**2022 S.T. Yau High School Science Award (Asia)**

**Research Report**

**The Team**

Registration Number: Math-169

Name of team member: Arhaan Garg
School: Step By Step school Noida
Country: India

Name of supervising teacher: Ms. Reetu Jain
Job Title: Mentor
School: Step by step
Country: India

**Title of Research Report**

Early detection of vocal disorders such as laryngeal cancer and dysphonia with the use of voice analysis and machine learning

**Date**

31 August 2021

2022 S.-T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

# Early detection of vocal disorders such as laryngeal cancer and dysphonia with the use of voice analysis and machine learning

**Arhaan Garg**

**Abstract**

Many serious disorders with our throat, such as laryngeal cancer, laryngitis, muscle tension dysphonia, vocal cord paralysis, and so on, are detected after the patient has become critically ill. These disorders can also be life-threatening, as I witnessed with my uncle. Looking at one of the most painful cancers, laryngeal cancer, I wanted to work on a remedy. This occurrence was crucial in directing my attention to this field of study. The majority of these disorders can be discovered early since the voice begins to change due to vocal cord disformations at an early stage. Smoking, drinking, bad eating habits, career, and other factors are all key contributors to these problems. The change in voice is typically the first sign of all of these disorders. People, on the other hand, have a tendency to disregard the very first symptom, which leads them deep into the problem. Voice irregularities, such as variations in frequency, may potentially be too deceiving to the human ear to be taken seriously. Voice disorders such as dysphonia and laryngeal cancer can be detected early using artificial intelligence and machine learning. I worked with Santosh Hospital to collect data and do background study on vocal problems and irregularities. Throughout the procedure, I collected 100+ minutes of audio data from individuals with laryngeal cancer while also researching approaches for detecting voice problems such as laryngoscopy. The project's goal is to distinguish between the voices of a healthy patient and a patient with a vocal cord disorder. A voice analysis comparison between a healthy patient and a patient with a vocal issue was used for this objective. 40 human voice parameters such as frequency, pitch, and zero crossing rate were retrieved using MFCCs and methods such as the discrete cosine transformation and the mel filter bank. A wrapper was used to pick the most important features in determining if the patient has a vocal problem or not. After that, the logistic regression model was used to train a machine learning model to determine if the audio sample was disordered or healthy. The instrument has an incredibly high accuracy of 88%, making it extremely efficient. This is a technology that assists patients at an early stage in order to keep therapy simple and cure cancer and other critical conditions faster. It also lessens stress on doctors and lowers medical costs while decreasing the effect of sedatives on patients. The technique is very simple to use and available in all places where competent doctors and proper equipment to detect such major voice problems are lacking.

2022 S.T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

**Keywords**: Voice Disorder, MFCC, Voice analysis, Machine learning, Dysphonia Detection, Laryngeal Cancer Detection

**Acknowledgement**

1. We would like to acknowledge Dr. Abhay Singh ENT specialist at Santosh Hospital for providing us background knowledge about voice disorders and the treatment procedures.
2. We would acknowledge Dr. Tripta Bhagat, chancellor at Santosh medical college for helping us in fetching data from patients suffering from Dysphonia, laryngitis and other diseases.
3. Walden, Patrick R (2020), "Perceptual Voice Qualities Database (PVQD)", Mendeley Data, v2, https://data.mendeley.com/datasets/9dz247gnyb/1

2022 S.-T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

**Commitments on Academic Honesty and Integrity**

We hereby declare that we

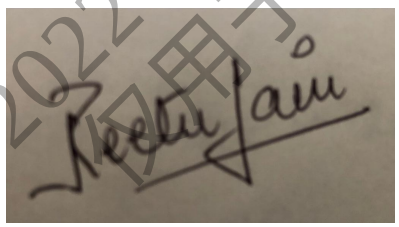
1. are fully committed to the principle of honesty, integrity and fair play throughout the competition.
2. actually perform the research work ourselves and thus truly understand the content of the work.
3. observe the common standard of academic integrity adopted by most journals and degree theses.
4. have declared all the assistance and contribution we have received from any personnel, agency, institution, etc. for the research work.
5. undertake to avoid getting in touch with assessment panel members in a way that may lead to direct or indirect conflict of interest.
6. undertake to avoid any interaction with assessment panel members that would undermine the neutrality of the panel member and fairness of the assessment process.
7. observe the safety regulations of the laboratory(ies) where we conduct the experiment(s), if applicable.
8. observe all rules and regulations of the competition.
9. agree that the decision of YHSA(Asia) is final in all matters related to the competition.

