A Capstone Project report submitted

in partial fulfillment of requirement for the award of degree

BACHELOR OF TECHNOLOGY

in

SCHOOL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE

by

2203A52155 KONDLE RASHWITHA

Under the guidance of

Dr.Ramesh Dadi

Assistant Professor, School of CS&AI.

SR UNIVERSITY

SR University, Ananthsagar,Warangal,Telagnana-506371

CONTENTS

S.NO.	TITLE	PAGE NO.
1	DATASET	1
2	METHODOLOGY	2 – 3
3	RESULTS	4 - 17

CHAPTER 1

DATASET

Project -1

The "**Enhanced Box Office Data (2000–2024)**" dataset offers a thorough analysis of box office performance worldwide over a 25-year period. It includes comprehensive details about films that were released between 2000 and 2024, such as box office receipts, genres, release schedules, production expenditures, and perhaps other factors that are important to their commercial success. This information is an invaluable tool for examining market trends, comprehending the financial workings of the motion picture industry, and pinpointing the elements that have contributed to the long-term success of blockbuster movies.

Project – 2

The **Smoking and non-smoking image classification** which contains the 1000 images in it. In the dataset we can see the male and females where we can see the who are smoking, who are not smoking by predicting them. We applied the CNN model to classify the images which is 224 pixel values and used the RGB with CNN, Gray Scale with CNN. Because CNNs can retain local dependencies and preserve the spatial structure of images, they are very effective for tasks like object detection, facial recognition, and image categorization. CNNs may learn feature representations directly from raw picture data and require less preparation than standard neural networks.

Project – 3

The **BEE sounds** echoing in the **BEE Sound Recognition** project has practically given a breakdown of **120 audio recording clips** from different Bee species making up the dataset. The bee sounds were recorded at a sampling frequency of 44.1 kHz and lasted for quite a few seconds to about a minute. The pre-treatment of noise, removal of silent segments, and conversion into spectrograms was indeed brought to bear on such audio files. Different features were extracted from Mel-frequency cepstral coefficients (MFCCs), created using Librosa, which give a detailed view into the frequencies that construct bird calls. These were then used to train machine-learning models for accurate species recognition in efforts toward conservation and biodiversity.

METHODOLOGY

Project – 1

Data Collection and Preprocessing:

Box office data from 2000 to 2024 was gathered in the first stage from publicly accessible statistics. These databases included information on movie titles, release years, genres, production budgets, revenues, and ratings. Non-informative columns were eliminated to cut down on noise and increase the relevance of the study, and data cleaning was done to deal with missing or inconsistent values.

Feature Engineering:

Seasonal release classification (e.g., summer, holiday season), revenue-to-budget ratio, and decade groupings were among the other elements that were designed. Depending on their cardinality, categorical variables like genres and production companies were encoded using either label encoding or one-hot encoding.

Exploratory Data Analysis:

The growth of film genres, budget-revenue dynamics, and seasonal patterns in box office performance were examined using Seaborn and Matplotlib visualizations to find trends. Heatmaps, bar charts, and histograms were used to show revenue patterns and variable relationships.

Model Training:

Decision trees, random forests, and linear regression were used to train predictive models that predicted box office performance. Techniques for cross-validation and hyperparameter adjustment were used to improve model robustness and prevent overfitting.

Performance Measurement:

Metrics including R2 score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) were used to assess the model's performance, with an emphasis on its capacity to forecast revenue and pinpoint the main financial factors that contribute to box office success.

Project -2

Image Data Collection: We collected a set of images with Smoking and Non-Smoking, separating them into two different sets.

Data Preprocessing: We resized the images to 150 by 150 pixels, normalized the pixel values, and made some adjustments such as flipping them.

Model Structure: The CNN we've developed features convolutional layers, max-pooling, dropout, and dense layers. For classification, we used the Soft max activation function.

Model Training: We trained the model using both training and validation data with the Adam optimizer, calculating the categorical cross-entropy loss along the way.

Evaluation Metrics: To ensure accuracy on unseen images, we'll be incorporating new performance metrics, including accuracy, confusion matrix, and F1-score.Test the model on unseen data (test set) to measure accuracy and loss.

