

2022 AI보안연구센터 인턴 미팅

Research REPORT

22년 10월 7일

20170622

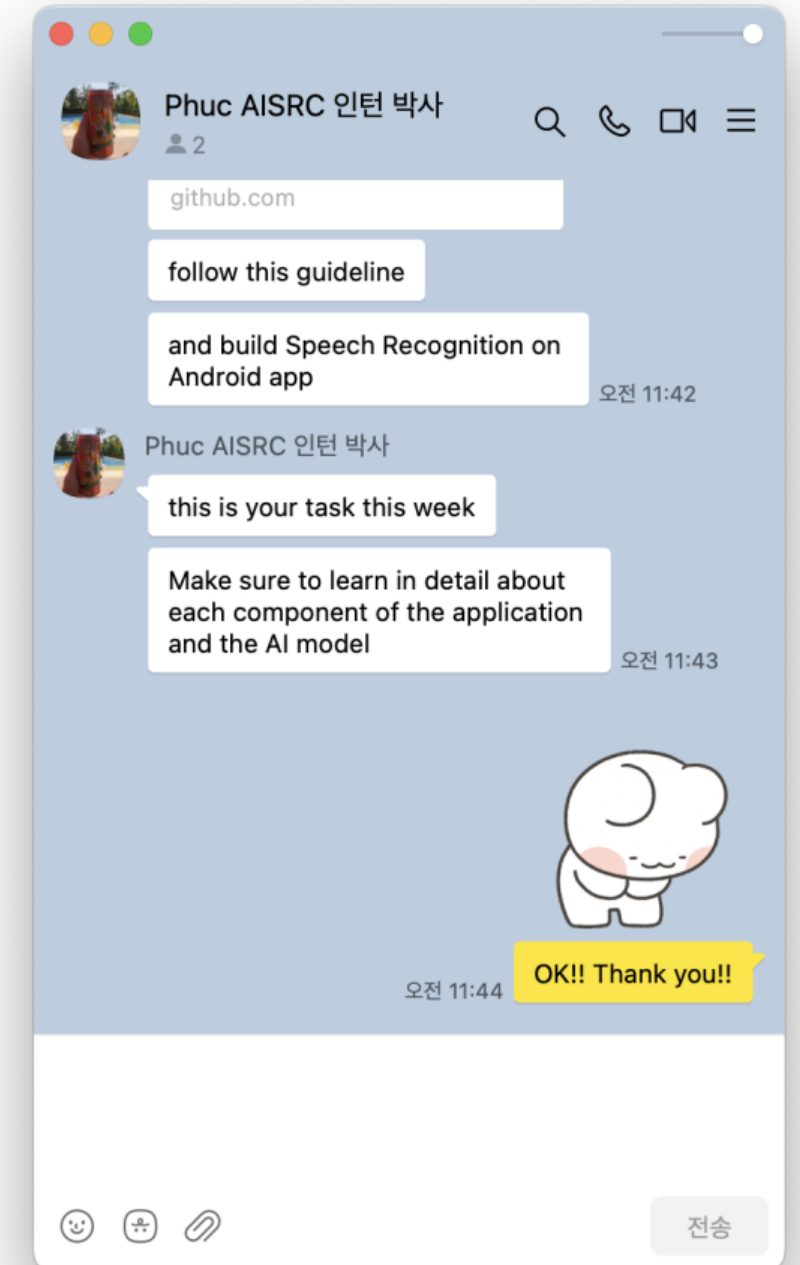
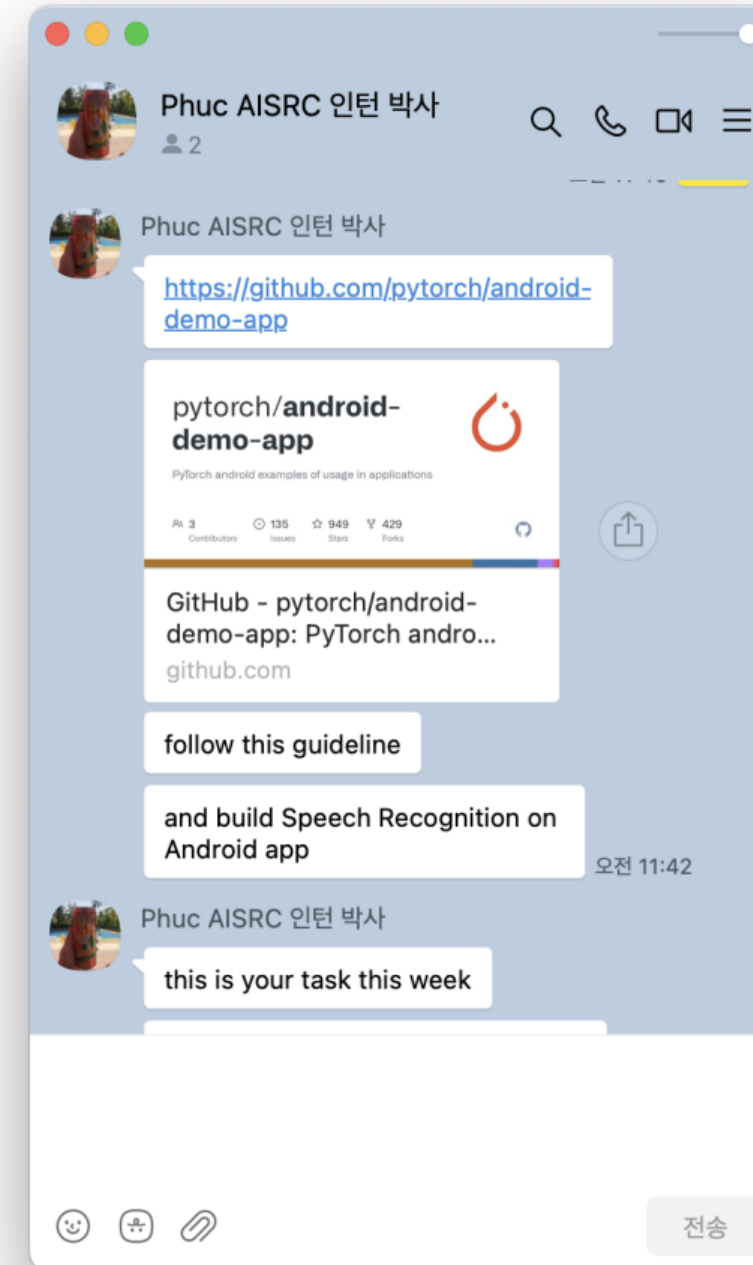
이종헌



