Disaster Response using Artificial Intelligence

By Rabindra Lamsal

Supervisor: Prof. T.V. Vijay Kumar
Thesis submitted on: 12 July, 2019

Below is the abstract of the M.Tech dissertation submitted by Rabindra Lamsal to the School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi.

Disasters, may it be natural or man-made, lead to significant loss of lives and destruction of millions worth of property. Floods, landslides, earthquakes, cyclones, etc., are the commonly occurring calamities in regions like Nepal and India. These calamities turn into major disasters when response systems and disaster risk reduction policies are not designed and implemented effectively. Therefore, dealing with post-disaster phase is of utmost importance when a calamity cannot be predicted before-hand. Disaster Response concerns with rescuing from immediate danger and stabilization of physical and emotional condition of survivors. This phase of a disaster consists of various activities including dissemination of warnings, performing evacuations, search and rescue operations, damage assessment, and providing immediate assistance and fulfilling basic needs of the affected people until a permanent solution is found. The focus of this dissertation is to design effective and efficient disaster response systems using Artificial Intelligence (AI).

Coastal areas are one of the most sensitive and altered ecosystems worldwide, as they are subject to many disasters and risks, including high winds resulting in cyclones, and underwater earthquakes initiating strong tidal waves or Tsunami. These generally occur for a short period of time and affect the community and society as a whole by way of widespread harm or damage to humans, wildlife, environment, infrastructure and economy in the coastal areas. It requires support that encompasses multidimensional processes in order to respond to such disasters. AI offers consummate solutions to develop systems that can play a major role in responding to coastal disasters. In this regard, an EWS, which is capable of predicting the maximum in-city horizontal inundation and accordingly raise the alarm in the areas that are most likely to get affected by it, is proposed. For this, an EWS is designed that is capable of predicting the extent of horizontal inundation, for a given new Tsunamic data instance, by classifying it into classes namely *MinorFloodCoverage*, *ModerateFloodCoverage* and *MajorFloodCoverage*. The classifier is trained, on the *Tsunami* dataset, using widely used classification-based machine learning algorithms, namely Naïve Bayes, Logistic Regression, Random Forests and Artificial Neural Network (ANN). Among

these, *ANN* outperformed all other classification algorithms in terms of Accuracy and F-Measure. The proposed system would enable timely dissemination of, the extent of horizontal inundation, information to communities at risk and the various governmental agencies so that effective and efficient response and adequate measures can be taken to reduce the potential harm or damage that can be caused by inundations.

Twitter, a microblogging platform, receives real-time information via informal conversations, and it has, accordingly, become the main source of data for research studies based on emergency situational awareness. Millions of tweets are posted on Twitter every day, and during disasters, the frequency of tweets relating to an on-going crisis event grows exponentially. This unprecedented increase in the number of tweets during disasters needs to be monitored, identified, processed and analyzed, so that necessary measures can be taken at the earliest to reduce the loss or damage during emergencies. However, due to large voluminous data being available during crisis hours, it is almost impossible for a human to perform these tasks in real-time. *NLP* and machine learning techniques can be used for processing Twitter data. In this regard, a semi-automated *AI*-based disaster response system for Twitter data is proposed. The proposed system monitors the Twitter feed for tweets specific to a crisis event, and whenever a new tweet is retrieved from the near real-time Twitter feed, the proposed disaster response system would be able to translate or transliterate the tweet, if required, pre-process the tweet and finally classify the tweet to the most appropriate class. The proposed system uses the Bag-of-words (BoW) model to create vector space models, based on *Word Count* term-weighting scheme and *TF-IDF* term-weighting scheme. The classifier used by the proposed system is trained, on *QCRI* dataset, using classification-based machine learning algorithms, namely Naïve Bayes, Logistic Regression, Random Forests, k-Nearest Neighbors and *ANN*. Among these, the Logistic Regression classifier outperformed others in terms of both Accuracy and F-Measure. The proposed disaster response system would be capable of extracting essential situational awareness information related to a disaster and would also be capable of sketching a tentative area of the critically affected population.

In the BoW model, the increment in the vocabulary size leads to sparsity in the tweet representation, which is computationally expensive, in terms of both time and space complexity. Also, the BoW model does not preserve the context in which each word is being written in a tweet. To address these limitations, four Deep learning models are trained using word embeddings to classify disaster-related tweets. Each Deep learning model comprises of three components, viz. Embedding layer, Deep network and Output layer. The choices considered for the embedding layer are word embedding training, using the embedding layer, or a pre-trained embedding model. The two variants of *RNN*, viz. *LSTM* and *GRU*, are considered as choices for the Deep network. The experimental results show that the *GRU* network learns faster than the *LSTM* network when the word embeddings are trained using the embedding layer. In the case of pre-trained word embeddings, the learning significantly improves within a few epochs of training, since the embedding layer does not require to learn the word

embeddings thereby requiring lesser number of parameters to be trained. Therefore, it can be inferred that the pre-trained word embeddings produce significant improvements in text classification tasks. Further, the Deep learning models using their own word embeddings were overfitting the training data compared to models using pre-trained word embeddings. Among the Deep learning models, the GRU deep network with pre-trained Embedding performed the best amongst all the other Deep learning models. Further, Deep learning models are shown to perform better than traditional machine learning models. This may be due to the fact that the Deep learning models use dense vector representation of words (word embeddings) in comparison to the sparse representation of words (BoW) used by traditional machine learning models. The use of word embeddings in Deep learning model is capable of preserving the context in which the word is used, thereby improving the semantic comprehensibility, which in turn may have resulted in achieving better accuracy. The proposed Twitter-based disaster response system would enable timely dissemination of information to various stakeholders, so that prompt response and proactive measures can be taken in order to reduce the severe consequences of disasters.