An Ensemble Based Method for the Classification of Flooding Event Using Social Media Data

Muhammad Hanif, Huzaifa Joozer, Muhammad Atif Tahir, Muhammad Rafi
National University of Computer and Emerging Sciences, Karachi Campus, Pakistan
{hanif.soomro,k190880,atif.tahir,muhammad.rafi}@nu.edu.pk

ABSTRACT

This paper presents the method proposed and implemented by team FAST-NU-DS, in "The Flood-related Multimedia Task at MediaEval 2020". The task includes data of tweets in Italian language, extracted during floods between 2017 and 2019. The proposed method has utilized text of the tweet and its relevant image for the purpose of binary classification, which identifies whether or not the particular tweet is about flood incident. The proposed method has designed an ensemble based method for the classification of tweets, on the basis of textual data, visual data and combination of both. For visual data, the proposed method has utilized the technique of data augmentation for oversampling of the minority class and applied stratified random sampling for the selection of input. Moreover, Visual Geometry Group (VGG16) convolutional neural network, pretrained on ImageNet and Places365 is utilized by the proposed method. For classification of textual data, the technique of Term Frequency Inverse Document Frequency (TF-IDF) is utilized for feature representation and Multinomial Naive-Bayes classifier is used for the prediction of class. The prediction of image and text are combined for the prediction of each instance. The evaluation of method revealed 36.31%, 20.76% and 27.86% F1-score for text, image and combination of both, respectively.

1 INTRODUCTION

In the last years, usage of social media is increased by many folds and social media data is emerged as a valuable resource of information. Data retrieved from social media is faster than other mass communication mediums, as its subscribers are available everywhere. Hence, it provides frequent and abundant data, which could be useful in solving various challenging problems, including flooding incidents.

The "Flood-related multimedia task at MediaEval 2020" has focused on usage of social media visual and textual contents for the classification of flooding incidents [2]. In visual data, it has been observed that the dataset related to flood requires appropriate method to deal with class imbalance problem [3–5]. Hence the proposed method has utilized stratified random samples from each class, which are equal in size. Moreover, it is also observed that the ImageNet weights used for pre-trained models may not produce effective results, as ImageNet classification focuses on object level information. In this regard, scene level information along with object level information may produce more accurate results. In textual data, the problem of class-imbalance remains a challenge in prediction of flooding incidents, along with feature representation and classifier selection challenges [1, 9, 11]. Therefore, the proposed method has utilized combination of weights from Places365 [10] and ImageNet [7] for the binary classification of "Flood related multimedia task" [2].

MediaEval'20, December 14-15 2020, Online

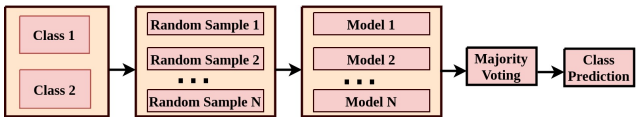

Figure 1: Data flow diagram showing visual data processing for flood related multimedia task

2 APPROACH

The method has been implemented in three phases, initially separate classification is performed by using visual data and text for this binary classification and finally results obtained from text and visual data are combined to generate final outcome. All three phases are discussed below:

2.1 Visual Data

The method of processing visual data for the task of "Flood related multimedia task" [2] is shown in Figure 1.

The dataset of the "Flood related multimedia task" [2] is higly imbalanced as it includes a quantity of 5,419 instances for development-set, out of which only 1126 are related to minority class, which shows indication of flood. Initially, proposed method has utilized technique of data augmentation to oversample the minority class. For that purpose, python based library "Augmentor" [6] has been used to create multiple copies of each image, by varying the parameters of rotate, flip and zoom. Three different copies of each image in minority class are created.

After oversampling of minority class, stratified random sampling is applied for the selection of input samples for training of model. From each of the class, equal quantity of samples (i.e. 3000) are randomly selected. There are (N=15) number of input samples are created.

