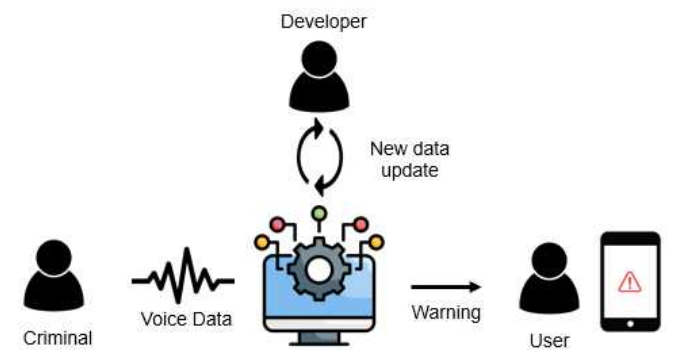


2023년 K-디지털 트레이닝 해커톤 아이디어 개발 기획서		
참가팀명		피싱방법대
제안 아이디어	명칭	딥러닝을 활용한 실시간 보이스피싱 감지 App
	소개	STT(Speech-To-Text) API를 활용해 보이스피싱을 AI 알고리즘으로 감지 후 실시간 알림을 주는 앱 서비스
1. 추진배경		▶ 해커톤에 응모하게 된 동기와 목표 및 아이디어 개요를 간략히 기술
<p>택배사나 공공기관을 사칭하는 등의 보이스피싱 수법이 점점 다양하고, 교묘한 방법으로 진화하고 있다. 이에 더해, 코로나19 이후 비대면 방식의 보편화로 온라인 뱅킹을 활용한 금융사기, 사칭피해 사례도 증가하고 있다. 예시로 최근에 은행과 동일한 구조를 가진 악성 어플을 제작해 다운로드 URL을 문자로 전송한 후, 은행 직원을 사칭해 전화를 걸어 어플 설치를 유도하는 수법이 있다.</p> <p>이렇게 발전하는 보이스피싱 수법은 기존에 보이스피싱 수법을 알고 있음에도 전화를 통해 피해자의 심리적 특성을 악용함으로써 피해 사례를 증가시킨다. 이러한 점들을 토대로, 보이스피싱의 각종 방식을 통해 피해자가 발생했을 때 심리적으로 위축된 피해자에게 실시간으로 알림을 줄 수 있는 생활 밀착형 솔루션이 필요하다고 결론을 내렸고 이를 '실시간 음성 분석 기술'을 활용해 구현하고자 한다. 해당 대응책을 통해 전화로 업무를 진행하는 금융 서비스 등의 분야에 다방면으로 활용 가능할 것이다.</p>		
2. 개발 목표 및 내용		▶ 아이디어 소개, 계획 등 간략히 기술 (필요 시 사진 등 첨부 가능)
<p>서비스 모델 구현 목표: 딥러닝을 활용한 실시간 보이스 피싱 App 개발</p> <p>개발 내용</p> <ul style="list-style-type: none"> ■ 알고리즘을 구현하기 전에 보이스피싱 데이터를 수집하여 EDA를 사용해 단어빈도를 파악하고, 토큰화된 단어에 대해 Softmax 함수를 활용하여 가중치를 부여, 이후 Binary Text Classification을 통해 보이스피싱 여부를 판별 ■ STT API를 활용해 음성파일을 텍스트로 변환하고, 이후 딥러닝을 활용해 학습한 모델을 파이썬 Streamlit을 통해 구현 ■ 보이스피싱 여부를 판단하는 모델을 학습시키기 위해 음성파일을 텍스트로 변환한 파일을 Tokenization 이후, BERT 과정을 거쳐 좀 더 정밀한 입력값을 설정 ■ 조건에 따른 성능 차이를 고려하여 여러 알고리즘을 비교 후, 모델 선택 ■ 보이스피싱 의심내용이라 판단되면 경고의 일종으로 휴대폰에 알람이 울리도록 구현 		

<ul style="list-style-type: none"> ■ 새로운 기법의 보이스피싱이 생길 경우, 알고리즘에 데이터를 추가하여 모델을 계속 업데이트할 계획 	
개발 범위	
<p>1. 시스템 구축</p> <ul style="list-style-type: none"> ■ 문제정의단계 : 유형별 보이스피싱을 조사하여 문제 정의서 작성 ■ 분석단계 : 보이스피싱 음성파일을 수집 STT API를 활용하여 텍스트 변환, 보이스피싱 단어 분석 및 Visualizaion ■ 설계단계 : 형태소 분석 및 Tokenization 시행 ■ 구현단계: Softmax 함수를 통해 단어의 가중치 추출 및 CNN 및 RNN 딥러닝을 활용하여 모델 개발 	
<p>2. 테스트 및 모델 안정화</p> <ul style="list-style-type: none"> ■ 테스트 <ul style="list-style-type: none"> - STT API별 성능을 비교하기 위해 정밀한 데이터셋 테스트 - CNN, RNN 등 여러 알고리즘을 활용하여 Text Classification에 대해 모델 성능을 비교 ■ 시스템 안정화 및 추적 <ul style="list-style-type: none"> - 신종 보이스피싱이 나타나면 기존 알고리즘에 데이터를 추가하여 모델을 업데이트 	
구조도	
	
3. 주요 특징 및 핵심 기술	
▶ 아이디어 컨셉, 핵심내용, 활용성, 특징 등 구체적으로 기술	

해당 기술은 사용자가 통화 중 실시간으로 수집한 음성 데이터를 STT(Speech-to-Text) API를 이용해 텍스트로 변환한 후, 일반적인 보이스피싱 수법을 인식하도록 학습된 딥러닝 알고리즘을 이용해 보이스피싱 공격의 가능성을 알려주어 경고 알람을 제공한다.

