2024 사이버보안트랙 캡스톤 디자인 최종보고서

**음성 위변조 탐지 포렌식**



지도교수: 김승천

팀명 : 거북이

팀장 : 2171529 신서형

팀원 : 2171262 박소희

2071480 남은희

1971323 이기정

**목차**

**Ⅰ. 프로젝트 소개** ........................................................... 3

프로젝트 개요 ........................................................................................................................ 3

프로젝트 배경 ........................................................................................................................ 3

유사 프로젝트와의 차별성 ............................................................................................... 5

**Ⅱ. 프로젝트 결과물** ...................................................... 6

시스템 구조도 .......................................................................................................................... 6

주요 기능 ................................................................................................................................... 7

사용 기술 .................................................................................................................................. 10

구현 결과 ..............................................................................................................................…. 12

**Ⅲ. 결론** ............................................................................ 20

기대 효과 .............................................................................................................................…. 20

개발 이슈 및 해결방안 ........................................................................................................ 21

프로젝트 수행 일정 .............................................................................................................. 25

수행 후기 …………………………………………………………………………………….................... 26

지원금 집행 내역 .................................................................................................................. 27

**Ⅳ. 참고자료 및 소스코드 URL ........................ 28**

Ⅰ. 프로젝트 소개

1. 프로젝트 개요

AI 기술이 발전하면서 음성변조 기술 또한 더 섬세하고 정교해지고 있다. 이러한 기술의 발전을 악용하여 법적 증거로 사용될 수 있는 녹취록 신뢰성 문제가 발생할 수 있다. 제출한 음성 파일이 변조 의심을 받는 상황에서 녹취록과 현장 녹음본의 음성 특징 정보를 추출하고, 획득한 정보를 토대로 딥러닝 학습을 통해 유사도를 판단하면 녹취록의 신빙성을 상승시킬 수 있다.

1. 프로젝트 배경

텍스트, 스크린샷, 폰트, 라인이(가) 표시된 사진

자동 생성된 설명

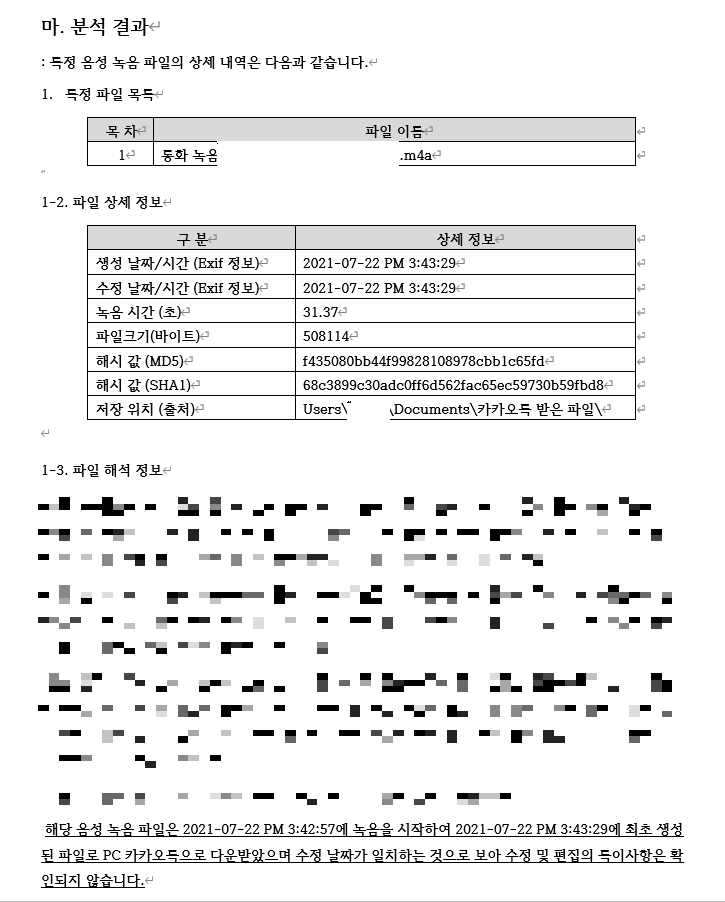
1. AI기술을 활용한 부정적인 사례

AI 기술의 발달로 인해 지인의 목소리를 변조하여 돈을 요구하는 보이스피싱이 등장했다. 가족이나 지인의 목소리로 돈을 빌려 달라는 요청에 별다른 의심없이 송금했지만, 확인해보니 지인의 목소리를 변조한 보이스피싱이었다.[[1]](#footnote-2) 현재의 음성 변조 기술은 단순 모방을 넘어서 진짜와 가짜를 구분하기 어렵고 핸드폰에 저장 되어있는 음성 사서함 메세지만으로도 쉽게 변조를 할 수 있 피해자들이 속수무책으로 당할 위험이 높아지고 있다.

1. 녹취록의 증거 채택 가능성

공개되지 아니한 타인 간의 대화를 녹음한 청취록은 증거능력이 없다. 이는 청취록의 신뢰성 문제(편집, 변조의 가능성), 압박이나 협박에 의한 녹음일 가능성 등이 있기 때문이다. 신뢰성 문제를 제외한 나머지를 수사를 통해 제거한 상태에서, 음성 파일을 통해 본인 여부를 입증할 수 있는 시스템이 개발된다면 녹음 파일은 해당 인물(용의자)의 발언이나 행동을 증명하는 데 필요한 신뢰할만한 증거로써 법정에서 사용될 수 있는 가능성을 열어준다.

1. 유사 프로젝트와의 차별성



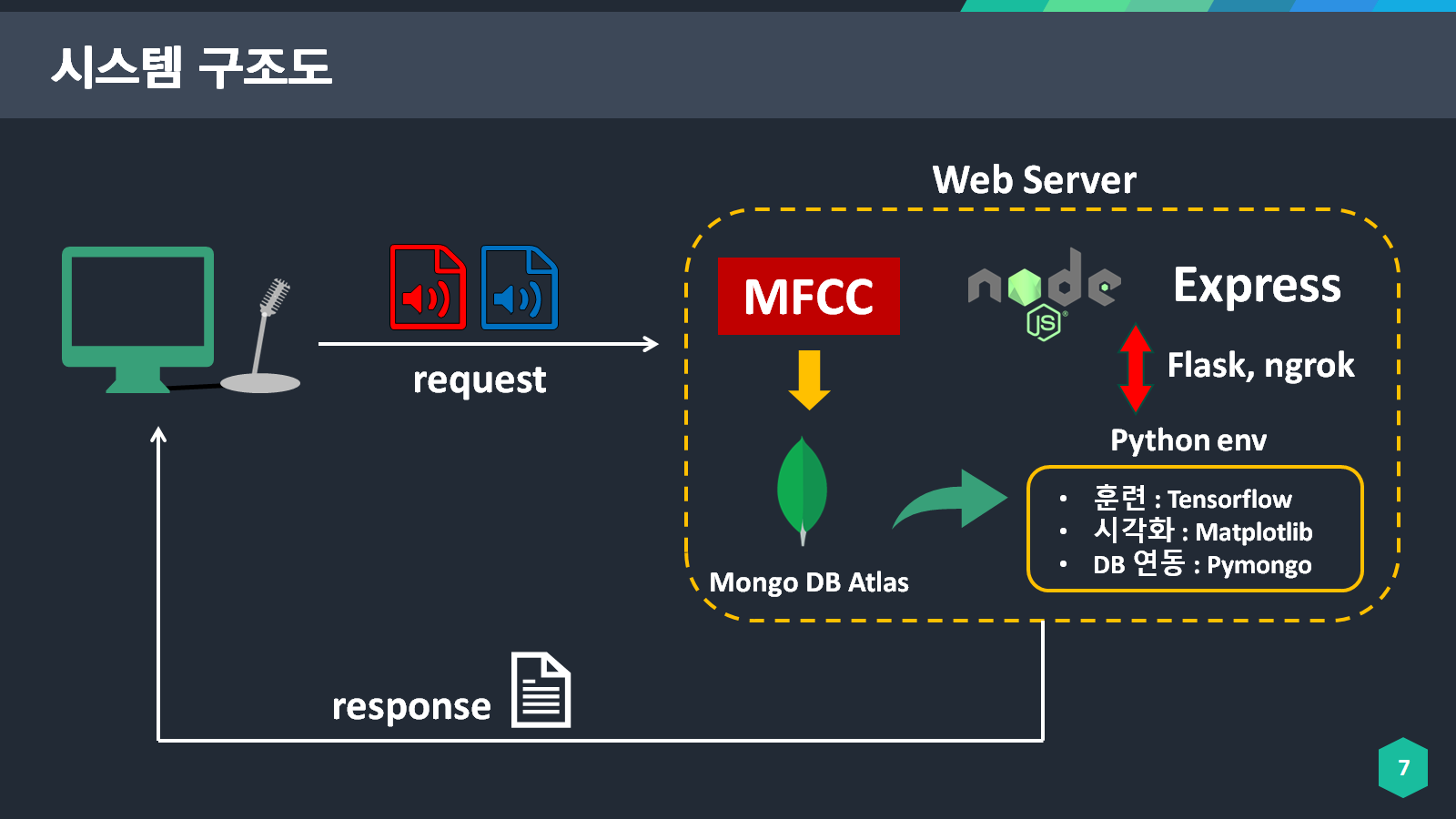
포렌식 기업에서 공개한 포렌식 분석결과지[[2]](#footnote-3)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 해석 정보 | 시각화 자료 | 자료 설명 | 포렌식 결과 | 인공지능 |
| 기존 프로그램 | O | X | X | O | X |
| 음성 위변조  탐지 포렌식 | O | O | O | O | O |

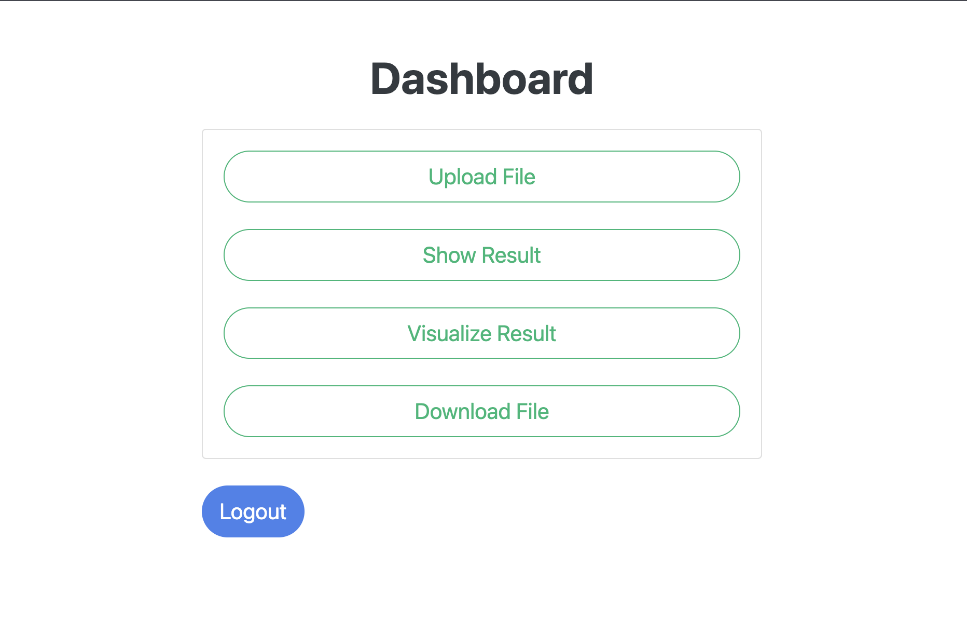
기존의 음성 파일 프렌식 분석 결과에서는 음성 변조 여부만을 판단할 수 있다. 반면 이 프로젝트는 딥러닝 모델을 활용하여 두 음성 간의 유사도를 평가하여 변조 가능성을 판단한다. 이를 통해 보다 정교하고 신뢰할 수 있는 결과를 제공할 수 있다. 더 나아가 분석 결과를 시각화하여 이미지로 첨부하고, 자료 아래에 추가적인 설명을 함께 제공하여 단순히 결과에 의존하지 않고 의뢰자 또한 포렌식 결과를 쉽게 이해할 수 있도록 구성한 점이 기존 프로그램과의 차별성이다.

