# Weekly Work Report 5/30/2025

**This Week:**

* **Minimum pupil data confidence levels evaluated: 0.6 and 0.8.**

### Data Processing Steps for Minimum Confidence Threshold Exploration

1. **Dataset Segmentation by Confidence Thresholds**
   * Pupil data were divided into three separate datasets based on minimum confidence levels: **Conf\_0.6**, **Conf\_0.7**, and **Conf\_0.8**.
2. **Pupil Data Preprocessing**
   * **Step 1: Signal Filtering**
     + Applied a low-pass Butterworth filter (< 5 Hz) to remove high-frequency noise.
   * **Step 2: Blink Artifact Removal**
     + Detected and interpolated blinks or spurious artifacts (e.g., drops to zero or spikes).
     + Baseline()
3. **Pupil Dilation Computation**
   * Calculated pupil dilation using a sliding window approach:
     + **Window size**: 1 second
     + **Step size**: 0.2 seconds

dilation = window\_data['pupil\_size\_processed'].max() - window\_data['pupil\_size\_processed'].min()

1. **Label Assignment (Aha vs. Non-Aha)**
   * Labeled pupil dilation values > 0.5 as **Aha moments** and others as **Non-Aha**.
2. **EEG-Pupil Label Alignment**
   * Synchronized EEG data with pupil data timestamps and assigned corresponding Aha/Non-Aha labels.
3. **Data Balancing 🡪 Balanced vs Imbalanced data (k-fold?) because in real word if always imbalanced len(non-aha) > len(aha) – increase subjects**
   * Balanced the EEG dataset to contain an equal number of Aha and Non-Aha samples.
4. **Classification**
   * Used a **Random Forest classifier** to perform subject-specific EEG classification.



### Literature review for key Eye Metrics & Cognitive States (Attention, Impasse, Aha! / Insight)

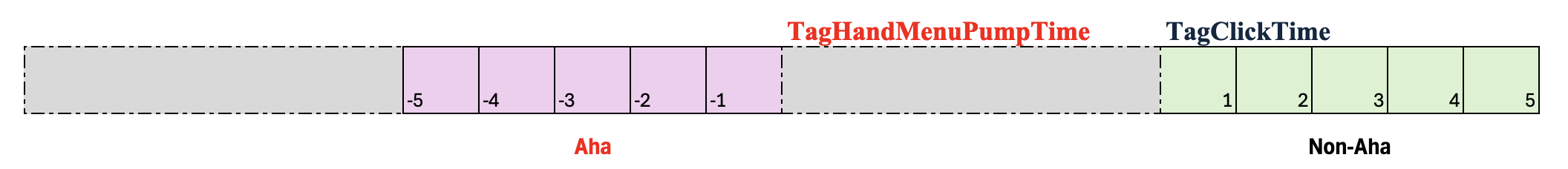
**A screenshot of a white box with black text

AI-generated content may be incorrect.**

**A screenshot of a white page

AI-generated content may be incorrect.**

**Aha! Moment Label Setting Exploration**

****

**A table of numbers and symbols

AI-generated content may be incorrect.**

**Q:** I aim to classify cognitive states such as **Attention, Impasse, and Aha! Moments,** or to explore **transitions between these states**. What is the best approach to label the data—should I rely on behavioral data, neural signals (e.g., EEG, pupil dilation), or a combination of both 🡪 raise hand + Pupil?

Blink – Aha: blink events after Dilation peak.

Rules for labeling: raise hand, blinks, dilation, Top K.

**Aha!**

* hand raise
* hand raise + pupil

**Impasse**

**Attention**

Guess, Tag 🡪 some subjects may just raise hand then guess to tag

Time delay 🡪 window

A screenshot of a computer

AI-generated content may be incorrect.