

S2W 데이터셋 학습 실험 결과 보고 (MAL + RAN)



20151322 이한용

실험 세팅

- dataset
 - total : 6013 sentences / 143571 tokens
 - train: 4009 sentences, test: 1002 sentences, dev: 1002 sentences
- Flair 제공 tagger 모델을 finetuning 함

S2W 데이터셋 학습 실험 결과 보고 - 2

실험 결과 1

Results:

- F1-score (micro) 0.7648

- F1-score (macro) 0.7834

By class:

LOC tp: 109 - fp: 1 - fn: 1 - precision: 0.9909 - recall: 0.9909 - f1-score: 0.9909

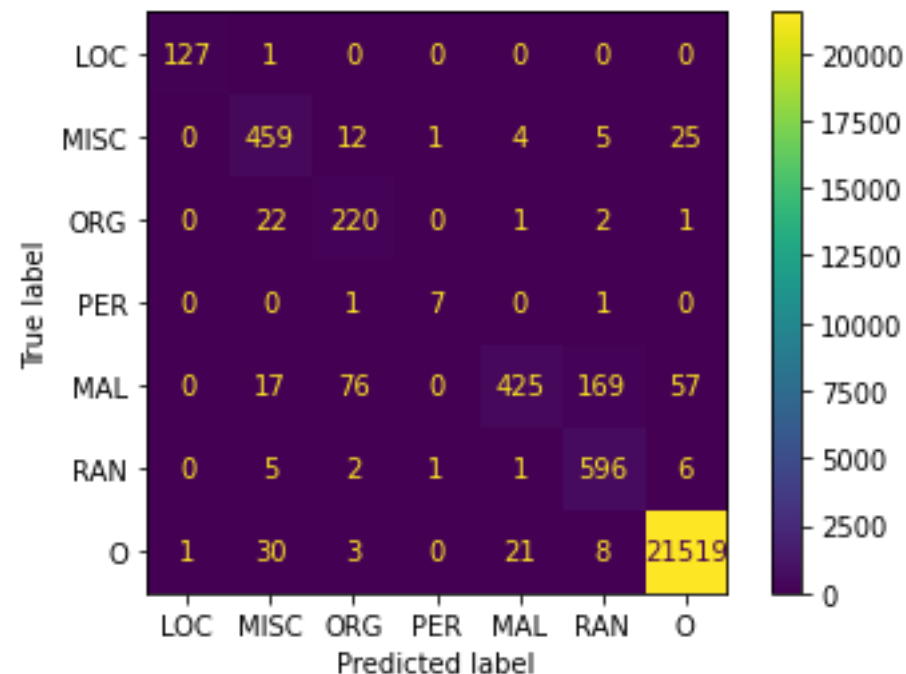
MAL tp: 245 - fp: 71 - fn: 296 - precision: 0.7753 - recall: 0.4529 - f1-score: 0.5718

MISC tp: 369 - fp: 61 - fn: 42 - precision: 0.8581 - recall: 0.8978 - f1-score: 0.8775

ORG tp: 130 - fp: 89 - fn: 21 - precision: 0.5936 - recall: 0.8609 - f1-score: 0.7027

PER tp: 6 - fp: 2 - fn: 2 - precision: 0.7500 - recall: 0.7500 - f1-score: 0.7500

RAN tp: 419 - fp: 182 - fn: 18 - precision: 0.6972 - recall: 0.9588 - f1-score: 0.8073



S2W 데이터셋 학습 실험 결과 보고 - 2

실험 결과 2

Results:

- F1-score (micro) 0.7707

- F1-score (macro) 0.7916

By class:

LOC tp: 56 - fp: 1 - fn: 2 - precision: 0.9825 - recall: 0.9655 - f1-score: 0.9739

MAL tp: 237 - fp: 53 - fn: 175 - precision: 0.8172 - recall: 0.5752 - f1-score: 0.6752

MISC tp: 361 - fp: 76 - fn: 63 - precision: 0.8261 - recall: 0.8514 - f1-score: 0.8386

ORG tp: 107 - fp: 61 - fn: 23 - precision: 0.6369 - recall: 0.8231 - f1-score: 0.7181

PER tp: 17 - fp: 7 - fn: 3 - precision: 0.7083 - recall: 0.8500 - f1-score: 0.7727

RAN tp: 447 - fp: 146 - fn: 119 - precision: 0.7538 - recall: 0.7898 - f1-score: 0.7714

