Predicting Individual Learning Trajectories in Mice via Early Behavior

Neta Thee & Pratit Kandel

Abstract

Understanding behavioral patterns in reinforcement learning paradigms is essential for studying animal learning mechanisms. In this project, we analyzed licking behavior of mice in a Pavlovian conditioning task. We trained a machine learning model on first-day licking features to predict each mouse's learning speed, defined by the number of days required to reach the learning criterion. While prediction accuracy was modest due to the small dataset, our analysis revealed structured changes in licking patterns and individual differences in learning trajectories. These results demonstrate the utility of data-driven approaches for quantifying and predicting behavioral strategies in animal learning.

Introduction

Reward learning is a fundamental process through which animals adapt their behavior based on the association between cues and outcomes. Initially, animals exhibit unconditioned responses but as learning advances over subsequent days, they begin to anticipate rewards based on predictive cues, resulting in a shift in licking times (Heffley & Hull, 2019). Anticipatory licking prior to reward indicates that the animal has learned to predict reward timing based on sensory cues. Data and direct observation revealed substantial individual variability in licking behavior from the Day 1. Some mice exhibited high exploratory behavior, licking frequently before any association was formed, while others licked sparsely in early trials. This variability led us to hypothesize that early licking patterns may predict the rate of learning.

Data Description

The dataset includes behavioral recordings of 18 mice performing a classical conditioning task. For each mouse, data from 5 days of training were analyzed; each day consisting of 200 trials, with event timestamps recorded in seconds relative to session start. Logged events include trial start and end, tone onset, reward delivery and licks. In each trial, a tone is played 3–5 seconds after trial onset, followed by a sucrose water reward delivered 500 ms later.

Methods

Feature Table: We first constructed a feature matrix for each mouse and for each day, quantifying licking behavior in relation to key events within each trial relative to tone and reward onset (*Figure 1*). For each bin, we calculated the average number of licks across all trials, resulting in a 13 D feature vector per mouse per day, with each 250 ms bin shifted by 10 ms after tone/reward onset to ensure licks reflecting the current trial are captured.

PCA and Clustering: To compare licking patterns among mice across days, we applied PCA to reduce dimensionality (2 D) and retain the first two components. K-means clustering (k = 3) was then used on the PCA-transformed data to identify groups of mice with similar licking profiles, capturing key behavioral strategies (e.g., anticipatory, delayed or minimal licking) while keeping the results interpretable.

Learning Trajectory Analysis: We computed the Euclidean distance in PCA space based on trajectory from Day 1 to Day 5 for each animal, using the first two principal components derived from binned behavioral data. This distance quantifies the magnitude of change in multivariate behavioral patterns over time.

Random Forest (RF) Model: RF performs well on small datasets and handles feature interactions robustly. It builds an ensemble of decision trees, each trained on random data subsets and aggregates their predictions to improve accuracy and reduce overfitting. We trained an RF classifier on Day 1 licking features to predict learning speed (fast = 1,2; normal = 3, 4; slow = 5), defined on the first day where a mouse showed a ≥ 0.5 difference in licking rate between control and prediction-to-reward bins. Features were scaled and reduced with PCA (4 components) and the model used 50 trees with max depth of 3. Performance was evaluated via 4-fold stratified cross-validation, reporting mean accuracy, standard deviation, precision, recall, F1-score and a confusion matrix. Feature importance was computed on the original features before PCA to identify which Day 1 behaviors contributed the most to the classification.

Results

The PCA plot (*Figure 2*) depicts the positions of individual animals in principal component space on Day 1 and Day 5, with dashed lines showing the trajectories across days. Clustering of ΔPC1 and ΔPC2 from Day 1 to Day 5 revealed three distinct groups of animals; most animals displayed positive ΔPC1 values, while a smaller subset had large negative ΔPC1 (*Figure 3*). Learning trajectory lengths varied substantially across animals, ranging from ~2 to over 8 units in PCA space, as shown in *Figure 4*. The RF model was able to identify learning speed with moderate overall accuracy (Accuracy = 0.725 ± 0.075). The model completely failed to identify fast and slow learners (Precision = 0.00, Recall = 0.00 for both groups), classifying all such cases as normal. For normal learners, precision was relatively high (0.72) and recall was perfect (1.00), meaning the model successfully detected all normal learners but also misclassified all other groups as normal. The confusion matrix reflects this strong bias towards predicting the normal category, at an expense of correctly identifying the extreme categories (*Figure 5*). The most important features for predicting learning speed on Day 1 (before PCA) were found to be Bin10, Bin4, Bin2 and Bin8, while the features like Bin3, Control and Reward contributed the least (*Figure 6*).

