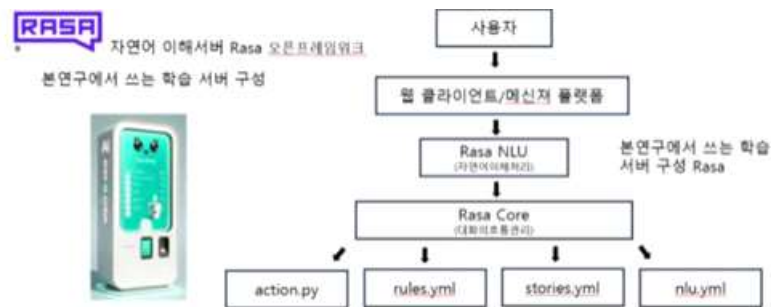


기업협업과제(캡스톤디자인 프로젝트) 최종결과보고서

사업명	[2-4-5] 경북 첨단산업 인재양성 사업																	
분야	<input type="checkbox"/> 바이오	<input type="checkbox"/> 미래형자동차	<input type="checkbox"/> 로봇	<input checked="" type="checkbox"/> ICT														
과제명	시설 안내용 생성형 AI 기반 음성 인식 키오스크 개발																	
과제기간	2024. 09. 01. ~ 2024. 10. 31.		사용지원금	16,810,000원														
연번	소속	성명	참여역할	비고														
책임교수	대구가톨릭대학교	김기성	기업협업과제 운영 총괄															
참여기업	(주)드림아이디어소프트	류지수	기업 역할 총괄/멘토															
	(주)드림아이디어소프트	이설기	실무 담당															
교육생	경북ICT첨단인재양성사업단	권나영	데이터셋 구축 및 전처리															
	경북ICT첨단인재양성사업단	백상조	모델 학습 및 NLP 시스템 개발															
	경북ICT첨단인재양성사업단	이보미	시스템 통합 및 테스트															
	경북ICT첨단인재양성사업단	정은경	프로젝트 기획 및 관리, 시스템 성능 최적화															
활동 실적																		
수행내용	<p>이 프로젝트는 세대 간 디지털 격차와 정보 접근성 문제를 해결하고자, ChatGPT와 유사한 고급 자연어 처리(NLP) 기술을 기반으로 사용자의 질문에 자연스럽게 응답할 수 있는 인공지능 시스템을 개발하여 누구나 쉽게 사용할 수 있는 대화형 키오스크를 제공하는 것을 목표로 한다. 이를 위해 자연어 처리 모델의 파인튜닝(Fine-tuning)과 강화학습을 통해 출입 말, 은어, 사투리 등 다양한 표현을 이해하고 이에 적절히 반응할 수 있도록 하여 사용자 요구를 효과적으로 지원한다. 데이터셋 구축 및 전처리, 모델 학습과 NLP 시스템 개발, 시스템 통합 및 테스트 등을 수행하며, 프로젝트 기획과 관리, 시스템 성능 최적화 작업을 통해 사용 편의성과 정확도를 높이는 것을 목표로 했다. 이러한 과정을 거쳐 완성된 AI 시스템은 디지털 접근성을 높이고 사용자들에게 자연스럽게 편리한 대화 경험을 제공한다.</p> <div><div>공공기관 및 교육기관 안내용 생성형 AI 기반 음성인식 키오스크 개발</div><div><div>“생성형 AI 기반 음성인식 키오스크” 과제 내용</div><table><tr><td>음성인식 알고리즘 최적화</td><td>사용자 인터페이스 개선</td><td>실용 분석을 통한 피드백 반영</td></tr><tr><td>사용자를 위한 맞춤형 음성인식 및 처리 알고리즘 개발</td><td>사용자 친화적인 최적화 인터페이스 구현</td><td>실용 시범(시범 운영 :기관 1개소) 확보 및 서비스 수행을 통한 활용성 평가</td></tr><tr><td>시설 접근성 및 사용 편의성 향상</td><td>자연어 처리 기능 통합을 통한 사용 경험 개선</td><td>서비스 확장을 위한 개선 인사이트 도출의 다양성을 내포한 유연한 시스템 구현</td></tr></table><div><div>시설의 시스템을 통합한 정보제공 솔루션 구축</div><div>생성형 AI 알고리즘 개발</div><div>실증서비스 운영 환경 구축과 안내 키오스크 활용의 혁신</div></div></div></div> <tr><td colspan="5">(그림 1) AI 기반 음성인식 안내 키오스크 목표 개념도</td></tr>				음성인식 알고리즘 최적화	사용자 인터페이스 개선	실용 분석을 통한 피드백 반영	사용자를 위한 맞춤형 음성인식 및 처리 알고리즘 개발	사용자 친화적인 최적화 인터페이스 구현	실용 시범(시범 운영 :기관 1개소) 확보 및 서비스 수행을 통한 활용성 평가	시설 접근성 및 사용 편의성 향상	자연어 처리 기능 통합을 통한 사용 경험 개선	서비스 확장을 위한 개선 인사이트 도출의 다양성을 내포한 유연한 시스템 구현	(그림 1) AI 기반 음성인식 안내 키오스크 목표 개념도				
음성인식 알고리즘 최적화	사용자 인터페이스 개선	실용 분석을 통한 피드백 반영																
사용자를 위한 맞춤형 음성인식 및 처리 알고리즘 개발	사용자 친화적인 최적화 인터페이스 구현	실용 시범(시범 운영 :기관 1개소) 확보 및 서비스 수행을 통한 활용성 평가																
시설 접근성 및 사용 편의성 향상	자연어 처리 기능 통합을 통한 사용 경험 개선	서비스 확장을 위한 개선 인사이트 도출의 다양성을 내포한 유연한 시스템 구현																
(그림 1) AI 기반 음성인식 안내 키오스크 목표 개념도																		

