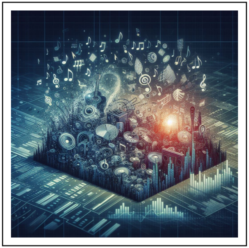
HARMONIZING SOUNDS

"Exploring Song Genre Classification through Audio Data Analysis"

ADARSH P (CEC22MCA-2002)

College of Engineering Cherthala
Department Of Computer Science and Engineering

June 14, 2024



Roll No: 2, S4, MCA

Guided by
Mrs. Vishnupriya G S

Table of Contents

INTRODUCTION

- **U**tilizes GTZAN dataset to classify 10 genres from classical to jazz to rock and electronic.
- **U**ses Mel Spectrograms for visual representation of audio files.
- **E**mploys Convolutional Neural Networks (CNNs) for genre classification.
- **M**ethodology includes experiments, feature extraction, model training, and performance evaluation.
- **F**indings are applied to recommendation systems and music organization.
- **A**ims to develop personalized recommendation algorithms for seamless navigation and discovery.
- **A**ims to deepen understanding of music's diversity and complexity.

OBJECTIVE

- **Understanding Audio Files:** Exploring the fundamental attributes of audio files through visualization techniques.
- **G**aining a deeper understanding of the structural and perceptual characteristics of audio signals.
- **Exploratory Data Analysis (EDA):** Employs Convolutional Neural Networks (CNNs) for genre classification.
- **C**onducting a thorough exploratory analysis of the dataset containing audio files.
- **I**dentifying trends, anomalies, and potential areas for further investigation.
- **Genres Classification:** Utilizing machine learning models to categorize audio files based on their genre.
- **L**everaging features extracted from provided CSV files containing statistical summaries of audio characteristics.
- **Recommender System:** Developing a recommender system capable of suggesting similar songs based on a given input.
- **E**nhancing music discovery and exploration for users.

LITERATURE SURVEY

An Empirical Study on Structured Dichotomies in Music Genre Classification (2015)

- Explores the application of ensemble learning and dichotomy-based methods for genre classification.
- Investigates various dichotomy structures of binary classifiers in music data.
- Utilizes base classifiers such as Support Vector Machines (SVM), Naive Bayes (NB), k-Nearest Neighbors (k-NN), and Logistic Regression (LR) in experiments.
- Findings suggest that dichotomy-based approaches do not significantly enhance genre classification.
- Discusses proposed heuristics for constructing Nested Dichotomy Trees (NDTs) and addresses the challenges and complexities of computational methods for genre classification in music.

An Empirical Study on Structured Dichotomies in Music Genre Classification (Contd.)

- **Advantages:**
  - Explores ensemble learning and dichotomy-based methods comprehensively.
  - Investigates various dichotomy structures and conducts empirical experiments on benchmarking music datasets.
  - Provides insights into future research directions for probabilistic analysis and alternative ranking algorithms.
- **Disadvantages:**
  - Initial results indicate that dichotomy-based approaches do not yield significant improvements in genre classification.
  - No ensemble methods demonstrate substantial performance gains over base classifiers.
  - Highlights challenges and complexities in computational methods for genre classification and acknowledges limitations of the proposed dichotomy-based approach.

An Empirical Study on Structured Dichotomies in Music Genre Classification (Contd.)
Problem Statement

- Despite extensive exploration of ensemble learning and dichotomy-based methods, the literature lacks a comprehensive understanding of their effectiveness in music genre classification. Therefore, there is a need for further research to address the limitations and challenges in improving genre classification accuracy using computational methods.

Music Genre Classification and Recommendation by Using Machine Learning Techniques (2018)

- Utilizes digital signal processing techniques for acoustic feature extraction.
- Employs machine learning techniques like convolutional neural networks and deep learning for genre classification.
- SVM algorithm shows the highest success rate in music genre classification.
- Utilizes the GTZAN dataset for experimentation and performance comparisons.

Music Genre Classification and Recommendation by Using Machine Learning Techniques (Contd.)

- **Advantages:**
  - Provides insights into digital signal processing techniques for acoustic feature extraction.
  - Uses machine learning methods like convolutional neural networks and deep learning.
  - Evaluates the performance of different algorithms and feature extraction methods.
  - Provides a standardized benchmark for comparison with other studies.
- **Disadvantages:**
  - Focuses solely on acoustic features, not considering lyrics or user preferences.
  - Only evaluates performance on the GTZAN dataset, not representative of all genres.
  - Doesn't provide a comprehensive comparison of different deep learning architectures or feature extraction methods.

