Comparison of different machine learning techniques on location extraction by utilizing geo-tagged tweets: A case study

Nazmiye Eligüzel[a,*], Cihan Çetinkaya[b], Türkay Dereli[a,c]

[a] *Gaziantep University, Industrial Engineering, 27310 Gaziantep, Turkey*
[b] *Adana Alparslan Turkes Science and Technology University, Department of Management Information Systems, 01250 Adana, Turkey*
[c] *Hasan Kalyoncu University, Office of the President, Gaziantep, Turkey*

ARTICLE INFO

ABSTRACT

In emergencies, Twitter is an important platform to get situational awareness simultaneously. Therefore, information about Twitter users' location is a fundamental aspect to understand the disaster effects. But location extraction is a challenging task. Most of the Twitter users do not share their locations in their tweets. In that respect, there are different methods proposed for location extraction which cover different fields such as statistics, machine learning, etc. This study is a sample study that utilizes geo-tagged tweets to demonstrate the importance of the location in disaster management by taking three cases into consideration. In our study, tweets are obtained by utilizing the "earthquake" keyword to determine the location of Twitter users. Tweets are evaluated by utilizing the Latent Dirichlet Allocation (LDA) topic model and sentiment analysis through machine learning classification algorithms including the Multinomial and Gaussian Naïve Bayes, Support Vector Machine (SVM), Decision Tree, Random Forest, Extra Trees, Neural Network, k Nearest Neighbor (kNN), Stochastic Gradient Descent (SGD), and Adaptive Boosting (AdaBoost) classifications. Therefore, 10 different machine learning algorithms are applied in our study by utilizing sentiment analysis based on location-specific disaster-related tweets by aiming fast and correct response in a disaster situation. In addition, the effectiveness of each algorithm is evaluated in order to gather the right machine learning algorithm. Moreover, topic extraction via LDA is provided to comprehend the situation after a disaster. The gathered results from the application of three cases indicate that Multinomial Naïve Bayes and Extra Trees machine learning algorithms give the best results with an F-measure value over 80%. The study aims to provide a quick response to earthquakes by applying the aforementioned techniques.

1. Introduction

Location information plays an important role in disaster situations. This information facilitates the decision-making process. In today's world, Twitter is the most popular platform, where corresponding information is shared. Provided that geospatial information is given for the tweets, a necessary intervention can be made in emergency management. Twitter posts with geographic locations are attributed to "geo-tagged tweets" [1]. However, there are very few tweets shared with geo-tagged. Tweet stream with geo-tagged provides a precious source for local event detection. Zhang et al. [2] presented an approach for detecting real-time local events effectively named GEOBURST +, support a new cross-modal authority measure to determine some pivots that extract to various geo-topical activities in the query window. It is an effective method as compared to the existing methods. The

utilization of this method is not limited to Twitter. Tsou et al. [1] developed a geo-targeted event observation that provides real-time social media messages monitoring, taking four functions into account which are geo-tagged tweets, interactive mapping function, text-tagging of messages and manual labeling, and spatial boundaries, date, and temporal search function. Murthy and Gross [3] proposed an article that seeks to analyze shifts in social media behavior using geo-tagged tweets provided before and after Hurricane Sandy. Bhuvaneswari and Valliyammai [4] proposed an information entropy-based event detection to explore the event and its location using clustered tweets according to the high-density ratio of tweets. In the study, geo-tagged tweets were extracted to determine the location of events during a disaster in cyber-social networks. Caragea et al. [5] performed a geo-mapped sentiment analysis on Twitter during Hurricane Sandy. Therefore, they indicated how Twitter users' sentiments vary according to not only the location of

users but also based on the distance from the disaster. Paule et al. [6] developed a location inference method based on a ranking approach integrated with a majority voting of tweets. Geo-tagged tweets were used from two cities, and satisfying results were provided in terms of accuracy and error distance. An emergency event detection framework with the social internet of things was proposed by Photographs et al. [7] by utilizing geo-tagged microblogs to determine the disaster events for a specified location and time. The paper focused on textual content in contrast to conventional approaches. For the search and rescue operations during an emergency event, location information is very important but geo-tagged tweets are highly rare. Therefore, Flizikowski et al. [8] developed a tool named "Text Analysis Tweet (Locator-TAT2) to enable additional hints about people's location on Twitter.

The contribution of this paper is the demonstration of the importance of location information in disaster management by utilizing the geo-tagged tweets. In our study, earthquake-related tweets are obtained and geo-tagged tweets are determined among these tweets. Therefore, it can be possible to focus on which locations are discussed in the case of an earthquake. LDA model and sentiment analysis are demonstrated by the application of machine learning classification algorithms. In our study, we propose a comparison of 10 different machine learning algorithms in order to provide reliable algorithms that can be utilized in the case of an earthquake. From our findings, there are two outstanding algorithms confirmed, which are Multinomial Naïve Bayes and Extra Trees for three cases by considering the achievement over 80% F-measure value. Therefore, we proved that current situation for earthquakes can be extracted effectively by the utilization of these two algorithms. It is also important to have an overall perspective on the general situation for disaster. Therefore, LDA is applied to have a better understanding of the attitude of society in the case of an earthquake.

The rest of the paper is organized as follows: Section 2 introduces related works. Section 3 describes the methodology by utilizing LDA and sentiment analysis, and performance evaluation techniques in the proposed study. In Section 4, we present the results and discussions. Finally, Section 5 presents the conclusion.

2. Literature review

The utilization of Twitter during emergencies has gained popularity because of obtaining real-time information. In our study, we focus on studies that include disaster, location extraction, and Twitter topics.

There are various techniques for location extraction from formal or informal texts such as geographical information systems [9–11], natural language processing techniques especially named entity recognition [12–15], machine learning-based approaches [16–20], gazetteer-based approaches [21,22], and rule-based approaches [23]. These different techniques can be integrated and then used. In the proposed study, geo-tagged tweets are evaluated to emphasize the importance of location information by utilizing machine learning algorithms.

In recent years, several studies have been performed related to geo-tagged tweets approaches for extracting location from tweets as follows:

A multi-elemental location inference method was developed [23] to separate geo-tagging tweets and estimate the location of tweets by utilizing the other related data elements. The user's profile location, textual concept, and place labeling were taken into account as the main location-attached elements. According to the results, 87% of the tweets' location, average distance error of 12.2 km, and the median distance error of 4.5 km were estimated in a good way. The proposed method is suitable for English tweets and may not be feasible for non-English languages. It estimated the location with extensive distance errors and provided significant improvement compared with the existed methods. Ogie and Forehead [24] proposed a study that refers to develop an understanding degree of geo-referenced locations of social media flood reports diverge from the real locations of the flood. Results showed that the accuracy of flood maps created with geo-referenced social media

data decreases with the rise in the size of the minimum mapping unit. OzCT geo-tagger was developed which automatically determines the aforementioned locations in the content of the tweets [25]. There are three probabilities: definite, uncertain, and no-location for this geo-tagger. Experimental results showed that the precision and recall for determination of certain locations are on average of 80% compared with geo-tagging by human decision. Rakesh et al. [26] demonstrated the significance of identifying location-specific tweets. Due to the challenges in obtaining geo-tagged tweets, they developed a novel model referred to Location Centric Word Co-occurrence that utilizes tweet content and network information of twitter users to provide location-specific tweets. The proposed model tested by utilizing geo-tagged tweets and compared with other weighting schemes. LDA was utilized to train the model. In this study, findings indicate users' network score occupies a significant place to determine location-specific tweets. Another study that uses the geo-tag tweets proposes a different application area that aims detection of underlying reasons for local criminal activity by using simple LDA to make clustering [27].

Machine learning algorithms have gained importance in disaster management through Twitter. Some of the disaster-related studies utilized machine learning approaches is demonstrated as follows:

Sadri et al. [28] investigated the evolution of different communication patterns by utilizing machine learning algorithms and identified user fears that came out with Hurricane Sandy such as storm location and time, media coverage, the behavior of political leaders and celebrities. Markov Chain Monte Carlo method and Gibbs sampling approaches were utilized to estimate the parameters of this model. Perplexity was utilized as a performance metric to measure the predictive capacity of the proposed model. A novel two phases of machine learning and a deep learning-based framework for power outage identification from Twitter were proposed by Mao et al. [17]. In the first phase, a probabilistic classification model was implemented by utilizing a bag of ngrams characteristics to find true power outage tweets. In the second phase, bidirectional long short-term memory networks as a novel deep learning approach were applied to obtain outage locations from context. Experimental results demonstrated that promising classification accuracy was provided with 86%. Jamali et al. [18] proposed a multi-step methodology for the post-disaster recovery phase of Hurricane Sandy. Determination of the people exposed to the disaster, their physical location, topics that discussed post-disaster, evaluation of the tract-level relationships between topics, and tract-level internal attributes are the outputs of this methodology. Major topics were faith-based, community, assets, and financial topics in the study. A Dirichlet regression model was used to determine important internal attributes and study the tract-level correlation between these attributes and determined discussion topics. Li et al. [29] focused on the limitation of supervised classifiers. Supervised classifiers learn only from the source disaster rather than the target disaster. To overcome this limitation, a domain adaptation approach that learns classifiers from unlabeled target data was proposed. Their approach includes the Naïve Bayes classifier with an iterative Self-Training strategy. They utilized crisis tweets data-set to evaluate their proposed approach. The experimental results demonstrated that the proposed domain adaptation approach is better as compared to the supervised classifiers learned exclusively from labeled source data. Study that deals with identifying disaster-related events via tweeter corpus by change point detection and brust corpus identification. In this study, k–means clustering is utilized for event clustering through similarity measure concerning the distance between user pairs calculated by Jaccard Coefficient Index [4].

In our study, machine learning algorithms are compared according to sentiment analysis. Some of the studies that deal with sentiment analysis by utilizing machine learning techniques are demonstrated as follows:

Almatrafi et al. [30] performed a location-based sentiment analysis through Indian general elections to determine the characteristics of tweets by utilizing machine learning techniques and natural language

processing. Naive Bayes algorithm was utilized for the classification. 70% accuracy was obtained together with 80% recall and 66% precision. They demonstrated sentiment change from one location to another. The study considers a point of interest that is a specific location point utilized on location-based social media implementations. Sentiments of tweets were investigated by utilizing both machine learning techniques such as conditional random fields, Naive Bayes, and Decision Tree techniques, and artificial neural network approaches. Results show that for the point of interest extraction problem location extraction solutions do not satisfy and accuracy of proposed conditional random field implementation has the best accuracy [31]. A study that deals with machine learning algorithms through tweets is applied eight different linguistic features with kNN, Naïve Bayes, and SVM [32]. In aforementioned study, game-with-a-purpose was utilized to crowdsource for identification of tweet topics and sentiments. Results show that SVM at each decision point provides better performance compared to rest of the applied algorithms. Reynard et al. [20] used machine learning techniques and geospatial to obtain geo-located tweets during Hurricane Irma in 2017. Sentiment analysis was applied for the classification of tweets to specify the multinomial logit model. Another study that proposes a technique to understand the correlated factors and their relevance through sentiment analysis according to the locations by using machine learning approaches such as Naive Bayes, SVM, and Random Forest classifier. In the study, SVM gives the best results with maximum efficiency [33].

The aforementioned studies consist of the location extraction through geo-tagged tweets, application of machine learning algorithms on Twitter for disaster management, and application of machine learning algorithms by utilizing sentiment analysis. However, there is no study compares several machine learning algorithms based on geo-tagged tweets. In our study, a comparison of 10 different machine learning algorithms is accomplished by making sentiment analysis on location-specific tweets for earthquake situation due to demonstrate the location information importance in disaster management by utilizing the geo-tagged tweets for fast responses towards situations. In addition, topic extraction is managed for location-specific tweets by using LDA. To the best of the authors' knowledge, there is no proposed study that deals with retrieving disaster-based location-specific tweets and demonstrates the effectiveness of most of the machine learning techniques on sentiment analysis in case of an earthquake.

3. Methodology

This section demonstrates the extraction of geo-tagged tweets by applying the LDA model and sentiment analysis. In Fig. 1, the model structure is presented.

3.1. Data collection, identifying of geo-tagged tweets, pre-processing stage, word frequencies

In our study, the tweets have been collected by utilizing the "earthquake" keyword from 3rd December to 9th December 2018 through Twitter API. Approximately a number of 3547 tweets have been obtained and evaluated. Among these tweets, a number of 931 tweets have been identified as geo-tagged. In Fig. 2, the numbers of geo-tagged tweets are demonstrated with respect to locations.

