Emotion Detection and Sentiment Analysis in Texts

Naman Pundir
University at Buffalo, New York, US
namanpun@buffalo.edu

Abstract

Emotion detection and recognition of sentiments from text is a recent field of research that is related to predict the sentiment in the texts. Sentiment analysis aims to detect positive, neutral or negative emotions from the given text and similarly on the other hand emotion detection aims to detect and recognize types of emotions through the expression of the texts, such as anger, fear, disgust, sadness, happiness and surprise. These 6 emotions are the most common emotions in human beings. There are plenty of useful application of detecting emotion and sentiments from the texts. In this paper we will be using some deep learning algorithms to detect emotions from the text using some datasets.

1 Introduction

There are 6 emotion categories that are widely used to describe humans' basic emotions, based on facial expression: anger, disgust, fear, happiness, sadness and surprise. Anger, disgust, fear, sadness is mainly associated with negative sentiment, with "Surprise" being the most ambiguous, as it can be associated with either positive or negative feelings and happiness as the only positive expression.

The question remains, however, how much of an emotion we can convey via text? How we can determine the emotion behind the text? Can we detect emotions in text like in images?

This is especially interesting since facial expression and voice intonation convey over 70% of the intended feelings in spoken language. Detecting emotions from the images or videos can be achieved by some computer vision and deep learning techniques. Here, we are exploring how we can achieve this task in text by neural networks approach and transformers, specifically using the deep learning technique. We will be using 1D-CNN (Convolution Neural Network), LSTM (Long Short-Term Memory) and BERT (Bidirectional Encoder Representations from Transformers) transformer to detect emotion in the texts and comparing these models based on their f1 score.

2 Datasets

To train these models we need some datasets which contain some sentences and labelled emotions. We are using three datasets to train our models which are Daily-Dialogue[link], Emotion-Stimulus[link], ISEAR[link].
All three datasets were merged and split into train and test dataset to get the variety of data to train the model.

Dataset	Year	Size
Daily-Dialogue	2017	102k lines
Emotion-Stimulus	2015	2.5k lines
ISEAR	1990	7.5k lines

Datasets are labeled with 5 labels sad, anger, joy, fear and neutral. Text in the datasets are simple dialogue or small texts. We will use these datasets to train our models and to test them.

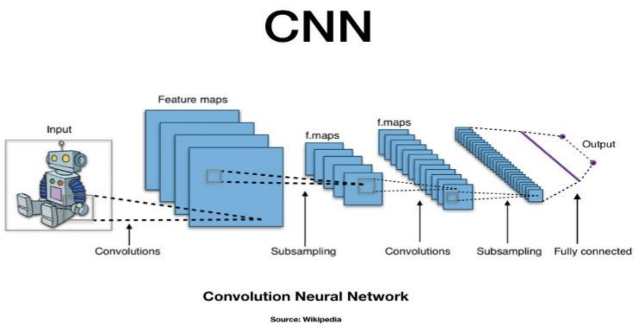
Anger	Text
neutral	There are tons of other paintin
sadness	Yet the dog had grown old and
fear	When I get into the tube or the
fear	This last may be a source of con
anger	She disliked the intimacy he sho
sadness	When my family heard that my
joy	Finding out I am chosen to colle
anger	A spokesperson said : ` Glen is fu
neutral	Yes .
sadness	When I see people with burns I
fear	I had gone to the hospital for m
sadness	One day I heard from a friend th

3 Models

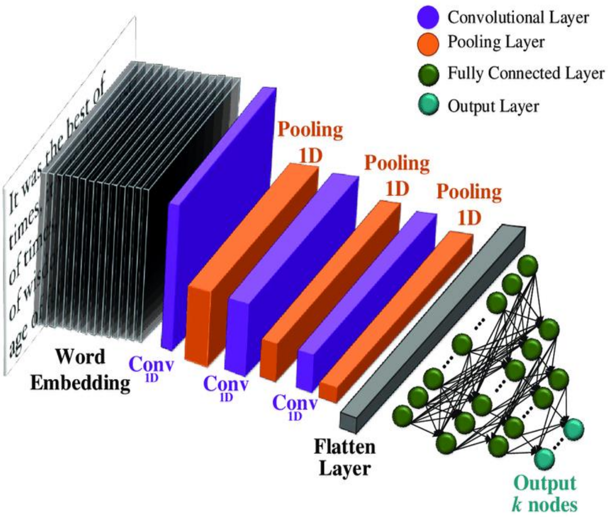
We are using three models CNN, LSTM(RNN) and BERT to detect the emotions in the text.

