

단어에 꼬리표 달기

NAVER Conversation
이기창

목차

- 과제 조망하기
- 데이터 살펴보기
- 모델 학습
- 인퍼런스

과제 조망하기

과제 조망하기

- 태스크 소개

개체명 인식 : 토큰 각각에 개체명 태그 붙여주기

과제 조망하기

- 태스크 소개

개체명 인식 : 토큰 각각에 개체명 태그 붙여주기

그 제품 삼성 건가요 → 그 → [0.97, 0.02, 0.01] → 개체명 아님
제품 → [0.96, 0.01, 0.03] → 개체명 아님
삼성 → [0.01, 0.97, 0.02] → 기관명
건가요 → [0.98, 0.01, 0.01] → 개체명 아님

과제 조망하기

- 모델 전체 구조

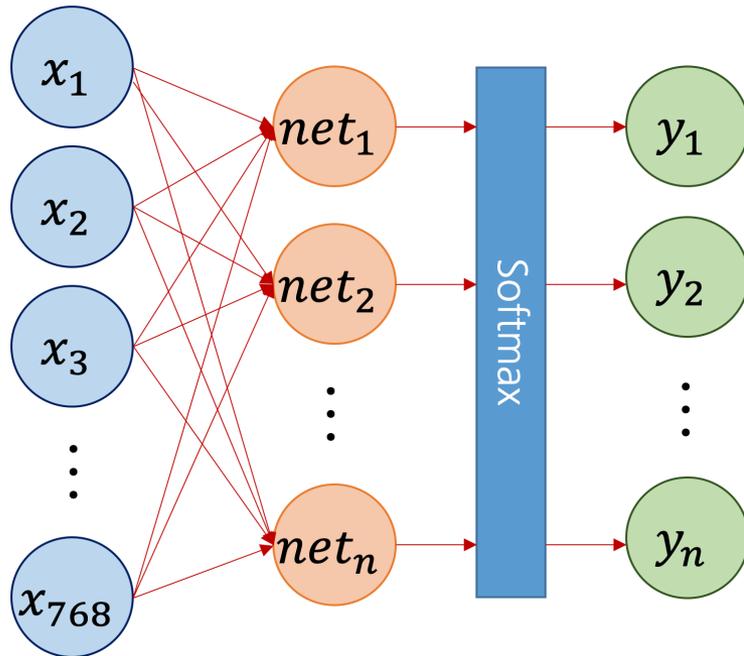
BERT + FeedForward Neural Network



과제 조망하기

- 모델 전체 구조

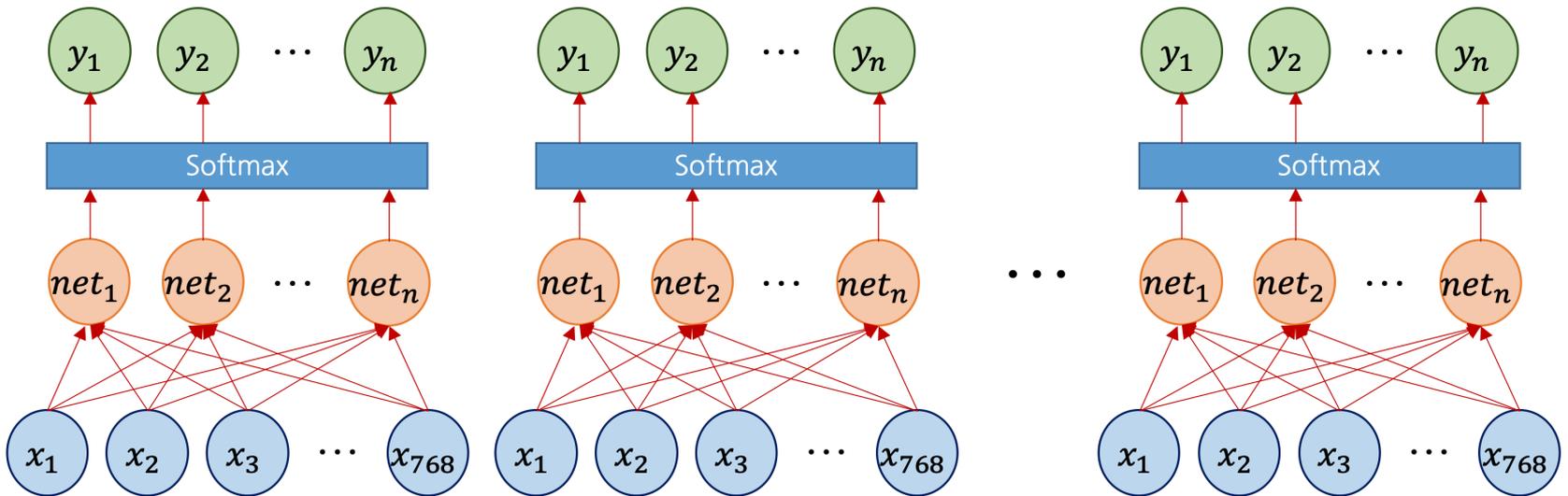
BERT + FeedForward Neural Network



과제 조망하기

- 모델 전체 구조

BERT + FeedForward Neural Network



데이터 살펴보기

데이터 살펴보기

- 데이터

한국해양대학교 자연언어처리 연구실 데이터 + 자체 제작

데이터 살펴보기

- 데이터

한국해양대학교 자연언어처리 연구실 데이터 + 자체 제작

태그	종류
PER	인명
LOC	지명
ORG	기관명
DAT	날짜
TIM	시간
DUR	기간
MNY	통화
PNT	비율
NOH	기타 수량표현
POH	기타

데이터 살펴보기

- 데이터

한국해양대학교 자연언어처리 연구실 데이터 + 자체 제작

—효진 역의 김환희(14)가 특히 인상적이었다.☒—〈효진:PER〉 역의 〈김환희:PER〉(〈14:NOH〉)가 특히 인상적이었다.