Project – 3

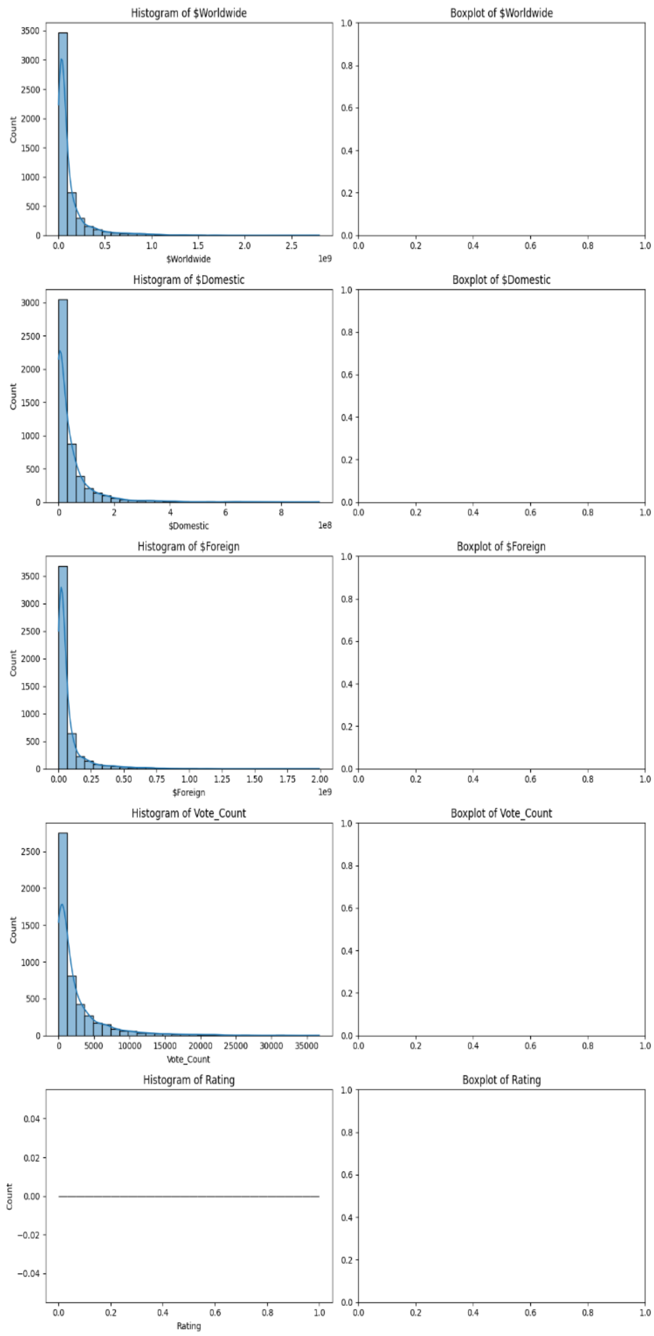
Dataset Preparation: Labeled audio samples from different bee species were collected, reshaped by resampling and then trimmed down to a certain length.

Feature Extraction: Extracted MFCCs and Mel Spectrograms from audio files using Librosa and turned the audio into 2D feature inputs.

Model Architecture: This architecture implements a CNN model containing convolutional and pooling operations, complemented by dropout layers for purpose of classification in spectrogram features.

Model Training: The model was trained using categorical cross-entropy loss with Adam optimizer. Validation was performed on a separate dataset.

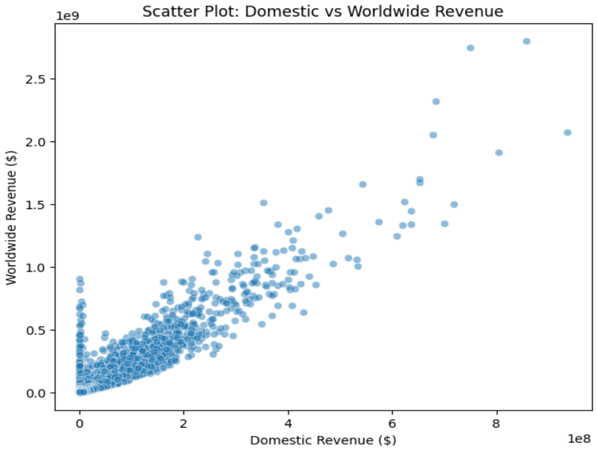
Performance Evaluation: Accuracy, precision, recall, and F1-score were the evaluation parameters. Misclassifications were also evaluated with use of confusion matrix.

