)	MMR	Examining emergent communities and social bots within the polarized online vaccination debate in Twitter	x		Pro-vaccine antivaccine Neutral	Logistic regression SVM kNN Nearest Centroid Naive Bayes		Social network analysis		2 months (Feb 2015–March 2015)	669,136 tweets Retweets relations

Table 1 (continued)

Paper	Disease(s)	Goal	SA	TM	Type SA	Method SA	Method TM	Other analysis	Data source	Period	Size
Furini (2021)	General (e.g. Measles, Autism, Rubella, Meningitis, Meningococcus, Polio, Tetanus)	Identifying psycho-linguistics signals of distrust toward vaccines to help health authorities to restore the trust toward vaccines in Italy						Psycho-linguistics and time-domain analyses	Facebook	2 years (Oct 2015–Oct 2017)	172,799 posts from ProVax and NoVax users
Jiang et al. (2021)	COVID-19	Understanding how vaccine favorability and specific vaccine-related concerns were articulated and transmitted by Twitter users from opposing ideological camps and with different follower scopes in USA	x	x	Favorable to vax, Unfavorable to vax, Side effect, Distrust of medical professionals, conspiracy theory	BERT	Structural Topic Modeling (STM)		Twitter	4 months (March 2020–June 2020)	16,959 tweets
Argyris et al. (2021)	General	Comparing discursive topics chosen by pro- and anti-vaccine advocates in their attempts to influence the public to accept or reject immunization in the engagement-persuasion spectrum	x	x	Pro-vaccine antivaccine Neutral	Logistic Regression	K-Means		Twitter	1 month (Nov 2019)	39,962 tweets

Table 1 (continued)

Paper	Disease(s)	Goal	SA	TM	Type SA	Method SA	Method TM	Other analysis	Data source	Period	Size
Wang et al. (2020)	General	Developing an automatic detector for antivaccine messages to counteract the negative impact that antivaccine messages have on the public health using images, texts, and hashtags	x		antivaccine Other	SVM LSTM VGG, ResNet, RNN, EAN, MVAE		OCR	Instagram	3 years, 10 months (Jan 2016–Oct 2019)	30,000 samples
Sear et al. (2020)	COVID-19	Using machine learning to quantify COVID-19 content among online pro-vaccines and anti-vaccines		x			LDA	Trend analysis	Facebook	2 months (Jan2020–Feb 2020)	8277 posts on Facebook pages
Cotfas et al. (2021)	COVID-19	Analyzing the dynamics of public opinion on Twitter in the first month after the start of the vaccination process in the UK, with a focus on COVID-19 vaccine hesitancy messages in connection with the major events in the analyzed period	x	x	In favor Neutral Against	Random Forest SVM Multinomial Naïve Bayes BERT RoBERTa	LDA	Trend analysis	Twitter	2 months (Dec 2020–Jan 2021)	5,030,866 tweets

Table 1 (continued)

Paper	Disease(s)	Goal	SA	TM	Type SA	Method SA	Method TM	Other analysis	Data source	Period	Size
Karami et al. (2021)	COVID-19	Identifying the sentiment of tweets using a machine learning rule-based approach, discovers major topics, explores temporal trend, and compares topics of negative and non-negative tweets using statistical tests, and discloses top topics of tweets having negative and non-negative sentiment (in the USA)	x	x	Negative Non-negative	LWIC VADER BrandWatch	LDA	Trend analysis	Twitter	3 months (Nov 2020–Feb 2021)	200,000 tweets
Abd Rahim and Rafie (2020)	General (e.g. Measles)	Developing a model that uses SVM classifier to classify the polarity of sentiments: positive, negative and neutral	x		Positive Negative Neutral	SVM			Twitter	6 months (Oct 2019–March 2020)	105,965 tweets
Tomaszowski et al. (2021)	HPV	Developing a systematic and generalizable approach to identifying false HPV vaccine information on social media	x	x	True Information False Information	SVM Naive Bayes CNN BiLSTM	DBSCAN		Twitter	4 years (2013–2017)	705,858 tweets

and then identify anti-vaccination tweets on COVID-19. They found that the BERT models outperformed the Bi-LSTM, SVM, and Naïve Bayes models in this task with an accuracy of 91.6%, a precision of 93.4%, a recall of 97.6%, an F1-score of 95.5%, and an AUC of 84.7%. Zhou et al. (2015) use 42,533 tweets in a 7-month period (October 2013 to March 2014) for examining if social connection information from tweets about HPV vaccines could be used to train classifiers that identify antivaccine opinions. In addition to the text of each tweet, they also used social connections between users (sources for people that the user follows, followers for people that follow the user) as additional features for training models. Using an SVM classifier, they found that for the task of classifying tweets about HPV vaccines as antivaccine or, otherwise, information about the social connections between users provided a useful addition to the content of what people write. They showed that it was possible to use information about the users that people follow online to help predict their opinions. The most accurate classifier achieved an accuracy of 88.6% on the test data set and used only social connection features. (Rodríguez-González et al., 2020) collected 1,651,687 tweets in a 4-year period (2015–2018) in Spain in order to identify sentiments in such messages by using different machine learning techniques, dealing with the unbalanced data problem. They used five machine learning techniques: SVM, Random Forest, Neural Networks, Bayesian GLM, Logit Boost and C5.0. They found that the model that provided the highest accuracy from all the studied possibilities was the one generated by the subset obtained from up-sampling with the ADASYN method and corresponding to the Random Forest technique. Wang et al. (2020) collected 30,000 samples of Instagram posts in a 3-year period (2016–2019) for developing an automatic detector for antivaccine messages to counteract the negative impact that antivaccine messages have on public health by using not only texts but also images and hashtags. They extracted textual contents from images by using OCR (Optical Character Recognition) methods. They proposed a deep learning network that leverages both visual and textual information with multiple deep learning models (LSTM, VGG, RNN, EAN, MVAE) and an SVM machine learning model. Their results demonstrated that the final network achieves above 97% testing accuracy and outperforms other relevant models, thus showing the possibility to detect a large number of antivaccine messages posted daily.

Finally, other studies also complement sentiment or stance analysis with topic modeling (TM) to automatically identify important topics related to vaccine hesitancy from social media's contents (Argyris et al., 2021; Cotfas et al., 2021; Jiang et al., 2021; Karami et al., 2021; Ma et al. 2021; Sear et al., 2020; Tomaszewski et al., 2021). Ma et al. (2021) used 3,403,166 tweets collected in a 3-month period (January 2021 to March 2021) to compare two different topic models (Top2Vec and LDA) and identify topics related to vaccine hesitancy in the USA. Their results demonstrated that Top2Vec is able to extract more relevant topics. Jiang et al. (2021) collected 16,959 tweets in a 4-month period (March 2020 to June 2020) in the United States to understand how vaccine favorability and specific vaccine-related concerns were articulated and transmitted by Twitter users from opposing ideological camps and with different follower scopes. They used the deep learning BERT model for tweet classification (favorable to vax, unfavorable to vax, side effect, distrust to medical professionals, and conspiracy theory). They use Structural Topic Modeling (STM) to identify important topics for each ideological camp. They found that the use of structural topic modeling could reveal distinct topical focuses among liberal and conservative users. Based on 39,962 tweets, Argyris et al. (2021) compared discursive topics selected by pro- and antivaccine advocates in their attempts to influence the public to accept or reject immunization in the engagement-persuasion spectrum. They used the Logistic Regression model for tweet classification (pro-vaccine, antivaccine, neutral) and K-means algorithm for topic identification. Their results indicated that antivaccine topics have greater intertopic distinctiveness (i.e.,

the degree to which discursive topics are distinct from one another) than their pro-vaccine counterparts. In addition, while antivaccine advocates the use of all four message frames known to make narratives persuasive and influential, pro-vaccine advocates have neglected having a clear problem statement. Sear et al. (2020) collected 8277 posts on pro-vaccine and antivaccine Facebook pages in a 2-month period (December 2020 to January 2021) to quantify COVID-19 content among online pro-vaccines and anti-vaccines with LDA topic modeling technique. They observed that the anti-vaccination community is developing a less focused debate around COVID-19 than its counterpart, the pro-vaccination community. However, the anti-vax community exhibits a broader range of "avors" of COVID-19 topics, and hence can appeal for a broader cross-section of individuals seeking COVID-19 guidance online (e.g. individuals wary of a mandatory fast-tracked COVID-19 vaccine or those seeking alternative remedies). Hence the anti-vaccination community looks better positioned to attract fresh support going forward than the pro-vax community. Cotfas et al. (2021) compiled 5,030,866 tweets in a 2-month period (December 2020 to January 2021) with a view to analyzing the dynamics of public opinion on Twitter in the first month after a vaccination campaign was launched in the UK. Particular focus was on COVID-19 vaccine hesitancy messages in connection with the major events that occurred in that period. Machine learning techniques (Random Forest, SVM, Multinomial Naïve Bayes) and deep learning techniques (BERT and RoBERTa) were leveraged by the authors to classify tweets according to their being in favor of, neutral, or against vaccination, while LDA technique served for identifying important topics within each group and perform a trend analysis to connect the findings with major events occurring in the same period. The number of tweets, as they observed, varied in accordance with the major events reported by the news in the corresponding days. This finding reveals that people use Twitter to disclose their opinion about the news concerning COVID-19 vaccination. The so-called "reaction" of the tweets to the news was in line with the previous research from the field. Karami et al. (2021) used 2,00,000 tweets in a 3-month period (November 2020 to February 2021) in the USA to identify the sentiment of tweets (negative and non-negative) through a machine learning rule-based approach (with existing tools or libraries such as LIWC, VADER and BrandWatch). They also identified major topics (with LDA technique) and the temporal trend, and could compare topics of negative and non-negative tweets using statistical tests. Top topics of tweets having negative and non-negative sentiment were therefore revealed. A bulk of 7,05,858 tweets collected between 2003 and 2005 were explored by Tomaszewski et al. (2021) to develop a systematic and generalizable approach for identifying false HPV vaccine information on social media. They made use of machine learning (SVM, Naïve Bayes) and deep learning (CNN, BiLSTM) techniques, then relied on DBSCAN clustering algorithms to identify the most important topics. Their finding revealed that the convolutional neural network model outperformed all other models when it comes to identifying tweets containing false HPV vaccine-related information (F-score = 91.95). Their proposed unsupervised causality mining models also identified HPV vaccine candidate effects for capturing risk perceptions of HPV vaccines.

Overall, even if most of these studies indicate having the vaccine hesitancy issue in their target, their sentiment or stance analysis is much more interested in the global opinion about vaccination (positive, negative, neutral) or in antivaccine concerns. Unlike these studies, we adopt a different classification approach in this paper by classifying social media messages into hesitant or non-hesitant with machine learning and deep learning methods.

