

2025년 새싹 해커톤(SeSAC Hackathon) AI 서비스 기획서

팀명	셋쌍이들
팀 구성원 성명	정건희, 주연우, 허준서

1. AI 서비스 명칭

청각 장애인을 위한 실시간 대화 및 환경 소리 비서 (Contextual Interpreter)

2. 활용 인공지능 학습용 데이터

	활용 데이터명	분야	출처
1	위급상황 음성/음향 데이터	재난/안전	AI Hub
2	소음 환경 음성인식 데이터	환경	AI Hub

3. 핵심내용

본 서비스는 환경 안전 인지와 대화 맥락 해석이라는 두 가지 핵심 AI 기능으로 구현됩니다.

첫째, 환경 소리 분류 및 안전 인지 기능입니다. 저희는 AI hub에 등록된 위급상황 음성/음향 데이터와 소음 환경 음성인식 데이터를 활용하여 위험 소리(ex. 사이렌, 경적, 비상벨)와 일상 필수 소리(ex. 초인종, 물 끓는 소리)를 정확하게 감지합니다. 여기에 음원 위치 추정(DOA) 기술을 결합하여, 단순한 소리 감지를 넘어 소리가 발생한 방향과 거리를 정확하게 시각화 합니다. 예를 들어, 사용자에게 "🚒 (오른쪽 3m) 경적 소리!"와 같이 구체적이고 즉각적인 안전 알림을 제공합니다.

둘째, 대화 및 맥락 해석 기능입니다. 고성능 자동 음성 인식(ASR) 모델을 통해 대화 내용을 실시간으로 텍스트로 변환하여 지연 없이 자막을 제공합니다. 여기서 나아가, 텍스트 기반 감정 분류 모델을 통해 상대방의 감정 톤을 분석하고 아이콘이나 경고 메시지(예:

"☹ (분노), 상대방이 격앙되었습니다")를 함께 표시하여 소통의 질을 높입니다. 또한, 긴 대화의 핵심 의도를 NLP 기반 요약 모델로 추출하여 청각 장애인의 인지적 피로를 낮추고 대화의 맥락을 놓치지 않도록 지원합니다.

4. 제안배경 및 목적

현대 사회에서 청각은 단순히 정보를 획득하는 수단을 넘어, 개인의 안전과 필수적인 사회적 연결을 위한 핵심 감각으로 기능합니다. 그러나 청각 장애인들은 이 필수적인 감각의 부재로 인해 만성적인 환경 및 소통 격차라는 구조적 문제에 지속적으로 직면하고 있습니다.

1. 정보 격차 및 안전 사각지대 심화

기존의 보청기나 인공 와우는 청각 장애인의 잔존 청력을 보조하는 데 집중되어 있습니다. 이로 인해 비언어적 환경 소리에 대한 인지 능력은 여전히 취약합니다. 특히 경적, 사이렌, 비상벨과 같이 생명과 직결된 소리의 방향, 긴급성, 그리고 맥락을 정확히 파악하는 데는 근본적인 한계가 존재합니다. 결과적으로 청각 장애인은 공공장소나 독거 환경에서 위험 상황 발생 시 안전의 사각지대에 놓여 즉각적인 대처가 어렵습니다.

2. 소통의 질적 한계 및 심리적 피로 가중

대화 상황에서 문자 통역이나 자막에 의존할 경우, 상대방의 감정 톤, 뉘앙스, 말의 숨겨진 의도 등 비언어적 정보가 누락됩니다. 이러한 정보의 결핍은 단순한 오해를 넘어, 원활한 사회적 상호작용 및 깊이 있는 관계 형성을 저해합니다. 또한, 누락된 정보를 채우기 위해 끊임없이 추론해야 하는 과정은 심리적 피로를 가중시키며 사회적 소외감을 느끼게 합니다.

