Technical Report : Natural Language Engineering Lab

Garima Jain

July 20, 2019

Abstract

This report contains the literature survey for the papers that I have read related to Emotion detection in text.

1 Corpus creation and emotion prediction for hindi english code mixed social media text"[1]

1.1 Proposed work

- 3 classes:
  - Happiness
  - Sadness
  - Anger
- Classes were annotated manually.To validate quality, calculated Inter-Annotator Agreement using Kohen's cappa coefficient
- 4 types of causal language annotations: Hindi, Eng, Mixed, Both
- Manual annotation again
- Total 2140 tweets
- 3 sub-processes:
  - Pre-processing
  - Feature identification and extraction
  - Classification using SVM. No explanation for model training
- Identified important features:
  - Char n-grams
  - Word n-grams
  - Punctuation marks
  - Emoticons
  - Upper case words

1.2 Future directions as mentioned:

- Use more classes
- Annotate corpus with PoS tags at word level

1.3 Results

1.4 Flaws with the work

- Sarcasm present in tweets can cause wrong annotation of tweets. Like they mention, that they have used emoticon as one of the features and that if a tweet contains :), then it expresses happiness, but it might not be the case.

Feature Eliminated	Accuracy
None	58.2
Emoticons	58.1
Char N-Grams	42.9
Word N-Grams	57.6
Repetitive Characters	58.2
Punctuation Marks	57.4
Upper Case Words	58.2
Intensifiers	58.2
Negation Words	58.2
Lexicon	57.9

Figure 1: Results

1.5 Some ideas:

- Dataset generation: Use their small dataset, train the model and further annotate new data using Active Learning. Thus contribute with bigger dataset.
- Model training: Train language model as pre-training step as researchers are exploring the same in other text classification tasks.

1.6 References

1. Vijay, Deepanshu, et al. "Corpus Creation and Emotion Prediction for Hindi-English Code-Mixed Social Media Text." Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop. 2018.

2 Emotion Detection in Text : A Review[1]

2.1 Introduction

- More important to recognise emotions and not just classifying positive and negative sentiments(sentiment analysis)
- Not all negative and positive sentiments are created equal
- Most works in emotion detection use [1] model for classifying emotion as one of 6 classes: sadness, happiness, anger, fear, disgust, and surprise.

2.2 Resources for detecting emotions in text:

- Labeled text
  - ISEAR survey dataset provided by SCA
  - EmotiNet knowledge base
  - Sem-Eval 2007: 1250 news headlines extracted from news websites - annotated with Ekman's 6 emotions
  - Alm's fairy tale dataset: 1580 sentences from children fairy tales-annotated with Ekman's 6 emotions
  - 2.5mn annotated Twitter tweets annotated with hashtags and emoticons[2]
- Emotion lexicon : when word based analysis is required
  - WordNet- Affect
  - Depeche Mood
  - Word embeddings
- Sentiment-specific word embeddings have been explored in [3]
- approaches for creating emotional word embeddings from scratch, or incorporating emotional information into pre-trained word vectors after the fact, might lead to better performances in emotion detection tasks, either in unsupervised methods, or as features for classification tasks using conventional machine learning, or deep learning

2.3 Supervised approaches:

4 used four pairs of opposite emotions in the Plutchik's wheel to create four binary classification tasks. With hashtags, emoticons, and emoji as labels for their data, they reached between 75% to 91% accuracy on a separate manually labeled dataset.

5 on Twitter data using SVM classifier reached 82% accuracy for classifying the emotion Happy in 10-fold cross validation, and 67% in classifying over the entire dataset for the same emotion, with emoticons as labels for the training set, and hashtags as labels for the test set.

- Concluded that hashtags and emoticons as labels is a promising labeling strategy and and alternative to manual labeling.

6 also used hashtags as label for tweets, and used support vector machines as a binary classifier for each emotion in Ekman's model.

7 collected tweets in 14 topics that "would frequently evoke emotion" and created a dataset where all seven emotions (Ekman + Love) were represented. Seven SVM binary classifiers were used to detect emotions in the dataset, resulting in the average F1-score of 0.66.