3 Data and methods

3.1 Data

Created in March 2006 by Jack Dorsey, Noah Glass, Biz Stone, and Evan Williams, Twitter is an online news and social networking service. Users, known as “twitterers”, can post, like, or retweet messages known as “tweets”. Initially restricted to 140 characters, the text width of a tweet can nowadays reach 280. Globally, this social network has around 100 million daily active users, who posts in average 500 million tweets per day. Registered users mostly rely on their mobile phone to interact with the platform (80%). Tweeter is a larger multinational source of English-language public comments in the social web.

The success of Twitter can be explained by many factors, the most important of which are audience increase, instant communication, real-time information, and direct support for response efforts (Martinez-Rojas et al., 2018). In addition, Twitter is very convenient for socializing. It is an easy way for finding people willing to support a specific cause. Considering these properties, Twitter can be a powerful mechanism for faster information dissemination in general and for vaccine hesitancy messages specifically.

Historical tweets were collected between 1 January and 30 June 2021 using the Twitter API. The original dataset includes 5 million tweets and retweets in the full version. Due to retweeting and copy tweeting, multiple copies of the same tweet can be collected from bots as well as many very similar tweets. Tweets were filtered to remove duplicates and tweets with identical content to another tweet after removing hashtags and @usernames.

We decided to use a clean version with no retweets, that is, around 1.5 million tweets.

The following keywords were used to extract relevant tweets: COVID vaccine OR COVID vaccine hesitancy OR COVID vaccine facts OR COVID Anti-vaccination OR vaccine hesitancy OR vaccine refusal OR vaccine acceptance OR vaccine resistant OR vaccination confidence OR vaccine uptake OR vaccine demand OR vaccine refusal OR Vaccination OR Health communication OR Vaccine resilience OR vaccination rumours OR vaccination trust OR vaccine misinformation OR Vaccine trust OR vaccine OR vaccinated OR vaccinate OR vaccinating OR immunized OR immunize OR immunization OR immunizing OR immunization programme OR Vaccine debate OR Pfizer-BioNTech OR Moderna vaccine OR AstraZeneca vaccine OR Spoutnik V vaccine OR Johnson vaccine OR Sinopharm vaccine OR Sinovac Biotech vaccine OR vaccines work.

To ensure that tweets posted across different times of the pandemic will be selected, a random sample of 20,000 tweets was extracted from the set of 1.5 million tweets for labelling. Tweets were labelled as either “Hesitant” or “Nonhesitant” (neutral, antivax, or ambiguous). Researchers worked in pairs to label the tweets. Differences in labelling were checked and decided by a third researcher. The following table gives a view of our tweets and their corresponding labels (Tables 2, 3 and 4).

3.2 Methods

Machine learning and deep learning (Henceforth, ML and DL) correspond to well-established methods that provide algorithms for computers to discover knowledge and make decisions by first learning from the given data. They are becoming increasingly popular within the field of text classification. In this specific domain, people use ML/DL to search for new data patterns and generate predictive models. Such patterns are used to improve future operational decisions

Table 2 Tweets and labels

Text	Vaccine_acceptance
News afternoon digest approval covid vaccine c...	Hesitant
Talking things missed mom revealed nothing mis...	Hesitant
One right mind would injected vax using totall...	Hesitant
Israelis found covid vaccine	Hesitant
One might agree accept cfr actually covid deat...	Hesitant
...	...
Opinion Nigerian vaccinated Europe let	Non Hesitant
Get second vaccination covid	Non Hesitant
Fully vaccinated zero complaints	Non Hesitant
Never thought would ever see world vaccination...	Hesitant
Wait covid got vaccinated covid need buy lotte...	Hesitant

Table 3 Additional numerical tweet features

	0	1	2	3	4	5	6	7
0	9.8	56.97	28.0	1.5555	111.0	111.0	18.0	18.0
1	10.5	55.41	31.0	1.5500	135.0	135.0	20.0	18.0
2	17.4	25.27	55.0	1.7742	214.0	214.0	31.0	30.0
3	3.7	75.88	6.0	1.4999	28.0	28.0	4.0	4.0
4	18.0	6.94	48.0	2.0869	175.0	175.0	23.0	20.0
...

(Cohen, 2018; Chen et al., 2020; Kusiak, 2020). The success of this family of prediction techniques can be explained by factors such as the improvement of computational processing, the availability of massive sources of data; the growing demand of data-driven decision-making, and the need for automating decision processes. ML/DL uses a few assumptions regarding the input and output variables and applies complex mathematical calculations to automatically produce models that can analyse large and complex data and produce fast and accurate results (Akyildirim et al., 2020). Most of the existing studies on text classification show that several different methods can perform well depending on the context or the dataset used. These techniques include: Artificial Neural Networks; Ensemble Learning (Random Forest, Gradient Boosting, Adaboost); Support Vector Classifier; K-Nearest Neighbors; Decision Tree; Logistic Regression; Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Bidirectional Encoder Representations from Transformers (BERT). For our specific study, we compare all these methods in the same context and with the same dataset. We present an overview of these selected techniques in the next subsections.

Table 4 Hyperparameters of machine and deep learning algorithms

ML algorithm	Hyperparameters chosen	Different values used
Logistic regression	solver penalty max_iter	'newton-cg', 'lbfgs', 'liblinear', 'sag', 'saga' 'l1', 'l2', 'elasticnet' 5000
Random forest	n_estimators criterion max_features	100, 200 "gini", "entropy" "auto"
Decision tree	max_features criterion min_samples_split	"auto" "gini", "entropy" 2
Ada boosting	Type of estimator (base_estimator) Decision tree max depth (max_depth) Number of estimators (n_estimators) Learning rate (learning_rate)	DecisionTreeRegressor 8, 32 100, 200, 250 0.001, 0.05, 0.1
Gradient boosting	Decision tree max depth (max_depth) Number of estimators (n_estimators) Learning rate (learning_rate) Loss function (loss)	8,32 100, 200, 250 0.001, 0.05, 0.1 'deviance', 'exponential'
K-nearest neighbors	Number of neighbors (n_neighbors) Neighbor's weight function (weights) Neighbour's algorithm (algorithm)	5, 30, 100 uniform, distance ball_tree, kd_tree, brute, auto
Support vector classifier	Intercept fitting (fit_intercept) Regularization parameter (C) Max number of iterations (max_iter) Loss function (loss)	True, False 1.0, 2.0, 3.0 1000, 2000 'hinge', 'squared_hinge'
Artificial neural networks	Network architecture (hidden_layer_sizes) Activation function (activation) Learning rate (learning_rate_init) Optimizer (solver)	150, (150,50), (50, 20) relu, logistic 0.001, 0.005, 0.1 adam, lbfgs
Ligth LSTM	Embedding size (embedding_size) Batch size (batch_size) Epoch (epoch) Optimizer (optimizer) Dropout (dropout) Units (units)	128, 300 16, 32, 64 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 'adam', 'rmsprop', 'Adadelta' 0.1, 0.2, 0.3 20, 30, 40
LSTM	Embedding size (embedding_size) Batch size (batch_size) Epoch (epoch) Optimizer (optimizer) Dropout (dropout) Units of the RNN layer Units of dense layers	128, 300 16, 32, 64 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 'adam', 'rmsprop', 'Adadelta' 0.1, 0.2, 0.3 25, 50 25, 50
Recurrent neural network	Embedding size (embedding_size) Batch size (batch_size) Epoch (epoch) Optimizer (optimizer) Units of dense layers	128, 300 16, 32, 64 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 'adam', 'rmsprop', 'Adadelta' 25, 50

3.2.1 Logistic regression (LR)

When the dependent variable to predict is dichotomous, the logistic regression is appropriate. It is predictive analysis used to describe the relationship between one dependent binary variable and nominal, ordinal, interval or ratio-level independent variables. The main task of the logistic regression analysis is to estimate the log odds of an event. Let's consider $S = \{(x_1, y_1), \dots, (x_N, y_N)\}$ a sample of N observations, where x_i is the set of p attributes of the individual i and y_i is its corresponding dichotomous variable ($y_i = 0$ or $y_i = 1$). LR assumes a linear relationship between the predictor variables and the log-odds, and estimates a multiple linear regression function defined as:

$$\text{Log}\left(\frac{P(y_i = 1)}{1 - P(y_i = 1)}\right) = \beta_0 + \beta_1x_{i1} + \beta_2x_{i2} + \dots + \beta_px_{ip}$$

The probability $P(y_i = 1)$ can be recovered by exponentiating the log-odds. After setting a threshold (a value between 0 and 1, e.g. 0.5) a new individual j will be assigned to the class 1 if $P(y_j = 1)$ is above the threshold.

3.2.2 Random forest (RF)

Based on decision trees, the random forest algorithm can be applied for a classification exercise (Breiman, 2001; Shalev-Shwartz & Ben-David, 2014). To construct a predictor for the output variable y using inputs inx , we consider the following training set $S = \{(x_1, y_1), \dots, (x_N, y_N)\}$. In Random Forest, the first step consists in drawing randomly with replacement from S , a sample $S_1 = \{(x_{11}, y_{11}), \dots, (x_{n1}, y_{n1})\}$ of n observations. From S_1 , another randomness is used to construct the decision tree T_1 as follows: during the construction of each node, p attributes are randomly selected from the set of P initial inputs and used to select the right node based on the information gain. At this iterative process, we obtain a decision tree. The process is repeated m times and leads to m decision trees T_1, \dots, T_m , as shown in Fig. 1 (with $m = 4$). For an observation of inputs x , the prediction of the output y is

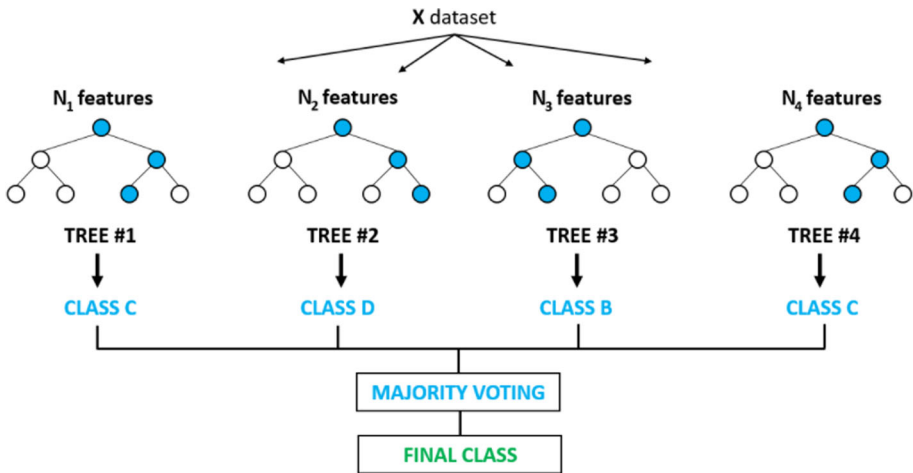


Fig. 1 Random forest. Source: Image courtesy

obtained by the majority voting from all individual trees T_1, \dots, T_m . Having the advantage of reducing the overfitting problem, RF is invariant to monotonic transformations of predictors and naturally accommodates categorical and numerical data in the same model. This model is able to approximate severe nonlinearities, since a tree of depth L can capture $(L - 1)$ -way interactions (Gu et al., 2020). However, RF is less interpretable than an individual decision tree. It has a high computational cost and a use of a great deal of memory which consequently lead to slow prediction speed.

