

네이버 Ai Burning Day 해커톤

1. 팀원 및 역할

- 가. 정의철: 풀스택, 머신러닝, 한국어 임베딩
- 나. 황동원: 프론트 개발
- 다. 박정규: 기획

2. 아이디어 구상

외국인들을 위한 생활(여행) 한국어 교육 웹 개발.

가. 구성

- 1) 먼저, 예문이 주어진다.
- 2) 외국인(이하 사용자)는 이를 한국어로 번역해야함. (음성 또는 텍스트 둘 다 가능)
- 3) 사용자가 작성한 문장과 답(데이터베이스에 저장되어 있는 답안)과 비교. 이를 스코어로 변환한다.
- 4) 특정 이상 스코어가 나와야만 Pass. 그 이하는 Fail로 판별한다. P/F결과와는 무관하게 정답을 보여준다. 정답의 문법요소와 사용자가 작성한 답의 문법요소를 각각 보여주며 비교할 수 있게 한다.

나. 구성 예시

- 1) Please translate this sentence to Korean: **I played tennis with tommy.**
- 2) 외국인 발화: 나는 토미랑 했어 테니스.
- 3) 외국인 발화를 스코어로 분석. 예) score: 70
- 4) Fail 한 상태.
- 5)

I	Played	tennis	with	tommy
나는	했다.	테니스	와	토미
Subject	Verb	---	---	---

표 1 분석 결과

다. 스코어 분석 방법

영어를 형태소 별로 끊어서 I -> 나는, Played -> 놀았다. 처럼 번역을 한다. 번역한 결과를 한국어 형태소 별로 분석을 해서 유사도가 가장 높은 것 끼리 매칭한다.

예시) [I, 나는], [Played, 했다.], [Tennis, 테니스]

또한 매칭된 단어 끼리의 코사인 거리를 구한다.

예시) $\sum \cos(i, j)$ 를 통해 점수 합을 구함.

해결해야 할 점.

- “나는 했었다. 토미랑 테니스를”과 같이 문장 순서가 맞지 않는 것은 어떻게 판별할 것인가?
- 어떻게 스코어를 보다 더 정확히 매길 수 있는가?