Proceedings of the Fourth International Conference on Smart Electronics and Communication (ICOSEC-2023)
IEEE Xplore Part Number : CFP23V90-ART ; ISBN : 979-8-3503-0088-8

Disaster Tweet classification using Machine Learning Model

1st Jagadishwari V
Dept of Computer Science and Engineering
CMR Institute of Technology
Bengaluru, India
jiswariv@yahoo.com

2nd Ganeshgouda H
Dept of Computer Science and Engineering
CMR Institute of Technology
Bengaluru, India
gahi19cs@cmrit.ac.in

3rd Harikrishna
Dept of Computer Science and Engineering
CMR Institute of Technology
Bengaluru, India
hana19cs@cmrit.ac.in

4th Pavan .H
Dept of Computer Science and Engineering
CMR Institute of Technology)
Bengaluru, India
pavh19cs@cmrit.ac.in

2023 4th International Conference on Smart Electronics and Communication (ICOSEC) | 979-8-3503-0088-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICOSEC58147.2023.10276240

***Abstract*—Social media is playing a vital role in communication, forums like Twitter and Instagram are where people socialize. Recently social media has played a crucial role during disasters. People could post messages and get the required help. If the messages posted can be classified into Normal and disaster posts it would be very helpful. In this work, the aim is to classify the Tweets posted on Twitter into Normal and disaster Tweets. The Multinomial Naive Bayesian, Passive Aggressive (PA) Classifier, and Support Vector Machine(SVM) were the models that were built and tested using Twitter data. Evaluation of these models with various measures and were able to conclude that the Passive Aggressive classifier outperforms the other two models.**

