Style Change Detection Based On Writing Style Similarity

Notebook for PAN at CLEF 2021

Zhijie Zhang, Zhongyuan Han*, Leilei Kong, Xiaogang Miao, Zeyang Peng , Jieming Zeng, Haojie Cao, Jinxi Zhang, Ziwei Xiao and Xuemei Peng

Foshan University, Foshan, China

Abstract

For the Style Change Detection task, the goal is to detect if the document has multiple authors (task1), then find out where the style changes have occurred (task2), and label the author identifier for each paragraph of the document (task3). This paper proposes a method of Style Change Detection based on Writing Style Similarity(SCDWSS). The style changes and the decision of author identifier are regarded as a binary classification task based on the similarity of writing style, and a pre-training model is utilized to estimate the similarity of writing style. Use the proposed SCDWSS, the three tasks of Style Change Detection can be achieved under a uniform framework. Finally, we obtained the F1 scores, which are 0.75, 0.75, 0.50 in task1, task2, task3, and ranked first in task2 and task3.

Keywords

Style Change Detection, Writing Style Similarity, Pre-training Model

1. Introduction

In modern society, the issue of text style change detection has always been important especially. Through the writing style testing, we can easily find whether the document's author is plagiarized and even find how many paragraphs are plagiarized and copied. Focused on the Style Change Detection, PAN 2021 Evaluation Laboratory organizes three related tasks. Among them, task1(Single vs. Multiple) requires us to judge whether one or more authors write the given text. Task2(Style Change Basic) is to find out the writing style changes for the multi-author documents. Task3(Style Change Real-World) needs to label the author identifier for each paragraph of the document, which involves the most critical and challenging task of multi-author documents, different from all the tasks in the past few years.

According to Ref. [1], the tasks of Style Change Detection are commonly recognized as separate tasks, and different models are implemented to solve the respective issues of each task. After analyzing the objectives of different tasks, we conceive of the Style Change Detection as discovering the similarity of writing styles between different text segments. Then, a classification method based on writing style similarity is proposed to address the issues of Style Change Detection. Specifically, we determined that a solution for task1 can be deduced from the solution for task2 and task3 can work with the model of task2. The style changes and the decision of author identifier are regarded as binary classification tasks based on the similarity of writing style. For estimating the writing style similarity, we adopt the popular pre-training model BERT to extract the paragraph features. In this way, we can build a model to do all three tasks simultaneously.

CLEF 2021 – Conference and Labs of the Evaluation Forum, September 21–24, 2021, Bucharest, Romania
EMAIL: zhangzhijie5454@gmail.com (A. 1); hanzhongyuan@gmail.com (A. 2)(*corresponding author); kongleilei@fosu.edu.cn (A. 3)
ORCID: 0000-0002-4854-0618 (A. 1); 0000-0001-8960-9872 (A. 2); 0000-0002-4636-3507 (A. 3)

2. Data

The Style Change Detection provided a data set[6]. It contains texts from a relatively narrow set of subjects related to technology. Statistics of the data set as shown in Table 1.

Table 1
Statistics of data set

Data set	Proportion	Number of texts	Number of authors	Number of paragraphs
Training set	70%	11,200	17,051	77,252
Validation set	15%	2,400	4,792	16,495
Test set	15%	2,400	---	---

3. Method

After analyzing the task definitions, we found that if the task2 label includes 1, the corresponding text will at least be two authors, and the corresponding task1 label will be 1. Otherwise, the task1 label will be 0. We believe that two paragraphs can be taken for similarity measurement. If the similarity is high, the writing style between the two paragraphs does not change. If the similarity is low, the writing style between the two paragraphs has changed. So, it can be a binary classification.

3.1. Identifying author label by binary classification

In task3, the paragraphs-author label includes 1, 2, 3, 4. In order to let three tasks can be achieved under a uniform framework, we convert the task3 label to the binary label, which is called the task3-binary label in this paper. In this way, we can solve task3 skillfully by using the writing style similarity.

In terms of task3-binary label, the principle of the converting is shown in Figure 1. In a document, four paragraphs are denoted as P0, P1, P2, and P3 separately. Then judge whether, for each paragraph and each of its preceding paragraphs, a style change occurs. In the task3-binary label, the label of P0 is always 1. P1 is compared with the P0, the label will be 1 if there is a change; otherwise, the label will be 0. If there are two changes when P2 is compared with preceding P0 and P1, the new label is [1, 1]; if there are no changes, the label is [0, 0]. It will be [1, 0] or [0, 1] if there is a change when P2 is compared with P0 or P1. Every next paragraph should be compared with its preceding paragraphs, and the new label will be 1 if it has changed; otherwise, the new label will be 0. Finally, the task3-binary label is obtained.

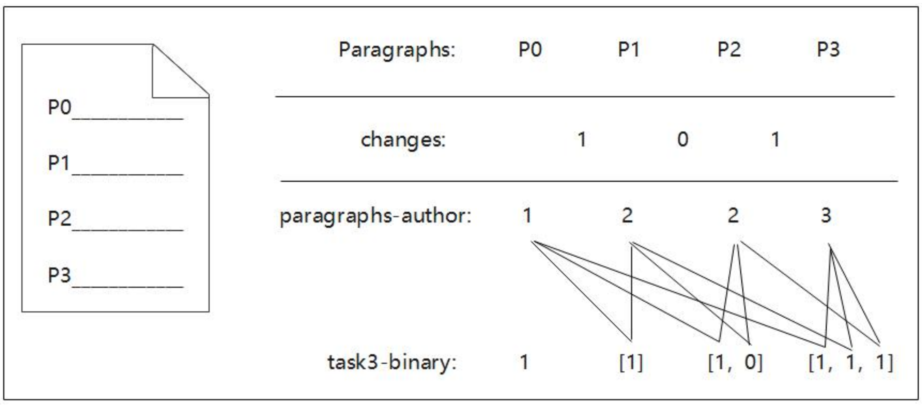

Figure 1: Convert the paragraphs-author label to the task3-binary label

3.2. Estimating Writing Style Similarity

What we use is a pre-training model BERT and Fully Connected Neural Network Classifier. As shown in Figure 2, we input two paragraphs of the document to the model and perform classification, then output the similarity labels.

