

Title:**Discussion of Moves Recognition of Scientific Documents under limited Samples**Zhang Xin^{1,2}, Xu Haiyun³

1 Chengdu Library and Information Center, Chinese Academy of Sciences, Chengdu, 640213

2 Department of Library, Information and Archives Management, University of Chinese Academy of Sciences, Beijing, 100190

3 Business School, Shandong University of Technology, Zibo, 255000

Abstract:

Moves Recognition refers to extracting semantic segments such as research purposes, objects, methods, results, and conclusions from unstructured abstracts. Moves recognition is an important approach for extracting structured information of papers and plays an important role in downstream text understanding tasks.

In this paper, aiming at the common few-shot labeling problem in actual move recognition task, the application of data augmentation and the prompt-based classification paradigm is discussed, and the LIME method is used to semantically explain the model results.

Prompt tuning on large model achieved higher accuracy than the fine-tuned small model on moves recognition task with less training cost, and the f1_score was improved by 2.5%, 4.1% and 3.9% on the three datasets. Combining the accuracy rate and interpretation results, the "method" and "result" moves recognition effect is better (f1_score about 90%), followed by "conclusion" (f1_score>75%), and the "background" and "method" moves are relatively poor (f1_score<70%).

Keyword: Few-shot, Data Augmentation, Prompt Turning, Explainable

Table1 Datasets Description

	Db100	Db500	Db1800
Total	100	500	1800
Background	9	48	234
Object	14	45	131
Method	22	180	581
Result	33	163	578
Conclusion	22	64	276

Table2 Moves Recognition Results (f1 score)

Database \ Classifier	Db100	Db-500	Db1800
Naïve Bayes	0.3380	0.5520	0.5920
SVM	0.3210	0.5350	0.6710
BERT	0.6810	0.7650	0.8000
Prompt(BERT-base)	0.6560	0.7700	0.7980
Prompt(T5-base)	0.7060	0.8060	0.8350

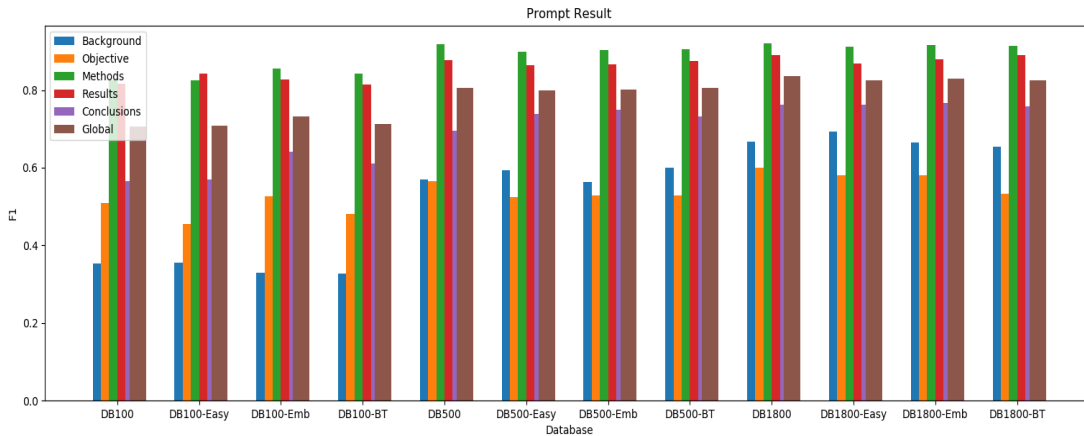


Figure1 Moves Recognition Results of Difference Moves

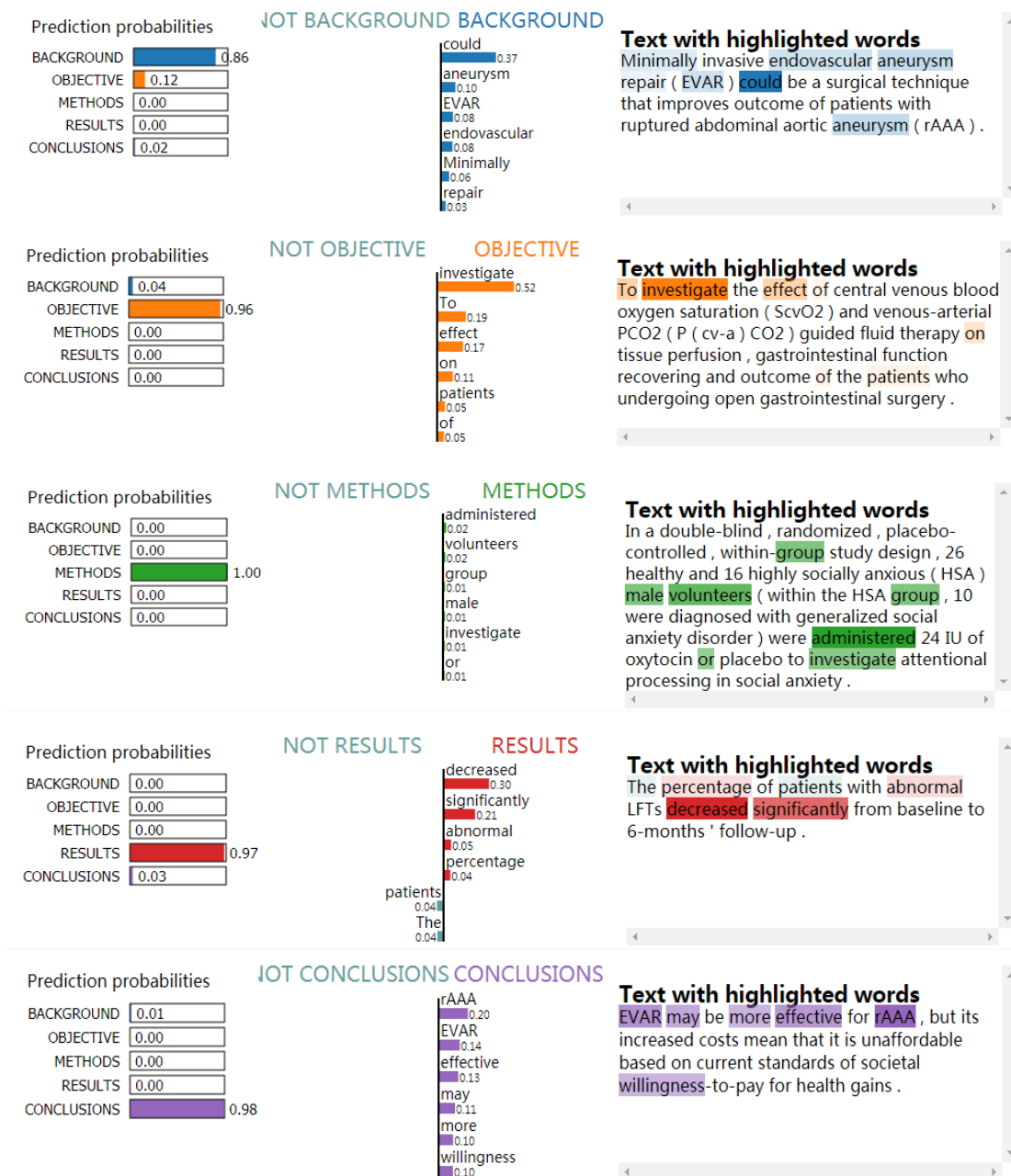


Figure2-6 LIME Explanation of “Background Objective Methods Results” and “Conclusions”