Comparison of support vector machine and random forest algorithms for classification of songs for relaxation purposes in individuals with stress disorders

Anusara Khamngam[1], *Wuttichai* Srisodaphol[1*], and *Prem* Junsawang[1]

[1]Department of Statistics, Faculty of Science, Khon Kaen University, Khon Kaen 40002 Thailand

Abstract. The research compares the performance of support vector machine (SVM) and random forest algorithms in identifying songs suitable for relaxation in patients with stress problems. The dataset comprises both Thai and international songs categorized into therapy and non-therapy groups. The results demonstrate that the support vector machine achieves an accuracy of 78%, outperforming the random forest with an accuracy of 72%. Precision and F1-score metrics further emphasize the superiority of the support vector machine in classification. Notably, the support vector machine has recall rates of 50% and 100% for therapy and non-therapy classes, respectively, while the random forest has recall from class therapy of 38% and class non-therapy of 100%. The findings suggest that providing individuals with stress issues the opportunity to listen to stress-reducing music can be a viable approach to reducing the need for psychiatric therapy. The support vector machine is a better algorithm than the random forest for classifying songs for relaxation because it is more accurate, precise, and has more even recall rates.

Keywords. Music therapy, random forest, support vector machine

1 Introduction

Currently, society, economy, science, and technology are changing rapidly, affecting individuals' mental health. For example, an uncertain or changing economic situation can lead to stress and anxiety, especially in cases of losing jobs or reduced income. In the past decade, the mental health situation of Thai people has been on an increasing trend based on the increased number of suicides from reports of patients receiving psychiatric treatment, from 5.9 cases per 100,000 people in 2010 to 7.37 cases per 100,000 people in 2020 [1]. In 2022, the number of people receiving psychiatric treatment for depression was ranked among the

*Corresponding author: wuttsr@kku.ac.th

top five, with 1.5 million people. However, regrettably, only 28 percent of these patients were able to obtain the necessary access to treatment [2]. Music therapy is an effective method for treating and enhancing mental well-being by directly stimulating brainwave activity with sound waves, thereby alleviating stress and anxiety. The appropriate rhythm of the music, or the rhythm that matches the beat of the heart rate, will create a state of balance and comfortable feelings. The internal components of relaxing music encompass various properties, including tempo, melody, beat, harmony, rhythm, complexity, and melodic range [3].

