ISSN 2277-2685
IJESR/June-2024/ Vol-14/Issue-2s/374-383
T. Jayarajan *et. al.,* / International Journal of Engineering & Science Research

A MACHINE LEARNING APPROACH TO MUSIC MOOD CLASSIFICATION FOR EMOTIONALLY INTELLIGENT PLAYLISTS

T. Jayarajan[1], Muchenthula Ashwini[2], S. Swetha[2], B. Yeshwanth Goud[2]
[1]Assist. Professor, [2]UG Scholar, [1,2]Department of Computer Science & Engineering (AI & ML),
[1,2]Kommuri Pratap Reddy Institute of Technology, Ghatkesar, Hyderabad, Telangana.

ABSTRACT:
Music is a fundamental aspect of human life and has the power to evoke various emotions and moods. Understanding and categorizing music based on its emotional content, also known as music mood classification, has become a prominent area of research in recent years. Analyzing the emotional aspect of music is crucial in applications such as personalized music recommendations, mood-based playlist generation, and emotion-aware music therapy. The traditional approach to music mood classification involved employing music experts to listen to each track and manually assign mood labels, such as happy, sad, calm, energetic, etc. This process was highly subjective and prone to inconsistencies due to individual biases. The annotated data would then be used to build handcrafted rule-based systems or simple statistical models to classify music into different mood categories. While these approaches provided some insights, they were limited in scalability, generalization, and accuracy. In addition, manual annotation is time-consuming, expensive, and lacks objectivity. Moreover, human listeners may not always agree on the emotional interpretation of a particular piece of music, leading to discrepancies in the labeled data. To overcome these challenges and enable large-scale mood analysis of music collections, there is a demand for automated and data-driven approaches. Machine learning techniques offer a promising solution to this problem by leveraging computational models to learn patterns and relationships from data, thus enabling the automatic classification of music based on its emotional content. Therefore, this project develops an emotion recognition-based music recommendation system, which performs the mood analysis first, and then recommend the music according to the detected mood of the users. The experiments on real data confirm that the proposed mood classification system can be integrated to any music recommendation engine.
Keywords: Machine Learning, Random Forest, Mood Detection, Music Classification, Pattern Recognition

1. INTRODUCTION

The pursuit of understanding and categorizing the emotional content of music, referred to as music mood classification, has emerged as a central focus in contemporary research, reflecting the profound impact of music on human emotions. The intricate interplay between music and mood is a fundamental aspect of the human experience, with the potential to elicit a myriad of emotions. This research endeavors to delve into the realms of machine learning to revolutionize the conventional methods of music mood classification, which historically relied on the subjective assessments of music experts. Traditionally, experts would meticulously listen to each musical track, manually ascribing mood labels such as happy, sad, calm, or energetic. However, this approach was fraught with subjectivity and inconsistencies, owing to individual biases and varying interpretations of emotional nuances.

The conventional method, although providing valuable insights, faced limitations in terms of scalability, generalization, and accuracy. The reliance on manual annotation not only proved to be time-consuming and expensive but also lacked the objectivity required for a standardized classification system. Furthermore, the inherent variability in human perception meant that different listeners might interpret the emotional content of a particular musical piece divergently, leading to discrepancies in the labeled data. Recognizing these challenges, there emerged a compelling need for automated, data-driven approaches that could overcome the limitations of traditional methods.