8 also used hashtags as their labels and created their features using the unigram model, removing any word from tweets which were not in their emotion lexicon (created using 28 basic emotion word in Circumplex model and extended with WordNet synsets). Four classifiers (Naive Bayes, SVM, Decision Tree, and KNN) achieved accuracies close to 90% in classifying four main classes of emotion categories in Circumplex model.

- In another paper, [9] created an automatic emotion detection system to identify emotions in streams of tweets. This approach included two tasks:
  - training an offline emotion classification model based on their 2014 paper,
  - and in the second part a two step classification to identify tweets containing emotions, and to classify these emotional tweets into more fine-grained labels using soft classification techniques.

10 created a large dataset (about 2.5 million tweets) using emotion related hashtags, and used two machine learning algorithms for emotion identification.

- They used [11] for mapping hashtags to emotions, and extending hashtag words to the total of 131 for the seven basic emotions.

- tried different combinations of features (e.g. different n-grams, position of n-grams, multiple lexicon, POS) with 250k of the training data to find the best set of features, with the best result for the combination of n-gram(n=1,2), LIWC lexicon, MPQA lexicon, WordNet-Affect, and POS.

- In another Twitter emotion classification task done by [12] manual labeling was used for around 8000 tweets, for six basic emotions in Ekman's model.

  - They used SVM multi-class classifier(?) with 11 features: Unigrams, Bigrams, Personal-pronouns, Adjectives, Word-net Affect lexicon, Word-net Affect lexicon with left/right context, Word-net Affect emotion POS, POS, POS-bigrams, Dependency-Parsing, and Emoticons resulting in an accuracy of 73.24%.

- In their paper, [13] attempted sentence level classification of emotion instead of document level.

  - They indicated that the two biggest problems in sentence level emotion classification is

    * firstly the fact that it is a multi-class classification, meaning that each sentence could have more than one label,

    * and secondly, the short length of a sentence, provides less content.

  - Considering these challenges they created a Dependence Factor Graph (DFG) based on two observations, label dependence, i.e. multiple labels for a sentence would be correlated to one another, like Joy and Love instead of Joy and Hate, and context dependence, i.e. two neighboring sentences, or sentences in the same paragraph might share the same emotion categories.

  - Using the DFG model, after learning they reached the accuracy of 63.4% with F1 of 0.37 showing significant improvement over previous methods

- comparative analysis done by [14]. They compared various classification features and comparedthem to EmotiNet, and concluded that the task of emotion detection can be best tackled using approaches based on commonsense knowledge. They showed that even with the small size of EmotiNet knowledge base they could produce comparative results to supervised learning methods with huge amount of training data.

- Methods proposed to tackle imbalanced data problem:

  - Changing the learning algorithm to adapt to this imbalance [15]

- Adding cost to majority classes during training [16]
- Sampling from trg data before learning to obtain balanced data [17]

18 proposed over-sampling based on word embeddings and recursive neural networks

2.4 Based on unsupervised approaches:

19 used an unsupervised method to automatically detect emotions in text, based on both categorical (anger, fear, joy and sadness), and dimensional models of emotions.

- They used three datasets, SemEval-2007 Affective Text, ISEAR, and childrens fairy tales.
- For categorical model, they used WordNet-Affect as the lexicon, and evaluated three dimensionality reduction methods: Latent Semantic Analysis (LSA), Probabilistic Latent Semantic Analysis (PLSA), and Non-negative Matrix Factorization (NMF).
- for the dimensional model, they used ANEW (Affective Norm for English Words) and WordNet-Affect as a means to extend ANEW.
- They assigned the emotion of the text based on closeness (cosine similarity) of its vector to the vectors for each category or dimension of emotions.
- Their study showed that NMF-based categorical classification performs best among categorical approaches, and dimensional model had the second best performance with highest F-measure of 0.73.

