Journal of Intelligent Computing

Print ISSN: 0976-9005
Online ISSN: 0976-9013

DLINE JOURNALS

JIC 2024;15 (4)
https://doi.org/10.6025/jic/2024/15/4/142-148

Application of Improved Boruta Algorithm in Music Emotion Research

Luyin Shao, Binshan Zhao*
Fine Arts School of Shandong
University, 250100, Jinan
Shandong, China
compdrzhao@sina.com

Received: 2 June 2024
Revised: 20 August 2024
Accepted: 29 August 2024

ABSTRACT

Music can be heard everywhere in the fast-paced and high-stress modern society. With the development of time, people's demand for music has shifted from simple leisure entertainment to seeking emotional resonance. To address this, the big data Boruta algorithm has emerged as an alternative to random forest classification algorithms, enabling more precise classification results. By re-implementing it, we applied the LightGBM algorithm to a Turkish music dataset and used random forest, XGBoost, and LightGBM algorithms for classification prediction. We found their better applicability to these datasets by evaluating these algorithms' accuracy, Kappa coefficient, and Hamming distance.

Keywords: Boruta, Light GBM, Feature Extraction, Music Emotion

1. Introduction

Music is a unique form of art that deeply influences our daily lives. Its various styles can convey the personality and ideas of the artists and allow the audience to experience a wide range of emotional experiences. According to the statistics from the "2022 China Online Music Industry Report," the total number of internet music users in our country has exceeded 658 million, indicating that netizens are paying increasing attention to and appreciating music. With the development of technology, many scholars are making efforts to explore this field in-depth. AnandN et al. categorized music genres into one using deep learning techniques [1]. Additionally, machine learning methods have been widely applied. KNN is a commonly used method that determines the category of a number by arranging the digits in descending order of their similarity based on the Euclidean distance [2]. However, due to the connection between KNN and fuzzy mechanisms, these algorithms may introduce significant decision risks and lead to incorrect judgments. To address this issue, SzulT et al. effectively controlled the weight of each feature by invoking the correlation of feature variables using KNN and fuzzy mechanisms, thereby significantly improving classification accuracy [3]. However, these approaches have high computational complexity, resulting in long processing times for high-dimensional information [4]. To solve this problem, we suggest using dimensionality reduction techniques to process feature variables. For example, Yinger O S's experiment on speech emotion recognition was achieved by utilizing typical similarities, leading to higher classification accuracy [5]. The Boruta algorithm can effectively reduce data loss by selecting a set of features correlated with the dependent variable and achieving dimensionality reduction while retaining data information, significantly different from traditional correlation analysis methods. Additionally, the Boruta algorithm provides stronger interpretability, allowing for effective interpretation of linear combinations of the original variables. By applying various classification techniques such as RF, XGBoost, and LightGBM, we can delve into the classification problem of music emotion [6]. By conducting feature selection on audio data in relevant music datasets and combining the Boruta algorithm with other improved algorithms, we can achieve classification prediction on audio and demonstrate the broader application of the improved Boruta algorithm on this dataset using multi-classification evaluation criteria [7].

2. Related Work

Each song has its unique emotions, and over time, these emotions can also change. Therefore, accurately capturing the overall emotion of each song becomes extremely complex. Additionally, people may have different perspectives on the same thing due to the strong subjectivity of individual emotions. In today's society, humans are increasingly relying on advanced artificial intelligence technology for fast and accurate music retrieval, making emotion-based music search an emerging field with potential future applications. Emotion analysis of sound may become an important field [8]. Since ancient times, with the development of artificial intelligence, the exploration of emotions has followed suit [9]. In this field, emotion analysis of music and model analysis are closely related, as they involve abstracting a complex concept to capture and understand its essence. With the advancement of scientific technology, many people have started using new methods to process audio data. One of the most common practices is to enhance data accuracy by using filtering and other special algorithms [10]. Additionally, mathematical models can be used to simulate different acoustic features to capture better and express the complexity of acoustics. Dimensional emotion analysis techniques have made significant progress, providing more reliable information. In audio search, their advantage lies in transforming complex audio information into easily understandable data. Researchers have made great strides in retrieving categories such as singers and lyrics of music and conducting audio searches from multiple perspectives based on the emotion categories of music. These studies are not limited to speech but encompass various types of audio, such as the semantic characteristics of music and the physical audio features of music, to explore music's emotional connotations better.