3. 기존 보조 기술의 한계

현재 시장에 존재하는 기술들은 대부분 단순한 소리 인식 또는 일대일 자막 변환 수준에 머물러 있습니다. 청각 장애인이 '주변 환경을 주도적으로 인지하고 이에 능동적으로 대처'할 수 있도록 돕는 지능형 맥락 해석 기능은 여전히 미흡한 상태입니다. 따라서 청각 장애인의 삶을 근본적으로 재설계하고 안전하고 자율적인 생활 환경을 보장하기 위해서는, AI 기반의 혁신적인 맥락 해석 솔루션 도입이 시급합니다.

5. 세부내용

- 활용 데이터 및 AI 모델

본 AI 서비스는 환경 안전 인지 및 대화 맥락 해석이라는 두 가지 핵심 기능을 구현하기 위해 AI-Hub에 등록된 두 가지 핵심 데이터셋을 활용합니다.

첫째, 위급상황 음성/음향 데이터는 사이렌, 경적, 비상벨 등 위험 소리를 포함한 다양한 음향 이벤트를 정확하게 분류하고 감지하는 소리 분류 모델 학습에 필수적으로 사용됩니다.

둘째, 소음 환경 음성인식 데이터는 카페나 길거리와 같은 일상 소음 환경 속에서도 대화 음성을 깨끗하게 분리하고 인식률을 높여주는 고성능 자동 음성 인식(ASR) 모델의 강건성을 확보하는 데 활용됩니다.

이 두 데이터셋을 기반으로 구현되는 AI 모델은 다음과 같습니다. 대화 내용을 실시간으로 텍스트 변환하는 고성능 ASR 모델이 기본적인 소통 환경을 마련합니다. 이 텍스트는 NLP 기반 감정 분류 모델로 전달되어 상대방의 감정 톤(분노, 기쁨 등)을 판별하고, 장문의 대화에서는 NLP 기반 요약 모델을 통해 핵심 의도를 추출합니다. 안전 인지 측면에서는 학습된 소리 분류 모델이 소리를 감지함과 동시에, 마이크 어레이와 연동된 음원 위치 추정(DOA) 기술을 적용하여 소리의 방향과 거리를 계산함으로써 사용자에게 시각화 된 안전 정보를 제공하게 됩니다.

- 세부내용

1. 서비스 아이디어 개요 및 서비스 방법

서비스는 사용자가 휴대하는 스마트폰이나 전용 웨어러블 디바이스의 마이크를 통해 주변 소리(오디오 스트림)를 수집하는 것으로 시작됩니다. 핵심은 AI 엔진이 소리를 환경 소리와 대화 음성으로 정확하게 분리하여 각각의 전문 AI 모듈로 전달하는 것입니다.

환경 소리는 Sound Event Detection (SED) 모델을 통해 경적, 사이렌, 초인종 등의 위험 및 필수 소리로 분류되며, 음원 위치 추정(DOA) 기술을 통해 소리가 발생한 방향까지 함께 분석됩니다.

대화 음성은 ASR(자동 음성 인식) 모델을 통해 텍스트로 변환되고, 이 텍스트는 NLP 기반 감정 분석 및 요약 모델을 거쳐 발화자의 감정 톤과 핵심 의도가 해석됨

니다.

이후, AI가 생성한 모든 맥락 정보는 단순한 자막을 넘어, 🚨 경고 아이콘, 방향 화살표, 감정 이모지, 차별화된 진동 패턴 등의 시각적 및 촉각적 알림으로 변환되어 사용자에게 실시간으로 출력됩니다. 이러한 다중 정보 전달 방식을 통해 청각 장애인은 환경에 대한 완전한 인지 능력과 통제감을 확보하게 됩니다.

2. 적용 기술 및 구현 가능성

본 서비스의 구현 가능성은 AI-Hub의 특화 데이터를 활용한 전이 학습과 기존 오픈소스 AI 모델의 융합을 통해 충분히 입증될 수 있습니다.

