Liquid Ensemble Selection for Continual Learning

Carter Blair, Ben Armstrong, Kate Larson

UNIVERSITY OF WATERLOO

Outline

1. Intro and Motivation
2. Liquid Democracy
3. Our Approach
4. Results
5. Future Directions

UNIVERSITY OF
WATERLOO

Motivation

Traditional machine learning

Training	Deployment

Motivation

Traditional machine learning

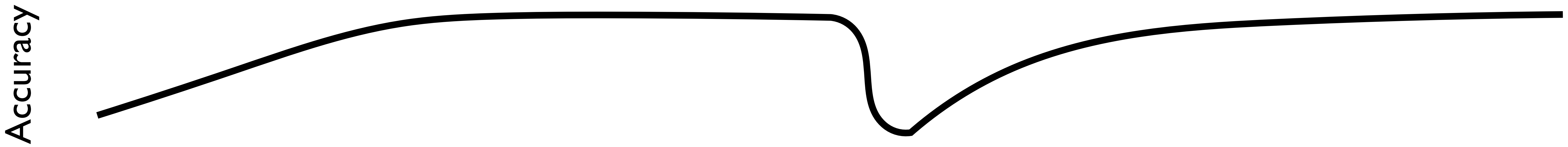
The world is non-stationary

UNIVERSITY OF WATERLOO

Motivation

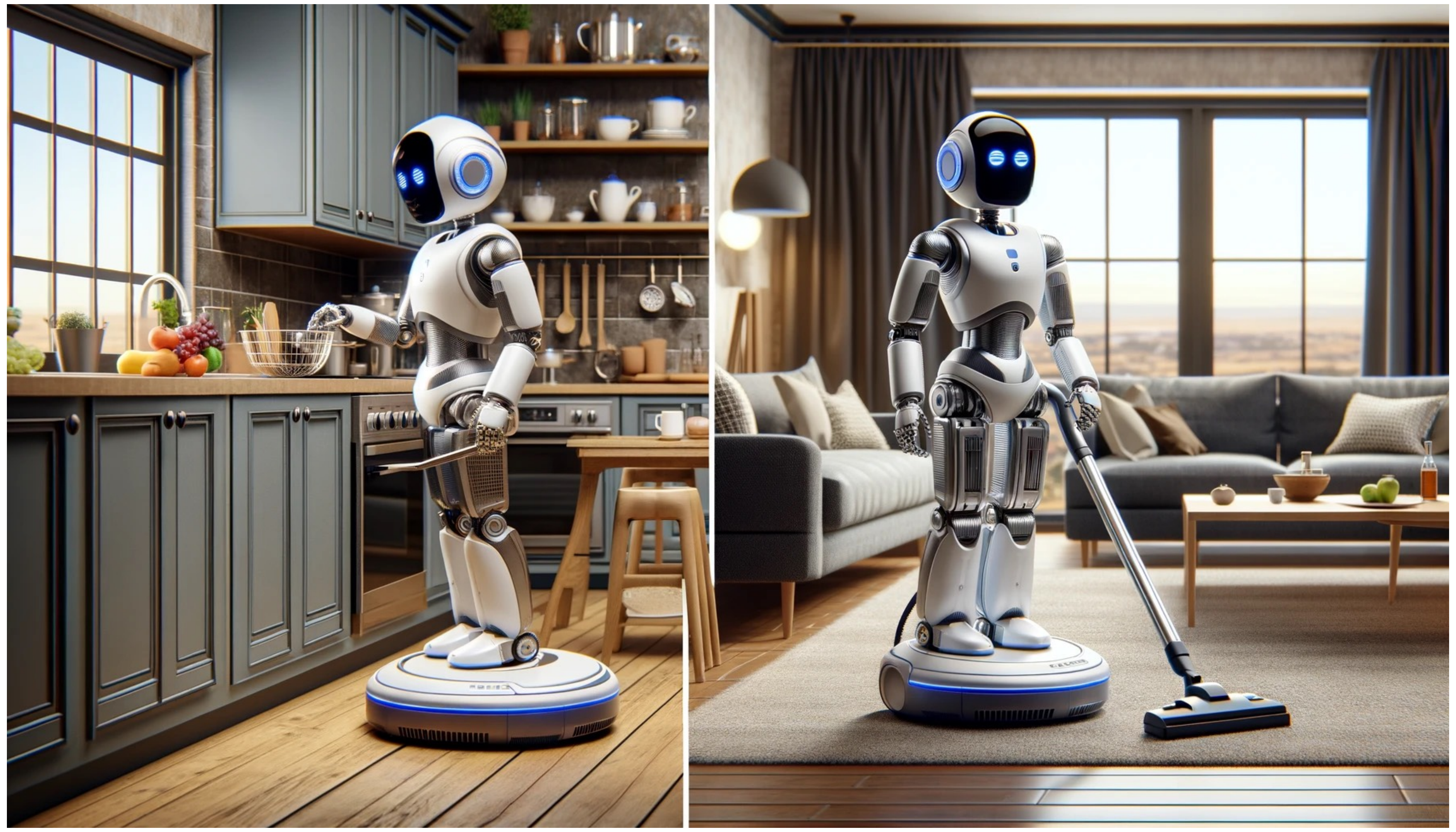
Motivation

Training

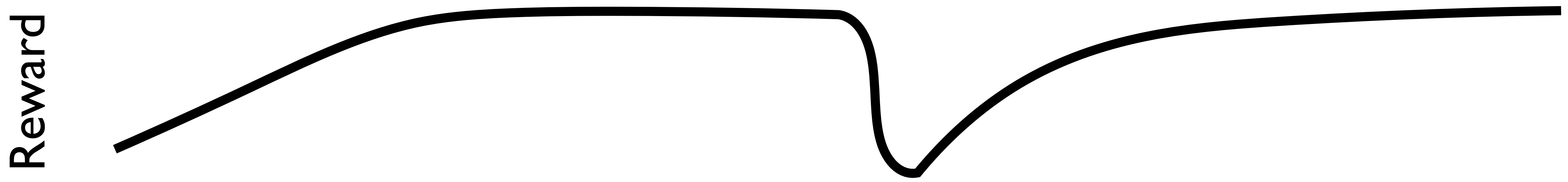


UNIVERSITY OF WATERLOO

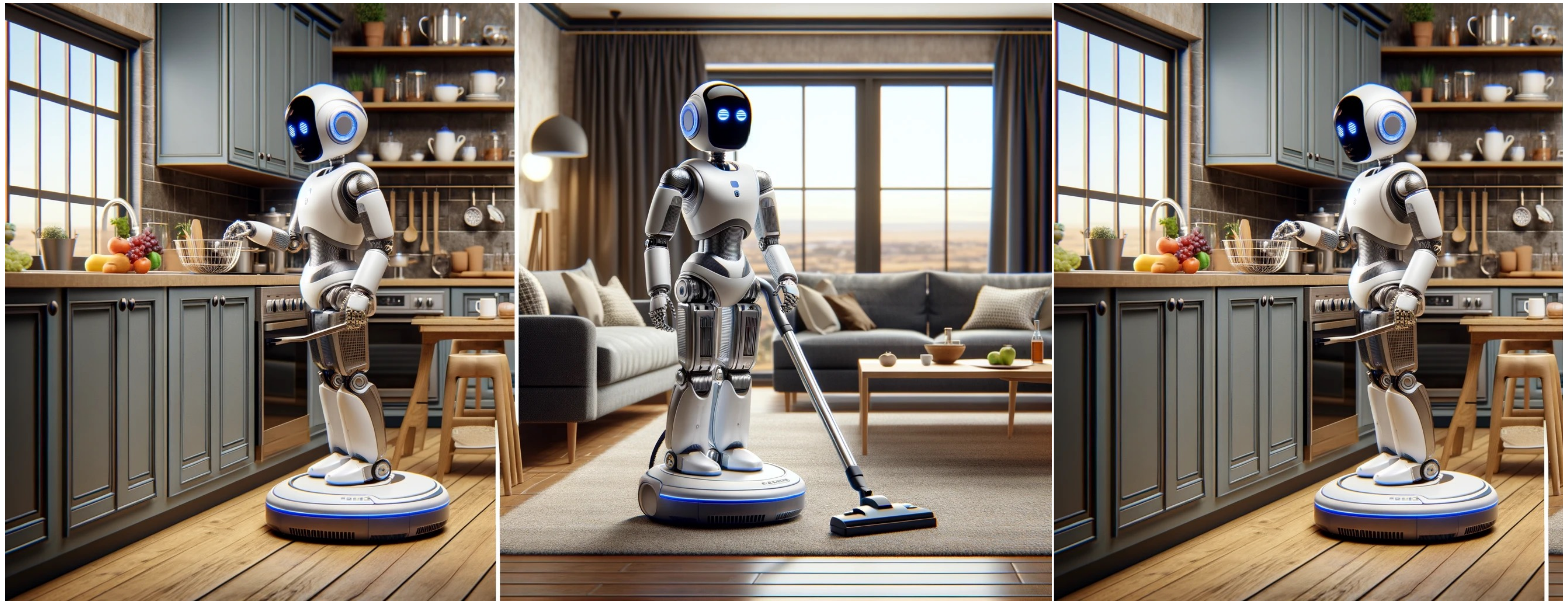
Motivation



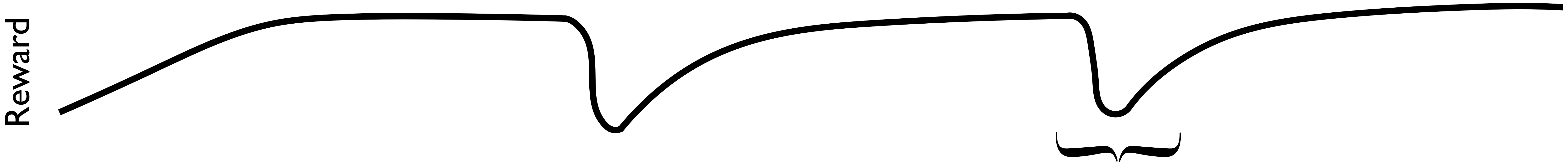
Training



Motivation



Training



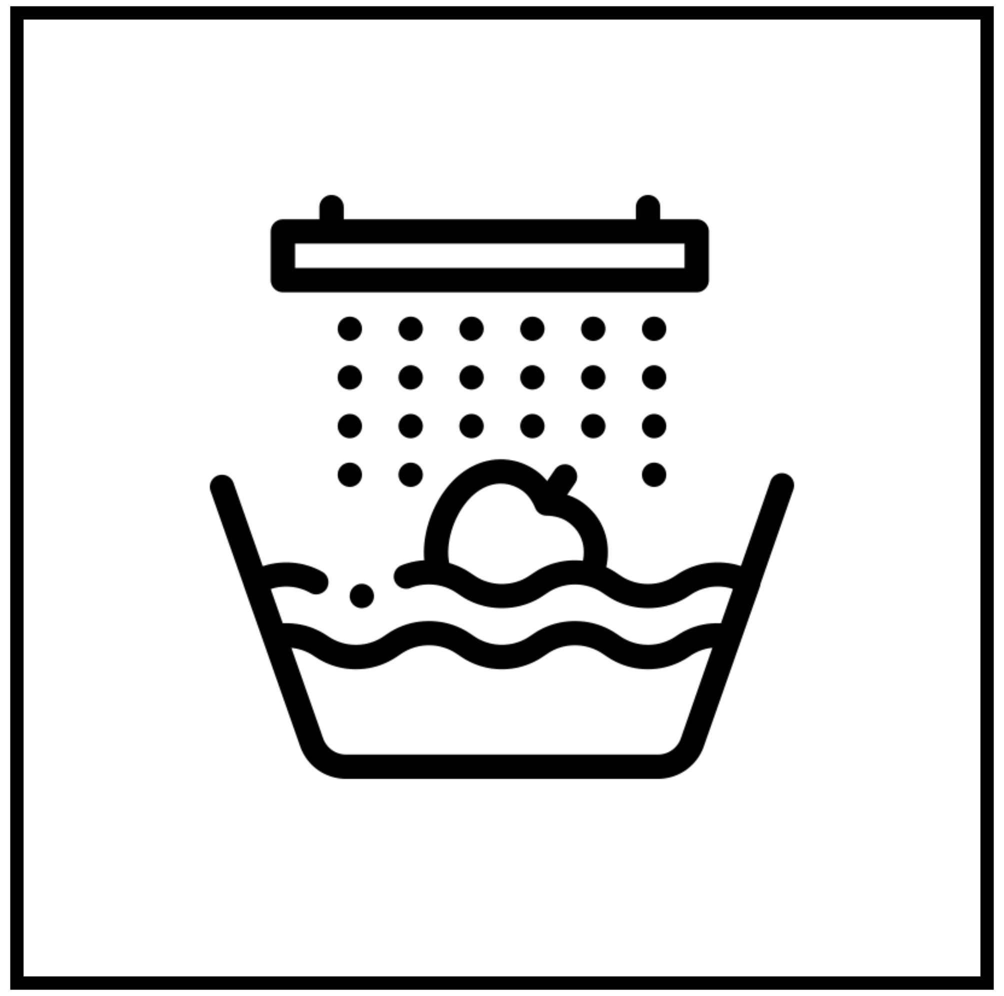
“Catastrophic Forgetting”

