

RESEARCH

Open Access



Predicting game-induced emotions using EEG, data mining and machine learning

Min Xuan Lim¹ and Jason Teo^{2,3*}

Abstract

Background Emotion is a complex phenomenon that greatly affects human behavior and thinking in daily life. Electroencephalography (EEG), one of the human physiological signals, has been emphasized by most researchers in emotion recognition as its specific properties are closely associated with human emotion. However, the number of human emotion recognition studies using computer games as stimuli is still insufficient as there were no relevant publicly available datasets provided in the past decades. Most of the recent studies using the Gameemo public dataset have not clarified the relationship between the EEG signal's changes and the emotion elicited using computer games. Thus, this paper is proposed to introduce the use of data mining techniques in investigating the relationships between the frequency changes of EEG signals and the human emotion elicited when playing different kinds of computer games. The data acquisition stage, data pre-processing, data annotation and feature extraction stage were designed and conducted in this paper to obtain and extract the EEG features from the Gameemo dataset. The cross-subject and subject-based experiments were conducted to evaluate the classifiers' performance. The top 10 association rules generated by the RCAR classifier will be examined to determine the possible relationship between the EEG signal's frequency changes and game-induced emotions.

Results The RCAR classifier constructed for cross-subject experiment achieved highest accuracy, precision, recall and F1-score evaluated with over 90% in classifying the HAPV, HANV and LANV game-induced emotions. The 20 experiment cases' results from subject-based experiments supported that the SVM classifier could accurately classify the 4 emotion states with a kappa value over 0.62, demonstrating the SVM-based algorithm's capabilities in precisely determining the emotion label for each participant's EEG features' instance.

Conclusion The findings in this study fill the existing gap of game-induced emotion recognition field by providing an in-depth evaluation on the ruleset algorithm's performance and feasibility of applying the generated rules on the game-induced EEG data for justifying the emotional state prediction result.

Keywords EEG-based emotion recognition, Data mining, Machine learning, Game-induced EEG

*Correspondence:

Jason Teo

jwteo@ums.edu.my

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

Emotion is a human's fundamental expression to deliver their thoughts of feelings such as happiness, sadness, anger, fear, hate and love. Human emotions are the feelings associated with a certain physical arousal, feeling feedback behavior reflected to the observed environment and inner awareness of those feelings. Emotion is essential to human existence since it substantially affects social interactions and life decisions. For the past decades, the appearance of electroencephalography (EEG) headsets that are commercially available at an affordable price has engaged researchers to study the relationship between each physiological signal and the emotion elicited from certain observed information in form of visual, audio and audio-visual (Soleymani et al. 2016; Bhatti et al. 2016; Khan and Rasool 2022).

The emotion models discovered in recent literature have identified that the human mind state could be expressed in discrete and dimensional space (Abdulrahman et al. 2022). The originally proposed discrete emotion models by Robert Plutchik have included eight primary emotions which have both the positive and negative emotion expression where the more complicated emotions could be produced from the combination of those basic emotions (Gupta et al. 2019). While the dimensional emotion model such as the arousal-valence model expresses human emotion as a coordinate in the arousal-valence plane where the arousal plane indicates the level of human brain activity such as from inactive to active status and the valence plane is indicative of the emotion category which belongs to either positive or negative (Gupta et al. 2019).

The recent emerging fields of human-computer interaction (HCI) and artificial intelligence machine learning algorithms have advanced emotion study by applying machine learning algorithms to recognize human emotion using the features of human facial expression, physiological information and personality traits (Saxena et al. 2022; Henia and Lachiri 2017; Khan and Rasool 2022). As a result, the emotion recognition studies performed by Saxena et al. (2022) and Bhatti et al. (2016) have demonstrated a great performance with the implementation of machine learning algorithms and those features. This implies there is a potential application of those algorithms in real life to identify the human's mental states for an appropriate medication schedule and stress reduction using a human-computer interaction system (Dhuheir et al. 2021; Mikuckas et al. 2014). According to a recent survey study by Alarcao and Fonseca (2019), human emotion is commonly assessed using the physiological signals acquired from the central and autonomic nervous system and functional neuroimaging techniques. Among each physiological signal utilized to recognize

human emotion, electroencephalography (EEG) signals obtained using functional neuroimaging techniques have gained great attention from researchers because of their non-invasive signal recording capabilities, availability of various types of EEG equipment, and unique ability to respond to the human emotional state. The EEG signals recorded by observing the scalp electrical activity have been divided into 5 distinct ranges that correspond to the human mental state (Alarcao and Fonseca 2019).

In this paper, the association between the potential EEG signal's features pattern and the emotion elicited by video game play will be determined by using data mining techniques. The association between the computer games genre and the EEG signal phase variation will be investigated for providing insight into how the EEG signal is altered when a human plays computer games with different genres. This paper also attempted to identify the most reliable supervised classifier built using different feature extraction techniques and different classification algorithms to recognize and predict the human emotion's category based on the arousal-valence model.

