



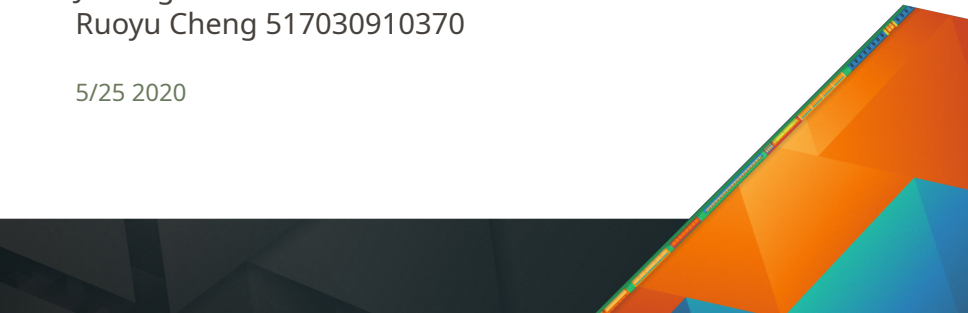
Paper review system


Yixiong Wang 517030910382

Jiasong Guo 517030910374

Ruoyu Cheng 517030910370

5/25 2020





INTRODUCTION

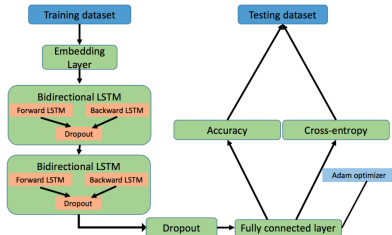
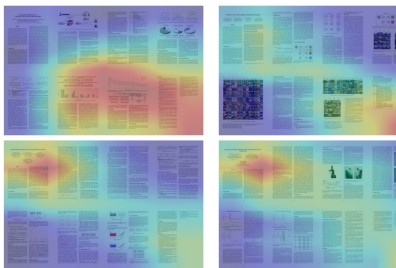
Background
motivation

Background

The tremendous increase in number of submissions to generalist AI conferences has led to a significant decrease of the reviewing quality.

During last IJCAI(2020), a panel has been organized. Summary reject is the most common idea proposed for reducing the reviewing load.

Existing Method



- "Deep Paper Gestalt."
- "An end-to-end learning solution for assessing the quality of Wikipedia articles"
- <https://github.com/Acemap/Acemap-Paper-X-Ray>

Motivation

Interpretability?

LGB + ResNet?


Quality Expectation-Variance Tradeoffs in Crowdsourcing Contests

1 / 7

1501.00267.pdf

1 / 35

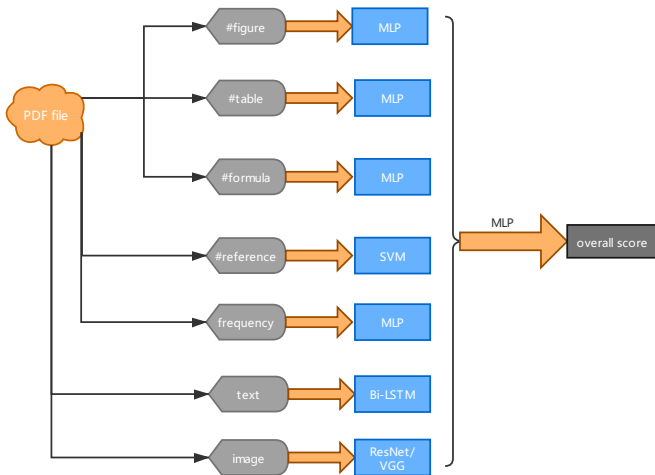
```
The score of (./input/val/conference-pdf/aaai18-12418.pdf) and (./input/val/conference-jpg/aaai18-12418.jpg) is 18.0  
Acc: 0.0  
The score of (./input/val/conference-pdf/aaai16-11924.pdf) and (./input/val/conference-jpg/aaai16-11924.jpg) is 20.3  
Acc: 0.0  
The score of (./input/val/conference-pdf/aaai18-16616.pdf) and (./input/val/conference-jpg/aaai18-16616.jpg) is 17.9  
Acc: 0.0  
The score of (./input/val/conference-pdf/aaai16-12023.pdf) and (./input/val/conference-jpg/aaai16-12023.jpg) is 20.0  
Acc: 0.0  
The score of (./input/val/conference-pdf/aaai14-8294.pdf) and (./input/val/conference-jpg/aaai14-8294.jpg) is 20.8  
Acc: 0.0  
The score of (./input/val/conference-pdf/aaai16-12080.pdf) and (./input/val/conference-jpg/aaai16-12080.jpg) is 20.0  
Acc: 0.0
```



