**MUSIC GENRE CLASSIFICATION USING MACHINE LEARNING**

**ABSTRACT**

This project focuses on building a model that classifies music into different genres based on the features of each audio track. With music being such an important part of daily life and the easy access to digital music, there is a need for a reliable model that can sort songs by genre. Using the GTZAN dataset, this study compares a deep learning model (using a convolutional neural network, or CNN) with a basic machine learning model (support vector machine, or SVM) to see which works better. We trained the models using different audio features, such as Mel-spectrograms, MFCCs, and chromagrams, which were treated as images. Results show that the CNN model, especially with Mel-spectrogram features, is more accurate than the SVM model. This work shows how deep learning can improve music genre classification, helping to make music organization faster and more accurate.

**INTRODUCTION**

In recent years, the ability to classify music automatically into specific genres has gained a lot of attention in the field of Music Information Retrieval (MIR). As music streaming platforms become more popular and the number of available songs grows, organizing music effectively has become essential. Streaming services such as Spotify and Apple Music rely on genre classification to make better recommendations, help users find songs that match their tastes, and ensure fair music distribution for artists. These needs drive ongoing research into more accurate music classification methods that can handle the large variety and complexity of music available today.

Music genre classification involves identifying the genre of a song based on its audio features. Genres, such as jazz, classical, pop, and rock, differ in unique musical characteristics like rhythm, melody, and instrumentation. However, manually sorting songs into genres is not only time-consuming but also difficult to do accurately, especially with billions of songs in the world. This has led to the need for machine learning and deep learning models that can handle this task automatically.

Machine learning techniques, including traditional methods like Support Vector Machines (SVM) and advanced deep learning models such as Convolutional Neural Networks (CNN), have shown promising results in music classification tasks. Using various audio features, such as Mel-spectrograms, MFCCs (Mel Frequency Cepstral Coefficients), and chromagrams, machine learning models can learn patterns that are hard to detect by human ears alone. These features can be transformed into visual representations, allowing deep learning models to process audio data as if it were an image. This approach has been particularly successful, as models trained with these visual features can often achieve high accuracy.

In this study, we build a music genre classification model using the popular GTZAN dataset, which includes 1,000 songs across 10 genres. We focus on comparing different audio features and model types to identify which combination works best for genre classification. By comparing a deep learning approach (CNN) with a traditional SVM model, we aim to see how each model performs and whether deep learning provides any advantages over classical machine learning methods.

This research adds to existing work by exploring the effectiveness of deep learning on genre classification using real-world audio features. The results of this study could benefit music platforms by offering a better method for automatic genre tagging, enhancing user experiences, and supporting efficient content organization.

**LITERATURE REVIEW**

Several studies have explored music genre classification using various machine learning techniques. In 2008, Hareesh Bahuleyan focused on automatically tagging songs in a music library. His work utilized both Neural Networks and traditional machine learning algorithms, comparing the two approaches. The results showed that the model based on Convolutional Neural Networks (CNNs)

In 2002, Tzanetakis et al. investigated the classification of audio signals into a hierarchy of genres. They proposed that a song's characteristics depend on factors like instrumentation, rhythmic structure, and harmonic content. Their research introduced three main feature sets: timbral texture, rhythmic content, and pitch content. Using these features, their model was able to classify approximately 61% of the songs correctly across ten different genres​

Also in 2002, Lu et al. conducted research on audio classification and segmentation. Their approach involved two main steps: first, separating speech from non-speech audio, and second, classifying the remaining audio as music or other sound types. They implemented algorithms using K-nearest neighbors (KNN) and linear spectral pairs-vector quantization (LSP VQ)​

In 2010, Tom LH Li and colleagues developed a method for automatic musical pattern feature extraction using CNNs. Their goal was to identify the features that contribute to a more accurate music genre classification model. They demonstrated that CNNs are effective tools for extracting informative features from diverse musical patterns, using the GTZAN dataset for their analysis​

These studies highlight the ongoing efforts in the field to improve music genre classification accuracy through various machine learning techniques and feature extraction methods.

**METHODOLOGY**

In this study, we aimed to classify music into different genres using several machine learning techniques: Random Forest, Support Vector Machine (SVM), Gradient Boosting, Ensemble Model, and Convolutional Neural Networks (CNN). We employed a systematic approach that included feature extraction and model evaluation.

