Team Advisors

Hello, Laxminarayana Ramachandran
Advisor Home My Info Logout

BACK TO HOME |

Mission Folder: View Mission for 'Psychologigals'

State	Illinois
Grade	9th
Mission Challenge	National Security and Safety
Method	Engineering Design Process
Students	Anjana Ramachandran (platypus777) Sruthi Kotlo (Potter11) Divya Lidder (sciencechamp)

Team Collaboration

(1) How was your team formed? Was your team assigned or did you choose to work with each other?

We have been friends and classmates since elementary school and all three of us love science. So, the opportunity to help the members of our community using science really interested us. We started participating in eCybermission in our 6th grade and have been a team ever since. We really liked our team dynamic and we thought the competition was a good way to allow us to spend some time together.

A major hurdle was finding the time to pursue our interest in the competition because of our varied interests. Our school's science curriculum and timelines do not match up very well to the eCybermission program. So, we have learned to manage our hobbies and passions in the arts and sports and still pursue the competition, by meeting outside of school hours. We are supportive of our different interests and do not see them as taking away from our shared goals on the project.

(2) Provide a detailed description of each team member's responsibilities and jobs during your work on the Mission Folder.

There weren't exactly clear-cut titles for each member, because we worked as a very cohesive team on most of our tasks. We had our strengths and supported each other as needed.

For example, one member who was the most comfortable with computer programming was able to teach the other team members the basics of programming and code specific to our ideas. Another member was interested in statistics and helped us all understand the science behind the predictions and accuracy. She was our team resource for a thorough analysis. One member delved into the science of sound and helped refine our ideas and approach for the project as a whole.

As a team, we have shared likes. All three of us enjoy using the internet for collaborative research. We also looked forward to meeting with experts to fill in the blanks and refine our understanding.

Interestingly however, the mission folder was an area where we took on more individual ownership. Since we had to use every minute of our time together to move forward with our project, the mission folder work was quieter, meaning it was more individual work at home, especially if an in-person meeting was not possible.

As the problem involved so many cycles of experiments and model builds, the work on the Mission Folder's Design, Build, and Test sections was challenging. We found a great way to address this. We would take turns interviewing each other. We used the questions to probe and deepen each other's understanding. The resulting answers strengthened our sections and our communication skills. Final editing and overall completeness was a shared responsibility.

(3) Did your team face any problems working together? If so, how did you solve them? If not, why do you think you were able to work together so well?

Since we've been a team for many years now, we have a good dynamic and we each have a good understanding about our different working styles.

However, if a problem would arise in the future, we have enough experience working together and sorting through issues that we are confident we would address it with just better communication of expectations within the team. We also support each other through challenges and do not let any one person struggle too long with a challenge alone.

Since our eCybermission project was not part of a school program, we had to find an alternative schedule for meeting. We met every Sunday morning for about 3 hours for discussions, brainstorming and developing our experiments and methods. Some meetings were shifted to Friday night to catch up on extra work. On the off chance that no one was available for a face to face meeting, we did our work on Google Drive so that we could be on a call and edit at the same time.

(4) What were some possible advantages to working together as a team on this project? How would working as individuals have made this project more difficult?

As our team chemistry was so good, we knew we would have far more fun working together than struggling alone on a project.

We have been an eCybermission team since 6th grade, and this has allowed us to bond and have a very strong knowledge of how to work together and an awareness of each others' strengths and weaknesses. We could use our strengths to an advantage when trying to split up jobs and get work done. Using this knowledge, we would delegate jobs that apply to each others strengths for an accurate and high-quality project. Finding time to handle all this work alone would have been next to impossible.

In our experimental design and hypotheses, we collaborated by assigning sections to each person so that we could divide and conquer the work. While completing experimentation, we would assign jobs to each team member so that we could have people to execute each part of the test. We would also rotate jobs as necessary. In external meetings and presentations, we would always switch around who answered the question. If someone knew the best answer and had studied that particular area, they would answer. If someone left out a part of the explanation, another member would jump in and add it.

Our team was very well balanced, as we loved to work together, but we were also able to sort out our problems if there were any. We are also classmates, which allowed us to discuss the project during the free time that we had. We were able to easily divide work among ourselves. As a team, we enjoyed meeting with experts and reaching out to our community. Going to visit professors from universities really interested us and helped us gain a better understanding of our project.

Engineering Design

Problem Statement

(1) What problem in your community will your team attempt to solve using the engineering design process? Why did your team choose this problem to try to solve?

In our school and in our local community, there is a rising issue of mental health. Suicides have increased substantially over the past decade. Depression threatens to be a significant problem in society and our research shows that it is the fourth most prevalent mental disorder according to the World Health Organization (WHO).(Mathers, D.M. et al 2008). Currently, more than 264 million people worldwide are affected by depression!

Low to middle income populations are hard hit. Almost 85% of people do not get a diagnosis or treatment due to a lack of resources and the negative social stigma, meaning many suffer without aid.

According to the National Institute of Mental Health (NIMH), there are currently 17.3 million adults in the United States, which is about 7.1% of the population, who are affected by depression. NIMH also stated that 2.3 million adolescents, just in 2017, suffered at least one major depressive episode. (https://www.nimh.nih.gov/health/statistics/index.shtml).

It was shocking when our research showed that, according to the Centers of Disease Control, the second leading cause of death of people from ages 10-34 is suicide!

Even though awareness and respect for mental health has increased, our research has shown that solutions for quick and effective early warnings are lacking. Doctor-patient conversations and different scale analyses are not accurately diagnosing the depression at early stages and people with depression are under representing the symptoms.

Traditionally, depression is diagnosed by patient–doctor conversations and scale analyses such as Beck Depression Inventory (BDI), the 8-and 9-point Patient Health Questionnaire (PHq-8 and 9). (K. Kroenke, R.L. Spitzer, (2002) K. Kroenke, T.W. (2009). However, these methods, to a certain extent, are undermined as patients may hide their true feelings.

Clinical psychologists have often faced difficulties in diagnosing depression due to their muffled speech. Many studies suggest that there is a close relationship between voice patterns and emotions and stress (D.R. Ladd, (1985) K.R. Scherer, (1986).; K.R. Scherer, (1991).

We found it interesting when our research showed that compared to normal people, depressed individuals exhibit slow, monotonous, and less fluent voice and speak anxiously. This gave us a hint that a voice-based extraction of acoustic features could be a useful tool to distinguish between normal and depressed states in a subject. The recent advances in voice pattern recognition gave us hope that a machine learning-based detection or diagnosis of depression could be a powerful tool in early detection.

We found that acoustic features associated with voice have been well studied and mainly include Zero Crossing Rate, Harmonics to Noise Ratio, and Mel-Frequency Cepstral Coefficients (MFCC's) and out of these MFCC's have been extensively used in speech recognition and considered key feature in voice analysis.

With the belief that detection of depression at early stages is critical for effective therapy, we decided on pursuing this project for the ECybermission competition.

(2) Research your problem. You must learn more about the problem you are trying to solve and also what possible solutions already exist. Find AT LEAST 10 different resources and list them here. They should include books, periodicals (magazines, journals, etc.), websites, experts, and any other resources you can think of. Be specific when listing them, and do not list your search engine (Google, etc.) as a resource.

When we met with our Principal, we learned that school districts in the mid-west have faced multiple tragic suicides in the past years, with mental health issues like anxiety and depression on the rise. Additionally, many of our friends and people we know have suffered with depression and suicide attempts. The death rate due to suicide is higher than ever should be, and this has affected our school and community in many ways.

When we met with Dr. Susarla, we understood that there are many reasons as to why one may commit suicide, and a major reason is due to how people conceal their feelings to themselves. Many people believe that their mental health may not need to be prioritized and often under represent their symptoms. Even when completing a standard written test for depression with a psychologist, they may not properly regard their symptoms due to the negative social stigma surrounding mental health.

Many of these unfortunate incidents could have been prevented if that connotation could be broken and the problem diagnosed. Society's view of mental health is constantly changing, leading to more and more people with mental health issues.

The urgency of this topic is overwhelming, and the faster we can solve this problem, the more lives we can potentially save.

We learned that traditionally, depression is diagnosed by patient–doctor conversations and scale analysis such as Beck Depression Inventory (BDI), 8-and 9-point Patient Health Questionnaire (PHq-8 and 9). (K. Kroenke, R.L. Spitzer, The PHq-9: a new depression diagnostic and severitymeasure, Psych. Annals 32 (9) (2002) 509–515.[10]K. Kroenke, T.W. Strine, R.L. Spitzer, J.B. Williams, J.T. Berry, A.H. Mokdad, The PHq-8 as a measure of current depression in the general population, J. Affect.Disord. 114

(1) (2009) 163–173).

However, these methods to a certain extent are undermined in correctly diagnosing depression as patients may hide their true feelings.

We organized meetings with many experts and thinkers in the field:

1. Greg Shakhnarovich Ph.D., Professor of Computer Science at the Toyota Technological Institute at the University of Chicago.
2. Dr. Vijaya Susarla, head physician at SITA f.c. (A Clinical Psychology practice that helped us with testing data).
3. Stephanie Posey, Principal of Naperville North High School.
4. Sankar Bharadwaj, Machine Learning and Data-Science expert working in management consulting.

=================================
Other research papers we used and referenced included:

https://search.proquest.com/openview/2d9976e90a019a50e1dc188b2d8172e4/1?pq-origsite=gscholar&cbl=41539 (Anjana)

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4671760/ (Anjana)

https://jech.bmj.com/content/61/7/619.short (Anjana)

https://www.tandfonline.com/doi/abs/10.1080/21635781.2015.1085928 (Anjana)

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5563010/ (Anjana)

https://psycnet.apa.org/record/1993-25780-001 (Anjana)

https://onlinelibrary.wiley.com/doi/full/10.1002/da.22890 (Divya)

https://www.ncbi.nlm.nih.gov/pubmed/19640579 (Divya) (General practitioners are bad at detecting depression)

https://www.nimh.nih.gov/health/statistics/mental-illness.shtml (Sruthi)

http://dcapswoz.ict.usc.edu/wwwutil_files/DAICWOZDepression_Documentation.pdf
This factsheet describes the audio clips we obtained from the DAIC-WOZ database. It contains 189 folders of audio files of clinical interviews with 77 normal and 31 depressed human subjects conducted by an animated virtual interviewer called Ellie. After the interviews, the Public Health Questionnaire (PHQ) 8 was used to categorize the samples into Non-depressed (control) and depressed individuals. A score of 1 to 10 is considered non-depressed and above 10 is depressed.

==============================

Further references included the following scientific papers:

Mathers, D.M. Fat, J.T. Boerma, The Global Burden of Disease: 2004 Update,World Health Organization, 2008

(https://www.nimh.nih.gov/health/statistics/index.shtml

J. Pedersen , J. T. M. Schelde , E. Hannibal, K. Behnke, B. M. Nielsen , and M. Hertz 1988. An ethological description of depression. Acta psychi-atrica scandinavica, vol. 78, no. 3, pp. 320330, 1988.

L. Fossi, C. Faravelli, and M. Paoli, The ethological ap- proach to the assessment of depressive disorders, The Jour- nal of nervous and mental disease, vol. 172, no. 6, pp. 332 341, 1984.

[4] P. Waxer et al., Nonverbal cues for depression. Journal of AbnormalPsychology, vol. 83, no. 3, p. 319, 1974).

