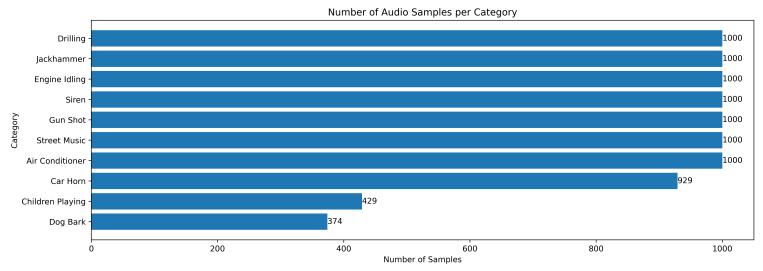
```
In [1]: import IPython.display as ipd
        import librosa
        import librosa.display
        import pandas as pd
        import os, time, warnings
        import seaborn as sns
        import numpy as np
        from tqdm import tqdm
        import matplotlib.pyplot as plt
        from sklearn.model selection import train test split
        from sklearn.preprocessing import LabelEncoder
        from tensorflow.keras.utils import to categorical
        from tensorflow.keras.models import Sequential
        from tensorflow.keras.layers import (
            Dense,
            Conv1D,
            MaxPooling1D,
            BatchNormalization,
            Dropout,
            Flatten,
            Conv2D,
            MaxPool2D,
        warnings.filterwarnings("ignore")
In [2]: log_cols = ["model", "accuracy", "train_time", "pred_time"]
        log = pd.DataFrame(columns=log cols)
In [3]: # reading the files
        audio dataset path = "C:/Users/vidit/Desktop/MS in DS/Projects/Audio Classification/urbansound8k/audio/"
        # Loading the csv
        meta data = pd.read csv("C:/Users/vidit/Desktop/MS in DS/Projects/Audio Classification/urbansound8k/UrbanSound8K.csv'
        meta data["class"] = meta data["class"].replace(
            to replace="air conditioner", value="Air Conditioner"
        meta data["class"] = meta_data["class"].replace(to_replace="car_horn", value="Car Horn")
        meta data["class"] = meta data["class"].replace(
            to replace="children playing", value="Children Playing"
```

Out[3]: slice_file_name fsID start end salience fold classID class 100032-3-0-0.wav 100032 5 Dog Bark 0.0 0.317551 3 **1** 100263-2-0-117.wav 100263 58.5 62.500000 1 5 2 Children Playing **2** 100263-2-0-121.wav 100263 5 60.5 64.500000 2 Children Playing **3** 100263-2-0-126.wav 100263 63.0 67.000000 5 2 Children Playing **4** 100263-2-0-137.wav 100263 68.5 72.500000 1 5 2 Children Playing

```
In [4]: meta_data.groupby("classID")["class"].unique()
```

```
Out[4]: classID
         0
               [Air Conditioner]
         1
                      [Car Horn]
         2
              [Children Playing]
         3
                      [Dog Bark]
         4
                      [Drilling]
         5
                 [Engine Idling]
         6
                      [Gun Shot]
         7
                    [Jackhammer]
         8
                         [Siren]
                  [Street Music]
         Name: class, dtype: object
```

```
In [5]: x = meta_data["class"].unique()
    y = meta_data["class"].value_counts(ascending=True)
    ind = np.arange(len(y))
    # plt.figure()
    fig, ax = plt.subplots(figsize=(15, 5))
    ax.barh(ind, y)
    ax.set_yticks(ind)
    ax.set_yticklabels(x)
    ax.bar_label(ax.containers[0])
    plt.gcf().set_dpi(500)
    plt.title("Number of Audio Samples per Category")
    plt.xlabel("Number of Samples")
    plt.ylabel("Category")
    plt.show()
```



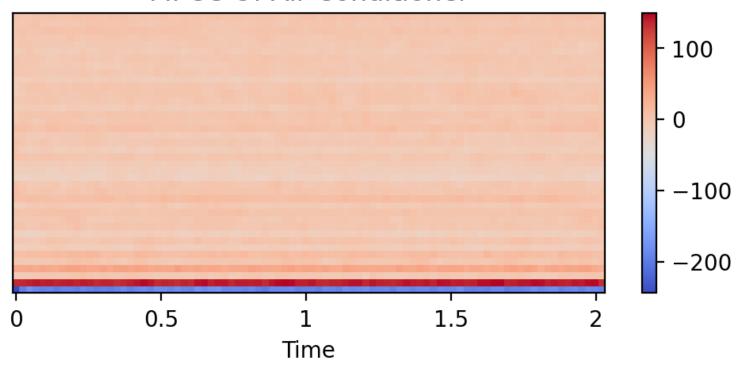
MFCC Visualization

```
In [6]: plt.rcParams["figure.figsize"] = (5, 2.5)
plt.rcParams["figure.dpi"] = 200
```

```
In [7]: audio_path = audio_dataset_path + "fold1/127873-0-0-0.wav"
    (xf, sr) = librosa.load(audio_path)
    mfccs = librosa.feature.mfcc(y=xf, sr=sr, n_mfcc=40)
    librosa.display.specshow(mfccs, x_axis="time")
    plt.colorbar()
    plt.tight_layout()
    plt.title("MFCC Of Air Conditioner")
    plt.show
```

Out[7]: <function matplotlib.pyplot.show(close=None, block=None)>

MFCC Of Air Conditioner

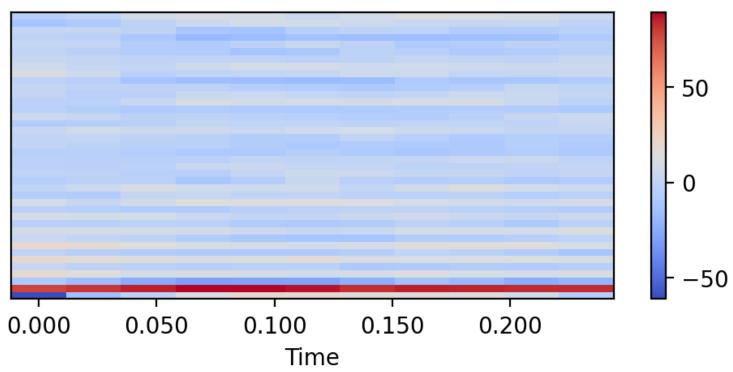


```
In [8]: audio_path = audio_dataset_path + "fold1/156194-1-0-0.wav"
  (xf, sr) = librosa.load(audio_path)
  mfccs = librosa.feature.mfcc(y=xf, sr=sr, n_mfcc=40)
```

```
librosa.display.specshow(mfccs, x_axis="time")
plt.colorbar()
plt.tight_layout()
plt.title("MFCC Of Car Horn")
plt.show
```

Out[8]: <function matplotlib.pyplot.show(close=None, block=None)>

MFCC Of Car Horn

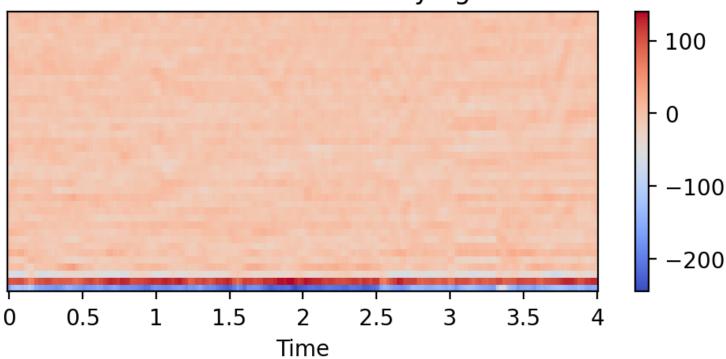


```
In [9]: audio_path = audio_dataset_path + "fold1/105415-2-0-1.wav"
    (xf, sr) = librosa.load(audio_path)
    mfccs = librosa.feature.mfcc(y=xf, sr=sr, n_mfcc=40)
    librosa.display.specshow(mfccs, x_axis="time")
    plt.colorbar()
    plt.tight_layout()
```

```
plt.title("MFCC Of Children Playing")
plt.show
```

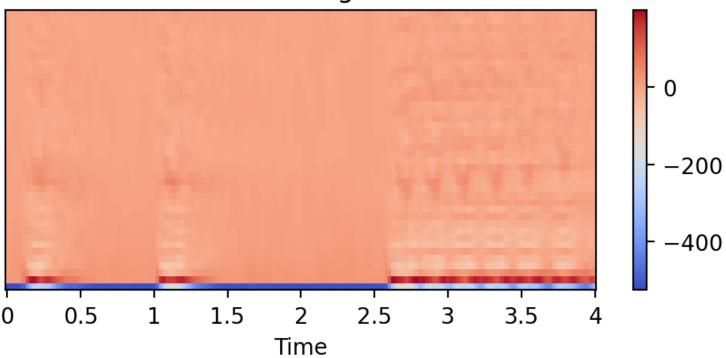
Out[9]: <function matplotlib.pyplot.show(close=None, block=None)>

MFCC Of Children Playing



```
In [10]: audio_path = audio_dataset_path + "fold1/101415-3-0-2.wav"
    (xf, sr) = librosa.load(audio_path)
    mfccs = librosa.feature.mfcc(y=xf, sr=sr, n_mfcc=40)
    librosa.display.specshow(mfccs, x_axis="time")
    plt.colorbar()
    plt.tight_layout()
    plt.title("MFCC Of Dog Bark")
    plt.show
```

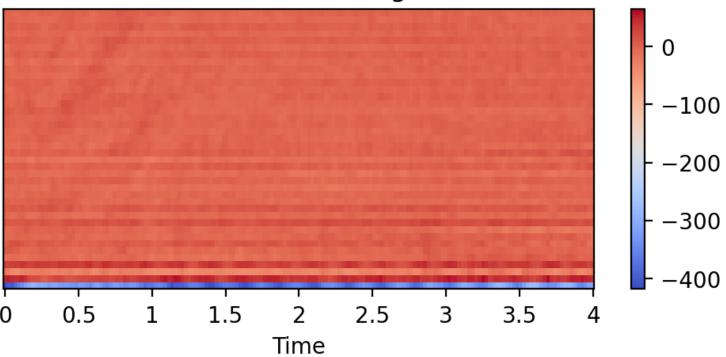
MFCC Of Dog Bark



```
In [11]:
    audio_path = audio_dataset_path + "fold1/14113-4-0-0.wav"
    (xf, sr) = librosa.load(audio_path)
    mfccs = librosa.feature.mfcc(y=xf, sr=sr, n_mfcc=40)
    librosa.display.specshow(mfccs, x_axis="time")
    plt.colorbar()
    plt.tight_layout()
    plt.title("MFCC Of Drilling")
    plt.show
```

Out[11]: <function matplotlib.pyplot.show(close=None, block=None)>

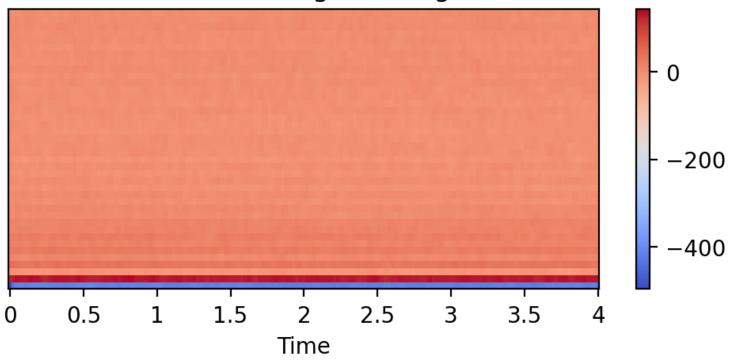
MFCC Of Drilling



```
In [12]:
    audio_path = audio_dataset_path + "fold1/103258-5-0-0.wav"
    (xf, sr) = librosa.load(audio_path)
    mfccs = librosa.feature.mfcc(y=xf, sr=sr, n_mfcc=40)
    librosa.display.specshow(mfccs, x_axis="time")
    plt.colorbar()
    plt.tight_layout()
    plt.title("MFCC Of Engine Idling")
    plt.show
```