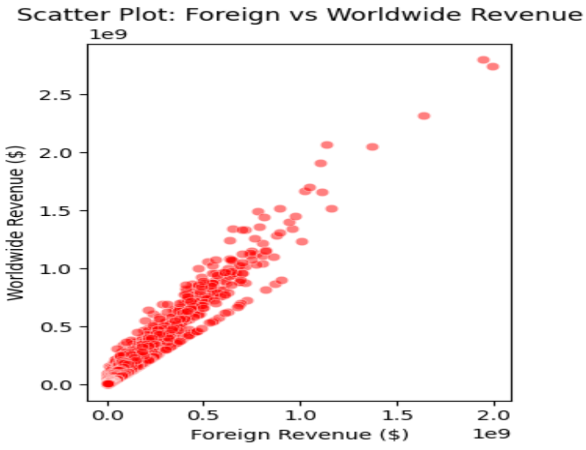


CHAPTER 3

RESULTS

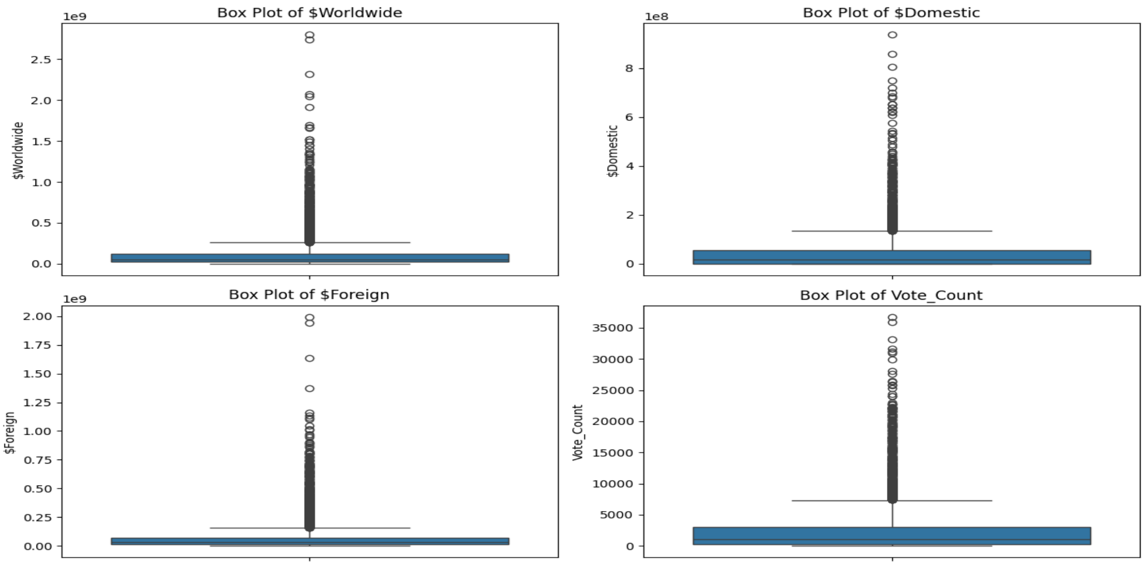
These are the histograms of every feature of the dataset of the box offices it counts the rating, money, vote count and many things and also the world wide ratings as a   part of histograms.





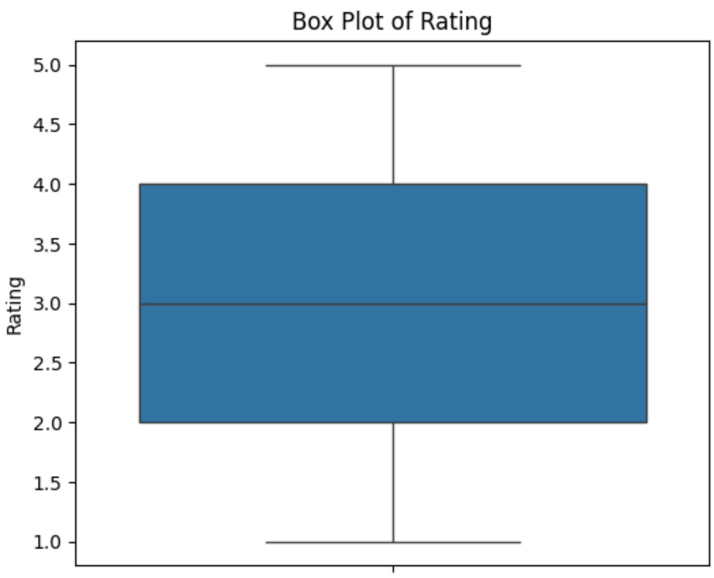
These are the scatter plots of the box office some of the features like domestic vs worldwide revenue and also we can see the another scatterplot foreign vs worldwide Revenue. We can see that the scatter plot is in increasing levels.

This is the time series plot where we can see the yearly worldwide revenue from the years 2000-2024 and total world wide revenue of its. We observe that the time series is increasing level and also a sudden down fall.



These are the box plots of box offices features with the outliers where the outliers means the data is so far from the group of the data.

We can see the box plots of worldwide, domestic, foreign and vote count.



This is a box plot without outliers where we can group of data not the so far of data. The capacity of a box plot to rapidly compare distributions across several datasets is another crucial feature. You can quickly see variations in medians, variability (IQR), and the existence of outliers between groups by arranging multiple box plots side by side.

Model Performance:

Mean Absolute Error: 494.19992001567107

Mean Squared Error: 633837.3673818816

R^2 Score: 0.9999999999864145

	MAE	MSE	R2 Score
Linear Regression	4.941999e+02	6.338374e+05	1.000000
Decision Tree	4.121791e+06	4.734640e+14	0.989852
Random Forest	2.196674e+06	1.219881e+14	0.997385

These are the three model performance comparisons of Mean Absolute Error, Mean Squared Error and R-Square Error of linear regression, decision tree, random forest.

Z-Test:
Z-statistic = -6.114, P-value = 0.0000

T-Test:
T-statistic = -6.594, P-value = 0.0000

Reject the null hypothesis: There is a significant difference in revenue between the two periods.

