Emotion Recognition in Text using Graph Similarity Criteria

Nadezhda Komarova, Inna Novalija, Marko Grobelnik
Jožef Stefan Institute
Jamova cesta 39, Ljubljana, Slovenia
nadezhdakomarova7@gmail.com

ABSTRACT

In this paper, a method of classifying text into several emotion categories employing different measures of similarity of two graphs is proposed. The emotions utilized are happiness, sadness, fear, surprise, anger and disgust, with the latter two joined into one category. The method is based on representing a text as a graph of *n*-grams; the results presented in the paper are obtained using the value of 5 for *n*: the *n*-grams were the sequences of 5 characters. The graph representation of the text was constructed based on observing which *n*-grams occur close together in the text; additionally, frequencies of their connections were utilized to assign edge weights. To classify the text, the graph was compared with several emotion category graphs based on different graph similarity criteria. The former relate to common vertices, edges, and the maximum common subgraphs. The evaluation of the model on the test data set shows that utilizing the construction of the maximum common subgraph to obtain the graph similarity measure results in more accurate predictions. Additionally, employing the number of common edges as a graph similarity criterion yielded more accurate results compared to employing the number of common vertices to measure the similarity between the two graphs.

KEYWORDS

emotion recognition, text classification, machine learning, graphs, graph similarity

1 INTRODUCTION

Emotion recognition is a problem that can be connected to different fields such as natural language processing, computer vision, deep learning, etc. [4] In this paper, the focus is on the task of recognizing emotions in texts.

In the literature, several approaches have been introduced that target this problem. Some of them employ vertex embedding vectors for emotion detection and recognition from text. The embedding vectors grasp the information related to semantics and syntax; however, a limitation of such approaches is that they do not capture the emotional relationship that exists between words. Some methods attempting to alleviate this issue include building a neural network architecture adopting pre-trained word representations. [3] Some text classification approaches employ *n*-grams to construct the text representation, e.g., to deal with the task of language identification. [9]

In this paper, the approach to emotion recognition employs *n*-grams to obtain graph representation of text. The text is viewed as a sequence of characters that is divided into *n*-grams, i.e., shorter overlapping sequences of characters as presented in Figure 1.

Information Society 2022, 10–14 October 2022, Ljubljana, Slovenia

In Section 2, it is further explained how the graph of *n*-grams is constructed for a given text and how an emotion label is assigned to the text based on the similarity with the emotion category graphs. Afterwards, in Section 3, the method is compared with related approaches.

In Section 4, an overview of results is focused on differences between the performance of the model when different graph similarity criteria are used. It is followed by the discussion of the model's limitations in Section 5.

2 PROPOSED METHOD

2.1 Constructing the Graph of *n*-grams

The method used in the paper to obtain text representation in the form of the graph of *n*-grams is the following.

- The given text was separated into *n*-grams of characters. Also, different values of *n* have been tested. The results in Section 4, use $n = 5$. The *n*-grams into which the given text was split were overlapping.
- The *n*-grams obtained in this way were utilized to represent the labels of vertices of the graph.
- The edges of the graph were created in the following manner. The ends of edges were the vertices that corresponded to *n*-grams that occurred close to each other in the text, e.g., the edge is connecting the first *n*-gram at the beginning of the text with the second *n*-gram (these two *n*-grams would overlap with each other), as seen in Figure 1. Different values have been tested for the maximal distance between the two vertices allowed for these two vertices to still be connected with the edge. The results in Section 4, use the value of 7.
- Performance of the model with both, the directed and the undirected graphs, has been tested.

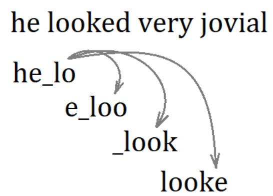

Figure 1: Constructing the edges between the 5-grams that occur close to each other

In Figure 2, it is depicted how the edges are constructed between the vertices labelled with *n*-grams. For the clarity of representation, each *n*-gram is shown connected to 3 other *n*-grams instead of 7. It is important to note that if the same *n*-grams occurred in the text more than once, there was still only one vertex with this *n*-gram as a label: the connections of the *n*-gram have been aggregated at a single vertex.

Additionally, the graph constructed is weighted. The weights of the edges are obtained utilizing the frequencies of connections of n-grams in the given text. In other words, the edge weights are initialized to 0. Then, when constructing the graph of n-grams for a text, every time a certain edge would be added, instead of adding it, the weight of the edge is increased by 1.

Afterwards, the edge weights are normalized to be in the range (0, 1); hence, the edge weights are more comparable among the graphs of n-grams for different texts.

