

Style Change Detection Based On Bi-LSTM And Bert

Notebook for PAN at CLEF 2022

Jiayang Zi^a, Ling Zhou^a, Zhengyao Liu^a

^aFoshan University, Foshan, China

Abstract

This article is an overview of PAN 2022's approach to the Style Change Detection task. The goal of this competition task is to identify the text positions for author switching in a given multi-author document. PAN helps us segment the tasks, respectively task1 finds paragraphs with two authors and cuts them into different texts; task2 finds paragraphs with two or more authors, finds the writing style happens changing location. task3 does a more fine-grained lookup on the task requirements of task2 to find where in the sentence the writing style changes. This paper designs a method to handle the task based on a model composed of convolutional neural networks, Bert and bidirectional long-short-term memory networks, using binary classification to judge style changes and author label problems. The f1 scores are obtained, which are 0.67 for task1 and 0.40 for task2 and 0.65 for task3.

Keywords

Style Change Detection, Bi-LSTM, Bert, CNN

1. Introduction

Today is an era that emphasizes intellectual property rights. There are many and secret means of plagiarism. It may be difficult to find out whether an article is suspected of plagiarism by manual work, and the labor efficiency is also very low. Using writing style detection makes the difficult task of detecting plagiarism much easier. By using writing style detection to screen articles, the problematic paragraphs in the article can be marked and then sent to manual detection, improving the detection efficiency and improving the detection accuracy to have the best of both worlds. In addition to detecting plagiarism, it can also classify according to different authors in the same article. This idea fits perfectly with task1 in Pan's 2022 Style Change Detection task [1]. Task 2 is to find all the positions where the writing style changes in the text written by two or more authors, and give the paragraph a corresponding author's number. Detecting author switching is to use the model to learn the characteristics of the training set and use this to judge whether the authors in the article have changed, and on this basis, count the numbers of different authors and assign them to the corresponding paragraphs. task3 is further screened and detected on the basis of task2. The position where the writing style changes may also appear between different sentences in the same paragraph. Task3 is to mark the position where the writing style changes between sentences [2][3].

After a comprehensive analysis of all tasks, this paper proposes a solution for the PAN 2022 Style Change Detection task, based on the BCG model, which consists of BERT, Bi-LSTM, CNN and GlobalMaxPooling. Use BERT to encode the features in the text and convert them into word vectors to better express the relationship between words. Bi-LSTM solves the problem of context dependence to extract relevant features. CNN further extracts the associated features, avoids overfitting and improves the generalization ability of the model, and finally uses the Pooling layer to improve the running speed and accuracy of the model. The data text given by this PAN task can be used to effectively identify and judge the author, and finally, it has achieved good results.

¹CLEF 2022 – Conference and Labs of the Evaluation Forum, September 5–8, 2022, Bologna, Italy

EMAIL: 1109618450@qq.com (A. 1); zhoulingfsu@gmail.com (A. 2);

ORCID: 0000-0002-4307-622X (A. 1); 0000-0001-6861-8980 (A. 2)



© 2022 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

2. Background

PAN 2022 Style Change Detection task is divided into three subtasks.

1. For a text written by two authors that contains a single style change only, find the position of this change.
2. For a text written by two or more authors, find all positions of writing style change.
3. For a text written by two or more authors, find all positions of writing style change, where style changes now not only occur between paragraphs, but at the sentence level.

In addition, PAN also provides three data sets for testing the algorithm, corresponding to three tasks, and each data set corresponds to three parts:

1. training set: Contains 70% of the whole dataset and includes ground truth data.
2. validation set: Contains 15% of the whole dataset and includes ground truth data.
3. test set: Contains 15% of the whole dataset, no ground truth data is given.

The method model in this paper is evaluated based on the above training set and validation set, and finally the test set is input into the model to obtain the experimental results. The generated results are submitted to the TIRA [7] platform, and finally the calculation and evaluation of the model in this paper is completed on the platform, and then the F1 score of the results is output.

