개발보고서

k-digital training’ 23’빅데이터 분석가 양성과정

파이널 프로젝트

Find your K-pop

2023년 12월 11일(월) ~ 2024년 1월 12일(금)

**화이팀**

조민우, 김수진, 서강석, 윤정인, 조은겸

**기안서**

|  |  |
| --- | --- |
| 제목 | Find Your K-Star |
| 일시 | 2023.12.11~2024.1.12 |
| 팀구성원 | 조민우 /  조은겸 / 김수진 / 서강석 / 윤정인 |
| * **프로젝트 개요**   외국인들에게 한국 아이돌에 대한 관심을 높이고 나아가 한국어를 학습하거나 한국의 문화에 대해 더 많이 이해할 수 있다. 아이돌의 이름과 관련된 정보를 찾는 과정에서 언어 및 문화적인 지식을 습득할 수 있다.   * **프로젝트 선정 이유**   외국인들의 한국 아이돌에 대한 관심을 이용해 한국에 대한 관심 증가 유도   * **업무 분담**   조민우 : 웹 구현, 텍스트 번역 및 음성출력  김수진 : 모델링, 데이터 수집  서강석 : 웹 크롤링, 데이터 수집  윤정인 : 모델링, 데이터 수집  조은겸 : 모델링, 데이터수집   * **데이터 개요**   사진이미지 : 웹크롤링(카카오 API 사용)  가사데이터 : 웹크롤링(네이버 API 사용)  번역 : 파파고 API 사용   * **개발환경**   Windows 10 Pro 64bit  코랩, 파이참 사용  python 3.10, keras 2.14.0, tensorflow 2.14.0, pandas 2.1.3, opencv 4.8.0, librosa 0.10.1  Django 4.1.13  모델링 : 이미지(CNN), 음성->파동이미지->이미지유사성분석(opencv) | |

목차

1. 서론(프로젝트 진행 배경)-------------------------------------------------------------------- 1
2. 프로젝트 개요 ------------------------------------------------------------------------------- 3
   * 프로젝트 요약 ----------------------------------------------------------------------- 3
   * 개발 환경 ---------------------------------------------------------------------------- 4
   * 팀원 구성 및 역할 ------------------------------------------------------------------ 4
   * 개발 일정 ---------------------------------------------------------------------------- 4
3. 웹 서비스 ------------------------------------------------------------------------------------ 5
   * 서비스 소개 ------------------------------------------------------------------------- 5
   * 서비스 구조 ------------------------------------------------------------------------- 6
4. 모델 ------------------------------------------------------------------------------------------ 8
   * 데이터 베이스 ---------------------------------------------------------------------- 8
   * 이미지 모델 ------------------------------------------------------------------------- 9
   * 사운드 모델 ------------------------------------------------------------------------ 12
   * Letter 모델 ------------------------------------------------------------------------- 15
5. 개선 사항 및 확장 방향 ------------------------------------------------------------------- 17
   * 모델의 품질 업그레이드 ---------------------------------------------------------- 17
   * 웹의 확장 -------------------------------------------------------------------------- 17
   * 서비스에 대한 고찰 --------------------------------------------------------------- 17
6. 프로젝트 진행 배경

# 01 쉽고 재밌게 다룰 수 있는 주제

팀 프로젝트의 성공과 실패는, 그 팀원들이 얼마나 주제에 관심을 갖고 적극적으로 과제를 수행 하느냐에 있을 것입니다. 우리는 프로젝트 주제의 선정 기준의 가장 첫 머리에 흥미와 재미를 넣었습니다. 그 이하로도 중요한 기준들을 세워 주제를 선정하게 되었습니다.

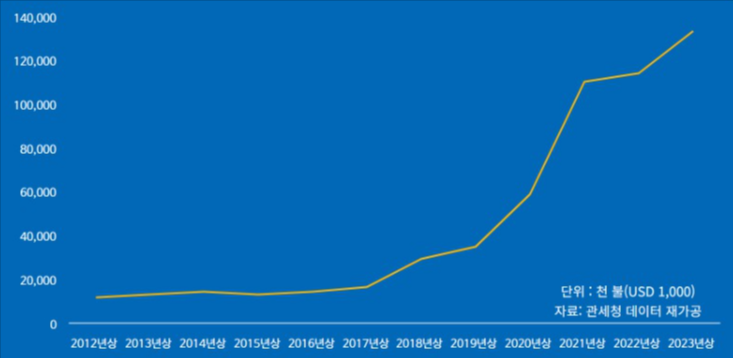
* 팀원 모두가 흥미와 재미를 느끼고 접근할 수 있는가?
* 해결할 수 있는 과제인가?
* 기존보다 더 나은 서비스를 제공할 수 있는가?
* 성장하는 산업을 타겟으로 삼았는가?