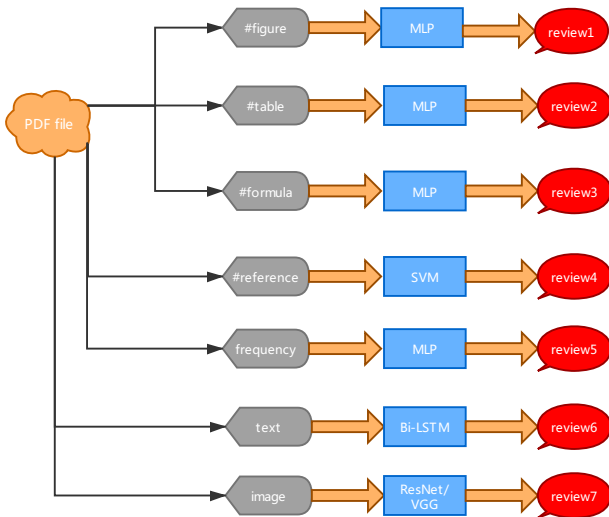
OUR WORK

Framework

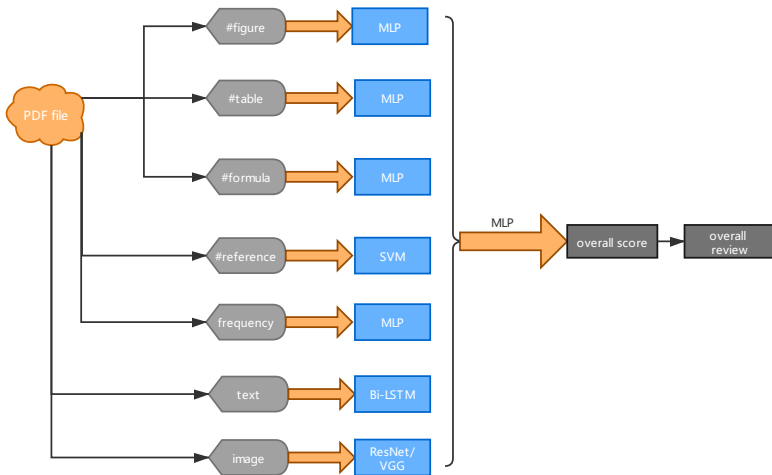
Framework

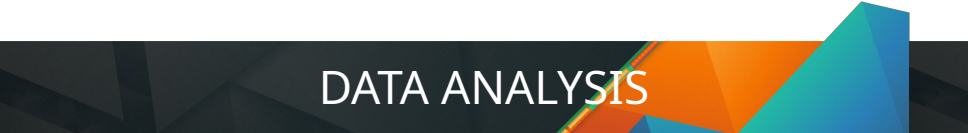


Review Generation



Review Generation



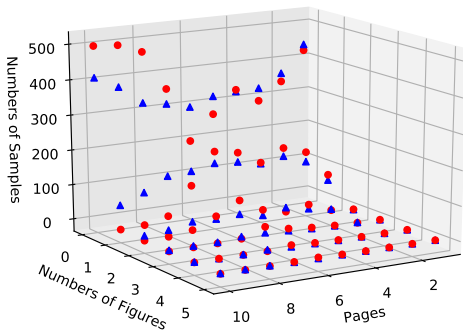


DATA ANALYSIS

Data Distribution

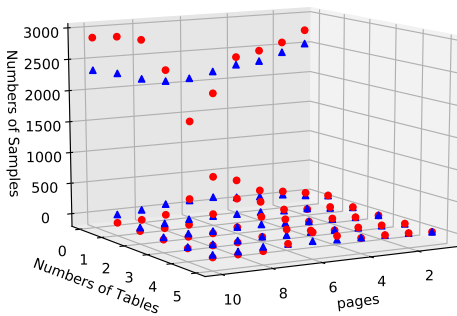
Figures

Distribution of figures



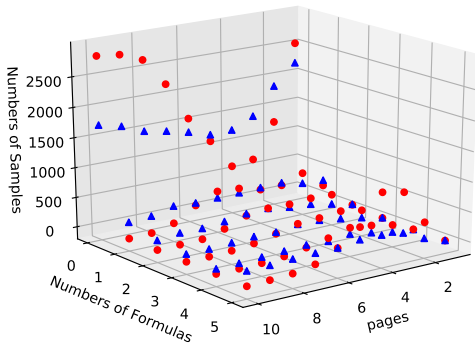
Tables

Distribution of tables



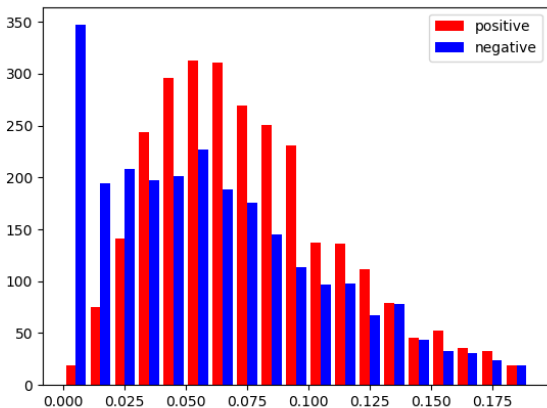
Formulas

Distribution of formulas



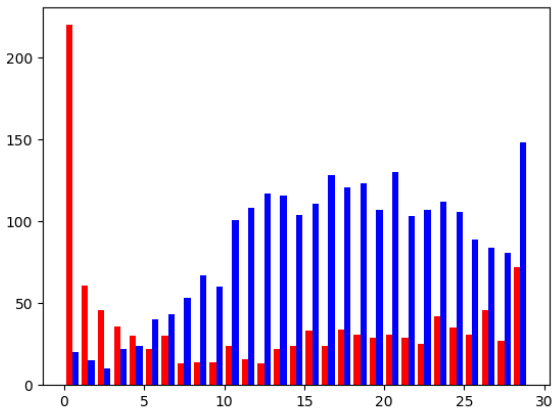
Key Word Frequency

Distribution of hottest word.



References

Distribution of references



Text

Text classification accuracy of Bi-LSTM

```
val_loss: 0.94951 - val_acc: 0.7042
```


Image

Image classification accuracy of ResNet-18(non-pretrained)

```
test_acc: 0.8751651254953765  
8.75975 0.87025 2.15781 0.87517 0.87583
```

Image classification accuracy of ResNet-18(pretrained)

```
test_acc: 0.9233817701453104  
45.68342 0.90905 10.21313 0.92338 0.93395
```

Image classification accuracy of VGG-16(pretrained)

```
test_acc: 0.8533685601056803  
76.3184 0.81562 15.79796 0.85337 0.86592
```



EXPERIMENTAL RESULT

Performance

	Image(VGG)	Image(ResNet)	Text	Reference	Frequency	Figure	Table	Formula
Acc	0.85	0.92	0.70	0.80	0.67	0.79	0.73	0.82

	Feature*	overall(VGG)	overall(Resnet)
Acc	0.90	0.98	0.99

Review

- Discriminative Analysis Dictionary Learning (AAAI)
- Service Function Chaining Resource Allocation: A Survey


Review

Result for "Discriminative Analysis Dictionary Learning":

