Emotion detection using EEG

By Kushagra Taneja

Dataset

- This is a dataset of EEG brainwave data that has been processed with the strategy of statistical extraction
- The data was collected from two people (1 male, 1 female) for 3 minutes per state - positive, neutral, negative.
- A Muse EEG headband was used which recorded the TP9, AF7, AF8 and TP10 EEG placements via dry electrodes.
- Six minutes of resting neutral data is also recorded, the stimuli used to evoke the emotions
- Type of Data Considered is Tabular Data (Structured Data)
- Represented as a structured table with rows (observations) and columns (features)
- Data Source (Kaggle) : EEG Brainwave Dataset: Feeling Emotions

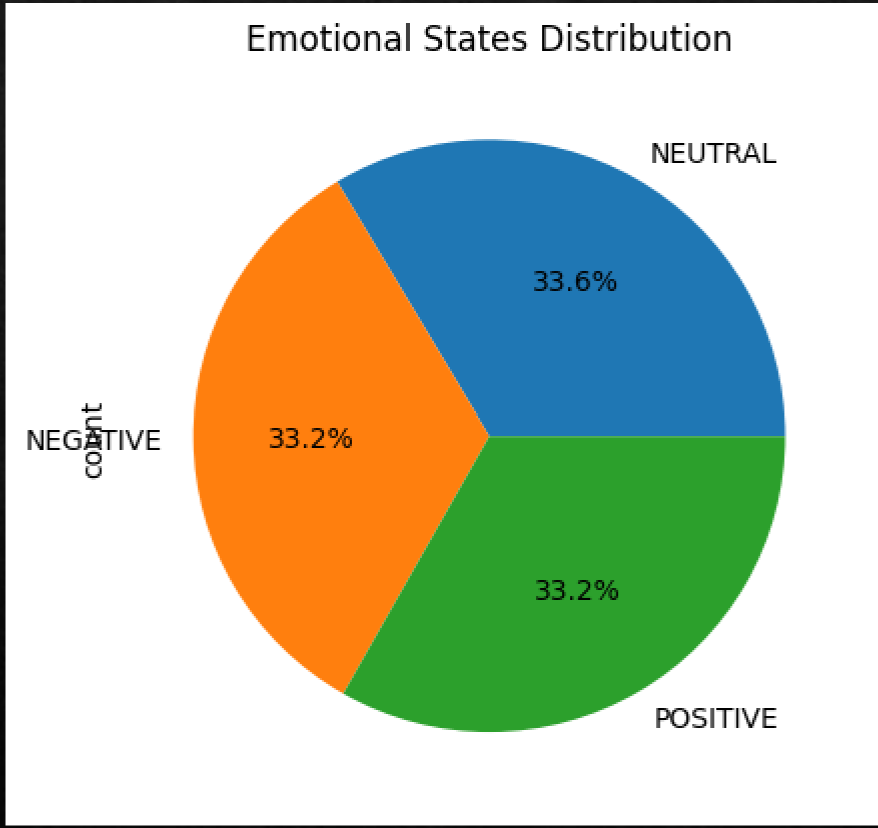
Number of Observations / Subjects

- Total Number of Samples: 2132
- Total Number of Columns: 2549
- Categories based samples :

  Neutral - 33.6%

  Negative - 33.2%

  Positive-33.2%



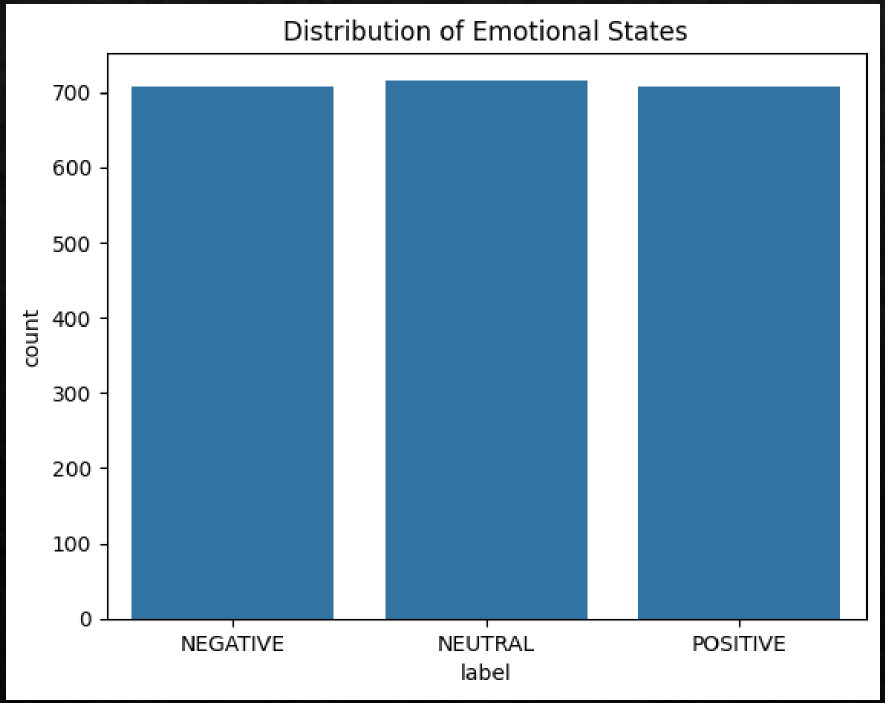
Project Type

- Classification
- The project involves predicting emotional states :

  1.NEGATIVE

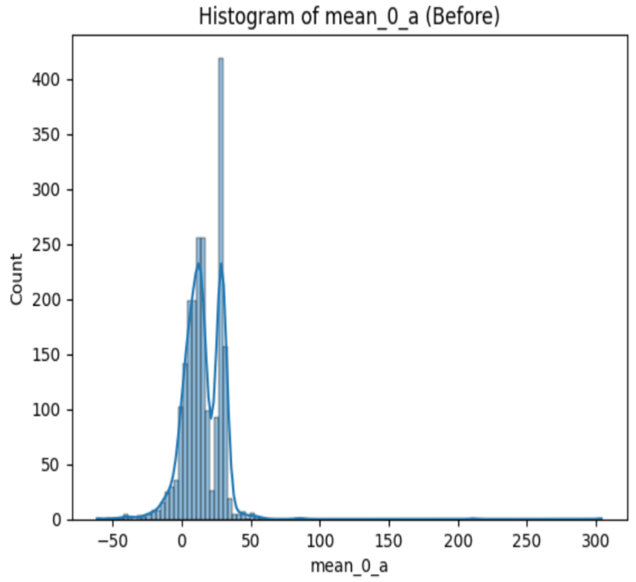
  2. NEUTRAL

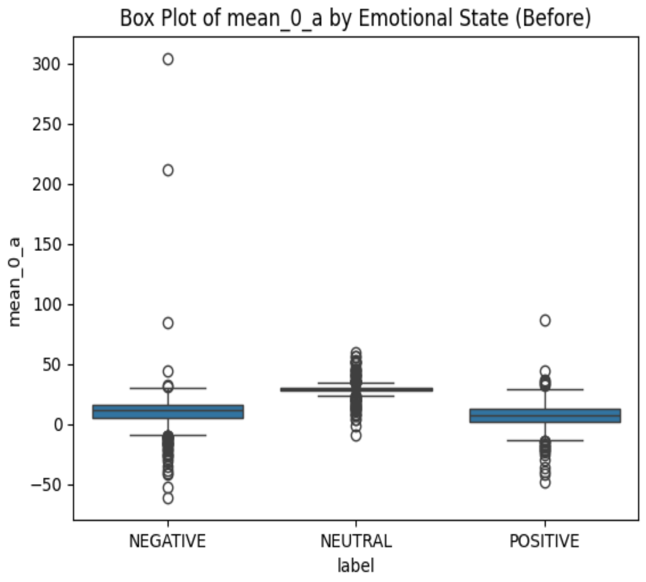
  3. POSITIVE

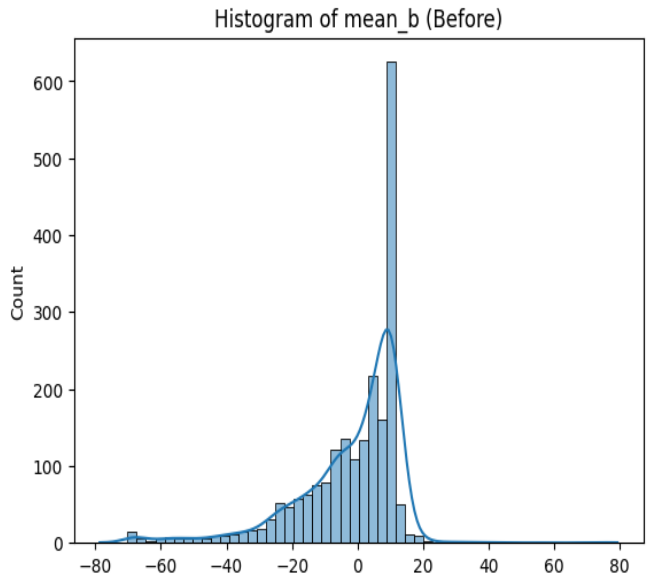


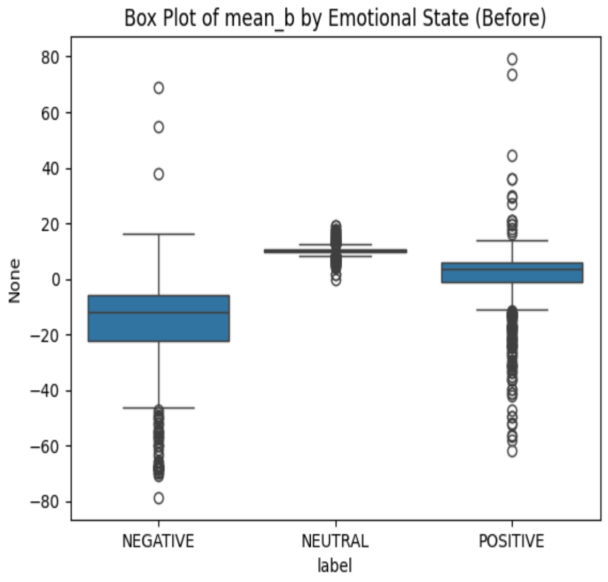
Data Representation Before Feature Extraction

- Basic statistical measures (mean, median, variance, standard deviation) and descriptive statistics are displayed, summarizing raw numerical features.
- Eight visualizations are used to represent the raw data distribution and relationships.









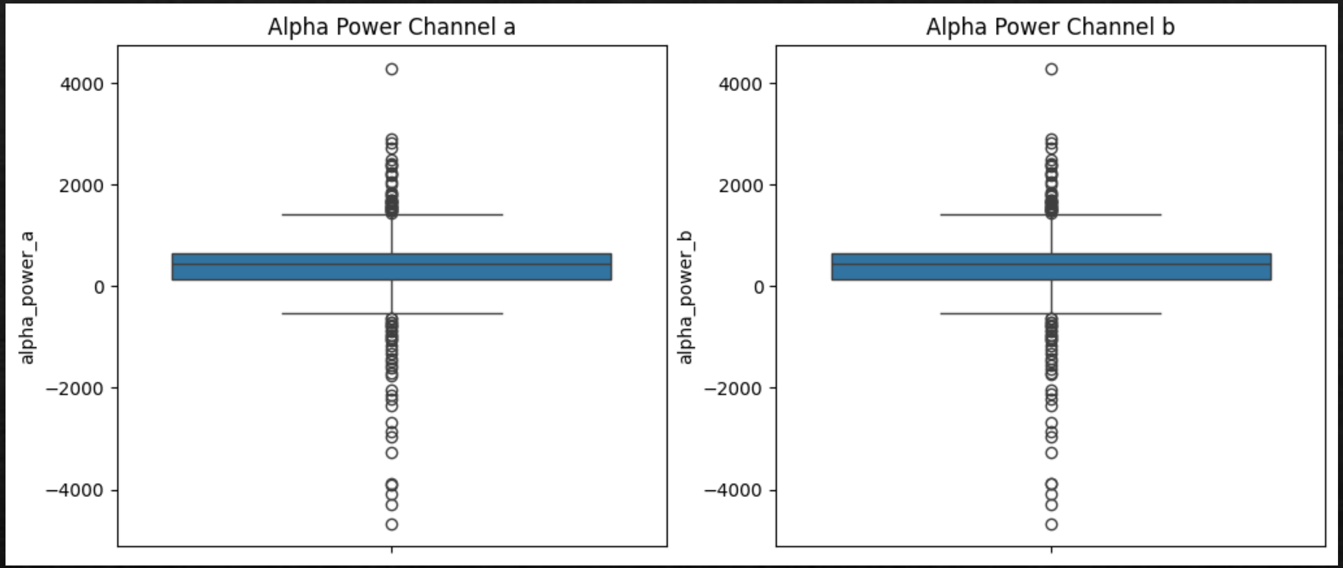
Feature Extraction / Creation Details

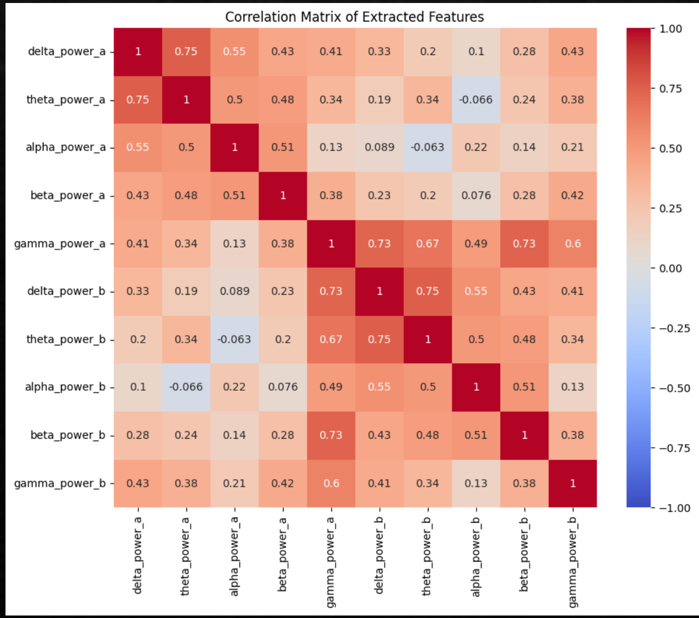
$$\text{Feature}_{band,channel} = \sum_{f=f_{\text{start}}}^{f_{\text{end}}} FFT(f)$$

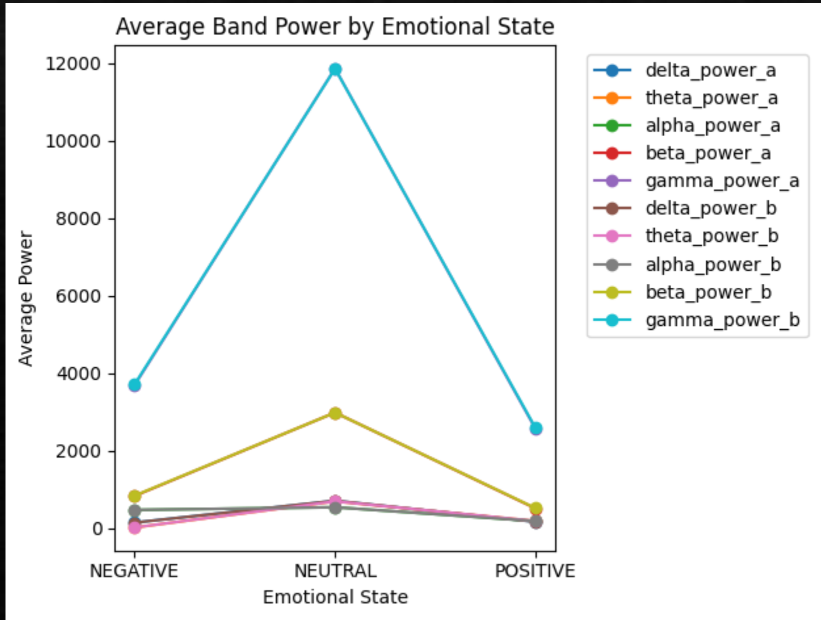
- Total Number of Features Extracted: 10.
- Here f is the frequency bin
- Delta Power (a and b) : f within 0–4 Hz, linked to deep sleep, useful for emotional baselines
- Theta Power (a and b): f within 4–8 Hz, helpful for neutral or calm emotions
- Alpha Power (a and b): f within 8–12 Hz, a relaxed yet alert state (positive or neutral emotions)
- Beta Power (a and b): f within 12–30 Hz, active concentration or anxiety, useful for negative or high-arousal states
- Gamma Power (a and b): f within 30–100 Hz, Indicates high-level cognitive processing, relevant for complex emotional processing

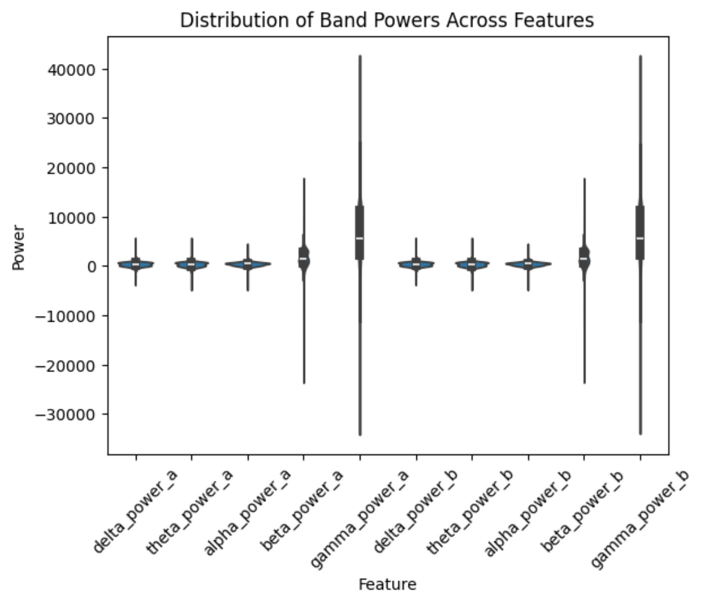
Data Representation After Feature Extraction

- Box Plot of Alpha Power
- Violin Plot of All Band Powers
- Heatmap of Correlation Matrix
- Line Plot of Average Band Power by Emotional State









Feature Selection Techniques Used

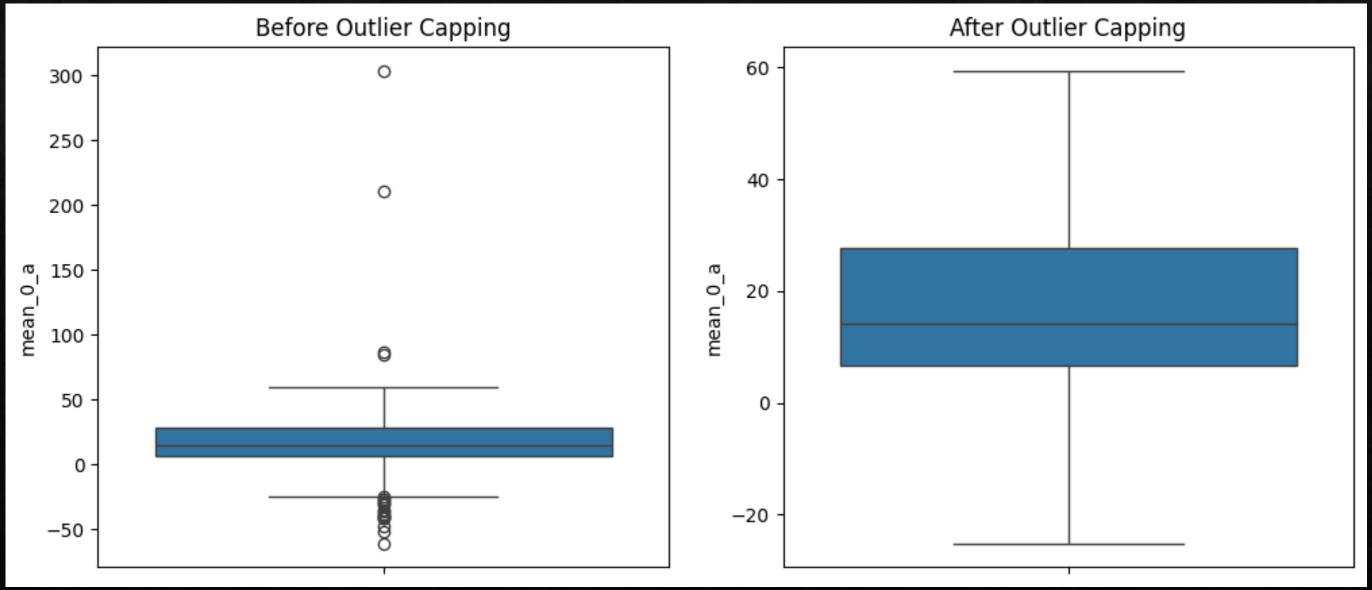
- RL implementation using a Q-learning algorithm with Logistic Regression as the reward function.
- The RL selects features by maximizing the accuracy of a logistic regression model on a validation set, iteratively adding or removing features. This adaptive approach ensures only the most predictive features are retained, reducing dimensionality while preserving classification performance
- Some of the features that were selected are:
  min_q_28_a, fft_725_b, fft_209_a, fft_564_a

Feature Transformation Techniques Used

- Standardization - Scales the selected features to have zero mean and unit variance, ensuring all features contribute equally to the model, which is crucial for algorithms like neural networks and Random Forest that are sensitive to feature scales.

$$x' = \frac{x - \mu}{\sigma}$$

- Outlier Capping - Implemented to mitigate the impact of outliers, which can skew statistical analyses, distort model performance.



Feature Reduction Techniques Used

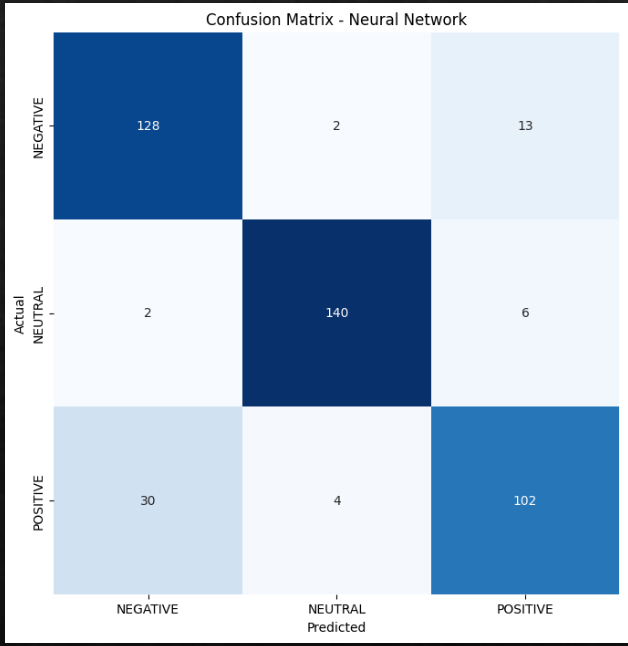
- An autoencoder reduces the dimensionality of the standardized features from the original number (2000+) very high which can lead to curse of dimensionality) to 50 latent dimensions.
- It learns a compressed representation by training the network to reconstruct the input, capturing the most significant variance.
- This reduces noise and computational complexity while retaining relevant information for emotion classification.
- Unlike linear methods like PCA, an autoencoder uses non-linear activation functions (e.g., ReLU) to learn complex patterns in the data.
- EEG signals often exhibit non-linear relationships due to brain dynamics, making an autoencoder better suited to capture these nuances for emotion classification compared to linear techniques.

Hypothesis Testing Methods Used

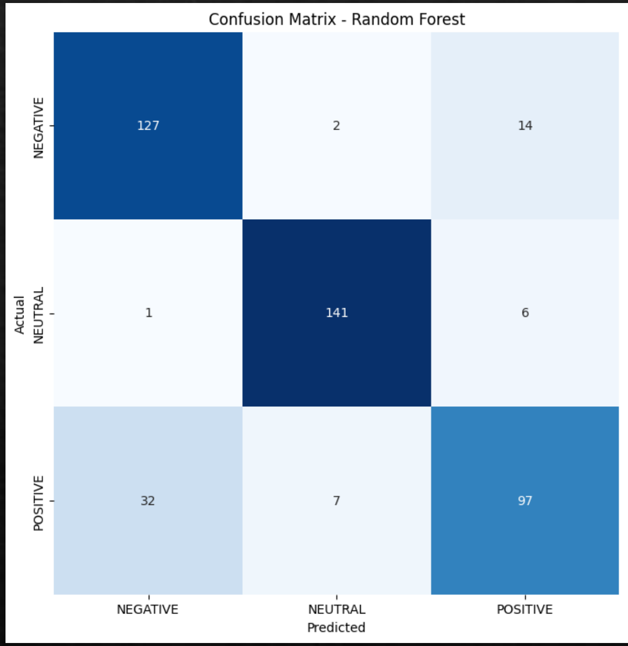
- Independent t-test: To compare the means of a feature (e.g., mean_0_a) between two emotional states (e.g., NEGATIVE vs. POSITIVE) to assess if there are significant differences, helping identify discriminative features.
- ANOVA (Analysis of Variance): To compare the means of a feature (e.g., alpha_power_a) across all three emotional states (NEGATIVE, NEUTRAL, POSITIVE) to determine if there are overall differences, useful for multi-class feature evaluation.
- Chi-Square Test: To assess the independence between the categorical label (label) and a discretized feature (e.g., binned alpha_power_a), verifying if emotional states are associated with feature distributions.

Models Employed

Neural Network



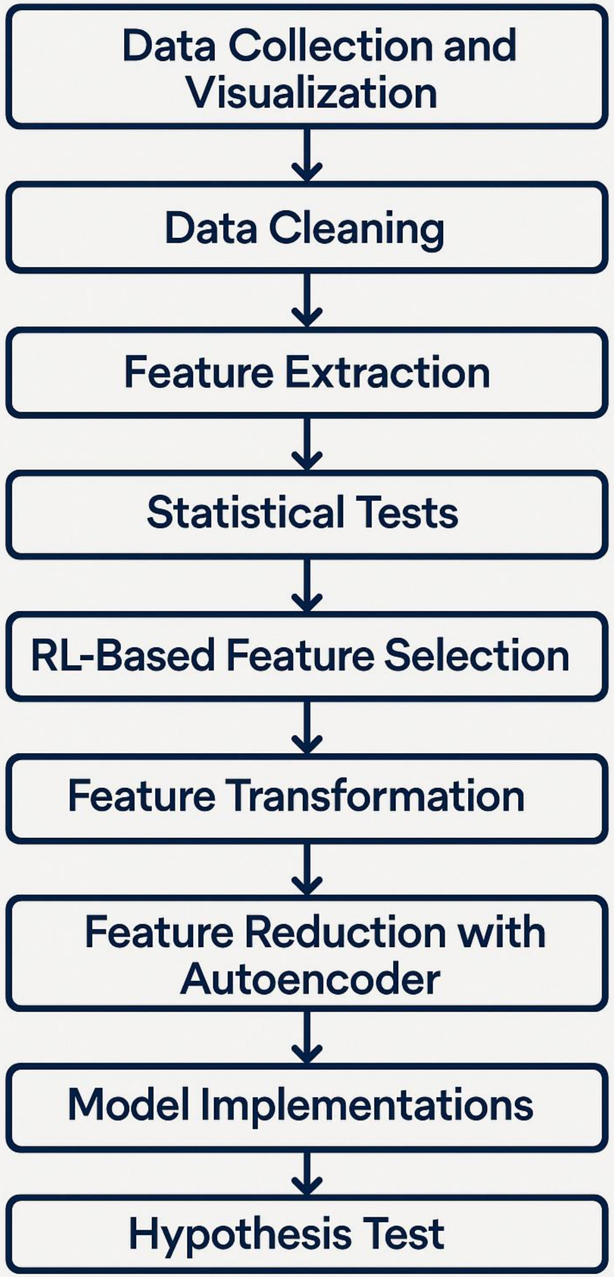
Random Forest



Best Model Selection Criteria

$$\chi^2 = \frac{(b-c)^2}{b+c}$$

- Neural Network Accuracy: 0.867
- Random Forest Accuracy: 0.855
- McNemar’s Test.
- The best model is identified by comparing the neural network and Random Forest using McNemar's test, which analyses prediction disagreements (correct vs. incorrect pairs) rather than just accuracy. A p-value < 0.05 indicates a significant difference, with the model having higher accuracy deemed better if significant
- McNemar's test statistic: 0.64
- p-value: 0.4237107971667936
- No significant difference between the two models.


Data Collection and Visualization
Data Cleaning
Feature Extraction
Statistical Tests
RL-Based Feature Selection
Feature Transformation
Feature Reduction with Autoencoder
Model Implementations
Hypothesis Test

References

- J. J. Bird, L. J. Manso, E. P. Ribiero, A. Ekart, and D. R. Faria, “A study on mental state classification using eeg-based brain-machine interface,”in 9th International Conference on Intelligent Systems, IEEE, 2018.
- J. J. Bird, A. Ekart, C. D. Buckingham, and D. R. Faria, “Mental emotional sentiment classification with an eeg-based brain-machine interface,” in The International Conference on Digital Image and Signal Processing (DISP’19), Springer, 2019.
- https://www.researchgate.net/publication/329403546_Mental_Emotional_Sentiment_Classification_with_an_EEG-based_Brain-machine_Interface
- https://www.researchgate.net/publication/335173767_A_Deep_Evolutionary_Approach_to_Bioinspired_Classifier_Optimisation_for_Brain-Machine_Interaction