



HAL
open science

Emotion Recognition in Sentences - A Recurrent Neural Network Approach

N. S. Kumar, Mahathi Amencherla, Manu George Vimal

► **To cite this version:**

N. S. Kumar, Mahathi Amencherla, Manu George Vimal. Emotion Recognition in Sentences - A Recurrent Neural Network Approach. 3rd International Conference on Computational Intelligence in Data Science (ICCIDS), Feb 2020, Chennai, India. pp.3-15, 10.1007/978-3-030-63467-4_1 . hal-03434798

HAL Id: hal-03434798

<https://inria.hal.science/hal-03434798v1>

Submitted on 18 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Emotion Recognition in Sentences- A Recurrent Neural Network Approach

N S Kumar^[0000-0001-8831-0625], Mahathi Amencherla^[0000-0001-7575-5457], and
Manu George Vimal^[0000-0001-7957-1890]

PES University, Bangalore, India nskumar@pes.edu, {[mahathiamencherla15](mailto:mahathiamencherla15@gmail.com),
[manugeorge004](mailto:manugeorge004@gmail.com)}@gmail.com

Abstract. This paper presents a Long short-term memory (LSTM) model to automate the tagging of sentences with a dominant emotion. LSTM is a type of artificial recurrent neural network (RNN) architecture. A labeled corpus of news headlines[1] and textual descriptions[3] are used to train our model. The data set is annotated for a set of six basic emotions: Joy, Sadness, Fear, Anger, Disgust and Surprise and also takes into consideration a 7th emotion, Neutral, to adequately represent sentences with incongruous emotions. Our model takes into account that one sentence can represent a conjunct of emotions and resolves all such conflicts to bring out one dominant emotion that the sentence can be categorized into. Furthermore, this can be extended to categorize an entire paragraph into a particular emotion. Our model gives us an accuracy of 85.63% for the prediction of emotion when trained on the above mentioned data set and an accuracy of 91.6% for the prediction of degree of emotion for a sentence. Additionally, every sentence is associated with a degree of the dominant emotion. One can infer that a degree of emotion means the extent of the emphasis of an emotion. Although, more than one sentence conveys the same emotion, the amount of emphasis of the emotion itself can vary depending on the context. This feature of determining the emphasis of an emotion, that is, degree of an emotion, is also taken care of in our model.

Keywords: Sentiment Analysis · Sentence Level Emotion Analysis · Natural Language Processing · Recurrent Neural Networks · Long Short-Term Memory · Emotion Recognition · Degree of Emotion

1 Introduction

A single distinct element of speech or writing is a word. Every word carries a meaning, a string of words can be combined to form a sentence that is not only meaningful but also expressive and can be associated with a particular state of mind, that is, emotion.

A lot of work has been carried out for the sentiment analysis of text wherein the text is classified into one of the three categories: positive, negative and neutral. This has been mostly achieved by the use of lexical resources such as the SentiWordNet that assigns each word with a positive, negative or neutral tag.

This approach only considers the individual words that constitute the sentence and not the relative positioning of the words amongst themselves, or to the context of the sentence. This process analysis text at a rudimentary level and also does not identify the particular emotion. As a sentence tagged negative encompasses a large variety of emotions such as sorrow, anger and fear and the above methodology does not differentiate amongst them and does not help us to discern the emotion being conveyed in a sentence. Although the sentiment analysis of text has its own relevance as a market analysis tool we opine that we could further benefit by an emotion annotation which has the potential to further strengthen the analysis and increase the effectiveness of these applications.

A similar lexical resource-based approach was taken to achieve emotional analysis of text using the WordNet-Affect List. The WordNet-Affect list contains words annotated with six emotional categories: joy, fear, anger, sadness, disgust and surprise. This too suffered a similar drawback wherein the the context of the sentence and the overall situation expressed in the sentence were not taken into consideration.

The emotion in a sentence need not necessarily be conveyed plainly, with the use of explicit words. For example, "I am happy I passed my exams" can clearly be identified as a sentence that expresses "Joy" due to the presence of the word, "happy". Whereas the sentence, "I finally passed my exam" although does not contain any words that express a single emotion explicitly; it can clearly be categorized as a sentence that conveys "Joy". This information was not extracted from the sentence based on the cue provided by a single word but rather by analyzing the situation expressed in the sentence through ones own past experience. Therefore, it can clearly be observed that the emotion of a sentence cannot be determined merely based on the words that form it, but also by analysing the underlying context evinced by a sentence. Furthermore, the extent to which an emotion is expressed can be understood by associating each emotion with a degree. We have given each emotion four levels characterized by one of the following values, 0.25, 0.50, 0.75 and 1; 0.25 being the lowest degree and 1 being the highest. Using these values, our model predicts the degree of emotion.