Literature review

EEG's feature extraction and feature selection techniques for EEG signals dataset

Abdulrahman et al. (2022) utilized the empirical mode decomposition (EMD) and variation mode decomposition (VMD) to decompose the EEG signals from the Gameemo dataset into 14 intrinsic mode functions (IMFs). The statistical features of each IMF included maximum, minimum and average value were then computed as the feature vector to be fed into the classification models. As a result, the decomposed dataset has led to a smaller data size which could ease the computational cost of performing the classification task. According to Khan and Rasool (2022), the ability of time domain analysis in providing a clear illustration of the EEG data, the ease of EEG data interpretation using frequency domain analysis and the widespread use of Discrete Wavelet Transform (DWT) for time-frequency domain analysis played a main role on their selection of feature extraction techniques. The hybrid domain was then formed by the combination of the features obtained from 3 different domains that could greatly affect the emotion classifiers' decision-making in classifying their EEG data into 4 basic types of emotion (happy, relaxed, bored and stressed). Abdulrahman and Baykara (2021) conducted research focusing on evaluating the shallow and deep learning classifiers' performance on classifying human emotions with 4 different feature extraction techniques applied on Gameemo dataset. They extracted the normalized EEG signals' frequency-domain and statistical characteristics and utilized those extracted features to perform both the

binary and 4-class arousal-valence emotion classification. Their findings indicated that the EEG features that were extracted by the combination of statistical features and wavelet packet decomposition method has led to the optimal classifiers' performance for both different classification tasks. However, the reliability of those classifiers' performance in repeatedly conducting the experiment using different training datasets and models' performance on other EEG signals datasets remained unknown. Abu et al. (2022) investigate the nature of human emotion based on the EEG signals elicited by multiple series of VR games. They have reviewed the major studies that adopted the Gameemo dataset were performing the emotion recognition task according to the game labels but not the self-reported valence-arousal level. In other words, those studies were classifying the EEG patterns into the emotion category that represented by the pre-defined game genres. To overcome this identified gap, they manually labelled each participant's EEG dataset with the arousal-valence emotion level according to the self-assessment manikins. To indicate the performance of the classifiers on distinguishing the EEG into the respective emotional state by separately following two different labels, the time domain features calculation of mean and standard deviation were utilized to extract the EEG features and fed the computed features into the classifiers for further classifiers' training and testing purposes.

Research conducted by Tuncer et al. (2022) suggested a lightweight emotion identification model using input characteristics generated via a unique textural feature generating approach inspired by the Tetris game. They implemented the DWT approach to obtain the low and high sub-bands for providing information on coarse input EEG signals and detailed EEG signals with precipitous transitions, respectively. Gupta et al. (2019) implemented the mRMR algorithm to select the most meaningful generated features to be fed in the classification model as input for high-performance classification task. Dogan et al. (2021) invented a novel emotion recognition approach with the implementation prime pattern and tunable q-factor wavelet transform (TQWT) techniques to segment the EEG data both low and high oscillatory sub-bands. The mRMR algorithm and SVM model were applied on the extracted features vectors to select the most informative features with low misclassification rate that could reveal the major EEG data trends. To promise the classifier's performance on recognizing the emotion with respect to arousal, dominance and valence aspects, the mRMR algorithm was implemented to select the most distinctive features as the classifier's input from the filtered feature vectors.

Hazarika et al. (2018) conducted an experiment investigating the effect of video games on the human inhibitory

control mechanism with respect to EEG signals. The frequency-domain statistical features regarding alpha, beta and gamma sub-frequency bands were extracted and denoised from the collected EEG signals using DWT and Wavelet based, respectively, after going through the pre-processing step. Their findings indicate that the alpha sub-frequency band actively influence the human response inhibition and the linear and nonlinear SVM classifiers' performance could be advanced by considering all the sub-frequency bands' statistical features as input.

Li et al. (2018) conducted an emotion recognition study examining the influence of using the different EEG frequency bands and vary number of channels in determining the emotion classifier's performance. To obtain the EEG features that could comprehensively explain the data trends regards to both spatial and temporal aspects, DWT method was computed on DEAP dataset to capture the time-frequency properties meanwhile to reduce the feature's dimensionality for ensuring the constructed classifier would not process the EEG features exhaustively. Their findings suggested that the combination of all EEG features obtained from each channel and gamma frequency band was associated with the valence-arousal emotion classification rate.

Methods

The methodology of this research study contains 6 stages which are EEG data acquisition, feature extraction on pre-processed EEG data, feature selection on extracted EEG signal data, construction of human emotion classifiers, performance evaluation of classifiers, and association rule mining analysis on relationship between EEG features' pattern and elicited emotion and computer game genre. Figure 1 depicts the methodology of this research.

