SECTION A

Q1. a)

$D = 1 \rightarrow$ Dividend issued
$\phantom{D = }0 \rightarrow$ Not issued

Q1.(a) Let $D$ represent the RV of a dividend being issued. Let $X$ represent RV of amount of profit ~~gained~~ increase.

$$P[X \mid D=1] = N(10, 6)$$

Standard Normal Gaussian; mean; $\rightarrow$ Std dev $= 6$ since variance $= 36\%$

$$= \frac{1}{\sqrt{2\pi 6^2}} e\left\{-\frac{1}{2}\frac{(x-10)^2}{6^2}\right\}$$

$$P[X \mid D=0] = N(0, 6)$$

$$= \frac{1}{\sqrt{2\pi(36)}} e\left\{-\frac{1}{2}\frac{x^2}{36}\right\}$$

$$P[D=1] = 0.8$$

$$P[D=0] = 1 - P[D=1] = 0.2$$

$$P[X] = P[X \mid D=1]\,P[D=1] + P[X \mid D=0]\,P[D=0]$$

$$= \frac{1}{\sqrt{72\pi}}\left(e^{-x^2/72}(0.8) + e^{-(x-10)^2/72}(0.2)\right)$$

Required result $\rightarrow P[D=1 \mid X=4]$

$\rightarrow$ Dividend issue given profits of 4% increase.

$$P[D=1 \mid X=4] = \frac{P[X=4 \mid D=1]\, P[D=1]}{P[X=4]}$$

$\downarrow$

Bayes Rule

$$= \frac{\frac{1}{\sqrt{72\pi}}\, e^{-36/72}\,(0.8)}{\frac{1}{\sqrt{72\pi}}\left(e^{-36/72}(0.8) + e^{-16/72}(0.2)\right)}$$

$$= \frac{0.8\, e^{-36/72}}{0.8\, e^{-36/72} + 0.2\, e^{-16/72}}$$

$$= \frac{0.485}{0.485 + 0.16} = 0.752$$

$\therefore$ There is a 75.2% chance, the company will issue a dividend.

Q1.b)

(b) Target Var → 'Attended ML Class' ($Y$)
Features → 'Class Time', 'Had Proper Sleep' & 'Weather'

'Class Time' → $X_1$, 'Had Proper Sleep' → $X_2$, 'Weather' → $X_3$

First Split → 1) $H(Y) - H(Y|X_1)$

$$H(Y) = -P[Y='Yes']\log_2 P[Y='Yes'] - P[Y='No']\log_2 P[Y='No']$$

$$= -\frac{7}{12}\log_2\frac{7}{12} - \frac{5}{12}\log_2\frac{5}{12}$$

$$= 0.454 + 0.526 = 0.980$$

$$H(Y|X_1) = P[X_1='Morning']\,H(Y|X_1='Morning') + P[X_1='Noon']\,H(Y|X_1='Noon') + P[X_1='Afternoon']\,H(Y|X_1='Afternoon')$$

$$= \frac{4}{12}\left(-\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4}\right) + \frac{4}{12}\left(-\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4}\right) + \frac{4}{12}\left(-\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4}\right)$$

$$= \frac{1}{3}\left(2 + 0.311 + 0.5\right) = 0.937$$

$H(Y) - H(Y|X_1) = 0.43$

2) $H(Y) - H(Y|X_2)$

$H(Y|X_2) = P[X_2 = \text{'Yes'}]\ H(Y|X_2 = \text{'Yes'})$
$+ P[X_2 = \text{'No'}]\ H(Y|X_2 = \text{'No'})$

$= \frac{6}{12}\left(-\frac{6}{6}\log\frac{6}{6} - \frac{0}{6}\log\frac{0}{6}\right)$

$+ \frac{6}{12}\left(-\frac{5}{6}\log\frac{5}{6} - \frac{1}{6}\log\frac{1}{6}\right)$

$= \frac{6}{12}\left(0 + 0.219 + 0.43\right)$

$= 0.3245$

$H(Y) - H(Y|X_2)$
$= 0.98 - 0.3245$
$= 0.6555$

3) $H(Y) - H(Y|X_3)$

$H(Y|X_3) = P[X_3 = \text{'Cool'}]\ H(Y|X_3 = \text{'Cool'})$
$+ P[X_3 = \text{'Hot'}]\ H(Y|X_3 = \text{'Hot'})$
$+ P[X_3 = \text{'Rainy'}]\ H(Y|X_3 = \text{'Rainy'})$

$$= \frac{5}{12}\left(-\frac{4}{5}\log\frac{4}{5} - \frac{1}{5}\log\frac{1}{5}\right) + \frac{2}{12}\left(-\frac{0}{2}\log\frac{0}{2} - \frac{2}{2}\log\frac{2}{2}\right)$$

$$+ \frac{5}{12}\left(-\frac{3}{5}\log\frac{3}{5} - \frac{2}{5}\log\frac{2}{5}\right)$$

$$= \frac{5}{12}\left(0.258 + 0.464 + 0.442 + 0.529\right)$$

$$= 0.705$$

$$H(Y) - H(Y|X_3)$$
$$= 0.275$$

Max Info Gain is for $X_2 \rightarrow$ 'Had Proper Sleep'

DT $\rightarrow$

Had Proper Sleep

Yes $\downarrow$ (In this branch all datapoints attended ML class – No more division required) $\downarrow$ $\boxed{\text{Attended ML Class}}$

No $\downarrow$ 5 Nos, 1 Yes $\downarrow$ Need to divide further

For 'Had No Proper Sleep'

| $X_1$ Class Time | $X_3$ Weather | $Y$ Attended ML Class |
|---|---|---|
| Morning | Rainy | No |
| Morning | Cool | Yes |
| Noon | Hot | No |
| Noon | Cool | No |
| Afternoon | Rainy | No |
| Afternoon | Hot | No |

$$H(Y) = -\frac{5}{6}\log\frac{5}{6} - \frac{1}{6}\log\frac{1}{6}$$

$$= 0.649$$

1) $H(Y) - H(Y|X_1)$

$$H(Y|X_1) = \frac{2}{6}\left(-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2}\right)$$

$$+\frac{2}{6}\left(-\frac{0}{2}\log\frac{0}{2} - \frac{2}{2}\log\frac{2}{2}\right)$$

$$+\frac{2}{6}\left(-\frac{0}{2}\log\frac{0}{2} - \frac{2}{2}\log\frac{2}{2}\right)$$

$$= \frac{2}{6} = 0.333$$

$$H(Y) - H(Y|X_1) = 0.316$$

2) $H(Y) - H(Y|X_3)$

$$H(Y|X_3) = \frac{2}{6}\left(-\frac{0}{2}\log\frac{0}{2} - \frac{2}{2}\log\frac{2}{2}\right)$$

$$+ \frac{2}{6}\left(-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2}\right)$$

$$+ \frac{2}{6}\left(-\frac{0}{2}\log\frac{0}{2} - \frac{2}{2}\log\frac{2}{2}\right)$$

$$= 0.333$$

$$H(Y) - H(Y|X_3) = 0.316$$

$\Rightarrow$ Same info gain, so, I randomly choose Weather to split on.

New Decision Tree →

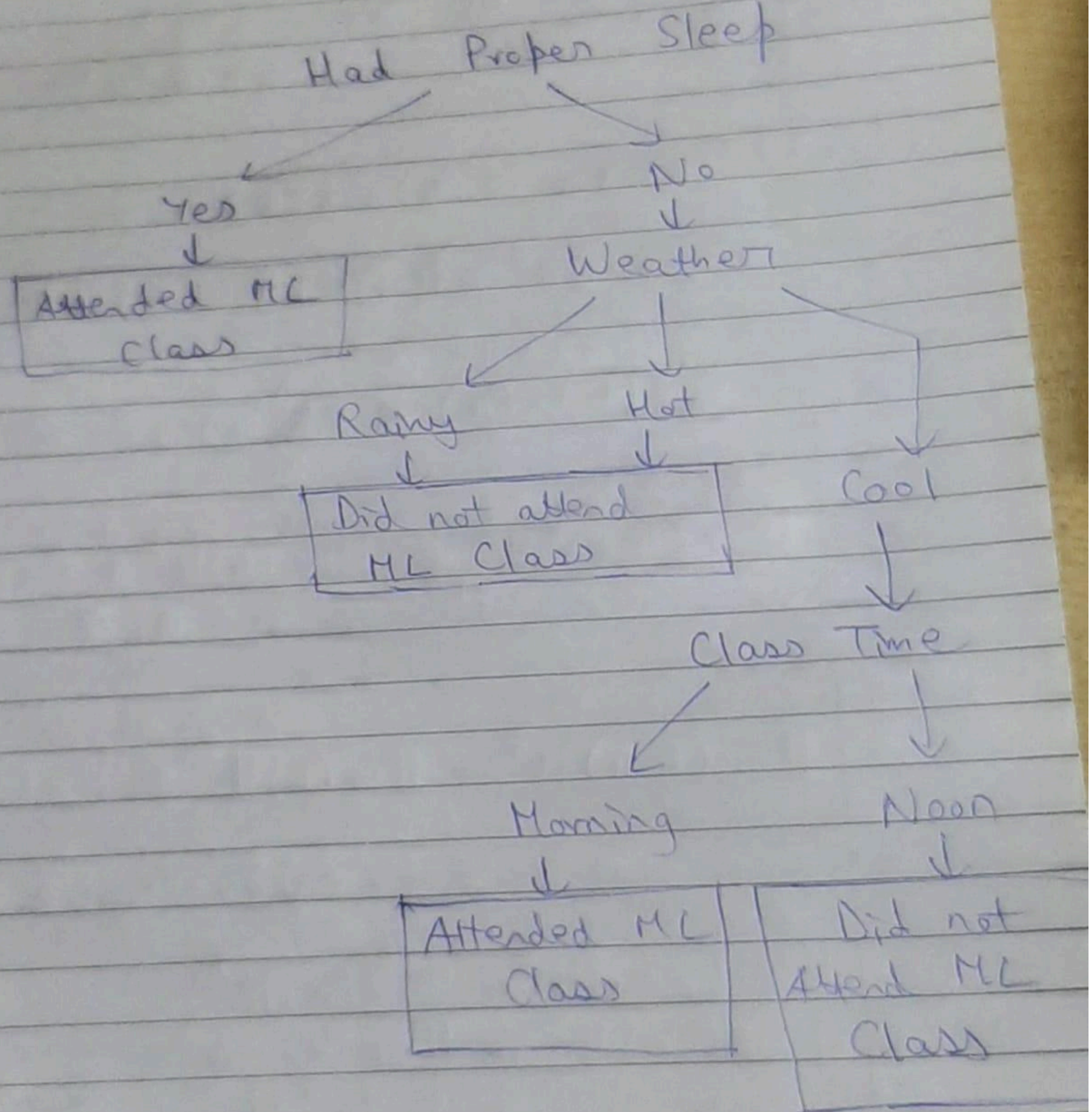
- Had Proper Sleep
  - Yes → Attended ML Class
  - No → Weather
    - Rainy → Did not Attend ML Class
    - Cool → 1Y, 1N → Need to split
    - Hot → Did not attend ML Class

⇒ For 'Not Had Proper Sleep' & 'Cool weather'

| Class Time ($X_1$) | Attended ML Class |
|---|---|
| Morning | Yes |
| Noon | No |

Since only 1 attribute left, we don't need to find info gain, So we make the final decision tree.

∴




Q1.c)

Q1. (c) Let's assume there's misclassification for $x(1), x(2), \ldots, x(n) \in C_1$

$w(n+1) = w(n) + x(n)$

$w(0) = 0$

$w(1) = x(0)$

$w(2) = w(1) + x(1)$

$w(n+1) = x(0) + x(1) + \ldots + x(n)$

Since $S$ is linearly separable (given in Q), there exists optimal $w_0$ such

that $w_0^T x(n) > \frac{\gamma}{2}$ for $x(1), x(2), \ldots, x(n) \in C_1$

$\alpha = \min_{x(n) \in C_1} w_0^T x(n) \Rightarrow \alpha > \frac{\gamma}{2}$

$\Rightarrow w_0^T w(n+1) = w_0^T x(1) + w_0^T x(2) + \ldots w_0^T x(n)$

$w_0^T w(n+1) >= n\alpha \geq \frac{n\gamma}{2}$

$\|w_0^T\|^2 \|w(n+1)\|^2 >= [w_0^T w(n+1)]^2$

↳ By Cauchy Schwarz

$\Rightarrow \|w_0^T\|^2 \|w(n+1)\|^2 >= \frac{n^2 \gamma^2}{4}$

$\Rightarrow \|w(n+1)\|^2 >= \frac{n^2 \gamma^2}{4 \|w_0^T\|^2}$

$w(k+1) = w(k) + x(k)$

$\|w(k+1)\|^2 = \|w(k)\|^2 + \|x(k)\|^2 + 2w^T(k)\, x(k)$

$w^T(k)\, x(k) < \frac{\gamma}{2}$ → since misclassification

$\|w(k+1)\|^2 \leq \|w(k)\|^2 + \|x(k)\|^2 + \gamma$

$\|w(k+1)\|^2 - \|w(k)\|^2 \leq \|x(k)\|^2 + \gamma$

Summing over $k = 0$ to $n$

$$\begin{aligned} &\|w(n+1)\|^2 - \|w(n)\|^2 \\ &+ \|w(n)\|^2 - \|w(n-1)\|^2 \\ &+ \\ &\vdots \\ &\|w(2)\|^2 - \|w(1)\|^2 \\ &\cancel{\|w(1)\|^2 - \|w(0)\|^2} \end{aligned} \leq \sum_{i=1}^{n} \left(\|x(k)\|^2 + \gamma\right)$$

↓ since $\|x(k)\|^2 = 1$ according to question

$$||w(n+1)||^2 \leq \underbrace{n + n\gamma}_{\text{Since } ||x(k)||^2 = 1}$$

Also, $$||w(n+1)||^2 \geq \frac{n^2\gamma^2}{4||w_0^T||^2}$$

Let us consider the max n can go till.
At this point, both limits equalize

$$||w(n+1)||^2 = n + n\gamma$$
$$||w(n+1)||^2 = \frac{n^2\gamma^2}{4||w_0^T||^2}$$

→ Here n is referring to $n_{max}$

$$\frac{n^2\gamma^2}{4||w_0^T||^2} = n(1+\gamma) \Rightarrow \frac{n\gamma^2}{4||w_0^T||^2} = 1+\gamma$$

$$\therefore n = \frac{(1+\gamma)\, 4\, \|w_0^T\|^2}{\gamma^2}$$

$\Rightarrow$ If margin is $(1-\varepsilon)\gamma$

Following same steps,

$$w_0^T w(n+1) \geq n(1-\varepsilon)\gamma$$

$$\|w(n+1)\|^2 \geq \frac{n^2 \gamma^2 (1-\varepsilon)^2}{\|w_0^T\|^2}$$

Similarly for lower bound,

$$\|w(n+1)\|^2 \leq n + n(2)((1-\varepsilon)\gamma)$$

$$= n + 2\gamma(1-\varepsilon)n$$

$$\frac{n_{max}^2 \gamma^2 (1-\varepsilon)^2}{\|w_0^T\|^2} = n_{max}(1+2\gamma(1-\varepsilon))$$

$$n_{max} = \frac{\|w_0^T\|^2 (1+2\gamma(1-\varepsilon))}{\gamma^2 (1-\varepsilon)^2}$$

$\hookrightarrow$ Close to 1

$$n_{max} = \frac{\|w_0^T\|^2 (2\gamma + 1 - 2\gamma\varepsilon)}{\gamma^2}$$

If $\varepsilon$ very small

$$n_{max} = \frac{\|w_0^T\|^2 (2\gamma + 1)}{\gamma^2}$$

$\therefore$ So, $n_{max}$ is polynomial in $\frac{1}{\gamma}$ & is $\leq \frac{\|w_0^T\|^2 (2\gamma + 1)}{\gamma^2}$

when margin is $(1-\varepsilon)\gamma$

For margin $= \frac{\gamma}{2}$ $\rightarrow$ $n \leq \frac{(1+\gamma)\, 4 \|w_0^T\|^2}{\gamma^2}$

Q1.(d).(a)

Q1. (d) (a) $\Rightarrow P[\text{"buy"} = 1 \mid \text{Spam}]$

$$= \frac{P[\text{"buy"} = 1 \cap \text{Spam}]}{P[\text{Spam}]}$$

$$= \frac{2/4}{2/4} = 1$$

$\Rightarrow P[\text{"buy"} = 0 \mid \text{Spam}] = 1 - P[\text{"buy"} = 1 \mid \text{Spam}]$
$= 1 - 1 = 0$

$\Rightarrow P[\text{"buy"} = 1 \mid \text{Not Spam}]$
$$= \frac{P[\text{"buy"} = 1 \cap \text{Not Spam}]}{P[\text{Not Spam}]}$$

Since they are exhaustive & exclusive.

$$= \frac{1/4}{2/4} = \frac{1}{2}$$

$P[\text{"buy"} = 0 \mid \text{Not Spam}]$
$= 1 - \frac{1}{2} = \frac{1}{2}$

$\Rightarrow P[\text{"cheap"} = 1 \mid \text{Spam}] = \frac{P[\text{"cheap"} = 1 \cap \text{Spam}]}{P[\text{Spam}]}$

$$= \frac{1/4}{2/4} = \frac{1}{2}$$

$P[\text{"cheap"} = 0 \mid \text{Spam}]$
$= 1 - P[\text{"cheap"} = 1 \mid \text{Spam}] = 1/2$

$\Rightarrow P[\text{"cheap"} = 1 \mid \text{Not Spam}] = \frac{P[\text{"cheap"} = 1 \cap \text{Not Spam}]}{P[\text{Not Spam}]}$

$$= \frac{1/4}{2/4} = 1/2$$

$P[\text{"cheap"} = 0 \mid \text{Not Spam}]$
$= 1 - \frac{1}{2} = \frac{1}{2}$

P["buy"=0|Spam]=0 ∴ P["buy"=1|Spam] = 1
P["buy"=0|Not Spam]=0.5 P["buy"=1|Not Spam] = 0.5
P["cheap"=0|Spam]=0.5 P["cheap"=1|Spam] = 0.5
P["cheap"=0|Not Spam]=0.5 P["cheap"=1|Not Spam] = 0.5

1.(d).(b)

(d) (b) Posteriors → P[Spam | "cheap"=1 ∩ "buy"=0]
P[Not Spam | "cheap"=1 ∩ "buy"=0]

$$P[\text{Spam} \mid \text{"cheap"}=1 \cap \text{"buy"}=0] = \frac{P[\text{"cheap"}=1 \cap \text{"buy"}=0 \mid \text{Spam}] \times P[\text{Spam}]}{P[\text{"cheap"}=1 \cap \text{"buy"}=0]}$$

By independence of features assumption in Naive Bayes

$$= \frac{P[\text{Spam}]\, P[\text{"cheap"}=1 \mid \text{Spam}]\, P[\text{"buy"}=0 \mid \text{Spam}]}{P[\text{"cheap"}=1 \cap \text{"buy"}=0]}$$

$$P[\text{Spam}] = \frac{\text{No. of Spam Mails}}{\text{Total No. of Mails}} = \frac{2}{4} = \frac{1}{2}$$

$$P[\text{Not Spam}] = 1 - P[\text{Spam}] = \frac{1}{2}$$

$P[\text{Spam} \mid \text{"cheap"}=1 \cap \text{"buy"}=0]$

$$= \frac{\frac{1}{2} \times \frac{1}{2} \times 0}{P[\text{"cheap"}=1 \cap \text{"buy"}=0]} = 0$$

$P[\text{Not Spam} \mid \text{"cheap"}=1 \cap \text{"buy"}=0]$

$$= \frac{P[\text{"cheap"}=1 \cap \text{"buy"}=0 \mid \text{Not Spam}]\; P[\text{Not Spam}]}{P[\text{"cheap"}=1 \cap \text{"buy"}=0]}$$

$$= \frac{P[\text{Not Spam}]\; P[\text{"cheap"}=1 \mid \text{Not Spam}]\; P[\text{"buy"}=0 \mid \text{Not Spam}]}{P[\text{"cheap"}=1 \cap \text{"buy"}=0]}$$

$$= \frac{0.5 \times 0.5 \times 0.5}{\left( P[\text{"cheap"}=1 \cap \text{"buy"}=0 \mid \text{Not Spam}]\; P[\text{Not Spam}] + P[\text{"cheap"}=1 \cap \text{"buy"}=0 \mid \text{Spam}]\; P[\text{Spam}] \right)}$$

Calculated in prev step

↳ Calculated at top of image

$$= \frac{0.5 \times 0.5 \times 0.5}{(0.5 \times 0.5 \times 0.5) + (0)} = 1$$

Posteriors ↓

$\therefore P[\text{Spam} \mid \text{"cheap"}=1 \cap \text{"buy"}=0] = 0$

$P[\text{Not Spam} \mid \text{"cheap"}=1 \cap \text{"buy"}=0] = 1$

$\therefore$ We declare email is <u>not spam</u>.

Q1.(d).(c)

(d)(c) The problem with zero probabilities is that if a particular tuple has never existed before, it can make ~~every~~ the probability ~~conditional likelihood~~ of the parameter 0, thus the posterior becomes 0, which isn't accurate since it's a small sample size most of the time and the lack of an entry shouldn't lead to the posterior being 0.

