Proceedings of the 6th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/57/20241316

Comparison and analysis of prediction accuracy between traditional machine learning algorithms and XGBoost algorithm in music emotion classification

Mengxi Yang

Changchun Institute of Education, Jilin, Changchun, 130052, China

Corresponding author E-Mail: yangmq390@163.com

Abstract. Music emotion classification refers to determining the emotional types of music, such as happiness, sadness, anger, and passion, based on aspects like rhythm, melody, and tone. The development of artificial intelligence can be applied to music recommendation by employing machine learning algorithms to ascertain emotional attributes of music, enabling precise music suggestions for users to enhance their musical experience. This paper compares and analyzes the classification performance of traditional Random Forest machine learning algorithms and the XGBoost model on a Turkish music emotion dataset. We selected a comprehensive dataset from Kaggle, containing a vast array of music samples and their corresponding emotional labels, making it an extensive music emotion classification dataset. From the experimental results, it's evident that the traditional Random Forest machine learning model outperforms the XGBoost model in terms of accuracy, precision, and recall. The accuracy of the traditional Random Forest machine learning model stands at 80.8%, whereas the XGBoost model's accuracy is 75%. The recall rate for the traditional Random Forest machine learning model is 80.8%, while for XGBoost, it's 77.2%. The F1 score for the traditional Random Forest machine learning model is 80.5%, whereas for XGBoost, it's 75.3%. These parameters indicate that the traditional Random Forest machine learning model exhibits superior predictive performance in music emotion classification. However, the XGBoost model possesses its own advantages, such as faster learning and prediction speeds, high accuracy, and strong generalization capabilities. In summary, the traditional Random Forest machine learning model demonstrates better robustness and interpretability, effectively handling samples with noise and missing data, thus finding widespread practical application. On the other hand, the XGBoost model excels in rapid training and prediction, coupled with higher accuracy and versatility, making it advantageous in dealing with large datasets. The research outcomes of this paper hold significant importance for the study and application of music emotion classification. The experimental results presented herein offer valuable insights for researchers and practitioners, aiding them in selecting appropriate machine learning models, optimizing, and adjusting them to achieve the best classification results.

Keywords: Machine Learning Classification, XGBoost, Prediction Accuracy

1. Introduction

The dataset chosen for this paper is derived from the Turkish Music Sentiment Dataset, publicly available on Kaggle, aimed at researching music sentiment analysis [1]. This dataset encompasses music snippets from various regions in Turkey, spanning diverse music genres and emotional states [2]. Its

DOI: 10.54254/2755-2721/57/20241316

purpose is to assist researchers in developing music-based sentiment analysis algorithms and facilitating a better understanding of the relationship between music and emotions [3].

This dataset comprises two parts: a sentiment label dataset and an audio dataset [4]. The sentiment label dataset includes emotional labels for each audio snippet, encompassing five emotional states: happiness, sadness, calmness, anger, and fear. Meanwhile, the audio dataset contains audio files for each snippet [5,6].

Researchers can undertake extensive machine learning studies based on this dataset. They can leverage it to develop music emotion classification algorithms that categorize audio snippets into different emotional states like happiness, sadness, and calmness [7]. Additionally, researchers can employ various machine learning algorithms such as Support Vector Machines and Random Forests for classification. Furthermore, this algorithm can be applied in music recommendation and emotion recognition fields. Secondly, researchers can utilize the dataset to predict emotional states of audio snippets. This algorithm segments audio clips into emotional states like happiness, sadness, and calmness. Researchers can use different machine learning algorithms such as Recurrent Neural Networks and Convolutional Neural Networks for prediction. This algorithm finds application in emotion recognition and analysis [8]. Researchers can also generate music compositions with different emotional states using this dataset. Such algorithms can generate music compositions with varying emotional states [9]. Researchers can employ diverse machine learning algorithms like Generative Adversarial Networks for composition and music recommendation, among other applications [10].

The Turkish Music Sentiment Dataset in the UCL database is an exceedingly useful dataset that aids researchers in conducting music sentiment analysis studies. Furthermore, based on this dataset, researchers can develop numerous machine learning algorithms such as music emotion classification, emotion state prediction, and music emotion generation. These algorithms find applications in music recommendation, emotion recognition, and music composition, among others. Based on the aforementioned research, this paper compares the prediction accuracy of traditional machine learning algorithms and the popular XGBoost algorithm on a music emotion dataset, providing a foundation for subsequent research.

