# Disaster Tweet Classification Using GloVe Embeddings and LSTM

Posted on November 4, 2026

View this essay online:
https://academic-master.com/disaster-tweet-classification-using-glove-embeddings-and-lstm/

## Abstract

Twitter provides a rapid stream of news and other information through short messages shared by its users. Tweets can report events before traditional mass media, but the volume and variability of this uncontrolled social-media data make it difficult to identify useful information. Deep learning offers a way to process these messages automatically. This paper presents a long short-term memory (LSTM) model for identifying disaster-related tweets. The system combines GloVe word embeddings, an embedding layer, spatial dropout and a sigmoid output for binary classification. After hyperparameter tuning, the study reports a validation accuracy of 83.33%. Smartphones enable people to share information about crises, epidemics and emergencies as events unfold, creating interest among organisations in monitoring and interpreting these messages. The proposed approach examines how word representations and sequence modelling can help distinguish relevant disaster reports from other tweets.

Keywords Disaster tweet classification, GloVe embeddings, LSTM, recurrent neural networks, deep learning.

## Introduction

During disasters and emergencies, individuals, public agencies, private organisations and local and global communities use microblogging sites to disseminate messages [1–2]. These platforms have attracted attention as channels for emergency communication because of their widespread availability, rapid messaging and accessibility across devices [3–4]. Social-media interactions are distributed, decentralised and often occur in real time, providing information that may be useful during an emergency [5]. Twitter developed around short messages, originally limited to 140 characters, through which users shared relevant information and breaking news. Emergency tweets may contain first-hand accounts or important information obtained from other sources [6].

Twitter can therefore support understanding of the circumstances of affected people [7]. Messages posted during a crisis may help agencies establish what is happening locally and improve their awareness of an unfolding situation. Such awareness is important when making timely and effective

emergency-response decisions. The role of social media in emergency management also depends on the ability to identify useful reports within a much larger stream of messages. The range of topics discussed on Twitter is extensive, and the volume of posts can become overwhelming in the days following a disaster. Increased information seeking and sharing contribute to this surge.

Microblogging allows information to be retrieved, produced and distributed through a rapid cycle of generation and consumption [8–9]. As tweets spread, automatic classification becomes useful for extracting relevant material. Accessible disaster information can contribute to more effective response systems and support efforts to reduce loss of life, property damage and other consequences of a crisis [10]. An automatic classification model must nevertheless handle messages containing both useful information and irrelevant characters or other forms of noise [11]. Informal writing and inconsistent language make this a challenging text-processing task.

The proposed pipeline uses GloVe vectors with an embedding layer and LSTM to classify tweets as disaster-related or not disaster-related [12]. The approach uses vector representations of words to support the learning of relationships within a message. The following sections review related methods, describe the model architecture and training procedure, present the reported results, and discuss future work.

# Related Work

Text-classification approaches discussed in the study include bag-of-words features combined with logistic regression, with accuracy reported at approximately 74% [13]. Bag-of-words represents a document through word occurrences but provides limited information about semantic relationships. A second approach combines term frequency–inverse document frequency (TF-IDF) with logistic regression, with accuracy reported at approximately 75%. TF-IDF assigns weights to terms according to their frequency within a document and their prevalence across the collection. It offers a useful feature representation, although it does not directly model the full context or order of words.

Word embeddings provide an alternative to sparse representations [14]. A large vocabulary can produce high-dimensional bag-of-words vectors, many of whose entries are zero. Embeddings instead represent words as dense vectors in a continuous space. Relationships learned from the contexts in which words occur help determine their positions in this space. The resulting representations can carry information useful for subsequent classification.

The position of a word in this learned space is its embedding. Word2Vec and GloVe are commonly used approaches [15]. This study uses 100-dimensional GloVe vectors to represent words. Related tweet-classification research has also explored bidirectional LSTM and convolutional neural-network architectures. Similar preprocessing and sequence-learning ideas appear in tweet-based sentiment analysis, although the present task concerns disaster relevance rather than sentiment.