Machine learning, as a powerful computational tool, presents a promising solution to the challenges posed by manual annotation and subjective interpretation. By harnessing the capabilities of machine learning algorithms, this project endeavors to develop an emotion recognition-based music recommendation system. The novel approach involves conducting mood analysis as a precursor to recommending music, thereby enhancing the personalization of music suggestions based on the detected mood of the users. The core principle revolves around leveraging computational models to learn intricate patterns and relationships inherent in the data, facilitating the automatic classification of music based on its emotional content. In the realm of traditional music mood classification, the reliance on human experts not only hindered scalability but also raised concerns regarding the replicability of results[1,2,3]. The subjectivity inherent in human perception introduced a level of uncertainty, as different experts might arrive at disparate conclusions when categorizing the emotional content of a given musical composition. The introduction of automated, data-driven approaches offers a transformative shift, where the focus shifts from subjective human judgments to the objective and consistent outcomes generated by machine learning models.[4]

2. LITERATURE SURVEY

The concept that people who listen to different types of music will change emotions and physical states has been widely accepted. Several studies have shown the relationship between psychology and physiology [5,6]. According to Bason et al. [7], the heart rate changes mainly for the following reasons. First, the heart rate is changed by the external auditory stimulation that leads to the neuron coupling into the cardiac centers of the brain, further arousing the sinus entrainment of rhythms. Another cause of changing the heart rate is the autonomic nervous system (ANS) that controls and sustains homeostasis in our body, such as blood pressure, body temperature, and sleep qualities. It mainly consists of the parasympathetic nervous system (PNS) and sympathetic nervous system (SNS). Additionally, it is typically distinguished by opposing characteristics. For instance, in an emergency state, the SNS increases the heart rate, but on the other hand, the PMS typically retards the heart rate in the static state. Some studies indicate that quality of life can be improved by different types of music, such as raising sleep quality, relieving pressure, supporting exercise, and enhancing brain liveliness [8,9,10]. In summary, we could further infer that there supposedly is a connection between music stimuli and heart rate. According to the above-mentioned factors, listening to different music genres changes one's emotions and heart rate, which is a kind of music therapy method. There are some benefits of music therapy, such as socialization, cognition, emotion, and neuron motor function [11]. Continuing music therapy research has led to many new and fascinating applications in sports and autistic and handicapped fields. According to Van Dyck E et al. [12], music rhythm can affect running cadence. In other words, the slower rhythm of music brings out a decrease in running cadence; on the other hand, the faster rhythm of music gives rise to an increase in running cadence. Moreover, another significant research of promoting exercise efficiency is proven by Karow et al. [13]. They provided extensive discussions of the importance of primary selected music. For instance, primary music could make humans more powerful and more stimulated during exercise. Moreover, it could effectively decrease the Rating of Perceived Exertion (RPE), which evaluates the degree of effort that a person feels by themselves. Consequently, music could draw attention away from uncomfortable feelings [14].

3. PROPOSED METHODOLOGY

3.1 Overview

In the domain of music mood classification, the progression from raw data to model predictions involves a structured research procedure comprising key steps such as dataset acquisition, data preprocessing, data splitting, existing XGBoost model training, proposed Random Forest with Adaboost model training, performance evaluation, and the final prediction from test data. Figure 3.1 shows the proposed system model. The journey commences with the acquisition of a comprehensive and diverse dataset. The dataset forms the bedrock of the research, encapsulating a broad spectrum of musical genres, styles, and emotional nuances. Collecting a representative dataset is crucial to ensure the model's ability to generalize across various music types and capture the intricate emotional patterns inherent in diverse musical compositions. This initial step sets the stage for robust and inclusive music mood classification.

Subsequent to dataset acquisition, the data preprocessing phase unfolds. This step involves the systematic cleansing and transformation of raw musical data to make it amenable for machine learning model training. In the context of music mood classification, data preprocessing may encompass tasks such as feature extraction from audio signals, considering elements like tempo, pitch, and spectral characteristics. Additionally, textual data, if available, could be processed to extract relevant information about song lyrics, contributing further to the holistic understanding of the emotional content of music. This preprocessing step ensures that the data fed into the models is structured, informative, and conducive to extracting meaningful patterns related to music mood.

