Special Project HW3

Few-shot Emotion Classification

Team A
吳冠緯 梁正

Experiment

- transfo, cnn, avg

- three sub-experiments

Experiment (1) -transfo

```
21/11/12 11:41:18, Save best model to /content/Meta-NLP/saved-runs/goemotions_filter_on_dailydialog_test_supervised_transfo/best
21/11/12 11:41:18, acc mean  0.7250 (std  0.1876), f1 mean  0.6653 (std  0.2115), mcc mean  0.6964 (std  0.2191),
transfo & mlp &  0.7250 \tiny $\pm  0.1876$ &  0.6653 \tiny $\pm  0.2115$ &  0.6964 \tiny $\pm  0.2191$
test_data
[0 6 0 ... 0 4 4] {0, 1, 2, 3, 4, 5, 6}
21/11/12 11:41:28, acc mean  0.2958 (std  0.2025), f1 mean  0.2376 (std  0.1812), mcc mean  0.1679 (std  0.2620),
transfo & mlp &  0.2958 \tiny $\pm  0.2025$ &  0.2376 \tiny $\pm  0.1812$ &  0.1679 \tiny $\pm  0.2620$
```

Experiment (2) -transfo

```
21/11/12 13:03:45, Save best model to /content/Meta-NLP/saved-runs/dailydialog_tmp_train_transfo_pretrain/best
21/11/12 13:03:46, acc mean  0.5617 (std  0.1774), f1 mean  0.4747 (std  0.1907), mcc mean  0.5052 (std  0.2165),
transfo & mlp &  0.5617 \tiny $\pm  0.1774$ &  0.4747 \tiny $\pm  0.1907$ &  0.5052 \tiny $\pm  0.2165$
test_data
[6 3 4 ... 6 4 4] {1, 2, 3, 4, 5, 6}
21/11/12 13:03:56, acc mean  0.6110 (std  0.1873), f1 mean  0.5331 (std  0.2038), mcc mean  0.5659 (std  0.2251),
transfo & mlp &  0.6110 \tiny $\pm  0.1873$ &  0.5331 \tiny $\pm  0.2038$ &  0.5659 \tiny $\pm  0.2251$
```

Experiment (3) -transfo

```
dailydialog_u_test
test_data
[6 3 4 ... 6 4 4] {1, 2, 3, 4, 5, 6}
21/11/12 13:08:03, acc mean  0.5867 (std  0.2772), f1 mean  0.5349 (std  0.3171), mcc mean  0.5510 (std  0.3157),
transfo & proto &  0.5867 \tiny $\pm  0.2772$ &  0.5349 \tiny $\pm  0.3171$ &  0.5510 \tiny $\pm  0.3157$
```

Experiment (1) -cnn

```
21/11/12 14:12:28, Save best model to /content/Meta-NLP/saved-runs/goemotions_filter_on_dailydialog_test_supervised_cnn/best
21/11/12 14:12:29, acc mean  0.7250 (std  0.1920), f1 mean  0.6644 (std  0.2214), mcc mean  0.6975 (std  0.2220),
cnn & mlp &  0.7250 \tiny $\pm  0.1920$ &  0.6644 \tiny $\pm  0.2214$ &  0.6975 \tiny $\pm  0.2220$
test_data
[0 6 0 ... 0 4 4] {0, 1, 2, 3, 4, 5, 6}
21/11/12 14:12:38, acc mean  0.2938 (std  0.2081), f1 mean  0.2273 (std  0.1849), mcc mean  0.1674 (std  0.2718),
cnn & mlp &  0.2938 \tiny $\pm  0.2081$ &  0.2273 \tiny $\pm  0.1849$ &  0.1674 \tiny $\pm  0.2718$
```

Experiment (2) -cnn

```
21/11/12 15:43:34, Save best model to /content/Meta-NLP/saved-runs/dailydialog_tmp_train_cnn_pretrain/best
21/11/12 15:43:34, acc mean  0.4767 (std  0.1434), f1 mean  0.3757 (std  0.1422), mcc mean  0.4077 (std  0.1813),
cnn & mlp &  0.4767 \tiny $\pm  0.1434$ &  0.3757 \tiny $\pm  0.1422$ &  0.4077 \tiny $\pm  0.1813$
test_data
[6 3 4 ... 6 4 4] {1, 2, 3, 4, 5, 6}
21/11/12 15:43:44, acc mean  0.6208 (std  0.1819), f1 mean  0.5342 (std  0.2059), mcc mean  0.5858 (std  0.2137),
cnn & mlp &  0.6208 \tiny $\pm  0.1819$ &  0.5342 \tiny $\pm  0.2059$ &  0.5858 \tiny $\pm  0.2137$
```

Experiment (3) -cnn

```
dailydialog_u_test
test_data
[6 3 4 ... 6 4 4] {1, 2, 3, 4, 5, 6}
21/11/12 15:45:03, acc mean  0.1710 (std  0.0682), f1 mean  0.1535 (std  0.0660), mcc mean  0.0055 (std  0.0856),
cnn & proto &  0.1710 \tiny $\pm  0.0682$ &  0.1535 \tiny $\pm  0.0660$ &  0.0055 \tiny $\pm  0.0856$
```