By analyzing and processing music signals and using different emotion modeling techniques, the accuracy and reliability of emotion recognition can be significantly improved. Shanghai Jiao Tong University and Beijing Institute of Technology have made significant breakthroughs in music emotion recognition. Scholars have delved into the complex correlation between music feature information and emotion cognition from a machine learning perspective and established an emotion-based Chinese ethnic music retrieval system, providing new insights for music emotion recognition [11,12]. EltamalyBota and others have developed a new music emotion recognition system that accurately captures and analyzes the emotional changes in music, including Gaussian models and Bayesian classification [13,14]. This system is mainly applied to popular music and employs techniques such as BP neural networks and self-organizing maps to automate the emotional classification research of Chinese ethnic music. MIREX and other international research institutions organize a series of global competitions yearly to promote music emotion recognition, enhance research levels in the field, and increase its influence in the academic community [15]. MediaEval has been committed to advancing music emotion recognition since 2008 by organizing various competition activities, providing rich data resources to participants, and promoting academic exchanges to achieve more in-depth research. Due to technological advancements, machine learning and deep learning have made significant progress in recent years, especially in music emotion recognition [16]. However, there is still room for improvement and further strengthening. Additionally, the continuous progress of artificial intelligence technology holds great significance for music emotion recognition [17,18]. Despite the extensive effort required in the training process of artificial intelligence, it has the potential to be widely applied in the field of music emotion analysis. With technological advancements, its potential will be further explored.

3. Boruta Algorithm in Music Emotion

3.1. Basics of the Boruta Algorithm

The Boruta algorithm aims to build a more efficient random forest classification algorithm that can accurately search all possible variables and provide more precise classification results. The final results are obtained by repeatedly screening irrelevant information to random detectors. XGBoost and GBDT have similar objective functions, but they have significant differences. GBDT achieves a more efficient objective function by adding a regularization term to the loss function, improving system performance and reliability.

$$\mathrm{Obj}_{\Theta} = \sum_{i}^{n} l\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{K} \Omega\left(f_k\right) \qquad (1)$$

In this formula, *yi* represents the accurate data of the i^{th} sample.

L(*yi, yi*) is a loss function that measures the difference between the sample and expected values.fk represents the model function added in the kth round, while (*fx*) represents the complexity of the kth tree.By incorporating a regularization term, XGBoost significantly reduces the variance of the objective function, thereby controlling the complexity of modeling and avoiding overfitting.Additionally, this method employs second-order Taylor expansion and first-order and second-order derivatives, which greatly enhances the decay rate of gradients.The birth of the XGBoost methodology greatly improved its performance by enabling fast and efficient iterations, thus saving considerable computational costs.It utilizes two methods, Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB), to achieve more efficient iterations and faster algorithm convergence.When the gradient is lower, the model becomes more capable of accurately identifying problems rather than discarding them, which means reevaluating their impact for more precise problem prediction.To achieve this, samples with large gradients are retained, while those with small gradients are randomly sampled during the training of weak learners, ultimately constructing a more robust, strong learner. This approach ensures accuracy and faster prediction, significantly increasing the model's generalization ability.LightGBM algorithm effectively addresses complexity by leveraging the lossless merging of information from various dimensions. Since sparse information often contains numerous unknown values, and each unknown value in each information can potentially be unknown, we can combine this unknown information with known information to solve complex information processing problems. We can efficiently improve the merge rate of information through the greedy algorithm. Unlike traditional feature sparsity optimization techniques, adopting the EFB algorithm can greatly reduce system memory consumption and accomplish tasks more quickly.There are various evaluation metrics for binary classification, among which the most important ones are Accuracy, Precision, Recall, F1 score, AUC, ROC curve, and PR curve. When considering multi-class problems, there are two approaches to obtain the best results: decomposing it into binary classifications conducting a comprehensive analysis, and using established evaluation metrics for more precise classification.

$$K = \frac{P_0 - P_e}{1 - P_e} \qquad (2)$$

The Kappa coefficient can range from negative to positive or even smaller values, which can be used to measure the agreement among evaluators. Additionally, Hamming distance can also be used to measure the accuracy of multi-class problems, with a range of [0,1]: when the value is close to 0, it indicates a small difference between the predicted and true results, whereas when the value is close to 1, it means a large difference between the predicted and true results.