이런 이유로 몇 가지 주제 후보가 있었지만 저희는 **외국인을 대상으로** **케이팝 정보를 제공**하는 프로젝트를 시작하게 되었습니다.

# 02 케이팝

케이팝은 2010년대부터 현재까지 급속도로 성장하는 산업입니다. 2023년 상반기에는 무려 1억 3,293만 4000달러를 수출했는데 전년 동기 대비 17.1% 증가했습니다. [그림 0-1]

[그림 0-1] 연도별 상반기 케이팝 수출 금액 추이



케이팝 산업은 정부에서도 적극적으로 지원하고 있습니다. 2024년 온라인 전문 공연장 운영 지원 예산은 약 70억원, 온/오프라인 콘텐츠 개발 지원액은 약 76억원으로 책정되었습니다. 케이팝이 포함되어 있는 콘텐츠 산업 분야 총 예산은 1조 5400만원으로 2023년에 비해서 18%나 증가했습니다. <부록 1, 기사, ‘정부,K-POP등 콘텐츠 산업 육성..내년 예산 1조22억 확정’>

케이팝은 해외 공연문화 시장에서 지속적으로 파이를 키우고 있습니다. 외국인들은 현재보다 미래에 케이팝 문화에 더 많은 관심을 가질 것으로 예측됩니다. 이는 우리가 제공하고자 하는 서비를 이용할 사람들이 많아진다는 의미입니다.

* 틈새는 어디인가?

그래서 외국인들이 케이팝 문화를 접할 때 어떤 어려운 점이 있는지 검토한 후 서비스 기획을 진행하게 되었습니다. 그리고 그 결론은 ‘**언어 장벽’**이었습니다.

외국인들이 흥미로운 케이팝 음악을 발견하거나, 좋아하는 스타의 얼굴만 아는 경우 검색해서 알아보는 일은 매우 번거롭습니다. 영어를 한글로 번역하는 작업이 선행되기 때문입니다.

우리 팀은 이러한 불편함을 해결할 수 있는 서비스를 만들 것입니다. 그리고 케이팝에 관심있는 외국인들이 언어장벽으로 좌절하지 않도록, 현재보다 더 쉽게 케이팝에 접근할 수 있는 웹 서비스를 제공할 것입니다.

1. 프로젝트 개요

# 01 프로젝트 요약

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 목적 | 외국인을 대상으로 한국을 알리는 서비스 제공 | | 목표 | 케이팝을 주제로 외국인들이 쉽게 이용하고 접할 수 있는 웹서비스 제작 |
| 제공  서비스 | * **사진으로 찾는 케이팝 스타** * 사진을 업로드하면 그 사진의 주인공이 어떤 스타인지 스타의 이름, 그룹 명, 앨범, 음악 등의 정보를 웹페이지로 보여주는 서비스 | | | |
| * **음악으로 찾는 케이팝 뮤직** * 알고 싶은 음악파일을 업로드 하면 그 음악의 제목, 가수, 가사 등의 정보를 웹페이지로 보여주는 서비스 | | | |
| * **한국어로 케이팝 스타에게 메시지 작성해보기** * 한국어를 모르는 외국인들이 영어로 메시지를 작성하면 한국어로 번역해주는 서비스 | | | |
| 기대 효과 | * 케이팝에 대한 관심이 증가하고 있으므로, 서비스 이용자가 앞으로 증가할 것이라고 예측 -> 광고 서비스 연계 * 케이팝을 좋아하는 외국인들이 더 쉽게 한국의 문화를 접하게 되어 연결 통로로서 한국을 알리는 서비스로의 성장을 기대 | | | |
| 주요  프로그램 설명 | 스타 사진 검색 | * 사진 업로드 기능 * 이미지 딥러닝 분석 기능 * 딥러닝 분석 결과 웹에 출력 기능 | | |
| 사운드 검색 | * 사운드 파일 업로드 기능 * 사운드 파일 딥러닝 분석 기능 * 딥러닝 분석 결과 웹에 출력 기능 | | |
| 한국어로 메시지 작성 | * 메시지 입력 기능 * 입력된 메시지 번역 기능 * 음원 녹음 후 분석 기능(미정) | | |