3개의 우승 트로피(정규리그 1회, FA컵 2회)를 친정팀에 안겨준 그는 올 시즌을 마치고 미련 없이 자리를 내놨다.☒〈3개:NOH〉의 우승 트로피(정규리그 〈1회:NOH〉, 〈FA컵:POH〉 〈2회:NOH〉)를 친정팀에 안겨준 그는 올 시즌을 마치고 미련 없이 자리를 내놨다.

현대는 최근 보수작업을 끝낸 한남대교의 경우도 안전여부를 재점검하기로 했다.☒〈현대:ORG〉는 최근 보수작업을 끝낸 〈한남대교:LOC〉의 경우도 안전여부를 재점검하기로 했다.

모델 학습

(실습)

모델 학습

- 형상

ratsgo's nlpbook

개체명 인식 (Named Entity Recognition)

주어진 문장 내에서 개체명을 판별합니다.

문장을 입력하세요

입력 문장
그 제품 삼성 건가요?

#	토큰	태그	확률
0	그	O	0.9998
1	제품	O	0.9997
2	삼성	기관명	0.9765
3	건가요	O	0.9985
4	?	O	0.9996

이 웹 데모가 어떻게 동작하는지 자세히 확인하고 싶으신 분은 [ratsgo's nlpbook](#)을 참고하세요. Copyright © 2020 [Gichang LEE](#). Distributed by an [CC BY-NC-SA 3.0 license](#).

모델 학습

- Colab 실습

https://colab.research.google.com/github/ratsgo/nlpbook/blob/master/examples/named_entity_recognition/train_colab.ipynb

인퍼런스
(실습)

인퍼런스

- Colab 실습

https://colab.research.google.com/github/ratsgo/nlpbook/blob/master/examples/named_entity_recognition/deploy_colab.ipynb