Keyword detection project proposal

Nika Čelan, Hubert Michalski

October 13, 2023

1 Introduction

In modern word, we use our voice not only for communication with other people but also for communication with our phones and even smart homes. If we want to turn on our google assistant the device has to recognise one specific line which is "Hi Google". That is how it all begins. If we want to successfully communicate with devices they also have to understand or detect keywords. This project aims to leverage federated learning techniques to keyword detection models while prioritizing user privacy.

1.1 Related work

There are already similar models build which we are going to look into and try to understand and improve. An example is "Simple audio recognition: Recognizing keywords" [1].

2 Application core

The application core encompasses the foundational elements of the project. Focus of our application is the model for recognising keywords. The main challenge of project will be preprocessing and working with the sound signals and implementing model into Andriod application.

2.1 Data

Data for the model training will be obtained from Google Speech Commands Dataset. We are going to use it's prerecorded commands/keywords. The Speech Commands Dataset has 65,000 one-second long utterances of 30 short words, by thousands of different people. It includes basic command words such as yes, no, up, down, left, right etc. Because of small amount of samples we will simulate users for federated learning by combining multiple users voice recordings and using them as one user.

2.2 Architecture and Technology

We are going to build best possible model for classification of speech commands. Additional to that we are going to build android application which will show us results of classifications. To build application we are going to use Android Studio with Kotlin, and for creating model and sound recognition we will use Pytorch library. For implementing Federated Learning into Android we are going to use library called Flower framework.

3 Work plan

Week	Nika	Hubert
(02) 09.10.	Create project proposal	
(03) 16.10.	Prepare database: make users and split it to train and test set.	
(04) 23.10.	Choosing a suitable federated learning model architecture for sound keyword recognition	
(05) 30.10.	Create base of frontend of Android app	Preprocess data for training
(06) 06.11.	Implement communication to Android app	Defining the Model
(07) 13.11.	Implement Flower client	Training the model
(08) 20.11.	Midsemester Presentation	
(09) 27.11.	Integrate the chosen federated learning framework for local training	
(10) 04.12.	Optimize application and its design	Optimize the model and federated learning process
(11) 11.12.	Deploy the Android app with federated learning	
(12) 18.12.	Fixing bugs and problems	
(13) 25.12.	Christmas holidays	
(14) 01.01.	Performance testing	
(15) 08.01.	Demo and presentation preparation + last minute project check (labs)	
(16) 15.01	Final presentation + report write-up	

4 Application Evaluation

To evaluate project we are going to split our dataset on test and train set for each user. Accuracy is going to be used for evaluation of our model.

References

[1] TensorFlow. *Simple audio recognition: Recognizing keywords*. 2023. URL: https://www.tensorflow.org/tutorials/audio/simple_audio.