Indian Institute of Technology, Guwahati

Department of Computer Science and Engineering

Project report

On

“Speech based Online Library System”

Based on

Speech recognition system

Course : CS 566(Speech Processing)

Submitted to

Prof. P.K.Das

Submitted by:

Sherin Abraham (214101052)
Rohith Reddy Sama (214101046)

TABLE OF CONTENTS

ABSTRACT

The project involves the development of a speech based Online Library System. The user can search the data base of the library for any book, both to borrow or to return using a speech based search system.

This report elaborates on the training and modeling of the speech system used and also details the flow of the project.

INTRODUCTION

The aim of the project is to develop a voice controlled online library system. In a normal search procedure, one keys in the search and uses the mouse clicks to navigate the website. However, speech systems can be helpful to avoid button clicks or typing the words. This was implemented in such a way that the user can use his voice to search for a book in the library. Apart from that, he can also borrow or return the book based on the availability of the book. This functionality can be useful in almost any online library system anywhere in the world.

An additional feature of live training is also added to this program. For example, when a new book arrives in the library, the librarian can add that to the book database so that when the user utters the name of the new book the system will be able to recognize it.

While this program may not be 100% accurate, it does a very good job in recognizing most of the words consistently. To achieve more accuracy, more data and advanced techniques such as machine learning or deep learning can prove useful.

SOFTWARE REQUIREMENTS

Basic requirements to develop this project are as follows:

1. Windows OS
2. Microsoft Visual Studio 2010
3. C++ 11 integrated with VS2010
4. Recording Module

The UI has been developed using Visual C++ Windows Forms Application and integrated with the HMM model in the backend.