Neural Networks have been employed to widely and achieved fairly significant results in the field of Natural Language Processing, such as translation, text summarization etc. A Recurrent Neural Network(RNN) is a type of Neural Network that remembers its previous states. But, a significant disadvantage of RNN is that it works only on a short-term basis. Meaning, when context comes into play, the desired outcome is not achieved. This issue has been resolved by the Long Short-Term Memory(LSTM) Neural Network. A typical/ vanilla RNN modifies the entire information as a whole and does not take into consideration important and non-important information. Whereas, LSTMs have a mechanism known as cell states in which information is modified with multiplications and additions. Hence, LSTM is used to predict emotion in sentences because it can selectively retrieve or drop information.

2 Data sets

2.1 SemEval 2007 Affect Sensing Corpus

The SemEval 2007 Affect Sensing Corpus data set[1] contains 1250 news headings along with a list of their corresponding emotions from the set of emotions- Joy, Sadness, Fear, Anger, Disgust, and Surprise. This data set is considered to be our gold standard because it not only complies with Ekman's basic emotion model[2] but also has been annotated previously by humans. An advantage of this data set is that each sentence also comes with a valence/ degree of emotion underlying in the sentence. Sentences in this data set have contextual meaning rather than an obvious emotion. For example, the sentence "Resolution approved for international games" is tagged as "Joy". But, this data set is skewed in the sense that sentences tagged as "Surprise" are very less compared to sentences tagged as "Joy". Hence, all emotions do not have equal number of sentences. In addition, each sentence is annotated with a degree varying between 0 and 1 for each of the six emotions, with 0 stating that the sentence does not carry the particular emotion and 1 stating that it definitely carries the particular emotion.

2.2 ISEAR

The ISEAR data set[3] contains 7666 situational sentences with a list of their corresponding emotions from the set of emotions- Joy, Sadness, Fear, Anger, Disgust, Shame and Guilt. For example, the sentence "When I pass an examination which I did not think I did well" is tagged as "Joy". Unlike the above mentioned data set, this data set is not skewed. All seven of these emotions have almost equal number of sentence examples. Having a symmetrical data set is important so as to have a uniform feature extraction in our Neural Network(LSTM). In addition to the Ekman's basic emotion model[2], where the emotions are Joy, Sadness, Fear, Anger, Disgust and Surprise, this data set has two emotions- Shame and Guilt. However, for our model, we decided against using these two emotions because they are not included in Ekman's basic emotion model[2]. Ekman proposed that only these six emotions are strongly expressed across all the cultures in the world. Since these sentences are based on situations, emotion tagging can be done based on context as well.

3 Literature Review

This section covers an extensive survey on the approaches previously used to detect the emotion of a sentence.

Dipankar Das and Shivaji Bandyopadhyay[4] employ a machine learning approach using Conditional Random Field(CRF). Their method is a two step process wherein, the first step is to create an emotion for each word in a sentence using WordNet's Affect List and the next step is to find the dominant emotion of each sentence using weight scores of each word in the sentence. The first step

uses CRF for word level annotation which searches for words in the sentence that are present in the Affect list and returns word level tagging of emotions. The second step is done using these word level emotions to find the overall sentence level emotion based on weight scores of each word in a sentence. Although this paper produces a high accuracy of 87.65%, it is limited by the Synonyms set(SynSet) in the Affect list. This model fails to account for words that do not exist in the said list. For example, "Serena misses Bangalore tournament" is a sentence in the SemEval 2007 data set[1]. If this model were to find the emotion of this sentence, it would fail to do so because the word "misses" indicates Sadness in this context but Sadness Affect list does not contain this word.