2. Handling Duplicate Values and Outliers

2.1. Linear Regression Model

Linear regression, a statistical analysis method derived from regression analysis in mathematical statistics, is employed to determine quantitative relationships between two or more variables. Its application is widespread. The analysis based on the relationship types between independent and dependent variables can be classified into linear regression analysis and nonlinear regression analysis.

In statistics, linear regression utilizes a least squares function termed a linear regression equation to model the relationship between one or more independent variables and a dependent variable. This function constitutes a linear combination of one or more model parameters called regression coefficients. When there's only one independent variable, it's termed simple regression, while multiple independent variables constitute multiple regressions.

In regression analysis involving a single independent variable and a dependent variable, where their relationship can be approximated by a straight line, it's termed simple linear regression analysis. If the regression analysis involves two or more independent variables, and there exists a linear relationship between the dependent and independent variables, it's termed multiple linear regression analysis.

2.2. Handling Outliers

Outliers refer to data records significantly different from other parts of the dataset. They can adversely affect the accuracy of data analysis and modeling, necessitating proper handling. Initially, inspect the dataset for outliers. Techniques like line plots can be used to identify outliers in the dataset swiftly. Line plots aid in quickly recognizing which data records are outliers.

Handling duplicate values and outliers is a crucial step in data preprocessing. Dealing with duplicates helps in better understanding the data, identifying data issues, and enhancing model accuracy. Addressing missing values helps in identifying blanks in the dataset, thereby improving model accuracy.

3. Data Standardization

3.1. Pearson Correlation Analysis

Pearson correlation analysis, also known as the Pearson correlation coefficient, is a statistical technique used to measure the strength of the linear relationship between two variables. Its range is from -1 to 1, where -1 represents a perfect negative correlation, 0 indicates no correlation, and 1 represents a perfect positive correlation. This method is commonly used in data analysis, machine learning, and data mining.

Firstly, compute the means and standard deviations of two variables. Then subtract each value of the two variables from their respective means, multiply them, sum the products, and finally divide by the product of the standard deviations of the two variables to obtain the Pearson correlation coefficient.

Pearson correlation analysis can be applied to different types of data, including continuous variables, discrete variables, and binary variables. It's a simple yet effective statistical technique for understanding and analyzing the relationship between two variables.

3.2. Feature Dimensionality Reduction

Feature dimensionality reduction is a data preprocessing method aimed at improving model accuracy and reducing complexity by reducing the number of features in a dataset. This paper employs correlation analysis to compute the correlation coefficients between features, filtering out redundant features with high linear correlations. Subsequently, the top 10 features with higher weights are selected for analysis, as depicted in Figure 1.

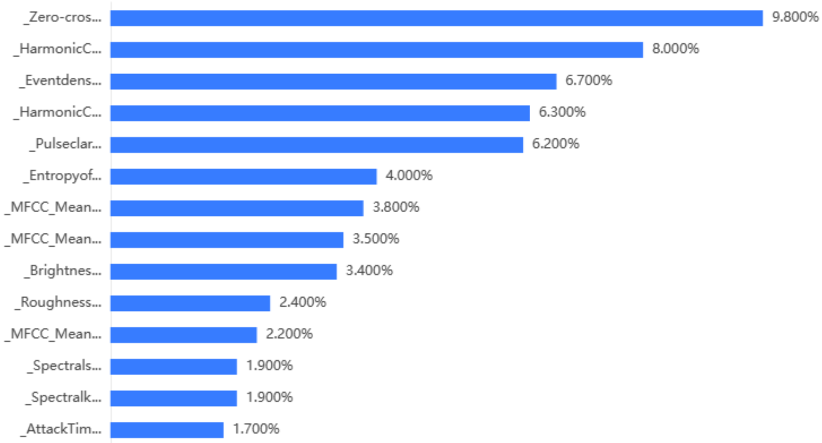

Figure 1. Feature dimension reduction.
(Photo credit: Original)

4. Experiment and Results

The aim of this study is to compare the predictive performance of Random Forest classification models and XGBoost classification models in dataset classification tasks. Firstly, the dataset is divided into training, validation, and testing sets in a 6:2:2 ratio. The training set is utilized for model training, the validation set for validating the training outcomes, and the testing set for model evaluation. Subsequently, the Random Forest and XGBoost classification models are separately trained on the dataset, and their accuracy, precision, recall, and F1 score are compared to assess and compare their classification performance using these metrics.