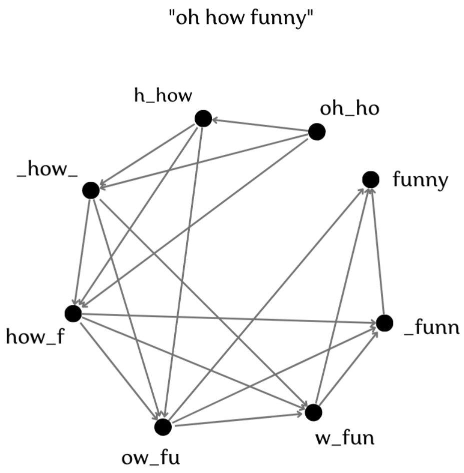

Figure 2: Constructing the edges between the 5-grams in the text fragment *"oh how funny"*

2.2 Constructing the Emotion Category Graphs

The core of the method is the construction of the graph of n-grams as described in Section 2.1. In the data set used to tune the model, there were shorter texts labelled with one of the following 5 emotions: happy, sad, surprised, fearful, or angry-disgusted. Overall, there were 1207 sentences included in the data set; out of this, the model was trained using 1086 sentences (to construct the emotion category graphs) and evaluated on 121 sentences (the split proportion is 90 : 10).

The process of obtaining the emotion category graphs is presented below.

(1) The data set was split into 5 parts containing only the text labelled with the same emotion.
(2) Then, the texts in each part of the data set were used to obtain 5 graphs corresponding to each emotion.
   (a) This process can be viewed as for each text labelled with a certain emotion, constructing the graph of n-grams as explained in Section 2.1.
   (b) Afterwards, merge these graphs separately for different emotions to obtain 5 larger graphs of n-grams; during the merging process, the edges are aggregated in such a way that there are not any two vertices in the emotion category graph sharing the same label (the character n-gram to which they correspond).

2.3 Assigning an Emotion to a Given Text

Utilizing the 5 emotion category graphs corresponding to different emotions, for a given text, it is determined, to which emotion the text most likely corresponds. For that, the pairwise similarity measures of the graph of the given text and of the 5 emotion category graphs are employed.

In other words, it is tested, to which of the 5 graphs the graph of the given text is most similar and the corresponding emotion is assigned to the given text.

Several similarity criteria of the two graphs have been explored.

(1) The number of *vertices* common to both graphs: the vertices are considered common if they share the same label (the n-gram they represent) in both graphs.
(2) The number of *edges* common to both graphs: the edge is considered common if the same vertices (vertices with the same labels) are the endpoints of the edge in both graphs and the edge weights are the same.
(3) The number of vertices in the *maximum common subgraph* (MCS) of the two graphs. Finding the maximum common subgraph is equivalent to finding a graph with the maximum number of vertices so that it is a subgraph of each of the two graphs. [8]
(4) The number of edges in the maximum common subgraph (MCS) of the two graphs.
(5) $z = \frac{m(m-1)}{2} - e$, where m denotes the number of vertices in the maximum common subgraph of the two graphs, and e denotes the number of edges in the maximum common subgraph.

3 RELATED WORK

In the literature describing related approaches to text classification and emotion recognition, deep learning models are often utilized to obtain high-quality predictions. [7]

Apart from the approaches that employ word embedding vectors [6], there are also methods that connect neural networks and graphs. Such approaches may be similar to the method described in this paper since the graph representation of text may be obtained in a similar way based on the semantic connections between words. One example of this kind of model is the graph neural network that is enhanced by utilizing BERT to obtain semantic features. [11]

The crucial part of the method in this paper is the graph similarity criterion that is used when comparing the graph of the given text with different emotion category graphs. The similar way as the construction of the maximum common subgraph is used in this method, it can be employed in combination with the probabilistic classifiers. [10]

The approach in this paper, on the other hand, does not employ probabilistic classifiers such as Bayes Classification or Support Vector Machine. [2] Instead, the emotion for which the similarity measure between the corresponding emotion category graph and the graph of the given text is maximised is assigned to the text.

Additionally, it is important to note that it is possible to incorporate alternative graph similarity criteria, e.g., related to subgraph matching, edit distance, belief propagation, etc. [5]

4 RESULTS

4.1 Experimental Setup

The data set used to train and evaluate the model was the one distributed by Cecilia Ovesdotter Alm. [1] It included the sentences each labelled with one of the following emotions: happiness, sadness, fear, surprise, anger, and disgust. The latter two emotions were joined into one category.

