Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 766–777

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

Emotion Cause Pair Extraction By Multi Task Learning on Enhanced English Dataset

Arunima Khunteta[a], Pardeep Singh[b]

[a]*National Institute of Technology, Hamirpur, Hamirpur, H.P. - 177005, India*
[b]*National Institute of Technology, Hamirpur, Hamirpur, H.P. - 177005, India*

Abstract

Emotion analysis and detection is a step further in sentiment analysis that aims to detect the emotion portrayed. A deeper task in mining opinions and emotions is to know the trigger behind the emotion or its stimuli and that is achieved through Emotion Cause Extraction. However, since it requires annotating sentences, limited research has been carried out in this area. Hence, there is a need to explore emotion cause extraction and expand its dataset in the most used language worldwide - English.

This paper aims to extract causes without needing annotation of emotion by the proposed approach of Emotion Cause Pair Extraction on English language. For this, we have constructed a dataset of about 6000 English emotionally annotated sentences spanning across six primary emotions - happy, sad, fear, guilt, shame, disgust. Added to this, we have proposed algorithms to extract the causes of emotions from these sentences. The experimental models of our proposed approach gives encouraging results (a ∼ 9% increase from the state of the art methods) and also uses a larger dataset as compared to existing research on English language. We conclude by sharing the results, and discussing future advancements in this field.

© 2023 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the International Conference on Machine Learning and Data Engineering

Keywords: emotion recognition; emotion cause extraction; natural language processing

1. Introduction

Natural Language Processing is the manipulation of speech and text by using a computer or software exploiting the language linguistics. It is focused on manipulation and interaction between computer and natural language has recently been a huge part of Artificial Intelligence and how it stands to change the world. With various new researches and

* Corresponding author. Tel.: +91-9882746078 ;
E-mail address: arunimakhunteta@gmail.com

1877-0509 © 2023 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the International Conference on Machine Learning and Data Engineering
10.1016/j.procs.2023.01.057

applications in this technological era, it has been finding its application in all the sectors, including chatbot assistant, text summarizations, machine translations, autocorrect, market translations, sentiment analysis etc.

Sentiment Analysis has been finding its foot at a much higher pace. In simple words, **sentiment analysis** is mining a sentiment from a text for business purposes. A direct example can be gathered from using sentiment analysis on various product reviews to know how your service is doing in the market. Sentiment Analysis helps brands make sense of this huge data by automatically understanding, processing, and finally tagging their sentiment.

Sentiment analysis models generally focus on polar emotions like positive, neutral, and negative. However, a paragraph of text may be far more than that. For instance, *"This product isn't friendly"* and *"I hate this bottle, it is the worst!!"* have negative sentiment in common. However, emotionally they are very different.

While sentiment analysis obtains a polarity from the text, it does not go into detail about the underlying reasons for the sentiment. Emotion Recognition aims to provide this additional analysis.

1.1. Emotion Recognition and Detection

An *emotion* is a feeling which points towards a persons state of mind. Various emotions that we feel in our everyday lives are happy, sad, joyous, disgust, angry, sad, fear, shocked, excited etc. Every textual, image, or any other information we exhibit, be it virtually or in real life results in an emotion. As computers in the face of artificial intelligence are occupying a huge part of our lives, it is only natural that we improve the human-computer interaction by including our emotions in the various tasks it does. However, the lack of common sense knowledge makes it much harder for the computer to recognize or detect human emotion especially implicit ones. When we automatically assign an emotion to a text from a collection of predefined emotion labels, it is referred to as emotion recognition in the text.

Emotion recognition, and most importantly implicit emotion recognition requires Natural Language Understanding (NLU) and is considered one of the difficult tasks in Natural Language Processing (NLP).

1.2. Emotion Cause Extraction

One of the most crucial elements for in-depth emotion analysis is emotion cause/trigger. The purpose of *Emotion cause extraction* is to recognize the reason behind an observed emotion.

Any event which triggers an emotion is provoked by underlying causes and characterized by certain emotion words. The context including the reasons and emotion words is then used to extract emotion causes as a passage. Emotion passages are divided into clauses, with each clause serving as an emotion cause candidate unit. Thus, the aim of Emotion Cause Extraction (ECE) is to extract potential causes that lead to emotion expressions in text. For important information extraction, the part of sentence that participates the most in yielding that information is Emotion-cause interaction.

Knowing that an emotion exists, for example, is rarely enough to predict future events or choose the best response. However, if the emotion cause is known in addition to the type of emotion, future events can be predicted more accurately, and potential consequences can be assessed more accurately. To put it another way, when emotion is treated as an event, the causal relationship is the most important one to figure out.