Many researchers have investigated the application of music in therapy. Kalapatapu et al. [4] investigated the influence of feature selection strategies on the classification accuracy of distinct Indian music genres using various classifiers. Four feature selection methods - genetic algorithm, forward feature selection, information gain, and correlation - were compared across four classifiers: Decision tree C4.5, K-Nearest Neighbors, neural network, and Support Vector Machines (SVM). Utilizing feature sets derived from preprocessed songs via the MIR Toolbox in MATLAB, encompassing rhythm-based, timbre-based, pitch-based, tonality-based, and dynamic features, the experiments focused on Carnatic, Hindustani, and Bollywood music styles. Results indicated that information gain-based feature selection consistently outperformed other methods, with neural networks and SVMs exhibiting superior classification performance. Leubner and Hinterberger [5] did a systematic analysis of 28 primary research studies that employed music or music therapy as a means of intervention for addressing depressive symptoms. A differentiation was established between passive engagement with music by listening and active involvement through singing, playing, or improvising with instruments. The experimental group exhibited a statistically significant decline in depression levels over time in comparison to the control or comparison groups. Music therapy, specifically, greatly enhanced the quality of life for elderly folks. Group settings were more commonly utilized than individual sessions, and the findings indicated a somewhat superior outcome in those cases. Tang et al. [6] performed a meta-analysis of randomized controlled trials (RCTs) to evaluate the impact of music therapy and music medicine on depression. They conducted a comprehensive search across many databases to identify pertinent studies and ultimately incorporated a total of 55 RCTs in their analysis. The findings demonstrated that music therapy, including music treatments within a therapeutic alliance, effectively diminished depression symptoms in comparison to the control group. Conversely, music medicine, primarily including the administration of prerecorded music by medical professionals, exhibited a more substantial impact on reducing symptoms of depression compared to the control group. Zhang [7] proposed a model employing one-dimensional convolution, which better aligns with the temporal expansion of audio signals. By utilizing convolutional gating mechanisms and residual connections, the model achieved high-level abstract features crucial for music classification. Experimental results on the GTZAN dataset demonstrated superior performance compared to SVM, Convolutional Neural Networks (CNN), and other models, particularly in capturing sound spectrum characteristics. The study explored the effectiveness of different pooling methods in aggregating global features, showing that combining multiple global pooling features enhances classification accuracy. The model's ability to adaptively determine relevant features during the learning process contributes to its effectiveness. Namsanit [8] employed a support vector machine to classify Thai music. Thai music encompasses various genres, such as Thai folk, Thai classical, and Thai traditional music. The classification of Thai songs was divided into two stages. Each stage utilized a radial basis function kernel with the parameters $C = 100$ and $\sigma = 1$. The experimental findings demonstrate that the support vector machine achieves a classification accuracy of 86.7%. Chaudhury et al. [9] studied various types of music features and audio signals. Their objective was to classify similar types of songs using suitable tags or indexes (genres) to efficiently provide comparable tracks to the user. They employed the GTZAN dataset, a publicly accessible resource on Kaggle that is frequently employed in studies to classify complicated musical genres. The dataset comprises 10,990 music clips spanning over 10 distinct genres, namely blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae, and rock. They employ a variety of scalable machine learning algorithms, including Naïve Bayes, decision trees, logistic regression, and random forest. The findings demonstrate that the random forest classifier is the most effective classifier for music genre categorization, surpassing other Apache Spark classifiers with an accuracy rate of 90%. Modran et al. [10] presented an experiment aiming to create a machine-learning model capable of predicting the therapeutic effects of specific songs on individuals based on their musical and emotional preferences. Acoustic features extracted from music, including spectral, temporal, melodic/harmonic, and rhythmic characteristics, are analyzed using Python libraries such as Librosa. The study focused on identifying relevant features for accurate prediction and employed deep learning algorithms for model training. The developed model achieved a high overall accuracy in predicting the therapeutic effects of songs across four mood categories: happy, sad, energetic, and calm.

Providing individuals with stress issues the opportunity to listen to stress-reducing music can potentially result in a reduction in the necessity for psychiatric therapy for depression. This paper aims to compare the support vector machine and random forest methods in identifying songs for relaxation purposes in patients with stress problems. The dataset utilized in this study comprised both Thai and international songs. The data is categorized into two distinct groups: therapy and non-therapy.

2 Materials and method

2.1 Dataset

The dataset utilized in this investigation is obtained from three Thai music therapists who used the songs for relaxation purposes in individuals with stress disorders. The songs are categorized into two distinct groups: The selected songs, curated by a music therapist, are categorized in the first group to treat patients with stress disorders. The total number of songs is 36, consisting of 25 Thai songs and 9 international songs (categorized as therapy). Within the second group, music therapists assess a selection of Thailand's Top Charts 100 songs from JOOK's Top 100 of 2022. We have 46 songs that are evaluated by 3 therapists and deemed unsuitable for therapy. The collection comprises 38 Thai songs and 8 international songs, specifically classified as non-therapy. Table 1 provides a comprehensive overview of the dataset's specifics.

Table 1. An overview of the dataset.

Group	Total	Thai songs		International songs	
		Count	Percent	Count	Percent
Therapy	36	27	75%	9	25%
Non - therapy	46	38	83%	8	17%
Total	82	65	79.27%	17	20.73%

2.2 Data preparation

The dataset is prepared for the classification of relaxation music in individuals with stress disorders. Five features (Mel-Frequency Cepstral Coefficients, Spectral Centroid, Spectral Bandwidth, Zero Crossings Rate, Tempo) are an important part of classifying relaxation music for relaxation purposes in individuals with stress disorders. By choosing a sound that can create calm and relaxation that the listener feels comfortable with. Utilizing the data set provided in Table 1, we proceed to prepare the data by following the subsequent steps:

Step 1 Convert music files to.mp3 by running the Python package named Librosa and using the import Librosa command.