**We understand and agree that failure to honor the above commitments may lead to disqualification from the competition and/or removal of reward, if applicable; that any unethical deeds, if found, will be disclosed to the school principal of team member(s) and relevant parties if deemed necessary; and that the decision of YHSA(Asia) is final and no appeal will be accepted.**

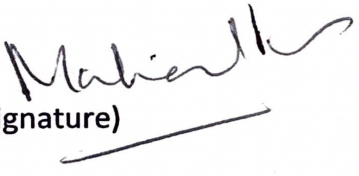
*(Signatures of full team below)*

Arhaan Garg
Name of team member:

Name of supervising teacher:
Mrs. Reetu Jain



Noted and endorsed by



(signature)

Name of school principal: Dr. Mahesh Prasad


2022 S.T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

**Table of Contents**



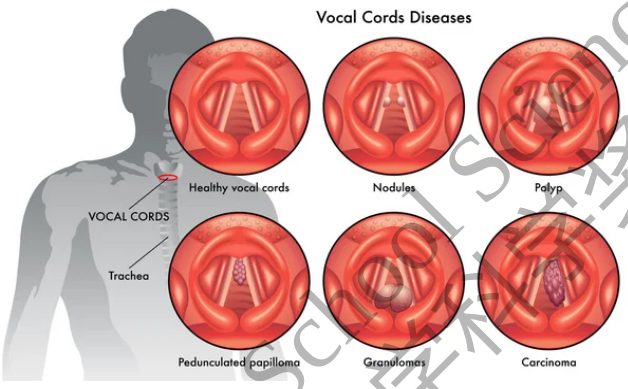
2022 S.-T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

## Introduction:

The cartilage, muscle, and mucous membranes that make up your voice box (larynx) are located near the top of your windpipe (trachea) and the base of your tongue. Your vocal cords are two flexible bands of muscle tissue located at the windpipe's entrance. When your vocal cords vibrate, they produce sound. This vibration is caused by air passing through your larynx, which brings your voice cords closer together. When you swallow, your vocal cords also help shut your voice box, keeping you from inhaling food or liquid.

If you have an issue with your voice's pitch, volume, tone, or other aspects, you may have a vocal disorder. If your vocal cords become inflamed, acquire growths, or become paralyzed, they will be unable to function normally.

People develop vocal issues for a variety of causes.




Doctors who specialize in ear, nose, and throat issues, as well as speech pathologists, are involved in the diagnosis and treatment of voice disorders.

The following are some of the most common examples of voice disorders:

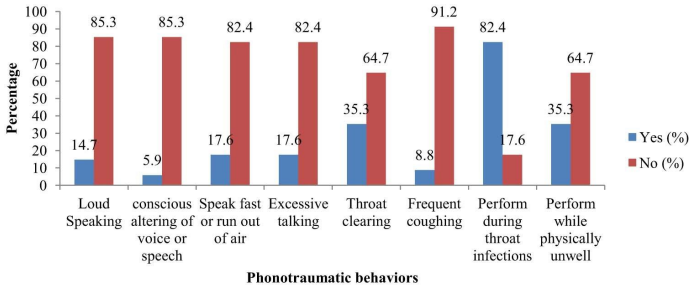
1. Laryngitis
2. Muscle tension dysphonia
3. Neurological voice disorders, such as spasmodic dysphonia
4. Polyps, nodules or cysts on the vocal cords (non cancerous lesions)
5. Precancerous and cancerous lesions
6. Vocal cord paralysis

Many risk factors include the following:

- Aging
- Alcohol use
- Allergies
- Illnesses, such as colds or upper respiratory infections
- Neurological disorders, Psychological stress
- Screaming
- Smoking
- Voice misuse or overuse

2022 S.T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

These factors are very common in humans, which put all humans at a high risk of getting the disease.




Possible causes can include:

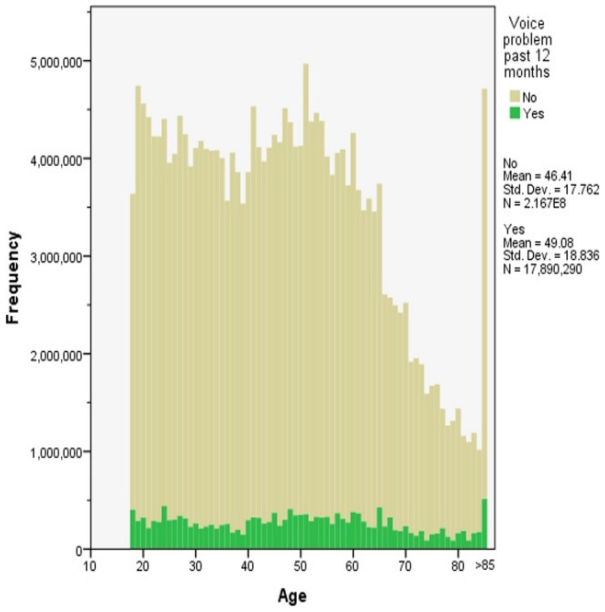
1. **Growth.** Extra tissue may grow on the vocal cords in some circumstances. This prevents the cords from functioning appropriately. Fluid-filled sacs called cysts, wart-like lumps called papilloma, and callus-like bumps called nodules are examples of growths. There may be areas of scar tissue or patches of damaged tissue termed lesions. Other types of growths include granulomas, which are small areas of chronic inflammation, and polyps, which are little blisters. Illness, injury, disease, and voice abuse are all possible causes of growth.
2. **Inflammation and swelling.** Swelling and inflammation Inflammation and swelling of the voice cords can be caused by a variety of factors. These include surgery, respiratory sickness or allergies, GERD (acid reflux), some medications, and chemical exposure, Smoking, alcoholism, and verbal abuse.
3. **Nerve problems.** Certain medical problems might cause nerves that regulate the voice chords to malfunction. Multiple sclerosis, myasthenia gravis, Parkinson disease, Amyotrophic lateral sclerosis (ALS), and Huntington illness are examples. Surgery or prolonged laryngeal inflammation can potentially cause nerve damage (laryngitis).
4. **Hormones.** Voice abnormalities can be caused by thyroid hormone, female and male hormones, and growth hormone deficiencies.
5. **Misuse of the voice.** Misuse of one's voice When speaking, using too much tension might strain the vocal cords. This might cause issues with the muscles of the throat, affecting the voice. A voice issue can also be caused by vocal abuse. Anything that stretches or hurts the vocal cords is considered vocal abuse. Excessive talking, shouting, and yelling are examples of vocal abuse. Smoking and a frequent clearing of the throat are also forms of vocal abuse. Voice cord abuse can result in the formation of nodes and polyps, which are calluses or blisters on the vocal cords. These alter the sound of the voice. Vocal abuse can cause a vocal chord to rupture in some situations. This causes the cord to bleed (hemorrhage) and can result in voice loss. Vocal cord hemorrhage must be treated as soon as possible.

Among the signs of voice abnormalities are:

1. Have a trembling sound,
2. Feel a rough or harsh (hoarseness) voice that is strained or choppy. Is it shaky, whispery, or breathy?
3. The pitch of the voice is too high or low, or it changes.
4. You may feel strain or soreness in your throat while speaking, or your voice box may be weary.

2022 S.T. Yau High School Science Award 仅用于2022丘成桐中学生科学奖公示

5. When swallowing, you may feel a "lump" in your throat or pain when you touch the outside of your throat.




## Literature Review:

- **Identification of Voice Disorders in Laryngeal Cancer Patients:**

  In this regard, the research presents a non-invasive voice illness diagnosis method for laryngeal cancer patients. Fifty-five laryngeal and fifty-five healthy cases of the sustained vowel /a/ were recorded. Because of the nonlinearity of the vocal cords, seven nonlinear parameters are retrieved together with biologically triggered 39 Mel-Frequency Cepstral Coefficients (MFCC). This documentation is a 110X46 laryngeal dataset. The wrapper technique is used to improve function selection and to enhance the discriminating capability of existing work. A tailored Support vector machine (SVM) with grid seek and random forest is used to achieve the type (RF). The current experiment demonstrated an enhanced accuracy of 76.56% with SVM and 80% in the case of random forest. The forward selection of capabilities, as well as the inclusion of non-linear functions, has played a significant role in the improved overall performance of the current gadget.

- **Noninvasive Detection of Potentially Precancerous Vocal Fold Lesions Using Glottal Wave Signal and SVM Methods**

  This research proposes an automatic identification of premalignant lesions based on human voice production theory. Premalignant lesions such as leukoplakia, Erythroplakia, Keratosis, and others are intimately associated with the vocal fold; hence, features retrieved from the glottal fold waveform can be relevant and significant. The suggested method is a non-intrusive methodology that extracts the vocal fold waveform from recorded utterances. The fundamental idea is to extract relevant information from the raw signal (glottal waveform). However, such a signal is not readily available. As a result, we begin by extracting the glottal waveform from recorded speech using Iterative Adaptive

2022 S.T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

Inverse Filtering (IAIF). IAIF is used on two databases with sustained vowel samples: the Massachusetts Eye and Ear Infirmary (MEEI) database and the Saarbrucken Voice Database (SVD). Relevant examples derived from glottal flow impulses are utilized to construct relevant characteristics. A thorough examination of collected features using statistical methods such as boxplot and Principal Component Analysis (PCA) results in the selection of the most crucial and relevant features. A distinction between normal and premalignant tumors is achieved utilizing the Support Vector Machine Tool as a classification technique. This study demonstrates that premalignant lesions can be detected with acceptable and reasonable sensitivity, specificity, precision, and accuracy. Such algorithms can aid in the detection of laryngeal cancer at an earlier stage.