However, there are various shortcomings in the current ECE task. The first and foremost is that cause extraction tasks require annotation of emotion which limits the real world scenarios and applications. Due to the complexity and subjectivity in annotation, the size of most common corpus are small especially in English Language. The most used English corpus is by [1] with 2403 instances and 4,858 emotion annotations. Along with this, the approaches that require annotation to be done prior to the extraction task ignore the mutually indicative relationship between cause and emotions.

2. Literature Review

Text based sentiment analysis and emotion analysis was getting a lot of coverage in 2010. However, Lee *et al.* [2] recognized the importance of extracting in-depth information on emotion - like cause, stimuli and result and coined the term 'Emotion Cause Extraction'. They started constructing a Chinese corpus with annotated emotion based on their proposed scheme. Then, by analysing the data, they observed seven groups of linguistic cues, compressing them into two sets of rules for emotion and cause detection. These linguistic cues were highly collocated with cause events and delivered a promising performance of 79.38% accuracy.

Exploring this further, Chen *et al.* [3] aimed to capture long-distance emotional information in the cause extraction task. They used the same dataset as Lee *et al.* [2] and tabulated the clause distribution of cause texts and a multi-clause distribution of cause text. They proposed a multi-label approach to detect emotion causes using an integrated system to detect causes by taking contextual information into account. They achieved a performance much higher than baseline model.

Taking this further and exploring with a different dataset, Gui *et al.* [4] first designed an emotionally annotated corpus from Chinese micro-blogging website Weibo. They compiled 25 manual rules for emotion cause detection and two machine learning based methods based on Support Vector Machine (SVM) and Conditional Random Field (CRF). The main task of these methods was to identify the cause candidate words including noun phrases and verb phrases to then determine if they contain the cause of the emotion. They employed a 5 fold validation for machine learning methods and had an accuracy of 61.98% for SVM, and 77.57% for CRF.

Extending rules based approaches, Gao *et al.* [5] proposed a rule based approach underlying the conditions that trigger the emotions from their emotional model and then extract the corresponding causes fine grained in these emotions. The performance of their model was calculated and the precision came out to be 82.50%, F-score at 73.09%.

In 2017, Gui *et al.* [6] proposed a novel method which considered emotion cause identification as a reading comprehension task in question answering approach instead of the traditional approach of considering it as an information extraction task. The network projected the text into a low-dimensional vector space based on the learned attention outcome . This vector is then used to generate an answer which is a 'Yes' or a 'No' depending on if the query correctly identifies the cause of the emotion.

Considering relationship between emotion clause and cause clause, Chen *et al.* [7] attempted to benefit the mutual benefits across emotion classification and emotion cause detection by proposing a neural network based joint approach. framework of their joint approach had two parts - a joint encoder (the lower part) which extracts feature representations for both EClass instancesand ECause instances, and a linear decoder (the upper part) which assigns labels to instances according to their representations.

From then, various researchers started using Deep learning based methods for the ECE task.

Yada *et al.* [8] proposed a bootstrapping method to acquire these phrases as textual cue patterns. Their proposed method takes into account the cue phrases between emotion and cause.Xu *et al.* [9] used rank to recognize emotion causes They looked at the pointwise, pairwise, and listwise ranking restrictions for creating successful emotion cause extraction ranking models in model training. Xiao *et al.* [10] formalized emotion cause extraction as a sequence labeling problem and proposed a Context-aware Multi-View attention networks (COMV). Their model learned multi-view clause representations and integrated context information. They achieved an F1 Score of 77.21% outperforming their baseline methods. Xia *et al.* Hu *et al.* [11] proposed to combine external sentiment knowledge in order to extract the emotion triggering event (cause) of a certain emotion. [12] in their paper aimed to extract the cause of emotions and proposed a two step approach. Firstly, they used a multi task learning network to extract emotion and cause clauses from document, and secondly by pairing cause and emotion by cartesian product which yields a candidate set of emotion-cause pairs, then training a filter to eliminate invalid pairs. Transforming this from two step approach, Singh *et al.* [13] proposed an end to end method for the same task.They used an English dataset adapting the NTCIR-13 ECE corpus and achieved an F1 Score of 65.63%.

Some more approaches and datasets with their performance has been discussed in [14].

This paper aims to develop an approach for extracting the clauses containing the causes of provoked emotion, thus yielding crucial information which can then be used in understanding people, interpreting emotions, and knowledge mining.

Given a document with clauses $c_1, c_2, c_3....c_n$, extract emotion cause pairs clauses, P = $...(e_1, c_1)...$

Example - He was a professional musician now , still sensitive and happy because he was doing something he loved.

The key emotion here is "happy" and the cause of the emotion is "he was doing something he loved".

Our proposed approach aims to break down the sentence in clauses, and find pairs of emotion and its cause i.e.