# 02 개발 환경

|  |  |  |  |
| --- | --- | --- | --- |
| **작업 환경 및 툴즈** | | | |
| OS | Windows10 64bit | IDE | Google Colab  Pycharm(2023.2.5) |
| Web Server | PythonAnywhere | DB | SQLlite |
| **주요 라이브러리** | | | |
| python | 3.10.0 | openai | 1.6.0 |
| Keras | 2.15.0 | gtts | 2.4.0 |
| tensorflow | 2.15.0 | numpy | 1.23.5 |

# 03 팀원 구성 및 일정

|  |  |  |  |
| --- | --- | --- | --- |
| **이름** | **역할** | **하위 프로젝트** | **업무** |
| 조민우 | 팀장 | 웹, DB, 메시지 | 팀의 대표, 웹 제작, 모델링과 서비스에 필요한 DB제작 |
| 조은겸 | 팀원 | 사운드 | 사운드 데이터 인코딩, 사운드 데이터 모델링 |
| 김수진 | 팀원 | 이미지 | 데이터 수집, 남자 이미지 모델링(모델명) |
| 윤정인 | 팀원 | 이미지 | 데이터 수집, 여자 이미지 모델링(모델명) |
| 서강석 | 팀원 | 이미지 | 데이터 수집, 이미지 모델링(모델명) |



2. 웹 서비스

# 01 서비스 소개

* + 동적 html : PC + 웹
  + 언어 : 영어
  + 서비스 제공 내용 : 스타 사진으로 이름 예측, 사운드 파일로 노래 제목 예측, 즉석 한국어 번역
  + 서버 : PythonAnywere

# 02 서비스 구조

메인 페이지인 Home에는 네비게이션 바 뿐만 아니라 페이지 하단에 영역에 각 서비스로 연결이 가능합니다. 메뉴의 항목을 클릭하면 페이지의 지정된 섹션으로 이동을 합니다. 섹션을 각 항목이라고 봤을 때, 총 4개의 항목이 있습니다.

[표 2-1]

|  |  |  |
| --- | --- | --- |
| **페이지명** | **기능** | **출력 화면** |
| Home | 메인 기능  네이게이션 기능 |  |
| Content | 케이팝 블로그  케이팝 스타들의 새로운 소식을 알리는 페이지 |  |
| Find Artist | 이미지의 스타가 누구인지 찾아주는 기능 |  |
| Find Music | 사운드 파일의 노래가 어떤 노래인지 찾아주는 기능 |  |
| Letter | 영어로 메시지를 작성하면 한국어로 변역해주는 기능  번역된 한국어를 음성으로 말해주는 기능 |  |
| Contact  팀원 소개 페이지 | 깃허브 링크 버튼  메일 보내내기 기능 |  |

3. 모델

# 01 데이터베이스

* SQLite
* Jango Models 셋팅
* Profile 테이블

[표 3-1] DB Profile 테이블

|  |  |  |
| --- | --- | --- |
| 변수명 | 타입 | 내용 |
| Ko\_name | CharField(max\_length=100) | 한글 이름 |
| name | CharField(max\_length=100) | 영어 이름 |
| birth | DateTimeField(auto\_now\_add=True) | 생일 |
| age | IntegerField() | 나이 |
| group | CharField(max\_length=100) | 그룹 |
| img\_url | TextField() | 이미지 |

* Music 테이블

[표 3-2] DB Music 테이블

|  |  |  |
| --- | --- | --- |
| 변수명 | 타입 | 내용 |
| title | CharField(max\_length=100) | 노래 제목 |
| group | CharField(max\_length=100) | 아티스트 |
| text | TextField() | 가사 |
| img | TextField() | 앨범 타이틀 이미지 |
| title\_upper | CharField(max\_length=100) | 타이틀 대문자 |

# 02 이미지 모델

이미지 모델의 종류가 다양하고, 그 모델의 입력 방식도 다양합니다. 본 프로젝트에서는 다양한 이미지의 모델들을 테스트하여 최종 1가지 모델로 서비스를 할 것입니다. Resnet계열, Deepface, 그리고 직접 구성한 CNN 모델을 사용했습니다. 사전에 수집된 이미지들을 google Drive의 공유 파일에 저장한 후 Colab Notebook에서 각 모델에서 요구하는 형식으로 처리해서 모델링을 했습니다.