- **A Survey on Throat Cancer Detection Using Machine Learning:**

  Medical applications in Machine Learning (ML) algorithms are well-being states that are based on assessing the various attributes that have a high impact on being sick. Cancer is one of the few human diseases for which experts are still searching for a perfect cure, and it is also unexpected. Cancer is a diverse disease, and treatment differs from type to type and might instill distinct stages. A tumor that spreads throughout the voice box (larynx), tonsils, or throat is referred to as throat cancer (pharynx). It is strongly advised to diagnose throat cancer and begin treatment as soon as possible. Deep Learning (DL) image processing techniques and machine learning (ML) approaches are utilized to predict throat cancer, specifically for supervised learning classification algorithms. This report examined the ML and DL-based research activities that have been conducted to categorize throat cancer.

- **Dysphonia Detection Index:**

  They introduce a novel marker, the dysphonia detection index, in this work, which might be integrated into a mobile health solution and help with the assessment of voice abnormalities. Four acoustic parameters are combined into a single marker to assess the overall condition of the voice and identify whether or not a vocal problem exists. A model tree regression technique was used to evaluate the link between these qualities, and the Youden analysis was used to estimate the threshold value to distinguish between a pathological and a healthy voice. The proposed index's accuracy, sensitivity, and specificity have been accurately classified in order to examine its dependability. To evaluate the performance of our proposed index, we used a dataset of 2003 voices drawn from the Massachusetts Eye and Ear Infirmary, the Saarbruecken Voice, and the VOice ICar fEDerico II databases. Our technique surpassed other algorithms in terms of performance, with an accuracy of 82.2 percent, sensitivity and specificity of 82 percent and 82.6 percent, respectively.

2022 S.T. Yau High School Science Award 仅用于2022丘成桐中学科学奖公开

- **Pattern recognition of speech spectra for dysphonia detection:**

  The neural network approach of Kohonen's self-organizing map was utilized to recognize the spectrum patterns of dysphonia. A team of speech pathologists assessed the speech samples, which included 17 men and 18 women speaking Finnish words with long [a:]. The speech sample judgements were then compared to the placements of the [a:] samples on a self-organized spectral feature map. The map distinguished between normal and dysphonic speech spectra in a statistically significant way. The most noticeable spectral element supporting the discrepancy was the energy ratio at 1-2 kHz and 7-9 kHz.

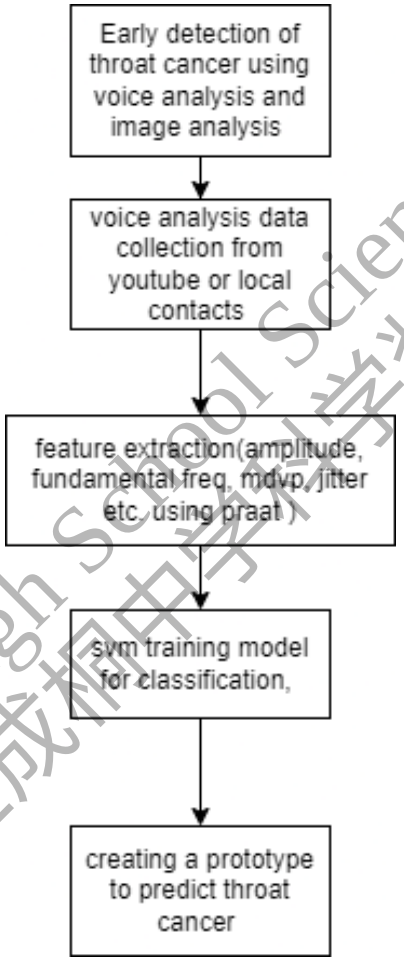
- **Early detection of speech and voice abnormalities in Parkinson's disease patients treated with deep brain stimulation to the subthalamic nucleus:**

  They previously documented the following phenotypes of speech and vocal impairments in Parkinson's disease (PD) patients undergoing STN-DBS: hypokinetic dysarthria, stuttering, breathy voice, strained voice, and spastic dysarthria. However, changes over time are still unknown. This study included 32 Parkinson's disease patients who were examined before and up to a year after surgery (PD-DBS). Eleven Parkinson's disease patients who were getting medication were also evaluated (PD-Med). Speech, voice, movement, and cognition functions were evaluated. All groups reported similar rates of stuttering (50 percent vs. 45 percent), breathy speech (66 percent vs. 73 percent), and strained voice (3 percent vs. 9 percent) at the start, as well as hypokinetic dysarthria (63 percent of PD-DBS vs. 82 percent of PD-Med). Only the PD-DBS group experienced a slight but significant decrease in speech intelligibility (p 0.001) and dysphonia grade (p = 0.001) at 1 year compared to baseline. Stuttering (9 percent vs. 18 percent) and breathiness (13 percent vs. 9 percent) emerged in both PD-DBS and PD-Med during the follow-up, while only strained voice (28 percent) and spastic dysarthria (44 percent) did not. When the stimulation was stopped, the majority of the patients' breathy and strained voices, as well as their spastic dysarthria, improved. These findings imply that the most common speech and vocal problems caused by DBS are strained voice and spastic dysarthria, and that STN-DBS may aggravate stuttering and breathy voice. A better understanding of these disorders may make it simpler to detect speech and voice impairment early on, leading to more effective treatments.