As seen in Fig. 2, the number of 152 Twitter users utilizes location information part in their profiles without location name. Their location name does not appear most likely because of occurrence of blank at the location name place. The number of 24 users from Washington Dc, 22 users from Anchorage Ak, 16 users from the United State, 11 users from Anchorage Alaska, 9 users from India are identified, respectively. Most of the remaining 697 location names are demonstrated as others because of too much variation in locations. In our study, the number of 51 tweets has been analyzed from the Washington location as a sample study. Tweets from Washington together with date and time can be seen
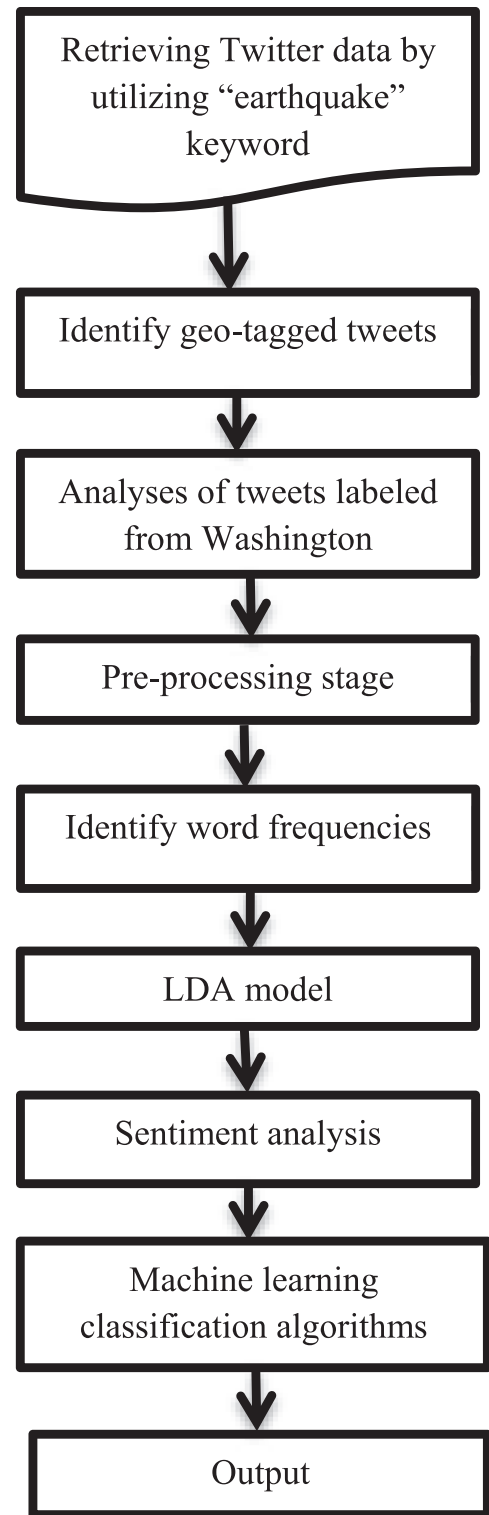
Fig. 1. Model structure of the proposed model.

in Appendix A in Table A.1.

To provide accurate information from tweets, the pre-processing stage is applied on data. In the pre-processing stage, URLs, hashtags, stop words, and punctuations are removed. The usernames which start with '@' are kept in the tweet to obtain location. All tweets are tokenized to evaluate topics and stemming process is applied to data-set.

After the pre-processing stage, all the words in the text are counted and word frequencies are determined. As seen in Table 1, 'earthquake',

Fig. 2. The number of geo-tagged tweets.

Table 1
Word counts of Twitter data-set.

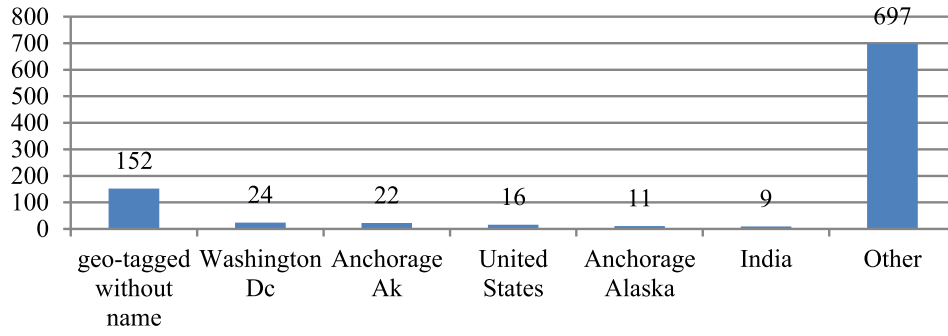| Word	Count
"earthquake"	39
"alaska"	13
"anchorage"	10
"wave"	5
"quake"	5
"help"	5
"tsunami"	4
"near"	4
"time"	4
"disaster"	3

Fig. 3. Structure of LDA model.

a topic z_n from θ_d, and select a word from topic z_n [35]. In our study, the number of topics (K) is considered as 50 and Dirichlet hyper-parameters are considered as $\beta = 0.1$ and $\alpha = 50/K$ in this model [36]. Topic distributions are generated for each document. These distributions demonstrate a relationship between each document and the topic. Some selected topics obtained by the LDA model can be seen in Table 2.

The number of 5 topics is selected randomly. The total scores of words for each topic can be seen in the bottom row of Table 2. The words on the same topic tend to be parallel. These associated words can be labeled with topic names. For example, topic 2 is about the earthquake in Alaska, topic 14 is about technical information related to the earthquake. Logarithmic probabilities can be seen in Fig. 4 for the Twitter data-set. A low log-probability may propose that the document may be an outlier.

The indices of the three documents with the lowest log-probability are 22, 27, and 44, respectively. These three tweets may be referred to as outliers.

'alaska', 'anchorage', 'wave', 'quake', 'help', 'tsunami', 'near', 'time', 'disaster' is the highest mentioned words within top 10 words.

As understood from the table, an earthquake has occurred in Alaska and it has been discussed with the dimension of the tsunami, time, help, etc. Frequency of words can give us an idea about the topics discussed. However, the LDA model is applied to evaluate which topics are discussed in detail.

3.2. Latent Dirichlet Allocation (LDA)

Powerful natural language tools can be considered under topic generative models which are utilized for divergence of data corpus topics [34]. Latent Dirichlet Allocation (LDA) is a generative probabilistic model of a corpus [35]. Underlying topics in a collection of documents is identified and probabilities of words in topics are deduced by LDA.

In Fig. 3, the outer box indicates documents, while the inner box indicates the repeated selection of topics and words inside a document. The probability of corpus can be seen in Eq. (1).

$$p(D|\alpha, \beta) = \prod_{d=1}^{M} \int P(\theta_d|\alpha) \left(\prod_{n=1}^{N_d} \sum_{z_{dn}} p(z_{dn} \mid \theta_d) p(w_{dn}|z_{dn}, \beta) \right) d\theta_d \tag{1}$$

where α is the Dirichlet-prior concentration parameter of the per-document topic distribution, β is the same parameter of the per-topic word distribution, θ_d is the topic distribution for document d, z_{dn} is the topic assignment for w_{dn}, w_{dn} is the nth word in the dth document, K is the number of topics, N is the number of words in the document, M is the number of documents to analyze, D is the corpus of collection M documents.

After the pre-processing stage for Twitter data, the LDA model is implemented by utilizing MATLAB software. LDA has two parts which are the words in a document and the probability of words in a topic. The generative model of LDA works as follows: choosing a multinomial distribution from a Dirichlet distribution for topic t with parameter β, choosing a multinomial distribution from a Dirichlet distribution for document d with parameter α, and for a word w_n in document d, select
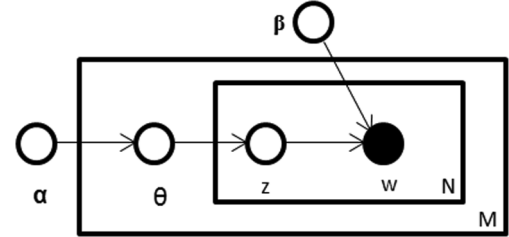
3.3. Sentiment analysis

Sentiment analysis is a method for finding polarity or understand emotions from words [37]. Sentiments and opinions are represented in the text through sentiment analysis. Variations can occur in opinions with respect to locations as needs of the people can vary depending on the location [32]. In addition, it is important to know people's idea for specific topics and issues. Therefore, sentiment analysis is applied. In our study, positive [38] and negative [39] word lists are utilized for determining sentiments of the tweets to deduce positive and negative words in our data-set. Meaning Cloud text analytic solution[1] and Python programming are utilized for the sentiment analysis. In Table A.2, sentiment analysis of the tweets can be seen.

For this analysis, 1 refers to positive meaning, -1 refers to negative meaning, and 0 refers to neutral meaning in the text. Considering these sentiments, machine learning algorithms are utilized to learn the new inputs of the system. Machine learning algorithms can learn from available inputs. Experience is converted to expertise or knowledge by the learning process [40].

---

[1] https://www.meaningcloud.com/

Table 2
Topic distribution of some selected topics.

Topic 2	Topic 6	Topic 9	Topic 14	Topic 20
"anchorage"	"alaska"	"wave"	"earthquake"	"ground"
"hit"	"earth"	"record"	"time"	"happen"
"moment"	"coast"	"newcaledonia"	"minute"	"recent"
"watch"	"break"	"volcano"	"aftershock"	"usar"
"america"	"warn"	"donaldtrumpv"	"great"	"need"
0.6827	0.6475	0.5462	0.8106	0.7015
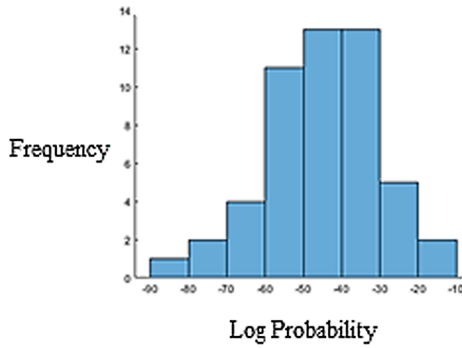
Fig. 4. Log-probabilities for the Twitter data.

Table 3
The best results of classification algorithms for Washington DC tweets.

Algorithm	Max-feature	ACC	Precision	Recall	F-measure
Naïve Bayes(Multinomial)	25	0.83	0.83	0.83	0.83
Naïve Bayes(Gaussian)	100	0.66	0.83	0.67	0.74
SVM	25	0.66	0.83	0.67	0.74
Decision Tree	25	0.66	0.83	0.67	0.74
Random Forest	1000	0.83	0.83	0.83	0.83
Extra Trees	100	0.83	0.83	0.83	0.83
Neural Network	100	0.83	0.69	0.83	0.76
kNN	25	0.83	0.83	0.83	0.83
SGD	100	0.83	0.69	0.83	0.76
AdaBoost	25	0.5	0.83	0.50	0.62

3.4. Machine learning classification algorithms

Machine learning is referred to as the automated finding of significant patterns in data [40]. Classification is referred to as assigning a class label to set of unclassified cases based on attributes, behavior or subject [41]. In the proposed study, for the classification, machine learning algorithms including the Naïve Bayes, Support Vector Machine (SVM), Decision Tree, Random Forest, Extra Trees, Neural Network, k Nearest Neighbor (kNN), Stochastic Gradient Descent (SGD), and Adaptive Boosting (AdaBoost) algorithms are utilized to classify tweets according to the sentiments.

3.4.1. Naïve Bayes

Naïve Bayes is one of the classification algorithms that is based on Bayes theorem with robust and naïve independence assumptions [42]. It corresponds to a Bayesian rule, as in Eq. (2).

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \tag{2}$$

where y denotes a class label and xi denotes a value of an attribute *Xi*. Based on the above equation, in the proposed study, the Multinomial Naïve Bayes algorithm and Gaussian Naïve Bayes algorithmare implemented. Multinomial Naïve Bayes algorithm is utilized in the text classification where the data are typically demonstrated as word vector counts. It considers assumptions. For the Gaussian Naïve Bayes

algorithm, the probability of the features is assumed to be Gaussian. Gaussian Naïve Bayes is utilized to handle continuous attributes. The reason behind the utilization of this algorithm is working well with the classifications of natural language processing (NLP) [43].

3.4.2. Support Vector Machine (SVM)

SVM is the binary classification algorithm that deals with the sample complexity challenge by exploration for "large margin" separators [40]. Hyperplanes are utilized as decision boundaries to classify the data points. SVM is developed from sound theory to its implementation and experiments. For the SVM, there are no parameters. Therefore, it is the non-parametric model [44]. In the study, Linear Support Vector Classification is utilized. The motivation behind using the SVM algorithm derives from its success in many NLP areas and its capability of handling large number of features. [45].

3.4.3. Decision trees

Decision Trees are a non-parametric method utilized for classification and regression. The number of nodes forms the decision trees by creating a rooted tree. That means it is a directed tree with a root node without incoming edges. The rest of the nodes have one incoming edge. Outgoing edges referred to as internal or test nodes. Finally, the other nodes are referred to as leaves [46]. Decision Tree learning aims to develop a model that estimates the target value variable considering several input variables [47]. It is a graphical model that describes decisions and their potential outcomes. They are ideally representative of ensemble methods. We choose the Decision Tree for the comparison because of its frequent utilization in sentiment classification [48].

3.4.4. Random forest

It is one of the classifier algorithms including a collection of decision trees, where each tree is established by implementing an algorithm. The estimation of the random forest is achieved by a majority vote over the estimations of the individual trees [40]. Each tree in the ensemble is constructed from a sample drawn with replacing from the training set. Decision Tree algorithm can lead overfitting problem. Therefore, Random Forest is utilized to solve this problem with its randomness. Due to the provide best randomness together with data and features, Random Forest algorithm is selected [49].

3.4.5. Extra trees

Extra Trees are extremely randomized trees based on ensemble method that reduces the poor generalization property and inclination to overfitting of conventional standalone decision trees [50]. This class applies a meta-predictor in which a number of randomized decision trees are fitted. Utilization of Extra Trees comes from the its differentiation from the Random Forest with randomly selected best feature. [51].

3.4.6. Neural network

As a supervised learning algorithm, a Neural Network is defined as a directed graph that its nodes related to the neurons and edges correspond to links between them [40]. A neural network is similar to the brain in two ways: through a learning process that knowledge is provided by the network from its environment, and through synaptic weights that store the obtained knowledge [52]. In the study, Multilayer Perceptron (MLP) classifier is implemented. Because of the generating own data statement in training phase, Neural Network is utilized for comparison [53].

