Exploration versus teaching

Scott Cheng-Hsin Yang and Patrick Shafto

May 15, 2016

1 Introduction

We formalize a simple scenario where can we compare the effectiveness of active exploration and teaching. We check that:

- teaching with the right assumption (about the learner) is more effective than teaching with wrong assumption
- teaching with right assumption is more effective than (Bayesian) active exploration
- active exploration can be more effective than teaching with very wrong assumption

and characterize the condition under which the third statement begins to hold.

2 Framework: Two-layer hierarchical model with deterministic label on a grid

The hypothesis space is hierarchical with two layers. The top layer specifies the configuration of compartments, and the bottom layer specifies the configurations with which the category labels populate the compartments. A uniform prior over this hypothesis space spreads probability evenly across the compartment configurations, and given each hypothesis, spread probability evenly across all possible label configurations.

Basic inference

The task here is to identify from which hypothesis is the task configuration drawn from. The learner's inference follows Baye's rule. Given some data

$\mathcal{D} = \{x_i, y_i\}$, the learner's joint posterior over h and f is

$$\begin{aligned}
\mathbb{P}_L(h, f|\mathcal{D}) &= \frac{\mathbb{P}(\mathcal{D}|h, f)\,\mathbb{P}_L(h, f)}{\sum_{k,j} \mathbb{P}(\mathcal{D}|h_k, f_j)\,\mathbb{P}_L(h_k, f_j)} \\
&= \frac{1}{Z}\mathbb{P}(\mathcal{D}|f)\,\mathbb{P}_L(f|h)\,\mathbb{P}_L(h) \\
&= \frac{1}{Z}\prod_i \mathbb{P}(y_i|x_i, f)\,\mathbb{P}_L(f|h)\,\mathbb{P}_L(h)\,. \qquad (1)
\end{aligned}$$

In the deterministic labelling scenario, $\mathbb{P}(y_i|x_i, f_j) \in \{0, 1\}$, and for simplicity, x can be at any of the compartment centres on a grid fine enough to specify all the compartment configurations. If the prior is uniform, $\mathbb{P}(f_j|h_k) = \frac{1}{N_k}$ where N_k is the number of label configurations in h_k, and $\mathbb{P}(h_k) = \frac{1}{N_h}$. This joint posterior can be used to obtain $\mathbb{P}_L(h|\mathcal{D}) = \sum_j \mathbb{P}_L(h, f_j|\mathcal{D})$ for accomplishing the task and $\mathbb{P}_L(f|\mathcal{D}) = \sum_k \mathbb{P}_L(h_k, f|\mathcal{D})$ for computing the predictive distribution.

Active exploration

The learner's exploration follows a Bayesian active learning scheme that myopically chooses queries to maximize an objective function $\mathcal{F}[\,]$. That is, the learner chooses x by

$$\arg\max_{x^*} \quad \langle \mathcal{F}[\mathbb{P}_L(h|\mathcal{D}, \{x^*, y^*\})] \rangle_{\mathbb{P}_L(y^*|x^*,\mathcal{D})} \qquad (2)$$

where

$$\mathbb{P}_L(y^*|x^*, \mathcal{D}) = \sum_j \mathbb{P}(y^*|x^*, f_j)\,\mathbb{P}_L(f_j|\mathcal{D})\,. \qquad (3)$$

The expectation operator indicates that the learner does not know exactly what label she will receive but maintains a predictive distribution about it.

[A curious computation observation: if I do

$$\mathbb{P}_L(h|\mathcal{D}, x^*) = \langle \mathbb{P}_L(h|\mathcal{D}, \{x^*, y^*\}) \rangle_{\mathbb{P}_L(y^*|x^*,\mathcal{D})}\,,$$

then $\mathbb{P}_L(h|\mathcal{D}, x^*)$ is the same for all x^*. Is this even correct? If so, this means that exploration works because the expectation is taken outside the $\mathcal{F}[\,]$, and that teaching works because the expectation is conditioned on one h at a time. What is the significance of this mathematical structure?]