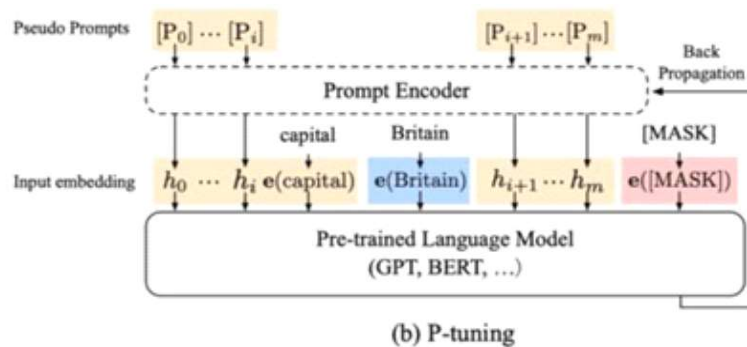
1.적절한 응답 생성 기능 구현



(그림 2) Rasa 기반의 AI 챗봇 개발 흐름도

자연어 처리 엔진과 생성형 AI를 이해하고 이를 연동하는 방법에 초점을 두고 이를 통해 챗봇이 사용자의 의도를 파악하고, 생성형 AI를 활용하여 적절한 응답을 생성하는 기능을 구현합니다.

2. 응답 정확도 향상을 위한 추가 학습 데이터 활용

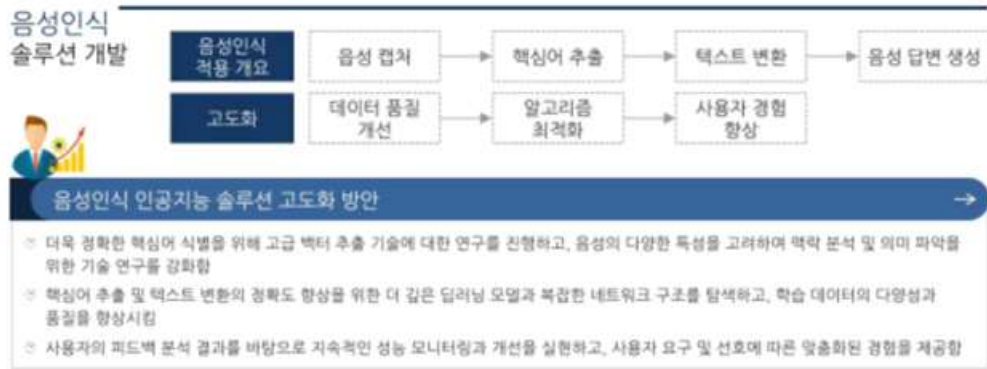


(그림 3) ChatGPT 거대모델 파인튜닝(fine-tuning)

ChatGPT 모델의 파인튜닝(fine-tuning) 과정을 통해 대학 환경에 특화된 응답 생성 능력을 향상시켰습니다. 먼저 기존의 사전 학습된 언어 모델(GPT, BERT 등)을 기반으로, 대학 시설 안내와 관련된 특수한 문맥과 용어를 학습할 수 있도록 추가 학습 데이터를 구축했습니다. 이 과정에서 Prompt Encoder를 활용하여 다양한 유형의 사용자 질문을 체계적으로 인코딩하고, Back Propagation을 통해 모델의 파라미터를 최적화했습니다.

특히 실제 사용 환경을 고려하여 다양한 상황별 시나리오를 설계하고, 이를 바탕으로 한 시뮬레이션 데이터를 생성하여 학습에 활용했습니다. 예를 들어, 같은 질문이라도 다른 표현 방식이나 문맥에서 사용될 수 있는 경우들을 포함시켰고, 불완전하거나 모호한 질문에 대해서도 적절한 응답을 생성할 수 있도록 학습 데이터를 확장했습니다. 구체적인 예시로 본관이 어디야? 라는 질문에는 본관은 어떻게 가, 본관 가는 길 알려줘, 본관은 여기서 몇 분 걸려?와 같은 다양한 방식의 표현들을 고려 하였습니다. 개발된 시스템의 성능 검증을 위해 Front-end 환경에서 다양한 테스트 케이스를 실행하고, 각 응답에 대한 신뢰도 점수 (Confidence Score)를 측정했습니다. 이를 통해 모델의 응답 품질을 정량적으로 평가하고, 낮은 신뢰도를 보이는 케이스들을 분석하여 추가 학습 데이터 구성에 반영했습니다. 이러한 반복적인 최적화 과정을 통해 시스템의 전반적인 응답 정확도와 신뢰성을 크게 향상시켰습니다.