For the training, Visual Geometry Group (VGG16) [13] Convolutional Neural Network (CNN) is selected. It is launched during ImageNet Large Scale Visual Recognition Challenge (ILSVRC), 2014 and produced effective results in image classification and image localization. The proposed method has utilized hybrid weights of Places365 [10] and ImageNet [7]. The dataset of ImageNet focuses on object based information and it is comprised of 1000 classes. While, images in dataset of Places365 are distributed into 365 classes and its focus is on scene level information. These weights from both ImageNet and Places365 are combined to form hybrid weights,

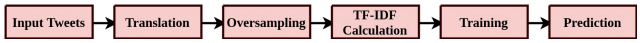

Figure 2: Data flow diagram showing text data processing for flood related multimedia task

having 1365 classes. The VGG16 pre-trained on Hybrid weights is retrained on the dataset for the task of "Flood related multimedia task". During training, different experiments are performed to find out the best combination of hyper-parameters for the retraining of VGG16 (Hybrid), including dropout and learning rate. The dropout value and learning rate are set as 0.3 and 10^{-6}, respectively. The activation function of last layer has been changed from softmax to sigmoid, which is more suitable for binary classification. The last four layers of the model have been retrained, however rest of the layers are freezed so that they could keep their original learning of hybrid dataset. For each of the N input samples, 70% of images are used for training and remaining 30% are used for validation of the model. For each of the training sample, separate model is trained. Later on, each model is utilized for separate prediction.

Each of the trained model is tested by providing unseen images. For each of the image in test-set, N number of predictions are generated from each of the model. Finally, majority voting is applied to find out the class prediction.

2.2 Text

For the processing of textual data, description of the tweets is considered for the classification of each tweet, as shown in Figure 2. The dataset for "Flood related multimedia task" [2] contains tweets in Italian language. So, googletrans [8] library has been used to translate each tweet from Italian language to English language. For that, all the symbols and punctuations are removed from the tweets. Then tweets were translate one by one into English language. Stop words have very little significance in classification of the tweets, so after translation, stop words are removed from each of the tweet. The problem of class imbalance is solved by oversampling the tweets of minority class and make them equal to the quantity of majority class. Subsequently, Term Frequency Inverse Document Frequency (TF-IDF) [12] is calculated for the tweets, and vocabulary of top 5000 is selected for the training. Later on, different classifiers are applied by distributing training set into 80% and 20% each for training and validation, including Support Vector Machine (SVM), Random Forest (RF) and Multinomial Naive Bayes. However, Multinomial Naive Bayes has produced the best results. The prediction of test data is performed through Multinomial Nave Bayes. The prediction result for each of the tweets is extracted as discrete 1 or 0 as well as in the form of probability. The discrete prediction is used for the submission of Run 2 (Textual) data. However, the probabilistic results are used during the average calculation of text and visual data, as shown in Figure 3.

2.3 An ensemble of Text and Visual Data

The probabilistic outcome is extracted from both image and text of each tweet in test-set and averaged to produce final prediction, as shown in Figure 3. For visual information, an ensemble of prediction by fifteen different models is extracted in form of probability

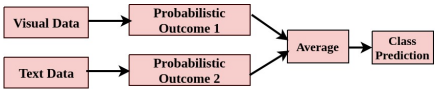

Figure 3: Data Flow diagram showing process of combining visual and textual data for flood related multimedia task

Table 1: Evaluation of proposed method on test-set

Run	F1-score(%)	Average F1-score(%)
Run 1: Text & Visual (Fusion)	27.86	14.15
Run 2: Text	36.31	39.38
Run 3: Visual	20.76	13.18

and averaged to produce probabilistic outcome. Similarly, the outcome received from Multinomial Naive-Bayes classifier is extracted in probabilistic form. Finally, outcomes from text and images are averaged for prediction of each tweet in test-set.

3 RESULTS AND ANALYSIS

An ensemble based method has achieved 20.76% F1-score for images, which is best in all submitted results for MediaEval, 2020 for visual data. It has been observed that in both classes of train set, very similar images are available which contains almost identical maps. This might confuse deep learning model to differentiate between flood and no-flood. Moreover, quantity of images in train-set could be increased by using more versatile techniques of data augmentation. For the text, proposed method has revealed 36.31% F1-score. Results may be improved using an ensemble of multiple classifiers. Moreover, different deep learning based techniques and pre-trained models of the text could be utilized to improve the feature representation and classification. The fusion of text and visual data has produced 27.86% F1-score, which is also best in all submitted results for the fusion of text and visual data. The results for visual, textual and combination of both visual and textual are summarized in table 1.