3.4.7. k-Nearest Neighbors (kNN)

It is a basic learning algorithm based on the assumption that "things that look alike must be alike." [40]. It is the non-parametric algorithm. We used the k-neighbors classification as the most commonly utilized method. The reason behind the utilization of this algorithm is capable of dealing with a small number of training models [43].
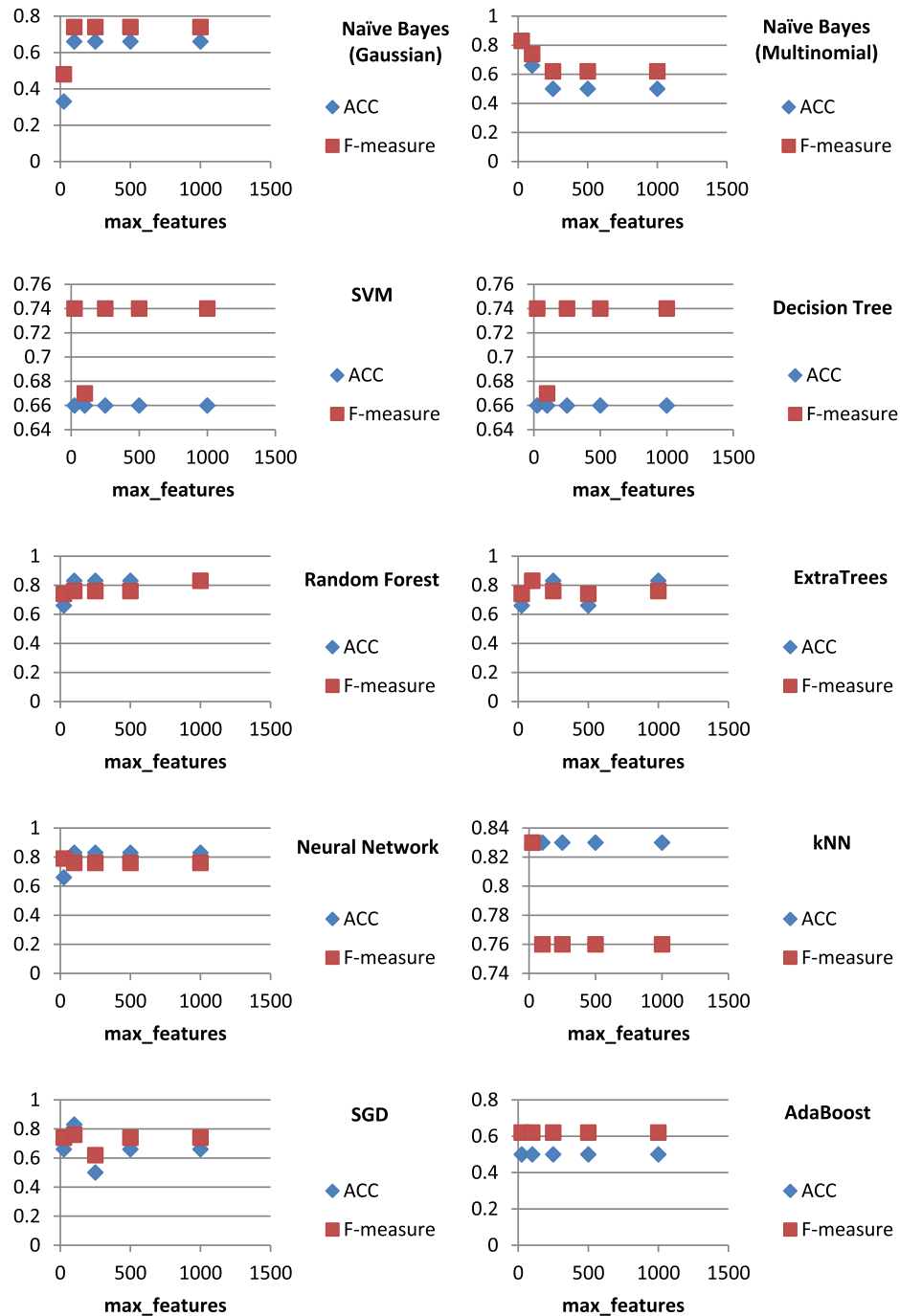
Fig. 5. The effects of max_features on classification algorithms considering F-measure and ACC for Washington DC tweets.

3.4.8. Stochastic Gradient Descent (SGD)

It is successfully implemented in text classification and natural language processing. SGD is referred to as a very influential method by utilizing the discriminative learning of linear classifiers below convex loss functions. Because of its common utilization in literature [54], SGD is included in rest of the algorithms.

3.4.9. Adaptive Boosting (AdaBoost)

The AdaBoost algorithm yields a hypothesis which is a linear combination of simple hypotheses [40]. The main principle of AdaBoost is to fit a sequence of poor learners on continuously changed versions of the data. It is the popular boosting algorithm. Due to having considerable attention from researchers and providing satisfying accuracy on imbalanced data, the AdaBoost algorithm is applied [55].

3.5. Performance evaluation

In the proposed study, the classification machine learning algorithms including the Naïve Bayes, Support Vector Machine (SVM), Decision Tree, Random Forest, Extra Trees, Neural Network, k Nearest Neighbor (kNN), Stochastic Gradient Descent (SGD), and Adaptive Boosting (AdaBoost) algorithms are utilized. Test size is taken as 0.10 by trial and error. Each classification algorithm has been run 5 times considering count vectorizer with maximum features (max_features). Max_features consider the term frequency and feature selection. The number of 25, 100, 250, 500, and 100 max_features are utilized for classification, respectively. Evaluation metrics are calculated by the given equations.

Classification accuracy (ACC): In the light of confusion matrixes,

Fig. 6. The effects of max_features on classification algorithms considering F-measure and ACC for first comparison.

Table 4
The best results of classification algorithms for first comparison.

Algorithm	Max-feature	ACC	Precision	Recall	F-measure
Naïve Bayes(Multinomial)	500	0.85	0.79	0.86	0.81
Naïve Bayes(Gaussian)	100	0.71	0.64	0.71	0.67
SVM	25	0.85	0.74	0.86	0.79
Decision Tree	25	0.78	0.85	0.79	0.79
Random Forest	25	0.78	0.85	0.79	0.79
Extra Trees	25	0.85	0.93	0.86	0.87
Neural Network	25	0.85	0.88	0.86	0.85
kNN	25	0.76	0.79	0.76	0.73
SGD	250	0.85	0.74	0.86	0.79
AdaBoost	25	0.85	0.93	0.86	0.87

accuracies of classifiers are evaluated.

$$ACC = \frac{TN + TP}{TP + FP + FN + TN} \tag{3}$$

where TP, FP, FN, TN represent the number of true positives, false positives, false negatives, and true negatives, respectively.

Precision: It is the ratio of the number of accurately predicted sentiments to the total number of predicted sentiments.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

Recall: It is the ratio of the number of accurately predicted sentiments to the total number of actual sentiments.
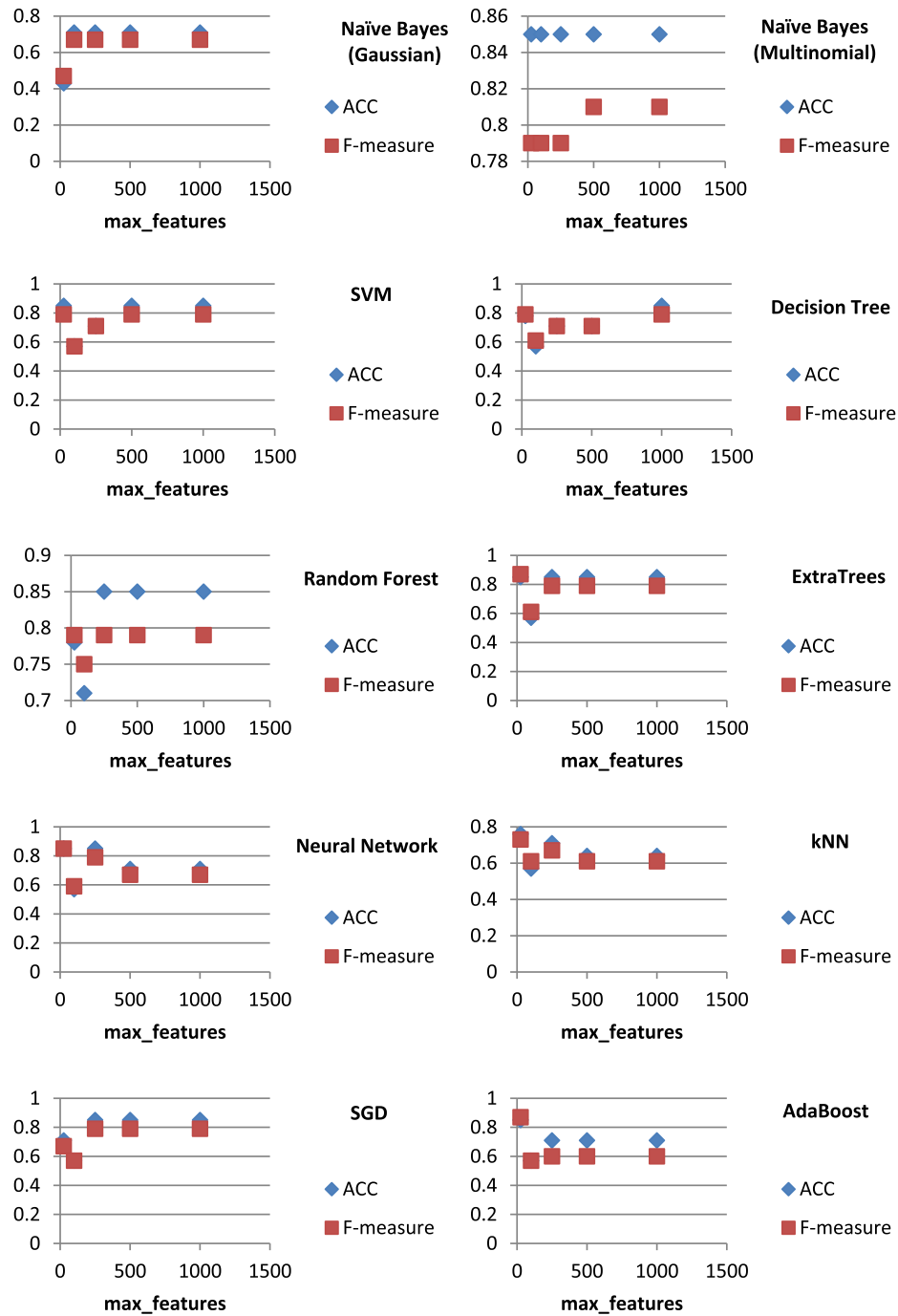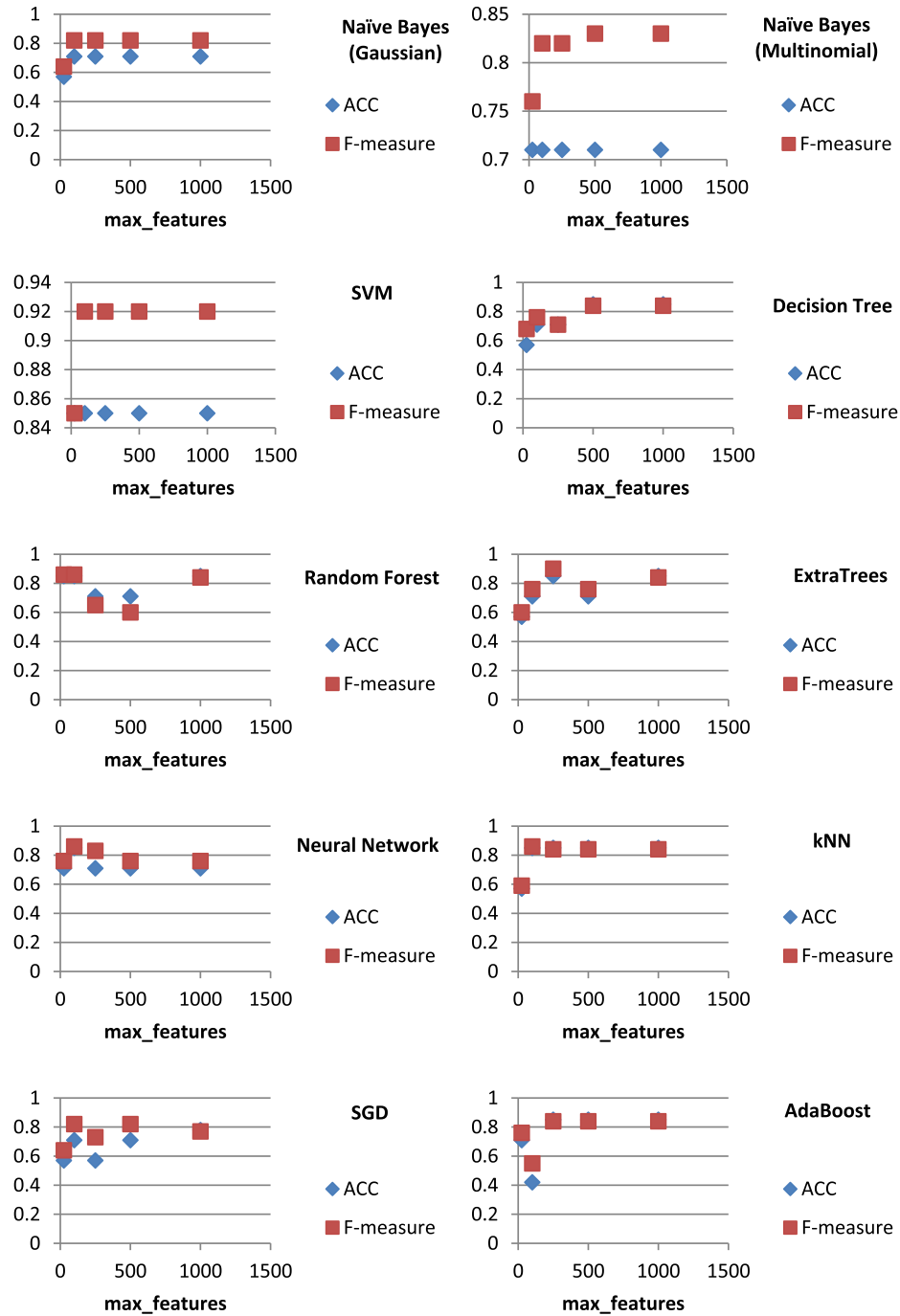
Fig. 7. The effects of max_features on classification algorithms considering F-measure and ACC for second comparison.

Table 5
The best results of classification algorithms for second comparison.

