자연어 처리 – AI 전문가 양성(기본) Day 3

02 Named Entity Recognition & Semantic Role Labeling

Jeong Minsu omicro03@gmail.com

INGLab

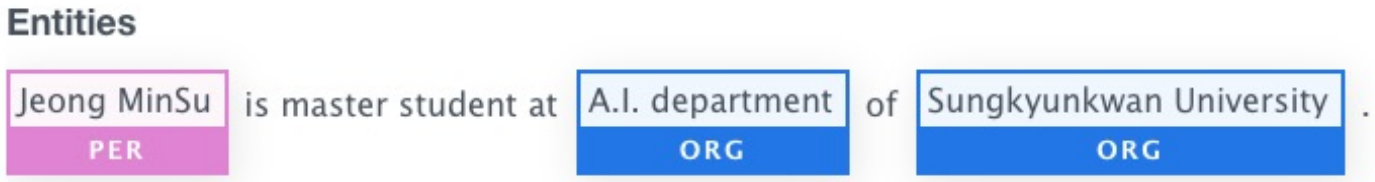
실습 소개

- **실습 목표**
  - Allen NLP 라이브러리를 이용한 NER & SRL Predictor 사용

- **실습 내용**
  - AllenNLP 라이브러리 설치
  - Predictor 불러오기
  - NER & SRL Predictor 적용

Annotate a sentence

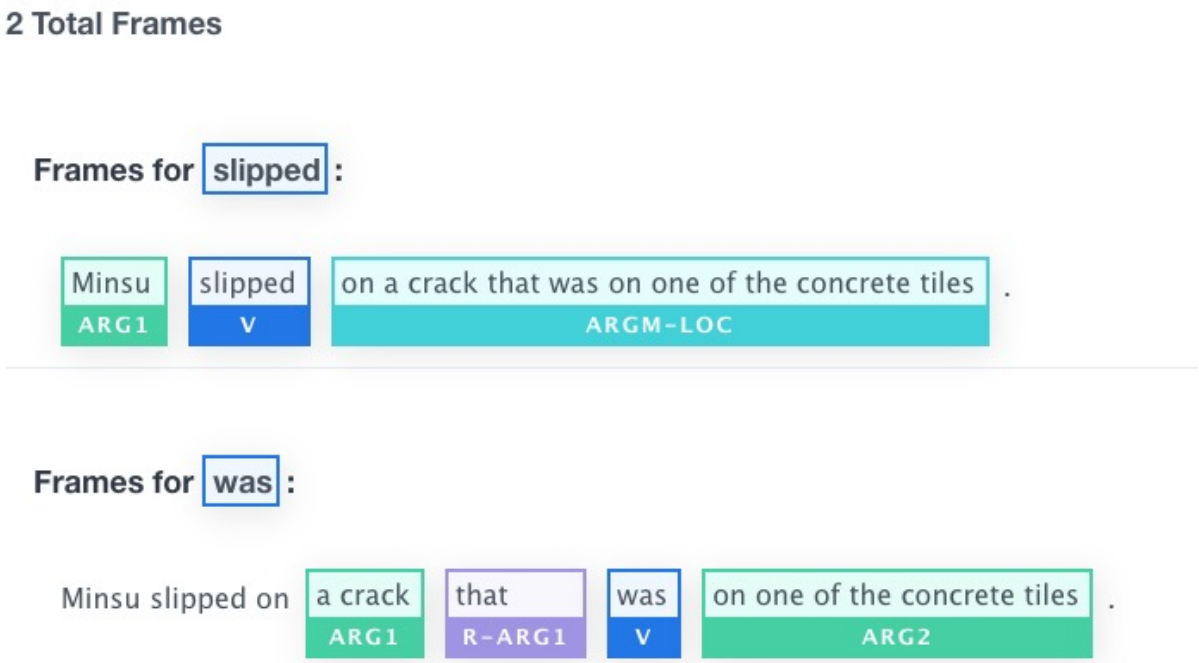
- Named Entity Recognition
  - 개체명인식은 텍스트에서 각 개체를 인식하는 task



https://demo.allennlp.org/named-entity-recognition/named-entity-recognition

Annotate a sentence

- Semantic Role Labeling
  - 의미역 결정은 문장 성분의 의미역을 결정하는 task.



https://demo.allennlp.org/semantic-role-labeling

Annotate a sentence

- AllenNLP Predictor
  - AllenNLP 라이브러리를 활용한 NER & SRL annotation
  - Predictor를 활용한 텍스트 데이터 정보 추가

```
from allennlp.predictors.predictor import Predictor

# NER과 SRL 모델 사용을 위한 predictor 정의
ner_predictor=Predictor.from_path("https://storage.googleapis.com/allennlp-public-models/ner-model-2020.02.10.tar.gz")
srl_predictor = Predictor.from_path("https://storage.googleapis.com/allennlp-public-models/structured-prediction-srl-bert.2020.12.15.tar.gz")

Sentence = ”MinSu is master student at A.I. department of Sungkyunkwan University.”

srl_predictor.predict(sentence=sentence)
ner_predictor.predict(sentence=sentence)
```

Annotate a sentence

- NER Token Replace
  - NER 태그 정보를 이용하여 텍스트 데이터의 고유 명사 처리

```
def ner_per_replace(sentence):
  ner_result = ner_predictor.predict(sentence=sentence)
  tags = ner_result['tags']
  words = ner_result['words']
  for i in range(len(tags)):
    if tags[i] == 'U-PER':
      words[i] = 'U-PER'
  sentence = ' '.join(words)
  return sentence
```

Annotate a sentence

- SRL Argument Parsing
  - SRL 태그 정보를 이용하여 동사별 argument 파싱

```
def srl_arg_replace(sentence):
  events = []
  srl_result = srl_predictor.predict(sentence=sentence)
  words = srl_result['words']
  for frame in srl_result['verbs']:
    events.append(frame['description'])
  return events
```