Final Project Rubric

Your slides should include:

- Title, Authors
- **(15%) Motivation:** Introduce your question and why the question is interesting. Explain what has been done before in this space. Describe your overall plan to approach your question. Provide a summary of your results.
- **(15%) Data:** Describe in detail the data you are using, including the source(s) of the data and relevant statistics.
- **(15%) Modeling:** Describe in detail the models (baseline + improvement over baseline) that you use in your approach.
- **(30%) Experiments:** Provide insight into the effect of different hyperperameter choices. Please include tables, figures, graphs to illustrate your experiments.
- **(10%) Conclusions:** Summarize the key results, what has been learned, and avenues for future work.
- **(15%) Code submission:** Provide link to your team's GitHub repo. The code should be well commented and organized.
- **Contributions**: Specify the contributions of each author (e.g., data processing, algorithm implementation, slides, etc.). Note that the final project grade is individual, based on each member's contribution and team size.

MUSIC GENRE CLASSIFICATION

DATA 207 FINAL PROJECT

Adithi Suresh, Annie Pang, Joshua Shin, Millie Kobayashi

MOTIVATION

Question: How to develop a music genre classification system that categorizes songs based on features?

Why Interesting: Automating music genre classification has the potential to change the music discovery process, making it easier for users to explore and find new music tailored to their tastes.

Previous Research:

- LSTM and SVM network to recognize unique patterns for each song (Link)
- LSTM based on Time and Frequency Domain Features (Link)

PLAN TO APPROACH & DATA PREVIEW

Data Pre-processing
- Handle missing values
- Encode categorical variables
- Feature engineering

EDA
- Visualization
- Feature Relationships

Model Selection, Training & Tuning
- Split into train/test
- Choose models
- Train and optimize

Model Evaluation
- Accuracy, Precision, Recall, F1-Score
- Confusion Matrix
- Cross- Validation

Test Model with Test Data
- Apply trained models

Model Interpretation
- Feature importance
- Model limitations

Presentation & Reporting
- Prepare presentation
- Document Findings

	instance_id	artist_name	track_name	popularity	acousticness	danceability	duration_ms	energy	instrumentalness	key	liveness	loudness	mode	speechiness	tempo	obtained_date	valence	music_genre
0	32894.0	Röyksopp	Röyksopp's Night Out	27.0	0.00468	0.652	-1.0	0.941	0.79200	A#	0.115	-5.201	Minor	0.0748	100.889	4-Apr	0.759	Electronic
1	46652.0	Thievery Corporation	The Shining Path	31.0	0.01270	0.622	218293.0	0.890	0.95000	D	0.124	-7.043	Minor	0.0300	115.002	4-Apr	0.531	Electronic
2	30097.0	Dillon Francis	Hurricane	28.0	0.00306	0.620	215613.0	0.755	0.01180	G#	0.534	-4.617	Major	0.0345	127.994	4-Apr	0.333	Electronic
3	62177.0	Dubloadz	Nitro	34.0	0.02540	0.774	166875.0	0.700	0.00253	C#	0.157	-4.498	Major	0.2390	128.014	4-Apr	0.270	Electronic
4	24907.0	What So Not	Divide & Conquer	32.0	0.00465	0.638	222369.0	0.587	0.90900	F#	0.157	-6.266	Major	0.0413	145.036	4-Apr	0.323	Electronic

	Unnamed: 0	track_id	artists	album_name	track_name	popularity	duration_ms	explicit	danceability	energy	...	loudness	mode	speechiness	acousticness	instrumentalness	liveness	valence	tempo	time_signature	track_genre
0	0	5SuOikwiRyPMVoIQDJUgSV	Gen Hoshino	Comedy	Comedy	73	230666	False	0.676	0.4610	...	-6.746	0	0.1430	0.0322	0.000001	0.3580	0.7150	87.917	4	acoustic
1	1	4qPNDBW1i3p13qLCt0Ki3A	Ben Woodward	Ghost (Acoustic)	Ghost - Acoustic	55	149610	False	0.420	0.1660	...	-17.235	1	0.0763	0.9240	0.000006	0.1010	0.2670	77.489	4	acoustic
2	2	1iJBSr7s7jYXzM8EGcbK5b	Ingrid Michaelson;ZAYN	To Begin Again	To Begin Again	57	210826	False	0.438	0.3590	...	-9.734	1	0.0557	0.2100	0.000000	0.1170	0.1200	76.332	4	acoustic
3	3	6lfxq3CG4xtTiEg7opyCyx	Kina Grannis	Crazy Rich Asians (Original Motion Picture Sou...	Can't Help Falling In Love	71	201933	False	0.266	0.0596	...	-18.515	1	0.0363	0.9050	0.000071	0.1320	0.1430	181.740	3	acoustic
4	4	5vjLSffimiIP26QG5WcN2K	Chord Overstreet	Hold On	Hold On	82	198853	False	0.618	0.4430	...	-9.681	1	0.0526	0.4690	0.000000	0.0829	0.1670	119.949	4	acoustic

DATA STATISTICS (Source: Links to Kaggle Included)

Prediction of Music Genre

"Data 1"

- 50,000 Songs
- 10 Genres

["popularity", "acousticness", "danceability", "duration_ms", "energy", "instrumentalness", "key_num", "liveness", "loudness", "mode_num", "speechiness", "valence"]

Spotify Tracks Dataset

"Data 2"

- 50,000 Songs
- 114 Genres

["popularity", "duration_ms", "danceability", "energy", "key", "loudness", "mode", "speechiness", "acousticness", "instrumentalness", "liveness", "valence", "tempo", "time_signature"]

SUCCESS/ FAILURE AND EVALUATION PARAMETERS

Success Criteria:

Accuracy Threshold: Achieve an accuracy rate of at least 80% on the test dataset, indicating a strong overall performance.

Failure Criteria:

F1-Score: Declare failure if the F1-Score remains below 0.5, indicating a lack of balance between precision and recall.

Cross-Validation Performance: A standard deviation in accuracy above 10% during cross-validation indicates instability in model performance and may signal overfitting or insufficient generalization.

Evaluation Parameters:

Accuracy, Confusion Matrix, ROC, AUC, Loss (Cross-Entropy)

MODELS

1) Random Forest
2) Transformer Model
3) K-Nearest Model
4) ANN
5) LSTM
6) Logistic Regression
7) CNN
8) CNN-LSTM
9) XG Boost
10) XG Boost Combined with CNN (to make hybrid model)

Random Forest – Baseline (Complete Graphs in Appendix)

Model Architecture:

- Basic Random Forest classifier without special preprocessing or feature engineering.
- Input dimension of the model: the number of features in the input data

Hyperparameters:

- Default parameters for the RandomForestClassifier (e.g., n_estimators=100)

Data Preprocessing:

- Standard train-test split with a test size of 20%

Accuracy for Data 1: 0.56
Classification Report for Data 1:

	precision	recall	f1-score	support
Alternative	0.44	0.36	0.40	1027
Anime	0.79	0.74	0.77	1032
Blues	0.61	0.55	0.58	1013
Classical	0.82	0.86	0.84	947
Country	0.56	0.58	0.57	995
Electronic	0.65	0.62	0.64	1007
Hip-Hop	0.35	0.39	0.37	1005
Jazz	0.55	0.53	0.54	972
Rap	0.34	0.32	0.33	997
Rock	0.49	0.65	0.56	1004
accuracy			0.56	9999
macro avg	0.56	0.56	0.56	9999
weighted avg	0.56	0.56	0.56	9999

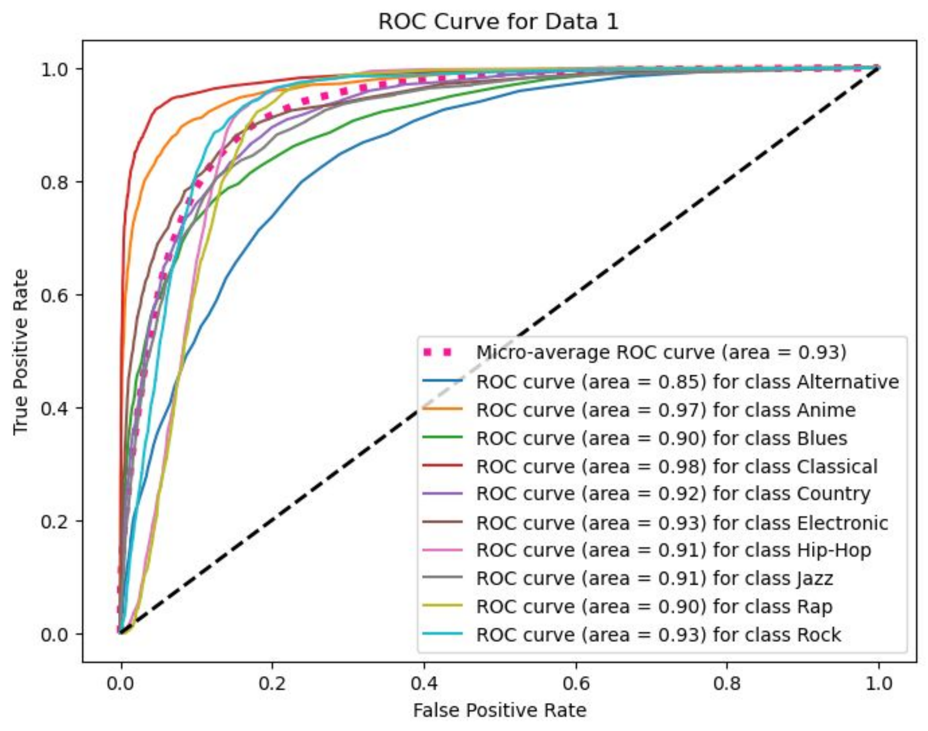
Baseline Model Data 1

Accuracy for Data 2: 0.34
Classification Report for Data 2:

