



Deep learning based sentiment analysis of public perception of working from home through tweets

Aarushi Vohra¹ · Ritu Garg¹

Received: 5 June 2022 / Revised: 1 August 2022 / Accepted: 8 August 2022

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

Abstract

Nowadays, we are witnessing a paradigm shift from the conventional approach of working from office spaces to the emerging culture of working virtually from home. Even during the COVID-19 pandemic, many organisations were forced to allow employees to work from their homes, which led to worldwide discussions of this trend on Twitter. The analysis of this data has immense potential to change the way we work but extracting useful information from this valuable data is a challenge. Hence in this study, the microblogging website Twitter is used to gather more than 450,000 English language tweets from 22nd January 2022 to 12th March 2022, consisting of keywords related to working from home. A state-of-the-art pre-processing technique is used to convert all emojis into text, remove duplicate tweets, retweets, username tags, URLs, hashtags etc. and then the text is converted to lowercase. Thus, the number of tweets is reduced to 358,823. In this paper, we propose a fine-tuned Convolutional Neural Network (CNN) model to analyse Twitter data. The input to our deep learning model is an annotated set of tweets that are effectively labelled into three sentiment classes, viz. positive negative and neutral using VADER (Valence Aware Dictionary for sEntiment Reasoning). We also use a variation in the input vector to the embedding layer, by using FastText embeddings with our model to train supervised word representations for our text corpus of more than 450,000 tweets. The proposed model uses multiple convolution and max pooling layers, dropout operation, and dense layers with ReLU and sigmoid activations to achieve remarkable results on our dataset. Further, the performance of our model is compared with some standard classifiers like Support Vector Machine (SVM), Naive Bayes, Decision Tree, and Random Forest. From the results, it is observed that on the given dataset, the proposed CNN with FastText word embeddings outperforms other classifiers with an accuracy of 0.925969. As a result of this classification, 54.41% of the tweets are found to show affirmation, 24.50% show a negative disposition, and 21.09% have neutral sentiments towards working from home.

Aarushi Vohra and Ritu Garg are contributed equally to this work.

✉ Aarushi Vohra
aarushivohra01@gmail.com

Ritu Garg
ritu.59@nitkr.ac.in

¹ Department of Computer Engineering, National Institute of Technology Kurukshetra, 136119 Kurukshetra, Haryana, India

Keywords Big data · Social media · Opinion mining · Deep learning · VADER

1 Introduction

An argument that is gaining popularity nowadays revolves around how employees will perceive their work environment in the times to come. Whether they will prefer to work on their office desks in a group setting or work virtually from the comforts of their home. Since the inception of the COVID-19 pandemic, the general outlook of people on working from home has changed. Work from home is the process of carrying out office-established work from the premises of home, using internet services (Islam, 2022; Tønnessen et al., 2021). The COVID-19 pandemic has forced organisations to adopt flexible professional engagements in the form of virtual working environments. There has been a remarkable growth in the use of digital technology for telework (Tønnessen et al., 2021). Figure 1 shows a graph obtained from Google Trends demonstrating a relative search interest in this regard.

Several studies have shown that the paradigm shift to the culture of working from home has positive impacts on the employees as it improves productivity, reduces stress, and leads to job satisfaction. A flexible schedule allows for more time with family. It has considerably reduced the time required to travel to the office (Kawakubo and Arata, 2022). Hence, it has led to a reduction in transport-related energy consumption, and therefore, limiting air pollution and carbon emissions (Jain et al., 2022). Work from home is also in the interest of organizations as the virtual operation has the potential to boost creativity and innovation among employees (Tønnessen et al., 2021). However, some researchers suggest that long hours of working from home are highly demanding which may reduce productivity and

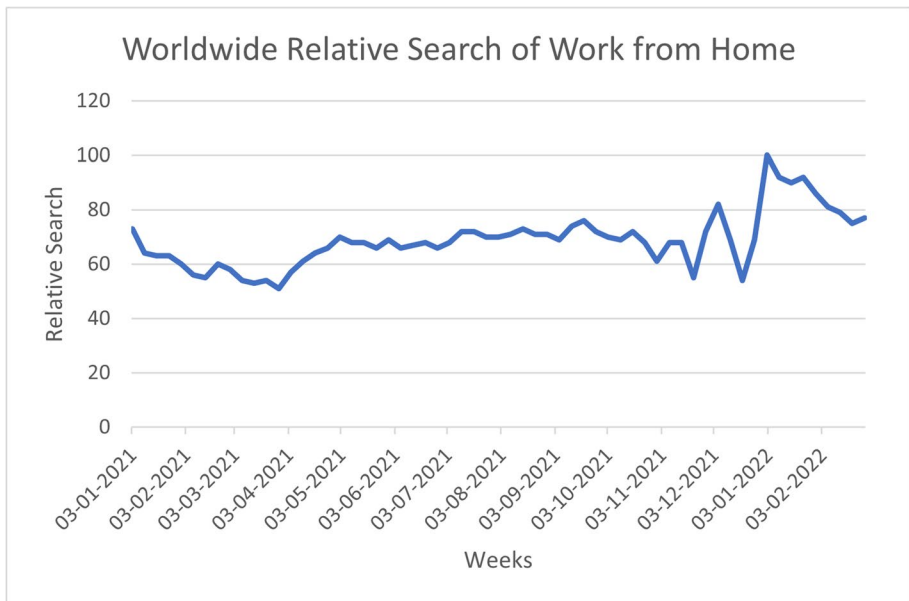


Fig. 1 Relative search interest of ‘Work from Home’ from January 2021 to February 2022 (Source: <https://trends.google.com/trends>)

disturb work-life balance. Limited in-person conversations and increased screen time may trigger anxiety. Moreover, interruptions from family members during working hours and lack of resources at home also create hindrances to work efficiently (Islam, 2022; Tønnesen et al., 2021; Kawakubo and Arata, 2022; Prodanova and Kocarev, 2021). These factors play a significant role in framing policies for employees. The success of an organisation largely depends on ability of its employees to perform efficiently. Thus, the orientation of employees towards working from home or office is crucial for deciding the future of work.

Due to the rise of social media users and the ubiquitous influence of social media, platforms such as Twitter, Facebook, YouTube, LinkedIn, Instagram have emerged as the principal contributors of big data. The escalation of social media data puts forward a wide range of opportunities in Natural Language Processing (NLP). When data from social media is analysed with different representation and modelling approaches, it gives a diverse idea of people's perception of social disciplines. This intense data helps researchers to elucidate people's opinions and develop novel prediction techniques for a variety of domains like decision making, stock market, and recommendation systems etc (Rani et al., 2022; Zachlod et al., 2022; Salim et al., 2022; Pathak et al., 2021).