Algorithm	Max-feature	ACC	Precision	Recall	F-measure
Naïve Bayes(Multinomial)	500	0.71	1	0.71	0.83
Naïve Bayes(Gaussian)	100	0.71	1	0.71	0.82
SVM	100	0.85	1	0.86	0.92
Decision Tree	500	0.85	0.88	0.86	0.84
Random Forest	25	0.85	0.90	0.86	0.86
Extra Trees	250	0.85	1	0.86	0.90
Neural Network	100	0.85	0.90	0.86	0.86
kNN	100	0.85	0.90	0.86	0.86
SGD	100	0.71	1	0.71	0.82
AdaBoost	250	0.85	0.88	0.86	0.84

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

F-measure: F- measure is the other widely utilized for performance evaluation of classification algorithms. It is the harmonic mean of the precision and recall of a classification algorithm. Higher values of F-measure mean better predictive performance. In this study, macro-averaged F-measure is used, which defines the average F-measure across all one-versus-all classes [56].

$$Macro - averaged F - measure = \frac{1}{n} \sum_{i=1}^{n} \frac{2 * Precisioni * Recalli}{Preciisoni + Recalli} \tag{6}$$

4. Results and discussions

Considering the location name entered by Twitter users, the tweets from Washington DC have been evaluated. Words with highest frequency are 'alaska', 'anchorage', 'wave', 'quake', 'help', 'tsunami', 'near', 'time', 'disaster ' in the text. So, the content of the text is an earthquake in Alaska. Donald Trump, technical information about the earthquake, etc. are some topics that are obtained through topic analysis by using the LDA model. Sentiment analysis is applied to the dataset. In particular, negative thinking is found more than positive thinking. After that, machine learning algorithms are utilized to classify tweets according to the sentiments. In Table 3, the best results of classification algorithms can be seen.

As it is shown in Table 3, experimental design is implemented by considering all classification algorithms. Naïve Bayes Multinomial classification algorithm, Random Forest, Extra Trees, and kNN algorithms provide the best performance with the accuracy, F-measure, precision, and recall of 0.83. Moreover, the Random Forest algorithm gives good results when the number of max_features is increased. The worst results are obtained with the AdaBoost model. In Fig. 5, the effects of max_features on classification algorithms considering F-measure and ACC can be seen. The model has been run five times for each selected max_features.

From Fig. 5, it can be said that when number of max_features exceeds 1000, F-measure and ACC either decrease or do not change. Multinomial Naïve Bayes, SVM, Decision Tree, kNN, and AdaBoost classification algorithms give the best results with the number of 25 max_features. Incline in max_features does not have a positive effect on performance of applied algorithms. The Random Forest, Extra Trees, and SGD algorithms are the most affected algorithms by altering number of max_features.

Two other case studies are demonstrated to measure the effectiveness of the algorithms. In the first comparison case, tweets are collected from various locations such as San-Francisco, Denver, Oregon etc. Tweets and sentiments can be seen in Appendix B in Table B.1 and B.2, respectively. In Fig. 6, results of classification algorithms for the first comparison case by F-measure and ACC is demonstrated with respect to max_features. The results from the comparison indicate that Extra Trees, AdaBoost, Neural Network, and Multinomial Naïve Bayes classifications give the better results as it is demonstrated in Table 4.

In Table 4, SVM, Decision Tree, and Random Forest algorithms also give good results with 79% F-measure value. However, Gaussian Naïve Bayes algorithm leads to worst solution.

For the second comparison case, tweets are gathered from geo-tagged without a name, namely Twitter users utilize location information part in their profiles without a location name. Tweets and sentiments can be seen in Appendix C in Table C.1 and C.2, respectively. In Fig. 7, results for application of algorithms on second comparison case is given and it is shown that SVM, Random Forest, Extra Tress, Neural Network, and kNN give the best results according to F-measure with over 86% value. However, only Random Forest classification provides the better result with the number of 25 max_features.

From the Table 5, it can be seen that all algorithms perform good results. SVM provides outstanding result with 92% F-measure value. When it comes to Gaussian Naïve Bayes and SGD, it can be said that they gives satisfying results but worst among the others. In addition, it can be expressed that Gaussian Naïve Bayes gives the worst results for the other two cases.

The gathered results from the application of three cases demonstrate that Multinomial Naïve Bayes and Extra Trees machine learning algorithms lead better results with F-measure value over 80%.

5. Conclusion

In a nutshell, geospatial information is an important aspect on Twitter to improve situation awareness in disaster management and rescue operations. Geo-tagged tweets provide significant information about an event such as "where the event location", "where the aid is needed" etc. In the proposed study, tweets are collected by utilizing the "earthquake" keyword to determine the location of Twitter users. Selected tweets based on Washington location with a specified date and times are analyzed. In addition, two cases with tweets from various locations such as San-Francisco, Denver, Oregon, etc. and geo-tagged tweets without a name are applied in order to make comparisons. The pre-processing stage (removing punctuation, stop words, and URLs, tokenizing and stemming levels) is implemented. After that, the topic extraction process is applied by utilizing the Latent Dirichlet Allocation (LDA) model to demonstrate important topics in the document. According to the sentiment analysis, positive, negative, and neutral tweets are identified and machine learning algorithms are utilized to classify tweets according to the sentiments. Results demonstrate that the Multinomial Naïve Bayes classification algorithm and Extra Trees provide the best results for the applied three cases. Multinomial Naïve Bayes algorithm gives 83%, 81%, and 83% F-measure values for all cases, respectively. Extra Trees algorithm leads 83%, 87%, and 90%. Our contribution can be expressed as finding the best two algorithms among the aforementioned algorithms deducted from our application. Moreover, a comparative study for most of the machine learning algorithms conducted on sentiment analysis with location-specific tweets. Earthquake related tweets are retrieved and the best two algorithms are extracted among the 10 different algorithms in order to provide effective sentiment analyses in a disaster situation. When the location information is known, it will be much easier to meet the needs of victims. Therefore, location information on Twitter is very important.

In the future study, besides geo-tagged tweets, all tweets can be evaluated and locations can be extracted from non-geo-tagged tweets by utilizing content analysis techniques.

Declaration of Competing Interest

Acknowledgment

Appendix A

See Table A.1 and A.2.

Table A.1
Tweets from Washington location.

Location	Date/Time	Tweet
Washington,DC	2018-12-11 (16 22 10)	Hot Spots in West Coast, US. Chances of strong earthquakes and volcanoes have gone up Mr Trump.
Washington,DC	2018-12-07 (05 52 30)	TIME is running out on the West Coast of America. #earthquake #forecast #terremoto #tsunami
Washington,DC	2018-12-06 (03 59 05)	DonaldTrump is responsible for making the Slow Slip in California worse and a man made disaster that is going
Washington,DC	2018-12-05 (01 55 04)	Where are the investments in SOLAR energy? Or the Truth about Animal Agriculture? Time is running.
Washington,DC	2018-12-04 (18 12 12)	#Anchorage soon hit by few major Shocks……. #earthquake @ReutersWorld @realDonaldTrump @POTUS @USATODAY Ignorance.
Washington,DC	2018-12-11 (05 28 58)	This is another ground motion visualization showing the motion of the ground recorded by the USArray during the Anchorage.
Washington,DC	2018-12-08 (23 22 19)	waves travel through the Earth at approximately 1/2 the speed of P waves. Stations close to the #earthquake record strong
Washington,DC	2018-12-08 (00 49 35)	How do we capture the motion of an #earthquake?
Washington,DC	2018-12-07 (23 54 15)	What is a seismic shadow zone and what can it teach us about earth's interior structure?
Washington,DC	2018-12-07 (01 56 41)	Moment Magnitude Explained Happened to the Richter Scale?
Washington,DC	2018-12-06 (23 54 20)	TEACHERS TOOLBOX - How Hard Does the Ground Shake During a Quake?
Washington,DC	2018-12-06 (13 56 29)	Watch the waves from the Anchorage earthquake roll across the USArray seismic network in the lower.
Washington,DC	2018-12-05 (04 02 41)	Amazing computer simulation from Dr. Carl Tape shows a surface view of 3D #earthquake waves from the M7.0 #Anchorageearthquak
Washington,DC	2018-12-04 (22 38 56)	What is an #earthquake travel time curve? How are they created?
Washington,DC	2018-12-04 (21 25 07)	Here's a great visual to help show the difference between a M7.0 #earthquake and a M9. (Watch the video)
Washington,DC	2018-12-04 (20 12 29)	How are seismic waves in the Earth like a drop of water falling into a pool?
Washington,DC	2018-12-04 (19 35 10)	What's the difference between a M7.0 and M9.0 #earthquake? Find out using pasta noodles in this crazy cool video!
Washington,DC	2018-12-04 (13 02 39)	Fascinating visualization of the Anchorage #earthquake by @IRIS_EPO Lots of interesting info on their Twitter page.
Washington, USA	2018-12-11 (02 40 09)	#Earthquake #Breaking Earthquake M7.4 - #SouthSandwichIslands
Washington,DC	2018-12-11 (00 51 18)	Despite the scope of the #earthquake near Anchorage, Alaska, there have been zero reported fatalities.
Washington,DC	2018-12-05 (10 56 14)	We recently awarded two grants to support recovery from the 2017 Iran-Iraq #earthquake. Learn about our grantees.
Washington,DC	2018-12-10 (14 45 47)	30 YRS AGO TODAY M6.9 #earthquake struck Soviet Union's #Armenia. @USAID deployed a disaster team including urban search.
Washington,DC	2018-12-07 (20 13 53)	.@MiamiDadeFire you were with us for our very first USAR #earthquake deployment to #Armenia, helping us make history.
Washington,DC	2018-12-07 (19 27 08)	3 reasons why we remember our response to #Armenia #earthquake after it struck 30 years ago today.
Washington,DC	2018-12-10 (13 10 10)	in #Chile we are well prepare for the @amazon datacenter. #earthquake are not a problem
Seattle, Washington	2018-12-08 (20 43 32)	amazon A variety of #45vinyl #vinylrecords #vinyl #records #soundtrack @Disneyland #MaryPoppins #MaryPoppinsRetur.
Washington DC-Singapore	2018-12-07 (21 09 44)	Why no one died in Alaska's #earthquake last week a major quake in 1964 changed the way the state designed its cities.
Washington,DC	2018-12-09 (23 24 16)	After learning to love her on TV, everybody in LA trusts @DrLucyJones for #earthquake
Washington,DC	2018-12-06 (02 14 43)	First images of slope failures triggered by the M = 7.0 #Alaska #earthquake.
Washington DC	2018-12-08 (14 15 49)	My sister in Eagle River, #Alaska needs your help after this 7.0 #earthquake. She's a #singlemom who is trying to piece everyt
Washington DC	2018-12-08 (02 42 04)	If you were impacted by the #earthquake in #Anchorage, #Alaska stay informed and follow these steps to help you and
Washington DC	2018-12-06 (17 06 59)	Before, during and after a quake, there are things you can do to prepare and be safe.
Washington DC	2018-12-06 (17 06 44)	AFTER A QUAKE Pay attention to how you and your loved ones are experiencing and handling stress. Promote emotional recover
Washington DC	2018-12-07 (15 55 58)	""[Alaska] Gov. Bill Walker said sometimes people, including himself, grouse about stringent #buildingcodes.

(*continued on next page*)

Table A.1 (*continued*)

Location	Date/Time	Tweet
Washington DC	2018-12-06 (22 07 58)	Check out some financial problems people experience after a natural disaster how to avoid them
Washington DC	2018-12-06 (15 18 24)	Did you all feel a possible #earthquake maybe 10 min ago? #Virginia #NBCWashington
Washington DC	2018-12-04 (18 12 12)	#Anchorage soon hit by few major Shocks……. #earthquake @ReutersWorld @realDonaldTrump
Washington DC	2018-12-06 (11 35 39)	Undersea #earthquake triggers #tsunami warning near #NewCaledonia
Washington DC	2018-12-05 (21 10 48)	#Tsunami possible within 1000 km (600 miles) of epicenter of #NewCaledonia #earthquake
Washington DC	2018-12-06 (00 33 15)	Nearly 1400 #Aftershocks Rattle #Alaska After 7.0 #Earthquake
Washington DC	2018-12-10 (19 10 14)	Skeletons found in a #Roman house in Cyprus, by David&Noel Soren in the mid 1980
Washington DC	2018-12-05 (20 19 24)	'Scientists are looking to #AI to help improve predictions for #earthquake aftershocks
Washington DC	2018-12-04 (18 17 12)	Wow And with only 6 h of daylight. #Earthquake #Alaska # nce #Rebuilding
Washington DC	2018-12-05 (06 08 28)	#BREAKING Third #earthquake in 15 min, preliminary 5.9, strikes off New Caledonia. #Tsunami warning remains in effe
Seattle Washington USA	2018-12-05 (03 50 18)	Alaska assesses damage following the magnitude 7.0 #earthquake
Whatcom County Washington	2018-12-05 (02 37 01)	Anchorage #earthquake #quake leaves its #library a bit unbound
Washington DC	2018-12-04 (23 40 21)	#Terrifying moment family flee 7.0 Magnitude #earthquake in #Alaska via @MailOnline
Washington DC	2018-12-04 (22 50 43)	The scene after a powerful #earthquake strikes near Anchorage, Alaska.
Washington DC	2018-12-04 (11 40 17)	Praying for the safety of those affected by the #Earthquake in Anchorage, Alaska. @ENERGY is monitoring the situation
Seattle Washington USA	2018-12-04 (06 44 19)	#NEW 2.5 #earthquake struck near Mount Vernon at 2 11p.m.
Washington DC	2018-12-03 (23 35 13)	For an overview on the structural damage caused by the #earthquake and aftershocks in #Alaska

Table A.2

Sentiment analysis of the Washington tweets.