3.2.3 Adaptive boosting and gradient boosting (AdaBoost)

To make a given learning technique system more efficient, the Adaptive Boosting (Henceforth, Ada Boosting or Ada Boost) is appropriated (Freund & Schapire, 1995). The Adaptive Boosting trains a weak learner (e.g. a decision tree) in several successive stages, on random samples formed by assigning significant weights to individuals who are difficult to classify. Figure 2 provides an illustration. More precisely, during the first step, a decision tree is produced from the sample. The model increases weights of individuals who are wrongly classified in order to form the sample for the next step. In the second step, a new decision tree is constructed from the resulting sample. The process is repeated a given amount of time. The final classifier is a majority voting of step classifiers weighted by coefficients related to their performances. Ada Boosting can be interpreted as an optimization algorithm on an exponential cost function. With its ability to enable optimization with other types of differentiable loss functions, Gradient Boosting is a generalization of boosting techniques. Ada Boosting and Gradient Boosting can achieve a very good accuracy levels with modest memory and limited runtime. They are appropriated with complex and high-dimensional data (Cui et al., 2018). Nevertheless, they are quite difficult to be interpreted. Also, their performances are poor when dealing with thousands of features with sparse values.

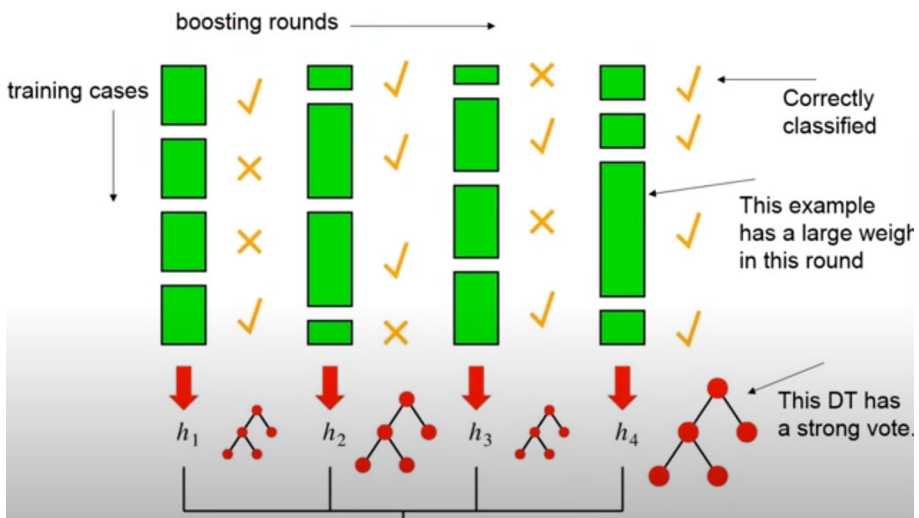


Fig. 2 Boosting. Source: Image courtesy

3.2.4 K-nearest neighbors (KNN)

The K-Nearest Neighbors (hereafter, KNN) is a non-parametric machine learning technique based on a local approximation and is deemed appropriate for classification (Cover & Hart, 1967; Devroye et al., 1996). It relies on the K-closest training examples of a new individual in the sample to generate as output a class membership. Let $S = \{(x_1, y_1), \dots, (x_N, y_N)\}$ be a sample of N observations, where x_i is the set of attributes of the individual i and y_i its corresponding class. Let us consider a new individual with inputs x . Using a distance metric, the K nearest neighbors of x in the sample will be observed. Let us call them $(x_{(1)}, \dots, x_{(K)})$. Due to the categorical nature of the output variable, the predicted class y will be the most common among the nearest neighbors and approximated by the mode of $(y_{(1)}, \dots, y_{(K)})$, as illustrated in Fig. 3. Weights are assigned to contributions of neighbors, so that the nearer neighbors contribute more to the predicted class than the more distant ones. An underlying assumption of this procedure is that the number of neighbors of K is known, which is not often the case in real applications. This number is approximated by minimizing Root Mean Square Error (RMSE). The KNN has the great advantage of being much faster than other algorithms that require training, since it does not derive any discriminative function from the

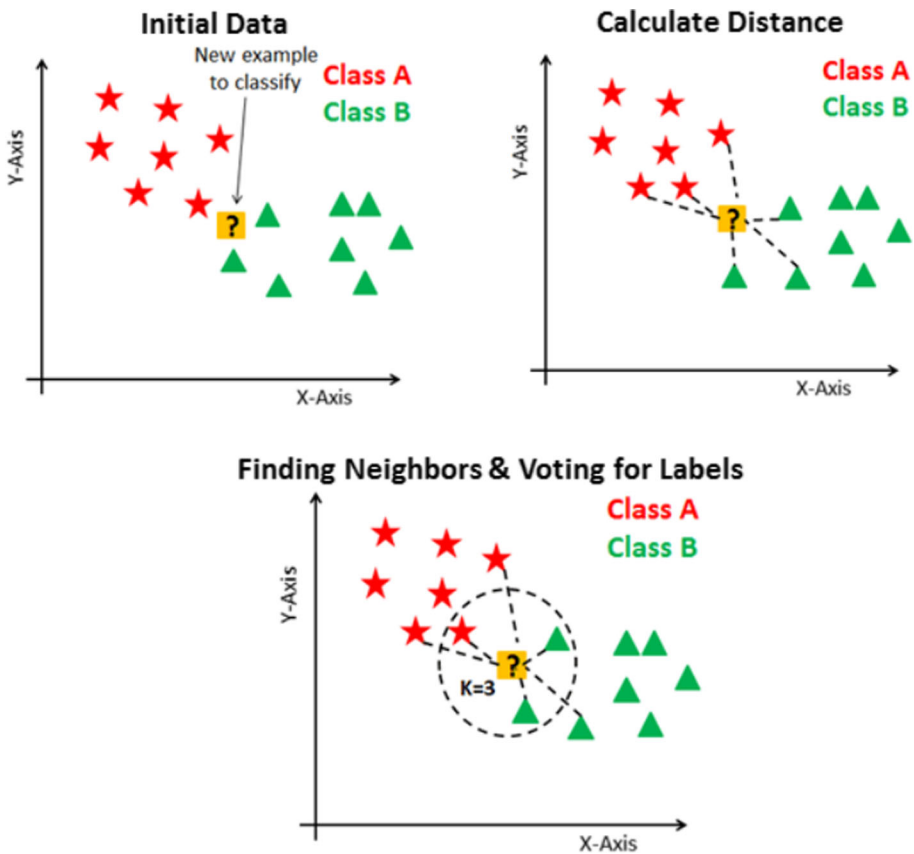


Fig. 3 K-nearest neighbors. Source: Image courtesy

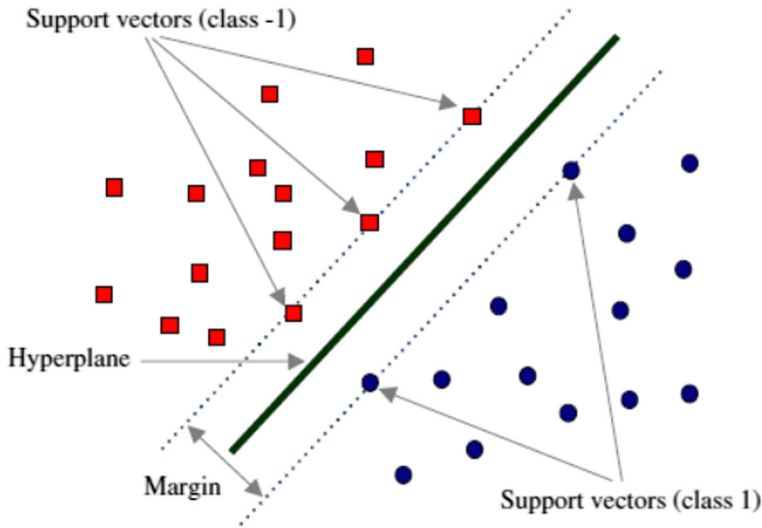


Fig. 4 Support vector classifier. *Source:* Image courtesy

training data. For its implementation, this method requires only two parameters to be set: the value of K , and the distance function. However, it performs poorly with a huge number of individuals or an important number of input variables.

3.2.5 Support vector classifier (SVC)

The SVC uses a hyperplane to best separate the data in classes. It maps in-sample items to points in space, to maximize the width of the gap between categories. To find the frontier between the categories to be separated, the SVC searches for the hyperplane that separates the training sample while maximizing the distance between the training points and this decision boundary: it maximizes the margin. The following figure is an illustration. Training points close to the border are called support vectors. It can happen in some cases that the training points are not linearly separable such that there is no hyperplane able to split the data. In such cases, the initial data should be transformed to allow the separation. This can be done by projecting the initial data into a larger dimensional space, where it becomes possible to find a linear separator (Shalev-Shwartz & Ben-David, 2014; Cortes & Vapnik, 1995; Boser et al., 1992). In the case of a nonlinear SVC, functional forms of transformation can be avoided by using a nonlinear kernel to get a nonlinear classifier without transforming the data at all. Because of the kernel function, using the SVC is a method with high flexibility, and there is a good out-of-sample generalization when the kernel tuning parameters are appropriately chosen. The SVC suffers from a lack of transparency like other non-parametric methods. Result interpretation can be facilitated using graphical visualization (Figs. 4 and 5).

