Assessment of Video Game Player Proficiency Through the Brain Activity Analysis Using EEG Signals

Context and Motivation

This research project is dedicated to the study of human brain activity in the context of video games. As a popular digital form of entertainment activity, video gaming has transformed into the rapidly growing electronic sport industry involving single players or teams competing in various tournaments.

According to the statistics for the current year the global gaming market size is nearly 3 hundred billion dollars ($300 billion). Almost half of the world's population plays games, and more and more casual gamers strive to make a career as eSports athletes.

The growing interest in competitive gaming rises to a number of challenges for players on how to progress from the amateur gaming level into the professional one and how one can characterize a professional eSports athlete and his/her skills.

More specifically, this project is focused on the analysis of the Counter-Strike game, as the most popular in its genre, whose number of active players is almost 2 million at the same time.

Video games have an inherently digital nature and provide easy access to a large amount of structured and unstructured data from various sources, which stimulates scientific interest.

To date, the vast majority of studies to characterize and assess the players are based on the study of gaming skills and performance metrics based on in-game data, as well as physical skills and physiological patterns based on data from peripheral sensors. Only a few works go deeper and consider the underlying brain activities that contribute to overall player behavior. The main research methods are the use of fMRI and EEG, but only the latter allows to track brain dynamics exactly during the game process.

Literature review

At the moment, there are only a few scientific papers using EEG to analyze the FPS of players. A brief overview of existing works is represented in Table 1.

Articles	Game (Game genre)	Object of study	EEG features used	Experimental protocol (N players, conditions recorded, tracking time)
Minchev et al. [1]	Project I.G.I (FPS)	reaction to game events (loss, win)	spectral	10 players, during game session
B. Meneses-Claudio & A. Roman-Gonzale [2]	DOTA 1 (MOBA)	player's proficiency	frequency	2 players, during game session
Anwar et al. [3]	TempleRunis (endless runner)	player's proficiency	temporal	20 players, during game session
Melentev et al. [4]	CS:GO (FPS)	player's proficiency and tiredness prediction	time-frequency	20 players, before and after game session
Smerdov et al. [5]	LoL (MOBA)	player's proficiency	time-frequency	10 players, during game session
Gostilovich et al. [6]	CS:GO (FPS)	player's winning prediction	temporal, time-frequency, spectral	24 players, visual search task, specific events

In [1] authors demonstrated that negative game events decrease EEG alpha rhythms power spectra frequencies, while positive game events increase it. The findings also showed the opposite relation to the theta rhythm. Work by B. Meneses-Claudio and A. Roman-Gonzale [2] identified that a professional player has fewer variations in the EEG signal than an amateur player, although only two participants were compared. In another research [3], authors trained ML models to predict player proficiency using features extracted from EEG band frequencies, and achieved 92% accuracy. It was also reported that Alpha and Beta rhythms for casual players change more than for the professionals. The opposite situation was for the Theta rhythm, as was found previously in [1]. Data were collected only before and after the experiment, which could potentially cause the loss of valuable observations. The other two studies [4, 5], although collecting EEG data, do not provide any relevant specifics or interpretation in the description of experiments and results. Moreover, in [4] it is mentioned that EEG data was not used to predict player skill at all.

Overall, it can be concluded that all mentioned works on gamers EEG analysis have their own peculiarities and drawbacks related to i) consideration of the EEG signal only during key game events or only some time before and after game sessions, which can cause the loss of relevant information during the game process; ii) the experimental protocol organization, when players are tested outside the real game context and only on the visual part of the reaction; iii) focus on different game genres. There are 2 most popular game genres, namely Multiplayer Online Battle Arena (MOBA) and tactical First-Person Shooter (FPS), which differ greatly in gameplay and respectively in the expressed brain activity. Therefore, the derived distinctive brain activity patterns of MOBA e-sportsmen may not be consistent with ones of FPS players. iv) repetitive methods for EEG signal processing.