Music Genre Classification and Recommendation by Using Machine Learning Techniques (Contd.)
Problem Statement

- Despite advancements in digital signal processing (DSP) and machine learning (ML) techniques, current methods predominantly focus on acoustic features and lack comprehensive consideration of other modalities such as lyrics or user preferences, potentially limiting the effectiveness of music genre classification and recommendation systems.

A Hybrid Model For Music Genre Classification Using LSTM And SVM (2018)

- Combines LSTM Neural Network and SVM classifiers for enhanced prediction accuracy.
- Implemented in two parts: training separately and combining results using sum rule.
- Uses GTZAN music database with 1000 music files for ten genres.
- Extracts 9 features from audio files for training models.
- Achieves 89 percentage accuracy in classifying ten genres, surpassing independent accuracies of LSTM and SVM classifiers.

A Hybrid Model For Music Genre Classification Using LSTM And SVM (Contd.)

- **Advantages:**
  - Combines LSTM Neural Network and SVM classifier for 89% accuracy in ten genres.
  - Leverages both classifiers' strengths for improved accuracy.
  - LSTM Neural Network handles long-term time dependencies in sequential data.
  - SVM classifier distinguishes between classes by creating a separation boundary.
- **Disadvantages:**
  - Doesn't discuss potential limitations like computational complexity or scaling challenges.
  - Lacks comprehensive analysis of potential failure modes.
  - Doesn't address computational resources required for training and deployment.
  - Doesn't discuss trade-offs in selecting and combining different machine learning models.

A Hybrid Model For Music Genre Classification Using LSTM And SVM (Contd.)
Problem Statement

- Despite achieving promising results, the hybrid model lacks discussion on potential limitations, failure modes, computational resources required, and trade-offs in model selection, hindering its applicability and generalization in real-world scenarios.

Exploring Data Augmentation to Improve Music Genre Classification with ConvNets (2018)

- Utilizes Convolutional Neural Networks (CNNs) for music genre classification.
- Uses data augmentation techniques to enhance accuracy.
- Experiments conducted using Latin Music Database (LMD).
- Data augmentation strategies included noise, pitch shifting, loudness variation, and time stretching.
- One-tone pitch shifting strategy significantly improved classification accuracy.
- Discusses the impact of different data augmentation strategies and compares results with state-of-the-art methods.

Exploring Data Augmentation to Improve Music Genre Classification with ConvNets (Contd.)

- **Advantages:**
  - Improves classification accuracy: Data augmentation techniques like noise, pitch shifting, loudness variation, and time stretching enhance accuracy in music genre classification using Convolutional Neural Networks (CNNs).
  - Enhances generalization capability: Data augmentation increases CNNs' ability to learn complex representations and hierarchical features.
  - Overcomes overfitting: Addresses thechallenge of overfitting in CNN-based classification tasks
- **Disadvantages:**
  - Specificity of Data Augmentation: Techniques used in other computer vision applications may distort important information in spectrograms.
  - Computational Complexity: CNNs using high-resolution images require more computational power and time.
  - Impact of Fusion Rules: Data augmentation depends on the fusion rule used to combine predictions from different patches.

Exploring Data Augmentation to Improve Music Genre Classification with ConvNets (Contd.)
Problem Statement

- Despite advancements in Convolutional Neural Networks (CNNs) and data augmentation techniques, challenges remain in effectively applying these methods to music genre classification tasks, including maintaining information fidelity during augmentation and addressing computational complexity.

Neural Network Music Genre Classification (2019)

- Machine learning techniques used for music genre classification.
- Success influenced by factors like song libraries, machine learning techniques, input formats, and neural network types.
- Different learning algorithms: supervised, unsupervised, semi-supervised, and reinforcement learning.
- Neural networks, especially convolutional neural networks (CNNs), extract critical features from complex datasets.
- Deep convolutional neural network implemented using TensorFlow for music genre classification.
- Importance of dataset preparation, model architecture, and optimization strategies.

Neural Network Music Genre Classification (Contd.)

- **Advantages:**
  - Automatic feature extraction: Neural networks, particularly CNNs, can automatically extract relevant features from spectrogram images.
  - Complex pattern recognition: They can learn complex patterns and relationships in music data, enabling accurate genre classification.
  - Scalability: Neural networks can accommodate a wide range of music samples for training and classification.
  - Adaptability: They can adapt to different music genres and styles, making them versatile for diverse music collections.
- **Disadvantages:**
  - Complexity: Designing and training neural networks can be time-consuming and require expertise in model architecture, hyperparameter tuning, and optimization.
  - Data requirements: Neural networks often require large amounts of labeled data for training, which may be challenging for certain genres.
  - Overfitting: Neural networks are prone to overfitting, necessitating techniques like dropout and regularization.