1. CNN (Convolution Neural Network)

Convolutional neural network (CNN), a class of artificial neural networks that has become dominant in various computer vision tasks, is attracting interest across a variety of domains, including radiology, medical image classification.

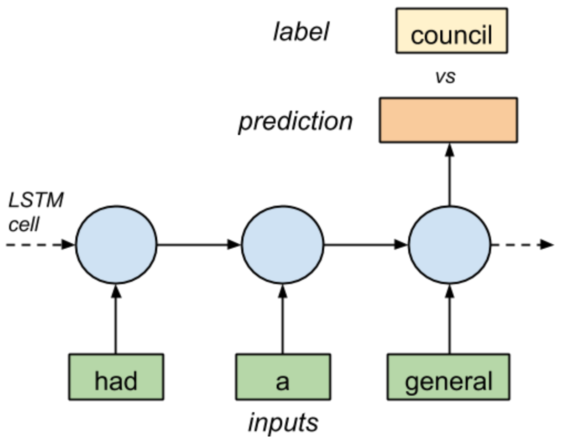


CNN is designed to automatically and adaptively learn spatial hierarchies of features through backpropagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers. CNN nowadays is used in many image and videos classification tasks including medical imaging and sports video analysis.

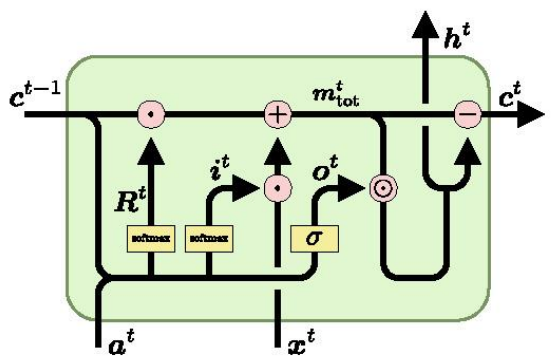


2. LSTM (Long Short-Term Memory)

Long Short-Term Memory networks – usually just called “LSTMs” – are a special kind of RNN (Recurrent Neural Network), capable of learning long-term dependencies. The basic idea behind LSTM is to keep the important context in the long-term memory for further use and less important contextual data in the short memory which might not be used in the future. They were introduced by Hoch Reiter & Schmid Huber (1997), and were refined and popularized by many people in following work. They work tremendously well on a large variety of problems and are now widely used.