Out[12]: <function matplotlib.pyplot.show(close=None, block=None)>

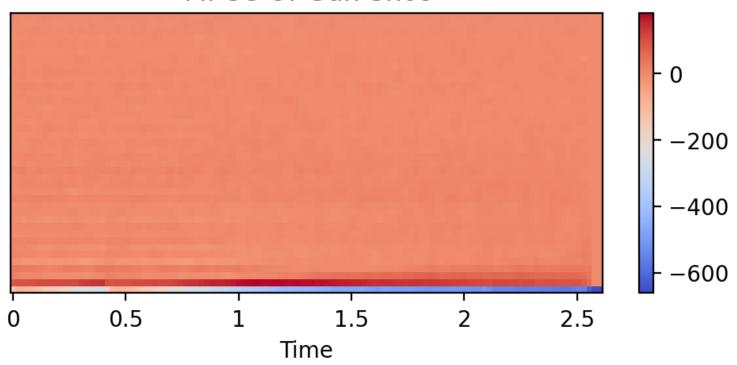
MFCC Of Engine Idling



```
In [13]:
    audio_path = audio_dataset_path + "fold1/102305-6-0-0.wav"
    (xf, sr) = librosa.load(audio_path)
    mfccs = librosa.feature.mfcc(y=xf, sr=sr, n_mfcc=40)
    librosa.display.specshow(mfccs, x_axis="time")
    plt.colorbar()
    plt.tight_layout()
    plt.title("MFCC Of Gun Shot")
    plt.show
```

Out[13]: <function matplotlib.pyplot.show(close=None, block=None)>

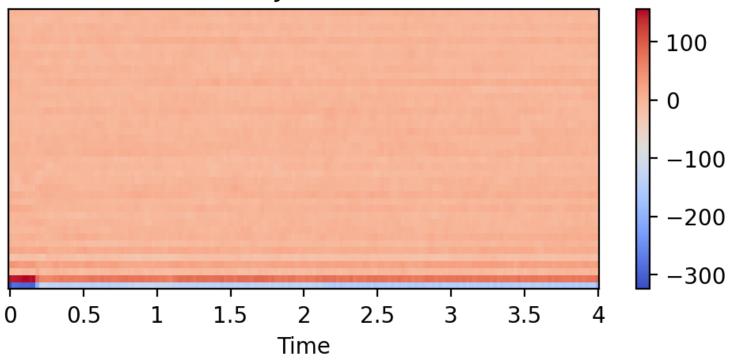
MFCC Of Gun Shot



```
In [14]:
    audio_path = audio_dataset_path + "fold1/103074-7-0-0.wav"
    (xf, sr) = librosa.load(audio_path)
    mfccs = librosa.feature.mfcc(y=xf, sr=sr, n_mfcc=40)
    librosa.display.specshow(mfccs, x_axis="time")
    plt.colorbar()
    plt.tight_layout()
    plt.title("MFCC Of Jack Hammer")
    plt.show
```

Out[14]: <function matplotlib.pyplot.show(close=None, block=None)>

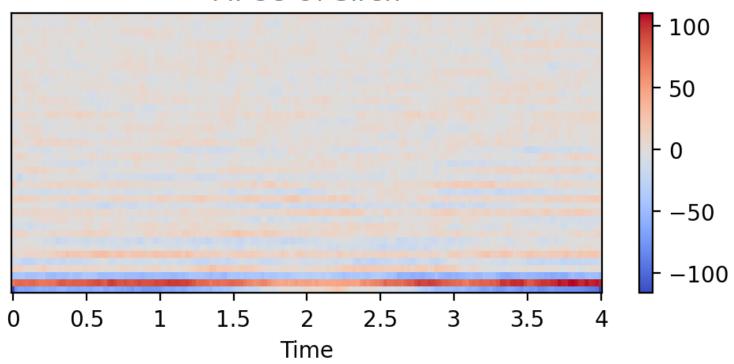
MFCC Of Jack Hammer



```
In [15]:
    audio_path = audio_dataset_path + "fold1/106905-8-0-0.wav"
    (xf, sr) = librosa.load(audio_path)
    mfccs = librosa.feature.mfcc(y=xf, sr=sr, n_mfcc=40)
    librosa.display.specshow(mfccs, x_axis="time")
    plt.colorbar()
    plt.tight_layout()
    plt.title("MFCC Of Siren")
    plt.show
```

Out[15]: <function matplotlib.pyplot.show(close=None, block=None)>

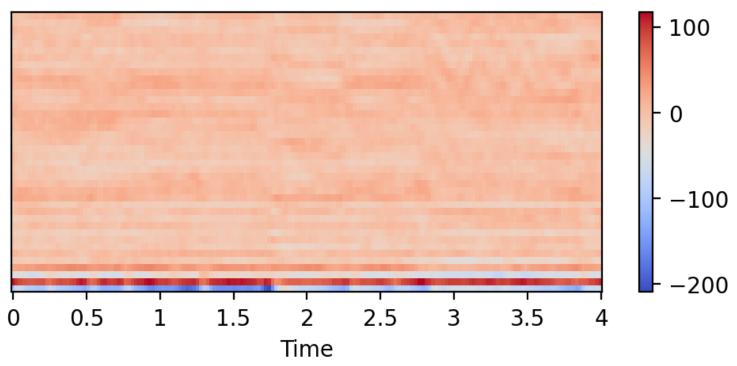
MFCC Of Siren



```
In [16]: audio_path = audio_dataset_path + "fold1/108041-9-0-11.wav"
    (xf, sr) = librosa.load(audio_path)
    mfccs = librosa.feature.mfcc(y=xf, sr=sr, n_mfcc=40)
    librosa.display.specshow(mfccs, x_axis="time")
    plt.colorbar()
    plt.tight_layout()
    plt.title("MFCC Of Street Music")
    plt.show
```

Out[16]: <function matplotlib.pyplot.show(close=None, block=None)>

MFCC Of Street Music



Feature Extraction and Database Building

- I have used Librosa to preprocess audio file.
- To do so, I will go through each fold and extract the data from each file using librosa's mfcc function.
- The extracted data is appended in a list and stored in a dataframe

```
In [17]: # list containing all the features
    extracted = []
# for each row in the csv
```

```
for index num, row in tqdm(meta data.iterrows()):
    # get the file
   file name = os.path.join(
       os.path.abspath(audio dataset path),
       "fold" + str(row["fold"]) + "/",
       str(row["slice_file_name"]),
   # get file label
   final class labels = row["class"]
   # load the audio file
   audio, sample rate = librosa.load(file name, res type="kaiser fast")
   # extract the features
   feature = librosa.feature.mfcc(y=audio, sr=sample rate, n mfcc=128)
   # feature scaling
   scaled_feature = np.mean(feature.T, axis=0)
   # store it in a list
   extracted.append([scaled feature, final class labels])
3553it [02:32, 23.61it/s]C:\Users\vidit\AppData\Roaming\Python\Python310\site-packages\librosa\core\spectrum.py:257:
UserWarning: n fft=2048 is too large for input signal of length=1323
 warnings.warn(
8326it [05:52, 31.85it/s]C:\Users\vidit\AppData\Roaming\Python\Python310\site-packages\librosa\core\spectrum.py:257:
UserWarning: n fft=2048 is too large for input signal of length=1103
  warnings.warn(
C:\Users\vidit\AppData\Roaming\Python\Python310\site-packages\librosa\core\spectrum.py:257: UserWarning: n fft=2048
is too large for input signal of length=1523
 warnings.warn(
8732it [06:08, 23.71it/s]
```

Data Preprocessing

Using a dataframe and pickle to save the extracted features array

```
In [18]: # create a new dataframe
    extracted_df = pd.DataFrame(extracted, columns=["feature", "class"])
# Storing the dataframe to pickle for further processing
    extracted_df.to_pickle("extracted_df.pkl")
    extracted_df.head()
```

Out[18]:		feature	class
	0	[-217.35526, 70.22338, -130.38527, -53.282898,	Dog Bark
	1	[-424.09818, 109.34077, -52.919525, 60.86475,	Children Playing
	2	[-458.79114, 121.38419, -46.52066, 52.00812,	Children Playing
	3	[-413.89984, 101.66371, -35.42945, 53.036358,	Children Playing
	4	[-446.60352, 113.68541, -52.402218, 60.302044,	Children Playing

Distribute the data to X and Y

```
In [19]: # create a new dataframe
final = pd.DataFrame(extracted, columns=["feature", "class"])
X = np.array(final["feature"].tolist())
y = np.array(final["class"].tolist())
```

Using LabelEncoder() to encode the string labels to an integer

```
In [20]: # Label encoding to get encoding
le = LabelEncoder()

# transform each category with it's respected Label
Y = to_categorical(le.fit_transform(y))
```

Split the data into train and test sets

```
Number of training samples = 6985
Number of testing samples = 1747
```

Model 1 - ANN

Building the model

```
In [22]: # Construct model

num_labels = Y.shape[1]
ANN_Model = Sequential()
ANN_Model.add(Dense(1000, activation="relu", input_shape=(128,)))
ANN_Model.add(Dense(750, activation="relu"))
ANN_Model.add(Dense(500, activation="relu"))
ANN_Model.add(Dense(500, activation="relu"))
ANN_Model.add(Dense(100, activation="relu"))
ANN_Model.add(Dense(50, activation="relu"))
ANN_Model.add(Dense(num_labels, activation="softmax"))
ANN_Model.summary()
```

Model: "sequential"

Layer (type)	Output	Shape	Param #			
dense (Dense)	(None,	1000)	129000			
dense_1 (Dense)	(None,	750)	750750			
dense_2 (Dense)	(None,	500)	375500			
dense_3 (Dense)	(None,	250)	125250			
dense_4 (Dense)	(None,	100)	25100			
dense_5 (Dense)	(None,	50)	5050			
dense_6 (Dense)	(None,	10)	510			
Total params: 1411160 (5.38 MB) Trainable params: 1411160 (5.38 MB) Non-trainable params: 0 (0.00 Byte)						

