

네이버, 창원대가 함께하는
NLP Challenge
18.11.16. Fri. — 12.14. Fri.

네이버, 창원대가 함께하는 NLP Challenge

차정원 - 창원대학교

행사 일정

- 4주간 NSML을 통해 온라인 진행
• 본선: 11월 16일(금)부터 12월 14일(금)까지
• 시상 및 워크샵: 12월 28일(금)
• 장소: 네이버 그린팩토리 카넥트원

참가 신청

<http://naver.me/1IF5AjtC>
접수: 11월 14일(수)부터 상시 가능

참가 대상

머신러닝을 이용한 한국어 NLP
문제 해결에 관심있는 누구나

시상 및 혜택

- 총 1300만원 상당의 상금 및 기념품
• 2개의 미션 수행, 각 미션 당 6팀 내외 수상,
심사 결과에 따라 변동 가능
• 수상팀은 네이버 제공 시 코딩 테스트
면제 등 혜택

상세내용

<https://github.com/naver/nlp-challenge>
아래의 QR코드를 통해 해당 페이지로
바로 들어 갈 수 있습니다



NAVER Clova

NSML

NAVER CLOUD PLATFORM

창원대학교

Adaptive Intelligence Research
적응지능 연구실



창원대학교

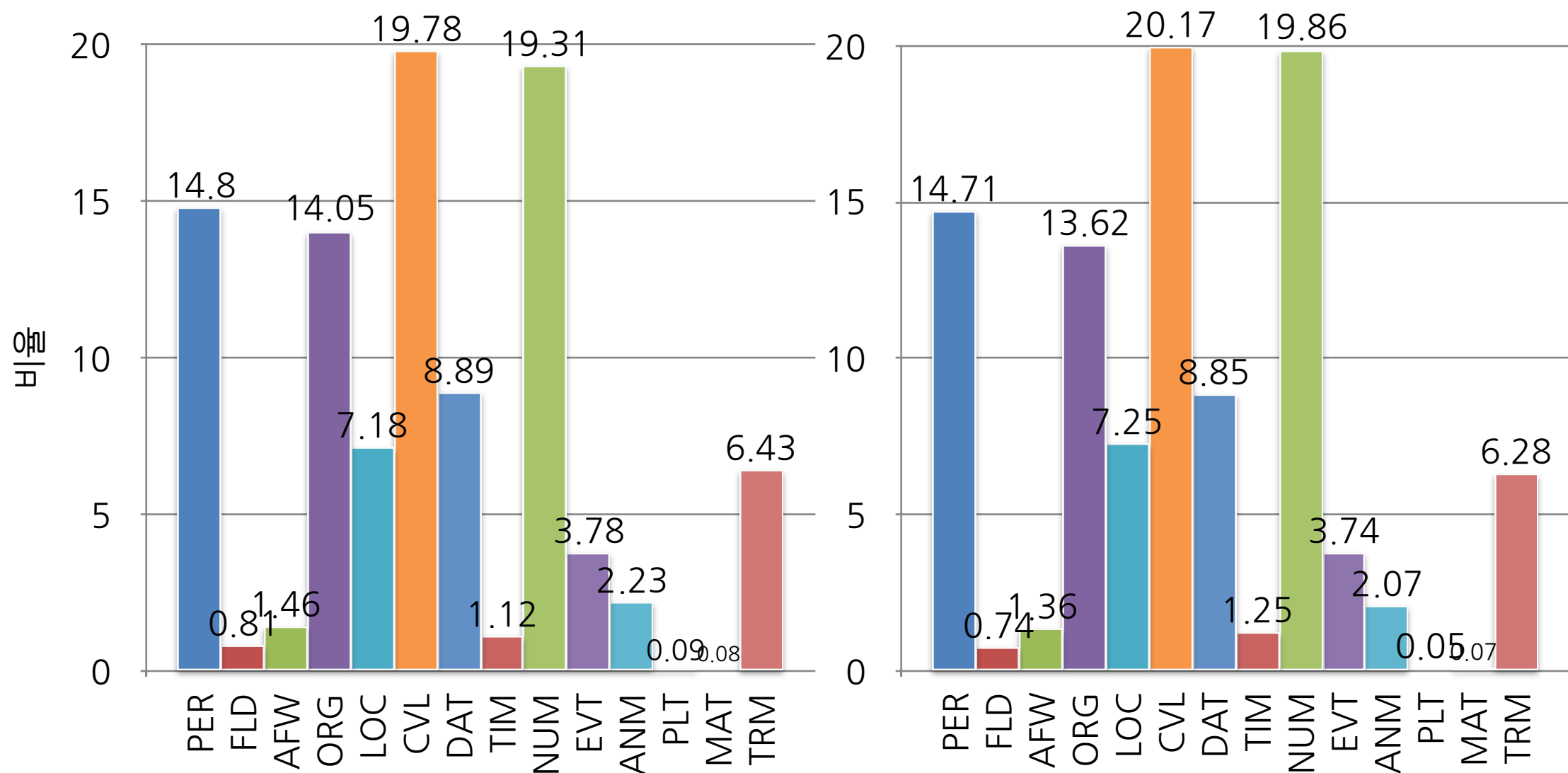


CONTENTS

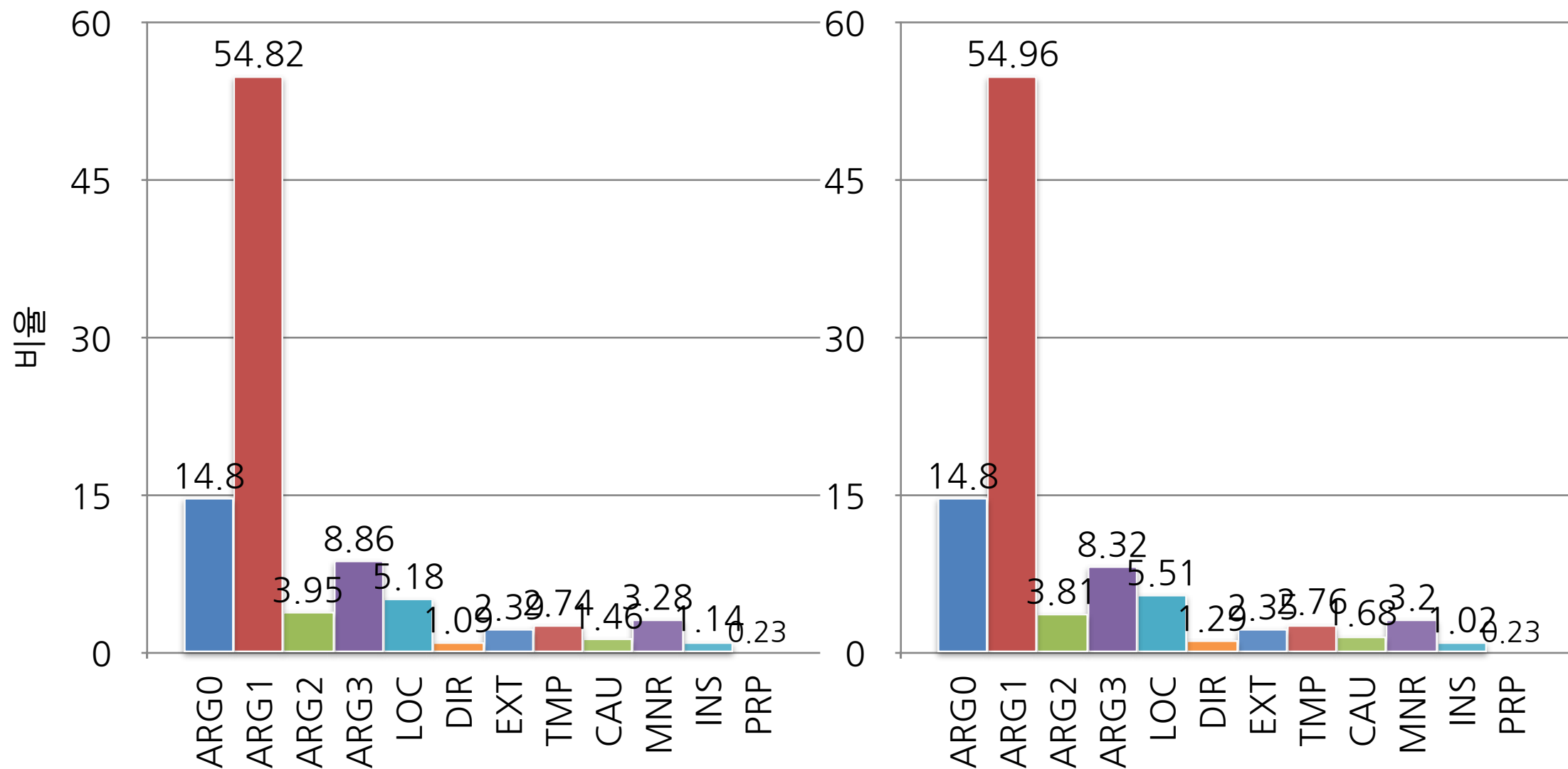
1. Challenge
2. 코퍼스 통계
3. 결과
4. 앞으로의 계획



	# Training data	# Validation data	# Test data
문장수	90,000		10,000
어절수	1,063,571		117,610



	# Training data	# Validation data	# Test data
문장수	34,857		3000
어절수	414,299		35,240





**BRIGHT
FUTURE
AHEAD**



개체명인식 with Naver

I 개체명 인식

‘개체명 인식 리더보드’는 개체명 추출을 위한 데이터 셋을 제공합니다. 개체명(Named Entity)은 인명, 기관명, 지명 등과 같이 문장 또는 문서에서 특정한 의미를 가지고 있는 단어 또는 어구를 말합니다. 이 때문에 개체명은 정보 검색 및 언어 이해를 위한 분석에서 주요한 대상으로 다루어지고 있습니다.