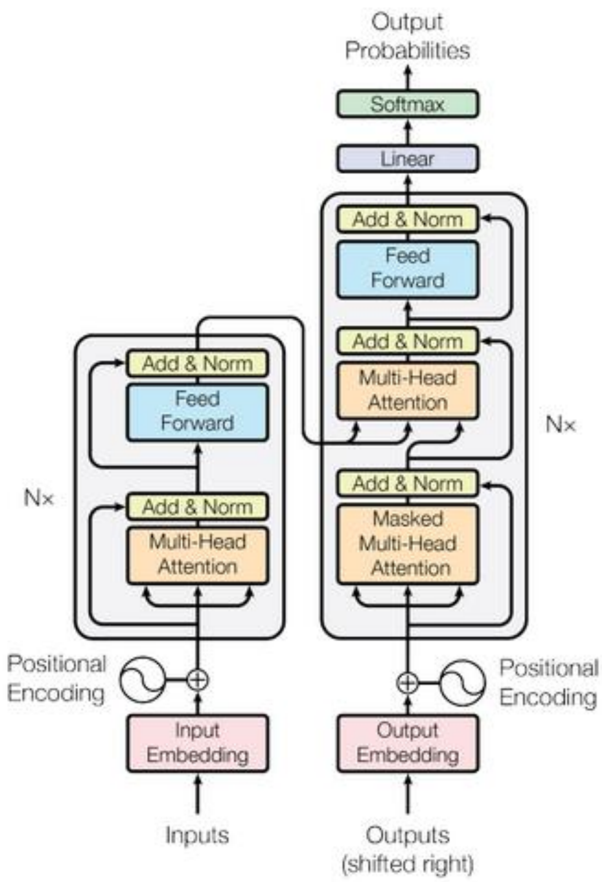


LSTMs are explicitly designed to avoid the long-term dependency problem. Remembering information for long periods of time is practically their default behavior, not something they struggle to learn. All recurrent neural networks have the form of a chain of repeating modules of neural network. In standard RNNs, this repeating module will have a very simple structure, such as a single tanh layer.

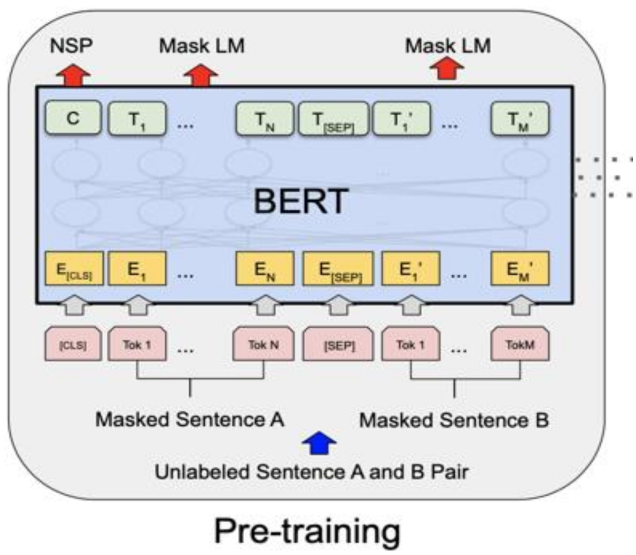


3. BERT (Transformer):

The Transformer in NLP is a novel architecture that aims to solve sequence-to-sequence tasks while handling long-range dependencies with ease. The Transformer was proposed in the paper Attention Is All You Need [link]. Attention is used in the transformers block and brought the revolutionary change in the NLP industry. This is the quote from the above-mentioned paper: "The Transformer is the first transduction model relying entirely on self-attention to compute representations of its input and output without using sequence-aligned RNNs or convolution."



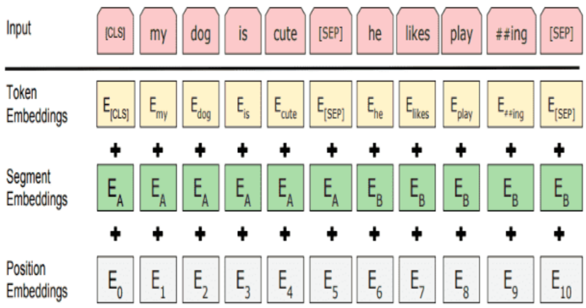
BERT (Bidirectional Encoder Representations from Transformers) is a recent paper published by researchers at Google AI Language. It has caused a stir in the Machine Learning community by presenting state-of-the-art results in a wide variety of NLP tasks. It is the most widely used transformer in the industry.



BERT's key technical innovation is applying the bidirectional training of Transformer, a popular attention model, to language modelling. This contrasts with previous efforts which looked at a text sequence either from left to right or combined left-to-right and right-to-left training.

As opposed to directional models, which read the text input sequentially (left-to-right or right-to-left), the Transformer encoder reads the entire sequence of words at once. Therefore, it is considered bidirectional, though it would be more accurate to say that it's non-directional. This characteristic allows the model to learn the context of a word based on all its surroundings (left and right of the word).

Bert use three embeddings to compute the input representations.



4 Working with CNN Model:

Architecture:

(X) Text

Embedding (W2V pretrained on wiki)

Deep Network (CNN 1D)

Fully connected (Dense)

Output Layer (SoftMax)

Emotion class (Y)

Embedding Layer:

Word Embedding is a representation of text where words that have the similar meaning have a similar representation. We will use 300-dimensional word vectors pre-trained on Wikipedia articles. We can also train the w2v model with our data, however our dataset is quite small and trained word vectors might not be as good as using pretrained w2v.