The authors in Paper[5] do an experimental analysis on different approaches to tackle the problem of sentence level emotion tagging. They use five approaches, which consist of four knowledge-based ideas and one corpus-based idea. The first approach in knowledge-based ideas, WordNet-Affect Presence, deals with annotating the emotions in a text simply based on the presence of words from the Word-Net Affect lexicon. The second approach, Latent Sentiment Analysis(LSA), the LSA similarity between the given text and each emotion is calculated and each emotion is defined as a vector of the word expressing that particular emotion. The third method, LSA emotion SynSet, is a method in which in addition to the word denoting an emotion, its synonyms from the WordNet SynSet are also used. Finally, the fourth method, LSA all emotion words, which aids the previous set by adding the words in all the SynSets labeled with a given emotion, as found in WordNet Affect List. What can be inferred from this paper[5], is that knowledge-based approaches are constrained by unknown words. Moreover, the corpus-based approach used in this paper[5] uses a machine learning classifier, Naive Bayes that is trained on Blog Posts to classify emotion in a labelled data set. This approach is a more practical approach and has been employed in a similar fashion in our model.

In paper[6], the authors discuss a method to find emotion labels using two methods: keyword spotting and lexical affinity. These methods use a ready-made lexical corpus to find words pertaining to a particular emotion of the above stated set of emotions. This paper[6] does not take into consideration negation words like not, neither, and never, which can give the polarizing emotion of the sentence. However, it tags an emotion to sentence based on the context rather than the word level emotion weights. Their process assigns an emotion to the sentence by weighing the relation between the different words and emotions present in it. Each word is looked up in the lists of emotional words(LEW) and if found, its value is obtained. Else, the ANEW list (Bradley and Lang, 1999) is consulted, and if found, its value is obtained. If the words are not found in either of the lists, the hypernyms of the word are retrieved from WordNet. The model for the above example, "Serena misses Bangalore tournament" successfully classifies it as "Sadness".

The model built by the authors in paper[7] employs a rule-based approach for the classification of the sentences with a particular emotion. A data set consisting of user reviews is used. A particularly novel approach explored in this paper is

to improve the performance of the categorization of the sentences by including additional emotion-related signals, such as emoticons, emotion words, polarity sifters, slang and negations, to detect and classify emotions efficiently in user reviews. One possible limitation is that it cannot incorporate domain-specific words because it is dependant on SentiWordNet. Another disadvantage of this model, is the fact that a rule-based approach requires an extensive research and possible linguistic experts.

The authors in paper[8] discuss Disgust. They find relations between the words used in the sentence to the emotion depicted by it and further the deeper contextual meaning imparted by each of the words to the sentence. Once these terms and the relations are found, they are generalized to construct a set of rules known as emotion recognition rules (ERRs). These rules are then used to recognize the emotion from any given text. However a possible drawback is that they use WordNet and ConceptNet to generalize their lexicon, hence any words not present in either of them would not be identified and classified and rules relating to them would not be constructed.

4 Model Architecture

4.1 Hybrid Data set

As discussed in section II, the two aforementioned data sets have their own pros and cons. An amalgam of the two data sets was used to produce a hybrid data set that is uniform and has equal sentences of each emotion. The data set now not only consists of news headlines but also situational descriptions, which enables the model to have an all round robust learning. The SemEval data set[2], which has sentences like "Bombers kill shoppers", does not train the model well because its sentences are vague. Therefore, combining it with the ISEAR data set[3] allows the model to understand context and implement emotion extraction and therefore, contextual learning as well as word level feature extraction is achieved. The advantage of the proposed Hybrid Data set is twofold- one being data from Semeval data set and ISEAR data set were chosen such that the final number of sentences in each category of emotion is the same thereby tackling the issue of skewedness; the other being, since now the data set consists of single sentence news headlines[1] which allow learning at the word level and multiple sentence situational descriptions[3] which allow learning on a contextual level. This data set was split into training and testing data, where 80% was used in training and 20% was used in testing.

4.2 Pre-processing

The first step before any work can be carried out is to pre-process the data to bring about uniformity. Since our data set contains sentences ranging from news headlines to situational descriptions, it is essential to conform these various sentences to a single template. The quality of data directly affects any models'

ability to learn. Hence, we convert all the text to lower case and remove punctuation marks such as colons and quotations. This also provides uniformity in one hot encoding of the words.