Neural Network Music Genre Classification (Contd.)
Problem Statement

- While neural networks, especially CNNs, show promise for music genre classification, their complex design, data requirements, and susceptibility to overfitting pose challenges in practical implementation and performance optimization.

Time-Frequency Analysis for Music Genre Classification by using Wavelet Package Decompositions (2019)

- Utilizes wavelet package decomposition (WPD) for music genre classification.
- Compares 1-D and 2-D WPD methods with Mel-Frequency Cepstral Coefficient (MFCC) approaches.
- Experiments show improved recognition rates with multiple singular values.
- Emphasizes WPD levels, sub-band selection, and singular value choice for superior performance.

Time-Frequency Analysis for Music Genre Classification by using Wavelet Package Decompositions (Contd.)

- **Advantages:**
  - Improved recognition rates compared to MFCC-based approach.
  - Multi-resolution characteristics and dimension reduction ability of singular value decomposition enhance effectiveness.
  - Offers valuable insights into time-frequency analysis using WPD for music genre classification.
  - Experiments conducted using ISMIR 2004 database enhance study's credibility.
- **Disadvantages:**
  - Lack of extensive discussion on potential limitations or challenges.
  - Lack of detailed comparison of computational complexity or resource requirements.
  - Absence of real-world implementation challenges.
  - Absence of potential drawbacks like sensitivity to noise or variability in music content.

Time-Frequency Analysis for Music Genre Classification by using Wavelet Package Decompositions (Contd.)
Problem Statement

- The literature lacks comprehensive discussion on the potential limitations, computational complexity, and real-world implementation challenges associated with time-frequency analysis methods like wavelet package decompositions for music genre classification.

Utilizes Deep Learning CNN for Multilingual Music Genre Classification (2020)

- Utilizes deep learning convolutional neural network to classify songs into metal, hip hop, and pop genres.
- Utilizes a dataset of sixty songs, each trimmed to five seconds, and converts them into spectrograms for analysis.
- Achieves a high classification accuracy of 93.3% for all twelve languages.
- Addresses bias in previous approaches by generating a multilingual dataset.
- Results section presents the image classifier's output and accuracy, emphasizing successful classification of multilingual songs.
- References related work in deep learning, spectrogram analysis, and music genre classification.

Utilizes Deep Learning CNN for Multilingual Music Genre Classification (Contd.)

- **Advantages:**
  - Achieved 93.3% classification accuracy for songs in twelve languages.
  - Utilized a diverse dataset of sixty songs, each trimmed to five seconds.
  - Addressed the gap in existing approaches by generating a multilingual dataset.
  - Extracted features from audio files using spectrograms and signal processing techniques.
- **Disadvantages:**
  - Low prediction confidence level indicated potential misclassifications.
  - Challenges in selecting a machine learning algorithm and converting sound files into images.
  - Research gap in multilingual music genre classification.
  - Requirement of significant computational resources and expertise.

Utilizes Deep Learning CNN for Multilingual Music Genre Classification (Contd.)
Problem Statement

- Despite achieving high classification accuracy, challenges remain in multilingual music genre classification, including low prediction confidence levels, algorithm selection, and computational requirements, indicating the need for further research in this area.

Combined Transfer and Active Learning for High Accuracy Music Genre Classification Method (2021)

- Introduces a novel approach for musical genre classification.
- Compares ATMGCM with traditional methods like SVM and RF.
- Details the feature engineering process, including DFT and music attributes.
- Discusses the application of transfer learning and active learning technologies for challenging music classification scenarios.

Combined Transfer and Active Learning for High Accuracy Music Genre Classification Method (Contd.)

- **Advantages:**
  - High accuracy in classifying music genres.
  - Significant decrease in niche genres.
  - Requires labeling only 10 to 15 percentage of unlabeled data.
  - Utilizes transfer learning and active learning techniques.
- **Disadvantages:**
  - No specific disadvantages mentioned.
  - Further research needed to understand limitations and challenges.

Combined Transfer and Active Learning for High Accuracy Music Genre Classification Method (2021)
Problem Statement

- Despite advancements in music genre classification methods, there is a need for approaches that can effectively handle niche genres and reduce the reliance on labeled data, especially in scenarios with limited annotated datasets.

A Hybrid Deep Learning Approach for Classification of Music Genres Using Wavelet and Spectrogram Analysis (2023)

- Utilizes Python and various libraries for music genre classification.
- Methodology includes CNN, transfer learning-based, multimodal training, and hybrid models.
- Performance evaluated using parameters like training accuracy, validation accuracy, training loss, precision, recall, F1-score, and support.
- Hybrid model outperforms other deep learning models in accuracy, achieving 81% and 71% accuracy using GTZAN and Ballroom datasets, respectively.

