UNIVERSITY OF THE PHILIPPINES VISAYAS
COLLEGE OF ARTS AND SCIENCES
DIVISION OF PHYSICAL SCIENCES AND MATHEMATICS

CMSC 198.1 Special Problem I
First Semester AY 2022-2023

Filipino Automatic Speech Recognition for Filipino Children Aging six to 14

Chapter 1
Assignment 2

Submitted by
Michael Vincent Dosado, 2019-04469;
Kevin Christian Lao, 2019-12279;
Allen Solomon Tam, 2019-03019
B.S. in Computer Science IV
College of Arts and Sciences
University of The Philippines Visayas
Miagao, Iloilo

Submitted to
Christi Florence C. Cala-or
Asst. Prof., Faculty of Computer Science
Division of Physical Sciences and Mathematics
College of Arts and Sciences
University of the Philippines Visayas
Miagao, Iloilo

Table of Contents

I. Background and Rationale of the Proposal

Speech recognition or speech to text, is a technology that recognizes speech data, allowing the voice as the main means of communication with the computer. Speech recognition apps have been a huge help for writers to boost their writing production and give opportunities for aspiring writers with physical impairements. Speech Recognition combines the fields of linguistics, math, and statistics. It detects human speech and translates it into text, and its aspects include a customized dictionary, acoustic modeling, pronunciation modeling, a decoder, and output - these are to be considered in creating Speech recognition technology (IBM Cloud Education, 2020).

Globally, youth internet users has increased by 40% in the prior pandemic 2019. Today, experts have suggested that children in the primary age have a relative screen time of that of an adult - spending about more than five hours per day on the screens (Scanlon, 2020). Simultaneously, Filipino children at the primary have trouble in reading. Nine out of 10 Filipino students have low proficiency in reading - and this has been looming the basic education system of our country (Oseña-Paez, 2022). Additionally, the Philippines placed the lowest in the ranking when it comes to reading comprehension of students, according to the 2018 Programme for Internation Student Assessment(San Juan, 2019).

The team found it timely and productive to help in addressing the learning crisis on our basic education system, specifically the lack of Filipino reading and writing skills,

thus the team aims to develop an English-Filipino Automated Speech Recognition for Filipino children in the primary age. The methodology of the project will be discussed further in the discussion.

Nowadays, there is increased production of speech recognition software with excellent performance for adult speech, however, the quality of speech recognition for children yields relatively poor results due to higher error rates in terms of variabilities that includes changes in vocal tract growth, the formant frequency, pronunciation, and range of their vocabulary (Hagen et al., 2007). Moreover, it's aside from the children's voice being thinner, their vocal chords are smaller, their vocal tracts are thinner and shorter, and the larynx is still developing. As a result, the speech patterns are very different from those that have fully grown vocal chords. Especially, the looming learning crisis in the Philippines that has relatively affect the reading and speaking abilities of Filipino youth in the primary age can contribute to the error rates of speech recognition software available in the internet. In addition, there has not been much evidence of speech recognition applications that focus on Filipino Children and the Filipino Language.

Furthermore, despite the fact that the younger generation is more native to digital technology. Parents and teachers are painfully aware of how difficult it can be for young "remote learners" to use the desktop computers or laptops - which are not efficiently designed for their growing physique but are much needed to fulfill the potential of educational technology. Another difficulty arises from the fact that children's voices, languages, and behaviors are much more complex than those of adults, and the speech recognition technology behind huge IT companies was never intended for them to use.

Therefore, the researchers have concluded that the project will be designed as a mobile application which is accessible and convenient to children in primary age. The team aims to provide a speech recognition mobile app for entertainment and educational purposes for Filipino children aging 6-14, and that is designed for their intonations and pronunciations. The mobile app framework will be discussed further.

II. Statement of the Problem

There has not been much evidence of speech recognition applications that focus on Filipino Children and the Filipino Language. A speech recognition software directed to detect, correct, and monitor Filipino vocabularies and pronunciations can greatly help in educational activities of students in school, especially in times of educational crisis in the Philippines. Today, the Philippines is under a learning crisis and yields low proficiency in reading and speaking abilities of Filipino students in the primary age (six to 14 years old). This is a threat in the quality of education in the Philippines and can reflect negatively to other countries if not addressed immediately. Thus, the team aims to help in addressing the problem by developing an English-Filipino Automated Speech Recognition for Filipino children in the primary age.