(still sensitive and happy) , (because he was doing something he loved)

The block diagram of the proposed emotion cause pair extraction task is given in Figure 1

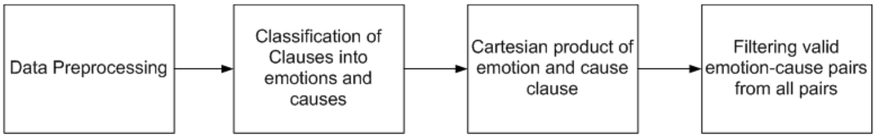

Fig. 1. Block diagram for emotion cause pair extraction task

3. Proposed Approach

The purpose of Emotion Cause Extraction is to find the emotion's matching cause clause. Traditionally this required annotating the emotions of the document to then determine the cause.

Hence, the aim of usual Emotion cause extraction task is to track the corresponding cause clauses when given the annotation of emotion: *happy*. However, the purpose of the proposed emotion cause pair extraction is to do this task without annotation and extract all valid pairs of emotion and cause. Thus, the output of our Emotion Cause Pair Extraction (ECPE) task is a pair of emotion - cause.

The proposed approach of ECPE first classifies all clauses in the document as 'emotion; or 'cause', and then conduct emotion-cause filtering using an improved version of the multi-task learning model for English Language.

The two crucial subtasks are -

Task 1 - Emotion Cause Classification

Here two classifiers are used to classify the clause into an emotion clause or a cause clause. At the end of this step, we get clauses in the Emotion-clause list and Cause-clauses list.

$$EC = (C_1^e, C_2^e..) \tag{1}$$

$$CC = (C_1^c, C_2^c..) \tag{2}$$

where C_1, C_2 represents clauses and C^e, C^c represents clauses having emotions and cause respectively. //

Task 2 - Emotion Cause Filtering

In this step, first a cartesian product is performed between emotion clauses and cause clauses of a particular sentence. Then the pairs are filtered according to if they have a causal relationship among them. The output is the correct emotion-cause pairs of the entire document.

The further sections will define dataset construction and each step in detail.

3.1. Dataset Construction

3.1.1. Corpora Source

As there is a lack of annotated English dataset, we constructed a corpora spanning across six emotions - happy, sad, fear, shame, disgust, guilt.

We included the dataset made by Ghazi *et al.* [15] in our corpora. This dataset considered Ekman's emotion model with 7 emotions across 820 sentences. A typical example in this dataset is

< *sad* >**It was hot , sunny and the water looked so cool and inviting that Maggie felt quite downcast** < *cause* > **about having no swimsuit** < *cause* > **.** < *sad* >

Since we required more sentences for machine learning training task, we use causes collected by International Survey on Emotion Antecedents and Reactions (ISEAR) [16].

These were a series of survey conducted by students and psychologists all around, where they were told to report events when they felt a particular emotion. In each case, the question covered the trigger and how they felt when they felt a particular emotion. The final data set thus contained reports on seven emotions each by close to 3000 respondents in 37 countries on all 5 continents [17].

We used the events reported in this survey as a cause and manually formed sentences with annotations of cause and emotion

3.1.2. Corpora Annotation Scheme

Each sentence in the dataset contains an emotion clause - where an emotion was expressed using an emotion keyword, a cause clause - where the reason of the emotion was expressed, and the annotation tags.

Emotion Clauses

Since there are many words in English vocabulary expressing a particular emotion, we have included synonyms to emote a particular feeling.

Some of them are - *shattered, revolted, disappointed, ecstatic, exhilerated, anxious, distraught, bored, disgusted, blissful, amazed* etc. These were appended in the pickle file. For annotation purposes, emotion tags are used at the start and end of the sentences.

<guilt> Rosie felt so regretful <cause> when she did not reply to her friend's letter in time. <\cause> <\guilt>

Fig. 2. Examples of Emotion Clauses

Figure 2 shows the emotion clause in red color text and the highlighted annotation tags.

Cause Clauses

Each sentence in the dataset also contains a cause of the sentence. The causes are manually annotated with the $< cause >$ and $< \backslash cause$ tag. Figure 3 shows the cause clause in red color text and the highlighted annotation tags.

<guilt> Rosie felt so regretful <cause> when she did not reply to her friend's letter in time. <\cause> <\guilt>

Fig. 3. Examples of Cause Clauses

3.1.3. Statistics of Dataset

The dataset contains 5947 sentences with the following classification category wise. The emotion wise distribution is given in the Table 1 below.

Table 1. Statistics of Dataset

happy	sad	disgust	fear	shame	guilt
1000	1000	1018	999	1000	930

3.2. Data Preparation and Preprocessing

The sentences with annotations are appended in the dataset document. A pickle file is made containing the synonym words for emotions.The number of sentences of each emotion is counted by an if statement that checks the tag.