- Another unsupervised approach to emotion detection can be seen in the paper by [20].
  - They start by extracting NAVA words (i.e. Nouns, Adjectives, Verbs,and Adverbs) from a sentence, and then extracting syntactic dependencies(?) between extracted words in each sentence to include contextual information in their model.
  - They then used semantic relatedness to compute emotion vectors for words, based on the assumption that the affect words (NAVA words) which co-occur together more often tend to be semantically related.
  - They use Point-wise Mutual Information (PMI) as the measure of semantic relatedness of two words(?)

21 used domain-specific lexicon that they created based on unigram mixture models to extract features and showed that their lexicon outperform methods like Point-wise Mutual Information, and supervised Latent Dirichlet Allocation.

22 used an unsupervised method to distinguish language pattern related to anxiety in online health forums. They define user behavioral dimension (BD) based on the LIWC lexicon focusing on its anxious word list.

2.5 Discussion and open problems:

- Difficulties due to complex nature of emotion expression in text: crucial to consider the contextual information in which the expression is occurring.
- inefficiency of current emotion detection models
- BoW representations consider the flow and composition of language
- Specific neural network designs and ensemble method approaches have shown to be successful in other NLP tasks
- New ways to increase the emotional qualities of embeddings and vector models could be beneficial in unsupervised methods, or be used as features in neural networks.
- lack of high quality data to be utilized by those models: any attempt to create a large balanced dataset, with high quality labels could provide a brighter future for the field.

2.6 General Observations

- Different approaches have created their own datasets and with different methods, achieving different accuracy on different emotions, owing to unbalanced distribution of tweets of different emotions
- Emotional vector embeddings can be explored

2.7 References

1. Seyeditabari, Armin, Narges Tabari, and Wlodek Zadrozny. "Emotion Detection in Text: a Review." arXiv preprint arXiv:1806.00674 (2018).
2. Paul Ekman. 1992. An argument for basic emotions. Cognition & emotion, 6(3-4):169–200.
3. Wenbo Wang, Lu Chen, Krishnaprasad Thirunarayan, and Amit P Sheth. 2012. Harnessing twitter" big data" for automatic emotion identification. In Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom), pages 587–592. IEEE.
4. Duyu Tang, Furu Wei, Nan Yang, Ming Zhou, Ting Liu, and Bing Qin. 2014b. Learning sentiment-specific word embedding for twitter sentiment classification. In ACL (1), pages 1555–1565.

5. Jared Suttles and Nancy Ide. 2013. Distant supervision for emotion classification with discrete binary values. In International Conference on Intelligent Text Processing and Computational Linguistics, pages 121–136. Springer.

6. Matthew Purver and Stuart Battersby. 2012. Experimenting with distant supervision for emotion classification. In Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, pages 482–491. Association for Computational Linguistics.

7. Saif M Mohammad. 2012a. emotional tweets. In Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation, pages 246–255. Association for Computational Linguistics.

8. Kirk Roberts, Michael A Roach, Joseph Johnson, Josh Guthrie, and Sanda M Harabagiu. 2012.Empatweet: Annotating and detecting emotions on twitter. In LREC, volume 12, pages 3806–3813. Citeseer.

9. Maryam Hasan, Elke Rundensteiner, and Emmanuel Agu. 2014. Emotex: Detecting emotions in twitter messages.

10. Maryam Hasan, Elke Rundensteiner, and Emmanuel Agu. 2018. Automatic emotion detection in text streams by analyzing twitter data. International Journal of Data Science and Analytics, Feb.

11. Wenbo Wang, Lu Chen, Krishnaprasad Thirunarayan, and Amit P Sheth. 2012. Harnessing twitter" big data" for automatic emotion identification. In Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom), pages 587–592. IEEE.

12. Phillip Shaver, Judith Schwartz, Donald Kirson, and Cary O'connor. 1987. Emotion knowledge: Further exploration of a prototype approach. Journal of personality and social psychology, 52(6):1061.

13. Rakesh C Balabantaray, Mudasir Mohammad, and Nibha Sharma. 2012. Multi-class twitter emotion classification: A new approach. International Journal of Applied Information Systems, 4(1):48–53.