Software tools and programming language

Jupyter Notebook is a web-based interactive development environment that support the user with a wide range of popular programming language in conducting the multiple data analysis and scientific programming tasks. The flexibility of Jupyter Notebook in providing the author with multiple ways in producing a variety of interactive outputs through the programming works. This could provide the author with more perspectives in conducting the data analysis tasks that involve the complex dimensional dataset. The Python programming language is deployed thorough this paper for conducting the human emotion classifiers construction and EEG features extraction tasks. Python language provides a variety of libraries that support EEG data feature extraction and data analysis could ease the author with the least coding effort in producing the

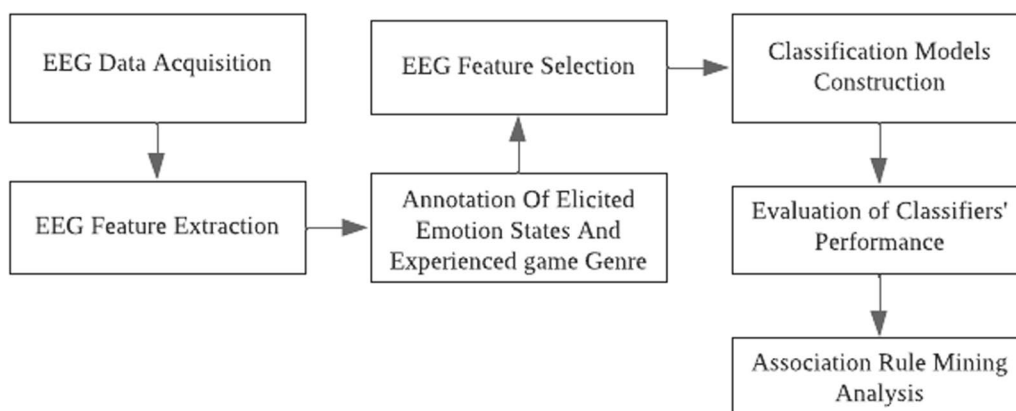


Fig. 1 Methodology of this research study

meaningful analysis result regarding EEG features. The R language implemented in the RStudio IDE is applied in this study as it supports the beginner-friendly data splitting functions and a variety of machine learning algorithm libraries including Ripper, RCAR and SVM classification algorithm.

EEG data acquisition

The publicly available Gameemo dataset is obtained from the Kaggle platform; its contributors are Alakus et al. (2020). The pre-processed EEG signals data using built-in 5th-order sinc filter collected from 28 subjects' 14 scalp locations during 4 different gameplays will be adopted in this study.

EEG feature extraction

Fast Fourier transform (FFT) and Welch's power spectral density (PSD) are adopted in this study to obtain the EEG features from frequency domain. The EEG features extracted using the FFT technique is in form of complex-value spectrum indicating the amplitudes and phases of each decomposed game-induced EEG signals. The EEG signal's characteristics that were described in terms of amplitudes and phases could help in identifying and explaining the appearance of the band powers that associate with the specific human emotion state among the certain segmented signals data. PSD technique is a signal segmentation method that is used to obtain the power distribution of each decomposed EEG signal. The features product of the Welch's PSD estimation could summarize the spectral content of the EEG signal in terms of power content that helps in analyzing the game-induced EEG's frequency properties. The sampling rate of 128 Hz was defined when computing PSD feature extraction.

Annotation of experienced game type and elicited emotion state

The datasets loaded into the Python environment under Jupyter Notebook IDE will first to be manually categorized with the elicited valence-arousal emotion state and the experienced game genre when eliciting the emotion state, respectively, based on the provided SAM ratings. Then, the series of if-else statements are deployed to assign the elicited emotion state along with the experienced game genre to each loaded EEG instance. Table 1 shows the list of attributes after annotation.

EEG feature selection

The maximum relevance–minimum redundancy (mRMR) algorithm is adopted to minimize the input

Table 1 List of attributes

Attributes	Type	Value
Elicited Valence-arousal emotion	Character	HAPV, HANV, LAPV, LANV
Video game's type	Character	HAPV, HANV, LAPV, LANV
AF3	Numeric	> -47 Hz
F7	Numeric	> -235 Hz
F3	Numeric	> -105 Hz
FC5	Numeric	> -84 Hz
T7	Numeric	> -71 Hz
O1	Numeric	> -105 Hz
O2	Numeric	> -88 Hz
P8	Numeric	> -67 Hz
T8	Numeric	> -58 Hz
FC6	Numeric	> -76 Hz
F4	Numeric	> -95 Hz
F8	Numeric	> -87 Hz
AF4	Numeric	> -72 Hz

dimensionality and computational cost of the human emotion state classifier meanwhile avoiding the overfitting classification result that was produced by the classifiers. The mRMR selector maximizes the correlation between the extracted EEG features and the target elicited human emotion labels meanwhile minimizing the correlation between each extracted EEG features. With this nature, the mRMR could select the top 50 most discriminative features that represent each subject’s EEG dataset with the least redundancy among each feature that engage in determining the human emotion.

Classification models construction

Data mining is a process to identify the underlying knowledge within the data, where the machine learning algorithms are applied as an approach to mine the meaningful insights lied within the data for resolving prediction, classification, and the other problems. In this study, two ruleset-based algorithms of repeated incremental pruning to produce error reduction (RIPPER) and regularized class association rules (RCAR) and a support vector machine (SVM) algorithm are applied for learning the EEG features in eliciting each emotion state from the prepared dataset.