Discussion

Some animals showed minimal displacement across days, while others exhibited large shifts, reflecting individual differences in learning trajectories. Clustering revealed three groups corresponding to distinct behavioral profiles (anticipatory, delayed, minimal licking). The positive ΔPC1 values for most animals suggest a common direction of change during training, whereas large negative ΔPC1 in a subset indicates an opposite pattern. The RF model partially predicted learning speed but was strongly limited by the small, imbalanced dataset, classifying all fast and slow learners as normal (4 fast, 1 slow, 13 normal). This bias led to perfect recall for normal learners but complete failure to detect the extremes. The most important features corresponded to time points with notable behavioral variation, supporting that early licking patterns can inform learning speed. Overall, early patterns are predictive, but larger, more balanced datasets and temporal feature engineering are needed to improve accuracy and capture individual differences more effectively.

Bibliography

Heffley, W., & Hull, C. (2019). Classical conditioning drives learned reward prediction signals in climbing fibers across the lateral cerebellum. *elife*, *8*, e46764.

Figures

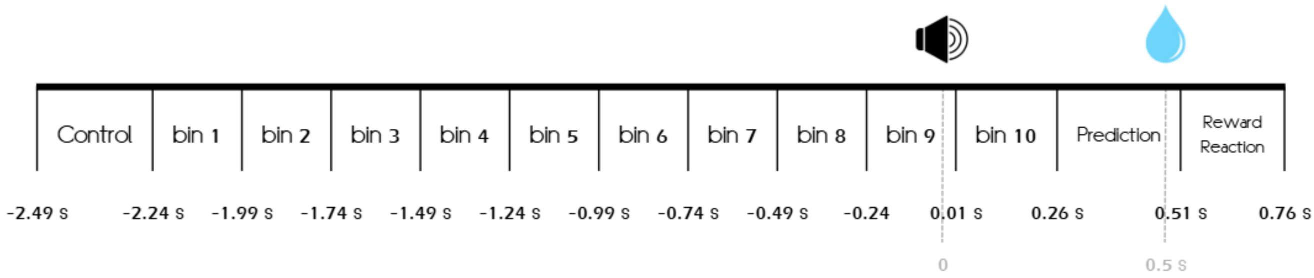

Figure 1: Bin timeline.

