**Cover Letter**

30th August, 2020

Dear Editors and Reviewers,

We would like to express sincere gratitude and appreciation toward the editors and peer reviewers who spent time reviewing our manuscript and provided critical comments to our paper. We have made significant changes to improve the quality of our paper to a level we believe will be satisfactory for the peer reviewer. All reviewers’ comments are responded, and revisions are made accordingly. For clarity, we have color-coded texts as follows:

* The reviewer’s comments in green;
* Our reply in **black;**
* Revised sentences in the manuscript in blue**;**

All changes made in the revised version of the manuscript are outlined in red.

Thank you for your consideration for this manuscript.

Sincerely,

Authors

**Response to the Reviewers’ Comments**

**Response to Reviewer 1**

**Comment 1:**

Using the traditional method, similar to clustering, on the same dataset can only get less than 20% accuracy. This confirms that our proposed method is better than traditional methods. -> The comparative work seems inadequate. Target related to clustering is not clear.

**Response 1:**

Thank you very much for your comment, we apologize that our aim lost its captivating style. We do agree that the description of the FDC method is not clear, and the presentation is not clear enough. In fact, we added a comparison method in section 2, which describes the principle of the FDC method. And we revised this sentence:

Since there are few related work that can use our data set, we almost explored it ourselves. There are very few methods that can be used on our data set, but we still found a traditional method called FDC analysis. We reproduced this method and used it on our data set. The final accuracy rate achieved was only 26.41. %. This accuracy is far less than the accuracy that DPNet can achieve.

Because the characteristics of our data are only diagnosis and prescription, they are completely discrete and without distance, and we can only find one type of FDC after consulting the data to process such data. So in such a situation, our method is novel, and related work is rare. Compared with FDC, the results are also quite excellent, so it can be considered that the effect shown by DPNet is good.

**Comment 2:**

References are too old: 2008 and 2009.

**Response 2:**

Thank you very much for your comment. We updated most of the old references, but one reference had to be left: cite 5. Because related work is very rare, and we only use the two-dimensional features of diagnosis and prescription, other methods are not applicable to our data . And the method in this reference is used by us for a comparison, so this reference needs to be kept. We update the reference:

We adopt five key delay propagation mechanisms focused on how the delay of an example train, train 1 travelling from station A to Station B propagates to some train 2, with the exception of delay propagation mechanism 1 which concerns train 1 exclusively. The pseudo code for the allocation of delay propagation minutes to train journeys for mechanisms 2 to 5 are detailed in Appendix A1 to A4..

**Comment 3:**

In Table 1, terms are not described: RN1, AKC196, BKE316 ...

**Response 3:**

Thank you very much for your comment. We apologize for the confusion and we understand that our claim of the importance of a tabular data format is irrelevant. We added the explanation in the bottom of the Table 1.

**Comment 4:**

L 134-141: Similar results, but one is that cross-validation verifies effect, and the other is that cross-validation shows little impact. The explanation is too short.

**Response 4:**

Thank you very much for your insight. What we want to express is that cross-validation has little effect on our experiment, and its influence on the experimental results is about 0.1%. We have written the description in more detail.

The main idea of ​​cross-validation is to divide the data into very small parts. In each cycle, most of them are selected as the training set, and the rest is the validation set. The model may show different results in different schemes. We used cross-validation to carry out the experiment, but it had little effect on the results, only affecting about 0.1% of ADAccurary.

**Comment 5:**

Descriptions for Table 4 and Table 5 are not enough.

Descriptions for Figure 4 and Figure 5 are not enough.

**Response 5:**

Thank you very much for your concern. We provide more details.

**Comment 6:**

A mixed notation such as Figure 4 or Fig. 5. The author's guide should be checked.

**Response 6:**

We have checked and unified the standard notaion.

**Comment 7:**

Figure 5 shows the system structure, but there are no explanations for each component: H5, Static resources, indicator, Diary, and so forth.

Readers may be not experts in this field.

**Response 7:**

**Comment 8:**

It needs to justify the novelty of the proposed scheme, through full explanations and result interpretation. Overall, it's hard to find out the contribution of research because of the abstraction and simplified expression.