To fix this, we may use Laplace Smoothing or m-estimate smoothing.

Laplace $\rightarrow P(A_i \mid C) = \dfrac{N_{ic}+1}{N_c + c}$

M-Estimate $\rightarrow \dfrac{N_{ic}+mp}{N_c + m}$

The additional terms 1, c, mp, m are all there so that the ~~likelihoods~~ probabilities are small but not 0 to not skew the result.

In prev example, let's do m-estimate smoothing

$m = 2 \quad p(\text{Spam}) = p(\text{Not Spam}) = 0.5$

$$P[\text{"buy"} = 0 \mid \text{Spam}] = \frac{0 + 0.5(2)}{2 + 2}$$

$$= \frac{1}{4} = 0.25$$

$$P[\text{"buy"} = 0 \mid \text{Not Spam}] = \frac{1 + 0.5(2)}{2 + 2} = 0.5$$

$$P[\text{"cheap"} = 1 \mid \text{Spam}] = \frac{1 + 0.5(2)}{2 + 2} = 0.5$$

$$P[\text{"cheap"} = 1 \mid \text{Not Spam}] = \frac{1 + 0.5(2)}{2 + 2} = 0.5$$

$P[\text{Spam} \mid \text{"cheap"} = 1 \cap \text{"buy"} = 0]$

$$= \frac{0.5 \times 0.5 \times 0.25}{0.5 \times 0.5 \times 0.25 + 0.5 \times 0.5 \times 0.5}$$

$$= \frac{0.25}{0.25 + 0.5} = \frac{1}{3}$$

$P[\text{Not Spam} \mid \text{"Buy"} = 0 \cap \text{"Cheap"} = 1]$

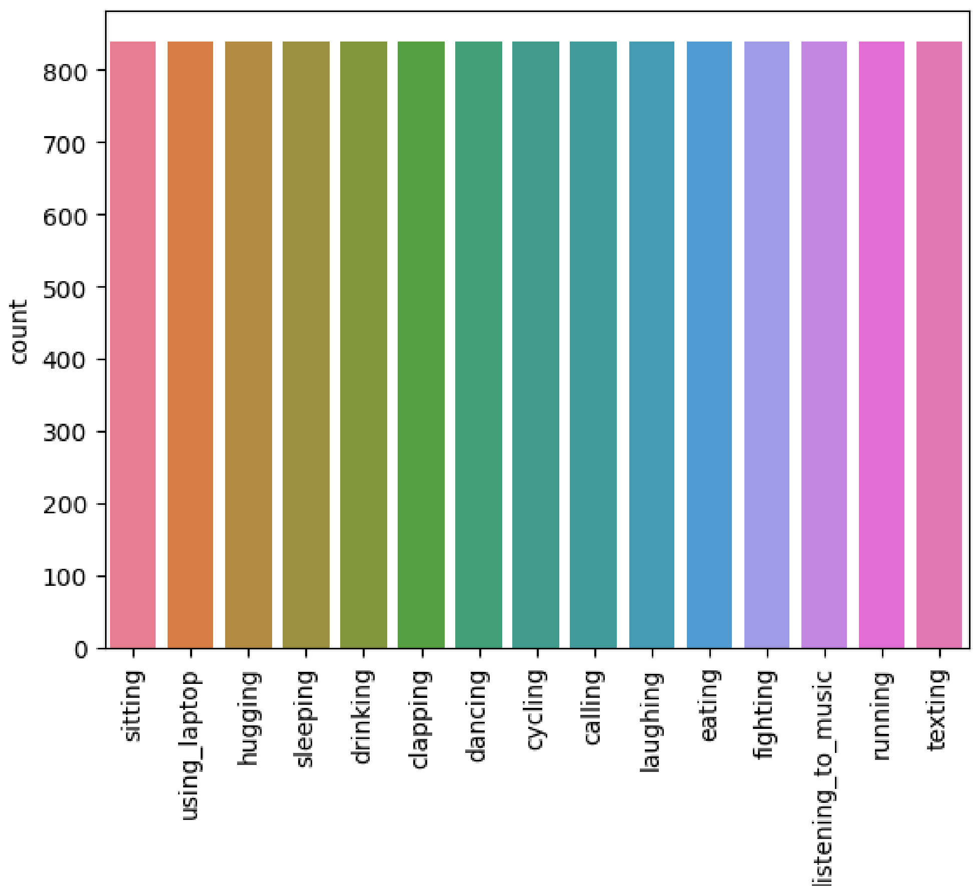
$$= \frac{0.5 \times 0.5 \times 0.5}{0.5 \times 0.5 \times 0.5 + 0.5 \times 0.5 \times 0.25}$$

$$= \frac{0.5}{0.5 + 0.25} = \frac{2}{3}$$

We choose & analyze that mail is not spam.

SECTION C

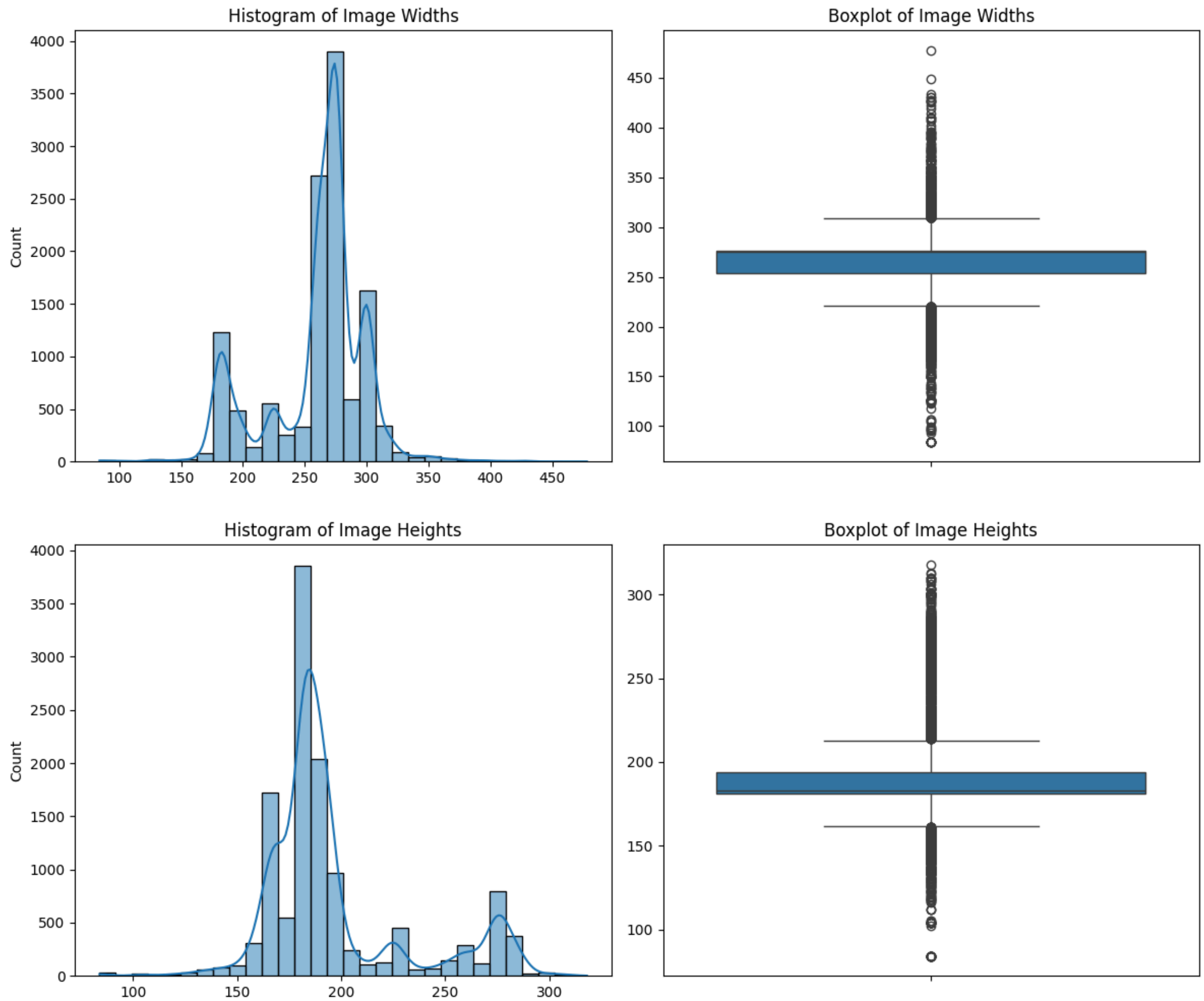
Q1.a)




**Distribution:-**

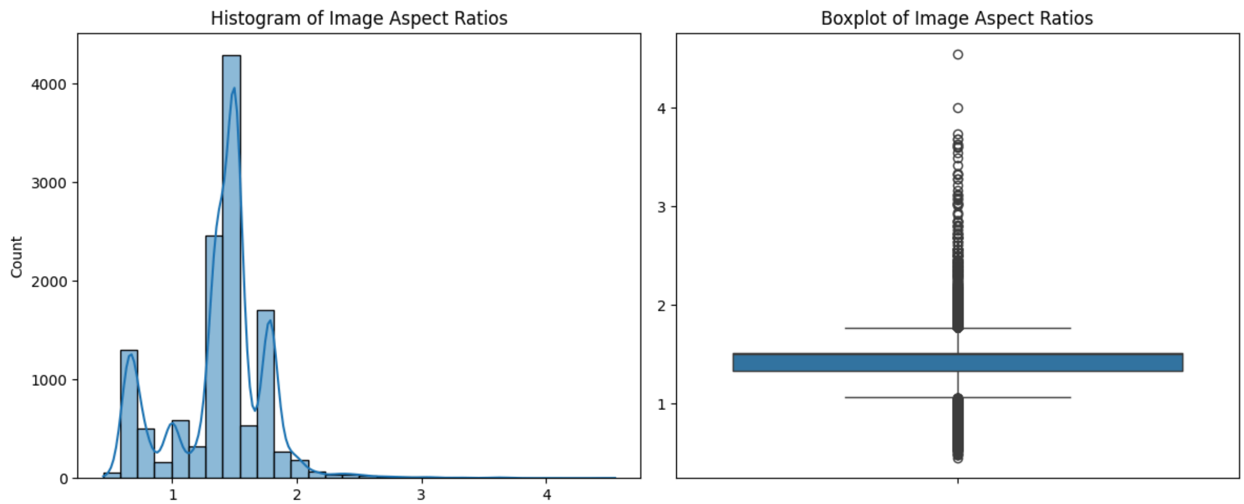
```
sitting               840
using_laptop          840
hugging               840
sleeping              840
drinking              840
clapping              840
dancing               840
cycling               840
calling               840
laughing              840
eating                840
fighting              840
listening_to_music    840
running               840
texting               840
```

The dataset is completely balanced in terms of examples from every label. This means that there won't be any bias in prediction of samples because of a larger/lesser number of datapoints from one class to another.



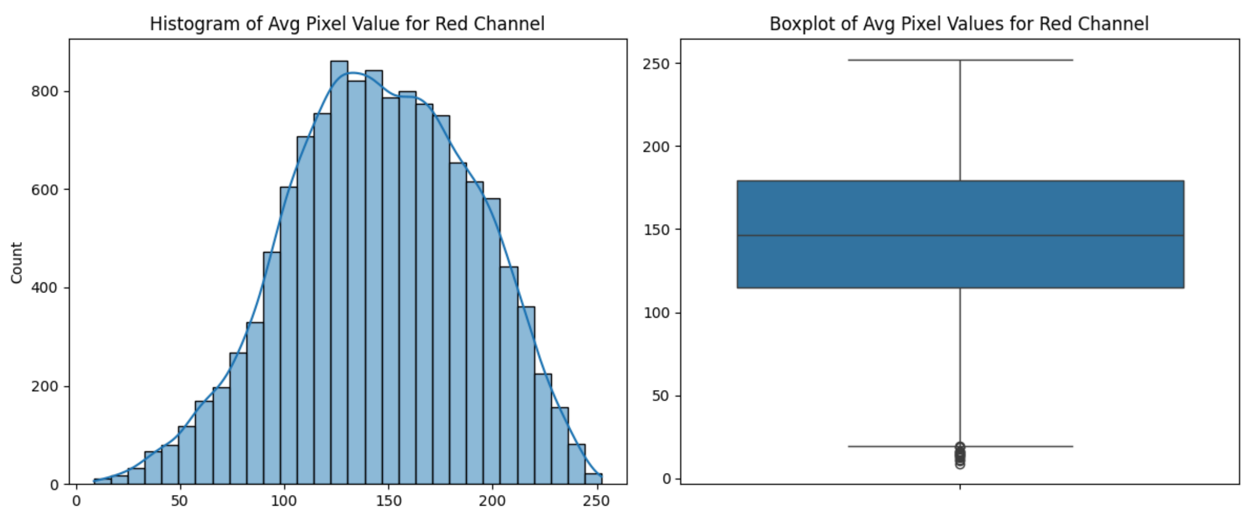

```
Mean of Image Heights -> 196.57439885723355 pixels
Variance of Image Heights -> 1244.5884057671092

Mean of Image Widths -> 260.379652408539 pixels
Variance of Image Widths -> 1593.3200335590473
```

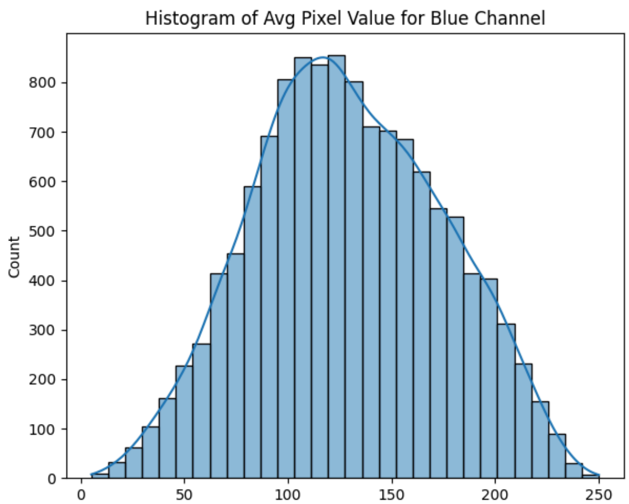
Clearly, most of the images are concentrated around having heights around 190 and widths around 270. There are some other heights and widths that are common, as shown by the spikes in the histograms.

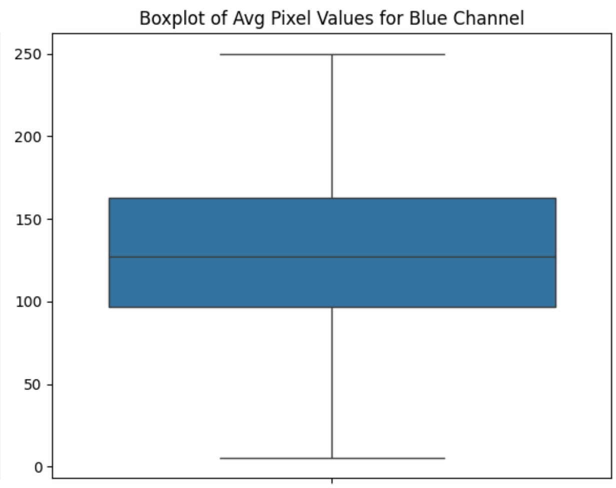


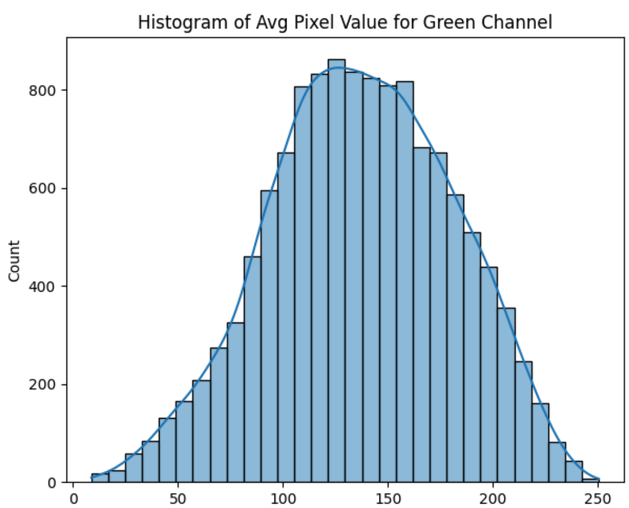

```
Mean of Image Aspect Ratios -> 1.3885801993320368
Variance of Image Aspect Ratios -> 0.14817010562549524
```

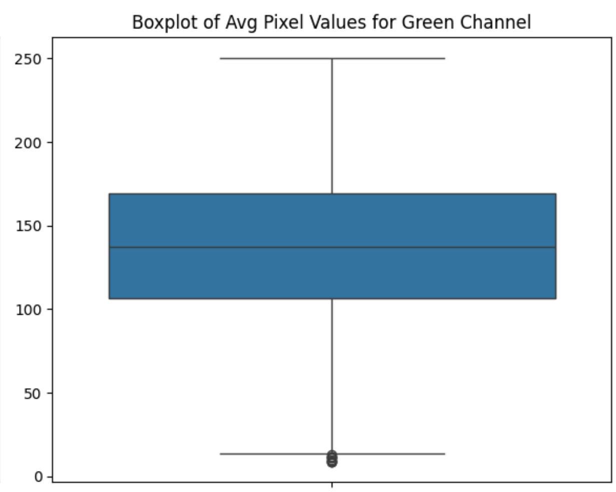
```
Significantly less variance than the heights and widths indicating that
the proportion of the corresponding heights and weights increase by the
same factor, relatively speaking.
```





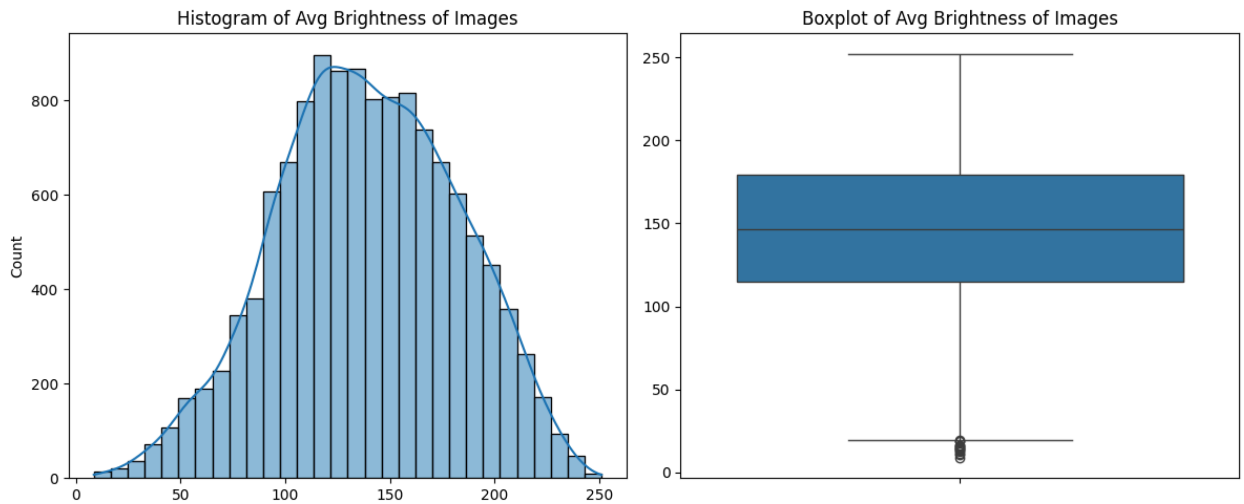




```
Variance of Means of Avg Pixel Values for Red Channel ->
1948.0849660626825

Variance of Means of Avg Pixel Values for Blue Channel ->
2084.3053194745257

Variance of Means of Avg Pixel Values for Green Channel ->
1912.8074663393675
```

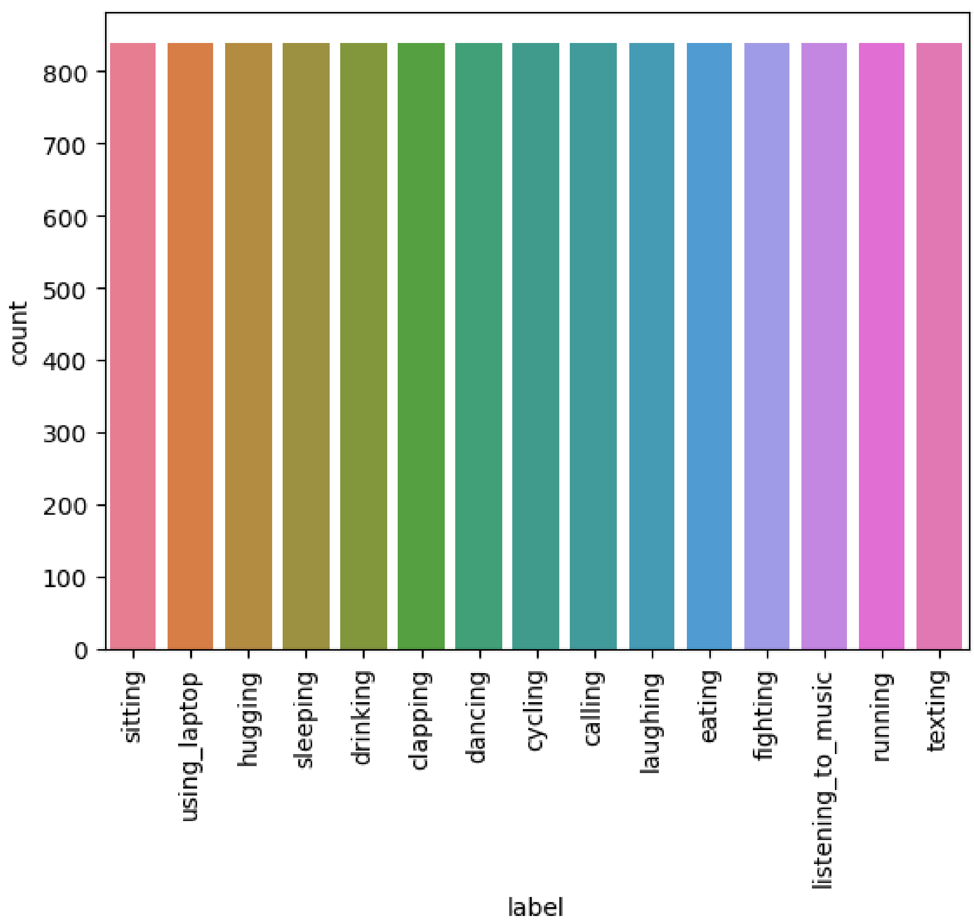
```
There is just a slighly higher amounts of red, then green and finally blue
in that order of avg pixel value for that channel.
```




Variance of Means of Avg Brightness of Images -> 1865.1596923745687

A lot of Variation in avg brightness in the images. It is almost averaging out into a standard normal Gaussian distribution.