* 데이터

[표 3-3] 이미지 모델 데이터 테이블

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **수집** | | | | |
| 출처 | 다음 검색 | | 구성 | 184명  1명당 500개 |
| 수집 방식 | API를 이용한 크롤링 | | 사이즈 | 20 GB |
| 저장된 위치 | Google Drive | | 파일명 규칙 | 그룹명\_이름.jpg |
| **처리** | | | | |
| 리사이징(크롭된) | (150 x 150), (160 x 160), (224 x 224), (152 x 152) | | | |
| 처리 | 샘플링 | 1명당 최종 평균 200개의 이미지 선택  500개 중에서 dlib 라이브러리를 사용하여 얼굴 인식이 안되거나 한 사진에 두 개 이상인 이미지는 삭제 | | |
| CV2를 사용해서 이미지에서 얼굴 영역만 크롭 | | |
| 채널 수 | 3 채널 (컬러) | | |

각 모델별로 다양한 튜닝을 통해 도출한 최적의 모델들 입니다. 각 모델에 맞게 이미지 데이터가 처리되었습니다. Deepface와 Resnet50은 사전학습된 모델로 facebook과 keras에서 제공하고 있습니다.

* 모델

[표 3-4] 이미지 모델 테이블

|  |
| --- |
| **테스트 모델 종류** |
| **Deepface** |
| * 딥 페이스 모델이란? * 2014년 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 에서 발표된 CNN 신경망. 특히 얼굴의 특징을 추출하고 학습 * 3차원 얼굴 모델링으로 얼굴을 정렬하고, 표준화된 이미지를 생성하여 인식의 정확도를 높임 * 핵심 기술   + 3D 얼굴 정렬 : 3D 모델을 사용하여 입력 이미지의 얼굴을 정렬. 다양한 각도에서 찍힌 얼굴 이미지도 효과적으로 처리   + 심층 학습 : 다층 컨볼루션 네트워크를 통해 얼굴의 복잡한 특징들을 학습하고 이를 기반으로 얼굴 인식을 수행   + 성능 : 사람과 유사한 수준의 얼굴 인식 정확도를 달성 * 학습 데이터 : 4백만개 이상 얼굴 이미지를 포함하는 대규모 데이터셋을 사용하여 훈련 * 정확도   + 훈련 : 73%   + 검증 : 65% |
| **CNN** |
| * 인간의 시신경 구조를 모방한 딥러닝 기술 / 주로 이미지와 같은 2D 데이터를 처리하는데 특화된 모델 * 합성곱 연산을 통해 특정 맵을 생산하는 필터를 학습할 수 있어 컴퓨터 비전 분야에서 우수한 성능을 보이고 있다. 이미지를 인식하기 위해 주로 패턴을 찾는데 활용되며 데이터를 입력 받아 패턴을 추출하고 이를 기반으로 이미지를 분류한다. * 본 프로젝트에서는 CNN을 랜드마크 모델과 결합하여 사용 * 이미지 모델   + Input shape : (150, 150, 3)   + 활성화 함수 : relu * 랜드마크 모델   + Input shape : (136, )   + 활성화 함수 : relu * 결합 모델   + [(34, 34, 128), (136, )]   + 활성화 함수 : softmax   + 옵티마이저 : SGD   + 훈련 정확도 : 99%   + 검증 정확도 : 82%   + 테스트 정확도 : 82% |
| **Resnet50** |
| * Resnet50이란? * 2015년 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 에서 발표된 딥러닝 알고리즘 * 핵심 개념   + 잔차 학습 : 출력에서 입력을 빼는 대신 입력을 출력에 더하는 방식으로, 네트워크가 학습해야 할 잔차 함수를 단순화   + 학습의 용이성 : 깊은 네트워크에서도 안정적으로 학습 가능   + 성능 개선 : 기존의 깊은 네트워크 대비 오류율을 상당히 줄임 * **주요 특징**   + 잔차 블록 : 이 블록들은 입력을 출력에 직접 추가하여 네트워크가 잔차 함수를 학습하게 함으로써 깊은 네트워크에서의 학습 문제를 해결   + 스킵 연결 : 각 잔차 블록은 입력을 블록의 출력에 더함으로써 스킵 연결을 제공합니다. 이는 깊은 네트워크에서의 그래디언트 소실 문제를 줄여줌 * 적용 모델   + 선정 이유 : 50개의 균형 잡힌 깊이, 중간 규모의 네트워크임에도 불구하고 고도로 정확한 결과를 제공   + 남자 모델에 적용 * **정확도**   + **훈련 : 99%**   + **검증 : 79%** |