Motivation

Model

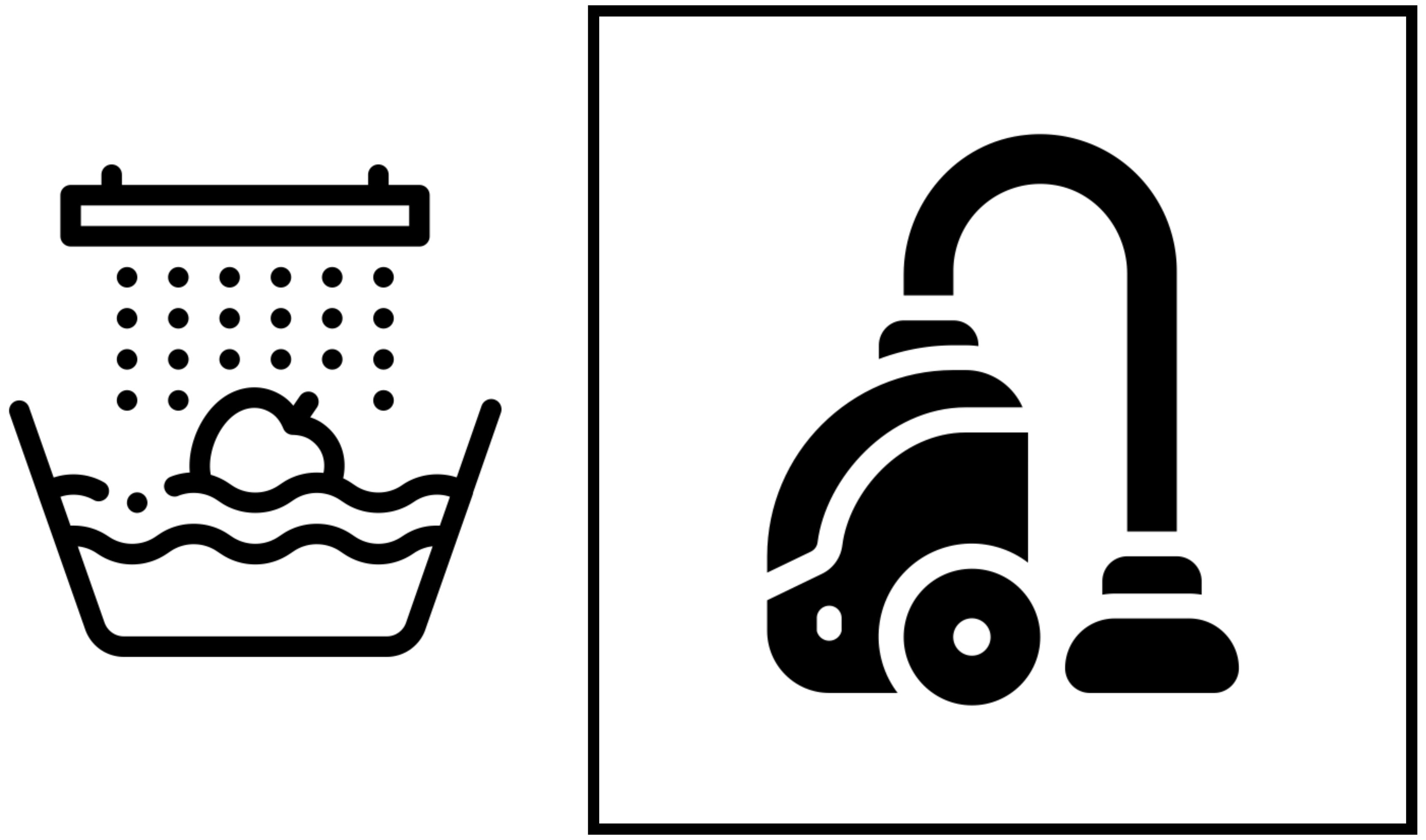
Motivation

Model



Motivation

Model



Continual Learning

Setting: Infinite stream of non-stationary data

Goal: Accumulate knowledge without forgetting what's already been learned

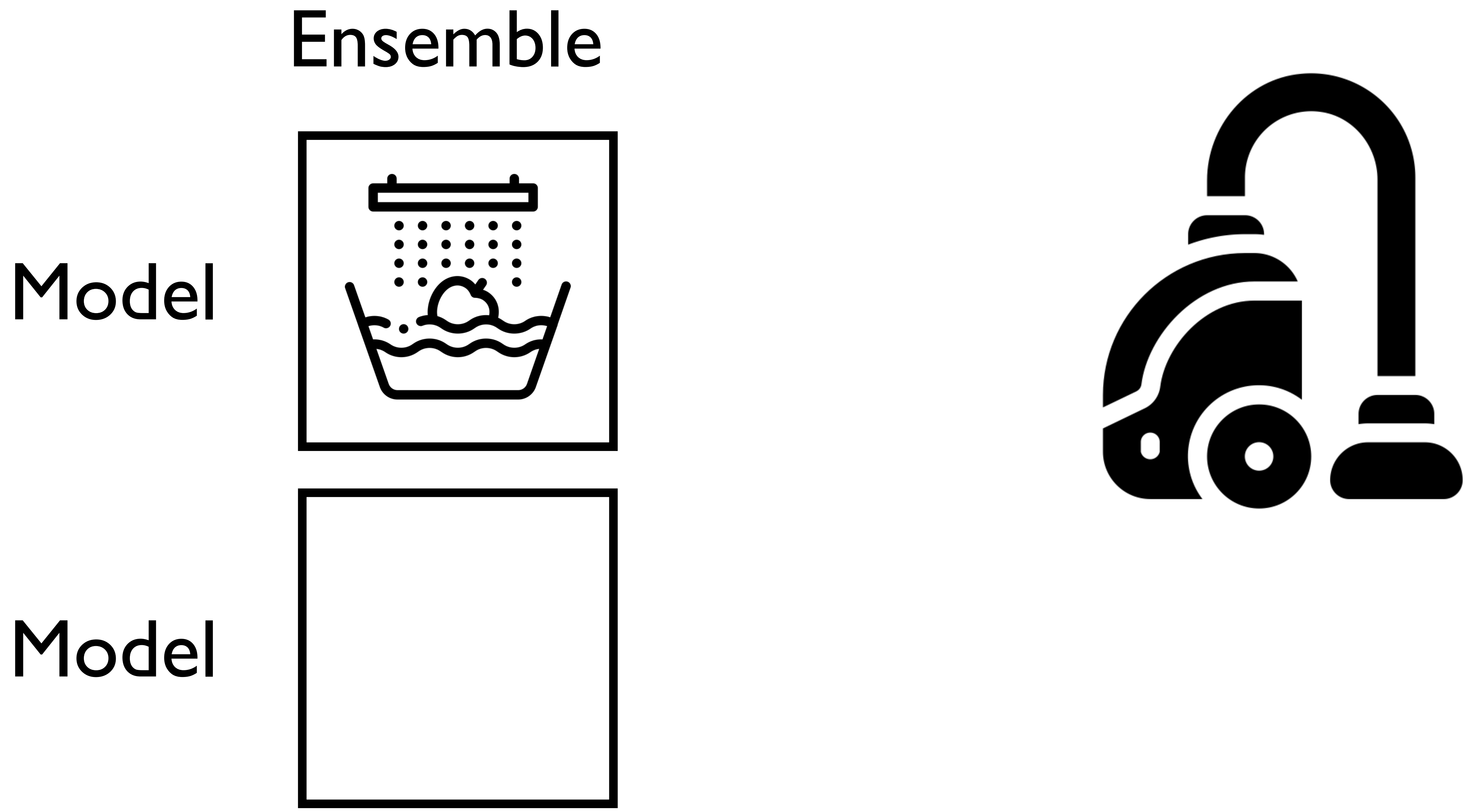
UNIVERSITY OF WATERLOO

Our Approach

Allow models to acquire new knowledge *without forgetting* old knowledge by selecting which members of an ensemble are learning at any given time

Our Approach

Allow models to acquire new knowledge *without forgetting* old knowledge by selecting which members of an ensemble are learning at any given time

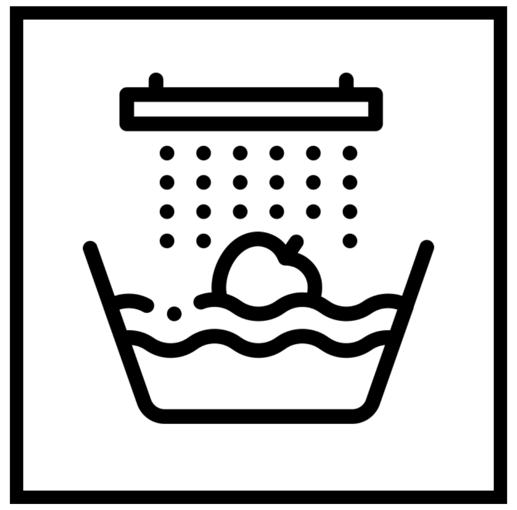


Our Approach

Allow models to acquire new knowledge *without forgetting* old knowledge by selecting which members of an ensemble are learning at any given time