Experiment (1) -avg

```
21/11/13 04:18:11, Save best model to /content/Meta-NLP/saved-runs/goemotions_filter_on_dailydialog_test_supervised_avg/best
21/11/13 04:18:11, acc mean  0.6933 (std  0.1662), f1 mean  0.6261 (std  0.1814), mcc mean  0.6626 (std  0.1981),
avg & mlp &  0.6933 \tiny $\pm  0.1662$ &  0.6261 \tiny $\pm  0.1814$ &  0.6626 \tiny $\pm  0.1981$
test_data
[0 6 0 ... 0 4 4] {0, 1, 2, 3, 4, 5, 6}
21/11/13 04:18:21, acc mean  0.3020 (std  0.2017), f1 mean  0.2367 (std  0.1831), mcc mean  0.1770 (std  0.2605),
avg & mlp &  0.3020 \tiny $\pm  0.2017$ &  0.2367 \tiny $\pm  0.1831$ &  0.1770 \tiny $\pm  0.2605$
```

Experiment (2) -avg

```
21/11/13 05:33:15, Save best model to /content/Meta-NLP/saved-runs/dailydialog_tmp_train_avg_pretrain/best
21/11/13 05:33:16, acc mean  0.6017 (std  0.1839), f1 mean  0.5214 (std  0.1950), mcc mean  0.5536 (std  0.2239),
avg & mlp &  0.6017 \tiny $\pm  0.1839$ &  0.5214 \tiny $\pm  0.1950$ &  0.5536 \tiny $\pm  0.2239$
test_data
[6 3 4 ... 6 4 4] {1, 2, 3, 4, 5, 6}
21/11/13 05:33:25, acc mean  0.6165 (std  0.1889), f1 mean  0.5432 (std  0.2026), mcc mean  0.5713 (std  0.2281),
avg & mlp &  0.6165 \tiny $\pm  0.1889$ &  0.5432 \tiny $\pm  0.2026$ &  0.5713 \tiny $\pm  0.2281$
```

Experiment (3) -avg

```
dailydialog_u_test
test_data
[6 3 4 ... 6 4 4] {1, 2, 3, 4, 5, 6}
21/11/13 05:34:51, acc mean  0.1725 (std  0.0685), f1 mean  0.1603 (std  0.0674), mcc mean  0.0073 (std  0.0847),
avg & proto &  0.1725 \tiny $\pm  0.0685$ &  0.1603 \tiny $\pm  0.0674$ &  0.0073 \tiny $\pm  0.0847$
```

Result record

(%)	Exp1	Exp2	Exp3
transfo	29.6 / 23.8 / 16.8	61.1 / 53.3 / 56.6	58.7 / 53.5 / 55.1
cnn	29.4 / 22.7 / 16.7	62.1 / 53.4 / 58.6	17.1 / 15.4 / 0.6
avg	30.2 / 23.7 / 18.0	61.7 / 54.3 / 57.1	17.3 / 16.0 / 0.7

Result Analysis

- simply compare transfo, cnn, avg

  (1) a>(~)t>(~)c / (2) c>(~)a>(~)t / (3) t>(>>)a>(~)c

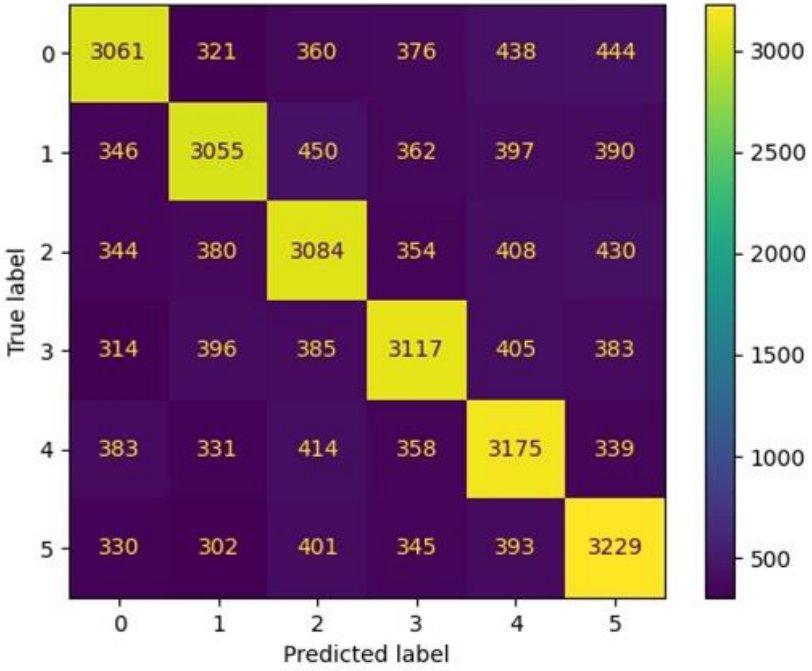
- anticipation

  exp 2 > exp 3, exp 1 result tested on DailyDialog may has low acc

- compare 1 and 3 (transfo_3 is high)
- whether 2 is the best (yes for that tested on DailyDialog)

Experiment (4-2) Meta Learned Model

need to modify /code/train/regular.py, /code/classifier/proto.py



Experiment (4-2) Supervised Model

need to modify /code/train/regular.py, /code/classifier/mlp.py