# 03 사운드 모델

사운드 모델의 초기 목표는 원음 또는 사람의 흥얼거림을 입력 받아 제목을 예측해 주는 것이었습니다. 그러나 여러 번의 시도에도 사람의 흥얼거림은 예측하기 어려웠습니다. 개선하는 과정에서 잡음 데이터를 추가하여 훈련시켰는데 개선점이 있어서 프로젝트의 결과물로 내놓을 수 있었습니다.

최종으로 선택된 ‘Augmented\_conv\_ver3’모델은 과적합이 해결되었고 앞으로 모델을 개선시킬 방향성을 보여주었습니다. 정확도가 80%대에 머무른 이유에 대해서는 훈련에 사용된 자료가 부족했거나 다양하지 못했기 때문이라는 잠정적 결론을 내놓겠습니다.

* 데이터

[표 3-5] 사운드 모델 데이터 테이블

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **수집** | | | | |
| 출처 | Youtube | | 구성 | 285개의 케이팝 음원 |
| 수집 방식 | API를 이용한 크롤링 | | 사이즈 | 20 GB |
| 저장된 위치 | Google Drive | | 파일명 규칙 | 가수명\_노래제목.wav |
| **처리** | | | | |
| 추출 | Librosa를 이용해서 비정형 데이터에서 배열 데이터(정형)를 추출 | | | |
| 처리 | 샘플링 | 원곡 : 1곡당 300개씩 샘플링  보컬 : 1곡당 100개씩 샘플링 | | |
| 스케일링 | 스탠다드 스케일링 : 개별 데이터의 특징마다 적용 | | |
| 잡음 처리 | 스팩클 잡음 처리 : 과대적합 방지를 위한 처리 | | |
| 샘플 사이즈 | 500 x 50  (배열) | 특징 : 50가지  한 구간 단위 : 500 프레임 | | |

* 모델

모델링은 2-step으로 진행되었습니다. 스텝은 데이터 셋 변화를 기준으로 했습니다. 첫 번째 스텝에서는 ‘원곡 285곡 (각 150샘플)’과 ‘보컬 285곡 (각 50샘플)’을 이용했습니다. 사실 샘플링은 원곡 300개, 보컬 100개로 샘플링을 진행했으나 데이터의 크기가 커져서 분석에 어려움이 있었습니다. 그래서 과감하게 1/2로 다시 샘플링하여 2-step으로 모델링을 진행했습니다.

첫 번째 스텝에서는 데이터 셋을 그대로 활용했고, 두 번째 스텝에서는 기존 데이터셋에 잡음을 추가한 데이터셋을 추가했습니다. 결과는 모델에 따라 상이했습니다. CNN모델은 두 번째 스텝의 모델이 과대적합이 해소되면서 더 좋은 모델로 선정되었고, Mobilenet의 경우 첫 번째 스텝에서 더 좋은 결과를 보여주었습니다. 아래 표에서 CNN모델은 두 번째 스텝의 세 번째 버전이고, Mobilenet은 첫 번째 스텝의 첫 번째 버전입니다.