Ensemble

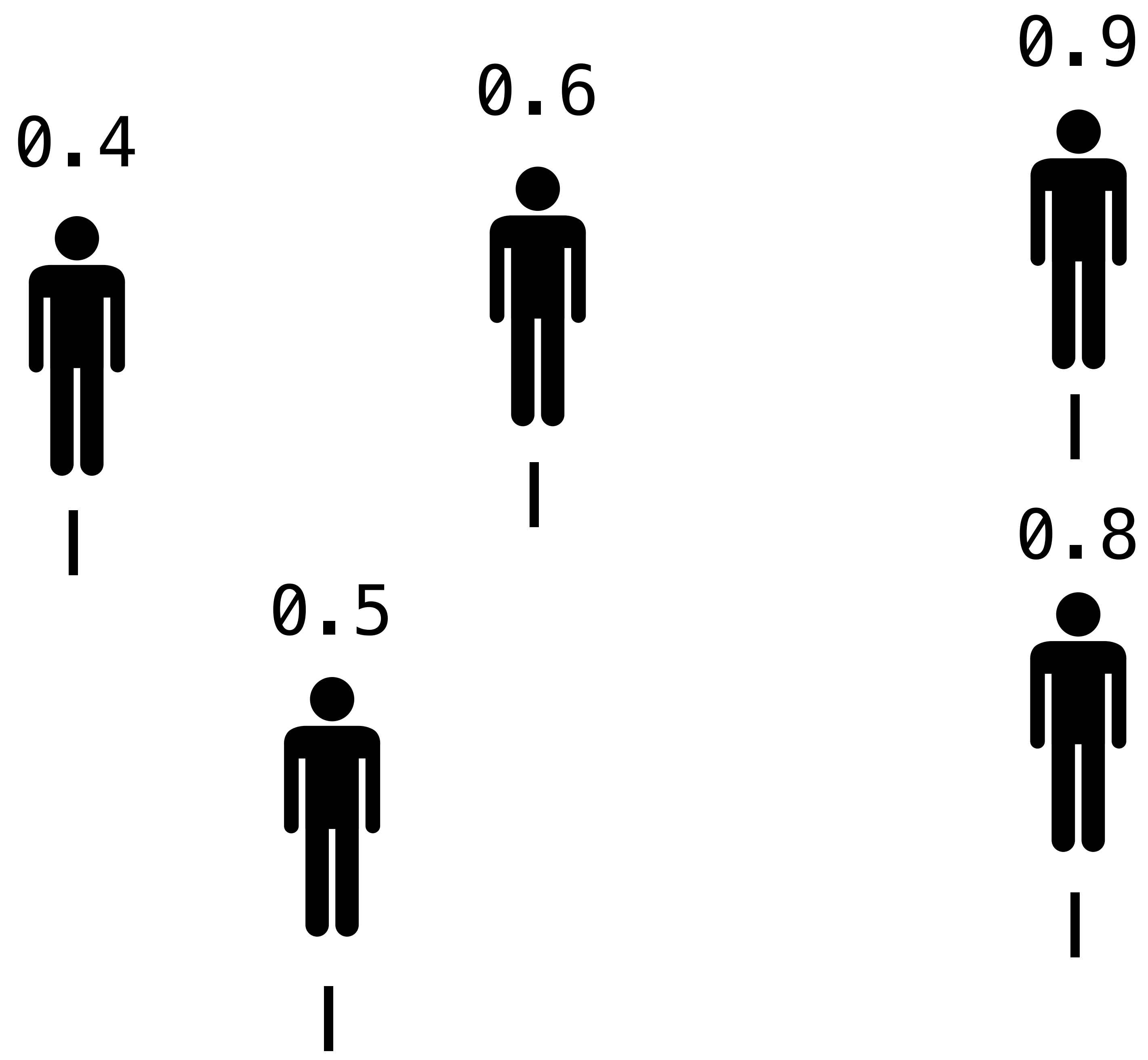
Model



Model

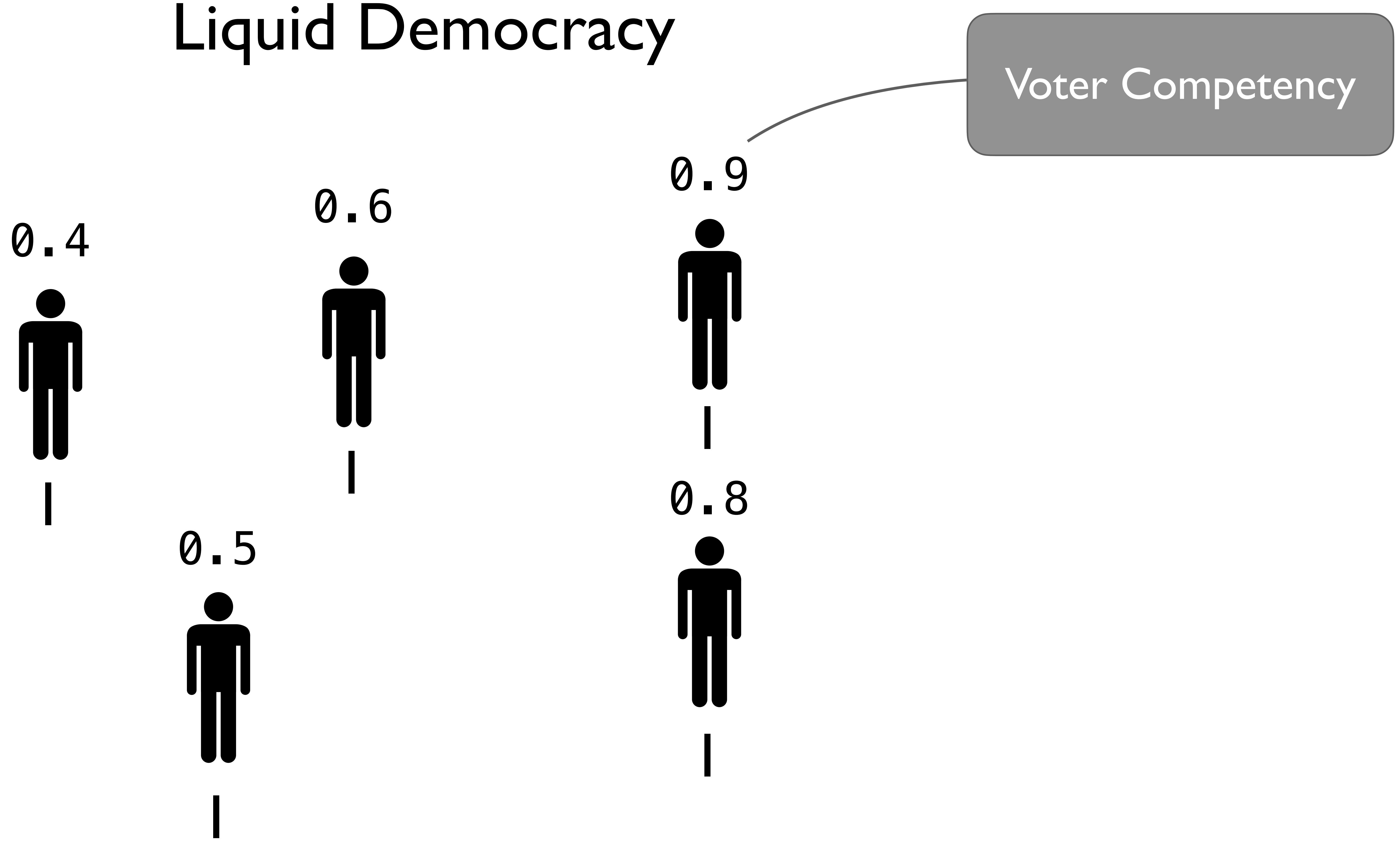
Liquid Ensemble Selection

Liquid Democracy



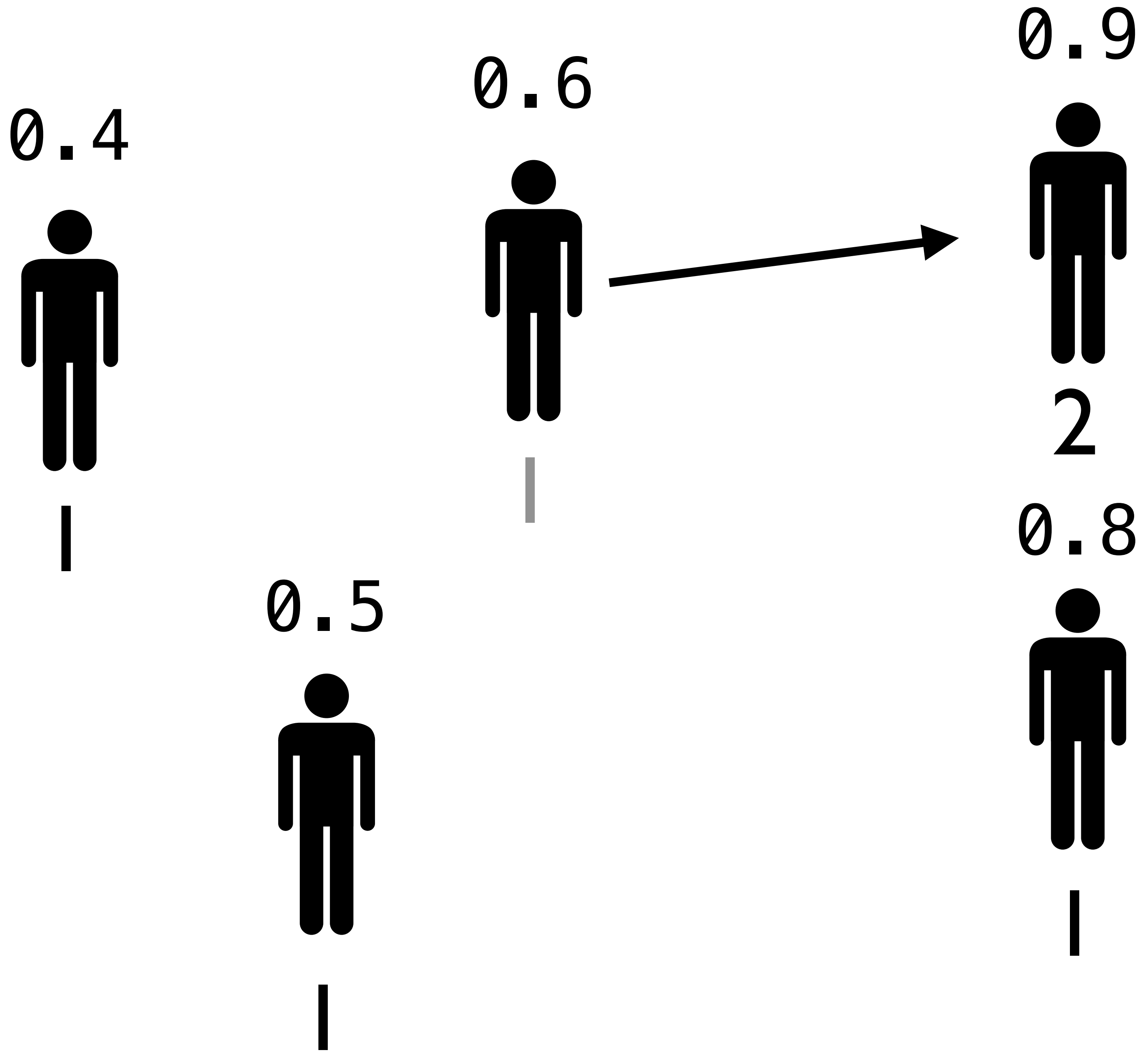
Liquid Ensemble Selection

Liquid Democracy



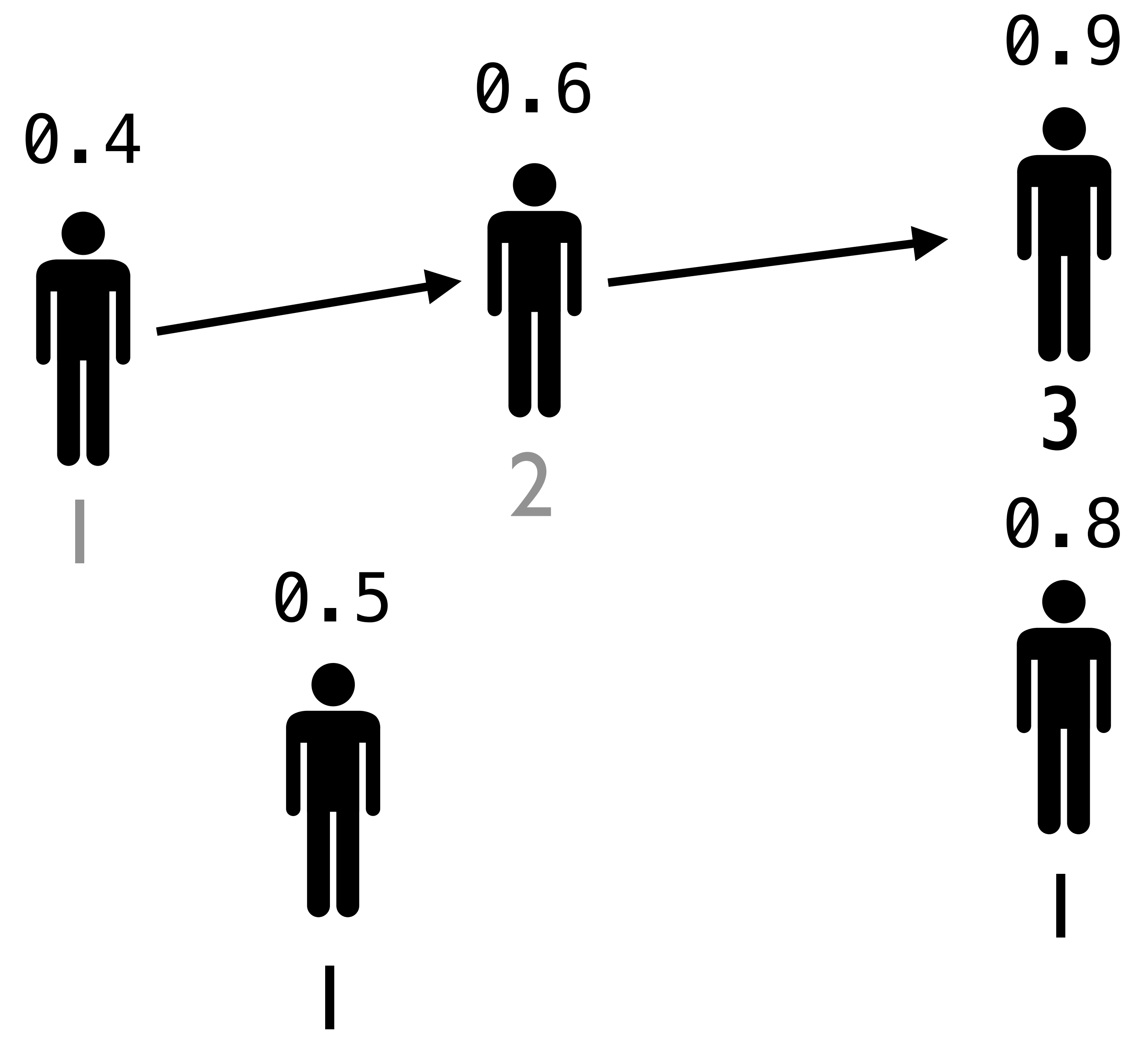
Liquid Ensemble Selection