4.3 Sentence Splitter

The next step is to check every sample for conjunctions, commas and full stops and split it accordingly. The common conjunctions include and, but, or, nor, for, so and yet. When a sample sentence is broken down into its several pieces, each piece can carry an emotion and hence, each emotion of each piece contributes to the overall emotion of the whole sentence. Taking the example, "I am sad that I didn't study for my exam but I am happy I passed", it can be seen that when the sentence is split into "I am sad that I didn't study for my exam" and "I am happy I passed" based on the conjunction "but", each of these pieces carry a different emotion, namely, Sadness and Joy. Therefore, this influences the final emotion of the entire sentence. The method used for this, is to check for the presence of a conjunction, comma(,) or full stop(.) and send these individual sentences to the subsequent step.

4.4 Obtaining Emotion Words

A very important aspect of emotion detection is the Part Of Speech(POS) tag. It can be concluded that most words which are either Nouns, Adjectives, Verbs or Adverbs have an underlying emotion. The sentence, "When I began school at UC, the pre-enrollment, the classes and the question of success scared me" has many words like the and when that do not contribute to the overall emotion of the sentence. It can be inferred that the words with a POS tag of Nouns, Adjectives, Verbs or Adverbs aids the sentence's emotion. In the above case, "question", "success" and most importantly, "scared" contributes to the overall emotion of "Fear" to the sentence. Hence, we decided to drop all words which have POS tags of anything other than the ones mentioned above.

4.5 Model

Prediction of emotion and degree is carried out by the use of a Long Short-Term Memory(LSTM) model. In the LSTM architecture there are three gates, forget gate, input gate and an output gate and a memory cell. The forget gate decides which information should be discarded and which should be retained. The input gate handles the the addition of relevant information to the cell state. The output gate deals with selecting useful information from the current cell state of all the available information to show as an output. LSTM in its core, preserves information from inputs that have already passed through it using the hidden state. More formally each gate can be modelled as follows

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i) \quad (1)$$

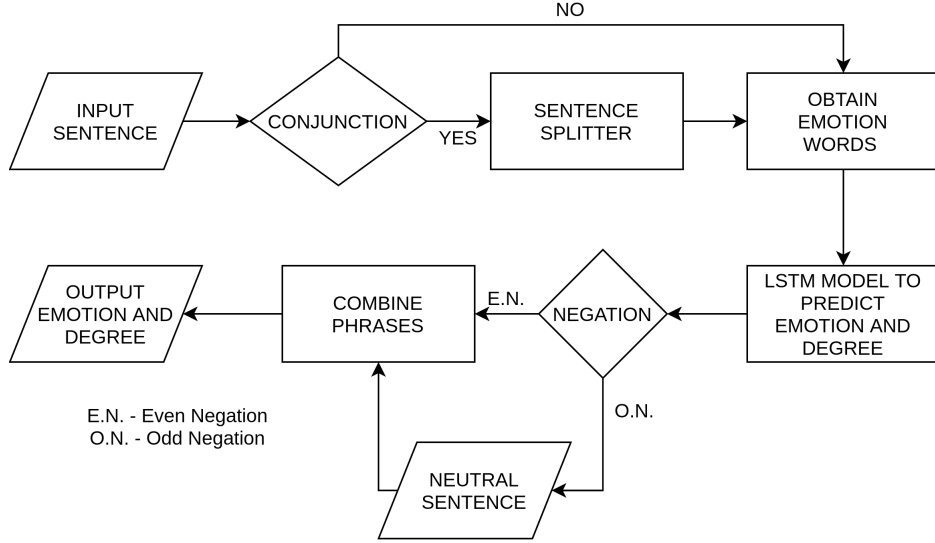


Fig. 1. System Architecture

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f) \quad (2)$$

$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o) \quad (3)$$

where,

i_t represents input gate, f_t represents forget gate, o_t represents output gate, σ represents sigmoid function, w_x represents weight of the neurons for the respective gate(x), h_{t-1} represents output of LSTM block at previous timestamp, x_t represents input at current timestamp, b_x represents biases for the respective gate(x).

Equations for the cell state, candidate state and the final output are as follows

$$c'_t = \tanh(w_c[h_{t-1}, x_t] + b_c) \quad (4)$$

$$c_t = f_t * c_{t-1} + i_t * c'_t \quad (5)$$

$$h_t = o_t * \tanh(c^t) \quad (6)$$

where,

c_t represents the cell state at current timestamp, c'_t represents the candidate for the cell state at current timestamp, h_t output after passing through the activation function.

