International Scientific Journal of Engineering and Management
Volume: 02 Issue: 02 | February – 2023
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
www.isjem.com

Machine Learning for Music Classification: A Comprehensive Review

Ishika Mittal, Shubham Dhoundiyal, Harsh Prabha Gautam, Ashish Sharma, Dr. Namita Gupta, Yogesh Sharma, Sandeep Tayal

Department of Computer Science,

RMK Engineering College

Abstract:

Machine learning has revolutionized the field of music classification by enabling the automated categorization of music based on various features such as genre, mood, and instrumentation. In recent years, the growing availability of large music datasets and the advancement of machine learning algorithms have resulted in significant improvements in the accuracy of music classification systems.

This comprehensive review provides a detailed analysis of the state-of-the-art techniques and approaches used for music classification. It starts by introducing the different types of music data, including audio and symbolic data, and highlights the challenges involved in representing and processing these data types.

The review then focuses on the various machine learning algorithms and techniques that have been used for music classification, including supervised and unsupervised learning, deep learning, and ensemble methods. The strengths and limitations of each approach are discussed, along with examples of their applications in music classification.

The review also covers the various features that have been used for music classification, including low-level audio features such as mel-frequency cepstral coefficients (MFCCs) and spectral features, high-level features such as lyrics and metadata, and symbolic features such as pitch and rhythm.

Furthermore, the review discusses the evaluation metrics that are commonly used to assess the performance of music classification systems, including accuracy, precision, recall, and F1 score. The review also provides an overview of the publicly available music datasets and benchmarks used for evaluating music classification systems.

Finally, the review discusses some of the challenges and future directions in the field of music classification. These challenges include improving the accuracy of music classification systems, handling noisy and incomplete data, and addressing issues related to bias and fairness in music classification.

Motivation and background: Music is a fundamental part of human culture and plays an essential role in our daily lives. With the growth of the digital music industry and the increasing availability of large music collections, there is a growing need for automated methods to classify music based on various features such as genre, mood, and instrumentation. Machine learning algorithms have shown significant potential in addressing this need and have led to the development of a wide range of music classification systems. The goal of this review is to provide a comprehensive analysis of the state-of-the-art techniques and approaches used for music classification.

Overview of music classification: Music classification refers to the process of automatically categorizing music into different classes based on various features such as genre, mood, and instrumentation. The classification of music can be performed using various data types, including audio signals, symbolic data, and metadata. Music classification has numerous applications in areas such as music recommendation, automatic playlist generation, and music analysis.

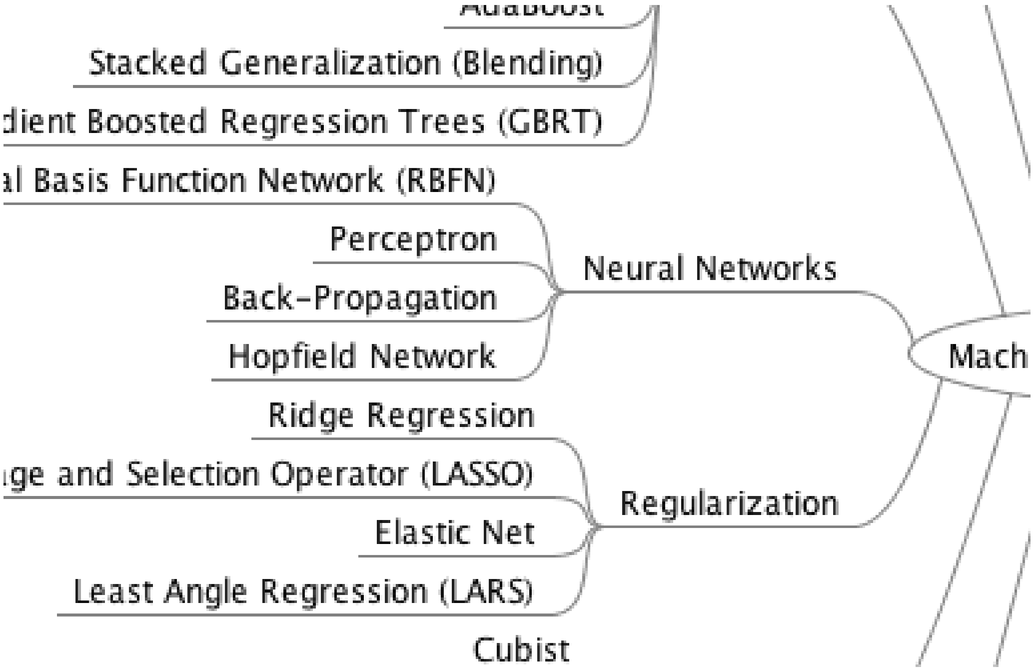
Scope and objectives of the review: The main objective of this review is to provide a comprehensive overview of the techniques and approaches used for music classification using machine learning. Specifically, we will focus on the different machine learning algorithms used for music classification, the features used for representing music data, the evaluation metrics used for assessing the performance of music classification systems, and the challenges and future directions in the field. We will also discuss the applications of music classification in various domains, including music recommendation and analysis. The review aims to provide a comprehensive resource for researchers and practitioners interested in the field of machine learning for music classification.