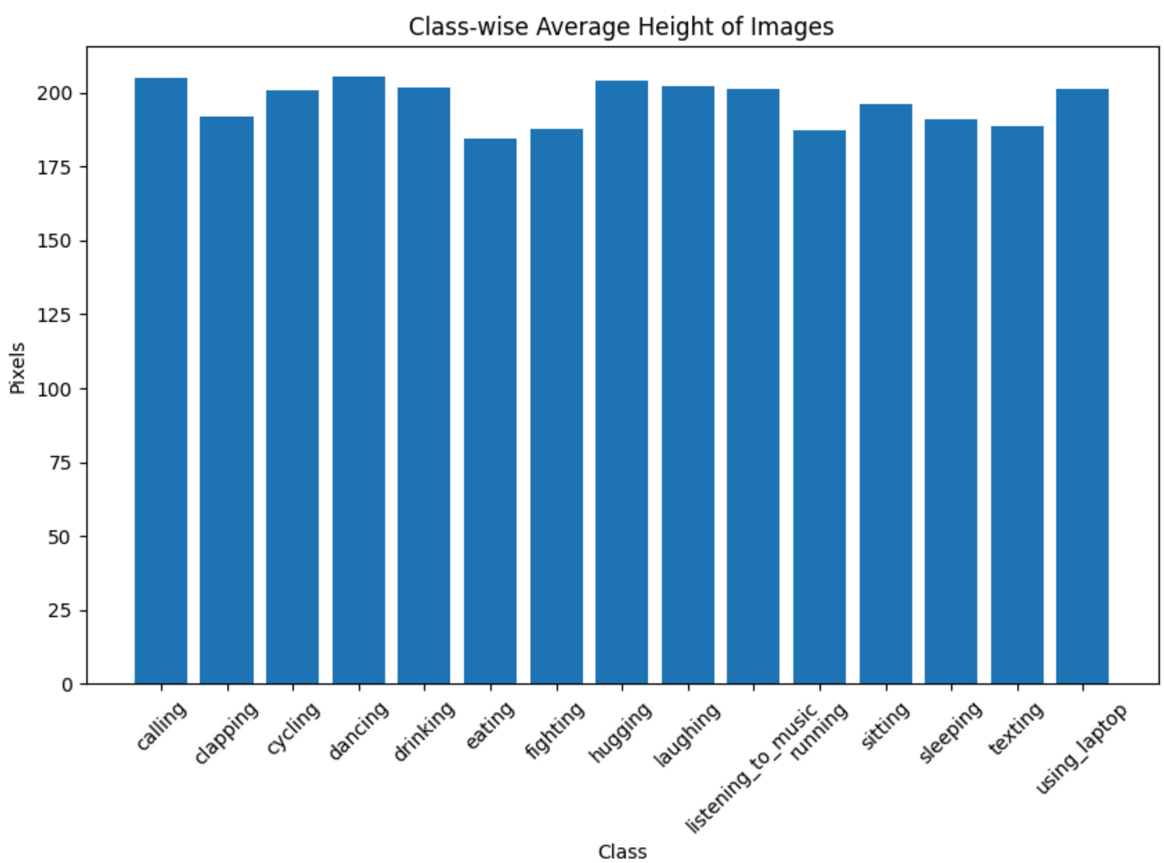
Q1.(b)

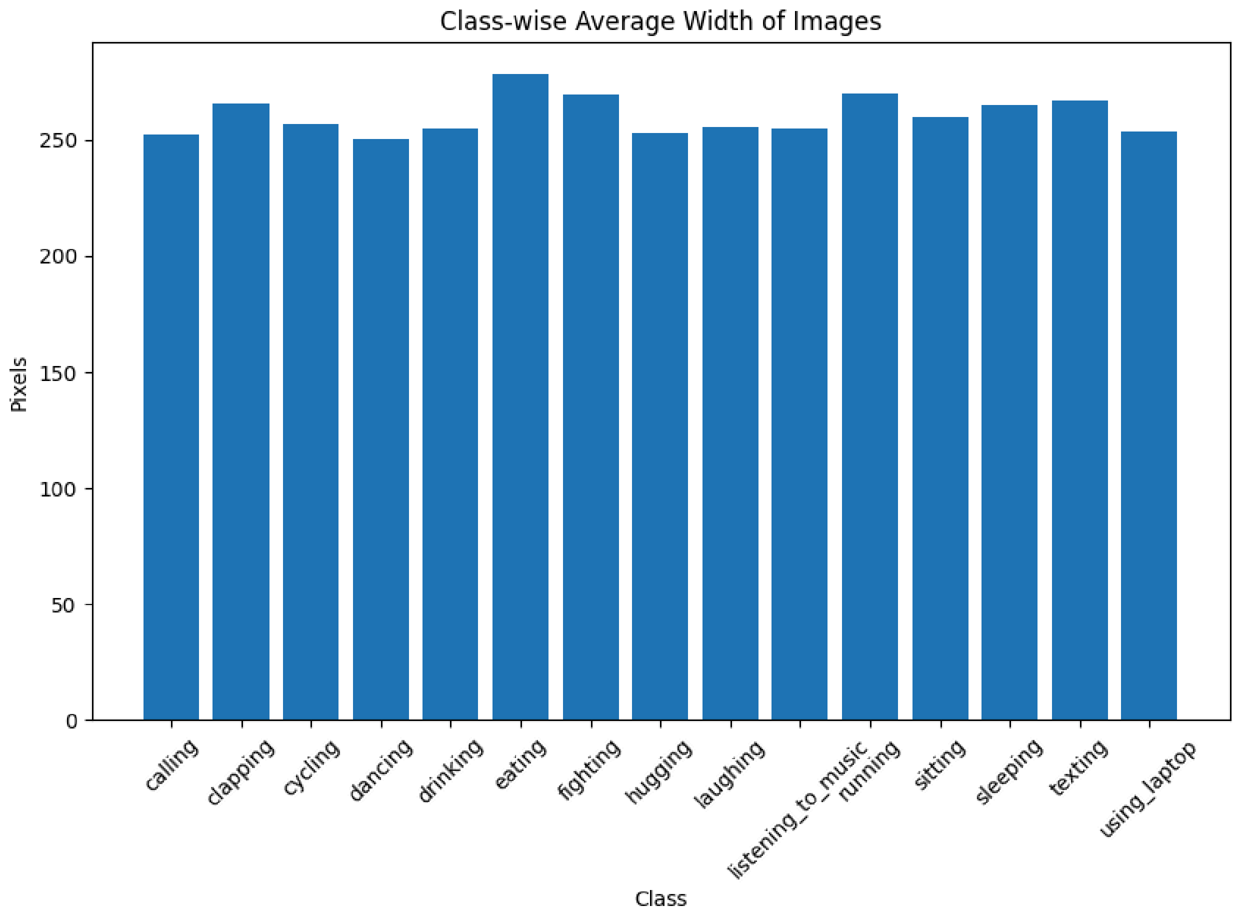


**Distribution:-**

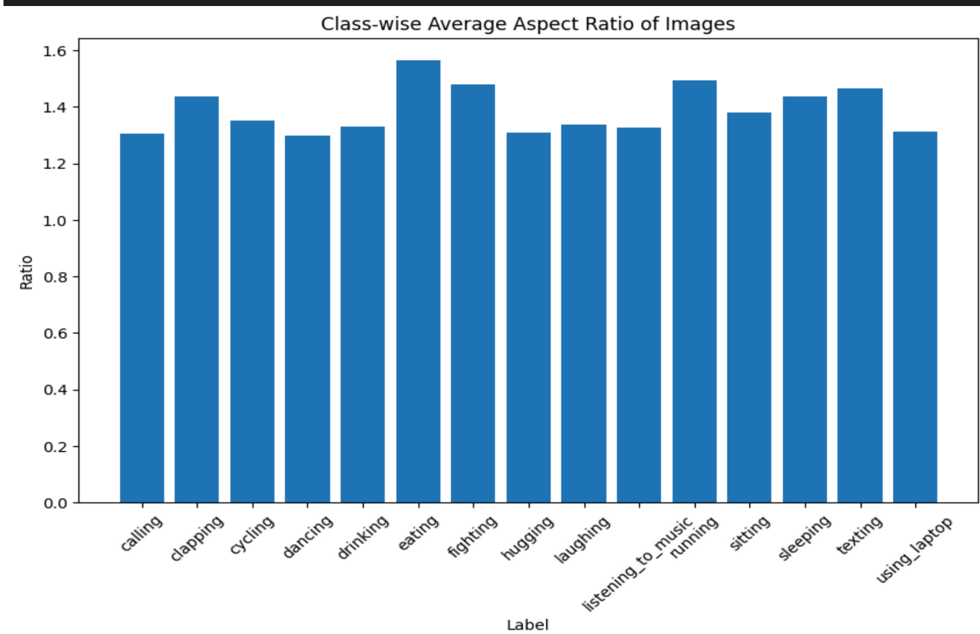
```
sitting                 840
using_laptop            840
hugging                 840
sleeping                840
drinking                840
clapping                840
dancing                 840
cycling                 840
calling                 840
laughing                840
eating                  840
fighting                840
listening_to_music      840
running                 840
texting                 840
```

The dataset is completely balanced in terms of examples from every label. This means that there won't be any bias in prediction of samples because of a larger/lesser number of datapoints from one class to another.

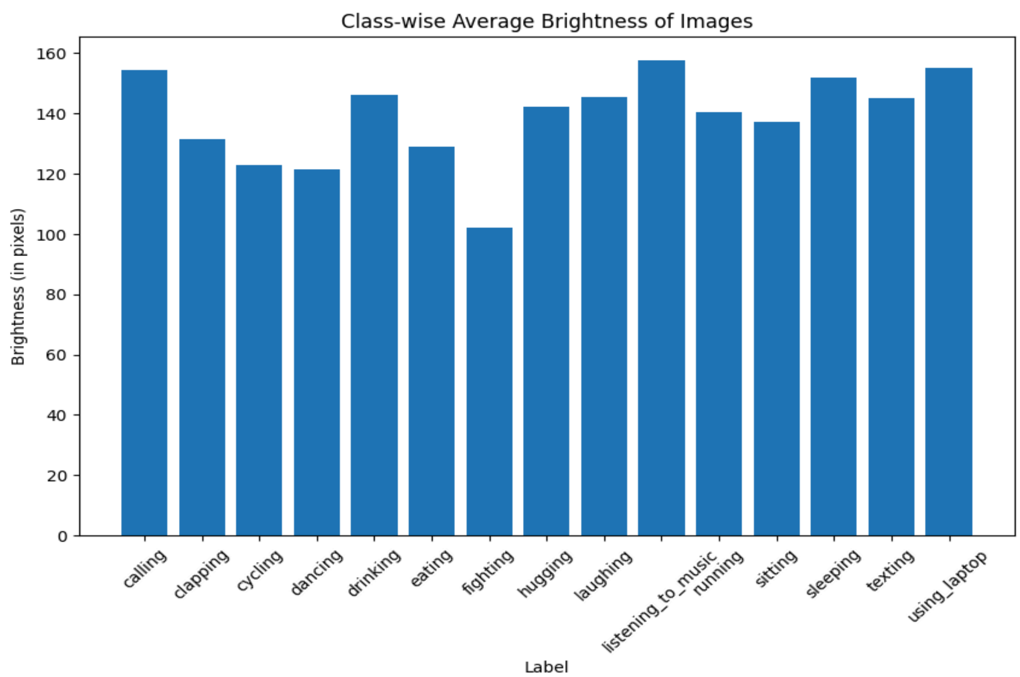





```
'Eating' has the highest Width but lowest Height, so clearly these images
are more in the landscape mode. 'Calling' has the highest Height but
lowest Width.
```



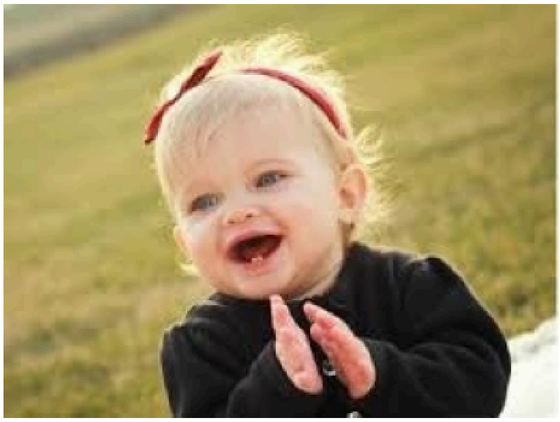
The higher the aspect ratio, the wider but shorter it is. Clearly 'eating' videos are like that, as are 'listening_to_music' videos. On the contrary, the images which are taller but have less width, so like that in portrait mode, belong to classes like 'dancing', 'hugging'.

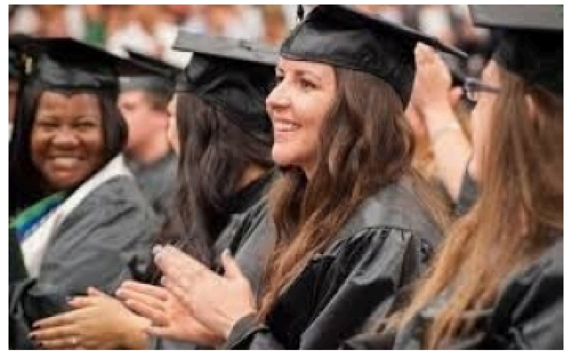



Fighting images have a much lower average brightness while 'listening_to_music', 'calling' and 'using_laptop' have much higher average brightnesses.

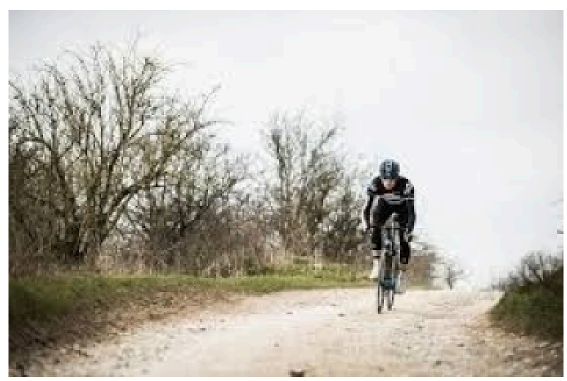
Label: calling

Label: clapping





Label: cycling

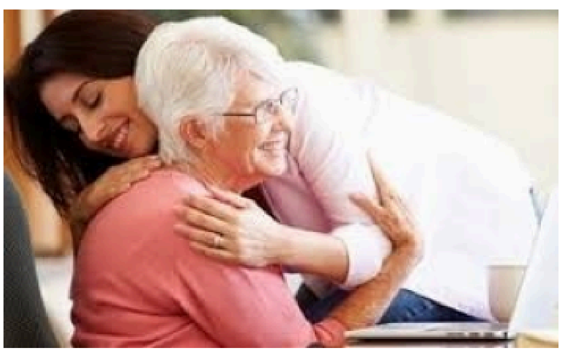
Label: dancing

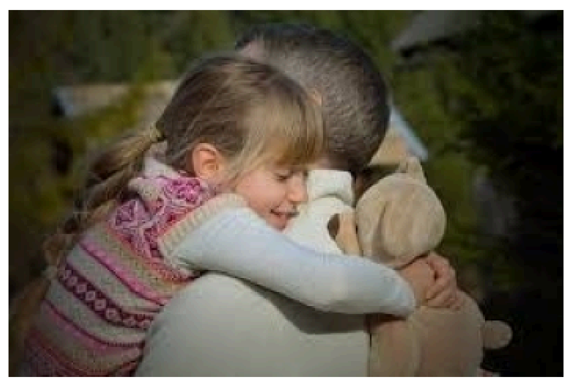
Label: drinking


Label: eating

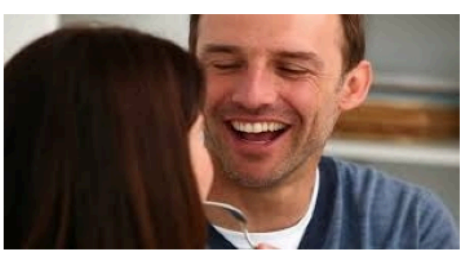
Label: fighting

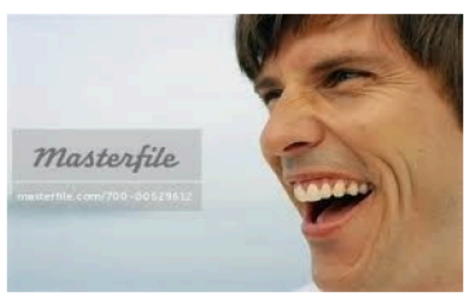
Label: hugging



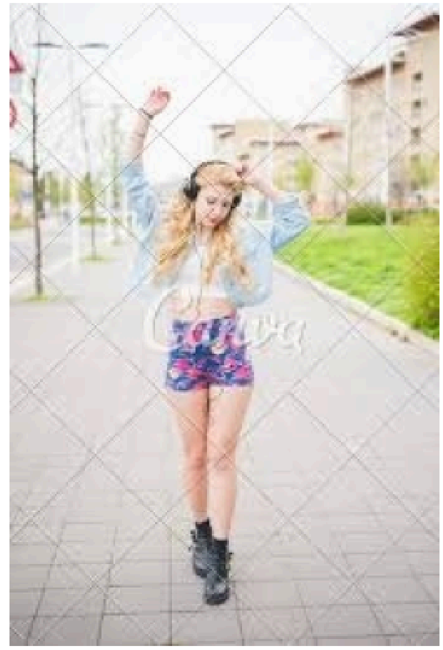


Label: laughing

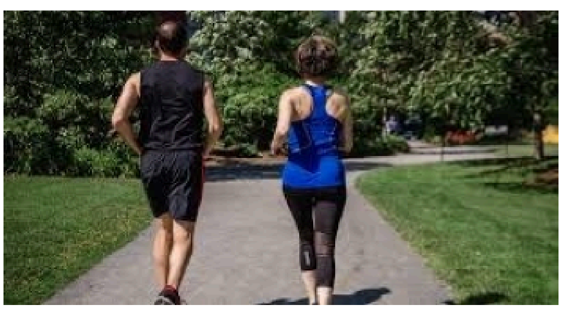




Label: listening_to_music



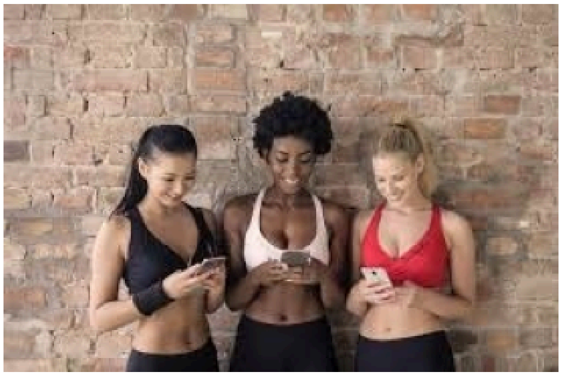
Label: running

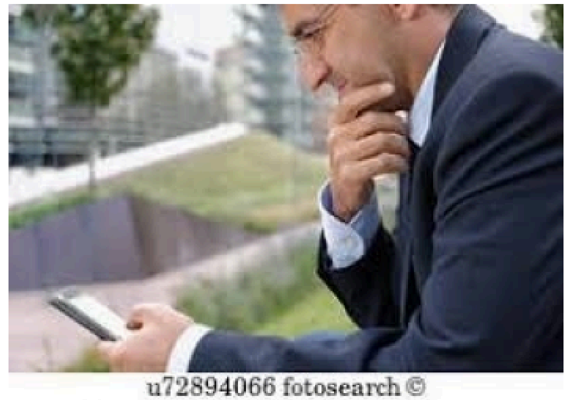


Label: sitting

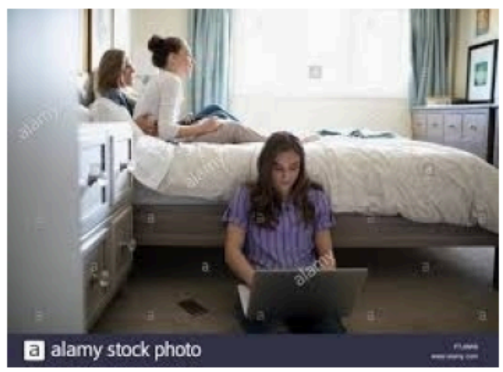

Label: sleeping

Label: texting






Label: using_laptop




**Observations:-** Most of these images have some objects that can identify the task. For example, a laptop or a headphone or even a gun. Some, like hugging always involve at least 2 people. Similarly, some other activities are activities that only involve the individual so, one can narrow done the class using that.

