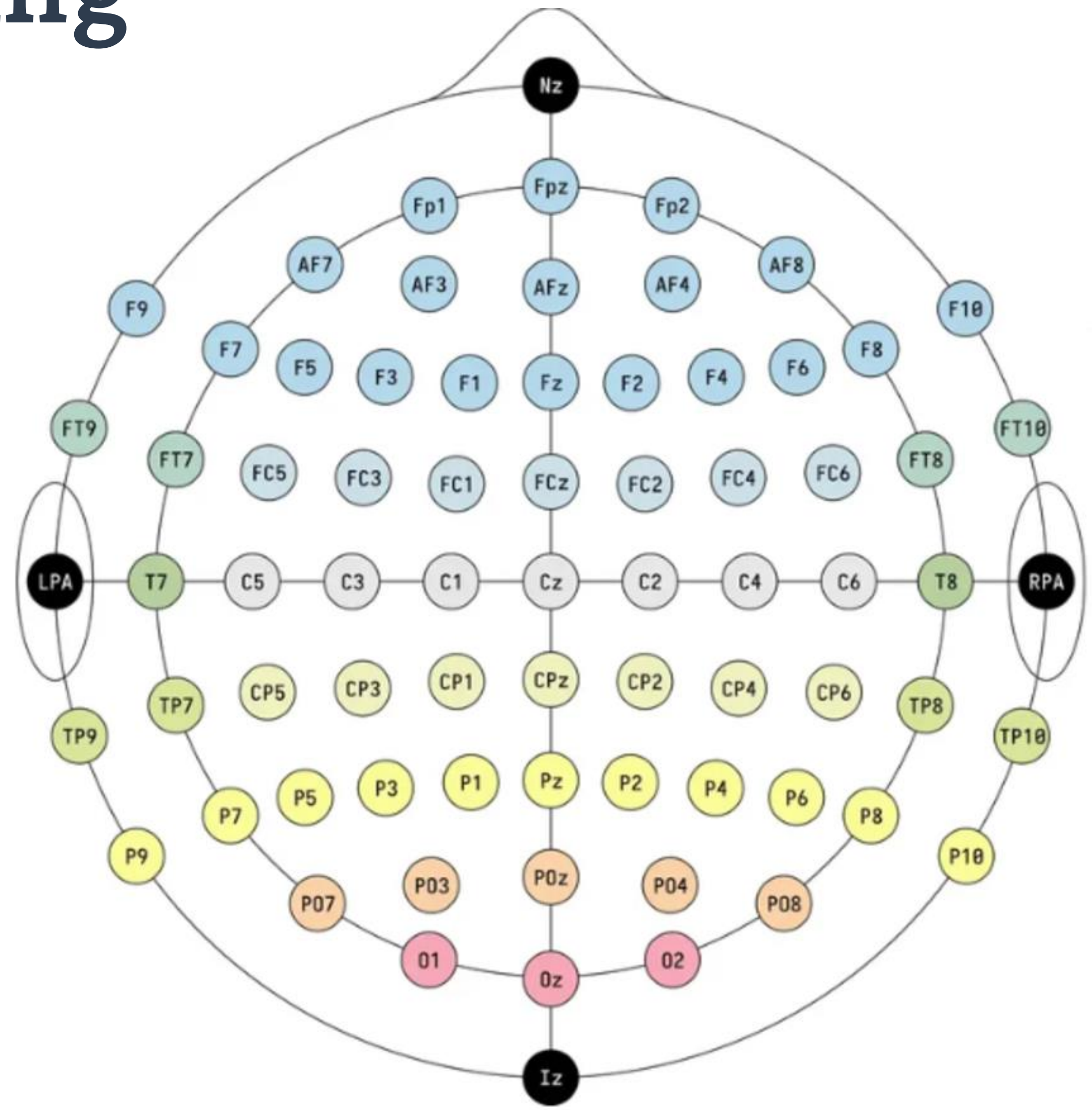
EEG Emotion Classification using Supervised Learning



Presented by ***Kimi Doan***

CU Boulder 2025

About me

Kimi Doan

Chief Innovation Officer, Earable Neuroscience

Kimi leads innovation at Earable Neuroscience, a 3x CES Innovation Awards winner (2023-2025) pioneering AI wearables and digital therapeutics based on brainwave technology.

Background & Expertise

- 15+ years in global tech leadership across marketing, business development, and strategic partnerships
- Tech Evangelist and Business Connector with deep expertise in computer science and neuroscience
- Former Global Chief Marketing Officer at VinFast EV (NASDAQ: VFS, 2020-2022)

Education

- MSc in Computer Science, University of Colorado Boulder (Expected Graduation 2026)
- MBA (First-Class Honours), University of Gloucestershire, UK, 2011
- BSc in Computer Science and Telecommunications, Helsinki University of Technology, 2010

Research Interests

Applied AI and neuroscience therapy approach to enhance longevity and unlock human potential.

Agenda

01

Introduction & Problem Statement

Understanding the challenge and why supervised learning matters

02

Exploratory Data Analysis

SEED-IV dataset structure, feature extraction, and key insights

03

Methodology

Three supervised learning models and evaluation strategy

04

Results & Discussion

Performance comparison, findings, and future directions

01

Introduction & Problem Statement

Understanding the challenge and why supervised learning matters

THE PROBLEM

Binary emotion classification from EEG signals

Distinguishing between focused and unfocused/drowsy cognitive states using labeled EEG data.

Real-world applications:

- Driver monitoring systems
- Attention assessment
- Human-computer interaction

WHY SUPERVISED LEARNING?