Ⅱ. 프로젝트 결과물

1. 시스템 구조도

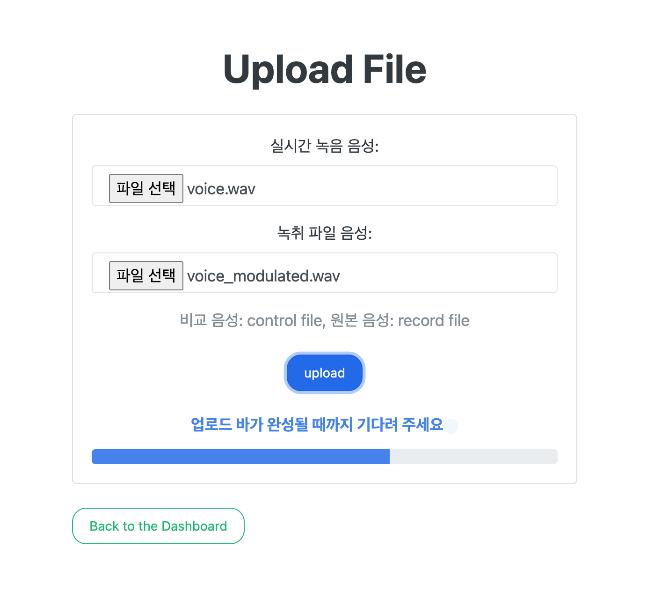


2. 주요 기능



1. **음성 포렌식 홈페이지 제작**

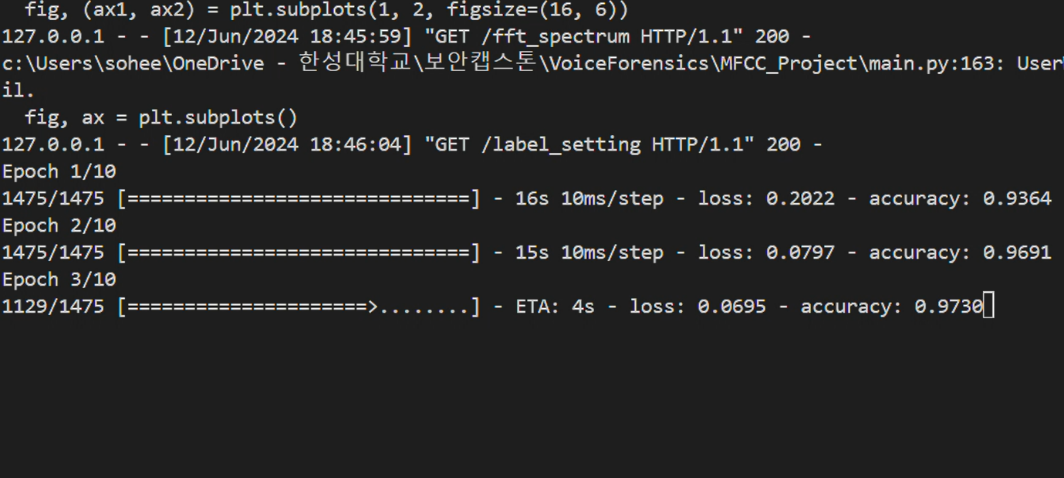
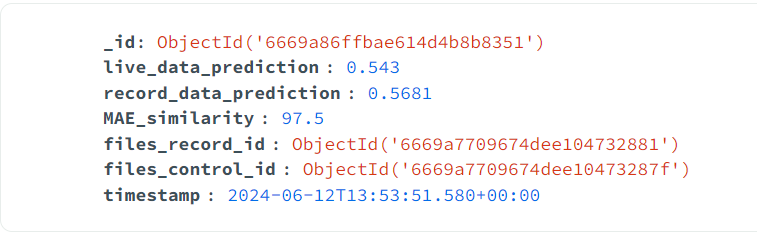
로그인 및 회원가입 기능을 구성하여 세션 기록 및 외부인 접근 통제가 가능하다. 추가적으로 파일 업로드, 결과 간단히 보기, 결과 시각화, 보고서 다운로드 기능을 구현하여 사용자에게 편리한 인터페이스를 제공한다.



1. **MFCC 계수 추출을 통해 화자의 발음 특성에 기반을 둔 음성 특징 값 추출**  
   음성 파일을 업로드하면 해당 음성 파일에 대한 MFCC 계수를 추출할 수 있다. 실시간 녹음 음성이란 피의자가 공개된 환경에서 녹음한 신뢰성 있는 음성 파일을 말하며, 변조 가능성이 존재하지 않는다. 녹취 파일 음성이란 피해자가 녹음한 피의자의 음성 녹취 파일이며, 변조 가능성이 있어 주의를 기울여야 한다.  
   MFCC 계수 추출을 위해 node-MFCC 라이브러리를 활용하였다. node-MFCC 라이브러리는 FFT(Fourier Transform), Mel Filter Bank, DCT(코사인 변환) 등의 과정을 거쳐 .wav 확장자의 음성 파일을 프레임 단위의 벡터 형태의 계수로 변환한다.
2. **녹취록/실시간 음성의 MFCC 계수 딥러닝 학습 및 예측/유사도 평가**  
   녹취록/실시간 음성의 화자가 동일 화자인지의 여부를 모르기 때문에 MFCC 계수들을 2차원 평면 상에 분포시킨 후 거리에 따라 2개의 군집으로 그룹화하여 레이블을 부여하는 클러스터링 알고리즘을 사용하였다. Tensorflow framework를 이용하여 Dense Layer를 순차적으로 쌓은 분류 모델을 구축하였고 마지막 Layer의 활성화 함수는 Sigmoid 함수를 이용하여 모델의 출력 예측값이 0~1 사이의 확률값으로 나타나도록 구현하였다. 훈련 횟수(epoch)은 10번, 손실함수는 이진 분류 모델이기 때문에 Binary Cross Entropy를 이용하였다.

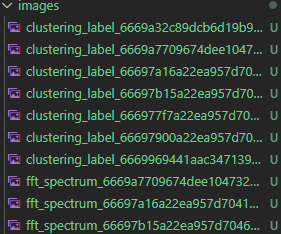
0~1 사이의 실시간/녹취록 테스트 MFCC 데이터셋에 대한 모델의 예측값을 이용하여 MAE(평균절대오차) 유사도를 계산하였는데 그 수식은 아래와 같다.

control\_prediction은 실시간 MFCC 테스트 데이터셋에 대한 모델의 평균 예측값, record\_prediction은 녹취록 MFCC 테스트 데이터셋에 대한 모델의 평균 예측값을 의미하며 두 값의 오차를 절댓값으로 표현하여 1에서 해당 오차를 뺀 값으로 유사도를 계산하였다. 여기서 1은 즉 오차의 절댓값이 최소 0, 최대 1의 값을 가질 수 있으므로 모델 예측값의 범위가 0~1 라고 가정할 경우 설정한 상수값이다. 위처럼 모델의 예측값과 유사도를 계산한 뒤, 해당 값들을 DB의 results 컬렉션에 저장하였다.



1. **시각화 그래프 구현 및 저장**

보고서에 첨부한 시각화 그래프는 모델의 정확도 및 손실 그래프, 클러스트링 분포도 그래프, MFCC 스펙트럼 그래프, FFT 주파수 대역 스펙트럼 그래프, MFCC 평균값 가시화 막대 그래프 이렇게 5개이다. 시각화 툴은 Python의 matplotlib 라이브러리를 이용하였고 이미지 저장은 Python 환경과 Node.js 환경이 같은 디렉토리에 있다는 점을 활용하여 프로젝트의 images 디렉토리에 저장되도록 하였다. 이미지 저장 시에는 file\_control\_id, 즉 실시간 음성의 id를 이미지 이름에 포함시켜 이후 보고서에서 이미지를 불러올 때, id를 참조하여 불러올 수 있도록 구현하였다.



1. **통계 자료 시각화 및 보고서 다운로드**



딥러닝 훈련 결과에 대한 시각화 페이지를 제공한다. 페이지는 모델 예측 결과, 음성 파일 간 유사도뿐만 아니라 일반인들도 이해할 수 있도록 MFCC 계수별 설명과 각종 시각화 그래프를 포함하고 있다. 홈페이지의 ‘Download File’ 버튼을 누르면 보고서를 pdf 형식으로 다운로드받을 수 있다.

3. 사용 기술

1. Express 프레임워크를 활용한 node.js 환경의 웹서버 구축
2. Wed socket을 활용하여 sever측에서의 진행 상황을 client에게 실시간으로 전달
3. MFCC 기술을 활용한 음성 특성 정보 추출
4. MFCC 계수별 설명  
   **MFCC1**:음성 신호의 전반적인 에너지 수준을 반영한다. 해당 계수를 통해 화자의 감정 상태, 강조 등을 측정할 수 있다.   
   **MFCC2:** 음성의 스펙트럼을 나타낸다. 음성 신호의 에너지가 집중되어 있는 위치를 알려준다. 이는 발성자의 성대의 조음 위치와 밀접한 관련이 있어, 음성의 높이변화를 포착하고, 발화자의 특징을 유추하는데 사용할 수 있다.  
   **MFCC3**: 음성의 특정 발음 특성을 반영하여 특정 모음의 발음 방식을 포착한다. MFCC3가 높은 값을 가지면 특정 발음 패턴이 강조된다는 것을 의미한다.  
   **MFCC4**:음성의 중간 주파수 대역에서 복잡성을 알려준다. 이는 발음의 강조 또는 음성의 변조가 포함되어 있을 수 있다.  
   **MFCC5**: 음성의 중주파수 범위에서의 세부적인 특성을 잡아낸다. 이 특징을 바탕으로 발음의 섬세함이나 발화의 독특함을 알 수 있다.  
   **MFCC6**: 음성 신호에서 더 미세한 중주파수 세부사항을 포착한다. 주로 음성의 흐름에서 복잡성이나 특정 발음의 빈번한 전환을 나타낼 수 있다.  
   **MFCC7**: 중간 주파수 대역의 최고 세부사항을 나타내며, 특히 발음의 미묘한 변화와 감정, 강조표현 등을 알 수 있다.  
   **MFCC8**: 고주파수 대역에서의 세부 변화를 포착하는데, 특히 입술과 혀의 빠른 움직임이나 미세한 발음의 변화를 반영하여 발음의 빠르기를 나타낸다.  
   **MFCC9**: 음성의 매우 높은 주파수 영역에서의 변화를 감지하여, 해당 계수값을 통해 발음의 섬세함 정도를 파악할 수 있다.