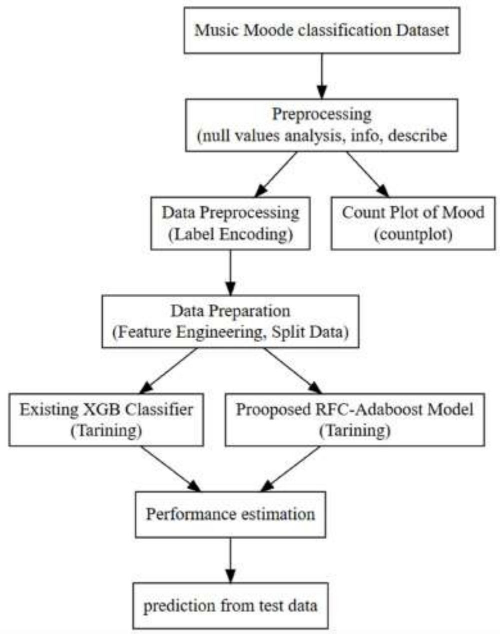


Figure 1. Proposed system model.

Introducing the proposed Random Forest with Adaboost model training, the research aims to leverage the strengths of ensemble learning techniques to enhance the music mood classification process. Random Forest, by aggregating the predictions of multiple decision trees, offers a more stable and less prone-to-overfitting model. Adaboost, as a boosting algorithm, focuses on correcting the errors of the base model, thus potentially mitigating the biases and limitations associated with the XGBoost model. The combination of Random Forest and Adaboost introduces a complementary approach that may capture nuanced emotional patterns in music more effectively.

The performance evaluation stage is pivotal in assessing the efficacy of both the existing XGBoost model and the proposed Random Forest with Adaboost model. A range of evaluation metrics is employed to gauge the models' performance, including accuracy, precision, recall, and F1 score. These metrics provide a multifaceted view of the models' ability to correctly classify different mood categories and their performance on both majority and minority classes. The evaluation process seeks to uncover insights into the strengths and weaknesses of each model, guiding decisions on their suitability for music mood classification tasks.

Finally, the research culminates in the prediction from the test data. Leveraging the trained models, predictions are generated for the mood labels of the test set. This phase serves as the ultimate validation of the models' generalization capabilities and their applicability to real-world scenarios. The predictions are compared against the true labels in the test set, shedding light on the models' accuracy and effectiveness in classifying music based on emotional content.

3.2 Random Forest Algorithm

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output. The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

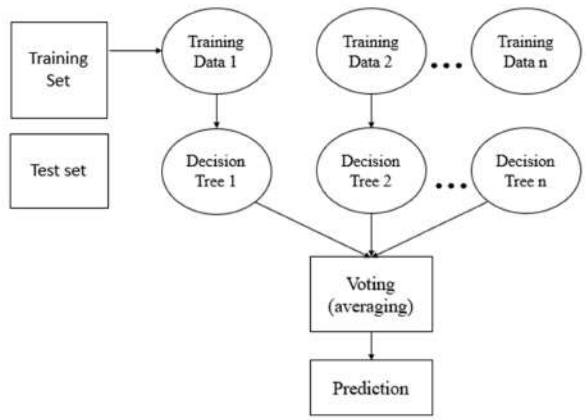


Fig 2: Random Forest algorithm.

3.2.1 Random Forest algorithm

Step 1: In Random Forest n number of random records are taken from the data set having k number of records.

Step 2: Individual decision trees are constructed for each sample.

Step 3: Each decision tree will generate an output.

Step 4: Final output is considered based on Majority Voting or Averaging for Classification and regression respectively.

3.2.2 Important Features of Random Forest

- **Diversity**- Not all attributes/variables/features are considered while making an individual tree, each tree is different.
- **Immune to the curse of dimensionality**- Since each tree does not consider all the features, the feature space is reduced.
- **Parallelization**-Each tree is created independently out of different data and attributes. This means that we can make full use of the CPU to build random forests.
- **Train-Test split**- In a random forest we don’t have to segregate the data for train and test as there will always be 30% of the data which is not seen by the decision tree.
- **Stability**- Stability arises because the result is based on majority voting/ averaging.