III. Significance of the Study

The study is made with the aim to provide Filipino automatic speech recognition (FASR) for Filipino children to maximize their learning experience. Specifically, the result of the study will benefit the following:

Students

This study will help students fine-tune their Filipino intonations and pronunciation and learn the Filipino language independently – allowing a meaningful way for them to

practice and determine the things they need to work on. Students can learn in a more independent approach with the help of FASR. For Filipino language learners, speech recognition allows a low-stress and free environment for them to learn difficult Filipino words without worrying about feeling embarrassed. Additionally, students will improve their writing and reading skills when using educational programs with a speech recognition process. In a study conducted by M. Raskind and E. Higgins (1999), 50 students with learning disabilities were randomly divided into speech recognition and computer instructions groups. They concluded that speech recognition group had significantly more improvement than the other group and their increased phonological processing results had significant differences with spelling and reading comprehension measures.

Schools

This study can help educational institutions such as primary-elementary schools, learning centers, and colleges to educate Filipino language learners. Filipino automatic speech recognition can provide a more interactive method for students, along with the aforementioned benefits of FASR in either remote or face-to-face setup. In an article by Oseña-Paez of Manila Times (2022), the Philippines is currently experiencing learning crisis, wherein nine out of 10 children are not proficient in reading and writing in Filipino language. The significance of speech recognition in the learning capabilities of students according to M. Raskind and E. Higgins along with new speech recognition algorithms can help curb the gaps in the Philippines' learning crisis. However, the team does not intend to fill in the role of the trainer or teacher, instead, our goal is to provide a tool that will aid the educators or to implement it in their class.

Software Programmers

This study aims to provide software programmers and researchers the reference concerned with speech recognition development, specifically in the field of Filipino linguistics. Also, the software developers can study the algorithm the team created and improve it from there on.

IV. Objectives

1. General

a. The purpose of this study is to design a system that will improve speech recognition for children in the Filipino language. To optimize the voice recognition capabilities of softwares to cater the unique needs and circumstances of Filipino children.

2. Specific

a. Optimize voice recognition capabilities of softwares to cater the unique needs and circumstances of Filipino children.

b. Implement .the Linear Prediction Cepstral Coefficient (LPCC) anda Mel-Frequency Cepstral Coefficient (MFCC) for creating the Filipino speech recognition.

c. Establish custom language models, pronunciation models, and tone models in Filipino for Filipino children.

V. Scope and Limitations

According to Elimat and AbuSeileek (2014), ASR yielded outstanding results in academic measures and provides a great opportunity in learning and teaching intonation than regular instructions.

The team proposed the project “Filipino Automatic Speech Recognition App” for the purpose of creating an software application that students and teachers can use to learn and teach Filipino (and English). Additionally, the purpose of the project is to create or improve accuracy and efficiency of speech recognition in translating to text (mostly Filipino), rectify mispronunciation, and create absolute Filipino model for the speech recognition. Wherein, to achieve this, the team included the following sub-sections: (a) Project Deliverables, (b) Acceptance Criteria, (c) Project Exclusion, and (d) Project Constraints.

(a) Project Deliverables

A Filipino automatic speech recognition app is essential for Filipino Students and teachers because of the following:

(i) Students can fine-tune their pronunciation.

(ii) Provides a flexible schedule for students to practice the Filipino language.

(iii) Students can freely learn the language in a low-stress environment which can lead to:

a) Positive feedbacks on the learning process of the students.

b) Increase the self-confidence and independence of the students.

c) Avoids anxiety of the student due to getting it incorrectly or getting embarassed.

By the end of the project development, it is expected that the team has delivered the following:

(iv) A speech recognition application that is able to analyze speech signals and extract features

(v) An acoustic model able to capture audio as input and convert it prababilities over set of characters in the Filipino alphabet.

(vi) A speech recognition application capable of providing safe and working back-end system for translating data.

(vii) A speech recognition that has an easy-to-use UI/UX for students.

(viii) A speech recognition application that is available for students and school to download or implement.

(b) Acceptance Criteria

The application will include only speech data during the time the automatic speech recognition is running. The speech recognition system secure the speech data after it is transcribed to text for machine learning. Identify Filipino words and transcribe to text with exceptional accuracy.