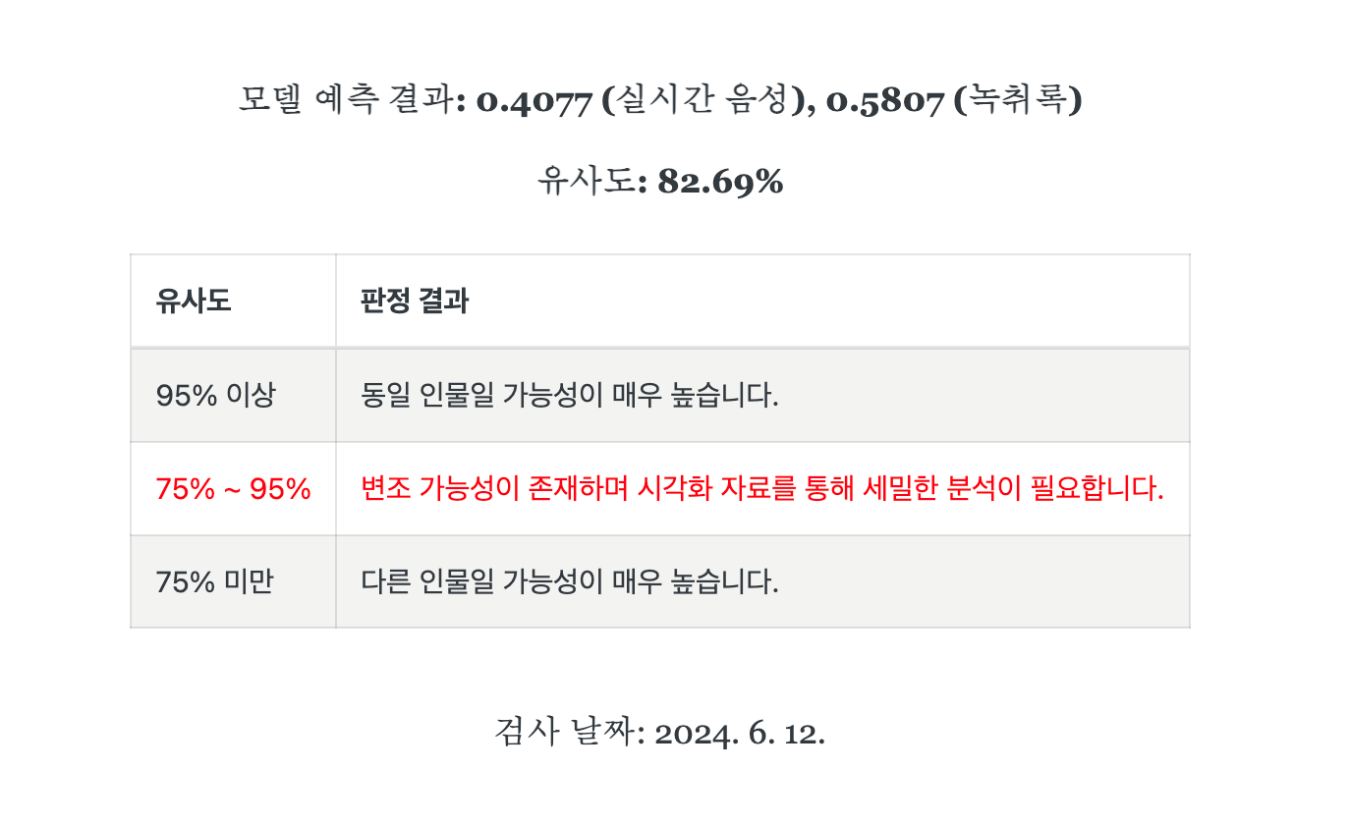
**MFCC10, MFCC11**: 음성의 가장 미세한 주파수 변화를 포착하고, 특히 높은 음역대에서의 발음 세부사항을 잘 나타내며 말하는 사람이 음성의 피치나 높이를 알 수 있다.

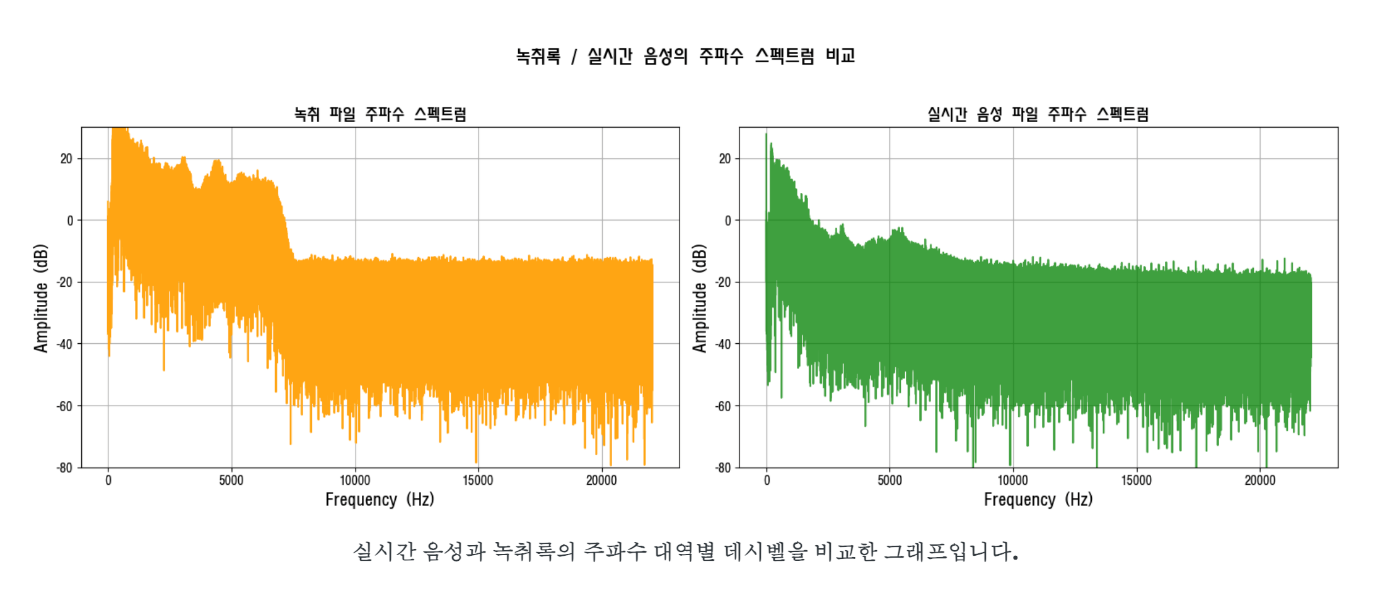
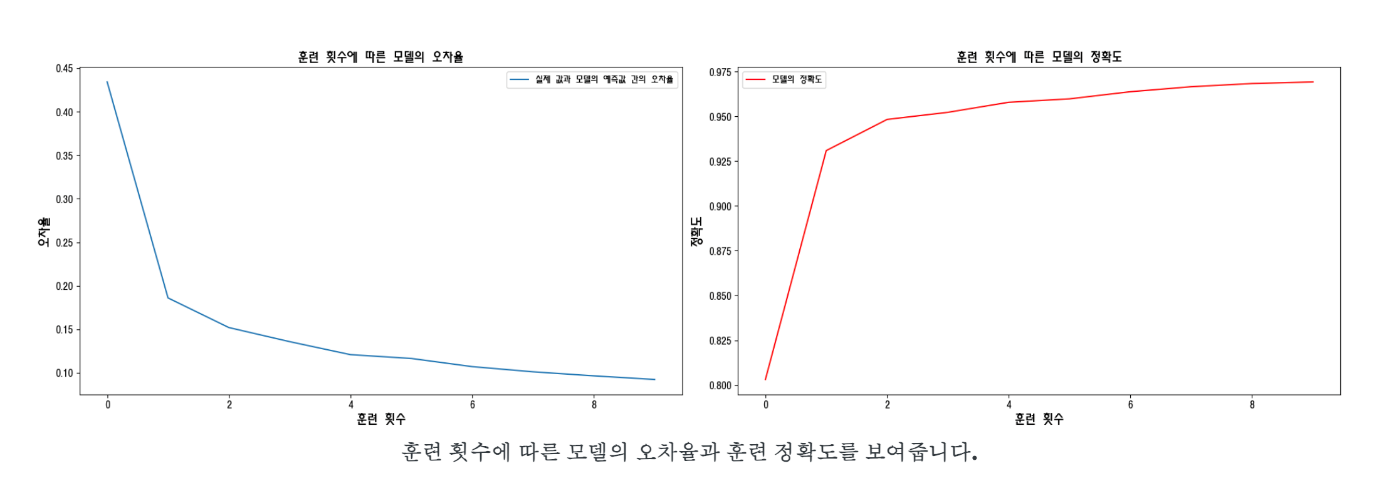
**MFCC12**: 음성의 가장 높은 주파수에서의 세밀한 특징을 포착해 음성의 질감, 미세한 발음의 변화, 그리고 발음의 매우 섬세한 부분을 포착한다.

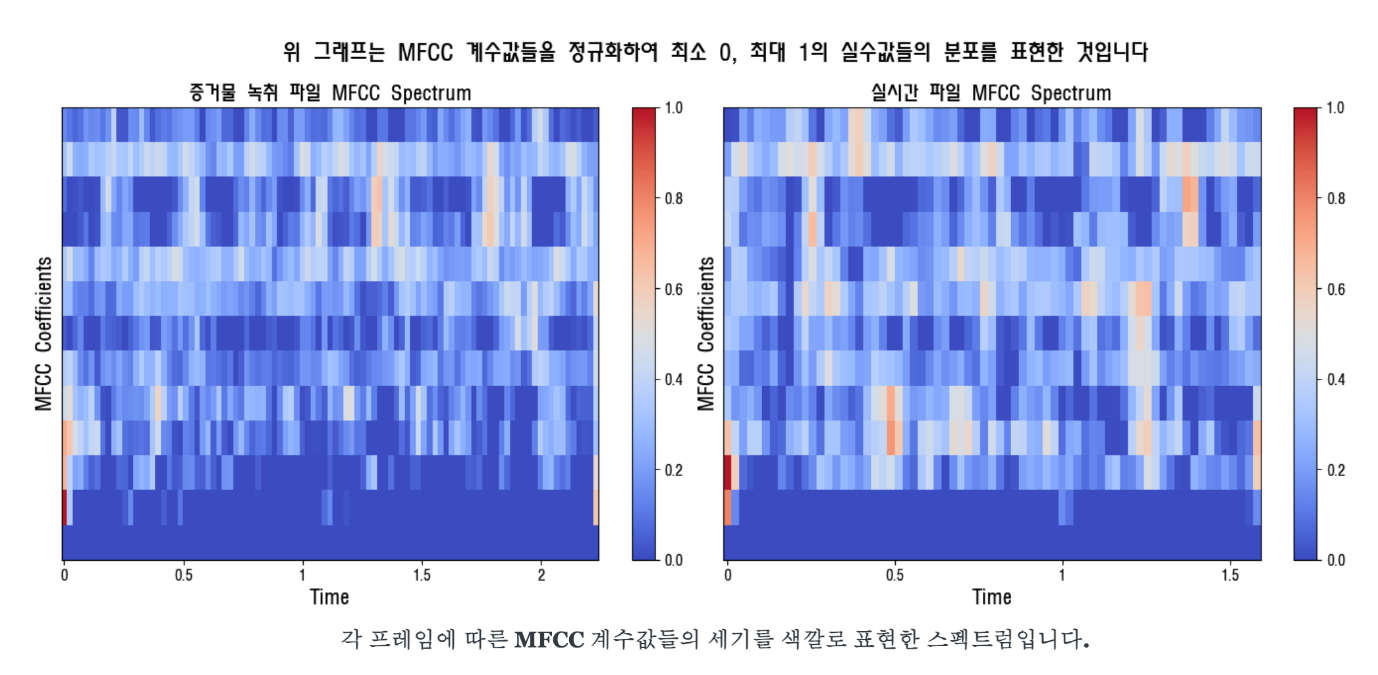
1. MFCC 계수 추출 및 모델 예측 결과값 등을 저장하는 mongoDB 클라우드 활용
2. 클러스트링을 통한 레이블을 부여 및 딥러닝 모델 훈련/예측
3. Python과 node.js 환경을 ngrok 터널링을 통해 연동
4. MFCC 스펙트럼, 딥러닝 모델 훈련 통계 자료 등을 포함한 보고서 작성 구현