4 CONCLUSION

The research has proposed an ensemble based approach which has combined deep learning and shallow learning based methods, for the classification of flooding incident by using text and visual data extracted from social media data. The proposed approach has utilized data augmentation for the oversampling of images in minority class. Also, stratified random sample selection is performed, along with transfer learning for the classification of visual data. For textual data, Multinomial Naive-Bayes algorithm is utilized along with the technique of TF-IDF. Both textual and visual results are averaged to produce the combined outcome. In future, method could be improved by using better techniques for increasing the quantity of images and textual data in the minority class of training-set.

ACKNOWLEDGMENTS

This research work was funded by Higher Education Commission (HEC) Pakistan and Ministry of Planning Development and Reforms under the National Center in Big Data and Cloud Computing.

REFERENCES

[1] Sheharyar Ahmad, Kashif Ahmad, Nasir Ahmad, and Nicola Conci. 2017. Convolutional Neural Networks for Disaster Images Retrieval.. In *Proc. of the MediaEval 2017 Workshop*.

[2] Stelios Andreadis, Ilias Gialampoukidis, Anastasios Karakostas, Stefanos Vrochidis, Ioannis Kompatsiaris, Roberto Fiorin, Daniele Norbiato, and Michele Ferri. 2020. The Flood-related Multimedia Task at MediaEval 2020. (2020).

[3] Bischke Benjamin, Helber Patrick, Zhao Zhengyu, Bruijn Jens de, and Borth Damian. 2018. The multimedia satellite task at mediaeval 2018: Emergency response for flooding events. (October 2018).

[4] Benjamin Bischke, Patrick Helber, Erkan Basar, Simon Brugman, Zhengyu Zhao, and Konstantin Pogorelov. 2019. The multimedia satellite task at mediaeval 2019: Flood severity estimation. In *Proc. of the MediaEval 2019 Workshop*.

[5] Benjamin Bischke, Patrick Helber, Christian Schulze, Srinivasan Venkat, Andreas Dengel, and Damian Borth. 2017. The multimedia satellite task at mediaeval 2017: Emergence response for flooding events. (2017).

[6] Marcus D Bloice, Christof Stocker, and Andreas Holzinger. 2017. Augmentor: an image augmentation library for machine learning. *arXiv preprint arXiv:1708.04680* (2017).

[7] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*. IEEE, 248–255.

[8] Suhun Han. 2020. googletrans 3.0.0. https://pypi.org/project/googletrans/. (2020). Accessed: 2020-11-1.

[9] Muhammad Hanif, Muhammad Atif Tahir, Mahrukh Khan, and Muhammad Rafi. 2017. Flood detection using Social Media Data and Spectral Regression based Kernel Discriminant Analysis. In *Proc. of the MediaEval 2017 Workshop, Dublin, Ireland*.

[10] Ryan Lagerstrom, Yulia Arzhaeva, Piotr Szul, Oliver Obst, Robert Power, Bella Robinson, and Tomasz Bednarz. 2016. Image classification to support emergency situation awareness. *Frontiers in Robotics and AI* 3 (2016), 54.

[11] Keiller Nogueira, Samuel G Fadel, Ícaro C Dourado, Rafael de Oliveira Werneck, Javier AV Muñoz, Otávio AB Penatti, Rodrigo Tripodi Calumby, Lin Li, Jefersson Alex dos Santos, and Ricardo da Silva Torres. 2017. Data-Driven Flood Detection using Neural Networks. In *Proc. of the MediaEval 2017 Workshop*.

[12] Gerard Salton and Christopher Buckley. 1988. Term-weighting approaches in automatic text retrieval. *Information processing & management* 24, 5 (1988), 513–523.

[13] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).