Ablation Study Experiment - Validation Metodology

To ensure model clinical robustness, a comparative study focused on the **Data Splitting Strategy** was conducted. The objective was to verify wheter the model was learning generalizable patterns of Burnout or merely memorizing the patient identities (*Data Leakage*).

Test A: Subject Isolation - Proposed

- **Methodology:** Sequential division where the last 20% of patients (previously unseen) are separeted exclusively for testing.
- **Error (Loss) Average:** 0.2122.
- **Average Accuracy:** 92.44% (σ = 0.97).
- **Observation:** Represents real performance in a clinical scenario (new patients).
- **Log:**

```
Initiating 5 Trainings Sessions.

Running 1/5 Training Session.
Registered Loss: 0.2029 | Acc: 93.75

Running 2/5 Training Session.
Registered Loss: 0.2095 | Acc: 91.71

Running 3/5 Training Session.
Registered Loss: 0.2042 | Acc: 92.66

Running 4/5 Training Session.
Registered Loss: 0.2298 | Acc: 91.03

Running 5/5 Training Session.
Registered Loss: 0.2144 | Acc: 93.07

================================================
Final Results (5 Sessions)
================================================
LOSS     -> Mean: 0.2122  | Std Dev: 0.0097
ACCURACY -> Mean: 92.44% | Std Dev: 0.97
------------------------------------------------
Raw Losses: [0.2029, 0.2095, 0.2042, 0.2298, 0.2144]
Raw Accs:   [93.75, 91.71, 92.66, 91.03, 93.07]
================================================
```

Test B: Without the Filter (Raw Data)

- **Methodology:** THe entire dataset was shuffle before the division. Time Epochs from the same patient appear in both training and test.
- **Error (Loss) Average:** 0.2752.

- **Average Accuracy:** 89.06% (σ = 0.40).
- **Observation:** Simulates a data leakage scenario by identity.
- **Log:**

```
Initiating 5 Trainings Sessions.

Running 1/5 Training Session.
Registered Loss: 0.2662 | Acc: 89.4

Running 2/5 Training Session.
Registered Loss: 0.2946 | Acc: 88.32

Running 3/5 Training Session.
Registered Loss: 0.2684 | Acc: 89.19

Running 4/5 Training Session.
Registered Loss: 0.2656 | Acc: 88.99

Running 5/5 Training Session.
Registered Loss: 0.2814 | Acc: 89.4

================================================
Final Results (5 Sessions)
================================================
LOSS     -> Mean: 0.2752  | Std Dev: 0.0113
ACCURACY -> Mean: 89.06% | Std Dev: 0.40
------------------------------------------------
Raw Losses: [0.2662, 0.2946, 0.2684, 0.2656, 0.2814]
Raw Accs:   [89.4, 88.32, 89.19, 88.99, 89.4]
================================================
```

Conclusion and Analysis

1. **Immunity to Identity Leakage:** Counterintuitively, the model performed slightly better in the isolated scenario (92%) than the mixed scenario (89%). In traditional neural networks, *Random Split* tends to inflate accuracy due to memorization of the Subject's Individual characteristics.

The fact that accuracy did not increase in Random Split demonstrates that the *Prototypical Network* (Few-Shot) architecture designed in this study has high **Identity Invariance**. The model focused on learning the neural signature of Burnout, ignoring the unique characteristics of each brain.

2. **Data Quality:** The superiority of test A suggests that the patients selected for the test set (the last ones of STEW Dataset) had clearer and more defined Burnout markers than the overall average of the population mixed in test B.

3. **Final Descision:** The final model will strictly adopt **Subject Isolation (test A)**. Althgouh the results are numerically close, this methodology is the only that garantees scientific validity for the real-world BCI applications, where the system must work for users never seen before.