```
$ python review.py --path="Discriminative Analysis Dictionary Learning.pdf"
[0,1,1,1,1,0.968,0.99608094]
visualizations not good, needs more figure to explain yourself
Sufficient proof for your result
Rigorous derivation of mathematical theory, provides a theoretical basis for your
algorithm, technically sound
The reference citations are considered acceptable
Well-structured context makes it clear to all readers
The paper is well-organized
The paper is clearly written
Advice: Accept
```

Result for "Service Function Chaining Resource Allocation: A Survey":

```
$ python review.py --path="Service Function Changing Resource AllocationASurvey.pdf"
[0,1,0,0,1,0.071,0.00267258]
visualizations not good, needs more figure to explain yourself
sufficient proof for your result
Need more derivation to proof your algorithm, not technically sound
Please cite more or less references from reliable sources
well-structured context makes it clear to all readers
Confusing page layout
Check for grammatical errors
Advice: Reject
```



FUTURE WORK

Future Work

- More data to test scalability.
- Implement multi-conference review.



WORK DIVISION

Work Division

- Jiasong Guo : Put forward the idea of these interpretable features and implement image part(ResNet & VGG)
- Yixiong Wang : pdf parsing, data analysis, interpretable multi-dimension feature extraction and data visualization
- Ruoyu Cheng: Implementation on MLPs and Bi-lstm model, and put forward the structure of overall score

THANKS FOR WATCHING!