The cell state considers what information it needs to discard from the previous state ($f_t * c_{t-1}$) and what it needs to include from the current candidate state ($i_t * c'_t$). Finally the state is passed through the activation function to obtain h_t which is the output of the current LSTM at timestamp t. This is then passed through the last softmax layer to get the predicted output.

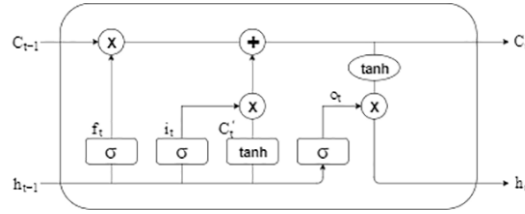


Fig. 2. LSTM Block

A bidirectional LSTM has been employed, so as to run the inputs in two ways, one from past to future and one from future to past. This method proved to be advantageous by relating to and understanding the context better, and showed a stark improvement in prediction when compared to unidirectional LSTM which only retained information of the past because the only inputs it has been given are from the past.

The first layer is an LSTM layer with 100 memory units which receives sequences of data rather than scattered data. A dropout layer is implemented to prevent overfitting of the model. The last layer is a fully connected layer with a softmax activation function.

The model has been trained using back propagation and the objective loss function used is the cross-entropy loss. The model is fit over 20 epochs with a batch size of 64.

The high accuracy attained by our model compared to standard machine learning and knowledge-based approaches can be attributed to the sequential learning nature of LSTM and its ability to discard and retain information to maximize output accuracy. In the emotion classification task, the input words are processed by LSTM networks sequentially and the last output of the LSTM represents the meaning of sentence, which is finally used to predict the emotion.

4.6 Post-processing

After categorizing the sentence under one emotion, the penultimate step is to check for the presence of negation words in the sentence. Given a sentence, "I am not happy", we see that the sentence is classified as Joy due to the presence of the word "happy". However, we know that this is not the case as the occurrence of "not" before the "happy" essentially means the opposite of happy. Therefore, it is imperative to check for negation words. Keeping in mind that most of the emotions of a text are subjective to the reader, if a sentence contains odd number of negation words then the predicted emotion is changed to a seventh emotion- Neutral because it is to be noted that, "not happy" does not directly correlate to another emotion in the Ekman's model[2] but rather depends on context again. The sentence "I am not happy" could imply disappointment, anger or sadness. The Neutral emotion is used to tackle such conflicts by making the overall emotion of a sentence neutral. But, if a sentence contains a double

negation, such as, "It is not that I'm not happy", it is implied that the negations cancel each other out, keeping the original emotion of the sentence intact. Hence, if a sentence has even number of negation words, the original emotion is not replaced with the Neutral emotion. If multiple sentences carrying varying set of emotions is considered, for example, "I am happy I reached the flight on time. But I am sad I spent a lot of money to go to the airport", we tackle the emotion of the sentences individually first and use Table 1 to get the overall emotion of the set of sentences. In the example, sentence 1 is Joy and sentence 2 is Sadness, so the overall emotion is Neutral. In general, if a set of sentences has conflicting emotions, the overall emotion for the set of sentences is categorized as Neutral as shown in Table 1.

Conflicting emotions are resolved as a post processing step and not within the prediction model itself because of two main reasons. Firstly, the data set does not include sentences tagged with a neutral emotion. So, the LSTM model does not learn what makes a sentence as neutral whereas since the data set has six other emotions, the model is able to learn what each of those emotions' mean. Secondly, as mentioned above if the sentence contains an even number of negation words then the sentence is left intact with its predicted emotion, only in the case wherein the sentence contains an odd number of negation words, it is classified as neutral. Therefore, it is essential that the sentence is first classified into one of Ekman's six emotions before being checked for negation, so that that in the event of an even negation the sentence can be left untouched and classified under its predicted emotion.

Emotion	Joy	Sadness	Fear	Anger	Disgust	Surprise	Neutral
Joy	Joy	Neutral	Neutral	Neutral	Neutral	Surprise	Joy
Sadness	Neutral	Sadness	Fear	Anger	Disgust	Neutral	Sadness
Fear	Neutral	Fear	Fear	Anger	Fear	Fear	Fear
Anger	Neutral	Anger	Anger	Anger	Anger	Anger	Anger
Disgust	Neutral	Disgust	Fear	Anger	Disgust	Surprise	Disgust
Surprise	Joy	Neutral	Fear	Anger	Surprise	Surprise	Surprise

Table 1. Dominant Emotion

If a sentence is found to be composed of multiple phrases joined by a conjunction, then each phrase is classified with a particular emotion. Since we require the entire sentence to be tagged with a single emotion, it is essential to merge the emotions of the phrases to produce a single dominant emotion. To achieve such a result, we have come up with a method to resolve ties between emotions. Table 1 shows the resulting emotion on combining two incongruous emotions. This table was obtained on careful observation and analysis of different conjunct sentences that expressed two or more contrary emotions.