Aim and Objectives

The aim of this project to investigate differences in brain activity between players of different skill levels and to assess the impact of high game proficiency on brain activity patterns.

The objectives run as follows:

1. Build a dataset consisting of cleaned, filtered, and class-separated EEG signals.
2. Determine the distinctive patterns of brain activity for casual and professional players.
3. Develop a reliable baseline for player skill prediction.

Methodology

Data description

The data used in this project was collected from 17 players with different gaming experience, including 4 professionals, during gaming sessions lasting 30-50 minutes.

EEG data were recorded using the wireless EEG headset Emotiv Epoc+ from 14 usable saline electrodes according to the 10–20 system (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) and 2 references on parietal sites (P3 and P4). For some reason, instead of O1, the Pz channel was recorded. The data stream from the headset represents automatically pre-filtered bands' power in time for 5 different frequency bands:

- theta (4-8Hz)
- alpha (8-12Hz)
- betaL (low beta, 12-16Hz)
- betaH (high beta, 16-25Hz)
- gamma (25-45Hz)

The power values are absolute, the unit is uV^2 / Hz.

More specifically, the band power is calculated every ~0.2 seconds, using the last 2 seconds of EEG data. The steps are:

1. High pass filter [5th order Sinc filter]
2. Remove DC offset (for each EEG channel, remove the mean value) [digital notch filters at 50Hz and 60Hz]
3. Hanning window
4. FFT
5. make the sum of the squares of the FFT values

Detailed technical documentation can be found at:
https://www.emotiv.com/products/epoc?srsltid=AfmBOopwvKfkjwR8bie8yF0JP3vozYoYsGhl8i8qXhS6ZkZ9kbPJqtsJ

Data preprocessing

Since the data undergoes built-in filtering and is already presented as frequency bands power, additional filtering is omitted.

Preprocessing consisted of:

1) resampling to a unified sampling rate. At the output it is intended to get short band power intervals of the same size for all subjects. But the sampling rate varies both within and between subjects' data. The average minimum time-points period is 136 milliseconds, the maximum is 157 ms. For resampling, the average sampling rate for all subjects was calculated and used, equal to 6,71 Hz that corresponds to 149 ms between recorded values.
2) removing outliers with help of Inter Quartile Range (IQR) method. It was found that the EEG records contained strong outliers (which appeared due to the subject's movement, changing and shifting of the EEG headset, changes in physiological conditions, etc.) and that there were noticeable fluctuations at the beginning and at the end of the recordings. To deal with this, the following was done: i) cutting off 2 minutes from the beginning and from the end of recording (2 minutes was selected by manual investigation), then ii) apply the interquartile range (IQR) based method to detect the remaining outliers, iii) and replace them with interpolated values.
3) windowing (splitting to short time intervals). The duration of the EEG records on average was 32 minutes. Since the game under consideration is very dynamic and involves a quick change in game conditions and the actions required from the player, it was not practical to consider whole single record interval. One could consider the event-related intervals, but the events in the game are replaced too quickly and often overlap with each other. And the extraction of individual specific EEG record which displayed the brain activity in response to this particular game event is very difficult. Therefore, the division of the entire time interval into short intervals was undertaken, as was done in existing research. In particular, 1-5 second intervals are most often selected. (1 second - [5,7,8], 2 seconds - [1,4], 3 seconds - [9], 4 seconds - [10], 5 seconds-[11]). As a result, the data were splitted by 2-second intervals without overlap.

Feature extraction

In this work, 2 subgroups of features are extracted - basic time-domain features and autonomic EEG indices (taken from studies on the rhythm analysis in relation to human behavior and preferences [10,12])

Time-domain features were extracted per channel and include Minimum, Maximum, Mean, Std, 1st Quartile, 2nd Quartile (Median), 3rd Quartile, Hjorth parameters (complexity, mobility), Skewness, Kurtosis, Root mean square, Decorrelation time, Peak-to-peak (PTP) amplitude, Katz Fractal Dimension, Area under the curve, Variance, Entropy.