But extracting valuable information from a huge volume of raw data and comprehending worthwhile insights is a challenge. Moreover, data storage is another problem to deal with Rani et al. (2022). Despite these concerns, sentiment analysis is a highly prevalent research area in NLP. The textual data from social media platforms can be analysed by lexicon based and machine learning based sentiment analysis methods (Li et al., 2020). Such techniques involve the use of algorithms like SVM, Naïve Bayes, Decision Tree, Random Forest etc. The machine learning approach can also be based on powerful Deep Neural Networks. They are found to give better results than conventional models because of their ability to detect features from a large amount of data (Joseph et al., 2022). Such a technique requires high-dimensional labelled data to train the model, but the training process may be a memory and time-consuming task. The quality of these factors may lead to a rise or drop in the overall performance of the model (Basiri et al., 2021). In the text-processing domain, CNNs are popularly used with word embeddings for classification and clustering. It is because of their ability to extract features from data with the convolution operation and measure the relationships in local patterns (Li et al., 2020; Basiri et al., 2021; Liao et al., 2017).

The analysis of Twitter data has become a powerful tool to capture dynamic information related to public perception, which can serve as useful information for decision makers (De Rosis et al., 2021). Some existing studies (De Rosis et al., 2021; Rakshitha et al., 2021; Yousefinaghani et al., 2021) on Twitter sentiment analysis are based on a lexicon-based approach. Such a technique, when combined with a statistical/machine learning approach, can lead to a more effective polarity classification (Cambria et al., 2017). Researchers in García-Ordás et al. (2021); Sasidhar et al. (2020) utilise a combination of both these techniques on a low sample size. Given these premises, we aim to address the emerging trend of working from home, by analysing public opinion and notions based on a rich dataset, using a combination of lexicon and machine learning based approaches.

The data set for the same is acquired for a period of 50 days from 22nd January 2022 to 12th March 2022 from Twitter, which is a popular medium of expression for various public interests. It is mostly used by intellectually aware people to discuss opinions in real-time (Ding et al., 2021; Liu and Liu, 2021). Hence, data-mining techniques are employed to extract a rich text corpus of around 450,000 English language tweets from Twitter, containing keywords like 'wfh', 'work from home' and 'working from home'. The tweets are pre-processed and are the labelled using VADER (Hutto and Gilbert, 2014) as three sentiment

classes, namely, positive, negative, and neutral. This textual data is then converted into word vectors using FastText word embeddings, that serves as an input to train the proposed CNN. The comparison with standard classifiers like SVM, Naïve Bayes, Decision Tree, and Random Forest validates the performance accuracy of our model.

The principal contributions of this research are as follows -

1. Collection of a rich text corpus containing around 450,000 English language tweets from Twitter.
2. Performing a series of state-of-the-art pre-processing tasks such as conversion into lowercase, removal of duplicates, handling emojis, usernames, URLs, and hashtags etc.
3. Proposing a fine-tuned CNN model for an effective sentiment analysis on the above-mentioned dataset.
4. A performance comparison of the CNN model with various machine learning classifiers, such as SVM, Naïve Bayes, Decision Tree, Random Forest to evaluate the best performing classifier on our dataset.
5. The findings of this study may help organisations and researchers in the ideation of a novel system to carry out office work.

The remainder of this study includes the following. A literature review in Section 2. In Section 3, we have discussed the methodology. Section 4 consists of our proposed CNN model. Section 5 is based on model predictions, analysis, and comparison with other machine learning classifiers. The paper is concluded in Section 6, with the final discussions and future scope.

2 Literature review

The extensive influx of social media data has been widely used for analysis. Previous studies have shown that this dynamic source of big data holds immense significance in a wide range of domains such as business, marketing, recommendation, politics, research, medicine and healthcare, opinion analysis, intelligence etc (Luo and Mu, 2022; Ansari et al., 202; Nezhad and Deihimi, 2022; Alamoodi et al., 2021; Rajalakshmi et al., 2017). A comprehensive sentiment analysis process includes the creation or extraction of textual data, storage of the text corpus into files, pre-processing the data, feature engineering and selection and finally, applying a sentiment analysis approach (Rajalakshmi et al., 2017).

Public opinions from around the world are discussed on social media platforms, like Twitter. A careful analysis of these facts can lend a constructive understanding for the development of novel strategies. Authors of Rachunok et al. (2022) emphasise the discussion of Twitter data on a large scale. They have extracted keyword and location-based tweets and analysed their metrics. A lexicon-based approach to know the polarity of data is much in trend. In paper (Rakshitha et al., 2021), TextBlob is used to analyse the polarity of customer reviews in five Indian regional languages. Researchers in Yousefinaghani et al. (2021) have used VADER to assign polarity to tweets relating to vaccine sentiments. The study (Ding et al., 2021) gives an insight into people's reaction during the initial weeks of the COVID-19 pandemic. In paper (Liu and Liu, 2021), a sentiment analysis is conducted on COVID-19 vaccines. These studies mainly employ a lexicon-based method for

sentiment analysis. Due to the effectiveness of classifying texts by using a lexicon-based approach, we use VADER to assign labels to our dataset.

Twitter can also be used for analysis in other domains other. Researchers in Neogi et al. (2021) focus on analysing tweets on the farmers' protest in India using Bag of Words and TF-IDF vectorizer along with some standard classifiers for prediction. Authors of Hidayat et al. (2022) propose the use of SVM and logistic regression to classify a dataset of a few thousand tweets. Paper (Ding et al., 2021) explains public interest in autonomous vehicles using data obtained from Twitter feeds. Researchers in García-Ordás et al. (2021) have proposed a novel neural network to classify variable-length audio in real time. Word2Vec CBOW (Continuous Bag of Words) model is used in study (Sasidhar et al., 2020), to detect emotions in Hindi-English code mix tweets. Authors of Fitri et al. (2019) have used Naïve Bayes, Decision Tree, and Random Forest algorithms for sentiment analysis of Anti-LGBT campaign on Twitter. These studies use a low sample size of data for modelling. Therefore, in this study we attempt to analyse more than 350,000 unique English tweets obtained from Twitter.

The survey (Liu and Liu, 2021) presents that supervised learning approaches are widely used for sentiment analysis. SVM and Naïve Bayes classifiers have been used in paper (Pavitha et al., 2022) for sentiment analysis. Deep neural networks are also used for sentiment analysis tasks (Joseph et al., 2022). Authors of Rani et al. (2022) have presented a CNN-LSTM Model to classify tweets into six sentiment classes. Researchers of Fiok et al. (2021) have performed a sentiment analysis on a five-level sentiment scale, based on tweets posted to a specific Twitter account. Paper (Chen et al., 2017) proposes a BiLSTM-CRF based approach to improve sentence type classification. The study (Ridhwan and Hargreaves, 2021) has used VADER with a deep learning method (RNN) to classify tweets related to COVID-19 in Singapore. Authors of Umair and Masciari (2022) have used TextBlob and BERT model to identify sentiments related to COVID-19 vaccines. With deep learning models, pre-trained word embeddings can also be used. In paper (Sharma et al., 2020), pre-trained Word2Vec embeddings have resulted in a better classification of small movie review sentences. Pre-trained GloVe word embeddings are used as initial weights in research (Basiri et al., 2021). The use of FastText word representations in study (Khasanah, 2021) has slightly improved the performance of the model. The research (Deb and Chanda, 2022) shows that contextual word embeddings have a better accuracy than context-free word embeddings. Advanced deep neural networks may achieve a higher accuracy with word embeddings. Other artificial neural networks, like bidirectional emotional recurrent units (Li et al., 2022) and improved graph convolutional networks (He et al., 2022), are also used for aspect-based sentiment analysis (Imani and Noforesti, 2022; Zhao et al., 2022). A sentiment analysis at finer classification levels to handle ambivalent emotions is proposed in paper (Wang et al., 2020). Hence, deep learning techniques have led to remarkable advancements in the field of sentiment analysis (Cambria et al., 2022). Therefore, in this research, we employ word embeddings with a deep neural network, and conduct different experiments to analyse the effect of use of word embeddings as initial weights to the model on our dataset.

