SNS PROJECT REPORT

SUBMITTED BY:

AYUSHI VERMA 2K19/EC/038

ARYAN 2K19/EC/032

SUBJECT: SIGNALS AND SYSTEMS (EC205)

SUBMITTED TO: Prof. M. S. Chowdhary

Department of Electronics & Communication Engineering Delhi Technological University,

Delhi Delhi-110042

November, 2020

Electronics & Communication Engg. Deptt.
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Bawana Road-Delhi-42

Certificate

I hereby certify that the Project titled "**BIRD DETECTION USING AUDIO SIGNALS THROUGH MATLAB** " which is submitted by, Department of Electronics & Communication Engineering, Delhi Technological University, Delhi is a record of the project work carried out by the students under my supervision.

Prof**.** M. S. Chowdhary
SUPERVISOR

Electronics & Communication Engg. Deptt.

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Bawana Road-Delhi-42

Candidate's Declaration

We, hereby, declare that the work embodied in this project entitled " **BIRD DETECTION USING AUDIO SIGNALS THROUGH MATLAB** " submitted to the Department of Electronics & Communication Engineering, Delhi Technological University, Delhi is an authentic record of our own bonafide work and is correct to the best of our knowledge and belief. This work has been undertaken taking care of engineering ethics.

Names of the Students:

AYUSHI VERMA 2K19/EC/038

ARYAN 2K19/EC/032

Electronics & Communication Engg. Deptt.
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Bawana Road-Delhi-42

Acknowledgement

We express our sincere thanks to Prof. Yogesh Singh, Vice-Chancellor, Delhi Technological University, Delhi. We pay our deep sense of gratitude to Prof. N. S. Raghava (HOD) of ECE Department, Delhi Technological University, Delhi to encourage us to the highest peak and to provide us the opportunity to prepare the project. We feel to acknowledge our indebtedness and deep sense of gratitude to our guide Prof**.** M. S. Chowdhary whose valuable guidance and kind supervision given to us throughout the course which shaped the present work as its show. We are immensely obliged to our friends for their elevating inspiration, encouraging guidance and kind supervision in the completion of this project.

NAME OF STUDENTS:

AYUSHI VERMA 2K19/EC/038

ARYAN 2K19/EC/032

Table of Contents

ABSTRACT

Given an unknown audio signal (which is a test signal), and a set of reference signals (which constitute the database), the task is to find which class of bird species/ animals (corresponding to the reference signal) does the given unknown signal matches the best. Concept of correlation is the key concept used to achieve this, as it gives the degree of similarity between 2 signals.

AIM: To identify a given unknown animal/bird species audio signal, through the concepts of cross-correlation.

ALGORITHM:

1. Import database audios using audio read
2. Resample all audios at same sampling frequency fsc = 30,000 Hz (common sampling frequency).
3. Extract that portion of the audio signal (of length = 0.5 seconds) which has sound of the animal/bird (by analysing it in time domain).
4. Import input audio which needs to be identified and resample it at fsc=30,000 Hz.
5. Also do time domain analysis of the input test signal and extract the portion of the signal of length=0.5 seconds which has sound of the bird/animal (which needs to be identified).
6. Convert all audio signals into frequency domain by fft function.
7. Normalise all frequency domain signals with range (1,100).
8. Now all signals (input and database signals) are ready for cross correlation as they all have common time length equivalent to 0.5 seconds which is analysed and common sampling frequency of 30k Hz and all are in frequency domain and plots are normalised with common range (1,100).
9. Take cross correlation of input test signal(unknown) with all database (known) signals.
10. The one whose xcorr at 0 lag gives maximum value is the one with which input signal matches the best.

THEORY OF CONCEPTS USED:

1. **Correlation:**

Correlation is a statistic that measures the degree to which two variables move in relation to each other.

When two sets of data are strongly linked togetherwe say they have a High Correlation. The word Correlation is made of Co- (meaning "together"), and Relation. Correlation is positive when the values increase together, and. Correlation is Negative when one value decreases as the other increases.

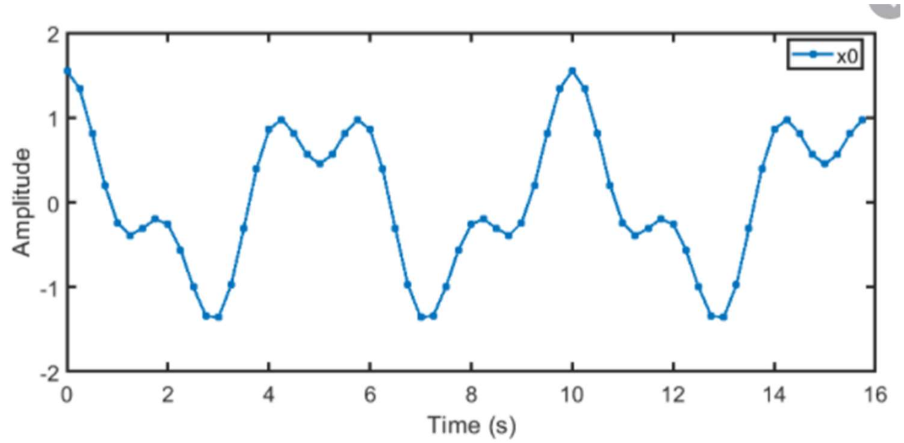
MATLAB function:

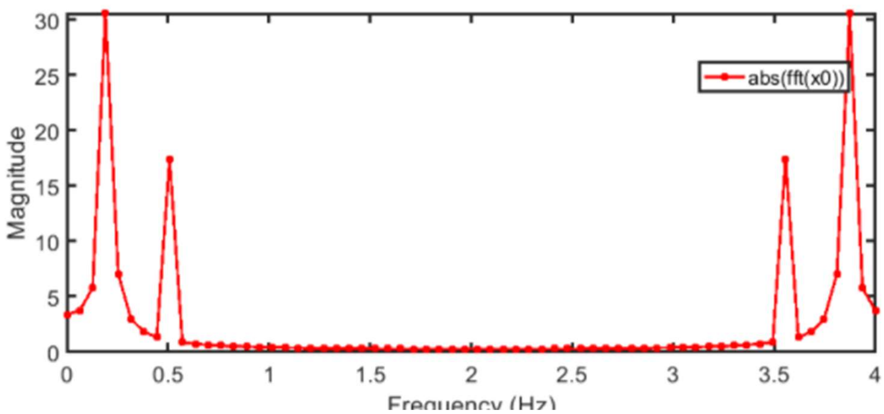
r = xcorr(x , y)

It returns the cross-correlation of two discrete-time sequences. Cross-correlation measures the similarity between a vector x and shifted (lagged) copies of a vector y as a function of the lag.

2. **Fourier transform:**

Fourier transform (FT) is a mathematical transform that decomposes a function (often a function of time, or a signal) into its constituent frequencies, such as the expression of a musical chord in terms of the volumes and frequencies of its constituent notes. The term Fourier transform refers to both the frequency domain representation and the mathematical operation that associates the frequency domain representation to a function of time.