Autonomic EEG indices tend to measure asymmetry in brain activity in specific frequency bands and express to human preferences and behavior. The following are used in this work:

- Approach-Withdrawal (AW) Index - measures the frontal alpha asymmetry and serves for the evaluation of engagement/disengagement towards the stimuli (in this case - game)
- Effort Index - reflects the activity level of the frontal theta in the prefrontal cortex and serves as the indication of cognitive processing. Higher theta activity associates with higher levels of task difficulty and complexity in the frontal area.
- Valence Index - measures the asymmetrical activation of the frontal hemisphere that correlates to player preferences.
- Choice Index - measures the frontal irregular fluctuations in beta and gamma, associated with the actual stage of decision-making.

All of the above are based on frontal area channels. Similar indices were measured for the temporal cortex.

Data analysis

Data analysis involved the following steps:

First, various ml models, namely (LR, SVM, RF, KNN, MLP) were trained on time-domain features, involving hyperparameter tuning with 5-fold cross validation. Tested and parameters are listed in Table 2.

Table 2. List of tested and best hyperparameters for the specified models.

Model	Tested parameter grid	Best parameters
Logistic Regression	solver: liblinear penalty: [l1, l2], (for LR try both L1 and L2 regularizations) C: [0.01, 1, 10], tol: 1e-6	C: 10, penalty: l1, solver: liblinear, tol: 1e-06
Random Forest Classifier	max_features: [sqrt, log2] ccp_alpha: [0.01, 0.001] max_depth: [3,4,5] criterion: [gini, entropy]	max_features: sqrt ccp_alpha: 0.001 max_depth: 5 criterion: gini
SVC	C: np.logspace(-3, -1, 10) kernel: [linear, rbf]	C: 0.1 kernel: linear
K Neighbors Classifier	n_neighbors: [3, 5, 7, 10, 15, 20, 30, 35]	n_neighbors: 15
MLP Classifier	max_iter: [150, 250] solver:[adam,lbfgs] alpha: [1e-3,1e-4] hidden_layer_sizes: [(200,),(300,)]	max_iter: 150 solver: adam alpha: 1e-4 hidden_layer_sizes: (200,)

The weighted F1-score was selected for model evaluation. The choice is explained by the fact that the dataset is quite unbalanced with ratio of majority to minority class is 2.707. And since it is desired to capture all possible distinctive patterns of pro players brain activity, but at the same time not falsely identify a beginner as a pro player, a harmonic mean of the precision and recall, which takes into account label Imbalance was used. Model scores on train and validation sets were also considered. The evaluation procedure involved the use of 5-fold cross validation.
Further, the selected best model at the previous step was trained on a mixed feature set when autonomic features were added to time-domain ones, and the validity of their use was evaluated.
In addition, methods were tested to improve the model performance, namely dimensionality reduction (PCA) and feature selection (univariate filter method SelectKbest), as well as the upsampling (SMOTE, ASDAYN) for balancing the class distribution.

Results and Discussion

Insight from data and comparison with existing works

Before model training, data exploration was conducted on how do the rhythm-channel values differ between subjects with different levels of game experience.

The hypothesis of distribution normality was tested and rejected; therefore, the nonparametric Mann Whitney test was used to calculate the statistical significance of differences.