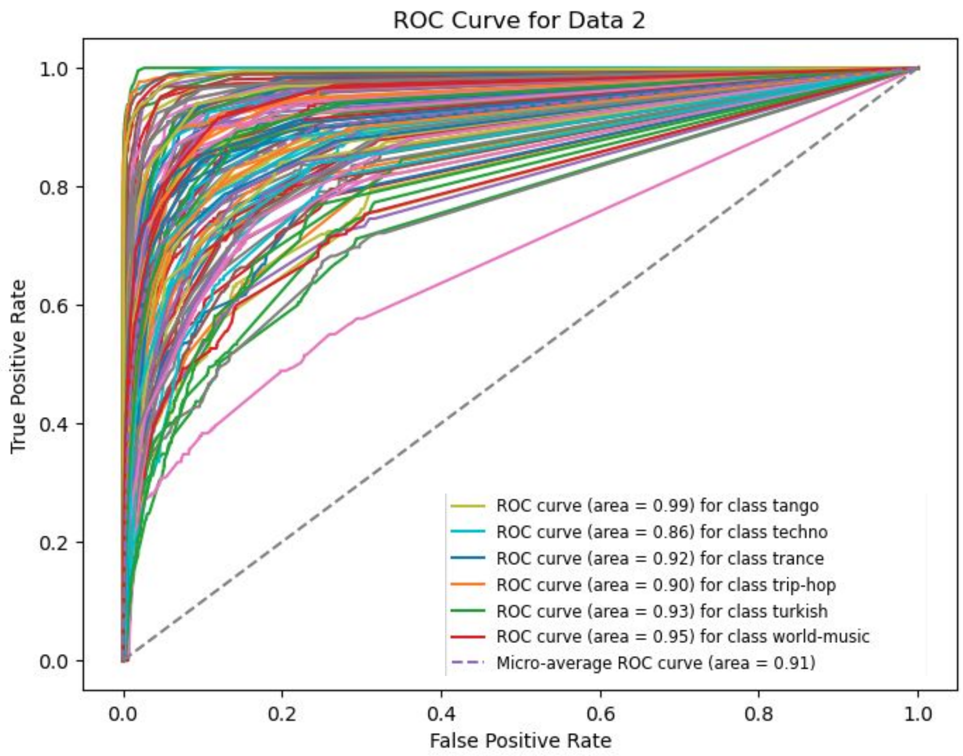
	precision	recall	f1-score	support
acoustic	0.28	0.24	0.26	213
afrobeat	0.36	0.37	0.37	203
alt-rock	0.06	0.06	0.06	215
alternative	0.11	0.13	0.12	184
ambient	0.35	0.35	0.35	197
anime	0.19	0.16	0.17	193
black-metal	0.58	0.65	0.61	210
bluegrass	0.41	0.60	0.49	205
blues	0.17	0.13	0.15	214
accuracy			0.34	22800
macro avg	0.33	0.34	0.33	22800
weighted avg	0.33	0.34	0.33	22800

Baseline Model Data 2

Random Forest – Results (Complete Graphs in Appendix)



Baseline Model Data 1



Baseline Model Data 2

Random Forest – Improved (Complete Graphs in Appendix)

Model Architecture:

- Utilized a Random Forest classifier with feature normalization using MinMaxScaler in a pipeline

Hyperparameters:

- Default parameters for the RandomForestClassifier (e.g., n_estimators=100)

Data Preprocessing:

- Feature normalization using MinMaxScaler applied in a pipeline

Accuracy for Data 1: 0.56
Classification Report for Data 1:

	precision	recall	f1-score	support
Alternative	0.40	0.38	0.39	856
Anime	0.78	0.79	0.78	919
Blues	0.60	0.54	0.57	873
Classical	0.86	0.86	0.86	881
Country	0.60	0.59	0.59	921
Electronic	0.67	0.59	0.62	931
Hip-Hop	0.35	0.36	0.35	905
Jazz	0.55	0.53	0.54	888
Rap	0.34	0.33	0.33	914
Rock	0.46	0.61	0.53	916
accuracy			0.56	9004
macro avg	0.56	0.56	0.56	9004
weighted avg	0.56	0.56	0.56	9004

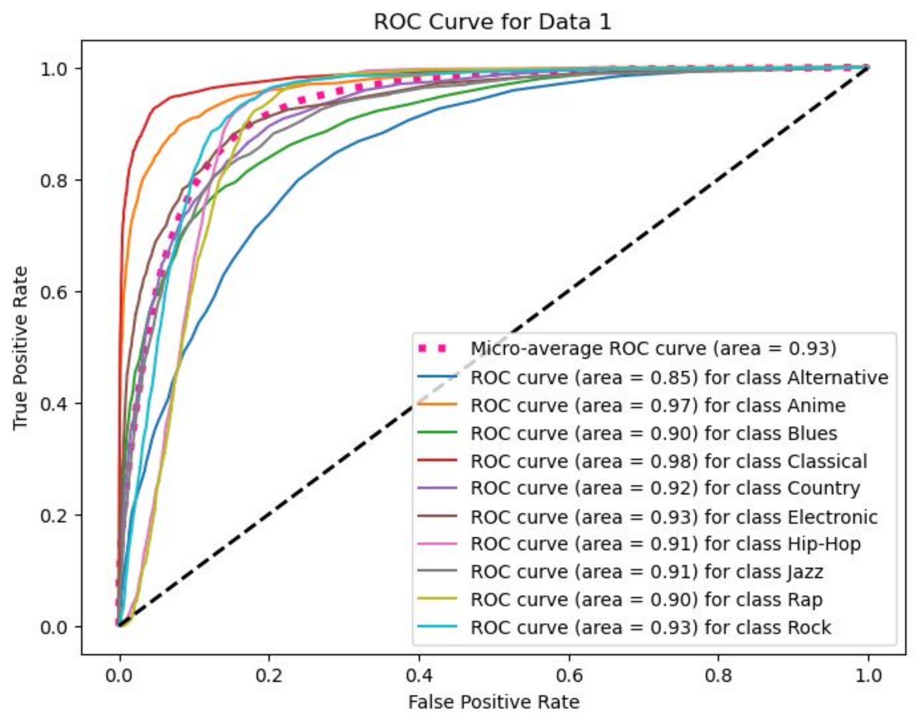
Improved Model Data 1

Accuracy for Data 2: 0.82
Classification Report for Data 2:

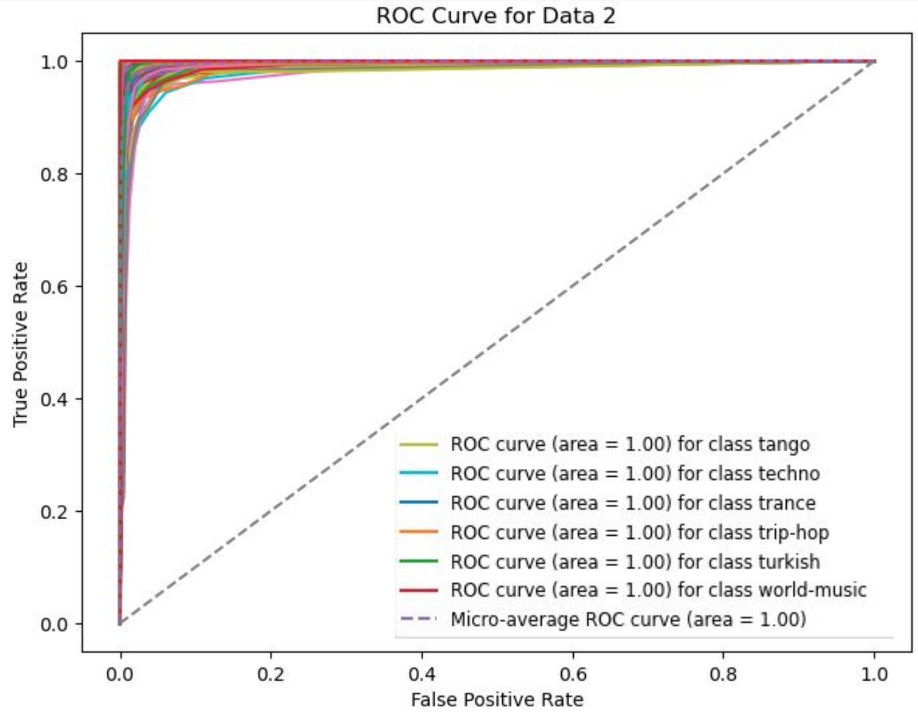
	precision	recall	f1-score	support
acoustic	0.99	1.00	0.99	213
afrobeat	0.95	0.97	0.96	203
alt-rock	0.73	0.79	0.76	215
alternative	0.70	0.74	0.72	184
ambient	0.93	0.93	0.93	197
anime	0.91	0.84	0.87	193
black-metal	0.95	0.92	0.93	210
bluegrass	0.93	0.97	0.95	205
blues	0.75	0.71	0.73	214
accuracy			0.82	22800
macro avg	0.82	0.82	0.82	22800
weighted avg	0.82	0.82	0.82	22800

Improved Model Data 2

Random Forest – Results (Complete Graphs in Appendix)



Improved Model Data 1



Improved Model Data 2

Transformer – Baseline

Model Architecture:

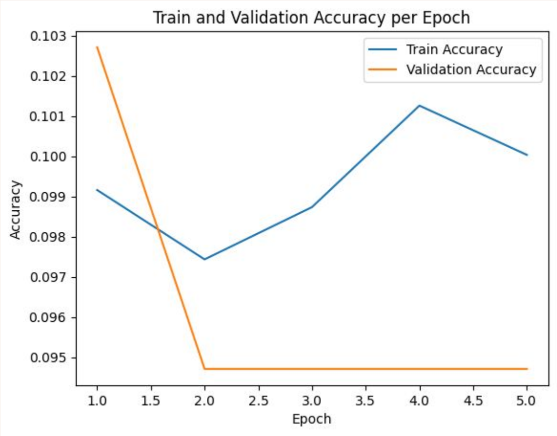
- Multiple layers of TransformerEncoderLayer
- Input dimension of the model: the number of features in the input data
- Employs a linear layer for embedding, followed by a transformer layer, and then a linear layer for classification
- Adam optimizer
- Cross Entropy Loss

Hyperparameters:

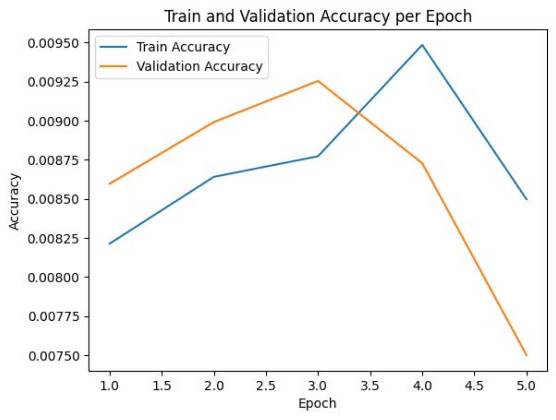
- Number of transformer layers: 3
- Number of attention heads: 8
- Hidden dimension size: 512
- Dropout rate: 0.1
- Number of Epochs: 5

Data Preprocessing:

- Standard train-test split with a test size of 20%.