This work revolves around analysing public opinions related to the present-day concept of working from home. The analysis is based on more than 450,000 English tweets obtained from Twitter. The authors of paper (Cambria et al., 2017) have proposed the use of knowledge-based techniques combined with machine learning approaches for polarity detection. Hence, we employ a lexicon-based approach as well as various machine learning models to get a thorough understanding of this trend. The entire discussion unfolds in the succeeding sections.

3 Methodology

For this study, we have acquired data (tweets) from Twitter. We have applied data pre-processing techniques to deal with emojis, usernames, hashtags, URLs, and remove duplicates. Then, the data is labelled, up sampled, and split into training, validation, and testing datasets. This data is fed to our model (a Convolutional Neural Network) for training and prediction. In this section, we present a discussion of the above steps in detail.

A diagrammatic representation of the proposed scheme is shown in Fig. 2.

3.1 Data acquisition

Python's Tweepy library is used to gather a rich corpus of tweets related to public opinion on working from home. Around 450,000 English language tweets are extracted from Twitter from 22nd January 2022 to 12th March 2022, by querying for a variety of related keywords such as 'wfh', 'work from home' and 'working from home'. The tweets thus collected consist of attributes like a Tweet ID (a unique identifier for the tweet), text (the textual contents of the tweet as posted by the user), and the date and time when the tweet was posted. The data obtained is stored in a CSV format. An exploratory analysis shows the presence of tweets of different polarities in the dataset.

3.2 Data pre-processing

A state-of-the-art pre-processing technique is applied on the raw data obtained from Twitter to get rid of any inconsistency or noise before the data is fed to the model. Pre-processing reduces the dimensionality of input data (Rajalakshmi et al., 2017). This, in turn, helps to achieve better model performance in sentiment analysis. To pre-process the extracted text corpus, we first convert the emojis present in the text into their CLDR (Common Locale Data Repository Project) short names using Python's Emoji module. Then the entire text data is converted to lowercase. Afterwards, all retweets, username tags, URLs and hashtag symbols are removed. The keyword 'wfh' is converted to work from home. Finally, we remove all the duplicate entries from the text. The above-mentioned pre-processing steps are carried out by using Python scripts. Hence, the dataset is reduced to a total of 358,823 unique tweets after pre-processing. Some sample examples of raw tweets obtained from Twitter and their pre-processed counterparts are shown in Table 1. Figure 3 represents a graph for the number of unique tweets corresponding to different days during the tweets' extraction period.

3.3 Data annotation

Due to the enormous volume of our dataset, it is infeasible to manually label the data. Hence, we have applied a popular lexicon-based approach for data annotation. We have used VADER to understand the semantic orientation of the tweets. VADER is attuned

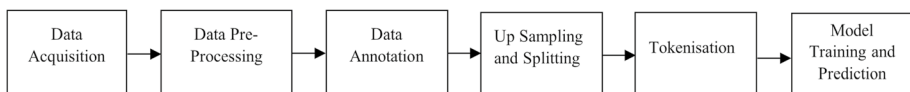


Fig. 2 Methodology

Table 1 Raw tweets and their pre-processed counterparts

Raw tweet	Tweet after pre-processing
I have never been so grateful that my company allows us to WFH, though I have a morning meeting at least I don't have to worry about morning commute https://t.co/RKICXx5s8K	i have never been so grateful that my company allows us to work from home though i have a morning meeting at least i don t have to worry about morning commute
A leave is not #WFH and WFH is not a LEAVE!!!! https://t.co/cmrAOUHxHA	a leave is not work from home and work from home is not a leave
Twitter is a weird space, it's very understandable that some people love WFH while some hate it	twitter is a weird space its very understandable that some people love work from home while some hate it
Ugh I'm just trying to work from home	ugh i m just trying to work from home