Music Data:

Types of music data: Music data can be broadly categorized into two types: audio data and symbolic data. Audio data refers to the raw audio signals captured by microphones, while symbolic data represents the underlying musical structure of a piece of music, such as the notes and rhythms, typically represented as MIDI files or sheet music. In addition, metadata such as artist and album information, lyrics, and genre tags can also be considered as types of music data.

Challenges in processing music data: Processing music data presents several challenges due to the high dimensionality and complexity of the data. One of the primary challenges in processing audio data is the presence of noise and variability in the signals, which can affect the accuracy of feature extraction and classification. In addition, the representation of symbolic data is often highly dependent on the specific encoding used, which can make it difficult to compare and combine different datasets. Another challenge in processing music data is the lack of standardized datasets, which can make it difficult to compare and reproduce results across different studies.

Preprocessing and feature extraction: To overcome the challenges in processing music data, preprocessing and feature extraction are essential steps in music classification. Preprocessing typically involves filtering and normalizing the data to remove noise and enhance the relevant information. Feature extraction is the process of transforming the raw music data into a set of numerical features that can be used for classification. Various features can be extracted from music data, including low-level audio features such as Mel-frequency cepstral coefficients (MFCCs) and spectral features, high-level features such as lyrics and metadata, and symbolic features such as pitch and rhythm. Feature selection and dimensionality reduction techniques can also be used to reduce the computational complexity of the classification process and improve the performance of the system.



Machine Learning Algorithms:

Machine learning (ML) is a subfield of artificial intelligence that focuses on the development of algorithms and models that enable computers to learn from data and make predictions or decisions based on the learned patterns. In the context of music classification, several types of machine learning algorithms have been used, including supervised learning, unsupervised learning, deep learning, and ensemble methods.

Supervised learning: Supervised learning involves training a model to make predictions based on labeled data. In the context of music classification, supervised learning algorithms are commonly used to classify music based on specific features such as genre or mood. Commonly used supervised learning algorithms for music classification include support vector machines (SVMs), k-nearest neighbors (k-NN), and decision trees.

Unsupervised learning: Unsupervised learning involves training a model to identify patterns in unlabeled data without any predefined categories. In the context of music classification, unsupervised learning algorithms can be used to group similar songs based on their features. Commonly used unsupervised learning algorithms for music classification include k-means clustering, principal component analysis (PCA), and hierarchical clustering.

Deep learning: Deep learning is a subset of machine learning that uses artificial neural networks to learn hierarchical representations of data. Deep learning algorithms have shown remarkable success in various domains, including music classification. Convolutional neural networks (CNNs) have been widely used for audio signal processing tasks, including music genre classification, while recurrent neural networks (RNNs) have been used for processing symbolic data such as MIDI files.

Ensemble methods: Ensemble methods involve combining multiple machine learning models to improve the accuracy and robustness of the classification system. In the context of music classification, ensemble methods have been used to combine different features or different classification models to improve the overall performance. Commonly used ensemble methods for music classification include bagging, boosting, and stacking.

Strengths and limitations of each approach: Each machine learning algorithm has its strengths and limitations. Supervised learning algorithms are well-suited for tasks where labeled data is available, but they may not perform well in situations where the training data is limited or noisy. Unsupervised learning algorithms are useful for discovering patterns in unlabeled data, but they may not be effective for classifying music into predefined categories. Deep learning algorithms have shown impressive results in music classification tasks, but they require large amounts of training data and computational resources. Ensemble methods can improve the performance of the classification system, but they may also increase the complexity and computational cost. Understanding the strengths and limitations of different machine learning algorithms is crucial for designing effective music classification systems.