# Proposed Deep Learning Architecture

The model contains successive processing layers that transform the input and pass learned features to the next stage. During training, predicted outputs are compared with the corresponding labels, and model weights are adjusted to reduce prediction error [16]. Figure 1 shows the overall sequence: dataset preparation, embedding, spatial dropout, LSTM processing and sigmoid-based binary classification.

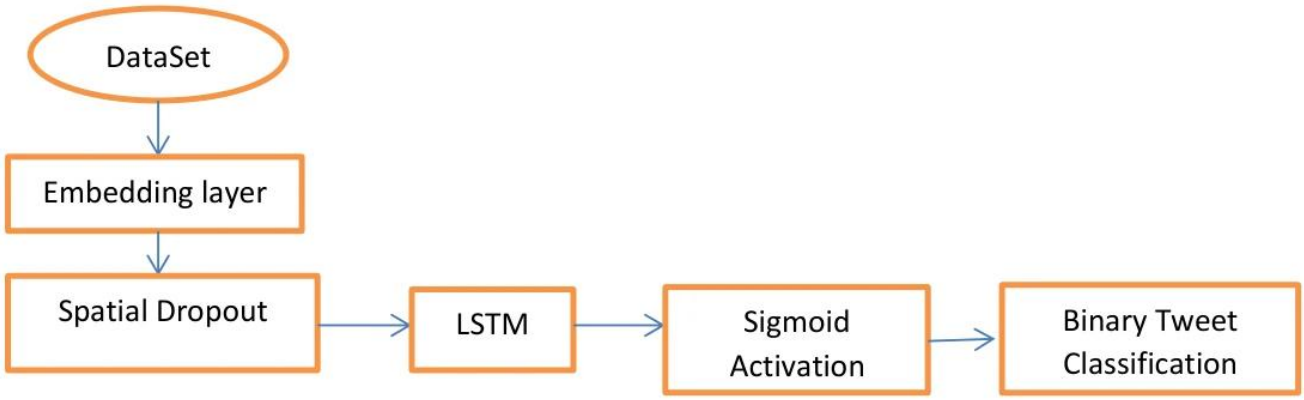



*Figure 1. Overall architecture of the disaster-tweet classification model.*

# Embedding Layer

Raw text must first be converted into numerical input. The Keras Tokenizer API assigns an integer identifier to each word, allowing a tweet to be represented as a sequence of integers. The embedding layer maps these identifiers to dense word vectors. The model uses GloVe vectors as word representations and describes updating the embeddings during training. The source also refers to random initialisation; Figure 2 preserves the supplied processing diagram. Embeddings provide the numerical representation passed to the later layers.

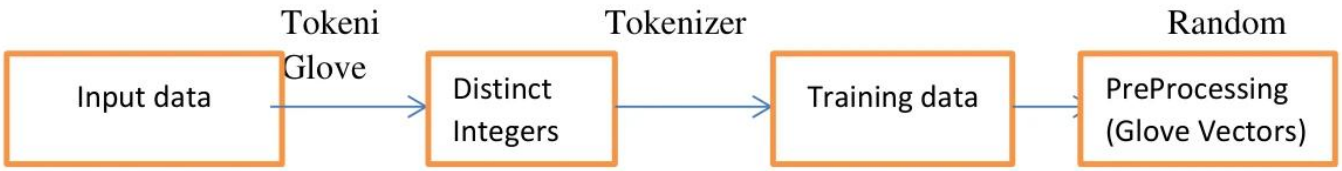



*Figure 2. Tokenization and embedding-layer processing.*

## Spatial Dropout

One-dimensional spatial dropout is used to reduce overfitting. During training, it drops entire feature maps rather than isolated elements. This encourages the model to avoid relying too heavily on particular features and supports independence among feature maps. Figure 3 illustrates the feature-map processing stage.