3.2. Improvements to the Boruta Algorithm

In this study, we conducted in-depth research on base classifiers and proposed three novel classifiers: RF-Boruta, XGB-Boruta, and LGB-Boruta. These classifiers not only exhibit good base classification performance but also effectively enhance classification accuracy. To assess the effectiveness of these three feature selection approaches, we need to apply them to similar datasets and then determine which approach is more effective based on the obtained information. Finally, we need to use various evaluation metrics to measure the performance of these approaches and determine their practical applications.

4. Experimental Design and Analysis

When the Boruta algorithm was applied to the NSL-KDD dataset, we observed its good infinite looping property. Therefore, we took multiple measures to improve the algorithm's performance, including analyzing and estimating random forest parameters and using the chi-square and Gini index as Z-scores. These measures help reduce infinite looping and enhance the algorithm's efficiency.

4.1. Data Description

The data is sourced from the TurkishMusicEmotion Dataset in the UCI database. This database contains four tones from Turkey: joy, reflection, anger, and sadness, with 100 audio clips for each tone, each lasting 30 seconds. Through this data, we can extract features such as Mel-Frequency Cepstral Coefficients (MFCC), tempo, color, spectrum, and harmonic features to understand the emotions behind the tones better. Additionally, the dataset is comprehensive and does not have any missing information. By performing feature selection using RF-Boruta, XGB-Boruta, and LGB-Boruta on the dataset, we identified key features and used them for classification prediction. In the end, we obtained nine different combinations. We calculated each combination's prediction accuracy, Kappa coefficient, and Hamming distance and found differences. With these results, we can compare the performance of these algorithms on this dataset.

4.2. Experimental Results and Analysis

Based on the above results, different feature selection methods yield different outcomes. Therefore, to evaluate the effectiveness of feature selection algorithms on this dataset, we need to ensure the stability of the classification prediction method and then make comparisons to assess their application effectiveness. According to Figure 1, we can evaluate the random forest.

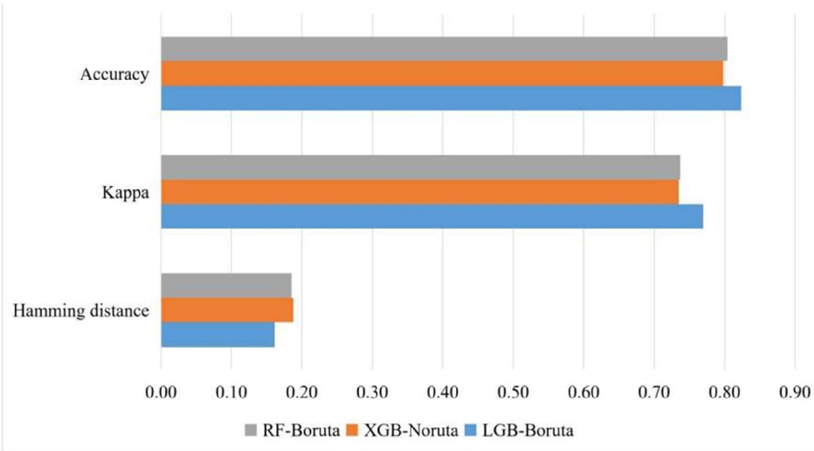

Figure 1. Evaluation of random forest prediction methods

When using the random forest algorithm for classification prediction, RF-Boruta and XGB-Boruta yield the same accuracy and Hamming distance results. However, the Kappa coefficient is significantly higher for RF-Boruta, indicating its superiority over XGB-Boruta. The LGB-Boruta algorithm demonstrates significantly higher accuracy thanthe XGB-Boruta, with a larger Kappa coefficient and a smaller Hamming distance. Therefore, when utilizing random forest as the controlled classification prediction method, the performance of the LGB-Boruta algorithm will be more outstanding. For the XGBoost prediction method, we can employ the evaluation criteria depicted in Figure 2.