Step 2 Extract 5 features (feature extraction); for each feature, the sampling rate is set to 44.10 kHz, and the length of the frame is to have 512 samples for each frame. Frames overlap 256 samples at a time. If the last frames with a size less than 512 are discarded and the average of each feature is found, except for the Tempo, so that the number of features for every song is the same.

- Mel-Frequency Cepstral Coefficients (MFCC) (specifying the order of MFCC in the first 20 sequences (subframes) in each frame and finding the average of the MFCC in the first 20 sequences of each frame). Fig. 1. illustrates the segmentation of the audio signal into subframes.

- Spectral Centroid

The spectral centroid determines the center of the audio signal in each frame using librosa.feature. spectral_centroid ()

- Spectral Bandwidth

The spectral bandwidth value is a value showing the width of the spectrum that covers the frequency in each frame of that signal using librosa.feature. spectral_bandwidth ()

- Zero Crossings Rate

The zero crossings rate is an expression of the frequency signals present in the music signal, which is between 0 and 1 using librosa.feature. zero_crossing_rate ().

- Tempo

The number of times a tempo occurs in one minute. It will be greater than 0 if the speed is 134 BPM, using librosa.feature. tempo ().

Step 3 Normalize all the features.

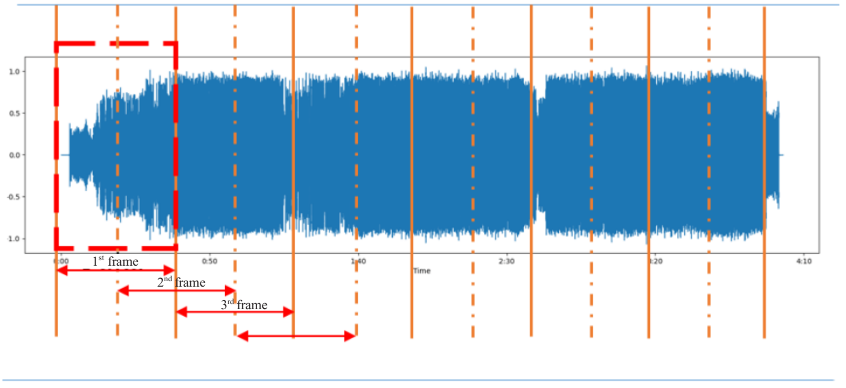

Fig. 1. The segmentation of the audio signal into subframes.

2.3 Model training

The dataset utilized in this study comprises 82 songs, which have been randomly partitioned into an 80% training set and a 20% test set. Fig. 2 illustrates the procedure.

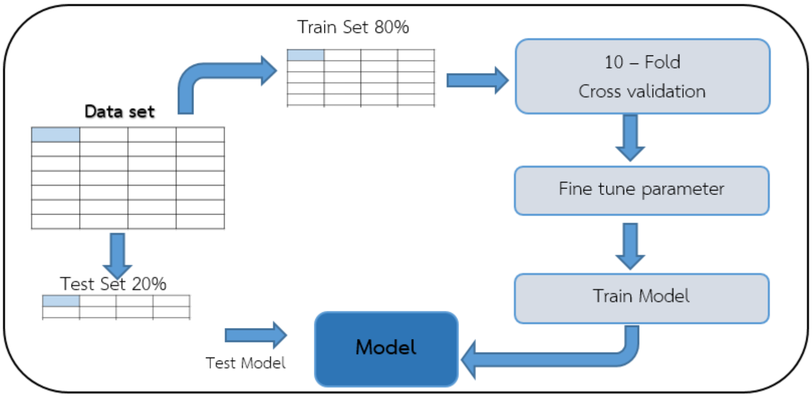

Fig. 2. The procedure of the prediction model.

The training set comprises a total of 64 songs, with 28 songs designated for therapy and 36 songs designated for non-therapy. The training set employs 10-fold cross-validation to determine the optimal parameters for classification using two algorithmic approaches.