During the evaluation stage, for each sentence, a corresponding emotion was predicted, e.g., the text *"then the servant was*

Table 1: Results of text classification using directed graphs

Similarity criterion	Accuracy	Precision	Recall	F1
Common vertices	0.488	0.506	0.332	0.323
Common edges	0.537	**0.683**	0.408	0.432
z	0.372	0.074	0.200	0.108
Vertices in the MCS	0.570	0.622	0.426	0.446
Edges in the MCS	**0.579**	0.625	**0.454**	**0.478**

Table 2: Results of text classification using undirected graphs

Similarity criterion	Accuracy	Precision	Recall	F1
Common vertices	0.488	0.506	0.332	0.323
Common edges	0.554	**0.669**	0.429	**0.460**
z	0.372	0.074	0.200	0.108
Vertices in the MCS	0.545	0.527	0.399	0.406
Edges in the MCS	**0.570**	0.581	**0.439**	0.453

greatly frightened and said it may perhaps be only a cat or a dog" was labelled fearful, while the text *"he looked very jovial did little work and had the more holidays"* was recognized to be related to the emotion of happiness.

The value of n that appeared to yield the best results and was also used to obtain the results in Tables 1 and 2 was 5. Furthermore, each 5-gram (except those at the end of the text) is connected to 7 5-grams further in the text.

In Tables 1 and 2, the "common edges" criterion means that the two edges from both graphs are considered common if they have the same weight and the same endpoints.

Additionally, in Table 1, z denotes the difference between the the actual number of edges in the maximum common subgraph and the number of edges in the complete graph with m vertices, where m is the number of vertices in the maximum common subraph.

In the trials that yielded the results in Table 1, the edges were directed and in the trials that yielded the results in Table 2, the edges were undirected.

4.2 Analysis

From the results in Table 1 and 2, it may be noticed that the highest accuracy on the test data set was achieved when the number of edges in the maximum common subgraph was used as the similarity measure. In Table 1, the second highest accuracy was achieved when the number of vertices in the maximum common subgraph was utilized.

From this, it may be observed that the construction of the maximum common subgraph reflects the similarity better in certain cases; possible reasons may be that deeper semantic relationships can be captured this way since connections between multiple n-grams are considered at the same time.

In Tables 3 and 4, the confusion matrices are presented for the trials when the number of edges in the maximum common subgraph was used as the criterion of graph similarity.

From the Tables 1 and 2, it is evident that this similarity criterion corresponded to the highest accuracy of predictions for both undirected and directed graphs. However, the accuracy corresponding to this similarity criterion is higher when the graphs are directed (0.579 compared to 0.570).

Table 3: Confusion matrix: directed graph, number of edges in the MCS as the similarity criterion

Actual/pred.	Happy	Fearful	Surpr.	Sad	Angry-Disg.
Happy	**43**	1	0	0	1
Fearful	7	**6**	1	3	0
Surprised	6	1	**2**	1	1
Sad	12	1	0	**12**	1
Angry-Disg.	11	2	0	2	**7**

Table 4: Confusion matrix: undirected graph, number of edges in the MCS as the similarity criterion

Actual/pred.	Happy	Fearful	Surpr.	Sad	Angry-Disg.
Happy	**42**	1	0	1	1
Fearful	8	**6**	1	2	0
Surprised	6	1	**1**	1	2
Sad	11	1	0	**13**	1
Angry-Disg.	11	2	0	2	**7**

Table 5: Confusion matrix: directed graph, number of common edges as the similarity criterion

Actual/pred.	Happy	Fearful	Surpr.	Sad	Angry-Disg.
Happy	**42**	1	0	2	0
Fearful	10	**4**	0	3	0
Surprised	6	0	**2**	3	0
Sad	13	0	1	**12**	0
Angry-Disg.	16	1	0	0	**5**

Table 6: Confusion matrix: undirected graph, number of common edges as the similarity criterion

Actual/pred.	Happy	Fearful	Surpr.	Sad	Angry-Disg.
Happy	**41**	1	0	2	1
Fearful	11	**4**	0	2	0
Surprised	6	0	**2**	3	0
Sad	12	0	1	**13**	0
Angry-Disg.	14	1	0	0	**7**

Furthermore, the accuracy corresponding to the similarity criterion being the number of the common edges (considering both the endpoints and the weight of the edge) is higher by 0.017 when the graphs are undirected than when the graphs are directed (0.554 compared to 0.537). When the graphs utilized are undirected, the model might be more flexible regarding the exact order of the words that occur together.

In Tables 5 and 6, confusion matrices are presented for the trials when the number of edges common to both graphs, considering the endpoints and the weights of the edges, was used as the the criterion of graph similarity.

5 DISCUSSION

A strength of the approach presented in this paper is the ability to capture the context of the given words on different levels; this is related to the process of constructing the edges of the graph by connecting n-grams that occur together in the text. Additionally, the breadth of the contextual frame considered may be varied by altering the number of *n*-grams with which a certain *n*-gram is connected when constructing the edges.