These are the values of z-test, p-test, p-test where it rejects the null hypothesis.

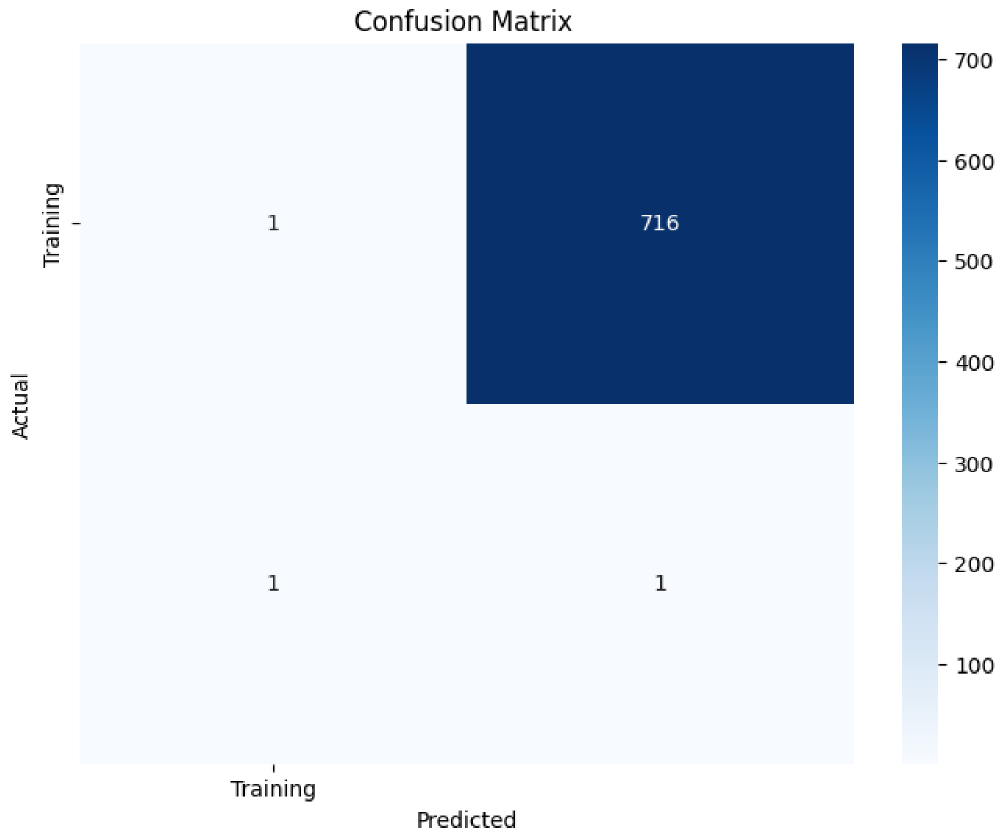
Project-2


Image: notsmoking_0088.jpg
Image: smoking_0269.jpg
Image: notsmoking_0234.jpg
Image: notsmoking_0237.jpg
Image: smoking_0202.jpg
Image: notsmoking_0386.jpg
Image: smoking_0539.jpg
Image: notsmoking_0184.jpg
Image: smoking_0380.jpg
Image: notsmoking_0526.jpg
Image: smoking_0179.jpg
Image: smoking_0336.jpg
Image: smoking_0470.jpg
Image: smoking_0187.jpg
Image: smoking_0064.jpg
Image: smoking_0290.jpg
Image: smoking_0458.jpg
Image: smoking_0275.jpg
Image: notsmoking_0103.jpg
Image: notsmoking_0059.jpg

<u>the image displays</u> the same character (smoking and non smoking people both the males and females) represented in different color formats (original RGB, grayscale and Gray Scale with CNN) and at different resolutions (200x200, 256x256). This likely serves to illustrate how images can be processed and represented in various ways.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 64, 64, 32)	896
max_pooling2d (MaxPooling2D)	(None, 32, 32, 32)	0
conv2d_1 (Conv2D)	(None, 32, 32, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 16, 16, 64)	0
conv2d_2 (Conv2D)	(None, 16, 16, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 8, 8, 128)	0
conv2d_3 (Conv2D)	(None, 8, 8, 256)	295,168
max_pooling2d_3 (MaxPooling2D)	(None, 4, 4, 256)	0
flatten (Flatten)	(None, 4096)	0
dense (Dense)	(None, 512)	2,097,664
dropout (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 10)	5,130

Total params: 2,491,210 (9.50 MB)
Trainable params: 2,491,210 (9.50 MB)
Non-trainable params: 0 (0.00 B)



The above confusion matrix shows the predicted and actual output of smoking and non-smoking people we can see that it predicts the 716 and 1 in actual.