The sentences are then split to form clauses. We've split the sentences from full stop, commas and conjunction like "when" and "because". Clauses are taken as an entity.All the processing of classification and filtering will be done on these clauses.

Since cause is a triggering "event" and in English language conjunctions are used before a causation, such splitting is done for accurate emotion-cause pairs. We've used "when" and "because" since most of the causes follow them.

Prepositions like ”to”, ”after”, ”while” are not used because there is an equal possibility that they’re used to denote time, reason, location or any other context.

The clauses are then cleaned.

```
Emotion - Cause Pairs and Labels

(['Octavia', 'felt', 'anxious'], ['when', 'she', 'saw', 'others',
'crying', 'she', 'felt', 'very', 'tense', 'and', 'as', 'time', 'went',
'by', 'her', 'fear', 'grew']) fear
```

Fig. 4. Emotion Cause pairs and labels

This is followed by classifying these clauses in emotion and cause clauses. Clauses having an emotion word from the pickle file of synonyms are appended in the emotion-clause list.Clauses ending with a closing tag $< \backslash cause >$ are appended in the cause-clause list. Each clause in the document is then labeled as an emotion or cause, by creating two separate label list.

Each clause in a sentence, if it is in emotion clause list, and the other one is in cause clause list, their pairs are formed and labeled with the emotion tag of the emotion clause. If they are not a valid pair, for example the second example of Figure 4 is not, since it is a (cause, emotion), it is labeled ’None’. The flowchart of complete preprocessing is shown in Figure 5.

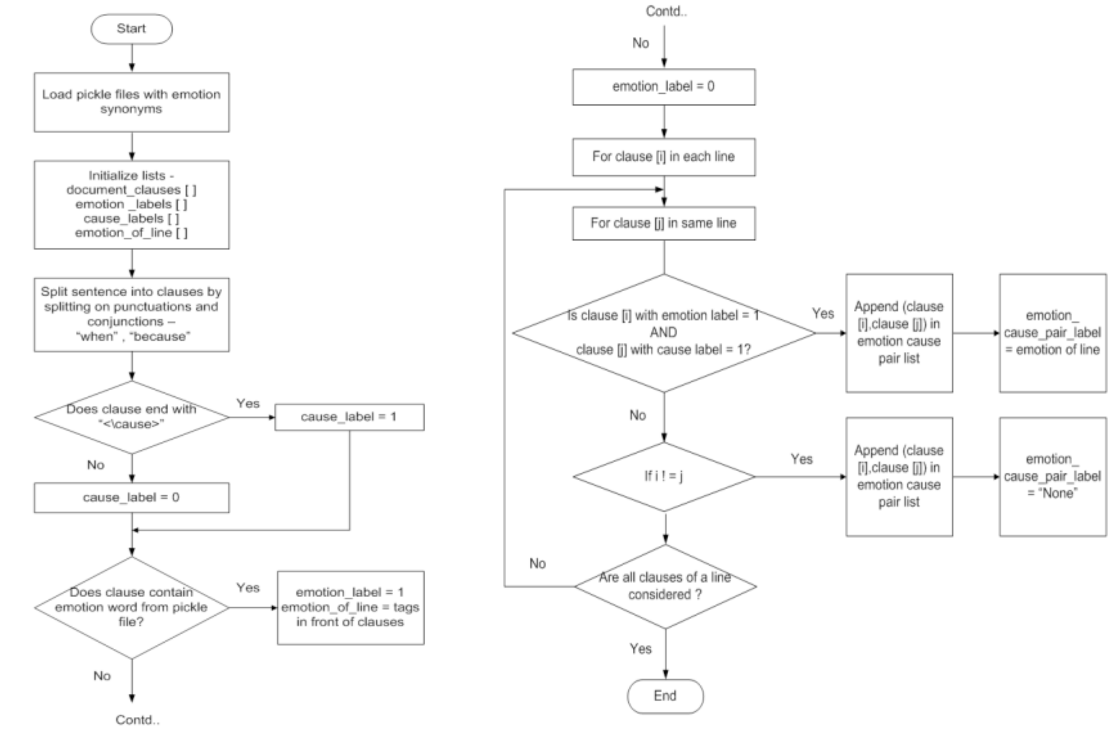

Fig. 5. Flowchart of preprocessing

3.3. Emotion Cause Classification

Firstly, the embedding layer was initialised using word2vec vectors.

We have used two models for the classification task, for the sake of comparison. One contains BiLSTM layers and the other contain BiLSTM + attention layers. Our deep learning model is a multi-task model i.e. two classifiers are trained in the same mode - one for emotion and the other one for cause.

The upper layer has two components for two associated tasks - emotion extraction and cause extraction. Both of them are Bidirectional-LSTM [18] on clause-level, receiving independent clause representations.