K. Kroenke, R.L. Spitzer, The phq-9: a new depression diagnostic and severity measure, Psych. Annals 32 (9) (2002) 509–515.[10]K.

Kroenke, T.W. Strine, R.L. Spitzer, J.B. Williams, J.T. Berry, A.H. Mokdad, Thephq-8 as a measure of current depression in the general population, J. Affect.Disord. 114 (1) (2009) 163–173).

D.R. Ladd, K.E. Silverman, F. Tolkmitt, G. Bergmann, K.R. Scherer, Evidence for the independent function of intonation contour type, voice quality, and f0 range in signaling speaker affect, J. Acoust. Soc. Am. 78 (2) (1985) 435–444.

K.R. Scherer, Vocal affect expression: a review and a model for future research,Psychol. Bull. 99 (2) (1986) 143.

K.R. Scherer, R. Banse, H.G. Wallbott, T. Goldbeck, Vocal cues in emotion encoding and decoding, Motiv. Emot. 15 (2) (1991) 123–148.).

Zhu Y, Kim YC, Proctor MI, Narayanan SS, Nayak KS. Dynamic 3-D visualization of vocal tract shaping during speech. IEEE Trans Med Imaging. 2013;32:838–48.

(3) What did you find out about your problem that you didn't know before? What kinds of possible solutions already exist? Be sure to put this in your OWN words, do not just copy And paste information. Also, be sure to cite your sources.

We didn't realize that depression was so frequently misdiagnosed and that it wasn't treated appropriately due to this. We also didn't know that people were purposely avoiding being diagnosed with depression, and it was also interesting that 18-25 year old's were the most depressed age group. We didn't know the statistics were so intense for veterans and that suicides were so prevalent. After seeing papers, we knew that using voices was a feasible way to diagnose depression. Previously, we didn't think depression could affect how you speak, and didn't think sound would play into how people individually express themselves this so much.

We knew that sound was a complicated science but we didn't know it could be so useful in studying people's minds. We knew that sound had values such as frequency and pitch and loudness that could be extracted and possibly analyzed, but we didn't realize their significance and we didn't realize it could be so important to analyzing features. We didn't know what MFCC's were and how accurate they could be in characterizing sound features.

Design Development

(4) What MUST be a part of your solution? These are called the criteria. Explain what criteria are needed to solve the problem. Make sure your criteria are measureable, connected to the problem, and related to your research.

The main criterion we have set for our solution is simply that we need to be able to detect depression from a set of voice samples with a 70% or above accuracy.

According to a 2009 meta-analysis of 50,000 patients published in the Lancet, a prestigious medical journal, it was said that depression was only correctly identified by general practitioners 47.3% of the time, and we aim to be able to have a higher degree of accuracy with our final model.

Additionally our model must be able to train with a very small amount of data, have reasonable accuracy, and be objective in diagnosis. Our solution should attempt to:

1. Reduce the chances for the misdiagnosis of depression.
2. Arrive at an accurate but also objective diagnosis method.
3. Have a reasonable accuracy, speed, and timeliness.
4. Use simple computer programming methods that are accessible to us (Python vs C++).
5. Use a program or model that should be able to identify multiple pointers of depression from the voice analysis, so errors could be reduced.

We hypothesize that certain acoustic features, if not all, will immensely help in distinguishing normal and depressed states. We hypothesize that extraction and thorough analysis of acoustic features from voices of control and depressed patients will significantly aid in the detection of early stages of depression preventing suicidal rates and aiming for effective therapeutic interventions.

(5) What limits are there on your solution? These are called constraints. Does it need to be a certain size? A certain weight? Is the cost a factor? Write down all of the limits on your solution.

An important constraint for our solution is that our model has to be able to train accurately without the thousands of voice samples that are usually needed for an accurate result. We have a limited number of depressed voice samples, and it will be difficult to gather an adequate number to optimize our model.

1. We don't have many other constraints, since our model doesn't cost anything, it doesn't need to fit a certain size, weight, etc.

2. We don't have access to the patients, so we have to approach trained professionals or online databases.

3. As high school students, we are limited in our access to patients with depression, so we'd need to find a clinically accepted standard of voice samples of depressed patients already categorized online.

We need objective, multi-factor determinants that can be replicated.

(6) Based on your criteria and constraints, what is your proposed solution to the problem you chose? Explain what it will look like and how it will work. If you can, include a detailed, labeled drawing.

Now that we have realized that we would like to analyze sound to identify characteristics, we have identified three methods that we think would be accurate to analyze voice recordings.

Refer to the attached “DESIGN DEVELOPMENT FLOW CHART” for a general overview of the experimentation on our project.

Also refer to the “GLOSSARY OF TERMS” for a general introduction to terms and concepts.

Data source:
We used DAIC-WOZ data base which is published free of charge for scientific use by University of Southern California (Gratch et al 2014). This database contains 189 folders of audio files of clinical interviews of 77 normal and 31 depressed human subjects conducted by animated virtual interviewer called Ellie. After interviews Public Health Questionnaire (PHq) 8 was used to categorize the samples into Non-depressed (control) and depressed individuals. A score of 1 to 10 is considered non-depressed and above 10 is depressed.

Processing of audio files:
Since these files are 28 min in length we cut these files programmatically to 6 min in length starting from 1st min (to avoid background noise) to 7th min using audio clip.

Extraction of acoustic features such as MFCC using Python –based programming:
Acoustic features from voice samples of non-depressed and depressed subjects were extracted using the python-based Librosa package.

Information Management:
Because of a limited number of samples, we want to use some of them to train our Machine Learning models (80% approximately) and save the remaining 20% to test the model accuracy of determinations.

We expect that MFCC values and visual spectrogram analysis would have the best chance of successful classification. To verify the machine learning model output, we will also create manual statistical analysis to corroborate the results.

1. We hypothesize that that without necessarily looking at the content of one's response in a depression questionnaire, the Machine Learning model and the statistical

analysis of MFCC data should be able to make a diagnosis solely based on one's voice characteristics.

2. Based on our research, our solution could be able to isolate some specific voice-based features which could differentiate between depressed patients and healthy patients.

3. Our research showed that one can use MFCC values and zero crossing rates to classify voice samples. [talk about how MFCC values are relevant and can be used to classify voice samples.

4. We hypothesize that MFCC values and a Zero-crossing rate value could be used to determine whether a voice is depressed.

5. We also want to analyze visual features of sound files, and our research elucidates that spectrogram representations of sound files can be used to classify them as well. Therefore our solution will involve analysis of these spectrograms.

6. We would also want to be able to study some graphs manually to be able to confirm and visually demonstrate the correlation (if it exists) between MFCC values and depression/non-depression determinations.

(7) How will you test your solution? The BEST way to test your solution is to build a working model or a prototype that you can actually use. Or you can guess how your solution will work BASED ON your research. Which method will you use and why?

To test our solution, we will start with a python-based analysis of the MFCC values extracted from the established database of voice samples.

These MFCC values will also come out as numbers on which we can run an independent mathematical statistical analysis. We would call this EXPERIMENT 1.

These would include multiple methods such as the mean, standard deviation, and area under curve, and potentially a regression analysis. Based on this analysis, we will look for and identify notable MFCC values that show significant differences between depressed and non-depressed samples.

=========

We would subsequently train and test a Sequential Machine Learning model on the same MFCC markers showing statistically significant variations. The values are extracted from the same samples. We will use 80% of the extracted data for training and 20% for testing in order to obtain an above 50% accuracy of detection. We would call this EXPERIMENT 2.

=========

We see promise in being able to use a visual representation of sound to understand the differences between the voice of a depressed and non-depressed person. After our meeting at the University of Chicago, we were able to understand how spectrograms are an incredibly useful tool in voice analysis by showing frequency and energy.

After using the python-based Sequential model analysis, run on the numbers, we will attempt a different approach with an EXPERIMENT 3 based on a visual analysis.

We will convert the audio Wav files into spectrograms. We will create a convolutional neural network (CNN) which will then analyze the spectrograms of each voice sample or set of samples to determine whether someone is depressed or healthy. We will corroborate results against the other two experiments.

=========

After determining the best method, we will proceed to validate our findings during a community clinic based validation exercise by collecting the voice samples from the SITA clinical psychology pracftice's clinic with depression and classify them using both methods and measure the percent accuracy of each one in order to develop our final model based on the most accurate method. Through these steps, we plan to be able to provide a more accurate diagnosis of depression through our computer program.

Build Model or Prototype

(8) If you built a prototype or model, explain how you built your prototype or model, step-by-step including ALL SAFETY PRECAUTIONS. If you guessed how your solution would work BASED ON your research, explain important information from your research that you used to prove how your solution would work and be sure to cite your sources.

EXPERIMENT 1 (Statistical Analysis):

1. First, we requested and gained access to the DAIC-WOZ database which is mentioned and defined in both the glossary and the reference section.

2. We downloaded the database, and removed every file except the patient interviews including the transcripts and other data provided.

3. Next, we clipped each of the 135 patient interviews (.wav file originally anywhere between 7-33 minutes) to a uniform 6 minutes long. This was because we wanted to make sure that the voice files were long enough to get an accurate representation of a depressed/nondepressed voice, but also short enough that the data would not be misleading.

4. Each patient's interview from the DAIC-WOZ database was accompanied by a score from the PHQ-9 questionnaire, as opposed to being labelled as 'depressed' or 'nondepressed'. According to the questionnaire, scores 10 and up are considered depressed, while 9 and below are nondepressed. We removed voice samples of those who scored 6,7,8, and 9 because due to our small dataset, it was imperative to minimize ambiguity. We needed our 'nondepressed' files to share few to no vocal characteristics of the depressed files to optimize our model, so we used a score of 1-5 as nondepressed and 10 and above as depressed. We ended up with 89 samples of depressed/nondepressed voices.

5. After selecting the 89 patients' voices, we used the Python packet 'Librosa' to extract audio features. These included 20 MFCC values, a Zero-crossing rate, Chroma STFT, Spectral centroid, spectral bandwidth, rolloff, and root means squared error for each of the files. The methods of this extraction were not determined by us, but for every 30 second interval of the file Librosa measured each of these values, and averaged them at the end. This provided us with one mean value which was representative

of the whole file. The initial section of code in the attachment between (a)"#generating a dataset" to "writer.writerow(to_append.split())?" demonstrates how we did this.

THE FIVE STEPS ABOVE ARE COMMON TO EXPERIMENTS 1 AND 2

For our first model, we decided to use conventional statistical methods to determine if any of the (26) values we trained on were significant. We calculated the mean of all the 26 values across all depressed voices and nondepressed voices separately and graphed them to determine a meaningful difference between the two. We then conducted a series of statistical tests to analyze the data further, which will be elaborated on in the testing section.

====================================

EXPERIMENT 2 (Sequential Neural Network):

1. The first five steps in Experiment 1 were common to Experiment 2 as well.

2. We then implemented our first method of analyzing this data. We wrote a neural network in Python using the high-level API 'Keras'.