The following findings were obtained with statistically significant differences (p-value <0.001):

- In more detail, it can be noted that in channel O2 for all rhythms the values are significantly higher for professionals, compared to casual players. This can be explained by the fact that O2 is located in the occipital lobe and is responsible for processing visual information and transmitting processed data to other brain areas. And for professionals, this area works better during the game due to their extensive experience.
- The theta rhythm in most channels is higher for professionals, which corresponds to the results of works [4,13] and is explained by the fact that professionals are more focused on the upcoming tasks.
- It can also be noted that for the area of the temporal cortex (T7, P7, T8, P8), the values of almost all rhythms (except theta) are higher for casual players. Considering that the temporal cortex is responsible for processing auditory signals, it can be assumed that pro players are more focused (theta rhythm) and less distracted by extraneous noises and game sounds compared to regular players, who are more entertained and engaged in the game.
- The values in the Pz channel for all rhythms are significantly higher for pro players. Since the parietal cortex area is responsible for motor functions, it can be assumed that for pro players the increased values are associated with their skills and experience.

Player proficiency prediction

The data (2-second intervals) was splitted in the following ratio: 20% test, 80% train+val, of which 20% val and 80% train.

First, the selected models with tuned hyperparameters were trained on time-domain features. The results of their evaluation with cross-validation are presented in Table 3. The weighted F1-score was calculated using the train+val set. When obtaining the model scores (accuracy), the training was on the train set, and the evaluation was performed on the val set.

Table 3. The classification results of players' proficiency prediction on time-domain features

Model	weighted F1-score	Accuracy
LR	0.971 ± 0.004	0.972
SVM	0.907 ± 0.008	0.906
KNN	0.726 ± 0.004	0.740
RF	0.972 ± 0.003	0.967
MLP	0.881 ± 0.009	0.879

It can be seen that Random Forest and Logistic Regression classifiers perform better than the others both in terms of F1-score and accuracy on the validation set. We select Random Forest for the further analysis as it shows a bit higher weighted F1-score compared Logistic Regression, has a bit lower std and also provides built-in feature importances.

In general, it is common practice that linear models (LogisticRegression and SVM) perform quite well on brain activity data. In this particular case, Random Forest also works well, which means that certain decision rules are well applicable to the considered data.

The Random Forest classifier results on the test sample: weighted F1 score = 0.965, and its feature importance analysis represents on Figure 1.

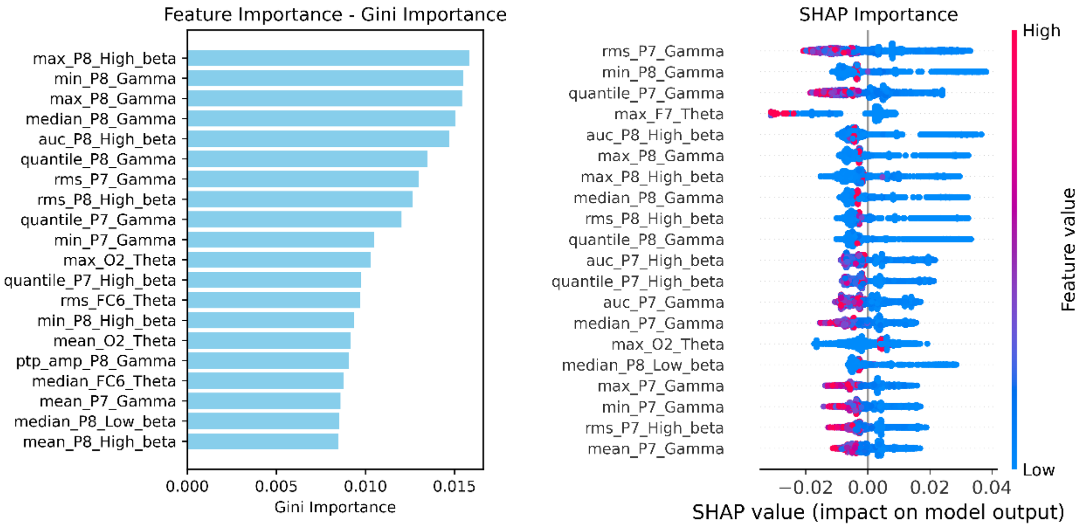

Figure 1. Feature importance analysis for Random Forest classifier on time-domain features using built-in impurity-based feature importance method and SHAP values.