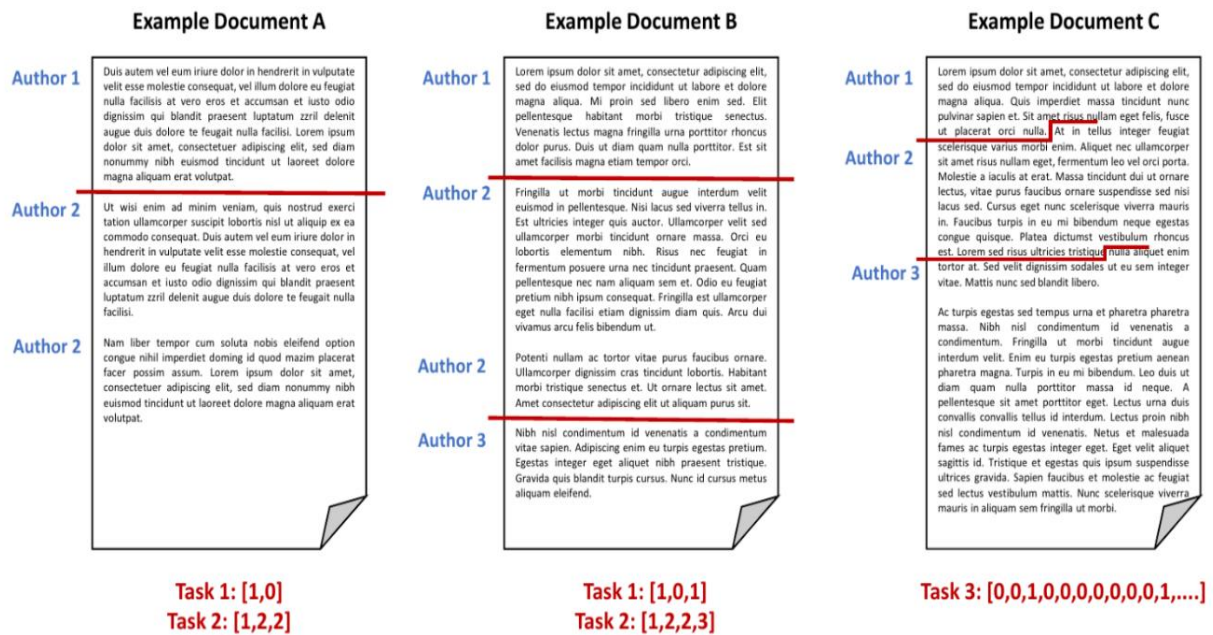


Figure 1: Expected output of simulation scenarios and tasks of Style Change Detection tasks [1]

3. Method

The neural network model proposed in the task is named BBCG model in this paper. It is a neural network structure composed of Bert [4][5], Bi-LSTM, one-dimensional convolution layer, pooling layer and full connection layer, as shown in Figure 2.

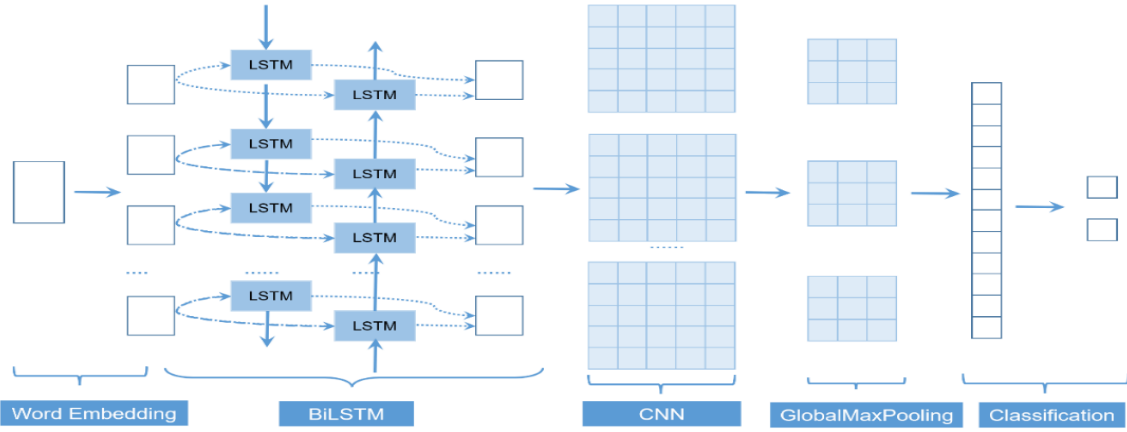


Figure2: Architecture diagram for model

3.1 Word embedding layer

Word embedding is used as the data input layer of the model, and in this task, the pre-training model Bert model proposed in the past four years is first used. This model emphasizes that the traditional single-item language model or the shallow splicing of two one-way language models is no longer used as usual, but the new MLM technology is used [6]. Bert is a deep bidirectional pre-trained language model using Transformers as a feature extractor. The model is used to train the word vector model, that is, it has a deeper number of layers and better parallelism, and has achieved quite good results in various language tasks. the result of. In this layer, it is mainly the process of encoding the words in the text and the corresponding features, and converting them into the form of word vectors for input. Through the mapped word vector form, the vector form is used to better express the relationship between words.

3.2 Bi-LSTM layer

Long short-term memory network (LSTM) is a variant of RNN, which is mainly used to solve the problem of context dependence. Bi-LSTM (Bidirectional Long Short-term Memory) is composed of forward and reverse LSTM. Compared with one-direction LSTM, a single LSTM can only get past information, but it can't get future information. Therefore, bidirectional LSTM can extract more comprehensive features than unidirectional LSTM. As Bi-LSTM is composed of two LSTMs.