METHODOLOGY

The sequence of steps followed in this project are

1) Pre-processing of speech data

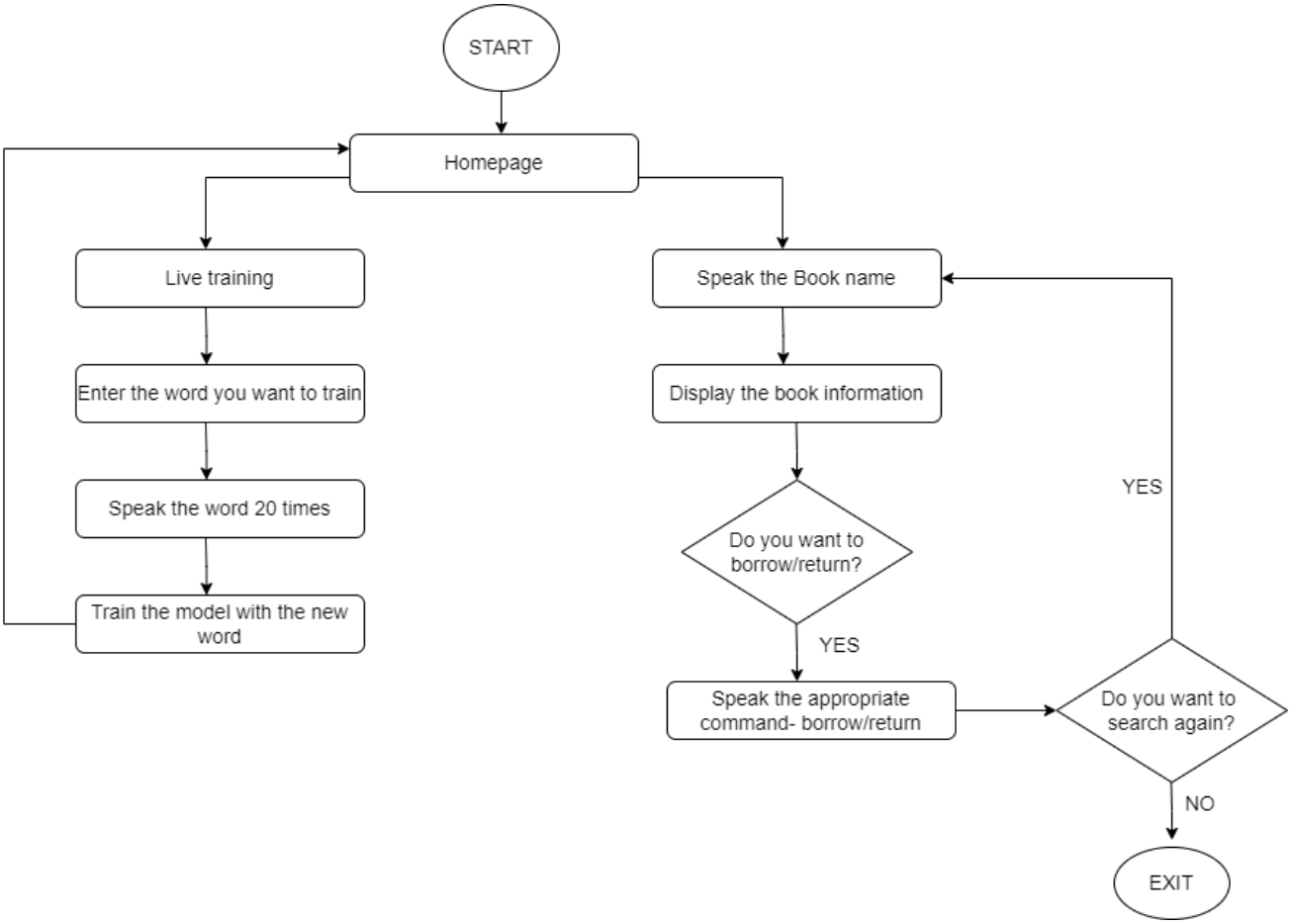
2) Feature extraction

3) Modelling of extracted feature

4) Enhancing model

Below is the flow chart for our project.

FLOW OF THE PROGRAM



EXPERIMENTAL SETUP

This project is divided into following modules:

1. Training Module
2. Testing Module
3. Live Training Module

1. Training

i. Record the training data which is 20 utterances of each word used in the project.

ii. Extract frames for every utterance. Use LBG to generate a codebook.

iii. Using local distance analysis (in vector quantization), calculate the observation sequence.

iv. Pass this observation sequence to HMM to generate a model.

v. Enhance the model using HMM re-estimation algorithm. Save the model in a folder.

The HMM model is now ready. Now onto the testing part.

2. Testing

i. The user speaks the word first. Pre-process that data.

ii. Load the model which was previously saved in a folder.

iii. Obtain the observation sequence and pass it to the model to obtain prediction.

iv. The word with the highest probability will be recognized as the uttered word.

3. Live Training

i. The user has to record 20 utterances of a new word of his choice.

ii. Apply the training process which was used above on the training data plus the new word.

iii. Save the final model. Now the training module is capable of detecting the new word.

RESULTS

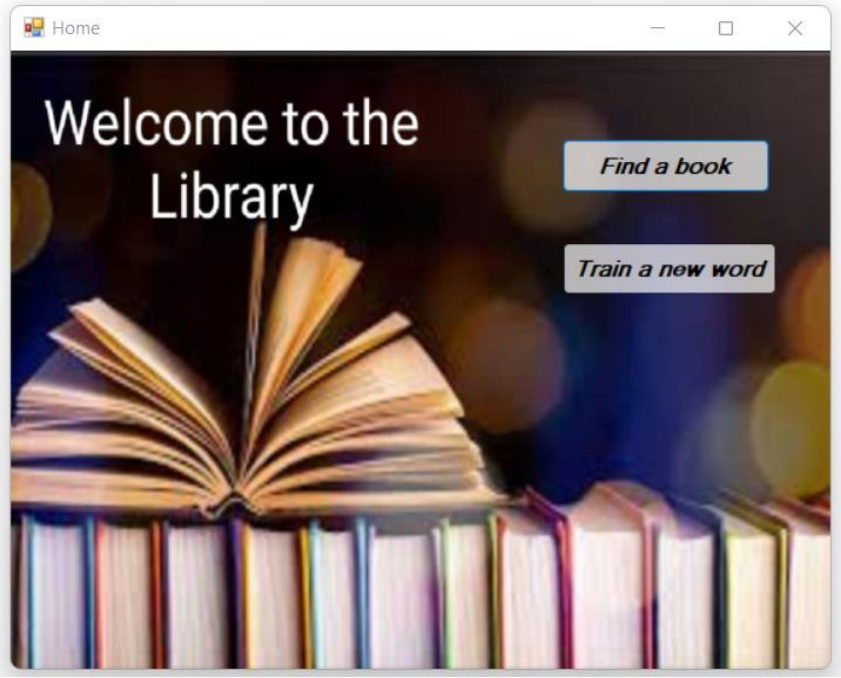


Fig: Homepage

Searching for a book, borrow or return.