Deep Network:

Though text data is one-dimensional, we can use 1D convolutional neural networks to extract features from our data. The result of each convolution will fire when a special pattern is detected. By varying the size of the kernels and concatenating their outputs, we are allowing ourselves to detect patterns of multiples sizes. Patterns could be expressions like "I love", "very less" and therefore CNNs can identify them in the sentence regardless of their position.

Fully Connected Layer:

The output from the convolutional layers represents high-level features in the data. While that output could be flattened and connected to the output layer, adding a fully connected layer is a (usually) cheap way of learning non-linear combinations of these features.

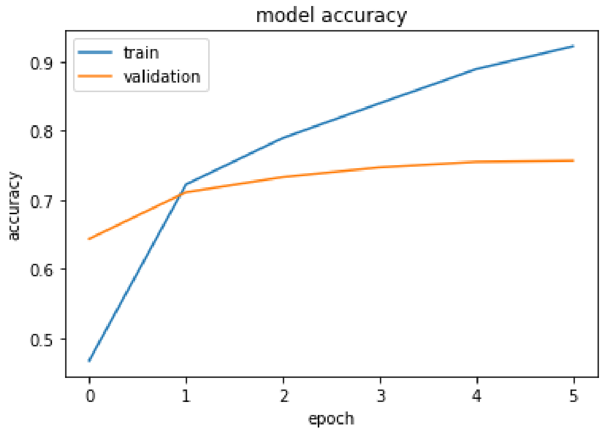
Output Layer:

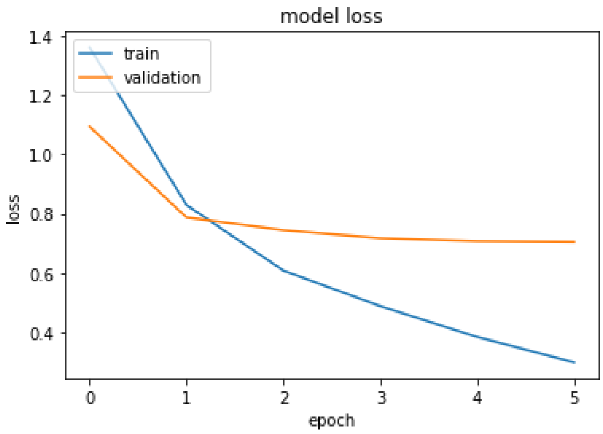
Based on the problem at hand, this layer can have either Sigmoid for binary classification or SoftMax for both binary and multi classification output. These are the activation functions which are responsible to help the network to learn the complex data.

Model:

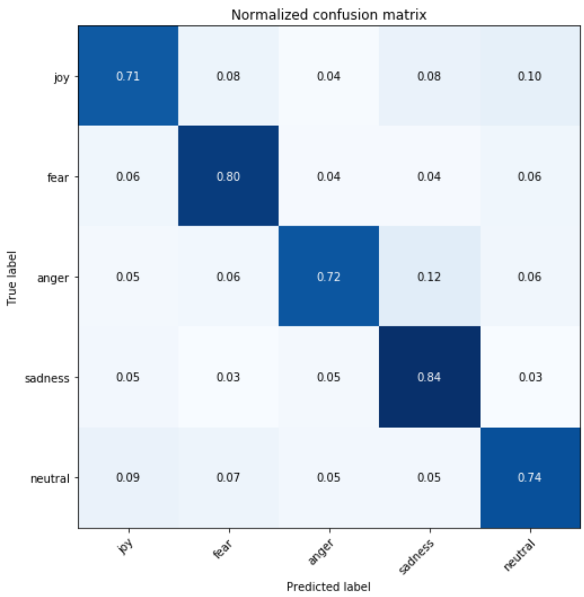
```
# Convolution
kernel_size = 3
filters = 256

model = Sequential()
model.add(embedd_layer)
model.add(Conv1D(filters, kernel_size, activation='relu'))
model.add(GlobalMaxPooling1D())
model.add(Dense(256, activation='relu'))
model.add(Dense(num_classes, activation='softmax'))
```