3.2.6 Artificial neural networks (ANN)

Artificial neural networks are inspired from the biological neural networks and mimic the human neural network. They are composed of nodes playing the role of artificial neurons. The connection between two nodes is ensured by edges, responsible for the transmission of

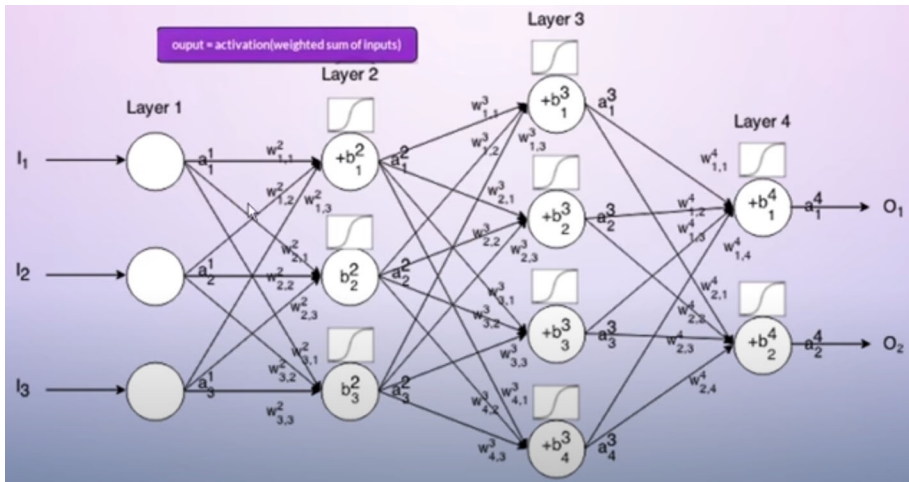


Fig. 5 Artificial neural networks. *Source:* Image courtesy

signals from one node to another. Signals transmitted through the network are assimilated to real numbers, and each node has a threshold above which the signal is significant. The importance of a given connection illustrated by a specific edge is measured using a weight. Many nodes can be joined together to carry out complex computations. As shown in the next figure, the architecture of a neural network can be summarized by a graph whose nodes are neurons and edges are connections between the output of some neuron to the input of another neuron (Shalev-Shwartz & Ben-David, 2014; Anthony & Bartlet, 1999; Kumar et al., 2020). The network consists of three different types of layers: the input layer, which receives the external data; the hidden layer, which performs nonlinear transformations of the inputs entered into the network; and the output layer, which contains all output values. More precisely, each node receives approximated numbers as signals coming from other nodes and computes its specific node output by combining those numbers with weights of all input edges and node bias adjustment using a transfer function. The process continues until the final output is obtained. Then, an observed error is computed after a comparison with the true value. For the case of a classification model whose output is a probability value between 0 and 1, a cross-entropy error can be used. Depending on the size of this error, edge weights and node biases are adjusted through the network and the output values are re-computed until a minimal error is obtained. Using approximation functions, the artificial neural network is a mathematical model that has the advantage of working with any data that can be made numeric. The artificial neural network performs well with nonlinear data and large numbers of inputs. However, this method is computationally expensive with a time-consuming training step. An ANN is often considered a black box, but its major drawback remains the unreadability of the learned knowledge, or the lack of an explanatory capability (Narazaki & Shigaki, 1999).

3.2.7 Recurrent neural network (RNN)

The Recurrent Neural Network is commonly used for ordinary or temporal problems, such as language translation, natural language processing, or speech recognition. The technology

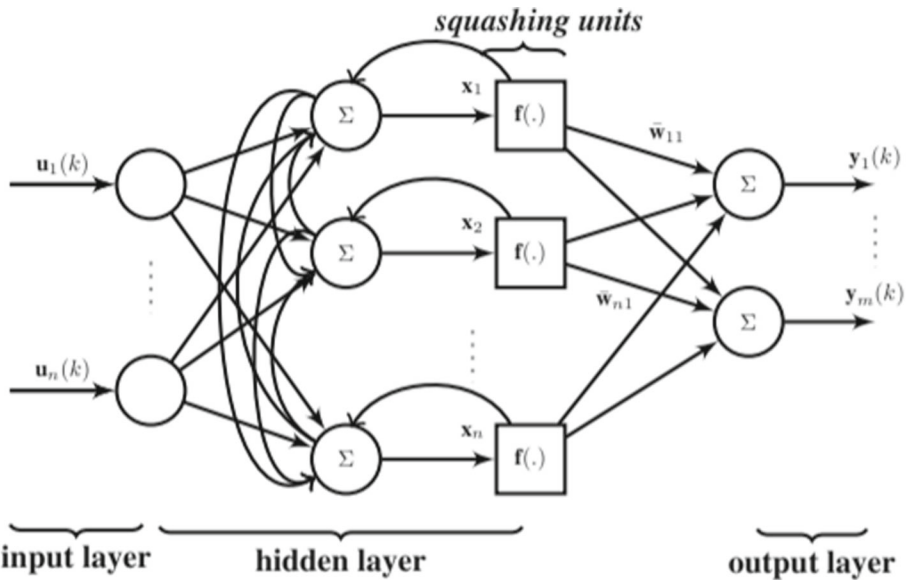


Fig. 6 Recurrent neural network. Reprinted from Graves et al. (2013)

behind the RNN is different from traditional feed-forward and convolutional neural networks. The specificity of this method is related to its memory, as it takes information from prior input to influence the current input and output. Some of these networks have embedded loops, thus enabling information persistence: the loop allows information transfer from one step to the next within the network (see Fig. 6). A recurrent neural network can appear as a set of multiple copies of the same network, each passing a message to a successor. Also, contrary to traditional neural networks, which have different weights across each node, recurrent neural networks share the same weight parameters within each layer of the network. To facilitate reinforcement learning, weights can be adjusted using backpropagation and gradient descent. RNN offers the possibility to process input of any length, with no increase in model size. Computation involving RNN considers historical information and timely sharing of weights. However, the computation time is slow while access to old information is difficult. Another drawback is that the RNN cannot consider any future input for the current state. This means that gradients vanish or explode more often, for instance, when it is hard to capture long-term dependencies, since the number of multiplicative gradients can decrease or increase exponentially according to the number of layers.

3.2.8 Long short-term memory (LSTM)

In traditional RNNs, the model is unable to accurately predict the current state in case of influence by the previous condition. This vanishing gradient problem is resolved by the LSTM, which is an advanced RNN capable of handling long-term dependencies: it memorizes the previous information and uses it for processing the current input (Eachempati et al., 2021). A standard RNN can be represented by a chain of repeating modules of a neural network with a very simple structure, such as a single tanh layer. A LSTM representation is slightly different with four neural network layers interacting in a very special way: a cell state runs straight

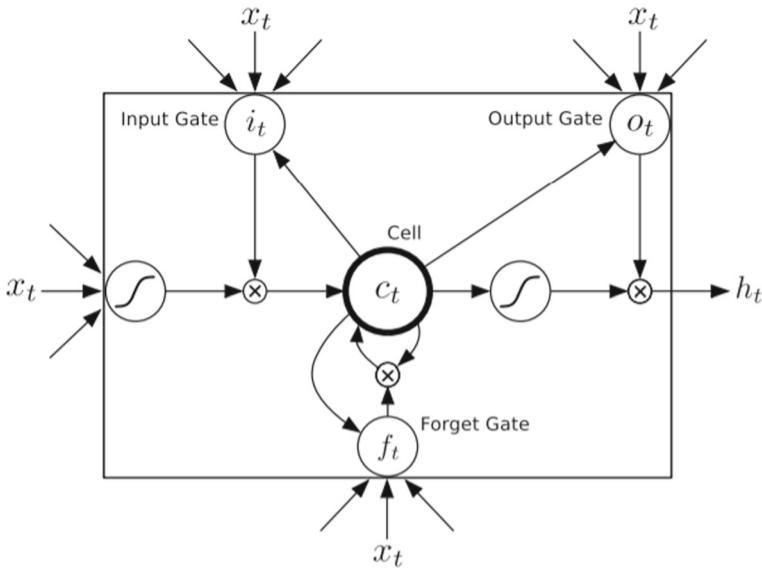


Fig. 7 Long short-term memory. Reprinted from Graves et al. (2013)

down the entire chain, with only some minor linear interactions. By means of structures called gates, the LSTM is able to optionally remove or add information to the cell state. Using a sigmoid layer called the “forget gate layer”, the LSTM decides what information will be thrown away from the cell state; in the next step, the LSTM decides which new information will be stored in the cell state by creating updates, using a combined sigmoid layer called the “input gate layer” (it chooses the values to update) and a tanh layer (which creates a vector for new candidate values). In the final step, the LSTM decides what is going to be the output, using a “output gate” which combines a sigmoid layer (to decide what parts of the cell state will be output) and a tanh layer (to push the values to be between -1 and 1) (Figs. 7, 8 and 9).

4 Experiments and results

The experimental process used in this study can be summarized by the figure below. As we can observe, the main steps are data preparation, model training with cross-validation using machine learning/deep learning techniques, and the model selection.

4.1 Data preparation

One of the most important steps in predictive modelling is data preparation. In practice, data cleaning and preparation is time-consuming and can take around 80% of the total data engineering effort (Zhang et al., 2003). Data preparation mainly includes data collection, data integration, data transformation, data cleaning, data reduction, and data discretization. It is a crucial step since machine learning, and deep learning algorithms require good quality

data to deliver high-quality patterns. Data preparation generates a dataset smaller than the original one, which can significantly improve the efficiency of algorithms.

After the labialization of our dataset, we removed Twitter handles, URLs, hyphens, numbers, and special characters. We removed stop words (e.g., this, have, you, is, that, has, a, do, etc.) from the tweets using a list of English stop words from the NLTK library (<https://www.nltk.org>). We also generated the canonical form of a word using the lemmatization process. At the end of the process, tweets with no content after being processed were removed.

We did some feature engineering in order to create some important variables. This step is crucial for machine learning models (Kumar et al., 2022). Since our goal is to apply machine learning algorithms, we needed numerical values. Thus, we used the TF-IDF library to measure the importance of a term with respect to a document or a collection of documents. We also added sentiment scores for each news item, thanks to the doc2vec library. We created additional numerical representations of each news item regardless of its length by measuring: the number of words, the average number of syllables, the number of characters and the number of unique terms. Thus, beyond TF-IDF and sentiment scores, our input dataset contains information on the modified FKRA grade, the modified FRE score, the number of syllables, the average number of syllables, the number of characters, the total number of characters, the number of words, and the number of unique terms. The table below illustrates these eight additional characteristics.

After creating all these features, we aggregated them all into a single matrix that we used to train the algorithms and test them. The data were split into two parts: training set (80%), and test set (20%). The training set was used to build the model, the performance of which was evaluated on the test set.

4.2 Training

The cleaned data set was first divided into two groups: 80% for the training set (corresponding to 16,000 tweets), and 20% for the testing set (a total of 4000 tweets). We implemented tenfold cross-validation for each of our machine and deep learning algorithms with a set of hyperparameters as given in the following table.

4.3 Performance evaluation

To assess the predictive performances of our Machine Learning and Deep Learning text classifiers, some evaluation metrics are needed. As they are commonly used in the literature (see, e.g. Tchuente & Nyawa, 2021, Kumar et al., 2016, 2018), the following measures need to be defined. Accuracy is the proportion of messages correctly predicted by the model among the total number of cases examined. Also called positive predictive value, precision is the proportion of hesitant tweets that are correctly predicted by the model over all vaccine-hesitant predictions. Recall refers to the proportion of hesitant tweets that are correctly predicted by the model over all vaccine-hesitant tweets.

- $F_1 \text{ Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
- $\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total number of Prediction}}$
- $\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$
- $\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$

4.4 Results

When we analysed tweets of people who hesitate to be vaccinated, we observed that they worried mainly about the safety of new COVID-19 vaccines. To be more precise, an analysis of the word cloud of those tweets reveals that hesitant twitterers have a lack of confidence about governments, vaccine efficacy or safety. They find available vaccines to be experimental given the hastily development of the Pfizer vaccine in the USA or vaccines developed in China or Russia. They worry about registered deaths from vaccinated individuals. The following word cloud summarizes their feelings and present most frequent words they used to express their selves. Those words play a determinant role in our machine and deep learning algorithms in order to identify hesitant tweets.