Liquid Democracy



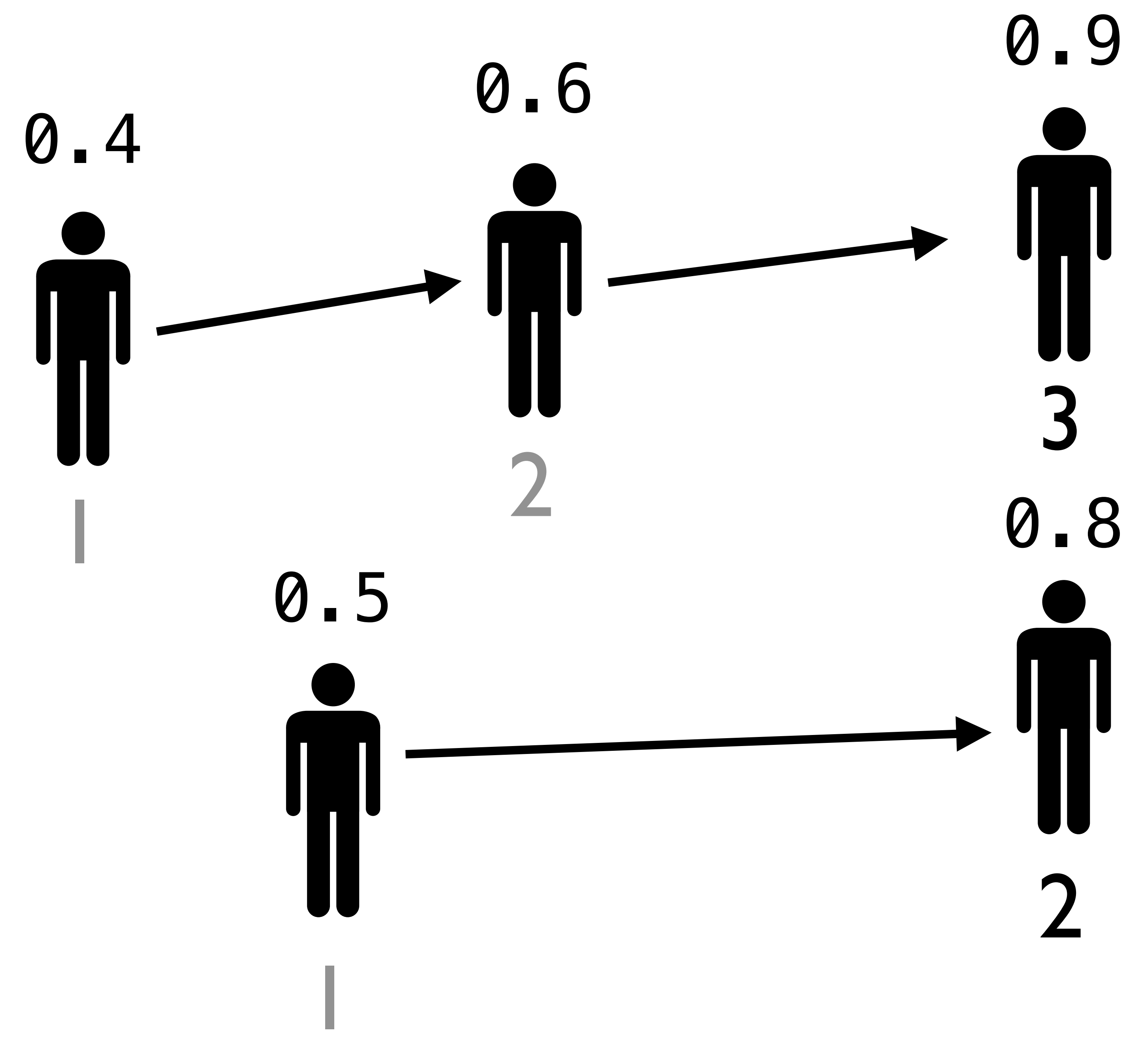
Liquid Ensemble Selection

Liquid Democracy



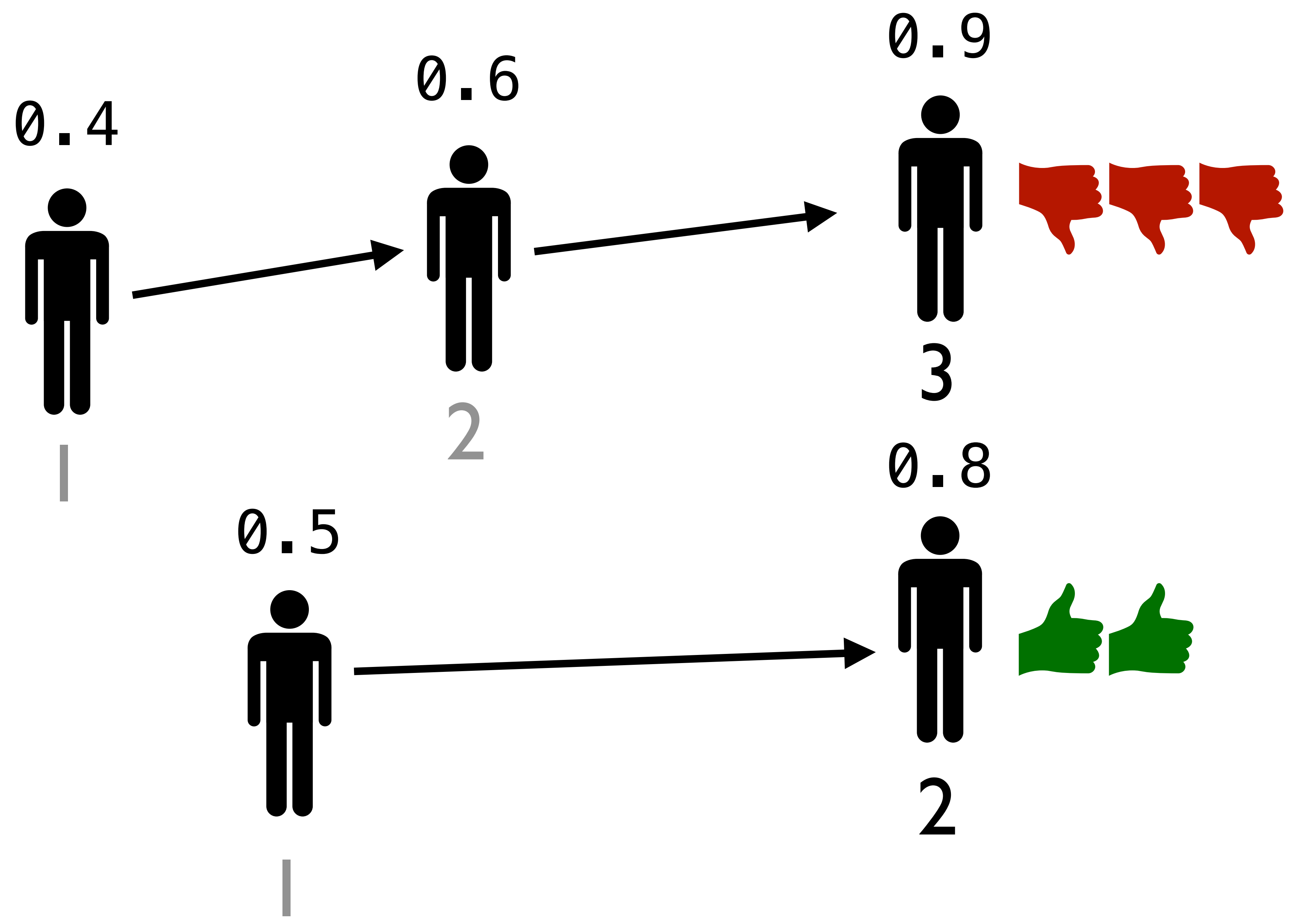
Liquid Ensemble Selection

Liquid Democracy



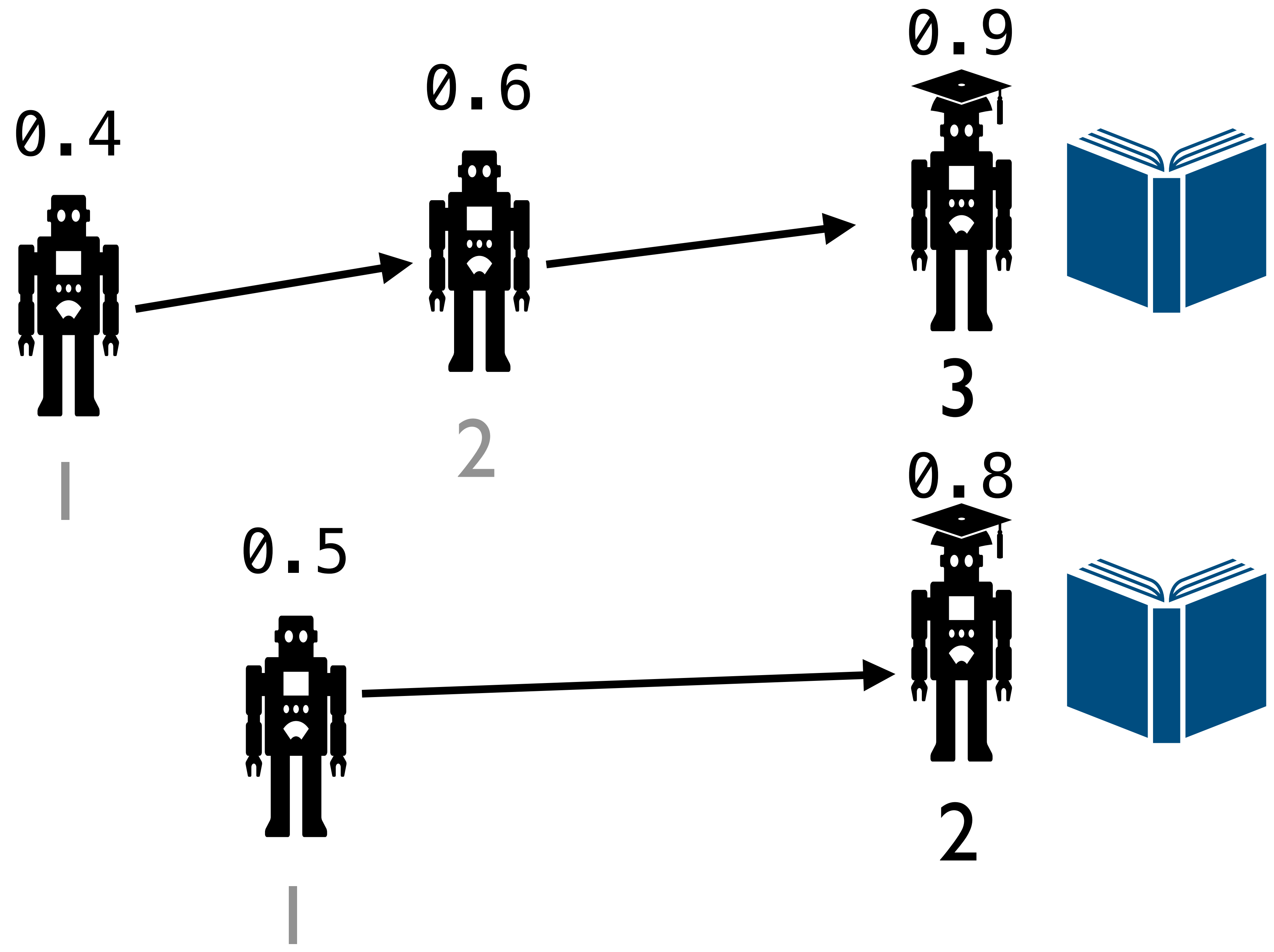
Liquid Ensemble Selection

Liquid Democracy



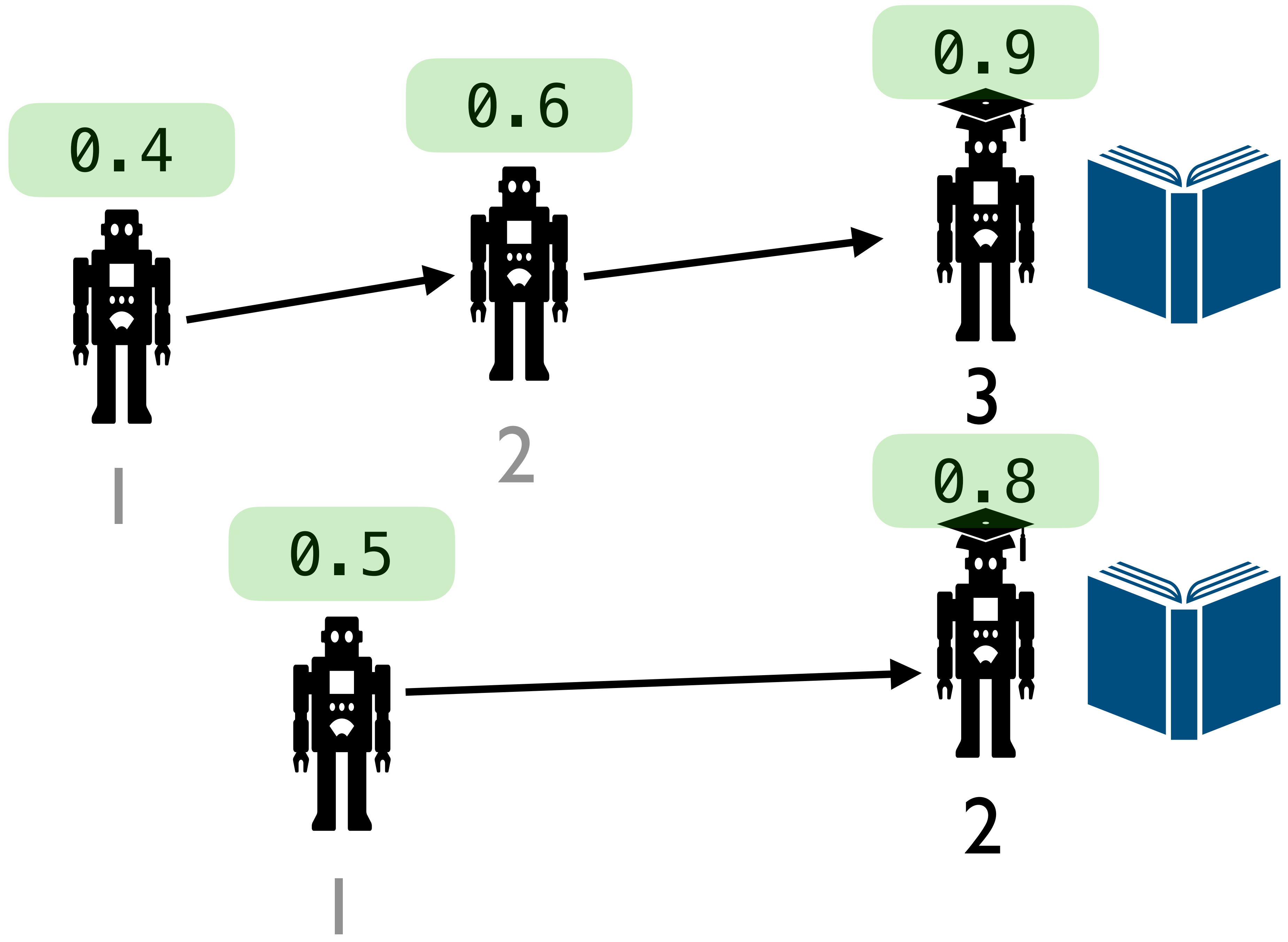
