A COMPUTERIZED APPROACH FOR EMOTION RECOGNITION USING VOICE

R.A.D.G.D.KUMARA

189463P

Faculty of Information Technology

University of Moratuwa

Sri Lanka

August 2021

Msc in IT, Faculty of IT, University of Moratuwa

I the undersigned announce that this project report entitled **"A COMPUTERIZED APPROACH FOR EMOTION RECOGNITION USING VOICE**" is based on my work done while studying under the direction of **Mrs. Indika Karunarathne** .I affirm the statements made and the conclusions drawn are the result of my research work. I continue to confirm that

II. The work contained in this report is first and foremost done by me under the supervision of my supervisor.

III. The work has not yet been submitted to any other institution for any other degree / diploma / certificate at this university or from any University of India or abroad.

IV. We have followed the guidelines provided by the university in writing this report.

V. Whenever we use building materials (data, theoretical analysis, and documents) from other sources, we provide them with the appropriate credit in the report text and provide their details in the index.

Student Name R.A.D.G.D Kumara

Student ID-189463P

Supervisor Name- Mrs. Indika Karunarathne

Supervisor Signature

Date

Acknowledgement

This project would not have been possible without the support of many people. Many thanks to my adviser, Mrs. Indika Karunarathne, who read my numerous revisions and helped make some sense of the confusion. I thank the University of Moratuwa, Faculty on Information Technology for giving me a Dissertation Completion Fellowship to complete this work. And finally, I thank the parents, and the many friends who endured this long process with me, always giving of love and affection.

ABSTRACT—

The purpose of this paper is providing guess information about emotion status of human using voice cut, previous machine learning researches for emotion status of human using voice cut and accuracy and recall of each research and comparison of current research and previous related researches. deeply discuss how to use machine learning for emotion status of human using voice. Firstly, explain ML using many categories Supervised learning algorithms, Unsupervised learning algorithms like that logistic regression, random forest, svm, gaussianNB, decision tree and k-nn .Python scipy provide facilities for analyze voice and provide frequency spectrum and amplitude of voice . fast Fourier transform (FFT) is providing maximum amplitude from all frequencies and frequency of maximum amplitude. The mel scale is a scale of pitches that human earshot usually observes to be intermediate from each other. Maximum mel value, maximum amplitude from all frequencies, frequency of maximum amplitude, age of user and gender are used to find out emotion status human. Part of this research is analyzing various type of voices according to emotional situation. Then train and test model using Machine learning for predict what emotional situation of is given voice. Elicited emotional speech database are creating emotional situation artificially by accumulating fact from the talker. emotional states are identified by researcher after features are trained and evaluated by machine learning models. The best recognition rate is reported by SER between 75% and 82% on random forest and recognition rate between 75% and 77% on decision tree. From this result, we can see that Decision-making algorithms (random forest and decision tree) often perform better with our data base. Therefore, we concluded that the K-Nearest Neighbor (accuracy 50%) , svm (46%), GaussianNB (46%) and logistic regression (46%). Rule based algorithms are good for recognize emotion status using voice. Random forest is best machine learning algorithm for recognize emotion status using voice according to our experiment. Random forest has good test accuracy (between 80% to 90%).

Table of contents

1. Introduction 1
2. Background and Motivation 5
3. Aim and Objectives 9
4. Related works 10
5. Methodology 14
   - Data preprocessing 14
     - Frequency spectrum 15
     - Fourier transform 15
     - fast Fourier transform (FFT) 18
     - Mel spectrogram 18
   - Machine learning classifiers 20
     - Logistic Regression 21
     - Decision Tree 24
     - K-Nearest Neighbor 26
     - Random Forest 28
     - GaussianNB 30
     - svm 32
6. Implementation 35
   - Voice clip name format 35
   - Voice clip features 36
     - Maximum amplitude 36
     - frequency’s amplitude 37
     - Frequency of maximum amplitude 38
     - frequency’s amplitude 39

Frequency of maximum amplitude 39
Mel value 40
Data features 41
Machine learning classifiers 42
Logistic Regression 42
Decision Tree 43
Random Forest 43
K-Nearest Neighbor 44
GaussianNB 44
svm 45
7. Results and analysis 46
test size 46
train size 46
Training accuracy 47
Test accuracy 47
8. Conclusion 49
9. Reference 50

List of figures

Figure 2. 1-face detection 5
Figure 2. 2-waveform 6
Figure 2. 3-sense of blood 6
Figure 2. 4 7

Figure 5. 1(approach) 14
Figure 5. 2-time domain 16
Figure 5. 3 16
Figure 5. 4-Fourier transform 17

Figure 5. 5-sigmoid function 22
Figure 5. 6-Decision tree example 25
Figure 5. 7 [15]-KNN process 27
Figure 5. 9 29
Figure 5. 10 [19] 31
Figure 5. 11 [22]-svm example 33

Figure 6. 1-file format decompose 35
Figure 6. 2 file format 36
Figure 6. 3-maximum amplitude 36
Figure 6. 4-Frequency spectrum 37
Figure 6. 5- frequency's amplitude 38
Figure 6. 6-Frequency of maximum amplitude 38
Figure 6. 7 39
Figure 6. 8- frequency's amplitude 39
Figure 6. 9- Frequency of maximum amplitude 40
Figure 6. 10-Mel value 40
Figure 6. 11 40
Figure 6. 12-Import Logistic Regression classifier 42
Figure 6. 13-Logistic Regression accuracy 43
Figure 6. 14-Import Decision Tree Classifier 43
Figure 6. 15-Decision Tree accuracy 43
Figure 6. 16-Import Random Forest Classifier 43
Figure 6. 17-Random Forest accuracy 44
Figure 6. 18-Import K-Neighbors Classifier 44
Figure 6. 19-KNeighbors Accuracy 44
Figure 6. 20-Import GaussianNB Classifier 45
Figure 6. 21-GaussianNB accuracy 45
Figure 6. 22-Import svm Classifier 45
Figure 6. 23-svm accuracy 45

figure 7. 1--Training accuracy example 01 47
Figure 7. 2-Training accuracy example 02 47
Figure 7. 3-Test accuracy 47
Figure 7. 4-Test accuracy 48