Journal of Industrial Information Technology and Application *ISSN(Online): 2586-0852*
Vol.8 No.1

Music Genre Classification of Philippine Music

Alden Robell M. de Loyola[1,*] and Robert R. Roxas[2]
[1,2] University of the Philippines Cebu, Cebu, Philippines

Abstract. This paper presents a study on Philippine music genre classification. The dataset was manually created by sampling audio features from 1,400 Philippine music tracks on seven genres. After classifying the data, the models were evaluated using accuracy (model analysis) and recall (genre analysis). Findings show that k-nearest neighbors, support vector machine, and random forest were the best-performing models, while decision tree was the worst-performing model. Rondalla was the most predictable genre, followed by Kulintang, Kundiman, and Rap. Pop Ballad, Rock, and Manila Sound were the difficult genres to predict. This study implies that popular machine learning models work well with the classification of Philippine music.

Keywords; music genre classification, machine learning, Philippine music, audio features

1. Introduction

Music genre classification automatically assigns a genre label to a music by extracting its audio features and feeding them into machine learning models. By extracting relevant features from the audio signals and using these features as inputs to a machine learning model, it is possible to achieve high levels of classification accuracy for a variety of music genres. These models have shown to vary across studies due to factors such as the dataset used, data preprocessing method [1], number of classes [2], model tuning (e.g., value of *k* in *k*-nearest neighbors) [3], model architecture [4], time duration of music input [4], and audio feature selection [5]. Factors in music genre

* Corresponding author: amdeloyola@alum.up.edu.ph
Received: Jan. 10, 2024; Accepted: Feb. 12, 2024; Published: Mar. 31, 2024

classification make it challenging to identify the best model for music genre classification.

Every audio can be represented in the form of an audio signal, which contains different features [5]. Literatures on music genre classification suggest that a combination of audio features from audio signals [6] can effectively classify music by genre. Traditional machine learning models outperform deep learning [5]. The widely used models for music genre classification include support vector machine, decision tree, logistic regression, k-nearest neighbors, and random forest. They are simpler than deep models since the latter is highly sensitive in architecture, which significantly affects the models' performance. Traditional models demonstrated good performance in classification tasks and had been widely used in various applications [7], including music recommendation systems and music platforms. The widely used models for music genre classification include support vector machine, decision tree, logistic regression, *k*-nearest neighbors, and random forest.

Philippine music tracks, in the past decade, are labeled as a merely single genre called 'OPM' across music platforms. This is a common observation among Filipinos when using applications like Spotify and YouTube. A good representation of Philippine music can be applied in relevant applications, where entire catalogs of Philippine music should be accurately labeled by their specific genres. Using machine learning, they can be represented well by their specific and complex genres. This study aims to analyze and compare the performance of machine learning models in classifying Philippine music according to its genre.

Specifically, this study aims to create a dataset of Philippine music by extracting 57 audio features from each of 1,400 samples of music tracks on genres such as *Rondalla*, *Kundiman*, *Kulintang*, *Pop Ballad*, *Manila Sound*, *Rap*, and *Rock*. Then this study will measure and analyze the performance of support vector machine, decision tree, linear regression, *k*-nearest neighbors, and random forest to classify Philippine music by genre with *accuracy* as the metric. Then it will compare and analyze the machine learning models with previous related studies and compare and analyze the music genres with *recall* as metric.

2. Methodology

This section discusses how the dataset was created. Then it is followed by a description of the activities performed in music genre classification and a description of how the model and genre analysis was done.

A. Dataset creation

To manually create a dataset of Philippine music, 1,400 pieces of 3-second tracks from seven genres were chosen: Rondalla, Kundiman, Kulintang, Manila Sound, Pop Ballad, Rock, and Rap. For time domain of the audio signal, root mean square energy, zero crossing rate, and tempo were extracted. For frequency domain, 20 samples of mel frequency cepstral coefficients, chromagram, spectral centroid, spectral bandwidth, spectral roll off, harmonic, and perceptual were extracted. Selection of these audio features was adopted from another popular music dataset GTZAN which was used in [3-5,7-9].

A total of 57 audio features were extracted using Librosa (python library for audio analysis) for each track, creating a dataset with size (1,400, 57).

B. Music Genre Classification

The dataset was preprocessed with normalization using *MinMaxScaler* and *fit_transfrom* from the *sklearn* library. Hyperparameter tuning was performed on the models using grid search with cross validation to find the best hyperparameters. The performance of the models was based how well each model performed on unseen data.

Table I. HYPERPARAMETER TUNING OF EACH MODEL

Models	**Grid Search with Cross Validation (folds = 10)**	
	Hyperparameter	***Values***
K-nearest neighbor	n_neighbors	1, 2, 3, 4, 5, 10
Decision tree	max_depth	1, 5, 10, 15, 20, 30, 40, 50, 60, 100
Random forest	max_depth	15, 20, 30, 40, 50
	n_estimators	100, 300, 500, 700, 900, 1,000
Support vector machine	C	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
	Kernel	poly, rbf, linear, sigmoid
Logistic regression	max_iter	1,000, 1,500, 2,000

K-nearest neighbors was used in [1-3,7-9], decision tree in [5-6,7,9], random forest in [3,7-8], support vector machine in [2-3,5-9], and logistic regression in [3- 4,7-8].