Therefore, LSTM is the state h_t at the current time from the value f_t of the forgetting gate, the value i_t of the memory gate and the cell state c_t , the temporary cell state g_t and the output gate o_t . Finally, Bi-LSTM outputs the h_t of two LSTMs in opposite directions respectively and combines them to obtain the output at time t The calculation formula is as follows:

$$f_{(t)} = \delta(W_f x_{(t)} + u_f h_{(t-1)} + b_f), \quad (1)$$

$$i_{(t)} = \delta(W_i x_{(t)} + u_i h_{(t-1)} + b_i), \quad (2)$$

$$g_{(t)} = \tanh(W_g x_{(t)} + u_g h_{(t-1)} + b_g), \quad (3)$$

$$c_{(t)} = f_{(t)} \times c_{t-1} + i_{(t)} \times g_{(t)}, \quad (4)$$

$$o_{(t)} = \delta(W_o x_{(t)} + u_o h_{(t-1)} + b_o), \quad (5)$$

$$h_t = o_{(t)} \times \tanh(c_t), \quad (6)$$

Where W represents the weight and b represents the offset.

3.3 CNN layer

Through the feature relationship extracted by Bi-LSTM, the convolution layer is used to further extract the features, thus reducing the parameter quantity and operation time of the model, and using this method to avoid over-fitting. Improve the generalization ability of the model.

3.4 Pooling layer

Pooling, also known as downsampling layer, is used to extract the average or maximum value in the region (called average pooling and max pooling), and global pooling, which is used to reduce dimensions, such as reducing 3 dimensions to 1 dimension. The most important purpose of the pooling layer is to reduce the dimension, reduce the number of parameters, improve the running speed and accuracy of the model, and avoid overfitting. In the task, global pooling is used, which is used to perform ensemble max-pooling over multiple data.

3.5 Full connection layer

In the fully connected layer, the nonlinear activation function "Relu" is used first, and the piecewise linear function is used for calculation. Finally, the "Softmax" function is used to calculate the output feature to obtain the probability of the input text sentence.

4. Result

Table 1

F1 score and accuracy of the trained model on the verification set

Validation set	Task1	Task2	Task3
Accuracy	0.7581	0.7445	0.6767
F1 score	0.6627	0.4295	0.6623

Table 2

Test set F1 score of the trained model

Test set	Task1	Task2	Task3
F1 score	0.6690	0.4012	0.6483

The result of task 1 is not bad, but task 2 still has a lot of room for improvement. It may be that the initial prediction of whether the author has changed or not has been wrong, resulting in errors in assigning author numbers and the accumulation of errors. The type of task that needs to be judged is sentences. Need to judge whether author changes have been made between shorter texts, and can be done better

5. Conclusion

This paper proposes a neural network-based model to deal with the PAN 2022 Style Change Detection task, trying to answer the three tasks proposed in the task. Although the results obtained based on the model in this paper are better than the baseline, it does not achieve the desired effect, which also shows that for tasks 1 and 2, the model in this paper still has a lot of room for improvement. In task 2, it is necessary to find texts with two or more authors. If there are multiple authors in each paragraph, it will undoubtedly increase the difficulty of detecting the change of writing style between the entire text paragraphs. In the future, we will also consider how to reduce the occurrence of misjudgments due to style approximation in the style detection of long texts. At the same time, it also shows that the Style Change Detection task can have more interesting challenges, such as detecting whether a piece of text is deliberately imitated by other authors, so as to confuse the machine inspection, these tasks will be more challenging.

6. References

- [1] E. Zangerle, M. Mayerl, M. Potthast, and B. Stein, “Overview of the Style Change Detection Task at PAN 2022,” in CLEF 2022 Labs and Workshops, Notebook Papers. CEUR-WS.org,2022.
- [2] Zhang Z, Han Z, Kong L, et al. Style Change Detection Based On Writing Style Similarity[J]. Training, 1970, 11: 17,051.
- [3] C. Zuo, Y. Zhao, R. Banerjee, Style Change Detection with Feed-forward Neural Networks, in: L. Cappellato, N. Ferro, D. Losada, H. Müller (Eds.), CLEF 2019 Labs and Workshops, Notebook Papers, CEUR-WS.org, 2019. URL: <http://ceur-ws.org/Vol-2380/>.
- [4] A. Iyer, S. Vosoughi, Style Change Detection Using BERT—Notebook for PAN at CLEF 2020, in: L. Cappellato, C. Eickhoff, N. Ferro, A. Névél (Eds.), CLEF 2020 Labs and Workshops, Notebook Papers, CEUR-WS.org, 2020. URL: <http://ceur-ws.org/Vol-2696/>.
- [5] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, ArXiv preprint arXiv:1810.04805 (2018).
- [6] E. Grave, P. Bojanowski, P. Gupta, A. Joulin, T. Mikolov, Learning word vectors for 157 languages, arXiv preprint arXiv:1802.06893 (2018).
- [7] M. Potthast, T. Gollub, M. Wiegmann, B. Stein, TIRA Integrated Research Architecture, in: N. Ferro, C. Peters (Eds.), Information Retrieval Evaluation in a Changing World, The Information Retrieval Series, Springer, Berlin Heidelberg New York, 2019. doi:10.1007/978-3-030-22948-1_5.