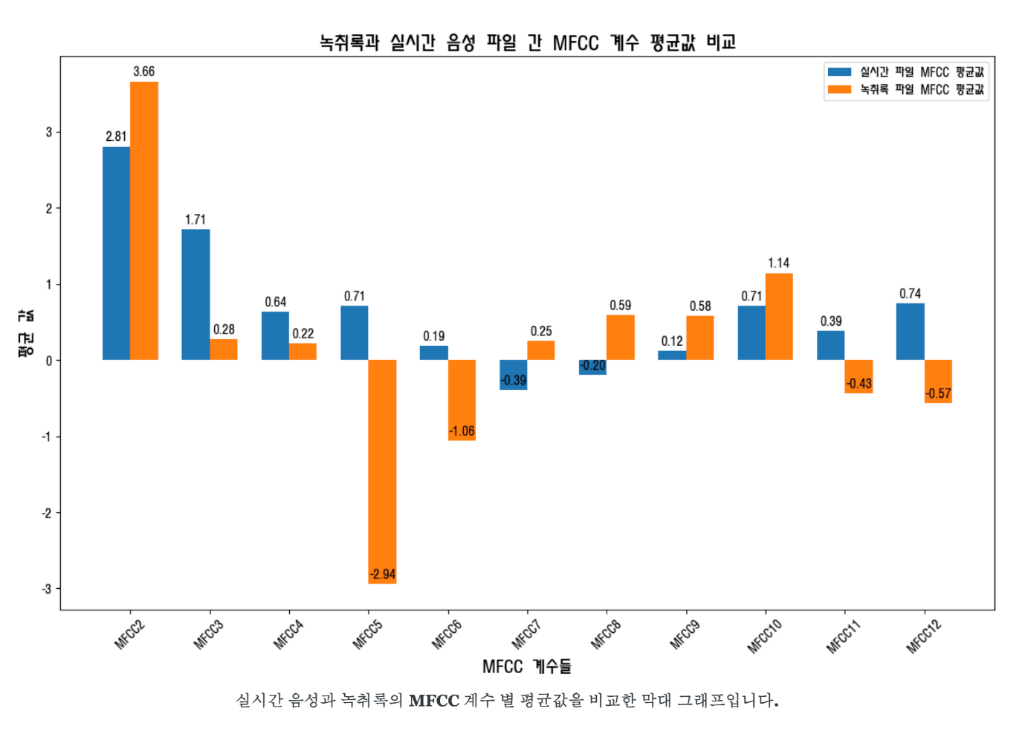
4. 구현 결과

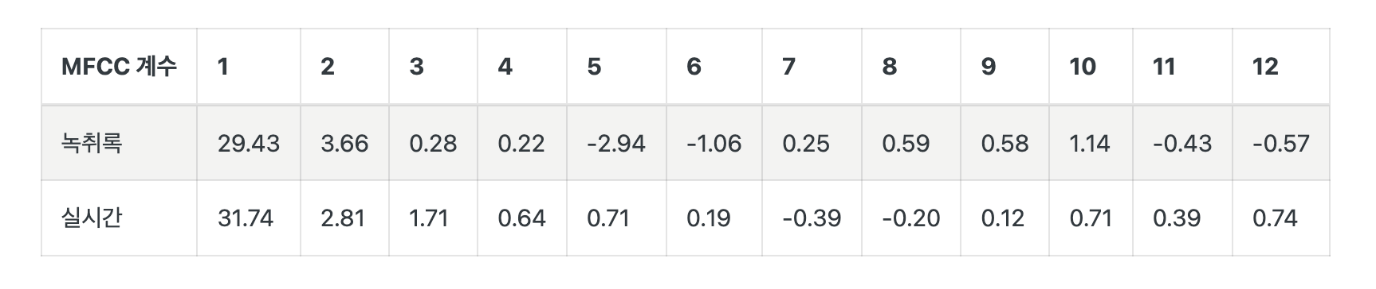
**1) 원본 음성(여성) vs 원본 AI 변조 음성**



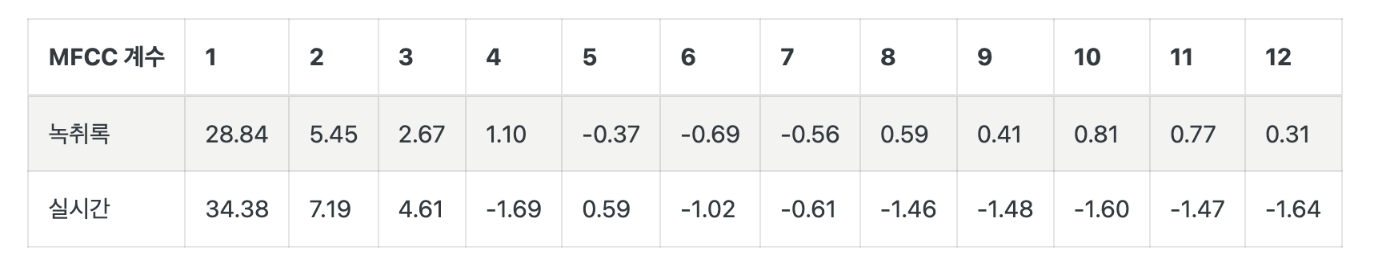
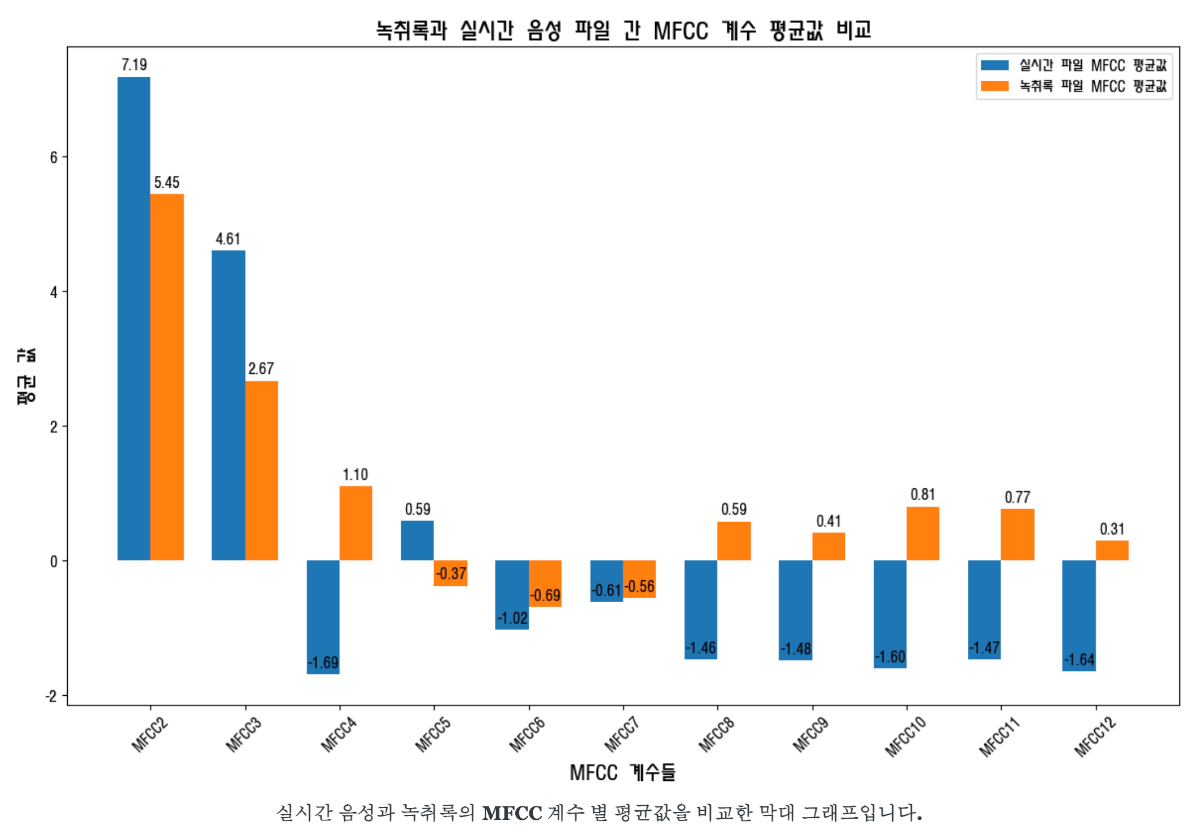