Common exploration objective functions include [1–4]:

- information gain: $\mathcal{F}[\mathbb{P}_L(\cdots)] = -\mathrm{H}[\mathbb{P}_L(\cdots)]$. If we write this as the difference of information gain before and after the query, ie. $\mathrm{H}[\mathbb{P}_L(h|\mathcal{D})] - \langle \mathrm{H}[\mathbb{P}_L(h|\mathcal{D}, \{x^*, y^*\})] \rangle_{\mathbb{P}_L(y^*|x^*,\mathcal{D})}$, one can identify that this is the mutual information between y^* and h. This means that there exists an alternative form, ie. $\mathrm{H}[\mathbb{P}_L(y^*|x^*, \mathcal{D})] - \langle \mathrm{H}[\mathbb{P}_L(y^*|x^*, \mathcal{D}, h)] \rangle_{\mathbb{P}_L(h|\mathcal{D})}$, which in some cases turn out to be more computationally efficient. [Here, since we are just counting, the two forms should have equal computational cost.]

- probability gain: $\mathcal{F}[\mathbb{P}_L(\cdots)] = \max[\mathbb{P}_L(\cdots)] \coloneqq \mathrm{p}_{(1)}$. In words, the algorithm goes like this. For each x, obtain a posterior over h. Pick out the highest probability, $\mathrm{p}_{(1)}$, in that distribution. Do this for all possible x and form a list of $\{x, \mathrm{p}_{(1)}\}$. Choose the x that has the highest $\mathrm{p}_{(1)}$.

- maximal margin: $\mathcal{F}[\mathbb{P}_L(\cdots)] = \max[\mathrm{p}_{(1)} - \mathrm{p}_{(2)}]$. This is to choose x that maximizes the probability difference between the most and second most likely h.

- KL distance: $\mathcal{F}[\mathbb{P}_L(\cdots)] = \mathrm{KL}[\mathbb{P}_L(h|\mathcal{D}, \{x^*, y^*\}) \, \| \mathbb{P}_L(h|\mathcal{D})]$.

The argument maximizing operator can be relaxed by a softmax function,

$$\mathbb{P}(x^*|\mathcal{D}) = \frac{\left[\langle \mathcal{F}[\mathbb{P}_L(h|\mathcal{D}, \{x^*, y^*\})]\rangle_{\mathbb{P}_L(y^*|x^*,\mathcal{D})}\right]^\alpha}{\sum_{x^*}\left[\langle \mathcal{F}[\mathbb{P}_L(h|\mathcal{D}, \{x^*, y^*\})]\rangle_{\mathbb{P}_L(y^*|x^*,\mathcal{D})}\right]^\alpha}, \tag{4}$$

where $\alpha \to \infty$ recovers the maximizing case. [The same approach can soften the max operator in $\mathcal{F}[\cdots]$ as well.] This explicitly accounts for the case where the objective function has multiple maxima and allows the learner to choose x probabilistically with potentially more robustness.

The myopic choice can also be relaxed by considering multiple x, that is, replacing x^* with $\{x_1^*, \cdots, x_n^*\}$ and y^* with $\{y_1^*, \cdots, y_n^*\}$ in Eq. 2 or 4. The number of $\{x^*, y^*\}$ pairs considered is then $(N_x N_y)^n$, which is the number of possible probe locations times the number of possible label value all to the power of step to look-ahead. In the example below, this would be $(16 * 2)^n$.

Teaching

The teacher can be helpful because he has access to the true hypothesis that the learner should infer. In a rational, cooperative setting, the learner assumes that the teacher is helpful, and the teacher knows that about the learner. The choice of teaching instances should thus involve a system of equations in which the teacher assumes the learner knows how he chooses teaching instances and the learner assumes the teacher knows her inference scheme [5].