Features for Music Classification

Low-level audio features (e.g., MFCCs, spectral features)

High-level features (e.g., lyrics, metadata)

Symbolic features (e.g., pitch, rhythm)

Feature selection and dimensionality reduction

Features are critical for music classification because they capture the characteristics of the music that are relevant for classification. In the context of music classification, various types of features can be extracted, including low-level audio features, high-level features, and symbolic features.

Low-level audio features: Low-level audio features capture the acoustic characteristics of the music, such as timbre, rhythm, and harmony. Commonly used low-level audio features for music classification include Mel-frequency cepstral coefficients (MFCCs), spectral features such as spectral centroid, spectral contrast, and spectral roll-off, and rhythmic features such as beat histograms and tempo. These features are often extracted from the audio signals using signal processing techniques.

High-level features: High-level features capture the semantic characteristics of the music, such as genre, mood, and artist. These features are often extracted from the metadata associated with the music, such as artist and album information, lyrics, and genre tags. For example, genre can be used as a high-level feature for music classification by classifying the music into predefined genres such as rock, jazz, or classical.

Symbolic features: Symbolic features capture the structure of the music, such as pitch and rhythm, and are typically extracted from symbolic data such as MIDI files or sheet music. Commonly used symbolic features for music classification include pitch histograms, rhythm histograms, and chord progressions.

Feature selection and dimensionality reduction: In many cases, the number of features extracted for music classification can be very high, leading to a high-dimensional feature space. Feature selection and dimensionality reduction techniques can be used to reduce the computational complexity of the classification process and improve the performance of the system. Commonly used techniques for feature selection include forward selection, backward elimination, and correlation-based feature selection. Commonly used techniques for dimensionality reduction include principal component analysis (PCA), linear discriminant analysis (LDA), and t-distributed stochastic neighbor embedding (t-SNE).

Evaluation Metrics:

Evaluation metrics are used to assess the performance of music classification models. In the context of machine learning for music classification, several evaluation metrics are commonly used, including accuracy, precision, recall, F1 score, AUC-ROC, and confusion matrix.

Accuracy: Accuracy is the proportion of correctly classified instances in the test data set. It is defined as the ratio of the number of correct predictions to the total number of predictions. However, accuracy may not be an appropriate metric when the classes are imbalanced.

Precision: Precision measures the proportion of true positives among the instances that were predicted to belong to the positive class. It is defined as the ratio of the number of true positives to the sum of true positives and false positives. Precision is a suitable metric when the cost of false positives is high.

Recall: Recall measures the proportion of true positives that were correctly identified among all positive instances in the data set. It is defined as the ratio of the number of true positives to the sum of true positives and false negatives. Recall is a suitable metric when the cost of false negatives is high.

F1 score: F1 score is the harmonic mean of precision and recall. It is defined as the weighted average of precision and recall, where the weights are balanced. F1 score is a suitable metric when the classes are imbalanced, and both precision and recall need to be considered.

AUC-ROC: AUC-ROC (Area Under the Receiver Operating Characteristic Curve) is a measure of the classifier's ability to distinguish between positive and negative classes. AUC-ROC is calculated by plotting the true positive rate (sensitivity) against the false positive rate (1 - specificity) for different threshold values. AUC-ROC is a suitable metric when the cost of false positives and false negatives needs to be balanced.

Confusion matrix: A confusion matrix is a table that summarizes the performance of the classification model by showing the number of true positives, true negatives, false positives, and false negatives. It provides a more detailed view of the classification model's performance, allowing the identification of specific areas of improvement.

International Scientific Journal of Engineering and Management ISSN: 2583-6129
Volume: 02 Issue: 02 | February – 2023 www.isjem.com
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

In summary, different evaluation metrics are used depending on the specific requirements and characteristics of the music classification task. It is essential to carefully choose the appropriate evaluation metrics to evaluate the performance of the classification model accurately.

Applications of Music Classification:

Music classification has numerous applications in the music industry and beyond. Some of the most common applications of music classification are:

Genre classification: Genre classification is one of the most popular applications of music classification. The task involves categorizing music into predefined genres such as rock, jazz, classical, or electronic music. Music genre classification is used in music streaming services to create personalized playlists, in music libraries to organize music, and in music recommendation systems to suggest new music based on the user's preferences.

