AI CUP

●●●

人工智慧論文機器閱讀競賽之論文分類

組員：施彥安、何鑑家、黃彥瑋

Outline

- Task

- Experiment

- Try new things

- Conclusion

Task

- Multi-label Classification
  - Theoretical
  - Engineering
  - Empirical
  - Others
- Train: 5400, Val: 600, Test: 1000
- Measure: macro-F1 score

Only DNN

- CNN
- RNN

Experiment - 1

- Paper
  - Comparative Study of CNN and RNN for Natural Language Processing
- All using **glove** pretrained embeding matrix
- Use **optimal** hyperparameter
  - hidden size
  - minibatch size
  - learning rate
  - maximal sentence length
  - filter size(CNN only)

Experiment - 1 - result

- CNN:
  - good with classification tasks
  - when the sequence gets longer, it become worse
  - a lot faster

- Bi-RNN: better with sequential information

Model	F1-score
Bi-RNN	67.349
CNN	66.715

Experiment - 1 - result

Engineering

```
[('network', 556),
 ('based', 429),
 ('using', 417),
 ('learning', 407),
 ('system', 274),
 ('deep', 256),
 ('neural', 246),
 ('image', 226),
 ('model', 217),
 ('data', 203)]
```

Theoretical

```
[('network', 502),
 ('learning', 309),
 ('based', 281),
 ('system', 216),
 ('model', 209),
 ('using', 187),
 ('analysis', 168),
 ('algorithm', 137),
 ('multi', 136),
 ('data', 135)]
```

Engineering

```
[('network', 2912),
 ('model', 2890),
 ('data', 2697),
 ('method', 2599),
 ('based', 2421),
 ('system', 2242),
 ('algorithm', 2068),
 ('paper', 1964),
 ('image', 1893),
 ('approach', 1821),
 ('using', 1799),
 ('result', 1677),
 ('problem', 1649),
 ('learning', 1610),
 ('proposed', 1608),
 ('performance', 1483),
 ('time', 1471),
 ('show', 1379),
 ('propose', 1160),
 ('feature', 1130)]
```

Theoretical

```
[('network', 2741),
 ('model', 2569),
 ('algorithm', 2314),
 ('problem', 2216),
 ('method', 2062),
 ('based', 1810),
 ('data', 1809),
 ('system', 1747),
 ('paper', 1644),
 ('result', 1591),
 ('show', 1453),
 ('approach', 1402),
 ('time', 1383),
 ('learning', 1317),
 ('proposed', 1234),
 ('using', 1163),
 ('information', 1112),
 ('two', 1102),
 ('also', 1056),
 ('one', 997)]
```

Experiment - 2

- Paper
  - Efficient Deep Learning Model for Text Classification Based on Recurrent and Convolutional Layers (IEEE 2017)

- Mixing CNN & RNN
- CNN
  - Pooling: loss **long-term** dependencies
  - Replace pooling layer with BRNN
- New model
  - Convolution + BRNN(1 layer)

Experiment - 2 - result

Model	F1-score
BRNN	67.349
CNN	66.715
C + BRNN	**67.964**

Try new things

- BRNN
  - maintain long-term dependencies
- CNN
  - good with extracting features
- Our model
  - BRNN(1 layer) + CNN(with pooling)

Model	F1-score
BRNN	67.349
CNN	66.715
C + BRNN	67.964
Our model	**68.866**

Conclusion

DNN for Natural language process
- Construct the DNN suits your task
  - RNN
    - Generally good for all NLP
  - CNN
    - Short texts
    - Classification

- Long sequence text classification
  - BRNN + CNN