Baseline Model Data 1



Baseline Model Data 2

Transformer – Improved

Model Architecture:

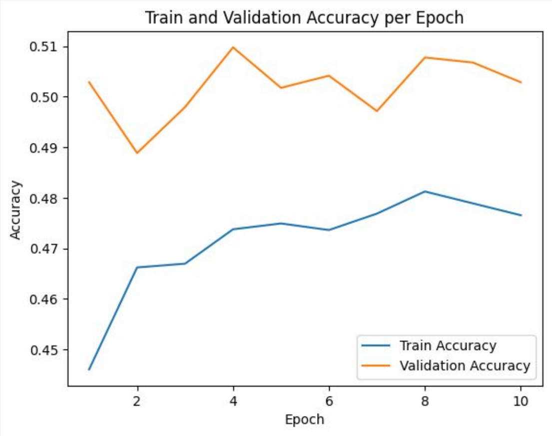
- Added L2 Regulation: dropout layer and a linear layer
- For regularization to prevent overfitting
- Class weights: handle class imbalance

Hyperparameters:

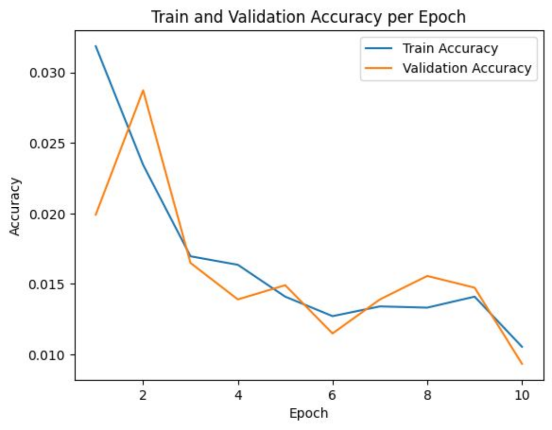
- Increased dim_feedforward in TransformerEncoderLayer to hidden_dim * 4.
- Added a dropout layer with a dropout rate of 0.1
- Added an L2 regularization layer with weight decay set to 0.01
- Increased the number of epochs from 5 to 10

Data Preprocessing:

- Selected 6 features based on absolute correlation values
- Applied standard scaling to normalize the input data



Improved Model Data 1



Improved Model Data 2

Transformer – Results

BASELINE

DATA 1	DATA 2
Training: 10.1%	Training: 0.95%
Test: 10.3%	Test: 0.93%

IMPROVED

DATA 1	DATA 2
Training: 48.1%	Training: 3.19%
Test: 50.8%	Test: 2.87%

```
Epoch 1/5
Train loss 2.3578872 accuracy 0.0991599
Val   loss 2.3179360 accuracy 0.1027103
Epoch 2/5
Train loss 2.3236991 accuracy 0.0974347
Val   loss 2.3144543 accuracy 0.0947095
Epoch 3/5
Train loss 2.3120693 accuracy 0.0987349
Val   loss 2.3149541 accuracy 0.0947095
Epoch 4/5
Train loss 2.3081996 accuracy 0.1012601
Val   loss 2.3076936 accuracy 0.0947095
Epoch 5/5
Train loss 2.3066860 accuracy 0.1000350
Val   loss 2.3057200 accuracy 0.0947095
```

Baseline Model Data 1

```
Epoch 1/5
Train loss 4.7895087 accuracy 0.0082129
Val   loss 4.7574063 accuracy 0.0085965
Epoch 2/5
Train loss 4.7480230 accuracy 0.0086405
Val   loss 4.7440438 accuracy 0.0089912
Epoch 3/5
Train loss 4.7420022 accuracy 0.0087721
Val   loss 4.7409478 accuracy 0.0092544
Epoch 4/5
Train loss 4.7407748 accuracy 0.0094849
Val   loss 4.7398903 accuracy 0.0087281
Epoch 5/5
Train loss 4.7399979 accuracy 0.0084980
Val   loss 4.7385654 accuracy 0.0075000
```

Baseline Model Data 2

```
Epoch 1/10
Train loss 1.4850932 accuracy 0.4460696
Val   loss 1.3236436 accuracy 0.5028503
Epoch 2/10
Train loss 1.4211544 accuracy 0.4662216
Val   loss 1.3353931 accuracy 0.4888489
Epoch 3/10
Train loss 1.4232043 accuracy 0.4669717
Val   loss 1.3145529 accuracy 0.4979498
Epoch 4/10
Train loss 1.4013639 accuracy 0.4737724
Val   loss 1.3002996 accuracy 0.5097510
Epoch 5/10
Train loss 1.3974236 accuracy 0.4749225
Val   loss 1.3099071 accuracy 0.5017502
Epoch 6/10
Train loss 1.3937568 accuracy 0.4736224
Val   loss 1.2879273 accuracy 0.5041504
Epoch 7/10
Train loss 1.3877345 accuracy 0.4768727
Val   loss 1.3003910 accuracy 0.4971497
Epoch 8/10
Train loss 1.3813889 accuracy 0.4812481
Val   loss 1.2822335 accuracy 0.5077508
Epoch 9/10
Train loss 1.3817808 accuracy 0.4788979
Val   loss 1.2897220 accuracy 0.5067507
Epoch 10/10
Train loss 1.3806878 accuracy 0.4765477
Val   loss 1.2831718 accuracy 0.5028503
```

Improved Model Data 1

```
Epoch 1/10
Train loss 4.5035145 accuracy 0.0318538
Val   loss 4.5287349 accuracy 0.0199123
Epoch 2/10
Train loss 4.5048833 accuracy 0.0234435
Val   loss 4.4010076 accuracy 0.0287281
Epoch 3/10
Train loss 4.6100597 accuracy 0.0169631
Val   loss 4.6313468 accuracy 0.0164912
Epoch 4/10
Train loss 4.6489790 accuracy 0.0163600
Val   loss 4.6766808 accuracy 0.0139035
Epoch 5/10
Train loss 4.6848995 accuracy 0.0141012
Val   loss 4.6744840 accuracy 0.0149123
Epoch 6/10
Train loss 4.7032847 accuracy 0.0127196
Val   loss 4.6955508 accuracy 0.0114912
Epoch 7/10
Train loss 4.7070285 accuracy 0.0134104
Val   loss 4.6934801 accuracy 0.0139035
Epoch 8/10
Train loss 4.7007198 accuracy 0.0133227
Val   loss 4.6777616 accuracy 0.0155702
Epoch 9/10
Train loss 4.6933939 accuracy 0.0141012
Val   loss 4.6809181 accuracy 0.0147368
Epoch 10/10
Train loss 4.7226529 accuracy 0.0105485
Val   loss 4.7374747 accuracy 0.0093421
```

Improved Model Data 2

KNN

BASELINE (No normalization, no feature selection)	
DATA 1	**DATA 2**
Training: 0.20555	Training: 0.05007
Test: 0.1858	Test: 0.03816

MinMaxScaler

IMPROVEMENT 1 (Normalization)	
DATA 1	**DATA 2**
Training: 0.4287	Training: 0.1848
Test: 0.4082	Test: 0.1694

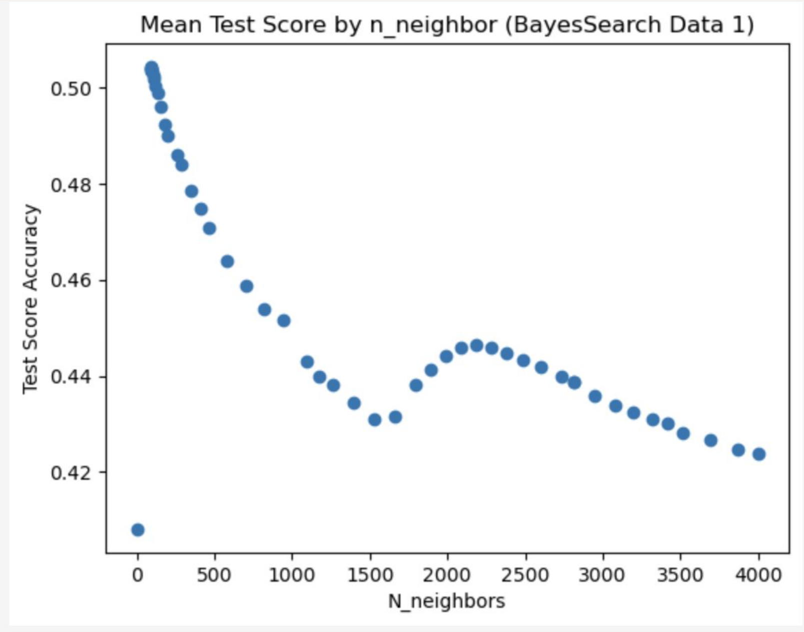
SequentialFeatureSelector

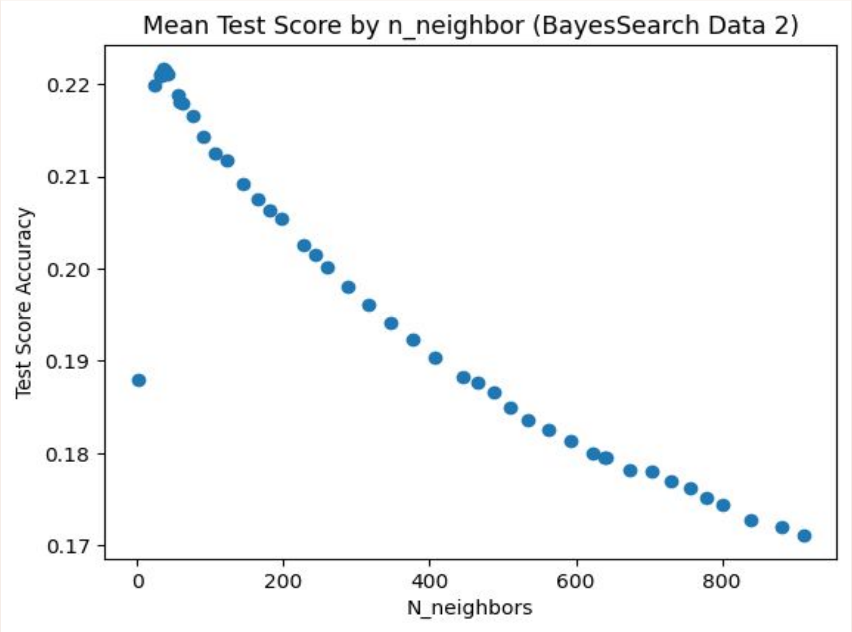
IMPROVEMENT 2 (Normalization + Feature Selection))	
DATA 1	**DATA 2**
Training: 0.50715	Training: 0.21167
Test: 0.4936	Test: 0.20105

Bayesian Search

IMPROVEMENT 3 (Normalization + Feature Selection + BayesianSearch)	
DATA 1	**DATA 2**
Training: 0.52778	Training: 0.28239
Test: 0.5102	Test: 0.22864