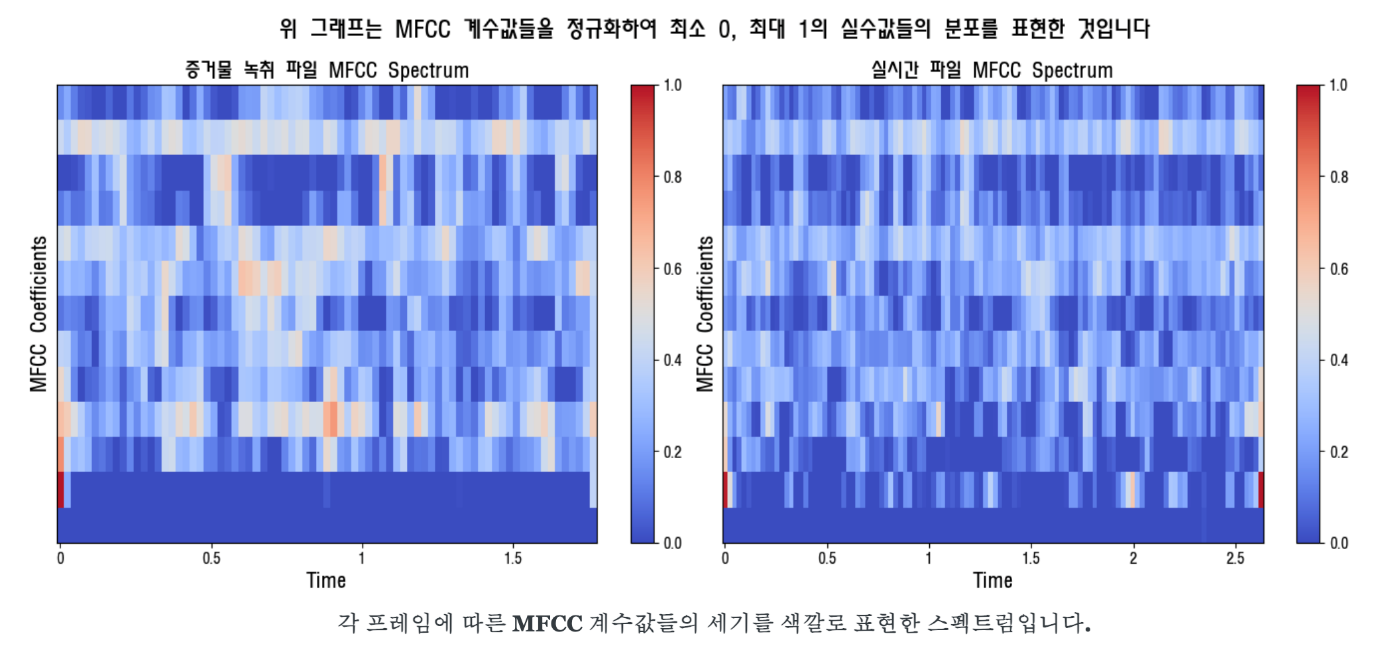
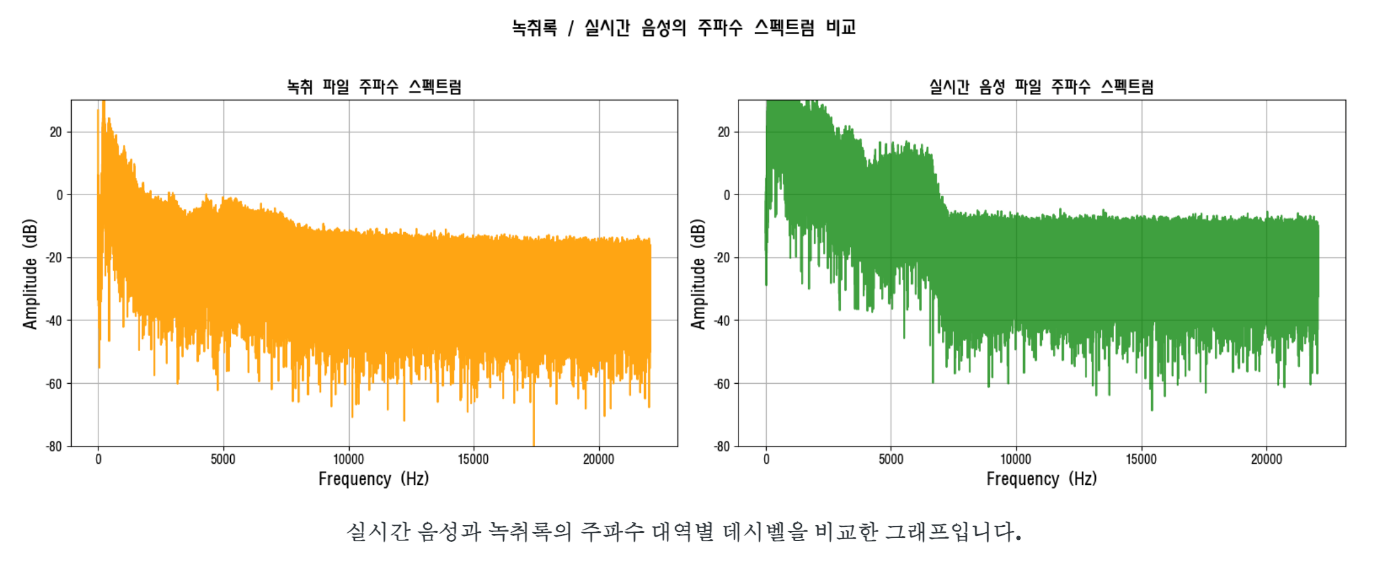
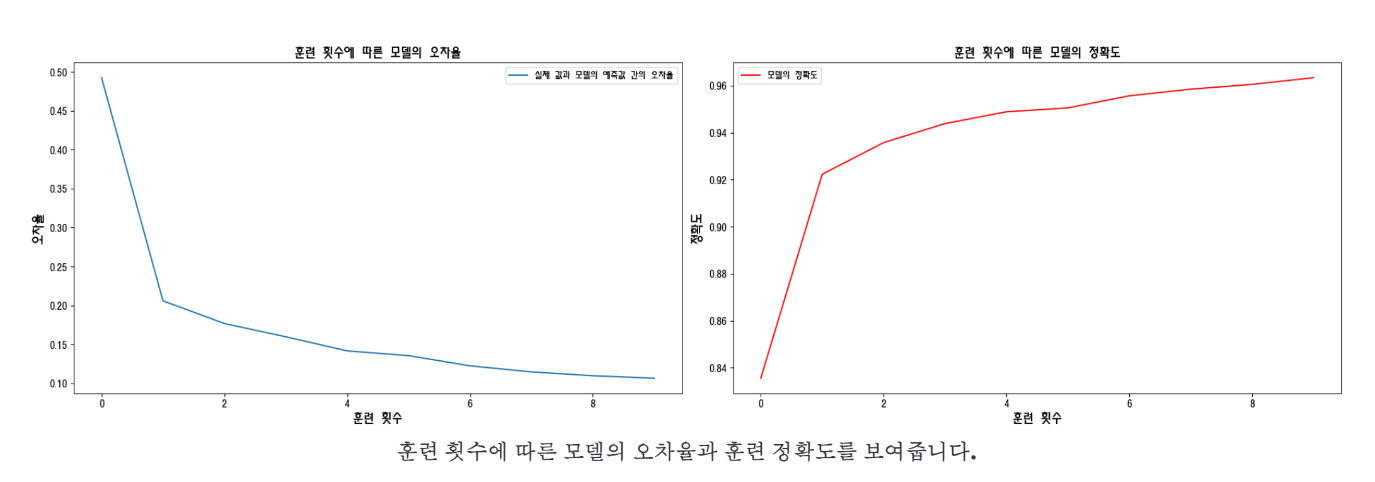
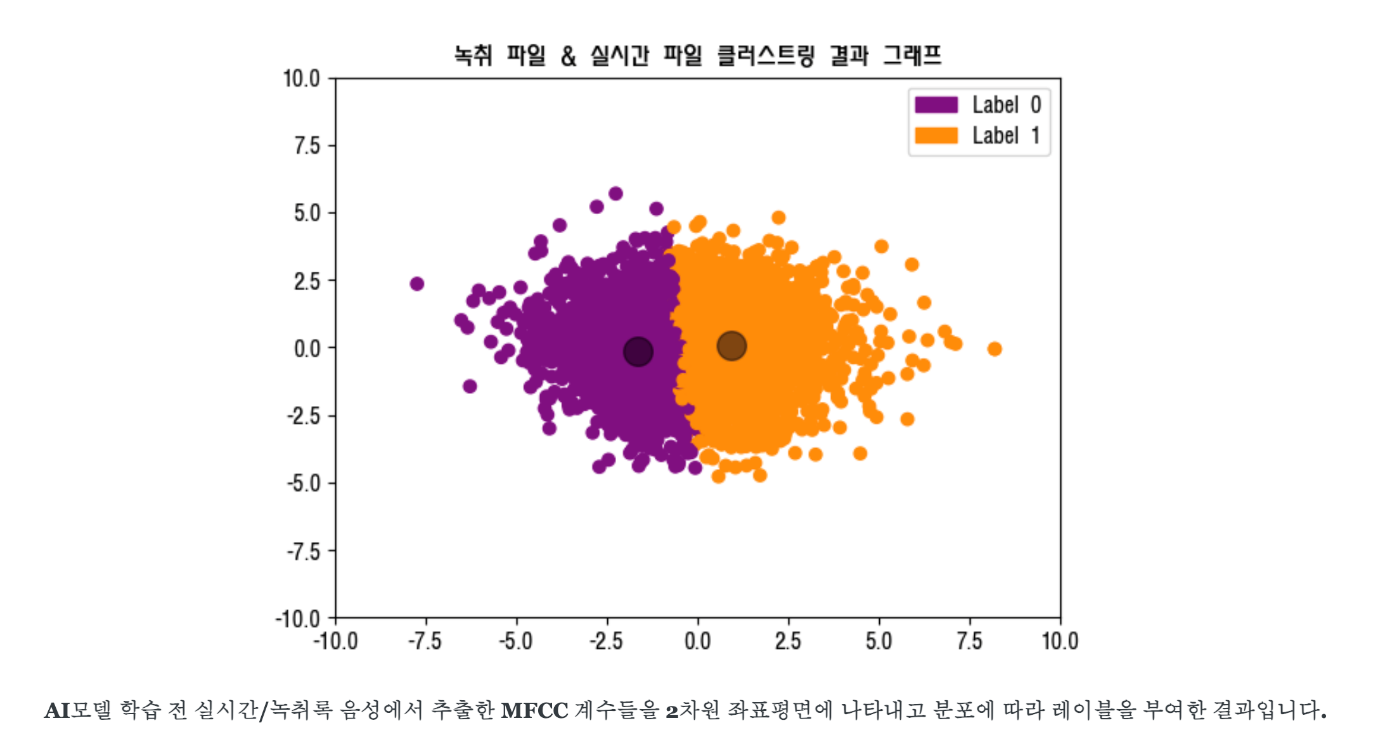
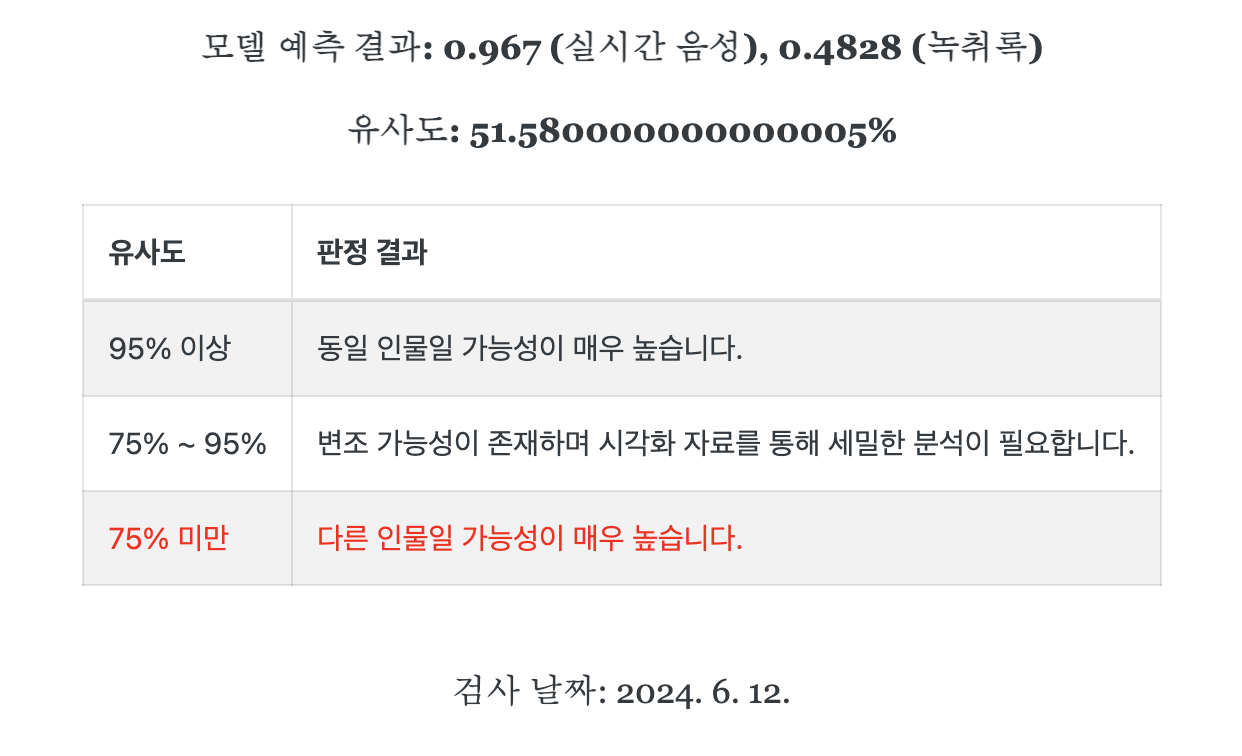