C. Model and Genre Analysis

Performance of the models was based on *accuracy* (1) to estimate how well each model performed on unseen data.

$$\text{Accuracy} = (\text{Correctly classified instances}) / (\text{Total instances}) \tag{1}$$

The model must predict songs according to their actual genre and not classify them as other genres since the cost of false negative was high. The *recall* metric (2), which provided insights into how well the model predicted the instances of a genre regardless of its performance in other genres, was most suitable in evaluating the genres.

$$\text{Recall} = (\text{True positives}) / (\text{True positives} + \text{False negatives}) \tag{2}$$

3. Results and Discussion

This section discusses the analysis on the machine learning models, comparison with related studies, and insights on the music genres.

D. Analysis on the Machine Learning Models

Table II shows the performance of the different models. K-nearest neighbors achieved the highest accuracy of 91.6%. Separable clusters due to the genres' distinctness, as shown in Fig. 1, made it easy for KNN with n_neighbors = 1 to assign the genres based on its nearest neighbors. It also had the shortest runtime of 1.88 seconds. In [4] and [5], KNN also ranked highest among SVM, random forest, and logistic regression.

Table II. PERFORMANCE OF THE MODELS

Models	Hyperparameters	Accuracy (%)	Runtime (s)
K-nearest neighbor	N_neighbors = 1	91.6	1.88
Support vector machine	C = 10 Kernel = rbf	90.62	3.84
Random forest	N_estimators = 900 Max_depth = 30	89.98	733
Logistic regression	Max_iter = 1,000 Multi_class = multinomial	82.67	28.05
Decision tree	Max_depth = 60	71.06	5.03

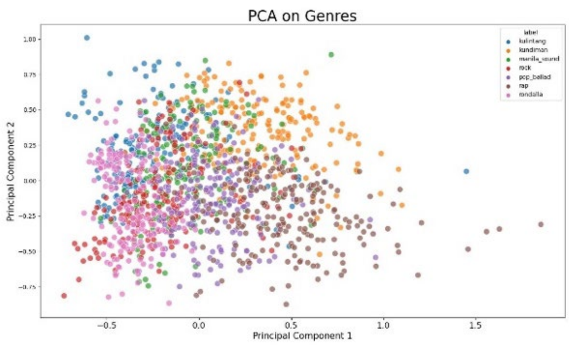

Figure 1. Principal component analysis

Decision trees created complex decision boundaries, which was a disadvantage for well-clustered data. This results in overfitting, yielding lower accuracy. This finding did not align with some previous literatures. In [1][2][6], decision tree scored greater than SVM, logistic regression, and KNN. Therefore, a lot of factors cannot generalize what model predicts music genres well. Random forest scored with high accuracy (89.98%), but it had the highest computation with 733-second runtime due to increased *max_depth.* What made SVM most notable was that it performed better than KNN in classifying the hardest genre (*Manila Sound*). SVM handled complex data better due to its kernel tricks and margin maximization. Logistic regression scored high accuracy, but it did not handle complex genres well. As shown in Fig. 2, the recall scores of logistic regression on the two hardest genres *Manila Sound* and *Rock* abruptly dropped.

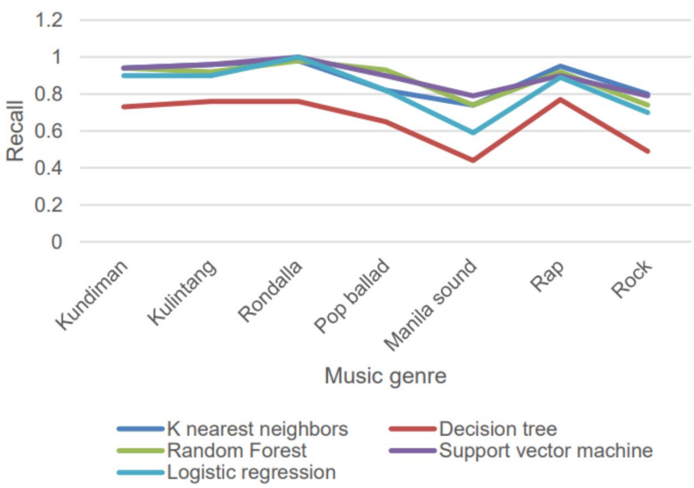

Figure 2. Recall scores of each music genre

E. Comparison with Related Studies

Table III shows the comparison of this study with other related studies. Findings in this study aligned well with [4] and [5], where KNN, SVM, and random forest were the highest scorers. The performance of decision trees from related literatures differed from this study. Several factors contributed to this, such as the dataset used in [6], or in [2] where three datasets were used for result variations, and lastly, experiment on data characterization in [1]. SVM scored 90% in handling this multiclass dataset. This finding, however, varied from findings in [2], where SVM poorly performed in handling multiclass dataset. Reference [2] also noted that random forest was a good classifier for time-limited resources, which varied in this study since it had a rough runtime of 700 seconds in this study.