2022 S.T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

**Our Solution:**

Our endeavor comprises developing a reliable approach for detecting vocal cord problems such as dysphonia and even laryngeal cancer early on. We accomplished this through the use of machine learning. We constructed a dataset and trained a machine learning model on it to properly predict whether or not a patient has a vocal cord disorder. It is particularly useful in detecting such disorders at an early stage. They are easily treatable if caught early. One big difficulty is late discovery, which causes many people to lose their lives. They can't perceive it because it causes changes in frequency and pitch of voice that the human ear can't detect.As a result, the machine learning model is trained on a model that recognizes aspects of the voice such as frequency and classifies the speech as normal or abnormal. This will aid in the saving of lives.



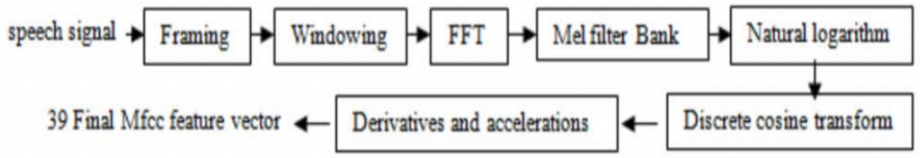

The procedure we used is as follows.

I obtained one-minute audio samples from a hospital with the assistance of an ENT doctor from 15 individuals suffering from vocal diseases such as dysphonia, laryngeal cancer, and vocal fold issues. The remaining 265 examples were obtained from an online dataset from a previous study on detecting distinct causes of dysphonia. In total, we collected data for 100+ minutes and trained a machine learning model on the data to predict whether or not a vocal cord issue exists. If so, what kind of disorder is it? We employed MFCCs, which are Mel frequency cepstral coefficients, for this. They are derived from an audio clip's cepstral representation (a nonlinear "spectrum-of-a-spectrum"). The mel-frequency cepstrum differs from the cepstrum in that the frequency bands in the MFC are equally spaced on the mel scale, which more closely approximates the human auditory system's response than the linearly-spaced frequency bands used in the normal spectrum. This frequency warping can improve sound representation. We utilize the following procedure to obtain the MFCCs. MFCCs are typically derived as follows:

1. Take the Fourier transform of (a windowed excerpt of) a signal.

2022 S.T. Yau High School Science Award 仅用于2022丘成桐中学生学术公示

2. Map the powers of the spectrum obtained above onto the mel scale, using triangular overlapping windows or alternatively, cosine overlapping windows.
3. Take the logs of the powers at each of the mel frequencies.
4. Take the discrete cosine transform of the list of mel log powers, as if it were a signal.
5. The MFCCs are the amplitudes of the resulting spectrum.



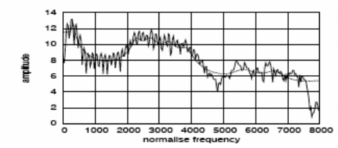

**Procedure of extracting MFCC:**

1. **Preemphasis:**
   The quantity of energy at the higher frequency is increased by pre emphasis. The energy at a higher frequency is significantly smaller than the energy at a lower frequency when we examine the frequency domain of the audio signal for voiced segments like vowels. Increasing the energy at higher frequencies will increase the model's performance by enhancing phone detection accuracy. The first order high-pass filter performs pre-emphasis as seen below. The audio signal's frequency domain for the vowel "aa," both before and after the pre emphasis, is shown below.
2. **Windowing:**
   The goal of the MFCC method is to extract characteristics from the audio stream that can be applied to the detection of phones in speech. However, because there will be multiple phones in the given audio signal, we will divide it into different segments, each with a width of 25ms and the signal spaced at 10ms, as indicated in the picture below. A person says three words per second on average with four phones, and each phone has three states, for a total of 36 states per second or 28ms per state, which is close to our 25ms window. We will extract 39 features from each section. Furthermore, if we directly chop the signal off at the edges of the signal while breaking it, the sharp drop in amplitude at the edges will cause noise in the high-frequency domain. So, rather than a rectangular window, we will utilize Hamming/Hanning windows to chop the signal, which will not cause noise in the high-frequency area.
3. **DFT (Discrete Fourier Transform):**
   We will use the dft transform to translate the signal from the time domain to the frequency domain. Analyzing audio signals in the frequency domain is simpler than in the time domain.
4. **Bank Mel-Filter:**
   The way our ears sense sound differs from how machines perceive sound. At lower frequencies, our hearing has greater resolution than at higher frequencies. So, if we hear sounds at 200 Hz and 300 Hz, we can easily distinguish them from those at 1500 Hz and 1600 Hz, even though both have a 100Hz difference present. The resolution of the machine, on the other hand, is constant across all frequencies. It has been discovered that simulating the human hearing property during the feature extraction step improves the model's performance. So we'll utilize the mel scale to convert the actual frequency to a frequency that humans can perceive. The mapping formula is shown below.