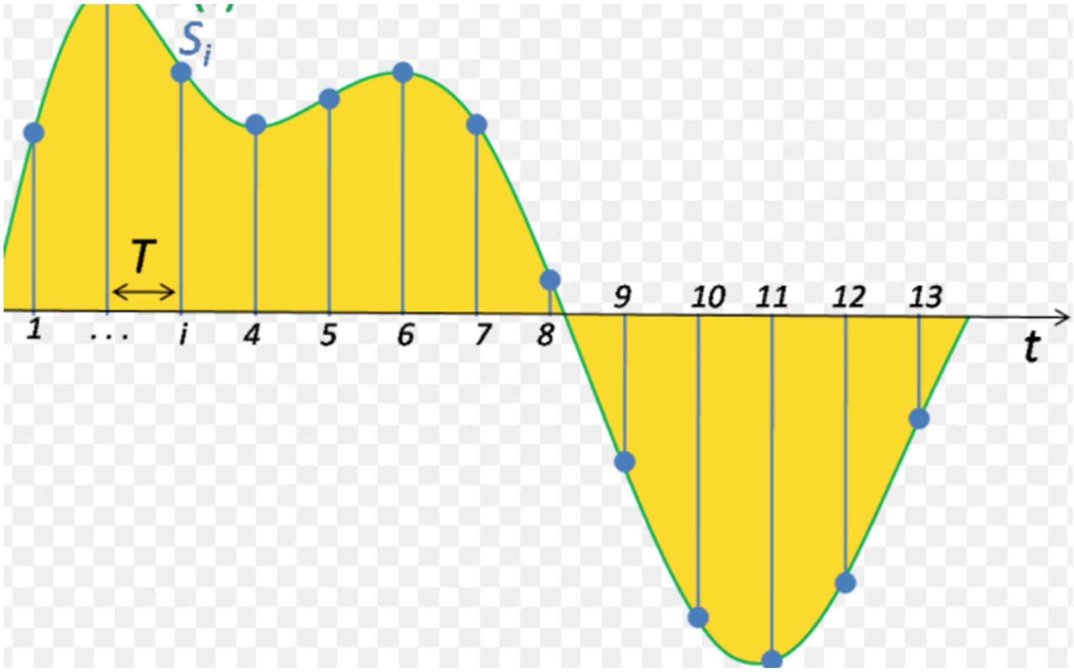
MATLAB CODE:

y = fft(x);

The fft function in MATLAB uses a fast Fourier transform algorithm to compute the Fourier transform of data.

3. **Sampling frequency:**

The sampling frequency (or sample rate) is the number of samples per second in a Sound. For example: if the sampling frequency is 44100 hertz, a recording with a duration of 60 seconds will contain 2,646,000 samples.



Common sampling frequency:

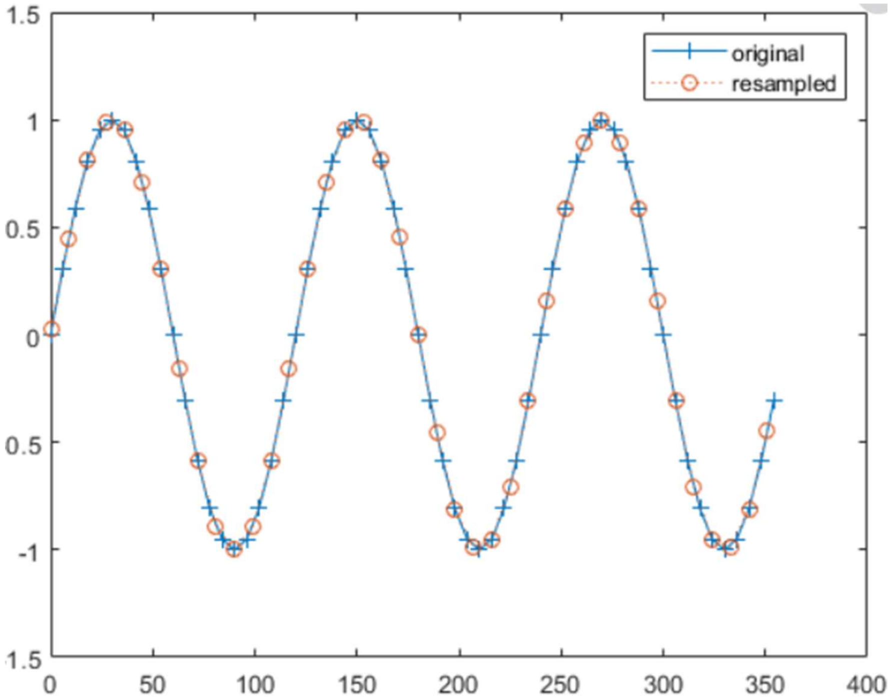
For correlation function to give correct results the output data of fft function must be sampled at same frequency. Hence, the sampling frequency and the length of input signal must be same.

So the signal must be resampled to same sampling frequency and the time length of input signal should be same (in the project, it's taken as 0.5 seconds).

Functions Used:

R = rat(X)

It returns the rational fraction approximation of X to within the default tolerance, 1e-6*norm(X(:),1). The approximation is a character array containing the truncated continued fractional expansion.

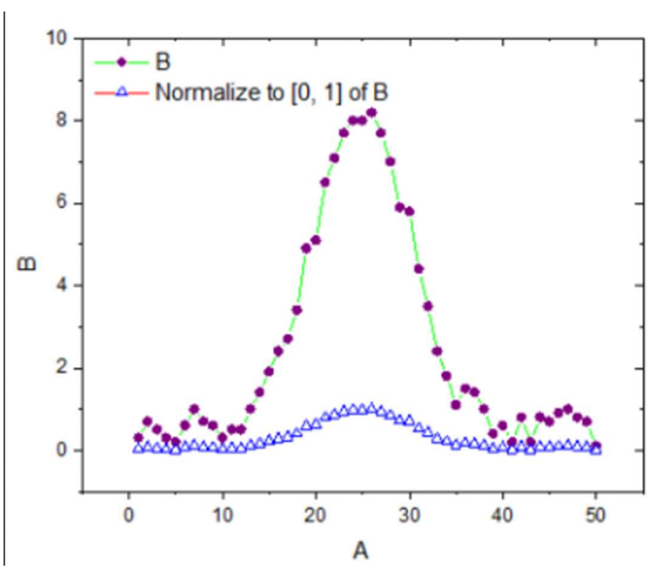


y = resample(x,p,q)

It resamples the input sequence, x, at p/q times the original sample rate.

4. **Normalisation:**

Normalization is used to scale the data of an attribute so that it falls in a smaller range, such as -1.0 to 1.0 or 0.0 to 1.0. It is generally useful for classification algorithms.



The output of Fourier transform of signal must be naormalised so that all the reference audios may have the same maximum amplitude.

SO that errors in classification due to different amplitudes of bird speaking may be reduced.

Functions used:

Nr = normalize(A,'range')

Scale A so that its range is in the interval [0,1].