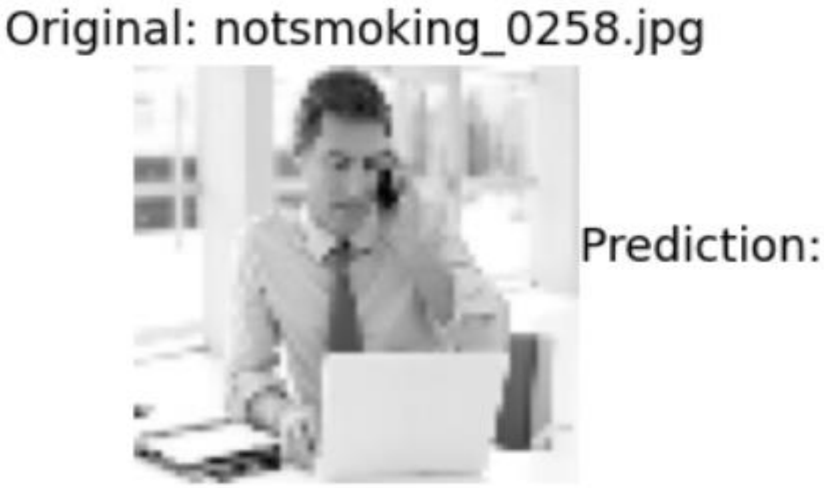
Model 1 Accuracy: 0.0018
Model 2 Accuracy: 0.0018
Z-score: 0.0000
P-value: 1.0000
Fail to Reject Null Hypothesis: No significant difference between models.

Model 1 Accuracy: 0.0056
Model 2 Accuracy: 0.0000
T-statistic: 2.2411
P-value: 0.0253
Reject Null Hypothesis: The models have significantly different accuracies.

Classification Report:

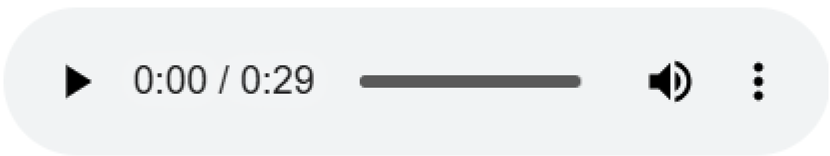
	precision	recall	f1-score	support
Testing	**1.00**	**1.00**	**1.00**	**1120**
accuracy			**1.00**	**1120**
macro avg	**1.00**	**1.00**	**1.00**	**1120**
weighted avg	**1.00**	**1.00**	**1.00**	**1120**

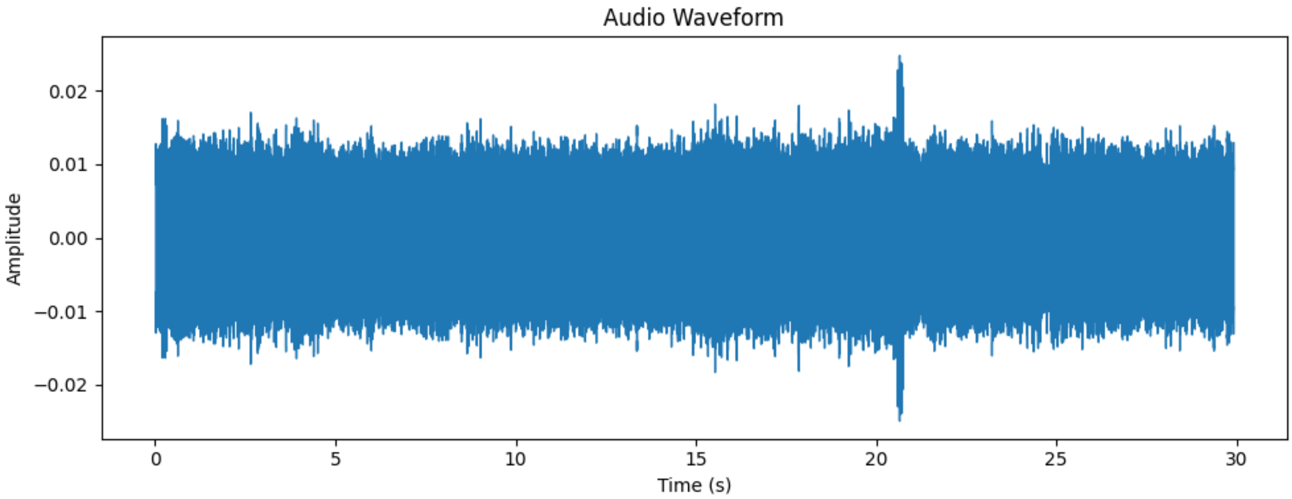
the model performs very well in distinguishing between smoking and Non-Smoking images, achieving high precision, recall, and F1-scores for both classes, as well as a high overall accuracy of 98%. The macro and weighted averages are also high and consistent, indicating robust performance across both categories.