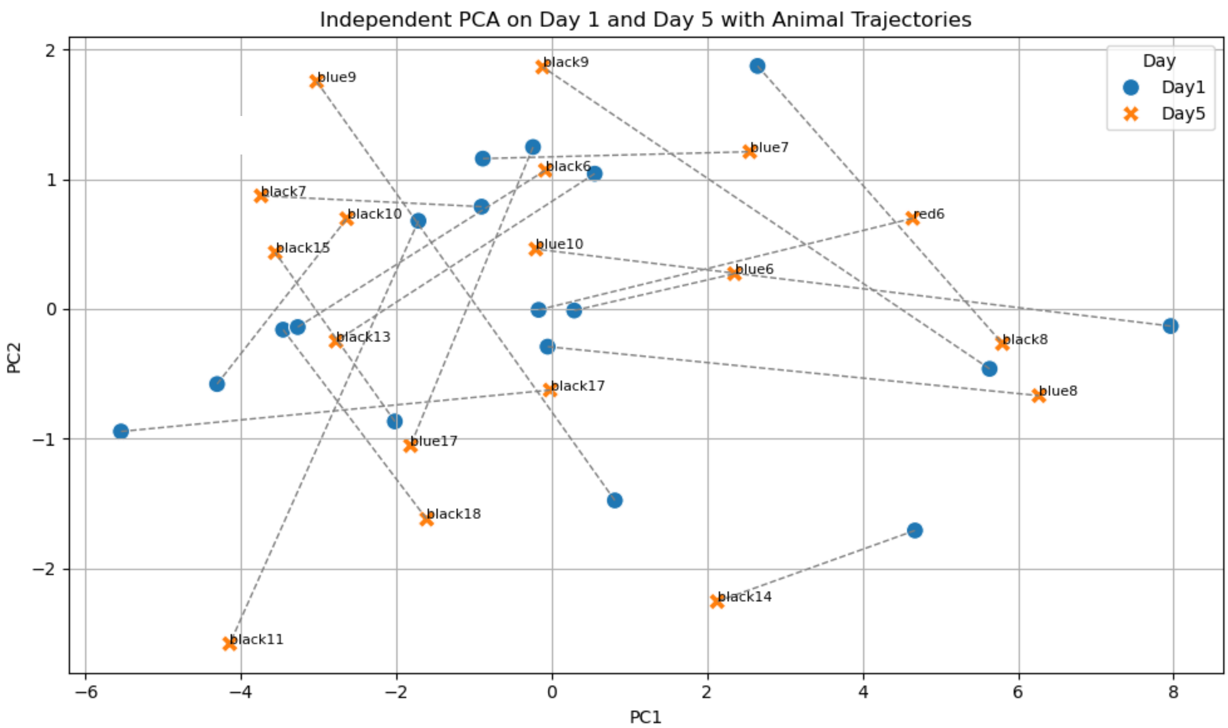

Figure 2: PCA trajectories from Day 1 to Day 5 for individual animals, showing shifts in principal component space.

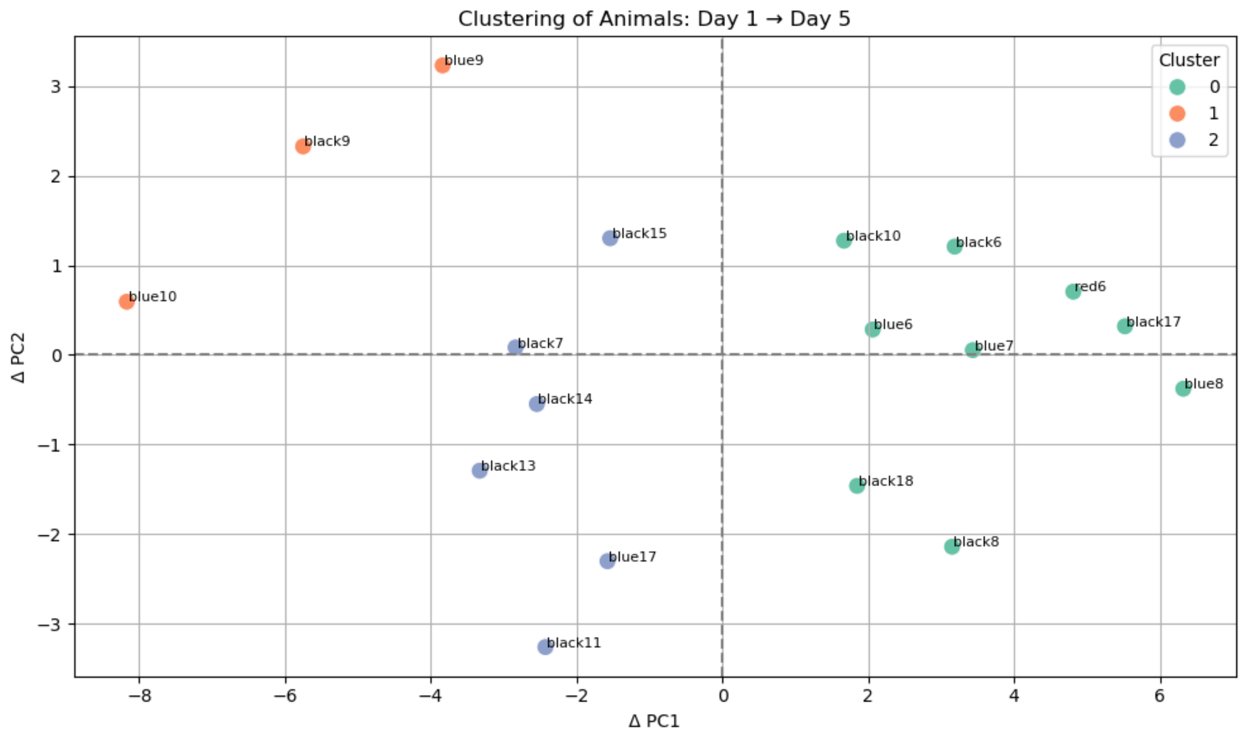

Figure 3: Animal PCA shifts from Day 1 to Day 5 reveal three distinct change patterns.