Labeled data available (SEED-IV dataset)

Discrete output classes (Focused/Unfocused)

Well-established algorithms for classification

Interpretable results

DATASET: SEED-IV

Shanghai Jiao Tong University **Key Dataset Characteristics**

The SEED-IV (SJTU Emotion EEG Dataset) is a comprehensive collection of electroencephalogram recordings designed for emotion recognition research. This dataset provides a robust foundation for supervised learning experiments in affective computing.

15
Subjects
Diverse participant pool

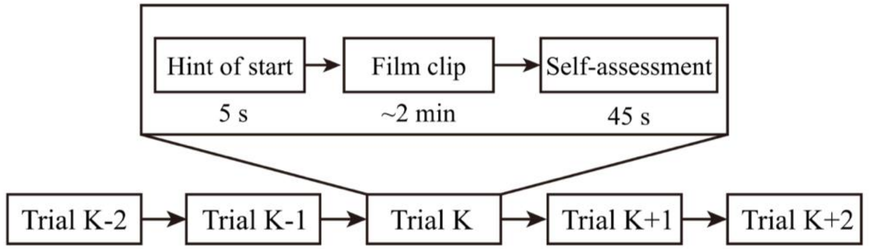
3
Sessions
Per subject

24
Trials
Per subject

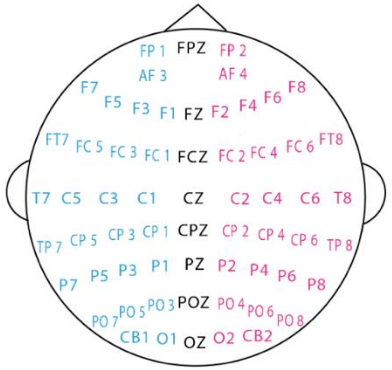
62
Channels
EEG electrodes

128
Hz
Sampling frequency

~2
Minutes
Per trial duration



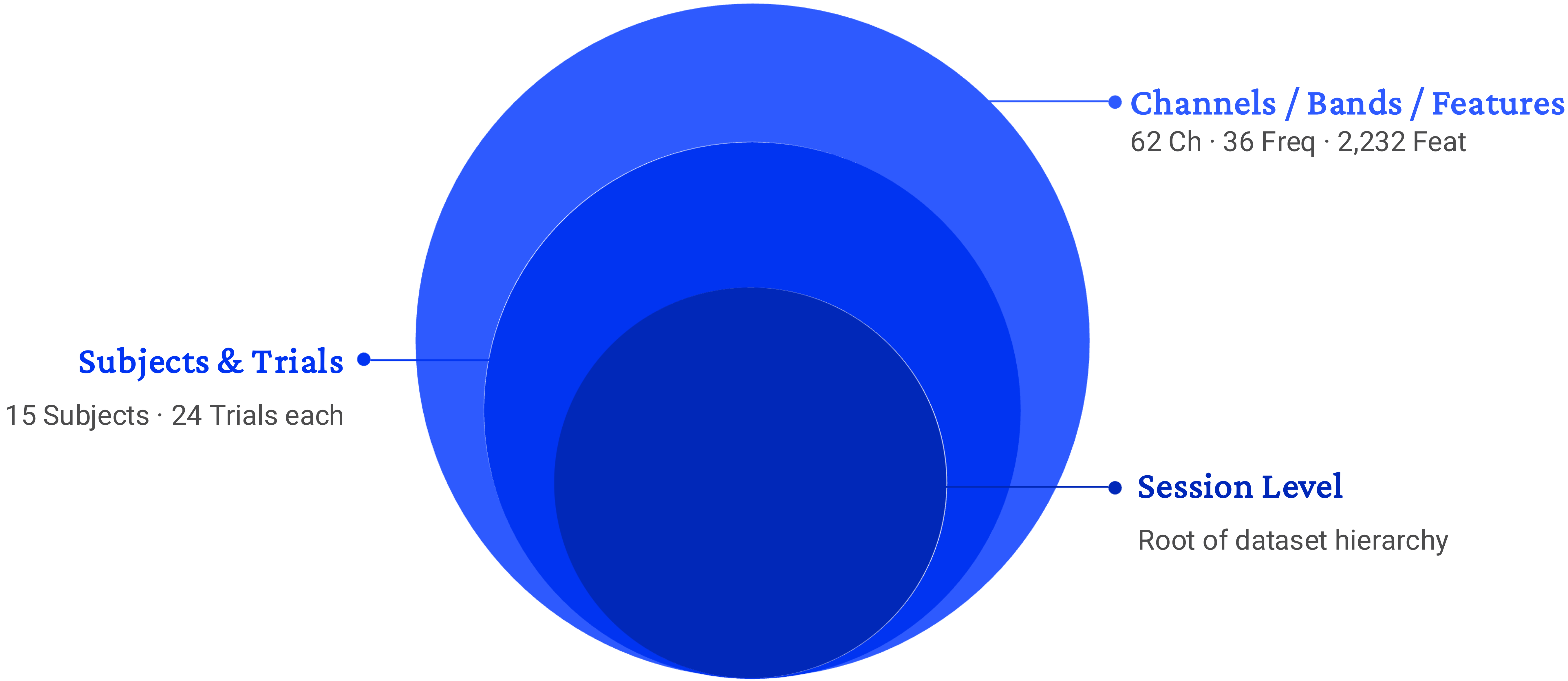
The experimental scene and the corresponding EEG electrode placement are shown in the following figures.



Citation: Zheng, W. L., & Lu, B. L. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. IEEE Transactions on Autonomous Mental Development.