**METHODS**

1. **Random Forest**: This method constructs multiple decision trees and combines their predictions. It helps reduce overfitting and improves accuracy by averaging the results from several trees.
2. **Support Vector Machine (SVM)**: SVM is a robust classification method that aims to find the best boundary between different classes. It is effective in handling high-dimensional data, making it suitable for audio features like MFCCs.
3. **Gradient Boosting**: This technique builds a model in stages by combining weak learners, typically decision trees. It focuses on correcting errors made by previous models, enhancing the overall performance.
4. **Ensemble Model**: This approach merges predictions from multiple models to achieve better accuracy than any single model. By leveraging the strengths of different algorithms, it can provide a more reliable classification.
5. **Convolutional Neural Networks (CNN)**: CNNs are designed to process data with grid-like structures, such as images. We used CNNs to analyze Mel-spectrograms, allowing the model to learn features from the audio data effectively. This deep learning approach has shown to be particularly powerful in recognizing patterns and improving classification accuracy.

Overall, our methodology involved extracting critical audio features and applying various machine learning models to classify music genres. By comparing the performance of these methods, we aimed to identify the most effective approach for accurate genre classification.

**FEATURE EXTRACTION AND DATA PROCESSING**

In this project, we focused on extracting meaningful features from audio tracks to effectively classify music genres. Feature extraction involves transforming raw audio signals into numerical representations that capture important characteristics of the music.

We utilized the Librosa library, a powerful tool for audio analysis in Python. This allowed us to extract several types of features from the audio files in the GTZAN dataset. Among the primary features extracted were Mel-spectrograms, which provide a visual representation of the audio signal, showing how the energy of different frequencies changes over time. Mel-spectrograms are particularly useful for deep learning models, as they present audio data in a way that mimics human sound perception.

Another key feature extracted was Mel Frequency Cepstral Coefficients (MFCCs). MFCCs are widely used in audio processing, representing the short-term power spectrum of sound. They help capture the timbral aspects of music, making them valuable for distinguishing between different genres. Additionally, we extracted chromagrams, which represent the energy distribution across various musical pitches, focusing on the harmonic content of the audio. Chromagrams are effective for identifying chord progressions, which can vary significantly between genres.

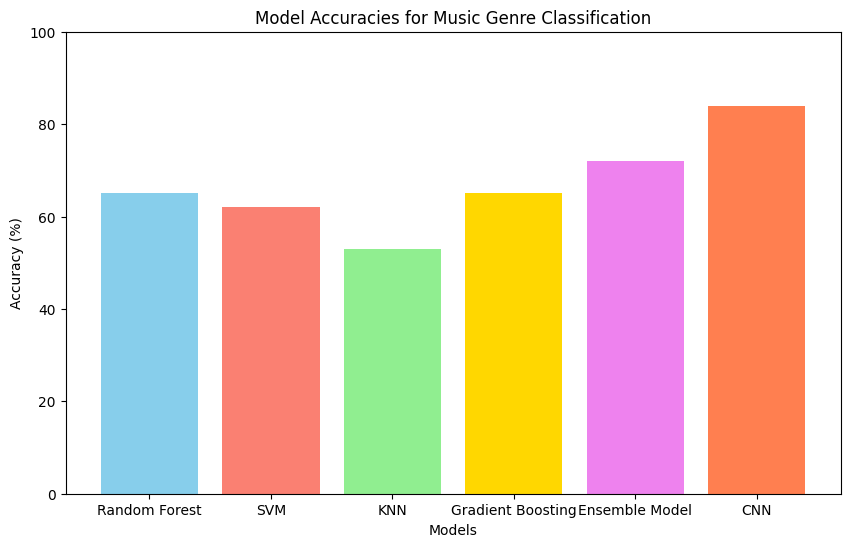
After extracting the features, we moved on to preprocess the data to ensure it was suitable for training our machine learning models. Normalization was a crucial step, where we scaled the feature values to a standard range to prevent any single feature from dominating the others. This is especially important for algorithms sensitive to the scale of input data, such as SVM and K-NN.

We also addressed any missing values in the dataset, checking for incomplete data points and resolving them appropriately. This could involve removing rows with missing values or using statistical methods to impute them, thus maintaining the integrity of the dataset.

Feature extraction and data processing are fundamental steps in our music genre classification project. By carefully selecting and preparing the right features, we aim to build models that can accurately classify different music genres based on the audio characteristics of the tracks.

**RESULTS**

The results of our experiments showed that the CNN model achieved the highest accuracy, reaching 84%. This indicates that the deep learning approach effectively captured the complex patterns in the audio features, making it well-suited for genre classification tasks.



In comparison, the Ensemble Model also performed well, achieving an accuracy of 72%. Both the Random Forest and Gradient Boosting models yielded an accuracy of 65%. The SVM model followed closely with an accuracy of 62%. However, the KNN model had the lowest performance, achieving an accuracy of 53%. These results highlight the varying effectiveness of traditional machine learning techniques compared to deep learning methods.