3. The code we wrote is an attachment called "SEQUENTIAL NEURAL NETWORK", and SEQUENTIAL NEURAL NETWORK - GET PREDICTIONS"

4. After creating the CSV (excel) file with all of the voice data, we preprocessed the date. First we assigned all the data to a variable named "data" and removed headings and labels. This left only raw numeric data in the variable. (b)(From "#reading dataset from csv?" to the second "data.head()"). (c)We then encoded each classification. Depressed voices were encoded as "0" and Nondepressed voices were represented by a "1". (D)Lastly, we normalized the data. "Normalizing is a process typically done to data for machine learning programs. It standardizes data so that its mean is 0 and standard deviation is 1. This allows programs to detect patterns without overexaggerating very large differences in data. For example, spectral centroid was found to have a range between 1,200 and 1,600 Hz, while MFCC 9 measured values between -2 and 9 Hz. This would be difficult for a neural network to analyze without normalization. Lastly we split the data into 80% training and 20% testing so we could judge the accuracy of our model.

5. We started to build our actual model. We used Keras' sequential model, the simplest of the models which it offered. We then decided that four layers would be the most optimal for the small dataset we had, since many layers require more data to improve a model and fewer layers would not be able to make meaningful connections with our data. Next, we used basic dense layers, where each neuron is connected to all of the neurons in the previous layers. This was because our smaller amount of data would make it difficult to justify using fancier layers. It is likely that our output would not change significantly by using a different type of layer. We assumed that what we learned from dense layers would best represent our ability to predict using the voice samples we had. Lastly, we used ReLu and Softmax as our activation functions because most of our research suggested it is overall a good choice to improve accuracy. For the same reasons, we chose Adam and sparse_categorical_crossentropy for our optimizer and loss functions respectively.

====================================

EXPERIMENT 3 (Convolutional Neural Network):

We returned to the original 6 minute voice files and converted them into colored and grayscale spectrograms using the matplotlib library and Python. With the help of an expert in the field, Mr. Sankar Bhardwaj, we wrote and used a 3-D convolutional neural network in Python to analyze the colored image data and provide a classification. We wanted to test this model to check its accuracy in comparison to our other models, and ultimately determine which model would be best using the limited training data we had. Our model was similar to our original neural network in that it consisted of an input, output and two hidden layers, but this model used an optimizer of SGD, a loss function of categorical-cross entropy and relu layers instead of softmax layers. This code is an attachment as well "CONVOLUTIONAL NEURAL NETWORK"

Test Model or Prototype

(9) Explain how you tested your prototype or model. Be sure to include every step of your testing including all safety precautions that were taken. If not stated it will be assumed no safety precautions were taken. If you are using research to guess how your solution will work, explain step-by-step how it will work and why.

EXPERIMENT 1 (Statistics):

We used a T-Test for statistical significance. The mean and standard deviation of all samples were calculated and student T test was applied to compare the differences . All tests were two tailed and results were considered significant when $P<0.05$ or 0.01.

Receiver operating Characteristic (area under curve) curve was used to evaluate the strength/probability of discriminating normal vs depression states. An excellent model has AUC near to the 1 which means it has good measure of separability. A poor model has AUC near to the 0 which means it has the worst measure of separability. In fact it means it is reciprocating the result. It is predicting 0s as 1s and 1s as 0s. And when AUC is 0.5, it means the model has no class separation capacity whatsoever.

Multiple regression analysis: Multiple regression was used to analyze how precisely the actual values of the data correlated with predicted values of the data and also to evaluate the relationship between independent variable (Normal vs depressed voice samples) and a dependent variable (MFCC).

EXPERIMENT 2 (SNN):

We tested our code by changing variables such as types of layers, optimizer, and loss function, and we found that our original assumption was correct. Our dataset was so small that any change we made did not have a measurable effect on our final accuracy. For this reason we did not stick with any changes we made, and decided to keep our original neural network.

We found that 30 epochs and a batch size of 128, however, significantly improved our results, so we changed our original code from 20 epochs and a batch size of 64.

First, we separated the voice samples from the DAIC-WOZ database randomly. We took 80% of the voice samples and used them as our training data, setting aside 20% of the voice samples in order to test the accuracy of our program in its diagnosis. We trained the model on the 80% of voice samples for the program to learn trends and then tested its accuracy on the other 20% of samples. After separating the data and feeding in the testing data, our program had an accuracy of 75%. However, our program needs as much training data as we can supply. The more data that the program has, the more accurate it can be. So, we added back the 20% of samples as testing data and used it to further train our model. Then, we used the depressed voice samples obtained from Dr. Susarla to test the program once more on the newly obtained data. Using Dr. Susarla's data was also helpful because it's outside the database, so it's from an outside source that would be useful to see if the program could also detect depressed voices outside of the database. Also, adding more data will help the computer program, because the more data the model has to learn from, the more accurate it can detect patterns in new data. From both runs of testing (the 20% and susarla's data), we extracted zero-crossing-rate values and 20 MFCC values of depressed and nondepressed voices to assess the possible differences and assess whether there are any significant differences between depressed and nondepressed voices with these values.

EXPERIMENT 3 (CNN):

We took 80% of the spectrograms of voice samples and used them as our training data, setting aside 20% of the spectrogram samples in order to test the accuracy of our program in its diagnosis. We trained the model on the 80% of voice samples for the program to learn trends and then tested its accuracy on the other 20% of samples using CNN. After separating the data and feeding in the testing data, our program had an accuracy of 63.64%. However, our program needs as much training data as we can supply. The more data that the program has, the more accurate it can be.

(10) What problems did you find with your solution? Be specific since you will need to redesign based on these problems.

EXPERIMENT 1 (Statistics): First, we tried using Microsoft excel to assess the data using statistics. The T-Test was successful in Microsoft excel, however, in order to perform other statistical methods such as ROC/AUC and Multiple Regression Analysis, Microsoft Excel was not useful. Thus, we used the Graph Pad Prism program to analyze these values using a variety of statistics to showcase our data. Compared to the R program, Graph Pad Prism was easier since the R program was tough for us to understand.

EXPERIMENT 2 (SNN): Obtaining audio files was tricky for us since it's difficult to obtain voice samples of depressed patients. Initially, we were not old enough to conduct interviews with patients or record the voices ourselves. For this reason, we gave a recording device to Dr. Susarla, so she could record her patient's voices. Thus, we needed to
EXPERIMENT 3 (CNN): We realized that we did not have the knowledge to write the code to analyze the spectrograms. Thus, we contacted an expert to help us analyze the spectrograms and write the code for us to do so.

(11) Describe all of the changes you made to your prototype or model (or proposed prototype) after your first test. Why will these changes improve your solution?

EXPERIMENT 1 (Statistics): After graphing each mean and standard deviation for each MFCC value, we were surprised with the significant differences of the average MFCC values between depressed and nondepressed voices for MFCC values 3, 5, and 16. These values also showed an area under the curve above 0.5, meaning that they have a class separation capacity. After noting the values, we made changes to our program by having it detect trends with obtained data based on these three significant MFCC values.

(12) Present the data you collected from your tests or from your research. If you tested a prototype or model then include all of the numbers you gathered during your testing and all observations you made. Use of graphs and charts is HIGHLY encouraged. If you used research to prove how your solution would work, be sure to include all of the numbers, charts, and graphs you used to make your case. Be sure that all data is related to your solution.

See attachment "Testing Results"

(13) What are your potential sources of error? Remember, this doesn't mean "Did everything work?", all tests have potential sources of error, so make sure you understand what that means. Explain how these sources of error could have affected your results.

EXPERIMENT 1 (Statistics):
When graphing the statistics using mean and standard deviation, we did not separate graphs based on whether the subjects were male or female. This may have resulted in a higher standard deviation as the characteristics of the voice vary, and perhaps if we had separated voices by gender we could have created a more accurate model.

EXPERIMENT 2 (SNN):

We would have used another data base to add more number and test different data sets. But due to copyright issues and other technical difficulties we could not access these files.

EXPERIMENT 3 (CNN):

We gave Sankar colored spectrograms to analyze for differences between depressed and nondepressed voices. However, we didn't do tests on black and white spectrograms. The color shows the loudness, so if we had also analyzed in black and white, perhaps the results may have been different because the AI would have analyzed the spectrograms on factors other than loudness, perhaps finding a different trend.

Drawing Conclusions

(14) What conclusions can you draw based on the data you gathered during your tests? Your conclusion should be related to your original problem and your testing, include the data you collected, and refer to your proposed solution.

We conducted our study to determine if acoustic features extracted from voice samples can help differentiate normal vs. depressed voices and used as biomarkers of depression.

We used sequential and convolutional neural networks to train and test the models.

Our findings from our first experiment suggest that mfcc 3, 5 and 16 were statistically significant when normal and depressed subjects were compared, out of total 20 mfcc values analyzed.
Also, we used receiver operating characteristics to determine area under curve, and multiple regression analysis to confirm the findings and discriminate between depressed and nondepressed samples.
This is a very notable and remarkable finding.

Our findings from the second experiment, the training and testing of the sequential neural network model, showed that the model was able to differentiate between depressed and nondepressed conditions at an accuracy level of 75%.

Our findings from the third experiment, the training and testing of the convolutional neural network model, showed that the model able to differentiate between depressed and nondepressed conditions at an accuracy level of 63%.

Overall, from the data, we believe that MFCC values 3, 5, and 16 are potential predictors of depression. To our knowledge, this is an important finding in the context of the present day with a large impact on early detection of depression to prevent suicide. While the models predicted at only 63% and 75% accuracy, we believe that testing with larger data sets will help reduce errors and increase accuracy rate. Compared to traditional methods of depression diagnosis with questionnaires and scales, our testing has shown that using deep learning, we are able to diagnose depression in an accurate way that can help reduce misdiagnosis of depression.

Uploaded Files:

- [**View**] **Experiment 1 Test Result - Statistical Analysis** (By: Advisor, 02/26/2020, .pdf)

  A statistical analysis is used to determine that significant differences can be determined in acoustic features that help distinguish depressed subjects from non-depressed subjects.
- [**View**] **Experiment 2 Test Result - Sequential Neural Network** (By: Advisor, 02/26/2020, .pdf)

  A Sequential Neural Network Model was built and tested on MFCC acoustic features extracted from voice samples. After 30 Epoch training runs, it was able to distinguish with a 75% test accuracy that acoustic features extracted from voice samples of subjects with depression were indeed distinguishable from the same values extracted from non-depressed subjects.
- [**View**] **Sequential Neural Network Code** (By: Advisor, 02/26/2020, .pdf)

  The SNN code
- [**View**] **Sequential Neural Network Code "Get Predictions"** (By: Advisor, 02/26/2020, .pdf)

  The code to get predictions
- [**View**] **Flow Chart for Design Development** (By: Advisor, 02/26/2020, .pdf)

  A flow chart explaining our design development is attached.
- [**View**] **Testing Results** (By: Advisor, 02/26/2020, .pdf)

  A comprehensive document of testing results from Experiments is as attached

---

Community Benefit

(1) Explain how investigating the problem your team chose will help the community. Be sure to include the impacts your research will have on individuals, businesses, organizations, and the environment in your community (if any). Make it very clear why solving this problem would help your community.

We focused on adolescent depression when investigating our problem because teenagers often underrepresented or lie about their symptoms in order to avoid attention or trips to social workers. Usually, schools administer questionnaires to students to alert school officials of students who are showing potential signs of major depression or suicide. However, the diagnosis is often inaccurate, and can lead to heightened suicide rates because students might not know how to deal with their mental health issues in a healthy way. Additionally there is a large veteran community that suffers from unrecognized and un-diagnosed depression; solving current issues of diagnosis by improving accuracy and early detection may benefit those who don't think they need help or are hesitant to reach out.

