HW5 Cross-lingual transferability of PLMs

吴冠緯 梁正

Outline

- Finetune on the udpos task (mbert / xlm)
  - base case
  - unfinetuned aligner
  - finetuned aligner
- Same process on XNLI (mbert)
  - base case
  - files being modified & files preparing
  - main problems during implementation
- Summary

Udpos – base case

```
language=en
f1 = 0.9552949641520263
loss = 0.19297418012016712
precision = 0.9554177232540607
recall = 0.9551722365919252
```

```
language=de
f1 = 0.865473288822649
loss = 0.7411949558807014
precision = 0.8712375919308764
recall = 0.8597847602150503
======================
language=fr
f1 = 0.8200726249548471
loss = 0.6608553826808929
precision = 0.8283881623168367
recall = 0.8119223746870706
======================
language=hi
f1 = 0.6201451549627305
loss = 1.4548488032932465
precision = 0.6227307048356916
recall = 0.6175809864886863
======================
language=zh
f1 = 0.5822884740410513
loss = 1.8537330541610717
precision = 0.5996652319464371
recall = 0.5658904233234047
```

Udpos – unfinetuned (de)

```
language=en
f1 = 0.9531022429362074
loss = 0.19101685046120426
precision = 0.9533962943283685
recall = 0.9528083728738074
```

```
language=de
f1 = 0.8715118023009448
loss = 0.721829998192693
precision = 0.8762917855524278
recall = 0.8667836837004322
==========
language=fr
f1 = 0.831570816923972
loss = 0.6623989191356596
precision = 0.8427003208225426 ↑
recall = 0.8207314548158187
==========
language=hi
f1 = 0.6617001768288691
loss = 1.2908983302231019
precision = 0.6655577879516485 ↑
recall = 0.657887025883119
==========
language=zh
f1 = 0.5729734747395048
loss = 1.8247180347442626
precision = 0.5897386901065303
recall = 0.5571351204982399
```

Udpos – unfinetuned (fr)

```
language=en
f1 = 0.9531196434693177
loss = 0.18877537853482237
precision = 0.953757225433526
recall = 0.9524829133763854
```

```
language=de
f1 = 0.8639427020084814
loss = 0.7118741574629429
precision = 0.8700191120117553
recall = 0.8579505813494631
======================
language=fr
f1 = 0.8247220343181624
loss = 0.6603469468325704
precision = 0.8350246951690076 ↑
recall = 0.8146705065221074
======================
language=hi
f1 = 0.666820079555541
loss = 1.2690848991847956
precision = 0.6733725060584935 ↑
recall = 0.6603939443268761
======================
language=zh
f1 = 0.5659947984395318
loss = 1.8386153841018678
precision = 0.5829426958767818
recall = 0.5500045130426934
```

Udpos – unfinetuned (hi)

```
language=en
f1 = 0.9479086344865402
loss = 0.1773738657237711
precision = 0.9482336001645555
recall = 0.9475838914678223
```

```
language=de
f1 = 0.8673524806499459
loss = 0.6432760564738152
precision = 0.8726350131490234
recall = 0.8621335193819986
====================
language=fr
f1 = 0.7831074535531705
loss = 0.7101902735757304
precision = 0.7983845732613627
recall = 0.768404013025392
====================
language=hi
f1 = 0.6474329701872269
loss = 1.134721028403594
precision = 0.6486389333681906
recall = 0.6462314829887678
====================
language=zh
f1 = 0.5569019653448076
loss = 1.6178467216491699
precision = 0.574554878341412
recall = 0.5403014712519181
```

Udpos – unfinetuned (zh)

```
language=en
f1 = 0.9520063751574509
loss = 0.1676143150518028
precision = 0.9524552499828544
recall = 0.9515579232258174
```

```
language=de
f1 = 0.8714259420280657
loss = 0.6148445724201798
precision = 0.8759988898565426
recall = 0.8669004902719067
====================
language=fr
f1 = 0.8319525611813944
loss = 0.5673612356799972
precision = 0.8428891550438512 ↑
recall = 0.8212961394394565
====================
language=hi
f1 = 0.6449645274402676
loss = 1.1898765710110848
precision = 0.6461764705882352 ↑
recall = 0.6437571219273971
====================
language=zh
f1 = 0.5744374186349265
loss = 1.6420903348922729
precision = 0.59233867101352
recall = 0.5575864247675782
```

Udpos – finetuned (de)

```
language=en
f1 = 0.877210429362074
loss = 0.199101685046120426
precision = 0.943962943283685
recall = 0.9428083728738074
```

```
language=de
f1 = 0.8715118023009448
loss = 0.721829998192693
precision = 0.8762917855524278
recall = 0.8667836837004322
====================
language=fr
f1 = 0.821570816923972
loss = 0.6823989191356596
precision = 0.8327003208225426 ↑
recall = 0.8007314548158187
====================
language=hi
f1 = 0.6417001768288691
loss = 1.2908983302231019
precision = 0.6655577879516485 ↑
recall = 0.655887025883119
====================
language=zh
f1 = 0.5729734747395048
loss = 1.837180347442626
precision = 0.5887386901065303
recall = 0.5571351204982399
```