Results: F1 Score: 75.6



(X) Text

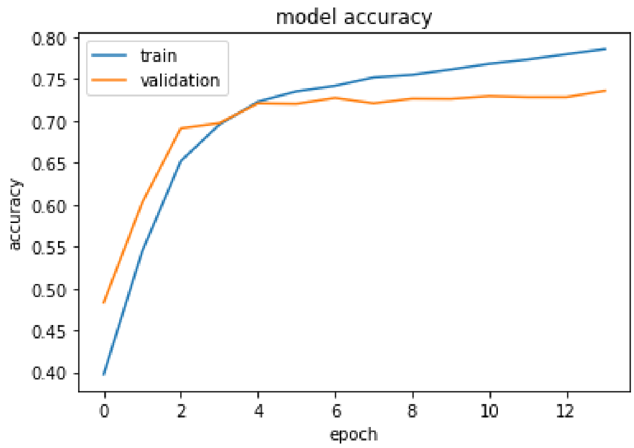
Embedding (W2V pretrained on wiki)

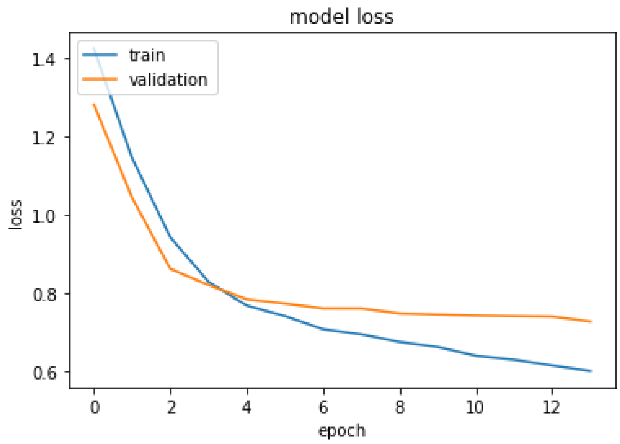
Deep Network (LSTM)

Fully connected (Dense)

Output Layer (SoftMax)

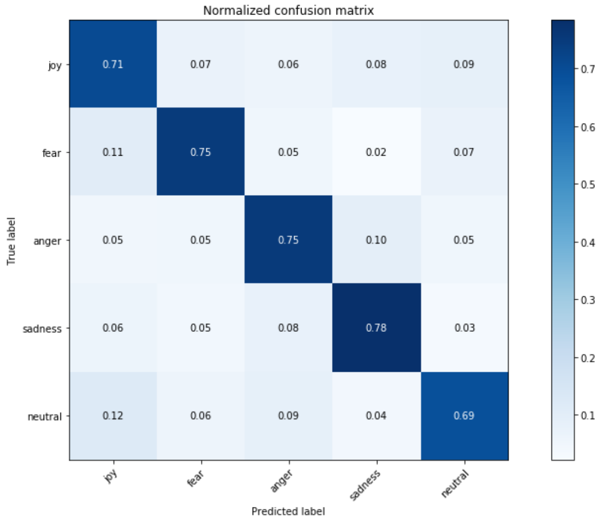
Emotion class (Y)

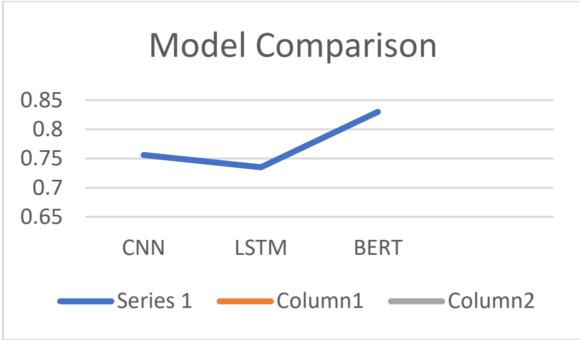




Results: F1 Score: 73.56

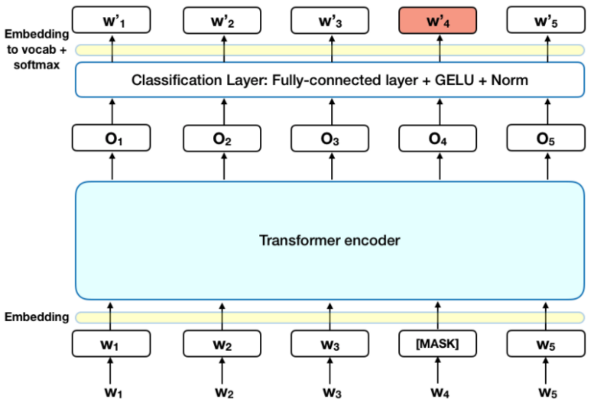
5 Working with LSTM Model:





6 Working with BERT Model:

Results: F1 Score: 0.83



7. Result Table:

Model	F1 Score
CNN	**0.756**
LSTM	**0.735**
BERT	**0.830**

We can observe that BERT transformer gives highest F1 score among all the models. Which clearly shows that the BERT model is the best model among CNN, LSTM and BERT to detect emotions in text.

8. Conclusion

As we know there are different kind of data i.e. 1D vector data, 2D data, Sequential data and 3D data. Text is kind of a sequential data and to process the sequential data we have three methods 1D CNN, LSTM, and Transformers, which we used with these datasets.

We can clearly see that BERT transformer is more accurate predicting emotions as compare to CNN and LSTM in the above three datasets.
BERT is better than other as it is pre trained on an absurd amount of data. BERT model comes in two sizes one is BERT-base (800million words) and BERT-large (2500million words). That makes BERT better than other models.
BERT is the first deeply bidirectional, unsupervised language representation, pre trained only a plain text corpus.

9. Applications

Emotion detection and sentiment analysis using transformers is one of the hottest topics in the tech industry.
Applications like **social media monitoring** which help companies about some of the most truthful points of view about products, services, and business because users offer their opinions unsolicited.

Sentiment analysis **Customer Support** reads regular human language for meaning, emotions, tone, and more, to understand customer requests, just as a human would.

Sentiment analysis can be used to gain insights from the **Customer Feedback** available and save employee hours and can also work with common mistakes like misspelled and misused words.

8. Scope of Work

Emotion detection and sentiment analysis can be used to detect sentiment behind the stock market prediction which can lead to better investment opportunity.

Lot of people put their views on social media platforms like Twitter and Facebook and detecting emotions behind these views by scraping the data and using models like BERT, which can give an insight about the stocks and future prediction. The movement of any stock depend on two factors one is the company work, policies and revenue and the other is public sentiment on that stock.
There is a scope of work of creating a model which can scrape the data from social media platform about the stocks and detect the major emotions behind the stock which can be used to predict the future performance of the stocks.

9. Reference:

1. Emotion Detection from Text: Shiv Naresh Shivhare and Prof. Saritha Khethawat
https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.301.9872&rep=rep1&type=pdf#:~:text=Emotion%20Detection%20in%20text%20documents,in%20emotion%20detection%20are%20discussed.
2. Text-based emotion detection: Advances, challenges, and opportunities https://onlinelibrary.wiley.com/doi/full/10.1002/eng2.12189
3. Emotion detection in Text: Armin Seyeditabari, Narges Tabari, Wlodek Zadrozny https://arxiv.org/abs/1806.00674
4. Classify text with BERT https://www.tensorflow.org/text/tutorials/classify_text_with_bert
5. Why and how to use BERT for NLP Text Classification? https://www.analyticsvidhya.com/blog/2021/06/why-and-how-to-use-bert-for-nlp-text-classification/
6. Daily Dialogue - https://arxiv.org/abs/1710.03957
7. Emotion-stimulus - https://metatext.io/datasets/emotion-stimulus
8. Isear(International Survey on Emotion Antecedents and Reactions)- https://paperswithcode.com/dataset/isear
9. Attention is all you need: https://arxiv.org/abs/1706.03762
10. Chowdhury GG. Natural language processing. *Ann Rev Inf Sci* https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/aris.1440370103?casa_token=cvvV-T3fVmQAAAAA%3A_1-M4d-lYvjNkdrhFeDDtwz08n_sUhVT_5Dd5rms6TT1PNynXTGDZxuOt5ZA6elTzz_Gp_I41oFdsg .