

# AI와 STT를 기반한 교육플랫폼 구현

이하리, 이준서, 정재현, 이영주  
한국폴리텍대학 서울정수캠퍼스 인공지능소프트웨어과  
e-mail: summerly.in.love@gmail.com

## Implement of the educational platform based on AI and STT

Hari Lee, Junseo Lee, Jaehyeon Jung, Youngju Lee  
Dept. of AI-Software, Seoul Jungsu Campus of Korea Polytechnics

### 요약

본 논문에서는 학습 보조 프로그램의 초기 구축단계로 딥러닝 기술과 STT(Speech To Text) 음성처리 기술을 이용하여 음성 형태의 수업 내용을 텍스트로 변환하여 서버에 전송 후 저장하고 변환된 데이터를 전송받은 서버에서는 웹과 앱에서 접근할 수 있도록 설계하였다. 각 강의실에 음성처리시스템을 설치하여 교육환경에 편리성을 제공하며 축적된 학생들의 데이터를 기반으로 학습 성취도를 측정하여 학생들이 어떠한 영향을 받는지 효과를 밝히고자 한다. STT 음성 변환 정확도 WER(Word Error Rate)는 82%로 측정되었으며 원인분석 및 향후 개선 방안을 제안하였다.

### 1. 서론

4차 산업혁명 시대를 맞이하여 급변하는 환경 속에서 COVID-19로 인한 생활 환경의 변화는 교육계에도 큰 영향을 미쳤다. 그로 인해 전 세계적으로 많은 나라들이 AI를 활용한 교육 시스템을 구축하고 있다.

교육부는 학교와 가정에서 인공지능 보조교사와 함께하는 수업은 AI 알고리즘을 통해 학생들의 수준을 진단 및 분석하는 맞춤형 학습활동을 제공함으로써, 수학뿐만 아니라 일상생활 속에서 인공지능 기술을 좀 더 친숙하게 활용할 수 있는 디지털 인재로의 성장 목표를 제시하였다[1].

최근 AI를 활용한 교육시장의 규모를 살펴보면 2021년 기준 북미는 20억 달러를 이미 초과했으며, 아시아는 약 25억 달러로 예상된다. 이는 2022년부터 향후 10년 후인 2030년까지 CAGR 45% 이상으로 성장할 것으로 전망된다[2].

현재 이러한 AI 기술에 대한 투자는 계속해서 증가하는 추세이며, 디지털 네이티브 리더와 Google, Amazon, Apple 등의 핵심 글로벌 대기업들이 선도해 나가고 있다. 그리고 캐나다에서는 교육 솔루션 AI 소프트웨어 플랫폼이 2021년에 약 70%의 시장점유율을 차지하고 있고, 기술 선진국 독일에서는 딥러닝 기술이 2030년에 약 50%에 도달할 것으로 예

측한다. 이를 통해 교사, 학생 및 사회의 다양한 조직들에게 더 넓은 협력학습 방법론을 충분히 제공할 수 있게 되었다. 영국에서는 고등 교육기관들을 중심으로 코스 추천, 행정 및 진로 옵션 지원을 위한 AI로의 전환이 빠르게 이루어지고 있고, AI 교육분야 활용의 세계화 추세에 따라 폐쇄적 국가인 중국에서도 ITS(Intelligent Tutoring System)를 앞선 다른 주요 나라와 마찬가지로 2030년까지 55%까지 늘어날 것으로 예상하고 있다[2].

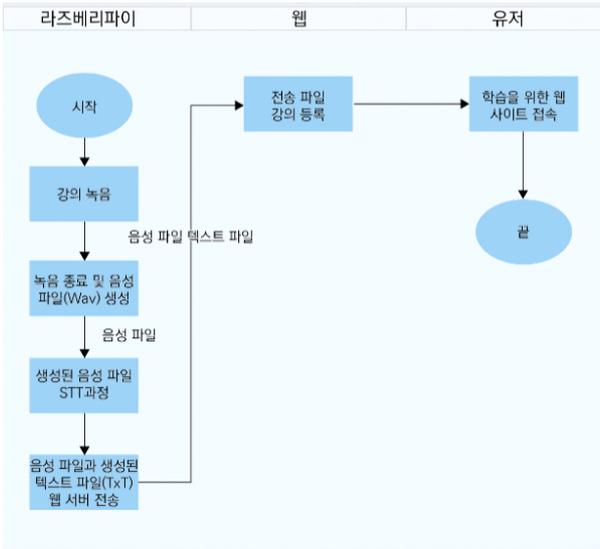
따라서, 본문에서는 학교 교육에서의 AI 및 에듀테크 활용 방안에 대한 구체적인 연구를 위해 해외 각국이 가진 이미 안정화되고 검증된 제품들과 기술을 접목시켜 K-Edu(한국형) AI플랫폼 구축방안을 제시하였다. 구축방안의 초기 단계로 AI와 STT를 이용한 학습보조 프로그램 활용방법과 개선방안을 제시하였다.