3. 음성 인식 및 데이터 학습



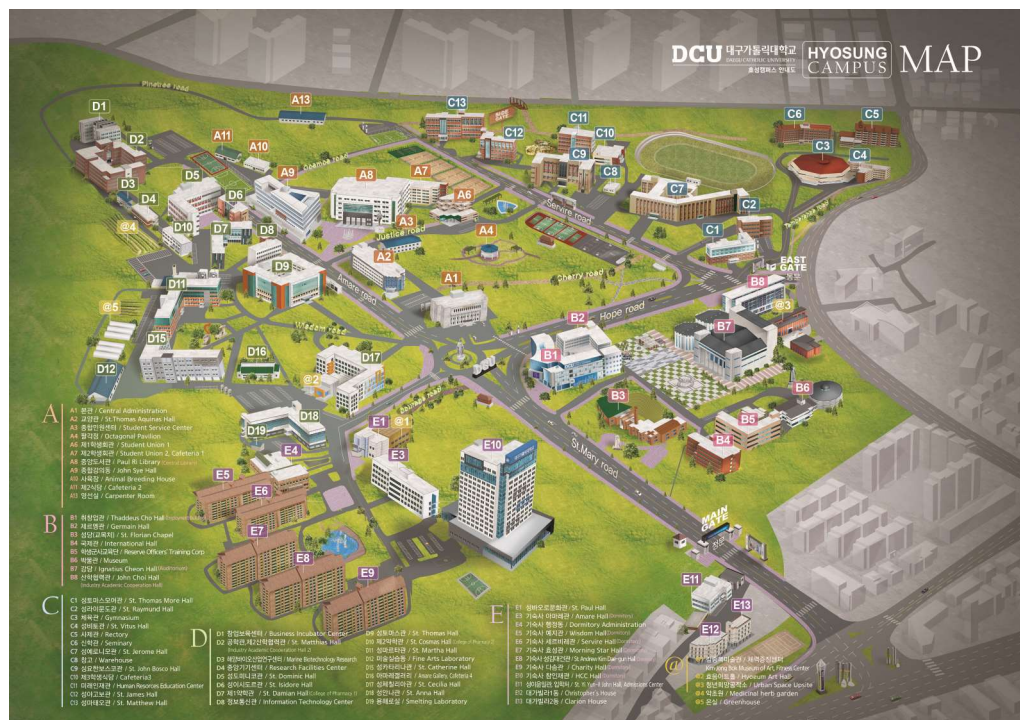
(그림 4) 음성인식 솔루션 개발 과정

음성 인식 후 핵심어를 추출하여 STT(Speech To Text) 및 TTS(Text To Speech)를 통해 음성 답변을 출력하도록 설계하고, 공공 시설 안내의 경우 대상 시설 및 정보 제공 주체에 따라 별도의 데이터셋을 구축해야 하므로 구축 대상을 확정된 후 시설 및 정보 제공 관련 질의-응답을 하나의 건수로 하여 학습용, 검증용 및 테스트용 데이터셋을 제작함. 실제 사용 상황에서의 데이터 학습이 중요하므로 실제 사용자(예: 대학생, 교수 등 학교 관계자)의 입장에서 수집된 대화 데이터셋의 내용을 검수 후 개발하며, 음성 인식 감지 후 신호 전처리 및 노이즈 제거, 후처리 및 신호 강화, 성능 모니터링 및 최적화의 순서로 음성 인식 정확도를 개선할 수 있습니다.

1. 데이터 수집 및 전처리

- 대구가톨릭대학교 캠퍼스 내 편의시설 및 길 안내 정보를 포함한 데이터를 수집합니다.
- 캠퍼스 내 정보를 수집하기 위해서 대구가톨릭대학교 지도와 학교 홈페이지를 참고했다.

수행방법



(그림 5) 대구가톨릭대학교 효성캠퍼스 지도

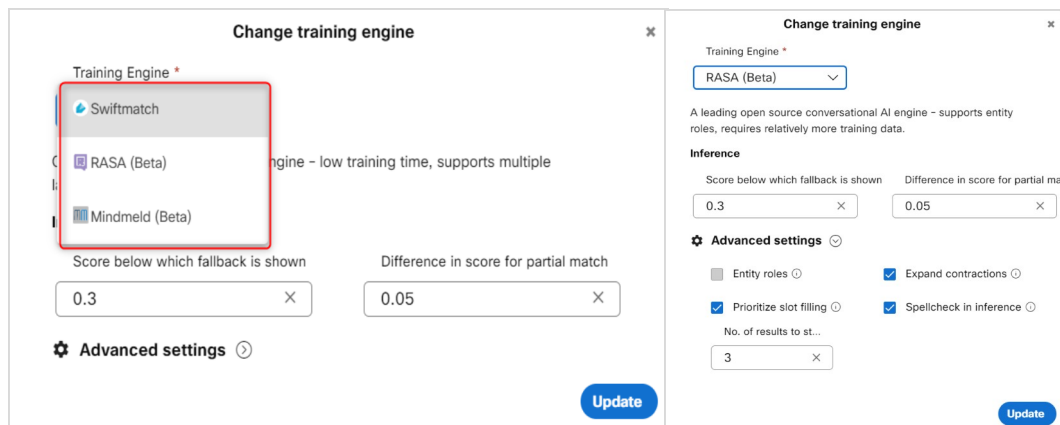


(그림 6) 대구가톨릭대학교 홈페이지

- 수집된 데이터를 Rasa 시스템에 맞게 전처리하여 음성 인식 및 응답에 최적화된 형식으로 준비합니다.