Our document is made of n clauses,

$$D = C_1, C_2, C_3, C_4....C_n \tag{3}$$

And each clause is made of many words.

We employ a hierarchical Bi-LSTM to capture such word-clause- document representation. The first set of layer contains word-level BiLSTM modules to capture the context information of the sentence. The output from these BiLSTM modules go into two components - for emotion and cause extraction.

Each of these are their own clause - level BiLSTM modules receiving representations obtained previously. The hidden states in these are context aware representations of the clauses. These are finally fed into softmax layer for prediction. In the BiLSTM + Attention Model, after the lower BiLSTM module there is an attention mechanism [19] to get clause representations according to relative importance between the words.

Figure 6 a) and b) shows the architecture of deep learning models used for emotion and cause classification using BiLSTM and BiLSTM + Attention mechanism respectively.

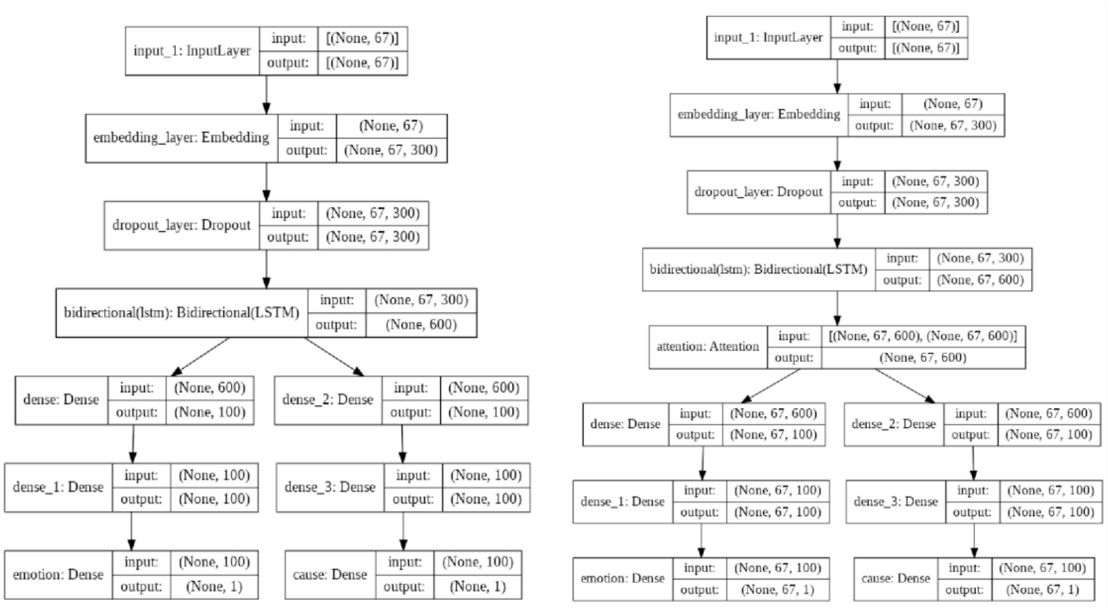

Fig. 6. (a) BiLSTM multi task model for classification; (b) BiLSTM + Attention multi task model for classification

3.4. Emotion Cause Filtering

After Step I, we have obtained a set of emotion clauses and a set of cause clauses.

$$E = c^e_1, c^e_2, c^e_3, c^e_4... \tag{4}$$

$$C = c^c_1, c^c_2, c^c_3, c^c_4... \tag{5}$$

Now, we performed a cartesian product between emotion and cause clauses of a particular sentence to form prospective pairs.

$$P = E \times C \tag{6}$$

$$P_{all} = ...c_i^e, c_j^c... \tag{7}$$

These are all emotion cause pairs that are possible. However, our task is now to filter the correct pairs out of this set. The pairs which will have a causal relationship among them will be correct.

This step is difficult than the last step as it requires semantic understanding to find the validity of a pair. In this step, firstly word2vec representations are used again as embedding for emotion and cause clauses. Then it is passed through a dropout and BiLSTM layer. After that, the output from the emotion and cause ends are concatenated, and flattened to then produce a final output showing if the pair is valid or not. In BiLSTM + Attention Model, an attention model is used right before concatenation.

Figure 7 a) and b) shows the Emotion Cause Filtering Model using BiLSTM and Emotion Cause Filtering Model using BiLSTM and Attention respectively.