RIPPER implemented a separate-and-conquer strategy to separate the 2/3 of the training dataset into a growing set while the remaining become the pruning set. It will iterate through the growing set to generate a series of rules that identify the correlation between each variable in determining the outcome based on the greedy heuristic methodology. Each generated rule will be repeatedly simplified to comfort the pruning set by removing a sequence of condition within it or whole rule until the latest simplified rule has found decreasing accuracy in determining the outcome with regards to the pruning set by comparing with the empty ruleset’s accuracy. To eliminate the generation of new rules, Cohen implemented an optimization process to construct two more alternative rules named of ‘Replacement’ and ‘Revision’ for each rule generated for selecting the best rule by obeying the minimum description length principle (MDL), which does not consider the rule that has a total description length that exceeded the defined maximum description length.

RCAR finds a collection of rules that meet the defined minimum support and confidence values for performing the binary and multiclass classifications. A priori algorithm was implemented to craft all possible combinations of rules with consideration of the predefined support and confidence thresholds. It adopted the nature of Lasso regularization as a core procedure that shrinks the coefficient of the logistic model’s parameter toward 0 for pruning the large dataset. The rules that have the retained

nonzero coefficient value after undergoing the Lasso regularization are used for constructing the classifier.

SVM finds the most optimal hyperplane that will maximize the boundary between both classes for the supervised classification task. The optimal hyperplane is formed by the maximum distance from each of the support vectors or critical points of different classes which are provided by the dataset. It can deal with linearly inseparable binary or multiclass by applying the most appropriate kernel function. In this study, the radial basis kernel is adopted for dealing with the high dimensionality of the prepared EEG data.

Classification models construction

The prediction performance of the constructed models based on the prepared dataset is assessed by the typical evaluation metrics, namely the accuracy, kappa value, precision, recall and F1-score. These evaluation metrics excluding the kapa value involve the true positive (TP), true negative (TN), false positive (FP) and false negative (FN) for calculation. Table 2 presents the formula of accuracy, precision, recall and F1-score.

Association rule mining analysis

The experiments conducted in this paper were divided into 2 different parts which are subject-based and cross-subject-based experiments. Each of the 84 subject-based experiments was conducted with a respective subject’s EEG dataset as an input for evaluating the robustness, reliability, and performance of the RCAR, RIPPER, and SVM classifiers by examining their kappa values and the stated classification performance metrics, while each of 3 cross-subject-based experiments was conducted with a unified dataset comprised of all subject’s EEG data as an input but using a different classification algorithm. To provide the analysis results about the patterns in eliciting each emotion state considering all involved subject’s data, the classification result of the outperforming classifier from the cross-subject-based experiment is considered. Thus, the top 10 association rule sets that were constructed by the ruleset-based algorithm, which has the best performance evaluated, that contributed to

Table 2 Formula of accuracy, precision, recall and F1-score

Performance metrics	Formula
Accuracy	$(TP + TN) / (TP + TN + FP + FN)$
Precision	$TP / (TP + FP)$
Recall	$TP / (TP + FN)$
F1-score	$(2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$

classifying the elicited human emotion will be further elaborated along with the explanation from the biological and physiological perspectives. The elaboration on the classified emotions could help the researchers understand the consideration of the algorithm to generate the classification results via the constructed association rule sets. Therefore, the 10 rule sets with the highest scores measured in terms of support, confidence, and coverage rate were used for the trend analysis in eliciting each human emotion state. The following section was constructed to elucidate the classification performance of each classification algorithm that was involved in both subject-based and cross-subject-based experiments as well as the analysis of the rulesets generated by the classifier that has the best performance evaluated in the cross-subject-based experiment.

Results

There are 84 experiment cases conducted for subject-based experiments. In each experiment, there were 28 subject's EEG data obtained from all 4 video games involved and each dataset was applied with the RIPPER, RCAR and SVM algorithms, respectively. Three more experiment cases were conducted for cross-subject-based experiments. All the 28 subjects' EEG data were combined into a whole EEG dataset and the combined dataset is applied in each cross-subject-based experiment case with the RIPPER, RCAR and SVM algorithms individually. The prediction modeling using three machine learning algorithms are implemented in RStudio IDE under R programming language. Each loaded dataset for both subject-based and cross-subject-based experiments is randomly separated into the training and testing datasets for models' training and performances' evaluation purposes, respectively, by following the ratio 70:30 ratio.

The experiments' procedure proceeds to construct three classifiers by utilizing the three algorithms individually based on the training dataset for learning the main EEG features in determining each elicited emotion state. The constructed classifier from each experiment case is evaluated based on its prediction performance on the testing dataset. For identifying the classifiers' performance when delivering the reliable prediction result, all

the subject-based experiment cases that have the classifier evaluated with over the 0.62 kappa value are averaged and recorded for further performance analysis. After evaluating each classifier, the top 10 association rules are obtained from the highest performance ruleset-based classifier for elucidating the relationship between the EEG frequency-based features and the experienced video gameplay to each elicited emotion state. The top 10 association rules are obtained based on the support measures in ascending order.