It is clear that the most important features are almost the same both with the built-in impurity-based feature importance analysis and with the calculation of the Shapley values. They also largely coincide with the findings at the exploratory data analysis stage. Specifically, the most distinctive are the brain activities in the gamma and beta frequency bands of the temporal cortex (P7, P8), as well as the theta rhythm of the occipital lobe (O2 channel), which is responsible for processing visual information. According to the SHAP analysis (built for pro class), it can be concluded that all the previously listed features (except max_Theta_O2) have an inverse relationship with professionalism (the lower their values, the more likely the considered time interval belongs to a pro player).
Also, gamma waves are thought to be a sign of the active exchange of information between the cerebral cortex and other areas, and usually generated in the brain when people are conscious and when the eyes move rapidly. Gamma and beta waves often overlap within the range of natural frequencies, and the exact boundary between these two frequency bands is not clear. Therefore, these 2 frequency bands are the most distinctive and are presented together.

Next, the feature set was expanded by adding channel-band based indexes. For the model selected in the previous step, the best hyperparameters were also searched along the same grid specified in Table 2. The weighted F1-score with 5-fold cross-validation calculated on train+val set is equal to 0.969±0.005, on test set weighted F1 score = 0.965. Feature importance analysis represents on Figure 2 and shows that autonomic EEG indices do not get in the top-30 important features, while the most significant ones are practically the same as in the previous case for both methods.

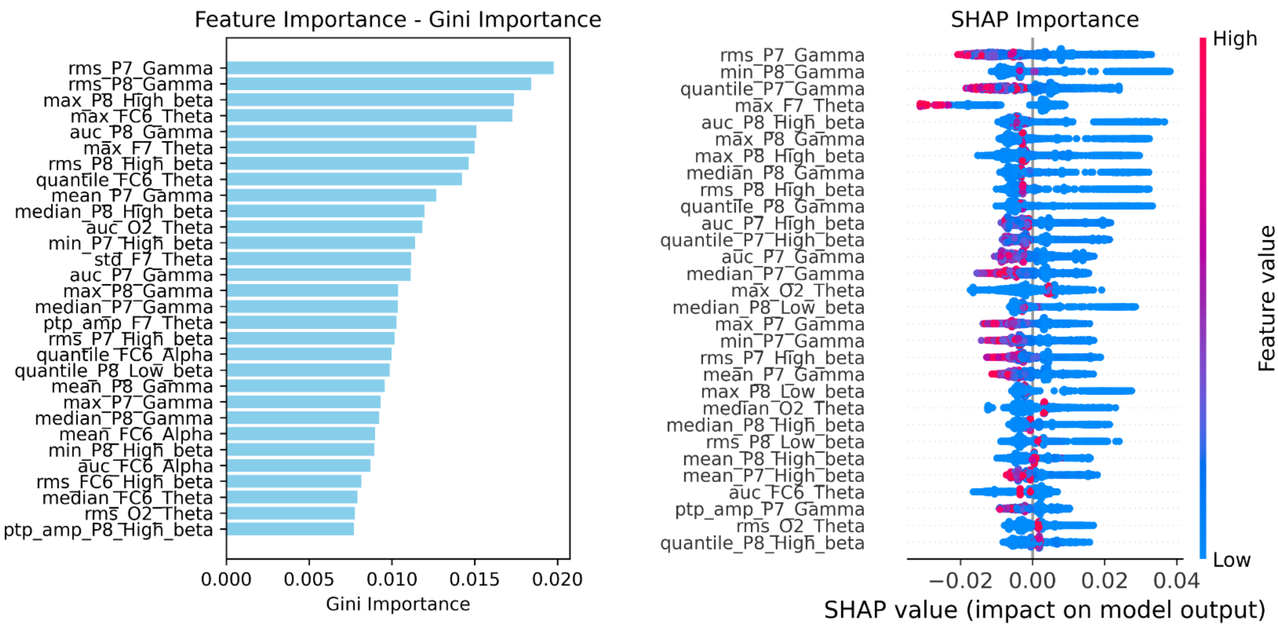

Figure 2. Feature importance analysis for Random Forest classifier on mixed features (time-domain and channel-band based indexes) using built-in impurity-based feature importance method and SHAP values.