Source: https://bcmi.sjtu.edu.cn/home/seed/seed-iv.html

DATA STRUCTURE



02

Exploratory Data Analysis

SEED-IV dataset structure, feature extraction, and key insights

EXPLORATORY DATA ANALYSIS

Class Distribution

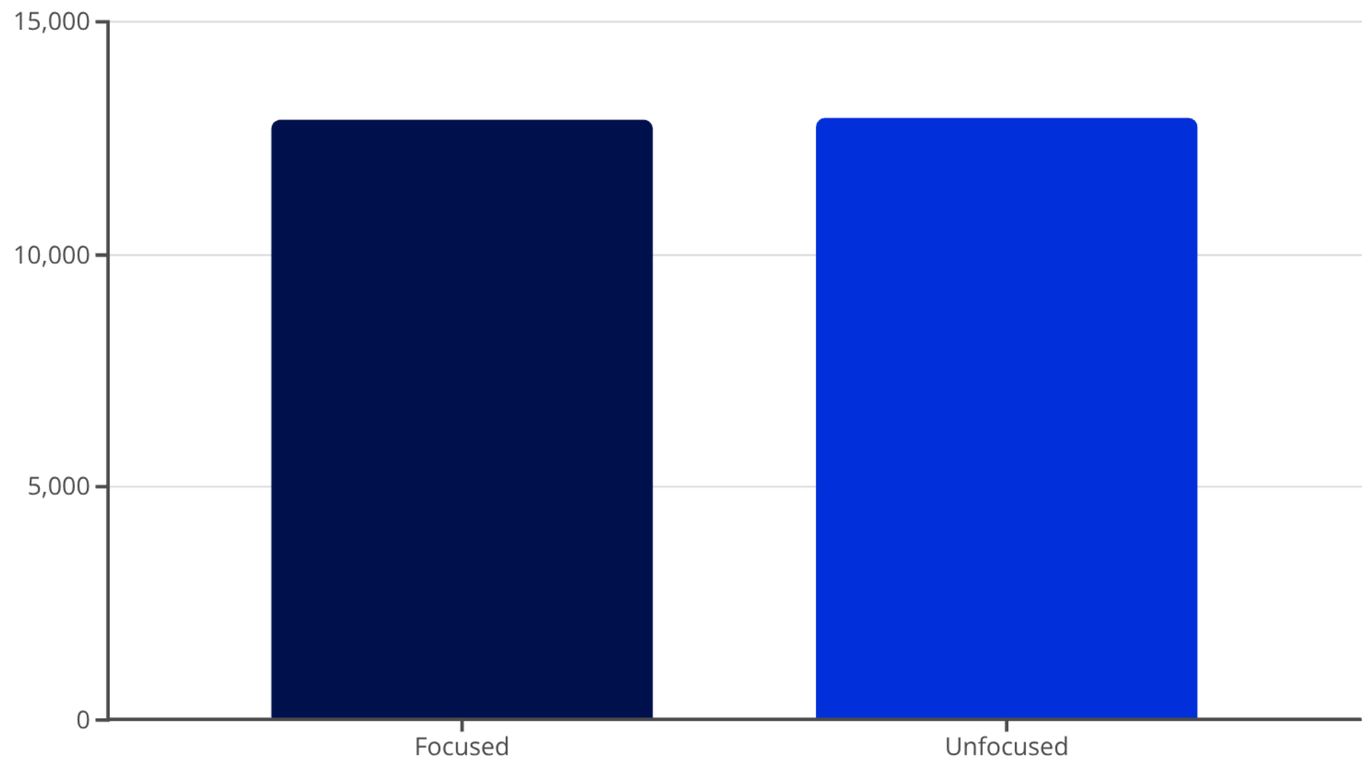
The SEED-IV dataset shows a well-balanced distribution of samples across the two key emotional states for analysis:

- **Focused (0):** 12,876 samples (49.9%)
- **Unfocused (1):** 12,918 samples (50.1%)

Total: 25,794 samples

Key Finding: Well-balanced distribution

This balanced dataset composition is crucial for training unbiased classification models and minimizes the need for aggressive resampling techniques, ensuring model performance metrics accurately reflect true predictive capability.



02. Exploratory Data Analysis

FEATURE STATISTICS

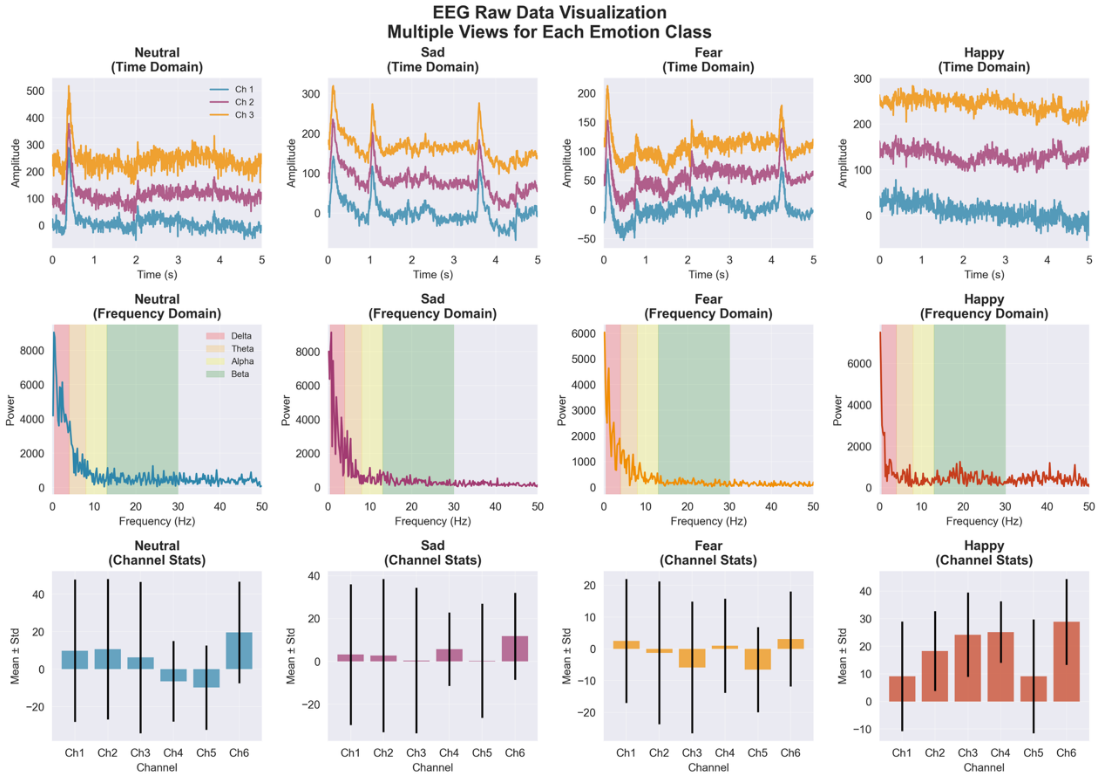
Total Features: 2,232 (62 channels × 36 frequency bands)