Discussion

Tables 3 and 4 show the averaged overall accuracy, kappa value, confidence interval's gap, precision, recall and F1-score of each classifier in subject-based experiments, respectively. Tables 5 and 6 show the overall accuracy, kappa value, confidence interval's gap, precision, recall and F1-score of each classifier in cross-subject-based experiments, respectively. Regarding Tables 3 and 5, the RCAR classifier is found that it achieved the highest performance across all the cross-subject and subject-based experiment cases in terms of averaged overall accuracy and averaged F1-score for HAPV and HANV states

Table 3 averaged overall accuracy, kappa value and confidence interval's gap of each classifier in subject-based experiments

Algorithm	Averaged overall accuracy (%)	Averaged kappa value (%)	Averaged confidence interval's gap (95%)
RIPPER	79.51	77.18	(0.9358, 0.9451)
RCAR	87.89	78.27	(0.9423, 0.9512)
SVM	87.29	79.44	(0.9402, 0.9493)

Table 4 averaged precision, recall and F1-score of each classifier in subject-based experiments

Algorithm	Elicited emotion state	Averaged precision (%)	Averaged recall (%)	Averaged F1-score (%)
RIPPER	HANV	84.81	79.97	76.14
	HAPV	85.01	85.89	79.24
	LANV	86.27	88.10	81.05
	LAPV	61.54	60.87	61.20
RCAR	HANV	86.97	79.61	85.44
	HAPV	84.01	89.27	87.07
	LANV	89.97	85.79	88.96
	LAPV	61.29	65.33	65.33
SVM	HANV	84.56	81.10	82.73
	HAPV	86.79	88.65	79.45
	LANV	90.59	89.51	89.95
	LAPV	76.68	80.43	78.51

Table 5 Overall accuracy, kappa value and confidence interval's gap of each classifier in cross-subject-based experiments

Algorithm	Overall accuracy (%)	Kappa value (%)	Confidence interval's gap (95%)
RIPPER	94.06	0.9083	0.93
RCAR	94.69	0.9180	0.89
SVM	94.49	0.9148	0.91

Table 6 Precision, recall and F1-score of each classifier in cross-subject-based experiments

Algorithm	Elicited emotion state	Precision (%)	Recall (%)	F1-score (%)
RIPPER	HANV	97.85	94.95	96.38
	HAPV	92.22	99.05	95.51
	LANV	95.94	91.56	93.70
	LAPV	52.07	24.23	33.07
RCAR	HANV	97.99	95.75	96.86
	HAPV	92.56	99.32	95.82
	LANV	96.57	92.33	94.40
	LAPV	70.19	28.08	40.11
SVM	HANV	98.71	94.99	96.81
	HAPV	92.54	99.25	95.78
	LANV	94.93	92.55	93.73
	LAPV	75.00	26.54	39.20

classification. Although the RIPPER achieved the slightly worst performance comparing to other 2 classifiers, it shows that it could provide the most stable prediction result with the lowest confidence interval's difference of 2.53% at its peak performance among each ruleset-based classifier's performance. Referring to Table 3, the SVM classifier could provide a more reliable emotion state prediction result with the highest kappa values achieved of 79.44% and a competitive averaged accuracy of 87.29% comparing to the RCAR's performance. Simultaneously, the SVM classifier achieved the most robustness and stable performance in terms of delivering the reliable prediction result in the greatest number of 20 subject-based experiment cases based on the Kappa value metric evaluated over 0.62 and the lowest confidence interval's difference achieved of 1.54%. Overall, the SVM classifier developed in this study demonstrated its strength in performing each subject-based experiment case with consistent and stable prediction result provided at most of the time. The findings support that the SVM's margin maximization strategy that mainly forms the optimal hyperplane in the data of high dimensionality could be well suited to effectively learn and distinguish the multiple properties of data as many as possible that belong to the same emotion state from the medium size of dataset around 10,000 EEG instances.

As depicted in Tables 5 and 6, all classifiers that deployed for each cross-subject-based experiment case are found to have a drastic increasing performance in predicting the emotion state for all subject's EEG instances that were never learnt by the classifiers before, comparing to their performances depicted in Tables 3 and 5. By comparing each classifiers' performance from

both subject-based and cross-subject-based experiments, RIPPER classifier is found to has its accuracy increased the most with 14.55% when comes to predict the emotion state for all subject's EEG data. The RCAR and SVM classifiers developed in the cross-subject-based experiments are found 6.8% and 7.2% greater than in the subject-based experiments, respectively. Simultaneously, each classifier's prediction result delivered in the cross-subject-based experiments are found more reliable and stable than the prediction results in each subject-based experiment where the classifier achieved Kappa values of over 90% and confidence interval's gap lower than 1%. However, each classifier's performance has greatly affected when comes to predicting the LAPV state to the large amount of EEG instances. Specifically, the RIPPER's, RCAR's and SVM's performance in terms of F1 score evaluated has dropped with 32.5%, 20.23% and 38.17% when intend to predict the LAPV state. It is believed that the small amount of data belongs to the LAPV state provided after combined all subjects' EEG data are insufficient to the classifiers in learning the key properties of EEG features and discriminating the difference of EEG properties for LAPV state and other state. Referring to Tables 3 and 4, the RCAR classifier achieved the highest performance among all the classifiers in all aspects of the performance evaluation stated in these two tables. This finding supports that the implementation of Lasso regularization within the RCAR algorithm for pruning the meaningless ruleset could greatly enhance the algorithm's decision-making mechanism in predicting the emotion state for the large amount of dataset that contains over 33,000 instances.