Table 4 contains the results of applying dimensionality reduction (PCA) and feature selection (SelectKbest) methods on time-domain features and mixed features (time-domain + autonomic indices). When evaluating the model performance, nested cross validation was used with a search for the best hyperparameters.

Table 4. The model results on time-domain features and mixed features with applying dimensionality reduction (PCA) and feature selection (SelectKbest) methods.

Method	features used	weighted F1-score
SelectKBest	time-domain	0.982 ± 0.003
	mixed	0.986 ± 0.003
PCA	time-domain	0.617 ± 0.007
	mixed	0.616 ± 0.007

The table shows that the SelectKBest feature selection method generally performs better than PCA. That is, direct removal of uninformative features for a given dataset is better than converting all features into principal components.
Having determined the best model hyperparameters (ccp=0.001, criterion=gini, max_depth=5, max_features=sqrt) and a specific SelectKBest setup (k=500) for mixed features evaluation was performed on test set showing f1-score 0.987. It can be concluded that the model slightly improved regarding f1 score.
Figure 3 provides the feature importance analysis for this case. The most important features remained almost the same as in the previous cases. At the same time, the dimension of the feature space decreased by 60% (from 1278 features to 500).

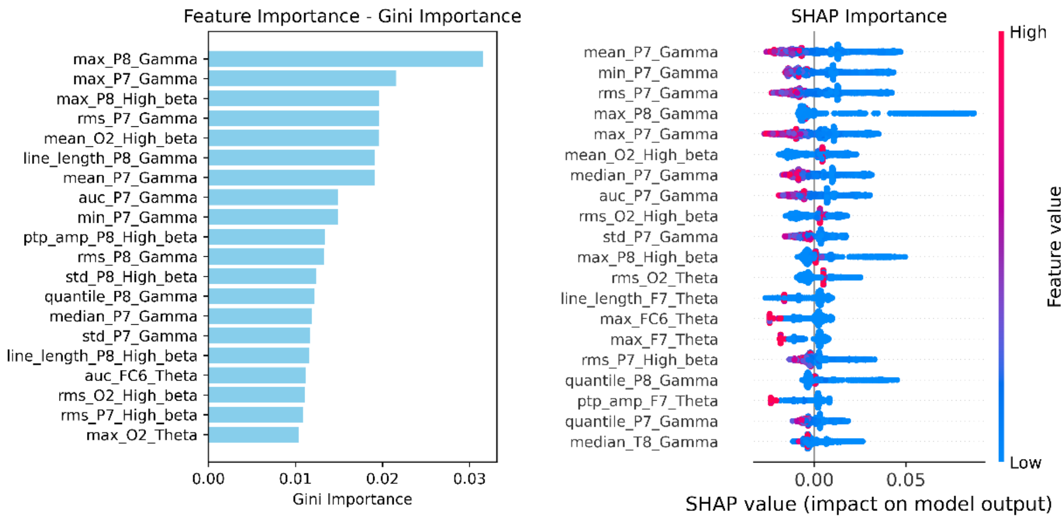

Figure 3. Feature importance analysis for Random Forest classifier on mixed features using SelectKBest method for selection the most important 500 features.

Next, for this particular setup (SelectKBest on mixed features), we test the use of upsampling methods SMOTE and ADASYN with an increase of minority class samples to the ratio N_minority/N_majority class samples = 0.65 or 0.85 (the initial ratio in the train set 0.37). Table 5 shows the results of applying upsampling methods using nested cross validation for evaluation.

Table 5. The model results on time-domain features with use of PCA with applying upsampling techniques.