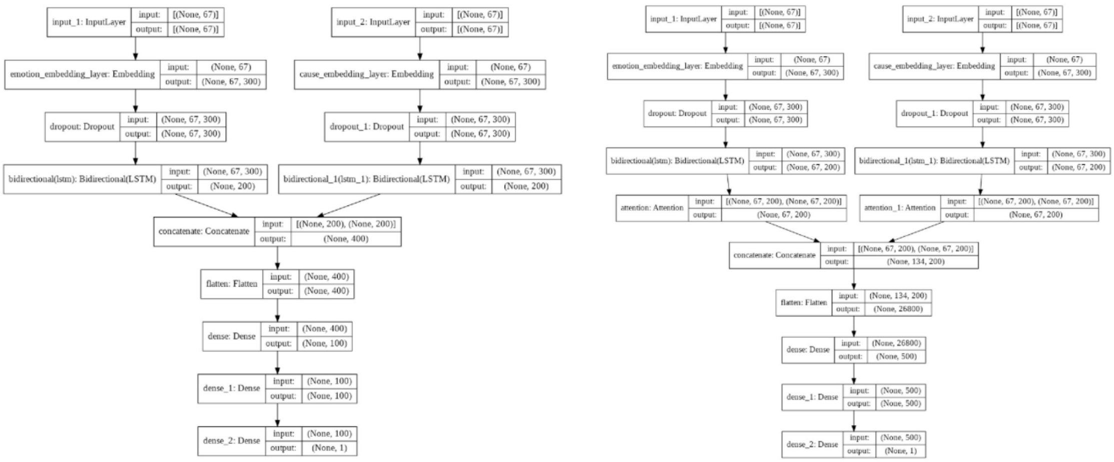

Fig. 7. (a) Emotion Cause Filtering Model using BiLSTM; (b) Emotion Cause Filtering Model using BiLSTM and Attention Model

3.5. Experimental Settings

The entire dataset is divided into Train Data 72%, Validation Data 8% and Test Data 20% by ensuring the distribution.

The maximum sequence length of clauses is 67. The size of the embedding layer 300 and so is the BiLSTM hidden units. The classification models all have BiLSTM with 2 hidden layers, activation function sigmoid, dropout of 0.8 and 100 units. The loss model is binary crossentropy, with Adam optimizer and a learning rate of 0.005. The batch size is 32. For filtering models, the BiLSTM again uses two hidden layers with sigmoid activation and 100 units. The dropout of emotion and cause embedding here is 0.01. It also uses a binary crossentropy model, with Adam optimizer and a learning rate of 0.001.

3.6. Evaluation Metrics

We've used precision, recall and F1 score for evaluating our models.

For precision, recall and F1-score, we first need to find certain operands.

1. **True Positive** - A true positive is when the model predicts the positive class properly and correctly. In our case, it consists of the pairs which are predicted by the model, and are actually valid.

2. **False Positive** - A false positive occurs when the model forecasts the positive class inaccurately. In our case, it consists of pairs which are predicted valid by the model, but do not have a causal relationship between them.

3. **False Negative** - A false negative occurs when the model predicts the negative class inaccurately. In our case, it is the set of pairs that are valid but not predicted by the model.

In our case, we have calculated the precision, recall and F-score by generating formulae and operands.

3.6.1. Precision

The formula for calculating the precision is

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{8}$$

In our approach, correct pairs refers to valid emotion cause pairs extracted from machine i.e. pairs that were TRUE and EXTRACTED.

Proposed pairs refers to all the pairs that were proposed by the machine as valid emotion cause pairs i.e. pairs that were EXTRACTED as true and may or may not be true.

Thus, precision for our approach is calculated by

$$Precision = \frac{CorrectPairs}{ProposedPairs} \tag{9}$$

3.6.2. Recall

The formula for calculating Recall is

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{10}$$

Again, correct pairs refers to valid emotion cause pairs i.e. pairs that were TRUE and EXTRACTED.

Annotated pairs refers to all the pairs that were extracted from the document which were valid i.e. all the emotion cause pairs which are TRUE (but may or may not be extracted).

Thus, recall for our approach is calculated by

$$Recall = \frac{CorrectPairs}{AnnotatedPairs} \tag{11}$$

3.6.3. F1 Score

For situations requiring both precision and recall, a model that optimises the F1 score formula can be chosen.

The formula for F1 Score is -

$$F1score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{12}$$

4. Observation and Results

The proposed approach aims at finding the emotion-cause pair list from a document without needing annotation of an emotion. This section discusses observations and results. We've obtained results in four ways.

- Classification Results
- Experimental Results of the proposed approaches
- Comparison of results with benchmark research

4.1. Classification Results

We used a multi-task model for classifying emotions and causes.

Table 2 shows the results of classification step from BiLSTM Model for emotions and Table 3 shows the results for classification of causes. The results show an accuracy of 95% for the classification task.

Table 4 shows the results of classification step from BiLSTM + Attention Model for emotions and Table 5 for causes. The results show a classification accuracy of 95%.