Discussion on classifiers' performance

Among each ruleset-based classifier, the RCAR is found to achieve the highest performance as stated in previous section. Thus, the association rules produced by the RCAR-based classifier implemented in the cross-subject-based experiments are used for further investigating the kind of relationship that existed in between the EEG features that elicited the certain human emotion state. There are three performance metrics of support, confidence and coverage adopted to examine the feasibility of the generated association rules. The support measurement of the association rules could assist the author in understanding how frequent the generated ruleset occurs within the dataset. The confidence measurement was involved to determine the reliability of the generated rulesets in terms of the conditional probability while the coverage measurement was applied to evaluate the proportion of transactions in the dataset that are applicable to a given rule. For each elicited emotion state, the 10 rules that achieved the highest measure of support are

selected. Tables 7, 8 and 9 show the top 10 association rules generated by the RCAR classifier for elicited HAPV, HANV and LANV state as their consequent, respectively.

Regarding Table 7, the findings suggest that there are 7 EEG features incorporated with the genre of experienced video games that greatly impact the classifier’s decision in

predicting the HAPV states. The 5 EEG features involved in the generation of the association rules were generated by the FFT technique and the remaining 2 features were generated by the PSD technique. However, all those rules achieved a below-moderate measure of support that is around 25–45%, which indicates those rules as

Table 7 Association rules for elicited HAPV state as consequent

Antecedent	Support (%)	Confidence (%)	Coverage (%)
game_play_type_HANV=[-Inf,0.5), game_play_type_LANV=[-Inf,0.5)	45.59	91.07	50.06
game_play_type_HANV=[-Inf,0.5),	44.63	61.08	73.07
DFT...output...complex...10_0_magnitude=[-Inf,3.33e+03)	38.57	64.06	60.20
game_play_type_HANV=[-Inf,0.5),	36.34	94.84	38.32
DFT...output...complex...9_98_magnitude=[-Inf,95.4)	35.52	94.46	37.61
game_play_type_HANV=[-Inf,0.5), PSD...nperseg...128_11_17=[-Inf,0.00549), game_play_type_LANV=[-Inf,0.5)	31.04	95.17	32.61
game_play_type_HANV=[-Inf,0.5), game_play_type_LANV=[-Inf,0.5),	27.86	64.06	43.50
DFT...output...complex...11_18_magnitude=[-Inf,117)	26.12	64.49	40.51
game_play_type_HANV=[-Inf,0.5), game_play_type_LANV=[-Inf,0.5),	26.12	96.06	27.19
PSD...nperseg...128_11_18=[-Inf,0.00356)	25.75	92.14	27.94

Table 8 Association rules for elicited HANV state as consequent

Antecedent	Support (%)	Confidence (%)	Coverage (%)
game_play_type_HANV=[0.5, Inf]	24.19	96.25	25.13
game_play_type_HANV=[0.5, Inf], game_play_type_HAPV=[-Inf,0.5)	24.19	96.25	25.13
game_play_type_HANV=[0.5, Inf], game_play_type_LAPV=[-Inf,0.5),	24.19	96.25	25.13
game_play_type_HAPV=[-Inf,0.5)	24.19	96.25	25.13
game_play_type_HANV=[0.5, Inf], game_play_type_LANV=[-Inf,0.5),	24.19	96.25	37.23
game_play_type_HAPV=[-Inf,0.5)	18.94	50.89	19.03
game_play_type_LANV=[-Inf,0.5), game_play_type_LAPV=[-Inf,0.5),	18.74	98.49	19.03
game_play_type_HAPV=[-Inf,0.5)	18.74	98.49	19.03
game_play_type_LANV=[-Inf,0.5), game_play_type_LAPV=[-Inf,0.5),	18.74	98.49	19.03
DFT...output...complex...11_18_magnitude=[-Inf,117)	18.74	98.49	19.03

Table 9 Association rules for elicited LANV state as consequent

Antecedent	Support (%)	Confidence (%)	Coverage (%)
game_play_type_LAPV=[-Inf,0.5), game_play_type_HAPV=[-Inf,0.5),	24.37	50.11	48.64
DFT...output...complex...10_0_magnitude=[-Inf,3.33e+03)	23.43	97.59	24.01
game_play_type_LANV=[0.5, Inf],	14.50	51.65	28.07
DFT...output...complex...10_0_magnitude=[-Inf,3.33e+03)	13.84	99.45	13.92
game_play_type_HANV=[-Inf,0.5), game_play_type_LAPV=[-Inf,0.5),	13.84	99.45	13.92
game_play_type_LANV=[0.5, Inf], game_play_type_HAPV=[-Inf,0.5),	13.84	99.45	13.92
DFT...output...complex...13_0_magnitude=[-Inf,191)	13.84	99.45	13.92
game_play_type_LANV=[0.5, Inf], game_play_type_LAPV=[-Inf,0.5),	13.78	99.47	13.86
DFT...output...complex...13_0_magnitude=[-Inf,191)	12.04	98.63	12.21
game_play_type_LANV=[0.5, Inf],	11.33	50.23	22.56

the identified patterns that are not commonly found in the EEG dataset. The selected rules were evaluated with a confidence score of around 69–96% indicating that the selected rules achieved a beyond-moderate reliability of around 69–96% in ensuring the HAPV state is always the main result when applying those rules on the matched instances. The coverage rate measured within 27–73% indicated that most of the rules including rules 4, 5, 6, 7, 8, 9, and 10 are not highly applicable in the EEG dataset while rules 1, 2 and 3 are found highly applicable in the EEG dataset. As indicated in rule 2 and rule 3, the classifier suggests that the computed FFT coefficients on the ninth and tenth subjects' EEG data at the first second and 98 s incorporating with their experienced game genre of HANV will have the chance of around 60–70% in determining the elicited human state fall in the HAPV category.

