

유사도 기반 컴퓨터공학 기술 면접 시스템의 설계

이동현^o, 김동현^{*}

^o동서대학교 소프트웨어학과,

^{*}동서대학교 소프트웨어학과

e-mail: lee588779@gmail.com^o, pusrover@dongseo.ac.kr^{*}

The Design of Technical Interview System for Computer Engineering based Similarity

Dong Hyun Lee^o, Dong Hyun Kim^{*}

^oDepartment of Software, Dongseo University,

^{*}Department of Software, Dongseo University

● 요약 ●

컴퓨터공학 분야 개발자를 채용할 때 대다수의 기업에서 일반 면접과는 달리 전공 분야 역량 파악을 위한 컴퓨터공학 기술 면접을 함께 진행한다. 컴퓨터공학 면접자의 기술 면접을 지원하기 위하여 이 논문에서는 컴퓨터공학 핵심 개념에 대한 면접자 답변의 정확도를 코사인 유사도를 이용하여 평가 후 결과를 알려주는 시스템을 제안한다. 제안한 시스템을 이용하면 개발자들의 컴퓨터공학 핵심 개념의 기술 면접 정확도를 향상시킬 수 있을 것으로 기대된다.

키워드: 기술면접(technical interview), 면접지원시스템(interview support system), 코사인 유사도(cosine similarity), CS 기술면접(computer science technical interview), 임베딩벡터(embedded vector)

I. 서론

컴퓨터공학 분야 개발자를 채용할 때 대다수의 기업들은 서류 심사 이후에 면접 과정을 진행하면서 전공 분야 역량을 파악하길 원한다. 특히, 개발자는 채용 후 즉시 개발 현장에서 소프트웨어를 개발하는 경우가 많기 때문에 코딩 테스트 외에 추가로 컴퓨터공학(Computer Science, CS) 기술 면접을 함께 진행하는 추세이기 때문에 CS 기술 지식에 대한 중요성이 높아지고 있다. 또한 실제 면접에서 발음을 불분명하게 이야기할 경우 면접에서 명확한 의미 전달이 실패할 수 있다[1].

CS 기술의 경우 핵심 개념이 누락되면 정답이 완전히 달라지는 경우가 있기 때문에 핵심 개념을 숙지하고 면접을 진행해야 한다. 만약 충분한 연습 과정을 거치지 않고 CS 기술 면접을 진행할 경우 답변을 논리적으로 전개하기 어렵기 때문에 반복 연습이 필요하다. 면접에서의 발음 또한 명확한 의미 전달을 위해 반드시 필요한 요소 중 하나이기 때문에 반복적인 답변 연습을 통해 발음을 개선해야 할 필요가 있다.

관련 연구로는 언어적 요소와 비언어적 요소를 분석하여 비대면 AI 면접을 대비하는 시스템[2]이 제시되었다. 텍스트 변환을 위하여 사용되는 4가지 모델에 대하여 성능 평가 후 비교 결과가 [3]에서

제시되었다.

이 논문에서는 면접자의 CS 기술 면접의 정확도를 측정 후 평가할 수 있는 면접시스템을 제안한다. 이를 위하여 사용자의 답변과 정답과의 유사도를 평가하여 CS 면접의 정확도를 향상한다. 유사도 평가 시 먼저 한국어 문장 임베딩 측정을 위하여 KorNLU, KorSTS 데이터 셋[4] 기반 학습 모델인 ko-sentence-transformers[5]을 사용한다. 그리고 면접자 답변 문장 토큰의 임베딩 벡터를 계산하고 문장 간의 코사인 유사도(cosine-similarity)를 평가한다.

이 논문의 구성은 다음과 같다. 2장에서는 관련 연구를 기술하고 3장에서 유사도 기반 면접시스템의 설계를 제안한다. 마지막으로 4장에서 결론을 기술한다.

II. 관련 연구

관련 연구로는 언어적 요소(문법적 습관, 답변 키워드, 발화 속도, 발화 높낮이)와 비언어적 요소(시선 안정성)를 분석하여 결과를 바탕으로 피드백을 제공하는 시스템이 제시되었다.[2] 한국어를 사용하는 외국인을 표기할 때 구글 텍스트 변환(Speech-To-Text, STT), 네이버

클로바 음성 인식, IBM 왓슨, 마이크로소프트 애저 STT의 4가지 STT 엔진의 성능 비교 결과가 [3]에서 제시되었다. 이 연구에서는 원본 문장과 텍스트 간의 유사성 비율을 확인한 뒤 평균 결과를 찾는 문장 변환용 파이썬(Sentence transformer Python) 프레임워크를 사용하여 비교하였다.