Tweets	Sentiments
Hot Spots in West Coast US Chances of strong earthquakes and volcanoes have gone up Mr Trump	−1
TIME is running out on the West Coast of America earthquake forecast terremoto tsunami	−1
DonaldTrump is responsible for making the Slow Slip in California worse and a man made disaster that is going	−1
Where are the investments in SOLAR energy Or the Truth about Animal Agriculture Time is running	1
#Anchorage soon hit by few major Shocks earthquake ReutersWorld realDonaldTrump POTUS USATODAY Ignorance	−1
This is another ground motion visualization showing the motion of the ground recorded by the USArray during the Anchorage	1
waves travel through the Earth at approximately 1/2 the speed of P waves Stations close to the earthquake record strong	−1
How do we capture the motion of an earthquake	−1
What is a seismic shadow zone and what can it teach us about earth's interior structure	1
Moment Magnitude Explained Happened to the Richter Scale	0
TEACHERS TOOLBOX How Hard Does the Ground Shake During a Quake	0
Watch the waves from the Anchorage earthquake roll across the USArray seismic network in the lower	−1
Amazing computer simulation from Dr Carl Tape shows a surface view of 3D earthquake waves from the M7.0 Anchorageearthquak	−1
What is an earthquake travel time curve How are they created	−1
Here's a great visual to help show the difference between a M70 earthquake and a M9 Watch the video	−1
How are seismic waves in the Earth like a drop of water falling into a pool	−1
Whats the difference between a M70 and M90 earthquake Find out using pasta noodles in this crazy cool video	−1
Fascinating visualization of the Anchorage earthquake by IRISEPO Lots of interesting info on their Twitter page	1
Earthquake Breaking Earthquake M74 SouthSandwichIslands	−1
Despite the scope of the earthquake near Anchorage Alaska there have been zero reported fatalities	1
We recently awarded two grants to support recovery from the 2017 IranIraq earthquake Learn about our grantees	1
30 YRS AGO TODAY M69 earthquake struck Soviet Unions Armenia USAID deployed a disaster team including urban search	−1
MiamiDadeFire you were with us for our very first USAR earthquake deployment to Armenia helping us make history	−1
3 reasons why we remember our response to Armenia earthquake after it struck 30 years ago today	−1
in Chile we are well prepare for the amazon datacenter earthquake are not a problem	1
amazon A variety of 45vinyl vinylrecords vinyl records soundtrack Disneyland MaryPoppins MaryPoppinsRetur	0
Why no one died in Alaskas earthquake last week a major quake in 1964 changed the way the state designed its cities	1
After learning to love her on TV everybody in LA trusts DrLucyJones for earthquake	1
First images of slope failures triggered by the M70 Alaska earthquake	−1
My sister in Eagle RiverAlaska needs your help after this 7.0 earthquake Shes a singlemom who is trying to piece everyt	−1
If you were impacted by the earthquake in Anchorage Alaska stay informed and follow these steps to help you and your love	1

Table A.2 (*continued*)

Tweets	Sentiments
Before during and after a quake there are things you can do to prepare and be safe	1
AFTER A QUAKE Pay attention to how you and your loved ones are experiencing and handling stress Promote emotional recover	1
Alaska Gov Bill Walker said sometimes people including himself grouse about stringent buildingcodes	−1
Check out some financial problems people experience after a natural disaster how to avoid them	1
Did you all feel a possible earthquake maybe 10 min ago Virginia NBCWashington	−1
Anchorage soon hit by few major Shocks earthquake ReutersWorld realDonaldTrump	−1
Undersea earthquake triggers tsunami warning near NewCaledonia	−1
Tsunami possible within 1000 km 600 miles of epicenter of NewCaledonia earthquake	−1
Nearly 1400 Aftershocks Rattle Alaska After 7.0 Earthquake	−1
Skeletons found in a Roman house in Cyprus by Davidamp Noel Soren in the mid 1980	0
Scientists are looking to AI to help improve predictions for earthquake aftershocks	−1
Wow And with only 6 h of daylight Earthquake Alaska nce Rebuilding	−1
BREAKING Third earthquake in 15 min preliminary 5.9 strikes off New Caledonia Tsunami warning remains in effe	−1
Alaska assesses damage following the magnitude 70 earthquake	−1
Anchorage earthquake quake leaves its library a bit unbound	−1
Terrifying moment family flee 70 Magnitude earthquake in Alaska via MailOnline	−1
The scene after a powerful earthquake strikes near Anchorage Alaska	1
Praying for the safety of those affected by the Earthquake in Anchorage Alaska ENERGY is monitoring the situation	1
NEW 2.5 earthquake struck near Mount Vernon at 2 11 pm	−1
For an overview on the structural damage caused by the earthquake and aftershocks in Alaska	−1

Appendix B

See

Table B.1

Tweets from various locations for first comparison.

Location	Date/Time	Tweet
San-Francisco	2018-12-10 12:33:26	@KPIXtv: Did you feel it? There was a magnitude 2.6 #earthquake near #Crockett that shook at 4:15AM.
San-Francisco	2018-12-09 19:48:00	@abc7newsbayarea: A magnitude 4.8 #earthquake hit near Anchorage just over a week after a larger quake damaged roads and bridges in the
San-Francisco	2018-12-07 19:50:48	@karaswisher Somewhat related: What happens when the grid goes offline for an extended period of time due to some
San-Francisco	2018-12-07 09:22:12	@elebrundamiens: Just arrived in #Alaska, impressed by how fast #Anchorage recovered from the 7.0 #earthquake of last Friday.
San-Francisco	2018-12-06 23:50:50	@NancyPelosi: Paul & I are praying for everyone impacted by the #earthquake in Alaska. Extremely thankful for all the first responders
San-Francisco	2018-12-06 19:38:30	@rabidmarmot: Public hearing on a mandatory seismic retrofitting policy on multi-unit housing in the City of @Oakland #resilience #earth
San-Francisco	2018-12-06 03:41:29	@DrDGrossman @davequast Hey Dave, when THOUSANDS of ppl die on the West Coast it will be YOUR fault for allowing
San-Francisco	2018-12-06 00:12:40	@kron4news: #Earthquake rattles California's Mojave Desert
San-Francisco	2018-12-05 09:08:14	@QuakesSF: 1.9 magnitude #earthquake. 11 mi from Livermore, #CA, United States
Denver	2018-12-11 16:46:49	@AshtonCBS4: Here are the details on the 2 earthquakes that woke up a lot of people near Glenwood Springs early this morning. No reports
Denver	2018-12-11 16:28:11	@ChrisCBS4: Two Early Morning Earthquakes Strike Glenwood Springs #COwx #4wx #Earthquake
Denver	2018-12-11 16:19:22	@BrittMorenoTV: Anyone shake this morning? @VisitGlenwood #earthquake #Colorado
Denver	2018-12-11 15:17:08	@9NEWS think that part of Colorado had been Californiaized yet #earthquake in the Rockies Glenwood Sorings'
Denver	2018-12-11 04:45:49	@CharleyRTaylor: A 7.3(downgraded to a 7.1)Magnitude Earthquake struck the South Sandwich Islands region at 9:26:32 pm EST.
Denver	2018-12-04 02:45:11	@spatialkey: Thanks to Alaska\'s stringent building codes, claims from Friday\'s #earthquake are mostly due to ""minor structural impacts""
Denver	2018-12-03 23:35:13	@westgov: For an overview on the structural damage caused by the #earthquake and aftershocks in #Alaska, read this by @GOVERNING
Oregon	2018-12-11 14:11:05	@Tonkala79: Two earthquakes an hour apart with another aftershock just after the 2nd. #earthquake #glenwoodsprings #Colorado'
Oregon	2018-12-06 04:50:24	@MultCoSO Fail to confirm w hold on, 3rd step. #alaskaearthquake spaghetti legs, duck and cover recommended.
Oregon	2018-12-05 20:57:45	@ChrisPietsch: Preparing for the worst: Experts recommend what to pack in your emergency kit. With story by @DylanJDarling
Oregon	2018-12-05 05:21:35	@OregonGovBrown Yo Yo.. Oregon you dont get a free pass!! Every State needs to do their due diligence in protect

Table B.1 (*continued*)

Location	Date/Time	Tweet
Oregon	2018-12-11 15:00:37	b'7.0 #earthquake hits Alaska, damages 2 hospitals
Oregon	2018-12-11 12:20:47	#earthquake #Messiah @LouisFarrakhan warns ""And to all of you #spiritual teachers, you #religious scientists:
Chicago	2018-12-11 02:23:14	@TheJennyLam: The ruins of Wuchang #Temple @ the epicenter of the 1999 #earthquake, which was 7.6–7.7 in magnitude:
Chicago	2018-12-10 02:30:36	@sonicviviana: so who feeling on ok all wrong help mind city kill stop time just who love sad tdendy? soon all Page 1–7 earth more.
Chicago	2018-12-09 21:46:40	Tune in for earthquake forecast tonight #earthquake #temblor #sismo #terremoto #japan #california #mexico #peru
California	2018-12-11 04:59:50	@UrgentNews911: @Viviurway UPDATE: No #Tsunami threat to #Australia from Mag. 7.3 #earthquake near South Sandwich Islands (58.7S 26.5 W)
California	2018-12-10 15:05:14	@QuakesInCA: 2.3 magnitude #earthquake. 2 mi from #OrovilleEast, CA, United States
California	2018-12-09 23:24:16	@EatTheCrust @theAGU @theIUGG After learning to love her on TV, everybody in LA trusts @DrLucyJones for #earthquake
California	2018-12-09 10:09:13	@istockhistory: Earthquake-risk maps pinpoint world most vulnerable areas #nature #wildlife #wildlifephotograp\xe2\x80\xa6′
California	2018-12-08 20:56:38	@foretellings30: 7Magnitude to 8 Magnitude earthquake soon hit in West..Venezuela or adjacent country JAPAN Argentina Peru Indonesia
California	2018-12-05 19:24:10	""@Toecephus_ @siminuteman1776 Very curious. There's a whole lot of shaking going on… Everywhere #D5 #earthquake"""
California	2018-12-05 17:15:28	@BriannaRoaach: .@SAConScene A 7.5 magnitude #earthquake strikes in the Pacific Ocean creating a tsunami warning.
California	2018-12-05 03:51:23	@SutterHealth: Sutter CPMC\xe2\x80\x99s Van Ness Campus is the 1st in North America to use a new technology that helps absorb an earthquake
California	2018-12-05 06:40:28	@psybrspcsuprstr: USGS	M 7.5\n 168 km ESE of Tadine, New Caledonia\nA #TSUNAMI THREAT EXISTS FOR PARTS OF THE PACIFIC LOCATED CLOSER
California	2018-12-05 06:36:10	@tiniskwerl: \ #Moments: #TsunamiWarning issued for #NewCaledonia #Vanuatu #Fiji #Earthquake Prelim M7.6 Rev M7.5
California	2018-12-05 04:33:42	@erikblair: Earthquake M7.6: Tadine, New Caledonia #earthquake #tsunamis possible
California	2018-12-04 21:31:54	@SeismicFox1: The seismic activity in Southern California is beginning to heat up again so far. #earthquakes #SouthernCalifornia
California	2018-12-04 15:31:47	@DLJCSS: Earthquake expert urges San Diego to be more ambitious about replacing old pipes, building retrofits #earthquake
California	2018-12-04 12:36:37	@MalibuSAR: We live in #earthquake country - be prepared. Food, water, clothing, and a good pre-plan.
California	2018-12-04 06:58:11	@BradLindenSoCal: Tsunami warning canceled after major earthquake in Alaska. My heart goes out to the people of #Alaska after the #earthquake
California	2018-12-03 23:13:22	The Well Runs Dry
Ohio	2018-12-11 17:25:37	@RVin31849659: #Planet X #Nibiru is #moving in #USA. @TradCatKnight reports on: ""ANOTHER 7.0+ (#earthquake) QUAKE!
Ohio	2018-12-04 19:22:40	@LLacrosse: Combining #coding #earthsci #AR in #earthquake simulation! Featuring #BOLT & 7 of his @Sphero pals to demo rate
Michigan	2018-12-11 17:25:53	@elenochle: That was a BANGER! most of the time its just straight lines with little fluctuations
Michigan	2018-12-10 23:12:04	@MichaELofYHWH: In my book, @SignsOfTimes777, I write about why the powerful #earthquakes, which the other nations of the world have
Michigan	2018-12-04 20:19:24	@CarstensNarissa: @JVonfrederick @One_Of_His @ApostleEric Apostolic Stewardship Returned to the Church:
Texas	2018-12-11 04:08:42	@GodCountryFami6: #earthquake in #Antarctica \n\nBombing #rothchild island and the #evil
Texas	2018-12-10 07:04:38	@POETreeOTIC: Dang ~ How does he do it? One day he's targeted by the #Mueller midgets as a crim cuz #Trump contemplated building
Texas	2018-12-07 14:31:59	@SamusAran2020: @Teresa13218380 @CactusFlower_1 @mitchellvii It's possible. A warning issued. But I hope there is no tsunami
Texas	2018-12-06 09:29:42	treeofevo: This was released 8 years ago, but it shows the CAVE earthquake data from around the world over the past 110 years in map
Texas	2018-12-05 19:24:10	""@Toecephus_ @siminuteman1776 Very curious. There's a whole lot of shaking going on… Everywhere #D5 #earthquake
Texas	2018-12-05 17:49:27	@dannyrellisjr: Around 30 #earthquake aftershocks in #NewCaledonia all over 4.5 mag since the 7.5 yesterday.
Texas	2018-12-05 04:30:16	@morganterry3: #Earthquake prelim. 7.6 near #Tadine #NewCaledonia apron depth 6.2 miles. #prayers'
Texas	2018-12-04 11:43:53	@courtneyknorris: HHS Secretary Alex Azar has declared a public health emergency in Alaska due to the #earthquake damage
India	2018-12-11 14:58:21	@ibtimes_india: Massive 7.1 magnitude earthquake hits Antarctica, volcanic Bristol island hit #Antartica #Earthquake
India	2018-12-11 09:23:34	@khalsaaid_india: Kids enjoying their hot meals after #nepal #earthquake provided by #khalsaaid #langar #programme.
India	2018-12-08 09:21:19	@VagariesWeather @mumbaimatterz Last tremors in Palghar was Dec 1 (here it was twice in a span of ten minutes)
India	2018-12-05 04:54:46	@kolkata24x7: BREAKING: Tsunami Warning As Massive Quake Strikes Off New Caledonia #BreakingNews #Earthquake

Table B.1 (*continued*)

Location	Date/Time	Tweet
India	2018-12-11 02:40:09	@dawn_morning: #Earthquake #Breaking@lookner. @Marfoogle Earthquake M7.4 #SouthSandwichIslands
India	2018-12-09 15:35:39	narendramodi @dpradhanbjp Sir i kindly request take immediate action and solve this problem…if emergency step
India	2018-12-08 07:54:50	@IndianStandards: the Indian Standards related to disaster resilience which is relevant to Sikkim which falls under Zone IV
India	2018-12-07 04:21:35	@IjoydeepRoy: A magnitude 7.0 aftershock has struck off the coast of small town of #Tadine on the Loyalty Islands,#NewCaledonia
India	2018-12-05 18:19:12	@LogicalIndians: #India #Latur #Armyman #MiracleBaby #Earthquake ""The disaster from 25 years ago connected us in a bond that will last
India	2018-12-05 14:01:41	@AparnaPawar13: Tsunami waves detected in #NewCaledonia, Vanuatu after quake - Egypt Today\n#earthquake #tsunami
India	2018-12-05 06:44:00	@DDNewsLive: #Tsunami alert issued after powerful #earthquake in #Alaska
India	2018-12-05 04:42:42	@fx16news: #Earthquake measuring 7.3 magnitude strikes off New Caledonia; hazardous tsunami waves possible near quake epicentre
India	2018-12-05 01:06:34	@satyug20201: Powerful M7.0 #earthquake shakes #Alaska Think who can save from these types of natural disasters.
India	2018-12-04 21:14:40	#Warning bell for mega #Himalayan #earthquake grows louder. \n#Himalaya #HimachalPradesh
India	2018-12-04 11:07:52	@73buzz_news: After a giant crack in the earth tore through Kenya, reports emerged saying it proves Africa will split into two continen
India	2018-12-04 10:13:09	@GEAG_India: Warning bell for mega Himalayan #earthquake grows louder: Could cause never-seen-before destruction in #Uttarakhand #himach

Table B.2
Sentiment analysis for first comparison.