1.c) There is no class imbalance, no need of re-sampling methods. In terms of augmentation, perhaps we could stabilize the brightness and size of all the images. However, if we made the brightness of each image same, the images which are naturally dark would become brighter (like night scenes). So, this isn't wanted, and could make us lose information. So, I shall only make all their sizes equal.

2. Features like HoGs have extremely high complexity especially with increase in parameters like no. of cells covered, orientation and other parameters. So, PCA or other dimensionality reduction could

be required. We also observe that we should take the color data separately and all the other features that don't depend on colour should be tested on the grayscale version.

I also look at LBP and hu moments for features which could give me the temporal information of the objects that could classify the images to the right class.

Q3.a and b)
Unoptimized Naive Bayes :- 17.9%
Optimized Gaussian Naive Bayes :- 22.7%
Parameters:- **`'var_smoothing': 0`**

Unptimized Perceptron :- 10.83%
Optimized :- 11.15%
Parameters:- **`{'alpha': 0.0001, 'n_jobs': -1, 'penalty': 'l1'}`**

Unoptimized Random Forests :- 26.07%
Optimized :- 32.3%
Parameters:- **`{'bootstrap': False, 'ccp_alpha': 0.0001, 'criterion': 'gini', 'max_depth': None, 'max_features': 'sqrt', 'n_estimators': 30}`**

Unoptimized Gradient Boosting :- 12.06%

I optimized the algorithms using gridsearch, and tried out a variety of things including reducing the number of features, increasing the number of iterations/trees, modifying the coefficients(like alpha, l1 ratio, etc.) and optimized each model. Then, I tried out stuff like PCA to see if the results weren't that bad so that I could go for a computationally more intensive model, because with 5000 features, it will not give great results. Finally, I tried taking an ensemble of models tried out in the past, since each are independently created. However, their accuracy weren't great.

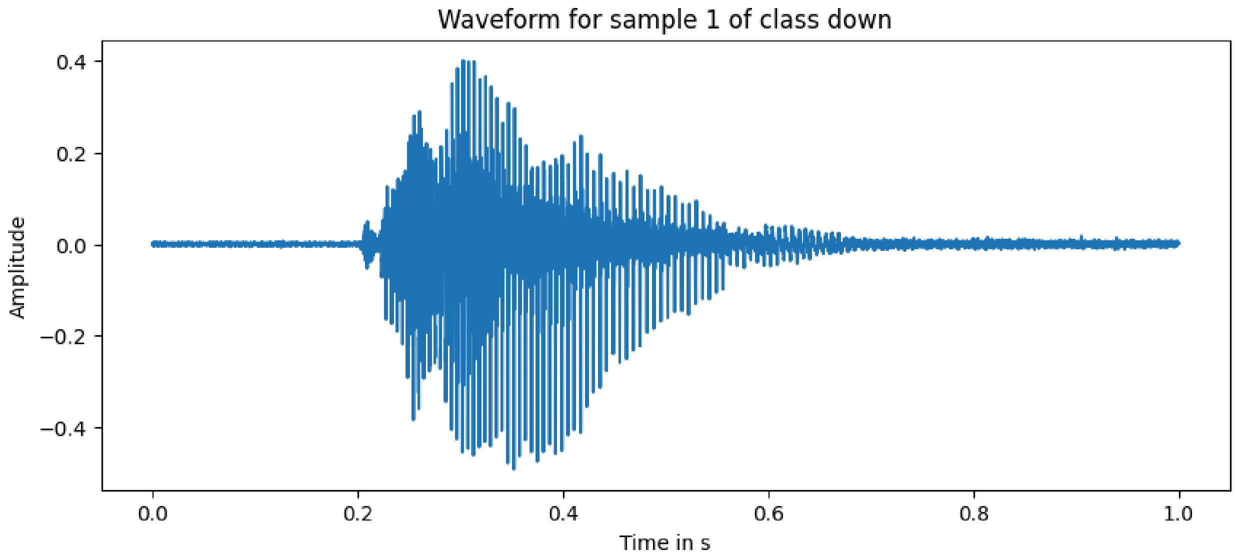
SECTION B

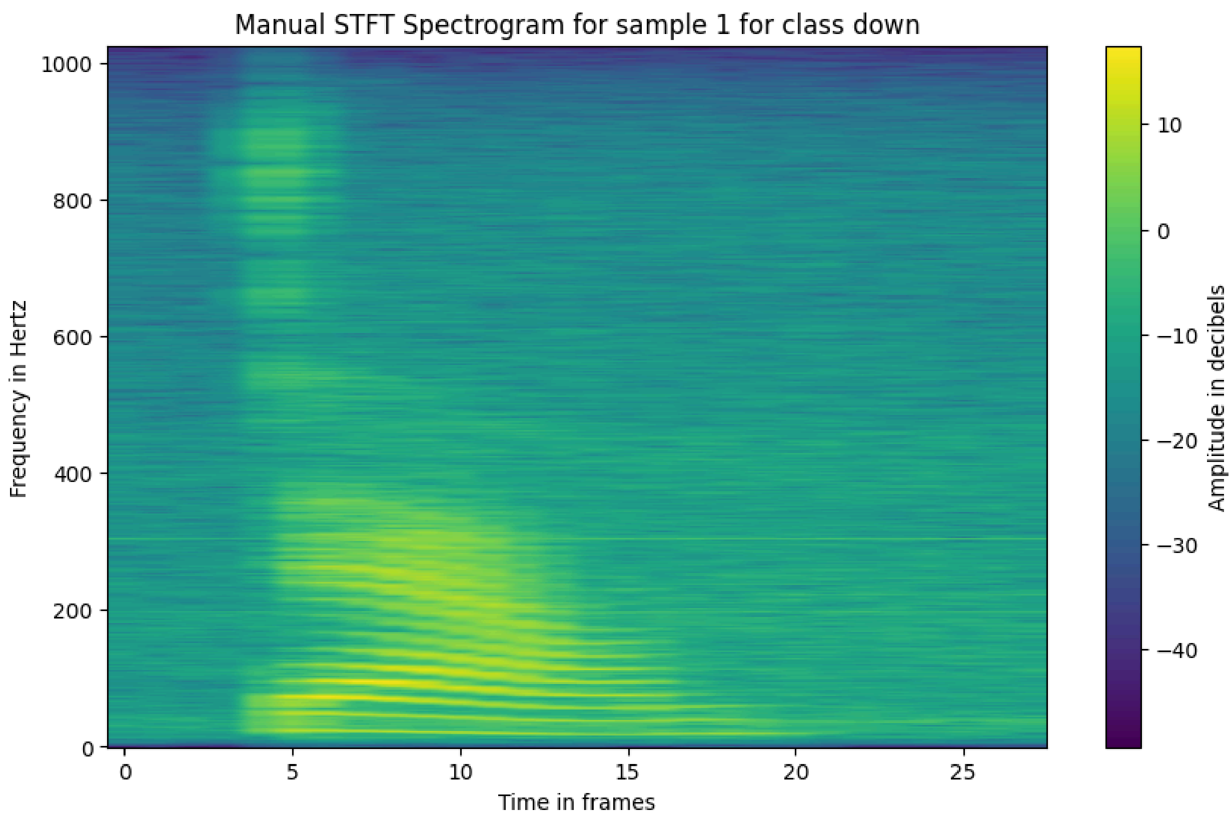
Q1.a) Plots are there in the notebook.

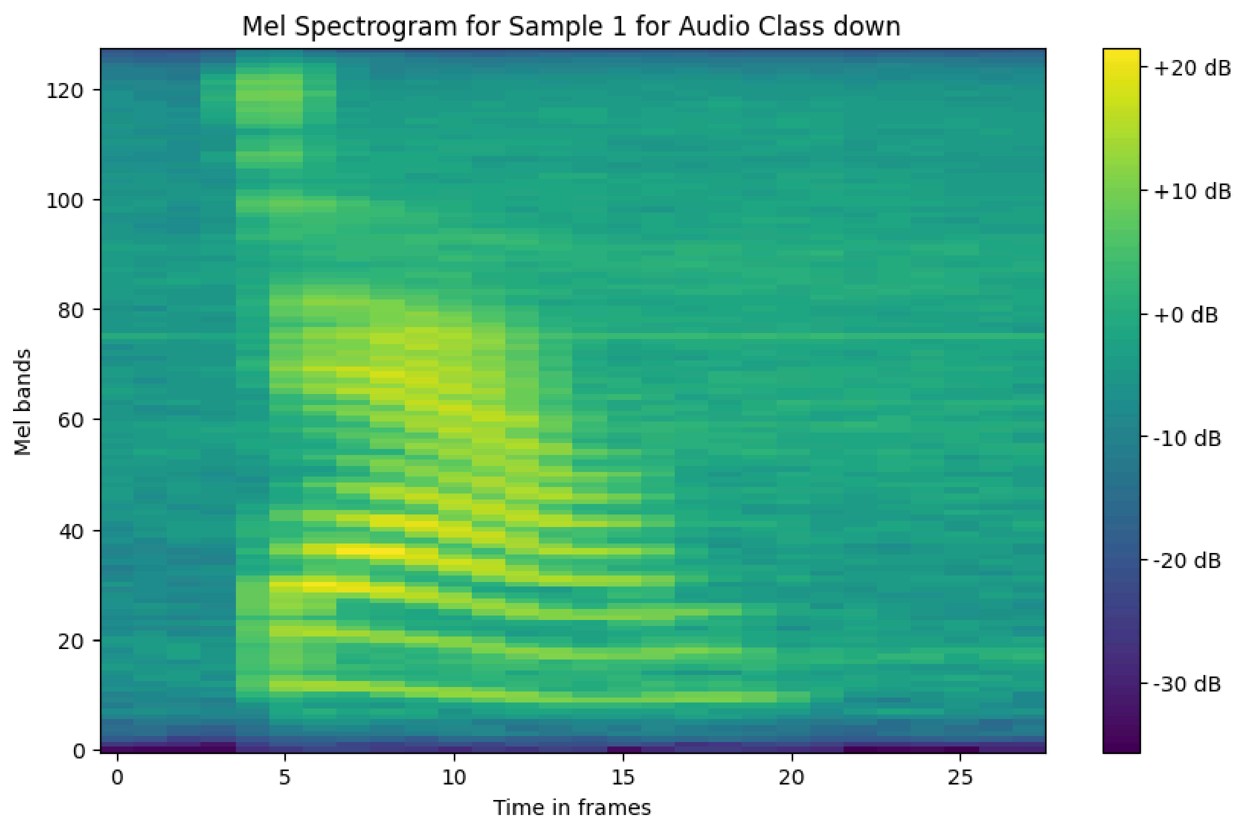
Observations:- Most audio classes have a mean amplitude around 0, and their durations are around 1s. The standard deviation for most of the classes is very small, meaning not a lot of variation.

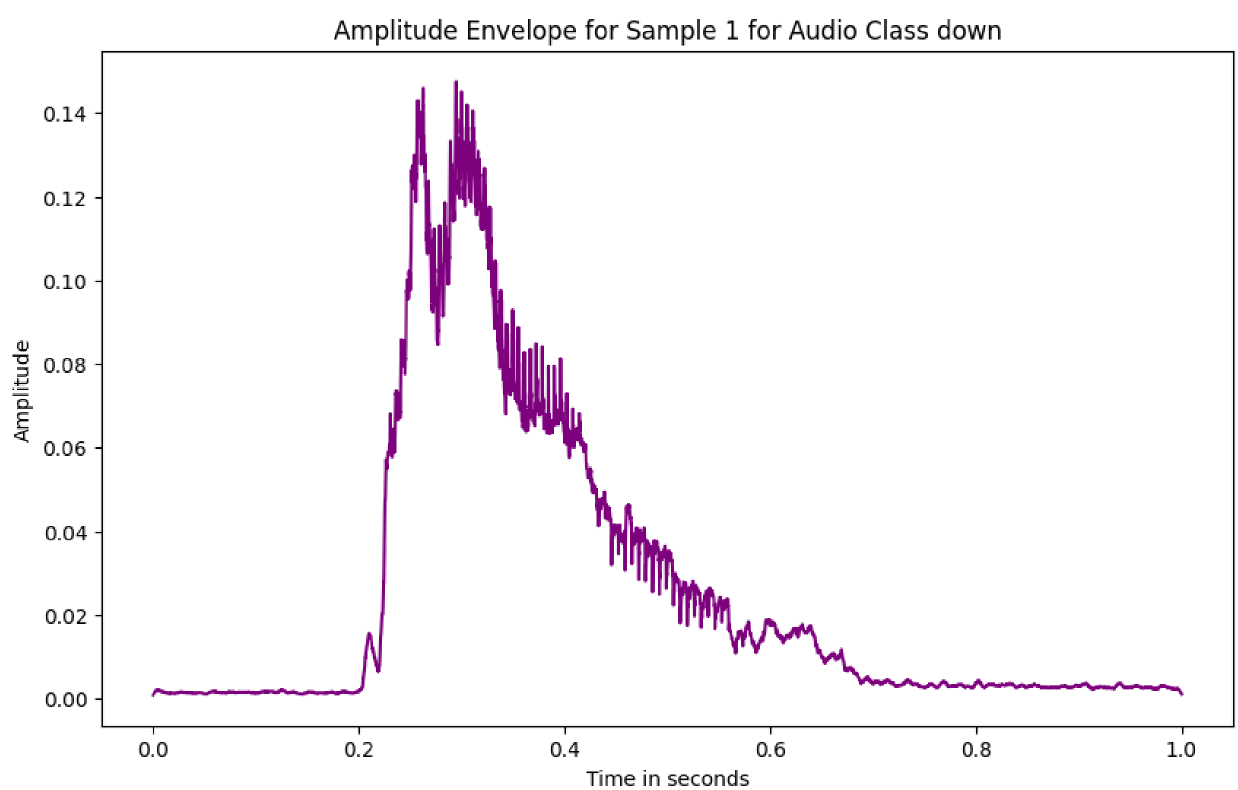
Some classes like forward have an amplitude that is tending towards greater than 0, whereas some like eight tend below 0. This could be due to the pitch and tone of the voice while saying these words, hence, the change. This naturally makes sense.

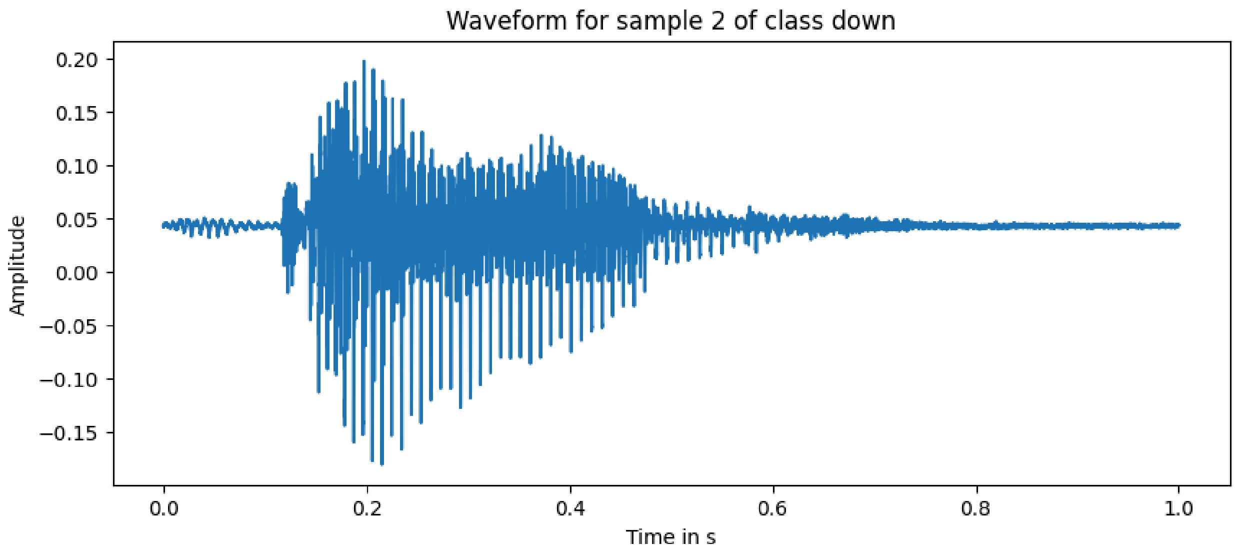
(b)

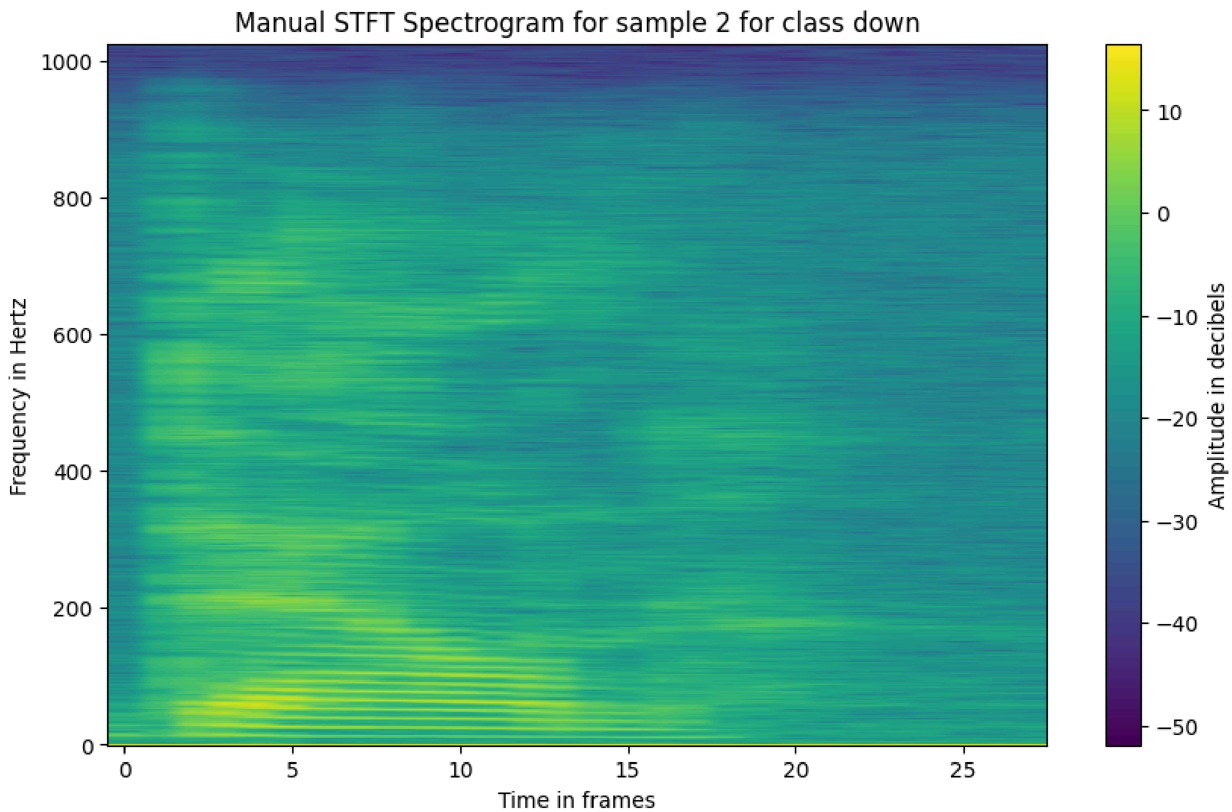

Waveform for sample 1 of class down
Amplitude
Time in s
0.4
0.2
0.0
−0.2
−0.4
0.0
0.2
0.4
0.6
0.8
1.0