145

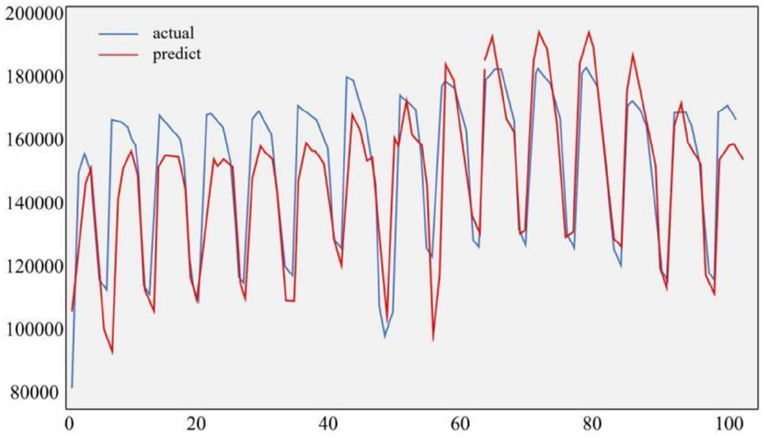

Figure 2. Criterion graph for evaluation under XGBoost prediction method

After applying the XGBoost method for classification prediction, RF-Boruta, XGB-Boruta, and LGB-Boruta algorithms exhibit the same level of accuracy, Kappa coefficient, and Hamming distance. This indicates their good consistency across these three evaluation criteria. However, when comparing the algorithms, XGB-Boruta shows the lowest accuracy (0.713), Kappa coefficient (0.617), and Hamming distance, indicating the poorest overall performance. On the other hand, LGB-Boruta outperforms XGB-Boruta in all three evaluation criteria. The RF-Boruta algorithm demonstrates the best performance among the three, indicating its stronger suitability for this dataset.

Using LightGBM as the underlying classification prediction technique, the results for the three main indicators (Kappa, Kappa coefficient, Hamming distance) are satisfactory. Among them, the results for the three indicators of XGB-Boruta (refer to Figure 3) are closer to the actual situation. When comparing the three evaluation criteria of XGB-Boruta, the overall performance is not satisfactory; the RF-Boruta algorithm performs slightly worse, while the LGB-Boruta algorithm achieves an accuracy exceeding 0.825 and a Kappa coefficient higher than 0.764, even surpassing 0.175. Therefore, the overall effect of the LGB-Boruta algorithm is remarkable and can be considered top-notch.

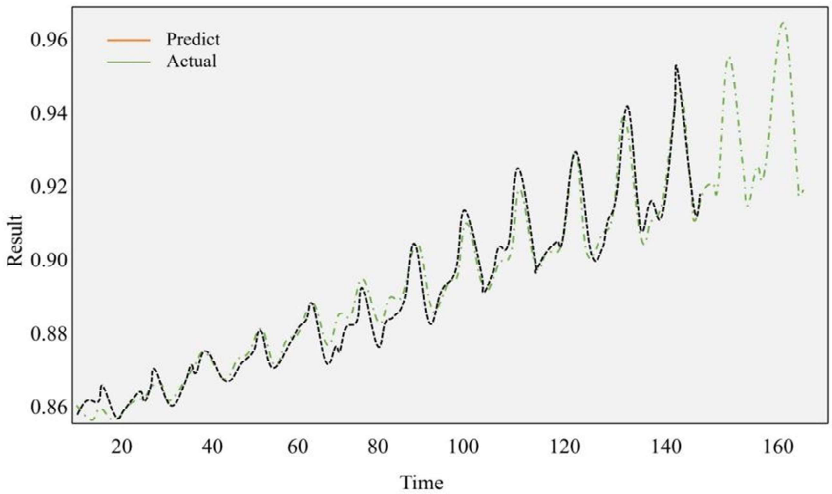

Figure 3. LightGBM provides a method for evaluating prediction criteria

When using the same classification prediction method, we assign scores based on the performance of the feature selection method. The best performance receives a score of 3, 2 for the next best, and 1 for the worst. Please refer to Figure 3 for the specific scoring details. It can be observed that in the original Boruta algorithm and its improved versions, where the random forest is used as the classification wrapper, the overall performance of the LGB-Boruta algorithm is superior, while XGB-Boruta fails to meet the goals of the improved algorithm and performs even worse. By comparing the LGB-Boruta algorithm with the random forest algorithm, it can be observed that the former excels in feature selection for Turkish music data, while the latter exhibits higher accuracy, resulting in better classification prediction performance.