Another interesting feature has been considered in our model known as degree of an emotion. Along with the dominant emotion, a sentence will have the extent

of the emotion conveyed as well. The difference in the following sentences, "I am happy I passed my exam", "I am very happy I passed my exam" and "I'm ecstatic that I passed my exam!" is apparent in the sense that each sentence conveys varying degrees of the emotion "Joy". So, the sequence is least to greatest degree of emotion in the above three sentences. Degree of every emotion has been categorised into four values, 0.25, 0.50, 0.75, 1.0 where the greatest value of 1.0 indicates highest degree and least value of 0.25 indicates the lowest degree. The LSTM model successfully categorised every sentence of the data set into one of these four categories. The SemEval 2007 data set[1] comes with a list of valences/ degrees for every emotion and consequently, every sentence. To begin with, the valences have been rounded off to the aforementioned categories of degree. This was used to train the Neural Network(LSTM) which achieved an accuracy of 91.6%. Examples from the tested data set are as follows- "Mother and daughter stabbed near school" received a degree of 1.0 of the Sadness emotion, "Hurricane Paul nears Category 3 status" received a degree of 0.75 of the Fear emotion, "UK announces immigration restrictions." received a degree of 0.50 of the Anger emotion and "Messi makes Barcelona squad return" received a degree of 0.25 of the Joy emotion.

5 Evaluation and Results

A comprehensive analysis has been conducted on the following data sets:- SemEval 2007 data set[1], ISEAR data set[3] and a hybrid of the previous two, to produce emotions. Four metrics are used to compare the model's performance on these data sets, namely, Accuracy, Precision, Recall and F1 score.

Accuracy can be given by,

$$(TP + TN)/(TP + FP + FN + TN) \quad (7)$$

Precision can be given by,

$$TP/(TP + FP) \quad (8)$$

Recall can be given by,

$$TP/(TP + FN) \quad (9)$$

F1 Score can be given by,

$$2 * (Recall * Precision)/(Recall + Precision) \quad (10)$$

where TP is number of true positives, TN is number of true negatives, FP is number of false positives and FN is the number of false negatives.

The following metrics have been tabulated and displayed in Table 2. It is inferred from the table that the combination of the SemEval 2007 data set and ISEAR data set improves the overall working of the LSTM model because it predicts not only based on situations but also on context. Table 3 shows how our deep learning model fairs with the state of the art classical models: Naïve Bayes Classifier and Knowledge Based Technique, in terms of accuracy. It is

Data set	Accuracy	Precision	recall	F1 Score
SemEval	68.45%	0.681	0.459	0.548
ISEAR	63.74%	0.729	0.356	0.478
Hybrid	85.63%	0.752	0.524	0.618

Table 2. Results

Model	SemEval	ISEAR	Hybrid
Deep Learning(ours)	68.45%	63.74%	85.63%
Naïve Bayes	63.16%	67.72%	71.23%
Knowledge Based Technique	51.13%	53.6%	56.18%

Table 3. Model Comparison

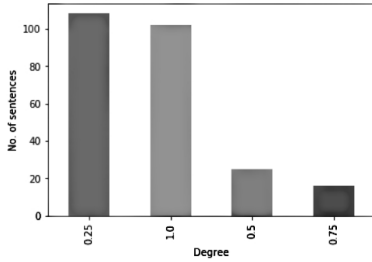


Fig. 3. Annotated Degree

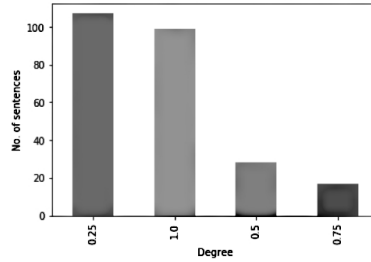


Fig. 4. Model Degree

evident that the model we have built performs more precisely than the classical models.