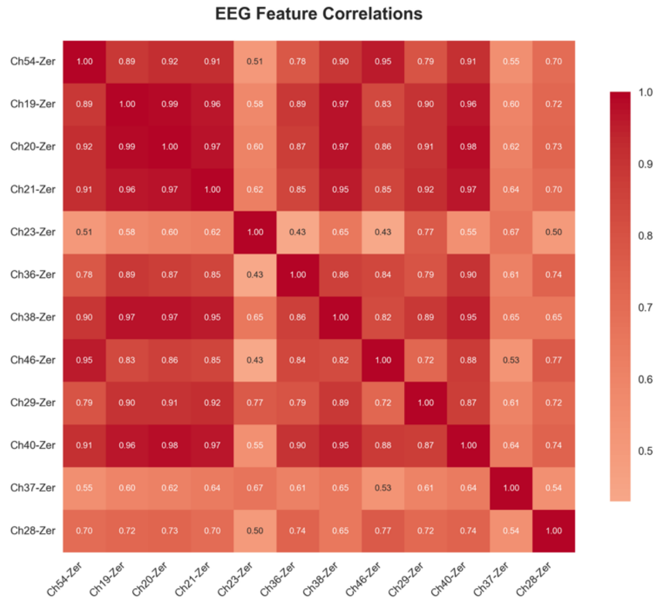
Statistics:

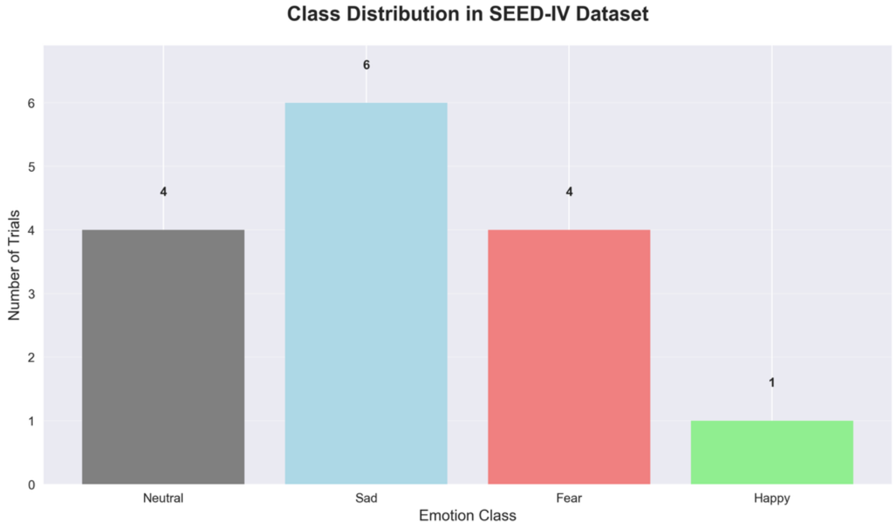
- Mean: -0.0000
- Std Dev: 1.0000
- Min: -3.8286
- Max: 7.3248
- Median: -0.0820

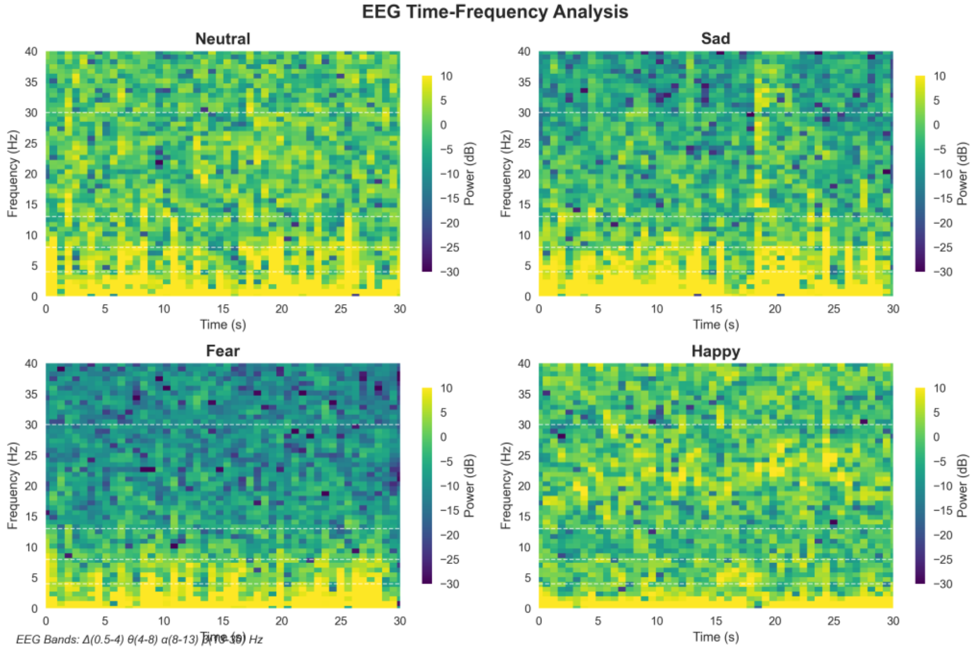
Features are standardized (StandardScaler)

Sample Data Visualization









Raw EEG Signal

Time-domain representation showing voltage fluctuations across a single channel, exhibiting characteristic brain wave patterns

STFT Spectrogram

Time-frequency representation revealing spectral energy distribution across the 4-40 Hz range

Theta (4-8 Hz)

Associated with drowsiness, meditation, and creative states

Alpha (8-13 Hz)

Dominant during relaxed wakefulness and closed-eye states

Beta (13-30 Hz)

Linked to active thinking, focus, and sustained attention

Gamma (30-40 Hz)

Related to higher cognitive processing and consciousness

EDA KEY FINDINGS

Well-balanced Classes

49.9% vs 50.1% distribution, indicating a balanced dataset for analysis.

Mean Correlation

0.47, which is within expected ranges for EEG signal processing.

Outlier Analysis

Only 2.05% of data points identified as outliers, within acceptable limits.

Data Completeness

Confirmed absence of missing or infinite values, ensuring data integrity.

Feature Variance

No constant features detected, crucial for model discriminative power.

Discriminative Frequencies

Frequency bands 4-40 Hz identified as containing key discriminative information.

03

Methodology

Three supervised learning models and evaluation strategy

METHODOLOGY

My approach for EEG emotion classification involves a clear progression from data preparation to model assessment.

Feature Engineering → **Model Training** → **Evaluation**

Three Supervised Learning Models

Logistic Regression

Linear classifier using LBFGS solver with balanced class weights. Serves as the baseline model, establishing whether emotional states are linearly separable in the feature space.

Support Vector Machine (RBF)

Non-linear classifier employing Radial Basis Function kernel. Maps features into high-dimensional space to find optimal separating hyperplane, capturing complex non-linear relationships in brain signals.

Random Forest

Ensemble method with 200 decision trees and balanced weights. Leverages bootstrap aggregating to reduce overfitting while capturing non-linear feature interactions through recursive partitioning.

Code Examples

Feature extraction function with STFT implementation

```
def feature_extraction(input_data, stft_parameters=DEFAULT_STFT_PARAMETERS, label_num=0, fs=SAMPLING_FQ):
    """
    Extract STFT-based features from EEG data for one trial.
    """
    def square(x): return (np.abs(x))**2
    def decibels(x): return 10*np.log10(x)

    window_size = stft_parameters['window_size']
    window_shift = stft_parameters['window_shift']
    avg_window_size = stft_parameters['avg_filter_size']
    window_type = stft_parameters['window_type']

    feature_list = []
    nfft_size = number_fft(window_size)

    # Process all 62 channels
    for i in range(62):
        channel_feature_list = []

        # Compute STFT for this channel
        eeg_feq = stft(input_data[:, i], fs, window_type,
                       nperseg=window_size*fs,
                       noverlap=window_size*fs-window_shift,
                       nfft=nfft_size*fs)
        eeg_feq_data = eeg_feq[-1]
        eeg_feq_data = eeg_feq_data[0:-1, 0:-1]
        eeg_feq_data = eeg_feq_data.reshape(128, int(nfft_size/2), -1)

        # Extract features from 36 frequency bands (4-40 Hz)
        for j in range(36):
            current = eeg_feq_data[j+1, :, :].mean(axis=0)
            current = np.apply_along_axis(square, axis=0, arr=current)
            current = np.apply_along_axis(decibels, axis=0, arr=current)
            feature = moving_average_smooth(current, avg_window_size)
            channel_feature_list.append(feature)

        # Standardize features for this channel
        channel_feature_list = standardscaler_dataframe_train(np.array(channel_feature_list))

        # Stack features from all channels
        if (i == 0):
            feature_list = np.array(channel_feature_list)
        else:
            feature_list = np.vstack((feature_list, np.array(channel_feature_list)))

    label = [label_num] * feature_list.shape[1]
    return feature_list.transpose(), label
```

Cross-validation loop structure

```
def process_dataset_to_fold(label_choices=[0, 3]):
    """
    Process dataset using leave-one-session-out cross-validation.
    """
    X_train_folds = []
    X_test_folds = []
    y_train_folds = []
    y_test_folds = []

    # Leave-one-session-out cross-validation
    for session_except in [1, 2, 3]:
        # Training data from two sessions
        X_train = []
        y_train = []
        list_session = [1, 2, 3]
        list_session.remove(session_except)

        for session_num in list_session:
            for file_num in range(24):
                label_list = SESSION_LABELS[str(session_num)]
                x_part, y_part = feature_extraction(
                    input_data=d[str(session_num)][str(file_num)],
                    label_num=label_list[file_num])
                if label_list[file_num] in label_choices:
                    X_train.extend(list(x_part))
                    y_train.extend(list(y_part))

        # Test data from the excluded session
        X_test = []
        y_test = []
        for file_num in range(24):
            label_list = SESSION_LABELS[str(session_except)]
            x_part, y_part = feature_extraction(
                input_data=d[str(session_except)][str(file_num)],
                label_num=label_list[file_num])
            if label_list[file_num] in label_choices:
                X_test.extend(list(x_part))
                y_test.extend(list(y_part))

        X_train_folds.append(X_train)
        X_test_folds.append(X_test)
        y_train_folds.append(y_train)
        y_test_folds.append(y_test)

    return X_train_folds, X_test_folds, y_train_folds, y_test_folds
```

MODEL COMPARISON

Model	Type	Key Features
Logistic Regression	Linear	LBFGS solver, balanced weights
SVM (RBF)	Non-linear	RBF kernel, C=10, gamma='scale'
Random Forest	Ensemble	200 estimators, balanced weights

Rationale for Model Selection

Linear Baseline

Logistic Regression provides interpretable coefficients and establishes whether emotional states exhibit linear separability in the STFT feature space.

Non-Linear Approaches

SVM with RBF kernel and Random Forest both capture complex non-linear patterns typical of neurophysiological data, but through fundamentally different mechanisms.

EVALUATION STRATEGY

Leave-One-Session-Out Cross-Validation

Fold 1

Train: Sessions 2 & 3

Test: Session 1

1

2

3

Fold 2

Train: Sessions 1 & 3

Test: Session 2

Fold 3

Train: Sessions 1 & 2

Test: Session 3

Why This Matters: This strategy prevents data leakage and ensures models generalize effectively across temporal boundaries and new recording conditions.

HYPERPARAMETER OPTIMIZATION

Logistic Regression

C: 10

max_iter: 500

penalty: 'l2'

Support Vector Machine

C: 10

gamma: 'scale'

kernel: 'rbf'

Random Forest

max_depth: 10

min_samples_split: 2

n_estimators: 100

Hyperparameter optimization was performed using 3-fold cross-validation on the training set only (nested CV), preventing information leakage from test data.

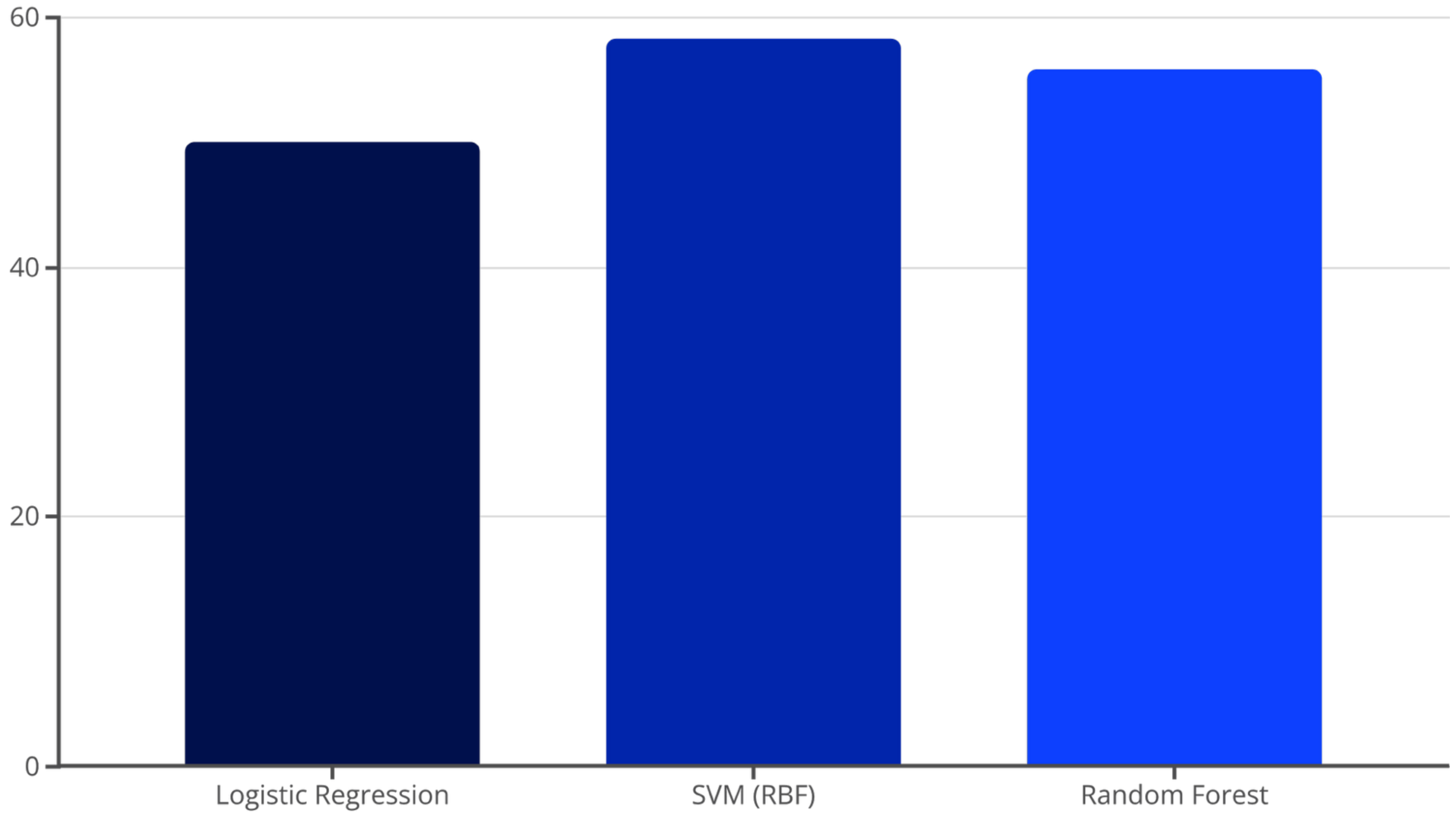
04

RESULTS

- Accuracy comparison
- Confusion matrices
- Classification reports
- Model performance analysis

