



King Fahd University of Petroleum & Minerals
College of Engineering Sciences
Electrical Engineering Department

EE 413 Applied Digital Signal Processing

Term **212**
Section **01**

Project Report

Arabic Speech Recognition

Prepared by:

Name: Ali Hassan Alnemer	ID#201627320
Name: Basil mashabi	ID#201617060
Name: Saud Ali Alharbi	ID# 201762110
Name: Hassan Bashammakh	ID#201692700

Instructor: Dr. Mudassir Masood

Submission Date: 8th of May 2022

I. Introduction

The aim of this project is to use the knowledge gained from applied signal processing course to build an Arabic speech recognition code that uses machine learning code that process a dataset to recognize Arabic spoken words irrespective of the speaker.

This report will first start with explaining the process of building the dataset made of twenty different Arabic spoken words. Then, this report will illustrate the training model used to build the machine learning code. Furthermore, this report will discuss the results and the accuracy of the built model. Finally, this report will sum up with a brief conclusion that will highlight the main ideas of this report.

II. Building the Dataset

Since this project is a word recognition machine learning program, we first started with building the dataset. The dataset used in this project includes Twenty different Arabic words, each one of them has 100 different samples. For the training model to understand our aim is to recognize spoken words irrespective of the speaker, the samples we collected are either from different speakers or recorded using different background noises.

We built a python code that is used to record the samples to obtain all samples in the same form of .wav and same sampling frequency of 44100, and of the same length of two seconds. This is done because we need to use samples of same specifications to build our training model.

We have chosen twenty Arabic words given in the following table:

Ahla	أهلاً	burtugal	برتقال
marhaba	مرحباً	fahad	فهد
laymon	ليمون	nemer	نمر
asd	أسد	asad	أسد
asfar	أصفر	salam	سلام
ahmar	أحمر	ghazal	غزال
enab	عنب	gadam	قدم
tufah	تفاح	raas	رأس

III. Building the training model

Use Support Vector Machine which should correctly classify the different between word, after the data extracting and processing. Build and train the SVM mode with linear kernel

```
accuracy: 0.5037037037037037
```

```
[[13 1 4 0 0 0 0 1 0 1 1 0 1 0 1 0 0 0 0 0]
 [ 1 10 1 0 1 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0]
 [ 0 1 9 0 0 0 1 0 1 1 1 0 0 0 1 0 0 0 0 0]
 [ 0 4 0 16 1 3 1 0 0 0 1 1 0 0 1 0 0 0 1 0]
 [ 1 0 2 0 12 0 1 0 0 1 0 0 1 0 0 0 0 0 0 0]
 [ 0 1 0 2 3 12 2 0 0 0 0 0 0 0 1 0 0 0 0 1]
 [ 1 3 0 1 1 1 4 0 1 0 0 0 1 0 0 0 2 0 0 0]
 [ 0 0 0 0 0 1 1 10 0 1 3 0 2 0 1 1 1 0 1 0]
 [ 1 0 0 0 0 2 0 0 12 0 1 0 0 0 0 2 0 1 0 0]
 [ 2 0 1 0 1 0 0 0 0 5 0 0 0 1 0 0 1 1 0 0]
 [ 1 0 1 1 1 0 0 4 0 0 8 0 0 0 0 1 0 0 1 0]
 [ 0 1 1 1 1 0 0 0 1 0 1 12 0 0 1 0 2 2 0 0]
 [ 2 1 3 0 1 0 0 0 0 2 1 0 9 0 1 0 1 0 2 0]
 [ 0 0 0 0 2 1 0 0 0 1 2 0 1 10 1 1 0 0 0 0]
 [ 0 0 1 1 1 1 0 1 0 2 0 0 1 1 15 1 0 0 0 0]
 [ 1 0 0 0 1 0 0 1 0 0 1 0 1 0 2 5 0 1 1 0]
 [ 1 4 4 1 0 0 1 0 0 0 1 0 0 0 0 0 10 0 0 0]
 [ 0 0 0 0 2 0 1 1 0 0 6 0 0 0 0 1 0 12 0 0]
 [ 1 2 1 1 0 0 1 3 0 0 3 0 1 0 0 3 0 0 8 0]
 [ 0 1 2 0 1 0 0 0 0 1 1 0 2 1 0 0 1 1 0 12]]
```

As can see from the figure, the linear model gives approximately 50.3% accuracy

When trying the data from one person, the accuracy is 99% because it was recorded in the same device, but when collecting the database from other people, which are 4 files, each with approximately 500 sounds, the accuracy decreased to 50%

IV. Conclusion

In conclusion, in this project (Arabic Speech Recognition) use one of the most famous unsupervised machine learning classifier called Supporting Vector Machine (SVM) to train the model of database for Arabic word and collected the data at different background noise