KNN


Mean Test Score by n_neighbor (BayesSearch Data 1)
Test Score Accuracy
0.50
0.48
0.46
0.44
0.42
0 500 1000 1500 2000 2500 3000 3500 4000
N_neighbors


Mean Test Score by n_neighbor (BayesSearch Data 2)
Test Score Accuracy
0.22
0.21
0.20
0.19
0.18
0.17
0 200 400 600 800
N_neighbors

KNN

Data 1 Baseline:

	precision	recall	f1-score	support
Alternative	0.15	0.15	0.15	1005
Anime	0.21	0.33	0.25	1057
Blues	0.17	0.10	0.13	985
Classical	0.33	0.36	0.34	967
Country	0.16	0.40	0.23	969
Electronic	0.17	0.10	0.12	1029
Hip-Hop	0.14	0.08	0.10	1012
Jazz	0.17	0.14	0.15	996
Rap	0.16	0.11	0.13	982
Rock	0.18	0.10	0.13	998
accuracy			0.19	10000
macro avg	0.18	0.19	0.17	10000
weighted avg	0.18	0.19	0.17	10000

Data 1 Improvement 3:

	precision	recall	f1-score	support
Alternative	0.40	0.33	0.36	1005
Anime	0.72	0.55	0.62	1057
Blues	0.55	0.45	0.49	985
Classical	0.77	0.85	0.81	967
Country	0.38	0.58	0.46	969
Electronic	0.62	0.54	0.58	1029
Hip-Hop	0.41	0.51	0.46	1012
Jazz	0.46	0.39	0.42	996
Rap	0.41	0.32	0.36	982
Rock	0.48	0.58	0.52	998
accuracy			0.51	10000
macro avg	0.52	0.51	0.51	10000
weighted avg	0.52	0.51	0.51	10000

Data 2 Baseline:

	precision	recall	f1-score	support
acoustic	0.00	0.00	0.00	214
afrobeat	0.02	0.01	0.01	202
alt-rock	0.01	0.01	0.01	206
alternative	0.03	0.05	0.04	201
ambient	0.00	0.00	0.00	192
anime	0.03	0.01	0.01	199
black-metal	0.02	0.01	0.01	195
bluegrass	0.00	0.00	0.00	178
blues	0.09	0.10	0.09	229
brazil	0.00	0.00	0.00	206
breakbeat	0.01	0.01	0.01	200
british	0.00	0.00	0.00	187
cantopop	0.01	0.00	0.01	222
chicago-house	0.05	0.15	0.07	203
children	0.05	0.08	0.06	201
chill	0.02	0.01	0.02	208
classical	0.13	0.05	0.08	201
club	0.04	0.01	0.02	209
comedy	0.00	0.00	0.00	216
country	0.03	0.09	0.04	209
dance	0.05	0.18	0.07	214
dancehall	0.04	0.08	0.05	192
death-metal	0.00	0.00	0.00	185
accuracy			0.04	22800
macro avg	0.03	0.04	0.03	22800
weighted avg	0.03	0.04	0.03	22800

Data 2 Improvement 3:

	precision	recall	f1-score	support
acoustic	0.10	0.13	0.11	214
afrobeat	0.22	0.15	0.18	202
alt-rock	0.05	0.05	0.05	206
alternative	0.09	0.10	0.10	201
ambient	0.25	0.27	0.26	192
anime	0.09	0.08	0.08	199
black-metal	0.30	0.43	0.36	195
bluegrass	0.17	0.44	0.25	178
blues	0.18	0.08	0.11	229
brazil	0.09	0.12	0.11	206
breakbeat	0.37	0.34	0.35	200
british	0.08	0.05	0.06	187
cantopop	0.12	0.09	0.10	222
chicago-house	0.39	0.48	0.43	203
children	0.20	0.17	0.18	201
chill	0.16	0.13	0.15	208
classical	0.43	0.52	0.47	201
club	0.32	0.16	0.22	209
comedy	0.89	0.83	0.86	216
country	0.17	0.24	0.20	209
dance	0.21	0.29	0.24	214
dancehall	0.16	0.20	0.18	192
death-metal	0.20	0.25	0.22	185
accuracy			0.23	22800
macro avg	0.22	0.23	0.21	22800
weighted avg	0.22	0.23	0.21	22800

ANN

BASELINE (SGD) (Hidden Layer 256, ReLU)	
DATA 1	**DATA 2**
Training: 0.1277	Training: 0.0091
Test: 0.1244	Test: 0.0073

BASELINE (Adam) (Hidden Layer 256, ReLU)	
DATA 1	**DATA 2**
Training: 0.1468	Training: 0.0091
Test: 0.1375	Test: 0.0073

MinMaxScaler

IMPROVEMENT 1 (Normalization)	
DATA 1	**DATA 2**
Training: 0.6057	Training: 0.3333
Test: 0.5726	Test: 0.3075

Hidden Layer

IMPROVEMENT 2 (Hidden Layer 256, 128, ReLU)	
DATA 1	**DATA 2**
Training: 0.6531	Training: 0.3671
Test: 0.5447	Test: 0.3152

IMPROVEMENT 3 (Hidden Layer 128, 64, ReLU)	
DATA 1	**DATA 2**
Training: 0.6152	Training: 0.3224
Test: 0.5724	Test: 0.3046

IMPROVEMENT 4 (Hidden Layer 128, 64, 32, ReLU)	
DATA 1	**DATA 2**
Training: 0.6229	Training: 0.3186
Test: 0.5654	Test: 0.3029

ANN

Data 1 Baseline:

	precision	recall	f1-score	support
0	0.44	0.02	0.04	1005
1	0.64	0.06	0.12	1057
2	0.52	0.05	0.08	985
3	0.79	0.07	0.12	967
4	0.33	0.06	0.10	969
5	0.69	0.04	0.08	1029
6	0.42	0.02	0.03	1012
7	0.10	0.93	0.18	996
8	0.36	0.06	0.10	982
9	0.52	0.08	0.14	998
accuracy			0.14	10000
macro avg	0.48	0.14	0.10	10000
weighted avg	0.48	0.14	0.10	10000

Data 1 Improvement 4:

	precision	recall	f1-score	support
0	0.44	0.36	0.39	1005
1	0.83	0.63	0.71	1057
2	0.54	0.59	0.56	985
3	0.82	0.83	0.83	967
4	0.55	0.51	0.53	969
5	0.64	0.61	0.63	1029
6	0.46	0.38	0.42	1012
7	0.53	0.48	0.51	996
8	0.44	0.55	0.49	982
9	0.49	0.73	0.58	998
accuracy			0.57	10000
macro avg	0.57	0.57	0.56	10000
weighted avg	0.57	0.57	0.56	10000

Data 2 Baseline:

	precision	recall	f1-score	support
0	0.00	0.00	0.00	214
1	0.00	0.00	0.00	202
2	0.00	0.00	0.00	206
3	0.00	0.00	0.00	201
4	0.00	0.00	0.00	192
5	0.00	0.00	0.00	199
6	0.00	0.00	0.00	195
7	0.00	0.00	0.00	178
8	0.00	0.00	0.00	229
9	0.00	0.00	0.00	206
10	0.00	0.00	0.00	200
11	0.00	0.00	0.00	187
12	0.00	0.00	0.00	222
13	0.00	0.00	0.00	203
14	0.00	0.00	0.00	201
15	0.00	0.00	0.00	208
16	0.00	0.00	0.00	201
17	0.00	0.00	0.00	209
18	0.00	0.00	0.00	216
19	0.00	0.00	0.00	209
20	0.00	0.00	0.00	214
21	0.00	0.00	0.00	192
...				
accuracy			0.01	22800
macro avg	0.00	0.01	0.00	22800
weighted avg	0.00	0.01	0.00	22800

Data 2 Improvement 4:

	precision	recall	f1-score	support
0	0.15	0.21	0.18	214
1	0.44	0.24	0.31	202
2	0.02	0.01	0.01	206
3	0.14	0.03	0.06	201
4	0.31	0.21	0.25	192
5	0.26	0.09	0.13	199
6	0.50	0.48	0.49	195
7	0.38	0.53	0.44	178
8	0.17	0.09	0.12	229
9	0.15	0.16	0.15	206
10	0.49	0.44	0.46	200
11	0.06	0.02	0.03	187
12	0.24	0.23	0.23	222
13	0.51	0.46	0.49	203
14	0.33	0.41	0.36	201
15	0.23	0.40	0.29	208
16	0.58	0.56	0.57	201
17	0.28	0.11	0.15	209
18	0.87	0.84	0.85	216
19	0.22	0.39	0.28	209
20	0.16	0.24	0.20	214
21	0.25	0.17	0.20	192
...				
accuracy			0.30	22800
macro avg	0.29	0.30	0.28	22800
weighted avg	0.29	0.30	0.28	22800

LSTM

BASELINE
(LSTM 256)

DATA 1	DATA 2
Training: 0.9367	Training: 0.7860
Test: 0.4794	Test: 0.2413

IMPROVEMENT 1
(LSTM 128)

DATA 1	DATA 2
Training: 0.7712	Training: 0.4958
Test: 0.5023	Test: 0.2646

MinMaxScaler

IMPROVEMENT 2
(Normalization)

DATA 1	DATA 2
Training: 0.7856	Training: 0.4836
Test: 0.5095	Test: 0.2733

Bidirectional

IMPROVEMENT 3
(Bidirectional LSTM)

DATA 1	DATA 2
Training: 0.6003	Training: 0.3472
Test: 0.5665	Test: 0.3065

Dropout

IMPROVEMENT 4
(Bidirectional LSTM with Dropout)

DATA 1	DATA 2
Training: 0.5916	Training: 0.3381
Test: 0.5655	Test: 0.3028

LSTM

Data 1 Baseline:

	precision	recall	f1-score	support
0	0.32	0.33	0.32	1005
1	0.70	0.65	0.68	1057
2	0.47	0.43	0.45	985
3	0.77	0.78	0.77	967
4	0.41	0.46	0.44	969
5	0.54	0.54	0.54	1029
6	0.37	0.33	0.35	1012
7	0.43	0.41	0.42	996
8	0.34	0.37	0.35	982
9	0.45	0.48	0.47	998
accuracy			0.48	10000
macro avg	0.48	0.48	0.48	10000
weighted avg	0.48	0.48	0.48	10000

Data 1 Improvement 3:

	precision	recall	f1-score	support
0	0.42	0.40	0.41	1005
1	0.73	0.71	0.72	1057
2	0.51	0.59	0.55	985
3	0.83	0.86	0.84	967
4	0.56	0.46	0.51	969
5	0.70	0.54	0.61	1029
6	0.45	0.44	0.45	1012
7	0.53	0.52	0.53	996
8	0.44	0.49	0.47	982
9	0.53	0.65	0.59	998
accuracy			0.57	10000
macro avg	0.57	0.57	0.57	10000
weighted avg	0.57	0.57	0.57	10000