Daily Number of Unique Tweets from 22nd January 2022 to 12th March 2022

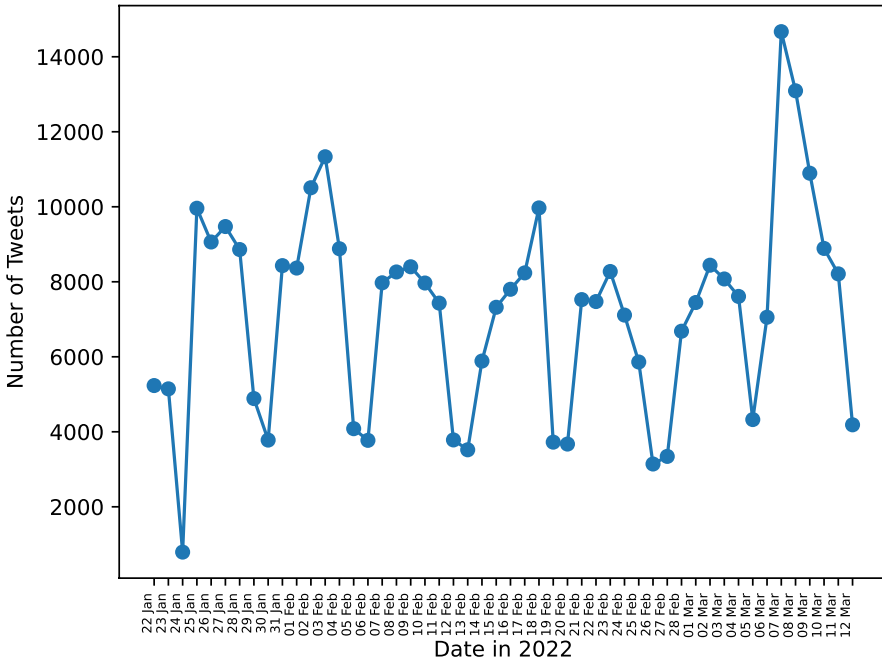


Fig. 3 Daily number of unique tweets

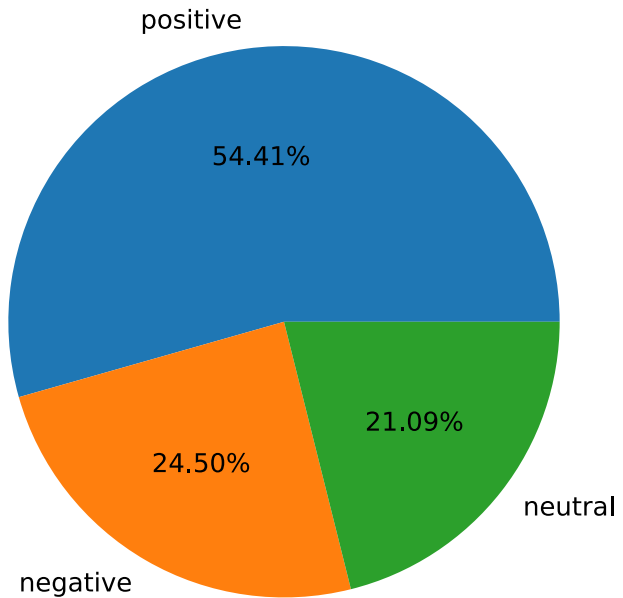
to work well on data consisting of abbreviations, short unconventional texts and slangs, catering to the needs of social media data. It is computationally fast and may exhibit higher accuracy than human annotators. We use this tool to categorise our data into three classes - positive, negative, and neutral, based on sentiment and polarity scores (Hutto and Gilbert, 2014). Examples of tweets labelled as positive, negative, and neutral are shown in Table 2. The distribution of tweets in these classes is shown in Fig. 4. A total of 1,95,233 tweets are found to be positive, 87,898 negative and 75,692 neutral. This gives us an overview of a public perception of working from home. A graph for polarity of unique tweets for each day during the data acquisition period is shown in Fig. 5.

Table 2 Tweets and their polarities as classified by VADER

Tweet	Polarity
i have never been so grateful that my company allows us to work from home though i have a morning meeting at least i don t have to worry about morning commute	Positive
twitter is a weird space its very understandable that some people love work from home while some hate it	Negative
i work from home and surprisingly it's more exhausting lol	Positive
a leave is not work from home and work from home is not a leave	Neutral
i need to buy a proper office chair for my house because of work from home long hours of sitting is killing my back	Negative

Fig. 4 Distribution of tweets into classes

Class Distribution of Sentiments



3.4 Data up sampling and splitting

From Fig. 4 it is evident that the data is highly unbalanced in the three classes - positive tweets dominating the data with 1,95,233 tweets and the number of negative and neutral tweets being 87,898 and 75,692 respectively. Hence, we have resampled the minority classes (negative and neutral) and injected it back to the original dataset, so that the model doesn't incline towards the majority class (positive). Then, we split the dataset into three parts - training (80%), validation (10%) and testing (10%) data. Data up sampling and splitting is done using Python scikit-learn. A precise outline of the number and class of tweets in the training, validation and testing dataset is shown in Table 3.

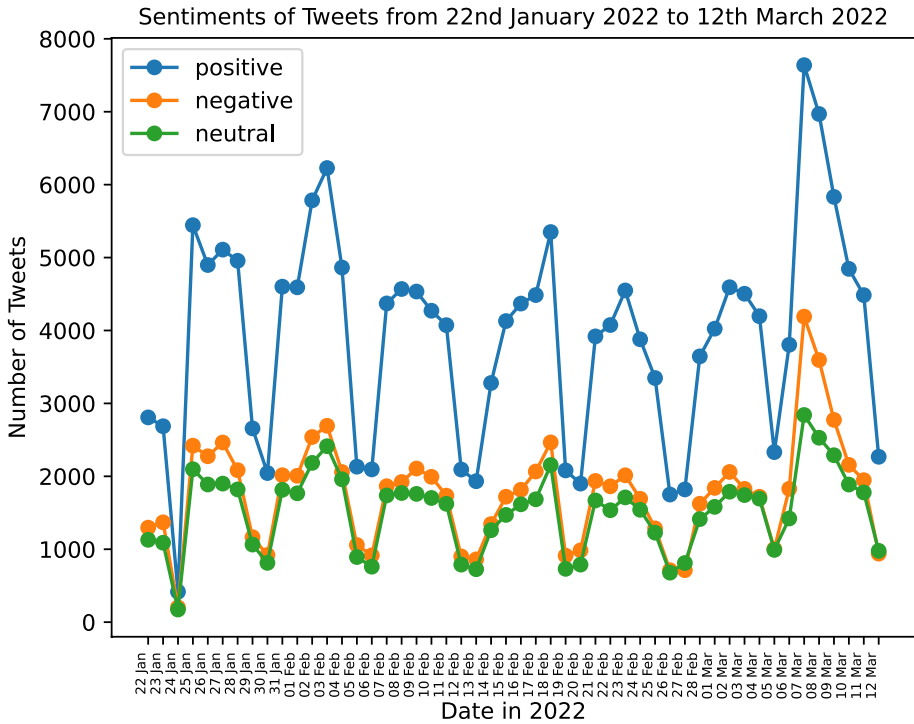


Fig. 5 Polarity of Unique Tweets for Each Day

3.5 Tokenisation, padding and word embeddings

To provide textual data (tweets) to our model in the form of training data, we have converted the textual data into vector form, known as word embeddings. This conversion is carried out using Keras Tokenizer in Python. The maximum sentence length of the text in training, validation and testing dataset is found to be 189, 255, and 188 respectively. We have limited the maximum sentence length to 190. Hence, word embeddings of sentences with length less than 190 are padded with zeros. Thus, the input shape of embeddings to our model is fixed to 190. Moreover, to train another instance of the model, FastText word embeddings are used as supervised word representations. These word embeddings are based on sub-word information. Hence, FastText can also handle out-of-vocabulary words. Further, the non-numerical labels (i.e., positive, negative, and neutral) are converted into categorical values (one-hot encoding) using Python scikit-learn Label Encoder.

Table 3 Number and polarities of tweets in training, validation, and testing datasets

	Positive	Negative	Neutral	Total
Training	156,071	156,137	156,351	468,559
Validation	19,527	19,511	19,532	58,570
Testing	19,635	19,585	19,350	58,570

4 The Model - Convolutional Neural Network

To perform sentiment analysis on our dataset, we have used a CNN. This section describes the model used in this study.

4.1 Model parameters

The proposed model consists of an input layer, an embedding layer, two sets of conv1d and max pooling layers, a dropout layer, and two dense layers.

Embedding layer: The word embeddings generated from the tweets' text are an input to this layer. It produces an output of shape (190, 300), with the embedding dimensions being 300. The maximum number of features to be considered are set to 30,000. In this research, we propose the use of pre-trained FastText word embeddings as weights to this layer.

Convolution layer: The proposed CNN model consists of two convolution layers. The first conv1d layer has 64 filters each of size 3, with ReLU activation and L2 regularizers. The input is down sampled by using a max pooling layer with the pool size and strides both being set as 2. Another conv1D layer, like the former, is also present.