AUDIO FILES USED:

These are the drive links of audio files used in the project.

TEST SIGNALS:

https://drive.google.com/drive/folders/17UpXk5EsPHbN53eAgzY0FpZbLQf-l7-Q?usp=sharing

REFERENCE SIGNALS:

https://drive.google.com/drive/folders/1CPwkclRSK8gi7v11Ikua_G7O9ML-DPz4?usp=sharing

CODE:

MAIN CODE:

```
clc;clear;

%DATA BASE SIGNALS
[ref1,fs1]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC431196 - Bar-headed Goose - Anser indicus.mp3');
[ref2,fs2]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\bat-noises.mp3');
[ref3,fs3]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\cow.mp3');
[ref4,fs4]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\elephant.mp3');
[ref5,fs5]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\parrot.wav');
[ref6,fs6]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\classic goat.mp3');
[ref7,fs7]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC369151 - Lesser Whistling Duck - Dendrocygna javanica.mp3');
[ref8,fs8]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC569023 - Snow Partridge - Lerwa lerwa.mp3');


%resample at 30,000 Hz sampling freq

fsc=30000; %fs_common

[P,Q] = rat(30000/fs1);
ref1 = resample(ref1,P,Q);
ref1=ref1(1:0.5*fsc,1); %extracting that portion of 0.5 seconds audio which has sound


[P,Q] = rat(30000/fs2);
ref2 = resample(ref2,P,Q);
ref2=ref2(1:0.5*fsc,1);



[P,Q] = rat(30000/fs3);
ref3 = resample(ref3,P,Q);
ref3=ref3(1:0.5*fsc,1);


[P,Q] = rat(30000/fs4);
ref4 = resample(ref4,P,Q);
ref4=ref4(0.5*fsc:1*fsc,1);


[P,Q] = rat(30000/fs5);
ref5 = resample(ref5,P,Q);
ref5=ref5(1:0.5*fsc,1);
```

```
[P,Q] = rat(30000/fs6);
ref6 = resample(ref6,P,Q);
ref6=ref6(1:0.5*fsc,1);


[P,Q] = rat(30000/fs7);
ref7 = resample(ref7,P,Q);
ref7=ref7(4.5*fsc:5*fsc,1);


[P,Q] = rat(30000/fs8);
ref8 = resample(ref8,P,Q);
ref8=ref8(2.2*fsc:2.7*fsc,1);


%TEST SIGNALS (the one which is to be used as test signal can be uncommented
and used)

[y_in,fs]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC431197 -
Bar-headed Goose - Anser indicus.mp3'); %take 0.5*fsc to 1*fsc

%[y_in,fs]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC460220 -
Bar-headed Goose - Anser indicus.mp3'); % 1 to 0.5*fs

%[y_in,fs]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC369150 -
Lesser Whistling Duck - Dendrocygna javanica.mp3');   %take 0.5*fsc to 1*fsc

%[y_in,fs]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC369151 -
Lesser Whistling Duck - Dendrocygna javanica.mp3'); %take 4.5*fs to 5*fs

%[y_in,fs]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC569023 -
Snow Partridge - Lerwa lerwa.mp3');      %take 2.2*fsc to 2.7*fsc

%[y_in,fs]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC569020 -
Snow Partridge - Lerwa lerwa.mp3');    %take 5*fsc to 5.5*fsc (correct)

%[y_in,fs]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS
DATASET\elephant.mp3');  %take 0.5*fsc to 1*fsc

%[y_in,fs]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS
DATASET\parrot_test.wav'); % 1 to 0.5*fsc (showing right)


%[y_in,fs]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS
DATASET\elephant8.mp3'); %0.5*fs to 1*fsc (showing wrong)
%[y_in,fs]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS
DATASET\parrot_sounds_short.wav'); % 1*fsc to 1.5*fsc (showing wrong)


%RESAMPLING OF TEST SIGNAL
[P,Q] = rat(30000/fs);
y_in = resample(y_in,P,Q);
```

```
%y_in = y_in(2.2*fsc:2.7*fsc,1);
y_in = y_in(0.5*fsc:1*fsc,1);
%y_in = y_in(4.5*fsc:5*fsc,1);
%y_in = y_in(1:0.5*fsc,1);
%y_in = y_in(5*fsc:5.5*fsc,1);
%y_in = y_in(1*fsc:1.5*fsc,1);




Y_in=abs(fft(y_in));
y_in=100*normalize(Y_in,'range');

Y1=abs(fft(ref1));
Y1=100*normalize(Y1,'range');

Y2=abs(fft(ref2));
Y2=100*normalize(Y2,'range');

Y3=abs(fft(ref3));
Y3=100*normalize(Y3,'range');

Y4=abs(fft(ref4));
Y4=100*normalize(Y4,'range');

Y5=abs(fft(ref5));
Y5=100*normalize(Y5,'range');

Y6=abs(fft(ref6));
Y6=100*normalize(Y6,'range');

Y7=abs(fft(ref7));
Y7=100*normalize(Y7,'range');

Y8=abs(fft(ref8));
Y8=100*normalize(Y8,'range');


plot(Y1);  % y axix = normalised amplitude, x axis= freq in Hz
ylim([1,100]);
title("Bar-headed Goose - Anser indicus");

plot(Y2);
ylim([1,100]);
title("BAT");

plot(Y3);
ylim([1,100]);
title("COW");
```

```
plot(Y4);
ylim([1,100]);
title("ELEPHANT");

plot(Y5);
ylim([1,100]);
title("PARROT");

plot(Y6);
ylim([1,100]);
title("GOAT");

plot(Y7);
ylim([1,100]);
title("Lesser Whistling Duck - Dendrocygna javanica");

plot(Y8);
ylim([1,100]);
title("Snow Partridge - Lerwa lerwa");


[corr_seq1 lags1] =xcorr(Y1,Y_in);
corr_seq1(lags1==0)
a=corr_seq1(lags1==0);


[corr_seq2 lags2] =xcorr(Y2,Y_in);
corr_seq2(lags2==0)
b=corr_seq2(lags2==0);


[corr_seq3 lags3] =xcorr(Y3,Y_in);
corr_seq3(lags3==0)
c=corr_seq3(lags3==0);


[corr_seq4 lags4] =xcorr(Y4,Y_in);
corr_seq4(lags4==0)
d=corr_seq4(lags4==0);

[corr_seq5 lags5] =xcorr(Y5,Y_in);
corr_seq5(lags5==0)
e=corr_seq5(lags5==0);

[corr_seq6 lags6] =xcorr(Y6,Y_in);
corr_seq6(lags6==0)
f=corr_seq6(lags6==0);


[corr_seq7 lags7] =xcorr(Y7,Y_in);
corr_seq7(lags7==0)
g=corr_seq7(lags7==0);
```

```
[corr_seq8 lags8] =xcorr(Y8,Y_in);
corr_seq8(lags8==0)
h=corr_seq8(lags8==0);


ans=[a,b,c,d,e,f,g,h];
max_val=max(ans)


if(max_val==a)
    display("the bird species is  Bar-headed Goose - Anser indicus");
end
if(max_val==b)
    display("the animal is a bat");
end
if(max_val==c)
    display("the animal is a cow");
end
if(max_val==d)
    display("the animal is an elephant");
end
if(max_val==e)
    display("the animal is a parrot");
end
if(max_val==f)
    display("the animal is a goat");
end
if(max_val==g)
    display("the bird species is Lesser Whistling Duck - Dendrocygna javanica
");
end
if(max_val==h)
    display("the bird species is Snow Partridge - Lerwa lerwa ");
end


%ans= max(all answers)
```