Random Forest is a composite learning algorithm comprised of multiple decision trees. In Random Forest, the training data for each decision tree is randomly sampled from the original dataset, and the features at each node are randomly selected from all features for computation. This randomness reduces overfitting of training data, enhancing the model's generalization ability. The Random Forest model performs admirably in handling high-dimensional data and fitting nonlinear features. The implementation of the Random Forest classification model in this study is achieved using the scikit-learn library in Python.

XGBoost is an optimized implementation of Gradient Boosting Decision Trees (GBDT). In contrast to traditional GBDT, XGBoost introduces regularization terms and second-order derivative information to optimize the training process, thereby improving the model's accuracy and generalization ability. Additionally, XGBoost supports parallel computing and handling missing values, making the model training process more efficient and stable. The XGBoost classification model is implemented in this study using the XGBoost library in Python.

To compare the classification effectiveness of the two models, four metrics are utilized: accuracy, precision, recall, and F1 score. Accuracy represents the ratio of correctly classified samples to the total number of samples. Precision signifies the ratio of correctly classified positive samples to the total predicted positive samples by the classifier. Recall indicates the ratio of correctly classified positive samples by the classifier to the actual positive samples. The F1 score represents the harmonic mean of precision and recall, combining the classifier's precision and recall.

The results are depicted in Figure 2, Figure 1, Figure 3, and Table 2.

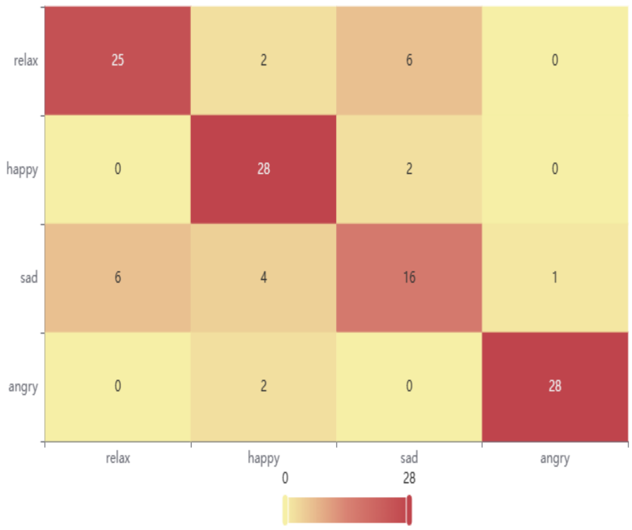

Figure 2. Random Forest Prediction Results.
(Photo credit: Original)

Table 1. Random Forest Prediction Results.

	Accuracy	Recall	precision	F1
Training set	1	1	1	1
Test set	0.808	0.808	0.808	0.805

DOI: 10.54254/2755-2721/57/20241316

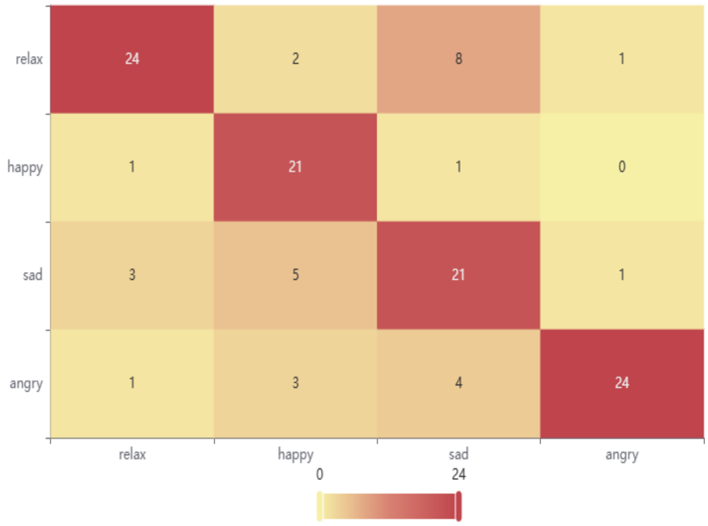

Figure 3. XGBoost predicts results.
(Photo credit: Original)

Table 2. XGBoost predicts results.