**Response 8:**

Thank you very much for your comment. We apologized for causing you confusing. We have revised the article to describe the work of this article as clearly as possible

**Response to Reviewer 2**

**Comment 1:**

The sub-section 3 2.1. Overview needs the dataset description should do before it, in order to understand the method. So, I suggest to have a sub-section before, describing the dataset.

**Response 1:**

Thank you very much for your comment, we apologize that our aim lost its captivating style. We do agree that the description of the FDC method is not clear, and the presentation is not clear enough. In fact, we added a comparison method in section 2, which describes the principle of the FDC method. And we revised this sentence:

**Comment 2:**

The model description should be re-structured. I suggest to firstly present the whole model, then present the layers.

**Response 2:**

Thank you very much for your comment, we apologize that our aim lost its captivating style. We do agree that the description of the FDC method is not clear, and the presentation is not clear enough. In fact, we added a comparison method in section 2, which describes the principle of the FDC method. And we revised this sentence:

**Comment 3:**

About the model: why does mean having a PRESCRIPTION layer at the end? Authors should explain it better.

**Response 3:**

Thank you very much for your comment, we apologize that our aim lost its captivating style. We do agree that the description of the FDC method is not clear, and the presentation is not clear enough. In fact, we added a comparison method in section 2, which describes the principle of the FDC method. And we revised this sentence:

**Comment 4:**

How the data as labeled? Manually?

**Response 4:**

Thank you very much for your comment, we apologize for not describing this problem clearly. We do agree that the description of How the data as labeled is not clear. We have added descriptions in the abstract and introduction and results of the article, and our data is unlabeled. Because the amount of data is large and there are few abnormal data, we can treat all data as normal data, so that the relationship between diagnoses and prescriptions that model learning is correct in terms of probability.

**Comment 5:**

Sub-section 3.2. Pre-processing is in the middle of results. Put it close to the dataset description.

**Response 5:**

Thank you for your suggest. We changed the order.

**Comment 6:**

Sub-sections 3.3.1. Diagnosis-to-Prescription and 3.3.2. Prescription-to-Diagnosis do not make any sense. Authors should explain what is the reason of these experiments and why they were delineated.

**Response 6:**

Thank y

**Comment 7:**

The description of the results presented in sub-section 3.4. Verification Experiments is very hard to understand. The same for sub-section 3.5. Comparative Experiments.

**Response 7:**

T

**Comment 8:**

There are some results in the discussion that were not presented before: "In most experiments, loss convergent 163 rapidly in the beginning, but in the middle and later stages its loss fluctuates between 10 and 30."

**Response 8:**

T

**Comment 9:**

The section 5. Application is not enough described and it seems not necessary.

**Response 9:**

T

**Comment 9:**

At the end of the paper, authors propose a FC layer based model, because they say that their proposal is still a bit inadequate. The main question is: why not present this new and probably better model instead this one?

**Response 9:**

T

**Response to Reviewer 3**

**Comment 1:**

a) doing a more thorough comparison with other methods of detecting prescription abnormality; one comparative experiment (presented in section 3.5) is really not enough, especially since it is not clear what this experiment is, nor why a "reduced dataset" was used. Why is ADAaccuracy used, if this is not a standard method in the field? Much more must be done to show that this method outperforms the other methods, as claimed.

**Response 1:**

Thank you very much for your comment, we apologize that our aim lost its captivating style. We do agree that the description of the FDC method is not clear, and the presentation is not clear enough. In fact, we added a comparison method in section 2, which describes the principle of the FDC method. And we revised this sentence:

**Comment 2:**

b) The authors should provide evidence (statements with references to relevant papers are sufficient) that prescription abnormalities are an important source of fraud.

**Response 2:**

Thank you very much for your comment, we apologize that our aim lost its captivating style. We do agree that the description of the FDC method is not clear, and the presentation is not clear enough. In fact, we added a comparison method in section 2, which describes the principle of the FDC method. And we revised this sentence: