BSE662: Decision Making and The Brain

Instructor: Dr. Arjun Ramakrishnan

Endsem Project Report

Ahmad Raza	220088
Anya Rajan	220191
Debarpita Dash	220328
Manasvi Nidugala	220613
Nischay Patel	220721
Poojal Katiyar	220770

EEG and Decision Game-Based Prediction of Trait Anxiety

Mentors: Peeusa Mitra, Ayan Das

Contents

EEG and Decision Game-Based Prediction of Trait Anxiety

`GitHub Repository: EEG_Data_Analysis`

1 Introduction and Key Questions

This project focuses on leveraging machine learning and deep learning techniques to predict trait anxiety levels using behavioral and physiological (EEG) data. The data was collected from two different experimental environments that varied in travel time during a foraging task:

- **Richer environment:** 3-second travel time
- **Poorer environment:** 10-second travel time

Behavioral data includes salivary cortisol and salivary alpha-amylase (SAA) levels collected before and after the task. **Physiological data** was obtained from EEG recordings using 29 scalp electrodes.

A total of 54 participants (18 females) took part in the study. Each participant completed approximately 100–120 epochs during the foraging game, where each epoch consisted of a binary decision to either *stay* or *leave* a patch.

For each epoch and each EEG channel, we extracted relevant features such as:

- Theta-Beta Ratio
- Event-Related Potentials (ERP)
- Entropy
- Fractal Dimensions
- Functional Connectivity metrics (e.g., PLV, coherence)

The central research questions of the project are as follows:

- **Trait Anxiety Prediction:**
  - Can we predict trait anxiety using theta-gamma coupling and total reward?
  - Can SAA (salivary alpha-amylase) levels be predicted using theta-gamma coupling?
  - Can alpha and theta power from frontal electrodes predict trait anxiety?
  - Can standard deviation of entropy predict patch switching frequency?
  - Can entropy-based features predict cortisol levels?
  - Can EEGNet applied to preprocessed epoch data (in .set and .fdt format) predict trait anxiety?
  - Can multimodal features (EEG + behavioral) be used to predict trait anxiety at the participant level?
- **Decision Prediction during Foraging:**
  - Can we predict stay/leave decisions using basic ML models applied to single-epoch data?
  - Can we improve decision prediction using multimodal data (EEG + behavioral)?
  - Can entropy-based features predict the decision to stay or leave?
  - Can reaction time during the task be predicted from entropy measures?

2 EEG Data Preprocessing Pipeline

The following preprocessing pipeline was implemented using **EEGLAB** in **MATLAB** for EEG data collected from 54 participants. Each EEG dataset was paired with behavioral data, and preprocessing was conducted to prepare the data for spectral and behavioral analyses, particularly in the context of trait anxiety prediction during a foraging task.

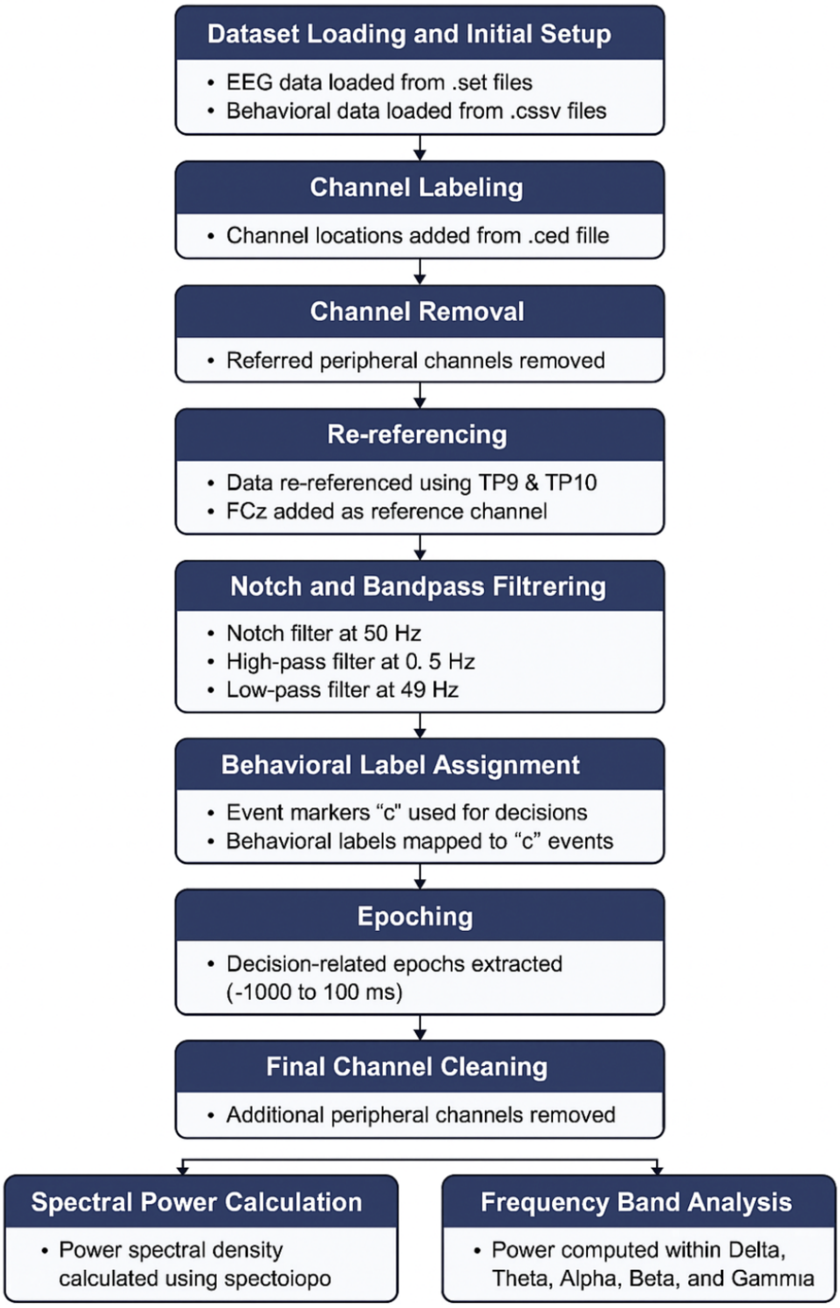

Figure 1: Data Preprocessing

3 Spectral Analysis

To demonstrate the results of preprocessing, spectral analysis was conducted on one preprocessed EEG dataset:

3.1 Spectral Power Calculation

- Power spectral density (PSD) was calculated using `spectopo`.
- This provided a frequency-wise distribution of power to assess cleaning efficacy and dominant frequency components.

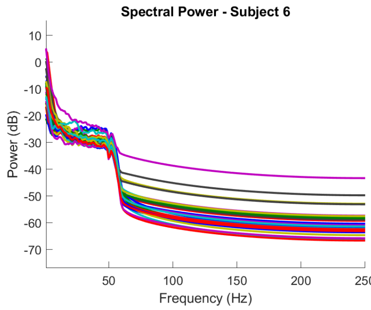

Figure 2: spectral_power_e2_p6

3.2 Frequency Band Analysis

- Power values were extracted across five standard EEG bands: **Delta(1-4Hz), Theta(4-8Hz), Alpha(8-12Hz), Beta(12-30Hz) and Gamma(30-50Hz)**
- Mean power values in each band were visualized using a bar plot, summarizing spectral characteristics.

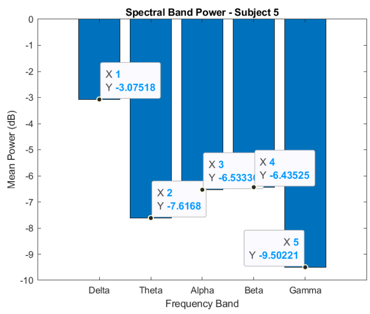

Figure 3: band_power_e2_p6

4 Feature Extraction

These are the features that we have extracted for analysis.

4.1 Sample Entropy

- EEG signals exhibit non-linear, complex dynamics—Sample Entropy (SampEn) captures these complexities better than traditional measures like mean or power.
- SampEn(m, r, N) quantifies unpredictability by comparing the number of matching patterns of length m and $m + 1$.
- Let A = number of matches of length $m + 1$, B = number of matches of length $m \Rightarrow$ **SampEn** = $-\ln(A/B)$.
- A greater drop in pattern matches indicates more irregularity, leading to higher entropy (less predictable signal).
- **Interpretation:** Higher SampEn $\rightarrow$ erratic brain activity (linked to anxiety); Lower SampEn $\rightarrow$ stable patterns (linked to relaxation).

4.2 ERP Components Analyzed

In this study, we analyzed a set of key Event-Related Potential (ERP) components using predefined latency windows and polarity-based amplitude extraction across all EEG channels. These components were selected due to their established roles in cognitive and affective processing, including sensory perception, attention, error monitoring, and emotional evaluation.

Table 1: ERP Components Extracted and Their Characteristics

ERP Component	Time Window (s)	Polarity	Functional Significance
P1	-0.05 to 0.05	Positive	Early visual sensory processing; reflects initial attentional responses to stimuli.
N170	-0.10 to 0.00	Negative	Specialized for face perception; sensitive to the structural encoding of facial features.
N200	-0.10 to 0.00	Negative	Associated with conflict monitoring and cognitive control; often enhanced during error detection.
P300	-0.10 to 0.10	Positive	Reflects stimulus evaluation, attention allocation, and decision-making under uncertainty.
FRN	0.00 to 0.10	Negative	Elicited by negative feedback; indicates sensitivity to performance errors or outcome evaluation.
LPP	0.00 to 0.10	Positive	Sustained attentional processing of emotionally salient stimuli; linked to affective evaluation.
MMN	0.00 to 0.10	Negative	Indicates automatic detection of auditory deviance, independent of conscious attention.

Epochs were cropped to these time windows, and both mean and peak amplitudes were computed per ERP component across all channels. Polarity guided the direction of peak amplitude extraction: for negative components (e.g., N200, FRN), the minimum value was used; for positive components (e.g., P300, LPP), the maximum value was taken. This approach ensured accurate quantification of neural responses for each ERP marker.

The extracted ERP features are foundational for subsequent participant-level analyses, including statistical comparisons and modeling of cognitive-affective dynamics.

4.3 Phase-Amplitude Coupling (PAC): Theta–Alpha/Beta/Gamma

- PAC measures how the phase of slower brain waves (e.g., theta) modulates the amplitude of faster waves (e.g., alpha, beta, gamma).
- This cross-frequency coupling reflects coordination across brain regions and efficient neural communication.
- In individuals with high trait anxiety, PAC patterns often differ—especially in how theta modulates faster oscillations associated with emotion and attention.
- PAC is thus a valuable and sensitive feature for predicting anxiety levels based on EEG signals.

4.4 Alpha Band Coherence in Frontal Regions

- Coherence quantifies functional connectivity between two EEG channels—how consistently they oscillate together at a given frequency (ranges from 0 to 1).
- Alpha band (8–12 Hz) coherence was measured between frontal channels (F3, F4, Fz).
- These regions are key in cognitive control, emotional regulation, and attentional processes—functions that are modulated by anxiety.
- Reduced alpha coherence may reflect impaired neural communication in anxious individuals.

4.5 Theta Band Coherence in Frontal Regions

- In this analysis, we computed theta-band coherence between key frontal electrodes (F3, F4, and Fz). These regions are associated with executive functioning and anxiety-related neural activity.

Increased or decreased coherence in this band may reflect how well the frontal cortex coordinates during foraging decisions under uncertainty, an aspect tightly linked to trait anxiety.

- Higher theta coherence may indicate more structured communication between frontal areas, while lower coherence can reflect neural dysregulation, often observed in high-anxiety individuals.

5 Feature Dendrograms

We created dendrograms for all the data that we had in order to visualize the relationships in the data. This analysis helps identify groups of related features, potentially enabling feature selection, redundancy removal, and understanding of the underlying structure in your dataset. The dendrogram provides a visual map of how features relate to each other, with similar features clustering together at lower heights

5.1 Participant-wise feature Hierarchy

In this case, we observed the following major clusters:

- Total_reward - theta_gamma - saa
- Patch_switching_freq - sd_ch(Entropy SD of all channels) - cort(Cortisol Level)
- Trait_anx - coherence

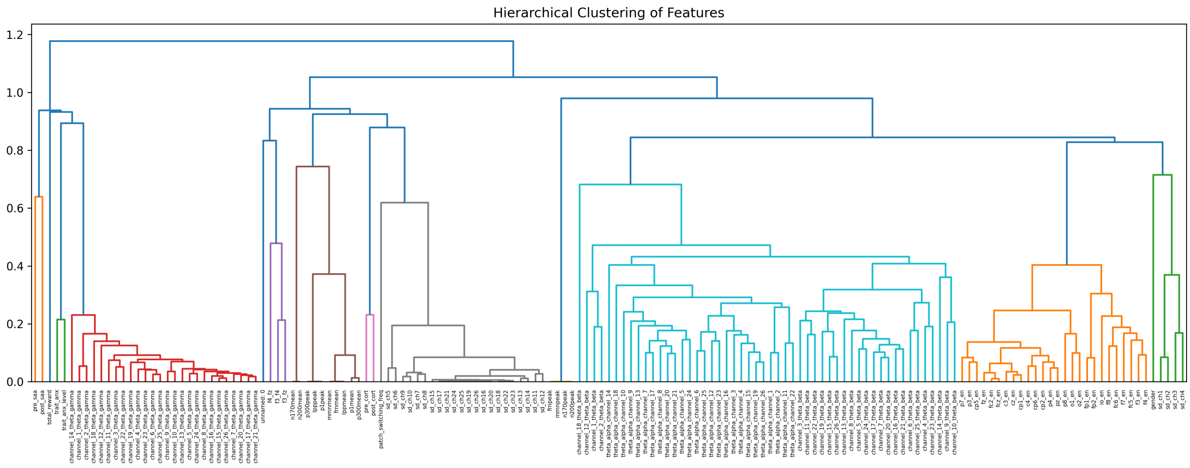

Figure 4: Participant-wise in Both Environments

5.2 All Epochs in Both Environments:

We see the following major clusters:

- Decision - Reward
- Train_anx - Reaction_time - Entropy
- Tree number - time_elapsed - total_cumulative_reward

We didn't see any major difference in the dendrograms of the 2 separate environments. By looking at the tree, we focused on theta-gamma coupling and which channels are important for reward gain. FC2 came as the most important channel and is located in the right frontal cortex, a region often tied to executive control, attention, and reward processing.

6 Mediation and Moderation

We examined how salivary alpha-amylase (SAA), a biomarker of sympathetic nervous system activity, interacts with theta-gamma coupling, linked to cognitive control and memory integration, to influence reward outcomes. This analysis tested both mediation (mechanism) and moderation (conditions) effects.

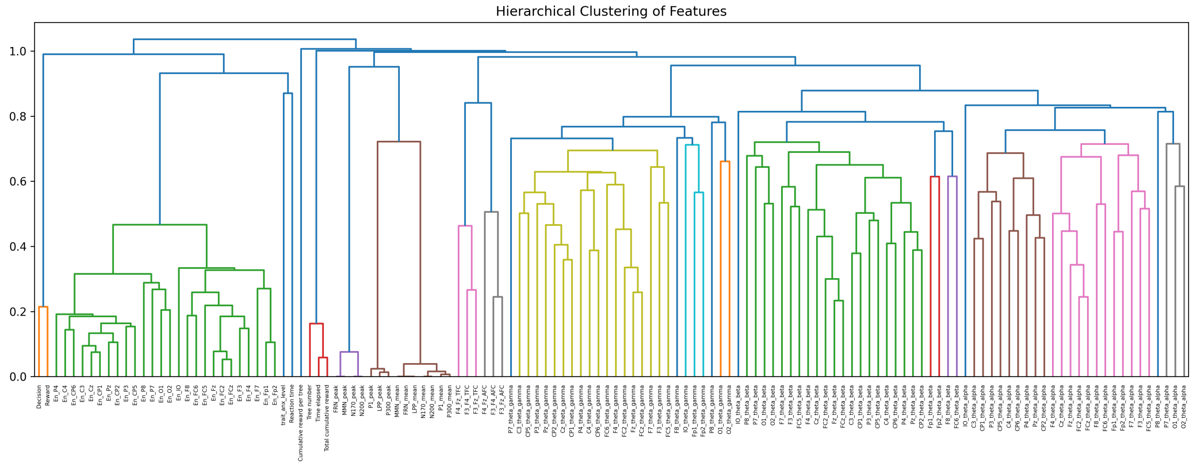

Figure 5: Epoch-wise in Both Environments

Rank	EEG Channel	Approx Brain Area	Importance
1	channel_19_theta_gamma	FC2 (Fronto-central R)	0.0829
2	channel_4_theta_gamma	F4 (Frontal R)	0.0644
3	channel_2_theta_gamma	FP2 (Prefrontal R)	0.0630
4	channel_15_theta_gamma	FZ (Frontal midline)	0.0584

Table 2: Top EEG channels with their approximate brain areas and importance scores.

6.1 Moderation Plot Interpretation

- **Different Slopes by SAA Level:**
  - **Low SAA (blue):** Slightly positive, relatively flat slope.
  - **Medium SAA (teal):** Moderately positive slope.
  - **High SAA (green):** Steepest positive slope.
- **Interaction Effect:** Higher pre-task SAA levels strengthen the relationship between theta-gamma coupling and reward outcomes.
- **Data Distribution:** Most data cluster between -5.5 and -5.0 on the theta-gamma scale, with fewer points at lower values (-7.0 to -6.0).
- **Variability:** Wide confidence intervals, especially in the High SAA group at lower theta-gamma values, indicate individual differences.

6.2 Analysis of Moderation Results: Theta-Gamma Coupling, SAA, and Reward

Statistical Model Performance and Technical Concerns

Reconciling with Visualization

- **Visual Slopes:** Low SAA (flat), Medium SAA (moderate), High SAA (steepest positive slope).
- **Wide Confidence Intervals:** Reflect substantial uncertainty, especially at lower theta-gamma values for High SAA.

Conclusion

Although visual trends suggest slope differences across SAA groups, statistical analysis finds no evidence of moderation. High variability and multicollinearity limit the interpretability of results.

Aspect	Details
Model Fit	Poor fit: R-squared = 0.015; adjusted R-squared = -0.016 (minimal explanatory power).
Model Significance	Non-significant overall: F-statistic = 0.4714, $p = 0.703$.
Predictor Significance	No significant predictors: Theta-gamma coupling ($p = 0.958$), Pre-SAA levels ($p = 0.703$), Interaction term ($p = 0.699$).
Interaction Term	Coefficient = 74.8126, $p = 0.699$ (no significant moderation).
Multicollinearity	High condition number (3.12e+03), indicating strong multicollinearity and unreliable coefficient estimates.

Table 3: Summary of Statistical Model Performance and Technical Concerns

7 Predicting Trait Anxiety Using ML Models

7.1 Using basic ML models

Dataset and Preprocessing:

- 54 participants (behavioral, physiological, EEG); Participants 1, 13, 23, 54 removed (data issues).
- Features: demographics, behavior (travel time, reward, switching), stress markers (SAA, cortisol), EEG band powers.
- No significant pre- vs. post-SAA differences (no state anxiety added).
- Imputation: Numerical (median), Categorical (mode); Target (`trait_anx_level`) encoded to binary (`trait_anx_level_enc`).
- PCA applied to EEG by band and brain region; selected PCs used.

Feature Engineering and Sampling:

- MinMaxScaler normalization (0–1 range); SMOTE for class balance.

Modeling and Evaluation:

- Random Forest with 5-fold stratified cross-validation.
- Confusion Matrix:

$$\begin{bmatrix} 24 & 20 \\ 6 & 50 \end{bmatrix}$$

- Mean CV Accuracy: 0.74 ([0.7, 0.85, 0.8, 0.6, 0.75]).

Model Comparison (80:20 split, no CV):

Model	Accuracy	F1 Score
Random Forest	0.90	0.898
Gradient Boosting	0.75	0.749
MLP Neural Network	0.65	0.651
KNN	0.65	0.637
SVM	0.55	0.547
Logistic Regression	0.45	0.446

Key Findings:

- Random Forest overfits without CV.
- Gradient Boosting (CV accuracy: 0.711) is most reliable.

7.2 Using EEGNet

We employed **EEGNet**, a compact convolutional neural network optimized for EEG-based classification. Its architecture leverages depthwise and separable convolutions to extract both temporal and spatial features from EEG data efficiently. Due to its lightweight design, EEGNet is well-suited for training on limited datasets and was effective in directly predicting trait anxiety levels from preprocessed EEG epochs in our study.

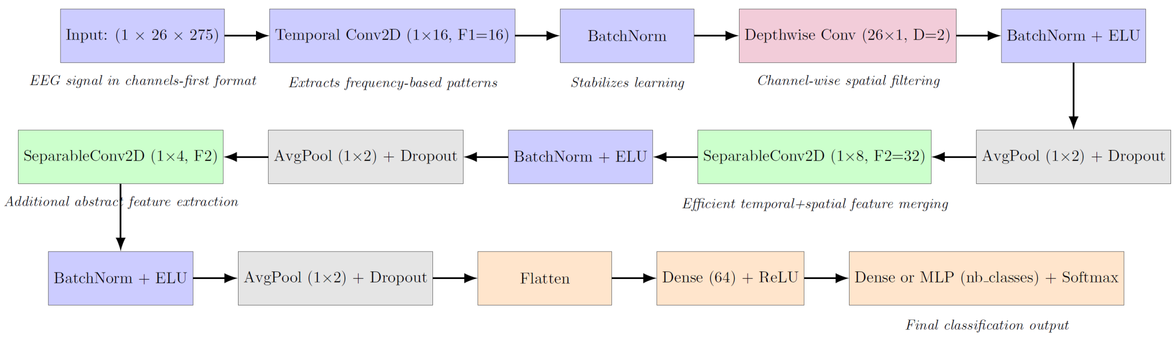

Figure 6: EEGNet Architechture Used

Table 4: Model performance across environments

Environment	Model	Accuracy	Loss	Learning Rate
Env1	Dense	0.7127	0.5872	4.88×10^{-7}
Env1	MLP	0.6679	0.6082	4.88×10^{-7}
Env2	Dense	0.7098	0.5515	1.95×10^{-6}
Env2	MLP	0.6724	0.6204	4.88×10^{-7}

8 Predicting 'Stay' - 'Leave'

8.1 Multimodal Approach

To exploit complementary information from both neural and behavioral signals, we implemented a multimodal classification model. The architecture consisted of a multi-branch neural network, where EEG and behavioral features were processed independently through separate branches before being concatenated and passed to a shared classifier.

The EEG branch consisted of a fully connected layer with 32 hidden units and ReLU activation, followed by a dropout layer with a rate of 0.2. Similarly, the behavioral branch used a fully connected layer with 16 units, ReLU activation, and dropout. The outputs of the two branches were concatenated and fed into a final classifier composed of two linear layers, with the last layer outputting logits for the binary decision classes: *stay* and *leave*.

To address the issue of class imbalance, class weights were computed using `compute_class_weight` from `scikit-learn` and incorporated into the loss function. The model was trained for 10 epochs using the Adam optimizer and a weighted cross-entropy loss.

On the held-out test set, the model achieved a final accuracy of **97.03%**. The predicted class distribution was as follows:

- Predicted *leave*: 306 instances
- Predicted *stay*: 1252 instances

These results support the effectiveness of a multimodal approach, where combining EEG and behavioral features leads to robust classification performance in the *stay/leave* decision task.

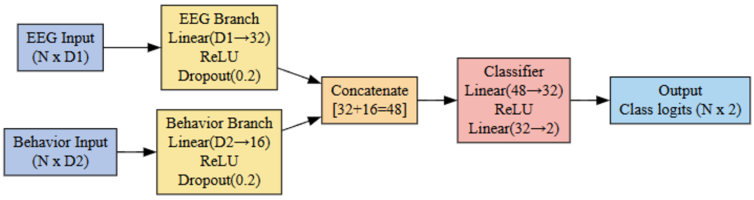

Figure 7: Multimodal architecture

8.2 Random Forest Classifier

This subsection describes the development of a machine learning model to classify participants' decision outcomes—specifically, whether a participant chose to *stay* or *leave* a patch—based on EEG-derived features and behavioral data. The provided dataset consisted of 1332 ¨leave¨decisions and 6457 ¨stay¨decisions, indicating a significant class imbalance.

To reduce bias, Reward and Cumulative Reward per tree features were excluded, while all other features were retained. Given the class imbalance (16.8% "leave" vs. 83.2% "stay"), SMOTE was used to generate minority samples. A Random Forest Classifier was chosen for its robustness and interpretability, with performance evaluated using 5-fold stratified cross-validation and F1-score as the primary metric.

8.2.1 Feature Importance Analysis

Feature importance was extracted from the trained Random Forest model to enhance interpretability:

- Feature importances were computed and visualized.
- The top 10 most influential features were identified for further analysis.

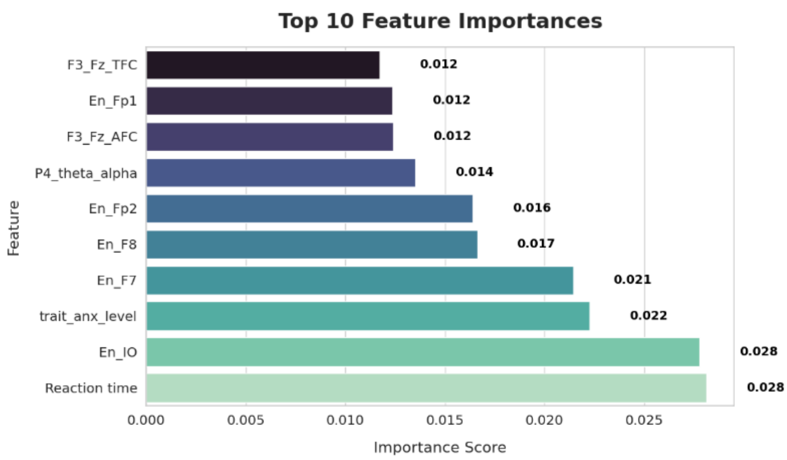

Figure 8: Feature importance scores extracted from the Random Forest model. The top 10 features are highlighted.

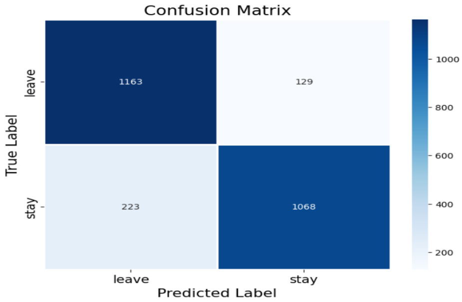

Figure 9: Feature importance scores extracted from the Random Forest model. The top 10 features are highlighted.

9 Predicting Non-EEG Features from EEG Data

This section explores the relationship between EEG-derived features and various non-EEG variables. The aim is to predict non-EEG features from EEG data based on the dendrogram constructed previously, specifically focusing on patch switching frequency, cortisol variation, and trait anxiety.

9.1 Patch Switching Frequency from EEG Entropy

In a foraging game setup, participants repeatedly decide whether to *stay* in a resource patch or *leave* it to find a new one. Patch switching frequency reflects how often a participant decides to leave their current patch.

Relationship between EEG Entropy and Patch Switching Frequency: EEG entropy measures the unpredictability or variability of brain signals. Higher entropy reflects greater neural variability, which is associated with more exploratory behavior, such as switching patches frequently. Conversely, lower entropy indicates more focus and a tendency to stay in the same patch. This suggests a potential link between neural flexibility and decision flexibility.

Model Selection and Prediction: Given the limited dataset (approximately 102 participants), multiple machine learning models were tested to predict patch switching frequency from EEG entropy. The MLP Regressor (a neural network) performed best, capturing complex non-linear patterns in EEG entropy.

Feature importance analysis using Random Forest highlighted that EEG entropy in channels such as Ch26 (FCz) and Ch23 (FC6) were most predictive of patch switching behavior, suggesting these regions play a crucial role in exploration vs. exploitation decisions.

Model	**R2**	**RMSE**	**CV_R2**
Lasso Regression	0.089003	8.153292	0.135614
Ridge Regression	0.085674	8.168172	0.131433
Random Forest	0.094422	8.129004	0.129208
MLP Regressor	0.104504	8.083628	0.089455
XGBoost	0.011825	8.491636	-0.094915
Linear Regression	-0.246215	9.536104	-0.826843

Table 5: Model evaluation metrics: Table 1

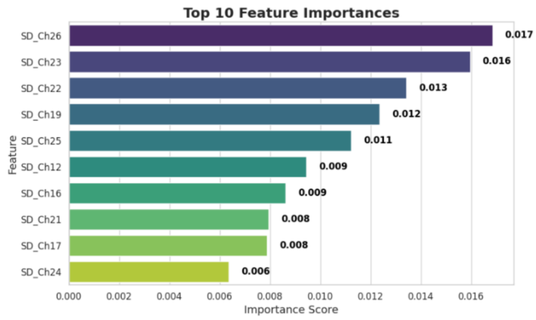

Figure 10: EEG Entropy's role in predicting patch switching behavior.

9.2 Variation in Cortisol from EEG Entropy

Target Variable and Feature Set: The target variable for this analysis was the variation in cortisol levels, defined as the difference between post-cortisol and pre-cortisol values. The feature set included EEG entropy (SD_Entropy) from multiple channels.

To capture non-linear relationships between entropy and cortisol variation, Polynomial Features were used to create interaction terms. However, overfitting and the small sample size hindered the performance of most models. Lasso Regression showed the best results but still provided weak predictions.

Conclusion: This analysis indicates that predicting cortisol variation from EEG entropy alone is weak, likely due to a limited correlation between these two variables. The results suggest that additional features (such as behavioral data or subjective stress ratings) or a larger dataset are needed to make meaningful predictions.

Model	**R2**	**RMSE**	**CV_R2**
Lasso Regression	0.152231	8.849052	-0.367240
Random Forest	-1.214946	14.303407	-0.832767
MLP Regressor	-2.124886	16.989280	-1.042197
Ridge Regression	-5.863752	25.179013	-2.550709
XGBoost	-5.459749	24.426752	-3.989432
Linear Regression	-109.910945	101.215055	-68.222976

Table 6: Model evaluation metrics: Table 2

9.3 Trait Anxiety from Frontal EEG Channels

Model Performance: The task of predicting trait anxiety from frontal EEG channels (alpha_frontal and theta_frontal) showed moderate performance, with cross-validation accuracies ranging from 0.43 to 0.56. This suggests that predicting trait anxiety from these features is challenging, likely due to insufficient data.

9.4 Reaction Time from EEG Entropy

The goal of this model is to predict reaction time based on EEG entropy features. This task is formulated as a regression problem where EEG entropy features extracted from multiple brain channels serve as the input, and the output is a continuous value representing the reaction time.

Model Architecture: The neural network model is a fully connected feedforward architecture consisting of multiple dense layers. Non-linear activation functions, normalization, and dropout are applied to improve convergence and prevent overfitting. Specifically:

- **Activation Function:** GELU (Gaussian Error Linear Unit) is used after each linear transformation to introduce non-linearity.
- **Regularization Techniques:** Batch Normalization is applied after the first linear layer, Layer Normalization after the second, and Dropout is used in early layers with probabilities of 0.5 and 0.3.

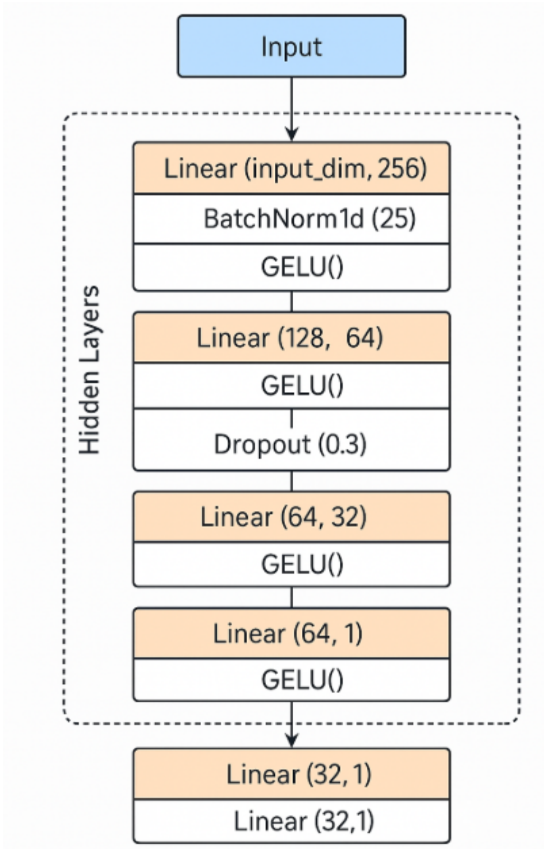

Figure 11: Model architecture

Evaluation Metrics: After training, the model is evaluated on the test set using Mean Squared Error (MSE), Mean Absolute Error (MAE), R^2 Score (Coefficient of Determination).

Environment	MSE	MAE	R^2
Env 2	0.0383	0.1397	0.2203
Env 1	0.0762	0.1820	0.1198
All Epochs	0.0621	0.1621	0.1754

Table 7: Final Test Metrics for Each Environment

9.5 Decision Outcomes: "Stay" vs "Leave" from Entropy

A deep learning model was developed to classify EEG-based decision outcomes—whether a participant decides to *stay* or *leave* a patch—using entropy features extracted from EEG signals. The model addresses class imbalance, normalizes the data, and evaluates performance using classification metrics.

Model Architecture: The `BetterNN` model is a feedforward, fully connected neural network designed for binary classification. It receives entropy features from EEG data as input and outputs a probability distribution over the two decision classes—*stay* and *leave*.

Network Layers:

- **Input Layer:** The input consists of entropy features, with the input dimension automatically determined from the training data.
- **Hidden Layers:** Three fully connected layers, each followed by:
  - Batch Normalization: To stabilize learning.
  - Activation Function: Leaky ReLU to introduce non-linearity.
  - Dropout: To prevent overfitting.

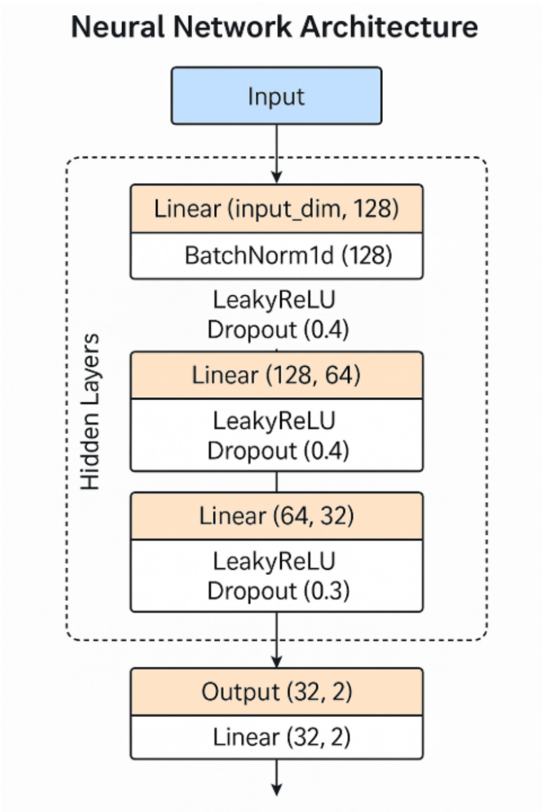

Figure 12: Model architecture

Model Evaluation: This table shows the accuracy metrics.

Setting	Class	Precision	Recall	F1-score	Support	Accuracy	Macro Avg F1
Test Set	0	0.83	0.86	0.84	674	2*0.75 (74.64%)	2*0.58
	1	0.34	0.30	0.32	166		
All Epochs	0	0.85	0.79	0.82	1292	2*0.72 (71.50%)	2*0.56
	1	0.25	0.35	0.29	266		
Env 2	0	0.86	0.74	0.79	1292	2*0.68 (68.04%)	2*0.55
	1	0.24	0.41	0.31	266		

Table 8: Model evaluation metrics (precision, recall, f1-score, support, accuracy, macro average F1) for each setting.

10 Results

This section presents the results of the prediction tasks performed using EEG features for various behavioral outcomes, including decision classification (stay/leave), trait anxiety prediction, and game-related parameter predictions. The performance of different Ml models are discussed.

10.1 Prediction of Stay and Leave Decisions

For predicting the "stay" vs "leave" decisions, the models were trained using epoch-wise data. The following models were tested:

Model	Accuracy	Type of Augmentation
Random Forest Classifier	86%	Weighted Mean
Multimodal Integration	94.53%	No

Table 9: Stay and Leave Decision Prediction Results

10.2 Prediction of Trait Anxiety

Trait anxiety was predicted using participant-wise combined data. The performance of different classifiers and their associated features are outlined below:

Model	Accuracy	Features
Random Forest	90%	No CV, No Data Augmentation
Gradient Boosting	71%	Yes CV, Data Augmentation
Random Forest (No Augmentation)	80%	No CV, No Data Augmentation

Table 10: Trait Anxiety Prediction Results

10.3 Using EEGNET on .set and .fdt Files

EEGNET was applied to .set and .fdt files to classify participants' behavioral outcomes, with the following results:

Environment	Classifier in Last Layer	Accuracy
Env1	Dense	71.27%
Env1	MLP	66.79%
Env2	Dense	70.98%
Env2	MLP	67.24%

Table 11: EEGNET Prediction Results

10.4 Prediction of Game Parameters and Behavioral Data from EEG Features

The performance of various models in predicting game parameters and behavioral data from EEG features is summarized below. All models were trained using participant-combined data.

Feature Predicted	Feature Used	Best Model	Performance Metrics
Patch Switching Frequency	SD of Entropy	MLP Regressor	$R^2 = 0.1045$, RMSE = 8.0836
Cortisol Variation	SD of Entropy	Lasso Regression	$R^2 = 0.1522$, RMSE = 8.8491
Trait Anxiety (Alpha Frontal)	Alpha Frontal	Logistic Regression	Accuracy = 56%
Trait Anxiety (Theta Frontal)	Theta Frontal	KNN	Accuracy = 56%

Table 12: Prediction of Game Parameters and Behavioral Data

10.5 Epoch-Wise Data Results

The results of predicting behavioral outcomes based on epoch-wise data are summarized below:

Feature Predicted	Feature Used	Performance Metrics
Reaction Time	Mean of Entropy	MSE: 0.0383, MAE: 0.1397, R^2: 0.2203
Decision	Mean of Entropy	Accuracy = 71.50%

Table 13: Epoch-Wise Data Prediction Results

11 Conclusion

In this study, we developed and applied an EEG analysis pipeline to investigate neural mechanisms underlying foraging decisions. Beginning with thorough preprocessing and spectral decomposition, we extracted a diverse set of features including nonlinear dynamics (sample entropy, fractal dimension), event-related potentials, phase–amplitude coupling between theta and gamma rhythms, and coherence in alpha and theta bands to capture distinct aspects of brain activity. Clustering these multivariate features revealed meaningful groupings that informed subsequent modeling choices.

We then examined how physiological stress indicators interacted with neural coupling patterns to shape reward processing, highlighting the modulatory role of stress on decision-related oscillatory dynamics. Leveraging these insights, a combination of conventional and deep-learning classifiers was employed to predict individual stay/leave behaviors and trait anxiety from EEG and behavioral inputs. Complementary regression analyses further demonstrated that entropy-based measures correlate with behavioral metrics such as switching frequency, salivary biomarkers, and reaction times.

Overall, our results underscore the value of combining spectral, nonlinear, and connectivity features with behavioral data to characterize individual differences in anxiety and decision making.

12 Individual Contributions

Everyone contributed equally to the project in understanding and implementing it.

- **Ahmad Raza:** EEG dataset preprocessing (for stay/leave and reward epochs), Sample Entropy Extraction (including mean and SD), Dataset Integration, Trait anxiety Prediction using EEGNet
- **Anya Rajan:** Dataset Preparation, ERP features, Prediction of trait anxiety using ML models, Multimodal approach to stay leave
- **Debarpita Dash:** Dataset Preparation, PAC Coupling, Constructing Feature Dendrogram, Statistical analysis on predicted features, Prediction of game parameters and behavioral data from EEG features
- **Manasvi Nidugala:**Dataset Preparation, PAC Coupling, Mediation Moderation Model, Prediction of stay and leave, Prediction of game parameters and behavioral data from EEG features
- **Nischay Patel:** EEG dataset preprocessing (for stay/leave and reward epochs), Sample Entropy Extraction (including mean and SD), Dataset Integration, Prediction of Patch Switching Frequency from SD_Entropy
- **Poojal Katiyar:** Theta frontal and alpha frontal coherence, ERP Features initial code, Prediction of trait anxiety using basic ML models (participant-wise)

References

[1] Lal, U., Chikkankod, A. V., & Longo, L. (2024). Fractal dimensions and machine learning for detection of Parkinson's disease in resting-state electroencephalography. *Neural Computing and Applications*, 36, 8257–8280. https://doi.org/10.1007/s00521-024-09521-4

[2] Ahn, J. S., Hong, H.-J., Lee, J. H., & Park, J. Y. (2024). Theta power reduction and theta–gamma coupling desynchronization are associated with working memory interference and anxiety symptoms in panic disorder: a retrospective study. *BMC Psychiatry*, 24, 875. https://doi.org/10.1186/s12888-024-06272-3

[3] Buckholtz, J. W., Treadway, M. T., Cowan, R. L., Woodward, N. D., Li, R., Ansari, M. S., Baldwin, R. M., Schwartzman, A. N., Shelby, E. S., Smith, C. E., Kessler, R. M., & Zald, D. H. (2010). Mesolimbic dopamine reward system hypersensitivity in individuals with psychopathic traits. *Biological Psychology*, 87(3), 408–419. https://doi.org/10.1016/j.biopsycho.2011.03.015

[4] Huang, Y., & Yu, R. (2014). The feedback-related negativity reflects "more or less" prediction error in appetitive and aversive conditions. *Frontiers in Neuroscience*, 8, 108. https://doi.org/10.3389/fnins.2014.00108

[5] Gu, R., Huang, Y.-X., & Luo, Y.-J. (2010). Anxiety and feedback negativity. *Psychophysiology*, 47(5), 961–967. https://doi.org/10.1111/j.1469-8986.2010.00997.x

[6] Seo, S. H., Lee, C. K., & Yoo, S. K. (2015). Neural correlates related to anxiety in attentional inhibition control: an ERP study. *Neurophysiology*, 46, 428–437. https://doi.org/10.1007/s11062-015-9469-2

[7] Gu, R., Huang, Y.-X., & Luo, Y.-J. (2010). Anxiety and feedback negativity. *Psychophysiology*, 47(5), 961–967. https://doi.org/10.1111/j.1469-8986.2010.00997.x

[8] Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2016). EEGNet: A Compact Convolutional Network for EEG-based Brain-Computer Interfaces. *arXiv preprint arXiv:1611.08024*. https://arxiv.org/abs/1611.08024v1

[9] Muhammad, F., & Al-Ahmadi, S. (2022). Human state anxiety classification framework using EEG signals in response to exposure therapy. *PLOS ONE*, 17(3), e0265679. https://doi.org/10.1371/journal.pone.0265679

[10] Lim, M. X., & Teo, J. (2024). Predicting game-induced emotions using EEG, data mining and machine learning. *Bulletin of the National Research Centre*, 48, 57. https://doi.org/10.1186/s42269-024-01200-7