- **환경 소리 감지 및 위치 추정:** AI-Hub의 위급상황 음성/음향 데이터로 훈련된 경량화 된 SED 모델을 활용하여 핵심 위험 소리 감지의 정확도를 높입니다. 여기에 음성인식 모델 학습에 사용되는 소음 환경 데이터를 추가적으로 활용하여 소음 속에서의 비상 소리 분리 성능을 극대화할 수 있습니다. 이미 검증된 DOA 기술과 결합하여 실시간 환경 인지 기능을 완성합니다.
- **음성 인식 및 NLP:** AI-Hub의 소음 환경 음성인식 데이터로 파인 튜닝(Fine-tuning)된 Whisper 또는 유사 오픈소스 ASR 모델을 활용하여 실제 소음 환경에서의 자막 정확도를 높여 안정적인 실시간 텍스트 변환을 확보합니다. 텍스트 감정 분석 및 요약은 Hugging Face Transformers 라이브러리 내의 사전 학습된 NLP 모델을 사용하여 빠르고 신속하게 맥락 정보를 제공합니다. 이러한 기술 구성은 AI-Hub의 목적성 높은 데이터를 기반으로 핵심 성능을 강화하고, 기존 기술을 융합하는 데 초점을 맞추어, 짧은 기간에도 핵심 기능의 동작을 성공적으로 시연할 수 있도록 합니다.

3. 서비스 아이디어의 창의성

본 서비스의 창의성은 단순히 소리를 문자로 바꾸는 기존의 보조 기술의 한계를 넘어, '소리의 맥락(Context)'을 해석하여 전달하는 데 있습니다.

- **맥락 융합을 통한 위험 감지:** 소리 감지, 방향 인지, 대화의 감정 톤을 동시에 융합하여 사용자에게 상황의 '긴급성'과 '출처'를 입체적으로 제공하는 것은 기존 보조 기술에서 찾아볼 수 없는 혁신적인 접근 방식입니다.
- **심리적 안정감 재설계:** 이 서비스는 청각 장애인의 삶을 둘러싼 정보적 불안 요소를 제거하고, 환경에 대한 완전한 통제감을 회복시켜 심리적 안정감을 증대시

킨다는 점에서 사회적 약자를 위한 AI 서비스의 새로운 패러다임을 제시합니다.

6. 기대효과

- **소리 공백 해소 및 독립적인 생활 환경 구축:** 청각 장애인이 혼자 있을 때 놓치기 쉬웠던 생활 필수 소리 (초인종, 전화 알림, 가스레인지 타이머, 물 끓는 소리 등)를 명확히 인지하게 됩니다. 이는 타인의 도움 없이도 일상 가사를 안전하고 자율적으로 수행하는 독립성을 획기적으로 향상시킵니다.
- **불안감 감소 및 환경 통제감 회복:** 소리를 듣지 못해 주변 환경에서 무슨 일이 벌어지는지 모를 때 느끼는 만성적인 불안감이 크게 줄어듭니다. 실시간으로 소리의 종류와 방향을 시각적으로 확인하면서, 환경에 대한 통제감을 회복하고 심리적 안정감을 얻습니다.
- **대화 맥락 이해를 통한 소통 피로도 감소:** 단순한 텍스트 변환을 넘어, AI가 대화의 감정톤 (기쁨, 불만, 긴급함)과 핵심 의도를 요약해 줌으로써, 청각 장애인은 비언어적 정보까지 이해하게 됩니다. 이는 오해를 줄이고, 대화를 따라가기 위해 집중해야 했던 인지적 피로를 크게 낮춥니다.
- **비장애인과 자연스러운 상호작용 촉진:** 대화 시 '무슨 말을 하는지'뿐만 아니라 '어떤 감정으로 말하는지'를 파악하게 되면서, 비장애인과 소통이 더욱 자연스럽고 깊이 있게 이루어집니다. 이는 청각 장애인이 소통에서 느끼는 소외감을 실질적으로 완화합니다.