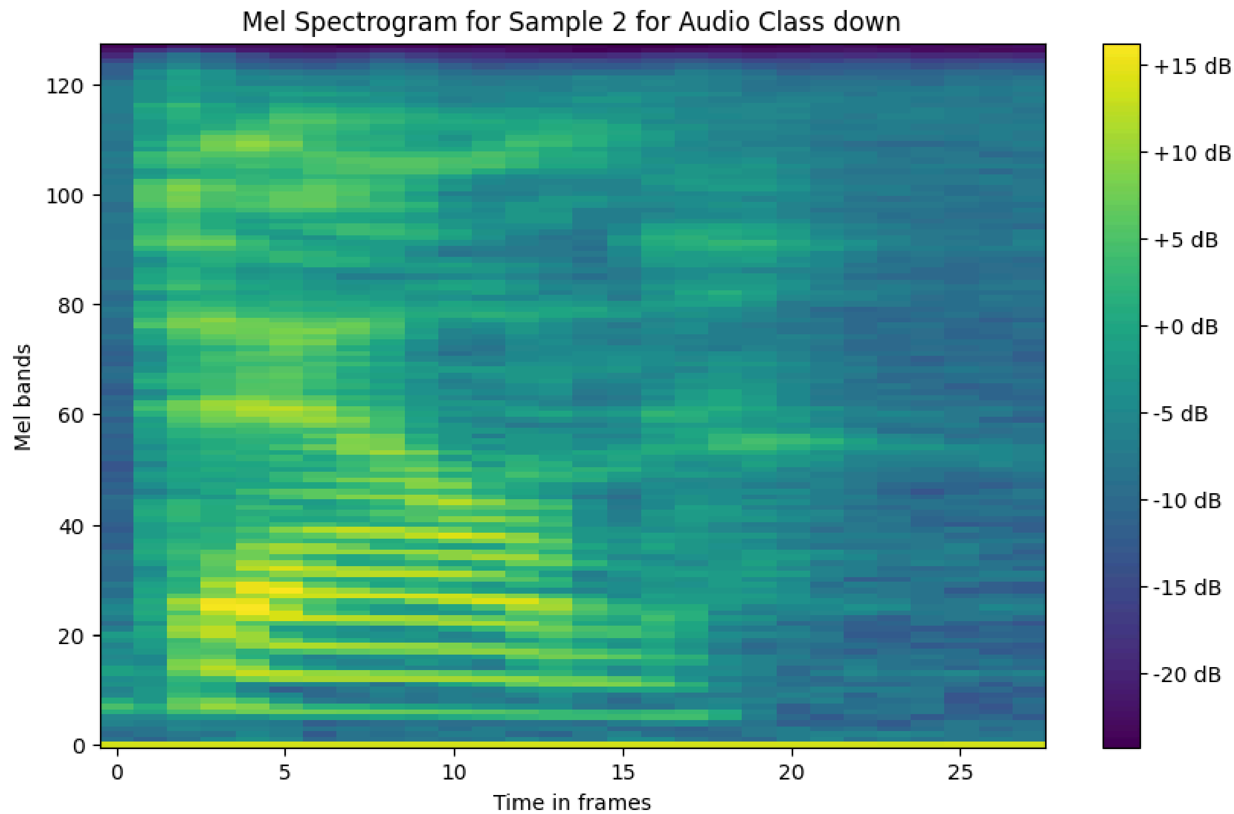

Manual STFT Spectrogram for sample 1 for class down
Frequency in Hertz
Time in frames
Amplitude in decibels
1000
800
600
400
200
0
5
10
15
20
25
−10
−20
−30
−40

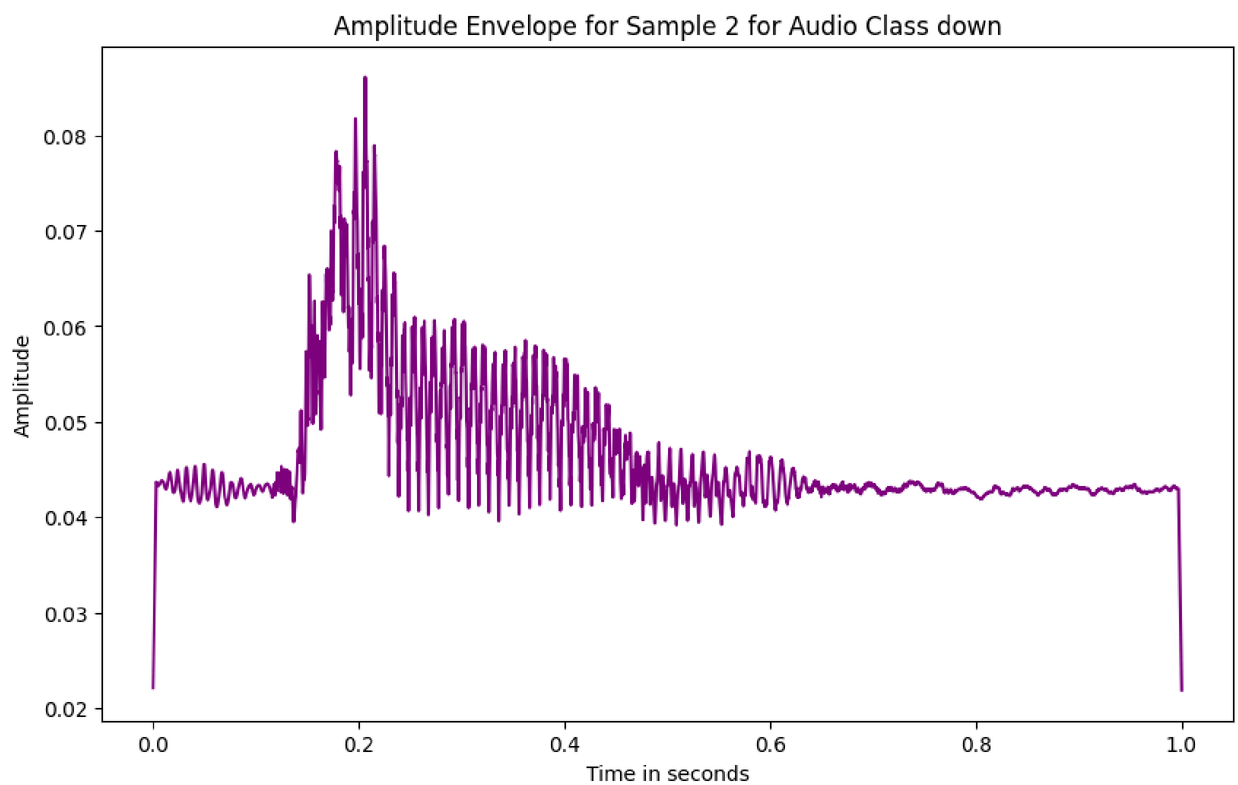

Mel Spectrogram for Sample 1 for Audio Class down
Mel bands
Time in frames
+20 dB
+10 dB
+0 dB
-10 dB
-20 dB
-30 dB

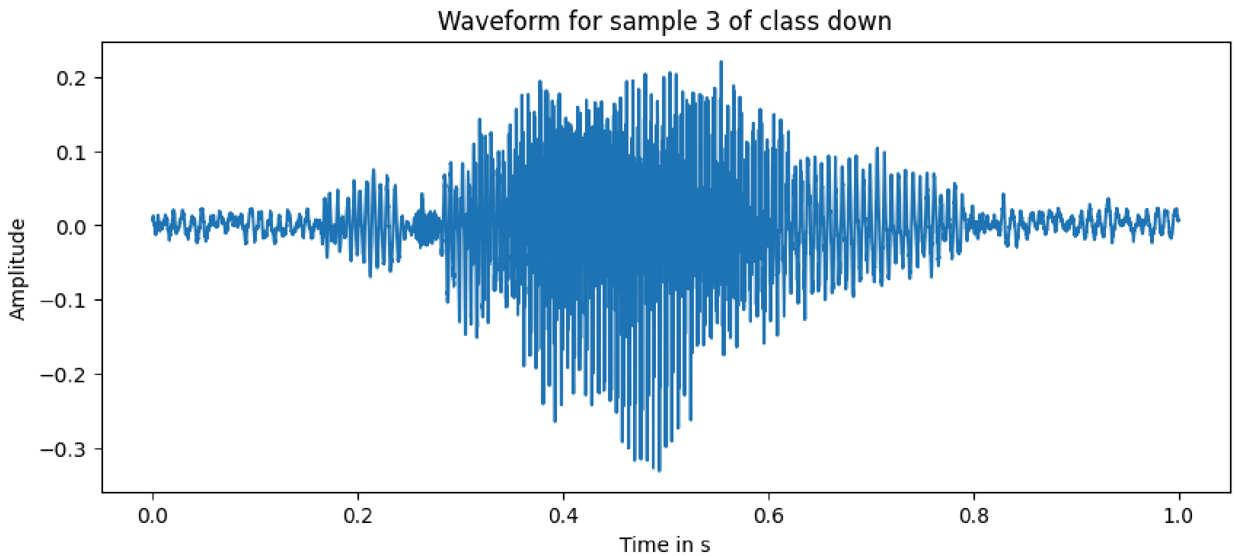

Amplitude Envelope for Sample 1 for Audio Class down
Amplitude
Time in seconds

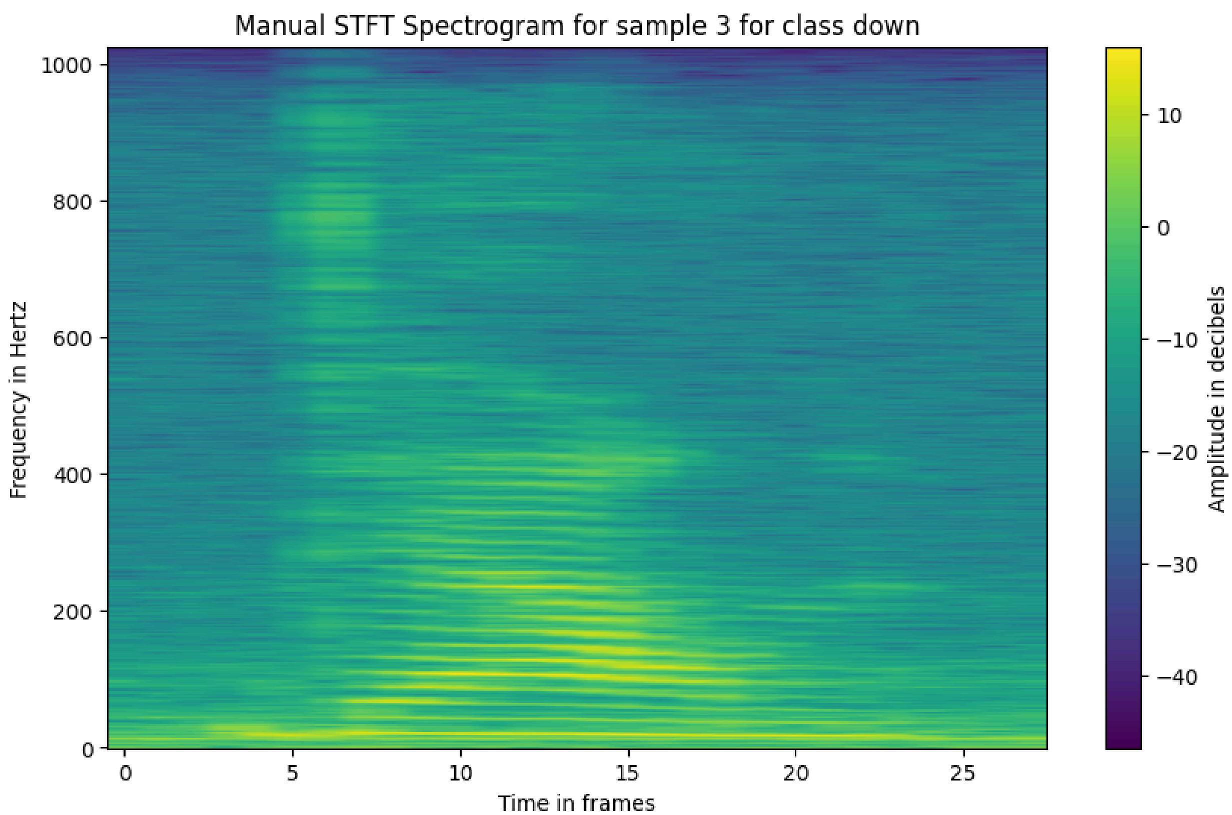

Waveform for sample 2 of class down
Amplitude
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
0.0
0.2
0.4
0.6
0.8
1.0
Time in s

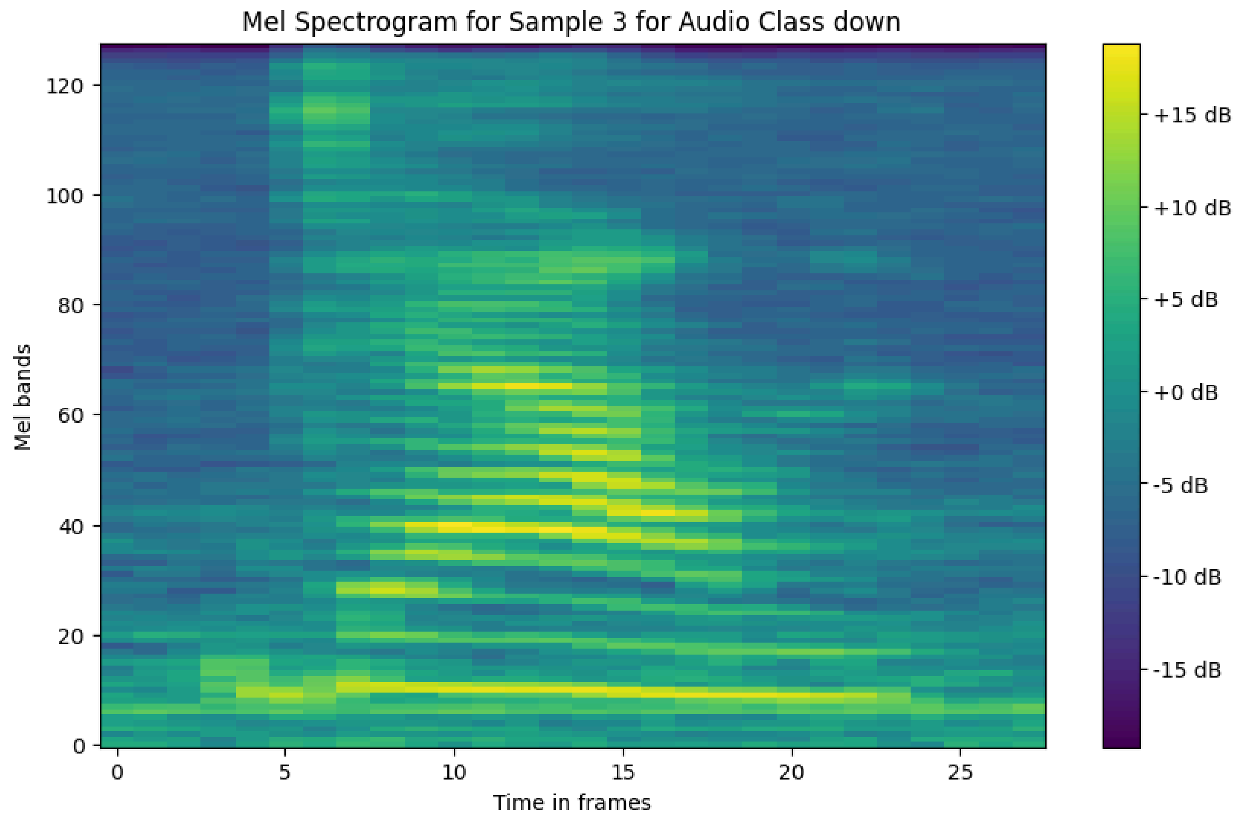

Manual STFT Spectrogram for sample 2 for class down
Frequency in Hertz
1000
800
600
400
200
0
0
5
10
15
20
25
Time in frames
Amplitude in decibels
10
0
−10
−20
−30
−40
−50

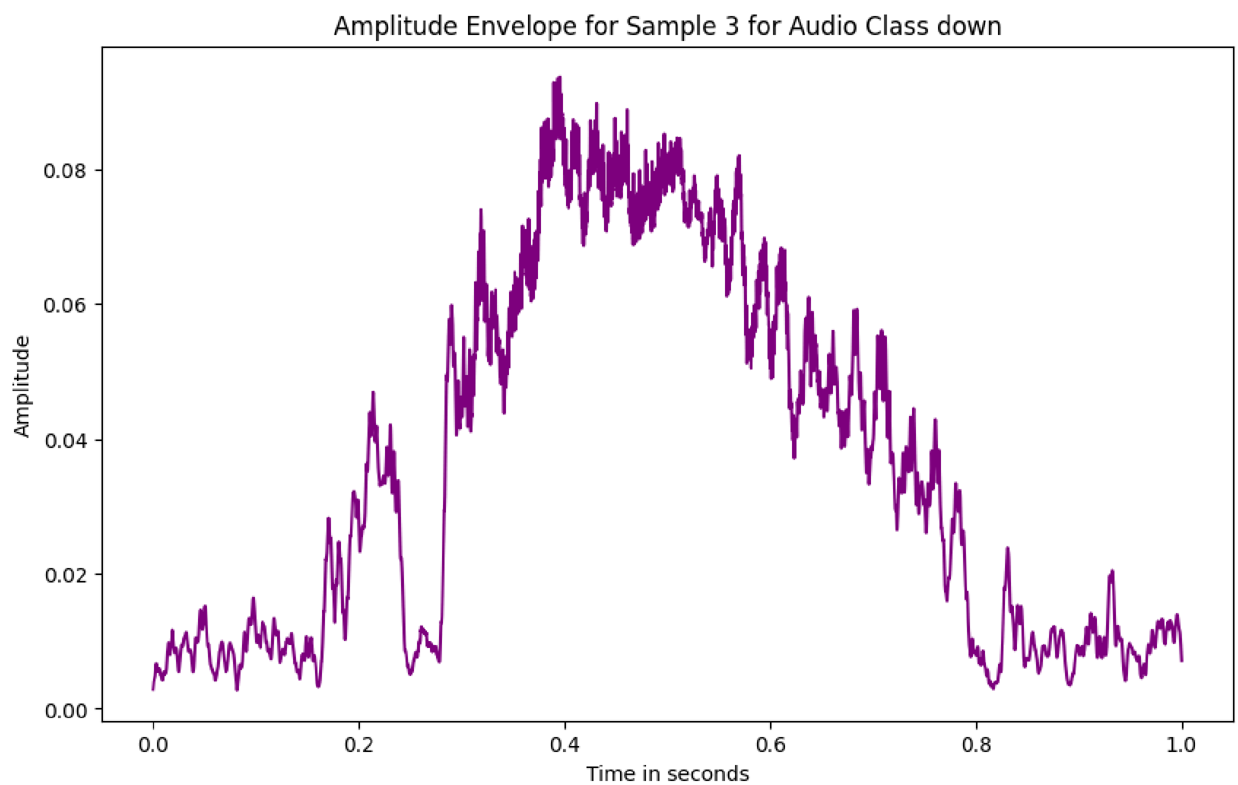

Mel Spectrogram for Sample 2 for Audio Class down
Mel bands
Time in frames
+15 dB
+10 dB
+5 dB
+0 dB
-5 dB
-10 dB
-15 dB
-20 dB

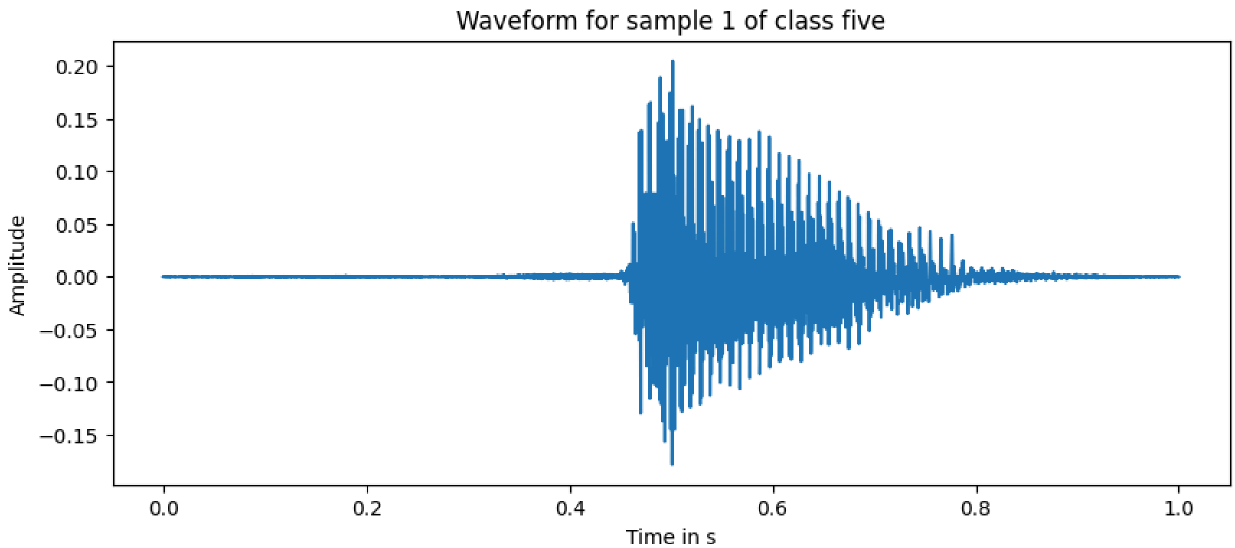

Amplitude Envelope for Sample 2 for Audio Class down
Amplitude
Time in seconds

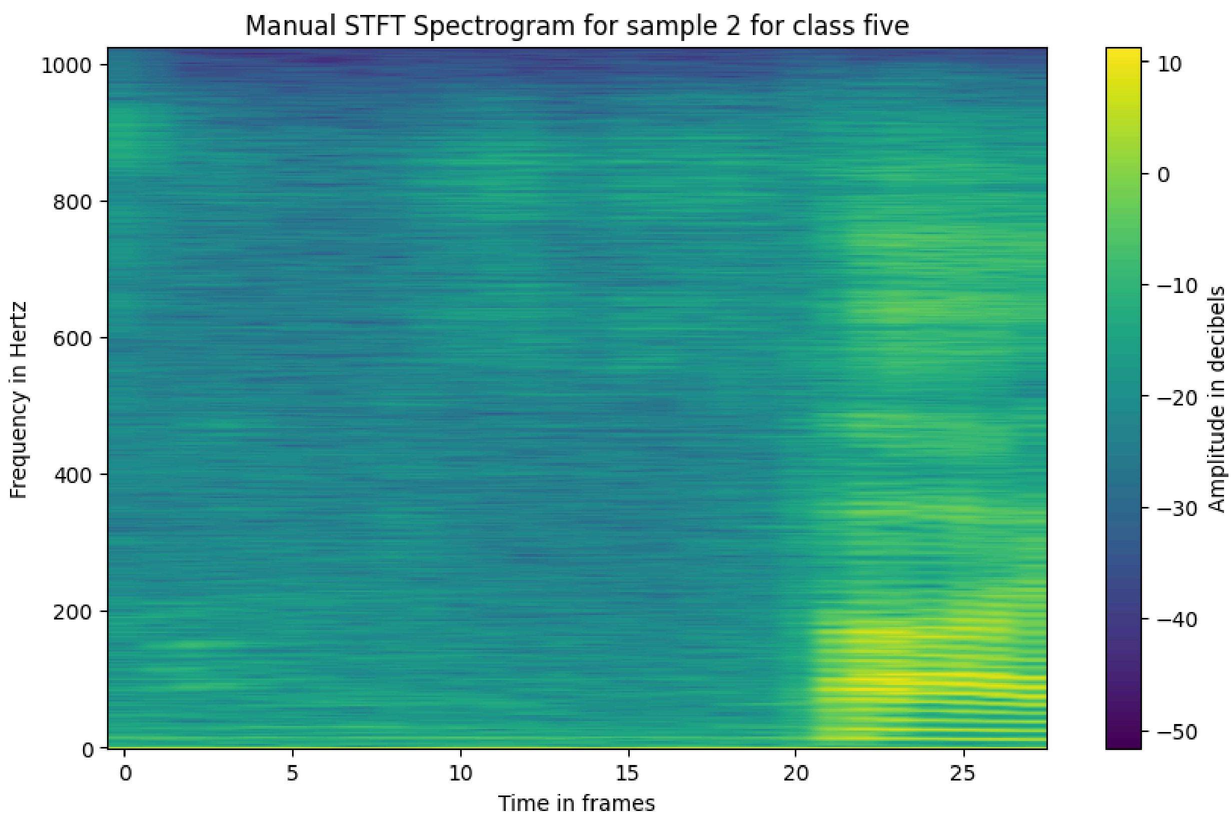

Waveform for sample 3 of class down
Amplitude
Time in s
0.2
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.2
0.4
0.6
0.8
1.0