### 2. 본론

본 연구의 각 주요 기능을 통한 학습의 성취도 향상을 측정하고 프로그램이 학생에게 주는 영향을 구체적으로 확인하는 것이 목적이다. 프로그램의 주요 기능으로는 음성 텍스트 변환 기능과 학습방법 추천 및 문제 출제기능이 있으며, 학습방법 추천 및 문제

출제기능은 추후 추가할 예정이다.

프로그램의 핵심 기술인 음성 텍스트 변환 기술은 구글의 STT API(Application Programming Interface)를 사용하여 구축하였다. J.Y. Chan은 다수의 STT 성능을 비교 평가하여 구글의 STT가 변환 기능 API 대비 86%로 매우 높은 정확도를 제시하였다[3].



[그림 1] 프로그램 순서도

전체 시스템 구축 환경은 STT기능을 포함한 음성 처리부, 웹 서버, 웹, 앱으로 이루어진다. 음성검출을 위해 소형 싱글보드컴퓨터인 라즈베리파이를 이용하였다.

그림 1은 프로그램 순서도를 보여주고 있다. 학습 보조 프로그램 처리 절차는 수업이 시작됨과 동시에 수업 음성을 녹음하고 이를 자연어 처리하여 교수자가 말한 내용을 텍스트로 변환하여 서버로 전송한다. 수업을 마친 학생은 웹과 앱의 두 가지 접근 방식으로 클라이언트 서버에 접속하여 수업을 찾아볼 수 있다. 또한 수업에서 많이 언급된 키워드를 추출하여 보여줌과 동시에 학습자의 학습 패턴에 맞는 최적 학습경로를 제공하여 자신만의 학습을 할 수 있도록 보조하는 역할을 수행시키려고 한다.

STT기능을 구현한 결과 위와 같이 음성을 텍스트로 변환할 수 있었고, 정확도는 J.Y. Chan에서 사용된 WER(Word Error Rate) 방법을 사용하였다[3, 5]. 그림 2는 샘플 데이터를 활용하여 시험 결과를 보여주고 있다. WER은 82%의 정확도로 계산된다.

박	범	복	군	은	범	범	꽃	놀이	를	가	고	방	범	복
양	은	낫	범	꽃	놀이	를	간	다	.					
박	범	복	구	는	범	범	꽃	놀이	를	가	고	방		
범	보	경	은	범	범	꽃	놀이	를	간	다	.			

[그림2] STT 구현 결과

이 결과는 J.Y. Chan 이 측정 한 86%의 정확도보다 낮은 결과를 확인할 수 있었다. 정확도 개선을 위해서 다양한 음성 파일의 음성 명료성을 향상시키는 방법이 있다[2][4]. 강의실 내 소음 유입에 따른 오차를 줄이기 위한 마이크 선정, 다양한 측정 환경을 변화하면서 불필요한 잡음 제거함으로 음성 명료성을 향상하여 정확도의 향상을 위한 추가 연구가 필요할것으로 사료된다.

### 3. 결론

본 논문에서는 학습 보조 프로그램의 초기 구축단계로 딥러닝 기술과 STT(Speech To Text) 음성처리 기술을 이용하여 음성 형태의 수업 내용을 텍스트로 변환하여 서버에 전송 후 저장하고 변환된 데이터를 전송받은 서버에서는 웹과 앱에서 접근할 수 있도록 설계하였다.

각 강의실에 음성처리시스템을 설치하여 교육환경에 편리성을 제공하며 축적된 학생들의 데이터를 기반으로 학습 성취도를 측정하여 학생들이 어떠한 영향을 받는지 효과를 밝히고자 하였다. 제안한 시스템을 활용하여 STT 음성 변환 정확도 WER(Word Error Rate)는 82%로 측정되었으며 정확도 향상을 위한 원인분석 및 향후 개선 방안을 제안하였다.

최근 음성 파일의 음성 명료성을 향상시키는 방법이 다양하게 연구되고 있어 향후 관련 기술 연구결과를 적용하여 명료성을 향상시키고 정확도와의 관계를 도출하는 연구를 진행할 예정이다.

### 참고문헌

- [1] 남궁양숙, “초등 3\_4학년도 인공지능 보조교사와 함께 수학 공부해요!”,교육부보도자료, 9월, 2022년.
- [2] “Artificial Intelligence (AI) in Education Market size 2022 - 2030”, Global Market Insights, 6월, 2022년.
- [3] J.Y. Chan, H.H. Wang, “Speech Recorder and Translator using Google Cloud Speech-to-Text and Translation”, Journal of IT in Asia, 11월, 2021년.
- [4] 임지원, 황용해, 김규현, “STT 효율 증대를 위한 음성 주파수 correlation 기반 노이즈 필터링 방안”, 한국방송미디어공학회:학술대회논문집, pp.176-179, 11월, 2021년.
- [5] 민소연, 이광형, 이동선, 류동엽, “한국어 특성 기반의 STT 엔진 정확도를 위한 정량적 평가방법 연구”, 넥스트지, 2020년.