Method	New class ratio	weighted F1-score
SMOTE	0.65	0.990 ± 0.001
	0.85	0.985 ± 0.003
ADASYN	0.65	0.992 ± 0.002
	0.85	0.990 ± 0.001

Using ADASYN (0.65) and the model with optimized hyperparameters (ccp=0.001, criterion=gini, max_depth=5, max_features=sqrt) weighted F1 score = 0.988. Feature importances are indicated in Figure 4 and mostly repeat Figure 3. Also note that autonomic indices do not listed as the most significant in any of the feature importance graphs, although using a mixed set of features with their inclusion provides improved performance compared to time-domain features alone. This supports their usefulness.

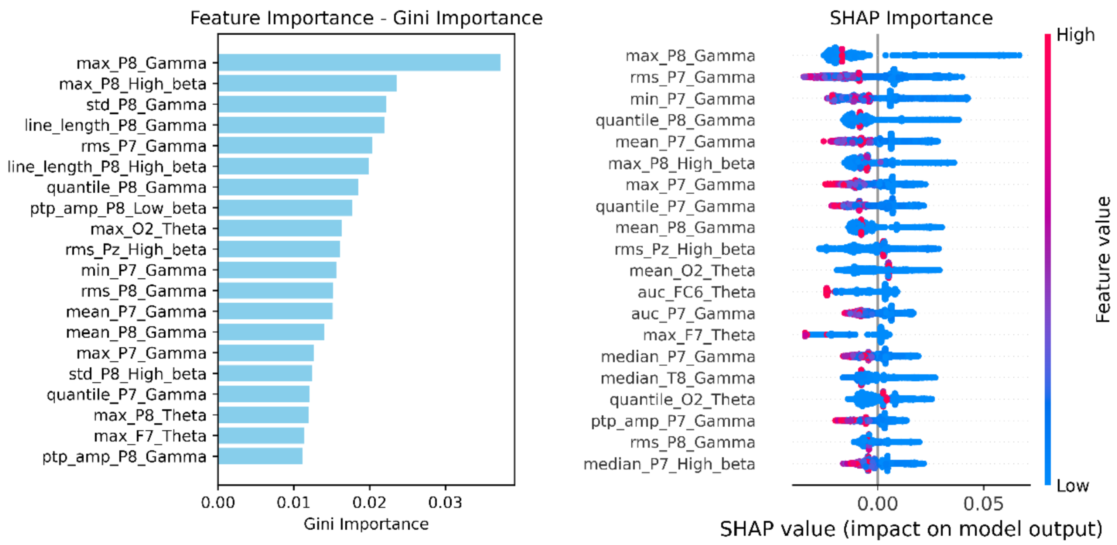

Figure 4. Feature importance analysis for Random Forest classifier on mixed features using SelectKBest (k=500) and ADASYN for upsampling minority class to 0.65 class ratio.

Conclusion

As a result, we found that the best prediction of players' skills is achieved using the Random Forest classifier with parameters (ccp=0.001, criterion=gini, max_depth=5, max_features=sqrt) on mixed features with preliminary application of SelectKBest to select 500 most significant features and subsequent use ADASYN upsampling method to increase the minority/majority class ratio to 0.65.

According to the feature importance analysis from Figure 4, it is evident that, as mentioned earlier, the most distinctive patterns of brain activity of casual and pro players are in the gamma, beta, and theta frequency bands, as well as in the temporal cortex, the frontal lobe, and the parietal cortex. Such findings are fully consistent with the previously obtained results at the exploratory data analysis stage, as well as existing articles [4,13].
In general, it can be concluded that i) increased concentration and focus on the sound environment and sound accompaniment in the game (channels T7, P7, T8, P8, rhythm bands Gamma and Beta) reflects a low level of gaming skill, and vice versa, concentration on processing visual stimuli in the game, and coordinated motor actions in response to them are characteristic of pro players. (channels Pz, O2, rhythm Theta).