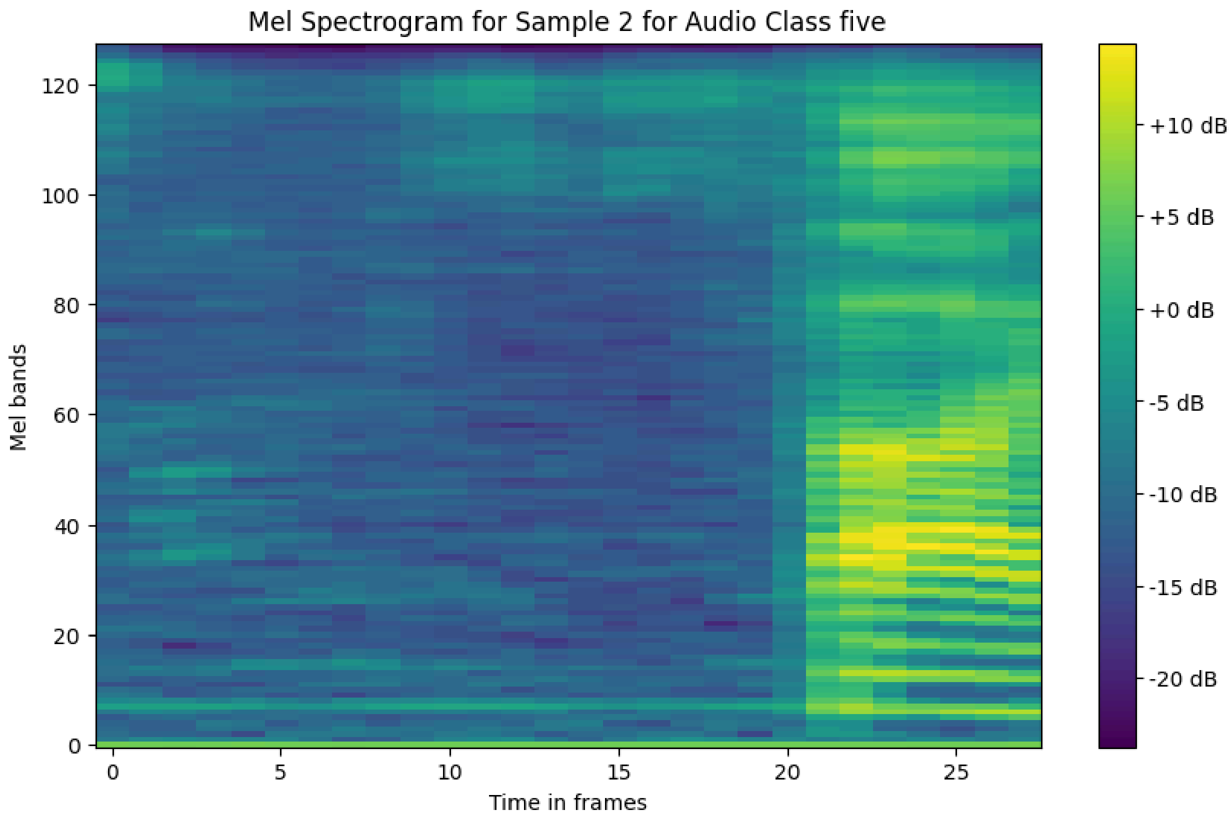

Manual STFT Spectrogram for sample 3 for class down
Frequency in Hertz
Time in frames
Amplitude in decibels
1000
800
600
400
200
0
5
10
15
20
25
−10
−20
−30
−40


Mel Spectrogram for Sample 3 for Audio Class down
Mel bands
Time in frames
0
20
40
60
80
100
120
0
5
10
15
20
25
+15 dB
+10 dB
+5 dB
+0 dB
-5 dB
-10 dB
-15 dB

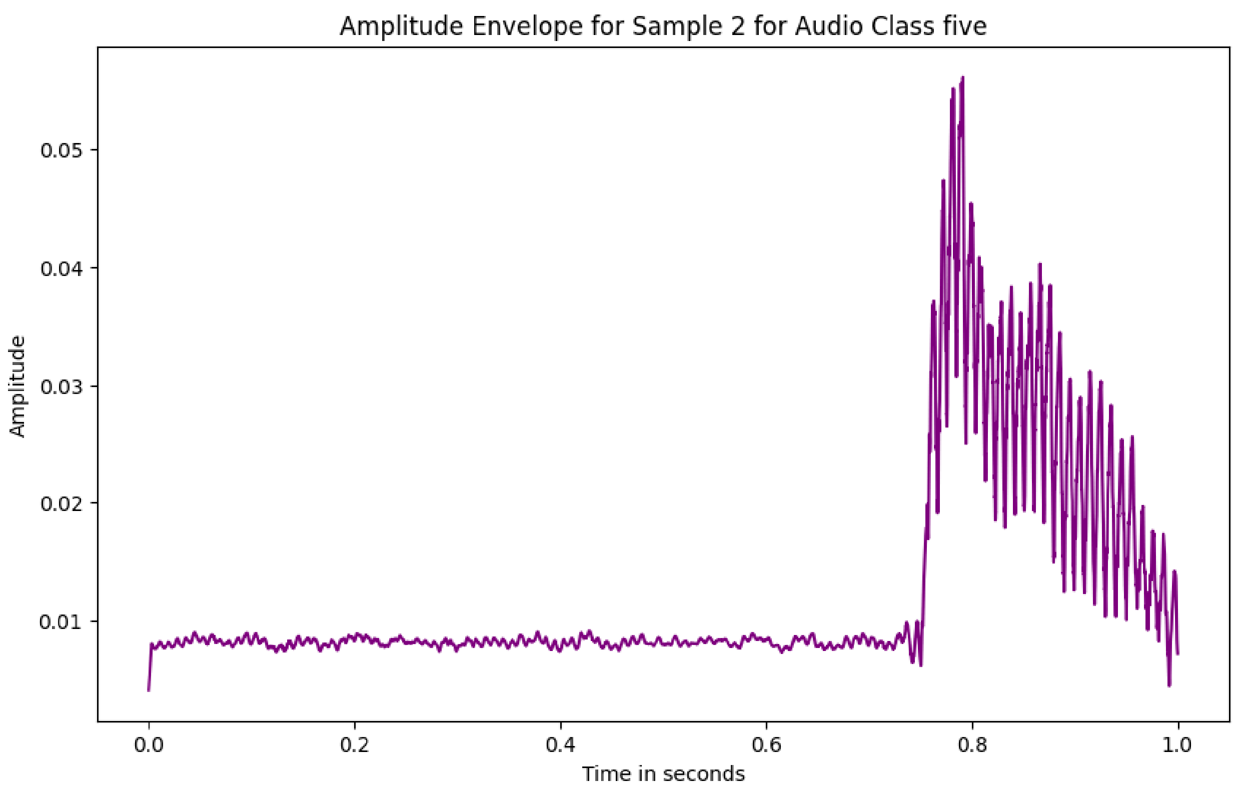

Amplitude Envelope for Sample 3 for Audio Class down
Amplitude
Time in seconds
0.00
0.02
0.04
0.06
0.08
0.0
0.2
0.4
0.6
0.8
1.0


Waveform for sample 1 of class five
Amplitude
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
0.0
0.2
0.4
0.6
0.8
1.0
Time in s

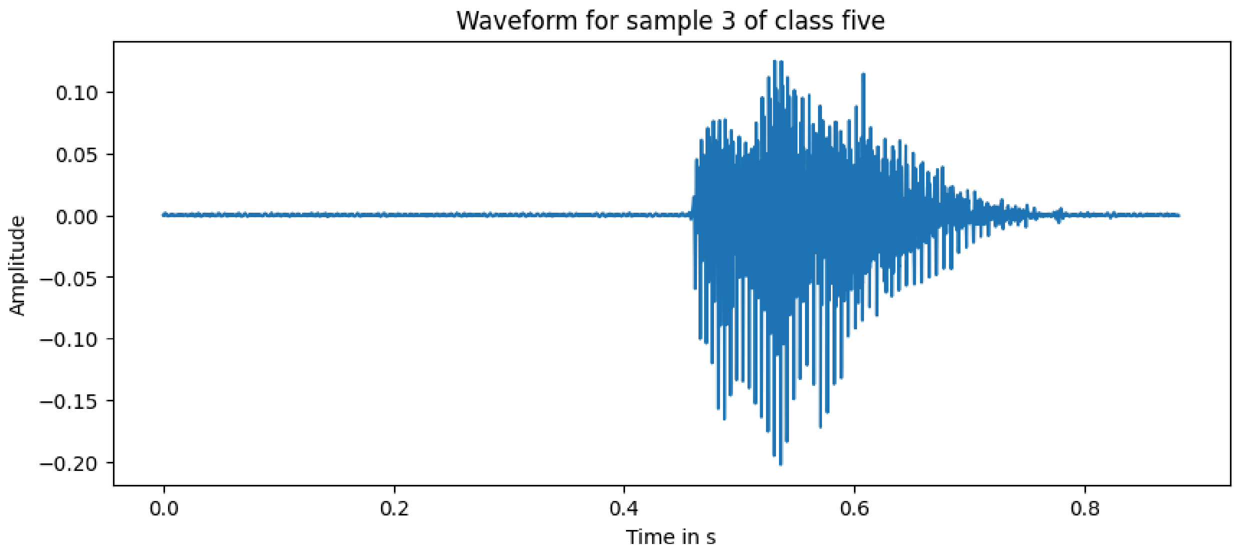

Manual STFT Spectrogram for sample 2 for class five
Frequency in Hertz
1000
800
600
400
200
0
0
5
10
15
20
25
Time in frames
Amplitude in decibels
10
0
−10
−20
−30
−40
−50

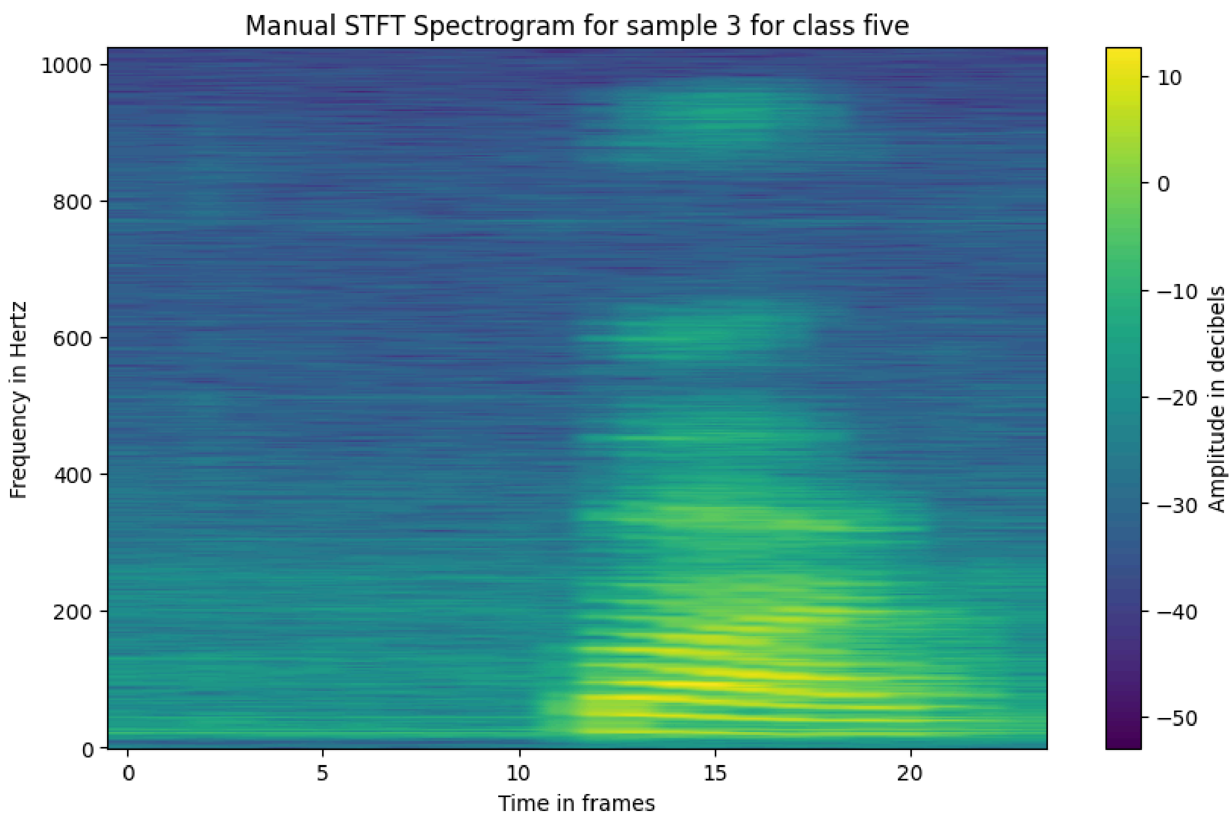


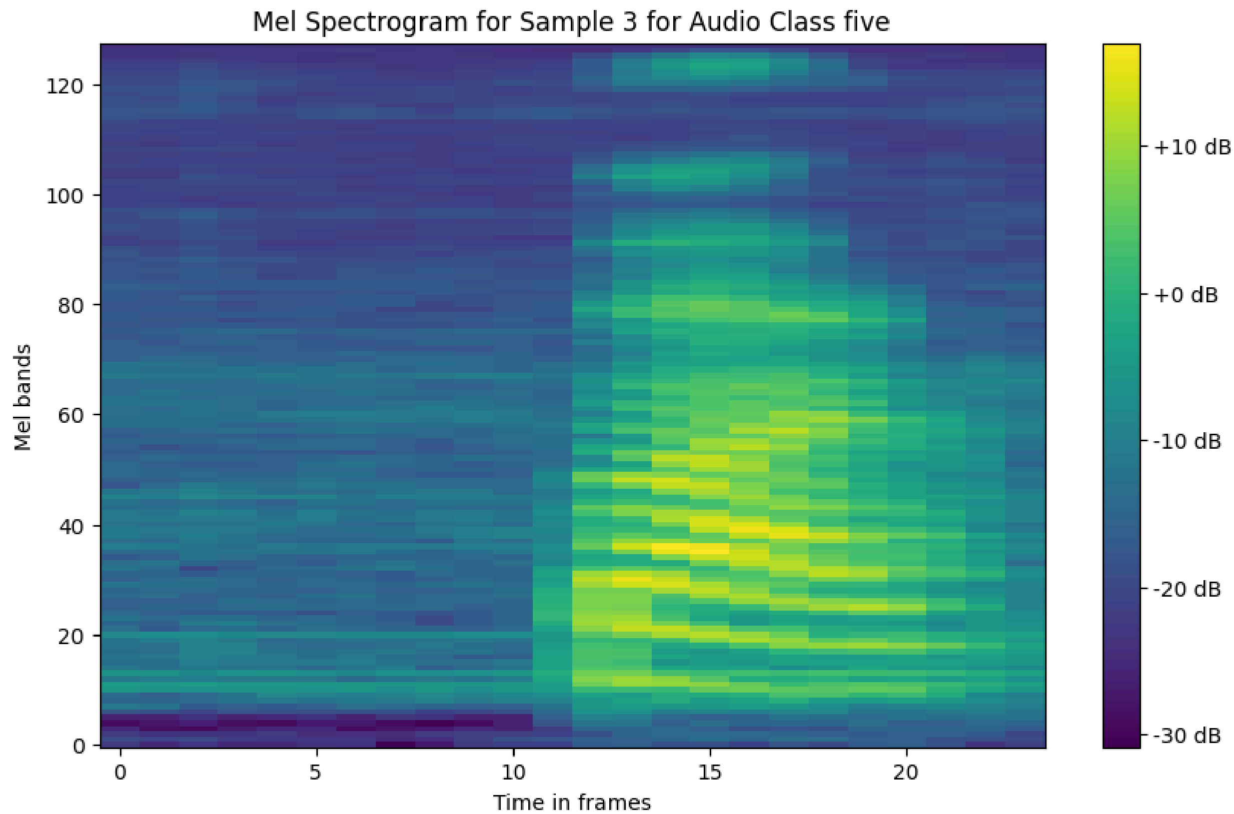
Mel Spectrogram for Sample 2 for Audio Class five
Mel bands
Time in frames
0
20
40
60
80
100
120
0
5
10
15
20
25
+10 dB
+5 dB
+0 dB
-5 dB
-10 dB
-15 dB
-20 dB

