

## 2021 국립국어원 인공 지능의 언어 능력 평가 대회

· 과제명: 판정 의문문(참/거짓 판단)

· 과제 제작: 에스케이텔레콤

· 자료 설명

· 내용

“판정 의문문(참/거짓 판단)”은 예시문(TEXT)과 그 예시문에 대한 ‘질문이나 명제’(QUESTION)가 주어졌을 때 ‘TRUE’나 ‘FALSE’로 답변하는 과제입니다. 이를 통해 모델의 독해 능력과 일반 언어상식을 테스트합니다. 본 자료는 슈퍼 글루(SuperGLUE)의 Boolean Questions(불리언 질문)를 참조하여 에스케이텔레콤에서 제작하였습니다.

· 데이터 구성

○ 데이터 셋 항목: 데이터는 ID, TEXT, QUESTION, ANSWER로 구성되어 있습니다.

- ID : 일련번호
- TEXT : Wikipedia에서 발췌. 주제별 분포를 고려하여 표제어를 선정한 다음, 표제어의 내용 중 충분한 정보를 담고 있는 단락을 선택.
- QUESTION : 위의 Wikipedia 단락의 내용을 토대로 예/아니오로 답할 수 있는 문장을 생성함, 단락의 내용에서 언급된 내용을 이용하여 문장을 만들어야 함.
- ANSWER: True(예)/False(아니오)로 구성. 질문 문장이 Wikipedia 단락의 내용을 올바르게 이해하고 있으면 True(예), 내용을 잘못 이해하고 있으면 False(아니오)로 태깅함.

TEXT: 태풍 하이선의 직격탄을 맞은 부산은 강풍과 비 피해가 잇따랐다. 부산 곳곳에서 침수나 산사태 사고가 발생했다

Q: 부산에 태풍이 왔나요? [직접적인 언급된 내용]

A: 1 (True)

TEXT: 산호세 샤크스는 2016년 스탠리컵 결승전에 1번 진출했지만, 피츠버그 펭귄스에 패배했다.

Q: 산호세 샤크스가 스탠리컵에서 우승한 적이 있나요? [결승전에 1번 진출했고 패했다는 문장으로 추론 가능한 내용]

A: 0 (False)

· 데이터 셋 생성 주요 조건

- Wikipedia 단락(TEXT)은 작업자들이 직접 문서를 확인하여 선별, 선택합니다. 다양한 분야(인문사회, 과학기술, 문화예술, 생활)의 문서를 균형 있게 선별하고 그 중 정보가 보편적인 단락을 선택합니다.
- QUESTION 문장은 반드시 Wikipedia 단락에서 제공된 내용을 토대로 생성해야 합니다. 직접적으로 언급된 문장을 활용하거나 내용에서 제공된 정보로 충분히 유추 가능해야 합니다.
- QUESTION 문장은 이해하기 쉽고 모호하지 않아야 합니다. TEXT의 내용과 상충되지 않아야 하며, 참/거짓 판단이 명료해야 합니다.

· 자료 규모 및 평가 지표

구분	훈련(Train)	개발(Dev)	시험(Test)	평가 지표
내용	3665셋	700셋	704셋	Accuracy

· 파일 형식: TSV(UTF-8 인코딩)

· 예시

ID	TEXT	QUESTION	ANSWER FALSE = 0, TRUE = 1
1	콜라독립 815(815 Cola)는 코카콜라를 대한민국에서 라이선스로 만들어 오던 범양식품이 1998년 4월 1일에 출시한 탄산음료다.	콜라독립 815는 코카콜라에서 출시하였는가?	0
2	주택임대차보호법은 주거용 건물의 임대차에 관하여 민법에 대한 특례를 규정함으로써 국민의 주거생활의 안정을 보장할 목적으로 제정된 법률이다. 토지의 경우 법률상의 주택이 아니므로 적용대상이 아니다.	토지는 주택임대차보호법에 적용되나요?	0
3	모질라 파이어폭스는 모질라 재단과 모질라 코퍼레이션이 개발하는 자유 소프트웨어 웹 브라우저로 윈도우, 리눅스, OS X 그리고 안드로이드에서 실행할 수 있다. 이 브라우저는 2014년 2월 기준, 전 세계 웹 브라우저 시장 점유율 약 12%~22%를 차지하고 있으며 지역에 따라 두 번째나 세 번째로 가장 많이 쓰이는 웹 브라우저이다.	안드로이드에서 모질라 파이어폭스를 실행할 수 있는가?	1
4	폴 고갱(Paul Gauguin, 1848년 6월 7일 ~ 1903년 5월 8일)은 프랑스의 탈인상주의 화가이다. 파리에서 태어났다. 생전에는 그리 평가받지 못하였으나 오늘날에는 인상주의를 벗어나 종합주의 색채론에 입각한 작품을 남긴 화가로 널리 인정받고 있다. 생의 마지막 10여년을 타히티를 비롯한 프랑스령 폴리네시아에서 생활하며 작업하였고 이 시기 작품이 그의 대표작이 되었다.	폴 고갱은 인상주의를 추구했던 화가이다.	0
5	고위공직자범죄수사처는 고위 공직자 및 대통령 친인척의 범죄행위를 상시적으로 수사·기소할 수 있는 독립된 국가기관이다.	고위공직자범죄수사처는 대통령 친인척을 수사할 수 있나?	1
6	숙종이 다시 환국을 일으켜 서인을 기용하고 남인을 실각하는 과정에서 폐위된 인현왕후가 복위하고 희빈 장씨가 왕비의 자리에서 내려와 빈(嬪)으로 강등되었다. 이때 희빈의 처우를 두고 서인 내의 노론과 소론이 갈등하였다. 희빈 장씨의 폐위를 반대하는 소론은 이후 왕세자(경종)를 지지하였고, 인현왕후의 복위를 지지하는 노론은 경종의 이복동생인 숙빈 최씨 소생의 연잉군(영조)을 지지하게 된다.	노론은 희빈 장씨의 폐위를 반대했다.	0
7	흑마늘은 단독으로 먹거나, 수프나 소스에 사용하거나, 마요네즈에 넣어 으깨거나, 야채 요리에 간단히 넣기도 한다. 비네그레트는 흑마늘, 세리 식초, 콩, 중성 오일, 그리고 디종 머스타드로 만들 수 있다. 마늘 속 수분 함유량에 따라 부드러움이 증가한다	흑마늘을 단독으로 먹기도 하나요?	1
8	캐나다로 입국하는 미국인은 반드시 차량 등록 문서와 보험 증빙 서류를 소지해야 한다.	미국인이 캐나다로 입국할 수 있나?	1