Tweets	Sentiments
Did you feel it There was a magnitude 2.6 earthquake near Crockett that shook at 4:15AM	0
A magnitude 4.8 earthquake hit near Anchorage just over a week after a larger quake damaged roads and bridges in the	−1
Somewhat related What happens when the grid goes offline for an extended period of time due to some	0
Just arrived in Alaska impressed by how fast Anchorage recovered from the 7.0 earthquake of last Friday	1
I are praying for everyone impacted by the earthquake in Alaska Extremely thankful for all the first responders	1
Public hearing on a mandatory seismic retrofitting policy on multi-unit housing in the City of Oakland resilience earthquake	1
Hey Dave when THOUSANDS of ppl die on the West Coast it will be YOUR fault for allowing	−1
Earthquake rattles California's Mojave Desert	−1
1.9 magnitude earthquake 11 mi from Livermore CA United States	0
Here are the details on the 2 earthquakes that woke up a lot of people near Glenwood Springs early this morning No reports	−1
Two Early Morning Earthquakes Strike Glenwood Springs Earthquake	−1
Anyone shake this morning VisitGlenwood earthquake Colorado	0
think that part of Colorado had been Californiaized yet earthquake in the Rockies Glenwood Sorings	−1
7.3 downgraded to a 7.1Magnitude Earthquake struck the South Sandwich Islands region at 9:26:32pm EST	−1
Thanks to Alaska's stringent building codes claims from Friday earthquake are mostly due to minor structural impacts	1
For an overview on the structural damage caused by the earthquake and aftershocks in Alaska read this by GOVERNING	−1
Two earthquakes an hour apart with another aftershock just after the 2nd earthquake glenwoodsprings Colorado	−1
Fail to confirm w hold on 3rd step alaskaearthquake spaghetti legs duck and cover recommended	−1
Preparing for the worst Experts recommend what to pack in your emergency kit With story	−1
Yo Yo Oregon you dont get a free pass Every State needs to do their due diligence in protect	−1
earthquake hits Alaska damages 2 hospitals	−1
earthquake Messiah And to all of you spiritual teachers you religious scientists	0
The ruins of Wuchang Temple the epicenter of the 1999 earthquake which was 7.6–7.7 in magnitude	−1
so who feeling on ok all wrong help mind city kill stop time just who love sad tdendy soon all Page 1–7 earth more	−1
Tune in for earthquake forecast tonight earthquake temblor sismo terremoto japan california mexico peru	−1
No Tsunami threat to Australia from Mag. 7.3 earthquake near South Sandwich Islands (58.7S 26.5W)	1
2.3 magnitude earthquake 2 mi from OrovilleEast CA United States	0
After learning to love her on TV everybody in LA trusts DrLucyJones for earthquake	1
Earthquake-risk maps pinpoint world most vulnerable areas nature wildlife wildlifephotograp	0
7Magnitude to 8 Magnitude earthquake soon hit in West Venezuela or adjacent country JAPAN Argentina Peru Indonesia	−1
Very curiousThere's a whole lot of shaking going on Everywhere D5 earthquake	−1
A 7.5 magnitude earthquake strikes in the Pacific Ocean creating a tsunami warning	−1
Van Ness Campus is the 1st in North America to use a new technology that helps absorb an earthquake	1
7.5 168km ESE of Tadine New Caledonia A TSUNAMI THREAT EXISTS FOR PARTS OF THE PACIFIC LOCATED CLOSER	−1
Moments TsunamiWarning issued for NewCaledonia Vanuatu Fiji Earthquake Prelim M7.6 Rev M7.5	−1
Earthquake M7.6 Tadine New Caledonia earthquake tsunamis possible	−1
The seismic activity in Southern California is beginning to heat up again so far earthquakes SouthernCalifornia	−1
Earthquake expert urges San Diego to be more ambitious about replacing old pipes building retrofits earthquake	−1
We live in earthquake country be prepared Food water clothing and a good pre-plan	1
Tsunami warning canceled after major earthquake in Alaska My heart goes out to the people of Alaska after the earthquake	1
The Well Runs Dry	−1
Planet X Nibiru is moving in USA TradCatKnight reports on ANOTHER 7.0 earthquake QUAKE	−1

(continued on next page)

Table B.2 (*continued*)

Tweets	Sentiments
Combining coding earthsci AR in earthquake simulation Featuring BOLT amp 7 of his Sphero pals to demo rate	1
That was a BANGER most of the time its just straight lines with little fluctuations	−1
In my book I write about why the powerful earthquakes which the other nations of the world have	0
Apostolic Stewardship Returned to the Church	0
earthquake in Antarctica Bombing rothchild island and the evil	−1
Dang How does he do it One day he's targeted by the Mueller midgets as a crim cuz Trump contemplated building	1
It's possible A warning issued But I hope there is no tsunami	−1
This was released 8 years ago but it shows the CAVE earthquake data from around the world over the past 110 years in map	−1
Very curious There's a whole lot of shaking going on Everywhere D5 earthquake	−1
Around 30 earthquake aftershocks in NewCaledonia all over 4.5 mag since the 7.5 yesterday	−1
Earthquake prelim 7.6 near Tadine NewCaledonia apron depth 6.2 miles prayers	1
HHS Secretary Alex Azar has declared a public health emergency in Alaska due to the earthquake damage	1
Massive 7.1 magnitude earthquake hits Antarctica volcanic Bristol island hit Antartica Earthquake	−1
Kids enjoying their hot meals after nepal earthquake provided by khalsaaid #angar programme.	1
Last tremors in Palghar was Dec 1 here it was twice in a span of ten minutes	−1
Tsunami Warning As Massive Quake Strikes Off New Caledonia BreakingNews Earthquake	−1
Earthquake Breaking Earthquake M7.4 SouthSandwichIslands	0
Sir i kindly request take immediate action and solve this problem if emergency step	−1
the Indian Standards related to disaster resilience which is relevant to Sikkim which falls under Zone IV	−1
A magnitude 7.0 aftershock has struck off the coast of small town of Tadine on the Loyalty Islands NewCaledonia	0
India Latur Armyman MiracleBaby Earthquake The disaster from 25 years ago connected us in a bond that will last	−1
Tsunami waves detected in NewCaledonia Vanuatu after quake Egypt Today earthquake tsunami	−1
Tsunami alert issued after powerful earthquake in Alaska	−1
Earthquake measuring 7.3 magnitude strikes off New Caledonia hazardous tsunami waves possible near quake epicentre	−1
Powerful M7.0 earthquake shakes Alaska Think who can save from these types of natural disasters	−1
Warning bell for mega Himalayan earthquake grows louder Himalaya HimachalPradesh	−1
After a giant crack in the earth tore through Kenya reports emerged saying it proves Africa will split into two continen	−1
Warning bell for mega Himalayan earthquake grows louder Could cause never seen before destruction in Uttarakhand himach	−1

Appendix C

See Table C.1 and C.2.

Table C1
"Geo-tagged tweets without name" for second comparison.

Location	Date/Time	Tweet
Geo-tagged without name	2018-12-11 16:10:35	@TheWatchers_: M7.1 earthquake hits South Sandwich Islands at intermediate depth earthquake #SouthSandwichIslan
Geo-tagged without name	2018-12-11 15:04:03	@WFP_ADAM: #Earthquake in #South Georgia and the South Sandwich Islands (14 m ago). 61 km NNE of Bristol Island. #Magnitude: 7.3
Geo-tagged without name	2018-12-04 18:14:53	@GB_Survival: The survival essentials you must have for under $50 total. Plus FREE Apps. #DisasterPreparedness
Geo-tagged without name	2018-12-04 03:25:56	@GB_Survival: 60% of people have not practiced or prepared for what to do in an emergency. Why Prepare Now? Because it's too late once
Geo-tagged without name	2018-12-11 13:31:17	@NERC_COMET: @ruthamey presents her work on a Fractal Future for #Earthquake Slip Inversions, make your way to Marriott Marquis
Geo-tagged without name	2018-12-11 13:02:03	@curmarshall: Natural Resources Canada says there's still not complete certainty the earthquake was caused by
Geo-tagged without name	2018-12-11 12:01:38	@NewsAndStocks: #BREAKINGNEWS: Magnitude 5.2 #earthquake strikes, #Molibagu, #Indonesia @USGS reports. @USGS says ""Did you feel it?
Geo-tagged without name	2018-12-11 04:16:58	@EconomicTimes: An #earthquake of magnitude 7.5 hits close to #Antarctica: United States Geological Survey
Geo-tagged without name	2018-12-04 11:41:45	@JaimeNelsonW: Some 1800 #aftershocks have been measured since Friday 7.0 #earthquake in #Alaska
Geo-tagged without name	2018-12-09 18:00:44	@NESTA_US: Satellite view of eruption Hawaii Kilauea volcano seen from @NASA satellites localitiesTo better orientation, scale
Geo-tagged without name	2018-12-11 04:35:03	@Clearing_Barrel: #seismologists of the world, I have a question. #earthquake #qanon What is causing this? This can't be normal or natural
Geo-tagged without name	2018-12-07 23:02:49	@mercenarygeo: M7.5 #earthquake 168 km ESE of Tadine, New Caledonia #USGS #Geology
Geo-tagged without name	2018-12-10 14:45:47	@theOFDA: 30 YRS AGO TODAY: M6.9 #earthquake struck Soviet Union's #Armenia. @USAID deployed a disaster team including urban search
Geo-tagged without name	2018-12-08 17:37:14	@USAID: When did USAID deploy its first Urban Search and Rescue Team? 30 years ago when a M6.9 #earthquake struck Soviet Union's #Armenia
Geo-tagged without name	2018-12-05 14:33:57	@simonealiprandi: Today in Pavia @GEMwrld is presenting the global #earthquake hazard and risk maps
Geo-tagged without name	2018-12-10 13:11:15	@MunichRe: A milestone for #earthquake risk assessment: The @GEMwrld global earthquake risk map, hot off the press.
Geo-tagged without name	2018-12-04 04:10:00	@Medikak: #AlaskaQuake #Alaska #earthquake this was a freaky one

Table C1 (*continued*)