5. Conclusions

We have discovered a new method for extracting emotional content from music signals by analysing the music dataset. The improvement lies in modifying the original random forest-based classification method to utilize XGBoost and LightGBM. With these improvements, we can predict the classification results of music signals more accurately and evaluate the applicability of these algorithms by calculating accuracy, Kappa coefficient, and Hamming distance. After the improvement, the LGB-Boruta algorithm has shown significant improvement in overall performance, thus achieving the improvement goals. Combining this feature selection algorithm with the random forest algorithm yields the best results, making it more reasonable to analyze the emotional content of music signals in the Turkish dataset.

References

[1] Subbiah, S., Anbananthen, K. S. M., Thangaraj, S., et al. (2022). Intrusion detection technique in wireless sensor networks using grid search random forest with Boruta feature selection algorithm. *Journal of Communications and Networks, 24*(2), 264–273.

[2] Anand, N., Sehgal, R., Anand, S., et al. (2021). Feature selection on educational data using Boruta algorithm. *International Journal of Computational Intelligence Studies, 10*(1), 27.

[3] Szul, T., Tabor, S., & Pancerz, K. (2021). Application of the BORUTA algorithm to input data selection for a model based on rough set theory (RST) to predict energy consumption for building heating. *Energies, 14*(10), 2779.

[4] Yinger, O. S., & Springer, D. G. (2022). Using psycholinguistic inquiry and content analysis to investigate emotions in memories of musical experiences. *Musicae Scientiae, 26*(2), 227–242.

[5] Barradas, G. T., & Sakka, L. S. (2022). When words matter: A cross-cultural perspective on lyrics and their relationship to musical emotions. *Psychology of Music, 50*(2), 650–669.

[6] Zhang, X. (2018). The power of arterial language in constructing a musical vocabulary of one's own: Inheriting the inspiration and gene of innovation in electroacoustic music from Chinese culture. *Contemporary Music Review, 37*(1–2), 126–134.

[7] Wang, J., Wang, Q., & Liu, H. (2021). Emotion recognition of musical instruments based on convolution long short-time memory depth neural network. *Journal of Physics: Conference Series, 1976*(1), 012015.

[8] Zhang, H., Meng, L., & Zhang, J. (2022). Why musical emotion can be induced by harmony? The effect and cognitive mechanism of musical consonance. *Advances in Psychological Science, 30*(4), 817–833.

[9] Eltamaly, A. M. (2021). A novel musical chairs algorithm applied for MPPT of PV systems. *Renewable and Sustainable Energy Reviews, 146*(17), 121–135.

[10] Ahmed, A. A. M., Deo, R. C., Feng, Q., et al. (2021). Deep learning hybrid model with

Boruta-Random Forest optimiser algorithm for streamflow forecasting with climate mode indices, rainfall, and periodicity. *Journal of Hydrology, 599*, 126350.

[11] Gómez-Cañón, J. S., Cano, E., Eerola, T., et al. (2021). Music emotion recognition: Toward new, robust standards in personalized and context-sensitive applications. *IEEE Signal Processing Magazine, 38*(6), 106–114.

[12] Xu, L., Wen, X., Shi, J., et al. (2021). Effects of individual factors on perceived emotion and felt emotion of music: Based on machine learning methods. *Psychology of Music, 49*(5), 1069–1087.

[13] Bota, P. J., Wang, C., Fred, A. L. N., et al. (2019). A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals. *IEEE Access, 7*, 140990–141020.

[14] Houssein, E. H., Hammad, A., & Ali, A. A. (2022). Human emotion recognition from EEG-based brain–computer interface using machine learning: A comprehensive review. *Neural Computing and Applications, 34*(15), 12527–12557.

[15] Daly, I., Williams, D., Hallowell, J., et al. (2015). Music-induced emotions can be predicted from a combination of brain activity and acoustic features. *Brain and Cognition, 101*, 1–11.

[16] Cont, A., Dubnov, S., & Assayag, G. (2006). Anticipatory model of musical style imitation using collaborative and competitive reinforcement learning. *Workshop on Anticipatory Behavior in Adaptive Learning Systems*. Berlin, Heidelberg: Springer
Berlin Heidelberg, 285–306.

[17] Tang, H., Zhang, Y., & Zhang, Q. (2022). The use of deep learning-based intelligent music signal identification and generation technology in national music teaching. *Frontiers in Psychology, 13*, 762402.

[18] Nag, S., Basu, M., Sanyal, S., et al. (2022). On the application of deep learning and multifractal techniques to classify emotions and instruments using Indian Classical Music. *Physica A: Statistical Mechanics and its Applications, 597*, 127261.