**Arabic Language**

**Dialect Identification**

**Software Design Document**

Group 6:

Karim Atwa

Mohamed Hamza

Nesma Badr

Ziad Osama

**Date: 20/3/2018**

[**1. Introduction**](#_77wa1v1gnqf2) **3**

[1.1 Purpose](#_d3ix1d1qqe2d) 3

[1.2 Scope](#_8jotpn7vsgar) 3

[1.3 Overview](#_tara0vxioii) 3

[1.4 Definitions and Acronyms](#_y9h57bnf1eea) 4

[**2. System Overview**](#_slltht827bue) **5**

[**3. System Architecture**](#_8mvoq6hx6872) **6**

[3.1 Architectural Design](#_93lgjja51jks) 6

[3.2 Decomposition Description](#_sairk8i5s6wz) 6

[3.3 Design Rationale](#_hf49jo4kjtcw) 7

[4.1 Overview of User Interface](#_2kmeyrn5luvd) 8

[**Appendices**](#_8kg5cq53gmg) **12**

# 1. Introduction

## 1.1 Purpose

The software design document presents a documentation of the system design of Arabic Language Dialect Identification. It’s written with the intention of using it as a reference for the development process of the system since it gives a detailed technical perspective and an insight to the system and its components.

This document is written having software architects, software engineers and software developers as the intended audience. They should comprehend and understand the document entirely in order to be able to use the system and proceed with its development.

## 1.2 Scope

The Arabic Language Dialect Identification system will provide an accurate way to detect different arabic dialects making it easier for Arabs to communicate with each other. The supported dialects are Egyptian, Levantine, Gulf and North African dialects.

## 1.3 Overview

This document is divided into 6 main sections. The first section is an introduction that is meant to outline the document and its content. The second and third sections contain the system overview and architecture which provides an abstract description of the system showing its architecture and main components. The fourth section is the component design section which highlights the different components in the system. The fifth section will feature the application interface and it will demo the different features of the system. The last section is an appendix and will have the sequence diagrams.

## 1.4 Definitions and Acronyms

|  |  |
| --- | --- |
| Acronym | Definition |
| EGY | Egyptian |
| LAV | Levantine |
| GLF | Gulf |
| NOR | North African |
| MVC | Model View Controller Architecture |
| GPU | Graphical Processing Unit |
| CNN | Convolutional Neural Network |

## 

# 2. System Overview

\ The Arabic Language Dialect Identification System is composed of:

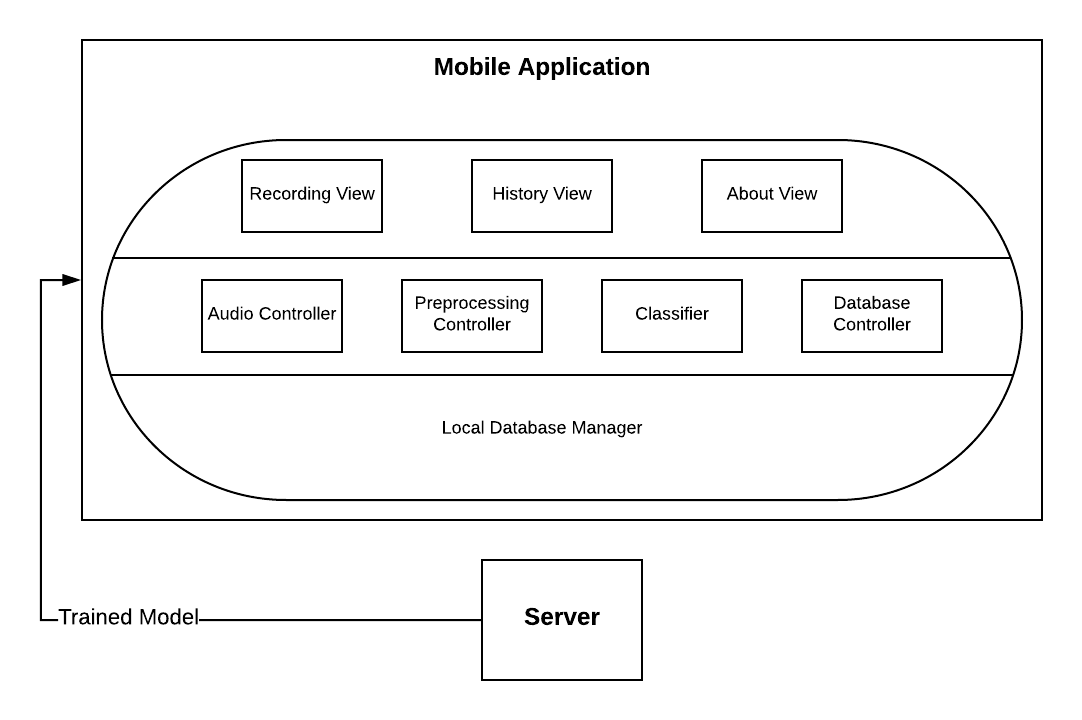
* A server on which the convolutional neural network (CNN) is trained on a dataset, which is composed of five-second audio samples of the four supported Arabic dialects converted to spectrogram images The dataset includes data from various sources; two premade datasets, and some samples were collected from YouTube videos.
* The Mobile Application where the five-second recorded user audio sample is converted to a spectrogram image, and preprocessing of the image is done.The pre-trained model, which is loaded on the device, is then used to make a prediction of the given audio sample. Finally, the prediction is locally saved to the user’s prediction history.