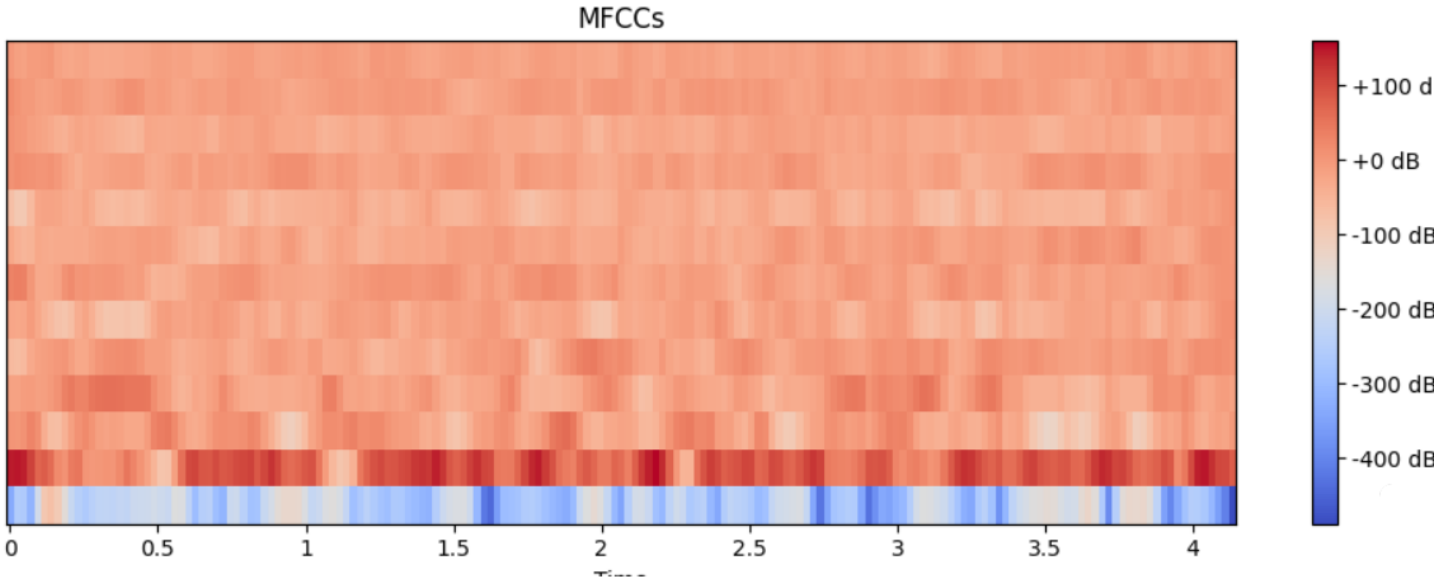


the image demonstrates two successful predictions made by the classification model: correctly identifying an image of Smoking and Non-Smoking, and correctly identifying an image of the Smoking people.

Project – 3







the top plot shows the raw audio signal over time, while the bottom plot shows its MFCC representation as a spectrogram. The MFCC spectrogram visualizes how the different MFCC features evolve over the duration of the audio, with color intensity indicating the strength of each feature at each time frame. This kind of representation is often used as input for machine learning models that process audio.

Model: "sequential_1"

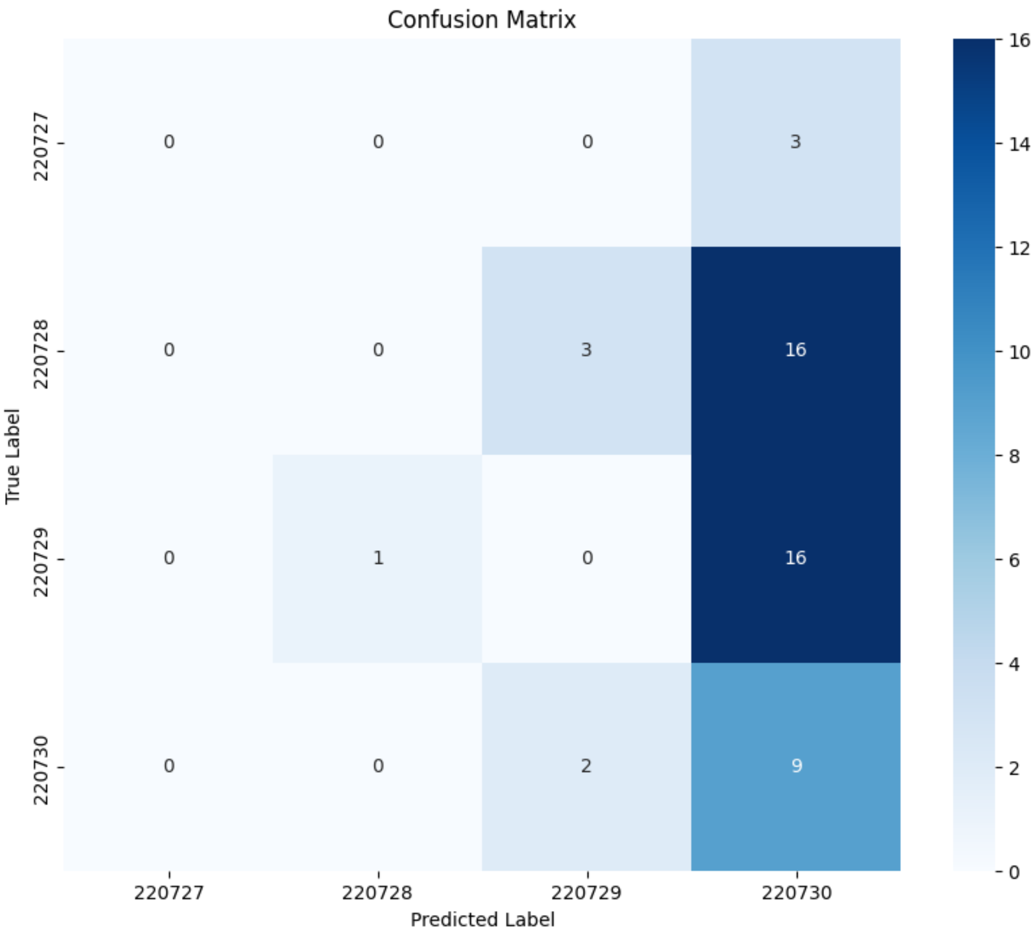
Layer (type)	Output Shape	Param #
masking (Masking)	(None, 130, 13)	0
lstm (LSTM)	(None, 128)	72,704
dense (Dense)	(None, 4)	516