Compiling the model

Fitting the model

```
In [24]: num_epochs = 250
num_batch_size = 32

t0 = time.time()

ANN_Results = ANN_Model.fit(
    X_train,
```

```
y_train,
batch_size=num_batch_size,
epochs=num_epochs,
validation_data=(X_test, y_test),
)

ANN_Model.save("Model1.h5")
print("ANN Model Saved")
train_hist_m1 = pd.DataFrame(ANN_Results.history)
train_m1 = round(time.time() - t0, 3)
```

```
Epoch 1/250
ccuracy: 0.5667
Epoch 2/250
ccuracy: 0.6938
Epoch 3/250
219/219 [============= - - 5s 22ms/step - loss: 0.7350 - accuracy: 0.7585 - val loss: 0.7043 - val a
ccuracy: 0.7653
Epoch 4/250
ccuracy: 0.8323
Epoch 5/250
ccuracy: 0.8706
Epoch 6/250
ccuracy: 0.8517
Epoch 7/250
ccuracy: 0.8609
Epoch 8/250
ccuracy: 0.8855
Epoch 9/250
ccuracy: 0.8729
Epoch 10/250
ccuracy: 0.8924
Epoch 11/250
ccuracy: 0.9170
Epoch 12/250
ccuracy: 0.8844
Epoch 13/250
ccuracy: 0.8975
Epoch 14/250
ccuracy: 0.9153
```

```
Epoch 15/250
ccuracy: 0.9204
Epoch 16/250
ccuracy: 0.9176
Epoch 17/250
ccuracy: 0.9273
Epoch 18/250
ccuracy: 0.9038
Epoch 19/250
ccuracy: 0.8998
Epoch 20/250
ccuracy: 0.9078
Epoch 21/250
ccuracy: 0.9307
Epoch 22/250
ccuracy: 0.9222
Epoch 23/250
ccuracy: 0.9244
Epoch 24/250
ccuracy: 0.9227
Epoch 25/250
ccuracy: 0.9244
Epoch 26/250
ccuracy: 0.9296
Epoch 27/250
ccuracy: 0.9216
Epoch 28/250
ccuracy: 0.9296
```

```
Epoch 29/250
ccuracy: 0.9273
Epoch 30/250
ccuracy: 0.9405
Epoch 31/250
ccuracy: 0.9204
Epoch 32/250
ccuracy: 0.9290
Epoch 33/250
ccuracy: 0.9473
Epoch 34/250
ccuracy: 0.9290
Epoch 35/250
ccuracy: 0.9284
Epoch 36/250
ccuracy: 0.9279
Epoch 37/250
ccuracy: 0.9290
Epoch 38/250
ccuracy: 0.9428
Epoch 39/250
ccuracy: 0.8861
Epoch 40/250
ccuracy: 0.9233
Epoch 41/250
ccuracy: 0.9330
Epoch 42/250
ccuracy: 0.9422
```

```
Epoch 43/250
ccuracy: 0.9084
Epoch 44/250
ccuracy: 0.9388
Epoch 45/250
ccuracy: 0.9279
Epoch 46/250
ccuracy: 0.9353
Epoch 47/250
ccuracy: 0.9222
Epoch 48/250
ccuracy: 0.9216
Epoch 49/250
ccuracy: 0.9376
Epoch 50/250
ccuracy: 0.9222
Epoch 51/250
ccuracy: 0.9290
Epoch 52/250
ccuracy: 0.9307
Epoch 53/250
ccuracy: 0.9393
Epoch 54/250
ccuracy: 0.9376
Epoch 55/250
ccuracy: 0.9410
Epoch 56/250
ccuracy: 0.9336
```

```
Epoch 57/250
ccuracy: 0.9388
Epoch 58/250
ccuracy: 0.9370
Epoch 59/250
ccuracy: 0.9342
Epoch 60/250
ccuracy: 0.9267
Epoch 61/250
ccuracy: 0.9210
Epoch 62/250
ccuracy: 0.9393
Epoch 63/250
ccuracy: 0.9336
Epoch 64/250
ccuracy: 0.9222
Epoch 65/250
ccuracy: 0.9405
Epoch 66/250
ccuracy: 0.9342
Epoch 67/250
ccuracy: 0.9365
Epoch 68/250
ccuracy: 0.9479
Epoch 69/250
ccuracy: 0.9347
Epoch 70/250
ccuracy: 0.9256
```

```
Epoch 71/250
ccuracy: 0.9319
Epoch 72/250
ccuracy: 0.9445
Epoch 73/250
ccuracy: 0.8975
Epoch 74/250
ccuracy: 0.9388
Epoch 75/250
ccuracy: 0.9445
Epoch 76/250
ccuracy: 0.9456
Epoch 77/250
ccuracy: 0.9428
Epoch 78/250
ccuracy: 0.9479
Epoch 79/250
ccuracy: 0.9244
Epoch 80/250
ccuracy: 0.9382
Epoch 81/250
ccuracy: 0.9210
Epoch 82/250
ccuracy: 0.9284
Epoch 83/250
ccuracy: 0.9307
Epoch 84/250
ccuracy: 0.9450
```

```
Epoch 85/250
ccuracy: 0.9273
Epoch 86/250
ccuracy: 0.9376
Epoch 87/250
ccuracy: 0.9410
Epoch 88/250
ccuracy: 0.9353
Epoch 89/250
ccuracy: 0.9433
Epoch 90/250
ccuracy: 0.9445
Epoch 91/250
ccuracy: 0.9450
Epoch 92/250
ccuracy: 0.9468
Epoch 93/250
ccuracy: 0.9479
Epoch 94/250
ccuracy: 0.9433
Epoch 95/250
ccuracy: 0.9267
Epoch 96/250
ccuracy: 0.9353
Epoch 97/250
ccuracy: 0.9141
Epoch 98/250
ccuracy: 0.9256
```

```
Epoch 99/250
ccuracy: 0.9365
Epoch 100/250
ccuracy: 0.9422
Epoch 101/250
ccuracy: 0.9365
Epoch 102/250
ccuracy: 0.9439
Epoch 103/250
ccuracy: 0.9456
Epoch 104/250
ccuracy: 0.9307
Epoch 105/250
ccuracy: 0.9405
Epoch 106/250
ccuracy: 0.9428
Epoch 107/250
ccuracy: 0.8935
Epoch 108/250
ccuracy: 0.9347
Epoch 109/250
ccuracy: 0.9353
Epoch 110/250
ccuracy: 0.9399
Epoch 111/250
ccuracy: 0.9370
Epoch 112/250
ccuracy: 0.9353
```

```
Epoch 113/250
ccuracy: 0.9428
Epoch 114/250
ccuracy: 0.9445
Epoch 115/250
ccuracy: 0.9428
Epoch 116/250
ccuracy: 0.9456
Epoch 117/250
ccuracy: 0.9445
Epoch 118/250
ccuracy: 0.9433
Epoch 119/250
ccuracy: 0.9456
Epoch 120/250
ccuracy: 0.9450
Epoch 121/250
ccuracy: 0.8849
Epoch 122/250
ccuracy: 0.9336
Epoch 123/250
ccuracy: 0.9353
Epoch 124/250
ccuracy: 0.9325
Epoch 125/250
ccuracy: 0.9450
Epoch 126/250
ccuracy: 0.9428
```

```
Epoch 127/250
ccuracy: 0.9542
Epoch 128/250
ccuracy: 0.9536
Epoch 129/250
ccuracy: 0.9519
Epoch 130/250
ccuracy: 0.9531
Epoch 131/250
ccuracy: 0.9491
Epoch 132/250
ccuracy: 0.9508
Epoch 133/250
ccuracy: 0.9508
Epoch 134/250
ccuracy: 0.9508
Epoch 135/250
ccuracy: 0.9496
Epoch 136/250
ccuracy: 0.8786
Epoch 137/250
ccuracy: 0.9388
Epoch 138/250
ccuracy: 0.9485
Epoch 139/250
ccuracy: 0.9445
Epoch 140/250
ccuracy: 0.9147
```

```
Epoch 141/250
ccuracy: 0.9365
Epoch 142/250
ccuracy: 0.9485
Epoch 143/250
ccuracy: 0.9491
Epoch 144/250
ccuracy: 0.9433
Epoch 145/250
ccuracy: 0.9473
Epoch 146/250
ccuracy: 0.9485
Epoch 147/250
ccuracy: 0.9542
Epoch 148/250
ccuracy: 0.9433
Epoch 149/250
ccuracy: 0.9531
Epoch 150/250
ccuracy: 0.9542
Epoch 151/250
ccuracy: 0.9525
Epoch 152/250
ccuracy: 0.9365
Epoch 153/250
ccuracy: 0.9279
Epoch 154/250
ccuracy: 0.9473
```

```
Epoch 155/250
ccuracy: 0.9428
Epoch 156/250
ccuracy: 0.9491
Epoch 157/250
ccuracy: 0.9462
Epoch 158/250
ccuracy: 0.9502
Epoch 159/250
ccuracy: 0.9496
Epoch 160/250
ccuracy: 0.9502
Epoch 161/250
ccuracy: 0.9491
Epoch 162/250
ccuracy: 0.9485
Epoch 163/250
ccuracy: 0.9496
Epoch 164/250
ccuracy: 0.9502
Epoch 165/250
ccuracy: 0.9491
Epoch 166/250
ccuracy: 0.9502
Epoch 167/250
ccuracy: 0.9485
Epoch 168/250
ccuracy: 0.9479
```

```
Epoch 169/250
ccuracy: 0.9473
Epoch 170/250
ccuracy: 0.8724
Epoch 171/250
ccuracy: 0.9336
Epoch 172/250
ccuracy: 0.9456
Epoch 173/250
ccuracy: 0.9284
Epoch 174/250
ccuracy: 0.9416
Epoch 175/250
ccuracy: 0.9479
Epoch 176/250
ccuracy: 0.9496
Epoch 177/250
ccuracy: 0.9508
Epoch 178/250
ccuracy: 0.9496
Epoch 179/250
ccuracy: 0.9508
Epoch 180/250
ccuracy: 0.9508
Epoch 181/250
ccuracy: 0.9491
Epoch 182/250
ccuracy: 0.9502
```

```
Epoch 183/250
ccuracy: 0.9479
Epoch 184/250
ccuracy: 0.9325
Epoch 185/250
ccuracy: 0.9176
Epoch 186/250
ccuracy: 0.9159
Epoch 187/250
ccuracy: 0.9336
Epoch 188/250
ccuracy: 0.9393
Epoch 189/250
ccuracy: 0.9473
Epoch 190/250
ccuracy: 0.9491
Epoch 191/250
ccuracy: 0.9479
Epoch 192/250
ccuracy: 0.9445
Epoch 193/250
ccuracy: 0.9491
Epoch 194/250
ccuracy: 0.9416
Epoch 195/250
ccuracy: 0.9422
Epoch 196/250
ccuracy: 0.9428
```

```
Epoch 197/250
ccuracy: 0.9433
Epoch 198/250
ccuracy: 0.9479
Epoch 199/250
ccuracy: 0.9468
Epoch 200/250
ccuracy: 0.9479
Epoch 201/250
ccuracy: 0.9485
Epoch 202/250
ccuracy: 0.9491
Epoch 203/250
ccuracy: 0.9473
Epoch 204/250
ccuracy: 0.9473
Epoch 205/250
ccuracy: 0.9473
Epoch 206/250
ccuracy: 0.9256
Epoch 207/250
ccuracy: 0.9290
Epoch 208/250
ccuracy: 0.9382
Epoch 209/250
ccuracy: 0.9416
Epoch 210/250
ccuracy: 0.9416
```

```
Epoch 211/250
ccuracy: 0.9273
Epoch 212/250
ccuracy: 0.9376
Epoch 213/250
ccuracy: 0.9164
Epoch 214/250
ccuracy: 0.9542
Epoch 215/250
ccuracy: 0.9479
Epoch 216/250
ccuracy: 0.9044
Epoch 217/250
ccuracy: 0.9250
Epoch 218/250
ccuracy: 0.9496
Epoch 219/250
ccuracy: 0.9416
Epoch 220/250
ccuracy: 0.9262
Epoch 221/250
ccuracy: 0.9456
Epoch 222/250
ccuracy: 0.9439
Epoch 223/250
ccuracy: 0.9405
Epoch 224/250
ccuracy: 0.9468
```

```
Epoch 225/250
ccuracy: 0.9456
Epoch 226/250
ccuracy: 0.9479
Epoch 227/250
ccuracy: 0.9496
Epoch 228/250
ccuracy: 0.9439
Epoch 229/250
ccuracy: 0.9262
Epoch 230/250
ccuracy: 0.9296
Epoch 231/250
ccuracy: 0.9279
Epoch 232/250
ccuracy: 0.9479
Epoch 233/250
ccuracy: 0.9399
Epoch 234/250
ccuracy: 0.9479
Epoch 235/250
ccuracy: 0.9479
Epoch 236/250
ccuracy: 0.9450
Epoch 237/250
ccuracy: 0.9256
Epoch 238/250
ccuracy: 0.9548
```

```
Epoch 239/250
ccuracy: 0.9473
Epoch 240/250
ccuracy: 0.9565
Epoch 241/250
ccuracy: 0.9559
Epoch 242/250
ccuracy: 0.9536
Epoch 243/250
ccuracy: 0.9554
Epoch 244/250
ccuracy: 0.9536
Epoch 245/250
ccuracy: 0.9531
Epoch 246/250
ccuracy: 0.9542
Epoch 247/250
ccuracy: 0.9554
Epoch 248/250
ccuracy: 0.9347
Epoch 249/250
ccuracy: 0.9227
Epoch 250/250
ccuracy: 0.9439
ANN Model Saved
C:\Users\vidit\AppData\Roaming\Python\Python310\site-packages\keras\src\engine\training.py:3079: UserWarning: You ar
e saving your model as an HDF5 file via `model.save()`. This file format is considered legacy. We recommend using in
stead the native Keras format, e.g. `model.save('my model.keras')`.
```