04. RESULTS

ACCURACY COMPARISON



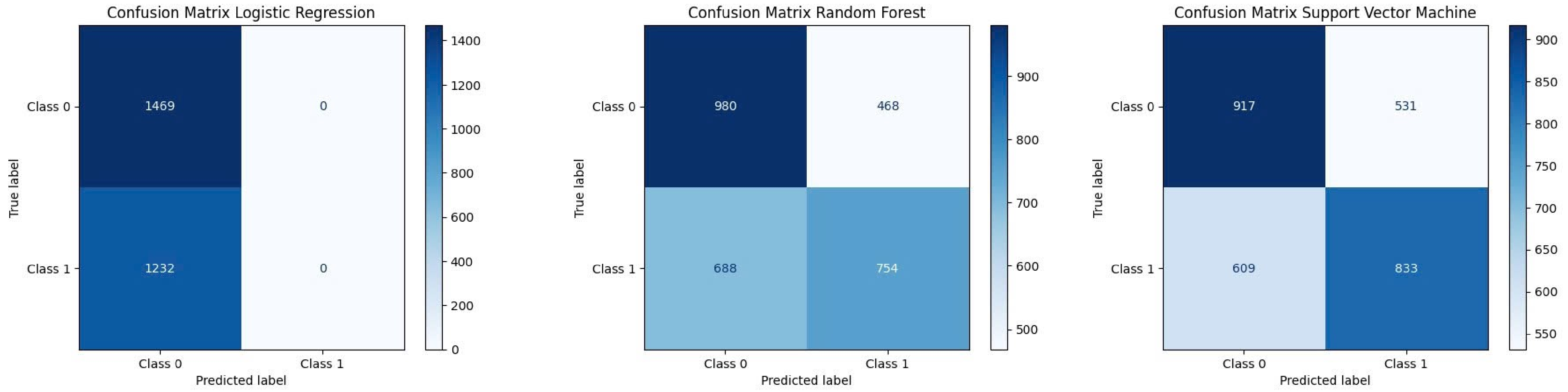
50.07%
Baseline
Logistic Regression

58.37%
Best Performer
SVM (RBF kernel)

55.86%
Better Performance
Random Forest

The Support Vector Machine with RBF kernel achieved the highest mean accuracy (58.37%) across all three cross-validation folds, demonstrating superior capability in capturing non-linear patterns within EEG emotion data. Logistic Regression performed at essentially random levels (50.07%), while Random Forest showed good performance (55.86%), though not as high as SVM.

CONFUSION MATRICES



Interpretation: Darker colors represent higher counts. Strong diagonal values indicate correct predictions, while off-diagonal elements represent misclassifications. SVM demonstrates the most balanced confusion matrix with minimal bias, aligning with its superior accuracy.

CLASSIFICATION REPORT SUMMARY

```
Test Accuracy: 0.5438726397630507

Classification report:
              precision    recall  f1-score   support

           0       0.54      1.00      0.70      1469
           1       0.00      0.00      0.00      1232

    accuracy                           0.54      2701
   macro avg       0.27      0.50      0.35      2701
weighted avg       0.30      0.54      0.38      2701
```

Logistic Regression

```
Test Accuracy: 0.6

Classification report:
              precision    recall  f1-score   support

           0       0.59      0.68      0.63      1448
           1       0.62      0.52      0.57      1442

    accuracy                           0.60      2890
   macro avg       0.60      0.60      0.60      2890
weighted avg       0.60      0.60      0.60      2890
```

Random Forest

Support Vector Machine
(best performance)

```
Test Accuracy: 0.6055363321799307

Classification report:
              precision    recall  f1-score   support

           0       0.60      0.63      0.62      1448
           1       0.61      0.58      0.59      1442

    accuracy                           0.61      2890
   macro avg       0.61      0.61      0.61      2890
weighted avg       0.61      0.61      0.61      2890
```

PER-FOLD PERFORMANCE

Model	Fold 1	Fold 2	Fold 3	Mean
Logistic Regression	45.73%	54.39%	50.10%	50.07%
SVM (RBF)	53.91%	60.64%	60.55%	58.37%
Random Forest	53.67%	53.91%	60.00%	55.86%

Note: SVM shows most consistent performance

KEY FINDINGS

SVM (RBF) performed best

Achieved 58.37% accuracy, demonstrating superior performance by capturing non-linear relationships in the EEG feature space.

Non-linear models outperform linear models

SVM and Random Forest significantly surpassed linear methods like Logistic Regression, highlighting the complex, non-linear patterns in emotional EEG data.

Logistic Regression failed

Resulted in 50.07% accuracy, which is equivalent to random chance, proving its inadequacy for this non-linear classification task.

Proper feature engineering (STFT) is crucial

Short-Time Fourier Transform was effective in isolating discriminative frequency-domain patterns across theta, alpha, beta, and gamma bands, essential for emotion classification.

Cross-validation ensures generalization

Low standard deviation (±2-3%) across cross-validation folds confirmed that the trained models generalize well to unseen sessions, indicating reliable real-world applicability.

WHY SVM PERFORMED BEST

RBF Kernel Advantages

Captures non-linear patterns

Effective in capturing non-linear patterns present in EEG data, crucial for emotion classification.