01 Task

Task

Pytorch의 공식 Github의
Android-demo-app
repository에 가서
Speech Recognition을 Android App으로 빌드



Speech Recognition of Android-demo-app

<https://github.com/pytorch/android-demo-app>

해당 링크에서 **Speech Recognition**으로 들어가

해당 **코드에 대한 분석**을 진행

☰ README.md

Speech Recognition on Android with Wav2Vec2

Introduction

Facebook AI's [wav2vec 2.0](#) is one of the leading models in speech recognition. It is also available in the [Huggingface Transformers](#) library, which is also used in another PyTorch Android demo app for [Question Answering](#).

In this demo app, we'll show how to quantize, trace, and optimize the wav2vec2 model, powered by the newly released torchaudio 0.9.0, and how to use the converted model on an Android demo app to perform speech recognition.

Prerequisites

- PyTorch 1.9.0 and torchaudio 0.9.0 (Optional)
- Python 3.8 (Optional)
- Android PyTorch library 1.9.0
- Android Studio 4.0.1 or later

Quick Start

1. Get the Repo

Simply run the commands below:

```
git clone https://github.com/pytorch/android-demo-app
cd android-demo-app/SpeechRecognition
```

Prepare the Wav2Vec2

Guideline에 따라 모델 생성 코드 실행
Pytorch와 Python을 통한 진행
모델은 Facebook AI팀의 모델

2. Prepare the Model

To install PyTorch 1.10.0, torchaudio 0.10.0 and the Hugging Face transformers, you can do something like this:

```
conda create -n wav2vec2 python=3.8.5
conda activate wav2vec2
pip install torch torchaudio transformers
```

Now with PyTorch 1.10.0 and torchaudio 0.10.0 installed, run the following commands on a Terminal:

```
python create_wav2vec2.py
```

This will create the PyTorch mobile interpreter model file `wav2vec2.ptl`. Copy it to the Android app:

```
mkdir -p app/src/main/assets
cp wav2vec2.ptl app/src/main/assets
```