Table 5 shows the performance of the top LR, RF, SVC, KNN, GD, DT, Gboost, AdaBoost, and ANN models that were evaluated on the test set. The RF model outperformed the other models with an accuracy rate of 83%. The second bests are the LR, ANN and Gboots with respectively 82.5%, 81.3% and 81.2% of accuracy. With respective precision rates of 65.5% and 68.6%, GD and KNN correspond to the machine learning methods with the lowest accuracy coefficients.

There are many reasons that can justify why RF outperforms the other competitors. First, some algorithms require extensive tuning for optimal performance. This is not the case with Random Forest, which can be advantageous in resource-limited scenarios. Second, due to randomness, the error rate of RF is low, which improves prediction results. Third, since the number of features is big, and many complex interactions between them could exist, RF is appropriate. Fourth, RF is easier to tune than ANN and Gboost. Fifth, RF will not overfit almost certainly if the data is neatly pre-processed and cleaned, which is not the case for other competitors.

We add to this comparison exercise deep learning algorithms such as the LSTM, RNN and a light version of the LSTM.

Table 6 shows the performance of the LSTM models on the validation set. We reported results for LSTM models with 50 units as these outperformed those with 25 units. In general, the performance of these 50-unit models was slightly different across learning rates and epochs. The top performer was the LSTM-50 model, that used a learning rate of 0.0001 and was trained for 10 epochs. The validation accuracy was 82.15% for this model, while the validation loss stood at 0.7576.

Table 5 Performance of machine learning models on the testing sets

	Accuracy	Precision	Recall	Micro F1	Macro F1	Weighted F1	F1 score
LR	0.825	0.607	0.604	0.797	0.725	0.798	0.605
RF	0.830	0.471	0.615	0.830	0.727	0.815	0.532
SVC	0.746	0.736	0.510	0.746	0.713	0.764	0.599
KNN	0.655	0.566	0.402	0.655	0.612	0.677	0.466
GD	0.686	0.674	0.448	0.686	0.654	0.707	0.534
DT	0.777	0.538	0.520	0.777	0.702	0.779	0.528
Gboost	0.812	0.581	0.620	0.812	0.749	0.816	0.600
AdaBoost	0.786	0.612	0.541	0.786	0.725	0.793	0.574
ANN	0.813	0.265	0.547	0.813	0.632	0.762	0.342

Table 6 Performance of LSTM models on the validation sets

Learning rate	Epoch	Training loss	Training accuracy	Validation loss	Validation accuracy
0.0001	10	0.2077	0.9402	0.7576	0.8215
	20	0.0858	0.9755	0.9843	0.8004
	30	0.0815	0.9774	1.0438	0.8089
	40	0.1290	0.9638	0.8218	0.8004
	50	0.0948	0.9718	1.0258	0.7709
	60	0.0770	0.9784	0.9674	0.8173
	70	0.0733	0.9807	0.9745	0.8046
	80	0.0871	0.9722	1.0035	0.8173
	90	0.0759	0.9760	1.0670	0.8089
	100	0.0832	0.9774	0.9803	0.8004
0.0005	10	0.0364	0.9901	1.7323	0.8089
	20	0.0315	0.9906	1.6416	0.7667
	30	0.0328	0.9882	1.6626	0.7793
	40	0.0396	0.9835	1.4843	0.8089
	50	0.0178	0.9939	2.0509	0.7878
	60	0.0293	0.9901	1.7441	0.7835
	70	0.0371	0.9868	1.6550	0.7371
	80	0.0175	0.9939	1.8112	0.7751
	90	0.0308	0.9911	1.6600	0.7920
	100	0.0426	0.9878	1.4149	0.7920
0.001	10	0.0247	0.9915	1.9479	0.8089
	20	0.0308	0.9906	1.7341	0.7709
	30	0.0295	0.9929	1.9028	0.8173
	40	0.0144	0.9953	2.1552	0.7751
	50	0.0235	0.9911	2.1632	0.8004
	60	0.0240	0.9920	1.8590	0.7751
	70	0.0208	0.9920	1.8483	0.8215
	80	0.0322	0.9896	1.7984	0.8173
	90	0.0324	0.9882	1.9341	0.7709
	100	0.0130	0.9948	2.3347	0.8004
0.01	10	0.0603	0.9760	2.2411	0.8173
	20	0.0539	0.9835	1.7254	0.7751
	30	0.0566	0.9831	2.0825	0.7835
	40	0.0483	0.9849	1.3854	0.7962
	50	0.0512	0.9807	2.0150	0.7751
	60	0.0432	0.9868	3.1149	0.8215
	70	0.0550	0.9807	1.4100	0.8131
	80	0.0346	0.9859	2.6846	0.7709
	90	0.0410	0.9864	1.9356	0.7835
	100	0.0525	0.9845	1.7164	0.8215

Performance on the validation set of the RNN models is shown in Table 7 below. With accuracies generally above 80%, all RNN models performed well. With a learning rate of 0.001 trained for 40 epochs, the top performer model was the RNN with 25 units. For this model, the validation accuracy was 83.28% and the validation loss 0.6310.

We also check what could be the performances on the validation set of a simpler LSTM model. Thus, we create a “light LSTM” model with no intermediary layers and only one dense final layer. The best performance of this last deep learning model is obtained for a learning rate of 0.0001, 50 epochs and 25 units. The corresponding accuracy is 87.22% and the resulting loss 0.5625 (Table 8).

This comparison exercise based on the validation sets indicates that the light LSTM dominates the other competitors with an accuracy of 87.22%. However, a fair comparison is supposed to be done on a completely independent data set. For this reason, we carried out another comparison exercise based entirely on a testing set.

Based on the testing set, the Table 9 shows that the LSTM model with the higher accuracy is obtained when the learning rate is equal to 0.0001, with the number of epochs of 30. Resulting performance metrics correspond to an accuracy rate of 86%, a precision rate of 80%, a recall rate of 82% and an F1-score of 80%.

Table 10 shows the performances of RNN models for different learning rates and epochs on the testing set. As we can observe, the best performing RNN model is obtained when the learning rate is equal to 0.005 and when the number of epochs is 10. Therefore, the leading performance measures appear as follows: accuracy, 83%; precision, 85%; recall, 83%; and F1 score, 72%.

On the testing set, the performances of the light LSTM models are quite stable over different hyperparameters. Contrary to its dominance over the different deep learning models on validation set, on the testing set, the light LSTM achieve an accuracy rate of 85% (smaller than the 87.22% obtained on the validation set), plus a precision rate of 84%, a recall rate of 85% and an F1 score of 84% (Table 11).

Putting everything together, it comes out that if we only care about predictive power on the testing set, then the top performing models within our selected machine and deep learning models are deep learning models. More specifically, LSTM models with 50 units completely dominate the other competing models with an accuracy rate of 86%. Its light version with 25 units comes second with 85% of accuracy. The RNN and the RF perform similarly, but we only compared their accuracy rates. In contrast, the RNN performs better if we consider their precision recall rates as well as their F1 scores.

The size of the available data firstly explains the outperformance of Deep Learning models: when the data size is large, Deep Learning most often outperforms other machine learning techniques (To et al., 2021). This is because deep learning algorithms need a large amount of data to understand it perfectly. Secondly, when there is a lack of domain understanding for feature introspection, Deep Learning techniques outshine others as you have to worry less about feature engineering. Thirdly, Deep Learning shines when it comes to complex problems such as text classification.

5 Discussion, implications, limitations, and future research directions

This study aims to evaluate the performance of the different machine and deep learning models' performance to identify vaccine-hesitant tweets published during the COVID-19 pandemic. Our findings showed that LSTM and RNN models outperformed traditional

Table 7 Performance of RNN models on the validation sets

Learning rate	Epoch	Training loss	Training accuracy	Validation loss	Validation accuracy
0.0001	10	0.3816	0.8900	0.5605	0.8225
	20	0.2641	0.9217	0.5599	0.8242
	30	0.1821	0.9932	0.8249	0.8225
	40	0.0126	0.9987	0.6310	0.8328
	50	0.1560	0.9992	1.0189	0.8225
	60	0.1472	0.9992	0.9325	0.8225
	70	0.1387	0.9996	1.0169	0.8208
	80	0.1337	0.9996	1.0942	0.8225
	90	0.1245	0.9996	1.1369	0.8191
	100	0.1171	1.0000	1.2516	0.8242
0.0005	10	0.1683	0.9966	0.7601	0.8286
	20	0.1222	0.9996	0.9602	0.8242
	30	8.6605e−04	1.0000	0.9204	0.8184
	40	3.4382e−04	1.0000	1.0501	0.8158
	50	0.0580	1.0000	1.0768	0.8140
	60	0.0486	1.0000	1.1318	0.8090
	70	4.9768e−05	1.0000	1.2876	0.6904
	80	4.7368e−05	1.0000	1.2700	0.8108
	90	2.4053e−05	1.0000	1.3996	0.8039
	100	1.9185e−05	1.0000	1.4046	0.6818
0.001	10	0.0046	0.9992	0.7631	0.8286
	20	0.0798	0.9996	0.8916	0.8140
	30	4.3638e−04	1.0000	1.1188	0.8225
	40	0.0338	1.0000	1.1959	0.8184
	50	6.2307e−05	1.0000	1.2890	0.8184
	60	0.0163	1.0000	1.2956	0.8090
	70	0.0115	1.0000	1.3688	0.8124
	80	0.0100	1.0000	1.1483	0.6988
	90	8.6857e−06	1.0000	1.5743	0.8022
	100	0.0045	1.0000	1.3327	0.8005
0.01	10	0.5895	0.7240	0.5906	0.8225
	20	0.5894	0.7240	0.5906	0.8225
	30	0.5894	0.7240	0.5907	0.8225
	40	4.8075e−06	1.0000	2.1718	0.6853
	50	0.5895	0.7240	0.5906	0.8225
	60	0.5896	0.7240	0.5906	0.8225
	70	0.5896	0.7240	0.5906	0.8225
	80	2.0296e−06	1.0000	2.0767	0.6954
	90	0.5894	0.7240	0.5907	0.8225
	100	6.6937e−07	1.0000	2.1622	0.6880