2. 모델 개발 및 학습

- Rasa NLU를 활용하여 사용자의 의도를 파악하는 음성 인식 모델과 자연어 처리(NLP) 모델을 학습하기 위해 데이터를 변환합니다.
- 변환된 데이터는 Action.py, rules.yml, stories.yml, nlu.yml로 나뉩니다.
- 학습된 모델은 사용자의 질문을 인식하고 적절한 응답을 제공할 수 있도록 설계됩니다.



(그림 7) RASA training engine 변경 화면

3. 시스템 통합 및 대화 흐름 설정

- 다양한 사용자 요구에 맞는 동작을 수행할 수 있도록 대화 흐름을 관리하고 통합합니다.
- Action.py, rules.yml, stories.yml, nlu.yml 파일을 활용하여 대화 규칙과 시나리오를 정의합니다.
- 4가지 파일들은 각기 다른 역할을 통해 음성 인식 키오스크가 원활한 대화형 시스템으로 작동할 수 있도록 지원하며, 사용자 요청에 맞는 정확한 응답을 제공합니다.

각 파일들의 기능과 설명은 다음과 같습니다.

(1) Action.py

```
from typing import Any, Text, Dict, List
from rasa_sdk import Action, Tracker
from rasa_sdk.executor import CollectingDispatcher
from rasa_sdk.events import SlotSet

class ActionStandardizeLocation(Action):
    def name(self) -> str:
        return "action_standardize_location"

    def run(self, dispatcher, tracker, domain):
        user_message = tracker.latest_message.get('text', '').strip()
        print(f"Received message: {user_message}")

        standardization_map = {
            "본관": "본관",
            "교양관": "교양관",
        }

        standardized_location = self.standardize_input(user_message, standardization_map)
        print(f"Standardized location: {standardized_location}")

        if standardized_location:
            response_message = {
                "본관": "본관은 정문에서 거리 394m에 위치해있으며, 도보로 총 7분 소요됩니다.",
                "교양관": "교양관은 정문에서 거리 576m에 위치해있으며, 도보로 총 11분 소요됩니다.",
            }.get(standardized_location)

            dispatcher.utter_message(text=response_message)
            return [SlotSet("location_response_given", True)] # 슬롯 업데이트

        dispatcher.utter_message(text="죄송합니다, 해당 위치를 찾을 수 없습니다.")
        return []
```

기능: 사용자와의 대화 중 특정 요청에 대해 실행할 행동(액션)을 정의합니다.

설명: 사용자의 의도에 맞는 기능 수행을 위한 행동을 설정하여 필요한 정보를 제공합니다.

2) rules.yml

```
version: "3.1"

rules:

# - rule: Say goodbye anytime the user says goodbye
#   steps:
#   - intent: goodbye
#   - action: utter_goodbye

# - rule: Say 'I am a bot' anytime the user challenges
#   steps:
#   - intent: bot_challenge
#   - action: utter_iamabot

# - rule: when got a english word
#   steps:
#   - intent: eng_word
#   - action: utter_eng_word

- rule: Ask the user to rephrase whenever they send a message with low NLU confidence
  steps:
  - intent: nlu_fallback
  - action: action_default_fallback
```

기능: 대화에서 특정 조건을 만족할 때 수행할 규칙(rule)을 정의합니다.

설명: 사용자가 특정 키워드를 말하거나 특정 상황이 발생했을 때 실행될 행동을 지정합니다.

(3) stories.yml

```
version: "3.1"

stories:
- story: dcu_route_a_1_path
  steps:
    - intent: dcu_route_a_1
    - action: action_standardize_location
    - action: utter_dcu_route_a_1
- story: dcu_route_a_2_path
  steps:
    - intent: dcu_route_a_2
    - action: action_standardize_location
    - action: utter_dcu_route_a_2
- story: dcu_route_a_3_path
  steps:
    - intent: dcu_route_a_3
    - action: action_standardize_location
    - action: utter_dcu_route_a_3
- story: dcu_route_a_4_path
  steps:
    - intent: dcu_route_a_4
    - action: action_standardize_location
    - action: utter_dcu_route_a_4
```

기능: 대화의 흐름(시나리오)을 구성합니다.

설명: 사용자가 질문을 이어갈 때, 문맥에 맞춰서 대화의 흐름을 관리합니다.

