Emotion detection for written texts: techniques, their limitations and general challenges

Finn Alberts
Student pre-Master's programme Artificial Intelligence
Open University of The Netherlands
Email: finnalberts@gmail.com

Abstract—**Emotion detection is a branch of artificial intelligence and focuses on identifying specific emotions such as happiness, sadness, or anger. This paper researches techniques for detecting emotions in written texts by studying literature from 2020 onwards. We zoom in on these different techniques, how they work and what their limitations are. We discuss the lexicon based approach, the rule based approach, machine learning, deep learning, transfer learning, transformer models, hybrid approaches, and the Multi-label Emotion Detection Architecture (MEDA). Apart from discussing limitations of each technique specifically, we discuss general challenges faced across all techniques. We especially look into challenges caused by a lack of datasets, both quantitatively and qualitatively. We also discuss challenges with detecting implicit emotions and sarcasm, as well as looking at the detection of multiple emotions for a single input.**

I. INTRODUCTION

Emotion detection is the branch of artificial intelligence (AI) that is focused on detecting specific emotions based on speech/voice, image, or textual data [1]. It is possible to detect emotion using multiple aspects, or just a single one, like written text. The latter does not take other elements into account such as facial expressions, and therefore makes it more difficult to correctly detect an emotion or multiple emotions [1].

This paper explores the many different approaches to detecting emotions in written text, by giving an overview of how they work, and highlighting their limitations. Furthermore, this paper aims to show the general challenges for emotion detection for written text, applicable to all discussed approaches. The explicit research question this paper aims to answer is *"What are the limitations of current techniques for emotion recognition in written text?"* To answer this research question, research has been conducted on existing literature solely.

The rest of this paper is built up as follows. Section 2 describes the research method used to gather relevant literature. The different techniques for detecting emotion and their limitations are discussed in section 3, and in section 4 general challenges for all techniques are presented. Section 5 describes related works, and the paper concludes with recommendations for future work in section 6.

II. RESEARCH METHOD

A. Search method

To find the works used in this paper, a comprehensive search method was used. First of all, a selection of search queries was used in Google Scholar. These queries were combined with the criterion that all works must be published from 2020 onwards, with an additional search for works from 2023 onwards, ensuring that no works present outdated techniques. An overview showing which query led to which source can be found in Table I.

The results of the queries were evaluated based on relevance for the topic of this paper. Here, we specifically looked for articles about emotion recognition in written text. Articles about emotion recognition in general, combining written text with other aspects such as audio and video, or articles about sentiment analysis were not deemed relevant.

TABLE I
OVERVIEW OF QUERIES AND RESULTING ARTICLES

Query	Article
emotion detection written text limitations	[1] [2] [3] [4]
emotion detection written text review	[1] [2] [3] [4]
emotion detection written text survey	[1] [2] [3] [4]
emotion recognition written text survey	[1] [2] [3] [4]
emotion recognition written text review	[1] [2] [3]
emotion recognition written text limitations	[1] [2]
emotion detection text limitations	[1] [2] [3] [4] [5] [6] [7]

B. Source evaluation

To evaluate the reliability of each source, we first looked at the number of citations for each article. Furthermore, the impact factor of the journal was taken into consideration. We also looked at the reliability of the main author of the articles, by looking at the h-index of the author, the amount of articles published by them, and the number of citations for all published articles combined.

III. EMOTION DETECTION TECHNIQUES AND THEIR LIMITATIONS

A. Lexicon based approach

In a lexicon based approach, also known as a keyword based approach, emotions are detected by looking for emotional keywords. The relations between the keywords and the associated emotions are captured within so called lexicons [3], [4], [7].

The lexicon based approach relies heavily on keywords in the given text that imply a certain emotion. This means that this approach is very limited for detecting implicit emotions. Furthermore, the lexicon based approach is reliant on the keyword being present in the lexicon. Lastly, a challenge for the lexicon based approach is context. The reason for this is that the meaning of a word can drastically change by the context of the sentence (e.g. "Does it seem like I am happy?") [4].

B. Rule based approach

The rule based approach encompasses the lexicon based approach and expands upon it. The key difference is that the rule based approach does not just look at specific keywords, but extracts rules based on linguistics, statistics, and computational concepts [2], [4], [7].

The same problem as seen by the lexicon based approach is seen here. The context of the sentence makes it possible that a sentence is misclassified [2]. Furthermore, rule based approaches are affected by the quality of the text. Additionally, detecting implicit emotions is only possible if a rule in the rule set represents it [4].