(c) Project Exclusion

The study will not cover children below six years old, since it is still an inadequate age to hold the responsibility of using a gadget, however, kids at this age ultimately needs the attention and immersive interactions with parents, peers, and siblings. It is better to play with other kids at their age

instead of immersing themselves with smartphones at a vulnerable age(Mary, Courage, & Mark, 2010).

In addition, children who are suffering with Dysarthria or difficulty in speaking are also not covered in this project, since it will require a different set of designs in order to cater their specific needs.

The project also focuses on creating a Filipino speech recognition app, there it will not include the following:

(i) Advertisement tracking. Personal information will neither be stored nor tracked in a speech recognition system. Therefore, there are no ways for this app to have advertisement services.

(ii) Storing of individial information. There is no need of getting the personal information from users as the application is mainly for speech recognition.

(d) Project Constraints

The proposed speech recognition application will only include English and Filipino Language Models The reason for choosing English and Filipino language as the medium for this project is because both English and FIlipino are the primary modes of instruction in primary and secondary schools in the entire Philippines. Hence, making it easier for Filipino children to easily adopt and utilize the finished product the moment it will be deployed.

In the speech recognition system, the constraints of this project are as following:

(i) The team will utilize the and Linear Prediction Cepstral Coefficient (LPCC) and Mel-Frequency Cepstral Coefficient (MFCC) , which are the most commonly used in speech recognition systems.

(ii) The team will use the flutter framework for developing the mobile application.

(iii) Due to limited workforce in the team, daily progress of the project may also be limited.

(iv) Future delays in the development may happen due to unprecedented factors.

References

Bjayou, F. (2020, February 6). *Advantages and drawbacks of Voice Recognition Technology*. InsideTelecom. Retrieved October 7, 2022, from https://insidetelecom.com/advantages-and-drawbacks-of-voice-recognition-technology/

Bouchrika, I. (2022, October 6). *What Age Should a Child Get a Smartphone: Pros and Cons of Early Phone Use*. Research.com. Retrieved October 8, 2022, from https://research.com/education/what-age-should-a-child-get-a-smartphone

Courage, M., & Howe, M. (2010, April 7). *To watch or not to watch: Infants and toddlers in a brave new electronic world*. ScienceDirect. Retrieved October 9, 2022, from https://www.sciencedirect.com/science/article/abs/pii/S0273229710000122?via%3Dihub

Elimat, A. K., & AbuSeileek, A. F. (2014). *Automatic speech recognition technology as an effective means for teaching pronunciation*. jaltcalljournal. Retrieved October 11, 2022, from https://files.eric.ed.gov/fulltext/EJ1107929.pdf

Hagen, A., Pellom, B., & Cole, R. (2007). *Highly accurate children's speech recognition for interactive reading tutors using subword units*. ResearchGate. Retrieved September 21, 2022, from

https://www.researchgate.net/publication/222574683_Highly_accurate_children's_speech_recognition_for_interactive_reading_tutors_using_subword_units

IBM Cloud Education. (2020, September). *Speech Recognition*. IBM. Retrieved September 21, 2022, from https://www.ibm.com/cloud/learn/speech-recognition#:~:text=Speech%20recognizers%20are%20made%20up,to%20determine%20the%20appropriate%20output.

Oseña-Paez, D. (2022, August 10). *Why 9 out of 10 Filipino children can't read*. The Manila Times. Retrieved October 11, 2022, from https://www.manilatimes.net/2022/08/10/opinion/columns/why-9-out-of-10-filipino-children-cant-read/1854046

Raskind, M., & Higgins, E. (1999, December). *Speaking to read: The effects of speech recognition technology on the reading and spelling performance of children with learning disabilities*. Springer Link. Retrieved October 11, 2022, from https://link.springer.com/article/10.1007/s11881-999-0026-9

San Juan, R. (2019, December 3). *Philippines lowest in reading comprehension among 79 countries*. Philstar. Retrieved October 8, 2022, from https://www.philstar.com/headlines/2019/12/03/1974002/philippines-lowest-reading-comprehension-among-79-countries

Scanlon, P. (2020, September 16). *Voice recognition technology doesn’t work for kids - and it's hurting them in the classroom*. WeForum. Retrieved October 8, 2022, from

https://www.weforum.org/agenda/2020/09/voice-assistants-speech-recognition-in-the-classroom/