Active Learning

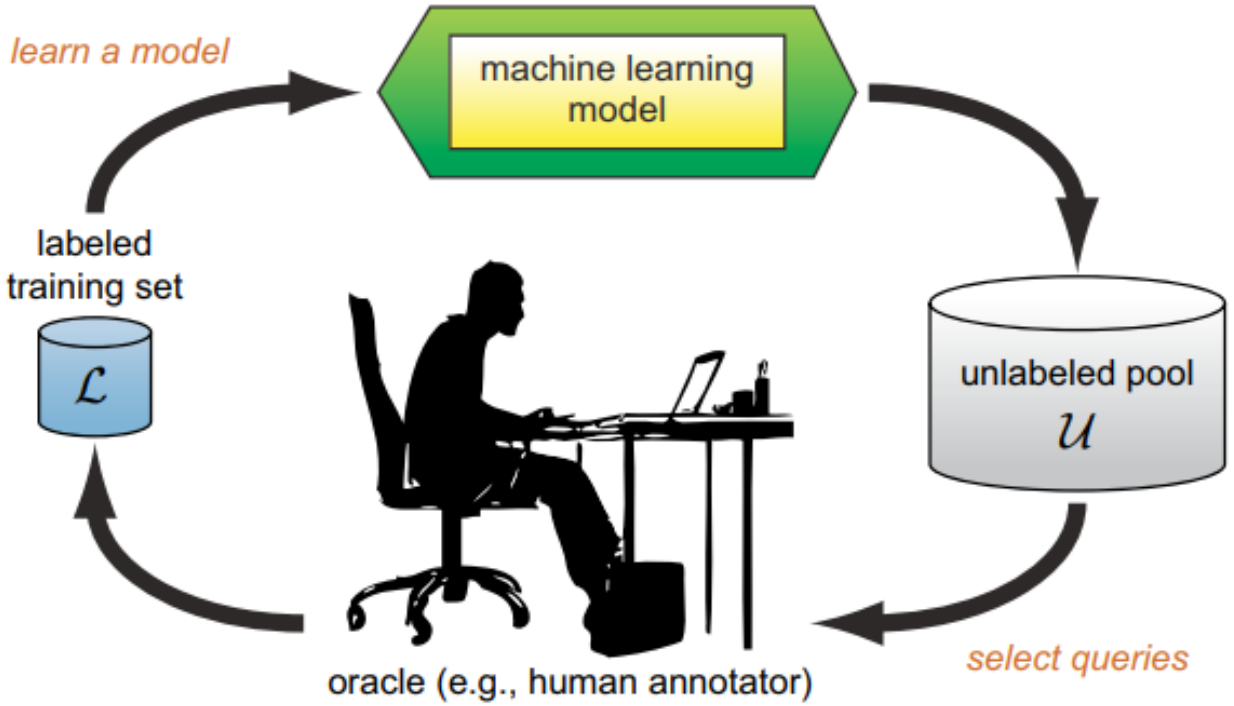
Topics in Trustworthy AI

Daksh Mittal

Spring 2025

Active Learning

- **Ground truth labels are costly:** Learn efficiently with less labels.
- Randomly querying X would be expensive.
- Adaptively choose which X to query next.

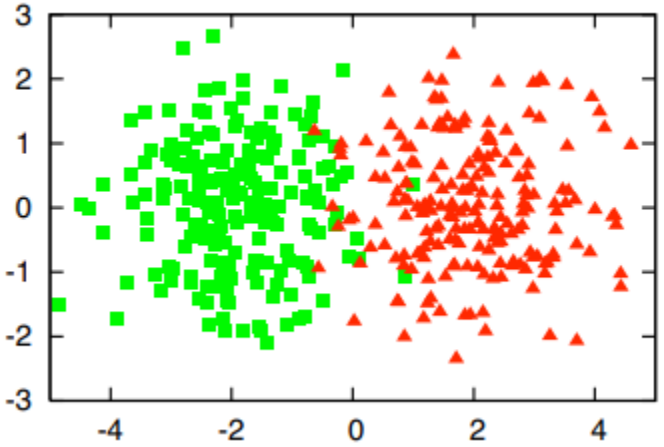


Efficient data collection for improving ML models

Active Learning: Decision problem under uncertainty

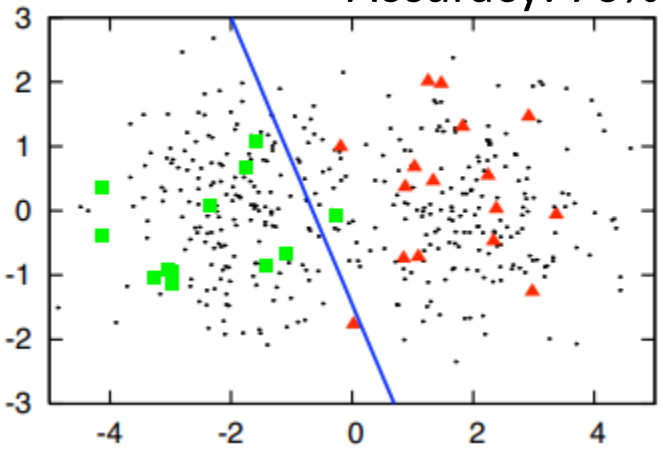
- Active Learning - Adaptive data collection for improving AI/ML models
  - ✓Uncertainty Quantification as a prerequisite
  - ✓How to select queries? Ideas closely linked to Multi-armed Bandits, and Bayesian optimization

An Example of Active Learning



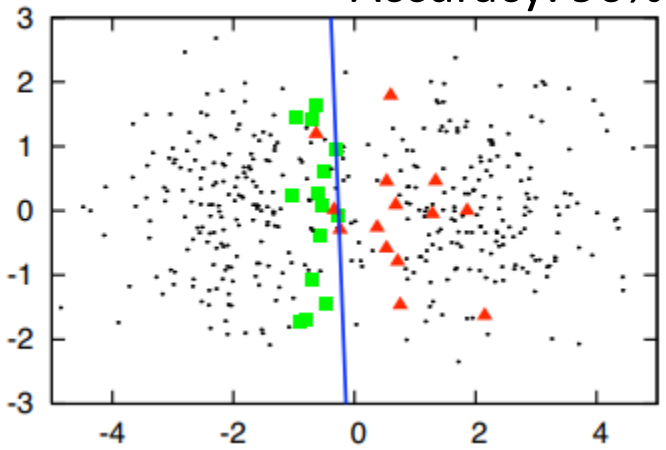
Dataset
(Two class Gaussians)

Accuracy: 70%



Randomly query
30 points

Accuracy: 90%



Actively query
30 points

Source: https://burrsettles.com/pub/settles.activelearning.pdf

Different Active Learning Scenarios

- **Query synthesis**
  - ✓ Generate a query (X) to be labeled
  - ✓ Potential issue - labeling arbitrary inputs
- **Stream-based selective sampling**
  - ✓ Sample an instance from input space – decide to query or discard
- **Pool-based sampling**
  - ✓ Sample a large pool of instances from input space – select the best query

Querying strategies

1. **Uncertainty based strategies:** Sample from the regions of the space with highest uncertainty.

2. **Representativeness-based strategies:** Query data that is representative of the underlying population – diversity and density based criteria.

3. **Performance-based strategies:** Sample to directly optimize the performance of the model.

Uncertainty based strategies

Uncertainty Sampling

1) **Least confident sampling:** Query the instance whose prediction is **least confident**.
Let $\hat{Y} = \text{argmax}_Y \, p_M(Y|X, D_{train})$,

$$X^* = \text{argmax}_X \left(1 - p_M\left(\hat{Y}|X, D_{train}\right)\right)$$

2) **Margin sampling:** Query the instance with **least difference** between probabilities of top two classes -