***Index Terms*—Tweet, Multinomial Bayes, Support Vector Machines, Passive Aggressive classifier**

I. Introduction

Social media has brought about a revolution in the way in which information is exchanged. Among the many forums available Twitter is a dominant one. It was initially designed to cater to messaging services for a small group, but today it has millions of users worldwide. It allows users to send Tweets which are essentially brief articles, reviews, and messages that are posted. The huge data generated data from social media forums are used in data analytics and the outcome of this analytics is useful in various domains. Analyzing the Sentiment of the Tweets reveals the way in which people's views on various issues. It has emerged as a vital medium of communication. In the case of natural calamities and disasters, the information flows seamlessly from the place of calamity to across the world. The accessibility of cellphones, laptops, and tablets has made it possible for people to share in real-time the occurrence of disasters they are experiencing. When people Tweet related to a disaster, more people get the information, and help will pour in from all over in whichever possible way. Since Twitter is a global platform and on average around 6,000 Tweets are posted every second, it may take some time to recognize and respond to disaster tweets. Hence it will be helpful to classify the tweets into normal and disaster tweets. Many data analytics firms are attempting to monitor and analyze Twitter data programmatically. News media and disaster aid organizations are specifically attempting to examine tweets in real-time to figure out the occurrence of calamities. This would aid in preventing the hazards of looming disasters for millions of ordinary people. People could take preventive action if they were alerted in real-time when disasters were occurring at a specific location. Before the situation goes out of control, government agencies can take measures to help the people in handling the disaster. At the time of crisis, social media messages serve as a critical source of information. As the information on forums like Twitter reaches people at lightning speed it may help recovery and disaster organizations to travel to the affected place on time and provide their support services. We can find an ample amount of research in classifying Tweets into Normal and disaster tweets. As most of us are aware that In 2012, Hurricane Sandy caused devastation in the United States. Extensive case studies were conducted on this calamity. An analysis of Tweets was done by collecting Twitter data in this period. In the work by Stowe et al work [1] Tweets for the event were captured by using the keywords such as DSNY, cleanup, debris, Franken storm, garbage, hurricane, hurricane sandy, lbi, occupysandy, perfect storm, sandy, sandycam, superstorm from October 23, 2012, to April 5, 2013. They have trained three models support vector machines (SVMs), maximum entropy (MaxEnt) models, and Naive Bayes. The SVM model is giving the best performance. Similarly, another work by Koshy et al [2] focused on the most devastating floods in Chennai, India. The tweets collected were analyzed with machine learning algorithms like Random Forests, Naive Bayes, and Decision Tree. By comparing the performances of all three it was concluded that Random Forests is the best algorithm that can be relied on, during a disaster. This work also targeted the sources of the Twitter messages to explore the most influential users of the Chennai flood. A similar work was carried out in [3]. In 2012, the Habagat flooding caused havoc in Metro Manila [4]. The Twitter data of this disaster was utilized in training the classifier models. A labeled data set was created by giving labels as either useful or useless. The tweets were classified with two ML algorithms the Naive Bayes and SVM [5]. Evaluation of these models were done with accuracy and precision. Another work that used Naive Bayes was done by Muppidi et al [6], In this work they classify the natural disaster-based tweets by using classification machine learning algorithms like Na¨ıve Bayes along with Logistic Regression, KNN, and Random Forest and determine the best machine learning algorithm was chosen on the basis of evaluation measures such as accuracy, kappa, etc. Multinomial naive Bayes was used by Gautam et al [3].This study proposes a novel decision diffusion method to categorize multimodal data pertaining to seven different natural disasters into informative and non-informative categories. The disasters include hurricanes, floods, earthquakes, and others. The methodology deployed in this work is that , the classification of tweets basis of their text has been broken into two distinctive approaches. The first approach uses N-gram feature representation which is fed to machine learning algorithms like Logistic Regression, Multinomial Naive Bayes, and Random Forest classifier. The second approach uses sophisticated deep-learning models for classification. An ample amount of work has been done by applying many deep-learning models to classify Tweets [7] [8] [9]. RNN and CNN was combined and used in the work by Xukun et al [10] [11]. Text processing was done with RNN and images were processed with CNN for and then they combine the two models into a multi-modal model. The dataset we utilized in this work is the CrisisMMD dataset [12] that was published by Alam Ofli. The dataset contains tweets crawled during seven disaster events. For each tweet, the text and image of the tweet were labeled as informative/non-informative. In the work by Zahera et al, BERT Model [13] was deployed for multiclass Tweet classification. They employed a fine-tuned BERT model with 10 BERT layers. They have used TREC-IS dataset which contains 17,682 tweets for training and 7,634 for testing. The data has been curated from different types of events (e.g. earthquakes, hurricanes, public or shootings). For each event, the tweets were collected using hashtags and keywords. Human annotators have labeled tweets into the multi-layer ontology of information types. Kabir et al [14] proposed a deep learning model combining attention-based BLSTM and CNN to classify the tweets. Image features along with the text features were also considered in the work proposed in [15]. It includes a model utilizing the text, a model utilizing the image, and also a fusion of both. Features were extracted with CNN in model utilizing the text and ANN is deployed for the classification. Features were extracted from the image with VGG-16 in the model that utilizes image. A fusion model by clubbing both the aforesaid models are built to label the Tweets.

II. METHODOLOGY

The methodology deployed in this work is depicted in Fig 1. The first phase is getting the data from Twitter and then

