Case Study: Emotional Cause-Pair Extraction

# Preamble

This case study is a summary of a recent technical publication, published in 2019 at the Annual Conference of the Association for Computational Linguistics (ACL 2019). The title of the paper is Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts, authored by Rui Xia and Zixiang Ding of Nanjing University of Science and Technology, China. This paper is about natural language processing (NLP), but does not cover much of the fundamentals covered in class; the paper primarily focuses on more advanced machine learning modeling techniques, but does relate directly to our learning in a few ways. First, the work depends on pre-processing to extract clauses out of text. This is assumed, but not touched upon directly. Secondly, the team uses several of the same evaluation techniques that we did in class, namely precision, recall, and F1 scores.

# The Problem

Sentiment analysis is a fundamental NLP task, and has a wide variety of applications. While the detection of emotion in text has come a long way, there is still a lot of work to be done on emotional cause extraction. Understanding the relationship between a textual “cause” and the resultant emotion in text can significantly improve sentiment analysis tools. Current state-of-the-art emotion cause extraction (ECE) approaches utilize a two-pass methodology. In step 1, the text is analyzed and the emotion in the text is annotated. During step two, these ECE methods then analyze the text to identify words or clauses that “cause” the annotated emotion.

There are a few problems with these approaches that this paper attempts to resolve. First, given the two pass approach, traditional ECE methods can be quite limited in real life scenarios, where emotions are not pre-annotated. The proposed methodology does the extraction of cause-effect pairs in a single step, eliminating the pre-annotation requirement. Second, the researchers believe that emotion extraction itself can also be improved, due to emotions and their causes not being mutually independent.

# Solution Overview

The solution to this problem is broken down into two main steps. It is worth noting, due to the intricacies of emotion, this work is done at the multi-word clause level rather than the individual word level. The primary innovation of this solution is the extraction operation performed in step one. In this step, the researchers extract both the emotion clauses and the cause clauses simultaneously. The researchers take advantage of the co-dependency of cause and effect in order to improve upon the efficacy of the extraction. To accomplish this, a custom designed Long Short Term Memory (LTSM) neural network model is implemented, baking in the co-dependency of the two types of clauses.

**1.** Set of m extracted emotion clauses. **2.** Set of n extracted cause clauses.

The second step of the ECPE technique is significantly simpler than the first. Step two uses the sets of extracted emotion clauses and cause clauses to construct pair-wise relationships between the two sets of extracted clauses. This is done using a Cartesian product of the two sets, followed by a logistic regression model acting as a filter.

# Methods and Procedures

Prior to step one, the document is first broken in to a set of clauses to be classified, denoted d, below. Here, the ith clause is denoted as ci­. The process of identifying clauses is outside of the scope of this paper and is not covered. The end goal of this process is to transform this list of |d| clauses into a set of pairs P, containing pairs of cause clauses cc and emotion clauses ce.

## Independent Extraction

The first step of this procedure is to take each clause in d, and classify it as an emotion clause or a cause clause. First, the researchers examine a base approach, where cause and emotion extraction is independent of one another. This is done using a “Bi-LTSM” model (blue), which accumulates contextual information of the clauses (green) as the model processes each word in the model. The researchers did not develop this model and do not touch upon the underlying theory in depth. Below, a diagram of the model used can be seen:

A close up of a map

Description automatically generated

The other method noted is ‘attention’ (also denoted blue). Attention in a machine learning technique used, especially in text analysis, to ensure that long sentences/clauses are processed by the neural network properly. Speaking broadly, this technique works by allowing the neural network to identify important words and phrases in a sentence, ignoring words with less information. This, like Bi-LTSM, is not a technique developed in this paper, but rather just utilized in their process.

The final layers of the model examines two hidden states of the BiLTSM. In orange, variables and denote the hidden, context-aware measurements of emotion and cause, respectively, for clause i. The values are fed into a softmax function (red), where the clause is identified as either an emotion clause, or a cause clause.

## “Interactive Multi-Task Learning”

The heart of this paper builds on the ideas of independent extraction, building in feedback between the clause and emotion branches of the model. While intimidating at first, this extension is actually quite simple. In essence, the researchers are simply applying the model twice to each the cause and emotion extraction, rather than once. We can understand this by examining just the left branch. The first, lowest four layers match exactly the left branch in the independent emotion clause extraction of the previous section. However, this value is used as an input to the next stage of the model.

In the second stage, and are used as inputs to the BiLSTM model again, where we extract a value to determine if the clause is a cause clause. This second step allows the model to include a dependency information between emotion an cause clauses. A very important detail worth noting is the memory inherent to the LSTM model. This memory allows the model to learn how subsequent clauses might effect one another (a previous cause clause can imply a subsequent emotion clause).

A close up of a map

Description automatically generated

## Emotion-Cause Pairing

The process of pairing the now extracted sets of clauses is much simpler. First, a cartesian product of all possible pairs is developed.

Next, each pair is represented by a feature vector, containing values from the extraction model, and a vector vd, which represents the distance between the two clauses.

The researchers train a logistic regression model on each x to detect if there is a causal relationship between the pairs of clauses. Any clause-pairs that the model does not show have a causal relationship are discarded, leaving only the expected cause-effect pairs.

# System Development

The use cases for this system largely overlap use cases for any sentiment analysis system, with some additional uses as well. For cases when even only emotion/sentiment extraction is needed, the results of this experiment show that this technique has superior results than state-of-the-art emotion extraction techniques. These tools have a wide variety of uses, but here are a few examples:

1. **Email Writing Assistant:** A sentiment analysis tool that analyzes emails live, prior to sending. This tool can gauge the emotion in an email, warning the user if the email appears to aggressive, or too <emotion>. This use cause will require a large amount of labeled text to train on, but many corpuses of this information is already publicly available. In addition to pure sentiment analysis, this tool could also aid to identify emotionally causal phrases to the user, who can then rephrase their email if necessary. This tool can also be used in other similar scenarios, such as a plug-in for Microsoft Word.
2. **Policy Guidance/Political Speech Writing:** This methodology can be used in the political sphere, to both guild political policy as well as speech writing. Data on public emotional responses to policy, speeches, and debates is already widely collected, and could be used to train this model. This could be done in a multitude of fashions, from simply providing insight to speech writing, or to analyzing the cause-effect relationship between political moves and public emotional response.

One of the key benefits to this kind of system over traditional sentiment analysis, is the ability to unearth latent sentiment in the causal phrases themselves.