Visual Knowledge Tracing

Neehar Kondapaneni, Pietro Perona, Oisin Mac Aodha

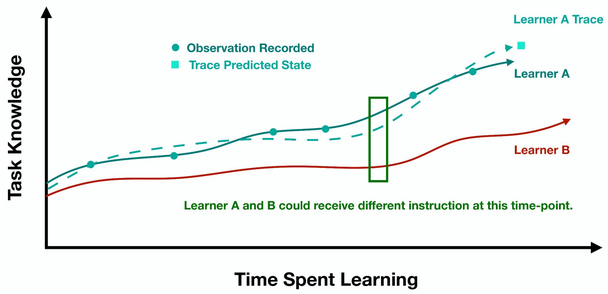
Caltech

The University of Edinburgh

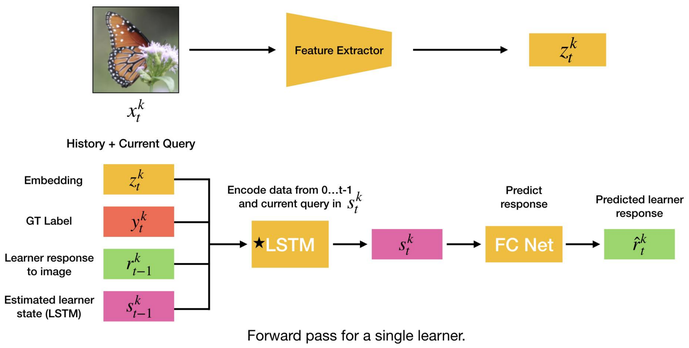
Motivation

Radiologists recognizing tumors, birdwatchers learning to disambiguate similar species, and crowd-workers learning to annotate new datasets all share a common goal -- they are trying to learn a novel visual categorization task.

In this work, we ask if we can trace how people learn these tasks by estimating their knowledge state as they learn a visual categorization task.



Tracing Model

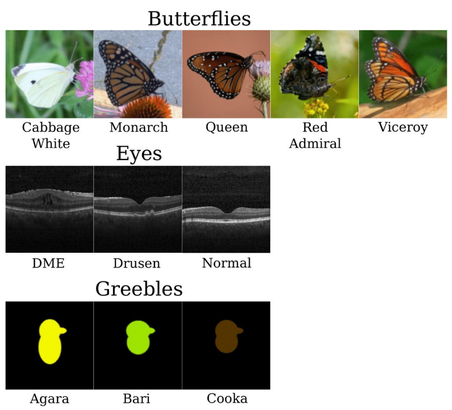


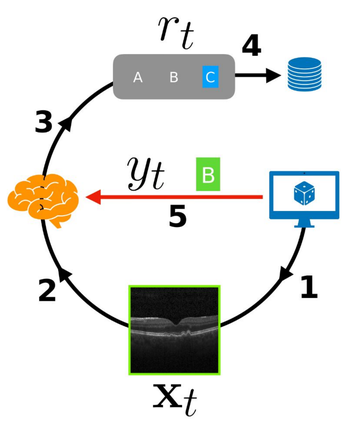
Forward pass for a single learner.

Top: A CNN produces a feature space used by all learners

Bottom: An LSTM predicts a knowledge state for an individual. The state is transformed into a linear a response prediction for a specific learner at a specific timestep.

