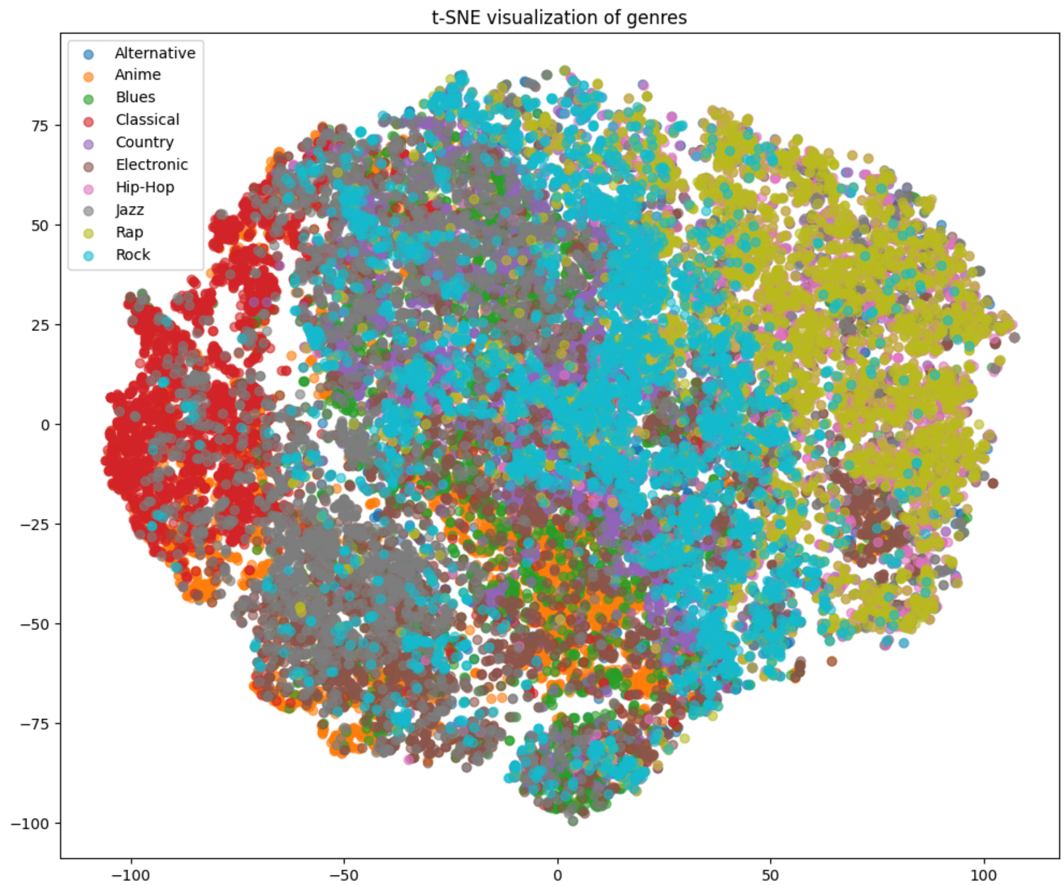
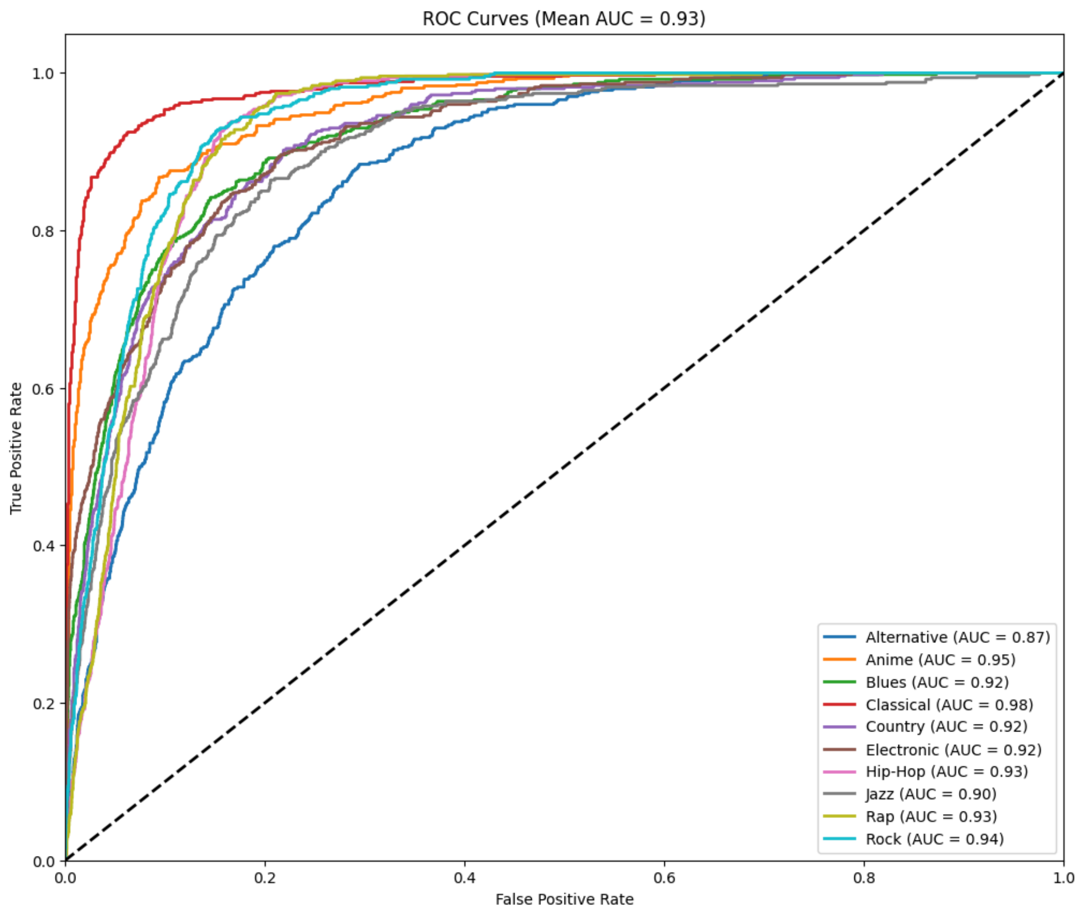
1. Intro:
   In this report, I will outline how I dealt with the challenges and built the model to train and fit the data to classify the Spotify songs into the correct genre and calculate my final ROC curve.
2. Preprocess the data
   First, I set the seed to my N-number and read musicData.csv. Notice that the dataset contains missing values in some features marked as “?” for tempo and “-1” for duration_ms. So, I set na_value or pandas to “?” and “-1” and drop the rows containing N/A, as they are relatively small compared to the whole dataset. Next, since some variables are not meaningful to include in the model as they are linguistic properties of the artist and song, we drop the “instance_id”, “artist_name”, “track_name”, and “obtained_date” columns in df. Then, we need to decode the variables “key” and “music_genre” from a string format to numerical values, and dummy code “mode” into 0 and 1. After handling the specific data, we split the data into the test set using 500 randomly picked songs for each genre, and the rest for the training set to prevent leakage. X will be our acoustic features, and y will be the music genre. Because different variables are on different scales, we normalize them except the “mode,” as it is a categorical value before reducing dimensionality.
3. Dimension reduction and Clustering
   Then, for good classification performance, we first do dimension reduction and a cluster step on these input features. Since the acoustic features are unlikely to be normally distributed and linear, we use t-SNE with n_components=2 and a standard perplexity=30 to fit and transform X_train and plot the graph. After that, we use the DBSCAN to fit and predict t-SNE results since it can identify clusters of arbitrary shapes and handles noise with eps=0.5 and min_samples=3 after finetuning the two hyperparameters, and plot the graph with clusters. From the graph, we found that **in the 2D projection of t-SNE, certain genres like Classical and Rap form distinct clusters, while other genres like Rock and Jazz overlap with each other.**
4. Classification model
   I implemented a deep neural network model for genre classification because of the complex and non-linear relationship between audio features and genres, and it can handle multi-class classification, not just binary. The deep neural network consists of two hidden layers, with 128 and 64 neurons respectively, with ReLU activation function. And we add two dropout layers, 0.3 and 0.2, to prevent overfitting. The input layer is the feature dimensions, and the output layer uses softmax activation for 10-class classification. During the training process, we use the cross-entropy loss function as it is appropriate for multi-class classification and the SGD optimizer with a learning rate of 0.01, L2 regularization to prevent overfitting, with 100 epochs and a batch size of 64.
5. AUC performance and evaulation
   Finally, the model performance is evaluated using ROC curves and AUC for each genre class on y_test. We get AUC for Alternative is **0.8705**, for Anime is **0.9550**, for Blues is **0.9212**, for Classical is **0.9798**, for Country is **0.9159**, for Electronic is **0.9190**, for Hip-Hop is **0.9280**, for Jazz is **0.8985**, for Rap is **0.9305**, for Rock is **0.9409**, the average AUC is 0.93, which is a decent number. **I think the most important factor that**