2022 S.T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

$$mel(f) = 1127\ln(1 + \frac{f}{700})$$

5. **Application of log:**
   Humans are less sensitive to changes in audio signal energy at higher vs lower energy levels. The log function has a similar trait in that it has a bigger gradient at low input values but a lower gradient at high input values. So, to simulate the human hearing system, we apply a log on the Mel-filter output.
6. **IDFT:**
   In this step, we perform the inverse transform of the preceding step's output. Before we can comprehend why we need to execute an inverse transform, we must first understand how humans make sound. The glottis, a valve that controls airflow in and out of the respiratory passageways, produces the sound. The sound is produced by the vibrating of the air in the glottis. The vibrations will occur in harmonics, and the smallest frequency produced is referred to as the fundamental frequency, with all subsequent frequencies being multiples of the fundamental frequency. The resulting vibrations will be transmitted into the vocal cavity.Based on the location of the tongue and other articulators, the vocal cavity selectively amplifies and dampens frequencies. Each sound will have a distinct position of the tongue and other articulators.The inverse of the log of the magnitude of the signal is called a cepstrum.




The fundamental frequency is at the rightmost peak in figure, and it will supply information about the pitch, while the frequencies at the rightmost will provide information about the phones. We shall disregard the fundamental frequency because it contains no information about phones.
After executing the idft procedures, the MFCC model takes the first 12 coefficients of the signal. It will use the energy of the signal sample as a feature in addition to the 12 coefficients. It will aid in the identification of the phones. The formula for the sample's energy is given below.

7. **Dynamic Features:**

Along with these 13 features, the MFCC approach will take into account the first and second order derivatives of the remaining 26 features.

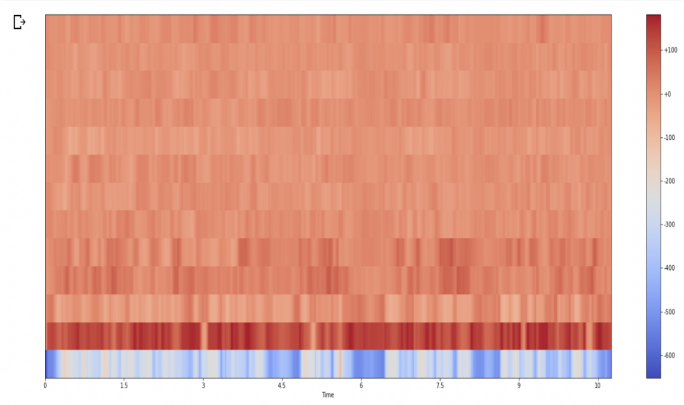
Derivatives are derived by subtracting these coefficients between audio signal samples, and they aid in understanding how the transition occurs.
As a result, the MFCC technique will generate 39 features from each audio signal sample, which will be fed into the model that recognizes whether the patient has a vocal disorder or not.