We reached out to a clinical psychologist, Dr. Vijaya Susarla PhD., about our project, and she was very impressed and offered to record voice samples of her patients for us to validate our deep learning models. During our meeting, she also mentioned that if given one patient, different psychiatrists would have different diagnoses depending on different scales and questionnaires used. Our solution could help solve this issue because it would help medical specialists come to a consensus on the diagnosis of a patient. Additionally, we met with our school principal to discuss the implementation of the solution in our school. This would help to give timely access to mental treatments to adolescents who may be showing signs of suicide but lie about symptoms or choose not to get help. We have reached out to the director of the Illinois Department of Veterans' Affairs for a meeting regarding the application of our solution in the field of veteran health.

Solving the issue of misdiagnosis of depression could decrease suicide rates in schools and in the community. Our model could be used to verify existing diagnoses of patients to see if they are receiving the proper treatment to address their needs. When meeting with Dr. Susarla, she mentioned that she left the patients alone to record for their comfort, which taught us that people may be more comfortable communicating with technology rather than humans for privacy. Additionally, veteran suicide rates are at an alarmingly high level, so being able to get these patients the care they need would be a big stride towards saving the lives of those who have been in combat. Both the adolescent community and the veteran community would be majorly benefitted by our solution to this problem.

Uploaded Files:

- [**View**] **Sita Clinic Consent Form** (By: Advisor, 02/26/2020, .pdf)

  A clinical psychologist has offered to help us with expanding our community benefit. We see it as a validation of the efficacy of the model. The attachment is a Consent Form that she is using with some of her patients to get us additional data to confirm our findings.
- [**View**] **Sita Clinic wav file** (By: Advisor, 02/26/2020, .pdf)

  This exhibit is a visual excerpt of a single patient voice sample converted to a wav file. It was sourced from the SITA clinic.

Mission Verification

(1) Does your Mission Folder project involve vertebrate testing, defined as animals with backbones and spinal columns (which include humans)? If yes, team must complete and attach an IRB approval form.

No

(2) Did your team use a survey for any part of your project? If yes, team must complete and attach a survey approval form.

No

(3) You will need to include an abstract of 250 words or less. As part of the abstract you will need to describe your project and explain how you used STEM (Science, Technology, Engineering and Mathematics) to improve your community

More than one in four high school students experience symptoms of depression, and 18-22 American veterans commit suicide daily. Doctor-patient conversations and scale analyses do not accurately diagnose depression at early stages, and depressed people are under-representing their symptoms. Hence, we want to utilize voice-based extraction of acoustic features to distinguish normal vs depressed states. We hypothesized that certain acoustic features will help distinguish normal and depressed voices, aiding in the detection of early stages of depression, preventing high suicide rates and aiming for effective intervention.

Using the DAIC-WOZ database published by the University of Southern California and python-based Librosa we extracted and analyzed vocal acoustic features by performing various statistical analyses to identify key acoustic features that distinguish depression from normal individuals. Further using Sequential Neural Network (SNN) and Convolutional Neural Network (CNN) programs we developed training and testing models that could be used to verify accuracy rates of depression detection.

Our findings suggested that MFCC values 3, 5, and 16 were significantly lower in depressed subjects as analyzed statistical tools. Out of three methods of analysis, the SNN provided the most accurate predictions at a rate of 75%. MFCC values 3, 5, and 16 in depressed individuals could be potential predictors of depression with our developed models.

In order to solve issues of misdiagnosis and under-representation of symptoms of depression, future steps include implementing SNN voice diagnosis into schools, veteran aid, and making it more accessible to the public, possibly in the form of an application.

RESULT OF EXPERIMENT 1 – STATISTICAL ANALYSIS OF EXTRACTED ACOUSTIC FEATURES

The Wav files generated from the audio files were used for a Python-based (a computer program) extraction of acoustic features such as Zero Crossing Rate, and Mel Frequency Cepstral Coefficients (MFCC). We analyzed the data statistically to determine if certain acoustic feature (s) would distinguishes depression from a normal state. Our data suggested that of 20 MFCC's extracted, only MFCC 3, 5, and 16 were significantly lower in depressed individuals compared to non-depressed subjects (red font and stars in Table 1 and figs 3A, 4A, 5A).

Other MFCC's were not significantly changed for example MFCC 10 (Fig 6 A).

Next, a Receiver Operating Characteristic (ROC) analysis was conducted. It revealed that Area Under Curve (AUC) of MFCC 3, 5, and 16 were 0.63, 0.61, and 0.69 respectively whereas MFCC 10 which is not significantly changed has AUC was 0.51 (Figs 3B, 4B, 5B, and 6B). AUC of 0.5 meant that the model has no discriminative ability between depressed and non-depressed subjects.

Finally we performed multiple regression analysis to analyze the correlation of actual values of the data with predicted values. The results showed that there is a correlation between the actual and predicted values of the MFCC 3. 5, and 16 with depressed and non-depressed individuals which were significant in the other tests. As expected the MFCC 10 which is not significantly different did not show any correlation (Figs 3C, 4C, 5C, and 6C).

Table 1: Analysis of MFCC in depressed and non-depressed subjects. ("*" and red font denotes statistically significant MFCC's)

MFCC values	Depressed (Mean ± S.D) → 31 samples	Non-depressed (Mean ± S.D) → 58 samples
1	-670.7182312 ± 57.61064359	-623.3501666 ± 47.63366157
2	101.268613 ± 17.04397142	99.74999592 ± 17.02158392
3 *	3.455296315 ± 9.202240751	9.616474198 ± 9.357447706
4	23.88700428 ± 8.718515099	24.55711574 ± 6.156947994
5 *	1.553210518 ± 5.21954796	4.513056596 ± 5.005780907
6	11.14655459 ± 6.437744967	10.64291618 ± 4.586626109
7	-2.512867958 ± 4.652040804	-0.7541774249 ± 4.440193508
8	3.416036087± 4.06051123	4.141875702 ± 2.744407829
9	2.627645195±3.351766665	2.843736975 ± 2.957794156
10	2.588276866 ± 2.671379279	2.596181298 ± 2.65466649
11	4.358039374 ± 2.300123381	4.158138447 ± 2.525268735
12	0.2464258112 ±2.25504841	0.5591302411± 2.429558409
13	3.892040574 ±2.408333609	4.471360458 ± 1.76368387
14	-0.4136840573 ±2.045423462	0.2008359595 ± 2.191492974
15	4.61324272 ±2.155055921	3.909769408± 1.661736673
16 *	0.5219774424 ± 2.324104492	1.367958285 ± 2.352055383
17	1.626248313 ± 1.895564749	1.563470854 ± 1.316335804
18	0.7194835087 ± 1.094509501	0.9991352697 ± 1.264191034
19	0.2773959628 ± 1.409168148	0.2141550713 ± 1.401578822
20	1.218331385 ± 1.636154629	1.214438957 ± 1.244544918

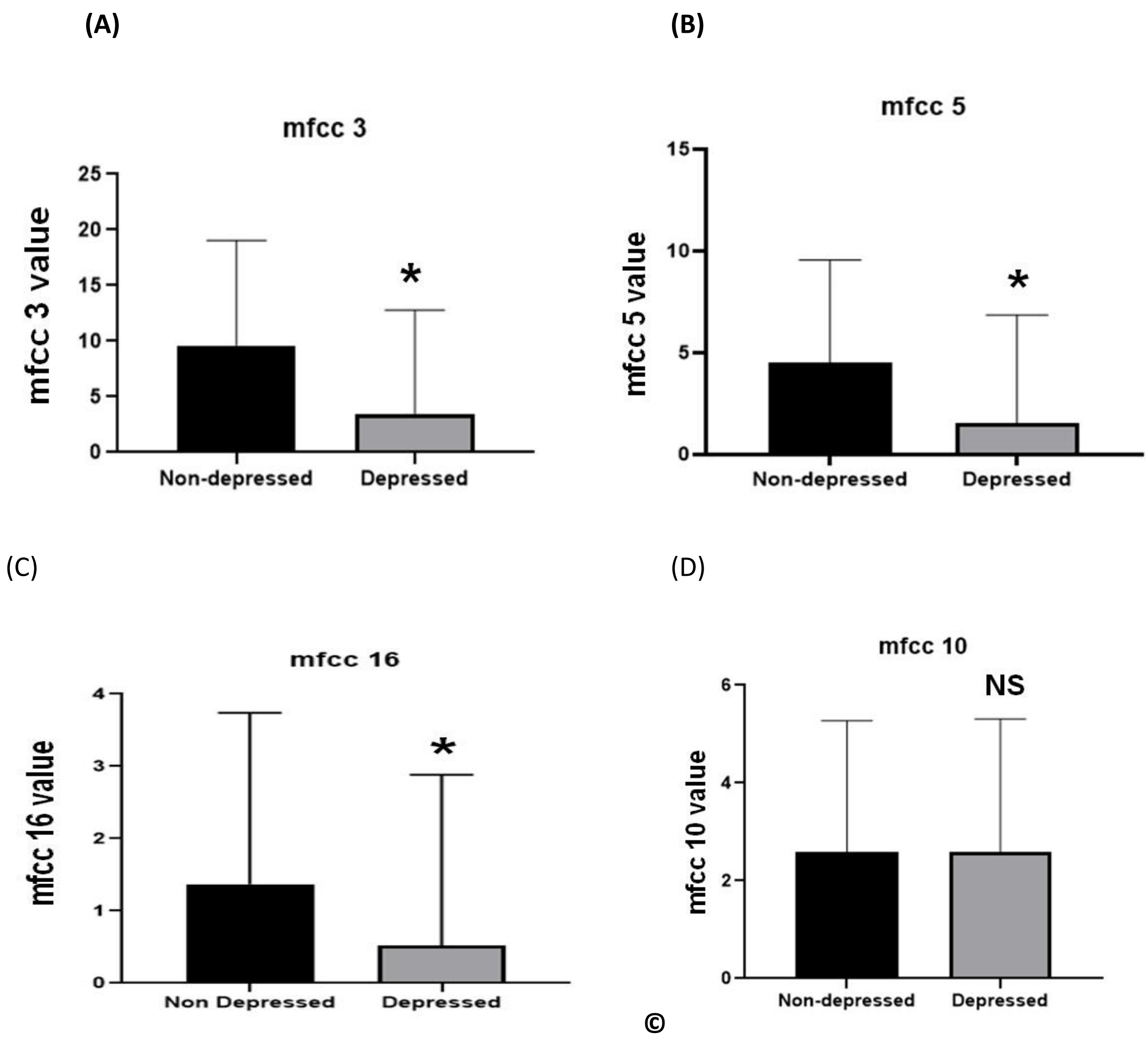
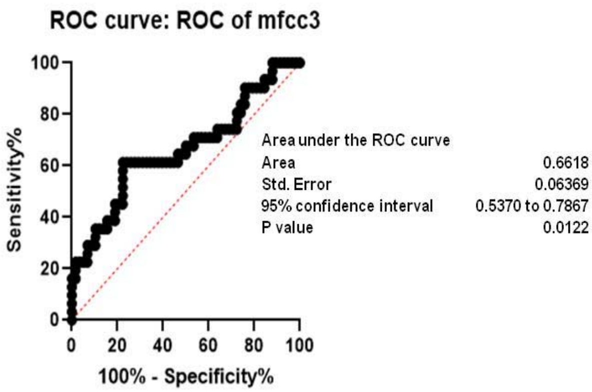
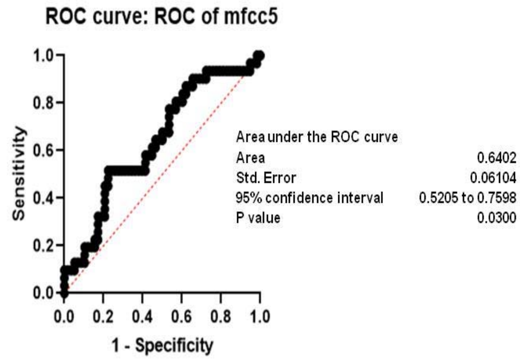


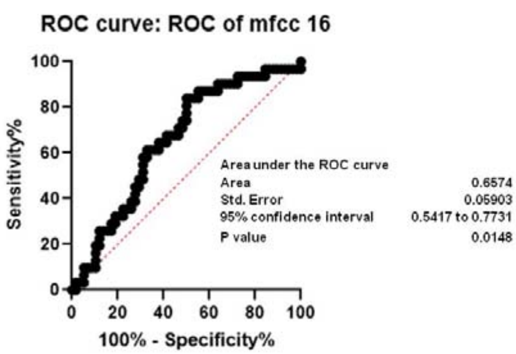
Figure 3. (A) MFCC 3, (B) MFCC 5, and (C) MFCC 16 are significantly lower in depressed individuals as analyzed with statistical T test (P,0.05) (D) MFCC 10 is not significantly changed between non-depressed and depressed individuals NS, non-significant.