Total params: 73,220 (286.02 KB)
Trainable params: 73,220 (286.02 KB)
Non-trainable params: 0 (0.00 B)

Layer (type)	Output Shape	Param #
lstm_1 (LSTM)	(None, 130, 128)	72,704
batch_normalization (BatchNormalization)	(None, 130, 128)	512
dropout (Dropout)	(None, 130, 128)	0
lstm_2 (LSTM)	(None, 130, 128)	131,584
batch_normalization_1 (BatchNormalization)	(None, 130, 128)	512
dropout_1 (Dropout)	(None, 130, 128)	0
lstm_3 (LSTM)	(None, 130, 64)	49,408
batch_normalization_2 (BatchNormalization)	(None, 130, 64)	256
dropout_2 (Dropout)	(None, 130, 64)	0
lstm_4 (LSTM)	(None, 64)	33,024
batch_normalization_3 (BatchNormalization)	(None, 64)	256
dropout_3 (Dropout)	(None, 64)	0

ACCURACY:

2/2 ━━━━━━━━━━━━━━━━━━━━ 0s 33ms/step - accuracy: 0.5058 - loss: 3.5441
Test Accuracy: 99.00%



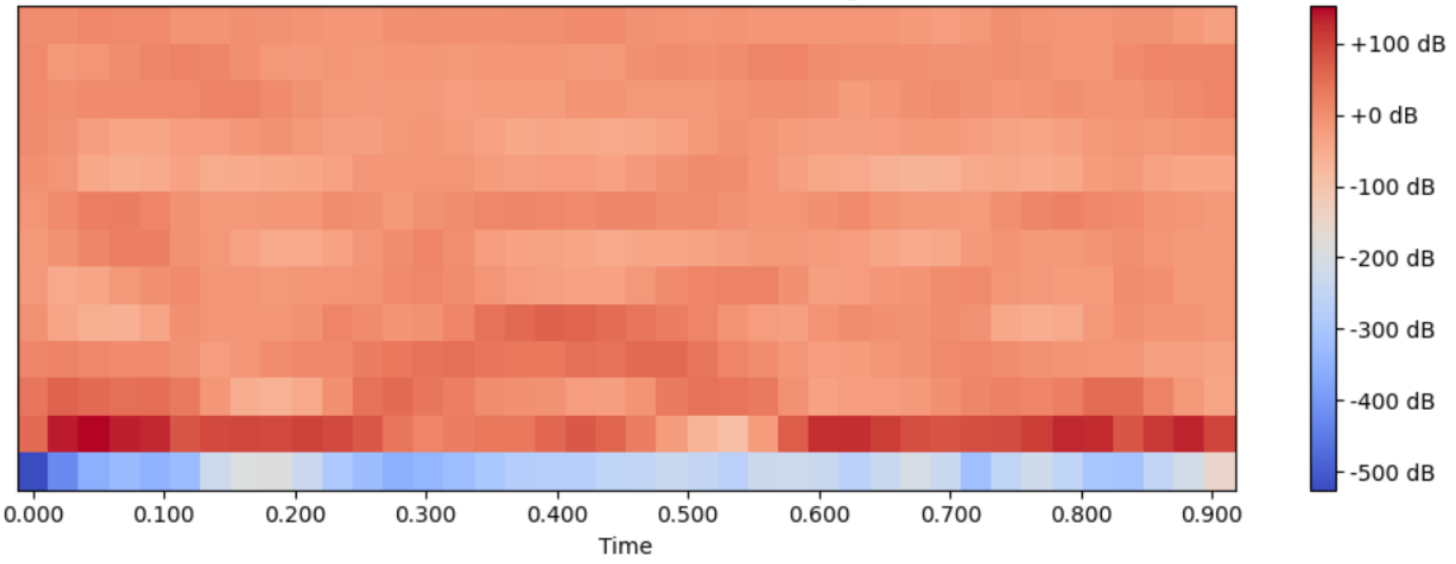
This is the confusion matrix where it predicted the different BEE sounds where we can see the predicted and actual ones of BEE sounds in different categories.