The parameters of the support vector machine and random forest are set as follows:

- Support vector machine
  Kernel set to 'liner', 'rbf' and 'poly'
  Gamma set to 0.001, 0.01, 0.1 and 1
- Random forest:
  criterion set to 'entropy' and 'gini'
  n_estimators set to 100, 200, 400, and 500
  maxdepth set to 5, 10, 15, 20, 25 and 30

After setting the parameters that yield the highest accuracy for the model in the training data, the model is retaught to identify the optimal model. The prediction model is subsequently evaluated using a test set. The test set comprises a total of 18 songs, with 8 songs designated for therapy and 10 songs designated for non-therapy. Then the results are presented as a confusion matrix. Accuracy, precision, recall, and F1-score are used as the criteria for comparing the performance of the algorithms.

3 Results

The support vector machine and random forest are used to classify the songs for treating individuals with stress disorders; the following results are shown:

In the results of the support vector machine using 10-fold cross-validation for training data, the optimal parameters are the kernel = "poly" and "gamma" = 0.001 with an accuracy of 66%, which is shown in Table 2.

Table 2. Results for finding the best parameter values with 10-fold cross-validation of SVM.

Kernel	'liner'		'rbf'		'poly'	
	Gamma	Accuracy	Gamma	Accuracy	Gamma	Accuracy
	0.001	0.562	0.001	0.562	0.001	0.660
	0.01	0.562	0.01	0.562	0.01	0.562
	0.1	0.629	0.1	0.562	0.1	0.562
	1	0.648	1	0.562	1	0.562

In the results of the random forest using 10-fold cross-validation for training data, the optimal parameters are “criterion” = “entropy”, “n_estimators” = 100 and “max_depth” = 5. with an accuracy of 64%, which is shown in Table 3.

Table 3. Results for finding the best parameter values with 10-fold cross-validation of random forest.

n_estimators	maxdepth	Accuracy	
		entropy	Gini
100	5	0.640	0.640
	10	0.640	0.640
	15	0.640	0.640
	20	0.640	0.640
	25	0.640	0.640
	30	0.640	0.640
200	5	0.638	0.607
	10	0.626	0.607
	15	0.626	0.607
	20	0.626	0.607
	25	0.626	0.607
	30	0.626	0.607
400	5	0.607	0.604
	10	0.607	0.590
	15	0.607	0.590
	20	0.607	0.590
	25	0.607	0.590
	30	0.607	0.590
500	5	0.607	0.590
	10	0.610	0.576
	15	0.610	0.576
	20	0.610	0.576
	25	0.610	0.576
	30	0.610	0.576

The optimal parameters of both classifiers are then assigned to the classifiers for the test set. The SVM shows an accuracy of 72%, and the random forest shows an accuracy of 61%, as presented in Table 4. Classification results are displayed in the confusion matrix presented in Fig. 3-4.

Table 4. Results of the SVM and random forest for classification.

Classifier	Accuracy	Precision		Recall		F1-score	
		Therapy	Non Therapy	Therapy	Non Therapy	Therapy	Non Therapy
SVM	0.780	1.000	0.710	0.500	1.000	0.670	0.830
Random forest	0.720	1.000	0.670	0.380	1.000	0.550	0.800

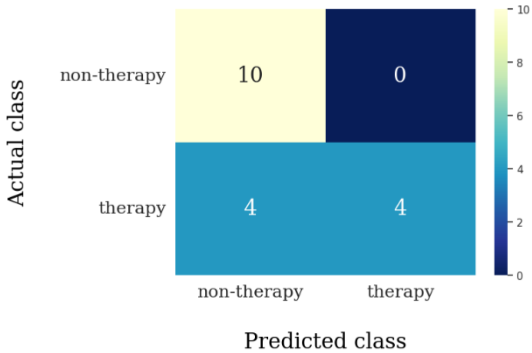

Fig. 3. Classification results in confusion matrix for support vector machine.