Table III. COMPARISON WITH RELATED STUDIES

Dataset	*KNN*	*SVM*	*Random forest*	*Logistic regression*	*Decision tree*
This study	91.6%	90.62%	89.98%	82.67%	71.06%
GTZAN [1]	67.5%	82.55%	-	67.5%	77.5%
MSD [2]	-	52%	62%	-	61%
GTZAN [4,5]	92.69%	74.72%	80.28%	67.52%	-
GTZAN [6]	-	68.9%	-		74.3%
Spotify [8]	68.40%	72%	-	-	-
GTZAN [9]	62.5%	72.39%	65.69%	-	55%

F. Insights on the Music Genres

Rondalla was easiest for the models to predict due to its limited variation, where songs follow consistent musical patterns. The same was true for *Kulintang* (0.90 recall) as it sounded distinctly due to its instruments that no other genres utilized (gong ensembles). *Kundiman*, with 0.89 recall, is more complex with its varied vocals but was classified by the models well. *Rap* paralleled with *Kundiman* due to its consistency in beats, percussive elements, and vocal cadences. *Pop Ballad*, *Rock*, and *Manila Sound* were the least predictable genres. *Pop Ballad* features varied range of melodic styles and tempo structures. Rock, the 2nd least predictable genre, was inherently diverse as it significantly varied in instrumentation, vocal style, and tempo. *Rock* being a “rebel” genre deviated from these standardized features. *Rock* was hard to classify [2].

Although the *Rock* genre scored poor relative to other genres, most of the misclassified tracks were concentrated only into *Manila Sound*. *Manila Sound* was the genre, where most *Rock* mispredictions were found. This finding aligned exactly with [4], where most of the misclassification of *Rock* was concentrated on the disco genre. *Manila Sound,* being the disco genre of the Philippines, and *Rock* shared the similarities with disco. *Manila Sound*, the least predictable genre, is a fusion of musical styles from Filipino folk, Western music, and Latin rhythms – making it diverse and hard to predict for the models. Selection of genres could affect the overall accuracy of the model [3][6].

The cultural and musical complexity of Philippine music, ranging from the sweet serenades of *Kundiman* to the rebellious *Rock* genre, could be represented well in applications like music streaming platforms, where entire catalogs of Philippine music are accurately represented by the specificity of their genres. Traditional machine learning models remained to be relevant and equally superior with other algorithms like deep learning.

4. Conclusion

This study classified Philippine music by genre using machine learning models. The results show that *k*-nearest neighbors, support vector machine, and random forest were the best-performing models, while decision tree was the worst-performing model. *Rondalla* was the most predictable genre, followed by *Kulintang*, *Kundiman*, and *Rap*. *Pop Ballad*, *Rock,* and *Manila Sound* were difficult genres to predict. Many of the results in this study aligned well with some existing studies.

Acknowledgment

I express my gratitude to the Commission on Higher Education (CHED) and National Grid Corporation of the Philippines (NGCP) for financially supporting my college studies.

References

[1] D. P. Kumar, B. J. Sowmya, Chetan, 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), 2016, pp. 190-194.

[2] W. Stokowiec, “A comparative study on music genre classification algorithms,” Studies in Big Data, 2016, pp. 123–132.

[3] E. N. Tamatjita and A. W. Mahastama, "Comparison of music genre classification using Nearest Centroid Classifier and k-Nearest Neighbours," 2016 International Conference on Information Management and Technology (ICIMTech), Bandung, Indonesia, 2016, pp. 118-123, doi: 10.1109/ICIMTech.2016.7930314.

[4] D. S. Lau and R. Ajoodha, “Music genre classification: A comparative study between deep learning and traditional machine learning approaches,” Proceedings of Sixth International Congress on Information and Communication Technology, 2021, pp. 239–247.

[5] N. Ndou, R. Ajoodha and A. Jadhav, "Music genre classification: A review of deep-learning and traditional Machine-Learning approaches," 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), Toronto, ON, Canada, 2021, pp. 1-6, doi: 10.1109/IEMTRONICS52119.2021.9422487.

[6] A. Ghildiyal, K. Singh, and S. Sharma, “Music genre classification using machine learning,” 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 1368-1372.

[7] T. Li, M. Ogihara, and Q. Li, “A comparative study on content-based music genre classification,” Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '03, 2003, pp. 282-289.

[8] D. R. I. M. Setiadi, D. S. Rahardwika, E. H. Rachmawanto, C. A. Sari, C. Irawan, D. P. Kusumaningrum, Nuri, and S. L. Trusthi, “Comparison of SVM, KNN, and NB classifier for genre music classification based on metadata,” 2020 International Seminar on Application for Technology of Information and Communication (iSemantic), 2020, pp. 12-16.

[9] A. Elbir, H. Bilal Cam, M. Emre Iyican, B. Ozturk, and N. Aydin, “Music genre classification and recommendation by using Machine Learning Techniques,” 2018 Innovations in Intelligent Systems and Applications Conference (ASYU), 2018, pp. 1-5.