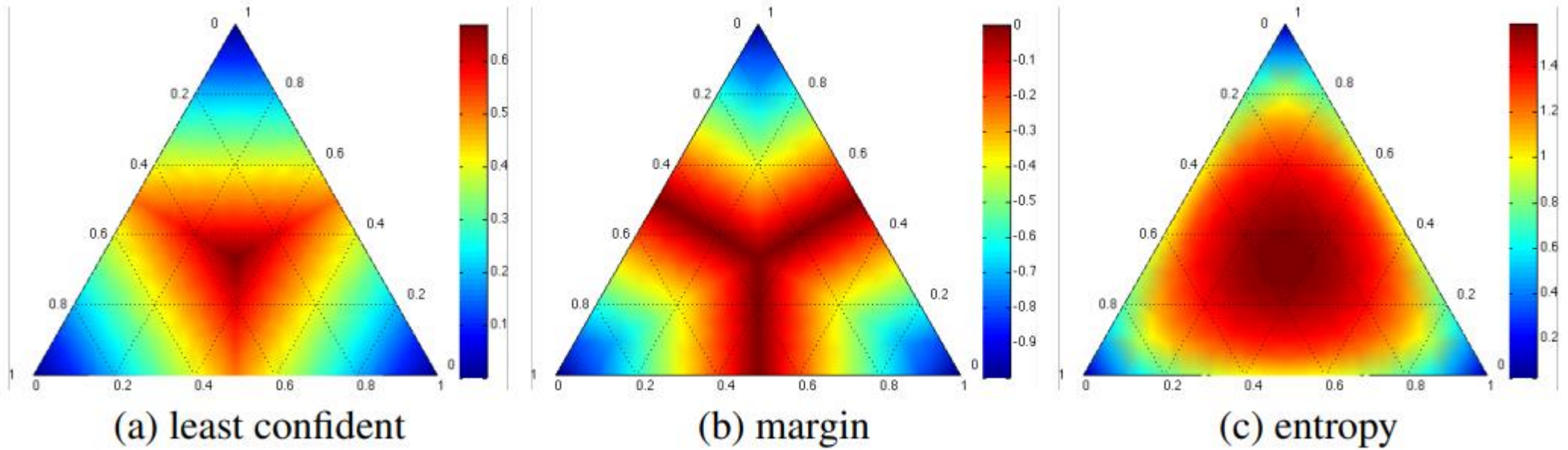
$$X^* = \text{argmin}_X\left(p_M(\hat{Y}_1|X, D_{train}) - p_M(\hat{Y}_2|X, D_{train})\right)$$

3) **Shannon entropy:** Query the instance with **highest entropy**

$$X^* = \text{argmax}_X\left(-\sum_c p_M\left(\hat{Y} = c|X, D_{train}\right)\log\left(p_M\left(\hat{Y} = c|X, D_{train}\right)\right)\right)$$

No differentiation between epistemic and aleatoric uncertainty

Visualizing uncertainty sampling



Source: https://burrsettles.com/pub/settles.activelearning.pdf

Query by committee

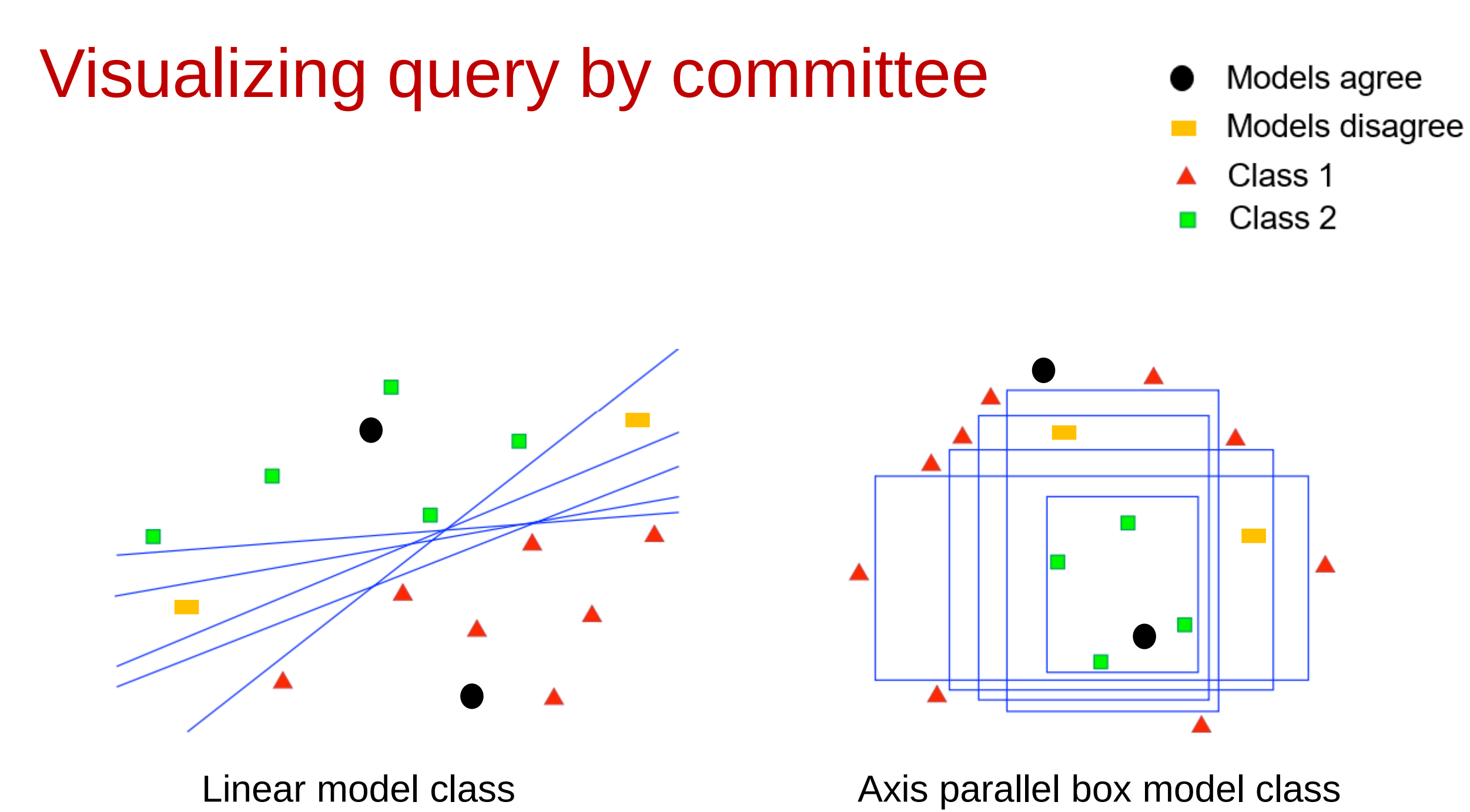
Maintain a committee $\mathcal{C} = \{\theta_1, \theta_2, \cdots, \theta_N\}$ of models – represent competing hypotheses.