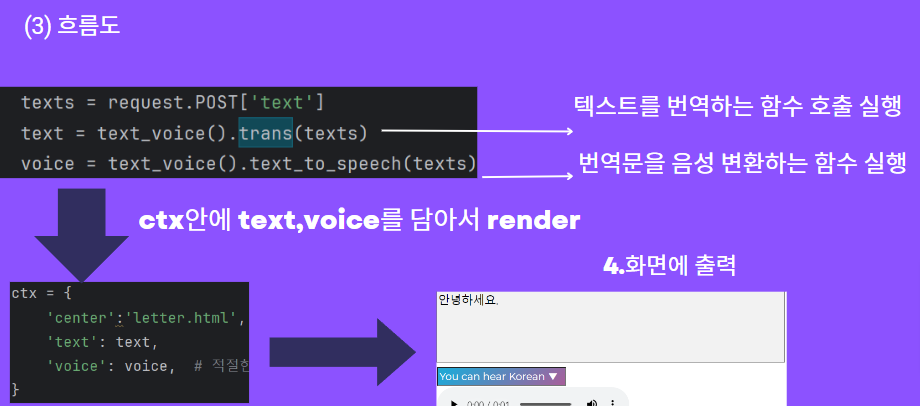
[표 3-6] 사운드 모델 테이블

|  |  |  |  |
| --- | --- | --- | --- |
| **테스트 모델 종류** | | | |
| CNN(컨볼루션 신경망)[부록] | | Mobilenet[부록] | |
| 구성 | 3층 컨볼루션 | 모바일넷 | - |
| 잡음 | 포함된 | 잡음 | 포함되지 않은 |
| Input shape | (500, 50, 1) | Input shape | (500, 50, 1)  1채널(그레이 스케일)의 이미지로 간주 |
| 활성화함수 | Leakyrelu | 활성화함수 | Relu |
| 확산 대비 | 각 컨볼루션 층마다 배치 정규화 | 훈련 시간 | Epoch당 30초 이내 |
| 총 파라미터 수 | 569,869 (2.17 MB) | 총 파라미터 수 | 4,570,013 (17.43 MB) |
| 훈련 정확도 | 82% | 훈련 정확도 | 99% |
| 검증 정확도 | 81% | 검증 정확도 | 93% |
| 테스트 정확도 | 86% | 테스트 정확도 | 92% |
| 정확도가 모바일넷과 비교해서 높지는 않지만 과대적합을 피했습니다. 튜닝도 비교적 자유로워서 앞으로 모델을 개선할 여지가 많이 있는 것으로 판단, 최종 모델로 선정했습니다.  모델은 작은 특징(커널 사이즈 10)부터 시작해서 사이즈가 증가하는 방향으로 특징을 잡고 최대(커널 사이즈 30)로 확장하는 방식으로 구성했습니다. | | 훈련 정확도부터 테스트 정확도까지 90%가 넘지만 과대 적합이고, 더 이상의 튜닝이 가능하지 않기 때문에 이 문제를 개선하기 어려웠습니다. 또 다른 이미지 모델들에 비해 경량이라고는 하지만 직접 제작한 컨볼루션 모델보다는 예측하는데 시간이 오래 걸리고 사이즈가 크다는 단점이 있어서 현 시점에서는 서비스 모델로는 적합하지 않다고 결론 내렸습니다. | |

# 03 메시지 모델

* 모델의 원리
  + 메시지 입력 : texts (영문으로 작성된 메시지)
  + 함수1 trans(texts) = 번역문 리턴 (‘text’)
  + 함수2 text\_to\_speech(texts) = 음성 리턴 (‘voice’)

[그림 4-1] 메시지 모델 흐름도



메시지 모델은 간단한 구성을 갖고 있습니다. 사용자는 영어만 하는 사람이라고 가정합니다. 사용자가 폼에 영문으로 메시지를 작성하면 ‘파파고’와 ‘TTS’를 함수를 사용하여 번역된 문장과 그것을 이용한 음성까지 화면에서 출력됩니다. 다음은 간단한 작동 과정입니다.

* 사용 예시
  + 1. John은 블랙핑크의 지수를 좋아합니다. 그래서 지수에게 한국어로 메시지를 보내고 싶습니다.
    2. ‘Find my k-pop star’사이트에 접속해서 영문으로 ‘Hi, I’m your big fan. Nice to meet you.’ 라고 작성하고 ‘submit’버튼을 클릭합니다.
    3. 잠시 뒤, ‘안녕하세요, 저는 당신의 열렬한 팬입니다. 만나서 반가워요.’라는 문구가 바로 다음 폼에 작성됩니다.
    4. 그리고 아래 방향으로는 번역된 한국어를 사운드로 재생할 수 있는 재생 버튼이 열립니다. 버튼을 클릭하면 출력된 문구가 재생됩니다.