14. Shoushan Li, Lei Huang, Rong Wang, and Guodong Zhou. 2015. Sentence-level emotion classification with label and context dependence. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume1: Long Papers), volume 1, pages 1045–1053.

15. Alexandra Balahur, Jes M Hermida, and Andres Montoyo. 2012. Detecting implicit expressions of emotion in text: A comparative analysis. Decision Support Systems, 53(4):742–753.

16. Yuchun Tang, Yan-Qing Zhang, Nitesh V Chawla, and Sven Krasser. 2009. Svms modeling for highly imbalanced classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 39(1):281–288.

17. Yanmin Sun, Mohamed S Kamel, Andrew KC Wong, and Yang Wang. 2007. Cost-sensitive boosting for classification of imbalanced data. Pattern Recognition, 40(12):3358–3378.

18. Nitesh V Chawla, Nathalie Japkowicz, and Aleksander Kotcz. 2004. Special issue on learning from imbalanced data sets. ACM Sigkdd Explorations Newsletter, 6(1):1–6.

19. Ruifeng Xu, Tao Chen, Yunqing Xia, Qin Lu, Bin Liu, and Xuan Wang. 2015. Word embedding composition for data imbalances in sentiment and emotion classification. Cognitive Computation, 7(2):226-240.

20. Sunghwan Mac Kim, Alessandro Valitutti, and Rafael A Calvo. 2010. Evaluation of Unsupervised Emotion Models to Textual Affect Recognition. In Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, CAAGET '10, pages 62–70, Stroudsburg, PA, USA. Association for Computational Linguistics.

21. Ameeta Agrawal and Aijun An. 2012. Unsupervised emotion detection from text using semantic and syntactic relations. In Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 01, pages 346–353. IEEE Computer Society.

22. Nicolas Rey-Villamizar, Prasha Shrestha, Farig Sadeque, Steven Bethard, Ted Peder- sen, Arjun Mukherjee, and Thamar Solorio. 2016. Analysis of anxious word usage on online health forums. EMNLP 2016, page 37.

3 Understanding emotions in text using deep learning and big data[1]

3.1 Introduction

- Deep learning based approach to detect emotions(4 classes : happy, sad, angry, other) in textual dialogues

- The essence of their approach lies in combining both semantic and sentiment based representations for more accurate emotion detection.

- proposed an end-to-end trainable deep learning model, called "Sentiment and Semantic Based Emotion Detector (SS-BED)" for detecting emotions in textual dialogues.

- By combining sentiment of different words with semantics of the sentence, emotion can be detected in various cases

- Their research is in the context of an online chatbot, designed for informal conversations with users

3.2 Proposed approach

- Model Architecture
  - Input user utterance is fed into 2 LSTM layers using 2 different word embedding matrices(one layer uses sentiment word embedding while other uses semantic word embedding)
  - These two layers learn semantic and sentiment feature representation and encode sequential patterns in the user utterance.
  - These two feature representations are then concatenated and passed to a fully connected network with one hidden layer which models interactions between these features and outputs probabilities per emotion class.

- Data collection
  - Collected 17.62 mn conversational pairs i.e tweets and their responses (Twitter Q-A pairs) extracted from Twitter Firehose
  - Followed 2 techniques for data collection (??)

- Choosing input embeddings
  - Semantic word rep : Word2vec[2], GloVe[3], FastText[4]
  - Sentiment word rep : SSWE(sentiment specific word embeddings) : encoding sentiment information in continuous representation of words[5]
  - chose GloVe and SSWE as their embedding for Semantic and Sentiment LSTM layer respectively.
- Tuning the network
  - Tried different combinations of hyper parameters
  - Performed 10 fold cross validation
- Model training
  - Used Microsoft Cognitive Toolkit
  - Used dropout
  - Cross-entropy with softmax as loss function
  - SGD learner

3.3 Experiment and Results

- dataset comprises of 2226 3-turn conversations along with their emotion class labels (Happy,Sad, Angry, Others) provided by human judges.