Table 8 Performance of light LSTM models on the validation sets

Learning rate	Epoch	Training loss	Training accuracy	Validation loss	Validation accuracy
0.0001	10	0.2925	0.8993	0.4976	0.8637
	20	0.1657	0.9647	0.5961	0.8637
	30	0.1579	0.9581	0.5990	0.8553
	40	0.1684	0.9638	0.5661	0.8553
	50	0.1813	0.9525	0.5625	0.8722
	60	0.1657	0.9605	0.6230	0.8637
	70	0.1597	0.9624	0.5922	0.8553
	80	0.1549	0.9605	0.6110	0.8637
	90	0.1843	0.9544	0.5814	0.8553
	100	0.1527	0.9675	0.5930	0.8637
0.0005	10	0.0257	0.9958	1.0331	0.7962
	20	0.0290	0.9962	1.0413	0.8173
	30	0.0323	0.9939	0.9663	0.8131
	40	0.0248	0.9962	1.1099	0.8131
	50	0.0297	0.9958	1.0660	0.8173
	60	0.0274	0.9962	1.0400	0.8004
	70	0.0384	0.9920	0.9841	0.8215
	80	0.0391	0.9929	0.9981	0.8468
	90	0.0259	0.9953	1.0317	0.8215
	100	0.0339	0.9929	1.0758	0.8215
0.001	10	0.0108	0.9972	1.2537	0.8257
	20	0.0167	0.9962	1.3142	0.7962
	30	0.0122	0.9962	1.4101	0.7793
	40	0.0096	0.9981	1.3247	0.8046
	50	0.0136	0.9962	1.2704	0.7962
	60	0.0189	0.9967	1.2359	0.7667
	70	0.0099	0.9976	1.2931	0.8046
	80	0.0142	0.9958	1.2112	0.8215
	90	0.0124	0.9972	1.2930	0.7920
	100	0.0101	0.9976	1.2897	0.7920
0.01	10	0.0025	0.9995	1.8656	0.8004
	20	0.0039	0.9991	1.7533	0.8300
	30	0.0080	0.9976	1.5326	0.8215
	40	0.0033	0.9981	1.4588	0.8342
	50	0.0028	0.9991	1.7429	0.8257
	60	0.0031	0.9995	1.7227	0.7878
	70	0.0037	0.9995	1.6904	0.7793
	80	0.0039	0.9986	1.6116	0.7962
	90	0.0059	0.9991	1.4797	0.8342
	100	0.0048	0.9981	1.6606	0.8173

Table 9 Performance of LSTM models on the testing sets

Learning rate	Epoch	Accuracy	Precision	Recall	F1 score
0.0001	10	0.83	0.80	0.83	0.69
	20	0.85	0.83	0.85	0.83
	30	0.86	0.80	0.82	0.80
	40	0.85	0.83	0.85	0.83
	50	0.81	0.81	0.81	0.81
	60	0.85	0.82	0.85	0.82
	70	0.82	0.81	0.82	0.81
	80	0.83	0.81	0.83	0.81
	90	0.83	0.81	0.83	0.82
	100	0.84	0.82	0.84	0.82
0.0005	10	0.83	0.80	0.83	0.81
	20	0.80	0.80	0.80	0.80
	30	0.82	0.81	0.82	0.81
	40	0.82	0.82	0.82	0.82
	50	0.83	0.80	0.83	0.81
	60	0.83	0.83	0.83	0.83
	70	0.82	0.82	0.82	0.82
	80	0.83	0.82	0.83	0.82
	90	0.82	0.80	0.82	0.81
	100	0.83	0.81	0.83	0.81
0.001	10	0.83	0.81	0.83	0.81
	20	0.83	0.82	0.83	0.82
	30	0.82	0.80	0.82	0.81
	40	0.82	0.80	0.82	0.81
	50	0.82	0.80	0.82	0.80
	60	0.82	0.82	0.82	0.82
	70	0.83	0.80	0.83	0.81
	80	0.84	0.82	0.84	0.83
	90	0.81	0.81	0.81	0.81
	100	0.81	0.69	0.81	0.80
0.01	10	0.82	0.80	0.82	0.80
	20	0.82	0.81	0.82	0.81
	30	0.69	0.69	0.69	0.69
	40	0.81	0.69	0.81	0.80
	50	0.68	0.69	0.68	0.69
	60	0.84	0.82	0.84	0.82
	70	0.82	0.69	0.82	0.69
	80	0.82	0.82	0.82	0.82
	90	0.80	0.68	0.80	0.69
	100	0.82	0.80	0.82	0.81

Table 10 Performance of RNN models on the testing sets

Learning rate	Epoch	Accuracy	Precision	Recall	F1 score
0.0001	10	0.82	0.62	0.82	0.71
	20	0.82	0.80	0.82	0.71
	30	0.82	0.62	0.82	0.71
	40	0.83	0.80	0.83	0.79
	50	0.82	0.52	0.82	0.71
	60	0.82	0.77	0.82	0.71
	70	0.82	0.71	0.82	0.71
	80	0.82	0.77	0.82	0.71
	90	0.82	0.59	0.82	0.71
	100	0.82	0.79	0.82	0.72
0.0005	10	0.83	0.85	0.83	0.72
	20	0.82	0.78	0.82	0.73
	30	0.82	0.79	0.82	0.80
	40	0.82	0.79	0.82	0.80
	50	0.81	0.75	0.81	0.75
	60	0.81	0.74	0.81	0.74
	70	0.79	0.78	0.79	0.78
	80	0.81	0.79	0.81	0.79
	90	0.80	0.78	0.80	0.78
	100	0.78	0.78	0.78	0.78
0.001	10	0.83	0.80	0.83	0.80
	20	0.81	0.77	0.81	0.77
	30	0.82	0.79	0.82	0.79
	40	0.82	0.78	0.82	0.78
	50	0.82	0.79	0.82	0.80
	60	0.81	0.78	0.81	0.78
	70	0.81	0.78	0.81	0.78
	80	0.80	0.78	0.80	0.78
	90	0.80	0.78	0.80	0.79
	100	0.80	0.78	0.80	0.78
0.01	10	0.82	0.52	0.82	0.71
	20	0.82	0.52	0.82	0.71
	30	0.82	0.52	0.82	0.71
	40	0.79	0.78	0.79	0.78
	50	0.82	0.52	0.82	0.71
	60	0.82	0.52	0.82	0.71
	70	0.82	0.52	0.82	0.71
	80	0.80	0.77	0.80	0.78
	90	0.82	0.52	0.82	0.71
	100	0.79	0.75	0.79	0.77

Table 11 Performance of light LSTM models on the testing sets

Learning rate	Epoch	Accuracy	Precision	Recall	F1 score
0.0001	10	0.84	0.81	0.84	0.79
	20	0.84	0.81	0.84	0.81
	30	0.83	0.81	0.83	0.82
	40	0.84	0.82	0.84	0.82
	50	0.85	0.83	0.85	0.83
	60	0.83	0.80	0.83	0.81
	70	0.84	0.82	0.84	0.82
	80	0.83	0.81	0.83	0.81
	90	0.83	0.81	0.83	0.82
	100	0.85	0.82	0.85	0.81
0.0005	10	0.83	0.82	0.83	0.82
	20	0.83	0.82	0.83	0.83
	30	0.83	0.82	0.83	0.82
	40	0.83	0.82	0.83	0.82
	50	0.83	0.82	0.83	0.82
	60	0.81	0.81	0.81	0.81
	70	0.81	0.80	0.81	0.81
	80	0.84	0.82	0.84	0.83
	90	0.83	0.82	0.83	0.82
	100	0.83	0.82	0.83	0.82
0.001	10	0.83	0.81	0.83	0.82
	20	0.82	0.81	0.82	0.82
	30	0.81	0.81	0.81	0.81
	40	0.82	0.82	0.82	0.82
	50	0.82	0.82	0.82	0.82
	60	0.82	0.82	0.82	0.82
	70	0.83	0.82	0.83	0.82
	80	0.85	0.84	0.85	0.84
	90	0.81	0.80	0.81	0.81
	100	0.82	0.82	0.82	0.82
0.01	10	0.80	0.80	0.80	0.80
	20	0.82	0.80	0.82	0.80
	30	0.83	0.82	0.83	0.82
	40	0.83	0.80	0.83	0.81
	50	0.82	0.81	0.82	0.81
	60	0.81	0.82	0.81	0.81
	70	0.83	0.82	0.83	0.82
	80	0.81	0.80	0.81	0.80
	90	0.83	0.81	0.83	0.82
	100	0.80	0.69	0.80	0.69

machine learning models across all performance metrics (accuracy, precision, recall, and F1 score). After deep learning models, the next top performers were RF and LR. The other classic machine learning models did not perform as well as for the identification of vaccine-hesitant tweets. With an accuracy of 86%, LSTM did very well on this text classification since other performance metrics were above 80%. This finding is consistent with other studies whose results show a better performance of deep learning-based models compared to classic machine learning methods on vaccine tweets sentiment classification (see, e.g., Du et al., 2020; Zhang et al., 2020; Tomaszewski et al., 2021; To et al., 2021).

5.1 Implications for research

Most existing studies that use social media to access vaccine hesitancy generally focus on analyzing global opinion about vaccination (positive, negative, neutral) or antivaccine concerns (see Table 1). The concept of vaccine hesitancy, which is our focus in this study, is different from strongly polarized opinions and is more subtle (Dubé et al., 2021). Vaccine hesitancy represents a shift from the dichotomous perspective on whether one is against or for a vaccine to an approach discussing the potential attitudes of people (active demand for vaccines, full rejection of all vaccines, etc.). Vaccine-hesitant individuals are a heterogeneous group along this continuum. Thus, identifying vaccine-hesitant content using machine learning or deep learning techniques is more challenging as the labelling step in particular must take into account all these nuances. Our study is different from existing studies on this specific point and provides a complementary approach that could improve the quality of information that inform strategies aimed at reducing hesitant-vaccination sentiments (Alamoodi et al., 2021; Karafillakis et al., 2021).

In the operations management context, this approach can serve as an efficient resource for disease surveillance, especially as regards communication during disease outbreaks (Anparasan & Lejeune, 2019; Kumar et al., 2021a, 2021b). It can also provide another way of reducing the impact of the pandemic on production systems by using digital technologies (Dubey et al., 2019a, 2019b; DuHadway et al., 2019; Fast et al., 2018; Griffith et al., 2019; Gupta et al., 2021; Singh et al., 2019; Wamba et al., 2019).

5.2 Limitations

This study has some limitations. Since access to tweets is restricted due to the Twitter data policy, the tweets used in this paper accounted for only a few percentage of daily tweets. Also, twitterers are not representative of the world population. Consequently, collected messages may not be representative from a global perspective. In addition, model fine-tuning has been limited to a few hyperparameters (learning rates, batch sizes, and the number of epochs, etc.), thereby ignoring some other parameters. The performance of these models might have been improved further if the tuning had been conducted more widely. However, we consider that the performance of LSTM models in this study was good enough to be used to identify vaccine hesitant tweets in future studies.

5.3 Implications for practice

Since hesitation about vaccination efficiently spread through social networks, it strongly impacts the population's decision to get vaccinated. As a result, a significant part of the

effort of governments, medical and social science researchers to develop the COVID-19 vaccine will end up wasted. Concerns about vaccination among the population should be automatically and efficiently detected in social media to monitor changes in vaccine-hesitancy behaviors. Machine and deep learning models developed in this paper provide such useful tools in this regard. Governments or institutions can rely on these algorithms to fight against negative influences of social media hesitant messages following the compilation of strategic information that could help reduce hesitant-vaccination sentiments.