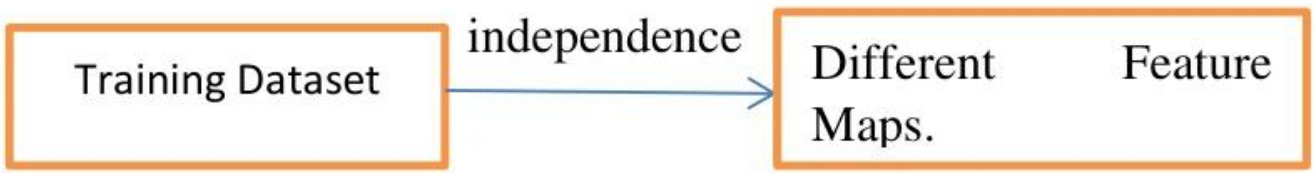



*Figure 3. Spatial dropout and feature-map processing.*

## Long Short-Term Memory

LSTM is a type of recurrent neural network designed to handle dependencies across a sequence. In a recurrent network, information from a previous step contributes to processing the current input. Conventional RNNs can struggle to retain useful information as the distance between related sequence elements grows. LSTM addresses this difficulty through a memory structure that can preserve information over longer intervals. Such networks are used in sequence classification, forecasting and language modelling [17]. In this model, LSTM processes the embedded tweet to learn relationships between its words, as outlined in Figure 4.

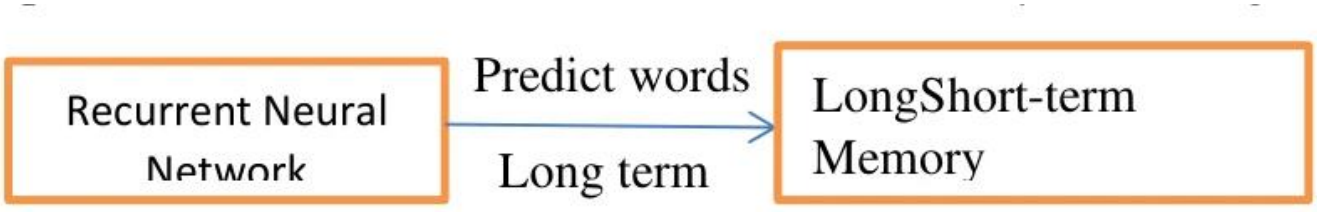



*Figure 4. Recurrent processing and long short-term memory.*

## Sigmoid Activation

The output layer uses sigmoid activation for binary tweet classification. The task is to distinguish disaster-related tweets from other messages. The sigmoid output provides a score for the binary decision. If the task were extended to several mutually exclusive categories, a softmax output could instead be considered. Figure 5 shows the supplied binary-classification stage.

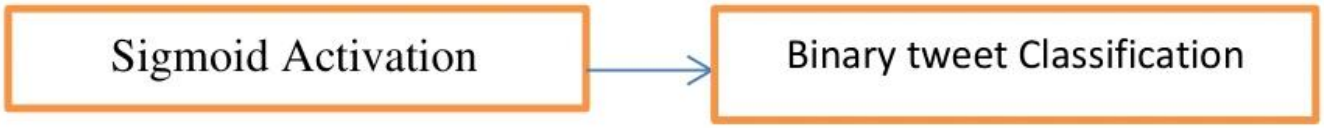



*Figure 5. Sigmoid activation for binary tweet classification.*

# Network Training

## Tweet Preprocessing

Tweets may contain incomplete sentences, informal expressions, misspellings, unusual vocabulary and words absent from a standard dictionary. Preprocessing prepares these messages for numerical representation. The described procedure removes HTTP addresses, hashtags, emoticons, stop words and punctuation marks. Figures 6 and 7 show examples before and after preprocessing. These transformations reduce some forms of input noise before tokenization and embedding.

| text |
| --- |
| Our Deeds are the Reason of this #earthquake M... |
| Forest fire near La Ronge Sask. Canada |
| All residents asked to 'shelter in place' are ... |
| 13,000 people receive #wildfires evacuation or... |
| Just got sent this photo from Ruby #Alaska as ... |
| #RockyFire Update => California Hwy. 20 closed... |
| #flood #disaster Heavy rain causes flash flood... |
| I'm on top of the hill and I can see a fire in... |
| There's an emergency evacuation happening now ... |
| I'm afraid that the tornado is coming to our a... |