Liquid Ensemble Selection

Liquid Democracy

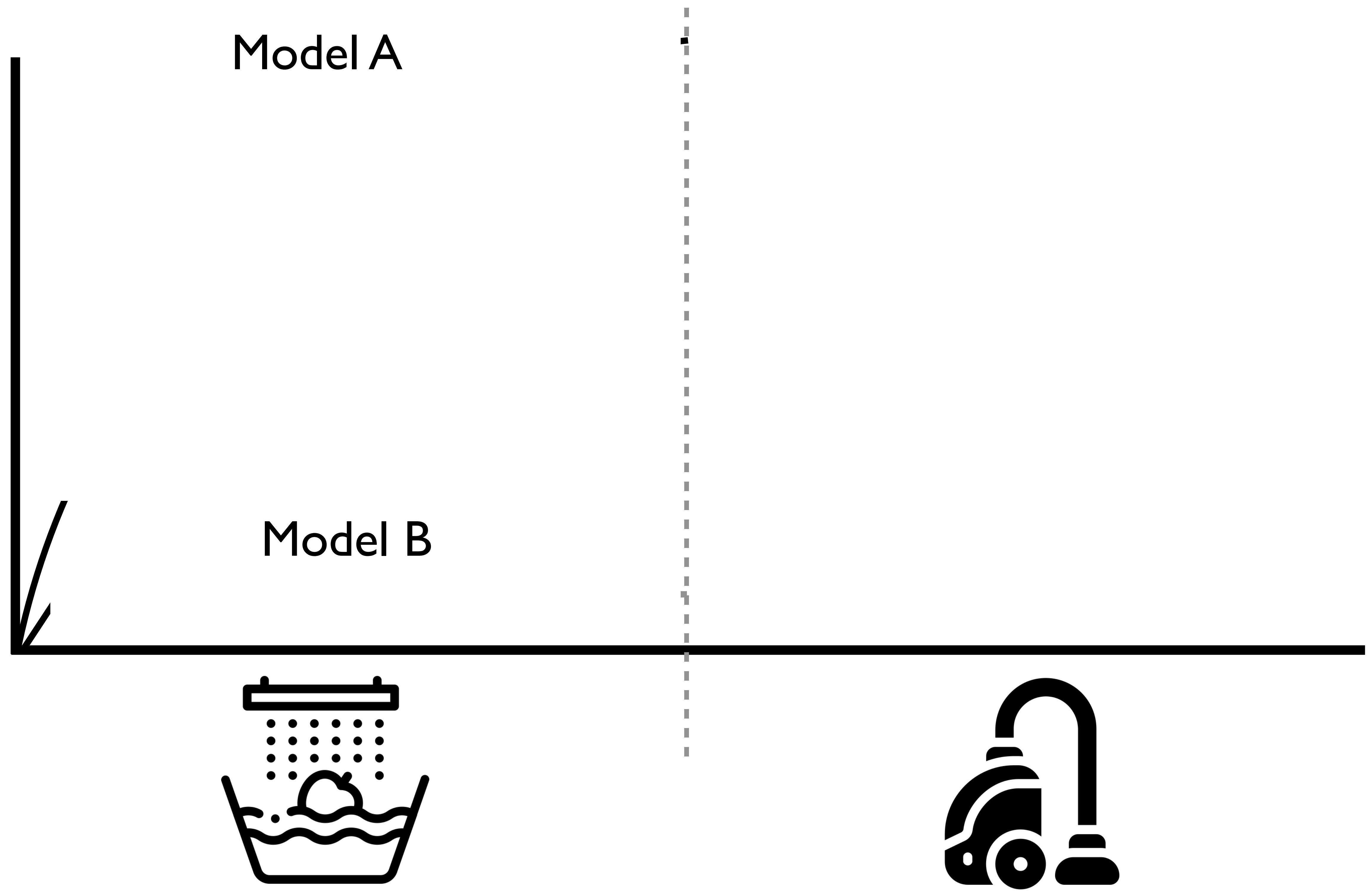


Liquid Ensemble Selection

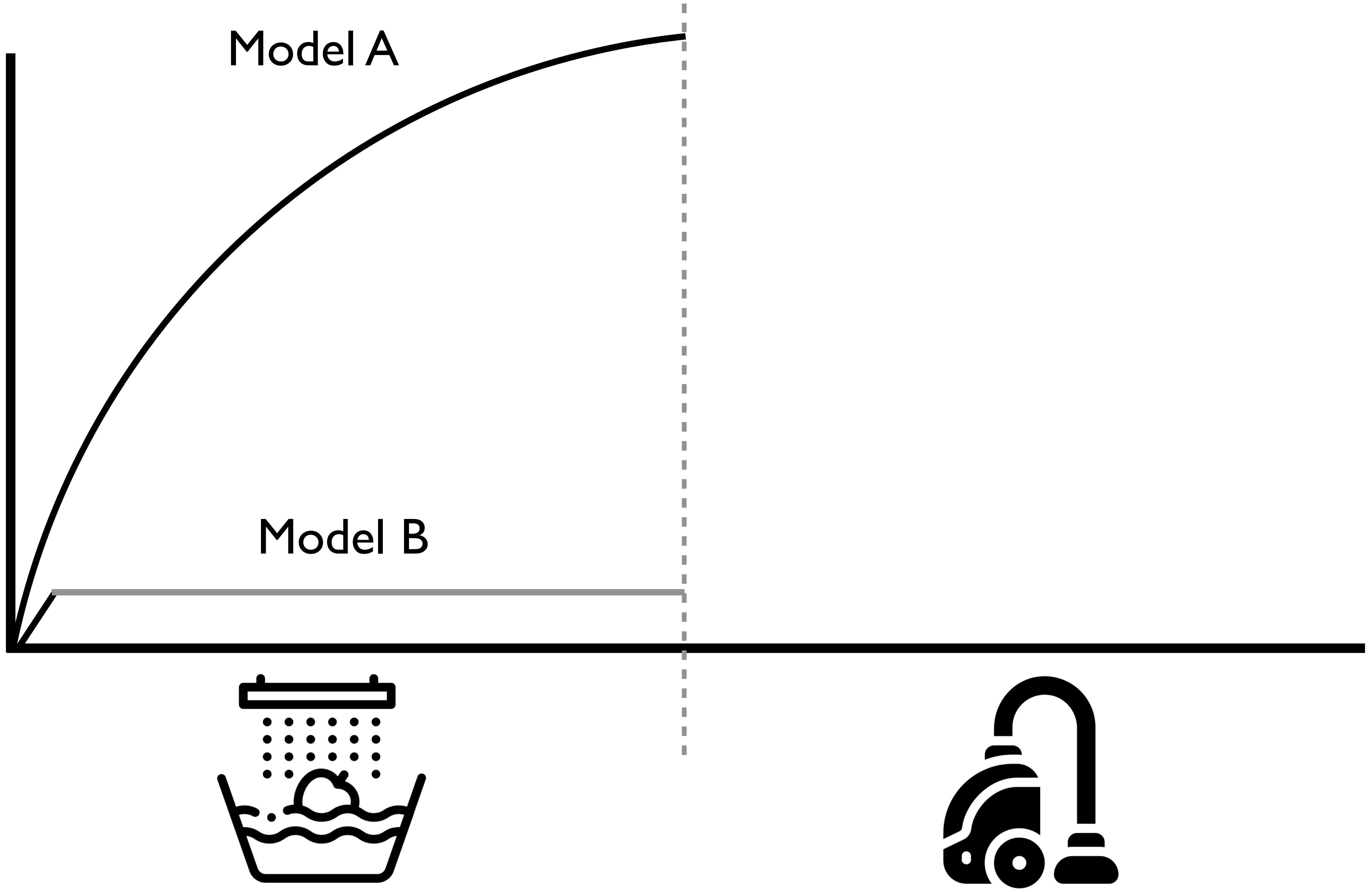
Liquid Democracy



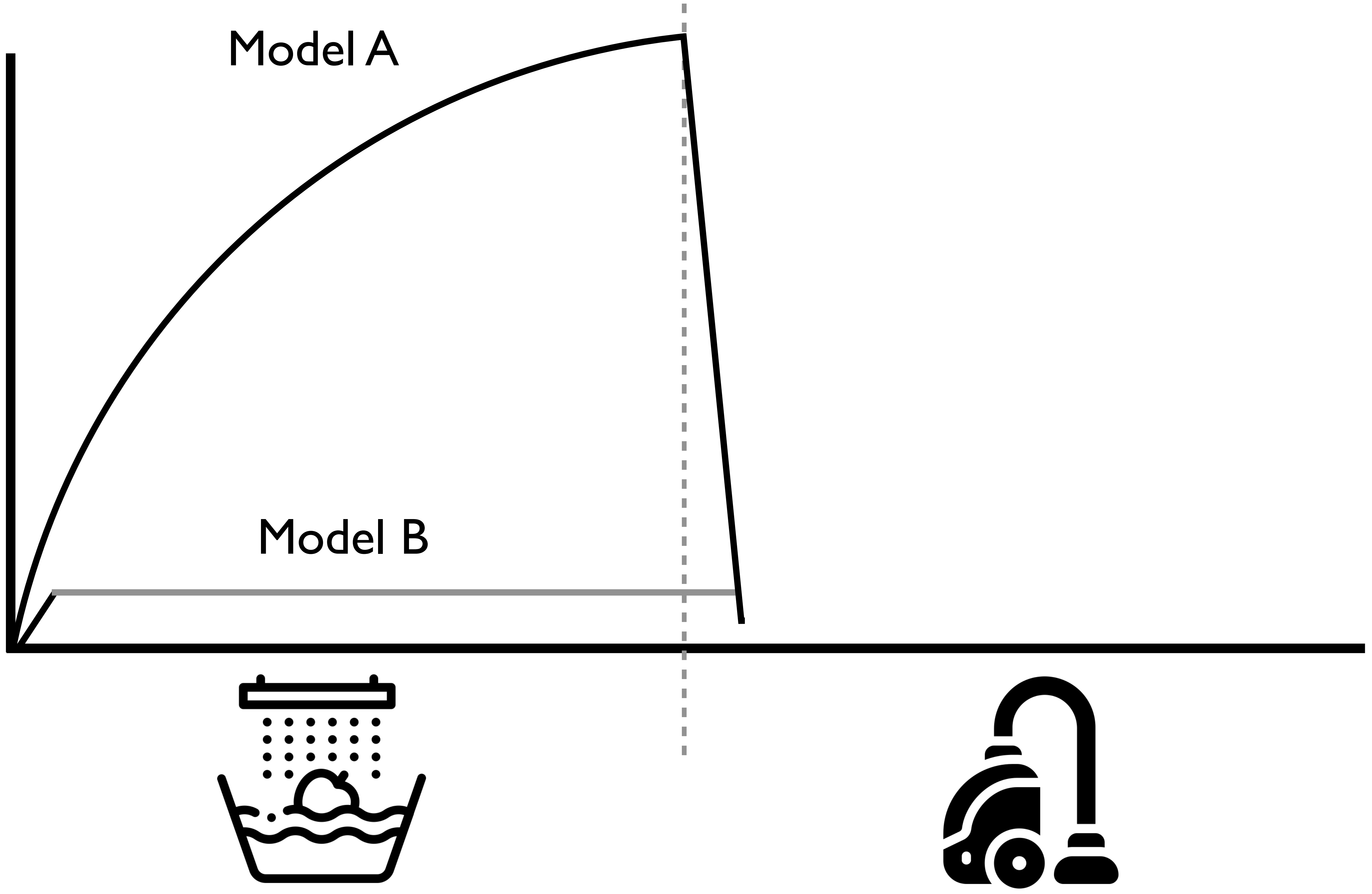
Liquid Ensemble Selection



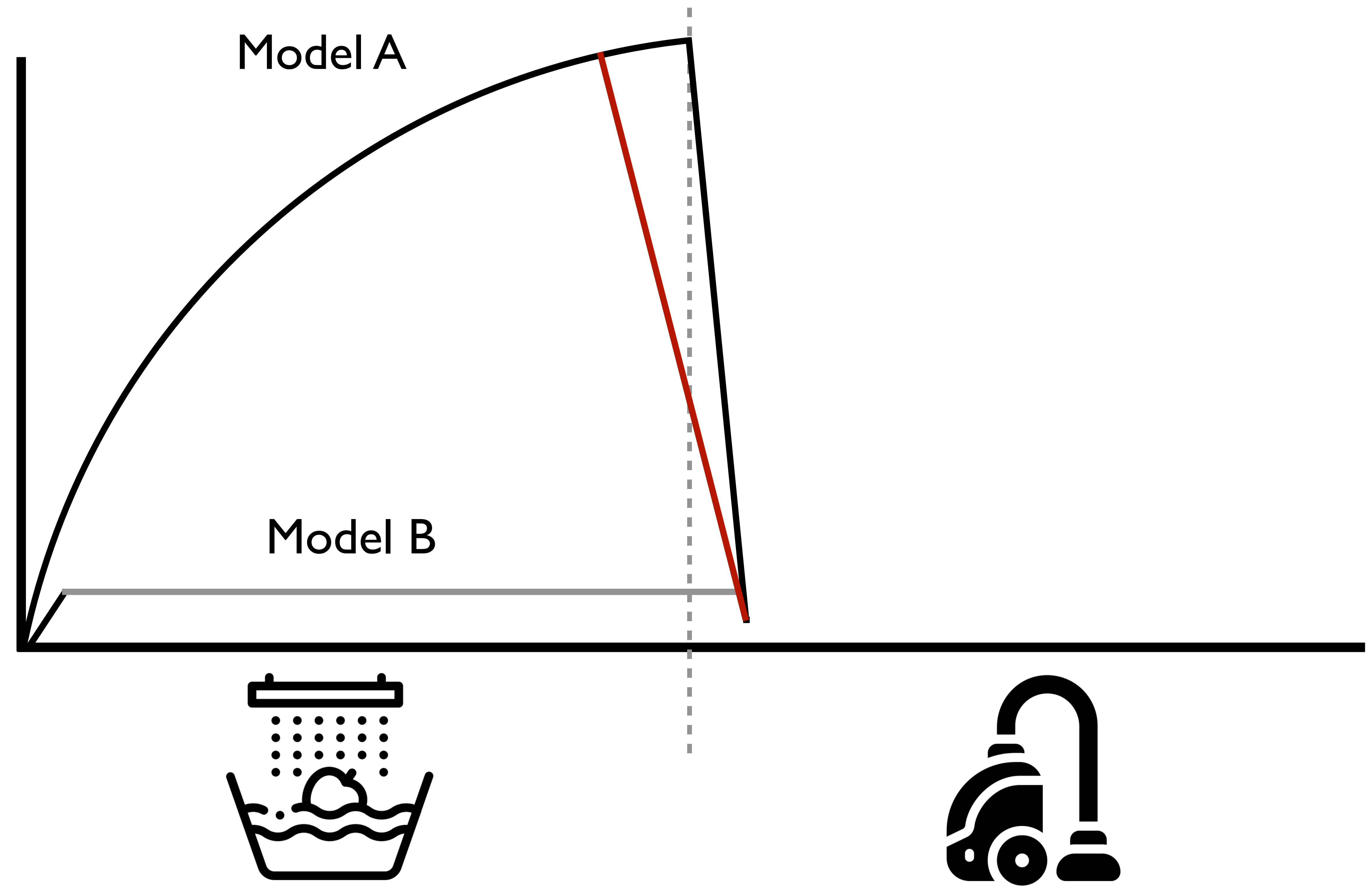
Liquid Ensemble Selection