- ✓Most informative query - instance about which the models most disagree.
- ✓How to get different models? Ensembles, Explicit prior and posterior distributions of model parameters etc.
- ✓How to measure disagreement? Vote entropy, Average KL divergence

$$\alpha_{VE}(X) = -\sum_{c} \frac{V(c)}{N} \log\left(\frac{V(c)}{N}\right)$$

Trying to capture some notion of epistemic uncertainty!

Visualizing query by committee



Linear model class

Axis parallel box model class

Bayesian Active Learning by Disagreement (BALD)

Maximize the mutual information between the model predictions and model parameters

$$\alpha_{BALD}(X) = I(Y, \theta | X, D_{train}) = \mathrm{H}(\mathrm{Y}|\mathrm{X}, D_{train}) - \mathbb{E}_{\theta \sim p(\theta|D_{train})}\big(H(Y|X, \theta)\big)$$

- ✓ **Term 1 :** X with high entropy in average output $\Big[P(Y|X, D_{train}) = \mathbb{E}_{\theta \sim p(\theta|D_{train})}\big(P(Y|X, \theta)\big)\Big]$
- ✓ **Term 2 :** Penalizes X, where many models are not confident – Inputs X on which models disagree.
- ✓ **Larger difference indicates :** Labeling X would provide more information about model's parameters.

Uncertainty based strategies – summary

- Query most uncertain input X
  - ✓notion of uncertainty differs among methods
- **Potential issue:** Prone to querying outliers
  - ✓Might not improve the performance of the model

Representativeness based strategies

- **Informative instances:** Uncertainty +“representative” of the underlying distribution

- **Density weighted methods:** average similarity to other instances

$$(\text{Uncertainty based metric}) \left(\frac{1}{U} \sum_{u \in [1,U]} sim(X, X^u) \right)^{\beta}$$

- **Other methods:** K-means clustering, Core set

  - ✓ **Main idea:** Least similar to labeled set + Most similar to unlabeled set

Performance-based strategies

- Expected error reduction:

Probability that label $Y = c$ for input X

$$Minimize_X \sum_c p(Y = c|X, \theta) \left(\sum_{u \in U} (1 - p_{\theta^{+(X,c)}}(\hat{Y}|X^u)) \right)$$

Error on the unlabeled pool U, under the updated model $[\theta^{+(X,c)}]$

- Expected log-loss:

$$Minimize_X \sum_c p(Y = c|X, \theta) \left(-\sum_{u \in U} \sum_{\hat{c}} p_{\theta^{+(X,c)}}(\hat{c}|X^u) \log \left(p_{\theta^{+(X,c)}}(\hat{c}|X^u) \right) \right)$$

Log loss on the unlabeled pool U, under the updated model $[\theta^{+(X,c)}]$

Other considerations!