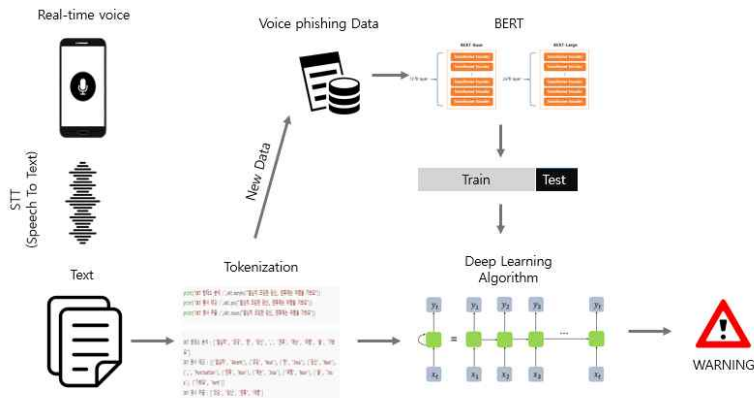
* 전체 개발 과정 :

STT API를 통해 변환한 텍스트 데이터를 Word Tokenization한 후, BERT 모델을 사용해 정확도를 높여 보이스피싱 여부를 판단하는 모델을 학습시키기 위한 데이터를 구성한다. 딥러닝 모델은 텍스트가 long-term인 경우 RNN을 사용하는 것이 더 성능이 높기에 CNN 모델 등과 비교해본 후 성능이 좋게 나온 모델을 선택한다. 성능을 기준으로 선택한 모델에 Softmax 함수를 활용한 Binary Text classification으로 데이터를 학습시킨 뒤 보이스피싱 여부를 판별한다.

* 특징 :

실시간으로 보이스피싱 여부를 판단하는 것이 중요하기에 모델의 처리속도를 높이는 것 또한 주요 과제이다. 한국어는 영어보다 텍스트 처리하는 것이 어렵기 때문에 엑소브레인 사업에서 한국어 특성을 반영한 특화 모델인 KoBERT를 사용해 기존보다 더 정확하게 텍스트를 처리할 수 있다.

또한, 통화 음성 데이터를 모두 수집하기에 막지 못한 신종 보이스피싱 데이터 발생 시, 해당 알고리즘에 학습되어 빠르게 신종 보이스피싱에 대한 피해를 막을 수 있다는 것이 특징이다.



4. 기대효과 및 활용방안 ▶ 경제적 · 기술적 · 사회적 파급효과, 고용창출 등을 자유롭게 기술

1. 보이스피싱의 실시간 예방

보이스피싱은 피해자가 심리적으로 불안한 상태에서 판단력이 흐려져 발생한다. 실시간으로 알람을 보내 경각심을 주어 보이스피싱 발생률을 감소시킨다.

2. 한국어 보이스피싱 범죄 패턴 분석 용이

딥러닝 기능에 일반화 성능을 높여주기 위하여 KoBERT를 사용해 변형되는 한국어 보이스피싱 시나리오 제작에 도움을 준다.

3. 신종 보이스피싱 예방

신종 수법의 보이스피싱 발생 시, 즉각적인 데이터 업데이트를 통해 해당 수법을 예방할 수 있는 방안을 찾아내는 속도를 높일 수 있다.

4. 취득 데이터를 기술 연구 개발에 활용

애플리케이션 사용자의 동의를 받을 시, 음성 및 텍스트 데이터 수집이 가능해 차후 음성 및 텍스트 관련 기술 개발에 도움을 줄 수 있음.

5. 개발 추진 체계

▶ 개발 목표 및 기간 등 전체 개발 추진 체계 기술

개발일정

분류	항목	1	2	3	4	5	6	7
개발 (최정상, 이현주)	STT(Speech To Text)							
	Text data processing							
	Deep learning Training							
	Test algorithms							
	Update new data							
기획 (한다현, 김노옥)	요구사항정의							
	기능개발영역							
	프로세스 정의							
	상세화면 기획							
디자인 (김태우, 한수빈)	원가 및 전개							
	디자인 컨셉 확인							
	디자인 시간 작업							
	시안 확정 및 Develop							
	전체화면 디자인 작업							

역할

개발자: 이현주, 최정상

기획자: 김노옥, 한다현

디자이너: 김태우, 한수빈