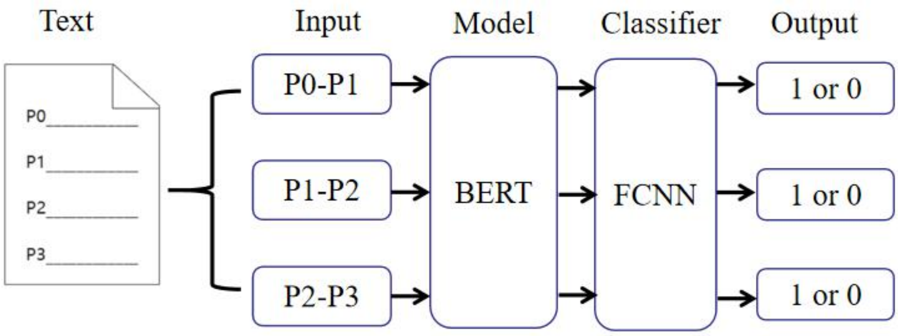

Figure 2: Architecture of the whole model

4. Experimental setting

In our method, we mainly used a popular pre-training model BERT[2]. Considering the resource limitations of TIRA[3], the Base version is adopted after careful consideration.

Before inputting two paragraphs to the model, we set the maximum length of paragraphs to 512 and 256 to analyze the effect. 512, 256 is the sum of two paragraphs in length. Each paragraph can be split into at least one sentence. Intuitively, when the maximum length is 512, it should retain as much information as possible to achieve a good classification effect. However, they were similar when we compared the results. It suggests that the paragraph binary classification task may not require too much information, and the length of the truncated paragraph should not be too long. Since the results are similar, we choose 256 because it saves running space and time.

For fine-tuning, the training set and validation set uses the task3-binary label and task 2 label separately. In this way, the model will be fine-tuned deeply because of the sufficient training data. Our goal is to use only one model to complete task1, task2, and task3, so the above fine-tuning solutions have considered integrating the scores of the three tasks. After training three epochs, it can achieve a better result.

5. Results

The trained model was used to evaluate the validation set, which results are shown in Table 2.

Table 2
The result of validation set

Data set	Task1.F1	Task2.F1	Task3.F1
Validation set	0.85542	0.75193	0.39669

Task3 is a difficult task because the predicted Task3 labels inevitably have a chain reaction. Once there is a predicted error in the labels, it may lead to all the subsequent labels being wrong. That is error accumulation, which leads to a lower result in Task3.

Finally, we obtained better test scores and ranked first in task2 and task3[4][5]. The results as shown in Table 3.

Table 3
Results

Team	Task1.F1	Task2.F1	Task3.F1
Zhang et al.(Our)	0.753	0.751	0.501
Strom	0.795	0.707	0.424
Singh et al.	0.634	0.657	0.432
Deibel et al.	0.621	0.669	0.263
Nath	0.704	0.647	---

6. Conclusion

This paper proposes a method of Style Change Detection based on Writing Style Similarity. The style changes and the decision of author identifier are regarded as a binary classification task based on the similarity of writing style, and a pre-training model is utilized to estimate the similarity of writing style. Use the proposed method, the three tasks of Style Change Detection can be achieved under a uniform framework. Finally, we obtained the F1 scores, which are 0.75, 0.75, 0.50 in task1, task2, task3, and ranked first in task2 and task3.

7. Acknowledgments

This work is supported by the National Natural Science Foundation of China (No.61806075 and No.61772177) and the Social Science Foundation of Guangdong Province (No. GD20CTS02).

8. References

[1] E. Zangerle, M. Mayerl, G. Specht, M. Potthast, and B. Stein, "Overview of the style change detection task at pan 2020." in CLEF (Working Notes), 2020.
[2] Devlin J., Chang M.W., Lee K., et al. Bert: Pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019, 1: 4171-4186
[3] M. Potthast, T. Gollub, M. Wiegmann, and B. Stein, "TIRA Integrated Research Architecture," in Information Retrieval Evaluation in a Changing World, ser. The Information Retrieval Series, N. Ferro and C. Peters, Eds. Berlin Heidelberg New York: Springer, Sep. 2019.
[4] E. Zangerle, M. Mayerl, , M. Potthast, and B. Stein, "Overview of the Style Change Detection Task at PAN 2021," in CLEF 2021 Labs and Workshops, Notebook Papers. CEUR-WS.org, 2021.
[5] J. Bevendorff, B. Chulvi, G. L. D. L. P. Sarracen, M. Kestemont, E. Manjavacas, I. Markov, M. Mayerl, M. Potthast, F. Rangel, P. Rosso, E. Stamatatos, B. Stein, M. Wiegmann, M. Wolska, , and E. Zangerle, "Overview of PAN 2021: Authorship Verification,Profiling Hate Speech Spreaders on Twitter,and Style Change Detection," in 12th International Conference of the CLEF Association (CLEF 2021). Springer, 2021.
[6] E. Zangerle, M. Mayerl, M. Tschuggnall, M. Potthast, and B. Stein, " Pan21 authorship analysis: Style change detection," 2021. https://zenodo.org/record/4589145#.YNfqkmhLhPY