A Hybrid Deep Learning Approach for Classification of Music Genres Using Wavelet and Spectrogram Analysis (Contd.)

Advantages:

- Deep learning models like CNN, transfer learning-based models, and multimodal models outperform traditional machine learning models.
- Python and its libraries provide a compatible environment for implementing the proposed work.
- The proposed hybrid model achieves superior performance with 81% and 71% accuracy using GTZAN and Ballroom datasets.
- Extract, and analyze identifying features of different music files.

Disadvantages:

- Overfitting is a recurrent problem with CNN for music genre classification.
- The computational time for the proposed hybrid model needs careful analysis.
- Average accuracy of around 70% indicates the need improvement.

A Hybrid Deep Learning Approach for Classification of Music Genres Using Wavelet and Spectrogram Analysis (Contd.)
Problem Statement

- Despite advances in deep learning techniques for music genre classification, challenges such as overfitting, computational time, and the need for improvement in accuracy persist, motivating the exploration of hybrid approaches integrating wavelet and spectrogram analysis.

Music Genre Classification Based on Fusing Audio and Lyric Information (2023)

- Proposes a unified framework for music genre classification.
- Explores three fusion strategies for fusing different modal information.
- Proposes a hybrid fusion method for effective use of multimodal information.
- Constructs a multimodal music dataset.
- Performs experiments to verify the effectiveness of the proposed method.
- Organized into sections for summarizing relevant work, presenting the framework, demonstrating the method's effectiveness, and concluding the work.

Music Genre Classification Based on Fusing Audio and Lyric Information (Contd.)

Advantages:

- Integrates audio and lyric information for comprehensive genre representation.
- Improves classification effectiveness through multimodal methods like feature concatenation, decision weighting, and hybrid fusion.
- Leverages complementarity of audio and lyric information for richer genre information.
- Uses BERT for lyric text processing and CNN for audio feature extraction.

Disadvantages:

- Larger parameter number in BERT model can lead to learning rate problems.
- Difficulty in vectorizing text data and capturing semantic information in lyrics.
- Requires more fusion methods at feature and decision level.

Music Genre Classification Based on Fusing Audio and Lyric Information (Contd.)
Problem Statement

- Despite advancements in multimodal fusion techniques, challenges remain in effectively leveraging audio and lyric information for accurate music genre classification, particularly concerning the handling of large model parameters, text data vectorization, and decision fusion methods.

Comparison Table

Paper-1	Method	Advantages	Disadvantages
An Empirical Study on Structured Dichotomies in Music Genre Classification (2015)	Ensemble learning and dichotomy-based methods	Explores various dichotomy structures, conducts experiments, outlines future research directions	Initial results show limited improvement, no ensemble methods outperform base classifiers

Comparison Table

Paper-2	**Method**	**Advantages**	**Disadvantages**
Music Genre Classification and Recommendation by Using Machine Learning Techniques (2018)	Digital signal processing for feature extraction, machine learning techniques like CNNs and deep learning	Provides insights into feature extraction, uses various machine learning methods, evaluates performance, provides benchmark	Focuses only on acoustic features, limited dataset, lacks comprehensive comparison of methods

Comparison Table

Paper-3	Method	Advantages	Disadvantages
A Hybrid Model For Music Genre Classification Using LSTM And SVM (2018)	Combines LSTM Neural Network and SVM classifiers	Achieves high accuracy, leverages strengths of both classifiers	Lacks discussion of limitations, no comprehensive analysis of failure modes, doesn't address computational resources

Comparison Table

Paper-4	**Method**	**Advantages**	**Disadvantages**
Exploring Data Augmentation to Improve Music Genre Classification with ConvNets (2018)	Convolutional Neural Networks (CNNs) with data augmentation	Improves accuracy, enhances generalization capability, overcomes overfitting	Specificity of data augmentation, computational complexity, impact of fusion rules

Comparison Table

Paper-5	Method	Advantages	Disadvantages
Neural Network Music Genre Classification (2019)	Machine learning techniques, especially neural networks	Automatic feature extraction, complex pattern recognition, scalability, adaptability	Complexity, data requirements, overfitting

Comparison Table

Paper-6	**Method**	**Advantages**	**Disadvantages**
Time-Frequency Analysis Methods for Music Genre Classification by using Wavelet Package Decompositions (2019)	Wavelet packet decomposition (WPD)	Improved recognition rates compared to MFCC, multi-resolution characteristics, experiment credibility	Lacks discussion of limitations, no detailed comparison of complexity, no real-world implementation challenges, no mention of drawbacks

