Literature Review

In the paper \citep{balabantaray2012multi}, the authors conducted an emotion analysis on twitter data. They have annotated the data with the help of five judges. The judge has to find whether the particular sentence belong to any emotion category and then label it accordingly. They introduced a neutral category of non-emotional sentence along with six different categories. Because of five different judge, the results for the same sentence can be subjective. Hence to overcome this, they used Cohen’s kappa metric to compare the consensus between each judge. They use SVM multiclass classifier for their problem. They used several features to identify the emotion bearing words. After preprocessing and model building, they finally achieved an accuracy of 73.24 percent.

@article{balabantaray2012multi,

title={Multi-class twitter emotion classification: A new approach},

author={Balabantaray, Rakesh C and Mohammad, Mudasir and Sharma, Nibha},

journal={International Journal of Applied Information Systems},

volume={4},

number={1},

pages={48--53},

year={2012},

publisher={Citeseer}

}

The authors in the paper \citep{ezhilarasi2012automatic}, developed an approach to create an emotion based ontology. The used UIUC children’s story set as a dataset or corpus. Their emotional ontology is based on WordNet. Their way of classification is to first develop an ontology of emotion class. This is semi-automated using WordNet. For reference, WordNet is a lexical dictionary with explains the relationship between two words in the form of conceptual and attribute. This is available as a network. After ontology creating. Their next step is into extract terms that can help in identifying emotional class. Finally using natural language processing, the text or query can be classified with the help of ontology.

@article{ezhilarasi2012automatic,

title={Automatic emotion recognition and classification},

author={Ezhilarasi, R and Minu, RI},

journal={Procedia Engineering},

volume={38},

pages={21--26},

year={2012},

publisher={Elsevier}

}

This is also related to emotion classification problem. But here, the author uses deep learning architectures for classification. In their paper \citep{illendula2019multimodal}, the authors have use multi model way to classify the emotion posts. Multi model way is basically using all forms modalities. Rather than only using textual or visual features. The combine all the available mode and were able to outperform the other unimodal approaches used. They had used Instagram for collecting data as it is a good source for multi-model data. Total of six different emotional categories are used for data collection. They used fast-text embedding model to encode textual data. Finally, after preprocessing, emoji features and textual features are collected. It is then passed to a classifier to produce an output. They have achieve around 71 percent accuracy. The total amount of training data used is 550 thousand.

@inproceedings{illendula2019multimodal,

title={Multimodal emotion classification},

author={Illendula, Anurag and Sheth, Amit},

booktitle={Companion Proceedings of The 2019 World Wide Web Conference},

pages={439--449},

year={2019}

}

In the paper \citep{esmin2012hierarchical}, the authors used twitter as a platform to solve a problem of classifying emotions. They use hierarchical approach to classify tweets. The hierarchy consists of three levels. First, they find whether the tweet are emotional or not. If the tweet contains any sentiment for emotion that means the tweet can be considered as emotional. After that, in the next step they derive the polarity of the particular tweet. This will be used for eliminating the emotions which represents negative emotions. Total of five emotional categories are used. The dataset used here is specific to the soccer matches played in Brazilian league in 2011. They had used SVM classifier as a machine learning model. The paper uses two approaches. The first one is flat classification and the other is the hierarchy based approach. They concluded that, the second approach performs significantly better in the classification problem.

@inproceedings{esmin2012hierarchical,

title={Hierarchical classification approach to emotion recognition in twitter},

author={Esmin, Ahmed AA and De Oliveira Jr, Roberto L and Matwin, Stan},

booktitle={2012 11th International Conference on Machine Learning and Applications},

volume={2},

pages={381--385},

year={2012},

organization={IEEE}

}

In the final paper \citep{sasidhar2020emotion}, the authors solved the emotion classification problems on English plus hindi code mix dataset. The code mix here mean, writing text as English but adding some words or sentences from hindi language. The dataset is collected from many social media websites like twitter, Facebook etc. Total of around 12000 values are collected for various emotions like happy, sad and anger. For training, they had used a bilingual pre-trained model. It was trained again using their corpus. In summary, they had used CNN-bi LSTM model. The CNN for generating meaningful word embedding and LSTM for classification. Finally they got an accuracy of 83.21 percent.

@article{sasidhar2020emotion,

title={Emotion detection in hinglish (hindi+ english) code-mixed social media text},

author={Sasidhar, T Tulasi and Premjith, B and Soman, KP},

journal={Procedia Computer Science},

volume={171},

pages={1346--1352},

year={2020},

publisher={Elsevier}

}