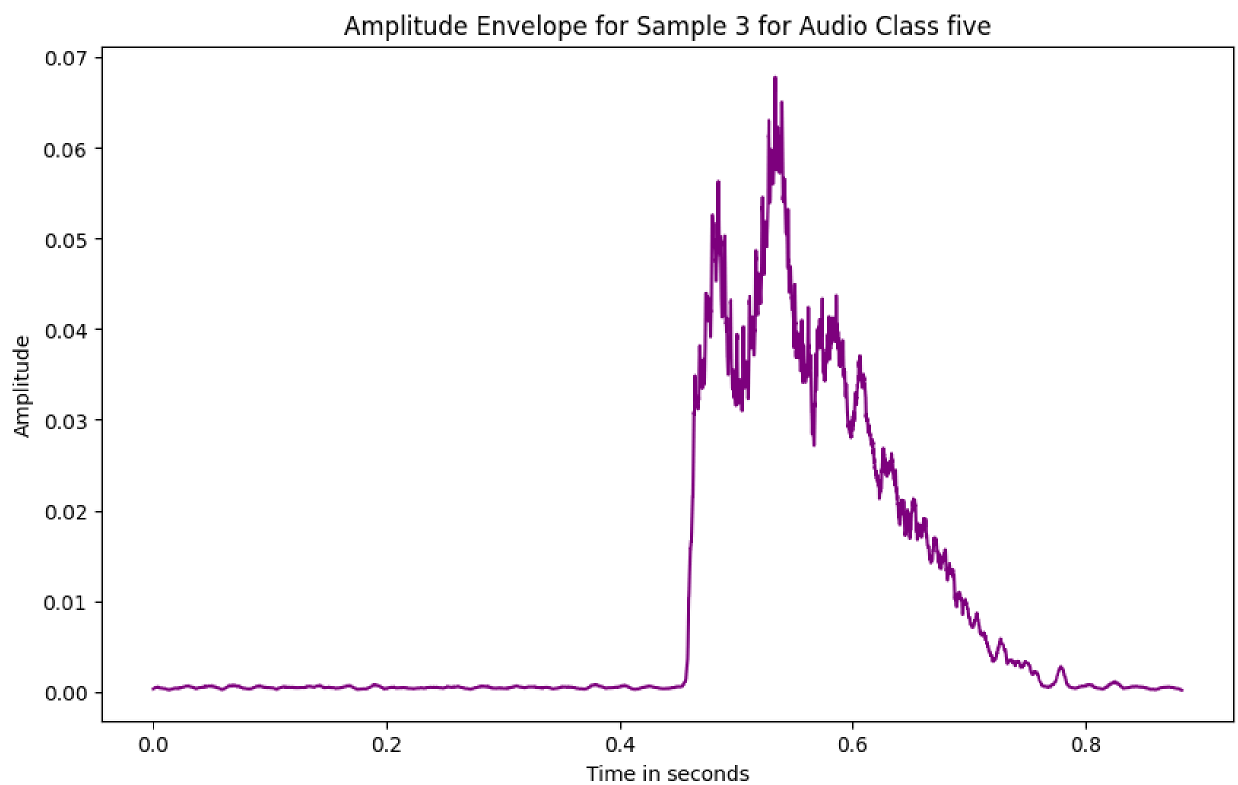



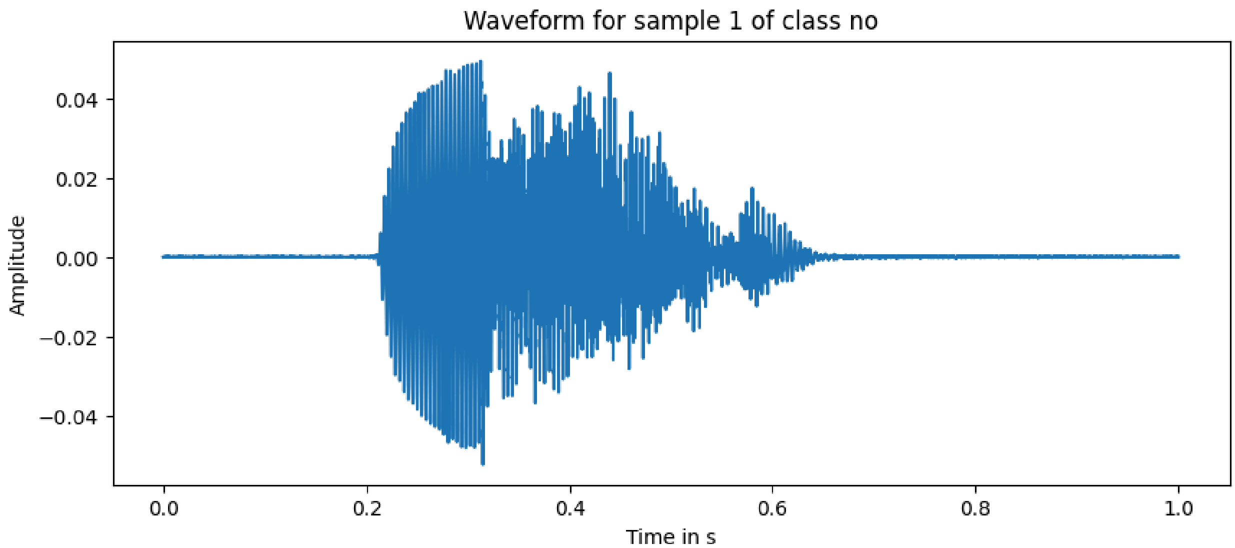
Amplitude Envelope for Sample 2 for Audio Class five
Amplitude
Time in seconds
0.01
0.02
0.03
0.04
0.05
0.0
0.2
0.4
0.6
0.8
1.0

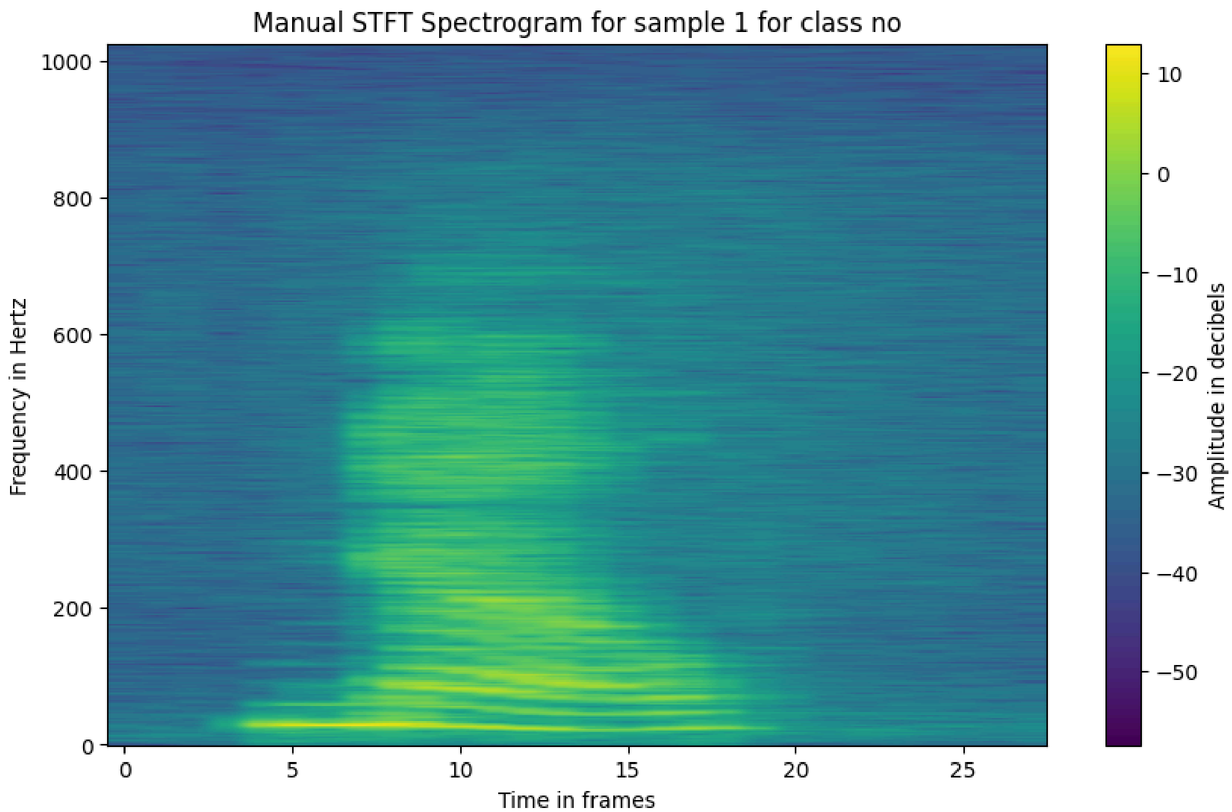

Waveform for sample 3 of class five
Amplitude
0.10
0.05
0.00
−0.05
−0.10
−0.15
−0.20
0.0
0.2
0.4
0.6
0.8
Time in s

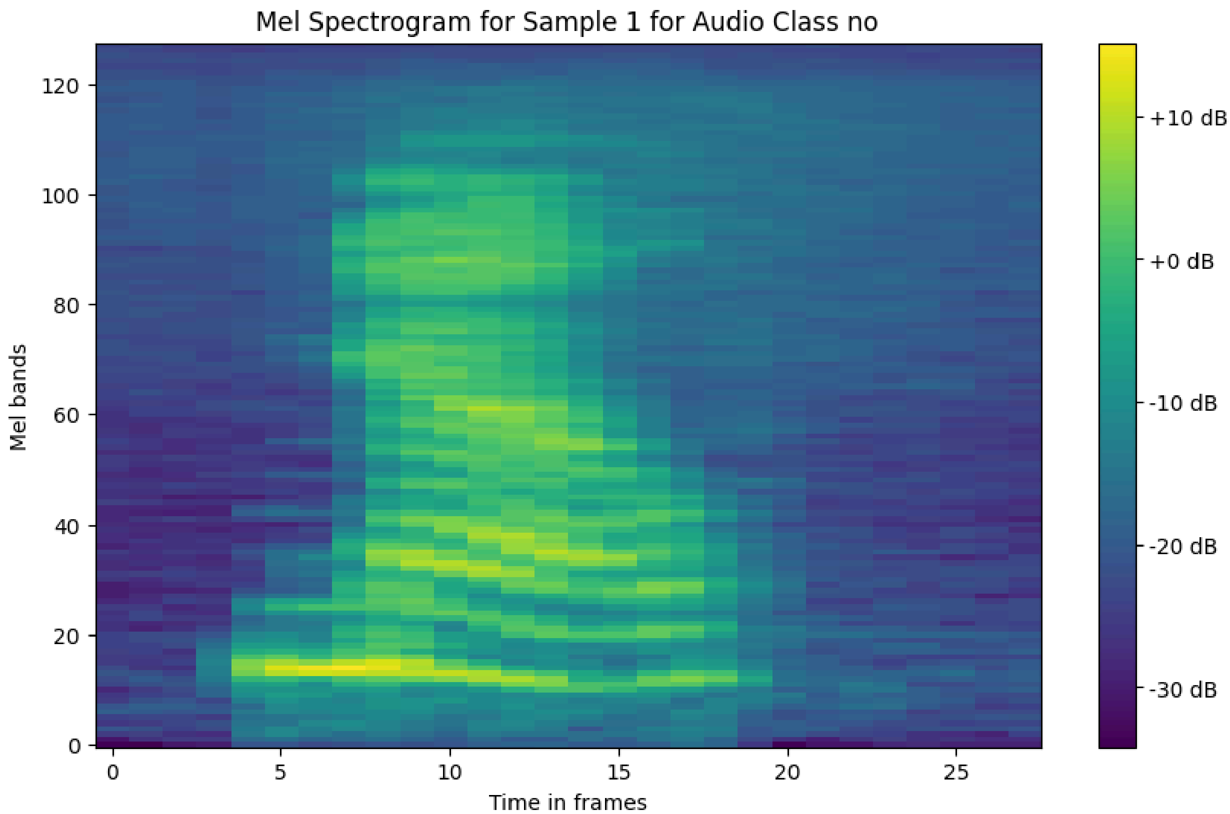

Manual STFT Spectrogram for sample 3 for class five
Frequency in Hertz
1000
800
600
400
200
0
0
5
10
15
20
Time in frames
Amplitude in decibels
10
0
−10
−20
−30
−40
−50


Mel Spectrogram for Sample 3 for Audio Class five
Mel bands
Time in frames
0
20
40
60
80
100
120
0
5
10
15
20
+10 dB
+0 dB
-10 dB
-20 dB
-30 dB

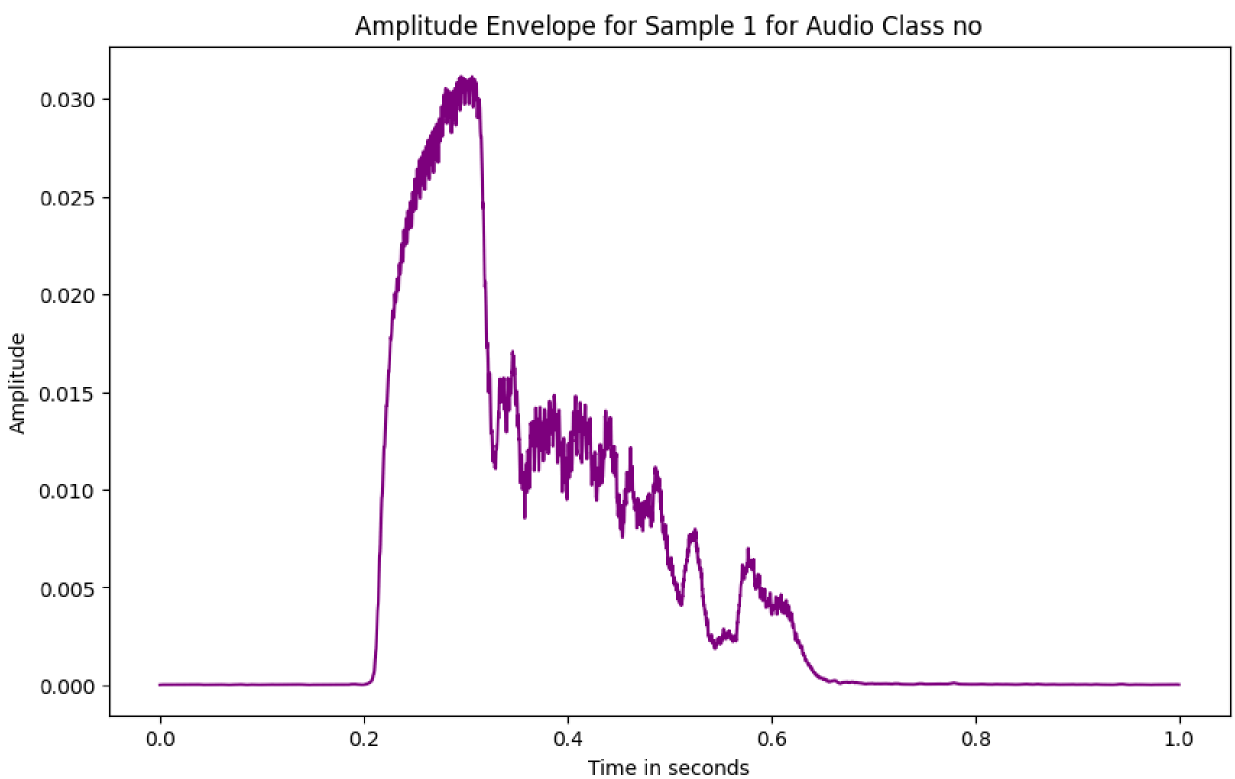

Amplitude Envelope for Sample 3 for Audio Class five
Amplitude
Time in seconds
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.0
0.2
0.4
0.6
0.8


Waveform for sample 1 of class no
Amplitude
0.04
0.02
0.00
−0.02
−0.04
0.0
0.2
0.4
0.6
0.8
1.0
Time in s

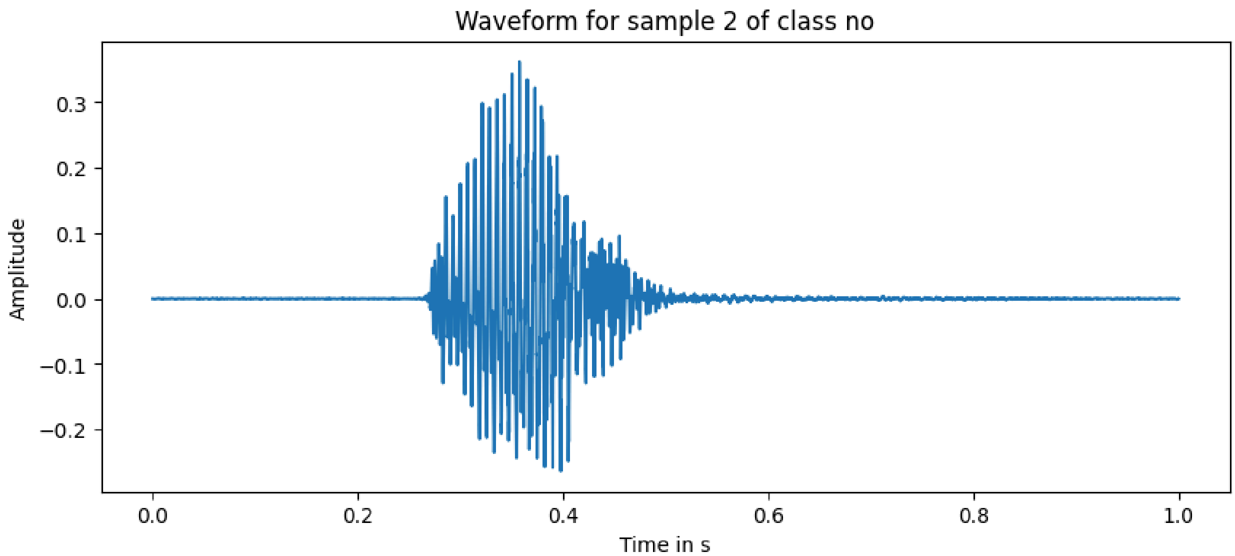

Manual STFT Spectrogram for sample 1 for class no
Frequency in Hertz
1000
800
600
400
200
0
0
5
10
15
20
25
Time in frames
Amplitude in decibels
10
0
−10
−20
−30
−40
−50

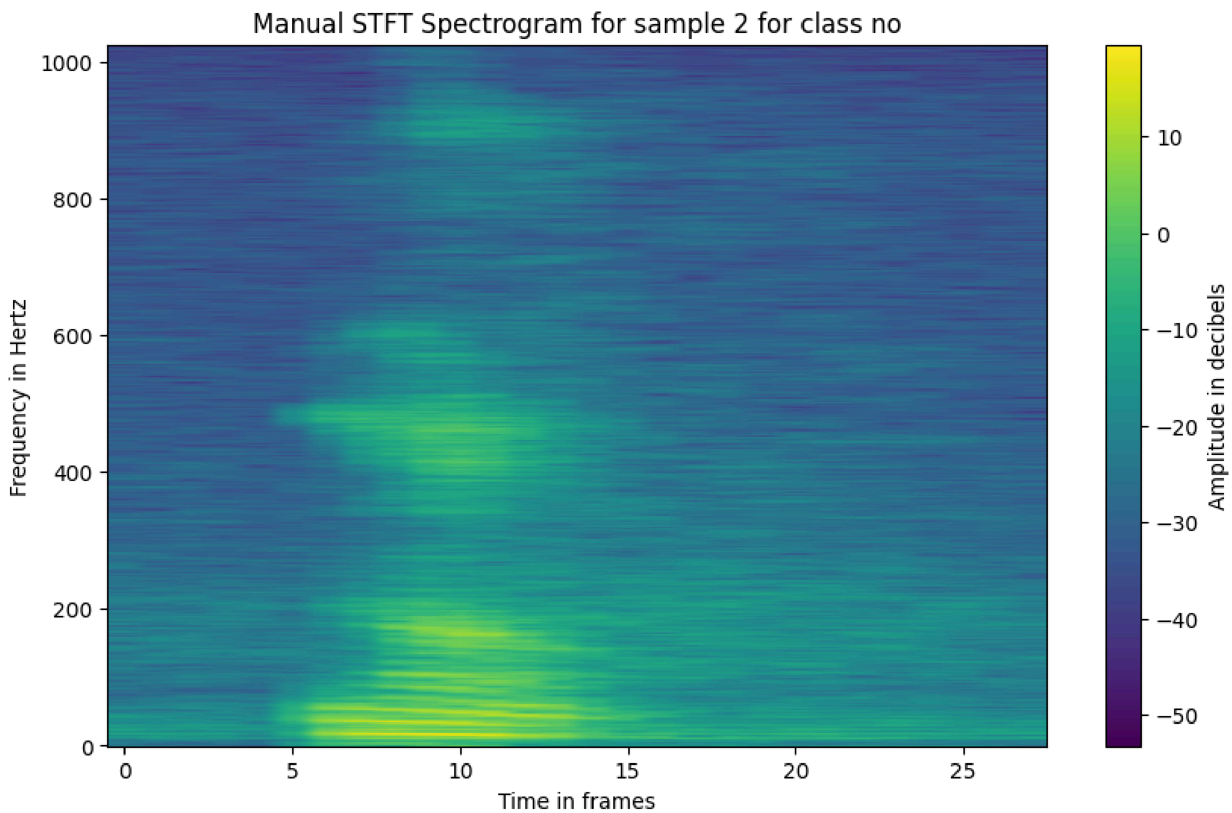


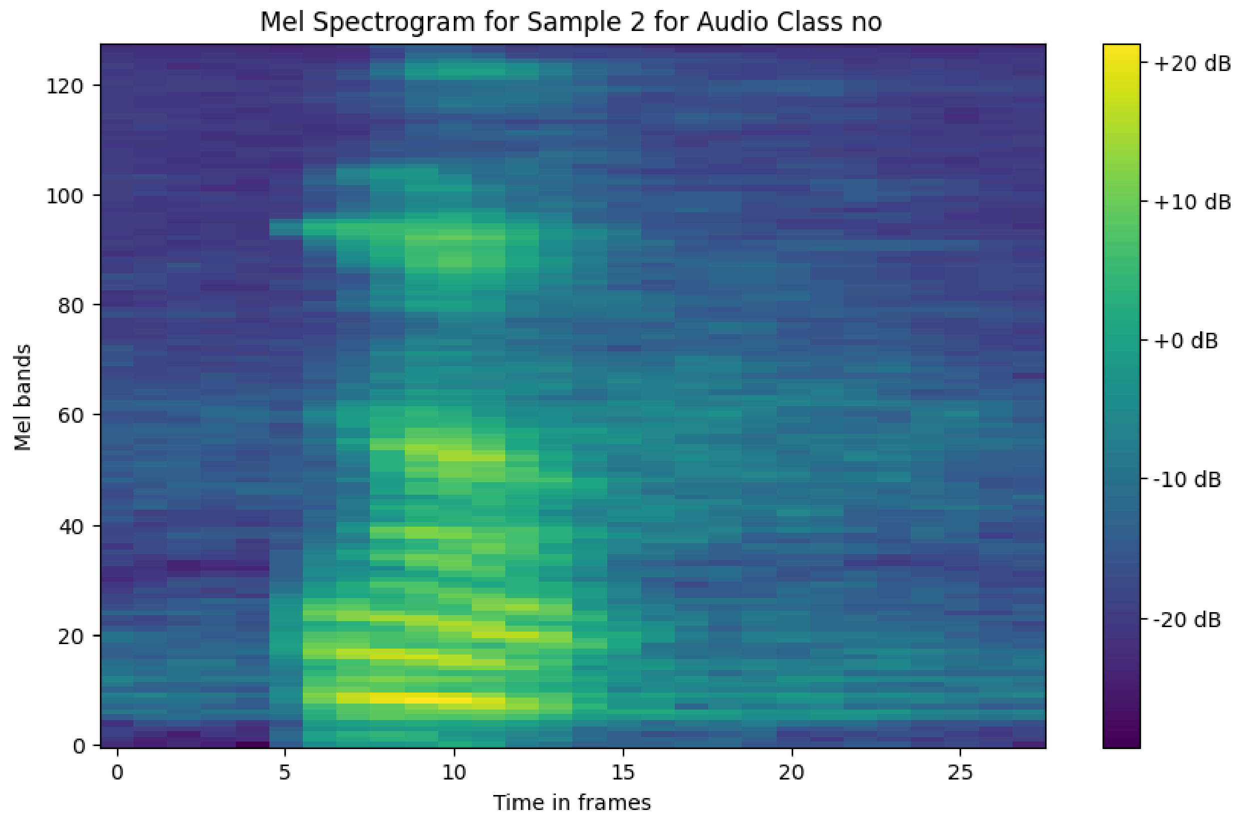
Mel Spectrogram for Sample 1 for Audio Class no
Mel bands
Time in frames
0
5
10
15
20
25
0
20
40
60
80
100
120
+10 dB
+0 dB
-10 dB
-20 dB
-30 dB