$$Energy = \sum_{t=t_1}^{t_2} x^2[t]$$

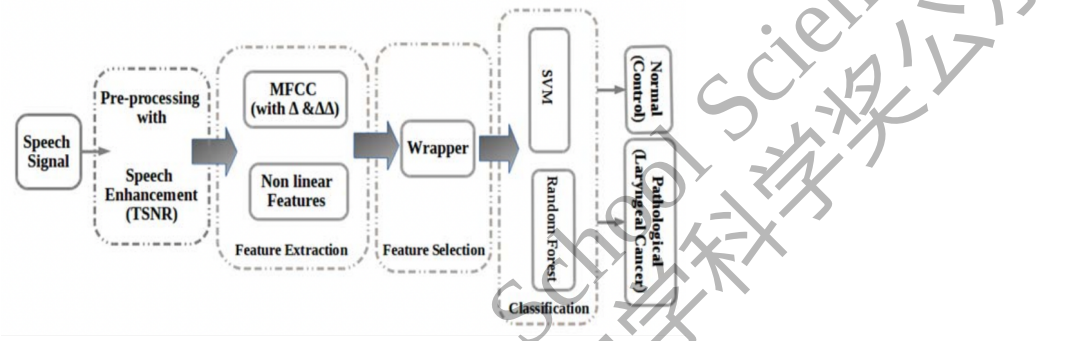
**MFCC extraction from dataset:**

2022 S.T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

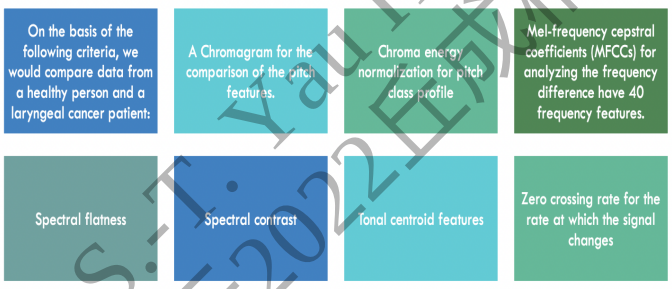
Voice features are extracted using librosa from an audio file for a healthy patient and one with vocal disorders. Further, the best feature is chosen to classify the disorders. This data is then used to train a machine learning model. These are the steps taken to derive the final MFCC features of a voice sample.



This is the MFCC Graphical output for the voice data.




This is the process algorithm used for finding whether the voice is pathological or normal.




These are the multiple features that are considered while training the model and getting the MFCC output.

```
array([[-6.51837097e+02, -6.09061890e+02, -5.46763672e+02, ...,
        -5.08928802e+02, -5.17013000e+02, -5.34831604e+02],
       [ 0.00000000e+00,  4.65155716e+01,  1.09424675e+02, ...,
         1.34649368e+02,  1.30927948e+02,  1.21156609e+02],
       [ 0.00000000e+00,  1.73918476e+01,  3.26933708e+01, ...,
         2.96752167e+01,  3.35306168e+01,  4.31818390e+01],
       ...,
       [ 0.00000000e+00,  2.71534967e+00,  2.90223265e+00, ...,
        -2.46921480e-01, -3.59625697e+00, -2.71077108e+00],
       [ 0.00000000e+00,  1.77788779e-01, -1.28865993e+00, ...,
        -1.19872248e+00, -4.37880182e+00, -2.22230148e+00],
       [ 0.00000000e+00, -4.39750314e-01, -3.03164244e+00, ...,
        -3.11957121e-01, -2.39689052e-01,  1.59559250e-01]], dtype=float32)
```

The MFCC array derived looks like the data presented above.

2022 S.T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

```
import os
import librosa
import librosa.display
import IPython.display as ipd
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
fn = []

def load_file_from_folder(folder):
    for filename in os.listdir(folder):
        fn.append(filename)

load_file_from_folder(r'/content/drive/MyDrive/Dataset/Audio Files')

extracteddata = []
fn.sort()
for i in fn:
    audio_file = r"/content/drive/MyDrive/Dataset/Audio Files/{}"
    audio_file = audio_file.format(i)
    #print(audio_file)
    signal, sr = librosa.load(audio_file)

    mfccs = librosa.feature.mfcc(y=signal, n_mfcc=40, sr=sr)
    mfccs_scaled_features = np.mean(mfccs.T,axis=0)

    extracteddata.append([mfccs_scaled_features])

dataframe = pd.DataFrame(extracteddata,columns=['features'])
print(dataframe)
```

The code above was used by us to extract the MFCCs from the audio dataset. In this we first load all essential libraries required. Then, we extract each audio file stored in the folder, and store the files in a list. After sorting the files according to their names, we extract 40 MFCC features using librosa and store the features in a dataframe, and align the arrays with their corresponding feature parameters such as age, diagnosis, and name. This dataframe is then used for training the model.

**Machine learning:**

We then split the feature array and created a dataset and classified each according to the diagnosis and age. Further, we encoded the data using feature mapping as shown.

Feature mapping

```
from sklearn.preprocessing import LabelEncoder

le=LabelEncoder()

for x in colname:
    data[x]=le.fit_transform(data[x])


    le_name_mapping = dict(zip(le.classes_, le.transform(le.classes_)))
    print('Feature', x)
    print('mapping', le_name_mapping)
```

Our next step was to scale the features and then split the dataset into test and train so our model could be accurately validated. We kept 80% as training data and 20% test data.
Then we used feature scaling to normalize the range of features in the data as shown.

2022 S.T. Yau High School Science Award 仅用于2022丘成桐中学科学奖公示

Feature scaling

```
from sklearn.preprocessing import StandardScaler

scaler=StandardScaler()

scaler.fit(X)
X = scaler.transform(X)
```

Our next step was to use logistic regression to train the model accurately.

Using logistic regression

```
from sklearn.linear_model import LogisticRegression
#create a model object
classifier = LogisticRegression()

#train the model object
classifier.fit(X_train,Y_train)

Y_pred=classifier.predict(X_test)
print(Y_pred)
print(list(zip(Y_test,Y_pred)))
```

After successfully training the model, we checked the accuracy of the model and found it to be efficient.