3.2.3 Assumptions for Random Forest

Since the random forest combines multiple trees to predict the class of the dataset, it is possible that some decision trees may predict the correct output, while others may not. But together, all the trees predict the correct output. Therefore, below are two assumptions for a better Random Forest classifier:

- There should be some actual values in the feature variable of the dataset so that the classifier can predict accurate results rather than a guessed result.
- The predictions from each tree must have very low correlations.

Below are some points that explain why we should use the Random Forest algorithm

- It takes less training time as compared to other algorithms.
- It predicts output with high accuracy, even for the large dataset it runs efficiently.
- It can also maintain accuracy when a large proportion of data is missing.

3.2.4 Types of Ensembles

Before understanding the working of the random forest, we must look into the ensemble technique. Ensemble simply means combining multiple models. Thus, a collection of models is used to make predictions rather than an individual model. Ensemble uses two types of methods:

Bagging– It creates a different training subset from sample training data with replacement & the final output is based on majority voting. For example, Random Forest. Bagging, also known as Bootstrap Aggregation is the ensemble technique used by random forest. Bagging chooses a random sample from the data set. Hence each model is generated from the samples (Bootstrap Samples) provided by the Original Data with replacement known as row sampling. This step of row sampling with replacement is called bootstrap. Now each model is trained independently which generates results. The final output is based on majority voting after combining the results of all models. This step which involves combining all the results and generating output based on majority voting is known as aggregation.

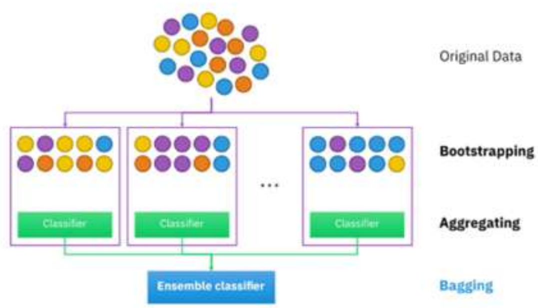
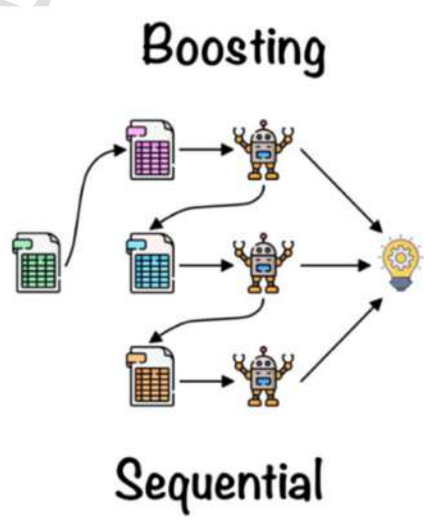


Fig. 3: RF Classifier analysis.

Boosting– It combines weak learners into strong learners by creating sequential models such that the final model has the highest accuracy. For example, ADA BOOST, XG BOOST.



ISSN 2277-2685
IJESR/June-2024/ Vol-14/Issue-2s/374-383
T. Jayarajan *et. al.,* / International Journal of Engineering & Science Research

Fig. 4: Boosting RF Classifier.

3.3 Ada Boost Algorithm

AdaBoost is a machine learning algorithm that belongs to the family of ensemble methods. Ensemble methods combine the predictions of multiple weaker models to create a more accurate and robust model. In the case of AdaBoost, these weaker models are often referred to as "weak learners" or "base classifiers."

Key Idea: The main idea behind AdaBoost is to sequentially train a series of weak learners on the data, assigning higher importance to the data points that are misclassified by the previous models. In other words, it adapts its learning strategy based on the errors made by the previous models.