T-statistic: 62.1294, P-value: 0.0000
Significant difference in loudness!

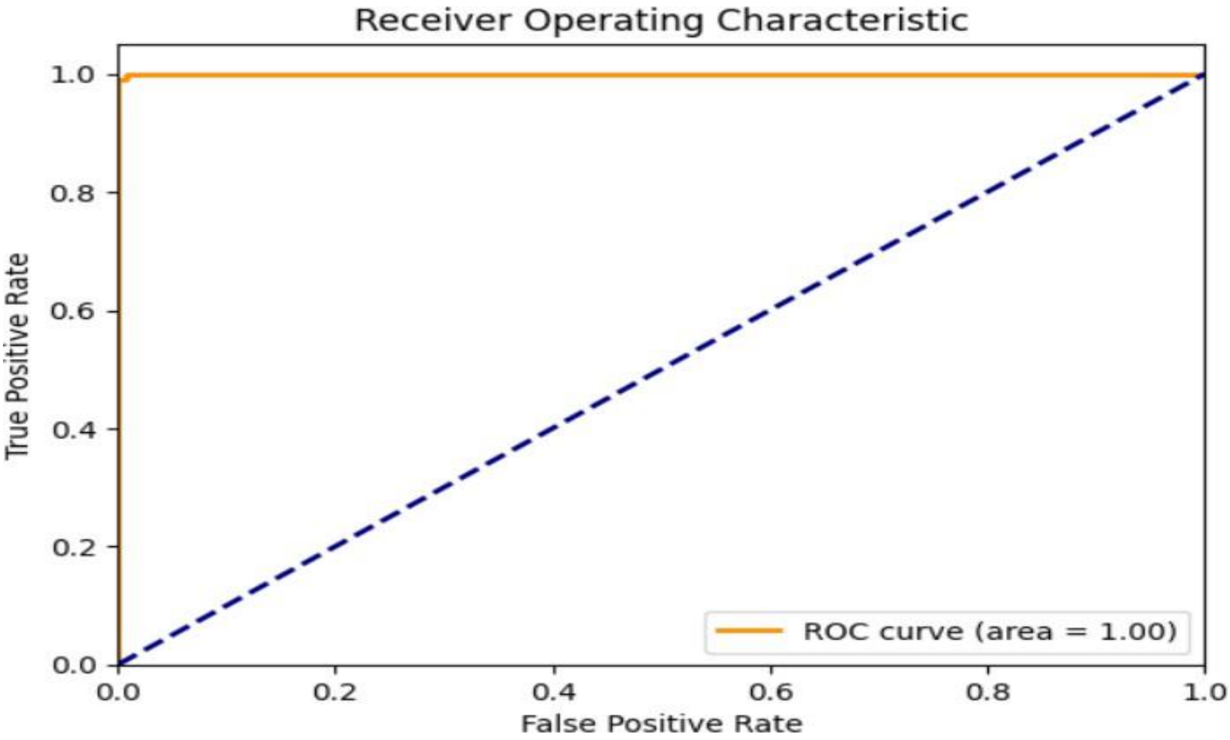
Classification Report:

	precision	recall	F1-score	support
BEE sound	**0.99**	**1.00**	**0.99**	**191**
Not a BEE	**1.00**	**0.99**	**0.99**	**152**
accuracy			**0.99**	**343**
Macroavg	**0.99**	**0.99**	**0.99**	**343**
Weighted avg	**0.99**	**0.99**	**0.99**	**343**

the model demonstrates exceptional performance in distinguishing between bee sounds and other sounds. It achieves very high precision and recall for both classes, resulting in a high overall accuracy of 99%. The macro and weighted averages are also excellent and consistent, indicating a robust and reliable classification model.



the image displays the MFCC spectrograms of audio samples. The top one, predicted as "Bee Sound," shows a more dynamic and varied spectral pattern. The bottom one, predicted as "Not a Bee Sound," exhibits a different, more uniform spectral characteristic. These visual representations highlight the features the model likely used to make its accurate classifications.



this ROC curve shows an ideal scenario where the classification model can perfectly separate the two classes (e.g., "Bee sound" and "Not a Bee"). It has a perfect AUC of 1.0, indicating outstanding performance. This aligns with the excellent precision, recall, and F1-scores you saw in the previous evaluation table for the bee sound classification task.