Comparison Table

Paper-7	Method	Advantages	Disadvantages
Utilizes Deep Learning CNN for Multilingual Music Genre Classification (2020)	Deep learning convolutional neural network	High accuracy for multiple languages, addresses bias, diverse dataset	Low prediction confidence, challenges in algorithm selection and data conversion, research gap, computational resource requirements

Comparison Table

Paper-8	**Method**	**Advantages**	**Disadvantages**
Combined Transfer and Active Learning for High Accuracy Music Genre Classification Method (2021)	Active transfer learning music genre classification method (ATMGCM)	High accuracy, reduces labeling needs, utilizes transfer and active learning	No specific disadvantages mentioned, requires further research on limitations

Comparison Table

Paper-9	**Method**	**Advantages**	**Disadvantages**
A hybrid deep learning approach for classification of music genres using wavelet and spectrogram analysis (2023)	Hybrid deep learning models (CNN, transfer learning, multimodal)	Outperforms traditional models, Python compatibility, superior accuracy	Overfitting issue with CNNs, computational time analysis needed, average accuracy indicates room for improvement

Comparison Table

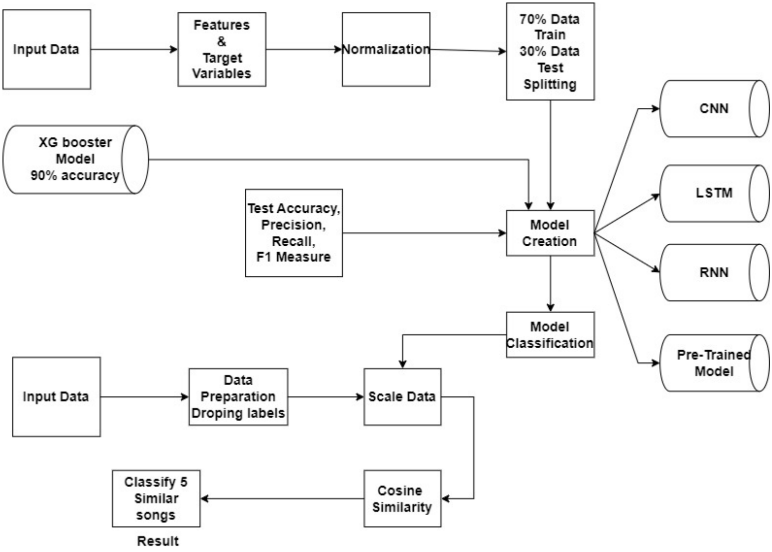
Paper-10	**Method**	**Advantages**	**Disadvantages**
Music genre classification based on fusing audio and lyric information (2023)	Fuses audio and lyric information	Comprehensive genre representation, improved effectiveness through multi-modal methods, leverages information complementarity, uses BERT and CNN	Large parameter number in BERT, difficulty in text vectorization and semantic capture, requires more fusion methods

PROPOSED SYSTEM

Exploring song genre classification through audio data analysis. Our system consists of several key components:

1. **U**tilizes advanced techniques for sound visualization and genre classification.
2. **C**aptures audio essence through meticulous data collection and feature extraction.
3. **U**ses machine learning models like CNN, LSTM, RNN, XGBoost, and Pre-trained Models for precision classification.
4. **U**ses evaluation metrics like accuracy, precision, recall, and F1-score to assess model performance.
5. **C**onducts cross-validation experiments to ensure model robustness and minimize overfitting.
6. **E**nhances understanding and appreciation of sound through systematic analysis.

ARCHITECTURE



ALGORITHM

- **Data Preparation**
  - Split the input data into two sets: training data (70%) and testing data (30%).
  - Preprocess the data by normalization and dropping labels.
- **Model Creation**
  - An XGBoost model is created using the training data.
  - A pre-trained convolutional neural network (CNN) model is used.
  - A Long short-term memory (LSTM) model is created using the training data.
  - A recurrent neural network (RNN) model is created using the training data.
- **Model Training**
  - The XGBoost model is trained on the training data.

- **Evaluation**
  - The accuracy of the XGBoost model is evaluated on the testing data.
  - The performance of the CNN, LSTM, and RNN models are evaluated on the testing data using metrics like precision, recall, and F1 measure.
- **Classification**
  - Classify 5 similar songs using cosine similarity.
- **Result**
  - The output is a list of the 5 most similar songs.