Table 8 depicts that the 2 types of EEG features computed using FFT and PSD techniques, respectively, have greatly influenced the RCAR classifier in generating the association rulesets for classifying the HANV state. The overall finding also depicted that all the 10 association rules are considered as rarely to be supported by the existing dataset with their support metric measured below 25%. However, once the EEG instances within the dataset have found that they could be suited to these rules, the results of these rules have achieved around 64–96% confidence in assigning the HANV state as the most likely human emotion state to that instance. This finding indicates that these obtained rules could have a high probability of considering an EEG instance belonging to the HANV state once the EEG instance has found could be fitted to apply those rules. Besides, rules 1–6 are found highly applicable on the unseen EEG dataset for determining the matched EEG instances as an induction to elicit the HANV state that is indicated with the coverage rates of between 25.13 and 37.23%, respectively. Among the 6 rules, rule 6 is found to explain that the EEG features computed using the FFT technique for the eleventh subject at 18 s incorporated with the experienced games' labels of LAPV and LANV states will have a rate of 37.23% in determining that a matched EEG instance will tend to have the HANV state as an elicited emotion state.

Table 9 demonstrates that there are 2 EEG features computed using FFT technique and 1 EEG features computed using PSD technique contributing to the RCAR classifier in explaining the relationship between the EEG features and experienced game's type in eliciting the LANV state. The finding shows that all the 10 rules are rarely could be found within the EEG dataset with the support metric of around 11.33–24.37%. Despite of the low support computed for each obtained rules, most of

the rules including rules 2, 4, 5, 6, 7, 8, and 9 achieved a great reliability with over 90% of confidence metric computed to ensure that the EEG instances that match to these rules will elicit the LANV state. Among these 10 rules, the rule 1 is found as the most applicable rule to the given EEG dataset as it was supported by the most of EEG instances comparing to other rules for resulting in LANV state elicitation, which is indicated by the coverage rate of 48.64%. Based on the discussed ruleset's evaluation metrics, rule 1 will be adopted for determining the relationship between the existing patterns in eliciting the LANV state. Rule 1 has suggested that the FFT coefficient computed on the EEG data of the tenth subject at 0 s incorporated with the experienced gameplay's genres of HAPV and LAPV will have 48.64% of chance to determine that a matched EEG instance will tend to have the LANV state as an elicited emotion state.

Conclusions

This research study has proposed the interpretable ruleset-based classifiers able to explain the consequences of the set of game-induced EEG features in eliciting the specific human emotion state. The classifier is proposed to fulfil the existing knowledge gap among the emotion recognition studies based on game-induced EEG data. The knowledge gap is identified from most of the relevant studies have not conducted an in-depth investigation on the patterns behind from the set of video game-induced EEG features that result in elicitation of certain human emotion state. The evaluation result of RCAR classifier constructed for the cross-subject experiment showed the highest performance of accuracy, precision and recall, F1-score that over 90% indicated that the RCAR algorithm could accurately identifying the valence-arousal human emotion states of HAPV, HANV and LANV from all the subject's EEG signals based on the generated condition rulesets. Among the ruleset-based classifier, 12 selected cases of the subject-based emotion recognition experiments indicated that a considerable robustness of RCAR-based classifier's in accurately identifying the actual emotion state elicited for each unseen EEG instance with the accuracy over 78% and kappa value greater than 0.62. As overall, the SVM-based classifier achieved the highest performance when performing the subject-based emotion recognition experiments as it delivered a reliable and accurate prediction result to most of the experiment cases involving the 20 subject's EEG data, respectively. The findings obtained from this study indicated the ruleset-based classifiers required a huge training time during the classifier's construction compared to the SVM-based classifier. Nonetheless, the ruleset-based classifiers were examined to require a small amount of time approximated to the execution

time required by SVM-based classifier to predict the emotion states on all subject's EEG data simultaneously and concurrently. Meanwhile, the advantages of utilizing the ruleset-based classifier are that the classifier could provide the researchers with a fine-grain level of understanding about the kind of EEG's features and the kind of relationships exist within each EEG feature that greatly associate with the elicited emotion state. In real life application, the association rules generated by the rule-based classifiers that frequently produce a reliable emotion state prediction result while ensuring a high accuracy maintained could help the medical professionals to further understand and allow them to have an early preparation to encounter the predicted patient's emotion changes.