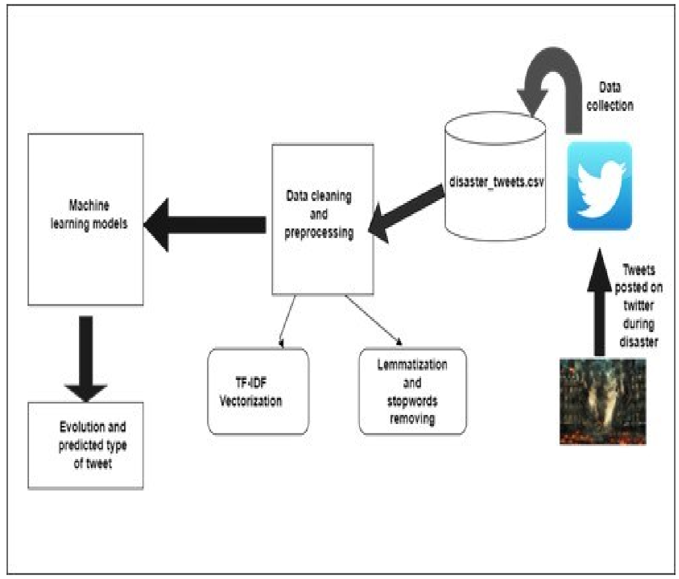

Fig. 1. Work Flow diagram

cleaning data which is explained in detail in the subsequent section. Training Machine learning is done using preprocessed data. The models are later tested. The models are given single tweets and they classify them as normal or disaster tweets. Evaluation measures of the model are discussed in the results section.

A. Description of the Data set and Preprocessing

B. Data preprocessing

The data was gathered from previous occurrences of disasters. The historical tweet data was collected from Twitter using web scraping. The Tweets were got for the time period between 1/1/2017 to 1/1/2023. The keywords like floods, earthquake, storm, fire, bomber, crash attack, wildfire, Hiroshima, nuclear, storm, Tsunami, crash, thunderstorm and etc, were used to extract tweets related to disasters. Around 7,000 tweets were scrapped in total, and these tweets were labeled as Normal and disaster tweets. The disaster tweets were given

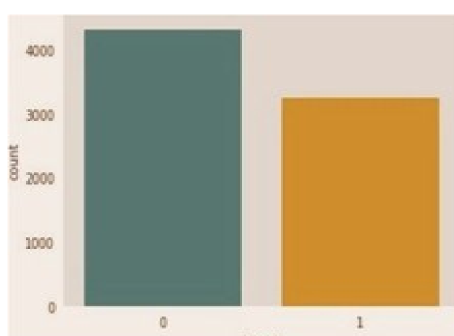

Fig. 2. Data set class

a label 1 and the normal tweets 0. Fig 2 illustrates the composition of the dataset, consisting of over 4000 disaster-related tweets and approximately 3000 normal tweets. In 3, the word cloud of disaster Tweets and Normal tweets is depicted. The following two-step procedure was applied to preprocess the data:

- Data cleaning: The data got from Twitter in the form of Tweets has to be preprocessed and converted into a format that is understood by the ML model. All stop words were removed , and then stemming or lemmatization was applied to the data. Stop word remaval and Lemmatization was done using the functions from the NLTK package in Python.
- Vectorization: The machine learning models can only handle numeric data in this step, the textual input was vectorized using TFIDF. With TFIDF two versions of vectorizers were developed one is the bi-grams and the other is trigrams. The ngram range =(1,2) is used in the first scenario, which implies it will accept unigram and bi-gram features from the text, ngram range =(1,3) will accept unigram, bi-gram, and tri-gram features.

C. *Machine learning classifiers*

Three machine learning classifiers namely Multinomial Naive Bayesian, Passive Aggressive (PA) Classifier, and Support Vector Machine(SVM) were trained to classify the Tweets into Disaster and Normal Tweets. These classifiers are described below:

- Multinomial Naive Bayesian (MNB): MNB, acronymn for Multinomial Naive Bayes, represents a variation of the Naive Bayes classifier. Bayes theorem is at work in this model. MNB is a widely employed probabilistic algorithm utilized for text classification tasks as well. The algorithm relies on Bayes' theorem to compute the probability of a text belonging to a specific class by considering the frequency of word occurrences within the text. MNB assumes that the number of times a word occurs in a text follows a multinomial distribution, which makes it well-suited for text classification tasks. MNB has been shown to perform well in text classification tasks, particularly in scenarios where the quantity of features (words) outweighs the number of samples.
- Passive Aggressive (PA) Classifier: This classifier falls into a group called the online learning algorithm. This is the best-suited algorithm for any kind of social media analytics. The reason behind it is that when the data is very huge, it becomes almost impossible to train the model at once with the entire data. The social media generates data at a very high velocity and the PA classifier will be trained on the go. As the data is presented to the classifier it classifies the data instance if the model has correctly classified the data then it remains passive and the data instance is discarded. If the prediction made by the model was wrong then it acts aggressively and makes the required modifications to the model. Hence the name Passive Aggressive algorithm. This model can be trained on the fly as and when data is available, hence it better suits Social Network analytics where data is generated dynamically.
- Support Vector Machine(SVM): SVM is a supervised technique deployed for classification problems.In comparison to regression, which takes a probabilistic approach, it can be described that MNB is a statistical approach. SVM is better suits the problems with small and complex data sets. It classifies the data by building a decision boundary called the hyperplane. To do this it uses a subset of the data points from the dataset, called support vectors and there by the name Support Vector Machines. SVM is a powerful algorithm that can handle non-linearly separable data by mapping the data points to a higher-dimensional space, where a linear separation can be found. SVM has been shown to perform well in text classification tasks, especially when combined with techniques like kernel methods if the hyperplane classifies the datasets linearly then the algorithm is called Support Vector Classifier(SVC) and the algorithm that separates the datasets by the non-linear approach is called SVM.

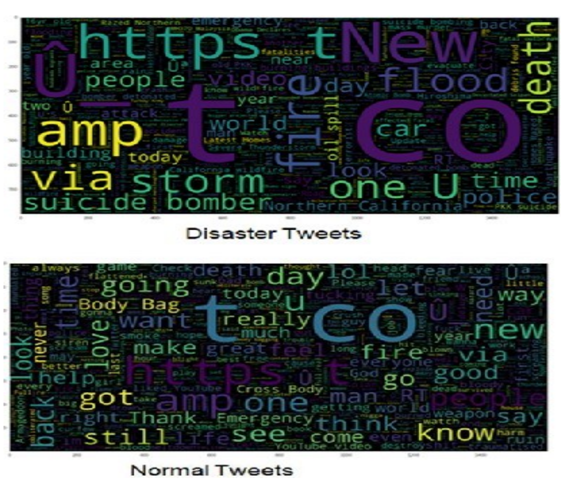

Fig. 3. Word cloud of labeled data

III. RESULTS AND DISCUSSION

Disaster Tweet classification was done by training the three models Multinomial Naive Bayesian, Passive Aggressive (PA) Classifier, and SVM described in the previous section. The models underwent training on 80% of the available data, while the evaluation of their performance was carried out using the remaining 20% of the data.. Models were trained with two versions of the vectorizer, namely the Bigram, and Trigram. Models evaluation were done with measures like accuracy, precision, sensitivity F1 score, and ROC. Table 1 shows the results. Inspecting the table it is evident that accuracy, precision, F1 score, and ROC for all models slightly vary but the difference is negligible. By considering the sensitivity metric, it becomes apparent that the Passive Aggressive model outperforms all other models. Sensitivity is a critical metric

Table 1. Model Evaluation Metrics

Model	Accuracy	Precision	Sensitivity	F1 Score	ROC
MBN Bigram	0.80	0.86	0.64	0.73	0.78
PA Bigram	0.78	0.74	0.75	0.75	0.77
MNB Trigram	0.79	0.87	0.63	0.73	0.77
PA Trigram	0.78	0.74	0.76	0.75	0.77
SVM Bigram	0.80	0.83	0.67	0.74	0.78
SVM Trigram	0.79	0.83	0.66	0.73	0.77