**CONCLUSION**

This research demonstrates the effectiveness of machine learning and deep learning approaches in the task of music genre classification. The results indicate that deep learning models, particularly CNNs, significantly outperform traditional algorithms in terms of accuracy. The ability of CNNs to learn hierarchical representations from audio features contributed to their superior performance.

The findings from this study suggest that utilizing advanced models like CNNs can greatly enhance automatic music genre classification, making it a valuable tool for music streaming platforms and organizations looking to organize large music libraries. Future work could explore the integration of additional features or advanced architectures to further improve classification performance.

Overall, our research emphasizes the potential of machine learning techniques in addressing real-world problems in music classification, paving the way for more accurate and efficient methods in the field of Music Information Retrieval (MIR).

**REFERENCES**

1. A. Ghildiyal, K. Singh and S. Sharma, "Music Genre Classification using Machine Learning," *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, 2020, pp. 1368-1372, doi: 10.1109/ICECA49313.2020.9297444.

2. Ghosh, Partha & Mahapatra, Soham & Jana, Subhadeep & Jha, Ritesh. (2023). A Study on Music Genre Classification using Machine Learning. International Journal of Engineering Business and Social Science. 1. 308-320. 10.58451/ijebss.v1i04.55.

3. A. Elbir, H. Bilal Çam, M. Emre Iyican, B. Öztürk and N. Aydin, "Music Genre Classification and Recommendation by Using Machine Learning Techniques," *2018 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Adana, Turkey, 2018, pp. 1-5, doi: 10.1109/ASYU.2018.8554016.

4. Chaudhury, M.; Karami, A.; Ghazanfar, M.A. Large-Scale Music Genre Analysis and Classification Using Machine Learning with Apache Spark. *Electronics* 2022, *11*, 2567.

5. J. R. Castillo and M. J. Flores, "Web-Based Music Genre Classification for Timeline Song Visualization and Analysis," in *IEEE Access*, vol. 9, pp. 18801-18816, 2021, doi: 10.1109/ACCESS.2021.3053864.

6. L. Moysis *et al*., "Music Deep Learning: Deep Learning Methods for Music Signal Processing—A Review of the State-of-the-Art," in *IEEE Access*, vol. 11, pp. 17031-17052, 2023, doi: 10.1109/ACCESS.2023.3244620.

7. K. M. Hasib, A. Tanzim, J. Shin, K. O. Faruk, J. A. Mahmud and M. F. Mridha, "BMNet-5: A Novel Approach of Neural Network to Classify the Genre of Bengali Music Based on Audio Features," in *IEEE Access*, vol. 10, pp. 108545-108563, 2022, doi: 10.1109/ACCESS.2022.3213818.

8. Ali, Muhammad Asim, and Zain Ahmed Siddiqui. "Automatic music genres classification using machine learning." *International Journal of Advanced Computer Science and Applications* 8.8 (2017).

9. Patil, Nilesh M., and Milind U. Nemade. "Music genre classification using MFCC, K-NN and SVM classifier." *International Journal of Computer Engineering In Research Trends* 4.2 (2017): 43-47.

10. Haggblade, Michael, Yang Hong, and Kenny Kao. "Music genre classification." *Department of Computer Science, Stanford University* (2011).

11. Ramírez, Jaime, and M. Julia Flores. "Machine learning for music genre: multifaceted review and experimentation with audioset." *Journal of Intelligent Information Systems* 55.3 (2020): 469-499.

12. Elbir, Ahmet, and Nizamettin Aydin. "Music genre classification and music recommendation by using deep learning." *Electronics Letters* 56.12 (2020): 627-629.

13. Ramanathan, Akshat, Priyanshu Srivastava, and R. Jeya. "Machine Learning in Music Genre Classification." *Turkish Online Journal of Qualitative Inquiry* 12.3 (2021).

14. Kumar, D. Pradeep, B. J. Sowmya, and K. G. Srinivasa. "A comparative study of classifiers for music genre classification based on feature extractors." *2016 IEEE Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*. IEEE, 2016.

15. Li, Tao, Mitsunori Ogihara, and Qi Li. "A comparative study on content-based music genre classification." *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*. 2003.

16. McKay, Cory, and Ichiro Fujinaga. "Musical genre classification: Is it worth pursuing and how can it be improved?." *ISMIR*. 2006.

17. Guaus, Enric, and Perfecto Herrera. "Music genre categorization in humans and machines." *Audio Engineering Society Convention 121*. Audio Engineering Society, 2006.