Speech Recognition Algorithm of Major Nigerian Languages (Yoruba, Hausa, Ibo) Using K-NN

Taiwo Mauyon KUPONU

LCU/PG/002632

Being a MSc Thesis Submitted to the Department of Computer Science, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Oyo State, Nigeria

In Partial Fulfilment of the Requirement for the Award of Master of Science (MSc) Degree in Information Science

2023

Certification

This is to certify that Taiwo Mauyon KUPONU with matriculation number LCU/PG/002632 carried out this research work titled "Speech Recognition Algorithm of Major Nigerian Languages (Yoruba, Hausa, Ibo) Using K-NN" in the Department of Computer Science, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Oyo State, for the award of Master of Science (MSc) in Computer Science and that this has not been previously submitted.

…….. .....

Dr. Wilson Sakpere **Date**

Supervisor

…….. .....

Dr. Wilson Sakpere **Date**

Head of Department

Dedication

This research work is dedicated to God, my parents, and my siblings.

Acknowledgement

Foremost, I would like to express my gratitude to the leadership of the Lead City University, Ibadan and also acknowledge the libraries used for creating a medium for us to acquire knowledge for self-reliance.

I acknowledge my supervisor Dr. Wilson Sakpere, for the continuous support for M.Sc study and research, for his motivation, enthusiasm, to guide me through the research. The immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life. Besides my supervisor, my sincere thanks go to the Head of Department Dr. Wilson Sakpere, the Postgraduate Coordinator Dr. Azeez Waheed, all other lecturers and staff members in the department of computer science for their guidance encouragement, and insightful comments.

I thank my fellow coursemates, Mr. Oladejo Samuel Adetunji, Mr. Ayomide Feyi-Robinson, Mrs. Temilade Fashina, Mr. Kingsley Efekodedo, Mr. Wasiu Olayinka and others numerous to mention. I also thank Mr. Oluseye Akinmoluwa and Mr. Faruq Ayoade for their kind support on this project.

I thank my parents Rt. Rev'd Dr. S.G Kuponu and Mrs Victoria Kuponu for their guidance, parental advice and commitment instilled in me towards learning. May you both live long in good health to enjoy and the reap fruits of your labour. My sincere thanks to my friends and loved ones for the understanding, sacrifice, and prayers during my study. Also, to my siblings Mrs. Funmi Oni, Mrs. Elizabeth Akinyanmi, Mrs. Esther Olowookere and Mrs. Kehinde Akinlabi for their continuous support and advice.

"Even though the above-mentioned institutions and persons have assisted in the process of this research work, I alone stand responsible for the errors, if any, found in the work"

Abstract

This study addresses the crucial need for enhanced voice recognition systems in the realm of human-machine interfaces, particularly with a focus on accent identification algorithms, and their application in the context of Nigerian English speakers. The research aims to improve the accuracy and efficiency of speech recognition for these major Nigerian languages using KNN to increase the efficiency and accuracy of accent identification by using a trained data with the three major Nigerian language. Voiced audio samples of speakers from these tribes speaking English and various platforms such as news media and radio recordings was scrapped and recorded and extracted. The audio data was then preprocessed and transformed from the time domain to the frequency domain using the Fourier transform. Matlab R2015A was employed for model training, encompassing input reading, window size and hop size definition, and noise reduction techniques such as high-pass filtering and spectral subtraction. For feature extraction, Mel Frequency Cepstral Coefficients (MFCC) were computed for each audio frame, subsequently aggregated to create fixed-length representations for each dialect sample for about sixty seconds in order to ensure uniformity in the inputs. The model underwent training with a classification algorithm KNN, followed by evaluation, which gave an accuracy rate of 84%. This result indicates that the model proficiently predicts the dialects within the context of English speech. The study's outcomes signify substantial progress in the development of an accent detection model tailored to the major Nigerian tribes: Yoruba, Hausa, and Igbo. The research is a significant stride toward more precise and effective voice recognition systems for Nigerian English speakers, contributing to the broader advancement of human-machine interfaces in an increasingly technology-driven world. It is recommended that future research explores alternative feature extraction techniques, particularly deep learning-based approaches capable of automatically learning relevant features from raw audio data.

Keywords: Accent. Accuracy, Algorithm, Dialect detection, Feature extraction, Fourier transform, Performance, Speech recognition, Voice recognition

Word Count: 298 Words

Table of Contents

List of Tables

List of Figures

List of Acronyms

AI-	Artificial Intelligence
AM-	Acoustic Model
ASR-	Automatic Speech Recognition
CMS-	Cepstral Mean Subtraction
DNN-	Deep Neural Networks
DTW-	Dynamic Time Warping
GMM-	Gaussian Mixture Model
HMM-	Hidden Markov Model
KNN -	K-Nearest Neighbour
LPC-	Linear Predictive Coding
LRE-	Language Resources and Evaluation
MFCC-	Mel Frequency Cepstral Coefficient
ML-	Machine Learning
MLP-	Multi-Layer Perceptron
NLP-	Natural Language Processing
RASTA-PLP-	Relative Spectral Transform-Perceptual Linear Prediction

SI-ASR-	Speaker-Independent Automated Speech Recognition
SVM-	Support Vector Machine
VQ-	Vector Quantization