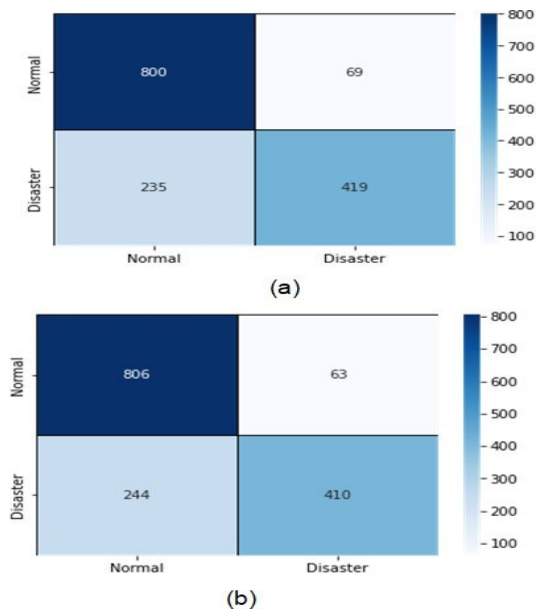

Fig. 4. Confusion Matrix of MNB: (a) BiGram (b) TriGram

when we have a balanced dataset. A considerably balanced dataset utilized in this study, as depicted in Fig 2. The Passive Aggressive model exhibits strong performance across various measures, indicating that it outperforms the other two models. Therefore, it can be concluded that the Passive Aggressive model surpasses the performance of the other two models. The confusion matrices of MNB, Passive aggressive, and SVM models for both the Bigram and Trigram versions are shown in Fig 4, Fig 5, and Fig 6 respectively. It can be seen that for the Multinomial Naive Bayesian model, the confusion matrices are almost similar for the Bigram and Trigram versions but they differ slightly for the Bigram and Trigram versions of the Passive Aggressive (PA) Classifier and SVM Classifier.

IV. Conclusion

In this work, we have shown how to train a machine-learning model for identifying disaster tweets from the Twitter