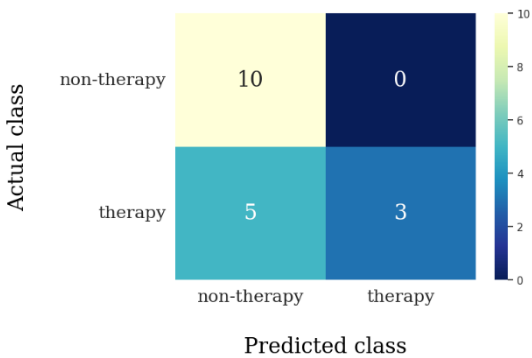

Fig. 4. Classification results in confusion matrix for random forest.

From Table 4, the results show that the support vector machine has an accuracy of 78% and the random forest has an accuracy of 72%. Based on precision, and F1-score, the support vector machine provides better algorithmic performance in classification. Recall from class therapy is 50% and class non-therapy is 100% based on the support vector machine. Based on the random forest, recall from class therapy is 38% and class non-therapy is 100%. It is indicated that the random forest identifies true positives of the therapy and non-therapy.

4 Conclusion and discussion

In conclusion, this study delved into the potential benefits of utilizing stress-reducing music as an alternative or complementary approach for individuals with stress-related issues, potentially reducing the reliance on psychiatric therapy for depression. The comparative analysis between support vector machine and random forest algorithms in identifying suitable relaxation songs for patients yielded intriguing insights.

The dataset utilized in this investigation is obtained from three Thai music therapists, encompassing both Thai and international songs, who used the songs for relaxation purposes in individuals with stress disorders and were meticulously categorized into therapy and non-therapy groups. This study has selected 5 important characteristics to classify: Mel-Frequency Cepstral Coefficients (MFCC), Spectral Centroid, Spectral Bandwidth, Zero Crossings Rate, and Tempo. The results of the classification algorithms revealed that the support vector machine exhibited superior performance with an accuracy of 78%, surpassing the random forest's accuracy of 72%. Precision, recall, and F1-score further emphasized the efficacy of the support vector machine in song classification, particularly in distinguishing between therapy and non-therapy categories.

The recall metrics highlighted a noteworthy difference between the two algorithms. The support vector machine demonstrated a recall of 50% for therapy and 100% for non-therapy, suggesting a balanced ability to identify songs in both categories. In contrast, the random forest exhibited a higher recall for non-therapy (100%) but struggled with identifying therapy songs (38%).

In essence, these findings suggest that the support vector machine is a more robust algorithm for accurately classifying stress-reducing songs, particularly those suitable for therapeutic purposes. This research contributes valuable insights into the intersection of music therapy and machine learning, opening avenues for further exploration and potential applications in personalized therapeutic interventions for individuals grappling with stress-related issues.

This research is fully supported by the Department of Statistics, Faculty of Science, Khon Kaen University.

References

1. Department of Mental Health. https://suicide.dmh.go.th/report/suicide/stat_prov.asp
2. Department of Mental Health. https://dmh.go.th/news-dmh/view.asp?id=31459
3. D. Elliott, R. Polman, R. McGregor, J. Music Ther. **48,** 3 (2011)
4. P. Kalapatapu, S. Goli, P. Arthum, A. Malapati, Procedia Comput. Sci. **98**, (2016)
5. D. Leubner, T. Hinterberger, Front. Psychol. **8**, (2017)
6. Q. Tang, Z. Huang, H. Zhou, P. Ye, PloS one. **15**, 11 (2020)
7. J. Zhang, Sci. Program. Volume 2021, Article ID 1651560 (2021)
8. A. Namsanit, *Thai Music Group Classification Using Support Vector Machines*, *M.S. Tech.Ed. Thesis,* (King Mongkul's Instute of Techonology North Bangkok, Thailand, 2006)
9. M. Chaudhury, A. Karami, M. A. Ghazanfar, Electronics. **11**, (2022)
10. H. A. Modran, T. Chamunorwa, D. Ursutiu, C. Samoila, H. Hedesiu, Sensors. **23**, 2 (2023)