- Batch mode Active learning
- Variable labeling costs
- Multi-task active learning – single query labeled for multiple tasks

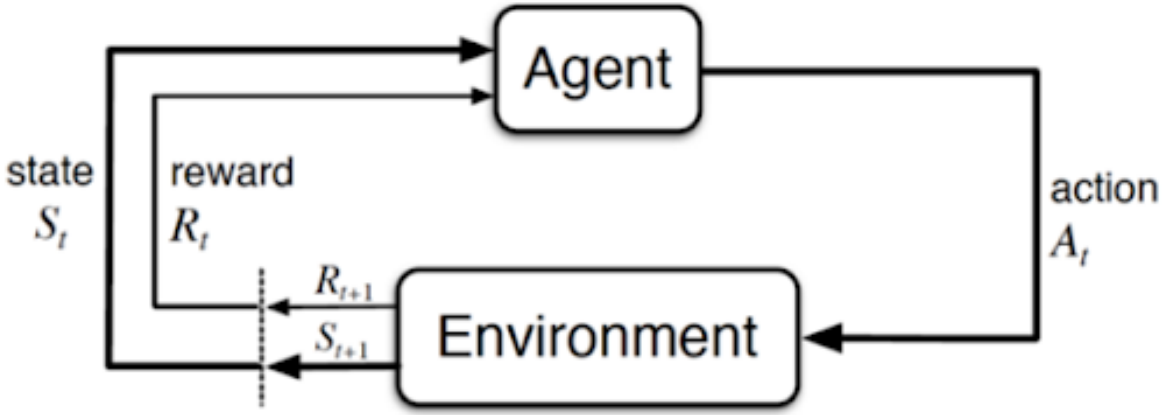
Critical components and Limitations

- **Uncertainty Quantification**: Requires methodologies that effectively differentiate epistemic and aleatoric uncertainty
- **Explicitly optimize for the objective under consideration:** model improvement over target population
- Heuristics try to balance trade-offs between uncertainty, representation - **How to approach in a principled manner?**
- Mostly caters to classification tasks and suffers highly under batching**.**

Unified Approach – Bayesian Adaptive MDPs

Markov Decision Processes (MDP)

- MDP: framework for sequential decision making
- State – $S_t \in \boldsymbol{\mathcal{S}}$
- Action – $A_t \in \boldsymbol{\mathcal{A}}$
- Dynamics – $P(\cdot\,|S, A) \in \boldsymbol{\mathcal{P}}(\boldsymbol{\mathcal{S}})$
- Reward – $R(S, A) \sim q(\cdot\,|S, A) \in \boldsymbol{\mathcal{P}}(\boldsymbol{\mathcal{R}})$
- MDP $\boldsymbol{\mathcal{M}} = (\boldsymbol{\mathcal{S}}, \boldsymbol{\mathcal{A}}, P, q)$ with objective: Maximize $\mathbb{E}(\sum_{t=0}^{T-1} \gamma^t R\,(S_t, A_t)|S_0 = s)$



Partially Observable MDP

- Complete states are not observable, instead observation O_t is observed at time step t
- State – $S_t \in \boldsymbol{\mathcal{S}}$, Action – $A_t \in \boldsymbol{\mathcal{A}}$, Dynamics – $P(\cdot\,|S, A) \in \boldsymbol{\mathcal{P}}(\boldsymbol{\mathcal{S}})$
- Observation $0 \in \boldsymbol{\mathcal{O}}$, where $0_t \sim \Omega(\cdot\,|S_{t+1}, A_t)$
- POMDP $\boldsymbol{\mathcal{M}} = (\boldsymbol{\mathcal{S}}, \boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{O}}, P, \Omega, q)$

MDP with unknown dynamics

- Dynamics – $\mathrm{P}_\theta(\cdot\,|S,A)$ with $\theta \sim \mu_0$
- Reformulation as Partially observable MDP
  - ✓ Define a new hyper-state $S' = (\theta, S)$ with observations $\mathrm{O}'_\mathrm{t} = S_t$
  - ✓ Dynamics over hybrid state

$$P'(S'_{t+1}|A_t, S'_t) = P'(S_{t+1}, \theta|A_t, S_t, \theta) = P_\theta(S_{t+1}|A_t, S_t)$$