04 Build and run with Android Studio

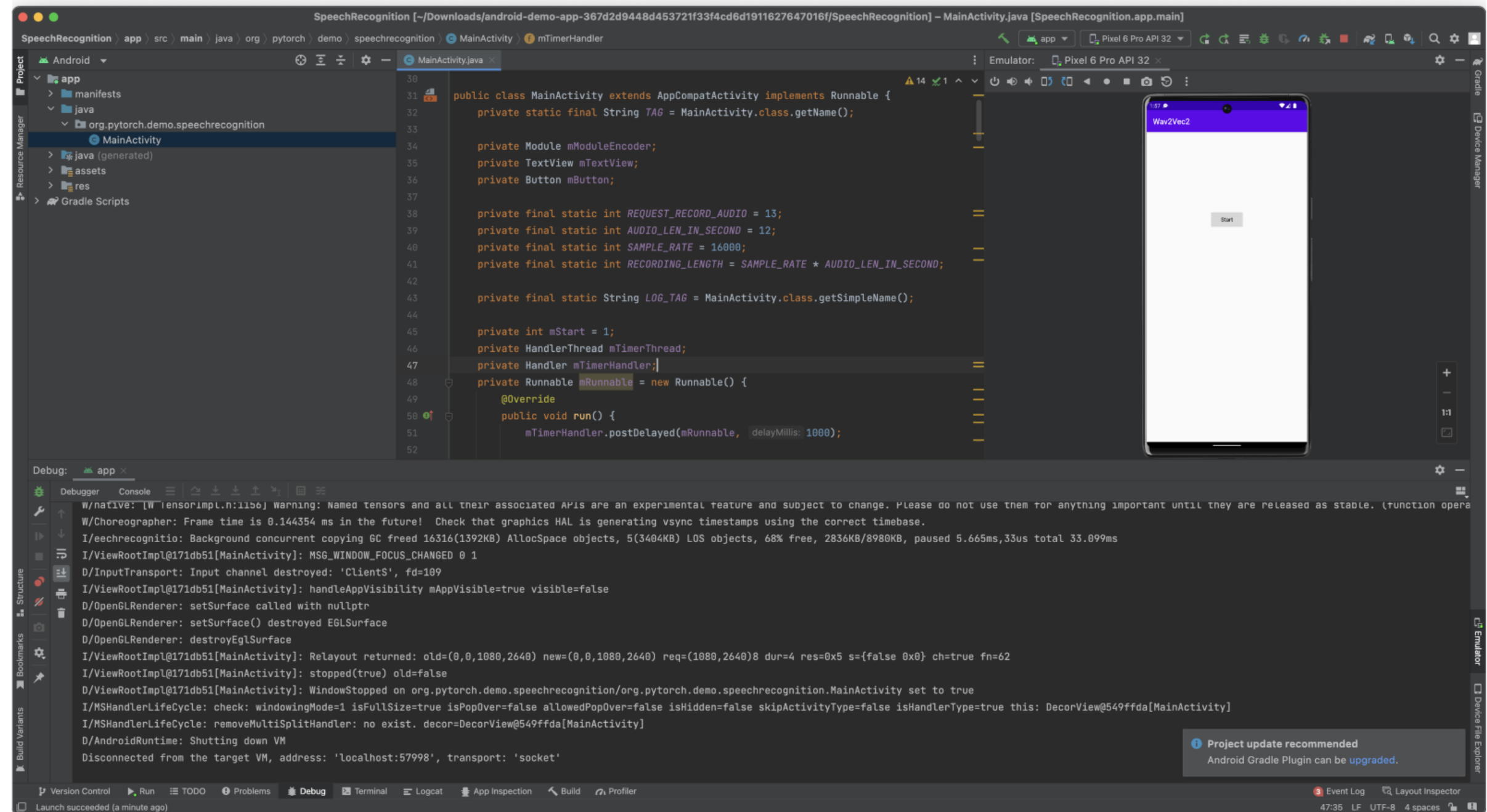
Build and run with Android Studio

Start 버튼을 클릭하면 녹음을 실행

12초가 지나면 녹음된 파일을

모델을 통해 분석 후

반환된 결과를 화면에 표시



Control record time & load AI model

1. Control record time

해당 변수의 값을 조절 함으로써 녹음시간 조절 Default : 12

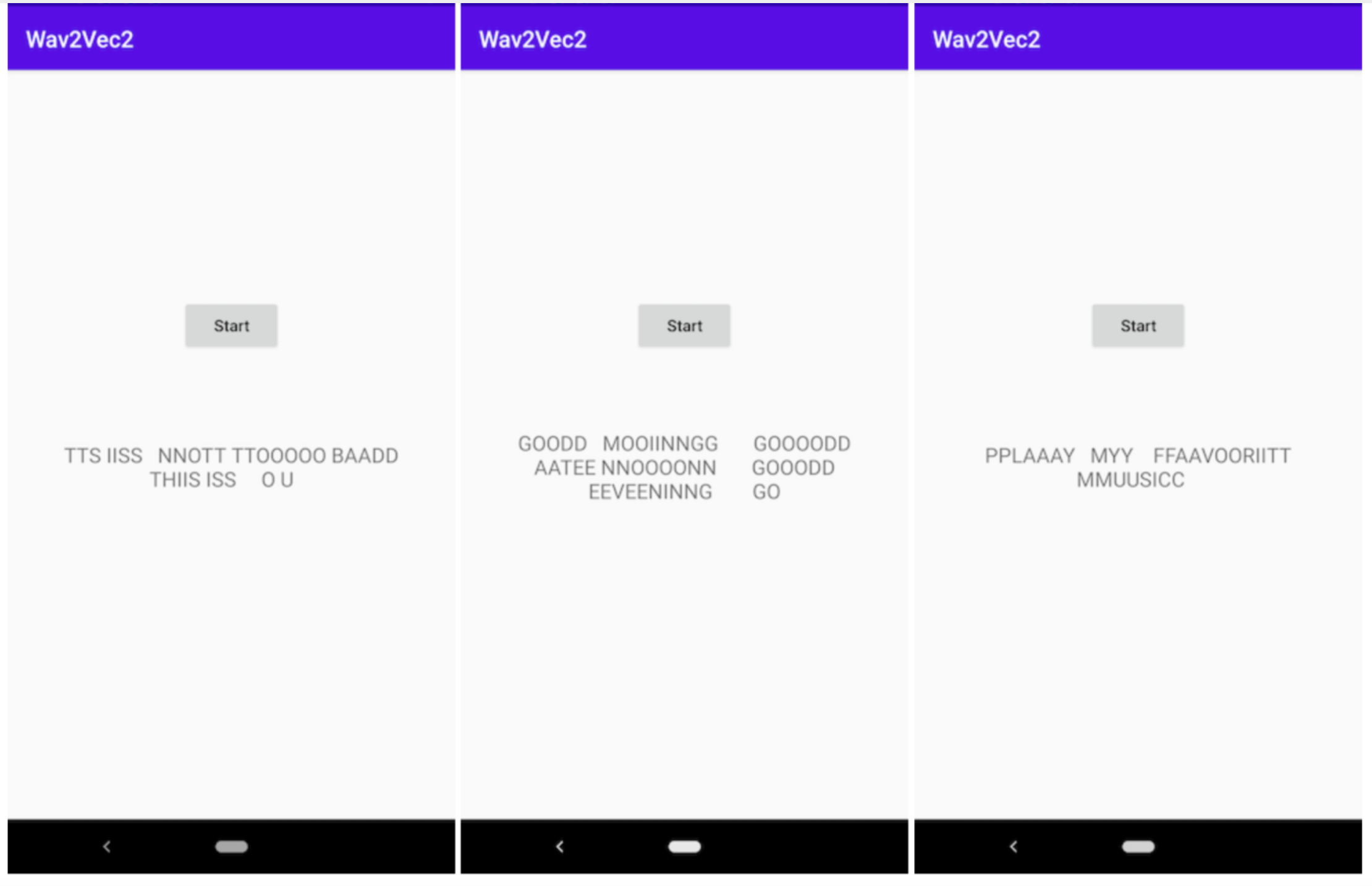
```
public class MainActivity extends AppCompatActivity implements Runnable {  
    private static final String TAG = MainActivity.class.getName();  
  
    private Module mModuleEncoder;  
    private TextView mTextView;  
    private Button mButton;  
  
    private final static int REQUEST_RECORD_AUDIO = 12;  
    private final static int AUDIO_LEN_IN_SECOND = 12;  
    private final static int SAMPLE_RATE = 16000;  
    private final static int RECORDING_LENGTH = SAMPLE_RATE * AUDIO_LEN_IN_SECOND;  
  
    private final static String LOG_TAG = MainActivity.class.getSimpleName();
```

2. Load AI model

assetFilePath 함수를 통한 pt파일 형식의 AI 모델 로드

```
private String recognize(float[] floatInputBuffer) {  
    if (mModuleEncoder == null) {  
        final String moduleFileAbsolutePath = new File(  
            assetFilePath(context, this, "wav2vec2.pt")).getAbsolutePath();  
        mModuleEncoder = Module.load(moduleFileAbsolutePath);  
    }  
}
```

Result - Android Studio



Result - On-device Test

On-device : Galaxy Z Flip 4