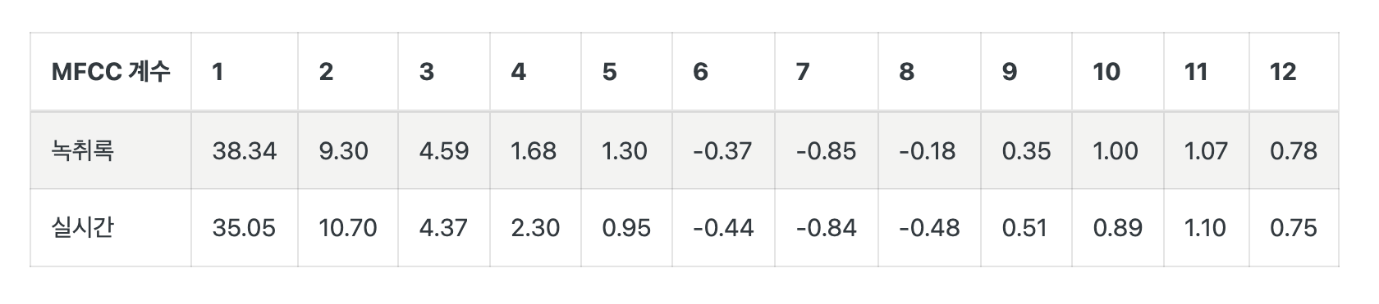
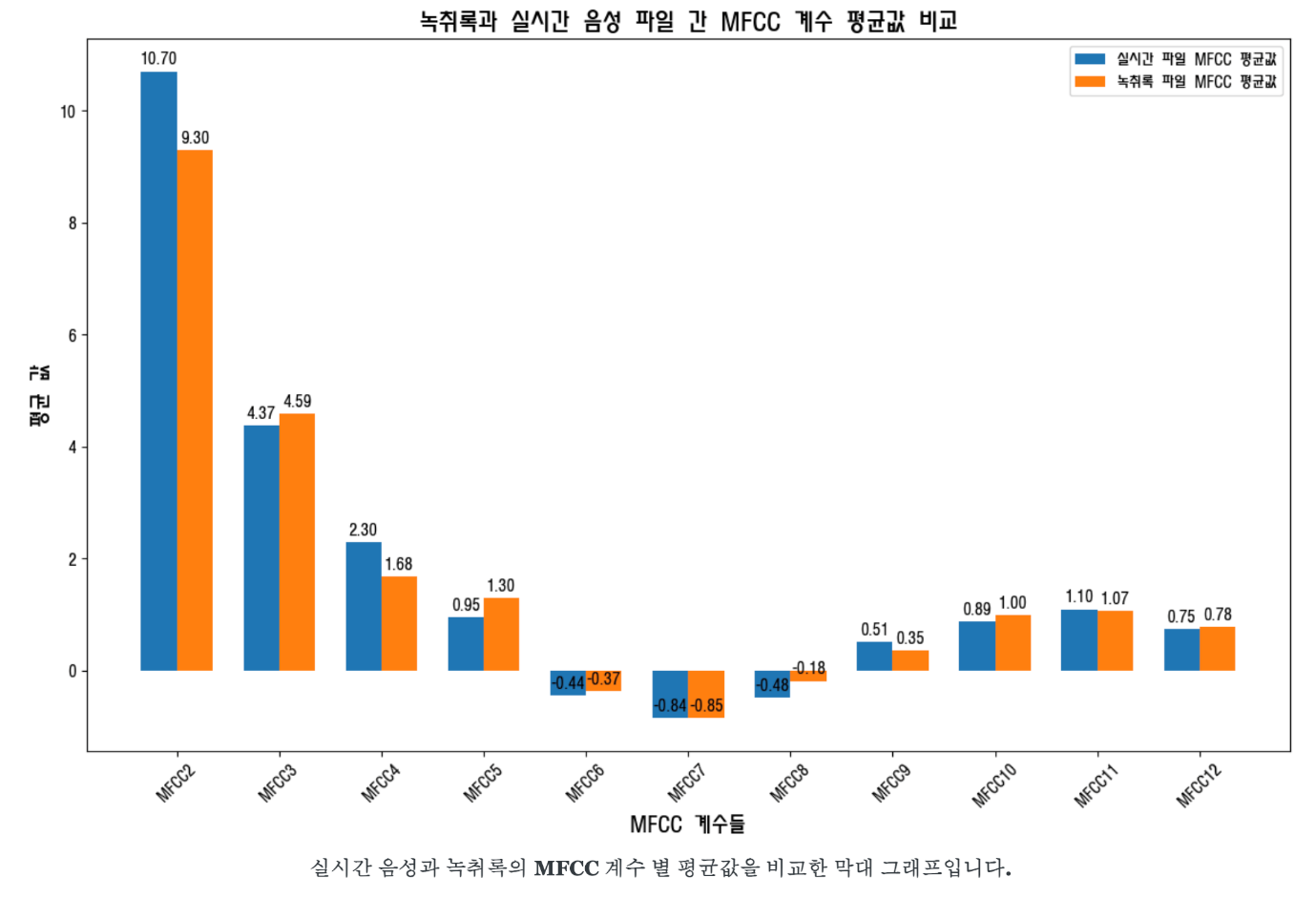
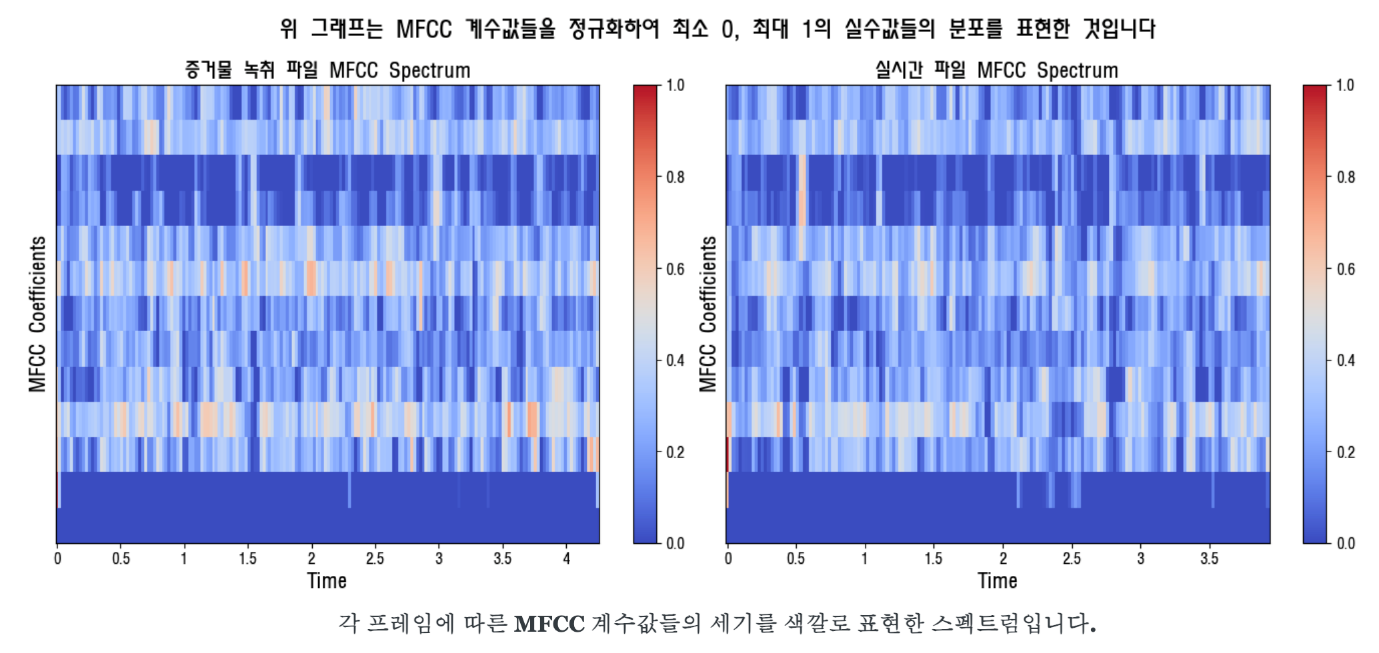
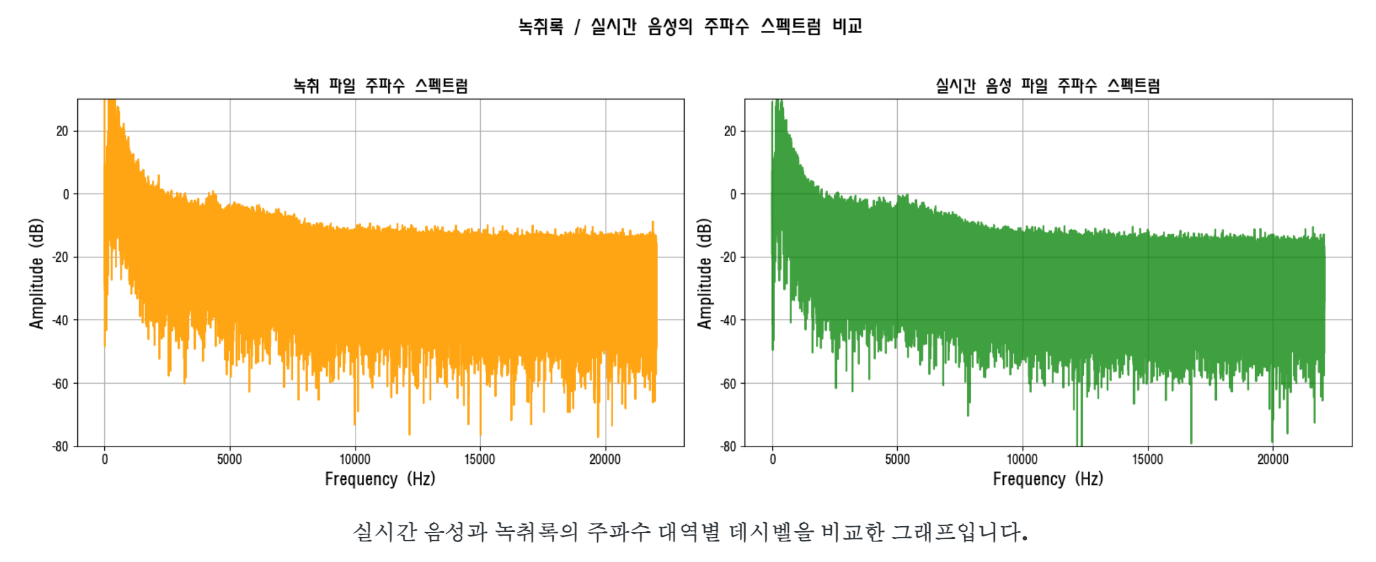
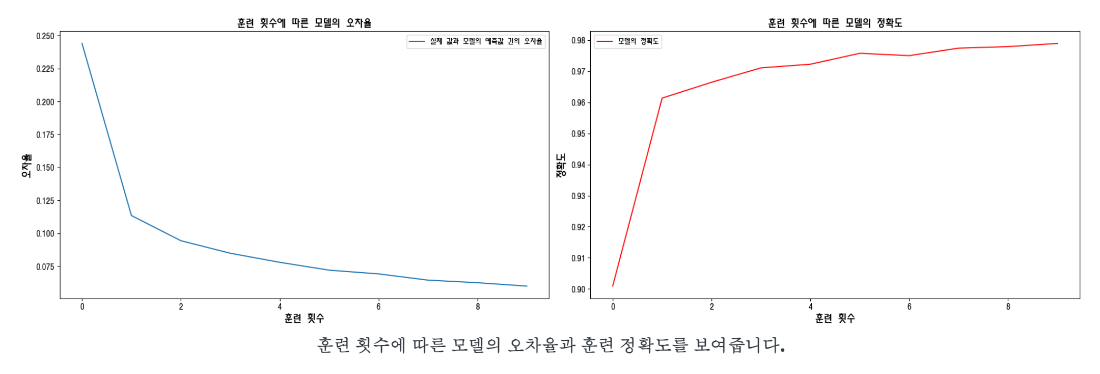
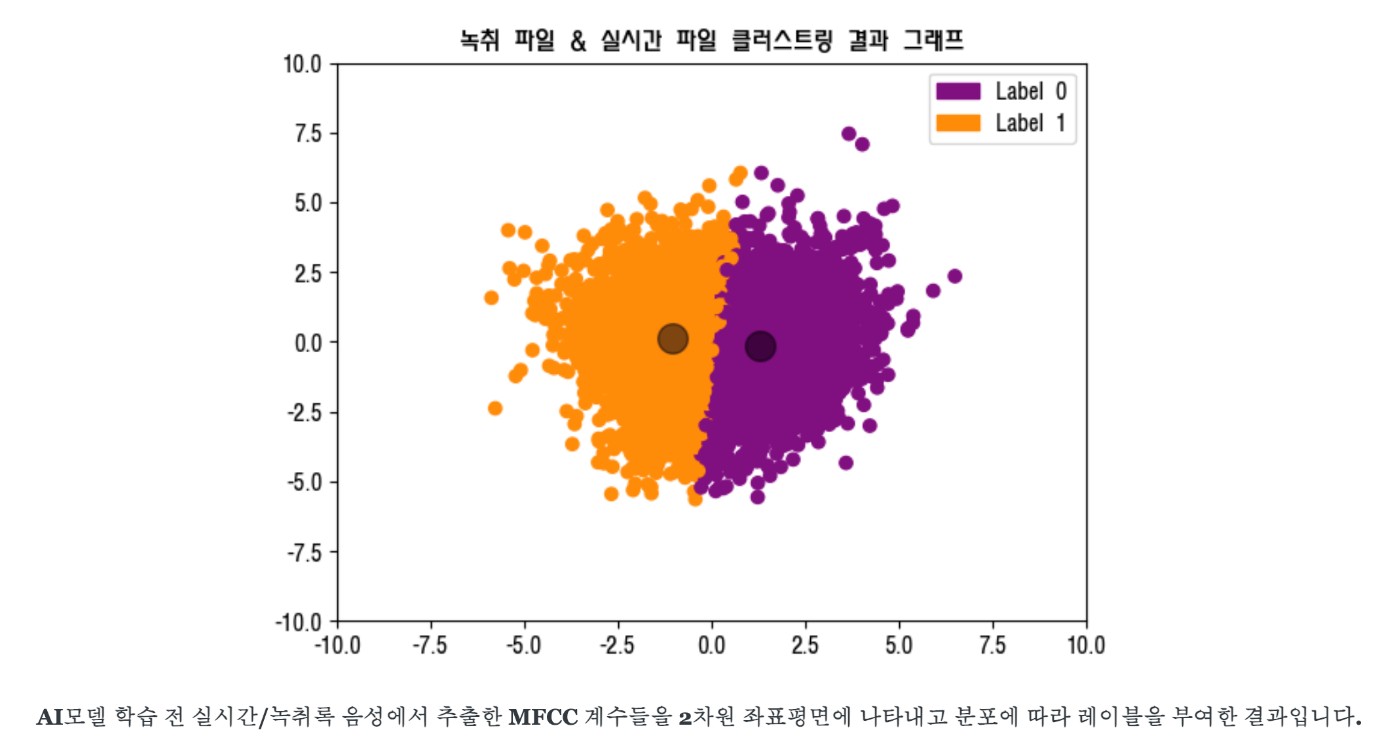
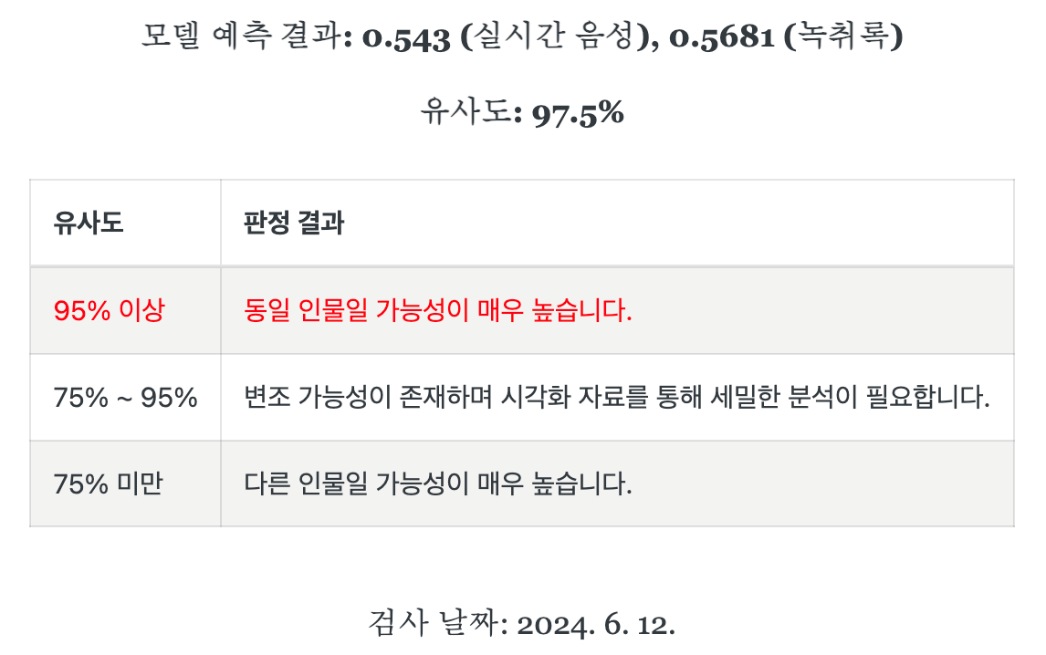