CODE WHICH USED TIME DOMAIN ANALYSIS OF THE SIGNAL TO EXTRACT THAT PORTION OF SIGNAL OF TIME LENGTH=0.5 SECONDS (MANUALLY) WHICH HAS SOUND OF THE BIRD:

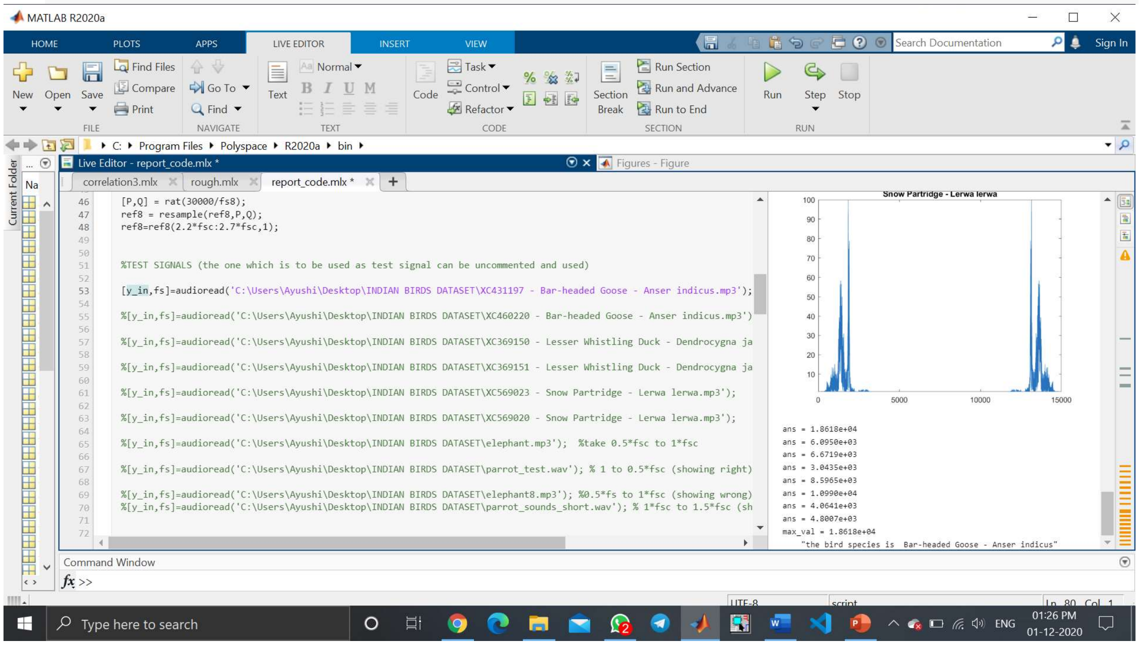
```
clc;clear;


%IMPORTING AUDIOS

[ref1,fs1]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC431196 - Bar-headed Goose - Anser indicus.mp3');

[ref2,fs2]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\bat-noises.mp3');

[ref3,fs3]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\cow.mp3');

[ref4,fs4]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\elephant.mp3');

[ref5,fs5]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\parrot_ref.aac');

[ref6,fs6]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\classic goat.mp3');

[ref7,fs7]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC369151 - Lesser Whistling Duck - Dendrocygna javanica.mp3');

[ref8,fs8]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\XC569020 - Snow Partridge - Lerwa lerwa.mp3');

[ref9,fs9]=audioread('C:\Users\Ayushi\Desktop\INDIAN BIRDS DATASET\parrot_sounds_short.wav');

%TIME DOMAIN ANALYSIS OF EACH AUDIO

t=0:1/fs1:(length(ref1)-1)/fs1;
ref=ref1(:,1);
ref=ref';
plot(t,ref);


t=0:1/fs2:(length(ref2)-1)/fs2;
ref=ref2(:,1);
ref=ref';
plot(t,ref);


t=0:1/fs3:(length(ref3)-1)/fs3;
ref=ref3(:,1);
ref=ref';
```

```
plot(t,ref);



t=0:1/fs4:(length(ref4)-1)/fs4;
ref=ref4(:,1);
ref=ref';
plot(t,ref);


t=0:1/fs5:(length(ref5)-1)/fs5;
ref=ref5(:,1);
ref=ref';
plot(t,ref);


t=0:1/fs6:(length(ref6)-1)/fs6;
ref=ref6(:,1);
ref=ref';
plot(t,ref);


t=0:1/fs7:(length(ref7)-1)/fs7;
ref=ref7(:,1);
ref=ref';
plot(t,ref);


t=0:1/fs8:(length(ref8)-1)/fs8;
ref=ref8(:,1);
ref=ref';
plot(t,ref);


t=0:1/fs9:(length(ref9)-1)/fs9;
ref=ref9(:,1);
ref=ref';
plot(t,ref);
```

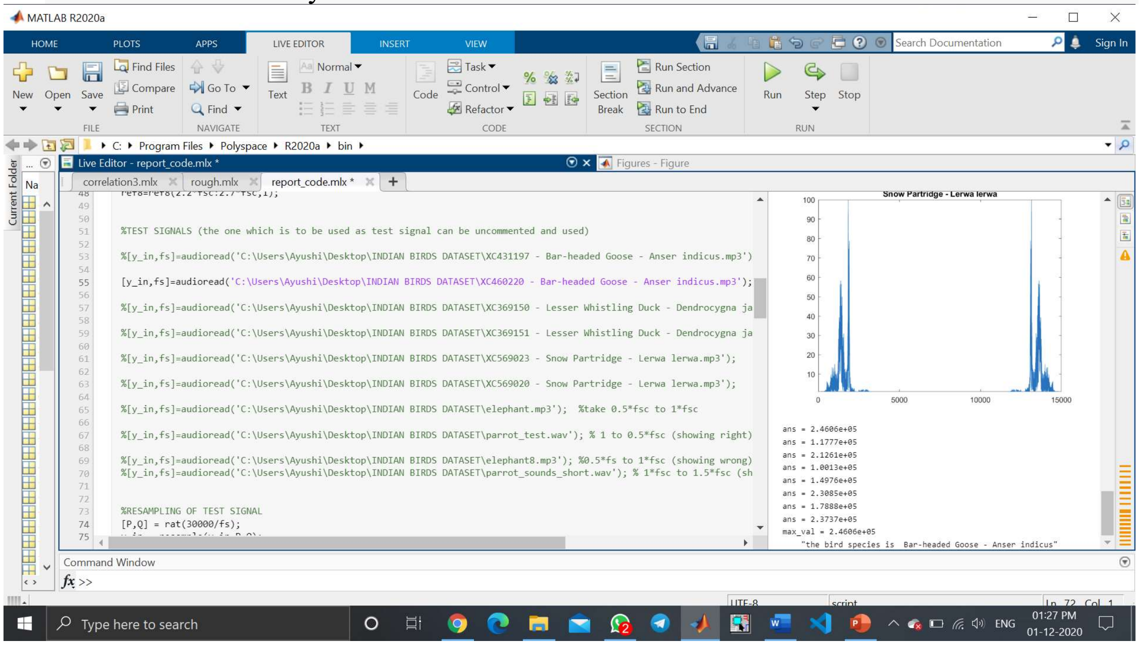
OUTPUTS:

MAIN CODE:

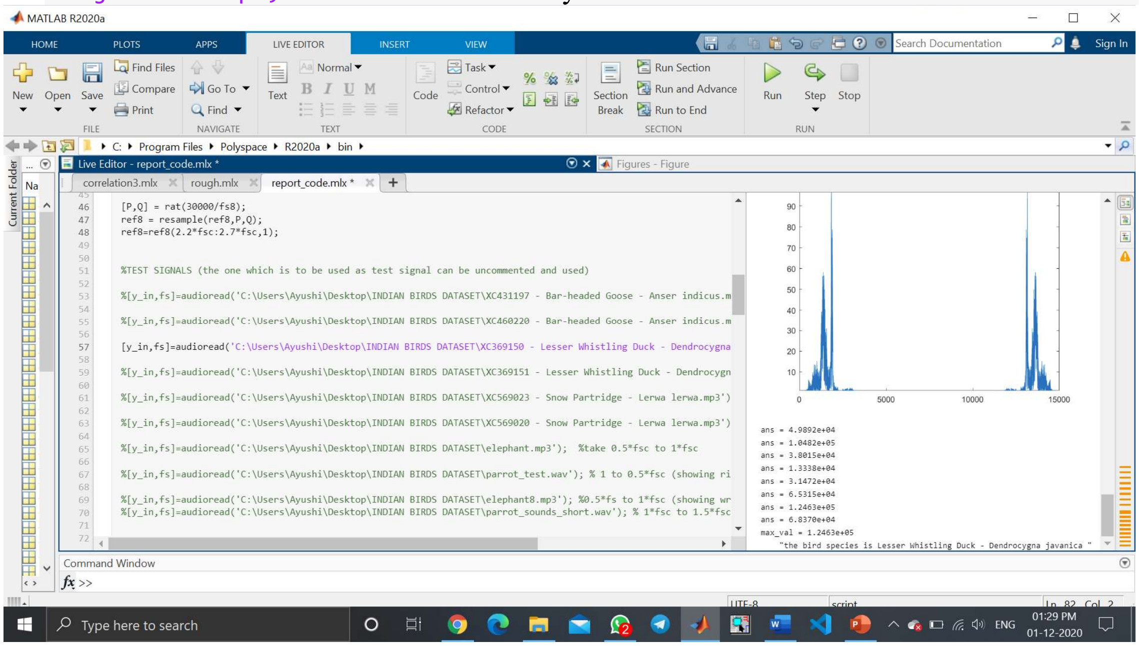
1. When test signal is `XC431197 - Bar-headed Goose - Anser indicus.mp3` , it is identified correctly.



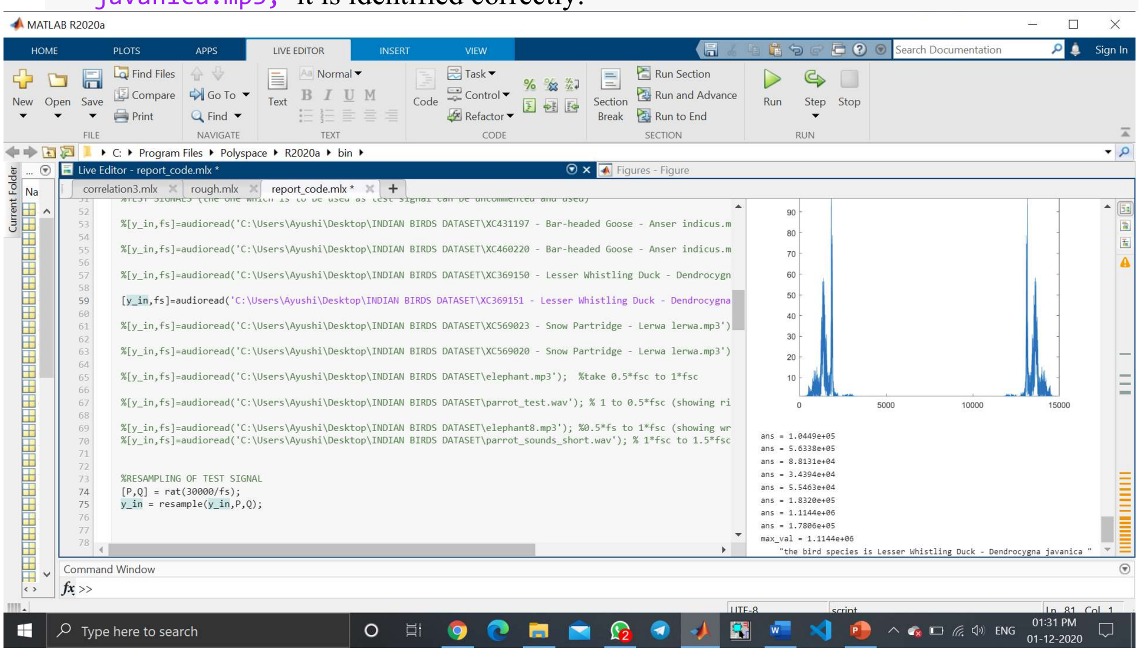
2. When test signal is `XC460220 - Bar-headed Goose - Anser indicus.mp3` , it is identified correctly.