Global max pooling: The convolution layers are followed by a global max-pooling operation. This gives the maximum value of all values across the entire input.

Dropout layer: Then, a dropout layer is used to prevent overfitting. The dropout rate is set to 0.5 which drops half of the input units at each step.

Dense layer: The model also has two fully-connected dense layers (in which neurons of the layer are deeply connected with the preceding layer). The first dense layer has 32 units (output space dimensions), ReLU activation, and regularizers. The other dense layer has 3 units (equal to the number of classes), SoftMax activation, and regularizers for kernel and bias.

The model is compiled with Categorical Cross entropy loss function for multi-label classification, as the labels are one-hot encoded. We have also used Adam optimizer with learning rate set to 0.001. The metrics are set to Categorical Accuracy. It calculates the percentage of correctly predicted labels for one-hot encodings. The entire model is implemented in Python using Tensorflow and Keras. All the above-mentioned parameters are tuned after careful experimentation. The architecture of the model is shown in Fig. 6.

4.2 Training

The model is trained on labelled data (supervised learning). The batch size is set to 128. Initially, the model is expected to train on 50 epochs, but an early stopping callback function is also applied to prevent overfitting of the model. The callback function monitors the validation loss and stops training the model when the metric ceases to decrease. Plots of categorical accuracy and loss for both training and validation datasets are shown in Figs. 7 and 8 respectively. Another instance of the same model is also trained using FastText pre-trained word embeddings as initial weights. Plots of categorical accuracy and loss for FastText word embeddings-based CNN model are shown in Figs. 7 and 8 respectively. It is found that the model that used pre-trained FastText word embeddings trained for 23 epochs, whereas the CNN model without FastText word embeddings model trained for 22 epochs, afterwards, the training is stopped by the early stopping function.

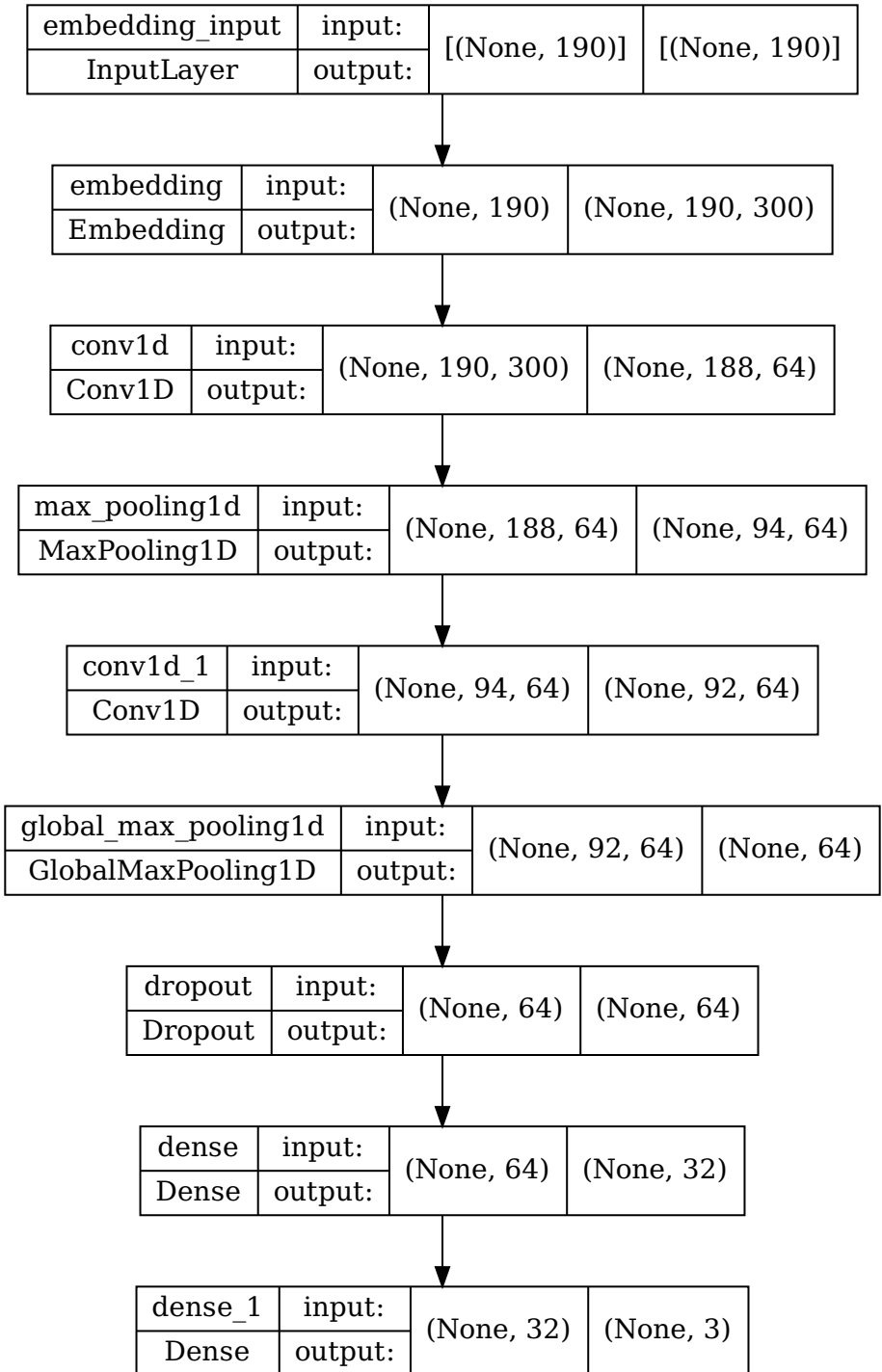


Fig. 6 Architecture of the proposed CNN model

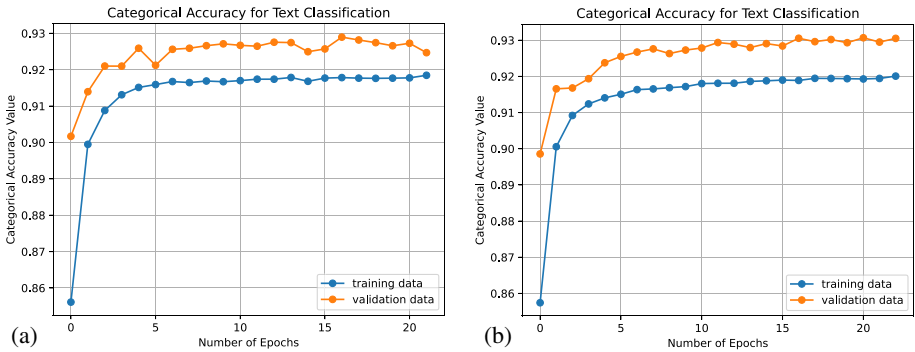


Fig. 7 Categorical Accuracy values for CNN (a) without FastText word embeddings (b) with FastText word embeddings