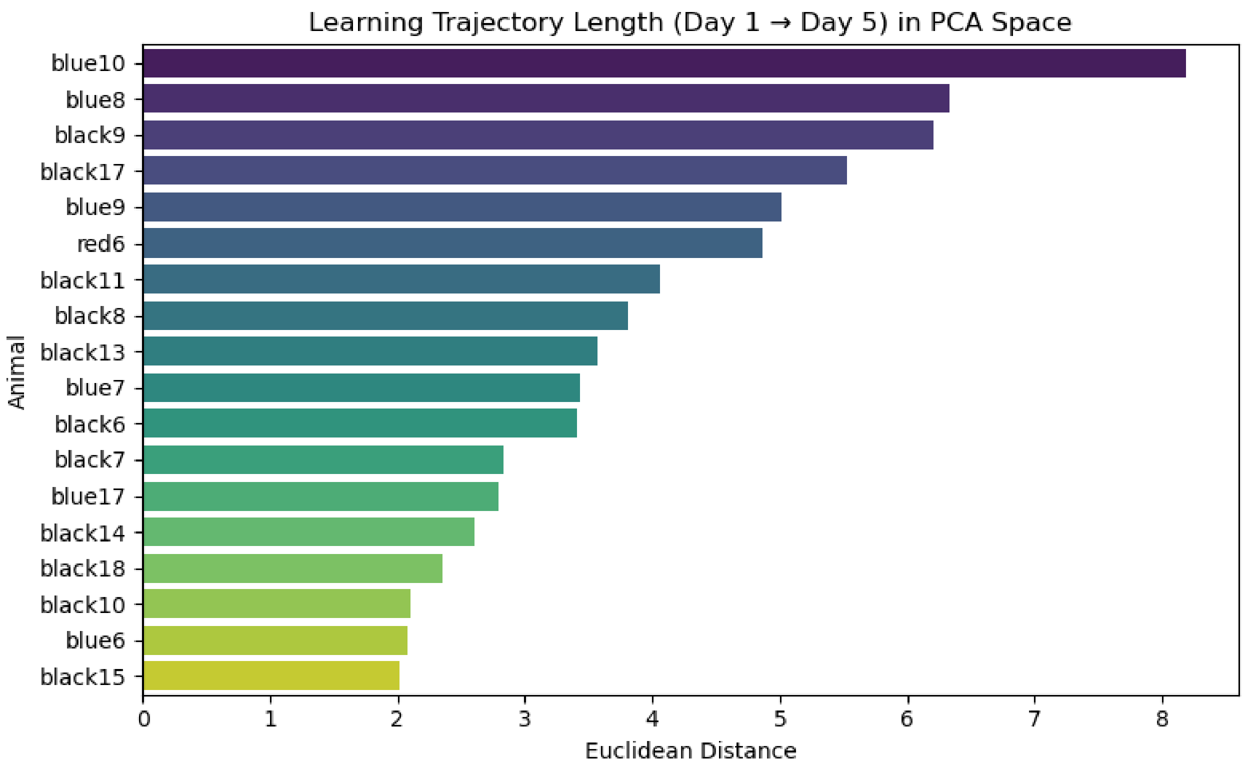

Figure 4: Euclidean distances in PCA space showing behavioral change from Day 1 to Day 5 for each animal.