C. Machine learning based approach

Another possible approach is the machine learning based approach. Both supervised learning and unsupervised learning are options for detecting emotions [2], [7]. The machine learning based approach can be implemented using one of the more traditional machine learning approaches [2], [3], [7]. According to Alswaidan and Menai, support vector machines (SVM) is the most used approach [4].

In recent years, it has been seen that deep learning has been more robust compared to traditional machine learning based approaches. Deep learning also shows to be outperforming these traditional approaches [2].

D. Deep learning based approach

When using deep learning for emotion detection, there are a few steps involved. First, an embedding layer is built. This layer converts the input into a vector representation. This embedding layer is then fed into a neural network. This neural network finally gives an output [4], [5].

According to Deng and Ren, there are two types of embeddings: typical word embeddings, which are more general and word embeddings specifically designed for emotion detection, called emotional word embeddings [5]. For the neural networks, there also are multiple options to choose from [5], [7]. According to Alswaidan and Menai, Long Short-Term Memory (LSTM) is the most used for emotion detection for written text [4].

E. Transfer learning based approach

Transfer learning is not an approach by itself, but helps to improve machine learning (and especially deep learning). The transfer learning approach allows one to reuse existing models, by using a model from a source domain and transferring it to a target domain. [3], [5].

The most frequently used method for transfer learning in natural language processing (NLP) is using sequential transfer learning, which consists of two stages. The first stage is the pre-training phase, where we try to train a model with universal knowledge of the natural language. In the second stage, the transfer phase, the model is finalized to be used for the target domain [5].

Another approach for transfer learning is multi-task learning (MTL). Here, target and source tasks are related and trained simultaneously. Compared to sequential transfer learning this method is less vulnerable to overfitting [5].

F. Transformer models based approach

Transformer models are a branch of deep learning where the importance of the input data is weighed differently, based on so-called attention vectors in an attention layer [1], [7].

There are many different transformer models, with Bidirectional Encoder Representations from Transforms (BERT) being the most researched transformer based model for emotion detection in written text as of 2021 [1].

When specifically looking at BERT, we can see that it is performing well on extracting contextual information. However, it can only detect emotions in a monolingual text. Furthermore, the size of the input sentence is limited. Lastly, it suffers from pragmatic inference, which means it can make assumptions based on previous knowledge, but is not necessarily true for the given input [1].

G. Hybrid approaches

In the hybrid approach, the rule based approach and machine learning based approach are combined into a unified model. It is also possible to combine the rule based approach with the deep learning based approach or transformer models based approach, as they are both machine learning techniques.

By combining the two techniques, the strengths of both techniques can be utilised, while also making the limitations of the individual techniques less of an issue. However, the hybrid based approach relies on the type of deep learning technique used and therefore requires a well-performing deep learning model to have satisfactory results [2].

H. Multi-label Emotion Detection Architecture (MEDA)

In most of the techniques mentioned before, emotions are seen as independent of each other and detected as such. However, emotions are closely related and impact each other. This means that emotions have correlations, which can be used to increase performance [6].

The Multi-Label Emotion Detection Architecture (MEDA) extracts both emotion-specified features and emotion correlations. The architecture consists of two parts: the Multi-Channel Emotion-Specified Feature Extractor (MC-ESFE) and the Emotion Correlation Learner (ECorL). In the MC-ESFE, each emotion is detected seperately in a dedicated channel. Then, in the ECorL, the correlations between emotions are taken into consideration to give the final prediction [6].

Although MEDA shows satisfactory results, one of its limitations is that weaker emotions can be suppressed by

stronger emotions. This can cause the architecture to not detect these emotions [6].

IV. GENERAL CHALLENGES

Apart from the limitations of each of the techniques mentioned in section 3, there are also some challenges for emotion detection for written text in general.

A main issues lies in the quality and quantity of (labeled) datasets. Issues arise here with spelling mistakes and incorrect grammar. Also new words and emoji are not yet represented well [3], [7]. Furthermore, most datasets are in English, with few other languages being available [2], [3], [4], [7]. The same issue of language goes for the domains the datasets are applicable to. Many domains are yet to be represented well in datasets [2], [3], [7]. Furthermore, many of the datasets show an imbalance in the representation of emotions in them [4], [5], [7].