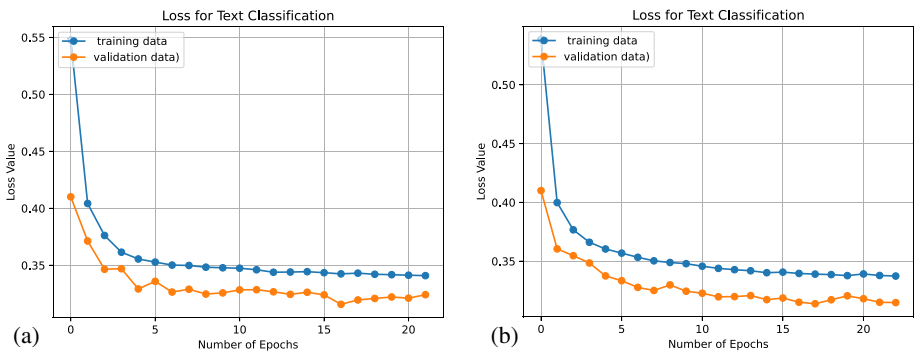


Fig. 8 Loss values for CNN (a) without FastText word embeddings (b) with FastText word embeddings

5 Evaluation and results

The model described in the previous section is evaluated on a testing dataset consisting of 58,570 tweets, of which 19,635 are positive, 19,585 are negative and 19,350 are neutral tweets. The performance of the proposed CNN model is also compared with existing machine learning classifiers like SVM (Pavitha et al., 2022), Naïve Bayes (Yang, 2018), Decision Tree (Swain and Hauska, 1977) and Random Forest (Biau and Scornet, 2016). Some standard metrics (Sasidhar et al., 2020) considered for this analysis are as follows -

1. Accuracy - It represents the proportion of the correctly predicted samples to the total number of samples.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Here,

- TP: the number of positive instances that the classifier correctly labels as positive.
- FP: the number of negative instances that the classifier incorrectly labels as positive.

- FN: the number of positive instances that the classifier incorrectly labels as negative.
 - TN: the number of negative instances that the classifier correctly labels as negative.
2. Precision - It represents the ratio between the correctly predicted positive samples to the total number of predicted positive samples.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

3. Recall - It represents the ratio between correctly predicted positive samples to the number of all positive samples.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

4. F1-score - It is a value between 0 and 1 which is calculated as the harmonic mean of precision and recall.

$$F1score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

5. Support - It is the actual number of samples of a class in the dataset.

5.1 Convolutional Neural Network model without FastText word embeddings

In this experiment, word embeddings generated using Keras tokenizer are used with the CNN model. The results are predicted on the testing dataset. The classification report generated using Python scikit-learn, based on evaluation metrics represented by (1), (2), (3), (4) is shown in Table 4. As a result, it is found that 17,819 positive, 18,557 negative and 17,778 neutral tweets are correctly classified into their respective classes. Figure 9 shows the confusion matrix for the classification of tweets into three classes. The accuracy of the CNN without the use of FastText word embeddings is 92.4603%.

5.2 Convolutional Neural Network model with FastText word embeddings

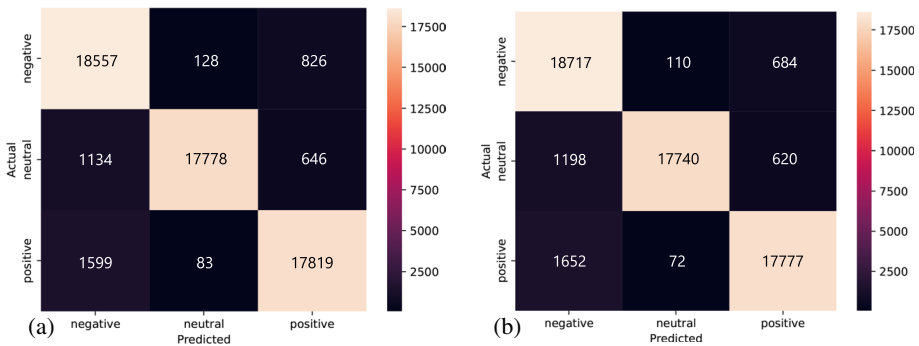
In this experiment, pre-trained FastText word embeddings are used as initial weights to the CNN model. After predicting labels on the testing dataset, the classification report generated using Python scikit-learn, based on evaluation metrics represented by (1), (2), (3), (4) is shown in Table 5. It is found that 17,777 positive, 18,717 negative and 17,740 neutral tweets are correctly classified into their respective classes. Figure 9 shows the confusion

Table 4 Classification Report of CNN model without FastText word embeddings

	Precision	Recall	F1-score	Support
Positive	0.923695	0.913748	0.918695	19501
Neutral	0.988271	0.908989	0.946973	19558
Negative	0.87163	0.951105	0.909635	19511
Accuracy				0.924603

Table 5 Classification Report of CNN model with FastText word embeddings as initial weights

	Precision	Recall	F1-score	Support
Positive	0.93166	0.911594	0.921518	19501
Neutral	0.989845	0.907046	0.946638	19558
Negative	0.867854	0.959305	0.911291	19511
Accuracy				0.925969

**Fig. 9** Confusion Matrix for CNN (a) without FastText word embeddings (b) with FastText word embeddings

matrix for the classification of tweets into three classes. The accuracy of the CNN with the use of FastText word embeddings is 92.5969%.

In the following series of experiments, other machine learning classifiers are also used for sentiment prediction on the same dataset.

5.3 SVM model

The performance accuracy of SVM on our data set is 0.864396. Table 6 shows the classification report for the SVM model generated using Python scikit-learn, based on evaluation metrics represented by (1), (2), (3), (4).

5.4 Naïve Bayes model

The performance accuracy of Naïve Bayes on our data set is 0.772759. Table 7 shows the classification report for the Naïve Bayes model generated using Python scikit-learn.

5.5 Decision Tree model

The performance accuracy of Decision Tree on our data set is 0.794007. Table 8 shows the classification report for the Naïve Bayes model generated using Python scikit-learn.

Table 6 Classification report of SVM model

	Precision	Recall	F1-score	Support
Positive	0.868624	0.834709	0.851344	19501
Neutral	0.882702	0.888881	0.885781	19558
Negative	0.842627	0.869647	0.855933	19511
Accuracy				0.864396

Table 7 Classification report of naïve bayes model

	Precision	Recall	F1-score	Support
Positive	0.83988	0.692507	0.759106	19501
Neutral	0.890919	0.734254	0.805035	19558
Negative	0.659951	0.89131	0.758378	19511
Accuracy				0.772759

Table 8 Classification report of decision tree model

	Precision	Recall	F1-score	Support
Positive	0.805072	0.579427	0.673862	19501
Neutral	0.81433	0.920235	0.86405	19558
Negative	0.767191	0.882538	0.820832	19511
Accuracy				0.794007

Table 9 Classification report of random forest model

	Precision	Recall	F1-score	Support
Positive	0.806344	0.847341	0.826334	19501
Neutral	0.874124	0.931088	0.901707	19558
Negative	0.952134	0.841926	0.893645	19511
Accuracy				0.873399

5.6 Random Forest model

The performance accuracy of Random Forest on our data set is 0.873399. Table 9 shows the classification report for the Naïve Bayes model generated using Python scikit-learn.