RESULT AND DISCUSSION

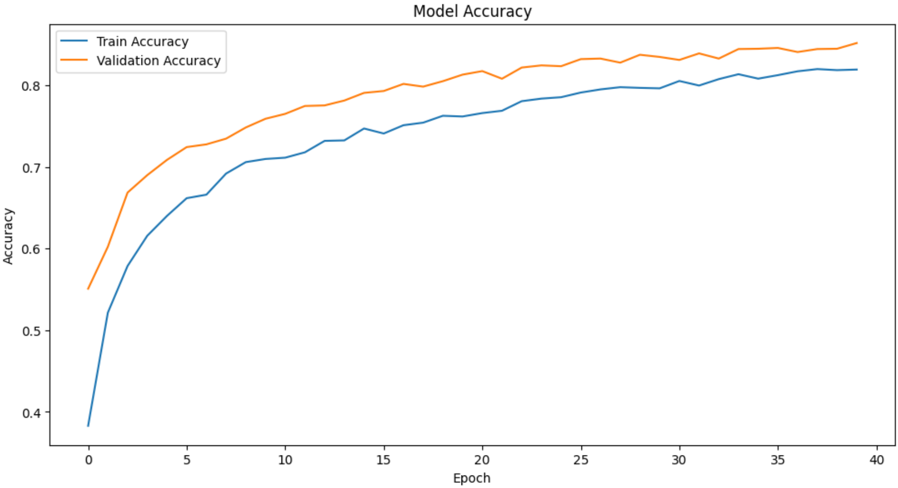
- **Dataset Used:**
  - The GTZAN dataset, created by George Tzanetakis and Perry Cook at the University of Victoria, is a primary source of information for training machine learning models.
  - The dataset comprises 1000 audio tracks, each 30 seconds long, sampled at a rate of 22050 Hz and stored in the .wav format.
  - Key features include genre diversity, audio content, annotated labels, standardized format, and its use in research.
  - The dataset covers a wide range of musical genres, providing a representative sample of music across different styles and categories.
  - The dataset includes manually annotated genre labels for supervised learning approaches.
  - citation : kaggle

Convolutional Neural Network (CNN)

- Trained with a batch size of 128 and 40 epochs.
- Achieved a test accuracy of approximately 86.89%.
- Classification Report

Genre	Precision	Recall	F1-score	Support
Blues	0.86	0.88	0.87	319
Classical	0.90	0.96	0.93	308
Country	0.82	0.79	0.80	286
Disco	0.84	0.79	0.81	301
Hip Hop	0.90	0.85	0.87	311
Jazz	0.84	0.87	0.85	286
Metal	0.88	0.91	0.90	303
Pop	0.85	0.91	0.88	267
Reggae	0.85	0.86	0.85	316
Rock	0.77	0.69	0.73	300
Accuracy			0.85	2997
Macro avg	0.85	0.85	0.85	2997
Weighted avg	0.85	0.85	0.85	2997

CNN (graph)

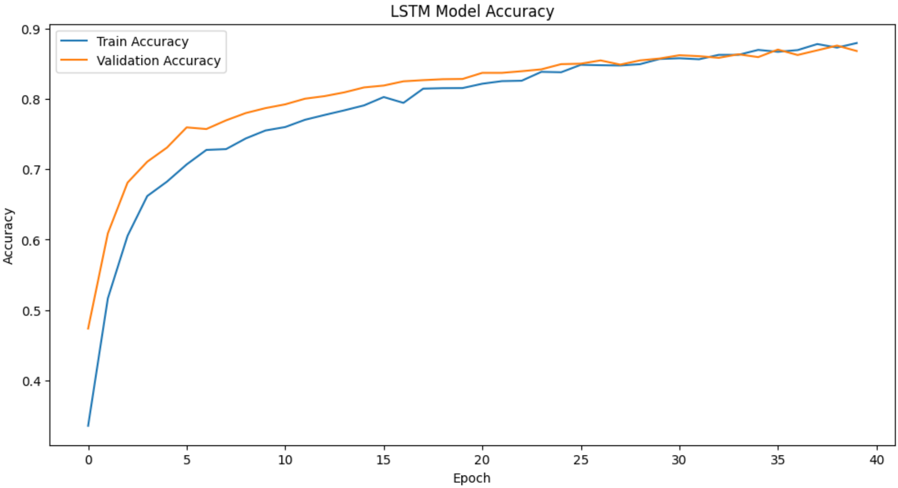

Model Accuracy
Train Accuracy
Validation Accuracy
Accuracy
0.8
0.7
0.6
0.5
0.4
0
5
10
15
20
25
30
35
40
Epoch

Long Short-Term Memory (LSTM)

- Trained with a batch size of 128 and 40 epochs.
- Achieved a test accuracy of approximately 86.89%.
- Classification report:

Genre	**Precision**	**Recall**	**F1-score**	**Support**
Blues	0.88	0.87	0.87	319
Classical	0.91	0.96	0.94	308
Country	0.82	0.85	0.83	286
Disco	0.80	0.82	0.81	301
Hip Hop	0.89	0.87	0.88	311
Jazz	0.89	0.88	0.89	286
Metal	0.91	0.92	0.92	303
Pop	0.86	0.91	0.88	267
Reggae	0.91	0.85	0.88	316
Rock	0.82	0.77	0.79	300
Accuracy	-	-	-	0.87

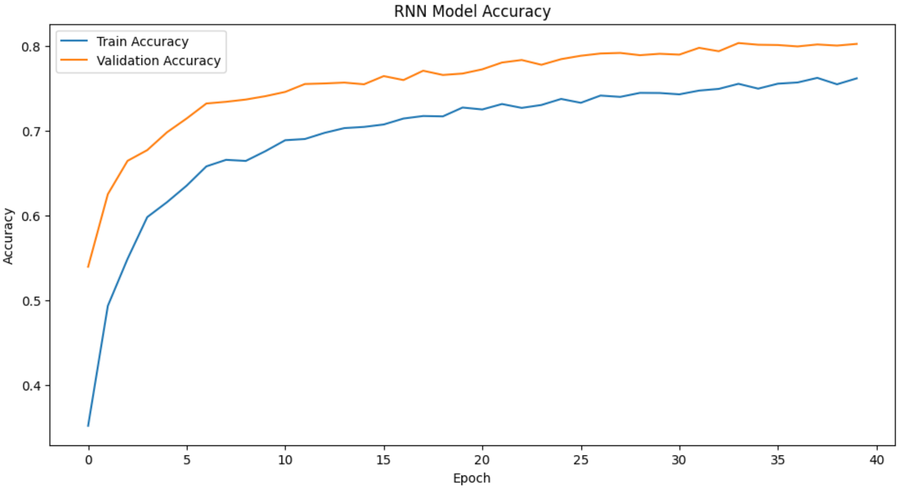
LSTM (graph)


LSTM Model Accuracy
Train Accuracy
Validation Accuracy
Accuracy
Epoch
0.9
0.8
0.7
0.6
0.5
0.4
0
5
10
15
20
25
30
35
40

- Trained with similar settings as LSTM.
- Achieved a test accuracy of approximately 80.41%.
- Classification report:

Genre	**Precision**	**Recall**	**F1-score**	**Support**
Blues	0.84	0.80	0.82	319
Classical	0.85	0.97	0.90	308
Country	0.74	0.77	0.75	286
Disco	0.71	0.77	0.74	301
Hip-hop	0.87	0.75	0.81	311
Jazz	0.86	0.83	0.84	286
Metal	0.86	0.90	0.88	303
Pop	0.80	0.90	0.85	267
Reggae	0.81	0.75	0.78	316
Rock	0.70	0.61	0.65	300
Accuracy	-	-	0.80	2997
Macro Avg	0.80	0.81	0.80	2997
Weighted Avg	0.80	0.80	0.80	2997

RNN (graph)


RNN Model Accuracy
Train Accuracy
Validation Accuracy
Accuracy
Epoch
0.4
0.5
0.6
0.7
0.8
0
5
10
15
20
25
30
35
40

Machine Learning Model

- Prediction on a random audio file: "Rock".
- Performance metrics for various algorithms:

Algorithm	Accuracy	Precision	Recall	F1-score
Naive Bayes	0.51952	0.534108	0.51952	0.501496
SGD	0.655322	0.659601	0.655322	0.629963
KNN	0.805806	0.813365	0.805806	0.806268
Decision Trees	0.635302	0.637027	0.635302	0.635315
Random Forest	0.814147	0.816764	0.814147	0.812406
SVM	0.754087	0.751228	0.754087	0.751084
Logistic Regression	0.697698	0.692071	0.697698	0.691998
Neural Nets	0.682015	0.677952	0.682015	0.678057
XGBoost	0.900901	0.901971	0.900901	0.900843
XGBoost (R F)	0.74708	0.757252	0.74708	0.745163

- Feature importance weights are available for the model

EQUATIONS

Accuracy: Accuracy is the ratio of correctly predicted instances to the total instances in the dataset. It provides an overall measure of how often the model makes correct predictions. Mathematically, accuracy is calculated as:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

Precision: Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. It focuses on the accuracy of the positive class predictions. Mathematically, precision is calculated as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Recall (Sensitivity): Recall, also known as sensitivity, is the ratio of correctly predicted positive observations to all observations in the actual class. It measures the ability of the model to identify all relevant instances. Mathematically, recall is calculated as:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

F1-score: The F1-score is the harmonic mean of precision and recall. It provides a balance between precision and recall, considering both false positives and false negatives. Mathematically, the F1-score is calculated as:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

These performance measures are essential for evaluating the classification models' effectiveness in correctly predicting the classes of instances. They provide insights into the model's strengths and weaknesses and help in making informed decisions about model selection and optimization strategies.