(A) **(B)**





(C) (D)



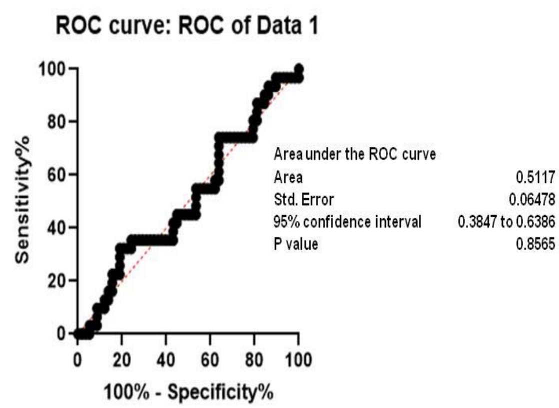


Figure 4. ROC curve analysis shows (A) 0.61 AUC of MFCC 3 B) 0.63 AUC of MFCC 5, and (C) 0.69 AUC of MFCC 16 significantly discriminates depressed vs non-depressed conditions. (D) 0.51 AUC of MFCC 10 is not significantly changed between non-depressed and depressed individuals.

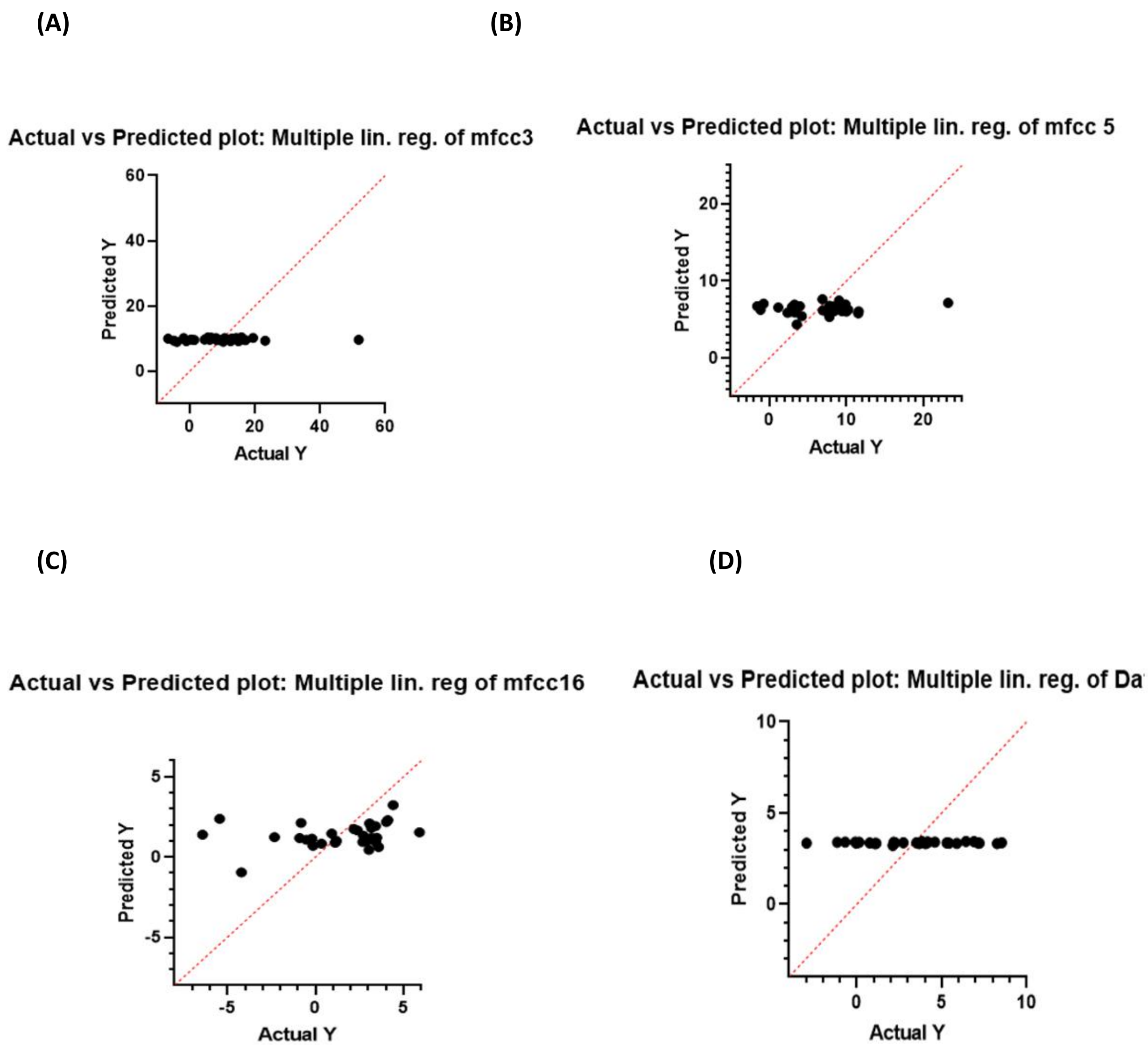


Figure 5. Multiple regression analysis confirms actual values of (A) MFCC 3, (B) MFCC 5, and (C) MFCC 16 correlate well with predicted values of 3 of non-depressed and depressed individuals. (D) MFCC 10 is not significantly changed between non-depressed and depressed individuals.

Experiment 2 Test Result - Sequential Neural Network Model

[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1
 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1]

Epoch 1/30

69/69 [==============================] - 0s 1ms/step - loss: 2.6171 - accuracy: 0.0000e+00

Epoch 2/30

69/69 [==============================] - 0s 29us/step - loss: 2.3848 - accuracy: 0.0000e+00

Epoch 3/30

69/69 [==============================] - 0s 29us/step - loss: 2.1903 - accuracy: 0.1739

Epoch 4/30

69/69 [==============================] - 0s 29us/step - loss: 2.0304 - accuracy: 0.4348

Epoch 5/30

69/69 [==============================] - 0s 29us/step - loss: 1.8936 - accuracy: 0.5507

Epoch 6/30

69/69 [==============================] - 0s 29us/step - loss: 1.7683 - accuracy: 0.6232

Epoch 7/30

69/69 [==============================] - 0s 14us/step - loss: 1.6470 - accuracy: 0.6667

Epoch 8/30

69/69 [==============================] - 0s 14us/step - loss: 1.5293 - accuracy: 0.6667

Epoch 9/30

69/69 [==============================] - 0s 14us/step - loss: 1.4131 - accuracy: 0.7246

Epoch 10/30

69/69 [==============================] - 0s 14us/step - loss: 1.2999 - accuracy: 0.7391

Epoch 11/30

69/69 [==============================] - 0s 29us/step - loss: 1.1913 - accuracy: 0.7391

Epoch 12/30

69/69 [==============================] - 0s 29us/step - loss: 1.0885 - accuracy: 0.7971

Epoch 13/30

69/69 [==============================] - 0s 15us/step - loss: 0.9923 - accuracy: 0.8116

Epoch 14/30

69/69 [==============================] - 0s 14us/step - loss: 0.9036 - accuracy: 0.8116

Epoch 15/30

69/69 [==============================] - 0s 14us/step - loss: 0.8225 - accuracy: 0.7971

Epoch 16/30

69/69 [==============================] - 0s 15us/step - loss: 0.7498 - accuracy: 0.7826

Epoch 17/30

69/69 [==============================] - 0s 29us/step - loss: 0.6868 - accuracy: 0.8116

Epoch 18/30

69/69 [==============================] - 0s 29us/step - loss: 0.6337 - accuracy: 0.8261

Epoch 19/30

69/69 [==============================] - 0s 15us/step - loss: 0.5898 - accuracy: 0.7971

Epoch 20/30

69/69 [==============================] - 0s 29us/step - loss: 0.5533 - accuracy: 0.7826

Epoch 21/30

69/69 [==============================] - 0s 29us/step - loss: 0.5216 - accuracy: 0.7826

Epoch 22/30

69/69 [==============================] - 0s 29us/step - loss: 0.4932 - accuracy: 0.8116

Epoch 23/30

69/69 [==============================] - 0s 15us/step - loss: 0.4671 - accuracy: 0.8551

Epoch 24/30

69/69 [==============================] - 0s 15us/step - loss: 0.4439 - accuracy: 0.8841

Epoch 25/30

69/69 [==============================] - 0s 29us/step - loss: 0.4234 - accuracy: 0.8986

Epoch 26/30

69/69 [==============================] - 0s 15us/step - loss: 0.4048 - accuracy: 0.9130

Epoch 27/30

69/69 [==============================] - 0s 14us/step - loss: 0.3864 - accuracy: 0.9275

Epoch 28/30

69/69 [==============================] - 0s 29us/step - loss: 0.3683 - accuracy: 0.9275

Epoch 29/30

69/69 [==============================] - 0s 29us/step - loss: 0.3510 - accuracy: 0.9275

Epoch 30/30

69/69 [==============================] - 0s 14us/step - loss: 0.3352 - accuracy: 0.9275