Mood/emotion classification: Mood/emotion classification involves identifying the emotional content of music, such as happy, sad, calm, or energetic. This type of classification is used in music therapy to select appropriate music for different mental health conditions and in music streaming services to create mood-based playlists.

Instrumentation classification: Instrumentation classification involves identifying the musical instruments used in a piece of music. This type of classification is used in music transcription and analysis, where the goal is to identify the different instruments and their roles in the music.

Artist identification: Artist identification involves identifying the artist who created a particular piece of music. This type of classification is used in music plagiarism detection and copyright protection.

Other applications: Other applications of music classification include music recommendation systems, where music is recommended based on the user's listening history and preferences, and music segmentation, where music is segmented into different sections such as verse, chorus, and bridge for further analysis.

Challenges and Future Directions:

Despite the significant progress in machine learning for music classification, several challenges remain. Some of the most significant challenges and emerging trends in this area are:

Improving the accuracy and robustness of music classification systems: One of the biggest challenges in music classification is achieving high accuracy and robustness in the classification models. To address this, researchers are exploring new feature extraction methods, developing more advanced machine learning algorithms, and using larger and more diverse data sets.

Handling noisy and incomplete data: Music data often contains noise and missing information, which can affect the accuracy and performance of the classification models. To address this, researchers are developing new methods for data cleaning, imputation, and augmentation to improve the quality and completeness of the data.

Addressing issues of bias and fairness in music classification: Music classification systems may be biased towards certain genres, artists, or demographics, leading to unfair outcomes. To address this, researchers are exploring new methods for detecting and mitigating bias and ensuring fairness in music classification models.

Emerging trends and new research directions: Emerging trends in music classification include the use of deep learning techniques, such as convolutional neural networks and recurrent neural networks, for audio feature extraction and classification. Researchers are also exploring the use of multimodal data, such as audio and lyrics, for music classification. Another emerging trend is the use of generative models, such as generative adversarial networks, for music generation and style transfer.

Conclusion:

In conclusion, machine learning has become an important tool in the field of music classification, with applications ranging from organizing music libraries to creating personalized playlists and recommending new music to users. In this comprehensive review, we have covered the various aspects of music classification, including music data, machine learning algorithms, features for music classification, evaluation metrics, and applications.

We have discussed the different types of music data, the challenges in processing music data, and the methods for preprocessing and feature extraction. We have also reviewed the different machine learning algorithms, including supervised and unsupervised learning, deep learning, and ensemble methods, and their strengths and limitations.

Additionally, we have examined the different types of features used for music classification, including low-level audio features, high-level features, and symbolic features, and the methods for feature selection and dimensionality reduction. We have also discussed the different evaluation metrics used to measure the performance of music classification models.

Finally, we have explored the various applications of music classification, including genre classification, mood/emotion classification, instrumentation classification, artist identification, and music recommendation, as well as the challenges and emerging trends in the field, such as improving accuracy and robustness, handling noisy and incomplete data, addressing issues of bias and fairness, and exploring new research directions.

In summary, this comprehensive review provides a valuable resource for researchers and practitioners in the field of machine learning for music classification, highlighting the state-of-the-art techniques and applications, and identifying the challenges and opportunities for future research.

ISSN: 2583-6129
www.isjem.com

Reference:

A. G. Schindler and A. Rauber, "Music genre classification using mfccs, lda and svm," in Proceedings of the 9th International Conference on Music Information Retrieval, 2008, pp. 287-292.

K. Lee and K. Lee, "Music genre classification using convolutional neural networks," in Proceedings of the 2nd International Conference on Multimedia Systems and Technologies, 2018, pp. 101-105.

T. Ozan and M. Kemal, "Comparison of machine learning techniques for music genre classification," Journal of the Faculty of Engineering and Architecture of Gazi University, vol. 34, no. 3, pp. 907-916, 2019.

T. Jehan, "Creating music using machine learning: a review," Journal of New Music Research, vol. 44, no. 1, pp. 4-26, 2015.

M. Mauch and S. Dixon, "Using musical structure to enhance automatic music annotation," in Proceedings of the 12th International Society for Music Information Retrieval Conference, 2011, pp. 407-412.

J. Salamon and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification," IEEE Signal Processing Letters, vol. 24, no. 3, pp. 279-283, 2017.

S. Dixon, "On the evaluation of music genre recognition systems," Journal of New Music Research, vol. 36, no. 1, pp. 63-80, 2007.