Further, Table 10 shows a comparative analysis of the performance of all the classifiers used for sentiment analysis on the dataset. It is found that the proposed CNN with FastText word embeddings outperforms other classifiers with a classification accuracy of 0.925969.

Thus, CNNs are found to have the best performance on the given dataset. Moreover, it is observed that the accuracy of CNN increases slightly with the use of pre-trained FastText word vectors. This is because of FastText's ability to hold sub-word information, which allows FastText to generate out of vocabulary words. Table 11 illustrates a few examples where the predicted value differed from the actual label. A word cloud representing some of the most commonly occurring words in our text corpus is shown in Fig. 10.

6 Conclusion

The huge volume of data from social media has immense potential for exploration. By performing a careful analysis, such a rich dataset of big data can be used for predictions for improvement in the field of decision making.

In this research, we have conducted a sentiment analysis on data obtained from Twitter. Opinion mining techniques are employed to gather a rich text corpus of more than 450,000 English tweets containing keywords related to work from home, over a period of 50 days. A series of state-of-the-art pre-processing techniques are carried out to handle emojis, usernames, URLs, hashtags, abbreviations, and inconsistencies in text. Polarities are assigned to the tweets using VADER. Further, we use a novel CNN to examine the tweets to infer public perception of working from home. The proposed deep learning model has multiple convolution and max pooling layers, dropout operation, and dense layers with ReLU and sigmoid activations. The use of FastText supervised word representations with our model has shown a promising performance on our dataset. Further, some standard machine learning classifiers - SVM, Naïve Bayes, Decision Tree, and Random Forest are used to validate the performance of the proposed CNN. As a result, it is shown that our CNN model with FastText word embeddings remarkably outperforms other classifiers, with a classification accuracy of 0.925969. Thus, we have addressed the emerging inclination of working from home on Twitter, using lexicon-based techniques and several machine learning classifiers, and found that 54.41% of tweets show affirmation for working from home, whereas 24.50% tweets show public dissatisfaction. However, 21.09% tweets have a neutral disposition on this present-day working trend.

Results from this study can be used to frame new flexible policies to give employees the freedom to choose their work settings. Moreover, a hybrid approach will considerably save the time and resources required to travel to the office. A limitation of this research includes the consideration of only English language tweets. In future, the proposed model could be used in conjunction with other classification models to classify the tweets into much finer classes to handle mixed emotions, and to improve accuracy and time-related metrics.

Acknowledgements Not applicable.

Author contributions All authors had the same contribution. All authors read and approved the final manuscript.

Funding No funding was received for conducting this study.

Data availability The data that support the findings of this study are available from Twitter. Data are however available from the authors upon reasonable request.

Declarations

Conflicts of interest The authors declare that they have no competing interests.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

- Alamoodi, A. H., et al. (2021). Multi-perspectives systematic review on the applications of sentiment analysis for vaccine hesitancy. *Computers in Biology and Medicine*, 139, 104957. <https://doi.org/10.1016/j.combiomed.2021.104957>
- Ansari, M. Z., et al. (2020). Analysis of political sentiment orientations on twitter. *Procedia Computer Science*, 167, 1821–1828. <https://doi.org/10.1016/j.procs.2020.03.201>
- Basiri, M. E., et al. (2021). Abcdm: An attention-based bidirectional cnnrnn deep model for sentiment analysis. *Future Generation Computer Systems*, 115, 279–294. <https://doi.org/10.1016/j.future.2020.08.005>
- Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Cambria, E., & et al. (2022). Senticnet 7: a commonsense-based neurosymbolic ai framework for explainable sentiment analysis. In Proceedings of LREC 2022 (pp. 3829–3839). <https://sentic.net/senticnet-7.pdf>
- Cambria, E., & et al. (2017). Affective computing and sentiment analysis. A practical guide to sentiment analysis (pp. 1–10). Springer. <https://doi.org/10.1007/978-3-319-55394-81>
- Chen, T., et al. (2017). Improving sentiment analysis via sentence type classification using bilstm-crf and cnn. *Expert Systems with Applications*, 72, 221–230. <https://doi.org/10.1016/j.eswa.2016.10.065>
- Deb, S., & Chanda, A. K. (2022). Comparative analysis of contextual and context-free embeddings in disaster prediction from twitter data. *Machine Learning with Applications*, 7, 100253. <https://doi.org/10.1016/j.mlwa.2022.100253>
- De Rosi, S., et al. (2021). The early weeks of the italian covid-19 outbreak: sentiment insights from a twitter analysis. *Health Policy*, 125(8), 987–994. <https://doi.org/10.1016/j.healthpol.2021.06.006>
- Ding, Y., et al. (2021). How are sentiments on autonomous vehicles influenced? an analysis using twitter feeds. *Transportation Research Part C: Emerging Technologies*, 131, 103356. <https://doi.org/10.1016/j.trc.2021.103356>
- Fiok, K., et al. (2021). Analysis of sentiment in tweets addressed to a single domain-specific twitter account: Comparison of model performance and explainability of predictions. *Expert Systems with Applications*, 186, 115771. <https://doi.org/10.1016/j.eswa.2021.115771>
- Fitri, V. A., et al. (2019). Sentiment analysis of social media twitter with case of anti-lgbt campaign in indonesia using naïve bayes, decision tree, and random forest algorithm. *Procedia Computer Science*, 161, 765–772. <https://doi.org/10.1016/j.procs.2019.11.181>
- García-Ordás, M. T., et al. (2021). Sentiment analysis in non-fixed length audios using a fully convolutional neural network. *Biomedical Signal Processing and Control*, 69, 102946. <https://doi.org/10.1016/j.bspc.2021.102946>
- He, Y., et al. (2022). An improved recommendation based on graph convolutional network. *Journal of Intelligent Information Systems*, 1–23. <https://doi.org/10.1007/s10844-022-00727-3>
- Hidayat, T. H. J., et al. (2022). Sentiment analysis of twitter data related to rinca island development using doc2vec and svm and logistic regression as classifier. *Procedia Computer Science*, 197, 660–667. <https://doi.org/10.1016/j.procs.2021.12.187>
- Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. <https://ojs.aaai.org/index.php/ICWSM/article/view/14550>
- Imani, M., & Noforesti, S. (2022). Aspect extraction and classification for sentiment analysis in drug reviews. *Journal of Intelligent Information Systems*, 1–21. <https://doi.org/10.1007/s10844-022-00712-w>
- Islam, A. (2022). Work-from/at-for-home: Covid-19 and the future of work-a critical review. *Geoforum*, 128, 33–36. <https://doi.org/10.1016/j.geoforum.2021.11.018>
- Jain, T., et al. (2022). Covid and working from home: Long-term impacts and psycho-social determinants. *Transportation Research Part A: Policy and Practice*, 156, 52–68. <https://doi.org/10.1016/j.tra.2021.12.007>
- Joseph, J., et al. (2022). A survey on deep learning based sentiment analysis. *Materials Today: Proceedings*, 58, 456–460. <https://doi.org/10.1016/j.matpr.2022.02.483>
- Kawakubo, S., & Arata, S. (2022). Study on residential environment and workers' personality traits on productivity while working from home. *Building and Environment*, 212, 108787. <https://doi.org/10.1016/j.buildenv.2022.108787>
- Khasanah, I. N. (2021). Sentiment classification using fasttext embedding and deep learning model. *Procedia Computer Science*, 189, 343–350. <https://doi.org/10.1016/j.procs.2021.05.103>
- Li, W., et al. (2022). Bieru: Bidirectional emotional recurrent unit for conversational sentiment analysis. *Neurocomputing*, 467, 73–82. <https://doi.org/10.1016/j.neucom.2021.09.057>