*Figure 6. Example tweet text before preprocessing.*

| text |
| --- |
| Our Deeds are the Reason of this earthquake Ma... |
| Forest fire near La Ronge Sask Canada |
| All residents asked to shelter in place are be... |
| 13000 people receive wildfires evacuation orde... |
| Just got sent this photo from Ruby Alaska as s... |
| RockyFire Update California Hwy 20 closed in ... |
| flood disaster Heavy rain causes flash floodin... |
| Im on top of the hill and I can see a fire in ... |
| Theres an emergency evacuation happening now i... |
| Im afraid that the tornado is coming to our area |

*Figure 7. Example tweet text after preprocessing.*

## Initialising GloVe Embeddings

GloVe uses global word co-occurrence information to learn vector representations. A co-occurrence

matrix records how often words appear together across the corpus. In the notation used here, $X_{ij}$ denotes the number of times word j appears in the context of word i. This global information differs from an approach that considers only a local context window when predicting a context word. Each sentence is divided into tokens, and each token is associated with a vector representation that reflects information learned from word contexts.

## Training Procedure and Hyperparameters

The model is trained using the numerical tweet sequences and their labels. The initial word representations are updated with the other model parameters through backpropagation [18]. Table 1 preserves the learning rate, batch size and number of epochs reported in the study. Hyperparameter tuning is used to improve the reported validation performance.

Table 1. Training hyperparameters

| Hyperparameter | Value |
|---|---|
| Learning rate | 1e-4 |
| Batch size | 4 |
| Number of epochs | 20 |

## Prediction

After training and validation, the model is applied to example tweets. The same tokenization and padding steps prepare an input message for prediction. Figure 8 reproduces the supplied code example and model output. The proposed application is to support information gathering or alerts within disaster-management and response systems. The example illustrates the prediction workflow; a single output does not establish the reliability of an operational alerting system.

```
tokL = pickle.load(open('tokenizer.pkl','rb'))
```

```
text = "heavy rains in madanapalle"
sequenceS=tokL.texts_to_sequences([text])
tweetPad=pad_sequences(sequenceS,maxlen=MAX_LEN,truncating='post',padding='post')
```

```
print(modelL.predict(tweetPad));
```

```
[[0.8227922]]
```