FUTURE SCOPE

- Enhance Model Performance: Explore advanced neural network architectures and training techniques.
- Real-time Music Generation: Investigate methods for enabling real-time music generation systems.
- Incorporate User Preferences: Develop mechanisms to incorporate user feedback and preferences.
- Collaboration and Co-Creation: Facilitate collaborative music composition by integrating features for multiple users.
- Integration with Music Production Tools: Integrate music generation models with existing music production software and tools.
- Expand Genre and Style Coverage: Expand the scope of music genres and styles supported by the model.
- Ethical and Social Implications: Investigate the ethical and social implications of AI-generated music.

GANTT CHART

<table>
<tr><th>Jan-24</th><th>Feb-24</th><th>Mar-24</th><th>Apr-24</th></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>Topic selection</td><td></td><td></td><td></td></tr>
<tr><td>Abstract</td><td></td><td></td><td></td></tr>
<tr><td colspan="3">Literature Review</td><td></td></tr>
<tr><td></td><td>Dataset Collection</td><td></td><td></td></tr>
<tr><td></td><td>Data preprocessing</td><td></td><td></td></tr>
<tr><td></td><td>Design</td><td></td><td></td></tr>
<tr><td></td><td colspan="3">Implementation</td></tr>
<tr><td></td><td>Unit test</td><td colspan="2">Unit test</td></tr>
<tr><td></td><td></td><td></td><td>Integration test</td></tr>
<tr><td>Documentation</td><td>Documentation</td><td colspan="2">Documentation</td></tr>
</table>

	Completed
	In Progress
	To be completed

CONCLUSION

- Explores diverse approaches and methodologies in music genre classification.
- Identifies the effectiveness of traditional machine learning algorithms like SVM and Random Forests, and modern deep learning techniques like CNNs and LSTMs.
- Highlights the importance of feature engineering, data augmentation, and hybrid model combinations for improving classification accuracy.
- Addresses challenges like overfitting, computational complexity, and dataset biases.
- Suggests the fusion of audio and lyric information and multimodal approaches for enhancing genre classification performance.
- Advocates for continuous advancement of machine learning techniques and innovative methodologies for music genre classification.

REFERENCES

- Tom Arjannikov; John Z. Zhang, 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA). Available: `https://doi.org/10.1109/ICMLA.2015.180` Accessed: January 2024.
- Ahmet Elbir; Hilmi Bilal Çam; Mehmet Emre Iyican; Berkay Öztürk; Nizamettin Aydin, 2018 Innovations in Intelligent Systems and Applications Conference (ASYU). Available: `https://doi.org/10.1109/ASYU.2018.8554016` Accessed: January 2024.
- Prasenjeet Fulzele; Rajat Singh; Naman Kaushik; Kavita Pandey, 2018 Eleventh International Conference on Contemporary Computing (IC3). Available: `https://doi.org/10.1109/IC3.2018.8530557` Accessed: January 2024.
- Rafael L. Aguiar; Yandre M.G. Costa; Carlos N. Silla, 2018 International Joint Conference on Neural Networks (IJCNN). Available: `https://doi.org/10.1109/IJCNN.2018.8489166` Accessed: January 2024.

- Nikki Pelchat; Craig M Gelowitz, 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE). Available: `https://doi.org/10.1109/CCECE.2019.8861555` Accessed: January 2024.
- Chih-Hsun Chou; Jun-Han Shi, 2018 1st IEEE International Conference on Knowledge Innovation and Invention (ICKII). Available: `https://doi.org/10.1109/ICKII.2018.8569119` Accessed: January 2024.
- Levi Ford; Sylvia Bhattacharya; Red Hayes; Wesley Inman, 2020 SoutheastCon. Available: `https://doi.org/10.1109/SoutheastCon44009.2020.9368270` Accessed: January 2024.
- Congyue Chen; Xin Steven, 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE). Available: `https://doi.org/10.1109/ICBAIE52039.2021.9390062` Accessed: January 2024.

- Jena, K.K., Bhoi, S.K., Mohapatra, S. et al. A hybrid deep learning approach for classification of music genres using wavelet and spectrogram analysis Available: https://doi.org/10.1007/s00521-023-08294-6 Accessed: January 2024.
- Li, Y., Zhang, Z., Ding, H. et al. Music genre classification based on fusing audio and lyric information. Available: https://doi.org/10.1007/s11042-022-14252-6 Accessed: January 2024.
- George Tzanetakis and Perry Cook. Kaggle. Available: https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification Accessed: January 2024.

Thank you!