Abbreviations

DWT	Discrete wavelet transform
EEG	Electroencephalography
FFT	Fast Fourier transform
HANV	High-arousal negative valence
HAPV	High-arousal positive valence
HCI	Human-computer interface
LANV	Low-arousal negative valence
LAPV	Low-arousal positive valence
PSD	Power spectral density
RCAR	Regularized class association rules
SVM	Support vector machine

Acknowledgements

The authors wish to thank the developers of GAMEEMO for making the dataset available for public access.

Author contributions

All authors have read and approved the manuscript. MXL was involved in writing—original draft. JT contributed to writing—review and editing and supervision.

Funding

No funding was received for this study.

Availability of data and materials

The datasets analyzed for this study can be found in the [KAGGLE] (<https://www.kaggle.com/datasets/sigfest/database-for-emotion-recognition-system-gameemo?resource=download>).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author details

¹Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia. ²Creative Advanced Machine Intelligence Research Centre, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia. ³Evolutionary Computing Laboratory, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia.

Received: 1 February 2024 Accepted: 16 April 2024

Published online: 24 May 2024

References

- Abdulrahman A, Baykara M (2021) Feature extraction approach based on statistical methods and wavelet packet decomposition for emotion recognition using EEG signals. In: 2021 International conference on innovations in intelligent Systems and applications (INISTA). <https://doi.org/10.1109/inista52262.2021.9548406>
- Abdulrahman A, Baykara M, Alakus TB (2022) A novel approach for emotion recognition based on EEG signal using deep learning. *Appl Sci* 12(19):10028. <https://doi.org/10.3390/app121910028>
- Abu M, Shin J, Islam MM, Abdullah A, Islam K (2022) Natural human emotion recognition based on various mixed reality (MR) games and electroencephalography (EEG) signals. <https://doi.org/10.1109/ecce53102.2022.9829482>
- Alakus TB, Gonen M, Turkoglu I (2020) Database for an emotion recognition system based on EEG signals and various computer games—GAMEEMO. *Biomed Signal Process Control* 60:101951. <https://doi.org/10.1016/j.bspc.2020.101951>
- Alarcoo SM, Fonseca MJ (2019) Emotions recognition using EEG signals: a survey. *IEEE Trans Affect Comput* 10(3):374–393. <https://doi.org/10.1109/taffc.2017.2714671>
- Bhatti AM, Majid M, Anwar SM, Khan B (2016) Human emotion recognition and analysis in response to audio music using brain signals. *Comput Hum Behav* 65:267–275. <https://doi.org/10.1016/j.chb.2016.08.029>
- Dhuheir M, Albaseer A, Baccour E, Erbad A, Abdallah M, Hamdi M (2021) Emotion recognition for healthcare surveillance systems using neural networks: a survey. *IEEE Xplore*. <https://doi.org/10.1109/IWCMC51323.2021.9498861>
- Dogan A, Akay M, Barua PD, Baygin M, Dogan S, Tuncer T, Dogru AH, Acharya UR (2021) PrimePatNet87: prime pattern and tunable q-factor wavelet transform techniques for automated accurate EEG emotion recognition. *Comput Biol Med* 138:104867–104867. <https://doi.org/10.1016/j.compbiomed.2021.104867>
- Gupta V, Chopda MD, Pachori RB (2019) Cross-subject emotion recognition using flexible analytic wavelet transform from EEG signals. *IEEE Sens J* 19(6):2266–2274. <https://doi.org/10.1109/jsen.2018.2883497>
- Hazarika J, Kant P, Dasgupta R, Laskar SH (2018) Neural modulation in action video game players during inhibitory control function: an EEG study using discrete wavelet transform. *Biomed Signal Process Control* 45:144–150. <https://doi.org/10.1016/j.bspc.2018.05.023>
- Henia B, Lachiri Z (2017) Emotion classification in arousal-valence dimension using discrete affective keywords tagging. <https://doi.org/10.1109/icemis.2017.8272991>
- Khan A, Rasool S (2022) Game induced emotion analysis using electroencephalography. *Comput Biol Med* 145:105441. <https://doi.org/10.1016/j.compbiomed.2022.105441>
- Li M, Xu H, Liu X, Lu S (2018) Emotion recognition from multichannel EEG signals using K-nearest neighbor classification. *Technol Health Care* 26:509–519. <https://doi.org/10.3233/thc-174836>
- Mikuckas A, Mikuckiene I, Venckauskas A, Kazanavicius E, Lukas R, Plauska I (2014) Emotion recognition in human computer interaction systems. *Elektron Elektrotech*. <https://doi.org/10.5755/j01.eee.20.10.8878>
- Saxena S, Tripathi S, Sudarshan TSB (2022) An intelligent facial expression recognition system with emotion intensity classification. *Cogn Syst Res* 74:39–52. <https://doi.org/10.1016/j.cogsys.2022.04.001>
- Soleymani M, Asghari-Esfeden S, Fu Y, Pantic M (2016) Analysis of EEG signals and facial expressions for continuous emotion detection. *IEEE Trans Affect Comput* 7(1):17–28. <https://doi.org/10.1109/taffc.2015.2436926>
- Tuncer T, Dogan S, Baygin M, Rajendra Acharya U (2022) Tetromino pattern based accurate EEG emotion classification model. *Artif Intell Med* 123:102210. <https://doi.org/10.1016/j.artmed.2021.102210>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.