WHO identifies vaccine hesitancy as one of the more important global health threats with a strongly negative impact on the vaccine demand. In the same light, vaccine hesitancy has been linked to the reduced vaccine acceptance rates and the recurrence of epidemics (Dubé et al., 2015). This clearly shows that vaccine hesitancy jeopardizes not only the hesitant individual's safety but also the safety of the entire community (Verelst et al., 2019). Population immunity, which is called "herd immunity", could only be achieved when a large proportion of the population acquires vaccination (Bhopal, 2020). According to WHO, incorporating vaccine hesitancy assessment into health policy-making is essential to help evaluate public opinions and behaviors in relation to vaccines (Domek et al., 2018; Kwok et al., 2021). The approach used in this paper can be a part of the solution in this specific context.

5.4 Some future research directions

Some directions for future research include further investigation of the role of social media in the accentuation of the COVID-19 vaccine hesitancy. We have already provided an appropriate identification tool. An in-depth study of such a role may be conducted through an additional content analysis using methods such as topic modeling, trend analysis or social network analysis (Cotfas et al., 2021; Karami et al., 2021; Yuan et al., 2019). In the same spirit, it could also be a plus to know the degree to which vaccine hesitant discourses influence vaccination attitudes and behaviors among the population, and how vaccine hesitancy information spreads within the social media user community.

6 Conclusion

Vaccine hesitancy has always existed, but the global antivaccine sentiment, the causes, consequences and impact of vaccination resistance have been the focus of much research over the past decade. This study aimed to emphasize hesitancy behavior with respect to vaccination by evaluating how performant machine and deep learning models can be in the identification of vaccine hesitant tweets during the COVID-19 vaccination campaigns across countries. We derived that LSTM and RNN models outperformed traditional machine learning models in detecting vaccine hesitant messages on social media, with 86% and 83% accuracy, respectively. Despite some few limitations related to both the sample representativeness of the world opinion or restrictions on parameters tuning, LSTM and RNN models achieved good performance and can be used to identify vaccine hesitant tweets in future studies. In this regard, the role of social media in the accentuation of the COVID-19 vaccine hesitancy may be deepened, as should also be conducted a study on the spread of discouraging information among different social media users.

References

- Abd Rahim, N. H., & Rafie, S. H. M. (2020). Sentiment analysis of social media data in vaccination. *International Journal*, 8(9).
- Akyildirim E., Corbet S., Efthymiou M., Guiomard C., O'Connell J., & Sensoy A. (2020). The financial market effects of international aviation disasters. *International Review of Financial Analysis*, 69, 101468.
- Alam, S. T., Ahmed, S., Ali, S. M., Sarker, S., & Kabir, G. (2021). Challenges to COVID-19 vaccine supply chain: Implications for sustainable development goals. *International Journal of Production Economics*, 239, 108193.
- Alamoodi, A. H., Zaidan, B. B., Al-Masawa, M., Taresh, S. M., Noman, S., Ahmaro, I. Y. Y., Garfan, S., Chen, J., Ahmed, M. A., Zaidan, A. A., Albahri, O. S., Aickelin, U., Thamir, N. N., Fadhil, J. A., & Salahaldin, A. (2021). Multi-perspectives systematic review on the applications of sentiment analysis for vaccine hesitancy. *Computers in Biology and Medicine*, 139, 104957.
- Anparasan, A., & Lejeune, M. (2019). Resource deployment and donation allocation for epidemic outbreaks. *Annals of Operations Research*, 283(1), 9–32.
- Anthony, M., & Bartlett, P. L. (1999). *Neural Network Learning: Theoretical Foundations*. Cambridge University Press.
- Argyris, Y. A., Monu, K., Tan, P.-N., Aarts, C., Jiang, F., & Wiseley, K. A. (2021). Using machine learning to compare provaccine and antivaccine discourse among the public on social media: Algorithm development study. *JMIR Public Health and Surveillance*, 7(6), e23105.
- Bag, S., Gupta, S., Choi, T.-M., & Kumar, A. (2021). Roles of innovation leadership on using big data analytics to establish resilient healthcare supply chains to combat the COVID-19 pandemic: A multimethodological study. *IEEE Transactions on Engineering Management*. <https://doi.org/10.1109/TEM.2021.3101590>
- Bar-Lev, S., Reichman, S., & Barnett-Itzhaki, Z. (2021). Prediction of vaccine hesitancy based on social media traffic among Israeli parents using machine learning strategies. *Israel Journal of Health Policy Research*, 10(1), 1–8.
- Bhopal, R. S. (2020). COVID-19 zugzwang: Potential public health moves towards population (herd) immunity. *Public Health in Practice*, 1, 100031.
- Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In: *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144–152).
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Budd, J., Miller, B. S., Manning, E. M., et al. (2020). Digital technologies in the public-health response to COVID-19. *Nature Medicine*, 26, 1183–1192.
- Chen, B., Bai, R., Li, J., Liu, Y., Xue, N., & Ren, J. (2020). A multiobjective single bus corridor scheduling using machine learning-based predictive models. *International Journal of Production Research*. <https://doi.org/10.1080/00207543.2020.1766716>
- Chevallier, C., Hacquin, A. S., & Mercier, H. (2021). COVID-19 vaccine hesitancy: Shortening the last mile. *Trends in Cognitive Sciences*, 25(5), 331–333.
- Cohen, M. C. (2018). Big data and service operations. *Production and Operations Management*, 27(9), 1709–1723.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
- Cotfas, L.-A., Delcea, C., & Gherai, R. (2021). COVID-19 vaccine hesitancy in the month following the start of the vaccination process. *International Journal of Environmental Research and Public Health*, 18(19), 10438.
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *Information Theory*, 13(1), 21–27.
- Cui, R., Gallino, S., Moreno, A., & Zhang, D. J. (2018). The operational value of social media information. *Production and Operations Management*, 27(10), 1749–1769.
- Devroye, L., Györfi, L., & Lugosi, G. (1996). *A probabilistic theory of pattern recognition*. Springer.
- Domek, G. J., O'Leary, S. T., Bull, S., Bronsert, M., Contreras-Roldan, I. L., Ventura, G. A., Kempe, A., & Asturias, E. J. (2018). Measuring vaccine hesitancy: Field testing the WHO SAGE working group on vaccine hesitancy survey tool in Guatemala. *Vaccine*, 36(35), 5273–5281.
- Du, J., Cunningham, R. M., Xiang, Y., et al. (2020). Leveraging deep learning to understand health beliefs about the Human Papillomavirus Vaccine from social media. *NPJ Digital Medicine*, 2–27.
- Dubé, E., Vivion, M., & MacDonald, N. E. (2015). Vaccine hesitancy, vaccine refusal and the antivaccine movement: influence, impact and implications. *Expert Review of Vaccines*, 14(1), 99–117.
- Dubé, É., Ward, J. K., Verger, P., & MacDonald, N. E. (2021). Vaccine hesitancy, acceptance, and anti-vaccination: trends and future prospects for public health. *Annual Review of Public Health*, 42, 175–191.
- Dubey, R., Altay, N., & Blome, C. (2019a). Swift trust and commitment: The missing links for humanitarian supply chain coordination? *Annals of Operations Research*, 283(1), 159–177.

- Dubey, R., Gunasekaran, A., & Papadopoulos, T. (2019b). Disaster relief operations: Past, present and future. *Annals of Operations Research*, 283(1), 1–8.
- DuHadway, S., Carnovale, S., & Hazen, B. (2019). Understanding risk management for intentional supply chain disruptions: Risk detection, risk mitigation, and risk recovery. *Annals of Operations Research*, 283(1), 179–198.
- Eachempati, P., Srivastava, P. R., Kumar, A., Tan, K. H., & Gupta, S. (2021). Validating the impact of accounting disclosures on stock market: A deep neural network approach. *Technological Forecasting and Social Change*, 170, 120903.
- ElonPoll. (2020). North Carolina willingness to take COVID-19 vaccine. The Charlotte Observer, The Durham Herald-Sun & The Raleigh News & Observer.
- Fast, S. M., Kim, L., Cohn, E. L., Mekaru, S. R., Brownstein, J. S., & Markuzon, N. (2018). Predicting social response to infectious disease outbreaks from internet-based news streams. *Annals of Operations Research*, 263(1), 551–564.
- Fornell, C., Larcker, D. F., & MacDonald, N. E. (2015). Vaccine hesitancy: Definition, scope and determinants. *Journal of Marketing Research*, 18(34), 4161–4164.
- Freeman, D., Waite, F., Rosebrock, L., Petit, A., Causier, C., East, A., & Lambe, S. (2022). Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. *Psychological Medicine*, 52(2), 251–263.
- Freund, Y., & Schapire, R. E. (1995). A decision-theoretic generalization of on-line learning and an application to boosting. In *European Conference on Computational Learning Theory*, pp. 23–37.
- Furini, M. (2021). Identifying the features of ProVax and NoVax groups from social media conversations. *Computers in Human Behavior*, 120, 106751.
- Germani, F., & Biller-Andorno, N. (2021). The anti-vaccination infodemic on social media: A behavioral analysis. *PLoS ONE*, 16(3), e0247642.
- Graves, A., Mohamed, A., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, no. 3, pp. 6645–6649.
- Griffith, D. A., Boehmke, B., Bradley, R. V., Hazen, B. T., & Johnson, A. W. (2019). Embedded analytics: Improving decision support for humanitarian logistics operations. *Annals of Operations Research*, 283(1), 247–265.
- Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33, 2223–2273.
- Gupta, M., Shoja, A., & Mikalef, P. (2021). Toward the understanding of national culture in the success of non-pharmaceutical technological interventions in mitigating COVID-19 pandemic. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-021-03962-z>
- Hussain, A., Tahir, A., Hussain, Z., Sheikh, Z., Gogate, M., Dashtipour, K., Ali, A., & Sheikh, A. (2021). Artificial intelligence-enabled analysis of public attitudes on facebook and twitter toward covid-19 vaccines in the united kingdom and the united states: Observational study. *Journal of Medical Internet Research*, 23(4), e26627.
- Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. *International Journal of Production Research*, 58(10), 2904–2915.
- Jiang, X., Su, M.-H., Hwang, J., Lian, R., Brauer, M., Kim, S., & Chin, J. (2021). Identifying false human papillomavirus (HPV) vaccine information and corresponding risk perceptions from twitter: advanced predictive models. *Journal of Medical Internet Research*, 23(9), 20563051211048412.
- Kapoor, K., Bigdeli, A. Z., Dwivedi, Y. K., & Raman, R. (2021). How is COVID-19 altering the manufacturing landscape? A literature review of imminent challenges and management interventions. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-021-04397-2>
- Karafilakis, E., Martin, S., Simas, C., Olsson, K., Takacs, J., Dada, S., & Larson, H. J. (2021). Methods for social media monitoring related to vaccination: Systematic scoping review. *JMIR Public Health and Surveillance*, 7(2), e17149.
- Karami, A., Zhu, M., Goldschmidt, B., Boyajieff, H. R., & Najafabadi, M. M. (2021). COVID-19 vaccine and social media in the US: Exploring emotions and discussions on twitter. *Vaccines*, 9(10), 1059.
- Kumar, A., Choi, T. M., Wamba, S. F., Gupta, S., & Tan, K. H. (2021a). Infection vulnerability stratification risk modelling of COVID-19 data: A deterministic SEIR epidemic model analysis. *Annals of Operations Research*, 4, 1–27.
- Kumar, A., Gopal, R. D., Shankar, R., & Tan, K. H. (2022). Fraudulent review detection model focusing on emotional expressions and explicit aspects: Investigating the potential of feature engineering. *Decision Support Systems*, 155, 113728.