In this teaching framework, the learner's inference given a new observation pair is

$$\mathbb{P}_L(h, f|\{x, y\}, \mathcal{D}) = \frac{\mathbb{P}(y|x, f)\, \mathbb{P}_T(x|h, \mathcal{D})\, \mathbb{P}_L(f, h|\mathcal{D})}{\sum_{j,k} \mathbb{P}(y|x, f_j)\, \mathbb{P}_T(x|h_k, \mathcal{D})\, \mathbb{P}_L(f_j, h_k|\mathcal{D})}. \tag{5}$$

[Is there a view where one works with $\mathbb{P}_T(x|f, h, \mathcal{D})$?]

If the teacher myopically chooses one teaching instance x at a time, he follows

$$\mathbb{P}_T(x|h, \mathcal{D}) = \frac{[\mathbb{P}_\ell(h|\mathcal{D}, x)]^\alpha}{\sum_x [\mathbb{P}_\ell(h|\mathcal{D}, x)]^\alpha} \tag{6}$$

$$\mathbb{P}_\ell(h|\mathcal{D}, x) = \frac{1}{Z}\left\langle \sum_j \mathbb{P}_\ell(h, f_j|x, y, \mathcal{D}) \right\rangle_{\sum_i \mathbb{P}(y|x,f_i)\mathbb{P}_\ell(f_i,h|\mathcal{D})}, \tag{7}$$

where $\mathbb{P}_\ell(\cdots)$ refers to the teacher's assumption about the learner. This system of equations can be calculated by iteration until convergence. A sensible initial condition is to use a uniform $\mathbb{P}_T(x|h,\mathcal{D}) = \frac{1}{N_x}$ for Eq. 7. Note first that Eq. 7 takes the form of Eq. 5, indicating that the iterated equations should incorporate the exact scheme of the learner's inference. Note second that Eq. 7 is conditioned on a particular h, so the predictive distribution part of Eq. 7 should be thought of as imagining the case when $h = h^+$. [What does taking the expectation of Eq. 5 with the normalizing constant describe?] [Can I argue through the iteration using this initial condition?]

Strictly following [5], Eq. 7 would be replaced by

$$\mathbb{P}_\ell(h|\mathcal{D},x) = \frac{\mathbb{P}_T(x|h,\mathcal{D})\,\mathbb{P}_\ell(h|\mathcal{D})}{\sum_k \mathbb{P}_T(x|h_k,\mathcal{D})\,\mathbb{P}_\ell(h_k|\mathcal{D})}, \tag{8}$$

where $\mathbb{P}_\ell(h|\mathcal{D})$ is computed according to the description after Eq. 1.

[Yet to understand how choosing multiple instances simultaneously is better than choosing the same number sequentially.]

Teaching with two-step look ahead.

Computation

Intuition, confusion, and questions

Trying to reconstruct and understand the teaching equations: The learner's basic inference given a new x goes like $\mathbb{P}_\ell(h|\mathcal{D},x^*) = \langle \mathbb{P}_\ell(h|\mathcal{D},\{x^*,y^*\})\rangle_{\mathbb{P}_\ell(y^*|x^*,\mathcal{D})}$. This form suggests that the learner is unaware of the teacher. How to put the awareness in? The learner knows that the teacher chose x according to some objective, say making the probability $h = h^+$, as large as it can be. So, she uses Bayes' rule and asks, what is the probability that x is chosen if that objective is fulfilled? The answer is the probability with which the teacher chose x. Here is a question: following this account, it feels like the learner should discard her prior belief $\mathbb{P}_\ell(h|\mathcal{D})$ in Eq. 7 if she trusts the teacher fully. Why is this not the case? Because one needs to assume some prior. (not having a prior is a prior.) the only one that makes sense, given that the learner trusts the teacher, is to extend that trust to knowing the prior.