Additionally, the degree of emotion was predicted for each emotion and had an accuracy of 91.6%. Figure 3 is a bar graph which indicates the data set’s annotated degree and figure 4 is a bar graph that indicates our model’s predicted degree. Taking a few examples, "I am happy" is given a degree of 0.25 and "I am very happy" is given a degree of 0.50.

6 Conclusion

In this paper three novelties are presented. The first is that a sentence can be a conjunct of emotions and it has to be amicably resolved. For example, in the sentence "I passed the examination but I am disappointed by my below par performance", although clearing the exam makes the speaker happy the general expression is that of sorrow as the speaker is disappointed with his/her performance.

The second is that, not all sentences can be classified into an emotion and some may express a neutral emotion. For example, "The sun sets in the west"

is not a sentence that bursts with a particular emotion but is a rather tame expression that just declares a fact in a neutral tone.

The third novelty is that our model also takes into consideration that a sentence belonging to a particular emotion can show varying levels of the emotion. For example, anger has two sub categories- hot anger and cold anger; where hot anger can be of the form rage and cold anger can be of the form passive aggressiveness. This is expressed by an additional output, that is, degree of the emotion that describes the extent of the emotion. The output degree has a set of four discrete values(0.25, 0.50, 0.75, 1).

This we conclude that our deep learning model fairs better than other state of the art classical models such as the Naïve Bayes Classifier and Knowledge Based Technique, and gives us a maximum accuracy of 85.63% to predict the emotion of text and an accuracy of 91.6% to predict the degree of the sentence.

7 Future Work

The application of such an analysis is multitudinous, ranging from differentiating the mental and psychological state of a person through his or her text messages to helping in the psychological analysis of the person to further developing an empathetic chat bot, that can dynamically gauge the emotional state of the customer/user and be better equipped to handle the situation to integrating a text-to-speech engine that can utilize the extracted emotion to synthesize speech that can incorporate the said emotion to emulate a dramatized speech. This dramatized speech can be used to generate more realistic audio books with varying emotion so as to mimic a human tone.

Furthermore, emotion recognition has helped opinion mining in social media of various reviews and political discussions progress one step further than sentiment analysis. Now, with the help of varying degree of emotion, progress can be made further on classifying what opinions social media users have and digitally analysing these opinions.

Our model has been developed which can effectively categorize sentences into one of the six basic emotions(Joy, Sadness, Fear, Anger, Disgust, Surprise) and an additional emotion category: Neutral. This additional information obtained can be passed along with the sentence to construct emotional speech. The model can also be expanded to include more emotions other than the above mentioned six. Rather than representing degree of an emotion by a set of discrete values, fine grained values, that is a continuous set of values that gives a greater control over the degree [0,100] too should be explored to further strengthen the emotional analysis of text.

References

1. Strapparava, Carlo, and Rada Mihalcea.: Semeval-2007 task 14: Affective text. In: Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007). pp. 70-74. (2007)

2. Ekman Paul.:An argument for basic emotions. *Cognition and Emotion*, vol 6(3-4). pp. 169-200. (1992)
3. International Survey on Emotion Antecedents and Reactions data set. *https : //www.unige.ch/cisa/index.php/download_file/view/395/296/*
4. Das, Dipankar, and Sivaji Bandyopadhyay.: Sentence level emotion tagging. In: 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops. pp. 1-6. IEEE (2009)
5. Strapparava, Carlo, and Rada Mihalcea.:Learning to identify emotions in text. In: Proceedings of the 2008 ACM symposium on Applied computing. pp. 1556-1560. (2008)
6. Francisco, Virginia, and Pablo Gervás.:Exploring the compositionality of emotions in text: Word emotions, sentence emotions and automated tagging. In: AAAI-06 workshop on computational aesthetics: Artificial intelligence approaches to beauty and happiness. (2006)
7. Asghar, M.Z., Khan, A., Bibi, A., Kundi, F.M. and Ahmad, H.: Sentence-level emotion detection framework using rule-based classification. *Cognitive Computation*, vol 9(6), pp.868-894. (2017)
8. Shaheen, Shadi, Wassim El-Hajj, Hazem Hajj, and Shady Elbassuoni.:Emotion recognition from text based on automatically generated rules. In: IEEE International Conference on Data Mining Workshop, pp. 383-392. (2014)