It is possible to create more datasets to increase the amount of available data, and have more data for currently under-represented emotions or domains. It should be noted that the labelling of the data can be subjective, based on the interpretation of the person doing the labelling [3], [5], [7].

This last point is not only applicable when labelling the data. Boundaries between emotions are fuzzy, making it not only hard for people to properly classify an emotion, but for emotion detection algorithms as well [5].

Another issue making detecting emotions difficult, is implicit emotions. Implicit emotions are emotions which are not explicitly stated in the input, but can logically be derived from it. This, however, asks for the algorithm to understand the linguistics and context of the text, which is a difficult task [4], [7]. The same is true for sarcasm and irony [3], [7].

This lack of context is an issue as well, when detecting emotions in dialogue. Most techniques presented in this paper detect emotions on a sentence level and therefore ignore the context of the rest of the conversation, which may include crucial information for detecting the correct emotion [5].

Lastly, even though we presented MEDA as a possible architecture for detecting multiple emotions, this remains a challenge. Emotions in a sentence can contradict each other and weaker emotions might be surpressed by stronger emotions [3], [6]. Also detecting the intensity of emotions remains a challenge [7].

V. RELATED WORKS

This paper is based on multiple different works. Nandwani and Verma presented a review on emotion detection for written text and gave insight into different emotion detection techniques. They also discussed challenges for emotion detection in written text in general [3]. Alswaidan and Menai dive deeper into some of the same approaches, as well as introducing some new ones. They also present some general challenges for the techniques presented [4]. Acheampong, Wenyu and Nunoo-Mensah also discuss some of these techniques. They too discuss general challenges for emotion detection in written text [2]. In a later paper they zoom in on transformer models for detecting emotions. They specifically zoom in on BERT based approaches [1]. Deng and Ren as well explored emotion detection techniques focused especially on the different approaches for deep learning [5]. They also developed an architecture (MEDA) for detecting multiple emotions and the correlations between them [6]. Lastly, Kusal et al. conducted a systematic literature review, where they discuss different techniques and general challenges [7].

This paper aimed to gather information from all these different sources, to create a single overview of all the techniques presented in these different papers. Furthermore, this overview explicitly states weakness and limitations for these techniques, as to create a complete overview of where there are open issues are for all emotion detection techniques.

VI. CONCLUSION

In summary, this paper discussed different techniques for detecting emotions in written text and their limitations by researching existing works. We looked into the lexicon based approach, the rule based approach, machine learning, deep learning, transfer learning, transformer models, hybrid approaches, and the Multi-label Emotion Detection Architecture (MEDA).

There are multiple limitations to these techniques. A main challenge is the lack of enough high-quality datasets. There are many issues related to this, including a lack of non-English datasets and an imbalance of emotions. Another main issue lies in detecting implicit emotions. Also, fuzzy boundaries between emotions poses an issue to properly classify written text. Lastly, despite the MEDA architecture, detecting multiple emotions remains an issue.

Further research is needed to solve these issues. Especially improvements for the datasets could have a big impact, as many issues are related to them. Furthermore, research in how we can improve detecting implicit emotions and multiple emotions is needed. Research in these areas could open up new doors for emotion detection in written text. This could create many new use-cases for these techniques.

REFERENCES

[1] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, "Transformer models for text-based emotion detection: a review of bert-based approaches," *Artificial Intelligence Review*, pp. 1–41, 2021.

[2] F. A. Acheampong, C. Wenyu, and H. Nunoo-Mensah, "Text-based emotion detection: Advances, challenges, and opportunities," *Engineering Reports*, vol. 2, no. 7, p. e12189, 2020.

[3] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," *Social Network Analysis and Mining*, vol. 11, no. 1, p. 81, 2021.

[4] N. Alswaidan and M. E. B. Menai, "A survey of state-of-the-art approaches for emotion recognition in text," *Knowledge and Information Systems*, vol. 62, pp. 2937–2987, 2020.

[5] J. Deng and F. Ren, "A survey of textual emotion recognition and its challenges," *IEEE Transactions on Affective Computing*, 2021.

[6] ——, "Multi-label emotion detection via emotion-specified feature extraction and emotion correlation learning," *IEEE Transactions on Affective Computing*, 2020.

[7] S. Kusal, S. Patil, J. Choudrie, K. Kotecha, D. Vora, and I. Pappas, "A systematic review of applications of natural language processing and future challenges with special emphasis in text-based emotion detection," *Artificial Intelligence Review*, vol. 56, no. 12, pp. 15 129–15 215, 2023.