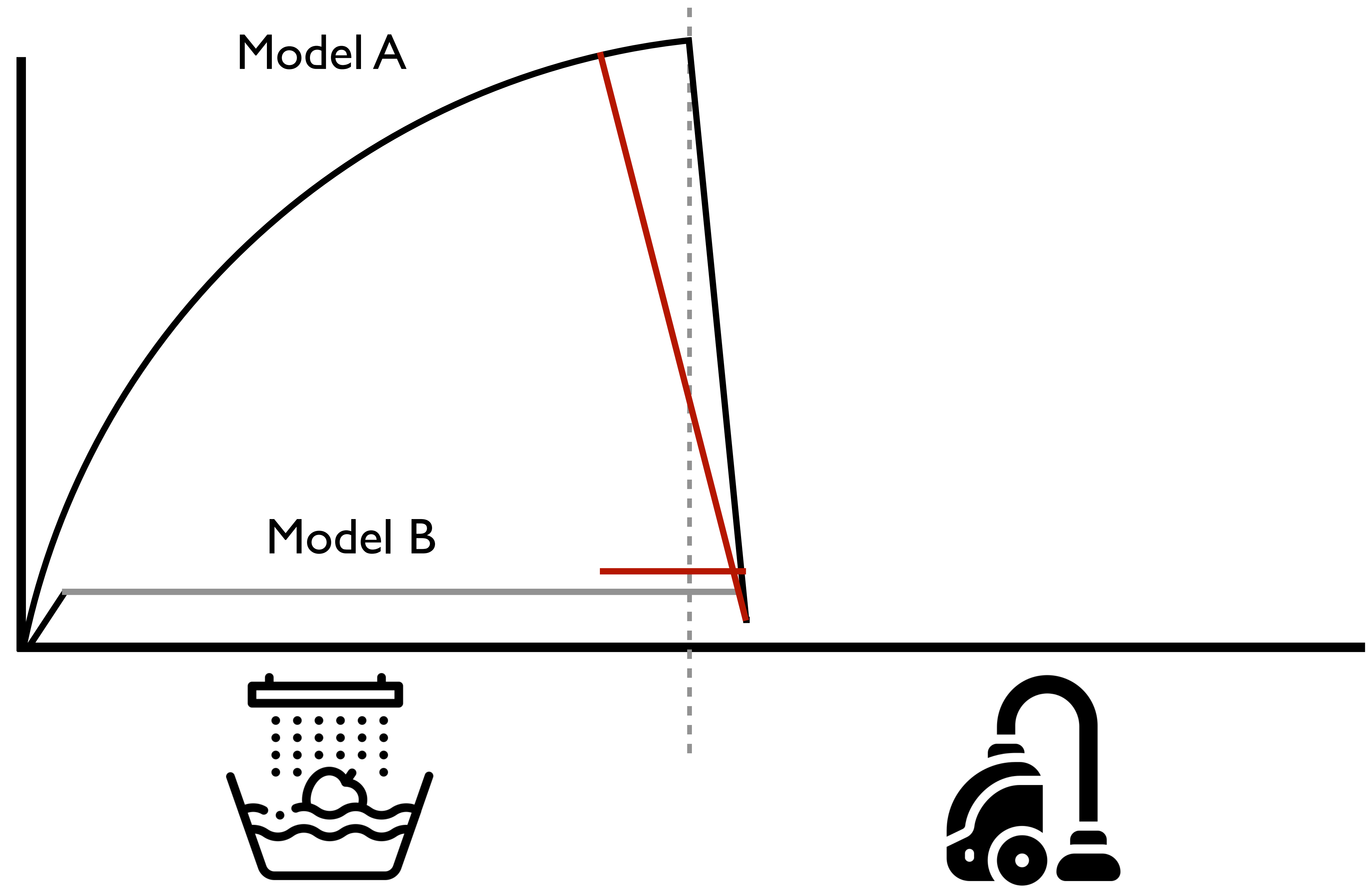
Liquid Ensemble Selection



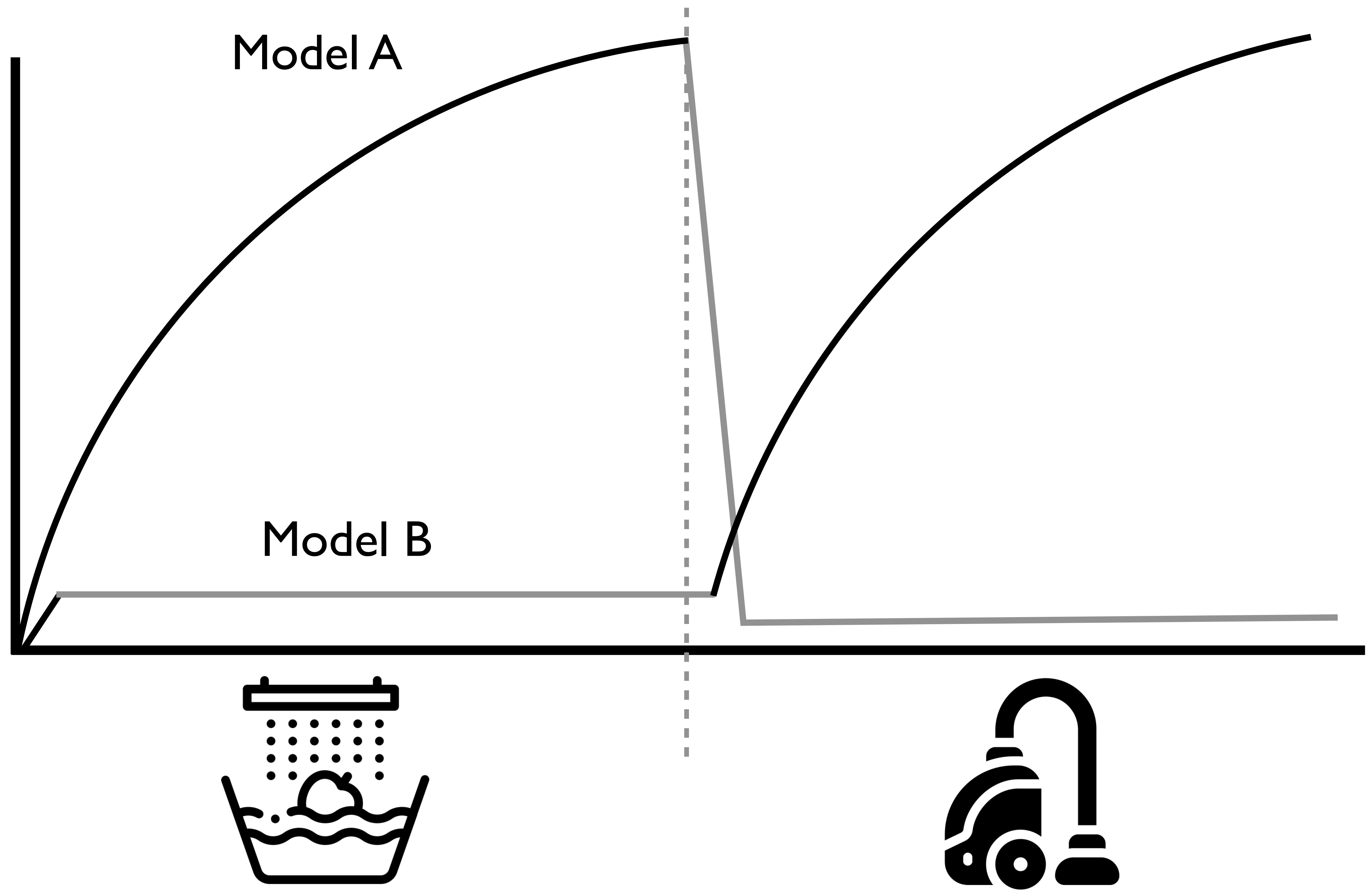
Liquid Ensemble Selection



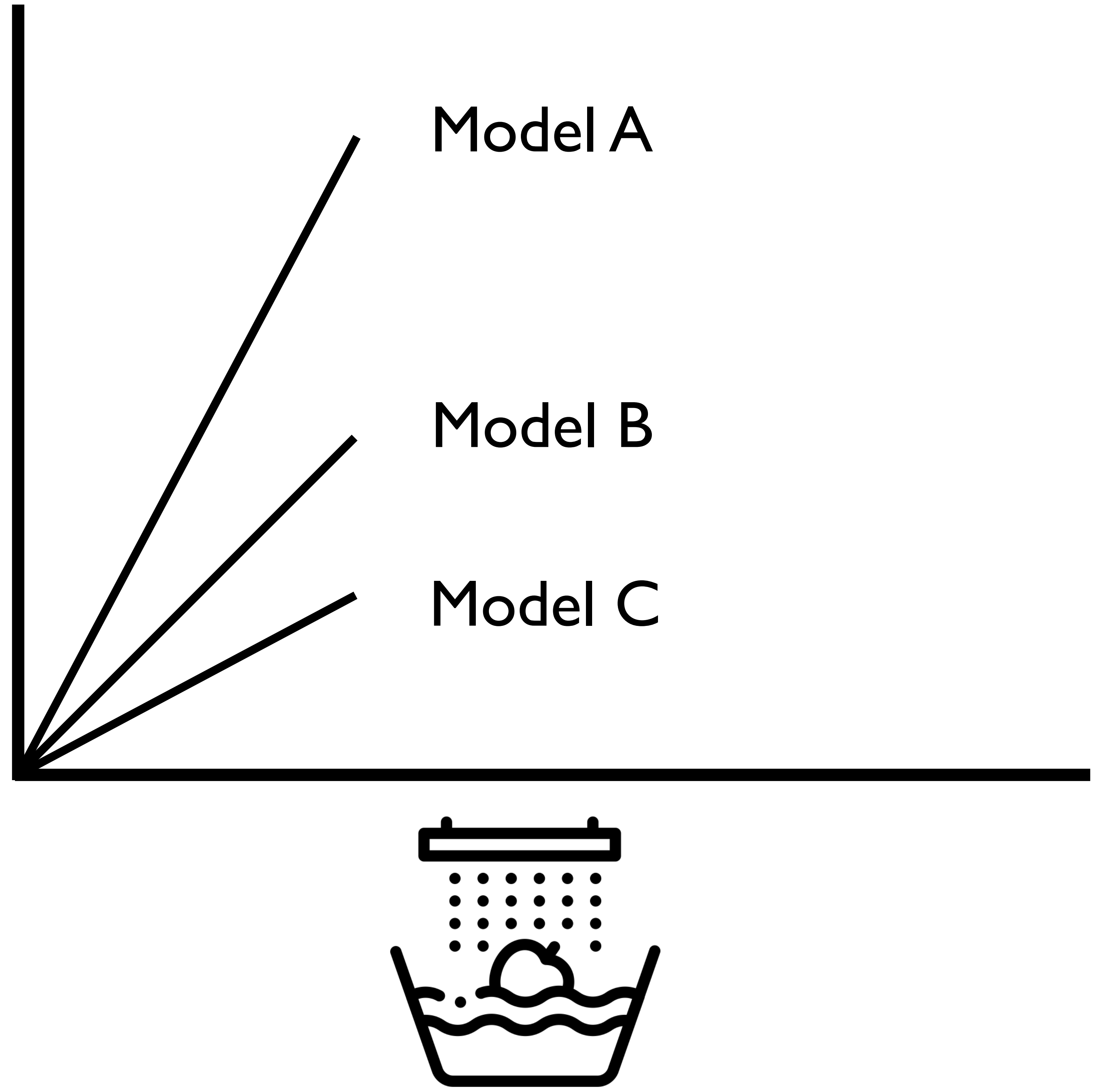
Liquid Ensemble Selection



Liquid Ensemble Selection



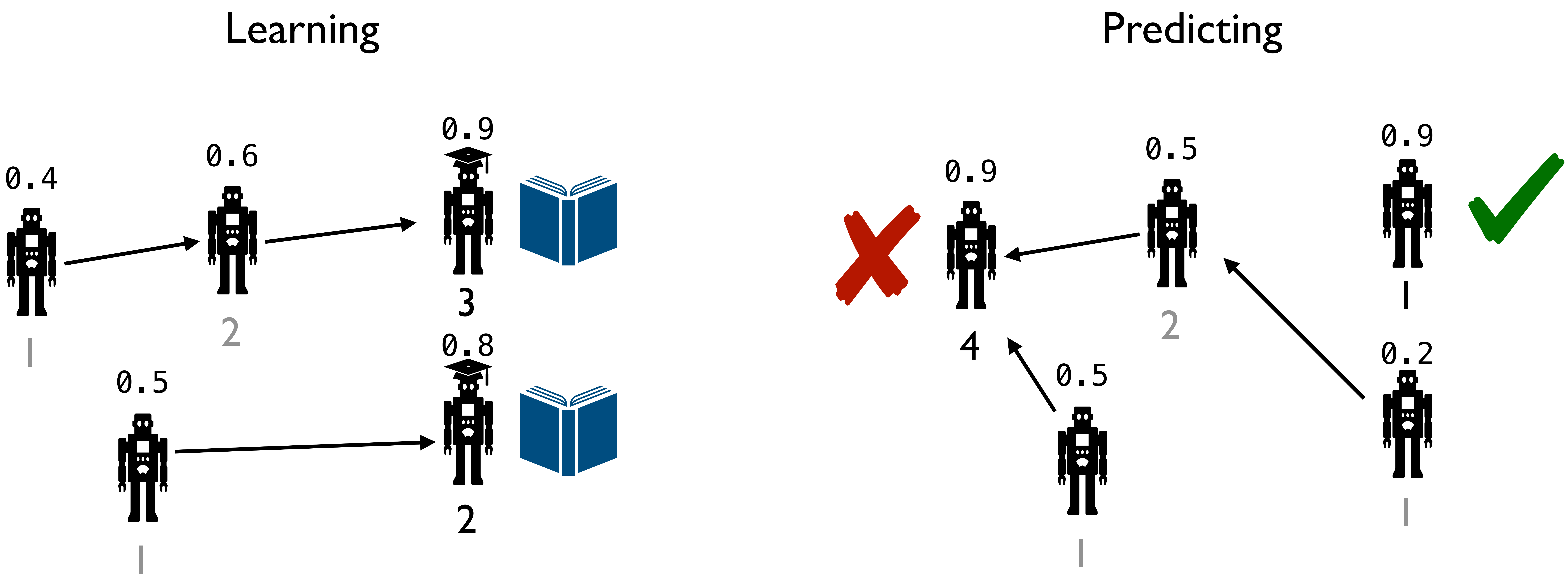
Liquid Ensemble Selection