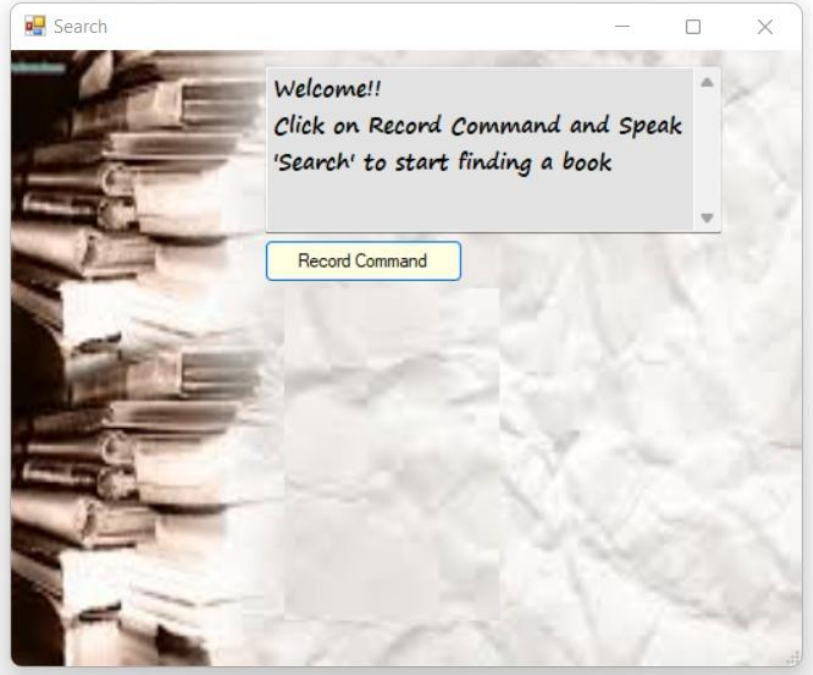
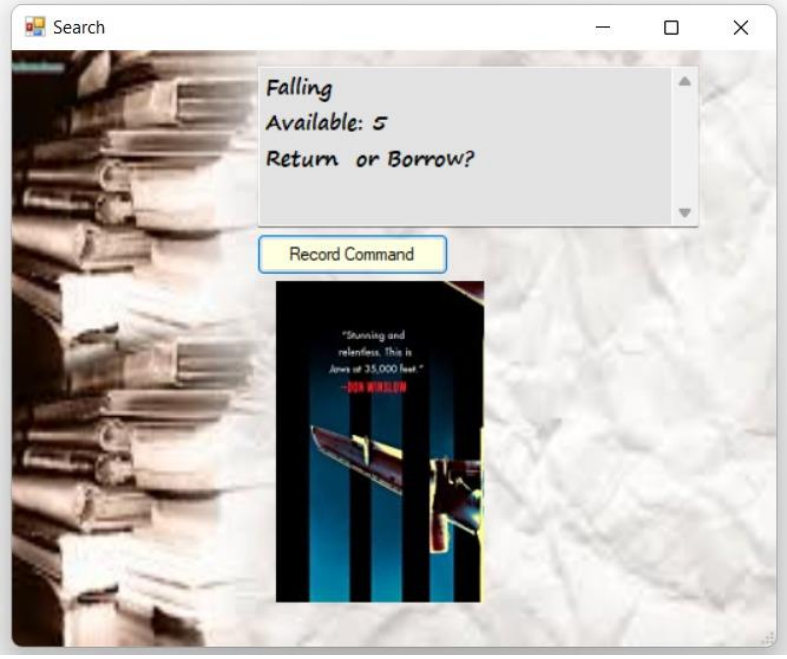