# 

# 3. System Architecture

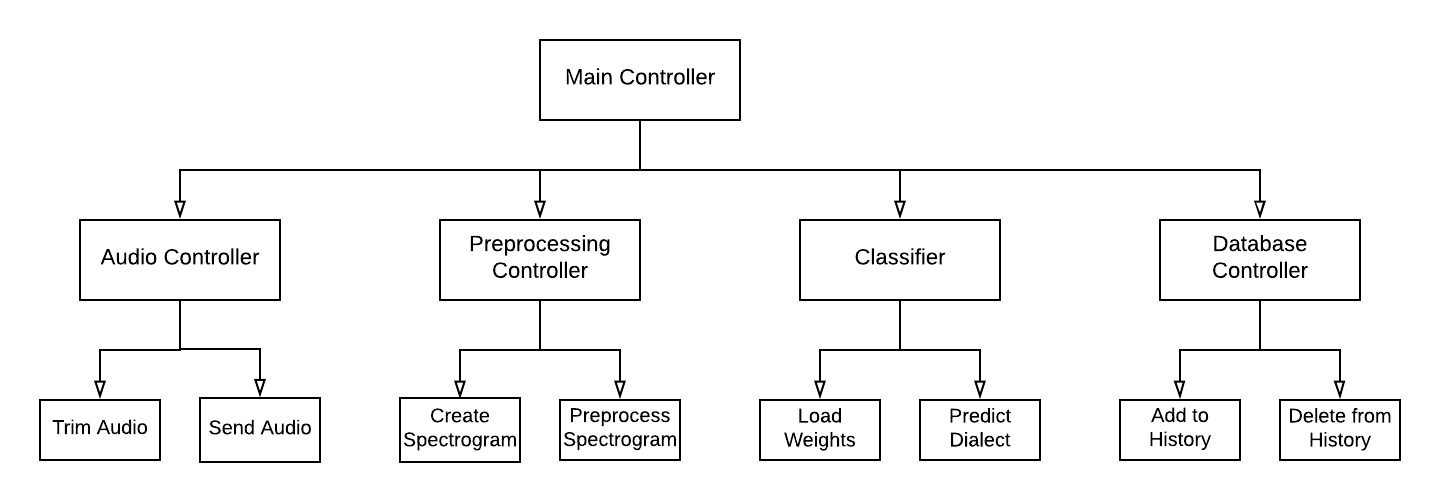
## 3.1 Architectural Design

The architecture of the system is composed of the server and the mobile application. The server is used to train the network. The trained model is embedded into the mobile application. The MVC components of the mobile application are shown below.