Data 2 Baseline:

	precision	recall	f1-score	support
0	0.11	0.10	0.10	214
1	0.24	0.22	0.23	202
2	0.07	0.08	0.07	206
3	0.12	0.14	0.13	201
4	0.23	0.26	0.24	192
5	0.11	0.09	0.10	199
6	0.45	0.48	0.47	195
7	0.31	0.37	0.33	178
8	0.15	0.14	0.14	229
9	0.11	0.09	0.10	206
10	0.32	0.40	0.36	200
11	0.07	0.07	0.07	187
12	0.20	0.15	0.17	222
13	0.41	0.36	0.39	203
14	0.29	0.29	0.29	201
15	0.10	0.09	0.09	208
16	0.44	0.55	0.49	201
17	0.19	0.14	0.17	209
18	0.86	0.82	0.84	216
19	0.41	0.36	0.38	209
20	0.33	0.29	0.31	214
21	0.20	0.20	0.20	192
...				
accuracy			0.24	22800
macro avg	0.24	0.24	0.24	22800
weighted avg	0.24	0.24	0.24	22800

Data 2 Improvement 3:

	precision	recall	f1-score	support
0	0.19	0.15	0.17	214
1	0.32	0.31	0.31	202
2	0.12	0.04	0.06	206
3	0.12	0.07	0.09	201
4	0.31	0.40	0.35	192
5	0.14	0.10	0.12	199
6	0.59	0.51	0.55	195
7	0.44	0.53	0.48	178
8	0.11	0.02	0.04	229
9	0.16	0.29	0.21	206
10	0.37	0.46	0.41	200
11	0.07	0.03	0.04	187
12	0.25	0.20	0.22	222
13	0.49	0.52	0.50	203
14	0.34	0.38	0.36	201
15	0.23	0.19	0.21	208
16	0.51	0.59	0.55	201
17	0.33	0.18	0.23	209
18	0.95	0.84	0.89	216
19	0.21	0.28	0.24	209
20	0.21	0.21	0.21	214
21	0.28	0.26	0.27	192
...				
accuracy			0.30	22800
macro avg	0.29	0.30	0.28	22800
weighted avg	0.29	0.30	0.28	22800

Logistic Regression

BASELINE
(No normalization, no feature selection)

DATA 1	DATA 2
Training: 0.2220	Training: 0.0089
Test: 0.2187	Test: 0.0081

MinMaxScaler →

IMPROVEMENT 1
(Normalization)

DATA 1	DATA 2
Training: 0.5291	Training: 0.1887
Test: 0.5232	Test: 0.1853

SequentialFeatureSelector →

IMPROVEMENT 2
(Normalization + Feature Selection)

DATA 1	DATA 2
Training: 0.52455	Training: 0.1975
Test: 0.5245	Test: 0.1936

K + 10 →

IMPROVEMENT 3
(Normalization + Feature Selection)

DATA 1	DATA 2
Training: 0.5278	Training: 0.2034
Test: 0.5244	Test: 0.2009

Logistic Regression

Classification Report for Data 1:

	precision	recall	f1-score	support
Alternative	0.38	0.29	0.33	1005
Anime	0.62	0.59	0.61	1057
Blues	0.50	0.48	0.49	985
Classical	0.76	0.80	0.78	967
Country	0.44	0.56	0.49	969
Electronic	0.61	0.59	0.60	1029
Hip-Hop	0.46	0.47	0.46	1012
Jazz	0.49	0.41	0.45	996
Rap	0.44	0.40	0.42	982
Rock	0.52	0.65	0.57	998
accuracy			0.52	10000
macro avg	0.52	0.52	0.52	10000
weighted avg	0.52	0.52	0.52	10000

Classification Report for Data 2:

	precision	recall	f1-score	support
acoustic	0.10	0.10	0.10	214
afrobeat	0.30	0.15	0.20	202
alt-rock	0.00	0.00	0.00	206
alternative	0.01	0.00	0.01	201
ambient	0.28	0.36	0.32	192
anime	0.02	0.01	0.01	199
black-metal	0.34	0.44	0.38	195
bluegrass	0.23	0.20	0.21	178
blues	0.00	0.00	0.00	229
brazil	0.06	0.01	0.02	206
breakbeat	0.18	0.18	0.18	200
british	0.25	0.01	0.02	187
cantopop	0.15	0.14	0.15	222
chicago-house	0.30	0.42	0.35	203
children	0.25	0.35	0.29	201
chill	0.15	0.14	0.14	208
classical	0.43	0.43	0.43	201
club	0.15	0.05	0.07	209
comedy	0.85	0.84	0.85	216
country	0.13	0.25	0.17	209
dance	0.09	0.07	0.08	214
dancehall	0.11	0.13	0.12	192
death-metal	0.22	0.23	0.22	185
deep-house	0.10	0.19	0.13	208
detroit-techno	0.36	0.45	0.40	201
disco	0.12	0.08	0.10	224
disney	0.30	0.18	0.22	203
reggae	0.09	0.04	0.06	180
reggaeton	0.07	0.06	0.07	213
rock	0.03	0.02	0.02	192
rock-n-roll	0.12	0.13	0.13	183
rockabilly	0.12	0.11	0.11	188
romance	0.37	0.62	0.47	191
sad	0.18	0.35	0.24	187
salsa	0.24	0.51	0.32	198
samba	0.15	0.14	0.15	195
sertanejo	0.28	0.44	0.34	227
show-tunes	0.20	0.13	0.16	216
singer-songwriter	0.08	0.04	0.05	187
ska	0.11	0.13	0.12	194
sleep	0.67	0.67	0.67	215
songwriter	0.09	0.04	0.05	197
soul	0.07	0.05	0.06	172
spanish	0.00	0.00	0.00	230
study	0.50	0.80	0.62	202
swedish	0.00	0.00	0.00	188
synth-pop	0.09	0.15	0.11	179
tango	0.39	0.62	0.48	210
techno	0.18	0.09	0.12	197
trance	0.11	0.13	0.12	166
trip-hop	0.19	0.09	0.13	191
turkish	0.09	0.07	0.08	193
world-music	0.13	0.25	0.17	207
accuracy			0.20	22800
macro avg	0.17	0.20	0.17	22800
weighted avg	0.17	0.20	0.17	22800

CNN - (Convolutional Neural Network)

BASELINE	
DATA 1	**DATA 2**
Training: 0.5592	Training: 0.3202
Test: 0.55919	Test: 0.3068

Dropout & Early Stoppage

IMPROVEMENT 1 & 2	
DATA 1	**DATA 2**
Training: 0.5193	Training: 0.3494
Test: 0.5424	Test: 0.31

CNN

Classification Report for Data 1:

	precision	recall	f1-score	support
Alternative	0.46	0.27	0.34	1005
Anime	0.74	0.66	0.70	1057
Blues	0.58	0.44	0.50	985
Classical	0.79	0.85	0.82	967
Country	0.53	0.47	0.50	969
Electronic	0.68	0.57	0.62	1029
Hip-Hop	0.39	0.30	0.34	1012
Jazz	0.46	0.52	0.48	996
Rap	0.42	0.63	0.51	982
Rock	0.46	0.73	0.56	998
accuracy			0.54	10000
macro avg	0.55	0.54	0.54	10000
weighted avg	0.55	0.54	0.54	10000

Classification Report for Data 2:

	precision	recall	f1-score	support
0	0.21	0.14	0.17	214
1	0.36	0.14	0.20	202
2	0.15	0.10	0.12	206
3	0.12	0.08	0.09	201
4	0.32	0.40	0.35	192
5	0.21	0.24	0.23	199
6	0.50	0.55	0.52	195
7	0.35	0.56	0.43	178
8	0.16	0.07	0.09	229
9	0.18	0.24	0.21	206
10	0.55	0.42	0.48	200
11	0.10	0.03	0.04	187
12	0.31	0.22	0.26	222
13	0.40	0.60	0.48	203
14	0.38	0.40	0.39	201
15	0.27	0.27	0.27	208
16	0.58	0.56	0.57	201
17	0.40	0.21	0.27	209
18	0.92	0.83	0.88	216
19	0.37	0.33	0.35	209
20	0.29	0.20	0.24	214
21	0.26	0.26	0.26	192
22	0.28	0.32	0.30	185
23	0.21	0.18	0.19	208
24	0.49	0.29	0.36	201
25	0.15	0.08	0.10	224
26	0.32	0.25	0.28	203
88	0.09	0.07	0.08	180
89	0.18	0.44	0.25	213
90	0.13	0.23	0.17	192
91	0.26	0.40	0.32	183
92	0.27	0.30	0.28	188
93	0.62	0.70	0.65	191
94	0.24	0.28	0.25	187
95	0.49	0.67	0.56	198
96	0.36	0.31	0.34	195
97	0.36	0.68	0.47	227
98	0.26	0.18	0.21	216
99	0.13	0.12	0.12	187
100	0.16	0.21	0.18	194
101	0.86	0.70	0.77	215
102	0.07	0.02	0.03	197
103	0.23	0.34	0.28	172
104	0.20	0.09	0.12	230
105	0.62	0.73	0.67	202
106	0.15	0.06	0.09	188
107	0.19	0.19	0.19	179
108	0.60	0.83	0.69	210
109	0.16	0.05	0.07	197
110	0.27	0.28	0.28	166
111	0.20	0.18	0.19	191
112	0.22	0.44	0.29	193
113	0.23	0.42	0.30	207
accuracy			0.30	22800
macro avg	0.29	0.30	0.29	22800
weighted avg	0.29	0.30	0.29	22800

CNN-LSTM

BASELINE	
DATA 1	**DATA 2**
Training: 0.5439	Training: 0.2736
Test: 0.5401	Test: 0.2588

Dropout & Early Stoppage

IMPROVEMENT 1 & 2	
DATA 1	**DATA 2**
Training: 0.5128	Training: 0.3466
Test: 0.546	Test: 0.3041

CNN-LSTM

Classification Report for Data 1:

	precision	recall	f1-score	support
0	0.45	0.27	0.34	1005
1	0.70	0.68	0.69	1057
2	0.58	0.45	0.50	985
3	0.79	0.85	0.82	967
4	0.46	0.58	0.51	969
5	0.67	0.56	0.61	1029
6	0.44	0.69	0.54	1012
7	0.51	0.47	0.49	996
8	0.43	0.21	0.28	982
9	0.46	0.71	0.56	998
accuracy			0.55	10000
macro avg	0.55	0.55	0.53	10000
weighted avg	0.55	0.55	0.53	10000