Training Process:

1. **Initialization:** Each data point is initially assigned an equal weight.
2. **Model Training:** A weak learner (e.g., a decision tree with limited depth) is trained on the data. It focuses on the samples that were misclassified by the previous models.
3. **Weight Update:** The algorithm assigns higher weights to the misclassified data points, making them more influential in the next iteration.
4. **Sequential Iteration:** Steps 2 and 3 are repeated for a predefined number of iterations or until a perfect model is achieved.

Combining Weak Learners: After all the weak learners are trained, their predictions are combined through a weighted sum to form the final strong model. Each weak learner contributes to the final prediction based on its accuracy, with more accurate models having a higher influence.

4 RESULTS AND DISCUSSION

Figure 5 shows that the sample dataset

	name	album	artist	id	release_date	popularity	length	danceability	acousticness	energy	instrumentalness
0	1999	1999	Prince	2H7PHVdQ3mXqEHXcvclTB0	1982-10-27	68	379266	0.866	0.13700	0.7300	0.000000
1	23	23	Blonde Redhead	4HIwL9ii9CcXpTOTzMq0MP	2007-04-16	43	318800	0.381	0.01890	0.8320	0.196000
2	9 Crimes	9	Damien Rice	5GZEeowhvSieFDiR8fQ2im	2006-11-06	60	217946	0.346	0.91300	0.1390	0.000077
3	99 Luftballons	99 Luftballons	Nena	6HA97v4wEGQ5TUClRM0XLc	1984-08-21	2	233000	0.466	0.08900	0.4380	0.000006
4	A Boy Brushed Red Living In Black And White	They're Only Chasing Safety	Underoath	47IWLfIKOKhFnz1FUEUIkE	2004-01-01	60	268000	0.419	0.00171	0.9320	0.000000
...	...	...	...	...	...	...	...	...	...	...	...
681	windcatcher	windcatcher	Leo Nocta	59VApBbrS2IADQk4mi5mdo	2020-06-19	36	123066	0.402	0.96100	0.2360	0.919000
682	yellow is the color of her eyes	yellow is the color of her eyes	Soccer Mommy	4D3nttJPU6L0M2epr7sld6	2019-11-19	5	435080	0.452	0.75700	0.5150	0.120000
683	you broke me first	you broke me first	Tate McRae	45bE4HXI0AwGZXfZtMp8JR	2020-04-17	87	169265	0.642	0.78600	0.3740	0.000000
684	you were good to me	brent	Jeremy Zucker	4CxFN5zON70B3VOPBYbd6P	2019-05-03	76	219146	0.561	0.91300	0.0848	0.000026
685	æfre	æfre	praam	2irbT1BSYalEF44PlyKaoM	2020-07-17	41	186331	0.377	0.99400	0.0156	0.881000

Figure 5. Sample Dataset.

```
Proposed Classifier classification_report:
              precision    recall  f1-score   support

       Happy       1.00      1.00      1.00        36
         sad       0.89      0.85      0.87        39
   energitic       0.70      0.73      0.71        22
        Calm       0.95      0.98      0.96        41

    accuracy                           0.91       138
   macro avg       0.88      0.89      0.89       138
weighted avg       0.91      0.91      0.91       138
```

Figure 6. Proposed method classification report.

Figure 6 is performance of Proposed RF-Adaboost model have a accuracy 91% which is very good model.

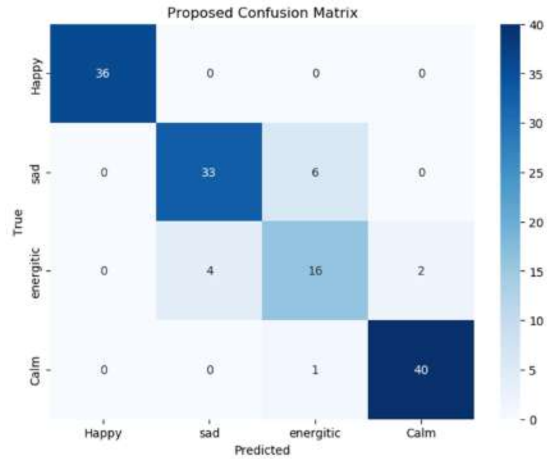


Figure 7. Proposed confusion matrix.