Trying to understand the iterative nature: The reasoning in words in [5] is a backward narrative, tracing from the final inference of the learner to the initial condition, which I guess can be either a $\mathbb{P}_T()$ or a $\mathbb{P}_\ell()$. A forward trace in words may sound like: Given P_0, the teacher would have chosen x_1. Given x_1, the learner would have inferred P_1. Given P_1, the teacher would have chosen x_2. And so it goes. So, iterate until convergence, that is $x_{t-1} = x_t$ or $P_{t-1} = P_t$ (one implies the other). The forward version is the teacher's version. backwards is the learners. In the forward version (from the teacher's perspective), the initial condition is a $\mathbb{P}_\ell(h|\mathcal{D},x)$, and I can imagine how to generate that. In the backward version (from the learner's perspective), the initial condition is really just an x, but it should be turned into a distribution. It's less apparent how I should do that.

It feels like Eq. 6 should involve the knowledge of the true hypothesis. Right now, it involves it only through the initial condition. Is this correct? And if the teacher's choice probability involves h^+, should the learner not make inference about the choice distribution rather than simply having access to it as Eq. 7 suggests? It is a bit confusing because it involves lots of perspective shifting. the teaching version (from the teacher's perspective) involves the correct hypothesis. but because the learner's inference is necessarily uncertain about the correct hypothesis, to get predictions about which data to choose, the teacher needs predictions for all hypotheses.

It feels like Eq. 6 should match Eq. 4 except for the distribution over which the expectation is taken, because both are about optimally choosing an x. So, how do I frame Eqs. 6 and 7 to use an objective function, ie. $\mathbb{P}_\ell(x^*|h^+, \mathcal{D}) \to \langle \mathcal{F}[\mathbb{P}_\ell(h^+|\mathcal{D}, \{x^*, y^*\})] \rangle_{\mathbb{P}_\ell(y^*|x^*, h^+)}$? Is it by changing just the initial condition? An objective-function based view would allow a fairer comparison between teaching and exploration (eg. both can be based on probability gain). In light of this, can one think of Eq. 6 as the counterpart to exploration with maximal probability gain? Let's talk through this, but these are really different not (only) in the examples but in their semantics. the choice of an example by a teacher is not just correct, it is "good" in some sense.

It feels like Eq. 6 should just be a softer, probabilistic version of $\arg\max_{x^*}$ as Eq. 4 is a softer version of Eq. 2, and by symmetry, Eq. 7 should be a softer, probabilistic version of $\arg\max_h$. Indeed, Eq. 7 can involve another exponent α_ℓ around $\mathbb{P}_T(\cdots)$ or around the whole numerator, but what is its interpretation? And is the iteration still sensible and somewhat Bayesian in the limit $\alpha_\ell \to \infty$? One way the strict maximizing version of teaching changes is that maximization cuts off the iteration. that is, we iterate to convergence, but if one chooses the max prob item on the first try, then one has converged.

Another symmetry-inspired question is that why does Eq. 7 look Bayesian while Eq. 6 does not? One could introduce a P(d) term in the numerator. this is an obtuse way to institute a cost function over the data.

If the teacher wants to teach a true task distribution, $\mathbb{P}^*(h)$, the teaching equations should now look like

$$\mathbb{P}_T(x^*|\mathbb{P}^*(\cdot), \mathcal{D}) = \frac{[\mathbb{P}_\ell(\mathbb{P}^*(\cdot)|\mathcal{D}, x^*)]^\alpha}{\sum_{x^*}[\mathbb{P}_\ell(\mathbb{P}^*(\cdot)|\mathcal{D}, x^*)]^\alpha}$$

$$\mathbb{P}_\ell(\mathbb{P}^*(\cdot)|\mathcal{D}, x^*) = \frac{\mathbb{P}_T(x^*|\mathbb{P}^*(\cdot), \mathcal{D})\,\mathbb{P}_\ell(\mathbb{P}^*(\cdot))}{\sum_k \mathbb{P}_T(x^*|\mathbb{P}^*(\cdot)_k, \mathcal{D})\,\mathbb{P}_\ell(\mathbb{P}^*(\cdot)_k)}.$$