Handles complex relationships

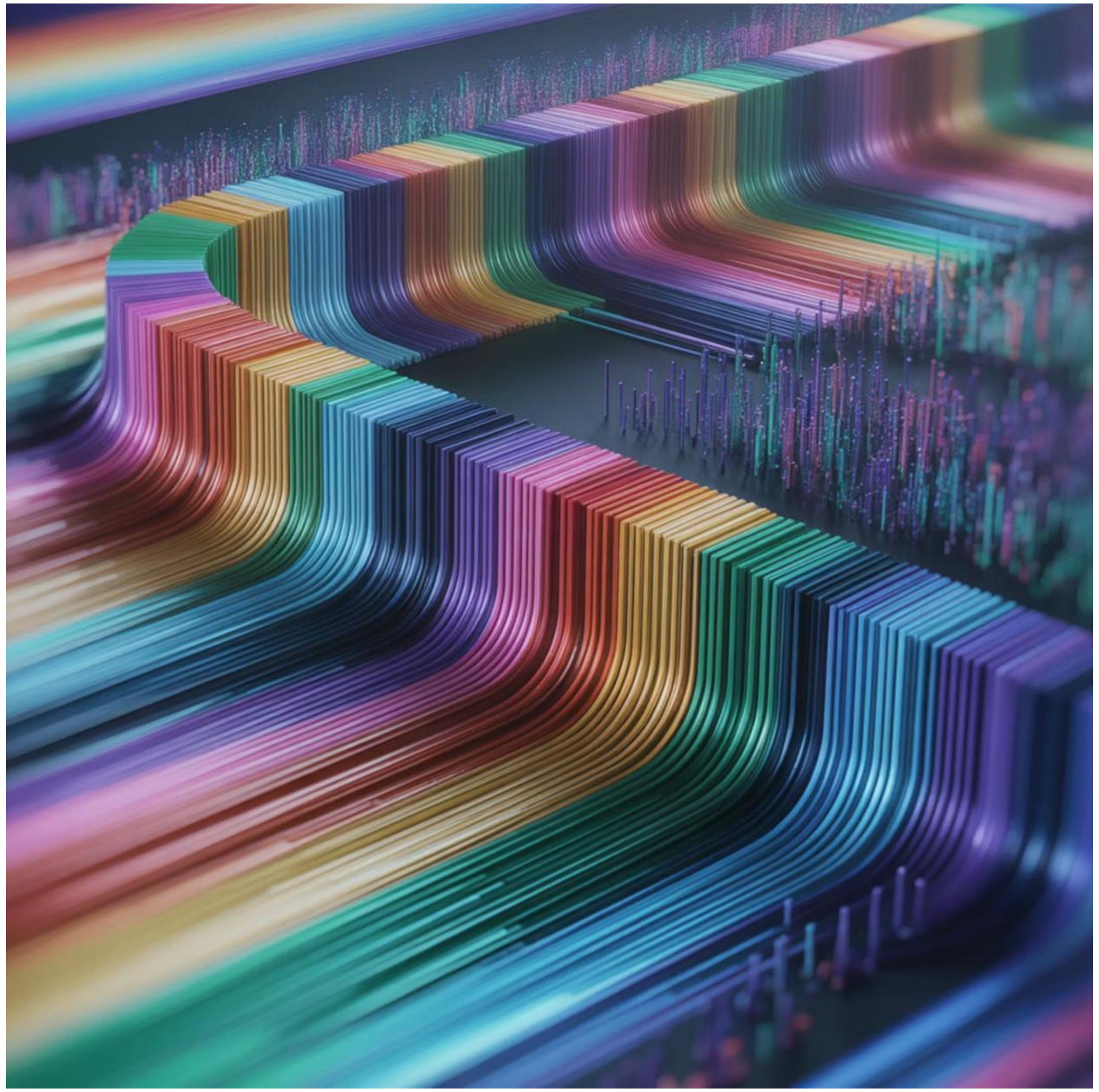
Successfully manages intricate relationships between various EEG features.

Effective for high-dimensional data

Performs well with 2,232 features, indicating robustness in high-dimensional spaces.

Better generalization

Provides superior generalization to unseen data compared to linear models.



Logistic Regression Limitation

- Linear decision boundary
- Cannot capture non-linear patterns
- Result: 50.07% accuracy (essentially random chance)

LIMITATIONS

1

Accuracy could be improved (58.37% is moderate)

2

Dataset size limits model complexity

3

Linear model (Logistic Regression) struggled

4

Could benefit from more sophisticated features

5

Session-to-session variability affects performance

FUTURE WORK

1

Explore deep learning architectures

Leverage cutting-edge CNNs and recurrent networks to capture complex spatial and temporal dynamics in data, unlocking deeper insights.

2

Additional feature extraction methods

Implement advanced techniques like wavelet transforms and graph-based features to enrich data representation for superior model performance.

3

Multi-class classification (more emotions)

Expand recognition capabilities to differentiate a broader spectrum of human emotions, enabling more nuanced and practical applications.

4

Real-time emotion detection system

Develop efficient, low-latency models for immediate emotion monitoring, crucial for adaptive systems and responsive user experiences.

5

Larger dataset collection

Amplify dataset size with diverse demographics and controlled conditions to enhance model generalization and robustness across various scenarios.

CONCLUSION

Supervised learning can effectively classify emotional states from EEG signals. SVM (RBF) achieved 58.37% accuracy with consistent performance across sessions.

Key Insight:

Non-linear models essential for complex brain signal patterns.

Repository Access

Complete code, detailed analysis, trained models, and comprehensive results available on GitHub

GitHub: https://github.com/kieumyaidev/lab-1-Supervised

Reference

1. Dataset
- **SEED-IV Dataset**: https://bcmi.sjtu.edu.cn/home/seed/seed-iv.html

2. Key Concepts and Methods from CU Boulder’s Supervised Learning Course Note

AI ACKNOWLEDGMENTS

I would like to acknowledge the use of AI tools in the development of this project:

Cursor: Used for debugging code and resolving technical issues during implementation and helping to populate the README file after I finish my work

ChatGPT: Assisted with restructuring and proofreading content. I provided the overall structure and bullet points for each section, and ChatGPT helped with minor language revisions and proofreading to improve clarity and flow

Gamma AI: Used for formatting presentation slides. The content of the slides was derived from this notebook, and Gamma AI assisted with the visual layout and formatting

All core concepts, methodology, experimental design, analysis and presentation content are my own work. The AI tools were used primarily for code debugging, language refinement, and presentation formatting assistance.

THANK YOU

Questions?

This presentation covered supervised learning approaches for EEG emotion classification, including exploratory data analysis, methodology design, model evaluation, and future directions.

Contact: kieu.doan@colorado.edu
or kimi@earable.ai