Location	Date/Time	Tweet
Geo-tagged without name	2018-12-09 20:19:38	@SatishKTM: #Earthquake hazard and risk as shown in the recently published Global Earthquake Model (#GEM)\ #nepal
Geo-tagged without name	2018-12-06 23:42:25	@SismoDetector: #earthquake reported by the users of the app Earthquake Network at 8 km from Anchorage, UnitedSt
Geo-tagged without name	2018-12-09 12:06:56	Electric Earthquakes in Alaska and Global Ringing Sounds gotminiiceage #miniiceage #littleiceage
Geo-tagged without name	2018-12-09 13:41:52	@KafuiDey: @NADMOGHANA How about some safety tips so we know what to do in the event of an #earthquake, #flooding or other #disaster?
Geo-tagged without name	2018-12-08 17:19:32	@Helperbit: #Earthquake in #Alaska, USA: 11 km far from #Anchorage. Magnitude: 7 Depth: 40 km Affected people: 255.000
Geo-tagged without name	2018-12-08 16:17:26	@mother_of_hera: Seismologists are saying us in #Alaska will be feeling #aftershocks for months… even years. Some of these have been
Geo-tagged without name	2018-12-08 15:24:25	#Earthquake Damage to #USA Military Infrastructure in #Alaska Revealed: @SputnikInt
Geo-tagged without name	2018-12-08 14:06:27	@the9numbers: MYSTERIOUS WAVES: can you imagine something moving so fast that it takes just 20 min to move half way round the world
Geo-tagged without name	2018-12-07 22:54:23	@gem_risk: Murray Journey from @NRCan demonstrates an excellent example of how #earthquake #risk assessment in #Canada is supporting
Geo-tagged without name	2018-12-07 20:40:09	@FMGlobal: A new report found that San Francisco high rises may not be fortified for #earthquake #resilience should there be another
Geo-tagged without name	2018-12-07 14:01:37	@JAMnewsCaucasus: #Armenia marks one of the saddest days in its modern history today, when the #Spitak #earthquake of 1988 took the live
Geo-tagged without name	2018-12-07 12:34:21	@shanonicity: is CERN really shut down #CERN #Earthquake
Geo-tagged without name	2018-12-07 12:18:21	@UlrikeZiemer: Still Recovering: #Armenia's Catastrophic #Earthquake, 30 Years Later
Geo-tagged without name	2018-12-07 09:58:17	@Shairi_M: @mofaganepal kicks off national #preparedness workshop on #disaster #recovery @NRANepal presentation on status of #post #disaster
Geo-tagged without name	2018-12-07 08:27:24	@RyanHobbsMMA: A drone shot of some more damage from the massive #earthquake in #Alaska #AlaskaEarthquake
Geo-tagged without name	2018-12-07 06:19:23	cassieschirmtv: KTVA newsroom felt the blow of the earthquake this morning. #anchorage #alaska #earthquake #weather
Geo-tagged without name	2018-12-06 21:52:50	cassieschirmtv: Just felt two very large aftershocks! My entire house shook. My poor Coco is still having PTSD from the big earthquake.
Geo-tagged without name	2018-12-06 21:49:19	@ApocalypticAct1: You are encouraged to file claims of damage to your property from an #earthquake even if you don't have earthquake
Geo-tagged without name	2018-12-06 17:52:15	@CITechRPM: #Geology #Environment: The #Earth grumbled! An #Earthquake Swarm!
Geo-tagged without name	2018-12-06 14:35:53	@theweatherjet: Alaska hit by more than 190 small earthquakes since Friday #news #weather #earthquake #worldnews #alert #alaska
Geo-tagged without name	2018-12-06 13:11:34	@DubesDenmark: The team from the Danish Emergency Management Agency (@BRSdk) has just returned to #Denmark after two-months in #Palu
Geo-tagged without name	2018-12-05 08:35:38	@MariettaPosts: 7.0 Earthquake This Morning In Alaska That Lasted 90 Seconds. Fortunately, No Injuries Have Been Reported but Heartbreak
Geo-tagged without name	2018-12-05 07:19:46	@MariettaPosts: Alaska Governor Has Just Declared The State a Disaster After 7.0 Earthquake Near Anchorage With 9 Aftershocks.
Geo-tagged without name	2018-12-06 08:49:48	#Earthquake risk in Delhi is high, but you can reduce it! As SEEDER Aakash explains @Honeywell_India CSR symposium
Geo-tagged without name	2018-12-06 08:42:47	@PMBreakingNews That #earthquake seeming as being of a #disastrous #magnitude.
Geo-tagged without name	2018-12-06 05:12:16	@ABC7: #EARTHQUAKE 4.2 magnitude quake hits Mojave, north of Barstow. Did you feel it?
Geo-tagged without name	2018-12-06 02:33:28	@paleoseismicity: Conference \xe2\x80\x9c30 years after #Spitak #Earthquake: Experience and Perspectives 3–7 Dec 2018, #Yerevan, #Armenia
Geo-tagged without name	2018-12-04 22:22:01	@EcoWatch: \xf0\x9f\x98\xa6After Alaska's 7.0 #earthquake 2 patients have life-threatening injuries as of Monday afternoon along with dozens of patient
Geo-tagged without name	2018-12-05 22:19:50	@outremers360 This is the evidence of submarine #nuclear #fracking!\Offshore or onshore fracking always means nuclear
Geo-tagged without name	2018-12-05 06:49:22	@MaKaElectric:7.5 Mag Quake Near New Caledonia Prompts #Tsunami Alert For Far Pacific
Geo-tagged without name	2018-12-05 14:31:00	@lifewithweather: Newly released footage inside Alaska classroom shows REAL scene of what it like during an earthquake
Geo-tagged without name	2018-12-05 14:15:47	@3Sjapan: A big #earthquake struck near New Caledonia. This is a sign of major earthquake in Japan soon.#earthquake #Japan #newcaledonia'
Geo-tagged without name	2018-12-05 11:50:35	@RT_com: #Tsunami possible within 1000 km (600 miles) of epicenter of #NewCaledonia #earthquake
Geo-tagged without name	2018-12-05 08:55:09	@NRATV: ""In a natural disaster you may call 911, and there may be nobody on the other line. Or they may tell you, \'Sorry,
Geo-tagged without name	2018-12-05 08:30:07	@CloseTheBase: Veterans for #Peace calls for US govt's GAO to reevaluate risks of landfill/construction at #Henoko #Okinawa #coralreef
Geo-tagged without name	2018-12-05 07:33:08	@dlitchfield: New Caledonia is having a pretty rough day
Geo-tagged without name	2018-12-05 07:18:17	@pennewstweet: 7.6 Magnitude earthquake jolts New Caledonia; Tsunami alerts issued #StrongEarthquake #Jolts #NewCaledonia #Tsunami
Geo-tagged without name	2018-12-05 05:17:56	@OrionMicroApps: New Caledonia M7.5 A TSUNAMI THREAT EXISTS FOR PARTS OF THE PACIFIC LOCATED CLOSER TO THE EARTHQUAKE.

Table C1 (*continued*)

Location	Date/Time	Tweet
Geo-tagged without name	2018-12-05 04:47:47	@edwardpeguero1: Big earthquake tune in for earthquake forecast #earthquake #temblor #sismo #terremoto #japan #california #mexico #peru
Geo-tagged without name	2018-12-05 05:14:57	@wordsandunwords: California Has Four Seasons Earthquake Fire Flood Throw Pillow
Geo-tagged without name	2018-12-05 05:05:17	@Terra_Miel: The Pacific Tsunami Warning Centre said tsunami waves could hit areas close to the French territory, which is about 1400 km
Geo-tagged without name	2018-12-05 01:20:41	@PasadenaGov: Thank you @USGS_ShakeAlert for giving emergency managers and science educators an update on the #earthquake early warning
Geo-tagged without name	2018-12-05 00:49:43	@SEIDirector: Say Thank You to your House and Senate Representatives for Reauthorizing NEHRB (National Earthquake Hazards Reduction Programme
Geo-tagged without name	2018-12-04 23:45:03	@DellEMCSLG: #Anchorage Counts Its Blessings After Dodging a Significant Catastrophe #Infrastructure #earthquake
Geo-tagged without name	2018-12-04 21:23:27	@KateInReallLife: Fun Fact: did you know that since Friday\xe2\x80\x99s 7.0 earthquake, we have had approx. 2100 aftershocks.
Geo-tagged without name	2018-12-04 17:44:39	@joinrocean: Strange waves rippled around the world, and nobody knows why: #earthquake #waves
Geo-tagged without name	2018-12-04 17:06:09	@cleanplanet_: Join us and be rewarded in #STEEM @realnedscott @planetenamek @andrarchy @VinnieLemon @NathanM
Geo-tagged without name	2018-12-04 13:02:39	@kaelbama: Fascinating visualization of the Anchorage #earthquake by @IRIS_EPO. Lots of interesting info on their Twitter page.
Geo-tagged without name	2018-12-04 08:39:39	@GeospatialNews1: Trimble Introduces Satellite-Based RTX Corrections to the Kestrel Seismogeodetic System for Earthquake, Volcano
Geo-tagged without name	2018-12-04 07:53:24	@blogruivismo: Alasca (2018) – 7.2 magnitude earthquake #alaska #earthquake #earthquakealasca #terremotoalaska #earthquakeus
Geo-tagged without name	2018-12-04 06:47:00	@tyrabanks I\xe2\x80\x99ve just got to say Friday I was in #anchoragealaska when a 7.2 #earthquake hit and lastnight I watched
Geo-tagged without name	2018-12-04 05:38:56	@sitecproject: After a full site inspection, we can confirm there was no damage to any structures or equipment as a result of yesterday
Geo-tagged without name	2018-12-04 04:37:04	The #USGS #Earthquake Hazards Program explained via @OpenAccessGov

Table C2

Sentiment analysis for second comparison.

Tweets	Sentiments
M 7.1 earthquake hits South Sandwich Islands at intermediate depth earthquake SouthSandwich Island	−1
Earthquake in South Georgia and the South Sandwich Islands 14 m ago 61 km NNE of Bristol Island Magnitude 7.3	0
The survival essentials you must have for under $50 total Plus FREE Apps Disaster Preparedness	−1
60% of people have not practiced or prepared for what to do in an emergency Why Prepare Now Because it s too late once	−1
ruthamey presents her work on a Fractal Future for Earthquake Slip Inversions make your way to Marriott Marquis	−1
Natural Resources Canada says there's still not complete certainty the earthquake was caused by	−1
BREAKINGNEWS Magnitude 5.2 earthquake strikes Molibagu Indonesia Did you feel it	0
An earthquake of magnitude 7.5 hits close to Antarctica United States Geological Survey	0
Some 1800 aftershocks have been measured since Friday 7.0 earthquake in Alaska	−1
Satellite view of eruption Hawaii Kilauea volcano seen from NASA satellites localities To better orientation scale	−1
seismologists of the world I have a question earthquake qanon What is causing this This can't be normal or natural	−1
M 7.5 earthquake 168 km ESE of Tadine New Caledonia USGS Geology	0
30 YRS AGO TODAYM 6.9 earthquake struck Soviet Union's Armenia USAID deployed a disaster team including urban search	−1
When did USAID deploy its first Urban Search and Rescue Team 30 years ago when a M 6.9 earthquake struck Soviet Union's Armenia	1
Today in Pavia is presenting the global earthquake hazard and risk maps	1
A milestone for earthquake risk assessment The global earthquake risk map hot off the press	1
AlaskaQuake Alaska earthquake this was a freaky one	−1
Earthquake hazard and risk as shown in the recently published Global Earthquake Model nepal	−1
earthquake reported by the users of the app Earthquake Network at 8 km from Anchorage UnitedSt	0
Electric Earthquakes in Alaska and Global Ringing Sounds gotminiiceage miniiceage littleiceage	−1
How about some safety tips so we know what to do in the event of an earthquake flooding or other disaster	1
Earthquake in Alaska USA 11 km far from Anchorage Magnitude 7 Depth 40 km Affected people 255.000	−1
Seismologists are saying us in Alaska will be feeling aftershocks for months even years Some of these have been	−1
Earthquake Damage to USA Military Infrastructure in Alaska Revealed	−1
MYSTERIOUS WAVES can you imagine something moving so fast that it takes just 20 min to move half way round the world	−1
Can demonstrates an excellent example of how earthquake risk assessment in Canada is supporting	0
A new report found that San Francisco high rises may not be fortified for earthquake resilience should there be another	−1
Armenia marks one of the saddest days in its modern history today when the Spitak earthquake of 1988 took the live	−1
is CERN really shut down CERN Earthquake	−1
Still Recovering Armenia's Catastrophic Earthquake 30 Years Later	−1
nepal kicks off national preparedness workshop on disaster recovery NRANepal presentation on status of post disaster	1
A drone shot of some more damage from the massive earthquake in Alaska AlaskaEarthquake	−1
KTVA newsroom felt the blow of the earthquake this morning anchorage alaska earthquake weather	−1
Just felt two very large aftershocks My entire house shook My poor Coco is still having PTSD from the big earthquake	−1
You are encouraged to file claims of damage to your property from an earthquake even if you don't have earthquake	1
Geology Environment The Earth grumbled An Earthquake Swarm	−1

Table C2 (*continued*)

Tweets	Sentiments
Alaska hit by more than 190 small earthquakes since Friday news weather earthquake worldnews alert alaska	−1
The team from the Danish Emergency Management Agency has just returned to Denmark after two months in Palu	1
7.0 Earthquake This Morning In Alaska That Lasted 90 Seconds Fortunately No Injuries Have Been Reported but Heartbreak	−1
Alaska Governor Has Just Declared The State a Disaster After 7.0 Earthquake Near Anchorage With 9 Aftershocks.	−1
Earthquake risk in Delhi is high but you can reduce it As SEEDER Aakash explains CSR symposium	1
That earthquake seeming as being of a disastrous magnitude	−1
EARTHQUAKE 4.2 magnitude quake hits Mojave north of Barstow Did you feel it	0
Conference years after Spitak Earthquake Experience and Perspectives 3 7 Dec 2018 Yerevan Armenia	1
After Alaska's 7.0 earthquake 2 patients have life threatening injuries as of Monday afternoon along with dozens of patient	−1
This is the evidence of submarine nuclear fracking Offshore or onshore fracking always means nuclear	−1
7.5 Mag Quake Near New Caledonia Prompts Tsunami Alert For Far Pacific	−1
Newly released footage inside Alaska classroom shows REAL scene of what it like during an earthquake	0
A big earthquake struck near New Caledonia This is a sign of major earthquake in Japan soon earthquake Japan newcaledonia	−1
Tsunami possible within 1000 km 600 miles of epicenter of NewCaledonia earthquake	−1
In a natural disaster you may call 911 and there may be nobody on the other line Or they may tell you Sorry	−1
Veterans for Peace calls for US govt's GAO to reevaluate risks of landfill construction at Henoko Okinawa coralreef	1
New Caledonia is having a pretty rough day	−1
7.6 Magnitude earthquake jolts New Caledonia Tsunami alerts issued Strong Earthquake Jolts NewCaledonia Tsunami	−1
New Caledonia M 7.5 A TSUNAMI THREAT EXISTS FOR PARTS OF THE PACIFIC LOCATED CLOSER TO THE EARTHQUAKE	−1
Big earthquake tune in for earthquake forecast earthquake temblor sismo terremoto japan california mexico peru	−1
California Has Four Seasons Earthquake Fire Flood Throw Pillow	−1
The Pacific Tsunami Warning Centre said tsunami waves could hit areas close to the French territory which is about 1400 km	−1
Thank you for giving emergency managers and science educators an update on the earthquake early warning	1
Say Thank You to your House and Senate Representatives for Reauthorizing NEHRB National Earthquake Hazards Reduction Programme	1
Anchorage Counts Its Blessings After Dodging a Significant Catastrophe infrastructure earthquake	1
Fun Fact did you know that since Friday 7.0 earthquake we have had approx. 2100 aftershocks	−1
Strange waves rippled around the world and nobody knows why earthquake waves	−1
Join us and be rewarded in STEEM	1
Fascinating visualization of the Anchorage earthquake Lots of interesting info on their Twitter page	1
Trimble Introduces Satellite Based RTX Corrections to the Kestrel Seismogeodetic System for Earthquake Volcano	1
Alasca 2018 7.2 magnitude earthquake alaska earthquake earthquakealasca terremotoalaska earthquakeus	0
just got to say Friday I was in anchoragealaska when a 7.2 earthquake hit and last night I watched	0
After a full site inspection we can confirm there was no damage to any structures or equipment as a result of yesterday	1
The USGS Earthquake Hazards Program explained via	0

References

[1] M. Tsou, et al., Building a real-time geo-targeted event observation (geo) viewer for disaster management and situation awareness, Lect. Notes Geoinf. Cartogr. (2017) 85–98, https://doi.org/10.1007/978-3-642-19214-2.