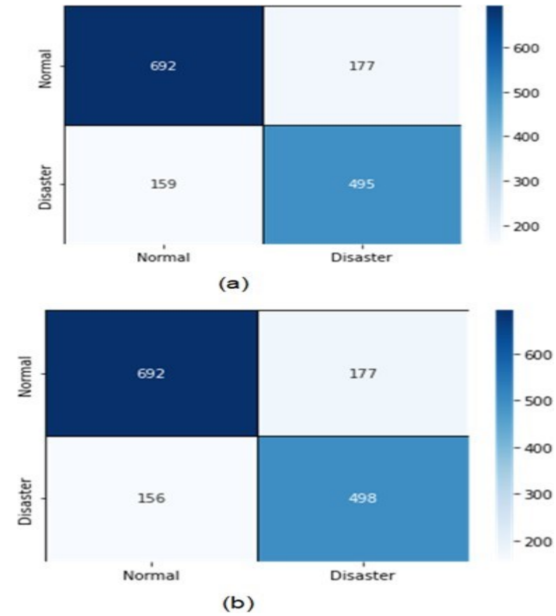

Fig. 5. Confusion Matrix of Passive Aggressive Model (a) BiGram (b) TriGram

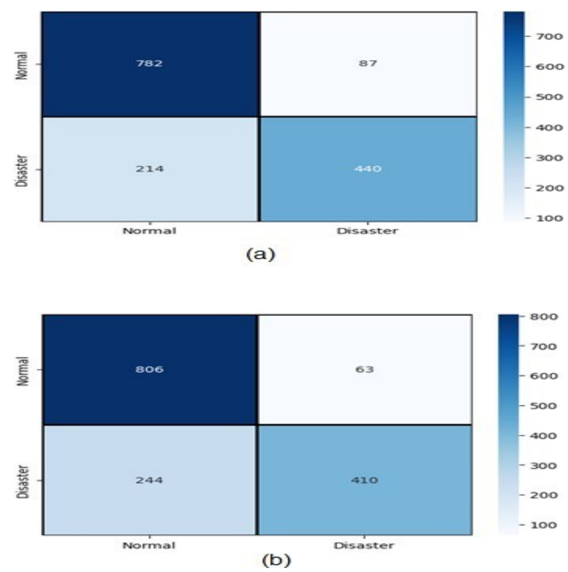

Fig. 6. Confusion Matrix of SVM: (a) BiGram (b) TriGram

dataset using natural language processing. In addition, we have shown various data pre-processing methods for cleaning the data. In the end, we discovered that the Passive Aggressive Classifier trained on Trigram performed the best for this use-case after training three machine learning models, namely Multinomial Naive Bayes , Passive Aggressive Classifier and SVM on Bi-gram and Tr-gram variants of TFIDF vectorized data. Last but not least, to evaluate the model's overall effectiveness, we also extracted key features from the model for the classes of disaster and regular tweets and made predictions on illustrative test words.

References

[1] K. Stowe, M. Paul, M. Palmer, L. Palen, and K. M. Anderson, "Identifying and categorizing disaster-related tweets," in *Proceedings of The fourth international workshop on natural language processing for social media*, 2016, pp. 1–6.

[2] R. Koshy and S. Elango, "Multimodal tweet classification in disaster response systems using transformer-based bidirectional attention model," *Neural Computing and Applications*, vol. 35, no. 2, pp. 1607–1627, 2023.

[3] A. K. Gautam, L. Misra, A. Kumar, K. Misra, S. Aggarwal, and R. R. Shah, "Multimodal analysis of disaster tweets," in *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)*. IEEE, 2019, pp. 94–103.

[4] B. E. Parilla-Ferrer, P. L. Fernandez, and J. T. Ballena, "Automatic classification of disaster-related tweets," in *Proc. International Conference on innovative engineering technologies (ICIET)*, vol. 62, 2014.

[5] A. Kumar, J. P. Singh, and S. Saumya, "A comparative analysis of machine learning techniques for disaster-related tweet classification," in *2019 IEEE R10 Humanitarian Technology Conference (R10-HTC)(47129)*. IEEE, 2019, pp. 222–227.

[6] S. Muppidi, P. S. Rao, and M. Rama Krishna Murthy, "Identification of natural disaster affected area using twitter," in *Advances in Decision Sciences, Image Processing, Security and Computer Vision: International Conference on Emerging Trends in Engineering (ICETE), Vol. 1*. Springer, 2019, pp. 792–801.

[7] R. ALRashdi and S. O'Keefe, "Deep learning and word embeddings for tweet classification for crisis response," *arXiv preprint arXiv:1903.11024*, 2019.

[8] A. Bhoi, S. P. Pujari, and R. C. Balabantaray, "A deep learning-based social media text analysis framework for disaster resource management," *Social Network Analysis and Mining*, vol. 10, pp. 1–14, 2020.

[9] H. Li, D. Caragea, and C. Caragea, "Combining self-training with deep learning for disaster tweet classification," in *The 18th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2021)*, 2021.

[10] L. Xukun and D. Caragea, "Improving disaster-related tweet classification with a multimodal approach," in *ISCRAM 2020 Conference Proceedings–17th International Conference on Information Systems for Crisis Response and Management*, 2020.

[11] S. Madichetty and M. Sridevi, "Detecting informative tweets during disaster using deep neural networks," in *2019 11th international conference on communication systems & networks (COMSNETS)*. IEEE, 2019, pp. 709–713.

[12] F. Alam, F. Ofli, and M. Imran, "Crisismmd: Multimodal twitter datasets from natural disasters," in *Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM)*, June 2018.

[13] H. M. Zahera, I. A. Elgendy, R. Jalota, M. A. Sherif, and E. Voorhees, "Fine-tuned bert model for multi-label tweets classification." in *TREC*, 2019, pp. 1–7.

[14] M. Y. Kabir and S. Madria, "A deep learning approach for tweet classification and rescue scheduling for effective disaster management," in *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2019, pp. 269–278.

[15] S. Madichetty, "Classifying informative and non-informative tweets from the twitter by adapting image features during disaster," *Multimedia Tools and Applications*, vol. 79, pp. 28 901–28 923, 2020.