README.md

[Real-world Sentence Boundary DetectionUsing Multitask Learning: A Case Study on French]

This repository is a submission "Real-world Sentence Boundary DetectionUsing Multitask Learning: A Case Study on French" for NAACL 2021. You can check the data in the "data" folder both for training and testing. We serve around 50 percent of our training data because of license issues. You can test three different models, namely 'Baseline', 'multi-task with XLM-Roberta', and 'multi-task with CamemBERT' models.

1. Dependencies

This is a list of packages that required to run the codes.

- Python 3.7 interpreter
- Pytorch 1.6.0
- transformers from Huggingface

1-1. Install Pytorch

- Windows, Linux

```
conda install pytorch-cpu torchvision-cpu -c pytorch
```

- MacOS

```
conda install pytorch torchvision -c pytorch
```

1-2. Install Required packages

- Windows, Linux

```
pip install -r requirements.txt
```

2. Training and testing

2-1. Baseline with XLM-Roberta

```
python baseline.py
```

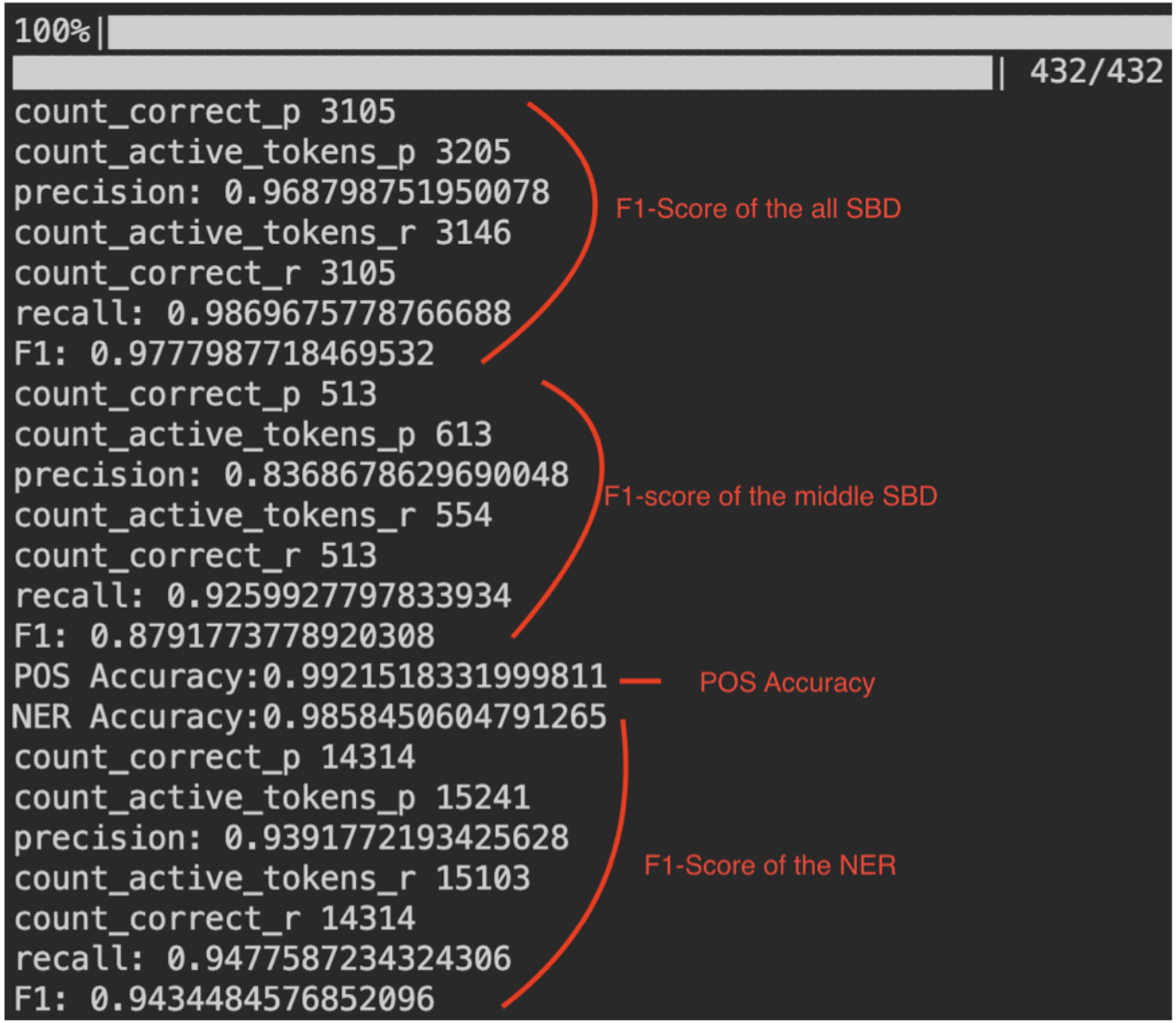
2-2. Multi-task with XLM-Roberta

```
python multi-task-roberta.py
```

2-3. Multi-task with CamemBERT

```
python multi-task-camembert.py
```

3. Performance


100%| | 432/432
count_correct_p 3105
count_active_tokens_p 3205
precision: 0.968798751950078
count_active_tokens_r 3146
count_correct_r 3105
recall: 0.9869675778766688
F1: 0.9777987718469532
F1-Score of the all SBD
count_correct_p 513
count_active_tokens_p 613
precision: 0.8368678629690048
count_active_tokens_r 554
count_correct_r 513
recall: 0.9259927797833934
F1: 0.8791773778920308
F1-score of the middle SBD
POS Accuracy:0.9921518331999811
POS Accuracy
NER Accuracy:0.9858450604791265
count_correct_p 14314
count_active_tokens_p 15241
precision: 0.9391772193425628
count_active_tokens_r 15103
count_correct_r 14314
recall: 0.9477587234324306
F1: 0.9434484576852096
F1-Score of the NER