[2] C. Zhang, et al., GeoBurst+: effective and real-time local event detection in geo-tagged tweet streams, ACM Trans. Intell. Syst. Technol. Artic. 9(34) (2018) https://doi.org/10.1145/3066166.

[3] D. Murthy, A.J. Gross, Social media processes in disasters: implications of emergent technology use, Soc. Sci. Res. 63 (2017) 356–370, https://doi.org/10.1016/j.ssresearch.2016.09.015.

[4] A. Bhuvaneswari, C. Valliyammai, Information entropy based event detection during disaster in cyber-social networks, J. Intell. Fuzzy Syst. 36 (5) (2019) 3981–3992, https://doi.org/10.3233/jifs-169959.

[5] C. Caragea, A. Squicciarini, S. Stehle, K. Neppalli, A. Tapia, Mapping moods: geo-mapped sentiment analysis during hurricane sandy, in: Proc. 11th Int. ISCRAM Conf., no. May, 2014, pp. 642–651 [Online]. Available: http://www.iscram.org/legacy/ISCRAM2014/papers/p29.pdf.

[6] J.D. Gonzalez Paule, Y. Moshfeghi, J.M. Jose, P. (Vonu) Thakuriah, On Fine-Grained Geolocalisation of Tweets, 2017, pp. 313–316 https://doi.org/10.1145/3121050.3121104.

[7] C. Photographs, A. Bhuvaneswari, C. Valliyammai, Social IoT-Enabled Emergency Event Detection Framework Using Geo-Tagged Microblogs and Crowdsourced Photographs, in: Advances in Intelligent Systems and Computing, vol. 814, no. March, Springer Singapore, 2019.

[8] A. Flizikowski, M. Przybyszewski, A. Stachowicz, T. Olejniczak, R. Renk, Text analysis tool Tweet locator - TAT2, in: Proc. ISCRAM 2015, 12th Int. Conf. Inf. Syst. Cris. Response Manag., 2015, pp. 171–181.

[9] A. Candelieri, F. Archetti, I. Giordani, G. Arosio, R. Sormani, Smart cities management by integrating sensors, models and user generated contents, WIT Trans. Ecol. Environ. 179 (2013) 719–730, https://doi.org/10.2495/SC130611.

[10] F. Ai, L.K. Comfort, Y. Dong, T. Znati, A dynamic decision support system based on geographical information and mobile social networks: a model for tsunami risk mitigation in Padang, Indonesia, Saf. Sci. 90 (2016) 62–74, https://doi.org/10.1016/j.ssci.2015.09.022.

[11] M.V. Sangameswar, M. Nagabhushana Rao, S. Satyanarayana, An algorithm for identification of natural disaster affected area, J. Big Data 4 (1) (2017) 1–11, https://doi.org/10.1186/s40537-017-0096-1.

[12] J. Gelernter, N. Mushegian, Geo-parsing messages from microtext, Trans. GIS 15 (6)

(2011) 753–773, https://doi.org/10.1111/j.1467-9671.2011.01294.x.

[13] J. Gelernter, S. Balaji, An algorithm for local geoparsing of microtext, Geoinformatica 17 (4) (2013) 635–667, https://doi.org/10.1007/s10707-012-0173-8.

[14] A. Hernandez-Suarez, et al., Using twitter data to monitor natural disaster social dynamics: a recurrent neural network approach with word embeddings and kernel density estimation, Sensors (Switzerland) 19 (7) (2019), https://doi.org/10.3390/s19071746.

[15] A. Agarwal, D. Toshniwal, Face off: travel habits, road conditions and traffic city characteristics bared using Twitter, IEEE Access 7 (2019) 66536–66552, https://doi.org/10.1109/ACCESS.2019.2917159.

[16] E. Fersini, E. Messina, F.A. Pozzi, Earthquake management: a decision support system based on natural language processing, J. Ambient Intell. Humaniz. Comput. 8 (1) (2017) 37–45, https://doi.org/10.1007/s12652-016-0373-4.

[17] H. Mao, G. Thakur, K. Sparks, J. Sanyal, B. Bhaduri, Mapping near-real-time power outages from social media, Int. J. Digit. Earth (2018) 1–15, https://doi.org/10.1080/17538947.2018.1535000.

[18] M. Jamali, A. Nejat, S. Ghosh, F. Jin, G. Cao, Social media data and post-disaster recovery, Int. J. Inf. Manage. 44(September) (2019) 25–37. https://doi.org/10.1016/j.ijinfomgt.2018.09.005.

[19] J.L.P. Barker, C.J.A. Macleod, Development of a national-scale real-time Twitter data mining pipeline for social geodata on the potential impacts of flooding on communities, Environ. Model. Softw. 115(January) (2019) 213–227. https://doi.org/10.1016/j.envsoft.2018.11.013.

[20] D. Reynard, M. Shirgaokar, Harnessing the power of machine learning: can Twitter data be useful in guiding resource allocation decisions during a natural disaster? Transp. Res. Part D Transp. Environ. (2019) 1–15 https://doi.org/10.1016/j.trd.2019.03.002.

[21] S. Koswatte, K. Mcdougall, X. Liu, Semantic location extraction from crowdsourced data, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 41 (July) (2016) 543–547, https://doi.org/10.5194/isprsarchives-XLI-B2-543-2016.

[22] T.B.N. Hoang, J. Mothe, Location extraction from tweets, Inf. Process. Manag. 54 (2) (2018) 129–144, https://doi.org/10.1016/j.ipm.2017.11.001.

[23] F. Laylavi, A. Rajabifard, M. Kalantari, A multi-element approach to location inference of Twitter: a case for emergency response, ISPRS Int. J. Geo-Information 5 (5) (2016) 56, https://doi.org/10.3390/ijgi5050056.

[24] R.I. Ogie, H. Forehead, Investigating the accuracy of georeferenced social media data for flood mapping: The PetaJakarta.org case study, in: Proc. 2017 4th Int. Conf. Inf. Commun. Technol. Disaster Manag. ICT-DM 2017, vol. 2018-Janua, no. 71984, 2018, pp. 1–6. https://doi.org/10.1109/ICT-DM.2017.8275672.

[25] L. Ghahremanlou, W. Sherchan, J.A. Thom, Geotagging twitter messages in crisis management, Comput. J. 58 (9) (2014) 1937–1954, https://doi.org/10.1093/comjnl/bxu034.

[26] V. Rakesh, C.K. Reddy, D. Singh, M.S. Ramachandran, Location-specific tweet detection and topic summarization in Twitter, in: Proc. 2013 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2013, 2013, pp. 1441–1444. https://doi.org/10.1145/2492517.2492583.

[27] S. Kitaoka, T. Hasuike, Where is safe: Analyzing the relationship between the area and emotion using Twitter data, in: 2017 IEEE Symp. Ser. Comput. Intell. SSCI 2017 - Proc., vol. 2018-Janua, pp. 1–8, 2018, https://doi.org/10.1109/SSCI.2017.8285210.

[28] A.M. Sadri, S. Hasan, S.V. Ukkusuri, M. Cebrian, Crisis Communication Patterns in Social Media during Hurricane Sandy, 2017, https://doi.org/10.1177/0361198118773896.

[29] H. Li, D. Caragea, C. Caragea, N. Herndon, Disaster response aided by tweet classification with a domain adaptation approach, J. Contingencies Cris. Manag. 26 (1) (2018) 16–27, https://doi.org/10.1111/1468-5973.12194.

[30] O. Almatrafi, S. Parack, B. Chavan, Application of location-based sentiment analysis using twitter for identifying trends towards indian general elections 2014, in: ACM IMCOM 2015 - Proc., 2015, https://doi.org/10.1145/2701126.2701129.

[31] T. Yilmaz, P. Karagoz, Y. Kavurucu, Exploring What Makes It a POI," in 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, 2017, pp. 0–5.

[32] M. Nik Bakht, T.E. El-Diraby, M. Hossaini, Game-based crowdsourcing to support collaborative customization of the definition of sustainability, Adv. Eng. Informatics 38 (September) (2018) 501–513, https://doi.org/10.1016/j.aei.2018.08.019.

[33] J. Vora, A.M. Chacko, Sentiment analysis of tweets to identify the correlated factors that influence an issue of interest, in: 2nd Int. Conf. Telecommun. Networks, TEL-NET 2017, vol. 2018-Janua, 2018, pp. 1–6. https://doi.org/10.1109/TEL-NET.2017.8343572.

[34] U.H. Govindarajan, A.J.C. Trappey, C.V. Trappey, Intelligent collaborative patent mining using excessive topic generation, Adv. Eng. Informatics 42(April) (2019) 100955. https://doi.org/10.1016/j.aei.2019.100955.

[35] D. Blei, M. Jordan, A.Y. Ng, Latent Dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022, https://doi.org/10.1162/jmlr.2003.3.4-5.993.

[36] Z. Tong, H. Zhang, A Text Mining Research Based on LDA Topic Modelling, 2016, pp. 201–210. https://doi.org/10.5121/csit.2016.60616.

[37] C.Y. Chu, K. Park, G.E. Kremer, A global supply chain risk management framework: an application of text-mining to identify region-specific supply chain risks, Adv. Eng. Informatics 45(December) (2020) 101053 https://doi.org/10.1016/j.aei.2020.101053.

[38] M. Hu, B. Liu, S.M. Street, Mining and summarizing customer reviews, in: Proceedings of the ACM SIGKDD International Conference on Knowledge, 2004, https://doi.org/10.1002/j.1532-2149.2013.00312.x.

[39] B. Liu, S.M. Street, S.M. Street, S.M. Street, Opinion Observer Analyzing and Comparing Opinions.pdf.

[40] S. Shalev-Shwartz, S. Ben-David, Understanding Machine Learning: From Theory to Algorithms, Cambridge University, New York, 2013.

[41] P. Sharma, A. Agrawal, A. Lalit, A. Garg, Challenges and Techniques in Preprocessing for Twitter Data, vol. 7(4), 2017, pp. 6611–6613.

[42] P. Kaviani, S. Dhotre, Short survey on naive bayes algorithm, Int. J. Adv. Eng. Res. 4(March) (2017) 607–611 [Online]. Available: http://www.ijaerd.com/papers/finished_papers/Short Survey on Naive Bayes Algorithm-IJAERDV04I1140826.pdf.

[43] A.M.D.E. Hassanein, M. Nour, A proposed model of selecting features for classifying Arabic text, Jordanian J. Comput. Inf. Technol. 05 (03) (2019) 275–290.

[44] V. Kecman, Basics of Machine Learning by Support Vector Machines, StudFuzz, no. May, 2014, pp. 49–103. https://doi.org/10.1007/11364160.

[45] Y. Benajiba, M. Diab, P. Rosso, Arabic Named Entity Recognition: A Feature-Driven Study, vol. 17, no. 5, 2009, pp. 926–934.

[46] L. Rokach, O. Maimon, Decision Trees, in: Data Mining and Knowledge Discovery Handbook, no. January, 2005, pp. 165–192.

[47] M.A. Jun, J.C.P. Cheng, Selection of target LEED credits based on project information and climatic factors using data mining techniques, Adv. Eng. Informatics 32 (2017) 224–236, https://doi.org/10.1016/j.aei.2017.03.004.

[48] K.J.M. Rifkie Primarthaa, Bayu Adhi Tamab, Azhary Arliansyaha, Decision tree combined with PSO-based feature selection for sentiment analysis Decision tree combined with PSO-based feature selection for sentiment analysis, 2019, https://doi.org/10.1088/1742-6596/1196/1/012018.

[49] G. Khanvilkar, D. Vora, Product recommendation using sentiment analysis of reviews: a random forest approach, Int. J. Eng. Adv. Technol., no. January, 2019.

[50] S. Galelli, A. Castelletti, Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling, Hydrol. Earth Syst. Sci. 17 (7) (2013) 2669–2684, https://doi.org/10.5194/hess-17-2669-2013.

[51] M.W. Ahmad, J. Reynolds, Y. Rezgui, Predictive modelling for solar thermal energy systems: a comparison of support vector regression, random forest, extra trees and regression trees, J. Clean. Prod. 203 (2018) 810–821, https://doi.org/10.1016/j.jclepro.2018.08.207.

[52] C. Simon Haykin (McMaster University, Hamilton, Ontario, "Neural Networks - A Comprehensive Foundation - Simon Haykin.pdf.", 2005, p. 823.

[53] J.T. Hancock, T.M. Khoshgoftaar, Survey on categorical data for neural networks, J. Big Data (2020), https://doi.org/10.1186/s40537-020-00305-w.

[54] A. Ravindran, K.R. Narayanan, G. Deepa, A Comparative Study of Stochastic Gradient Descent and Naïve Bayes Multinomial for Text Classification on Spam Words, vol. 29, no. 03, 2020, pp. 5393–5398.

[55] A. Taherkhani, G. Cosma, T.M. Mcginnity, Neurocomputing AdaBoost-CNN: An adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning, vol. 404, 2020, pp. 351–366, https://doi.org/10.1016/j.neucom.2020.03.064.

[56] A. Onan, Biomedical text categorization based on ensemble pruning and optimized topic modelling, Comput. Math. Methods Med. 2018 (2018), https://doi.org/10.1155/2018/2497471.