다음 표는 작동에 사용된 API와 관련 기술입니다.

[표 3-7] 메시지 모델에 사용된 API와 함수

|  |
| --- |
| **파파고** |
| 파파고는 다국어 언어 처리에 대한 네이버의 기술과 경험을 번역 엔진에 적용해 보다 정확한 번역 결과를 제공하는 서비스입니다. 파파고가 제공하는 RESTful 형태의 API를 사용하면 서비스에 번역 기능을 간단하게 적용할 수 있습니다.  파파고가 제공하는 API는 다음과 같습니다.   * + Papago 번역: 파파고의 인공 신경망 기반 기계 번역 기술(NMT, Neural Machine Translation)로 텍스트를 번역한 결과를 반환하는 RESTful API입니다.   + 언어 감지: 입력된 텍스트의 언어를 감지해 주는 RESTful API입니다.   + 한글 인명-로마자 변환: 한글로 된 이름을 로마자 표기로 변환한 결과를 반환하는 RESTful API입니다. |
| **TTS(Text-to-Speech)** |
| TTS는 텍스트를 음성으로 변환하는 기술입니다. 한국어도 사용 가능합니다. TTS시스템은 다음과 같은 두 가지 주요 구성 요소로 이루어져 있습니다.   * + 텍스트 처리 : 입력된 텍스트를 읍성으로 변환하기 위한 처리입니다. 여기에 텍스트 정규화, 문장 부호 처리, 그리고 발음 변환 등이 포함됩니다.   + 음성 합성 : 이 단계에서는 처리된 텍스트를 실제 음성으로 변환합니다. 자연스럽고 다양한 음성 톤과 감정을 표현할 수 있습니다. |

4. 개선 사항 및 확장 방향

# 01 모델의 품질 업그레이드

현재는 프로젝트의 초기 버전으로 모델이 완전하지 않습니다. 현 시점을 버전 1이라고 상정한다면 버전 2에서는 이미지 모델, 사운드 모델의 정확도를 더 올리는 작업이 진행되야 합니다. 버전1에서 얻은 모델링 인사이트에서 버전2의 방향성을 정리했습니다.

[표 4-1] 모델 ver2에 대한 고찰

|  |  |  |
| --- | --- | --- |
| 개선 요구사항 | 이미지 모델 | 사운드 모델 |
| 정확도 | 더 많은 훈련용 데이터 수집 필요  훈련용 데이터 사전 처리의 다양화 시도(스케일링, 그레이스케일 등)  사전학습된 모델 이외에 직접 구성한 알고리즘으로 심도 있는 튜닝 | 다양한 사운드 자원을 수집(콘서트 환경, 매장에서 흘러나오는 음악, 다양한 사람들의 보컬 등) |
| 텐서의 사용 | 이미지 처럼 대용량 파일의 경우 텐서를 이용하면 모델링을 하는데 gpu 사용의 효율을 올릴 수 있습니다. 다음 버전부터는 텐서를 사용해 볼것으로 예상합니다. | 배열로 딥러닝을 계산하는데 불편한 점들이 있었습니다. 채널을 조절하는 과정 등, 텐서를 사용하면 이런 점들이 개선될 것으로 예상합니다. |

# 02 웹의 확장

현재 서비스의 반응을 참고하여 웹 서비스를 개선할 수 있습니다. 디자인적 요소, UI, 그리고 더 개선된 환경의 서버를 찾을 수도 있을 것입니다. 그러기 위해서는 현제 버전의 서비스에 관심을 갖고 지속적으로 개선하고자 하는 의지가 필요하고 아쉬운 점들을 하나씩 해결해 나갈 것입니다.

# 03 서비스에 대한 고찰

* 서비스로 수익 창출 가능성

본 프로젝트는 결과물을 웹으로 표현합니다. 웹으로 표현한다는 것은 다수의 사용자가 서비스를 이용할 수 있다는 뜻이기도 합니다. 그리고 서비스할 모델들은 시간과 인적 자원이 들어간 가치 있는 결과물입니다. 그러나 우리는 이 것들을 돈을 받고 운영하고자 하는 의도는 없습니다. 대신에 사용자가 증가하면서 페이지에 광고를 달 수 있고 거기서 수익을 창출할 가능성이 있다고 기대합니다.

5. 참고자료