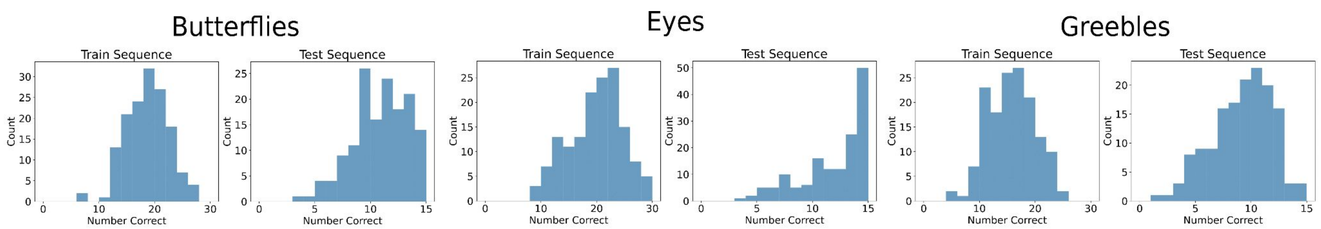
Dataset & Learner Training



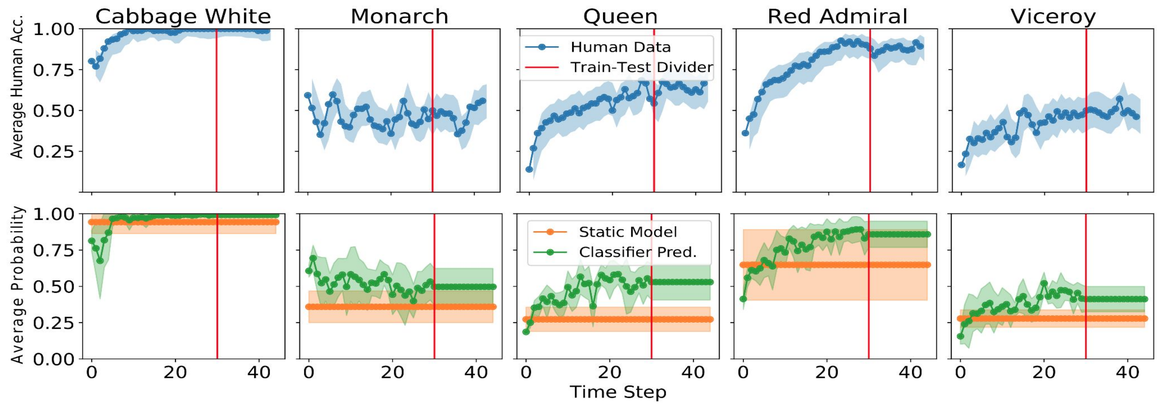


1) Randomly select an image.
2) Present the image to the learner.
3) The learner responds.
4) The response is recorded.
5) The correct label for the image is given as feedback.

After training, the test histograms show a right skew, indicating that students were able to learn the task.



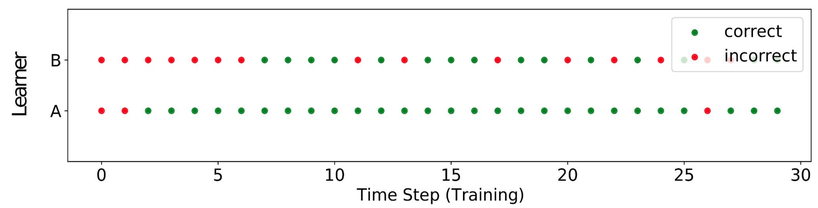
Results



Top: The average accuracy of human learners over training steps (butterflies dataset). At the end of 30 training steps, we present 15 test examples with no feedback.

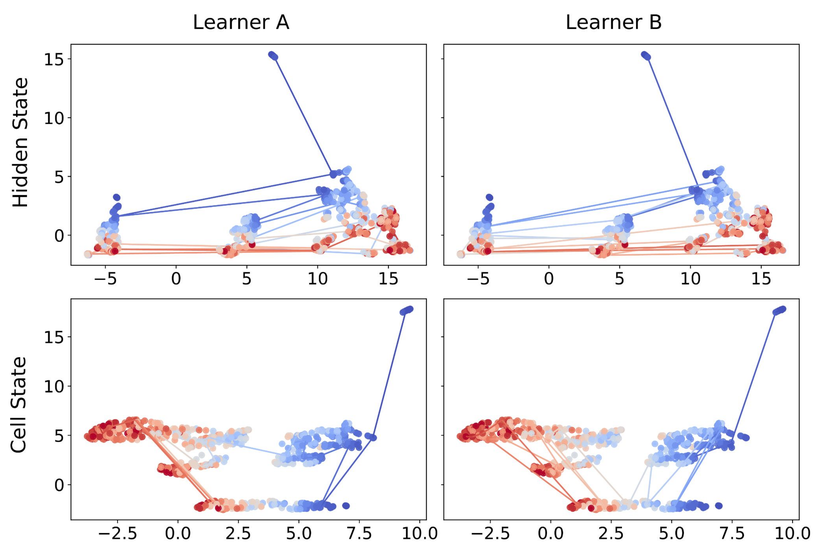
Bottom: Two different model's average predicted probability of a student correctly predicting a class. The LSTM model that outputs a classifier outperforms simple baseline methods.

Individual Learner Analysis



We consider two learners with different skill levels.

Learner A seems to have prior knowledge of the subject material.



The hidden state encodes which class the model is conditioned on.

Learner A's cell state moves to a high probability of correct earlier, reflecting higher skill.