saving api.save model(

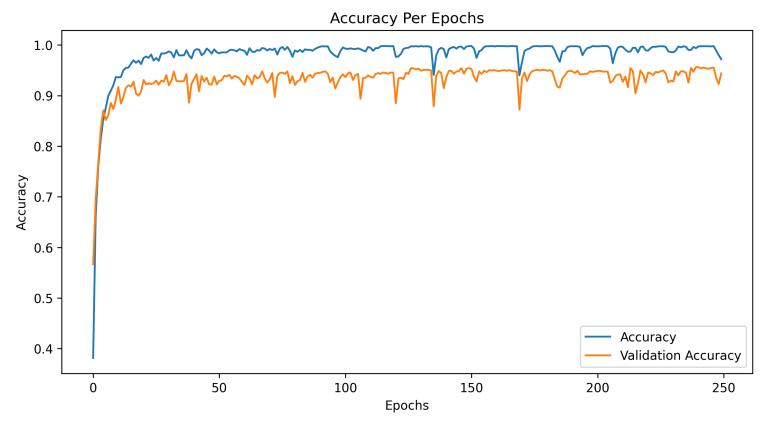
Results

```
In [25]: plt.figure(figsize=(10, 5), dpi=300)
    plt.plot(train_hist_m1[["loss", "val_loss"]])
    plt.legend(["Loss", "Validation Loss"])
    plt.title("Loss Per Epochs")
    plt.xlabel("Epochs")
    plt.ylabel("Loss")
    plt.show()
```

Loss Per Epochs 2.00 Loss Validation Loss 1.75 1.50 1.25 S 1.00 0.75 0.50 0.25 0.00 50 100 150 200 250 0 **Epochs**

```
In [26]: plt.figure(figsize=(10, 5), dpi=300)
    plt.plot(train_hist_m1[["accuracy", "val_accuracy"]])
```

```
plt.legend(["Accuracy", "Validation Accuracy"])
plt.title("Accuracy Per Epochs")
plt.xlabel("Epochs")
plt.ylabel("Accuracy")
plt.show()
```



```
log = log.append(log_entry)

C:\Users\vidit\AppData\Local\Temp\ipykernel_46916\3928430697.py:8: FutureWarning: The frame.append method is depreca
ted and will be removed from pandas in a future version. Use pandas.concat instead.
    log = log.append(log_entry)
```

ANN Prediction Function

```
In [28]: # function to predict the feature
         def ANN Prediction(file name):
             # Load the audio file
             audio_data, sample_rate = librosa.load(file_name, res_type="kaiser_fast")
             # get the feature
             feature = librosa.feature.mfcc(y=audio data, sr=sample rate, n mfcc=128)
             # scale the features
             feature scaled = np.mean(feature.T, axis=0)
             # array of features
             prediction_feature = np.array([feature_scaled])
             # get the id of label using argmax
             predicted vector = np.argmax(ANN Model.predict(prediction feature), axis=-1)
             # get the class label from class id
             predicted_class = le.inverse_transform(predicted_vector)
             # display the result
             print("ANN has predicted the class as --> ", predicted class[0])
```

Testing the Model on Sample audio

Model 2 - CNN1D

Preprocessing

```
In [30]: xTrainval, xTest, yTrainval, yTest = train test split(
             X, Y, test_size=0.1, stratify=y, random_state=387
         xTrain, xvalid, yTrain, yvalid = train_test_split(
             xTrainval, yTrainval, test_size=0.2, stratify=yTrainval, random_state=387
         print("\nNumber of samples for Train set :", xTrain.shape[0])
         print("Number of samples for Validation set :", xvalid.shape[0])
         print("Number of samples for Test set :", xTest.shape[0])
         xTrain = np.expand dims(xTrain, axis=2)
         xvalid = np.expand_dims(xvalid, axis=2)
         print("Shape of X Train", xTrain.shape)
         print("Shape of X Test", xTest.shape)
         Number of samples for Train set : 6286
         Number of samples for Validation set : 1572
         Number of samples for Test set: 874
         Shape of X Train (6286, 128, 1)
         Shape of X Test (874, 128)
```

Building the CNN1D Model

```
5,
        strides=1,
        padding="same",
        activation="relu",
        input_shape=(xTrain.shape[1], 1),
CNN1D Model.add(BatchNormalization())
CNN1D_Model.add(MaxPooling1D(3, strides=2, padding="same"))
CNN1D Model.add(Conv1D(256, 5, strides=1, padding="same", activation="relu"))
CNN1D Model.add(Dropout(0.3))
CNN1D_Model.add(MaxPooling1D(3, strides=2, padding="same"))
CNN1D Model.add(Conv1D(128, 5, strides=1, padding="same", activation="relu"))
CNN1D Model.add(Dropout(0.3))
CNN1D_Model.add(MaxPooling1D(3, strides=2, padding="same"))
CNN1D_Model.add(Conv1D( , 5, strides=1, padding="same", activation="relu"))
CNN1D Model.add(Dropout(0.3))
CNN1D_Model.add(MaxPooling1D(3, strides=2, padding="same"))
CNN1D_Model.add(Flatten())
CNN1D Model.add(Dense(units=1024, activation="relu"))
CNN1D Model.add(Dropout(0.3))
CNN1D_Model.add(Dense(units=10, activation="softmax"))
CNN1D Model.summary()
```

Model: "sequential_1"

/None 120 256)	
(None, 128, 256)	1536
(None, 128, 256)	1024
(None, 64, 256)	0
(None, 64, 256)	327936
(None, 64, 256)	0
(None, 32, 256)	0
(None, 32, 128)	163968
(None, 32, 128)	0
(None, 16, 128)	0
(None, 16, 64)	41024
(None, 16, 64)	0
(None, 8, 64)	0
(None, 512)	0
(None, 1024)	525312
(None, 1024)	0
(None, 10)	10250
	(None, 128, 256) (None, 64, 256) (None, 64, 256) (None, 64, 256) (None, 32, 256) (None, 32, 128) (None, 32, 128) (None, 16, 128) (None, 16, 64) (None, 16, 64) (None, 8, 64) (None, 512) (None, 1024) (None, 109)

```
Total params: 1071050 (4.09 MB)
Trainable params: 1070538 (4.08 MB)
Non-trainable params: 512 (2.00 KB)
```