(4) nlu.yml

```
- intent: duc_route_c_1
  examples: |
    - [성토마스모여관](location)은 [어디에](word) 있나요?
    - [성토마스모여관](location)의 위치가 [어디](word)예요?
    - [성토마스모여관](location)으로 가려면 어느 쪽으로 가야 하나요?
    - 여기서 [성토마스모여관](location)까지 [얼마나](word) 걸리나요?
    - [성토마스모여관](location)이 있는 방향을 [알려줘](word).
    - [성토마스모여관](location) 어떻게 가?
    - [성토마스모여관](location) [어디야](word)?
    - [성토마스모여관](location) 찾고 있는데 도와줘.
    - [성토마스모여관](location)이 어딘지 헛갈려.
    - [성토마스모여관](location) 위치와 가는 방법.
    - [성토마스모여관](location) 경로 [알려줘](word)
    - [성토마스모여관](location) 몇 분 걸려?
    - [성토마스모여관](location) 몇 미터 남았어?
    - [[성토마스모여관](location)](location) [얼마나](word) 걸려?

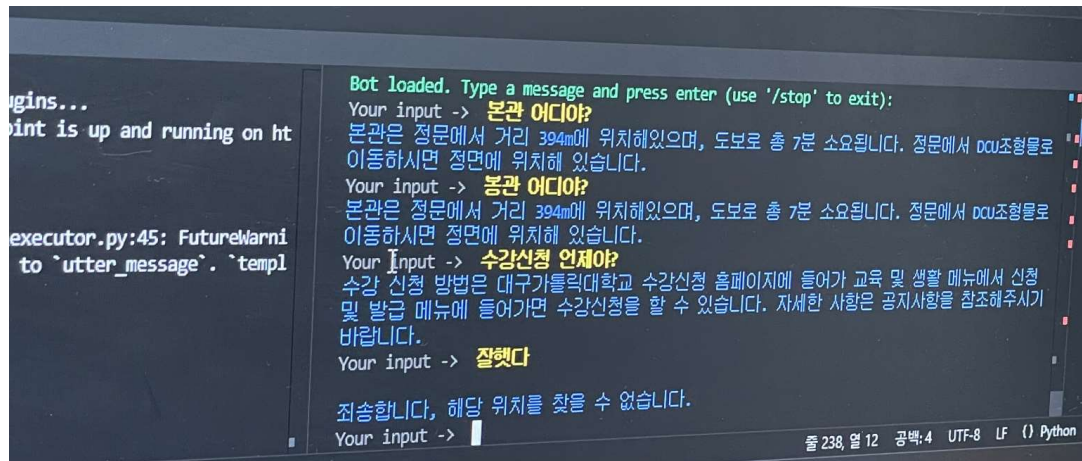
- intent: duc_route_c_2
  examples: |
    - [성라이문도관](location)은 [어디에](word) 있나요?
    - [성라이문도관](location)의 위치가 [어디](word)예요?
    - [성라이문도관](location)으로 가려면 어느 쪽으로 가야 하나요?
    - 여기서 [성라이문도관](location)까지 [얼마나](word) 걸리나요?
    - [성라이문도관](location)이 있는 방향을 [알려줘](word).
    - [성라이문도관](location) 어떻게 가?
    - [성라이문도관](location) [어디야](word)?
    - [성라이문도관](location) 찾고 있는데 도와줘.
    - [성라이문도관](location)이 어딘지 헛갈려.
    - [성라이문도관](location) 위치와 가는 방법.
    - [성라이문도관](location) 경로 [알려줘](word)
    - [성라이문도관](location) 몇 분 걸려?
    - [성라이문도관](location) 몇 미터 남았어?
    - [성라이문도관](location) [얼마나](word) 걸려?
```

기능: 사용자 의도를 파악하고, 중요한 엔터티를 추출합니다.

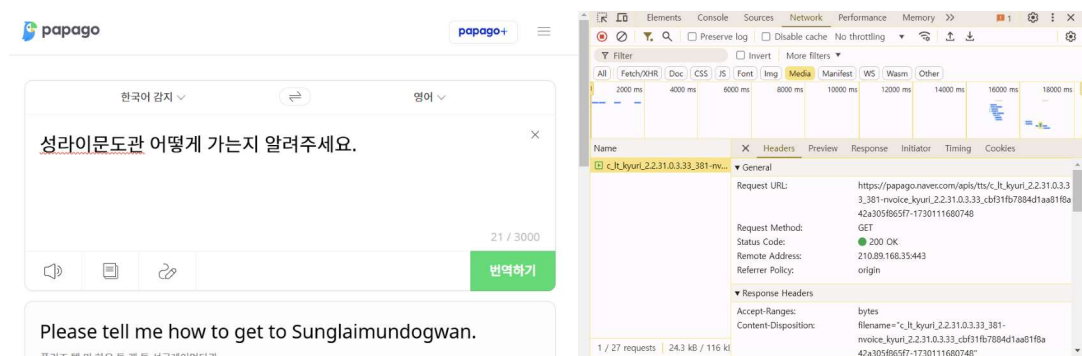
설명: 사용자가 말한 문장을 분석하여 의도를 이해하고 필요한 정보를 인식합니다.

4. 음성 인식과 TTS 시스템 테스트

- STT(음성-텍스트 변환) 및 TTS(텍스트-음성 변환) 기술을 이용하여 음성 인식 기능을 테스트합니다.
- 음성 테스트는 구글 파파고(papago)를 활용하여 테스트를 수행하고, 사용자 질문에 대한 음성 응답이 정확하게 출력되도록 개선 및 오류를 수정합니다.



(사진 1) 터미널에서 테스트 수행



(그림 8) STT 및 TTS 테스트

5. 결과 검토 및 최적화

- 사용자 피드백을 바탕으로 시스템 성능을 평가하기 위해 답변 중에 높은 confidence를 가진 답변을 채택합니다.
- 질문에 대한 정확한 답변을 하기 위해 시스템의 응답 정확도 및 대화 흐름을 지속적으로 개선합니다.



(그림 9) 답변별 confidence 목록

수행 성과는 다음과 같습니다.