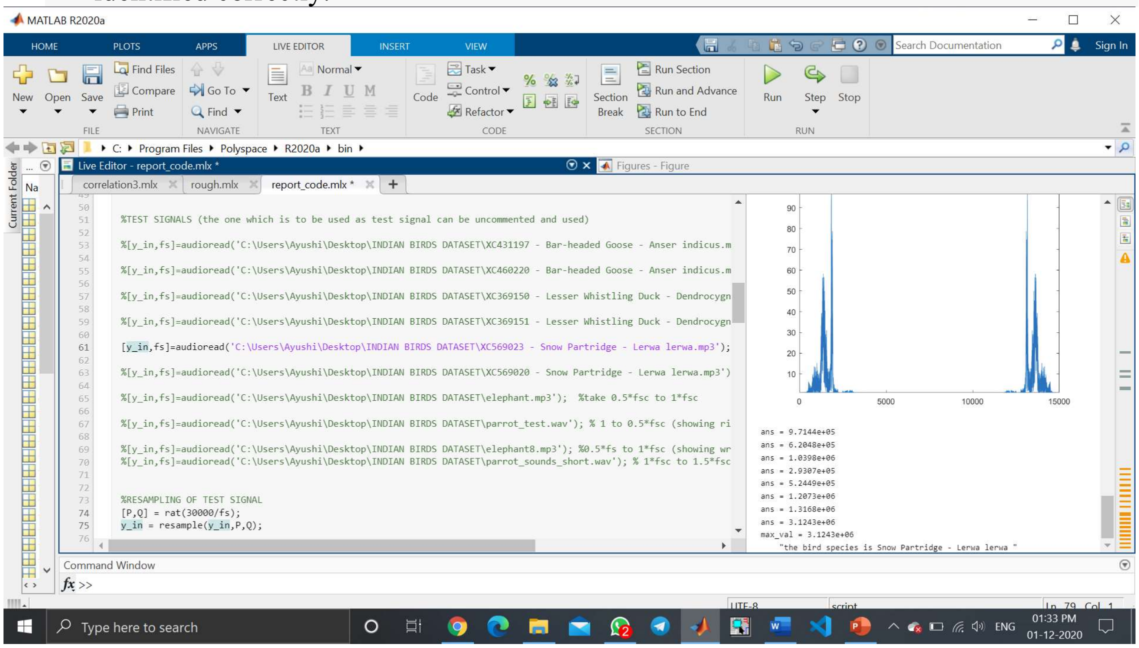
3. When test signal is `XC369150 - Lesser Whistling Duck - Dendrocygna javanica.mp3`, it is identified correctly.



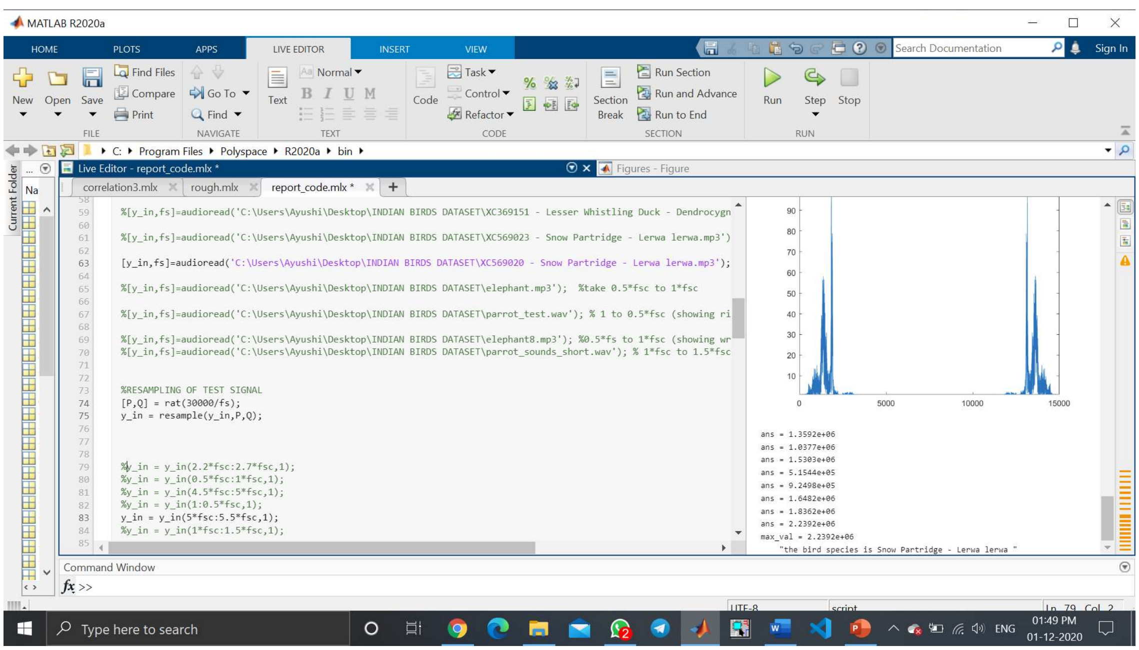
4. When test signal is `XC369151 - Lesser Whistling Duck - Dendrocygna javanica.mp3`, it is identified correctly.



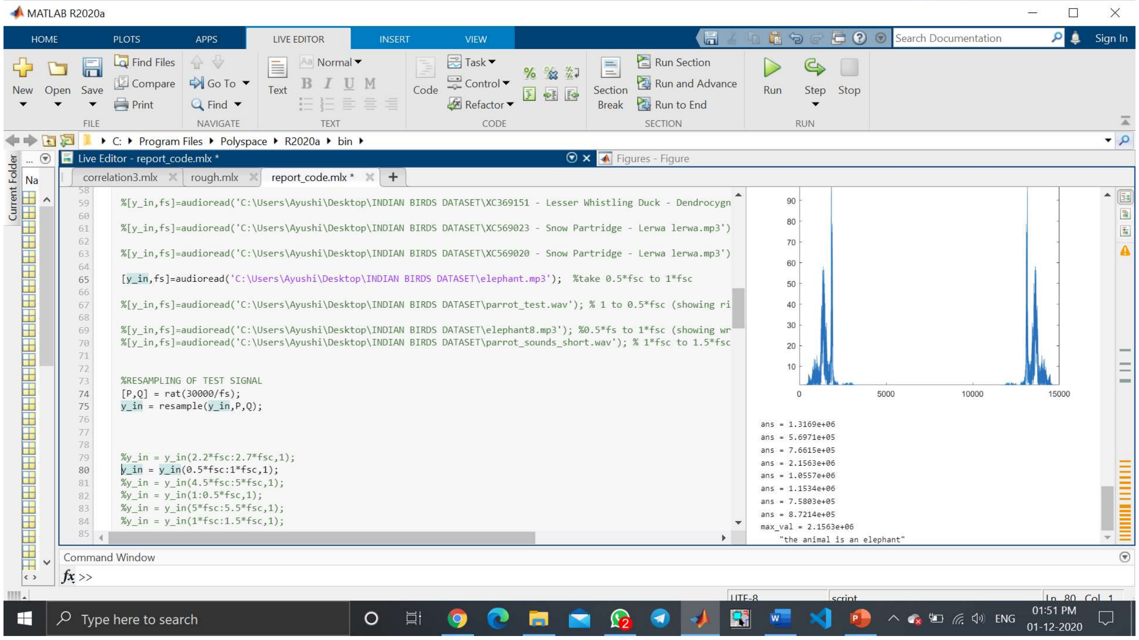
5. When test signal is `XC569023 - Snow Partridge - Lerwa lerwa.mp3` , it is identified correctly.