Fig: Search page



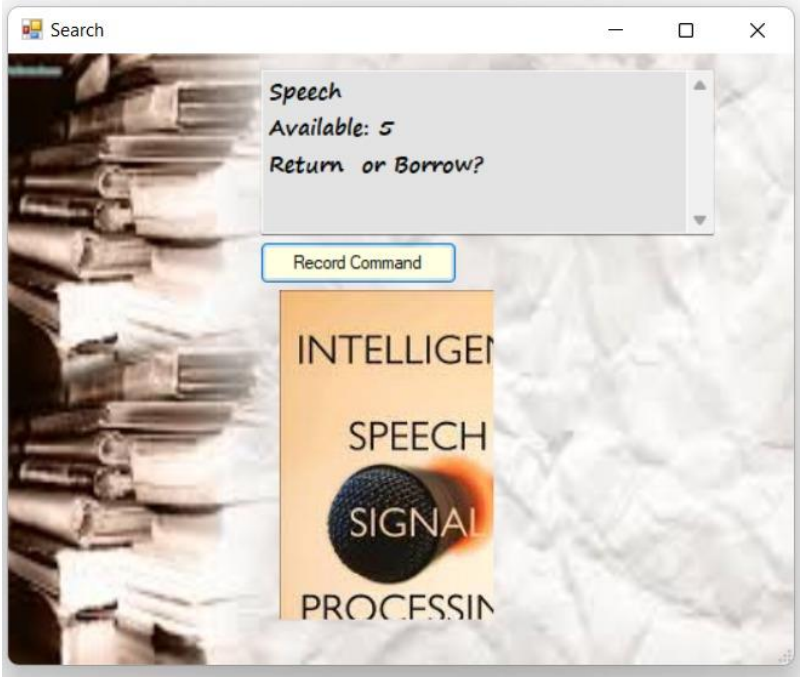


Fig: Searching for a book

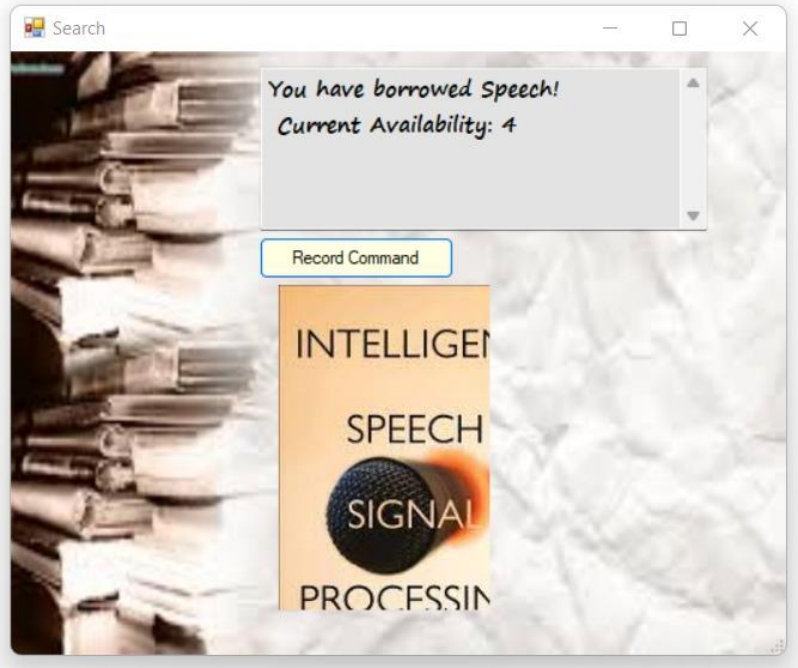


Fig: Borrowing a book

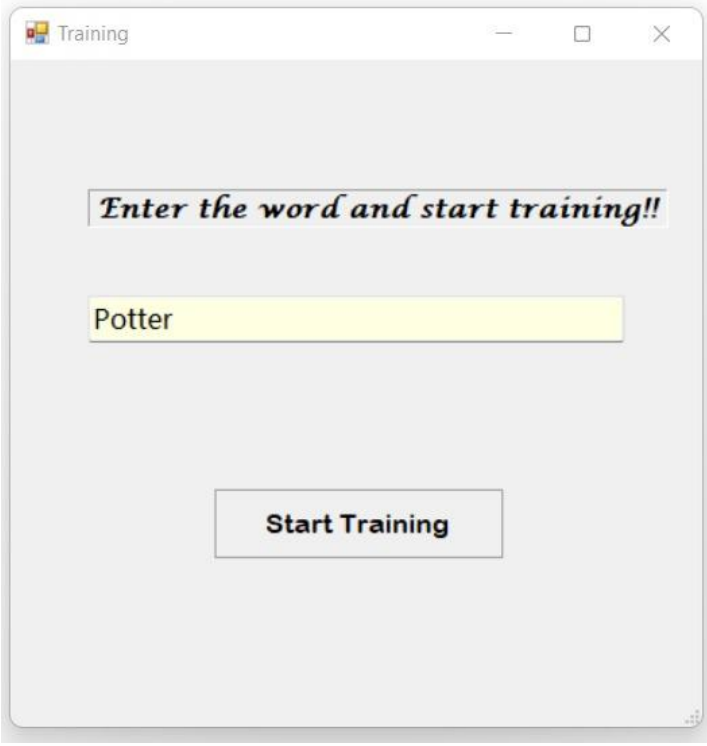
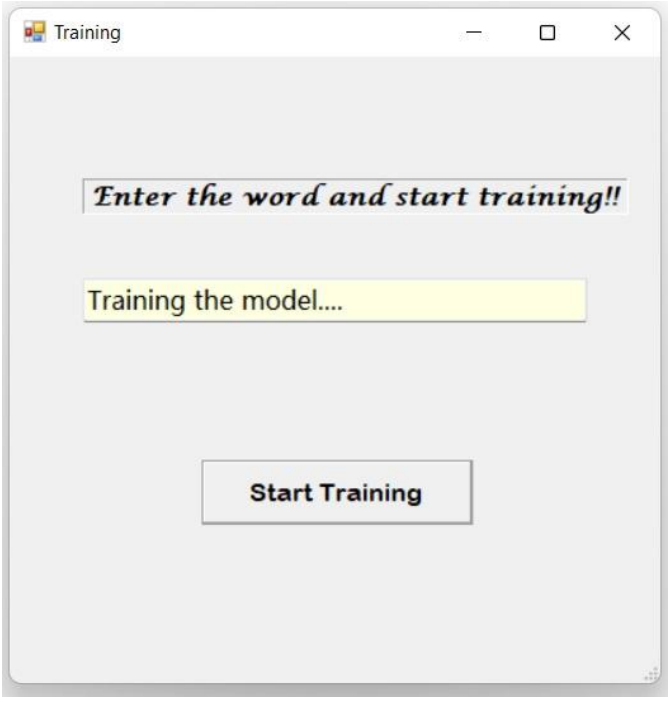
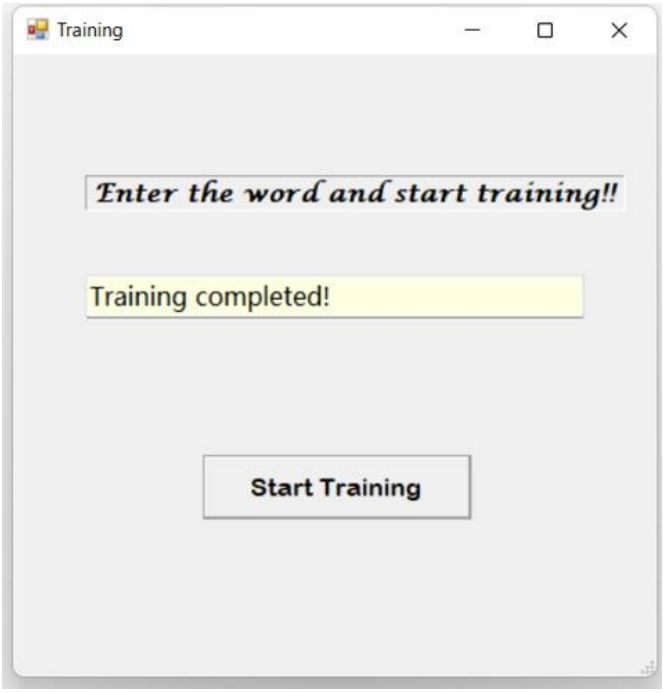


Fig: Training to add a new Book name to the database

Once the recording of the word is complete, the model is re-trained.





CONCLUSION

We are able to display the book information accurately based on the voice input. Additional features- borrow and return are working perfectly.