Compiling the Model

Fitting the Model

```
In [33]: t0 = time.time()

CNN1D_Results = CNN1D_Model.fit(
    xTrain, yTrain, batch_size=64, epochs=250, validation_data=(xvalid, yvalid)
)

CNN1D_Model.save("Model2.h5")
print("CNN1D Model Saved")
train_hist_m2 = pd.DataFrame(CNN1D_Results.history)
train_m2 = round(time.time() - t0, 3)
```

```
Epoch 1/250
99/99 [===========] - 15s 139ms/step - loss: 1.8370 - accuracy: 0.3446 - val loss: 1.6371 - val a
ccuracy: 0.4561
Epoch 2/250
99/99 [===========] - 13s 133ms/step - loss: 1.2854 - accuracy: 0.5495 - val loss: 1.2840 - val a
ccuracy: 0.6457
Epoch 3/250
ccuracy: 0.7131
Epoch 4/250
ccuracy: 0.7424
Epoch 5/250
ccuracy: 0.7742
Epoch 6/250
99/99 [===========] - 14s 138ms/step - loss: 0.7776 - accuracy: 0.7297 - val loss: 0.7613 - val a
ccuracy: 0.7990
Epoch 7/250
ccuracy: 0.8149
Epoch 8/250
99/99 [============] - 13s 132ms/step - loss: 0.6075 - accuracy: 0.7940 - val loss: 0.6252 - val a
ccuracy: 0.8314
Epoch 9/250
ccuracy: 0.8422
Epoch 10/250
ccuracy: 0.8384
Epoch 11/250
ccuracy: 0.8505
Epoch 12/250
99/99 [===========] - 14s 137ms/step - loss: 0.4751 - accuracy: 0.8358 - val loss: 0.5264 - val a
ccuracy: 0.8601
Epoch 13/250
ccuracy: 0.8550
Epoch 14/250
ccuracy: 0.8492
```

```
Epoch 15/250
99/99 [===========] - 13s 136ms/step - loss: 0.3845 - accuracy: 0.8667 - val loss: 0.4773 - val a
ccuracy: 0.8645
Epoch 16/250
99/99 [===========] - 13s 133ms/step - loss: 0.3818 - accuracy: 0.8711 - val loss: 0.4608 - val a
ccuracy: 0.8626
Epoch 17/250
ccuracy: 0.8556
Epoch 18/250
ccuracy: 0.8798
Epoch 19/250
99/99 [===========] - 14s 144ms/step - loss: 0.3345 - accuracy: 0.8863 - val loss: 0.4791 - val a
ccuracy: 0.8588
Epoch 20/250
99/99 [===========] - 14s 139ms/step - loss: 0.3599 - accuracy: 0.8737 - val loss: 0.4615 - val a
ccuracy: 0.8702
Epoch 21/250
99/99 [==========] - 13s 130ms/step - loss: 0.3163 - accuracy: 0.8910 - val loss: 0.3847 - val a
ccuracy: 0.8976
Epoch 22/250
99/99 [===========] - 13s 133ms/step - loss: 0.2999 - accuracy: 0.8988 - val loss: 0.4058 - val a
ccuracy: 0.8849
Epoch 23/250
99/99 [===========] - 14s 138ms/step - loss: 0.2972 - accuracy: 0.9003 - val loss: 0.3677 - val a
ccuracy: 0.8963
Epoch 24/250
ccuracy: 0.8925
Epoch 25/250
ccuracy: 0.8912
Epoch 26/250
99/99 [===========] - 13s 130ms/step - loss: 0.2716 - accuracy: 0.9063 - val loss: 0.3502 - val a
ccuracy: 0.8982
Epoch 27/250
ccuracy: 0.8830
Epoch 28/250
ccuracy: 0.8798
```

```
Epoch 29/250
99/99 [===========] - 13s 133ms/step - loss: 0.2478 - accuracy: 0.9170 - val loss: 0.3664 - val a
ccuracy: 0.8855
Epoch 30/250
99/99 [===========] - 13s 131ms/step - loss: 0.2296 - accuracy: 0.9246 - val loss: 0.3406 - val a
ccuracy: 0.9008
Epoch 31/250
ccuracy: 0.9020
Epoch 32/250
ccuracy: 0.8976
Epoch 33/250
ccuracy: 0.9008
Epoch 34/250
99/99 [===========] - 14s 138ms/step - loss: 0.2466 - accuracy: 0.9149 - val loss: 0.3232 - val a
ccuracy: 0.9039
Epoch 35/250
ccuracy: 0.9071
Epoch 36/250
ccuracy: 0.8944
Epoch 37/250
ccuracy: 0.9027
Epoch 38/250
ccuracy: 0.9027
Epoch 39/250
ccuracy: 0.9122
Epoch 40/250
99/99 [===========] - 14s 141ms/step - loss: 0.1818 - accuracy: 0.9365 - val loss: 0.3378 - val a
ccuracy: 0.9059
Epoch 41/250
ccuracy: 0.8957
Epoch 42/250
ccuracy: 0.8963
```

```
Epoch 43/250
99/99 [===========] - 13s 136ms/step - loss: 0.2147 - accuracy: 0.9295 - val loss: 0.3091 - val a
ccuracy: 0.9033
Epoch 44/250
99/99 [===========] - 14s 140ms/step - loss: 0.2497 - accuracy: 0.9213 - val loss: 0.3454 - val a
ccuracy: 0.8976
Epoch 45/250
ccuracy: 0.9192
Epoch 46/250
ccuracy: 0.9148
Epoch 47/250
ccuracy: 0.9167
Epoch 48/250
99/99 [===========] - 13s 134ms/step - loss: 0.1640 - accuracy: 0.9464 - val loss: 0.2979 - val a
ccuracy: 0.9109
Epoch 49/250
ccuracy: 0.9090
Epoch 50/250
99/99 [===========] - 13s 135ms/step - loss: 0.1738 - accuracy: 0.9432 - val loss: 0.3020 - val a
ccuracy: 0.9141
Epoch 51/250
ccuracy: 0.9243
Epoch 52/250
ccuracy: 0.9173
Epoch 53/250
ccuracy: 0.9154
Epoch 54/250
99/99 [===========] - 13s 134ms/step - loss: 0.1589 - accuracy: 0.9494 - val loss: 0.3252 - val a
ccuracy: 0.9039
Epoch 55/250
ccuracy: 0.9059
Epoch 56/250
ccuracy: 0.9052
```

```
Epoch 57/250
99/99 [===========] - 14s 142ms/step - loss: 0.1690 - accuracy: 0.9467 - val loss: 0.3287 - val a
ccuracy: 0.9097
Epoch 58/250
99/99 [===========] - 13s 133ms/step - loss: 0.1897 - accuracy: 0.9427 - val loss: 0.3430 - val a
ccuracy: 0.9001
Epoch 59/250
ccuracy: 0.9109
Epoch 60/250
ccuracy: 0.9109
Epoch 61/250
ccuracy: 0.9198
Epoch 62/250
99/99 [===========] - 14s 138ms/step - loss: 0.1392 - accuracy: 0.9550 - val loss: 0.3374 - val a
ccuracy: 0.8995
Epoch 63/250
99/99 [==========] - 14s 142ms/step - loss: 0.1460 - accuracy: 0.9529 - val loss: 0.3093 - val a
ccuracy: 0.9033
Epoch 64/250
99/99 [===========] - 13s 133ms/step - loss: 0.1696 - accuracy: 0.9526 - val loss: 0.3041 - val a
ccuracy: 0.9128
Epoch 65/250
ccuracy: 0.9167
Epoch 66/250
ccuracy: 0.9116
Epoch 67/250
ccuracy: 0.9198
Epoch 68/250
99/99 [===========] - 14s 138ms/step - loss: 0.2206 - accuracy: 0.9322 - val loss: 0.2963 - val a
ccuracy: 0.9167
Epoch 69/250
ccuracy: 0.9141
Epoch 70/250
ccuracy: 0.9179
```

```
Epoch 71/250
99/99 [===========] - 14s 142ms/step - loss: 0.1401 - accuracy: 0.9574 - val loss: 0.3131 - val a
ccuracy: 0.9052
Epoch 72/250
99/99 [===========] - 14s 141ms/step - loss: 0.1865 - accuracy: 0.9440 - val loss: 0.2580 - val a
ccuracy: 0.9262
Epoch 73/250
ccuracy: 0.9186
Epoch 74/250
ccuracy: 0.9294
Epoch 75/250
ccuracy: 0.9128
Epoch 76/250
99/99 [===========] - 14s 140ms/step - loss: 0.1481 - accuracy: 0.9566 - val loss: 0.3370 - val a
ccuracy: 0.9103
Epoch 77/250
ccuracy: 0.9014
Epoch 78/250
ccuracy: 0.9154
Epoch 79/250
ccuracy: 0.9160
Epoch 80/250
ccuracy: 0.9268
Epoch 81/250
ccuracy: 0.9122
Epoch 82/250
99/99 [===========] - 14s 139ms/step - loss: 0.1610 - accuracy: 0.9532 - val loss: 0.3036 - val a
ccuracy: 0.9090
Epoch 83/250
ccuracy: 0.9173
Epoch 84/250
ccuracy: 0.9116
```

```
Epoch 85/250
99/99 [===========] - 14s 141ms/step - loss: 0.1307 - accuracy: 0.9590 - val loss: 0.2928 - val a
ccuracy: 0.9243
Epoch 86/250
99/99 [===========] - 16s 161ms/step - loss: 0.1285 - accuracy: 0.9617 - val loss: 0.2772 - val a
ccuracy: 0.9173
Epoch 87/250
ccuracy: 0.9211
Epoch 88/250
ccuracy: 0.9192
Epoch 89/250
ccuracy: 0.9160
Epoch 90/250
99/99 [===========] - 14s 137ms/step - loss: 0.1242 - accuracy: 0.9602 - val loss: 0.2993 - val a
ccuracy: 0.9027
Epoch 91/250
99/99 [===========] - 13s 136ms/step - loss: 0.1551 - accuracy: 0.9531 - val loss: 0.2833 - val a
ccuracy: 0.9173
Epoch 92/250
99/99 [===========] - 13s 133ms/step - loss: 0.1156 - accuracy: 0.9620 - val loss: 0.3148 - val a
ccuracy: 0.9097
Epoch 93/250
ccuracy: 0.9243
Epoch 94/250
ccuracy: 0.9097
Epoch 95/250
ccuracy: 0.9275
Epoch 96/250
99/99 [===========] - 13s 131ms/step - loss: 0.1264 - accuracy: 0.9599 - val loss: 0.2546 - val a
ccuracy: 0.9281
Epoch 97/250
ccuracy: 0.9154
Epoch 98/250
ccuracy: 0.9141
```

```
Epoch 99/250
99/99 [===========] - 13s 132ms/step - loss: 0.1588 - accuracy: 0.9542 - val loss: 0.2823 - val a
ccuracy: 0.9128
Epoch 100/250
99/99 [===========] - 13s 134ms/step - loss: 0.1730 - accuracy: 0.9518 - val loss: 0.3349 - val a
ccuracy: 0.9078
Epoch 101/250
ccuracy: 0.9275
Epoch 102/250
ccuracy: 0.9218
Epoch 103/250
ccuracy: 0.9160
Epoch 104/250
99/99 [===========] - 14s 138ms/step - loss: 0.1339 - accuracy: 0.9594 - val loss: 0.3116 - val a
ccuracy: 0.9141
Epoch 105/250
ccuracy: 0.9288
Epoch 106/250
99/99 [===========] - 13s 134ms/step - loss: 0.1071 - accuracy: 0.9669 - val loss: 0.2420 - val a
ccuracy: 0.9237
Epoch 107/250
ccuracy: 0.9275
Epoch 108/250
ccuracy: 0.9319
Epoch 109/250
ccuracy: 0.9205
Epoch 110/250
99/99 [===========] - 13s 132ms/step - loss: 0.1292 - accuracy: 0.9633 - val loss: 0.2848 - val a
ccuracy: 0.9154
Epoch 111/250
ccuracy: 0.9173
Epoch 112/250
ccuracy: 0.9179
```

```
Epoch 113/250
99/99 [===========] - 13s 133ms/step - loss: 0.1649 - accuracy: 0.9542 - val loss: 0.3224 - val a
ccuracy: 0.9135
Epoch 114/250
99/99 [===========] - 13s 128ms/step - loss: 0.1557 - accuracy: 0.9585 - val loss: 0.3209 - val a
ccuracy: 0.9154
Epoch 115/250
ccuracy: 0.9135
Epoch 116/250
ccuracy: 0.9071
Epoch 117/250
ccuracy: 0.9198
Epoch 118/250
99/99 [==========] - 13s 132ms/step - loss: 0.1609 - accuracy: 0.9539 - val loss: 0.2736 - val a
ccuracy: 0.9243
Epoch 119/250
ccuracy: 0.9294
Epoch 120/250
99/99 [===========] - 13s 130ms/step - loss: 0.1435 - accuracy: 0.9583 - val loss: 0.2953 - val a
ccuracy: 0.9103
Epoch 121/250
ccuracy: 0.9218
Epoch 122/250
ccuracy: 0.9307
Epoch 123/250
ccuracy: 0.9205
Epoch 124/250
99/99 [===========] - 13s 130ms/step - loss: 0.1060 - accuracy: 0.9707 - val loss: 0.2779 - val a
ccuracy: 0.9230
Epoch 125/250
ccuracy: 0.9262
Epoch 126/250
ccuracy: 0.9256
```

```
Epoch 127/250
99/99 [==========] - 13s 133ms/step - loss: 0.1308 - accuracy: 0.9636 - val loss: 0.2562 - val a
ccuracy: 0.9275
Epoch 128/250
99/99 [===========] - 13s 130ms/step - loss: 0.1194 - accuracy: 0.9613 - val loss: 0.3438 - val a
ccuracy: 0.9097
Epoch 129/250
ccuracy: 0.9262
Epoch 130/250
ccuracy: 0.9167
Epoch 131/250
ccuracy: 0.9192
Epoch 132/250
99/99 [===========] - 13s 132ms/step - loss: 0.1002 - accuracy: 0.9687 - val loss: 0.2827 - val a
ccuracv: 0.9281
Epoch 133/250
ccuracy: 0.9268
Epoch 134/250
99/99 [===========] - 13s 134ms/step - loss: 0.1129 - accuracy: 0.9683 - val loss: 0.2920 - val a
ccuracy: 0.9192
Epoch 135/250
ccuracy: 0.9198
Epoch 136/250
ccuracy: 0.9173
Epoch 137/250
ccuracy: 0.9224
Epoch 138/250
99/99 [===========] - 20s 202ms/step - loss: 0.1189 - accuracy: 0.9658 - val loss: 0.2845 - val a
ccuracy: 0.9256
Epoch 139/250
ccuracy: 0.9262
Epoch 140/250
ccuracy: 0.9256
```

```
Epoch 141/250
99/99 [===========] - 19s 188ms/step - loss: 0.1158 - accuracy: 0.9701 - val loss: 0.2612 - val a
ccuracy: 0.9256
Epoch 142/250
99/99 [==========] - 13s 133ms/step - loss: 0.1200 - accuracy: 0.9663 - val loss: 0.2790 - val a
ccuracy: 0.9268
Epoch 143/250
ccuracy: 0.9154
Epoch 144/250
ccuracy: 0.9122
Epoch 145/250
ccuracy: 0.9256
Epoch 146/250
99/99 [===========] - 19s 188ms/step - loss: 0.1243 - accuracy: 0.9663 - val loss: 0.2920 - val a
ccuracy: 0.9186
Epoch 147/250
99/99 [===========] - 13s 135ms/step - loss: 0.0950 - accuracy: 0.9701 - val loss: 0.3005 - val a
ccuracy: 0.9224
Epoch 148/250
99/99 [===========] - 13s 132ms/step - loss: 0.1094 - accuracy: 0.9682 - val loss: 0.2817 - val a
ccuracy: 0.9237
Epoch 149/250
ccuracy: 0.9332
Epoch 150/250
ccuracy: 0.9275
Epoch 151/250
ccuracy: 0.9332
Epoch 152/250
ccuracy: 0.9268
Epoch 153/250
ccuracy: 0.9243
Epoch 154/250
ccuracy: 0.9154
```

```
Epoch 155/250
99/99 [===========] - 13s 132ms/step - loss: 0.1204 - accuracy: 0.9685 - val loss: 0.2657 - val a
ccuracy: 0.9230
Epoch 156/250
99/99 [===========] - 13s 133ms/step - loss: 0.1062 - accuracy: 0.9690 - val loss: 0.2623 - val a
ccuracy: 0.9256
Epoch 157/250
ccuracy: 0.9173
Epoch 158/250
ccuracy: 0.9135
Epoch 159/250
ccuracy: 0.9256
Epoch 160/250
99/99 [==========] - 13s 132ms/step - loss: 0.0881 - accuracy: 0.9734 - val loss: 0.2574 - val a
ccuracy: 0.9256
Epoch 161/250
ccuracy: 0.9294
Epoch 162/250
99/99 [===========] - 13s 133ms/step - loss: 0.0855 - accuracy: 0.9773 - val loss: 0.2340 - val a
ccuracy: 0.9427
Epoch 163/250
ccuracy: 0.9249
Epoch 164/250
ccuracy: 0.9319
Epoch 165/250
ccuracy: 0.9313
Epoch 166/250
99/99 [===========] - 13s 133ms/step - loss: 0.1234 - accuracy: 0.9690 - val loss: 0.2707 - val a
ccuracy: 0.9281
Epoch 167/250
ccuracy: 0.9313
Epoch 168/250
ccuracy: 0.9319
```

```
Epoch 169/250
99/99 [==========] - 13s 131ms/step - loss: 0.1392 - accuracy: 0.9666 - val loss: 0.2553 - val a
ccuracy: 0.9256
Epoch 170/250
99/99 [===========] - 13s 135ms/step - loss: 0.0784 - accuracy: 0.9739 - val loss: 0.2636 - val a
ccuracy: 0.9281
Epoch 171/250
ccuracy: 0.9326
Epoch 172/250
ccuracy: 0.9256
Epoch 173/250
ccuracy: 0.9300
Epoch 174/250
99/99 [===========] - 13s 135ms/step - loss: 0.1044 - accuracy: 0.9734 - val loss: 0.2846 - val a
ccuracy: 0.9249
Epoch 175/250
99/99 [==========] - 13s 131ms/step - loss: 0.1492 - accuracy: 0.9620 - val loss: 0.2644 - val a
ccuracy: 0.9256
Epoch 176/250
99/99 [===========] - 13s 130ms/step - loss: 0.1686 - accuracy: 0.9586 - val loss: 0.3330 - val a
ccuracy: 0.9230
Epoch 177/250
ccuracy: 0.9230
Epoch 178/250
ccuracy: 0.9237
Epoch 179/250
ccuracy: 0.9249
Epoch 180/250
99/99 [===========] - 13s 130ms/step - loss: 0.1003 - accuracy: 0.9714 - val loss: 0.2792 - val a
ccuracy: 0.9345
Epoch 181/250
ccuracy: 0.9300
Epoch 182/250
ccuracy: 0.9288
```

```
Epoch 183/250
99/99 [===========] - 13s 130ms/step - loss: 0.1101 - accuracy: 0.9715 - val loss: 0.2668 - val a
ccuracy: 0.9307
Epoch 184/250
99/99 [===========] - 13s 131ms/step - loss: 0.0903 - accuracy: 0.9750 - val loss: 0.3044 - val a
ccuracy: 0.9281
Epoch 185/250
ccuracy: 0.9313
Epoch 186/250
ccuracy: 0.9294
Epoch 187/250
99/99 [===========] - 13s 129ms/step - loss: 0.1274 - accuracy: 0.9658 - val loss: 0.3131 - val a
ccuracy: 0.9224
Epoch 188/250
99/99 [==========] - 13s 130ms/step - loss: 0.1297 - accuracy: 0.9669 - val loss: 0.3446 - val a
ccuracy: 0.9205
Epoch 189/250
ccuracy: 0.9179
Epoch 190/250
99/99 [===========] - 18s 183ms/step - loss: 0.0929 - accuracy: 0.9730 - val loss: 0.2907 - val a
ccuracy: 0.9211
Epoch 191/250
ccuracy: 0.9224
Epoch 192/250
ccuracy: 0.9211
Epoch 193/250
ccuracy: 0.9281
Epoch 194/250
ccuracy: 0.9179
Epoch 195/250
ccuracy: 0.9307
Epoch 196/250
ccuracy: 0.9249
```

```
Epoch 197/250
99/99 [===========] - 13s 131ms/step - loss: 0.1129 - accuracy: 0.9683 - val loss: 0.3095 - val a
ccuracy: 0.9218
Epoch 198/250
99/99 [===========] - 18s 183ms/step - loss: 0.1188 - accuracy: 0.9744 - val loss: 0.3391 - val a
ccuracy: 0.9160
Epoch 199/250
ccuracy: 0.9300
Epoch 200/250
ccuracy: 0.9249
Epoch 201/250
ccuracy: 0.9288
Epoch 202/250
99/99 [===========] - 18s 185ms/step - loss: 0.1255 - accuracy: 0.9715 - val loss: 0.3062 - val a
ccuracy: 0.9211
Epoch 203/250
99/99 [==========] - 13s 135ms/step - loss: 0.1144 - accuracy: 0.9699 - val loss: 0.2753 - val a
ccuracy: 0.9249
Epoch 204/250
ccuracy: 0.9262
Epoch 205/250
ccuracy: 0.9192
Epoch 206/250
ccuracy: 0.9211
Epoch 207/250
ccuracy: 0.9326
Epoch 208/250
ccuracy: 0.9288
Epoch 209/250
ccuracy: 0.9281
Epoch 210/250
ccuracy: 0.9383
```

```
Epoch 211/250
99/99 [===========] - 13s 131ms/step - loss: 0.0835 - accuracy: 0.9777 - val loss: 0.2978 - val a
ccuracy: 0.9256
Epoch 212/250
99/99 [==========] - 13s 135ms/step - loss: 0.1192 - accuracy: 0.9699 - val loss: 0.2730 - val a
ccuracy: 0.9256
Epoch 213/250
ccuracy: 0.9332
Epoch 214/250
ccuracy: 0.9345
Epoch 215/250
ccuracy: 0.9389
Epoch 216/250
99/99 [===========] - 13s 130ms/step - loss: 0.1206 - accuracy: 0.9717 - val loss: 0.2550 - val a
ccuracv: 0.9383
Epoch 217/250
99/99 [===========] - 15s 147ms/step - loss: 0.0961 - accuracy: 0.9755 - val loss: 0.3389 - val a
ccuracy: 0.9256
Epoch 218/250
ccuracy: 0.9300
Epoch 219/250
ccuracy: 0.9256
Epoch 220/250
ccuracy: 0.9249
Epoch 221/250
ccuracy: 0.9109
Epoch 222/250
99/99 [==========] - 13s 129ms/step - loss: 0.0966 - accuracy: 0.9730 - val loss: 0.3226 - val a
ccuracy: 0.9128
Epoch 223/250
ccuracy: 0.9307
Epoch 224/250
ccuracy: 0.9179
```

```
Epoch 225/250
99/99 [===========] - 13s 132ms/step - loss: 0.0868 - accuracy: 0.9757 - val loss: 0.2783 - val a
ccuracy: 0.9326
Epoch 226/250
99/99 [===========] - 13s 130ms/step - loss: 0.1081 - accuracy: 0.9739 - val loss: 0.3118 - val a
ccuracy: 0.9179
Epoch 227/250
ccuracy: 0.9224
Epoch 228/250
ccuracy: 0.9377
Epoch 229/250
ccuracy: 0.9415
Epoch 230/250
99/99 [===========] - 13s 130ms/step - loss: 0.0932 - accuracy: 0.9757 - val loss: 0.3078 - val a
ccuracy: 0.9307
Epoch 231/250
99/99 [===========] - 18s 181ms/step - loss: 0.1987 - accuracy: 0.9570 - val loss: 0.3128 - val a
ccuracy: 0.9205
Epoch 232/250
ccuracy: 0.9224
Epoch 233/250
ccuracy: 0.9332
Epoch 234/250
ccuracy: 0.9313
Epoch 235/250
ccuracy: 0.9319
Epoch 236/250
ccuracy: 0.9345
Epoch 237/250
ccuracy: 0.9281
Epoch 238/250
ccuracy: 0.9319
```