- Kumar, A., Shankar, R., & Aljohani, N. (2020). A big data driven framework for demand-driven forecasting with effects of marketing-mix variables. *Industrial Marketing Management*, 90, 493–507.
- Kumar, A., Shankar, R., Choudhary, A., & Thakur, L. (2018). A big data MapReduce framework for fault diagnosis in cloud-based manufacturing. *International Journal of Production Research*, 27, 428–439.
- Kumar, A., Shankar, R., & Thakur, L. (2016). A big data driven sustainable manufacturing framework for condition-based maintenance prediction. *Journal of Computational Science*, 54, 7060–7073.
- Kumar, S., Xu, C., Ghildayal, N., Chandra, C., & Yang, M. (2021b). Social media effectiveness as a humanitarian response to mitigate influenza epidemic and COVID-19 pandemic. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-021-03955-y>
- Kusiak, A. (2020). Convolutional and generative adversarial neural networks in manufacturing. *International Journal of Production Research*, 58(5), 1594–1604.
- Kwok, S. W. H., Vadde, S. K., & Wang, G. (2021). Twitter speaks: An analysis of Australian twitter users' topics and sentiments about COVID-19 vaccination using machine learning. *Journal of Medical Internet Research*. <https://doi.org/10.2196/26953>
- Lazarus, J. V., Ratzan, S. C., Palayew, A., Gostin, L. O., Larson, H. J., Rabin, K., Kimball, S., & El-Mohandes, A. (2021). A global survey of potential acceptance of a Covid-19 vaccine. *Nature Medicine*, 27, 225–228.
- Ma, P., Zeng-Treitler, Q., & Nelson, S. J. (2021). Use of two topic modeling methods to investigate covid vaccine hesitancy. *International Conference on ICT, Society and Human Beings*, 384, 221–226.
- Majumdar, P., Biswas, A., & Sahu S. (2020). COVID-19 pandemic and lockdown: Cause of sleep disruption, depression, somatic pain, and increased screen exposure of office workers and students of India. *Chronobiology International*, 37(8), 1191–1200.
- Martinez-Rojas, M., del Carmen Pardo-Ferreira, M., & Rubio-Romero, J. C. (2018). Twitter as a tool for the management and analysis of emergency situations: A systematic literature review. *International Journal of Information Management*, 43, 196–208.
- Mishra, N., & Singh, A. (2018). Use of twitter data for waste minimisation in beef supply chain. *Annals of Operations Research*, 270(1), 337–359.
- Moorhead, S. A., Hazlett, D. E., Harrison, L., Carroll, J. K., Irwin, A., & Hoving, C. (2013). A new dimension of health care: Systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of Medical Internet Research*, 15(4), e1933.
- Narazaki, H., & Shigaki, I. (1999). A machine-learning approach for a sintering process using a neural network. *Production Planning and Control*, 10(8), 727–734.
- Odone, A., Gianfredi, V., Sorbello, S., Capraro, M., Frascella, B., Vigezzi, G. P., & Signorelli, C. (2021). The use of digital technologies to support vaccination programmes in Europe: State of the art and best practices from experts' interviews. *Vaccines*, 9, 1126.
- Palamenghi, L., Barello, S., Boccia, S., & Graffigna, G. (2020). Mistrust in biomedical research and vaccine hesitancy: The forefront challenge in the battle against Covid-19 in Italy. *European Journal of Epidemiology*, 35(8), 785–788.
- Piedrahita-Valdés, H., Piedrahita-Castillo, D., Bermejo-Higuera, J., Guillem-Saiz, P., Bermejo-Higuera, J. R., Guillem-Saiz, J., Sicilia-Montalvo, J. A., & Machío-Regidor, F. (2021). Vaccine hesitancy on social media: Sentiment analysis from June 2011 to April 2019. *Vaccines*, 9(1), 28.
- Pujawan, I. N., & Bah, A. U. (2021). Supply chains under COVID-19 disruptions: literature review and research agenda. In *Supply Chain Forum: An International Journal*, pp. 1–15. Taylor & Francis.
- Qayyum, A., Razzak, I., Tanveer, M., & Kumar, A. (2021). Depth-wise dense neural network for automatic COVID19 infection detection and diagnosis. *Annals of operations research*. <https://doi.org/10.1007/s10479-021-04154-5>
- Queiroz, M. M., Ivanov, D., Dolgui, A., & Wamba, S. F. (2020). 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*. <https://doi.org/10.1007/s10479-020-03685-7>
- Quyen, T., To, K. G., Huynh, V. -A. N., Nguyen, N. T. Q., Ngo, D. T. N., Alley, S. J., Tran, A. N. Q., Tran, A. N. P., Pham, N. T. T., Bui, T. X., et al. (2021). Applying machine learning to identify anti-vaccination tweets during the COVID-19 pandemic. *International Journal of Environmental Research and Public Health*, 18, 4069.
- Reinhardt, A., & Rossmann, C. (2021). Age-related framing effects: Why vaccination against COVID-19 should be promoted differently in younger and older adults. *Journal of Experimental Psychology: Applied*, 27(4), 669–678.
- Researcher, C. Q. (2020). *Issues for debate in american public policy: Selections from CQ researcher* (pp. 22). SAGE Publications.
- Rodríguez-González, A., Tuñas, J. M., Prieto Santamaría, L., FernándezPeces-Barba, D., Menasalvas Ruiz, E., Jaramillo, A., Cotarelo, M., ConejoFernández, A. J., Arce, A., & Gil, A. (2020). Identifying polarity

- in tweets from an imbalanced dataset about diseases and vaccines using a meta-model based on machine learning techniques. *Applied Sciences*, 10(24), 9019.
- Sear, R. F., Velásquez, N., Leahy, R., Restrepo, N. J., El Oud, S., Gabriel, N., Lupu, Y., & Johnson, N. F. (2020). Quantifying COVID-19 content in the online health opinion war using machine learning. *IEEE Access*, 8, 91886–91893.
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge University Press.
- Singh, J. P., Dwivedi, Y. K., Rana, N. P., Kumar, A., & Kapoor, K. K. (2019). Event classification and location prediction from tweets during disasters. *Annals of Operations Research*, 283(1), 737–757.
- Skinner, G. (2020). *Who's least likely to say they'll get a Covid-19 vaccine?* Ipsos Mori. <https://www.ipsos.com/ipsosmori/en-uk/whos-least-likely-say-theyll-get-covid-19-vaccine>.
- Tavoschi, L., Quattrone, F., D'Andrea, E., Ducange, P., Vabanesi, M., Marcelloni, F., & Lopalco, P. L. (2020). Twitter as a sentinel tool to monitor public opinion on vaccination: An opinion mining analysis from September 2016 to August 2017 in Italy. *Human Vaccines & Immunotherapeutics*, 16(5), 1062–1069.
- Tchuenté, D., & Nyawa, S. (2021). Real estate price estimation in French cities using geocoding and machine learning. *Annals of Operations Research*, 308(1), 571–608.
- Thelwall, M., Kousha, K., & Thelwall, S. (2021). Covid-19 vaccine hesitancy on English-language Twitter. *Profesional de la información*, 30(2), e300212.
- To, Q. G., To, K. G., Huynh, V. A., Nguyen, N. T., Ngo, D. T., Alley, S. J., Tran, A. N., Tran, A. N., Pham, N. T., Bui, T. X., & Vandelanotte, C. (2021). Applying machine learning to identify anti-vaccination tweets during the COVID-19 pandemic. *International Journal of Environmental Research and Public Health*, 18(8), 4069.
- Tomaszewski, T., Morales, A., Lourentzou, I., Caskey, R., Liu, B., Schwartz, A., & Chin, J. (2021). Identifying false human papillomavirus (HPV) vaccine information and corresponding risk perceptions from twitter: Advanced predictive models. *Journal of Medical Internet Research*, 23(9), e30451.
- Verelst, F., Kessels, R., Delva, W., Beutels, P., & Willem, L. (2019). Drivers of vaccine decision-making in South Africa: A discrete choice experiment. *Vaccine*, 37(15), 2079–2089.
- Wagner, A. L., Huang, Z., Ren, J., Laffoon, M., Ji, M., Pinckney, L. C., Sun, X., Prosser, L. A., Boulton, M. L., & Zikmund-Fisher, B. J. (2020). Vaccine hesitancy and concerns about vaccine safety and effectiveness in Shanghai, China. *American Journal of Preventive Medicine*, 60(1), S77–S86.
- Wamba, S. F., Edwards, A., & Akter, S. (2019). Social media adoption and use for improved emergency services operations: The case of the NSW SES. *Annals of Operations Research*, 283(1), 225–245.
- Wang, Z., Yin, Z., & Argyris, Y. A. (2020). Detecting medical misinformation on social media using multimodal deep learning. *IEEE Journal of Biomedical and Health Informatics*, 25(6), 2193–2203.
- Xiong, Y., Lam, H. K. S., Kumar, A., Ngai, E. W. T., Xiu, C., & Wang, X. (2021). The mitigating role of blockchain-enabled supply chains during the COVID-19 pandemic. *International Journal of Operations & Production Management*, 41(9), 1495–1521.
- Yuan, X., Schuchard, R. J., & Crooks, A. T. (2019). Examining emergent communities and social bots within the polarized online vaccination debate in Twitter. *Social Media + Society*, 5(3), 205630511986546.
- Zhang, S., Zhang, C., & Yang, Q. (2003). Data preparation for data mining. *Applied Artificial Intelligence*, 17, 375–381.
- Zhang, L., Fan, H., Peng, C., Rao, G., & Cong, Q. (2020). Sentiment analysis methods for HPV vaccines related tweets based on transfer learning. *Healthcare*, 8, 307.
- Zhou, X., Coiera, E., Tsafnat, G., Arachi, D., Ong, M.-S., & Dunn, A. G. (2015). *Using social connection information to improve opinion mining: Identifying negative sentiment about HPV vaccines on Twitter*.