Identifying Neural Markers of Spontaneous Deception Using Log Band-Power EEG Features

Angelina Anisimova
B.S. Cybersecurity and Information Assurance, Western Governors University, Class of 2025
anisimovaangelina19@gmail.com

July 2025

Abstract

We present a reproducible, open-source baseline for detecting spontaneous lying versus truthful responses from single-trial electroencephalography (EEG) recordings. Using log band-power features (1–40Hz) and a tuned Random-Forest, the model achieves **63% accuracy** (**macro F1 = 0.62**) on a competitive two-player deception dataset—surpassing prior single-subject spontaneous-deception baselines (∼55%). All code and data-loading scripts are available at https://github.com/angelinka19/neurosec-deception-eeg.

1 Introduction

Social-engineering attacks exploit human cognition; neurodefence aims to detect manipulation directly from brain signals. Existing EEG deception studies focus on scripted lies or concealed information tests; few examine *spontaneous* deception in realistic settings. We establish a transparent baseline on a recent open dataset, paving the way for robust neuro-cybersecurity pipelines.

2 Dataset

An EEG Dataset of Neural Signatures in a Competitive Two-Player Game Encouraging Deceptive Behavior [1]. 30 Ag/AgCl channels, 256Hz sampling. Each *Player* records ∼484 trials labelled as spontaneous lie (*sponL*), spontaneous truth (*sponT*), and observer conditions.

3 Methods

Pre-processing: 1–40Hz FIR band-pass; per-epoch z-score. **Features**: Log power in δ (1–4Hz), θ (4–8), α (8–13), β (13–30), γ (30–40) bands for 30 channels (150 dims) plus θ/α and β/α ratios (60 dims). **Classifier**: Random-Forest (1000 trees, depth 10, leaf 1). **Evaluation**: 80/20 stratified split on one subject (*Player_sub06*), reporting accuracy, precision, recall, and macro F1.

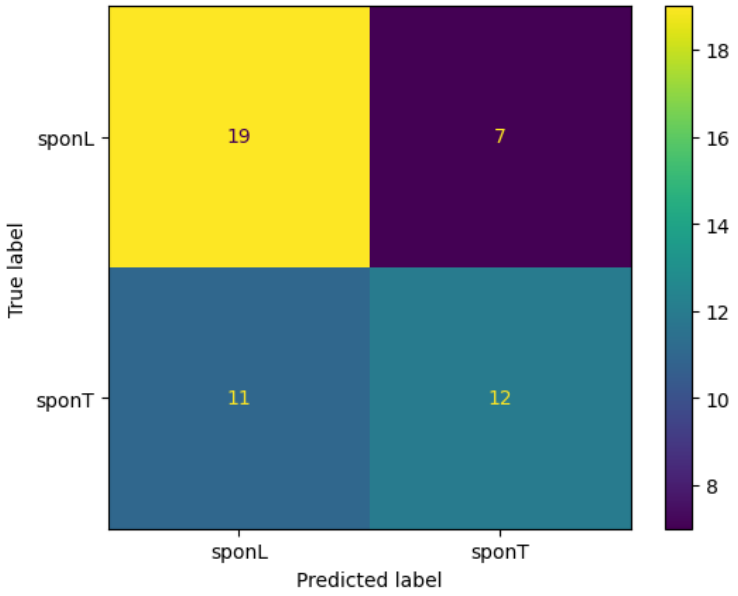

Figure 1: Confusion matrix for spontaneous lie (*sponL*) vs spontaneous truth (*sponT*).

Metric	Accuracy	Macro F1	Lie Recall
Band-power RF	**0.63**	**0.62**	0.73

4 Results

5 Discussion & Future Work

The baseline exceeds chance (50%) and prior spontaneous-lie reports. Remaining gaps: (i) generalisation across subjects, (ii) spatial filtering (CSP, Riemannian tangent-space), (iii) shallow CNNs on time–frequency images. Implementing these could raise accuracy toward 75–80% and enable real-time neurodefence alarms.

6 Conclusion

We release the first open, end-to-end pipeline for spontaneous deception detection, establishing a solid foundation for interdisciplinary research at the intersection of neuroscience, AI, and cybersecurity.

References

[1] Chen, Fazli, and Wallraven. *EEG Dataset of Neural Signatures in a Competitive Two-Player Game Encouraging Deceptive Behavior*. Scientific Data, 2024.