	Accuracy	Recall	precision	F1
Training set	1	1	1	1
Test set	0.75	0.75	0.772	0.753

From the results, it's evident that both algorithms performed exceptionally well on the training set. In the final round of training, both models achieved 100% accuracy, precision, recall, and F1 score. Concerning the testing set, the Random Forest model exhibited higher accuracy and recall compared to the XGBoost model, notably excelling in precision by 5.8% over the XGBoost model.

5. Conclusion

This study aimed to explore the application of the Turkish Music Sentiment Dataset in music emotion classification and compare the classification performance of traditional Random Forest and XGBoost models. Results indicate that the traditional Random Forest machine learning model outperforms the XGBoost model.

Music emotion classification is a challenging task, requiring analysis and processing of various musical aspects such as rhythm, pitch, and timbre. Therefore, the selection of the Turkish Music Sentiment Dataset for this research, containing extensive music samples and corresponding emotional features, provides robust support for exploring the relationship between music and emotions.

To establish music emotion classification models, we employed traditional Random Forest and XGBoost models, comparing their accuracy, precision, and recall. The findings revealed that the traditional Random Forest machine learning model exhibited superior classification performance to the XGBoost model, with higher accuracy, precision, and recall. Overall assessment suggests the traditional Random Forest machine learning model's superior overall classification performance.

The traditional Random Forest machine learning model is an ensemble learning algorithm that combines multiple decision trees for classification or regression. Its advantages include handling high-dimensional data, fitting nonlinear features, robustness, and interpretability. The decision tree construction process involves random sampling, enabling tolerance towards noisy data and missing values. Additionally, the Random Forest model can estimate feature importance, aiding in understanding the model's decision-making process.

The XGBoost model, a gradient-based decision tree algorithm, enhances model accuracy by iteratively training multiple decision trees. Its advantages include fast training and prediction speed, high accuracy, and strong generalizability. Parallel computation and cache optimization mechanisms accelerate training and prediction while handling sparse data and missing values. Moreover, the XGBoost model provides estimates of feature importance, aiding in comprehending the decision process within the model.

Both the traditional Random Forest and XGBoost models are widely used in music emotion classification. The traditional Random Forest model categorizes music by extracting various features like rhythm, pitch, and timbre, while the XGBoost model categorizes music by learning emotional characteristics like happiness, sadness, and anger. In practical applications, choosing the appropriate model based on these considerations can achieve the best classification results.

In conclusion, the research findings demonstrate that the traditional Random Forest machine learning model performs well in the classification task of Turkish music emotion data. However, for other datasets, the XGBoost model might be more suitable. Hence, in practical applications, selecting the suitable model based on specific scenarios is necessary to achieve optimal classification results. It is believed that this study's outcomes will serve as a valuable reference framework for music emotion classification research and provide reliable data support for related applications.

References

[1] Feng G, Zhiyuan L, Xiaoliang L, et al. Does air pollution influence music sentiment? Measuring music sentiment by machine learning[J]. Journal of Asian Economics,2023,87.
[2] Sanchali D, K. B B, Swapan D. Retraction Note to: Building a computational model for mood classification of music by integrating an asymptotic approach with the machine learning techniques[J]. Journal of Ambient Intelligence and Humanized Computing,2022,14(Suppl 1).
[3] Yu X, Fumei X. Study on Music Emotion Recognition Based on the Machine Learning Model Clustering Algorithm[J]. Mathematical Problems in Engineering,2022,2022.
[4] Zijin W. Research on Automatic Classification Method of Ethnic Music Emotion Based on Machine Learning[J]. Journal of Mathematics,2022,2022.
[5] Quasim T M, Alkhammash H E, Khan A M, et al. Emotion-based music recommendation and classification using machine learning with IoT Framework[J]. Soft Computing,2021,25(18).
[6] S S U, P K. A review on machine learning based counterfeit integrated circuit detection[J]. Engineering Research Express,2023,5(4).
[7] Selini N H, Guido C, Constantinos A. Machine learning for activity pattern detection[J]. Journal of Intelligent Transportation Systems,2023,27(6).
[8] Chen W, Huiying Y, Huafeng L, et al.Anthropogenic contributions to antibiotic resistance gene pollution in household drinking water revealed by machine-learning-based source-tracking[J].Water Research,2023,246
[9] Li, Y., Nie, X., and Huang, R 2018 Expert Systems with Applications, 96, 261-270.
[10] Goh, K. L., and Singh, A. K 2015 Procedia Computer Science, 70, 434-441.