Table 2. BiLSTM Emotion Classification results

Measure	0	1	**Accuracy**	**Macro Avg**
Precision	0.9661	0.9504	0.9564	0.9582
Recall	0.9227	0.9786	0.9564	0.9506
F1-score	0.9439	0.9643	0.9564	0.9541

Table 3. BiLSTM Cause Classification results

Measure	0	1	**Accuracy**	**Macro Avg**
Precision	0.9593	0.9448	0.9536	0.9520
Recall	0.9637	0.9382	0.9536	0.9509
F1-score	0.9615	0.9415	0.9536	0.9515

Table 4. BiLSTM + Attention Emotion Classification results

Measure	0	1	**Accuracy**	**Macro Avg**
Precision	0.9508	0.9507	0.9508	0.9508
Recall	0.9241	0.9684	0.9508	0.9462
F1-score	0.9373	0.9595	0.9508	0.9484

Table 5. BiLSTM + Attention Cause Classification results

Measure	0	1	**Accuracy**	**Macro Avg**
Precision	0.9493	0.9590	0.9530	0.9541
Recall	0.9740	0.9213	0.9530	0.9476
F1-score	0.9615	0.9398	0.9530	0.9506

4.2. Experimental Results of the proposed approaches

In this section, we discuss results after the filtering process as well. For the sake of comparison, we have experimented with four combinations of approaches. These are -

- BiLSTM Classification and BiLSTM Filtering
- BiLSTM + Attention Classification and BiLSTM + Attention Filtering
- BiLSTM + Attention Classification and BiLSTM Filtering
- BiLSTM Classification and BiLSTM + Attention Filtering

Table 6 shows the result comparisons of all these approaches on the basis of evaluation metrics, and the number of pairs they found out. All the approaches have a precision of 1.0 meaning all the pairs that were found were correct. The recall and F1 Score however vary over the approaches with ”BiLSTM for classification and BiLSTM + Attention for filtering” scoring the highest amongst them. The column ”Annotated Pairs” shows the total correct emotion-cause pairs in the document which is 1221 for everyone. The ”Proposed Pairs” are the number of pairs which were proposed by our approach. The ”Correct Pairs” are the number of pairs which were valid out of the one proposed. As the proposed and correct pair count is same in all the approach, it means that all proposed pairs are correct. That is why the precision is 1.0.

Table 6. Experimental Results

Proposed Approaches	Precision	Recall	F1 - score	Annotated Pairs	Proposed Pairs	Correct Pairs
BiLSTM Classification and BiLSTM Filtering	1.0	0.58	0.74	1221	719	719
BiLSTM + Attention Classification and BiLSTM + Attention Filtering	1.0	0.57	0.73	1221	700	700
BiLSTM Classification and BiLSTM + Attention Filtering	1.0	0.61	0.76	1221	753	753
BiLSTM + Attention Classification and BiLSTM Filtering	1.0	0.57	0.73	1221	704	704

4.3. Comparison of results with benchmark research

We've used three existing benchmarks for comparison. Xia *et al.* [12] first coined the Emotion Cause "Pair" Extraction task and performed it on a Chinese Dataset having around 13,000 instances and achieved an F1 Score of 61.28% Fan *et al.* [20] performed emotion cause extraction using a knowledge regularized hierarchical structure on English dataset having 2145 instances and achieved an F1 Score of 59.75% Singh *et al.* [13] performed an end to end Emotion cause extraction on English Dataset having 2843 instances and achieved an F1 score of 65.63%.

The tabular comparison is shown in Table 7.

Table 7. Comparison with benchmark research

Reference	Approach	F1 score	Dataset Language	No. of instances
Xia *et al.* [12]	Two Step Emotion Cause Pair Extraction	61.28%	Chinese	13,000
Fan *et al.* [20]	Knowledge Regularized Hierarchical Neural Network	59.75%	English	2145
Singh *et al.* [13]	End to End Emotion Cause Pair Extraction	65.63%	English	2843
Proposed Approach	**BiLSTM Classification and BiLSTM + Attention Filtering**	**76.2%**	**English**	**5947**

5. Conclusion and Future Works

In this paper, we proposed an enhanced English dataset for the training task of deep learning. In addition to this, we used multi-task deep learning based approach for Emotion Cause Pair extraction. For the purpose of experimentation, 5947 sentences spanning across six emotions - happy, sad, fear, disgust, shame and guilt, in English language was prepared. Expanding the existing benchmark research on emotion cause pair extraction, we have used BiLSTM and BiLSTM + Attention Networks for classification and filtering purposes. As per our knowledge, most of the existing research bifurcates the sentences into clauses on punctuations. We have bifurcated the sentences into clauses on conjunctions that are mostly used before a "cause" phrase. We have classified the clauses into emotion and cause causes. We have then, performed a cartesian product on those clauses and filtered out the valid pairs, that is, pairs having a causal relationship among them. This method takes into account the causal relationship of emotions and their causes, as well as does not require emotion to be annotated when cause needs to be extracted. From our different models for experimentation, we got the highest performance score with BiLSTM classification model and BiLSTM along with Attention layer for filtering model. It achieved an F1 Score of 76.2% on the yet largest available dataset which we constructed, outperforming the existing benchmark research. In the future, expansion of this dataset to include more instances and much wider emotion range can be done. As the accuracy of first step could affect the filtering task and cartesian product exponentially increases the number of pairs, this task can be condensed in one step.