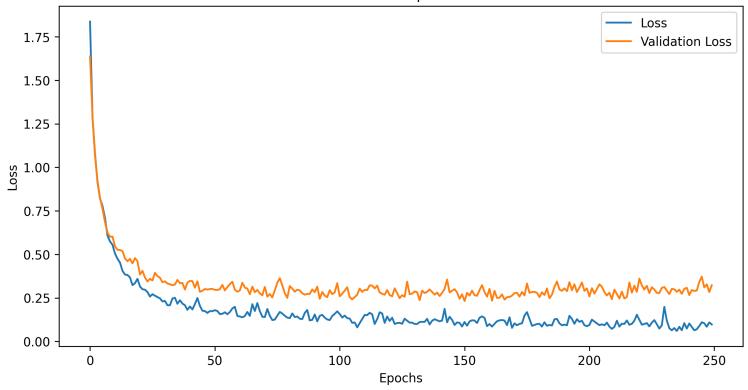
```
Epoch 239/250
99/99 [===========] - 13s 130ms/step - loss: 0.1053 - accuracy: 0.9747 - val loss: 0.3025 - val a
ccuracy: 0.9338
Epoch 240/250
99/99 [===========] - 13s 131ms/step - loss: 0.0742 - accuracy: 0.9788 - val loss: 0.3084 - val a
ccuracy: 0.9332
Epoch 241/250
ccuracy: 0.9281
Epoch 242/250
ccuracy: 0.9288
Epoch 243/250
ccuracy: 0.9313
Epoch 244/250
99/99 [===========] - 13s 130ms/step - loss: 0.0702 - accuracy: 0.9812 - val loss: 0.2925 - val a
ccuracv: 0.9288
Epoch 245/250
ccuracy: 0.9237
Epoch 246/250
ccuracy: 0.9198
Epoch 247/250
99/99 [===========] - 13s 132ms/step - loss: 0.1046 - accuracy: 0.9734 - val loss: 0.3102 - val a
ccuracy: 0.9294
Epoch 248/250
ccuracy: 0.9281
Epoch 249/250
ccuracy: 0.9268
Epoch 250/250
99/99 [===========] - 13s 132ms/step - loss: 0.0980 - accuracy: 0.9744 - val loss: 0.3230 - val a
ccuracy: 0.9205
CNN1D Model Saved
C:\Users\vidit\AppData\Roaming\Python\Python310\site-packages\keras\src\engine\training.py:3079: UserWarning: You ar
e saving your model as an HDF5 file via `model.save()`. This file format is considered legacy. We recommend using in
stead the native Keras format, e.g. `model.save('my model.keras')`.
```

saving_api.save_model(

Results

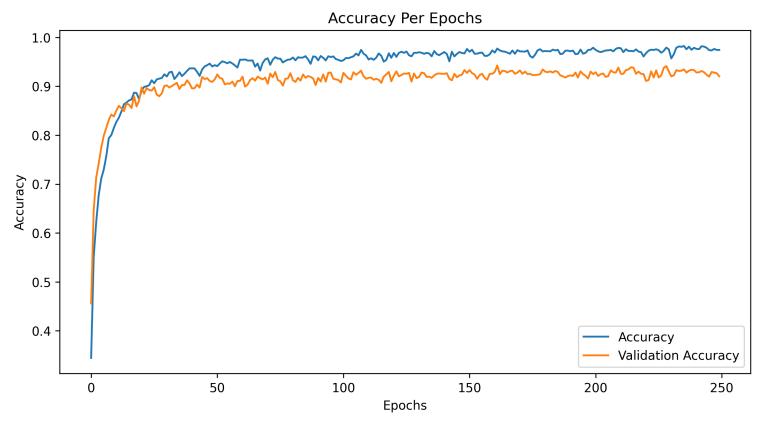
```
In [34]: plt.figure(figsize=(10, 5), dpi=300)
    plt.plot(train_hist_m2[["loss", "val_loss"]])
    plt.legend(["Loss", "Validation Loss"])
    plt.title("Loss Per Epochs")
    plt.xlabel("Epochs")
    plt.ylabel("Loss")
    plt.show()
```