1. 디지털 격차 해소

대구가톨릭대학교 내 디지털 소외계층을 포함하여 음성 인식 기반 길 안내 및 학교 정보 제공 서비스를 구축하여 정보 접근성을 높였습니다. 복잡한 메뉴 구조를 이를 통해 고령자 및 신체적 제약이 있는 사용자가 편리하게 서비스를 이용할 수 있는 환경을 마련했습니다. 이를 통해 사회적 포용성을 강화하고 디지털 격차 문제 해결에 기여하였습니다.

2. 기술 적용 및 실용성 검증

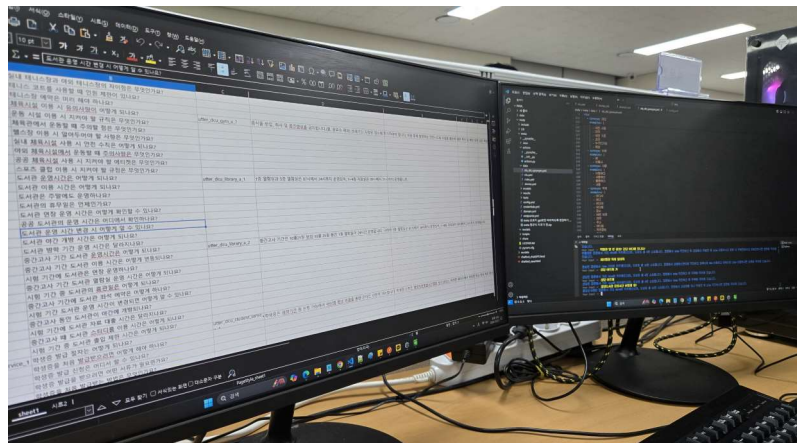
Rasa NLU 및 NLG, STT/TTS 등의 기술을 통해 자연어 이해와 음성 안내를 제공하는 AI 시스템을 구축하여 사용자 의도를 정확히 파악하고 응답할 수 있는 대화형 서비스를 완성하였습니다. 대화 흐름과 시나리오를 관리하여 사용자 친화적인 인터페이스를 구성하였습니다.



(사진 2) 자연어 이해 및 NLU 학습하는 모습

3. 데이터셋 구축 및 학습

대구가톨릭대학교의 주요 정보와 편의시설 데이터를 수집하여 자연어 처리 모델에 반영, 사용자 의도에 맞는 안내와 서비스 정보를 제공했습니다. 또한, 사용자와의 상호작용을 위한 데이터셋 구축과 전처리 작업을 통해 시스템의 정확성과 신뢰도를 높이는 성과를 달성했습니다.



(사진 3) 데이터셋 구축 작업 화면



(사진 4) 데이터셋 구축 작업 모습

4. 프로토타입 개발 및 테스트 성과

프로젝트의 체계적인 검증을 위해 프로토타입을 개발하고 다각도의 테스트를 수행했습니다. 개발 환경 설정부터 시작하여 상세한 테스트 케이스를 작성하고, 다양한 사용 시나리오에 따른 기능 검증을 진행했습니다. 또한 사용자들로부터 받은 피드백을 바탕으로 반복적인 프로토타입 개선을 수행하여 시스템의 완성도를 높였습니다. 이러한 체계적인 테스트와 개선 과정을 통해 실제 사용 환경에서의 안정성과 효용성을 확인하였습니다.

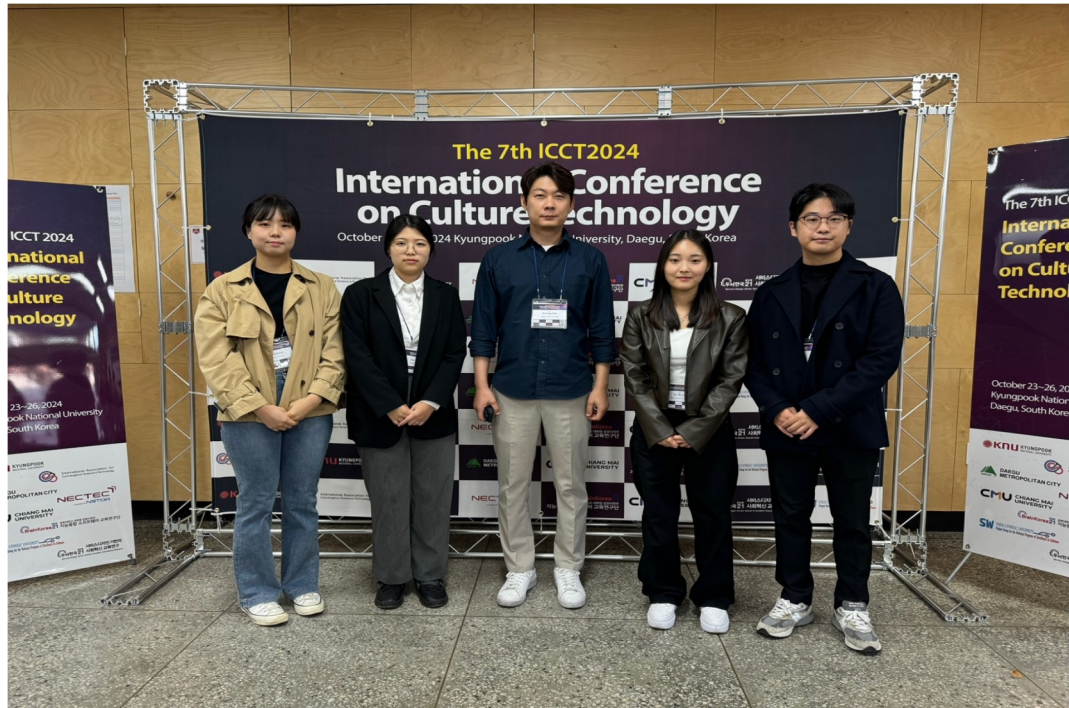
5. 산학 협력을 통한 인재 양성 기여

기업과 교육생이 협력하여 AI 음성 인식 시스템의 실무 개발 경험을 제공하고, 이를 통해 학생들이 실제 산업 문제를 해결하는 경험을 쌓았습니다.