**2) 원본 음성(여성) vs 비교 음성(남성) -> 다른 화자인 경우**



**3) 원본 여성 vs 원본 여성 -> 동일 화자인 경우**



저희는 테스트 케이스 구성시, 총 3가지 경우의 수로 나누어 포렌식 상황을 가정했습니다. 각 테스트 케이스의 음성 파일은 팀장 신서형과 팀원 이기정의 목소리를 이용하여 녹음되었습니다.   
 첫 번째로, 녹취록이 AI 변조 음성이고 실시간 파일이 원본 성인 경우입니다. 변조 음성은 갤럭시 스마트폰의 ‘텍스트로 전화받기‘ 기능을 활성화하여 원본 음성에 AI 변조 기능을 적용하여 제작하였습니다. 이 경우의 유사도는 약 82.69%로, 변조 음성 파일은 듣기에 어색한 부분도 존재하지만 특정 발음이 새어 나오는 소리, 화자와 유사한 피치 등 정밀한 모사 능력으로 인해 유사도가 다소 높게 도출되었습니다. 하지만 동일 화자인 경우보다는 유사도가 상대적으로 낮을 뿐만 아니라, 그래프 상에서도 두드러진 차이점이 드러나 유의미한 결과라고 판단할 수 있었습니다.

두번째로, 음정, 주파수, 발음 특성 등이 확연하게 다른 여성과 남성의 경우를 포렌식하였습니다. 이 경우의 녹취록 예측값 0.4828, 실시간 예측값 0.967, 유사도 약 51.58%로 유사도가 낮고 예측값에서도 확연하게 차이가 있음을 파악할 수 있었습니다

세 번째로, 녹취 음성과 실시간 음성의 화자가 동일한 경우입니다. 이 경우의 녹취록 예측값

0.5681, 실시간 음성 예측값 0.543, 유사도 97.5%로 두 음성에 대한 모델의 예측값이 매우 유사하고 유사도 또한 높은 수치임을 확인하였습니다.

Ⅲ. 결론

1. 기대효과

1. 법정 집행 혼란 감소

증거에 해당하는 음성 파일을 조작하거나 녹취록이 변조된 음성일 경우 법정 집행에 혼란을 줄 수 있다. 음성 변조 탐지 포렌식 프로그램을 통해 녹취록을 사전에 검사하여 재판이 조작된 증거로 인해 오판되는 사례를 줄일 수 있다.

1. 녹취록의 신뢰성 상승

AI 기술로 사람의 목소리를 쉽게 조작하여 녹음파일을 만들 수 있어 녹취록의 신뢰도를 문제삼아 증거 채택에 어려움이 발생하는 경우 결과 보고서를 통해 녹취록과 발화자의 음성 유사도를 보여줘 증거로 제출된 증거물의 신뢰도를 상승시킨다.

2. 개발 이슈 및 해결 방안

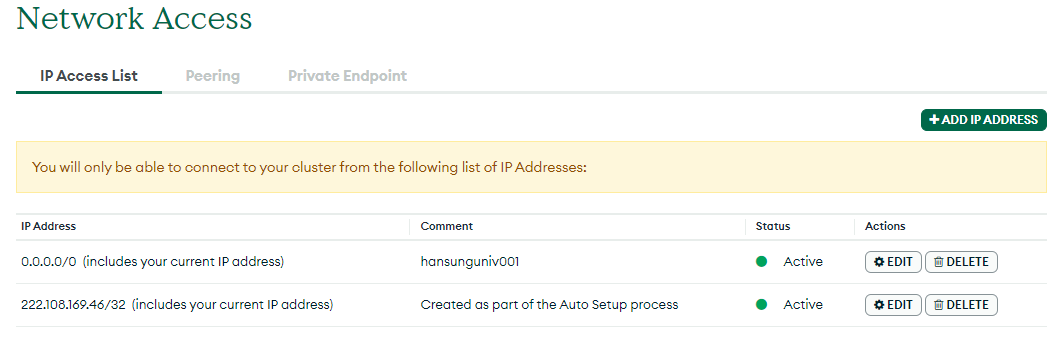
1. **MFCC의 원리 이해와 라이브러리 활용**

프로젝트 주제 선정 당시 MFCC 원리에 대해 익숙치 않은 상태였다. 지도 교수님의 권유로 MFCC 계수들의 추출 과정을 조사하여 세미나를 진행하면서 MFCC가 FFT(푸리에 변환), Mel Filter Bank(청각 시스템을 반영한 필터), DCT(코사인 변환) 등을 거쳐 일종의 벡터 형태로 반환된 계수라는 것을 알게 되었다.

또한, github에서 MFCC 알고리즘을 javascript언어로 구현한 오픈 소스 코드를 발견했는데 계수값들을 추출하기에 앞서 명령어 옵션 등의 라이브러리 활용법에 대한 분석이 요구되었다. 팀원들과 함께 조사한 결과, -w 옵션이 파일 경로라는 점, 음성 파일이 8bit mono 형식이라는 점, -n 옵션을 주어서 음성 주파수의 샘플링 횟수를 지정할 수 있다는 점 등을 알게 되었다.

1. **Mongo DB 클라우드 연동 및 접근 권한 설정**

Node.js 환경에서 로컬에 설치된 Mongo DB에 접근할 수 있었지만 팀원들 간 데이터를 mongoExport를 통해 넘겨주고 mongoImport를 통해 받는 과정이 요구되어 번거로웠다. 이렇게 각자 로컬 환경에서 DB를 구축하여 진행하던 도중, Python 환경에서 mongoDB에 저장된 MFCC 계수들을 불러와서 데이터 프레임을 생성하고 학습해야 하는 일이 있었다. 간편하게 DB와 연동할 수 있는 방법을 고민하다가 mongo DB Cloud 버전인 mongo DB Atlas라는 툴을 통해 공유 DataBase를 생성할 수 있으며 connection string(url 문자열)만으로도 사용자를 인증하고 DB에 접근할 수 있음을 알게 되었다. 처음에는 접속이 잘 되지 않아 당황했지만 IP 주소를 아래 사진과 같이 Atlas의 네트워크 권한 탭에서 설정해주니 모든 팀원들이 정상적으로 접근할 수 있게 되었다.