How do I incorporate fancier objective function into these? In this case, fancier objectives, such as $\langle \mathrm{KL}[\mathbb{P}_\ell(h|\mathcal{D}, \{x^*, y^*\}) \| \mathbb{P}^*(h)] \rangle_{y^*}$, seem to be quite reasonable. Also, what would the corresponding equations for exploration look like? Another way to ask your question is to ask what loss function is implied by the standard model. the answer is that it assumes 0/1 loss. relaxing to a distribution yields a k-l divergence as a natural loss function.

Teaching with uncertain assumption

Teaching with perfect knowledge about the learner is implausible, and teaching with the wrong assumption can be disastrous. Therefore, we consider the case in which both the teacher and learner know that the teacher is uncertain about the learner. We introduce this uncertainty via a hyper prior distribution over the learner's induction biases, that is over the learner's joint prior $\mathbb{P}_\ell(h, f)$. Indexing the setting of the learner's joint prior by B, the teacher proceeds by generating $\mathbb{P}_T(x|h, \mathcal{D}, B)$ according to Eqs. 6-7 for each setting of the joint prior, then marginalizing out the settings to obtain $\mathbb{P}_T(x|h, \mathcal{D}) = \sum_i \mathbb{P}_T(x|h, \mathcal{D}, B_i)\,\mathbb{P}(B_i)$.

3 Examples

Exploration vs. Teaching

We start with a simple scenario with 2 labels ($y \in \{1, 2\}$), 4 compartment configurations, and deterministic labelling given the label configuration. The 4 compartment configurations are: horizontal split (h_1), vertical split (h_2), 2-by-2 split (h_3), and 4-by-4 split (h_4). [Can extend to probabilistic labelling by assigning labels from mixture of Gaussians that populate the compartments.] For each hypothesis, all compartments after the split have equal size, and the label configurations are all the unique configurations, or distinct permutations, in which the number of category 1 and 2 labels balance. For example, with the 4-by-4 split, there are $\binom{16}{8}$ configurations. Here the probing grid is 4-by-4 since this is the most coarse grid that contains all configurations.

[Figure showing configurations.]

For uniform priors over h and $f|h$, the probing choices for optimal exploration is:

- with no observation: random
- with 1 observation pair: anywhere in the two adjacent quadrants
- with 2 pairs: start to depend on the actual observation

The probing choices for teaching is:

- with no observation: random
- with 1 observation pair:
  - if $h^+ = h_1$: anywhere in the adjacent quadrants, $\mathbb{P}_T(x|h) = [\frac{1}{2}, \frac{1}{2}, 0, 0]$
  - if $h^+ = h_2$: same as hypo 0, $\mathbb{P}_T(x|h) = [\frac{1}{2}, \frac{1}{2}, 0, 0]$
  - if $h^+ = h_3$: anywhere in the diagonal quadrant, $\mathbb{P}_T(x|h) = [0, 0, 1, 0]$
  - if $h^+ = h_4$: anywhere in the same quadrant, $\mathbb{P}_T(x|h) = [0, 0, 0, 1]$
- with 2 pairs: no 3rd choice needed, because two observation pairs at adjacent quadrants completely distinguishes between h_1 and h_2, and $\mathbb{P}_T(x|h)$ reveals h^+ with certainty after one observation pair if $h^+ = h_3$ or h_4.

Interestingly, if Eq. 7 is replaced with a simpler version, $\frac{1}{Z}\mathbb{P}_T(x|h,\mathcal{D})\,\mathbb{P}_\ell(h|\mathcal{D})$, as in [5], the probing choices with one observation pair is different for h_1 and h_2. For h_1, the choice becomes anywhere in the adjacent horizontal quadrants with $\mathbb{P}_T(x|h) = [1,0,0,0]$. For h_2, the choice becomes anywhere in the adjacent vertical quadrant with $\mathbb{P}_T(x|h) = [0,1,0,0]$.