References

1. Z. Minchev, G. Dukov, and S. Georgiev, “Eeg spectral analysis in serious gaming: An ad hoc experimental application,” International Journal Bioautomation, vol. 13, 12 2009.
2. B. Meneses-Claudio and A. Roman-Gonzalez, “Study of the brain waves for the differentiation of gamers category between a newbie and a hardcore in the game dota 1,” in 2018 Congreso Argentino de Ciencias de la Inform´atica y Desarrollos de Investigaci´on (CACIDI), 2018, pp. 1–4.
3. S. M. Anwar, S. M. U. Saeed, M. Majid, S. Usman, C. A. Mehmood, and W. Liu, “A game player expertise level classification system using electroencephalography (eeg),” Applied Sciences, vol. 8, no. 1, 2018.
4. N. Melentev, A. Somov, E. Burnaev, I. Strelnikova, G. Strelnikova, E. Melenteva, and A. Menshchikov, “esports players professional level and tiredness prediction using eeg and machine learning,” in 2020 IEEE SENSORS, 2020, pp. 1–4.
5. A. Smerdov, A. Somov, E. Burnaev, B. Zhou, and P. Lukowicz, “Detecting video game player burnout with the use of sensor data and machine learning,” IEEE Internet of Things Journal, vol. 8, no. 22, pp. 16 680–16 691, 2021.
6. S. Gostilovich, A. Kotliar Shapirov, A. Znobishchev, A.-H. Phan, and A. Cichocki, “Biomarkers of professional cybersportsmen: Event related potentials and cognitive tests study,” PLOS ONE, vol. 18, no. 8, pp. 1–37, 08 2023.
7. Fellrath, Julia & Mottaz, Anaïs & Schnider, Armin & Guggisberg, Adrian & Ptak, Radek. (2016). Theta-band functional connectivity in the dorsal fronto-parietal network predicts goal-directed attention. Neuropsychologia. 92. 10.1016/j.neuropsychologia.2016.07.012.
8. Puszta, András & Katona, Xénia & Bodosi, Balázs & Pertich, Ákos & Nyujtó, Diána & Braunitzer, Gabor & Nagy, Attila. (2018). Cortical Power-Density Changes of Different Frequency Bands in Visually Guided Associative Learning: A Human EEG-Study. Frontiers in Human Neuroscience. 12. 10.3389/fnhum.2018.00188.
9. Diaz, Hernan & Gárate, Elías & Maureira Cid, Fernando & Flores Ferro, Elizabeth. (2022). Hurst entropy profiles for Beta Low and Beta High EEG sub-bands Part I: Intragroup comparison.. Procedia Computer Science. 199. 1416-1423. 10.1016/j.procs.2022.01.179.
10. Aldayel, Mashael & Ykhlef, Mourad & Alnafjan, Abeer. (2021). Recognition of Consumer Preference by Analysis and Classification EEG Signals. Frontiers in Human Neuroscience. 14. 10.3389/fnhum.2020.604639.
11. Mateos, Diego & Krumm, Gabriela & Filippetti, Vanessa & Gutierrez, Marisel. (2022). Power Spectrum and Connectivity Analysis in EEG Recording during Attention and Creativity Performance in Children. NeuroSci. 3. 347-365. 10.3390/neurosci3020025.
12. Khondakar, Md. Fazlul Karim & Sarowar, Md & Chowdhury, Mehdi & Majumder, Sumit & Hossain, Md. Azad & Dewan, M. & Hossain, Quazi Delwar. (2024). A systematic review on EEG-based neuromarketing: recent trends and analyzing techniques. Brain Informatics. 11. 10.1186/s40708-024-00229-8.
13. M. Doppelmayr, T. Finkenzeller, and P. Sauseng, "Frontal midline theta in the pre-shot phase of rifle shooting: Differences between experts and novices," Neuropsychologia, vol. 46, pp. 1463–7, 05 2008.