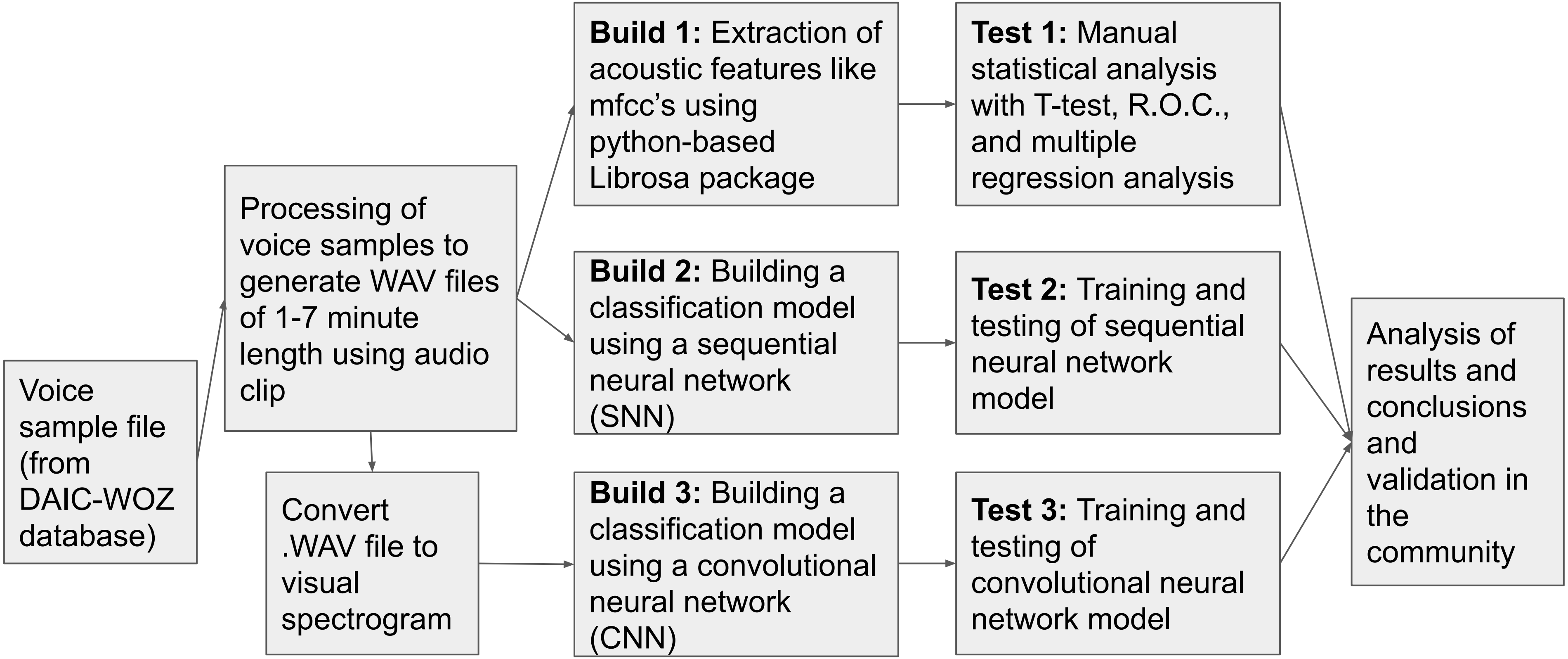
20/20 [==============================] - 0s 750us/step

test_acc: 0.75

```python
import librosa
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import os
import csv
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder, StandardScaler
import keras
from keras import models
from keras import layers

os.chdir('E:')

header = 'filename chroma_stft rms spectral_centroid spectral_bandwidth rolloff zero_crossing_rate'
for i in range(1, 21):
    header += f' mfcc{i}'
header += ' label'
header = header.split()

file = open('finaldep.csv', 'w', newline='')
with file:
    writer = csv.writer(file)
    writer.writerow(header)
genres = 'DP NDP'.split()
for g in genres:
    for filename in os.listdir(f'E:/Voix/{g}'):
        songname = f'E:/Voix/{g}/{filename}'
        y, sr = librosa.load(songname, mono=True, duration=30)
        chroma_stft = librosa.feature.chroma_stft(y=y, sr=sr)
        rms = librosa.feature.rms(y=y)
        spec_cent = librosa.feature.spectral_centroid(y=y, sr=sr)
        spec_bw = librosa.feature.spectral_bandwidth(y=y, sr=sr)
        rolloff = librosa.feature.spectral_rolloff(y=y, sr=sr)
        zcr = librosa.feature.zero_crossing_rate(y)
        mfcc = librosa.feature.mfcc(y=y, sr=sr)
        to_append = f'{filename} {np.mean(chroma_stft)} {np.mean(rms)} {np.mean(spec_cent)} {np.mean(spec_bw)} {np.mean(rolloff)} {np.mean(zcr)}'
        for e in mfcc:
            to_append += f' {np.mean(e)}'
        to_append += f' {g}'
        file = open('finaldep.csv', 'a', newline='')
        with file:
            writer = csv.writer(file)
            writer.writerow(to_append.split())

data = pd.read_csv('finaldep.csv')
data.head()

data = data.drop(['filename'],axis=1)
data.head()

genre_list = data.iloc[:, -1]
encoder = LabelEncoder()
y = encoder.fit_transform(genre_list)
print(y)

scaler = StandardScaler()
X = scaler.fit_transform(np.array(data.iloc[:, :-1], dtype = float))

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=20)

model = models.Sequential()
model.add(layers.Dense(256, activation='relu', input_shape=(X_train.shape[1],)))
```

```
model.add(layers.Dense(128, activation='relu'))

model.add(layers.Dense(64, activation='relu'))

model.add(layers.Dense(2, activation='softmax'))

model.compile(optimizer='adam',
              loss='sparse_categorical_crossentropy',
              metrics=['accuracy'])

history = model.fit(X_train,
                    y_train,
                    epochs=30,
                    batch_size=128)

test_loss, test_acc = model.evaluate(X_test,y_test)
print('test_acc: ',test_acc)

predictions = model.predict(X_test)
np.argmax(predictions[0])

model.save('finmodel.h5')
```

```
import librosa
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import os
import csv
# Preprocessing
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder, StandardScaler
#Keras
import keras
from keras import models
from keras import layers
from keras.models import load_model

os.chdir('E:')
model= load_model('finmodel.h5')

# generating a dataset
header = 'filename chroma_stft rms spectral_centroid spectral_bandwidth rolloff zero_crossing_rate'
for i in range(1, 21):
    header += f' mfcc{i}'
header += ' label'
header = header.split()

#input.csv will have input from an actual patient
file = open('input.csv', 'w', newline='')
with file:
    writer = csv.writer(file)
    writer.writerow(header)
for filename in os.listdir(f'./input/'):
    songname = f'./input/{filename}'
    y, sr = librosa.load(songname, mono=True, duration=30)
    chroma_stft = librosa.feature.chroma_stft(y=y, sr=sr)
    rms = librosa.feature.rms(y=y)
    spec_cent = librosa.feature.spectral_centroid(y=y, sr=sr)
    spec_bw = librosa.feature.spectral_bandwidth(y=y, sr=sr)
    rolloff = librosa.feature.spectral_rolloff(y=y, sr=sr)
    zcr = librosa.feature.zero_crossing_rate(y)
    mfcc = librosa.feature.mfcc(y=y, sr=sr)
    to_append = f'{filename} {np.mean(chroma_stft)} {np.mean(rms)} {np.mean(spec_cent)} {np.mean(spec_bw)} {np.mean(rolloff)} {np.mean(zcr)}'
    for e in mfcc:
        to_append += f' {np.mean(e)}'
    to_append += f' foo'
    file = open('input.csv', 'a', newline='')
    with file:
        writer = csv.writer(file)
        writer.writerow(to_append.split())

#reading dataset from csv

data = pd.read_csv('input.csv')
data.head()

# Dropping unneccesary columns
data = data.drop(['filename'],axis=1)
data.head()

genre_list = data.iloc[:, -1]
encoder = LabelEncoder()
y = encoder.fit_transform(genre_list)
print(y)
```


Voice sample file (from DAIC-WOZ database)
Processing of voice samples to generate WAV files of 1-7 minute length using audio clip
Convert .WAV file to visual spectrogram
Build 1: Extraction of acoustic features like mfcc's using python-based Librosa package
Build 2: Building a classification model using a sequential neural network (SNN)
Build 3: Building a classification model using a convolutional neural network (CNN)
Test 1: Manual statistical analysis with T-test, R.O.C., and multiple regression analysis
Test 2: Training and testing of sequential neural network model
Test 3: Training and testing of convolutional neural network model
Analysis of results and conclusions and validation in the community

Testing the model

RESULTS of experiment 1

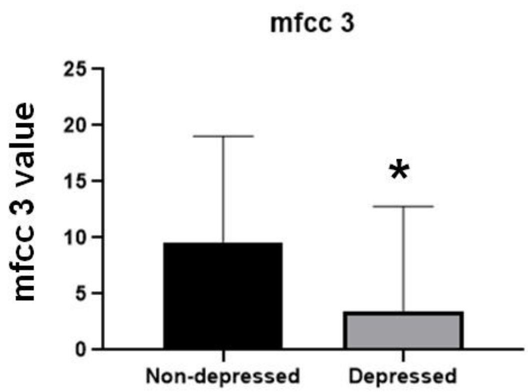
The Wav files generated from the audio files were subjected to Python-based (a computer program) extraction of acoustic features such as Zero cross rate, mel frequency cepstralcoefficients (mfcc). We analyzed the data statistically to determine which feature (s) distinguishes depression from normal state. Our data suggested that only mfcc 3, 5, and 16 were significantly lower in depressed individuals compared to non-depressed subjects (red font and stars in Table 1 and fig 1 A to D). Other mfccs were not significantly changed for example mfcc 10 (Fig 6 A).

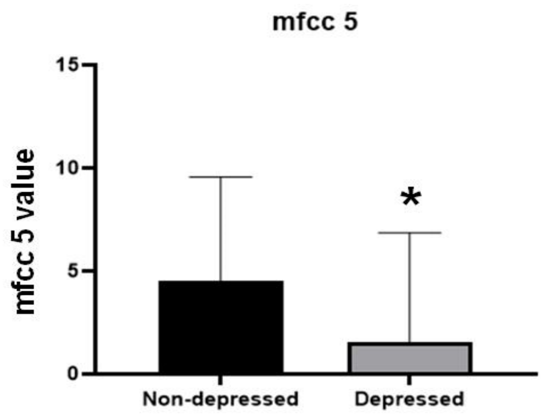
Table 1: Analysis of mfcc in depressed and non-depressed subjects. (*a nd red font denotes statistically significant mfccs)

MFCC values	Depressed (Mean ± S.D) → 31 samples	Non-depressed (Mean ± S.D) → 58 samples
1	-670.7182312 ± 57.61064359	-623.3501666 ± 47.63366157
2	101.268613 ± 17.04397142	99.74999592 ± 17.02158392
3 *	3.455296315 ± 9.202240751	9.616474198 ± 9.357447706
4	23.88700428 ± 8.718515099	24.55711574 ± 6.156947994
5 *	1.553210518 ± 5.21954796	4.513056596 ± 5.005780907

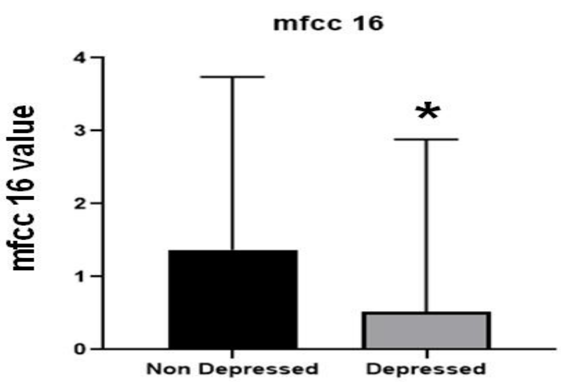
6	11.14655459 ± 6.437744967	10.64291618 ± 4.586626109
7	-2.512867958 ± 4.652040804	-0.7541774249 ± 4.440193508
8	3.416036087± 4.06051123	4.141875702 ± 2.744407829
9	2.627645195±3.351766665	2.843736975 ± 2.957794156
10	2.588276866 ± 2.671379279	2.596181298 ± 2.65466649
11	4.358039374 ± 2.300123381	4.158138447 ± 2.525268735
12	0.2464258112 ±2.25504841	0.5591302411± 2.429558409
13	3.892040574 ±2.408333609	4.471360458 ± 1.76368387
14	-0.4136840573 ±2.045423462	0.2008359595 ± 2.191492974
15	4.61324272 ±2.155055921	3.909769408± 1.661736673
16 *	0.5219774424 ± 2.324104492	1.367958285 ± 2.352055383
17	1.626248313 ± 1.895564749	1.563470854 ± 1.316335804
18	0.7194835087 ± 1.094509501	0.9991352697 ± 1.264191034
19	0.2773959628 ± 1.409168148	0.2141550713 ± 1.401578822
20	1.218331385 ± 1.636154629	1.214438957 ± 1.244544918