Checking the accuracy of the model

```
from sklearn.metrics import confusion_matrix, accuracy_score, classification_report

cfm=confusion_matrix(Y_test,Y_pred)
print(cfm)

print("Classification report:")

print(classification_report(Y_test,Y_pred))

acc=accuracy_score(Y_test,Y_pred)
print("Accuracy of the model: ",acc*100,'%')
```

```
[[19  6]
 [ 4 51]]
Classification report:
              precision    recall  f1-score   support

           0       0.83      0.76      0.79        25
           1       0.89      0.93      0.91        55

    accuracy                           0.88        80
   macro avg       0.86      0.84      0.85        80
weighted avg       0.87      0.88      0.87        80

Accuracy of the model:  87.5 %
```

2022 S.T. Yau High School Science Award 仅用于2022丘成桐中学科学奖公示

**Conclusion:**

This method will allow people to discover whether they have a vocal disease, such as laryngeal cancer, at an early stage, preventing significant problems from developing. These concerns are typically disregarded in the early stages since changes in voice, such as frequency, pitch, and breath, are not detectable by humans. As a result, individuals are ignorant and tend to ignore these difficulties, thus complicating the situation. Certain circumstances are highly dangerous and cannot be remedied. In certain regions of the world, such illness detection tools do not even exist, and the number of specialists in the field is also relatively restricted. This would be available to everyone and would help save many lives. It will allow the population to receive treatment on time, making it beneficial in saving lives. Many lives could be spared around the world if early detection is used. It would also drastically cut the cost of checkups because it would be a publicly accessible application. Additionally, it would ease stress on doctors, and patients would not have to undergo rigorous detection techniques such as laryngoscopy, which can be uncomfortable and require the use of sedatives. This app would employ a straightforward methodology, requiring only that the patient's voice be recorded in order to determine whether he or she has a vocal issue. It would be accomplished by a simple strategy including the usage of a basic application. This would be done using the machine learning model, which has a high accuracy of 87.5%. The model is capable of detecting all kinds of voice disorders including laryngeal cancer. The main advantage of this technology is that it does not exist and has not been worked on previously, therefore we will present people with a new perspective on typical health-related concerns.

**Future Scope:**

One of our primary goals would be to expand the dataset and improve model accuracy. We would also focus on additional uses, such as establishing prediction models that are effective at forecasting what type of person in a specific job or doing certain activities could suffer from based on their histories. We will add breathing rate to our model, which will be useful in detecting other voice-related disorders as well. Other vocal cord-related issues could be diagnosed using the extensive frequency data we have. Major studies on vocal cord issues could be conducted.

We will also work on the mobile application and its user interface to make it more user-friendly and interactive.Then we'd need to raise greater knowledge about the app and its benefits so that more people may use it and benefit on a broad scale.

2022 S.T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示

**References**

1. Pham, Minh & Lin, Jing & Zhang, Yanjia. (2018). Diagnosing Voice Disorder with Machine Learning. 5263-5266. 10.1109/BigData.2018.8622250.
2. Mittal, Vikas and R. K. Sharma. "Deep Learning Approach for Voice Pathology Detection and Classification." *IJHISI* vol.16, no.4 2021: pp.1-30. http://doi.org/10.4018/IJHISI.20211001.oa28
3. Hu H, Chang S, Wang C, Li K, Cho H, Chen Y, Lu C, Tsai T, Lee O, Deep Learning Application for Vocal Fold Disease Prediction Through Voice Recognition: Preliminary

Development Study,J Med Internet Res 2021;23(6):e25247 URL: https://www.jmir.org/2021/6/e25247 DOI: 10.2196/25247

4. Won Ki Cho, Seung-Ho Choi, Comparison of Convolutional Neural Network Models for Determination of Vocal Fold Normality in Laryngoscopic Images, Journal of Voice, 2020,ISSN 0892-1997, https://doi.org/10.1016/j.jvoice.2020.08.003. (https://www.sciencedirect.com/science/article/pii/S0892199720302927)
5. Jonathan Reid, Preet Parmar, Tyler Lund, Daniel K. Aalto, Caroline C. Jeffery, Development of a machine-learning based voice disorder screening tool, American Journal of Otolaryngology, Volume 43, Issue 2, 2022, 103327, ISSN 0196-0709, https://doi.org/10.1016/j.amjoto.2021.103327. (https://www.sciencedirect.com/science/article/pii/S0196070921004282)
6. H. Wu, J. Soraghan, A. Lowit and G. Di Caterina, "Convolutional Neural Networks for Pathological Voice Detection," *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 1-4, doi: 10.1109/EMBC.2018.8513222.

2022 S.-T. Yau High School Science Award
仅用于2022丘成桐中学科学奖公示