III. 유사도 기반 면접 시스템 설계

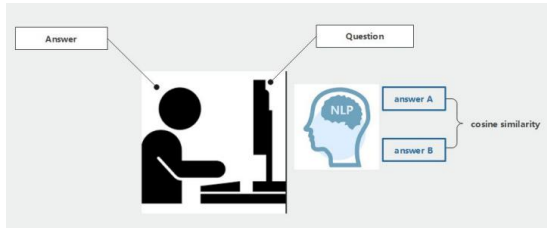


Fig. 1. CS 기술 면접 시스템 개념도

그림 1은 CS 기술 면접 시스템의 개념도를 보여준다. 사용자가 음성 데이터를 입력하면 음성을 텍스트로 먼저 변환한다. 그리고 사용자의 답변과 문제 정답과의 정확도를 측정하여 측정 결과를 사용자에게 보여준다. 사용자는 측정 결과를 이용하여 정답 대비 미비한 점을 확인할 수 있다. 또한, 사용자의 음성 데이터가 저장되어 있는 wav 파일을 ETRI Open API에서 제공하는 발음 평가 API[6]를 사용하여 사용자의 발음 점수를 출력한다.

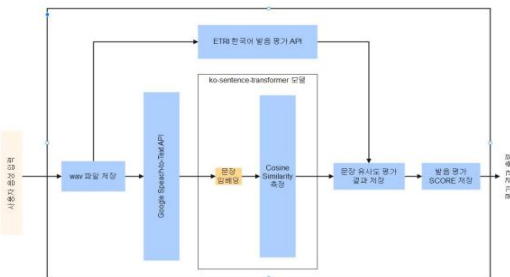


Fig. 2. 정확도 측정 절차

그림 2는 면접자 답변의 정확도를 측정하기 위한 절차를 보여준다. 먼저, 면접자 음성 답변은 구글 STT 모델을 사용하여 문장으로 변환한다. 변환된 문장과 저장되어 있는 정답 문장을 koBERT 모델이 제공하는 토큰라이저를 사용해 토큰화를 진행한다. 그리고 모델에 입력하여 얻은 임베딩 벡터의 평균을 구하여 문장 벡터를 얻는다. 문장 벡터를 사용하여 변환된 문장과 정답 문장 간의 코사인 유사도를 계산한다. 자카인 및 유클리디안 유사도 평가에 비하여 문장의 길이나 형태에 영향을 받지 않기 때문에 이 논문에서는 코사인 유사도 기법을 활용한다.

IV. 결론

일반 면접에 비하여 CS 기술면접은 전공 분야 지식이 필요하기 때문에 기술 내용에 대한 정확한 답변이 필요하다. 이를 위하여 이 논문에서는 면접자의 CS 기술 면접 답변의 정확도를 코사인 유사도를 이용하여 평가 후 면접자에게 알려주는 면접 시스템을 제안하였다. 향후 연구는 koBERT 모델 기반으로 문장 간 코사인 유사도를 평가하는 시스템을 구현하는 것이다.

ACKNOWLEDGEMENT

본 연구는 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음(2019-0-01817)

REFERENCES

- [1] 잡코리아, “IT 취업 가이드 개발자 면접 준비 방법“, https://www.jobkorea.co.kr/GoodJob/Tip/View?News_No=18307&cmpid=vm_viral_post&utm_source=&utm_medium=&utm_campaign=
- [2] 정윤정, 함종현, 이희준, 김지미, 진재형, 양준, 이성현, 이민구, “리허설: 비대면 AI 면접 대비 시스템 설계”, 한국HCI학회 학술대회, pp.556-558, 2021.
- [3] A. B. Wahyutama and M. Hwang, "Performance Comparison of Open Speech-To-Text Engines using Sentence Transformer Similarity Check with the Korean Language by Foreigners," 2022 IEEE International Conference on Industry 4.0, Artificial Intelligence and Communications Technology (IAICT), pp. 97-101, 2022.
- [4] kakaobrain, kor-nlu-datasets, <https://github.com/kakaobrain/kor-nlu-datasets>
- [5] Jhgan00, ko-sentence-transformers, <https://github.com/jhgan00/ko-sentence-transformers>
- [6] AI API·DATA, “발음평가 기술”, <https://aiopen.etri.re.kr/guide/pronunciation>