## 3.2 Decomposition Description

The Mobile Application has four main controllers; the audio controller, the preprocessing controller, the classifier and the database controller. The audio controller is responsible for trimming the user’s audio sample and sending it to the preprocessing controller, which then converts it to a five-second spectrogram pre-processed for the classification. The classifier loads the pre-trained weights to use them in the dialect prediction process. The final controller ーthe database controllerー adds or deletes the predictions on the user’s local history saved on the device.



## 3.3 Design Rationale

The Architecture of the system is divided into a server and a mobile application. The rationale behind having a server with a powerful GPU is for training the neural network. Then, the trained weights are transferred to the mobile application, which is based on the MVC architecture chosen to result in a standard, well-known, and faster development process. The mobile app is meant to make the system more accessible and user-friendly.

4. Human Interface Design

## 4.1 Overview of User Interface

The mobile application is called “Lahga “ and it has three views: The *Record* view, the *History* view and the *About* view where the supported dialects are stated. The user can freely move through the three screens depending on what they wish to do

4.2 Screen Images

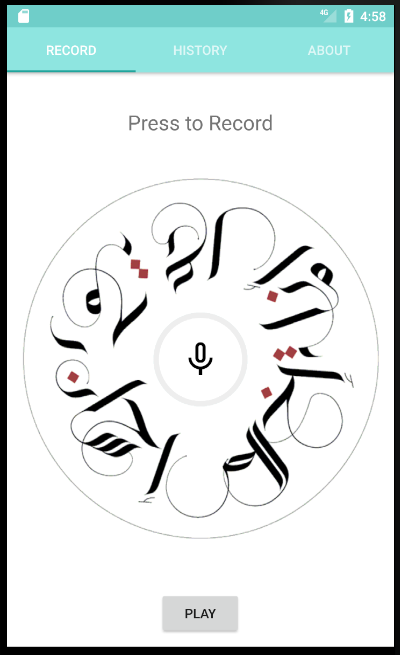


Figure 1: Record View

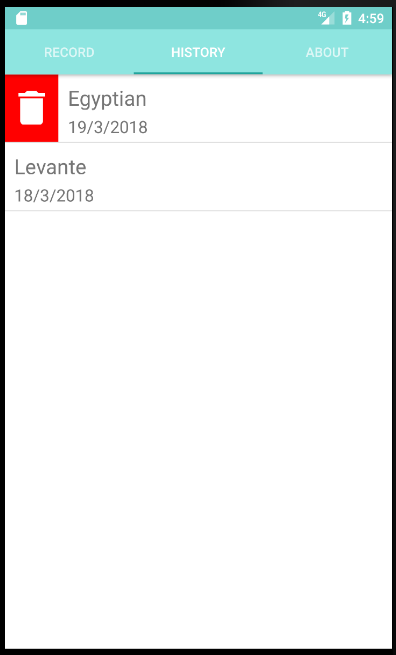


Figure 2: History View

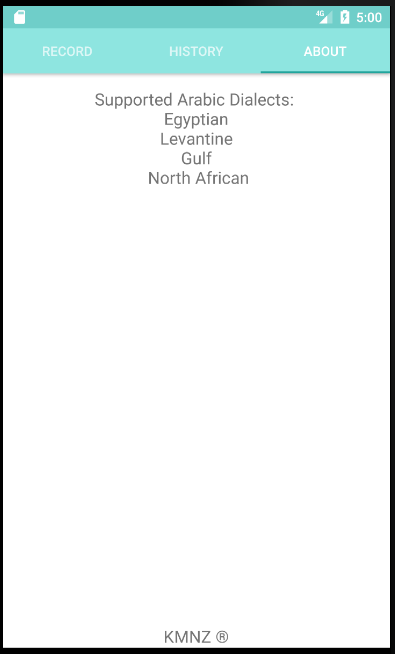
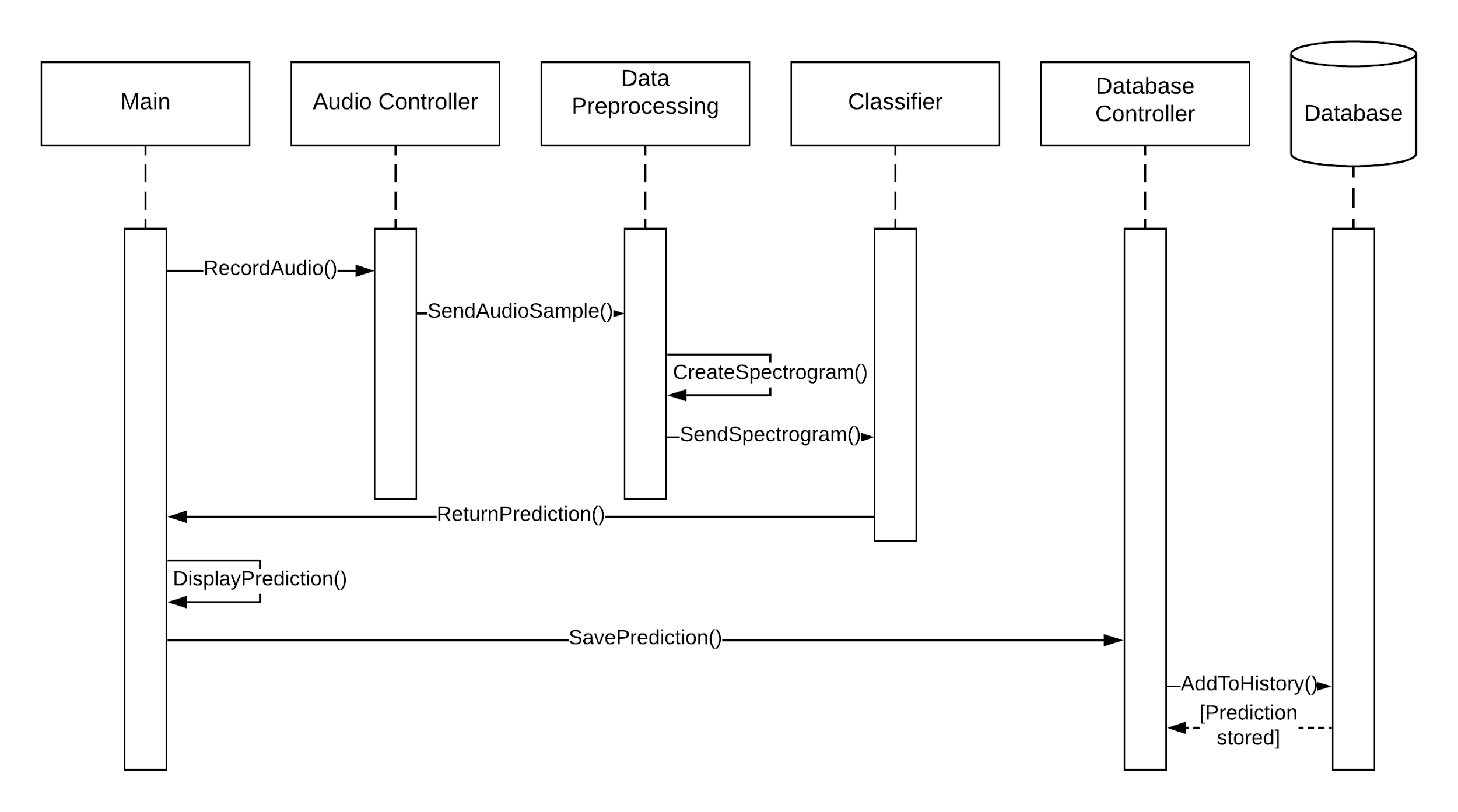


Figure 3: About View

# Appendices



Sequence Diagram: Predicting audio sample