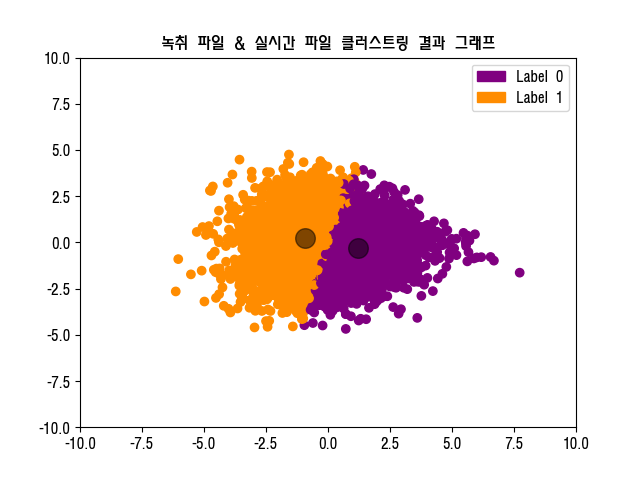


1. **웹 서버와 데이터베이스 연동**  
   사용자가 음성 파일을 업로드하면 해당 메타데이터 값들을 저장하고 파일 mfcc.js를 실행한 결과값을 지정한 모델(스키마)의 형태로 저장하는 과정에서 많은 시행착오가 있었다. 문제를 해결하기 위하여 설계한 모델에 추출한 mfcc 벡터 값들을 저장하기 위해 문자열을 parsing하고, 에러를 처리하는 소스 코드를 구현하였다.
2. **Socket 프로그래밍**  
   수만 라인 이상의 프레임을 mongoDB atlas 환경에 전송하는 데 시간이 오래 걸릴 것을 대비해 upload Progress bar를 추가하는 과정에서 Socket을 이용했다. 유저 Session으로 생성된 Socket을 사용 범위(dashboard, upload)에 한정시키고, 프론트엔드와 백엔드를 소켓으로 연결하여 Server side에서 뷰(upload.pug)에게 업로드 진행 상황을 전달하는 형태로 구현하였다. 이러한 과정에서 구현의 어려움, 환경 설정 문제에 직면했다.
3. **Node.js 비동기 프로그래밍**  
   많은 양의 데이터를 처리하기 위하여 await, async, Promise 등의 비동기 프로그래밍 API를 사용하였다. 이러한 과정에서 비동기 프로그래밍과 관련 메소드 사용법에 대해 이해하는 과정이 필요했고, 적지 않은 에러가 발생했다. 또한, Progress bar를 올바르게 구현하기 위해 의도적으로 병렬 실행을 다시 직렬화하기도 하였다.
4. **레이블이 불분명한 데이터셋 클러스트링 및 모델 구축**

3분 가량의 음성 파일에서 MFCC 계수를 추출하면 아래 사진과 같이 몽고 DB에 MFCC 계수 12개 값을 포함하는 레코드가 약 60000개 기록되었다.

실시간 / 녹취록 파일에서 추출한 약 60000개의 MFCC 계수들을 이용하여 딥러닝 모델을 설계하기전, 가장 큰 난제는 동일 화자인지 아닌지 모르는 상태에서 어떻게 레이블을 부여할 것인지에 대한 고민이었다.

이러한 난제를 갖고 약 2주가량 고민한 결과, “클러스트링” 이라는 비지도 학습 알고리즘을 도입하게 되었다. 실시간/녹취록 음성에서 추출한 MFCC 계수들에서 test 데이터셋을 제외한 나머지 데이터들을 2차원 평면에 두고 포인트 간의 거리를 기반으로 2개의 그룹으로 그룹핑하도록 하였다. 그리고 해당 그룹을 기반으로 레이블 0, 1을 부여하여 Dense 분류 모델에 학습시켰다. 그 결과, 실시간 test 데이터셋에 대한 모델의 예측값과 녹취록 test 데이터셋에 대한 모델이 동일 화자와 다른 화자를 잘 예측함을 확인하였다.



1. **Python 패키지 관리 및 환경 설정**

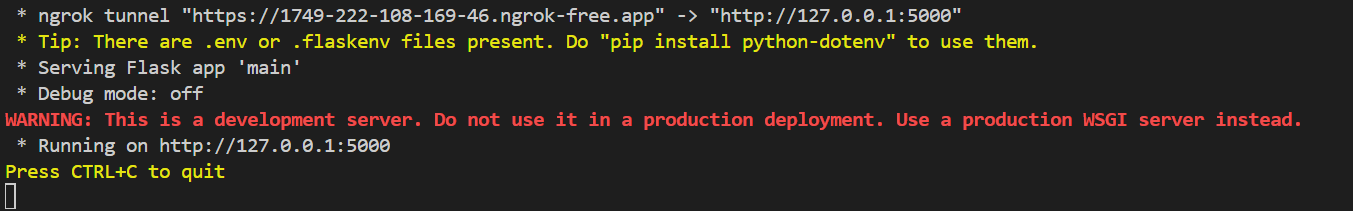
프로젝트 제안 발표 시점에는 Node.js 환경에서 딥러닝을 구현하기 위한 프레임워크인Tensorflow.js를 활용하여 java script 언어로 모델 구축 및 훈련을 진행하고자 하였다. 하지만 tensorflow.js 를 설치하는 도중, 다른 패키지 들에 너무 많은 의존성이 있었고 심지어는 node.js의 버전을 다운그레이드 해야하는 번거로움도 존재하였다. 이 점을 해결하고자 파이썬 언어를 웹상에서 쉽게 작성할 수 있고 고성능의 GPU를 대여할 수 있는 Colab이라는 사이트에서 딥러닝 모델 설계 및 훈련을 진행하였다.

여기서 발생한 또 다른 문제는 다양한 패키지와 모듈 설치가 요구되었는데 코랩 특성상 런타임이 끊길 때마다 일일이 재설치해야 하는 번거로움이 있었다.

따라서 Anaconda라는 툴을 통해 제한된 환경 안에서 필요한 모듈과 패키지들을 미리 설치해놓은 다음에 프로젝트 폴더에 연동시켜놓았다. 이 결과, 실행 버튼 하나만으로 작성한 모든 코드를 간편하게 실행시킬 수 있었다.

1. **Python과 node.js 간 연동 문제**

Node.js 환경에서 Python 언어로 작성한 쉘을 실행시키기 위해서는 node.js와 python간의 연동이 필수적이다. 따라서 Python의 flask 툴을 통해 파이썬 내부 앱(서버)를 실행시키고 데이터셋 불러오기(import\_dataset), 클러스트링(label\_setting), 훈련(training), 스펙트럼 시각화(fft\_spectrum)등 7개의 라우터를 설정하여 코드를 작성하였다. Node.js에서는 클라이언트에게 각 라우터가 호출될 때마다 진행 상황을 통지하는 비동기 프로그래밍을 구현하였다. 이때, Flask 앱 서버와 node.js 간 통신을 위해서 ngrok를 통해 토큰을 등록 및 인증하고 터널링을 통해 각자의 서버에 접근하는 방식으로 구현하였다.



3. 프로젝트 수행 일정

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명

|  |  |
| --- | --- |
| 신서형 | 박소희 |
| 음성녹음 샘플 제작  MFCC 계수 분석  결과 보고서 시각화 자료 제작 및 구성  웹페이지 레이아웃 디자인 | MFCC 원리 및 개념 조사  딥러닝 모델 훈련 및 예측값 DB에 저장  Python 환경 구축 및 node.js와 연동  프로젝트 기능 시연, 유지 보수 |
| 남은희 | 이기정 |
| 데이터베이스 설계 및 구현  딥러닝 모델 조사 및 훈련 | 웹 서버(프론트엔드/백엔드) 설계 및 구현  데이터베이스 설계 및 구현  프로젝트 기능 시연, 유지보수 |

4. 수행 후기

**신서형**

이번 캡스톤 디자인 과목을 통해 프로젝트를 진행하는 방식에 대해 배울 수 있었습니다. 과목을 수강하기 전에는 개인의 실력이 중요하다고 생각했습니다. 하지만 과제를 진행하면서 개인의 실력도 중요하지만 협력도 중요하다는 걸 알았습니다. 한 학기 동안 미팅을 하면서 소통과 협력의 방식을 배워 매우 소중한 시간이었습니다. 마지막으로 저와 함께해 준 팀원들과 팀 전체가 더 나은 방향으로 나아갈 수 있게 아낌없는 조언을 해주신 교수님께 감사드립니다.

**박소희**

약 3개월 가량의 프로젝트를 진행하면서 화자 인식에 활용되는 MFCC 기술, 원격 서버에 접근하기 위한 ngrok 터널링, Python의 flask 웹 프레임워크 등을 실제로 설계해보고 구현할 수 있어 유익한 경험이었다. 또한 팀원들과 매주 회의를 통해 생각을 구체화하고 서로의 의견에 피드백하며 의견을 조율해나가는 과정의 중요성을 알게 되었다. 처음 설계했던 방향과 달라진 점도 있었지만 팀원들의 아이디어가 잘 조립되어 구현까지 잘 완료할 수 있어 뿌듯하다. 저희 프로젝트에 관심 가져주시고 피드백해주신 지도 교수님께 감사의 인사를 드리고 싶다.

**남은희**

이번에 캡스톤을 통해 지금까지 대학교를 다니면서 배운 것들에 대해 잘 정리할 수 있었던 기회가 되었던 것 같습니다. 이 프로젝트를 준비하는 시간 동안 아쉬운 점도 많았지만, 사회로 나가기 전 마지막으로 도움이 되는 것들을 정말 많이 느끼고 배우며 많은 경험들을 얻을 수 있었던 좋은 시간들이였습니다. 지금까지 배우고 느낀 것들을 잘 활용해 앞으로 한 단계 더 성장한 개발자로서 활약할 수 있을 것 같습니다.

**이기정**

제대로 된 웹 개발 기반 지식과 경험이 없는 상태에서 익숙하지 않은 프로그래밍 언어와 프레임워크를 이용해 웹 애플리케이션을 개발하는 데 다양한 어려움에 직면했다. 하지만 이러한 어려움을 극복하는 과정에서 팀원과 협업하는 환경에서 이루어지는 전반적인 웹 개발 과정에 대해 더 깊이 있게 이해할 수 있었다. 단순히 코드를 작성하는 것을 넘어 설계와 구현부터 배포에 이르기까지 웹 애플리케이션의 동작과 흐름에 대해 더 자세히 알게 되는 뜻깊은 경험이었다.

5. 지원금 집행내역

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명

마이크 구매(142,400원)+챗지피티 구독 2달(56,115원)+코랩 프로 구독(14,859원)으로 총 213,374원 사용했습니다.

Ⅳ. 참고자료 및 소스 코드 URL

1. 참고 자료

**1) 참고 도서**

파이썬으로 배우는 음성인식/다카시마 료이치 지음 / 출판사 비제이 퍼블릭

파이썬을 이용한 기초통계학/박철근 지음 / 자유아카데미

백견불여일타 Node.js로 서버 만들기 / 저자 : 박민경 / 출판사 : 로드북

몽고 DB 완벽 가이드 / 저자 : 크리스티나 초도로 / 출판사 : 한빛미디어

실시간 모니터링 시스템을 만들며 정복하는 MEVN / 저자 : 주홍철 / 출판사 : 비제이 퍼블릭

모두의 데이터 분석 with 파이썬 / 저자 : 송석리, 이현아 / 출판사 : 길벗

**2) 참고 사이트**

Valerio Velardo (The sound of AI) , Mel-Frequency Cepstral Coefficients Explained Easily

<https://www.youtube.com/watch?v=4_SH2nfbQZ8>

Slide 9, Mel Filter Banks 사진 자료 및 개념 참고

<https://hyunlee103.tistory.com/46>

한 번에 끝내는 Node.js, 제주코딩베이스캠프

<https://www.youtube.com/watch?v=pK2IKP3M5Ls&list=PLkfUwwo13dlWJpPLxJRYVrHx3kRCiPRb2>

자바스크립트 비동기 개념에 익숙해지기 3편

<https://goldenrabbit.co.kr/2023/07/28/%EC%9E%90%EB%B0%94%EC%8A%A4%ED%81%AC%EB%A6%BD%ED%8A%B8-%EB%B9%84%EB%8F%99%EA%B8%B0-%EA%B0%9C%EB%85%90%EC%97%90-%EC%9D%B5%EC%88%99%ED%95%B4%EC%A7%80%EA%B8%B0-3%ED%8E%B8/>

[Node.js] Socket 모듈 (.emit .on)

<https://velog.io/@rzee/Node.js-Socket.io-%EB%AA%A8%EB%93%88-.emit-.on>

2. 소스 코드 URL

* <https://github.com/modelable/VoiceForensics>

1. 양다훈, "나 아들인데 돈 보내줘" 보이스피싱에 목소리 변조 AI기술 악용,세계일보,2023년 3월 7일,"https://m.segye.com/view/20230307514317"(2024년 6월 12일 접속) [↑](#footnote-ref-2)
2. 디지털 포렌식 센터,”<https://m.blog.naver.com/dfcforensic/223149697929?view=img_5>”,(2024년 6월12일 접속) [↑](#footnote-ref-3)