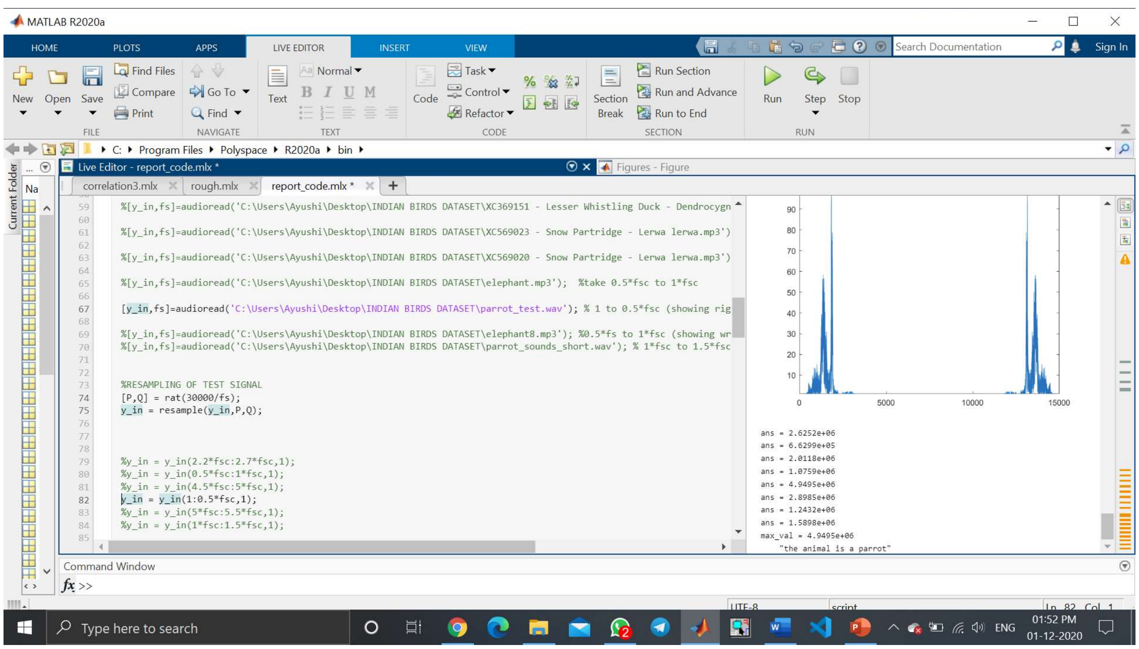
6. When test signal is `XC569020 - Snow Partridge - Lerwa lerwa.mp3`, it is identified correctly.



7. When test signal is `elephant.mp3`, it is identified correctly.

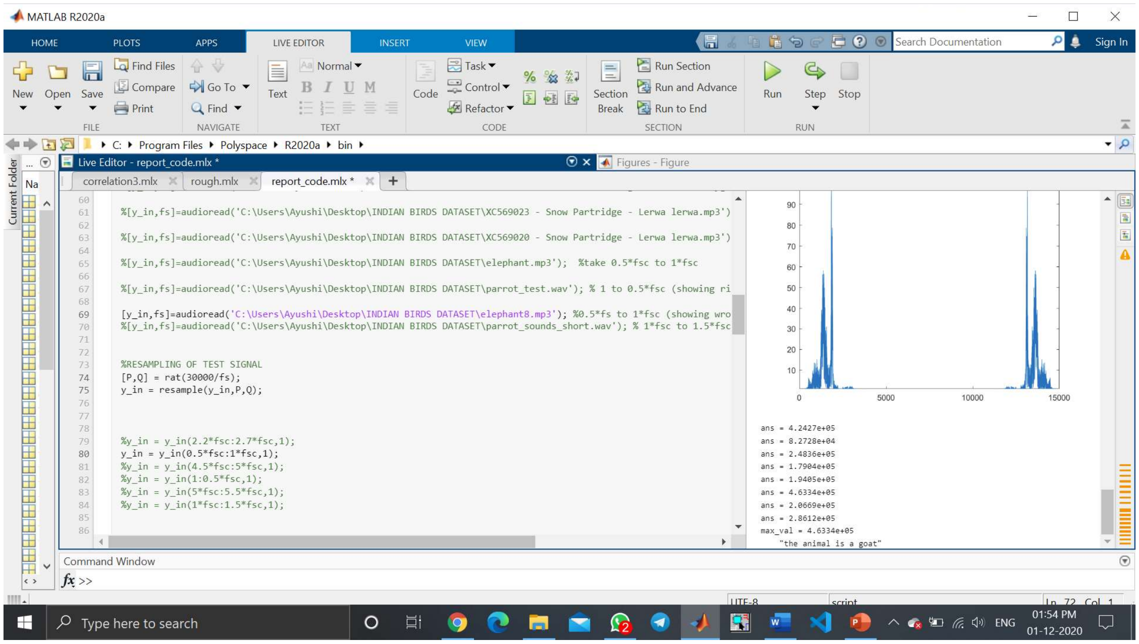


8. When test signal is `parrot_test.wav`, it is identified correctly.



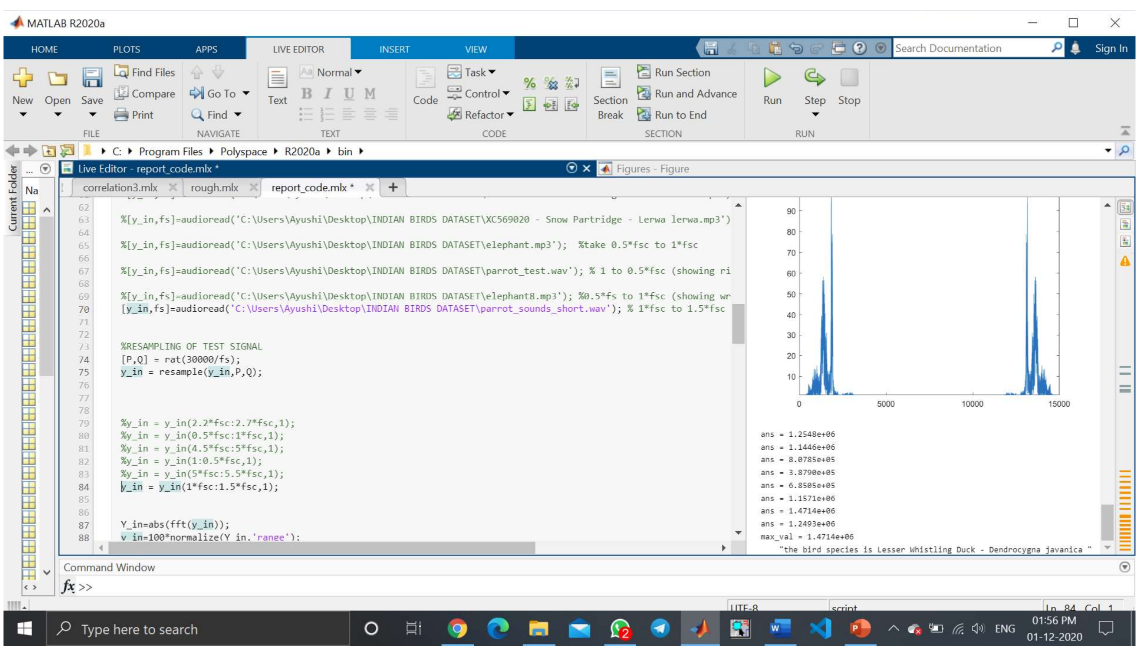
9. When test signal is `elephant8.mp3`, it is identified incorrectly. (since no classification model is 100% accurate)

Here elephant is getting recognised as a goat.