Classification Report for Data 2:

	precision	recall	f1-score	support
0	0.17	0.22	0.19	214
1	0.26	0.26	0.26	202
2	0.08	0.04	0.06	206
3	0.19	0.07	0.10	201
4	0.34	0.34	0.34	192
5	0.20	0.12	0.15	199
6	0.49	0.55	0.52	195
7	0.35	0.50	0.41	178
8	0.26	0.20	0.23	229
9	0.15	0.11	0.13	206
10	0.49	0.46	0.47	200
11	0.05	0.02	0.03	187
12	0.33	0.23	0.27	222
13	0.46	0.51	0.49	203
14	0.42	0.31	0.35	201
15	0.18	0.20	0.19	208
16	0.58	0.57	0.58	201
17	0.31	0.23	0.27	209
18	0.86	0.85	0.86	216
19	0.29	0.35	0.32	209
20	0.23	0.36	0.28	214
21	0.22	0.30	0.25	192
22	0.32	0.31	0.32	185
23	0.19	0.36	0.25	208
24	0.51	0.18	0.27	201
25	0.18	0.14	0.16	224
26	0.34	0.23	0.27	203
88	0.09	0.07	0.08	180
89	0.18	0.44	0.25	213
90	0.13	0.23	0.17	192
91	0.26	0.40	0.32	183
92	0.27	0.30	0.28	188
93	0.62	0.70	0.65	191
94	0.24	0.28	0.25	187
95	0.49	0.67	0.56	198
96	0.36	0.31	0.34	195
97	0.36	0.68	0.47	227
98	0.26	0.18	0.21	216
99	0.13	0.12	0.12	187
100	0.16	0.21	0.18	194
101	0.86	0.70	0.77	215
102	0.07	0.02	0.03	197
103	0.23	0.34	0.28	172
104	0.20	0.09	0.12	230
105	0.62	0.73	0.67	202
106	0.15	0.06	0.09	188
107	0.19	0.19	0.19	179
108	0.60	0.83	0.69	210
109	0.16	0.05	0.07	197
110	0.27	0.28	0.28	166
111	0.20	0.18	0.19	191
112	0.22	0.44	0.29	193
113	0.23	0.42	0.30	207
accuracy			0.30	22800
macro avg	0.29	0.30	0.29	22800
weighted avg	0.29	0.30	0.29	22800

XG Boost

- Utilized to assist with goal of developing a robust music genre classification system.
- In the case of music genre classification, precision and accuracy and important components that XG Boost provides us with
- Feature Importance Analysis: scores allowing us to determine which features contribute most to genre classification
- Handling missing values: model is able to use available data for training without extensive preprocessing
- Baseline Model:
  - Data 1 Accuracy: 0.581
  - Data 2 Accuracy: 0.324
- Improved model (GridSearchCV):
  - Data 1 Accuracy: 0.596
  - Data 1 Accuracy: 0.622
  - `'learning_rate': 0.05, 'max_depth': 5, 'n_estimators': 300`

Accuracy for Data 1: 0.5807580758075808
Classification Report for Data 1:

	precision	recall	f1-score	support
0	0.46	0.41	0.43	1027
1	0.83	0.74	0.78	1032
2	0.63	0.57	0.60	1013
3	0.85	0.86	0.85	947
4	0.58	0.60	0.59	995
5	0.69	0.65	0.67	1007
6	0.38	0.38	0.38	1005
7	0.55	0.56	0.55	972
8	0.37	0.38	0.37	997
9	0.52	0.69	0.60	1004
accuracy			0.58	9999
macro avg	0.59	0.58	0.58	9999
weighted avg	0.59	0.58	0.58	9999

Baseline Model Data 1

Best parameters for Data 1: {'learning_rate': 0.05, 'max_depth': 5, 'n_estimators': 300}
Accuracy for Data 1: 0.596059605960596
Classification Report for Data 1:

	precision	recall	f1-score	support
0	0.50	0.39	0.44	1027
1	0.83	0.74	0.78	1032
2	0.64	0.56	0.60	1013
3	0.85	0.86	0.85	947
4	0.59	0.62	0.61	995
5	0.68	0.65	0.67	1007
6	0.42	0.43	0.43	1005
7	0.56	0.55	0.55	972
8	0.42	0.42	0.42	997
9	0.52	0.76	0.62	1004
accuracy			0.60	9999
macro avg	0.60	0.60	0.60	9999
weighted avg	0.60	0.60	0.60	9999

Accuracy for Data 2: 0.32394736842105265
Classification Report for Data 2:

	precision	recall	f1-score	support
0	0.26	0.20	0.23	213
1	0.35	0.33	0.34	203
2	0.08	0.07	0.08	215
3	0.13	0.15	0.14	184
4	0.31	0.34	0.33	197
5	0.19	0.18	0.19	193
6	0.59	0.58	0.58	210
7	0.47	0.53	0.50	205
8	0.15	0.14	0.15	214
9	0.09	0.11	0.10	197
10	0.51	0.44	0.47	199
11	0.10	0.05	0.07	214
12	0.28	0.26	0.27	193
13	0.57	0.59	0.58	206
14	0.58	0.46	0.51	214
15	0.22	0.22	0.22	198
16	0.51	0.55	0.53	198
17	0.31	0.27	0.29	191
18	0.90	0.84	0.87	191
19	0.64	0.57	0.60	208
20	0.38	0.44	0.41	206
accuracy			0.32	22800
macro avg	0.32	0.33	0.32	22800
weighted avg	0.32	0.32	0.32	22800

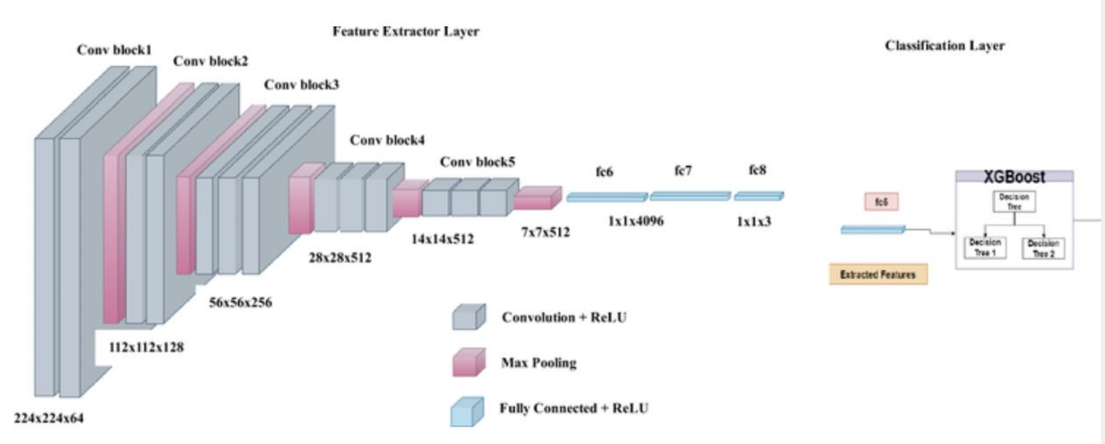
Baseline Model Data 2

Accuracy for Data 2 subset: 0.622
Classification Report for Data 2 subset:

	precision	recall	f1-score	support
0	0.46	0.46	0.46	173
1	0.51	0.54	0.52	179
2	0.62	0.62	0.62	226
3	0.54	0.53	0.53	203
4	0.55	0.62	0.58	214
5	0.54	0.61	0.57	199
6	0.93	0.90	0.92	211
7	0.67	0.76	0.71	202
8	0.48	0.41	0.44	208
9	0.57	0.57	0.57	204
10	0.65	0.56	0.60	194
11	0.74	0.70	0.72	203
12	0.86	0.79	0.83	204
13	0.45	0.37	0.41	209
14	0.44	0.43	0.44	194
15	0.76	0.81	0.79	193
16	0.71	0.78	0.75	202
17	0.92	0.94	0.93	200
18	0.46	0.49	0.47	194
19	0.52	0.48	0.50	188
accuracy			0.62	4000
macro avg	0.62	0.62	0.62	4000
weighted avg	0.62	0.62	0.62	4000

Improved Model Data 1 & 2

XG Boost & CNN




```
accuracy = accuracy_score(y_test, xgb_predictions)
print("Accuracy: %.2f%%" % (accuracy * 100.0))
```

- Decided to created a combined model using XG Boost and CNN (Convolutional Neural Network) to leverage the strengths of both models
  - Feature Extraction with CNN: using to capture more details within the data
  - Integration with XG Boost: once the features are extracted from CNN, they can be fed into XG Boost for further classification
  - Ensemble Learning
  - Powerful tool to extract hierarchical features from structured data

XG Boost and CNN Continued

```
Epoch 1/5
1250/1250 [==============================] - 15s 11ms/step - loss: 1.4313 -
accuracy: 0.4546 - val_loss: 1.3242 - val_accuracy: 0.5007
Epoch 2/5
1250/1250 [==============================] - 13s 11ms/step - loss: 1.2329 -
accuracy: 0.5270 - val_loss: 1.2126 - val_accuracy: 0.5350
Epoch 3/5
1250/1250 [==============================] - 13s 11ms/step - loss: 1.1916 -
accuracy: 0.5404 - val_loss: 1.1792 - val_accuracy: 0.5455
Epoch 4/5
1250/1250 [==============================] - 13s 11ms/step - loss: 1.1626 -
accuracy: 0.5486 - val_loss: 1.1621 - val_accuracy: 0.5488
Epoch 5/5
1250/1250 [==============================] - 13s 10ms/step - loss: 1.1382 -
accuracy: 0.5576 - val_loss: 1.1429 - val_accuracy: 0.5537
```

Data 1 Training and Validation Accuracy per Epoch and Cross Entropy Loss

```
Epoch 1/5
2850/2850 [==============================] - 40s 14ms/step - loss: 3.3245 -
accuracy: 0.1913 - val_loss: 2.9657 - val_accuracy: 0.2463
Epoch 2/5
2850/2850 [==============================] - 39s 14ms/step - loss: 2.8602 -
accuracy: 0.2641 - val_loss: 2.7793 - val_accuracy: 0.2746
Epoch 3/5
2850/2850 [==============================] - 39s 14ms/step - loss: 2.7075 -
accuracy: 0.2913 - val_loss: 2.6889 - val_accuracy: 0.2948
Epoch 4/5
2850/2850 [==============================] - 39s 14ms/step - loss: 2.6173 -
accuracy: 0.3080 - val_loss: 2.6574 - val_accuracy: 0.2981
Epoch 5/5
2850/2850 [==============================] - 39s 14ms/step - loss: 2.5495 -
accuracy: 0.3185 - val_loss: 2.6054 - val_accuracy: 0.3116
```