- Li, W., Zhu, L., et al. (2020). User reviews: Sentiment analysis using lexicon integrated two-channel cnn-lstm family models. *Applied Soft Computing*, 94, 106435. <https://doi.org/10.1016/j.asoc.2020.106435>
- Liao, S., et al. (2017). Cnn for situations understanding based on sentiment analysis of twitter data. *Procedia Computer Science*, 111, 376–381. <https://doi.org/10.1016/j.procs.2017.06.037>
- Liu, S., & Liu, J. (2021). Public attitudes toward covid-19 vaccines on englishlanguage twitter: A sentiment analysis. *Vaccine*, 39(39), 5499–5505. <https://doi.org/10.1016/j.vaccine.2021.08.058>
- Luo, M., & Mu, X. (2022). Entity sentiment analysis in the news: A case study based on negative sentiment smoothing model (nssm). *International Journal of Information Management Data Insights*, 2(1), 100060. <https://doi.org/10.1016/j.ijime.2022.100060>
- Neogi, A. S., et al. (2021). Sentiment analysis and classification of indian farmers' protest using twitter data. *International Journal of Information Management Data Insights*, 1(2), 100019. <https://doi.org/10.1016/j.ijime.2021.100019>
- Nezhad, Z. B., & Deihimi, M. A. (2022). Twitter sentiment analysis from iran about covid 19 vaccine. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 16(1), 102367. <https://doi.org/10.1016/j.dsx.2021.102367>
- Pathak, A. R., et al. (2021). Topic-level sentiment analysis of social media data using deep learning. *Applied Soft Computing*, 108, 107440. <https://doi.org/10.1016/j.asoc.2021.107440>
- Pavitha, N., et al. (2022). Movie recommendation and sentiment analysis using machine learning. *Global Transitions Proceedings*, 3(1), 279–284. <https://doi.org/10.1016/j.gltp.2022.03.012>
- Prodanova, J., & Kocarev, L. (2021). Is job performance conditioned by workfrom-home demands and resources? *Technology in Society*, 66, 101672. <https://doi.org/10.1016/j.techsoc.2021.101672>
- Rachunok, B., et al. (2022). Is the data suitable? the comparison of keyword versus location filters in crisis informatics using twitter data. *International Journal of Information Management Data Insights*, 2(1), 100063. <https://doi.org/10.1016/j.ijime.2022.100063>
- Rajalakshmi, S., & et al. (2017). A comprehensive survey on sentiment analysis. In 2017 fourth international conference on signal processing, communication and networking (ICSCN) (pp. 1-5). <https://doi.org/10.1109/ICSCN.2017.8085673>
- Rakshitha, K., et al. (2021). Sentimental analysis of indian regional languages on social media. *Global Transitions Proceedings*, 2(2), 414–420. <https://doi.org/10.1016/j.gltp.2021.08.039>
- Rani, S., et al. (2022). An efficient cnn-lstm model for sentiment detection in# blacklivesmatter. *Expert Systems with Applications*, 193, 116256. <https://doi.org/10.1016/j.eswa.2021.116256>
- Ridhwan, K. M., & Hargreaves, C. A. (2021). Leveraging twitter data to understand public sentiment for the covid-19 outbreak in singapore. *International Journal of Information Management Data Insights*, 1(2), 100021. <https://doi.org/10.1016/j.ijime.2021.100021>
- Salim, S., & et al. (2022). Data analytics of social media 3.0: Privacy protection perspectives for integrating social media and internet of things (sm-iot) systems. *Ad Hoc Networks*, 128, 102786. <https://doi.org/10.1016/j.adhoc.2022.102786>
- Sasidhar, T. T., et al. (2020). Emotion detection in hinglish (hindi+ english) code-mixed social media text. *Procedia Computer Science*, 171, 1346–1352. <https://doi.org/10.1016/j.procs.2020.04.144>
- Sharma, A. K., et al. (2020). Sentimental short sentences classification by using cnn deep learning model with fine tuned word2vec. *Procedia Computer Science*, 167, 1139–1147. <https://doi.org/10.1016/j.procs.2020.03.416>
- Swain, P. H., & Hauska, H. (1977). The decision tree classifier: Design and potential. *IEEE Transactions on Geoscience Electronics*, 15(3), 142–147. <https://doi.org/10.1109/TGE.1977.6498972>
- Tønnessen, Ø., et al. (2021). Digital knowledge sharing and creative performance: Work from home during the covid-19 pandemic. *Technological Forecasting and Social Change*, 170, 120866. <https://doi.org/10.1016/j.techfore.2021.120866>
- Umair, A., & Masciari, E. (2022). Sentimental and spatial analysis of covid-19 vaccines tweets. *Journal of Intelligent Information Systems*, 1–21. <https://doi.org/10.1007/s10844-022-00699-4>
- Wang, Z., et al. (2020). Multi-level fine-scaled sentiment sensing with ambivalence handling. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 28(04), 683–697. <https://doi.org/10.1142/S0218488520500294>
- Yang, F.-J. (2018). An implementation of naive bayes classifier. In 2018 international conference on computational science and computational intelligence (CSCI) (pp. 301-306). <https://doi.org/10.1109/CSCI46756.2018.00065>
- Yousefinaghani, S., et al. (2021). An analysis of covid-19 vaccine sentiments and opinions on twitter. *International Journal of Infectious Diseases*, 108, 256–262. <https://doi.org/10.1016/j.ijid.2021.05.059>

- Zachlod, C., et al. (2022). Analytics of social media data-state of characteristics and application. *Journal of Business Research*, 144, 1064–1076. <https://doi.org/10.1016/j.jbusres.2022.02.016>
- Zhao, Z., et al. (2022). Graph convolutional network with multiple weight mechanisms for aspect-based sentiment analysis. *Neurocomputing*, 500, 124–134. <https://doi.org/10.1016/j.neucom.2022.05.045>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.