A simpler way to think about how the teacher generates optimal guidance is by thinking in terms of the "most representative" features of each concept, defined as the features that have the maximum prior probability given that concept—$\arg\max_f \mathbb{P}(f|h = h^*)$.

For h_4, a smaller but rich enough set of configurations is all unique combinations of two horizontal strips plus all unique combinations of two vertical strips. The teacher's guide for this new h_4 is to choose the diagonal square in the same quadrant that contains the first observation.

[Figure showing performance for random, optimal exploration, and teaching.]

Note that performance under exploration does not approach 1 because certain configurations are shared across hypotheses in this setup. However, teaching breaks this limitation because the learner takes advantage of the fact that the teacher knows the true answer. In fact, with teaching, the ideal learner gets to the right answer with certainty after only two observations.

One perspective is that the teacher and learner have established a "code" via the teacher's choice of x, where the decoding is via cooperative and rational reasoning.

[Run some proper simulations to show that this is or isn't the case. This extraordinary improvement comes at a cost: if the teacher assumed a wrong prior for the learner, the learner would be misled to the wrong answer with certainty. Get more precise conditions, such as perturbed learner's prior vs uniform assumed prior, uniform learner's prior vs perturbed assumed prior, both priors perturbed, etc. In other words, the benefit of teaching is fragile with respect to the alignment of the learner's prior with that assumed by the teacher.]

Teaching with softer cooperation

It is implausible that real learner can perform the full, ideal cooperative inference used by the model. What are some principled ways we can soften this?

Teaching with wrong assumption

Teaching with uncertainty

Because teaching with uncertainty involves repeating teaching $Nsetting$ times, we introduce a hypothesis-compartment space with fewer configurations.

The hyper prior on each setting is specified by

4 Methods

Participants

The experiment was run on Amazon Mechanical Turk. There were a total of 60* participants, 30* for each of the two experimental conditions. One of the participants....

IRB stuff

Stimuli

All patterns are composed of black and white squares on a 4-by-4 grid. The concepts are defined by the set of patterns shown in Fig. XX. Altogether, there are 3 concepts (named H1, H2, H3) and 6 different patterns.

show figure of concept-pattern space

Procedure

There are two experimental conditions: one composed of a learning phase followed by a self-exploration phase, while the other has the same learning phase followed by a teaching phase. Details of the different phases are described below.

Each phase contains 4 rounds of 18 trials. Each set of 18 trials contains 6 of each concepts. The patterns in concept 1 and 2 are each shown 3 times while those in concept 3 shown only once so that both the frequency of concepts and the frequency of patterns given a concept are uniform. The order of each set is randomly shuffled.

At the beginning of each phase before entering the trials, a page of instructions was shown to the participants to explain what they would see (ie. a pattern and 3 concept choices), what they should do (eg. click on buttons to choose a concept), and when the phase would terminate.

Learning phase. For each trail of this phase, the participants were shown the entire pattern and asked to respond by choosing one of the concepts. Upon making a choice, a feedback of "Correct" or "Incorrect" was given.

The participants were instructed that the goal of this phase is to learn which pattern associates which concept and at what frequency. They were explicitly told that some patterns are shared by multiple concepts. They were also told that they would proceed to the next phase once they reached an accuracy of 72% or when they used up all the 72 available trials. This accuracy was the running average of proportion correct in the last 18 trials and was shown as score bar through this phase.

The accuracy threshold for ending the phase early was meant to be an incentive for the participants to learn the concept-pattern association well. It was set to be between the accuracy that an ideal observer model would achieve by using the optimal decision rule (78%) versus a probability-matching rule (67%). Only XX of the 60 participants were able to reach the accuracy threshold in less than 72 trials.

