

In this paper, the authors are concerned about recognizing faculty homepages, which is valuable for obtaining the relevant information of researchers and tracing hot research topics. Recognizing faculty homepages in university websites automatically is difficult, because it faces two issues. Firstly, it is a multimodal classification problem that consists of several data modes, including textual contents, layout, and images. Secondly, recognizing faculty homepages is also a class-imbalanced problem. In order to address the first problem, a multimodal fusion framework is proposed by authors. Then, the Multimodal Generative Adversarial Nets is introduced to rebalance the dataset. The authors conducted extensive experiments on the dataset crawled from several university official websites in the United States to prove the effectiveness of the proposed method. Moreover, the faculty-oriented search engine based on the proposed approach has been designed and implemented.

Strong points:

1. This paper is the good mix of theory and practical applications.
2. The Multimodal Generative Adversarial Nets and the Gated Fusion Network proposed in this paper is effective for recognizing faculty homepages.
3. The paper is well organized.
4. Based on the approach proposed in the paper, the authors designed and implemented a vertical search engine, and put it into use. In addition, the authors published the source code for recognizing faculty homepages.

Weak points:

Comment 1: In the experiment of model comparison, the experimental result is not very convincing because there is no comparison with the classification approaches for the class-imbalanced.

Response: Much appreciate for this invaluable suggestion. Due to the limited pages, we did not cover all of our comparison work. However, we would like to add more details in the revised version. In addition, we have released a full version of our work which has not been submitted to anywhere else for publication. This full version can be accessed at https://github.com/mrspider520/gated_fusion_network/blob/master/paper/paper.pdf. It provides extra experiments about model tuning and model comparison with different data sets in terms of scale and imbalance degree. The proposed framework shows the advantage to deal with larger and more imbalanced data with multimodal interactions. We also released the data and source code with this publication, in which the details about model training and tuning were presented.

Comment 2: If the authors disclose the experimental dataset, the experimental results will be more convincing.

Response: We have disclosed the source code along with the experimental dataset via Github website which can be accessed at https://github.com/mrspider520/gated_fusion_network.git.

Comment 3:

The text in pictures is too small so that they cannot be seen clearly after printing.

In Eq(9), the term " $f_i w_i$ " should be " $h_i w_i$ ".

Some grammar and style issues. For example, in page#6 Section 5, "It mainly consists of the a Multimodal Generative Adversarial Nets (MGANs) and a Gated Fusion Network (GFN)."

Response: Much appreciate for the invaluable comments. The updated manuscript and full version have been professionally edited and proofread by the language specialists from American Journal Experts. The released full version has not been submitted to anywhere else for publication. This full version can be accessed at https://github.com/mrspider520/gated_fusion_network/blob/master/paper/paper.pdf.