*Figure 8. Example tweet prediction using the trained model.*

# Results and Analysis

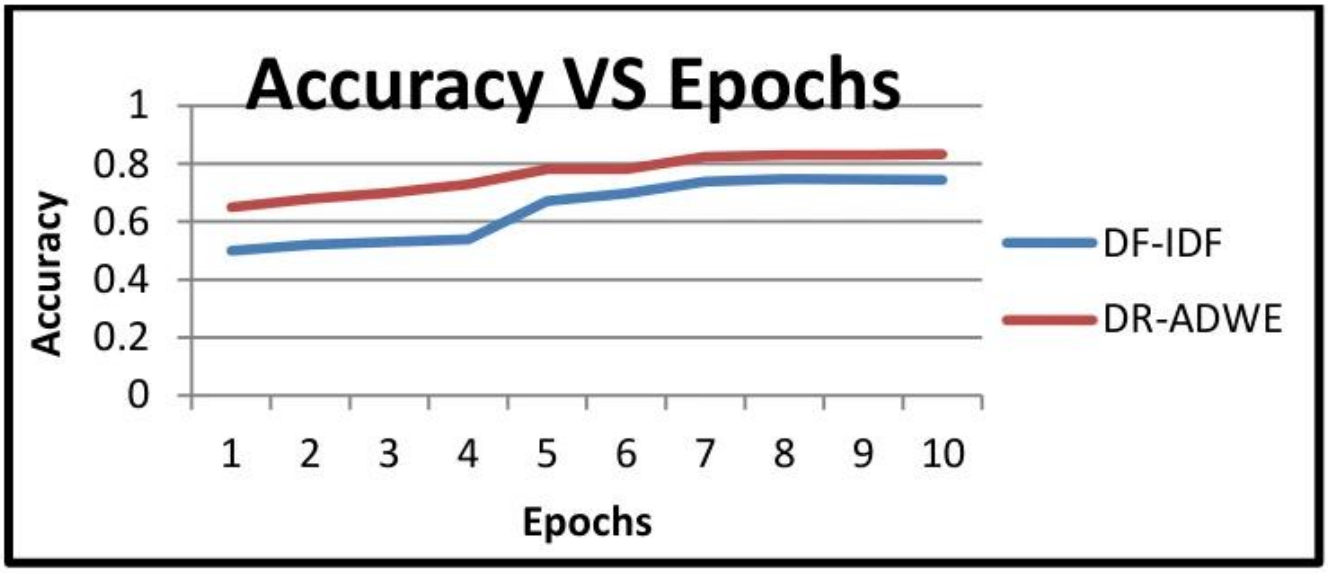



*Figure 9. Accuracy by epoch for the TF-IDF comparison and DR-ADWE. The original chart labels TF-IDF as "DF-IDF".*

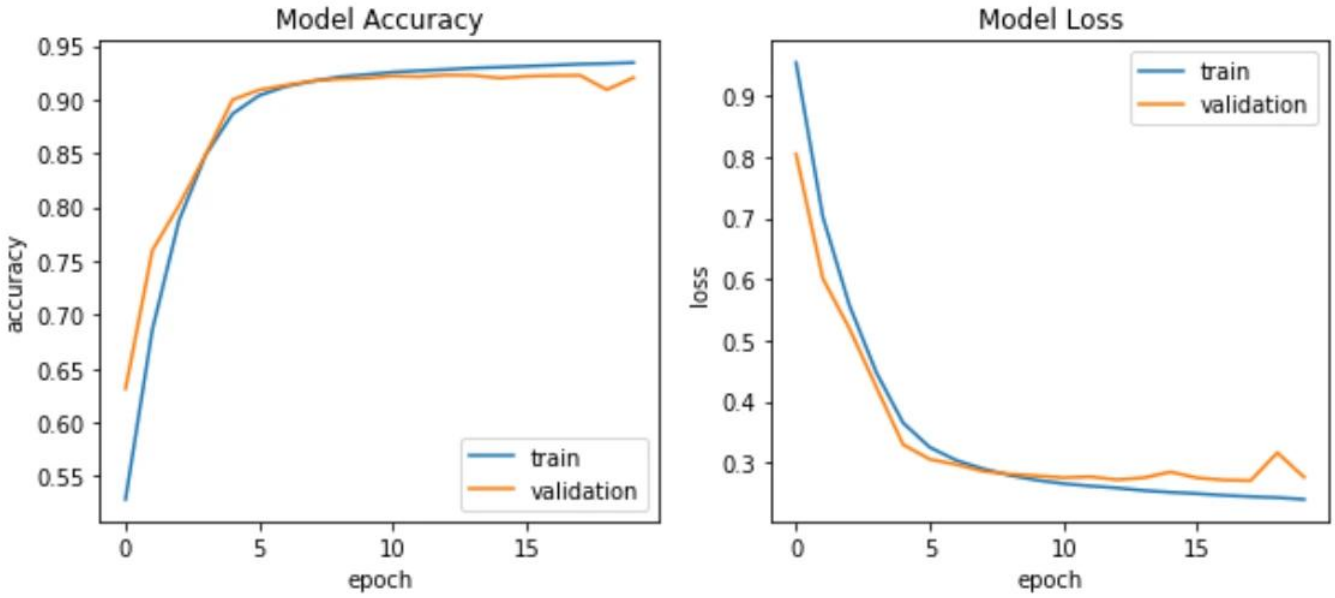



*Figure 10. Training and validation accuracy and loss curves.*

Figure 9 compares the accuracy trends of the TF-IDF representation and the proposed approach,

labelled DR-ADWE in the source. TF-IDF assigns a weighted representation to the terms in each message. It provides useful independent text features, but its representation offers limited semantic information. The embedding-based approach seeks to capture more information about relationships between words. The study calls this approach deep learning processing–atomic word embedding (DR-ADWE) and reports a validation accuracy of 83.33%.

Figure 10 presents the supplied training and validation accuracy and loss curves. These figures are retained as reported in the manuscript. The stated hyperparameter setting is 20 training epochs, while the comparison in Figure 9 displays ten epoch positions. The curves and the stated validation accuracy should therefore be read as the reported experimental material, without assuming that every figure describes an identical evaluation run. The manuscript does not provide enough dataset and split detail to establish independent reproducibility or performance across unseen disaster events.

# Conclusion and Future Work

This paper presents a disaster-tweet classification model using GloVe embeddings for word representation, followed by spatial dropout, LSTM sequence processing and a sigmoid output for binary classification. The architecture is intended to learn relationships between words and identify messages relevant to disasters. The study reports a validation accuracy of 83.33% after hyperparameter tuning. Its prediction example demonstrates how a new message can be processed by the trained model, while the reported figures show the supplied comparisons and learning curves.

Future work could examine tweets from particular subject areas and extend deep learning to review-based analysis. Bidirectional LSTM networks and transformer models such as BERT are also proposed as possible extensions. Further evaluation across clearly documented datasets and disaster contexts would help establish how well the classifier generalises beyond the reported validation setting.

# References

[1]. Yuan Q, Gascó M. Citizens' use of microblogging and government communication during emergencies: A case study on water contamination in Shanghai. First Monday. 2018 Apr 1.

[2] Vieweg, Sarah. "Microblogged contributions to the emergency arena: Discovery, interpretation and implications." Computer Supported Collaborative Work (2010): 515-516.

[3] Wong A, Ho S, Olusanya O, Antonini MV, Lyness D. The use of social media and online communications in times of pandemic COVID-19. Journal of the Intensive Care Society. 2021 Aug;22(3):255-60.

[4] Palen, Leysia, and Sarah Vieweg. "The emergence of online widescale interaction in unexpected

events: assistance, alliance & retreat." Proceedings of the 2008 ACM conference on Computer supported cooperative work. 2008.

[5] Ventola CL. social media and health care professionals: benefits, risks, and best practices. Pharmacy and therapeutics. 2014 Jul;39(7):491.

[6] Castillo, Carlos, Marcelo Mendoza, and Barbara Poblete. "Information credibility on twitter." Proceedings of the 20th international conference on World wide web. 2011

[7] Hart M, Stetten N, Islam S, Pizarro K. Twitter and public health (Part 2): qualitative analysis of how individual health professionals outside organizations use microblogging to promote and disseminate health-related information. JMIR public health and surveillance. 2017 Oct 4;3(4): e6796.

[8] Zahra K, Imran M, Ostermann FO. Automatic identification of eyewitness messages on twitter during disasters. Information processing & management. 2020 Jan 1;57(1):102107.

[9] Tim Y, Pan SL, Ractham P, Kaewkitipong L. Digitally enabled disaster response: the emergence of social media as boundary objects in a flooding disaster. Information Systems Journal. 2017 Mar;27(2):197-232.

[10] Yates, Dave, and Scott Paquette. "Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake." International journal of information management 31.1 (2011): 6-13.

[11] Palen, Leysia, et al. "A vision for technology-mediated support for public participation & assistance in mass emergencies & disasters." ACM-BCS Visions of Computer Science 2010 (2010): 1-12.

[12] Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. "Glove: Global vectors for word representation." Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). 2014.

[13] Das, Bijoyan, and Sarit Chakraborty. "An improved text sentiment classification model using TF-IDF and next word negation." arXiv preprint arXiv:1806.06407 (2018).

[14] Mikolov, Tomas, et al. "Efficient estimation of word representations in vector space." arXiv preprint arXiv:1301.3781 (2013).

[15] ALRashdi, Reem, and Simon O'Keefe. "Deep learning and word embeddings for tweet classification for crisis response." arXiv preprint arXiv:1903.11024 (2019)

[16] Collobert, Ronan, and Jason Weston. "A unified architecture for natural language processing: Deep neural networks with multitask learning." Proceedings of the 25th international conference on Machine learning. 2008.

[17] Sundermeyer, Martin, Ralf Schlüter, and Hermann Ney. "LSTM neural networks for language modeling." Thirteenth annual conference of the international speech communication association. 2012

[18] Mnih, Andriy, and Koray Kavukcuoglu. "Learning word embeddings efficiently with noise-contrastive estimation." Advances in neural information processing systems 26 (2013).

[19] GeeksforGeeks. Introduction to Recurrent Neural Network.