- Final inter-annotator agreement based on fleiss' kappa value[6] is found to be 0.59. This kappa value, while slightly less then desirable, indicates the difficulty in judging textual conversations due to ambiguities

- Experimented with 2 sets of features:

  – N-grams and emoticons

  – combination of n-grams, wordNet Affect-presence, SSWE, PoS, Emoticons, Misc., Negations

emotions DL.png emotions DL.png

Table 6
Comparison of various models on evaluation dataset. SS-BED results are statistically significant with $p < 0.005$.

	Happy			Sad			Angry			MACRO	MICRO
	PRECISION	RECALL	F1	PRECISION	RECALL	F1	PRECISION	RECALL	F1	F1	F1
NB (Feat-1)	41.35	50.46	45.45	70.87	68.22	69.52	38.16	32.22	34.94	49.97	50.81
SVM (Feat-1)	66.67	25.69	37.09	86.49	59.81	70.71	85.42	45.56	59.42	55.74	56.59
GBDT (Feat-1)	75.76	22.94	35.21	89.47	63.55	74.31	86	47.78	61.43	56.98	58.49
NB (Feat-2)	43.27	57.36	49.33	70.83	69.1	69.95	68.26	42.96	52.73	57.34	57.42
SVM (Feat-2)	73.33	33.79	46.26	87.02	61.23	71.88	86.73	46.33	60.39	59.51	60.42
GBDT (Feat-2)	78.46	25.00	37.92	94.25	58.92	72.51	88.98	50.26	64.23	58.22	58.95
CNN-NAVA	63.32	42.29	50.71	79.37	68.69	73.64	67.42	45.79	54.54	59.63	60.15
CNN-SSWE	67.69	40.37	50.57	77.45	73.83	75.6	80.95	37.77	51.51	59.23	60.97
CNN-GloVe	52.29	52.29	52.29	93.72	67.29	74.61	67.82	65.55	66.66	64.52	64.93
LSTM-SSWE	70.69	37.61	49.1	83.87	72.89	78	73.24	57.77	64.6	63.9	64.77
LSTM-GloVe	64.18	39.45	48.86	72.88	80.37	76.44	72.15	63.33	67.45	64.25	65.26
SS-BED	69.51	52.29	**59.68**	85.42	76.63	**80.79**	87.69	63.33	**73.55**	**71.34**	**71.4**

Figure 2: Results

3.4 References

1. Chatterjee, Ankush, et al. "Understanding emotions in text using deep learning and big data." Computers in Human Behavior 93 (2019): 309-317.

2. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., Dean, J. (2013). Distributed representations of words and phrases and their compositionality. Advances in neural information processing systems (pp. 3111–3119). .

3. Pennington, J., Socher, R., Manning, C. D. (2014). Glove: Global vectors for word representation. EMNLP. Vol. 14. EMNLP (pp. 1532–1543).

4. A. Joulin, E. Grave, P. Bojanowski, T. Mikolov, Bag of tricks for efficient text classifica- tion, arXiv preprint arXiv:1607.01759.

5. Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T., Qin, B. (2014). Learning sentiment specific word embedding for twitter sentiment classification. Proceedings of the 52nd annual meeting of the association for computational linguistics (pp. 1555–1565). .

6. P. E. Shrout, J. L. Fleiss, Intraclass correlations: Uses in assessing rater reliability., Psychological Bulletin Vol. 86, page 420.