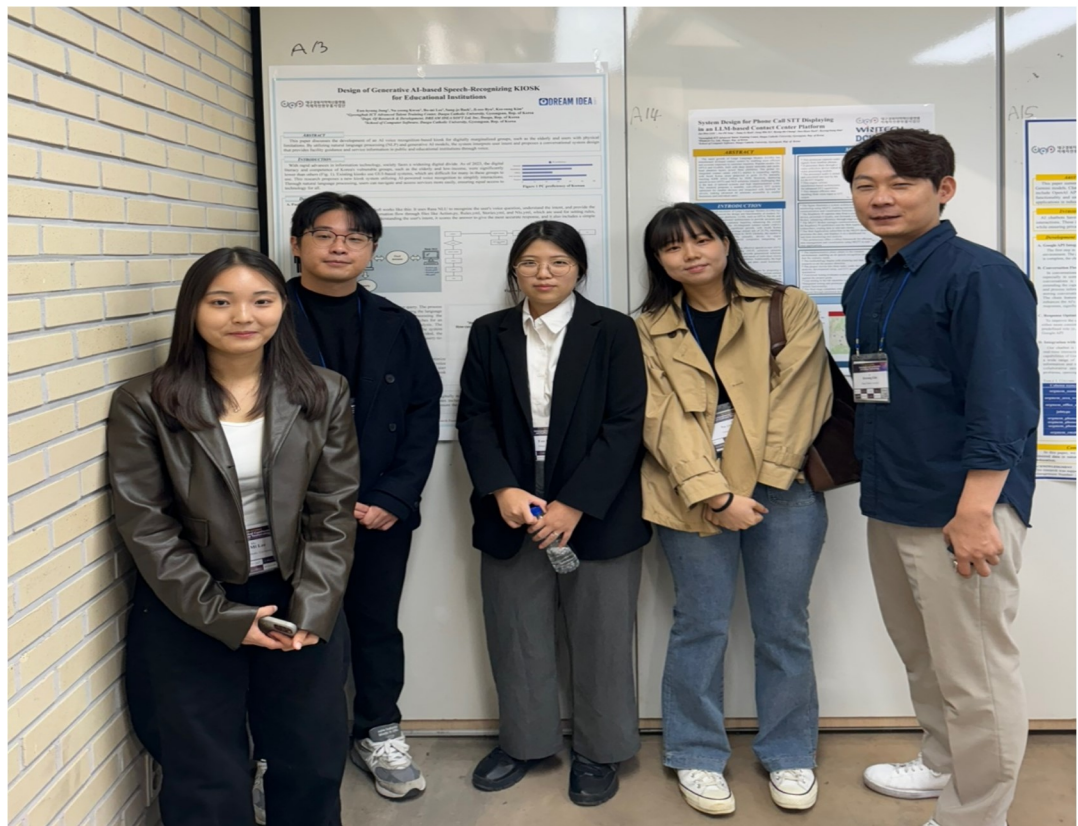
6. 학술 활동 및 연구 성과

2024년 International Conference on Culture Technology (ICCT 2024)에서 포스터 세션을 통해 본 프로젝트의 연구 성과를 성공적으로 발표했습니다. 특히 Rasa NLU를 활용하여 사용자의 자연스러운 발화 의도를 정확히 파악하고 맥락에 맞는 응답을 생성하는 지능형 음성 키오스크 시스템의 구현 방법과 성과를 중점적으로 소개했습니다. 또한 이러한 기석자들로부터 긍정적인 관심을 받았습니다. 학술대회 포스터 발표를 통해 연구 성과를 공유하고 검증받음으로써, 디지털 포용성 증진을 위한 기술 발전에 의미 있는 기여를 했습니다. 특히 국제 학술대회에서 포스터 발표는 우리의 연구를 학술적 관점에서 체계화하고 발전시키는 계기가 되었으며, 프로젝트의 학술적인 가치를 논의할 수 있는 기회였습니다.