  - ✓ $\Omega'(O_{t+1}|S'_{t+1}, A_t) = \Omega(O_{t+1}|S_{t+1}, \theta, A_t) = (\mathrm{S}_{\mathrm{t+1}}\ \mathrm{w.p.}\ 1)$
  - ✓ POMDP $\mathcal{M}' = (\mathcal{S}', \mathcal{A}, \mathcal{O}', P', \Omega', q)$

Bayesian Adaptive MDP

- Let $\mu_t = \mu(\theta|H_t, \mu_0)$ be the posterior over the parameter θ given the history $H_t = (S_0, A_0, \cdots, A_{t-1}, S_t)$
- Define an MDP as follows
  - ✓ Hyper-state with posteriors $\tilde{S}_t = (S_t, \mu_t)$
  - ✓ Transition dynamics over the hybrid state space $\tilde{P}\left(\tilde{S}_{t+1}\middle|\tilde{S}_t, A_t\right)$

    $$\tilde{P}(S_{t+1}, \mu_{t+1}|S_t, \mu_t, A_t) = \mathbb{1}(\mu_{t+1}|S_{t+1}, S_t, A_t, \mu_t)\left(\int P_\theta(S_{t+1}|S_t, A_t)\, d\mu_t(\theta)\right)$$

  - ✓ Reward function $\tilde{q}\left(\cdot\middle|\tilde{S}, A\right) = \tilde{q}(\cdot|S, \mu, A) = q(\cdot|S, A)$
  - ✓ Bayesian Adaptive MDP $\widetilde{\mathcal{M}} = (\tilde{\mathcal{S}}, \mathcal{A}, \tilde{P}, \tilde{q})$

Generalizing to unknown reward functions

- Reward $R(S_t, A_t) \sim q_\alpha(\cdot\,|S_t, A_t)$ with $\alpha \sim \nu_0$
- Define $\mu_t = \mu(\theta|H_t, \mu_0)$ and $\nu_t = \nu(\alpha|H_t, \nu_0)$ be the posterior over the parameter θ and α given the history H_t
- Define an MDP as follows
  - ✓ Hyper-state with posteriors $\tilde{S}_t = (S_t, \mu_t, \nu_t)$
  - ✓ Transition dynamics over the hybrid state space and rewards

$$\tilde{P}\left(\tilde{S}_{t+1}, R_t \middle| \tilde{S}_t, A_t\right) = \tilde{P}(S_{t+1}, \mu_{t+1}, \nu_{t+1}, R_t | S_t, \mu_t, \nu_t, A_t)$$

$$= \mathbb{1}(\mu_{t+1}|S_{t+1}, S_t, A_t, \mu_t)\left(\int P_\theta(S_{t+1}|S_t, A_t)\, d\mu_t(\theta)\right) \mathbb{1}(\nu_{t+1}|R_t, S_t, A_t, \nu_t)\left(\int q_\alpha(R_t|S_t, A_t)\, d\nu_t(\theta)\right)$$

  - ✓ Bayesian Adaptive MDP $\widetilde{\mathcal{M}} = (\widetilde{\mathcal{S}}, \mathcal{A}, \tilde{P})$