4 Learning Sentiment Specific word embedding for Twitter sentiment classification.[1]

4.1 Introduction

- most existing word representations only model syntactic context of words but ignore sentiment of text
- problematic for sentiment analysis as they usually map words with similar syntactic context but opposite sentiment polarity, such as good and bad, to neighboring word vectors
- addressed this issue by learning SSWE (sentiment specific word embedding)
- developed 3 neural networks to incorporate supervision from sentiment polarity of text in their loss functions from massive distant-supervised tweets without any manual annotations
- evaluated on benchmark dataset in SemEval 2013
- yields 84.89% in macro-F1 only by using SSWE as feature
- concatenating SSWE with existing feature set, got 86.58% (state-of-the-art)
- released SSWE learned from 10mn tweets
- learning continuous representations for sentiment classification
  - ,[2] initialize the word embedding by Latent Semantic Analysis and further represent each document as the linear weighted of ngram vectors for sentiment classification.
  - ,[3] model each word as a matrix and combine words using iterated matrix multiplication.
  - ,[4]explore Stacked Denoising Autoencoders for domain adaptation in sentiment classification.
  - ,[5] propose Recursive Neural Network (RNN) (2011b), matrix-vector RNN (2012) and Recursive Neural Tensor Network (RNTN) (2013b) to learn the compositionality of phrases of any length based on the representation of each pair of children recursively
  - ,[6] present Combinatory Categorial Autoencoders to learn the compositionality of sentence, which marries the Combinatory Categorial Grammar with Recursive Autoencoder.

4.2 Proposed Approach

- extended existing word embedding learning algorithm [7]
- it does not explicitly capture sentiment information in texts
- Sentiment specific word embedding
  - Models
    - Model 1
      - slide the window of n-gram across the sentence(sentences might be of varying lengths)
      - predict the sentiment polarity based on each n-gram using a shared neural Network
      - utilise continuous vector of top layer to predict sentiment distribution of text
      - Assuming there are K labels, they modified the dimension of top layer in C&W model as K and add a softmax layer upon the top layer
      - further maths part, not understood
    - Model 2
      - constraint of Model 1 too strict, this is it's relaxed version
      - predicted positive sentiment score of a tweet is expected to be greater than predicted negative sentiment score if the tweet's sentiment polarity is positive and vice-versa
      - maths part, not understood
    - Model 3(Unified of above 2)
      - captures sentiment information of tweets as well as syntactic context of words
      - predicts 2d vector for each input n-gram : the 2 scalars are Language model score and sentiment score
      - 2 training objectives: original n-gram should obtain higher lang model score than corrputed n-gram and sentiment score of original ngram should be more consistent with the gold polarity annotation(??) of sentence than corrupted ngram.
      - loss function given, not understood
- Model training
  - train the 3 models using backpropagation wrt all parameters[7]
  - used Ada-grad to update parameters[8]
  - various hyper-parameters are specified
  - learnt embeddings for unigram, bigram, trigram

4.3 References

1. Tang, Duyu, et al. "Learning sentiment-specific word embedding for twitter sentiment classification." Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2014.

2. Dmitriy Bespalov, Bing Bai, Yanjun Qi, and Ali Shokoufandeh. 2011. Sentiment classification based on supervised latent n-gram analysis. In Proceedings of the Conference on Information and Knowledge Management, pages 375–382. Dmitriy Bespalov, Yanjun Qi, Bing Bai, and Ali Shokoufandeh. 2012. Sentiment classification with supervised sequence embedding. In Machine Learning and Knowledge Discovery in Databases pages 159–174. Springer.

3. Ainur Yessenalina and Claire Cardie. 2011. Compositional matrix-space models for sentiment analysis. In Proceedings of Conference on Empirical Methods in Natural Language Processing, pages 172–182.

4. Xavier Glorot, Antoine Bordes, and Yoshua Bengio.2011. Domain adaptation for large-scale sentiment classification: A deep learning approach. Proceedings of International Conference on Machine Learning.

5. Richard Socher, Cliff C Lin, Andrew Ng, and Chris Manning. 2011b. Parsing natural scenes and natural language with recursive neural networks. In Proceedings of the International Conference on Machine Learning, pages 129–136.

6. Karl Moritz Hermann and Phil Blunsom. 2013. The role of syntax in vector space models of compositional semantics. In Proceedings of the Annual Meeting of the Association for Computational Linguistics, pages 894–904.

7. Ronan Collobert, Jason Weston, Leon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa.2011. Natural language processing (almost) from scratch. Journal of Machine Learning Research, 12:2493–2537.

8. John Duchi, Elad Hazan, and Yoram Singer. 2011.Adaptive subgradient methods for online learning and stochastic optimization. The Journal of Machine Learning Research, pages 2121–2159.