Loss Per Epochs



```
In [35]: plt.figure(figsize=(10, 5), dpi=300)
    plt.plot(train_hist_m2[["accuracy", "val_accuracy"]])
```

```
plt.legend(["Accuracy", "Validation Accuracy"])
plt.title("Accuracy Per Epochs")
plt.xlabel("Epochs")
plt.ylabel("Accuracy")
plt.show()
```



```
log = log.append(log_entry)

C:\Users\vidit\AppData\Local\Temp\ipykernel_46916\3300022719.py:8: FutureWarning: The frame.append method is depreca
ted and will be removed from pandas in a future version. Use pandas.concat instead.
    log = log.append(log_entry)
```

CNN1D Prediction Function

```
In [37]: # function to predict the feature
         def CNN1D Prediction(file name):
             # Load the audio file
             audio_data, sample_rate = librosa.load(file_name, res_type="kaiser_fast")
             # get the feature
             feature = librosa.feature.mfcc(y=audio data, sr=sample rate, n mfcc=128)
             # scale the features
             feature scaled = np.mean(feature.T, axis=0)
             # array of features
             prediction_feature = np.array([feature_scaled])
             # expand dims
             final prediction feature = np.expand dims(prediction feature, axis=2)
             # get the id of label using argmax
             predicted_vector = np.argmax(CNN1D_Model.predict(final_prediction_feature), axis=-1)
             # get the class label from class id
             predicted class = le.inverse transform(predicted vector)
             # display the result
             print("CNN1D has predicted the class as --> ", predicted_class[0])
```

Testing the Model on Sample audio

Model 3 - CNN2D

Preprocessing

```
In [40]: xtrain = xTrain.reshape(xTrain.shape[0], 16, 8, 1)
    xtest = xTest.reshape(xTest.shape[0], 16, 8, 1)

print("The Shape of X Train", xtrain.shape)
    print("The Shape of Y Train", yTrain.shape)
    print("The Shape of X Test", xtest.shape)
    print("The Shape of Y Test", yTest.shape)

The Shape of X Train (6286, 16, 8, 1)
    The Shape of Y Train (6286, 10)
    The Shape of Y Test (874, 16, 8, 1)
    The Shape of Y Test (874, 16)
```

Building the CNN2D Model

Model: "sequential_2"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 16, 8, 64)	640
<pre>max_pooling2d (MaxPooling2 D)</pre>	(None, 8, 4, 64)	0
conv2d_1 (Conv2D)	(None, 8, 4, 128)	73856
<pre>max_pooling2d_1 (MaxPoolin g2D)</pre>	(None, 4, 2, 128)	0
dropout_4 (Dropout)	(None, 4, 2, 128)	0
flatten_1 (Flatten)	(None, 1024)	0
dense_9 (Dense)	(None, 1024)	1049600
dense_10 (Dense)	(None, 10)	10250
Total params: 1134346 (4.33 Trainable params: 1134346 (4 Non-trainable params: 0 (0.0	.33 MB)	

Compiling the Model

Fitting the model

```
In [43]: t0 = time.time()
CNN2D_Results = CNN2D_Model.fit(
```

```
xtrain, yTrain, epochs=250, batch_size=50, validation_data=(xtest, yTest)
)

CNN2D_Model.save("Model3.h5")
print("CNN2D Model Saved")
train_hist_m3 = pd.DataFrame(CNN2D_Results.history)
train_m3 = round(time.time() - t0, 3)
```

```
Epoch 1/250
ccuracy: 0.7792
Epoch 2/250
ccuracy: 0.8009
Epoch 3/250
ccuracy: 0.8684
Epoch 4/250
ccuracy: 0.8444
Epoch 5/250
ccuracy: 0.8867
Epoch 6/250
ccuracy: 0.8627
Epoch 7/250
ccuracy: 0.8764
Epoch 8/250
ccuracy: 0.8776
Epoch 9/250
ccuracy: 0.8867
Epoch 10/250
126/126 [=============] - 3s 28ms/step - loss: 0.1110 - accuracy: 0.9617 - val loss: 0.4444 - val a
ccuracy: 0.8982
Epoch 11/250
ccuracy: 0.8799
Epoch 12/250
ccuracy: 0.8879
Epoch 13/250
ccuracy: 0.8764
Epoch 14/250
ccuracy: 0.8787
```

```
Epoch 15/250
ccuracy: 0.9039
Epoch 16/250
ccuracy: 0.9073
Epoch 17/250
ccuracy: 0.8753
Epoch 18/250
ccuracy: 0.8787
Epoch 19/250
ccuracy: 0.8787
Epoch 20/250
ccuracy: 0.9062
Epoch 21/250
ccuracy: 0.9142
Epoch 22/250
ccuracy: 0.8959
Epoch 23/250
ccuracy: 0.9062
Epoch 24/250
ccuracy: 0.8970
Epoch 25/250
ccuracy: 0.8970
Epoch 26/250
ccuracy: 0.9027
Epoch 27/250
ccuracy: 0.9096
Epoch 28/250
ccuracy: 0.8993
```

```
Epoch 29/250
ccuracy: 0.8844
Epoch 30/250
126/126 [============] - 3s 28ms/step - loss: 0.0391 - accuracy: 0.9866 - val loss: 0.4904 - val a
ccuracy: 0.9062
Epoch 31/250
ccuracy: 0.9050
Epoch 32/250
ccuracy: 0.8890
Epoch 33/250
ccuracy: 0.8627
Epoch 34/250
ccuracy: 0.9142
Epoch 35/250
ccuracy: 0.9119
Epoch 36/250
ccuracy: 0.8913
Epoch 37/250
ccuracy: 0.8970
Epoch 38/250
ccuracy: 0.8936
Epoch 39/250
ccuracy: 0.8970
Epoch 40/250
ccuracy: 0.9199
Epoch 41/250
ccuracy: 0.9096
Epoch 42/250
ccuracy: 0.8947
```

```
Epoch 43/250
ccuracy: 0.9050
Epoch 44/250
ccuracy: 0.8970
Epoch 45/250
ccuracy: 0.8982
Epoch 46/250
ccuracy: 0.9073
Epoch 47/250
ccuracy: 0.9142
Epoch 48/250
ccuracy: 0.9085
Epoch 49/250
ccuracy: 0.9027
Epoch 50/250
ccuracy: 0.9073
Epoch 51/250
ccuracy: 0.9050
Epoch 52/250
ccuracy: 0.9027
Epoch 53/250
ccuracy: 0.9130
Epoch 54/250
ccuracy: 0.9142
Epoch 55/250
ccuracy: 0.9165
Epoch 56/250
ccuracy: 0.9096
```

```
Epoch 57/250
ccuracy: 0.9108
Epoch 58/250
ccuracy: 0.9027
Epoch 59/250
ccuracy: 0.9073
Epoch 60/250
ccuracy: 0.8844
Epoch 61/250
ccuracy: 0.9027
Epoch 62/250
ccuracy: 0.8890
Epoch 63/250
ccuracy: 0.8867
Epoch 64/250
ccuracy: 0.9039
Epoch 65/250
ccuracy: 0.9005
Epoch 66/250
ccuracy: 0.9096
Epoch 67/250
ccuracy: 0.9279
Epoch 68/250
126/126 [============] - 3s 28ms/step - loss: 0.0090 - accuracy: 0.9968 - val loss: 0.4823 - val a
ccuracy: 0.9199
Epoch 69/250
ccuracy: 0.9153
Epoch 70/250
ccuracy: 0.9108
```

```
Epoch 71/250
ccuracy: 0.9096
Epoch 72/250
ccuracy: 0.9176
Epoch 73/250
ccuracy: 0.8924
Epoch 74/250
ccuracy: 0.9085
Epoch 75/250
ccuracy: 0.9176
Epoch 76/250
ccuracy: 0.8970
Epoch 77/250
ccuracy: 0.8970
Epoch 78/250
ccuracy: 0.8970
Epoch 79/250
ccuracy: 0.9016
Epoch 80/250
ccuracy: 0.9199
Epoch 81/250
ccuracy: 0.8959
Epoch 82/250
ccuracy: 0.9130
Epoch 83/250
ccuracy: 0.9119
Epoch 84/250
ccuracy: 0.9096
```

```
Epoch 85/250
ccuracy: 0.9199
Epoch 86/250
ccuracy: 0.9199
Epoch 87/250
ccuracy: 0.9291
Epoch 88/250
ccuracy: 0.9199
Epoch 89/250
ccuracy: 0.9211
Epoch 90/250
ccuracy: 0.9199
Epoch 91/250
ccuracy: 0.9165
Epoch 92/250
ccuracy: 0.9130
Epoch 93/250
ccuracy: 0.8799
Epoch 94/250
ccuracy: 0.8696
Epoch 95/250
ccuracy: 0.9085
Epoch 96/250
ccuracy: 0.9245
Epoch 97/250
ccuracy: 0.9073
Epoch 98/250
ccuracy: 0.8913
```

```
Epoch 99/250
ccuracy: 0.9153
Epoch 100/250
ccuracy: 0.9108
Epoch 101/250
ccuracy: 0.9108
Epoch 102/250
ccuracy: 0.9119
Epoch 103/250
ccuracy: 0.8913
Epoch 104/250
ccuracy: 0.8970
Epoch 105/250
ccuracy: 0.9016
Epoch 106/250
ccuracy: 0.8993
Epoch 107/250
ccuracy: 0.9153
Epoch 108/250
ccuracy: 0.9130
Epoch 109/250
ccuracy: 0.9176
Epoch 110/250
ccuracy: 0.9142
Epoch 111/250
ccuracy: 0.8959
Epoch 112/250
ccuracy: 0.8970
```

```
Epoch 113/250
ccuracy: 0.9050
Epoch 114/250
ccuracy: 0.9096
Epoch 115/250
ccuracy: 0.9176
Epoch 116/250
ccuracy: 0.9119
Epoch 117/250
ccuracy: 0.9153
Epoch 118/250
ccuracy: 0.9211
Epoch 119/250
ccuracy: 0.9130
Epoch 120/250
ccuracy: 0.8993
Epoch 121/250
ccuracy: 0.8982
Epoch 122/250
ccuracy: 0.9096
Epoch 123/250
ccuracy: 0.9050
Epoch 124/250
ccuracy: 0.9039
Epoch 125/250
ccuracy: 0.9073
Epoch 126/250
ccuracy: 0.8993
```

```
Epoch 127/250
ccuracy: 0.8993
Epoch 128/250
ccuracy: 0.9050
Epoch 129/250
ccuracy: 0.9142
Epoch 130/250
ccuracy: 0.9027
Epoch 131/250
ccuracy: 0.8879
Epoch 132/250
ccuracy: 0.9027
Epoch 133/250
ccuracy: 0.9165
Epoch 134/250
ccuracy: 0.9062
Epoch 135/250
ccuracy: 0.9096
Epoch 136/250
ccuracy: 0.9130
Epoch 137/250
ccuracy: 0.8947
Epoch 138/250
ccuracy: 0.8947
Epoch 139/250
ccuracy: 0.9039
Epoch 140/250
ccuracy: 0.9153
```

```
Epoch 141/250
ccuracy: 0.9153
Epoch 142/250
ccuracy: 0.9005
Epoch 143/250
ccuracy: 0.9153
Epoch 144/250
ccuracy: 0.9027
Epoch 145/250
ccuracy: 0.9233
Epoch 146/250
126/126 [============= - - 5s 37ms/step - loss: 0.0200 - accuracy: 0.9933 - val loss: 0.7005 - val a
ccuracy: 0.9073
Epoch 147/250
ccuracy: 0.9039
Epoch 148/250
ccuracy: 0.9027
Epoch 149/250
ccuracy: 0.8993
Epoch 150/250
ccuracy: 0.9119
Epoch 151/250
ccuracy: 0.9130
Epoch 152/250
ccuracy: 0.9039
Epoch 153/250
ccuracy: 0.8982
Epoch 154/250
ccuracy: 0.9130
```

```
Epoch 155/250
ccuracy: 0.9176
Epoch 156/250
ccuracy: 0.9027
Epoch 157/250
ccuracy: 0.9142
Epoch 158/250
ccuracy: 0.9016
Epoch 159/250
ccuracy: 0.9085
Epoch 160/250
ccuracy: 0.9027
Epoch 161/250
ccuracy: 0.9142
Epoch 162/250
ccuracy: 0.9027
Epoch 163/250
ccuracy: 0.9165
Epoch 164/250
ccuracy: 0.9050
Epoch 165/250
ccuracy: 0.9050
Epoch 166/250
ccuracy: 0.9176
Epoch 167/250
ccuracy: 0.9153
Epoch 168/250
ccuracy: 0.9096
```

```
Epoch 169/250
ccuracy: 0.9016
Epoch 170/250
ccuracy: 0.8799
Epoch 171/250
ccuracy: 0.9153
Epoch 172/250
ccuracy: 0.9039
Epoch 173/250
ccuracy: 0.9073
Epoch 174/250
ccuracy: 0.9073
Epoch 175/250
ccuracy: 0.9016
Epoch 176/250
ccuracy: 0.9165
Epoch 177/250
ccuracy: 0.8924
Epoch 178/250
ccuracy: 0.8879
Epoch 179/250
ccuracy: 0.8947
Epoch 180/250
ccuracy: 0.9073
Epoch 181/250
ccuracy: 0.9016
Epoch 182/250
ccuracy: 0.9096
```

```
Epoch 183/250
ccuracy: 0.9130
Epoch 184/250
ccuracy: 0.9119
Epoch 185/250
ccuracy: 0.9153
Epoch 186/250
ccuracy: 0.9085
Epoch 187/250
ccuracy: 0.9130
Epoch 188/250
ccuracy: 0.9268
Epoch 189/250
ccuracy: 0.9130
Epoch 190/250
ccuracy: 0.9199
Epoch 191/250
ccuracy: 0.8993
Epoch 192/250
ccuracy: 0.9096
Epoch 193/250
ccuracy: 0.8993
Epoch 194/250
ccuracy: 0.9027
Epoch 195/250
ccuracy: 0.8947
Epoch 196/250
ccuracy: 0.9073
```

```
Epoch 197/250
ccuracy: 0.8993
Epoch 198/250
ccuracy: 0.8993
Epoch 199/250
ccuracy: 0.9005
Epoch 200/250
ccuracy: 0.9085
Epoch 201/250
ccuracy: 0.8982
Epoch 202/250
ccuracy: 0.9027
Epoch 203/250
ccuracy: 0.9073
Epoch 204/250
ccuracy: 0.9027
Epoch 205/250
ccuracy: 0.8936
Epoch 206/250
ccuracy: 0.8924
Epoch 207/250
ccuracy: 0.8947
Epoch 208/250
ccuracy: 0.9119
Epoch 209/250
ccuracy: 0.8970
Epoch 210/250
ccuracy: 0.8993
```

```
Epoch 211/250
ccuracy: 0.9050
Epoch 212/250
ccuracy: 0.9085
Epoch 213/250
ccuracy: 0.8982
Epoch 214/250
ccuracy: 0.9176
Epoch 215/250
126/126 [============ - - 5s 36ms/step - loss: 0.0119 - accuracy: 0.9960 - val loss: 0.6690 - val a
ccuracy: 0.9176
Epoch 216/250
ccuracy: 0.9108
Epoch 217/250
ccuracy: 0.8936
Epoch 218/250
ccuracy: 0.9130
Epoch 219/250
ccuracy: 0.9050
Epoch 220/250
ccuracy: 0.8879
Epoch 221/250
ccuracy: 0.8764
Epoch 222/250
ccuracy: 0.9062
Epoch 223/250
ccuracy: 0.9073
Epoch 224/250
ccuracy: 0.9027
```

```
Epoch 225/250
ccuracy: 0.9027
Epoch 226/250
ccuracy: 0.9016
Epoch 227/250
ccuracy: 0.9130
Epoch 228/250
ccuracy: 0.9153
Epoch 229/250
ccuracy: 0.8993
Epoch 230/250
ccuracy: 0.9039
Epoch 231/250
ccuracy: 0.9108
Epoch 232/250
ccuracy: 0.9027
Epoch 233/250
ccuracy: 0.9005
Epoch 234/250
ccuracy: 0.9039
Epoch 235/250
ccuracy: 0.8856
Epoch 236/250
ccuracy: 0.8993
Epoch 237/250
ccuracy: 0.9119
Epoch 238/250
ccuracy: 0.9050
```