However, overall, the accuracy values noted in Tables 1 and 2, were not very high possibly indicating that the training data set was not large enough. Moreover, the data set did not include texts corresponding to different emotions in even proportions resulting in an imbalance which could have also had a detrimental influence on the quality of predictions. The confusion matrices (Tables 3, 4, 5, and 6) indicate, e.g., that the texts were often falsely assigned the emotion of happiness since it was the most abundant class in the data set.

One of the limitations of the design of the model described it that although it may be reasonable to expect that to obtain more accurate predictions on the test data set, training the model (obtaining the emotion category graphs) on a larger corpus of texts is needed, this may bring a significant rise in computational complexity since the category graphs would possess significantly larger amounts of vertices and edges.

This is especially important if the maximum common subgraphs are constructed when obtaining a similarity measure, since for each text in the test data set, a maximum common subgraph would have to be constructed several times: between the graph of *n*-grams for a given text and each emotion category graph (5 such graphs in this case).

A possible solution to the problem of having too large category graphs might be reducing the length of *n*-grams, i.e., using smaller values of *n*, and hence reducing the number of vertices in the graph.

Also, reducing the number of *n*-grams with which a certain *n*-gram is connected when constructing the edges of the graph may be investigated as a possible solution. However, if this value is too low, too much contextual information may be lost; therefore, it appears necessary that for each value of n, the optimal number of *n*-grams with which a certain *n*-gram is connected is determined experimentally.

6 CONCLUSION

In this paper, the model that utilizes graph similarity criteria to classify a given text into one of the emotion categories is described. The core of the method is to construct a graph of *n*-grams for a given text and to compare this graph to each of the emotion category graphs. The text is classified into the emotion category, the graph of which yielded the highest similarity value when compared to the graph of the given text.

From the results of the trials noted in Tables 1 and 2, it may be concluded that among the graph similarity criteria described, that number of edges in the maximum common subgraph resulted in the highest quality of predictions.

Furthermore, it may also be noted that employing the number of edges common to both graphs resulted in higher prediction accuracy than using the number of common vertices (0.537 and 0.488 accuracy for the directed graphs).

This may appear to be intuitively reasonable as using edges may seem to incorporate more contextual information. Additionally, it may be important to investigate the effect of the difference between the size of the graph of *n*-gram for the given text and the size of the emotion category graph on the probability that the same connections between the two *n*-grams are found in both graphs. Moreover, it may be more probable that the same vertices (vertices labelled with the same *n*-gram) are contained in both graphs resulting in more noisy data.

To conclude, the future work on the task of emotion recognition related to the proposed method may, on the one hand, be focused on employing alternative graph similarity measures in addition to those described in this paper, e.g., those connected to deriving the edit distance or to the belief propagation. [5] Furthermore, clustering algorithms may be used to obtain the patterns characteristic to the emotion categories and further employ them for the emotion recognition task. To this end, both, the vertex clustering algorithms as well as the clustering of graphs as objects, might be utilized. Additionally, graph neural network architecture may be built along with incorporating the graphs of *n*-grams as the input for the network.

7 ACKNOWLEDGEMENTS

This work was supported by the Slovenian Research Agency under the project J2-1736 Causalify and the European Union through Odeuropa EU H2020 project under grant agreement No 101004469.

REFERENCES

[1] Alm, E. C. O. Affect in text and speech, 2008.
[2] Bahritidinov, B., and Sanchez, E. Probabilistic classifiers and statistical dependency: The case for grade prediction. pp. 394–403.
[3] Batbaatar, E., Li, M., and Ryu, K. H. Semantic-emotion neural network for emotion recognition from text. *IEEE Access 7* (2019), 111866–111878.
[4] Guo, J. Deep learning approach to text analysis for human emotion detection from big data. *Journal of Intelligent Systems 31*, 1 (2022), 113–126.
[5] Koutra, D., Ramdas, A., Parikh, A., and Xiang, J. Algorithms for graph similarity and subgraph matching, 2011.
[6] Li, S., and Gong, B. Word embedding and text classification based on deep learning methods. *MATEC Web of Conferences 336* (01 2021), 06022.
[7] Prasanna, P., and Rao, D. Text classification using artificial neural networks. *International Journal of Engineering and Technology(UAE) 7* (01 2018), 603–606.
[8] Quer, S., Marcelli, A., and Squillero, G. The maximum common subgraph problem: A parallel and multi-engine approach. *Computation 8*, 2 (may 2020), 48.
[9] Tromp, E., and Pechenizkiy, M. Graph-based n-gram language identification on short texts. *Proceedings of Benelearn 2011* (01 2011), 27–34.
[10] Violos, J., Tserpes, K., Varlamis, I., and Varvarigou, T. Text classification using the n-gram graph representation model over high frequency data streams. *Frontiers in Applied Mathematics and Statistics 4* (2018).
[11] Yang, Y., and Cui, X. Bert-enhanced text graph neural network for classification. *Entropy (Basel) 23* (11 2021).