(A) **(B)**





(C) (D)



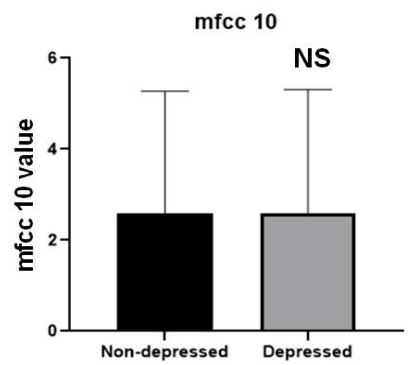
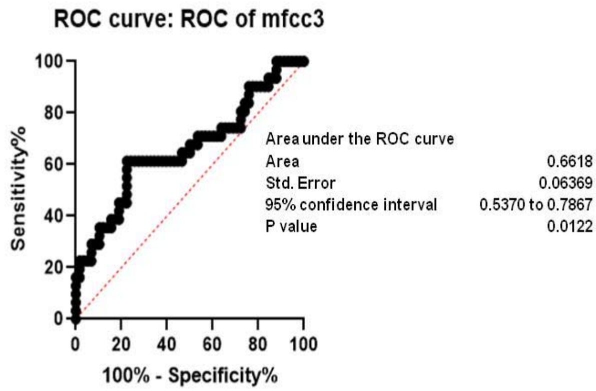
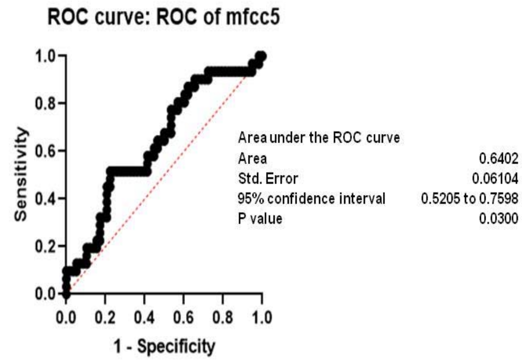


Figure 1. (A) mfcc 3, (B) mfcc 5, and (C) mfcc 16 are significantly lower in depressed individuals as analyzed with statistical T test (P,0.05) (D) mfcc 10 is not significantly changed between non-depressed and depressed individuals NS, non significant.

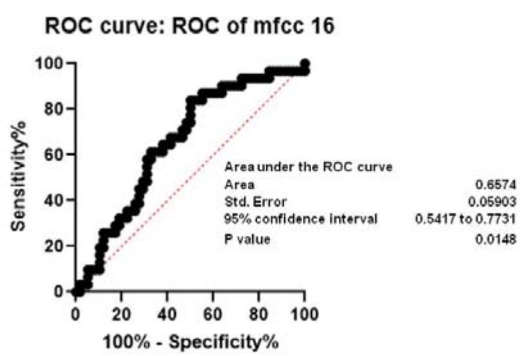
<u>*ROC analysis*</u>: ROC analysis revealed that AUC of mfcc 3, 5 , and 16 were 0.63, 0. 61, and 0.69 respectively whereas mfcc 10 which is not significantly changed has AUC was 0.51 (Figs 2 A to D). AUC of 0.5 means model has no discriminative ability between depressed and non-depressed subjects

(A) **(B)**





(C) (D)



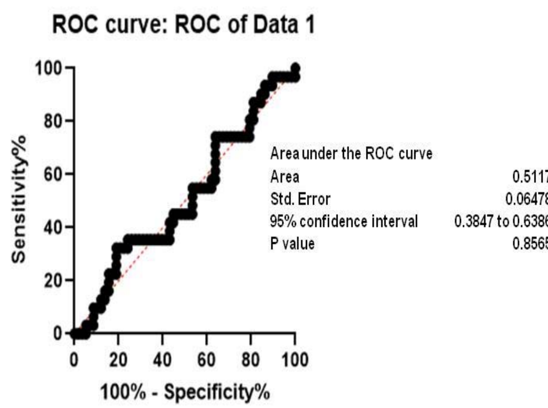
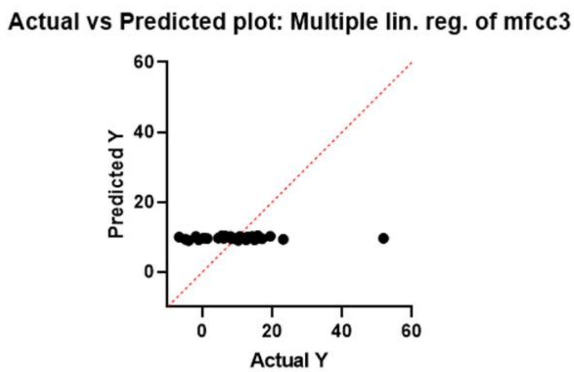


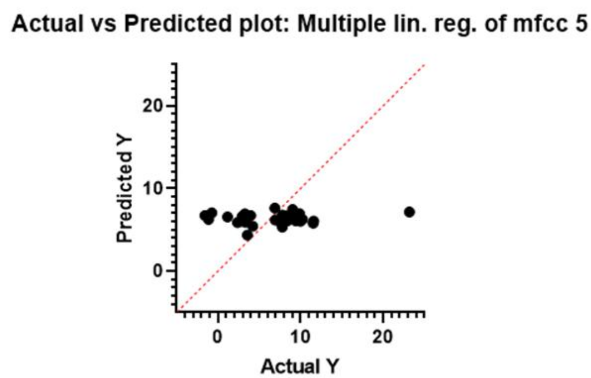
Figure 2. ROC curve analysis shows (A) 0.61 AUC of mfcc 3 B) 0.63 AUC of mfcc 5, and (C) 0.69 AUC of mfcc 16 significantly discriminates depressed vs non-depressed conditions. (D) 0.51 AUC of mfcc 10 is not significantly changed between non-depressed and depressed individuals.

Multiple regression analysis: Finally we performed multiple regression analysis to analyze the correlation of actual values of the data with predicted values. The results showed that there is a correlation between the actual and predicted values of the mfcc 3. 5, and 16 with depressed and non-depressed individuals which were significant in the other tests. As expected the mfcc 10 which is not significantly different did not show any correlation (Fig 3A to D).

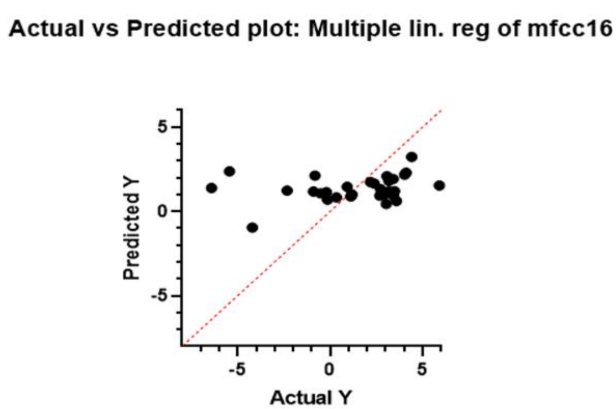
(A)



(B)



(B)



(D)

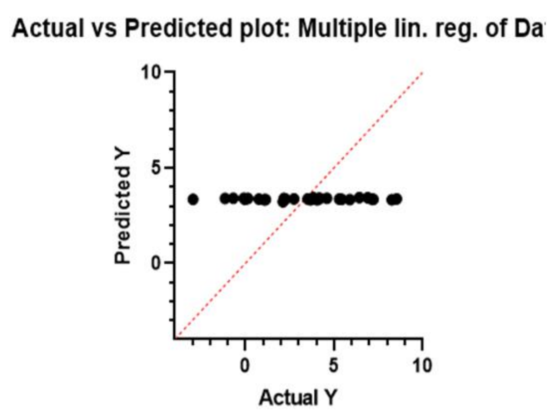


Figure 3. Multiple regression analysis confirms actual values of (A) mfcc 3, (B) mfcc 5, and (C) mfcc 16 correlate well with predicted values of 3 of non-depressed and depressed individuals. (D) mfcc 10 is not significantly changed between non-depressed and depressed individuals.

Results of experiment 2: The accuracy rate of this model as shown below is 75%

```
Epoch 1/30
69/69 [==============================] - 0s 1ms/step - loss: 2.6171 - accuracy: 0.0000e+00
Epoch 2/30
69/69 [==============================] - 0s 29us/step - loss: 2.3848 - accuracy: 0.0000e+00
Epoch 3/30
69/69 [==============================] - 0s 29us/step - loss: 2.1903 - accuracy: 0.1739
Epoch 4/30
69/69 [==============================] - 0s 29us/step - loss: 2.0304 - accuracy: 0.4348
Epoch 5/30
69/69 [==============================] - 0s 29us/step - loss: 1.8936 - accuracy: 0.5507
Epoch 6/30
69/69 [==============================] - 0s 29us/step - loss: 1.7683 - accuracy: 0.6232
Epoch 7/30
69/69 [==============================] - 0s 14us/step - loss: 1.6470 - accuracy: 0.6667
Epoch 8/30
69/69 [==============================] - 0s 14us/step - loss: 1.5293 - accuracy: 0.6667
Epoch 9/30
69/69 [==============================] - 0s 14us/step - loss: 1.4131 - accuracy: 0.7246
Epoch 10/30
69/69 [==============================] - 0s 14us/step - loss: 1.2999 - accuracy: 0.7391
Epoch 11/30
69/69 [==============================] - 0s 29us/step - loss: 1.1913 - accuracy: 0.7391
Epoch 12/30
69/69 [==============================] - 0s 29us/step - loss: 1.0885 - accuracy: 0.7971
Epoch 13/30
69/69 [==============================] - 0s 15us/step - loss: 0.9923 - accuracy: 0.8116
Epoch 14/30
69/69 [==============================] - 0s 14us/step - loss: 0.9036 - accuracy: 0.8116
Epoch 15/30
69/69 [==============================] - 0s 14us/step - loss: 0.8225 - accuracy: 0.7971
Epoch 16/30
69/69 [==============================] - 0s 15us/step - loss: 0.7498 - accuracy: 0.7826
Epoch 17/30
69/69 [==============================] - 0s 29us/step - loss: 0.6868 - accuracy: 0.8116
Epoch 18/30
69/69 [==============================] - 0s 29us/step - loss: 0.6337 - accuracy: 0.8261
Epoch 19/30
69/69 [==============================] - 0s 15us/step - loss: 0.5898 - accuracy: 0.7971
Epoch 20/30
69/69 [==============================] - 0s 29us/step - loss: 0.5533 - accuracy: 0.7826
Epoch 21/30
69/69 [==============================] - 0s 29us/step - loss: 0.5216 - accuracy: 0.7826
Epoch 22/30
69/69 [==============================] - 0s 29us/step - loss: 0.4932 - accuracy: 0.8116
Epoch 23/30
69/69 [==============================] - 0s 15us/step - loss: 0.4671 - accuracy: 0.8551
Epoch 24/30
69/69 [==============================] - 0s 15us/step - loss: 0.4439 - accuracy: 0.8841
```

Epoch 25/30
69/69 [==============================] - 0s 29us/step - loss: 0.4234 - accuracy: 0.8986
Epoch 26/30
69/69 [==============================] - 0s 15us/step - loss: 0.4048 - accuracy: 0.9130
Epoch 27/30
69/69 [==============================] - 0s 14us/step - loss: 0.3864 - accuracy: 0.9275
Epoch 28/30
69/69 [==============================] - 0s 29us/step - loss: 0.3683 - accuracy: 0.9275
Epoch 29/30
69/69 [==============================] - 0s 29us/step - loss: 0.3510 - accuracy: 0.9275
Epoch 30/30
69/69 [==============================] - 0s 14us/step - loss: 0.3352 - accuracy: 0.9275
20/20 [==============================] - 0s 750us/step
test_acc: 0.75

Results of experiment 3

The accuracy rate of this model as shown below is 63.64%

[INFO] training network...