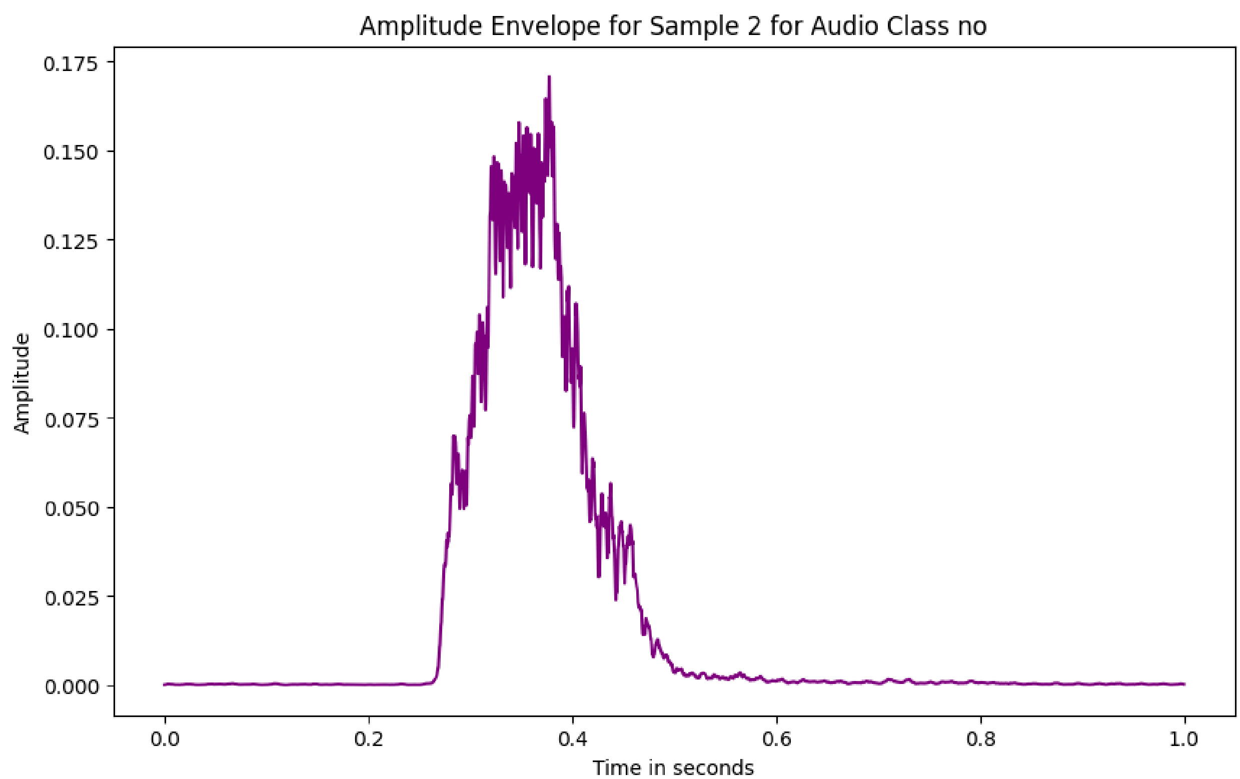



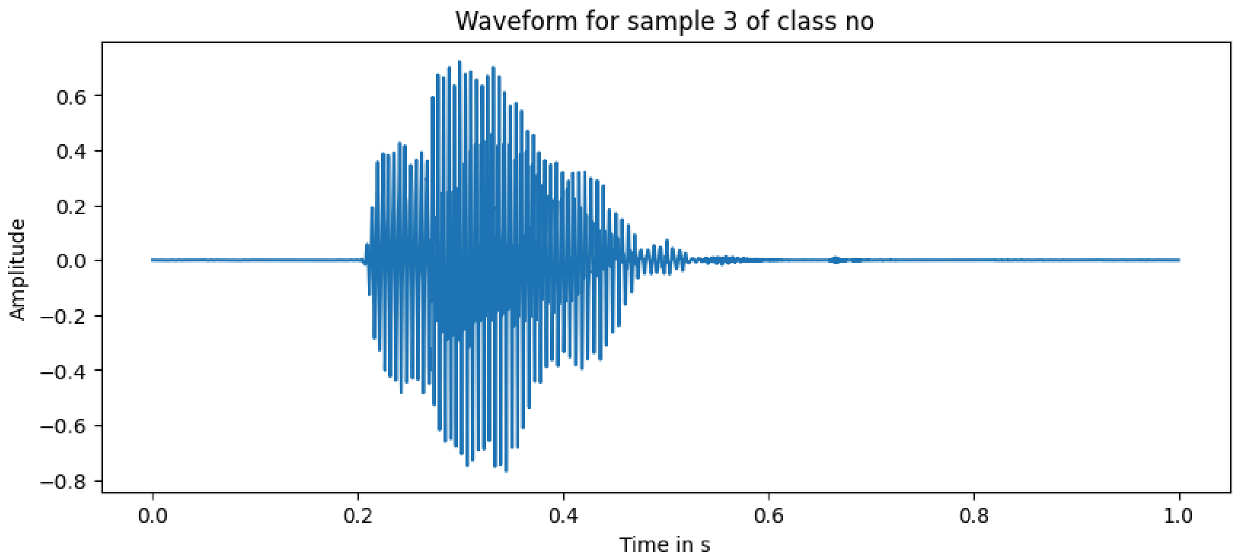
Amplitude Envelope for Sample 1 for Audio Class no
Amplitude
Time in seconds
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.0
0.2
0.4
0.6
0.8
1.0

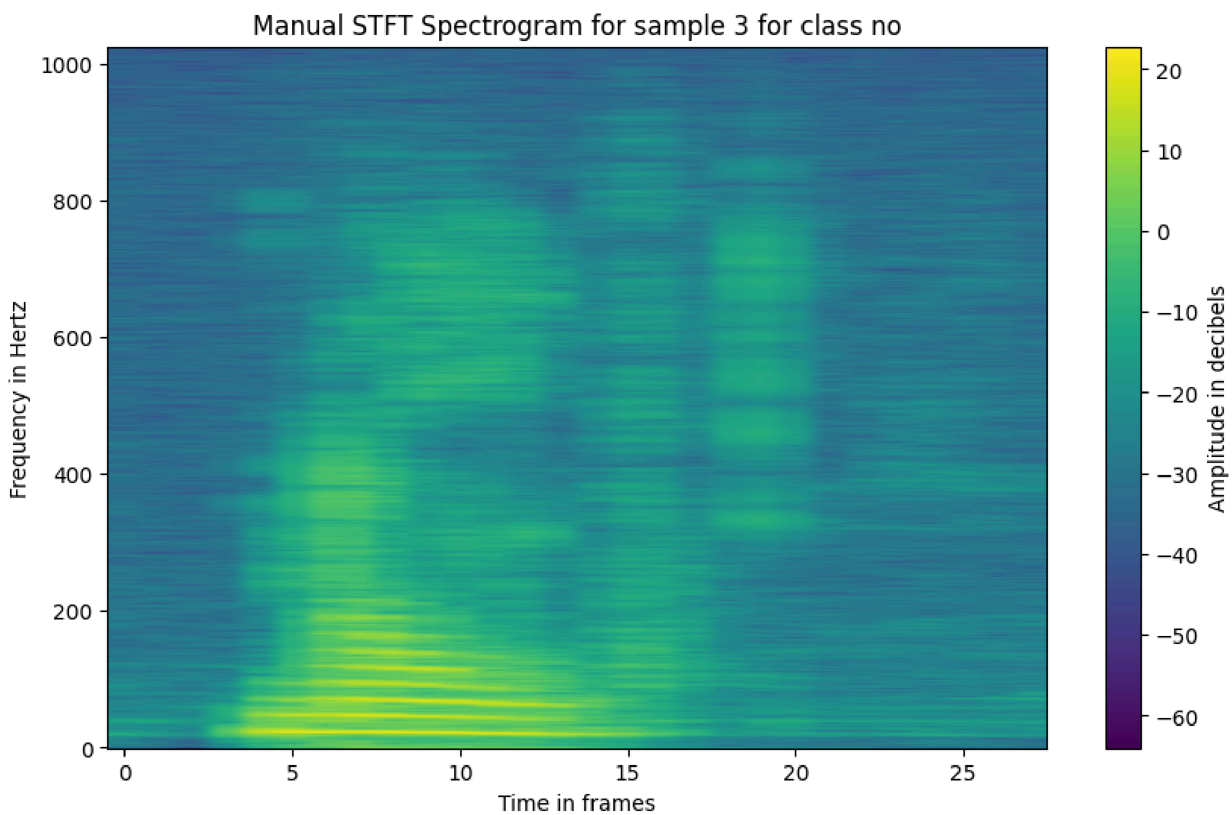

Waveform for sample 2 of class no
Amplitude
0.3
0.2
0.1
0.0
−0.1
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
Time in s

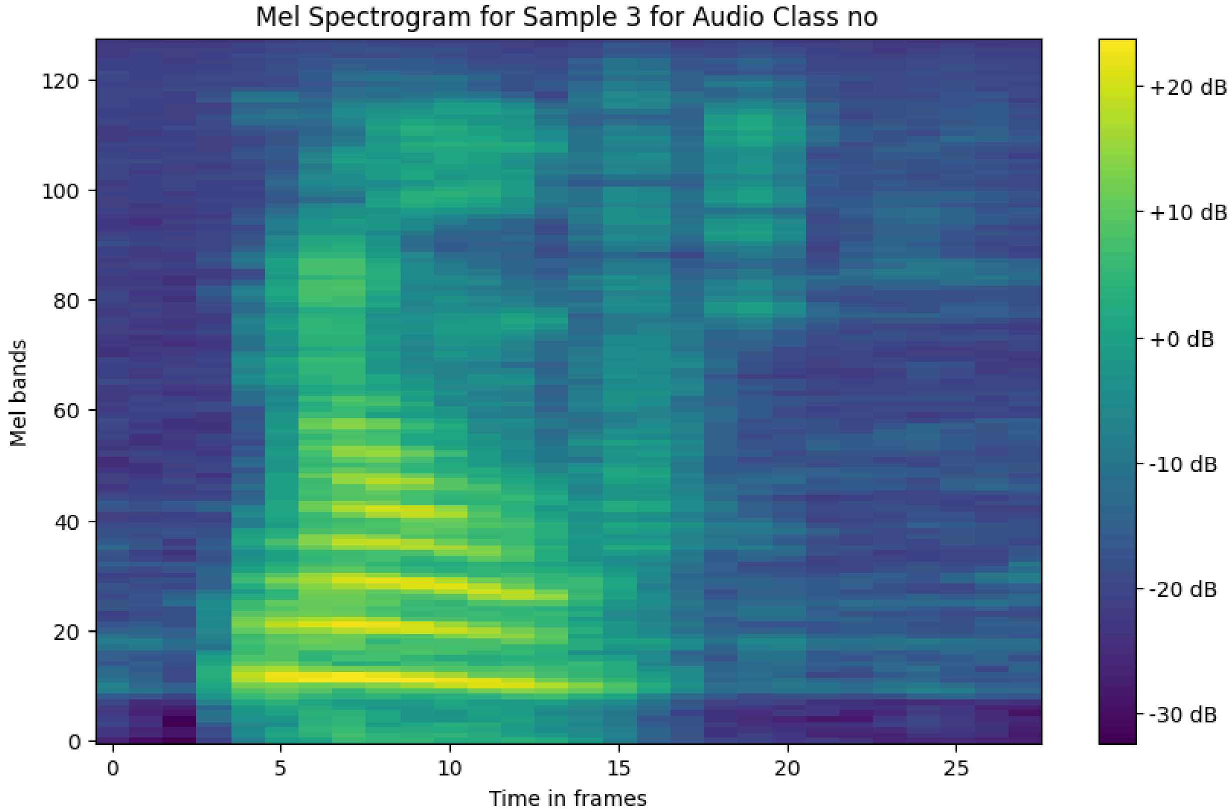

Manual STFT Spectrogram for sample 2 for class no
Frequency in Hertz
1000
800
600
400
200
0
0
5
10
15
20
25
Time in frames
Amplitude in decibels
10
0
−10
−20
−30
−40
−50

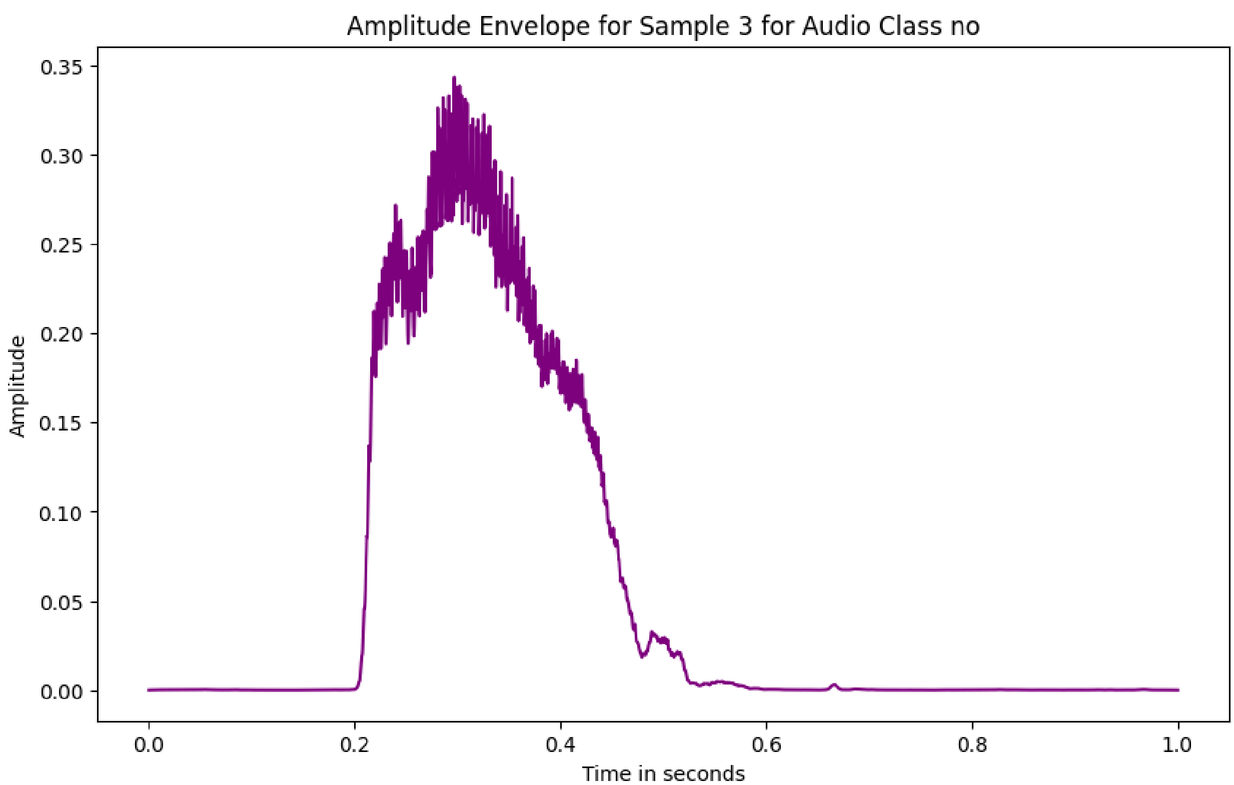

Mel Spectrogram for Sample 2 for Audio Class no
Mel bands
Time in frames
120
100
80
60
40
20
0
0
5
10
15
20
25
+20 dB
+10 dB
+0 dB
-10 dB
-20 dB


Amplitude Envelope for Sample 2 for Audio Class no
Amplitude
Time in seconds
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0.0
0.2
0.4
0.6
0.8
1.0


Waveform for sample 3 of class no
Amplitude
Time in s
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
0.0
0.2
0.4
0.6
0.8
1.0


Manual STFT Spectrogram for sample 3 for class no
Frequency in Hertz
Time in frames
Amplitude in decibels
1000
800
600
400
200
0
5
10
15
20
25
−10
−20
−30
−40
−50
−60

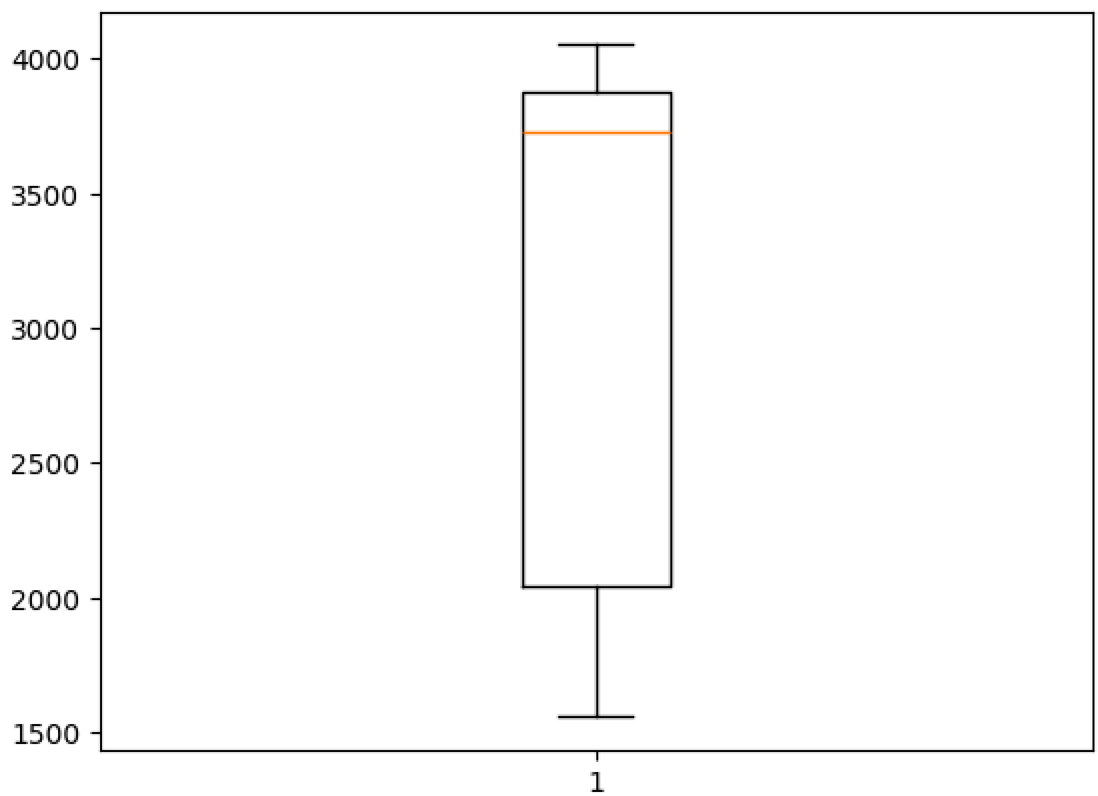

Mel Spectrogram for Sample 3 for Audio Class no
Mel bands
Time in frames
+20 dB
+10 dB
+0 dB
-10 dB
-20 dB
-30 dB


Amplitude Envelope for Sample 3 for Audio Class no
Amplitude
Time in seconds

Observations:- The amplitude envelopes show the obvious fact that there's a lot of silence in a lot of these clips which needs to be cut out, as that is just a waste of compute. All these samples show that there is no consistent pronouncing of words. Only 1 chunk and then silence. The spectrograms show that there is a bit of difference between even 2 samples of the same class, where the intensity of one differs dramatically from the other. This indicates that the way 2 words are said at least in terms of intensity, isn't indicative of its class.

(c) I resampled the data such that those that had less number of samples got oversampled(by duplicating entries at random) so that they reach the mean no. of samples per class. And those that had more, got undersampled by removing entries at random.

Before adjusting class imbalances:-




```
Mean Number of Samples for each Class -> 3023.6857142857143
Min Number of Samples for each Class -> 1557
Max Number of Samples for each Class -> 4052
Median Number of Samples for each Class -> 3728.0
```

(d) I removed the silences from the audio segments by having an amplitude threshold below which it is considered as silence. Those clips that had a prolonged period of silence where removed. I judged the threshold to be the mean, as it had good results.

Examples can be found by listening to the new .wav file formed.

Q2. I extracted the spectral features from the spectrogram.