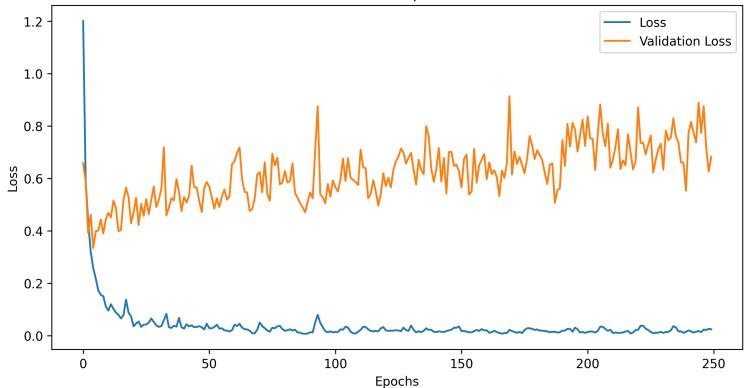
```
Epoch 239/250
ccuracy: 0.9085
Epoch 240/250
ccuracy: 0.9245
Epoch 241/250
ccuracy: 0.8993
Epoch 242/250
ccuracy: 0.9130
Epoch 243/250
ccuracy: 0.9108
Epoch 244/250
ccuracy: 0.9016
Epoch 245/250
ccuracy: 0.8982
Epoch 246/250
ccuracy: 0.9016
Epoch 247/250
ccuracy: 0.8959
Epoch 248/250
ccuracy: 0.9142
Epoch 249/250
ccuracy: 0.9005
Epoch 250/250
ccuracy: 0.9005
CNN2D Model Saved
C:\Users\vidit\AppData\Roaming\Python\Python310\site-packages\keras\src\engine\training.py:3079: UserWarning: You ar
e saving your model as an HDF5 file via `model.save()`. This file format is considered legacy. We recommend using in
stead the native Keras format, e.g. `model.save('my model.keras')`.
```

saving api.save model(

Results

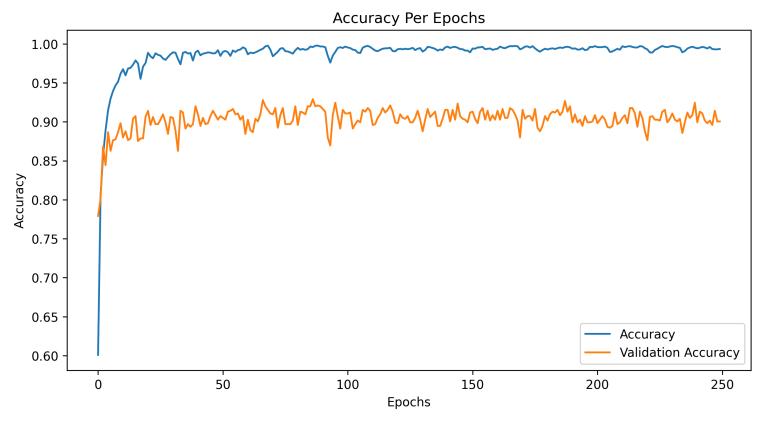
```
In [44]: plt.figure(figsize=(10, 5), dpi=300)
    plt.plot(train_hist_m3[["loss", "val_loss"]])
    plt.legend(["Loss", "Validation Loss"])
    plt.title("Loss Per Epochs")
    plt.xlabel("Epochs")
    plt.ylabel("Loss")
    plt.show()
```

Loss Per Epochs



```
In [45]: plt.figure(figsize=(10, 5), dpi=300)
   plt.plot(train_hist_m3[["accuracy", "val_accuracy"]])
```

```
plt.legend(["Accuracy", "Validation Accuracy"])
plt.title("Accuracy Per Epochs")
plt.xlabel("Epochs")
plt.ylabel("Accuracy")
plt.show()
```



```
C:\Users\vidit\AppData\Local\Temp\ipykernel_46916\2068890745.py:8: FutureWarning: The frame.append method is depreca
ted and will be removed from pandas in a future version. Use pandas.concat instead.
  log = log.append(log_entry)
```

CNN2D Prediction Function

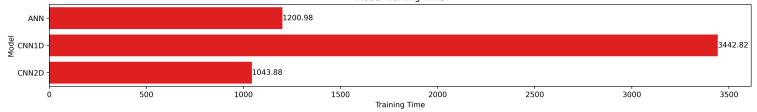
```
In [47]: # function to predict the feature
         def CNN2D Prediction(file name):
             # load the audio file
             audio data, sample rate = librosa.load(file name, res type="kaiser fast")
             # get the feature
             feature = librosa.feature.mfcc(y=audio data, sr=sample rate, n mfcc=128)
             # scale the features
             feature scaled = np.mean(feature.T, axis=0)
             # array of features
             prediction feature = np.array([feature scaled])
             # reshaping the features
             final prediction feature = prediction feature.reshape(
                 prediction feature.shape[0], 16, 8, 1
             # get the id of label using argmax
             predicted_vector = np.argmax(CNN2D_Model.predict(final_prediction_feature), axis=-1)
             # get the class label from class id
             predicted_class = le.inverse_transform(predicted_vector)
             # display the result
             print("CNN2D has predicted the class as --> ", predicted class[0])
```

Testing the Model on Sample audio

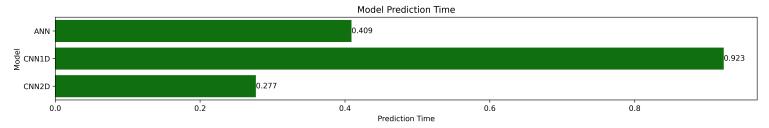
Comparative Analysis

```
In [53]: plt.rcParams["figure.figsize"] = (17, 2)
          plt.rcParams["figure.dpi"] = 550
In [54]: ax = sns.barplot(x="accuracy", y="model", data=log, color="b")
          ax.bar label(ax.containers[0])
          plt.xlabel("Accuracy")
          plt.ylabel("Model")
          plt.title("Model Accuracy")
          plt.show()
                                                                    Model Accuracy
          CNN1D -
                                                                                                                          92.0483
                                                                                                                        90.0458
           CNN2D
                                                             40
                                       20
                                                                                                           80
                                                                                    60
                                                                      Accuracy
In [55]: ax = sns.barplot(x="train_time", y="model", data=log, color="r")
          ax.bar_label(ax.containers[0])
          plt.xlabel("Training Time")
          plt.ylabel("Model")
          plt.title("Model Training Time")
          plt.show()
```





```
In [56]: ax = sns.barplot(x="pred_time", y="model", data=log, color="g")
    ax.bar_label(ax.containers[0])
    plt.xlabel("Prediction Time")
    plt.ylabel("Model")
    plt.title("Model Prediction Time")
    plt.show()
```



```
In []:
```