Data.ly에서는 개체명 코퍼스를 제공하여 연구에 도움을 드리고자 하며, 공개적인 리더보드를 통해 많은 분들의 연구 동향을 논의/공유하고자 합니다.

제공되는 코퍼스는 Data.ly에서 제작한 것으로, 연구 및 리더보드를 위한 학습으로 사용 가능하며 상업적인 목적으로 사용될 수 없음을 알려드립니다. 추가문의가 필요한 경우 메일을 남겨주세요.

I Corpus

개체명 추출 리더보드에서 제공되는 코퍼스는 문장에 나타난 개체명을 14개 분류 카테고리 주석 작업이 되어있습니다.

I Leaderboard [CLOSED]

Rank	Date	Model	F1
1	2018-12-14 07:45:43	모델번호: 14 팀명: State_Of_The_Art	90.4219
2	2018-12-14 00:29:34	모델번호: 22 팀명: cheap_learning	90.2417
3	2018-12-14 22:46:11	모델번호: 1 팀명: nlp_pln	89.783
4	2018-12-14 15:18:28	모델번호: 3 팀명: Sogang_Alzzam	88.8506
5	2018-12-14 23:51:09	모델번호: 0 팀명: ner_master	88.5818
6	2018-12-13 19:28:32	모델번호: 370 팀명: bible	88.3348
7	2018-12-14 18:12:42	모델번호: 22 팀명: adullam	88.1113
baseline	2018-11-19 21:05:54	모델번호: 1 팀명: nsm1team	88.0977
8	2018-12-14 18:46:35	모델번호: 0 팀명: nlp___newbie	87.8633
9	2018-12-06 01:35:06	모델번호: 0 팀명: zerocoke	87.4329

의미역결정 with Naver

I 의미역 결정

‘의미역 결정 리더보드’는 의미역 결정(Semantic Role Labeling)을 위한 데이터 셋을 제공합니다. 의미역 결정은 문장내에 포함된 서술어와 해당 서술어의 수식을 받는 논항간의 의미관계를 인식하고, 그 역할을 분류하는 것을 말합니다. 논항은 문장의 구조가 바뀌어도 의미 논항(행위주, 피동작주)은 유지되기 때문에 올바른 의미역 결정을 한다는 것은 문장의 의미를 이해하고, 나아가 문서 또는 대화의 의미를 이해하기 위한 처리에서 주요한 역할을 합니다.

Data.ly에서는 의미역 결정 코퍼스를 제공하여 연구에 도움을 드리고자 하며, 공개적인 리더보드를 통해 많은 분들의 연구 동향을 논의/공유하고자 합니다.

제공되는 코퍼스는 Data.ly에서 제작한 것으로, 연구 및 리더보드를 위한 학습으로 사용 가능하며 상업적인 목적으로 사용될 수 없음을 알려드립니다. 추가문의가 필요한 경우 메일을 남겨주세요.

I Corpus

의미역 결정 리더보드에서 제공되는 코퍼스는 문장에 나타난 논항을 다음과 같이 분류합니다.

I Leaderboard [CLOSED]

Rank	Date	Model	F1
1	2018-12-14 22:01:42	모델번호: 4 팀명: Sogang_Alzzam	77.6628
2	2018-12-14 17:44:24	모델번호: 0 팀명: KANE_team	76.3328
3	2018-12-10 19:53:58	모델번호: 17 팀명: cheap_learning	76.2543
4	2018-12-08 22:06:30	모델번호: model_weights/ 팀명: OnlyOne	75.4308
5	2018-12-14 17:43:00	모델번호: 5 팀명: nlp_pln	74.8749
6	2018-12-14 09:17:23	모델번호: 13 팀명: kozistr_team	74.7695
7	2018-12-14 18:15:26	모델번호: 40 팀명: fortunity	74.4579
8	2018-12-14 10:42:41	모델번호: 16 팀명: DATABF	74.4404
baseline	2018-11-22 20:38:14	모델번호: 16 팀명: nsmlteam	73.0739
9	2018-12-04 20:17:38	모델번호: 9 팀명: Jayten	73.0231
10	2018-12-10 01:16:31	모델번호: 10 팀명: zerocoke	72.9294



Google 애널리틱스 보고서

12월 12일 - 12월 18일 vs 12월 05일 - 12월 11일...

시간 흐름에 따른 사용자

291

+170 (+140.50%)

— 현재 기간 — 비교 기간



Winner!



추가 리더보드

| 분야

복합명사 분리, 품사태깅, 구문분석, 감성분석
(Sentiment Analysis) 분야의 리더보드를
2019년 1월 1일에 오픈할 예정입니다.

| Leaderboard

Rank	Date	Model	F1
baseline	2019-01-01 00:00:00	모델번호: - 팀명: -	00.0

Learning Data Creation and Management Tools
for Deep Learning

인공지능과 딥러닝을 위한 데이터
우리가 만들겠습니다!

감사합니다.