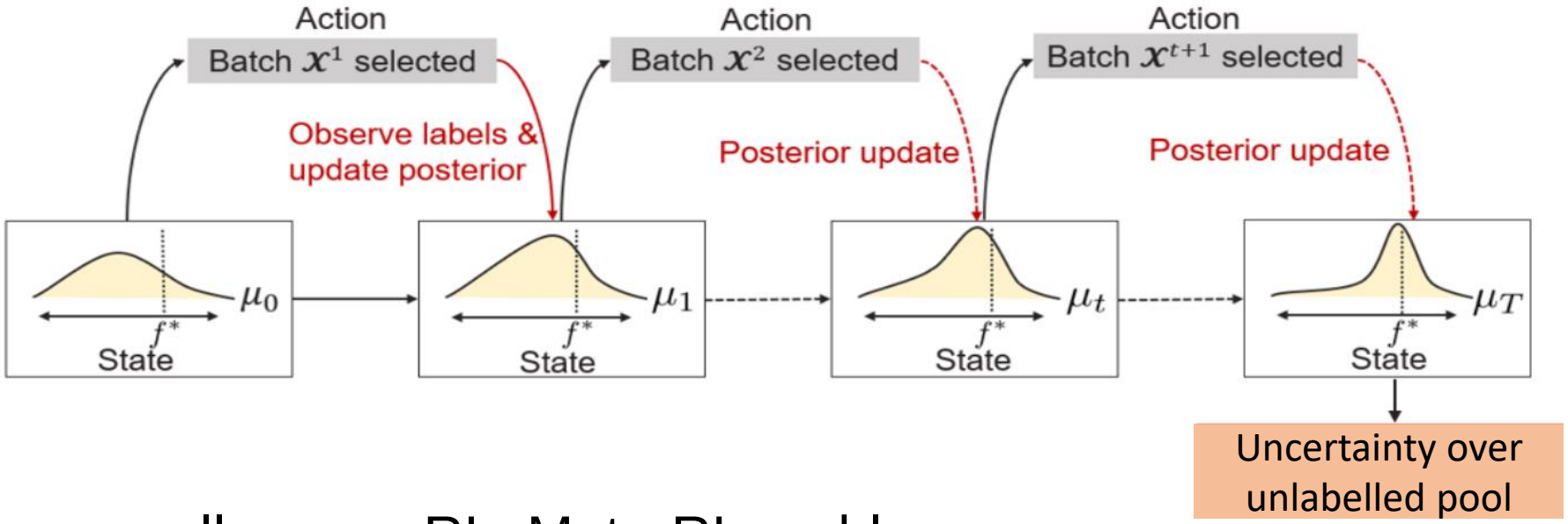
Connecting back…

- Bayesian bandits as an MDP with unknown reward function
  - ✓ K arms with means $\alpha = (\alpha_1, \cdots, \alpha_K)$ with $\alpha \sim \nu_0$ (Prior)
  - ✓ Actions A_t – which arm to pull
  - ✓ States $S_t = \phi$, with $P(S_{t+1} = \phi | S_t, A_t) = 1$
  - ✓ Reward $R_t = q_\alpha(\cdot\,|A_t, S_t) = q_\alpha(\cdot\,|A_t)$

- Bayesian bandits as BAMDP
  - ✓ Maintain a posterior ν over α, and define hyper-state $\tilde{S}_t = (S_t, \nu_t) = \nu_t$
  - ✓ $\tilde{P}\left(\tilde{S}_{t+1}, R_t \middle| \tilde{S}_t, A_t\right) = \mathbb{1}(\nu_{t+1} | R_t, A_t, \nu_t)\left(\int q_\alpha(R_t|A_t)\, d\nu_t(\alpha)\right)$
  - ✓ Bayesian Adaptive MDP $\widetilde{\mathcal{M}} = (\tilde{\mathcal{S}}, \mathcal{A}, \tilde{P})$

Connecting back…

- Similarly, we can express other decision-making problems as BAMDP
  - ✓ Bayesian Optimization
  - ✓ Adaptive data collection



  - ✓ More generally – any RL, Meta-RL problem

Conventional Solution Approaches

- Offline Value Approximation – (approximately solves BAMDP) – intractable for most domains
- Online near-myopic value approximation
  - ✓ Bayesian Dynamic Programming: extends Thompson sampling to BAMDP – sample a model θ from μ_t and take best action A_t according to θ (solve MDP θ).
  - ✓ Value of information heuristic: extends knowledge gradient ideas to BAMDP
- Online Tree search approximation
  - ✓ Perform a forward search in the space of hyper-states.

Two key challenges

- How to solve the BAMDP?
  - ✓ Continuous state space due to posteriors
  - ✓ Number of possible hyper-states increases exponentially with horizon

- How to get reliable posteriors?
  - ✓ Posterior updates are intractable for most domains
  - ✓ Further the current UQ methodologies (such as BNN etc.) suffer from following challenges
    - ✓ How much we should sharpen our belief as we see more data points – requires learning priors to effectively quantify uncertainty
  - ❖ Potential solution: UQ through meta learned sequence models

References

- Active Learning Literature Survey. Burr Settles (2010). https://burrsettles.com/pub/settles.activelearning.pdf
- A Survey on Deep Active Learning: Recent Advances and New Frontiers. Li et. al. (2024) https://arxiv.org/pdf/2405.00334
- Active Learning: A survey. Aggarwal et. al. https://charuaggarwal.net/active-survey.pdf
- Bayesian Reinforcement Learning: A Survey, Ghavamzadeh et. al. (2016) https://arxiv.org/pdf/1609.04436
- Deep Bayesian Active Learning with Image Data. Gal et. al. (2017) https://arxiv.org/abs/1703.02910