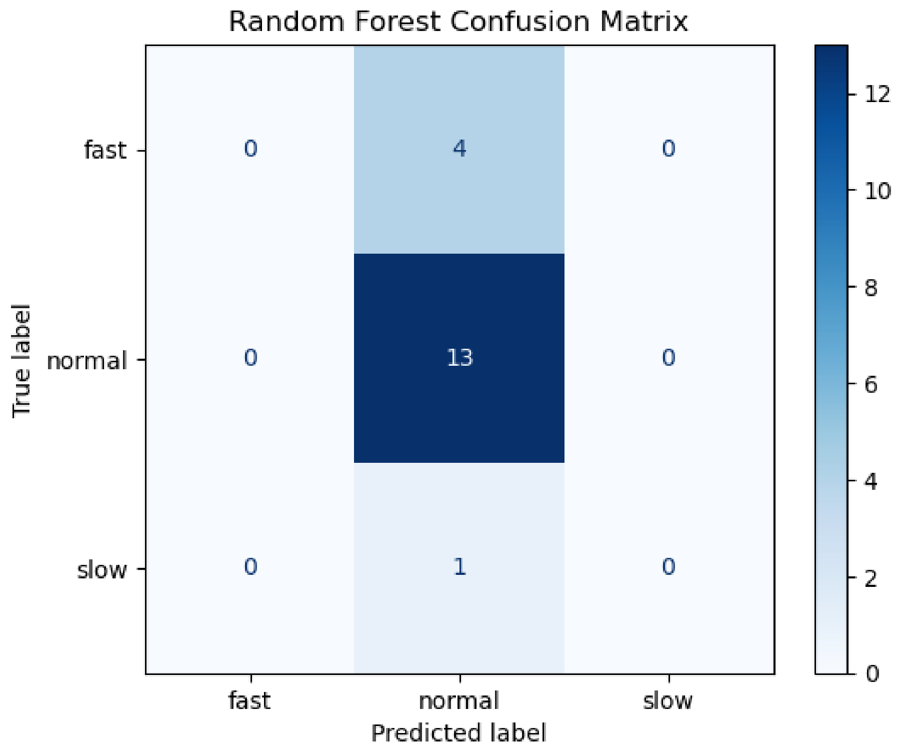

Figure 5: Random Forest confusion matrix.

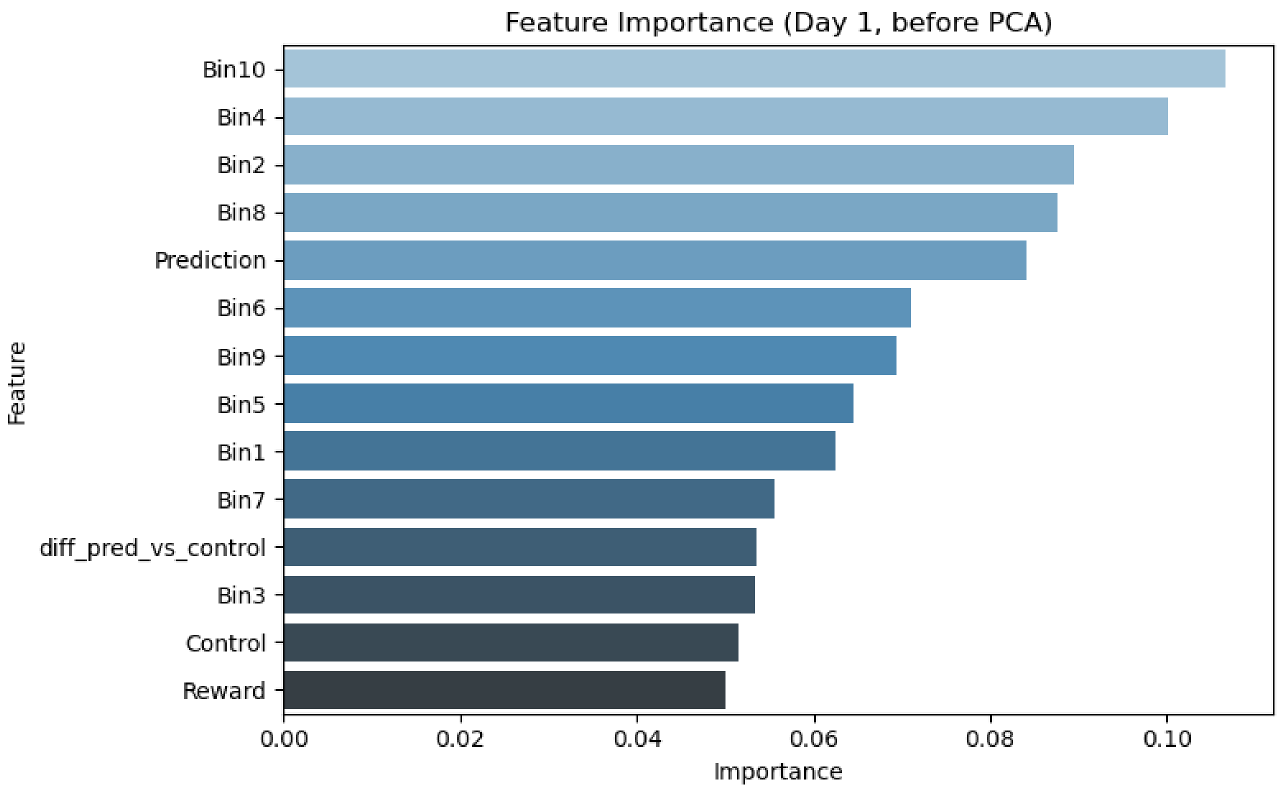

Figure 6: Feature importance in the RF model.