Train on 65 samples, validate on 22 samples

Epoch 1/30

16/65 [======>.......................] - ETA: 5s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 2s - loss: 5.7500 - accuracy: 0.6250

48/65 [=====================>........] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

64/65 [============================>.] - ETA: 0s - loss: 5.0312 - accuracy: 0.6719

65/65 [==============================] - 4s 54ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 2/30

16/65 [======>.......................] - ETA: 1s - loss: 3.8333 - accuracy: 0.7500

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 6.0694 - accuracy: 0.6042

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 3/30

16/65 [======>.......................] - ETA: 0s - loss: 6.7083 - accuracy: 0.5625

32/65 [=============>................] - ETA: 0s - loss: 5.7500 - accuracy: 0.6250

48/65 [=====================>........] - ETA: 0s - loss: 5.4305 - accuracy: 0.6458

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 27ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 4/30

16/65 [======>.......................] - ETA: 0s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 0s - loss: 4.3125 - accuracy: 0.7188

48/65 [=====================>........] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 5/30

16/65 [======>.......................] - ETA: 1s - loss: 4.7916 - accuracy: 0.6875

32/65 [=============>................] - ETA: 0s - loss: 5.7500 - accuracy: 0.6250

48/65 [=====================>........] - ETA: 0s - loss: 5.1111 - accuracy: 0.6667

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 28ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 6/30

16/65 [======>.......................] - ETA: 1s - loss: 3.8333 - accuracy: 0.7500

32/65 [=============>................] - ETA: 0s - loss: 3.8333 - accuracy: 0.7500

48/65 [=====================>........] - ETA: 0s - loss: 3.8333 - accuracy: 0.7500

64/65 [============================>.] - ETA: 0s - loss: 5.0312 - accuracy: 0.6719

65/65 [==============================] - 2s 27ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 7/30

16/65 [======>.......................] - ETA: 0s - loss: 7.6666 - accuracy: 0.5000

32/65 [=============>................] - ETA: 0s - loss: 4.3125 - accuracy: 0.7188

48/65 [=====================>........] - ETA: 0s - loss: 4.4722 - accuracy: 0.7083

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 8/30

16/65 [======>.......................] - ETA: 1s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 0s - loss: 5.7500 - accuracy: 0.6250

48/65 [=====================>........] - ETA: 0s - loss: 5.4305 - accuracy: 0.6458

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 9/30

16/65 [======>.......................] - ETA: 1s - loss: 2.8750 - accuracy: 0.8125

32/65 [=============>................] - ETA: 0s - loss: 3.8333 - accuracy: 0.7500

48/65 [=====================>........] - ETA: 0s - loss: 5.1111 - accuracy: 0.6667

64/65 [============================>.] - ETA: 0s - loss: 5.0312 - accuracy: 0.6719

65/65 [==============================] - 2s 27ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 10/30

16/65 [======>.......................] - ETA: 1s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

48/65 [=====================>........] - ETA: 0s - loss: 4.1528 - accuracy: 0.7292

64/65 [============================>.] - ETA: 0s - loss: 5.0312 - accuracy: 0.6719

65/65 [==============================] - 2s 32ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 11/30

16/65 [======>.......................] - ETA: 1s - loss: 7.6666 - accuracy: 0.5000

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 5.4305 - accuracy: 0.6458

64/65 [============================>.] - ETA: 0s - loss: 5.0312 - accuracy: 0.6719

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 12/30

16/65 [======>.......................] - ETA: 0s - loss: 3.8333 - accuracy: 0.7500

32/65 [=============>................] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

48/65 [=====================>........] - ETA: 0s - loss: 4.4722 - accuracy: 0.7083

64/65 [============================>.] - ETA: 0s - loss: 5.0312 - accuracy: 0.6719

65/65 [==============================] - 2s 25ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 13/30

16/65 [======>.......................] - ETA: 1s - loss: 4.7916 - accuracy: 0.6875

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 5.1111 - accuracy: 0.6667

64/65 [============================>.] - ETA: 0s - loss: 5.0312 - accuracy: 0.6719

65/65 [==============================] - 2s 25ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 14/30

16/65 [======>.......................] - ETA: 1s - loss: 6.7083 - accuracy: 0.5625

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 5.1111 - accuracy: 0.6667

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 15/30

16/65 [======>.......................] - ETA: 1s - loss: 4.7916 - accuracy: 0.6875

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 5.1111 - accuracy: 0.6667

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 27ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 16/30

16/65 [======>.......................] - ETA: 1s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

48/65 [=====================>........] - ETA: 0s - loss: 5.4305 - accuracy: 0.6458

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 17/30

16/65 [======>.......................] - ETA: 1s - loss: 7.6666 - accuracy: 0.5000

32/65 [=============>................] - ETA: 0s - loss: 6.2291 - accuracy: 0.5938

48/65 [=====================>........] - ETA: 0s - loss: 6.0694 - accuracy: 0.6042

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 18/30

16/65 [======>.......................] - ETA: 1s - loss: 4.7916 - accuracy: 0.6875

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 5.4305 - accuracy: 0.6458

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 27ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 19/30

16/65 [======>.......................] - ETA: 0s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 5.7500 - accuracy: 0.6250

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 25ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 20/30

16/65 [======>.......................] - ETA: 1s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

48/65 [=====================>........] - ETA: 0s - loss: 5.1111 - accuracy: 0.6667

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 25ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 21/30

16/65 [======>.......................] - ETA: 0s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 0s - loss: 6.2291 - accuracy: 0.5938

48/65 [=====================>........] - ETA: 0s - loss: 5.7500 - accuracy: 0.6250

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 22/30

16/65 [======>.......................] - ETA: 1s - loss: 6.7083 - accuracy: 0.5625

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 5.1111 - accuracy: 0.6667

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 23/30

16/65 [======>.......................] - ETA: 1s - loss: 4.7916 - accuracy: 0.6875

32/65 [=============>................] - ETA: 0s - loss: 4.3125 - accuracy: 0.7188

48/65 [=====================>........] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 25ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 24/30

16/65 [======>.......................] - ETA: 0s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 5.4305 - accuracy: 0.6458

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 25ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 25/30

16/65 [======>.......................] - ETA: 1s - loss: 3.8333 - accuracy: 0.7500

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 5.4305 - accuracy: 0.6458

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 26/30

16/65 [======>.......................] - ETA: 1s - loss: 1.9167 - accuracy: 0.8750

32/65 [=============>................] - ETA: 0s - loss: 3.8333 - accuracy: 0.7500

48/65 [=====================>........] - ETA: 0s - loss: 5.4305 - accuracy: 0.6458

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 27/30

16/65 [======>.......................] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 5.4305 - accuracy: 0.6458

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 28/30

16/65 [======>.......................] - ETA: 1s - loss: 6.7083 - accuracy: 0.5625

32/65 [=============>................] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

48/65 [=====================>........] - ETA: 0s - loss: 4.4722 - accuracy: 0.7083

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 25ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 29/30

16/65 [======>......................] - ETA: 1s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

48/65 [=====================>........] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 25ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

Epoch 30/30

16/65 [======>......................] - ETA: 0s - loss: 5.7500 - accuracy: 0.6250

32/65 [=============>................] - ETA: 0s - loss: 4.7916 - accuracy: 0.6875

48/65 [=====================>........] - ETA: 0s - loss: 5.4305 - accuracy: 0.6458

64/65 [============================>.] - ETA: 0s - loss: 5.2708 - accuracy: 0.6562

65/65 [==============================] - 2s 26ms/step - loss: 5.1897 - accuracy: 0.6615 - val_loss: 5.5757 - val_accuracy: 0.6364

[INFO] evaluating network...

Print accuracy ...

Accuracy: 63.64%

Process finished with exit code 1

Self Improvement Through Awareness pc (SITA)

Dr. Viji Susarla PhD
Clinical Psychology
drviji@sitapc.com

1809 N Mill Street, Suite C
Naperville, IL 60563
Phone: 630 246 6810
Fax: 630 246 6809

180 N Michigan Ave, Suite 2210
Chicago, IL 60601
Phone: 312 781 1139

02/10/2020

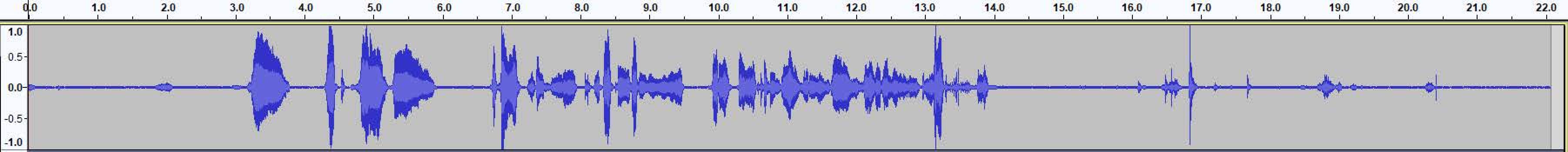
Consent to participate in the research study:
'Detection of Depression by Voice Acoustic Features and Convolution Neural Network-based Machine Learning'

Dear client,
Three high school students from Naperville North High School – Sruthi Kotlo, Anjana Ramachandran and Divya Lidder, are working on a research project to develop a computer-aided program that can detect depression from voice samples. They approached me to help them in data collection. With this study, they hope to create a better and more accurate method of early detection of depression. For this project they need voice samples of those with and without depression. They will extract and analyze acoustic features of voice samples to distinguish between normal and depression states. Further, they will generate a computer-based model to diagnose early stages of depression based on voice analysis. To collect voice samples SITA staff will either ask you answers to some questions or ask you to read a passage. We are looking to generate a 2- or 3-minute voice sample. We would greatly appreciate your participation in doing this study and your willingness to further research in the field of medicine.

Thank you.

_______________________________ _______________________________
Dr. Viji Susarla PhD Date

_______________________________ _______________________________
Signature Date


0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
12.0
13.0
14.0
15.0
16.0
17.0
18.0
19.0
20.0
21.0
22.0
1.0
0.5
0.0
-0.5
-1.0