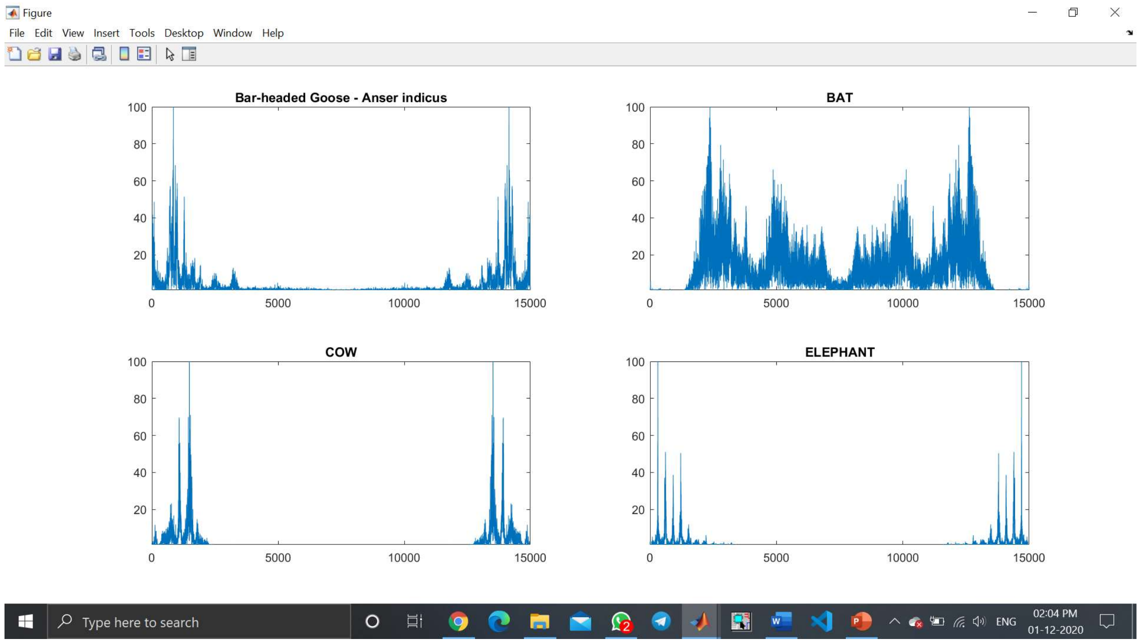


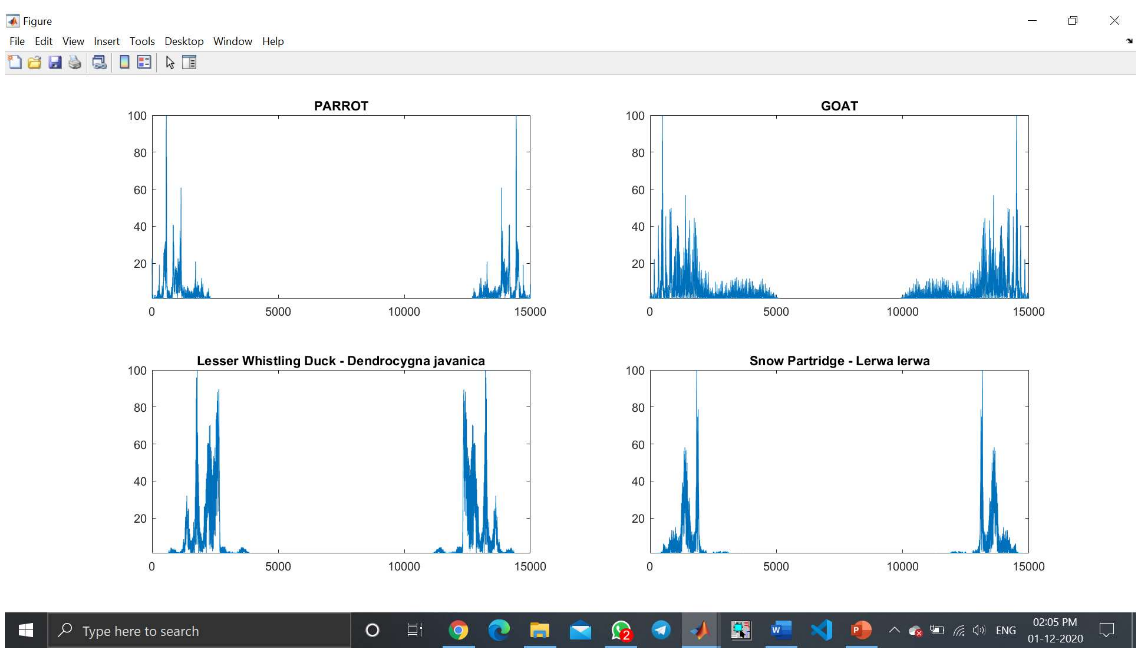
10. When test signal is `parrot_sounds_short.wav`, it is identified incorrectly. (since no classification model is 100% accurate)

Here parrot is getting recognised as lesser whistling duck.



11. NORMALISED FREQUENCY PLOTS OF ALL ANIMALS IN REFERENCE DATABASE





SOURCES OF ERRORS:

- Errors due to noise.
- Errors due to different species of the same bird (eg: there are so many species of same bird parrot).
- Errors due to sound made by single bird and group of birds (if both are considered as same).
- Errors due do similar frequency of different birds (or birds of different species).

CONCLUSION:

We were able to identify some of the test audio signals of the unknown bird species correctly, while 2 test signals (out of 10) were identified incorrectly, as no classification model can be 100% accurate. Correlation is used as a key concept for finding the similarity of input test signal with the database signals (from which the test signal is to be matched and identified). The database signals consisted of 8 animal/bird species, namely:

- Bar-headed Goose - Anser indicus
- bat
- cow
- elephant
- a parrot
- a goat
- Lesser Whistling Duck - Dendrocygna javanica
- Snow Partridge - Lerwa lerwa

And we tested the code with 10 different test signals of birds/animal species whose audio is present in the database, and for 80% cases we got the correct output.

REFERENCES:

1. Site referred for dataset of audio signals:

- https://www.xeno-canto.org/explore?query=area%3A%22asia%22%20cnt%3A%22India%22

2. Sites referred for theory of concepts used:

- https://www.mathsisfun.com/data/correlation.html
- https://www.statisticshowto.com/probability-and-statistics/normal-distributions/normalized-data-normalization/#:~:text=Normalization%20usually%20means%20to%20scale,a%20standard%20deviation%20of%201
- https://en.m.wikipedia.org/wiki/Resampling_(statistics)
- https://www.mathworks.com/help/matlab/math/fourier-transforms.html