Other future advancements in emotion cause extraction and analysis can be -

- Taking implicit emotion expressions into account and extracting its causes.

- Explore effective ways to model discourse relationship among clauses.
- Exploring fine grained methods such as span representations instead of coarse grained clause representation.
- Designing effective methods to include appropriate linguistic knowledge in neural network.
- Include more semantic and contextual knowledge in the task.

References

[1] Q. Gao, J. Hu, R. Xu, G. Lin, Y. He, Q. Lu, and K.-F. Wong, "Overview of ntcir-13 eca task," in *Proceedings of the NTCIR-13 Conference*, 2017.

[2] S. Y. M. Lee, Y. Chen, and C.-R. Huang, "A text-driven rule-based system for emotion cause detection," in *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pp. 45–53, 2010.

[3] Y. Chen, S. Y. M. Lee, S. Li, and C.-R. Huang, "Emotion cause detection with linguistic constructions," in *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, pp. 179–187, 2010.

[4] L. Gui, L. Yuan, R. Xu, B. Liu, Q. Lu, and Y. Zhou, "Emotion cause detection with linguistic construction in chinese weibo text," in *CCF International Conference on Natural Language Processing and Chinese Computing*, pp. 457–464, Springer, 2014.

[5] K. Gao, H. Xu, and J. Wang, "A rule-based approach to emotion cause detection for chinese micro-blogs," *Expert Systems with Applications*, vol. 42, no. 9, pp. 4517–4528, 2015.

[6] L. Gui, J. Hu, Y. He, R. Xu, Q. Lu, and J. Du, "A question answering approach to emotion cause extraction," *arXiv preprint arXiv:1708.05482*, 2017.

[7] Y. Chen, W. Hou, X. Cheng, and S. Li, "Joint learning for emotion classification and emotion cause detection," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 646–651, 2018.

[8] S. Yada, K. Ikeda, K. Hoashi, and K. Kageura, "A bootstrap method for automatic rule acquisition on emotion cause extraction," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 414–421, IEEE, 2017.

[9] B. Xu, H. Lin, Y. Lin, Y. Diao, L. Yang, and K. Xu, "Extracting emotion causes using learning to rank methods from an information retrieval perspective," *IEEE Access*, vol. 7, pp. 15573–15583, 2019.

[10] X. Xiao, P. Wei, W. Mao, and L. Wang, "Context-aware multi-view attention networks for emotion cause extraction," in *2019 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pp. 128–133, IEEE, 2019.

[11] J. Hu, S. Shi, and H. Huang, "Combining external sentiment knowledge for emotion cause detection," in *CCF International Conference on Natural Language Processing and Chinese Computing*, pp. 711–722, Springer, 2019.

[12] R. Xia and Z. Ding, "Emotion-cause pair extraction: A new task to emotion analysis in texts," *arXiv preprint arXiv:1906.01267*, 2019.

[13] A. Singh, S. Hingane, S. Wani, and A. Modi, "An end-to-end network for emotion-cause pair extraction," *arXiv preprint arXiv:2103.01544*, 2021.

[14] A. Khunteta and P. Singh, "Emotion cause extraction-a review of various methods and corpora," in *2021 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC)*, pp. 314–319, IEEE, 2021.

[15] D. Ghazi, D. Inkpen, and S. Szpakowicz, "Detecting emotion stimuli in emotion-bearing sentences," in *International Conference on Intelligent Text Processing and Computational Linguistics*, pp. 152–165, Springer, 2015.

[16] K. Scherer and H. Wallbott, "International survey on emotion antecedents and reactions (isear)," 1990.

[17] K. R. Scherer and H. G. Wallbott, "Evidence for universality and cultural variation of differential emotion response patterning.," *Journal of personality and social psychology*, vol. 66, no. 2, p. 310, 1994.

[18] J. Lafferty, A. McCallum, and F. C. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," 2001.

[19] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[20] C. Fan, H. Yan, J. Du, L. Gui, L. Bing, M. Yang, R. Xu, and R. Mao, "A knowledge regularized hierarchical approach for emotion cause analysis," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 5618–5628, 2019.