underlies the classification success is the prevention of overfitting and the correct use of the non-linear model to capture the characteristics of the data.

Extra Credit:
We use the correlation coefficient to find the top 3 most important features for each genre classification, which are
Electronic: ['acousticness', 'instrumentalness', 'energy']
Anime: ['popularity', 'danceability', 'speechiness']
Jazz: ['acousticness', 'instrumentalness', 'energy']
Alternative: ['loudness', 'energy', 'acousticness']
Country: ['instrumentalness', 'mode', 'speechiness']
Rap: ['popularity', 'speechiness', 'danceability']
Blues: ['popularity', 'valence', 'speechiness']
Rock: ['popularity', 'speechiness', 'instrumentalness']
Classical: ['loudness', 'acousticness', 'energy']
Hip-Hop: ['speechiness', 'popularity', 'danceability']


t-SNE visualization of genres
Alternative
Anime
Blues
Classical
Country
Electronic
Hip-Hop
Jazz
Rap
Rock
75
50
25
0
−25
−50
−75
−100
−100
−50
0
50
100



ROC Curves (Mean AUC = 0.93)
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Alternative (AUC = 0.87)
Anime (AUC = 0.95)
Blues (AUC = 0.92)
Classical (AUC = 0.98)
Country (AUC = 0.92)
Electronic (AUC = 0.92)
Hip-Hop (AUC = 0.93)
Jazz (AUC = 0.90)
Rap (AUC = 0.93)
Rock (AUC = 0.94)