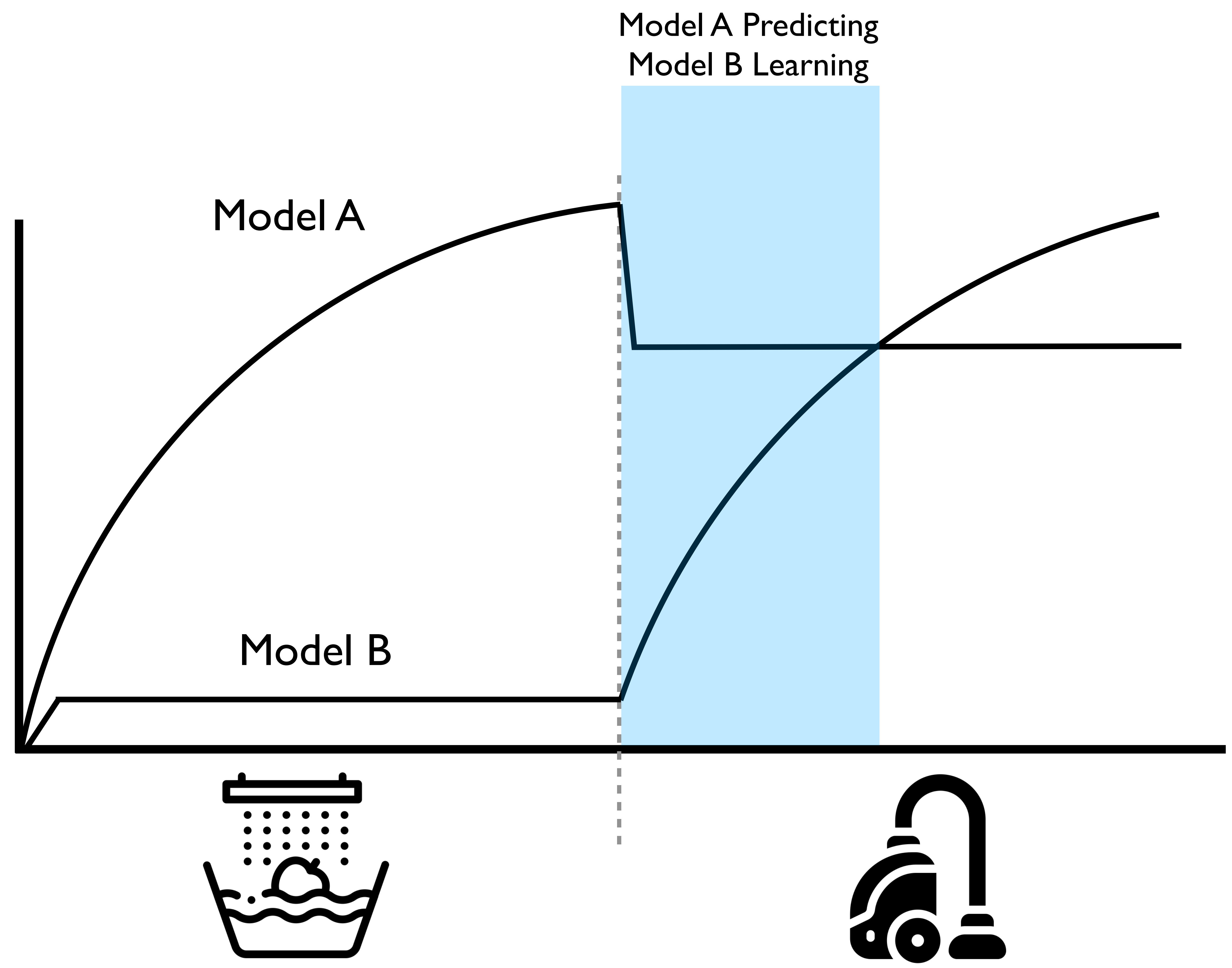
- Neural networks lose plasticity over time
- In expectation, models that have learned more will have a lower learning rate (less plasticity)
- We should delegate the learning of new tasks to models that have learned the least

Lyle, Clare, Zeyu Zheng, Evgenii Nikishin, Bernardo