Figure 7 is confusion matrix of the RFC-adaboost model have a nice performance (using heat map)

	name	album	artist	release_date	popularity	length	danceability	acousticness	energy	instrumentalness	liveness	valence	loudness	speechiness
482	482	535	517	20130101	0	166560	0.684	0.905000	0.0646	0.000019	0.0789	0.334	-18.052	0.0522
23	23	500	447	20031110	39	219573	0.429	0.000019	0.9820	0.018500	0.3120	0.583	-2.962	0.0617
189	189	185	515	20191122	47	240000	0.500	0.671000	0.1850	0.000000	0.0827	0.159	-13.209	0.0302
47	47	57	175	20170201	0	290736	0.399	0.848000	0.4430	0.000406	0.0899	0.265	-9.168	0.0272
363	363	205	477	19790101	74	295400	0.586	0.058900	0.7000	0.001250	0.0318	0.897	-9.558	0.0363
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
365	365	323	457	20180518	25	250373	0.665	0.336000	0.5600	0.085700	0.3270	0.414	-7.301	0.0667
362	362	16	278	19990222	72	169026	0.494	0.001290	0.9460	0.000000	0.3980	0.741	-2.757	0.0637
378	378	304	525	20160603	66	237973	0.581	0.893000	0.4000	0.576000	0.1610	0.433	-8.739	0.0291
262	262	104	258	20191025	57	203946	0.432	0.807000	0.3400	0.000000	0.0864	0.449	-7.977	0.0824
666	666	641	11	20200710	49	239500	0.634	0.819000	0.1930	0.000000	0.1130	0.159	-9.503	0.0277

138 rows × 17 columns

```
predictions = adaboost_model.predict(X_test)
predictions
```

```
array([3, 1, 3, 3, 2, 1, 3, 1, 3, 1, 0, 0, 1, 1, 3, 1, 3, 0, 3, 1, 2, 0,
       1, 0, 1, 1, 1, 0, 3, 3, 1, 3, 2, 3, 0, 2, 1, 0, 1, 2, 2, 0, 1, 3,
       3, 0, 1, 1, 1, 0, 3, 1, 3, 0, 3, 0, 1, 3, 0, 1, 2, 3, 1, 0, 3, 2,
       1, 0, 1, 3, 1, 1, 2, 2, 2, 3, 0, 1, 0, 0, 0, 0, 3, 2, 2, 3, 0, 2,
       1, 0, 3, 3, 2, 1, 1, 1, 0, 3, 0, 3, 1, 3, 0, 0, 3, 1, 0, 0, 2, 1,
       0, 0, 2, 1, 0, 3, 2, 1, 3, 2, 3, 0, 2, 2, 3, 3, 0, 3, 1, 2, 0, 0,
       3, 3, 1, 3, 3, 3])
```

Figure 8. Prediction from test data.

Figure 8 shows that the prediction output using adaboost model

5. Conclusion

In conclusion, the exploration of machine learning approaches for music mood classification represents a significant stride toward unraveling the intricate relationship between music and human emotions. The journey from traditional, manual methods relying on human experts to automated, data-driven approaches, particularly leveraging algorithms like XGBoost and Random Forest with Adaboost, showcases the evolving landscape of this field. The dataset, with its wealth of music-related attributes, serves as a valuable resource in this pursuit, enabling a nuanced analysis of songs and their emotional impact.

The dataset's attributes, ranging from fundamental details like song name and artist to more complex musical characteristics such as danceability, valence, and instrumentalness, contribute to a holistic understanding of each musical composition. The inclusion of a mood attribute emerges as a pivotal element, providing a categorical lens through which songs can be classified based on their emotional or thematic essence. This categorization lays the foundation for developing models that can predict and classify music moods, offering applications in personalized music recommendations, mood-based playlist generation, and even extending into the realm of emotion-aware music therapy.