Exploration phase. For each trial of this phase, the pattern was initially fully covered by gray squares. The participants were instructed to open 2 squares, after which the concept choices would appear, and the participants could make their choice. Unlike the learning phase, the feedback in this phase only acknowledged that a choice was made but was not corrective. Also, the score was not shown until the very end of the phase.

The accuracy threshold for ending the phase early was 67%, which is also what an ideal observer with the optimal decision and query rules would achieve. Only XX participants were able to reach this threshold within 72 trials.

Teaching phase. For each trail of this phase, the pattern was also initially fully covered. Two squares were automatically revealed in sequence, followed by the appearance of the concept choices. As in the exploration phase, the feedback was neutral, and the score was shown at the end of the phase.

The sequence of openings followed the rule generated by the teaching equations (Eqs.XX-XX). Explain the rules?

To encourage epistemic trust,

[Teaching is done by highlighting one by one the square that the participants should open. Variations of include: 1) Highlight the square to be chosen and have the participant click on it. 2) Highlight all the squares that the teacher would choose from and have the participant click on one. The problem with this is that it gives more information to the learner than that modeled by the formalism. On another note, this can be used for guided learning. 3) One and a half revealing, ie. the second revealing is only a highlighted position without a label. This gives information about $\mathbb{P}_T(x|h,\mathcal{D})$ when compared to the performance with two revealings.]

5 Results

Performance

Is teaching better than self-exploration?

Mixture of binomial analysis

Unlike a coin flip experiment where every trial uses the same coin, here some patterns (and partial patterns) provide more information about the associated concepts than other. For example, a learner who sees two different labels within the same quadrant should know with certainty that the answer is concept 4. In contrast, a learner who sees two labels from adjacent quadrant would still be quite uncertain about the correct answer. Therefore, these Bernoulli trials are independently but not identically distributed, and one should account for this in order to properly evaluate the expected performance induced by teaching relative to optimal exploration.

5.1 Variance

Test the double-edged sword hypothesis of teaching by looking at variances.

5.2 Optimality

Empirical distributions

Decision rule, lapse rate, and concept prior Maximum-likelihood analysis?

Extend of trust and meta-reasoning

6 Characterizing conditions

One way to characterize the wrongness of the teacher's assumption is $\mathrm{KL}[\mathbb{P}_L(h)\,||\mathbb{P}_\ell(h)]$.

To assess the effectiveness of exploration and teaching, we consider several metrics as a function of data number and then quantify the effectiveness based on some feature and summary statistics of these curves. Metrics include:

- performance: $\mathbb{P}_L(h^+|\mathcal{D})$
- KL distance: $\mathrm{KL}[\mathbb{P}_L(h|\mathcal{D})\,||\mathbb{P}^*(h)]$
- more?

[A cool generalization of the framework: the teacher can teach by asking to student to choose from a range of x. When the range is a particular x, this is the traditional teaching. When the range is all possible x, this is allowing the learner to explore freely while the teacher observes her and makes inference about her belief. Any range of x in between is like a granny's method. The natural extension from this is to include uncertainty in the teacher's assumption and put this in a POMDP framework for choosing the best action.]

References

[1] Settles, B. Active Learning Literature Survey. Technical report, (2010).

[2] Nelson, J.D., McKenzie, C.R.M., Cottrell, G.W. & Sejnowski, T.J. Experience matters: Information acquisition optimizes probability gain. *Psychol Sci* **21**, 960–969 (2010).

[3] Markant, D.B., Settles, B. & Gureckis, T.M. Self-Directed Learning Favors Local, Rather Than Global, Uncertainty. *Cognitive Sci*, 1–21 (2015).

[4] Yang, S.C.H., Lengyel, M. & Wolpert, D.M. Active sensing in the categorization of visual patterns. *eLife* (2016).

[5] Shafto, P., Goodman, N.D. & Griffiths, T.L. A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology* **71**, 55–89 (2014).