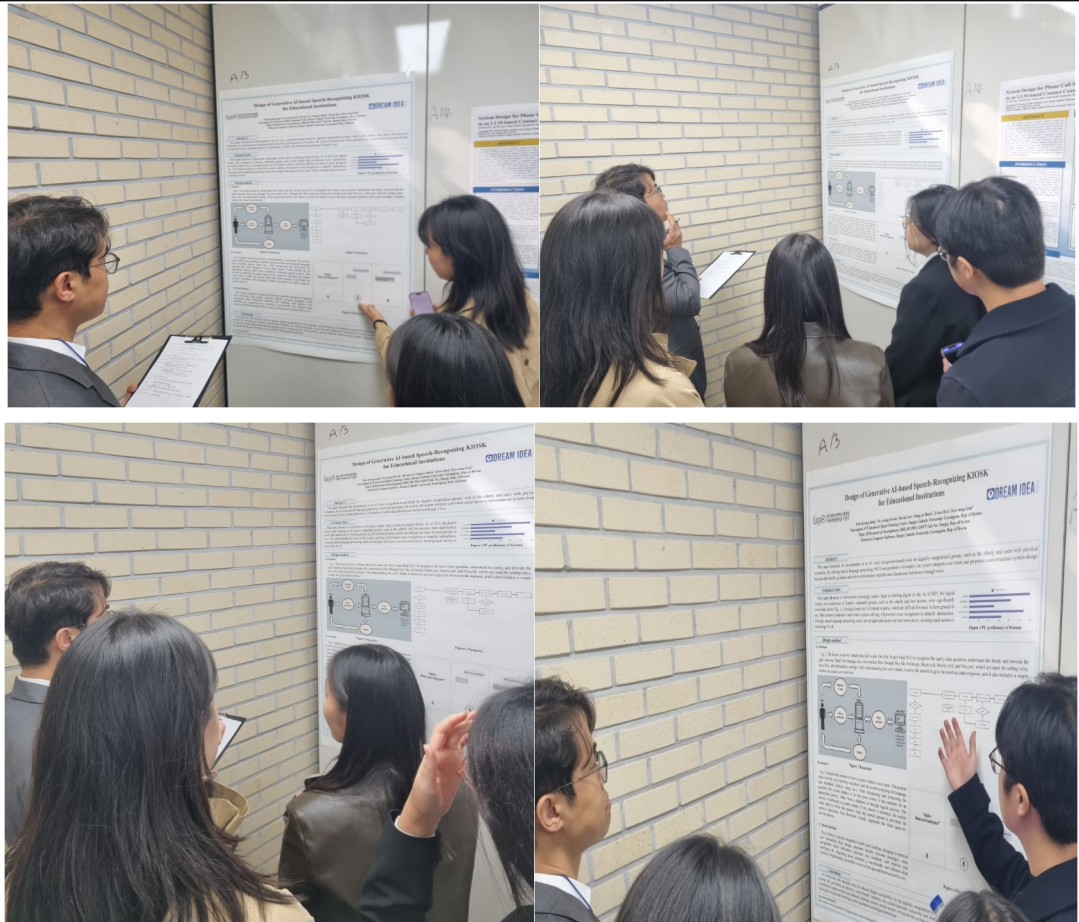
• ICCT 2024 국제학술대회 포스터 세션 참가



(사진 5) ICCT 2024 국제학술대회 포스터 세션 참가 모습 - 1



(사진 6) ICCT 2024 국제학술대회 포스터 세션 참가 모습 - 2



(사진 7) ICCT 2024 연구 성구 성과 발표 모습

※ [붙임] 지원금 사용 내역

위와 같이 기업협업과제(캡스톤디자인 프로젝트)를 신청합니다.

2024. 10. 28.

책임교수 : 김기성 (서명/인)

미래차전환부품사업단장 귀하

【붙임】 지원금 사용 내역

재료비					
품명	규격	단가	수량	금액(원)	활용내역
키오스크 프로세서	AMD Ryzen 9 5900x	470,000	1	470,000	사용자 요청을 해석하고 실행하기 위함
키오스크 딥러닝 코어	INNO3D 지포스 RTX 4090 OC D6X	3,900,000	1	3,900,000	사용자 요청을 AI를 활용하여 실행 연산을 하기 위함
키오스크 보드	X570S AORUS Elite	393,000	1	393,000	사용자 요청의 데이터 흐름을 처리하기 위함
키오스크 메모리	Kingston BEAST Memoria DDR4	390,000	1	390,000	사용자의 요청을 기억하고 처리하기 위함
터치 패널	32인치 적외선 IR 터치패널 USB타입	2,600,000	1	2,600,000	사용자의 UX 개선을 위해 터치 디스플레이를 입력하기 위함
키오스크 영상입력장치	영상입력장치 모듈 USB 13MP 4K	380,000	1	380,000	사용자의 얼굴을 인식하여 답변을 출력하기 위함
키오스크 음성출력장치	BOSE SURROUND SPEAKERS	450,000	1	450,000	사용자의 질문에 음성 답변을 출력하기 위함
키오스크 음성입력장치	RODE PODMIC USB	317,000	1	317,000	사용자의 질문 음성을 입력하기 위함
키오스크 금형	450(W)×157(D)×1790(H)mm	4,200,000	1	4,200,000	사용자의 UX 개선을 위해 장치 외형을 구현하기 위함
재료비 소계				13,100,000	

기타 지원금	
항목	금액(원)
학회등록비	2,500,000
자문료	1,000,000
기타 지원금 소계	3,500,000

회의비			
회의일자	회의내용	참석자	금액(원)
24.10.22	음성안내 AI 개발을 위한 회의	김기성, 류지수, 이설기, 권나영, 이보미, 백상조, 정은경	105,000
24.10.30	키오스크 제작을 위한 회의	김기성, 류지수, 이설기, 권나영, 이보미, 백상조, 정은경	105,000
회의비 소계			210,000
지원금 합계			16,810,000