Data 2 Training and Validation Accuracy per Epoch and Cross Entropy Loss

Accuracy for Data 1: 0.6170212765957447
Classification Report for Data 1:

	precision	recall	f1-score	support
0	0.46	0.50	0.48	962
1	0.80	0.77	0.78	1033
2	0.55	0.52	0.54	960
3	0.56	0.62	0.59	1021
4	0.65	0.58	0.61	1018
5	0.57	0.57	0.57	1019
6	0.75	0.82	0.78	977
7	0.00	0.00	0.00	107
accuracy			0.62	7097
macro avg	0.54	0.55	0.54	7097
weighted avg	0.61	0.62	0.61	7097

Data 1 Results

Data Accuracy = 0.617%

Possible AI Fairness Biases/Constraints/Limitations

- Possible biases include selection bias in the datasets and model biases that result from training algorithms
- Given that the data is sourced from Kaggle, there is potential for this data to underrepresented genres of music
- Can result in skewed predictions, causing the model to be less effective at classifying genres
- Main constraints include data availability and computational resources.
- Limitations of our study include:
  - Data quality and diversity
  - Generalization
  - Interpretability

Standards/Comparisons/Future Work

- Standard: Python 3.10.9, sklearn, keras, matplotlib, pandas
- Constraint:
  - Limited Time for models with more hyperparameter choices
  - Limited Hardware Capacity: Deepnote, Google Colab, VScode all experienced “Kernel Timeout Error” when running certain models, so had to switch to another platform
  - Limited Money for access to machines with higher capacity and bigger dataset
- Evolving nature of music genres and subjectivity can pose as potential challenges
- In comparing our models to existing approaches in the market, this diverse ensemble approach promises a better mechanism of capturing the complex relationships present within music data in comparison to single model systems
- Future additions could include: diversification, user centric adaptations, and model explainability.

KEY RESULTS OF MODELS

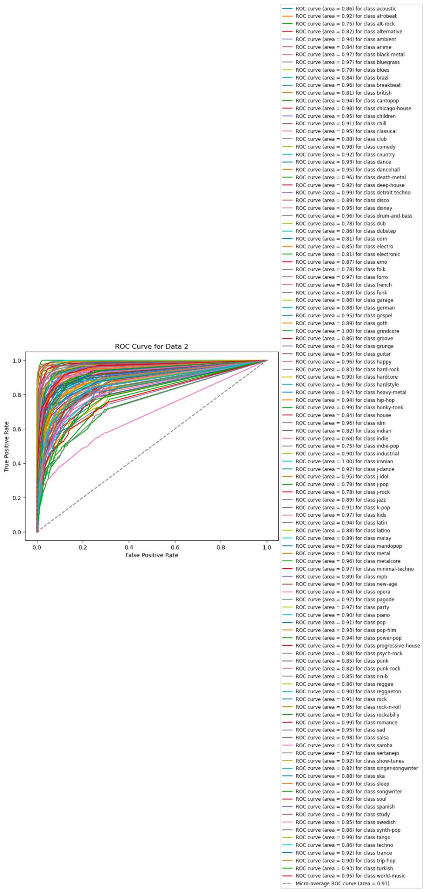
Success:

1) Random Forest – Data 2 (82%) accuracy.

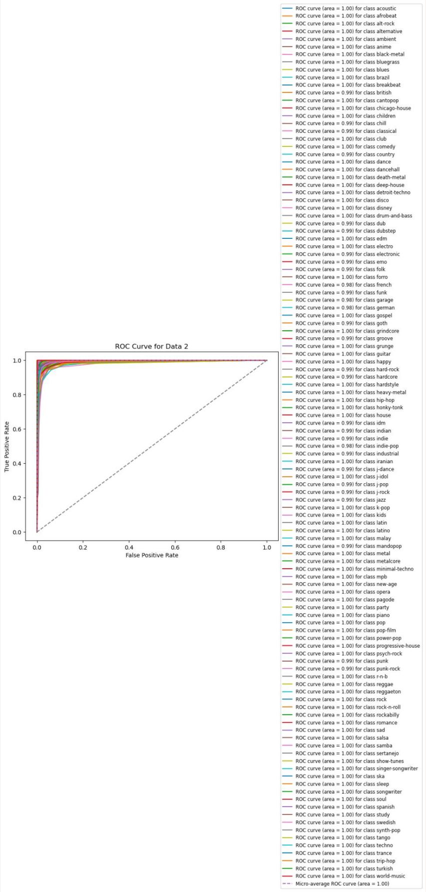
Did not meet threshold:

2) Transformer Model
3) K-Nearest Model
4) ANN
5) LSTM
6) Logistic Regression
7) CNN
8) CNN-LSTM
9) XGBoost
10) XGBoost-CNN Hybrid

Appendix – Results



Random Forest Classifier Data 2 Baseline Model



Random Forest Classifier Data 2 Improved Model

Appendix – Results

Accuracy for Data 2: 0.34
Classification Report for Data 2:

	precision	recall	f1-score	support
acoustic	0.28	0.24	0.26	213
afrobeat	0.36	0.37	0.37	203
alt-rock	0.06	0.06	0.06	215
alternative	0.11	0.13	0.12	184
ambient	0.35	0.35	0.35	197
anime	0.19	0.16	0.17	193
black-metal	0.58	0.65	0.61	210
bluegrass	0.41	0.60	0.49	205
blues	0.17	0.13	0.15	214
brazil	0.04	0.04	0.04	197
breakbeat	0.57	0.50	0.53	199
british	0.17	0.09	0.12	214
cantopop	0.32	0.33	0.32	193
chicago-house	0.57	0.58	0.57	206
children	0.53	0.51	0.52	214
chill	0.27	0.27	0.27	198
classical	0.56	0.56	0.56	198
club	0.37	0.24	0.29	191
comedy	0.89	0.85	0.87	191
country	0.61	0.61	0.61	208
dance	0.37	0.44	0.40	206
dancehall	0.27	0.27	0.27	204
death-metal	0.29	0.40	0.33	191
deep-house	0.17	0.24	0.20	186
detroit-techno	0.58	0.56	0.57	191
disco	0.27	0.21	0.23	189
disney	0.52	0.42	0.46	197
drum-and-bass	0.69	0.69	0.69	199
dub	0.10	0.09	0.10	214
dubstep	0.12	0.15	0.14	212
edm	0.06	0.06	0.06	196
electro	0.26	0.22	0.24	204
electronic	0.08	0.04	0.05	210
emo	0.26	0.28	0.27	197
folk	0.16	0.12	0.14	224
forro	0.39	0.60	0.48	194
french	0.34	0.26	0.30	193
funk	0.33	0.33	0.33	214
garage	0.21	0.16	0.18	201
german	0.31	0.19	0.23	204
gospel	0.30	0.41	0.34	198
goth	0.15	0.08	0.11	215
grindcore	0.86	0.87	0.86	225
groove	0.11	0.07	0.09	213
grunge	0.21	0.22	0.22	188
guitar	0.34	0.30	0.32	191
happy	0.45	0.47	0.46	211
hard-rock	0.11	0.10	0.11	206
hardcore	0.32	0.33	0.32	203
hardstyle	0.50	0.50	0.50	199
heavy-metal	0.33	0.40	0.36	198
hip-hop	0.30	0.33	0.32	186
honky-tonk	0.80	0.81	0.81	203
house	0.14	0.13	0.13	220
idm	0.68	0.48	0.56	213
indian	0.17	0.16	0.16	196
indie	0.09	0.07	0.08	227
indie-pop	0.13	0.11	0.11	208
industrial	0.22	0.19	0.20	194
iranian	0.71	0.85	0.77	194
j-dance	0.41	0.48	0.44	203
j-idol	0.51	0.62	0.56	194
j-pop	0.12	0.10	0.11	212
j-rock	0.10	0.09	0.10	207
jazz	0.45	0.41	0.43	224
k-pop	0.35	0.39	0.37	204
kids	0.62	0.67	0.65	202
latin	0.21	0.22	0.21	204
latino	0.12	0.08	0.10	213
malay	0.26	0.19	0.22	205
mandopop	0.25	0.30	0.27	193
metal	0.14	0.14	0.14	192
metalcore	0.32	0.37	0.34	188
minimal-techno	0.36	0.41	0.38	227
mpb	0.10	0.13	0.11	190
new-age	0.55	0.49	0.52	229
opera	0.40	0.49	0.44	185
pagode	0.43	0.57	0.49	186
party	0.49	0.59	0.54	198
piano	0.45	0.42	0.43	199
pop	0.28	0.30	0.29	179
pop-film	0.30	0.47	0.37	204
power-pop	0.42	0.46	0.44	222
progressive-house	0.25	0.30	0.27	171
psych-rock	0.22	0.17	0.19	199
punk	0.09	0.10	0.09	196
punk-rock	0.07	0.07	0.07	194
r-n-b	0.14	0.10	0.12	194
reggae	0.09	0.11	0.10	188
reggaeton	0.12	0.12	0.12	212
rock	0.31	0.28	0.29	199
rock-n-roll	0.28	0.39	0.33	193
rockabilly	0.29	0.22	0.25	184
romance	0.64	0.85	0.73	201
sad	0.29	0.37	0.32	191
salsa	0.51	0.78	0.62	190
samba	0.37	0.42	0.39	209
sertanejo	0.45	0.57	0.50	196
show-tunes	0.36	0.29	0.32	196
singer-songwriter	0.04	0.05	0.05	179
ska	0.21	0.25	0.23	212
sleep	0.90	0.85	0.88	174
songwriter	0.04	0.04	0.04	196
soul	0.43	0.37	0.40	197
spanish	0.20	0.09	0.12	188
study	0.60	0.88	0.71	182
swedish	0.20	0.12	0.15	176
synth-pop	0.35	0.23	0.27	200
tango	0.70	0.86	0.77	207
techno	0.15	0.13	0.14	197
trance	0.41	0.42	0.42	184
trip-hop	0.29	0.24	0.26	189
turkish	0.26	0.27	0.27	212
world-music	0.37	0.44	0.40	177
accuracy			0.34	22800
macro avg	0.33	0.34	0.33	22800
weighted avg	0.33	0.34	0.33	22800