Udpos – finetuned (fr)

```
language=en
f1 = 0.9530166836616653
loss = 0.1695713852052024
precision = 0.953482390864510 8
recall = 0.952551431165316 3
```

```
language=de
f1 = 0.8574882185192313
loss = 0.6672720898492669
precision = 0.8640641327273076
recall = 0.8510116396170008
==========
language=fr
f1 = 0.8271354776106625
loss = 0.5976060783633819
precision = 0.8361800346220427 ↑
recall = 0.8182844881133887
==========
language=hi
f1 = 0.6552870687384778
loss = 1.2394560793271432
precision = 0.6626826913473383 ↑
recall = 0.6480546964024092
==========
language=zh
f1 = 0.5766262659034553
loss = 1.6801170015335083
precision = 0.5926261134663936
recall = 0.5614676414838884
```

Udpos – finetuned (hi)

```
language=en
f1 = 0.9467921974454087
loss = 0.1985513582422673
precision = 0.9470112421168083
recall = 0.9465732540810908
```

```
language=de
f1 = 0.8665333868993083
loss = 0.6983338017845352
precision = 0.8717562861433272
recall = 0.8613726982002318
====================
language=fr
f1 = 0.7747908854917797
loss = 0.766212597325608
precision = 0.791875479040151
recall = 0.7584279180077926
====================
language=hi
f1 = 0.6646075403245575
loss = 1.1898581637785985
precision = 0.6672135450846568
recall = 0.6620218134461989
====================
language=zh
f1 = 0.567325789436996
loss = 1.6751320524215698
precision = 0.5847023324871881
recall = 0.550952252008304
```

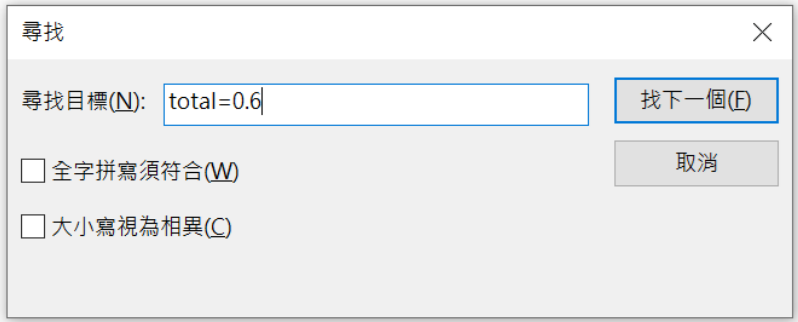
Udpos – finetuned (zh)

```
language=en
f1 = 0.9519827977623384
loss = 0.1866866157457403
precision = 0.9522193658954584
recall = 0.9517463471453776
```

```
language=de
f1 = 0.8714652262842292
loss = 0.6683833907460968
precision = 0.8760847371107708
recall = 0.8668941764031785
======================
language=fr
f1 = 0.7980234477100803
loss = 0.6925392635576018
precision = 0.8090209276236896
recall = 0.7873209479172549
======================
language=hi
f1 = 0.6504762216476286
loss = 1.2559053958035433
precision = 0.6539436017241946 ↑
recall = 0.6470454175484291
======================
language=zh
f1 = 0.5796966811436004
loss = 1.6769686288833618
precision = 0.5971577587444376
recall = 0.5632277281343081
```

XNLI – base case

======= Predict using the model from checkpoint-2200:
ar=0.618562874251497
bg=0.6099800399201597
de=0.5924151696606786
el=0.6225548902195609
en=0.5101796407185629
es=0.5554890219560878
fr=0.5870259481037924
hi=0.6842315369261477
ru=0.6043912175648702
sw=0.6445109780439122
th=0.7718562874251497
tr=0.6229540918163673
ur=0.6736526946107785
vi=0.603193612774451
zh=0.6135728542914172
total=0.6209713905522288



XNLI – files being modified & files preparing

- files being modified

- files preparing

XNLI – files being modified

- in train_xnli.sh -> run **run_classify.py** :

  <1> add the argument **aligned_suffix**

  <2> **aligned_suffix** is considered while obtaining train_examples

- goes to **xnli.py** -> in function **get_examples** :

  <1> need to read in the **aligned_suffix** and implement reordering

  <2> reorder using -> in **reorder_with_ind.py** -> function reorder_2_sent

XNLI – files preparing

- also, while processing on XNLI, there's not **corresponding training-datas**

  <1> download the translation of other languages for training-datas

  <2> repeatedly combined **train-en.tsv** already being located under the path **/xtreme/download/xnli/** with downloaded files in the **XNLI-MT 1.0** including **multinli.train.[fr|de|hi|zh]**

  <3> the way to combine files is to write each corresponding translations of each sentences one by one

XNLI – main problems during implementation

- while running the base case, there is **FileNotFoundError** sometimes (No such file or directory) **->** refresh and try again usually work
- while running XNLI task, there is **RuntimeError: CUDA error: CUBLAS_STATUS_EXECUTION_FAILED ->** general (free) version
- XNLI-MT 1.0 may not be successfully downloaded

- other problems：argument error while running xlm base case, path error while simply running scripts/train.sh, xlm dev_result ...

Summary

- the results using aligner perform a little better than that without using
- the finetuned results perform a little better than those unfinetuned
- finetuned results are especially better while being the target language
- roughly the same, doesn’t change much