The training and comparison of machine learning models, specifically XGBoost and Random Forest with Adaboost, illuminate the strengths and limitations inherent in these algorithms for music mood classification. XGBoost, known for its efficiency and predictive power, excels in learning intricate patterns from data. However, challenges such as potential overfitting, sensitivity to imbalanced datasets, and limited interpretability underscore the need for alternative models. The introduction of Random Forest with Adaboost, leveraging ensemble learning techniques, presents a promising avenue for addressing these challenges. The combination of these algorithms offers a more stable and potentially less biased approach, enhancing the robustness of music mood classification systems.

The performance evaluation phase, encompassing metrics such as accuracy, precision, recall, and F1 score, provides a comprehensive assessment of the models' effectiveness. These metrics go beyond mere predictive accuracy, offering insights into how well the models perform across different mood categories and their ability to handle imbalanced datasets. The evaluation process serves as a critical benchmark, guiding decisions on model selection and further refinement, ensuring that the chosen algorithm aligns with the specific requirements of music mood classification tasks.

References

[1] Lane, A.M.; Lovejoy, D.J. The effects of exercise on mood changes: The moderating effect of depressed mood. J. Sports Med. Phys. Fit. 2001, 41, 539–545.

[2] Warburton, D.E.; Nicol, C.W.; Bredin, S.S. Health benefits of physical activity: The evidence. CMAJ Can. Med. Assoc. J. 2006, 174, 801–809. [Green Version]

[3] Pedersen, B.K.; Fischer, C.P. Beneficial health effects of exercise—The role of IL-6 as a myokine. Trends Pharmacol. Sci. 2007, 28, 152–156.

[4] Geirsdottir, O.G.; Arnarson, A.; Briem, K.; Ramel, A.; Tomasson, K.; Jonsson, P.; Thorsdottir, I. Physical function predicts improvement in quality of life in elderly Icelanders after 12 weeks of resistance exercise. J. Nutr. Health Aging 2012, 16, 62–66.

[5] Bernardi, L.; Porta, C.; Sleight, P. Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: The importance of silence. Heart 2006, 92, 445–452.

[6] Trappe, H.J. The effects of music on the cardiovascular system and cardiovascular health. Heart 2010, 96, 1868–1871.

[7] Bason, P.; Celler, B. Control of the heart rate by external stimuli. Nature 1972, 238, 279–280.

[8] Karageorghis, C.I.; Terry, P.C.; Lane, A.M.; Bishop, D.T.; Priest, D.l. The BASES Expert Statement on use of music in exercise. J. Sports Sci. 2012, 30, 953–956.

[9] Johnson, J.E. The use of music to promote sleep in older women. J. Community Health Nurs. 2003, 20, 27–35.

[10] Cooke, M.; Chaboyer, W.; Schluter, P.; Hiratos, M. The effect of music on preoperative anxiety in day surgery. J. Adv. Nurs. 2005, 52, 47–55.

[11] Karow, M.C.; Rogers, R.R.; Pederson, J.A.; Williams, T.D.; Marshall, M.R.; Ballmann, C.G. Effects of preferred and nonpreferred warm-up music on exercise performance. Percept. Mot. Skills 2020, 127, 912–924.

[12] Wijnalda, G.; Pauws, S.; Vignoli, F.; Stuckenschmidt, H. A personalized music system for motivation in sport performance. IEEE Pervasive Comput. 2005, 4, 26–32.

[13] Van Dyck, E.; Moens, B.; Buhmann, J.; Demey, M.; Coorevits, E.; Dalla Bella, S.; Leman, M. Spontaneous entrainment of running cadence to music tempo. Sports Med.-Open 2015, 1, 15. [Green Version]

[14] Gallego, M.G.; García, J.G. Music therapy and Alzheimer’s disease: Cognitive, psychological, and behavioural effects. Neurología 2017, 32, 300–308.