Random Forest Classifier Data 2 Baseline Model

Appendix – Results

Accuracy for Data 2: 0.32394736842105265
Classification Report for Data 2:

	precision	recall	f1-score	support
0	0.26	0.20	0.23	213
1	0.35	0.33	0.34	203
2	0.08	0.07	0.08	215
3	0.13	0.15	0.14	184
4	0.31	0.34	0.33	197
5	0.19	0.18	0.19	193
6	0.59	0.58	0.58	210
7	0.47	0.53	0.50	205
8	0.15	0.14	0.15	214
9	0.09	0.11	0.10	197
10	0.51	0.44	0.47	199
11	0.10	0.05	0.07	214
12	0.28	0.26	0.27	193
13	0.57	0.59	0.58	206
14	0.58	0.46	0.51	214
15	0.22	0.22	0.22	198
16	0.51	0.55	0.53	198
17	0.31	0.27	0.29	191
18	0.90	0.84	0.87	191
19	0.64	0.57	0.60	208
20	0.38	0.44	0.41	206
21	0.27	0.28	0.28	204
22	0.31	0.40	0.35	191
23	0.19	0.26	0.22	186
24	0.54	0.50	0.52	191
25	0.27	0.24	0.25	189
26	0.46	0.44	0.45	197
27	0.66	0.66	0.66	199
28	0.08	0.08	0.08	214
29	0.16	0.20	0.18	212
30	0.07	0.08	0.07	196
31	0.25	0.22	0.23	204
32	0.09	0.05	0.07	210
33	0.25	0.19	0.22	197
34	0.12	0.09	0.10	224
35	0.47	0.52	0.49	194
36	0.25	0.22	0.23	193
37	0.33	0.36	0.34	214
38	0.22	0.21	0.21	201
39	0.29	0.22	0.25	204
40	0.30	0.38	0.34	198
41	0.17	0.15	0.16	215
42	0.91	0.82	0.86	225
43	0.15	0.11	0.13	213
44	0.19	0.23	0.21	188
45	0.35	0.39	0.37	191
46	0.43	0.46	0.44	211
47	0.12	0.11	0.11	206
48	0.32	0.31	0.32	203
49	0.51	0.45	0.48	199
50	0.32	0.33	0.33	198
51	0.31	0.33	0.32	186
52	0.82	0.80	0.81	203
53	0.18	0.18	0.18	220
54	0.50	0.49	0.49	213
55	0.13	0.16	0.14	196
56	0.07	0.05	0.06	227
57	0.13	0.10	0.12	208
58	0.21	0.19	0.20	194
59	0.66	0.67	0.66	194
60	0.44	0.47	0.45	203
61	0.56	0.56	0.56	194
62	0.11	0.10	0.10	212
63	0.11	0.10	0.10	207
64	0.43	0.40	0.41	224
65	0.37	0.37	0.37	204
66	0.67	0.61	0.64	202
67	0.20	0.22	0.21	204
68	0.11	0.09	0.10	213
69	0.23	0.20	0.21	205
70	0.25	0.34	0.29	193
71	0.13	0.14	0.13	192
72	0.29	0.30	0.29	188
73	0.38	0.36	0.37	227
74	0.14	0.19	0.16	190
75	0.50	0.40	0.44	229
76	0.43	0.43	0.43	185
77	0.43	0.50	0.46	186
78	0.48	0.51	0.50	198
79	0.42	0.38	0.40	199
80	0.25	0.30	0.28	179
81	0.35	0.42	0.38	204
82	0.44	0.44	0.44	222
83	0.23	0.29	0.26	171
84	0.24	0.25	0.25	199
85	0.07	0.07	0.07	196
86	0.07	0.09	0.08	194
87	0.17	0.14	0.16	194
88	0.07	0.08	0.08	188
89	0.09	0.08	0.09	212
90	0.21	0.22	0.22	199
91	0.30	0.36	0.33	193
92	0.28	0.29	0.28	184
93	0.68	0.72	0.70	201
94	0.25	0.30	0.27	191
95	0.60	0.65	0.63	190
96	0.39	0.39	0.39	209
97	0.46	0.58	0.51	196
98	0.35	0.36	0.36	196
99	0.06	0.06	0.06	179
100	0.22	0.24	0.23	212
101	0.88	0.84	0.86	174
102	0.06	0.06	0.06	196
103	0.36	0.31	0.33	197
104	0.21	0.17	0.19	188
105	0.68	0.75	0.71	182
106	0.22	0.20	0.21	176
107	0.34	0.26	0.29	200
108	0.77	0.72	0.74	207
109	0.12	0.13	0.12	197
110	0.36	0.35	0.35	184
111	0.30	0.33	0.31	189
112	0.30	0.36	0.33	212
113	0.33	0.33	0.33	177
accuracy			0.32	22800
macro avg	0.32	0.33	0.32	22800
weighted avg	0.32	0.32	0.32	22800

70	0.25	0.34	0.29	193
71	0.13	0.14	0.13	192
72	0.29	0.30	0.29	188
73	0.38	0.36	0.37	227
74	0.14	0.19	0.16	190
75	0.50	0.40	0.44	229
76	0.43	0.43	0.43	185
77	0.43	0.50	0.46	186
78	0.48	0.51	0.50	198
79	0.42	0.38	0.40	199
80	0.25	0.30	0.28	179
81	0.35	0.42	0.38	204
82	0.44	0.44	0.44	222
83	0.23	0.29	0.26	171
84	0.24	0.25	0.25	199
85	0.07	0.07	0.07	196
86	0.07	0.09	0.08	194
87	0.17	0.14	0.16	194

XGboost Data 2 Baseline Model

Appendix – Results

Accuracy for Data 2: 0.82

Classification Report for Data 2:

	precision	recall	f1-score	support
acoustic	0.99	1.00	0.99	213
afrobeat	0.95	0.97	0.96	203
alt-rock	0.73	0.79	0.76	215
alternative	0.70	0.74	0.72	184
ambient	0.93	0.93	0.93	197
anime	0.91	0.84	0.87	193
black-metal	0.95	0.92	0.93	210
bluegrass	0.93	0.97	0.95	205
blues	0.75	0.71	0.73	214
brazil	0.76	0.87	0.82	197
breakbeat	0.94	0.94	0.94	199
british	0.78	0.68	0.73	214
cantopop	0.82	0.93	0.87	193
chicago-house	0.95	0.95	0.95	206
children	0.87	0.87	0.87	214
chill	0.85	0.85	0.85	198
classical	0.92	0.85	0.88	198
club	0.91	0.76	0.83	191
comedy	0.97	0.90	0.93	191
country	0.91	0.83	0.87	208
dance	0.87	0.90	0.88	206
dancehall	0.87	0.92	0.89	204
death-metal	0.86	0.95	0.90	192
deep-house	0.82	0.88	0.85	186
detroit-techno	0.97	0.98	0.98	194
disco	0.75	0.78	0.76	194
disney	0.87	0.90	0.88	199
drum-and-bass	0.90	0.91	0.90	194
dub	0.50	0.48	0.49	204
dubstep	0.59	0.64	0.61	226
edm	0.71	0.67	0.69	200
electro	0.62	0.73	0.67	207
electronic	0.65	0.55	0.60	200
emo	0.74	0.78	0.76	193
folk	0.53	0.56	0.55	212
forro	0.79	0.97	0.87	205
french	0.68	0.60	0.64	194
funk	0.78	0.68	0.73	218
garage	0.66	0.64	0.65	192
german	0.79	0.63	0.70	196
gospel	0.79	0.92	0.85	212
goth	0.78	0.70	0.74	223
grindcore	0.97	0.99	0.98	226
groove	0.74	0.67	0.70	212
grunge	0.74	0.77	0.75	202
guitar	0.76	0.88	0.82	189
happy	0.85	0.85	0.85	199
hard-rock	0.60	0.55	0.57	185
hardcore	0.72	0.73	0.72	205
hardstyle	0.85	0.89	0.87	210
heavy-metal	0.92	0.95	0.94	198
hip-hop	0.86	0.78	0.82	202
honky-tonk	0.94	0.96	0.95	205
house	0.75	0.76	0.76	230
idm	0.95	0.87	0.91	217
indian	0.51	0.58	0.54	202
indie	0.36	0.33	0.34	212
indie-pop	0.36	0.36	0.36	209
industrial	0.86	0.79	0.82	211
iranian	0.96	0.97	0.96	181
j-dance	0.92	0.83	0.87	213
j-idol	0.95	0.91	0.93	190
j-pop	0.66	0.62	0.64	224
j-rock	0.60	0.66	0.63	196
jazz	0.84	0.74	0.79	213
k-pop	0.89	0.84	0.86	213
kids	0.89	0.91	0.90	198
latin	0.82	0.85	0.84	210
latino	0.81	0.81	0.81	193
malay	0.77	0.79	0.78	189
mandopop	0.86	0.85	0.86	187
metal	0.69	0.78	0.73	186
metalcore	0.93	0.90	0.91	206
minimal-techno	0.95	0.95	0.95	208
mpb	0.71	0.84	0.77	191
new-age	0.93	0.89	0.91	199
opera	0.87	0.88	0.87	189
pagode	0.90	0.93	0.91	192
party	0.90	0.91	0.91	205
piano	0.95	0.78	0.85	197
pop	0.78	0.83	0.80	166
pop-film	0.81	0.86	0.83	213
power-pop	0.90	0.85	0.88	223
progressive-house	0.85	0.89	0.87	191
psych-rock	0.85	0.75	0.80	196
punk	0.61	0.59	0.60	221
punk-rock	0.54	0.61	0.57	192
r-n-b	0.71	0.65	0.68	198
reggae	0.79	0.78	0.79	194
reggaeton	0.84	0.83	0.84	199
rock	0.82	0.86	0.84	205
rock-n-roll	0.79	0.81	0.80	189
rockabilly	0.83	0.82	0.83	200
romance	0.98	0.96	0.97	213
sad	0.96	0.99	0.97	177
salsa	0.92	0.98	0.95	174
samba	0.93	0.91	0.92	189
sertanejo	0.95	0.99	0.97	188
show-tunes	0.96	0.91	0.94	198
singer-songwriter	0.60	0.66	0.63	194
ska	0.89	0.86	0.87	209
sleep	0.99	0.99	0.99	197
songwriter	0.61	0.64	0.62	187
soul	0.89	0.88	0.89	210
spanish	0.95	0.85	0.90	193
study	0.97	1.00	0.99	192
swedish	0.87	0.83	0.85	188
synth-pop	0.92	0.91	0.91	185
tango	0.97	0.99	0.98	184
techno	0.97	0.97	0.97	200
trance	0.97	0.98	0.98	191
trip-hop	0.95	0.98	0.96	183
turkish	0.99	0.95	0.97	185
world-music	0.99	0.99	0.99	188
accuracy			0.82	22800
macro avg	0.82	0.82	0.82	22800
weighted avg	0.82	0.82	0.82	22800

Random Forest Classifier Data 2 Improved Model