Avila Pires, Razvan Pascanu, and Will Dabney. "Understanding Plasticity in Neural Networks." In International Conference on Machine Learning, pp. 23190-23211. PMLR, 2023.

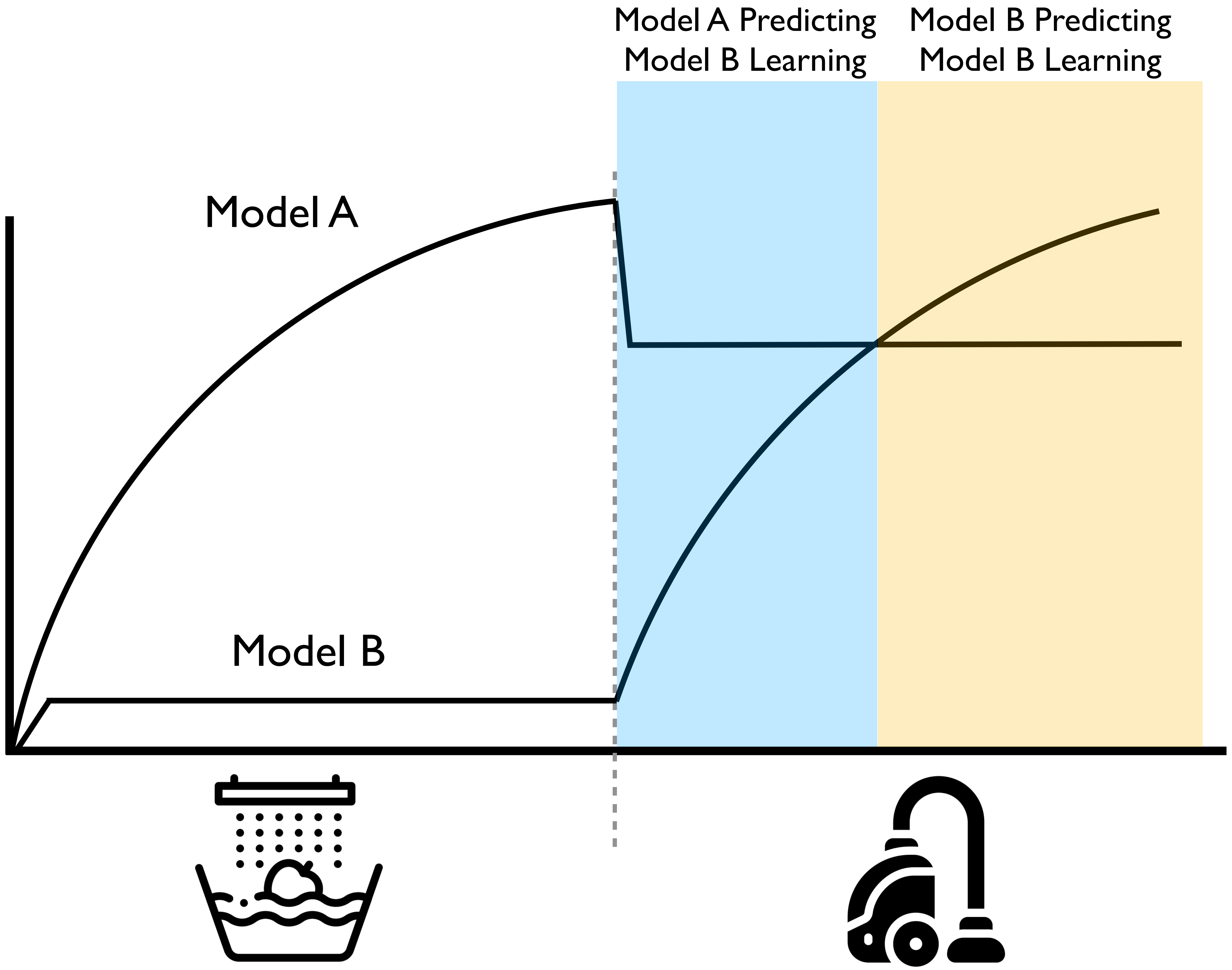
Liquid Ensemble Selection



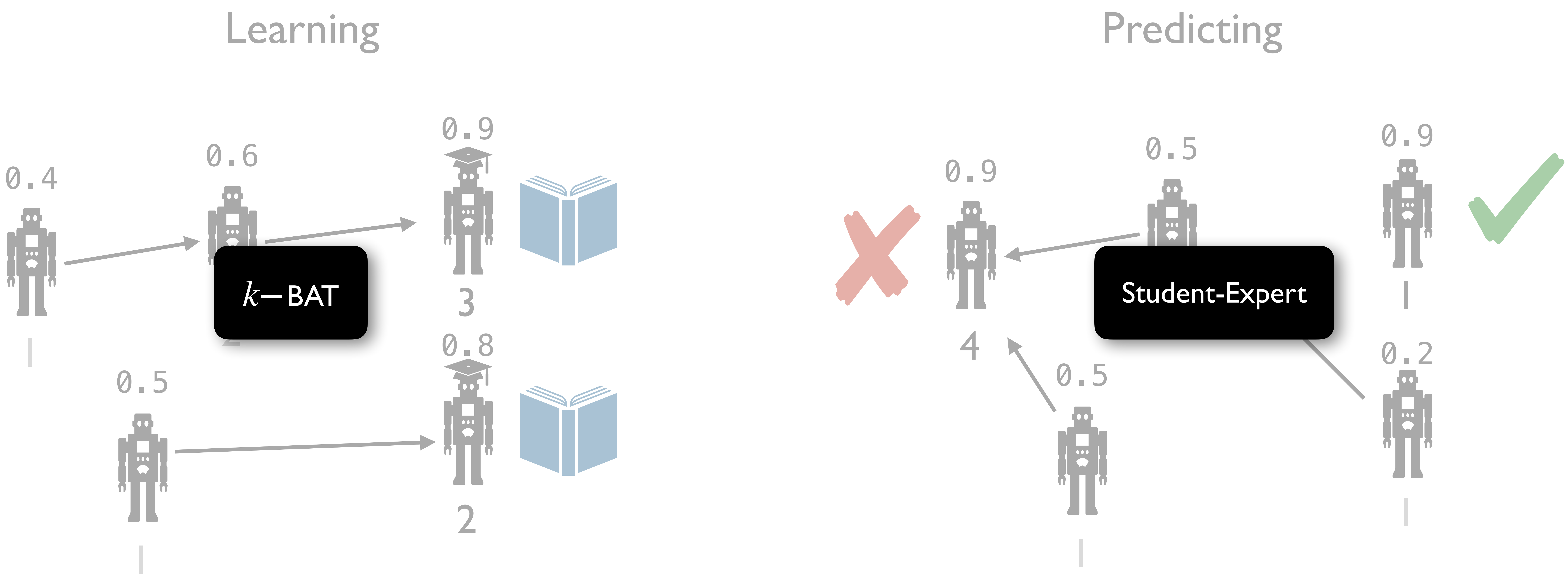
Liquid Ensemble Selection



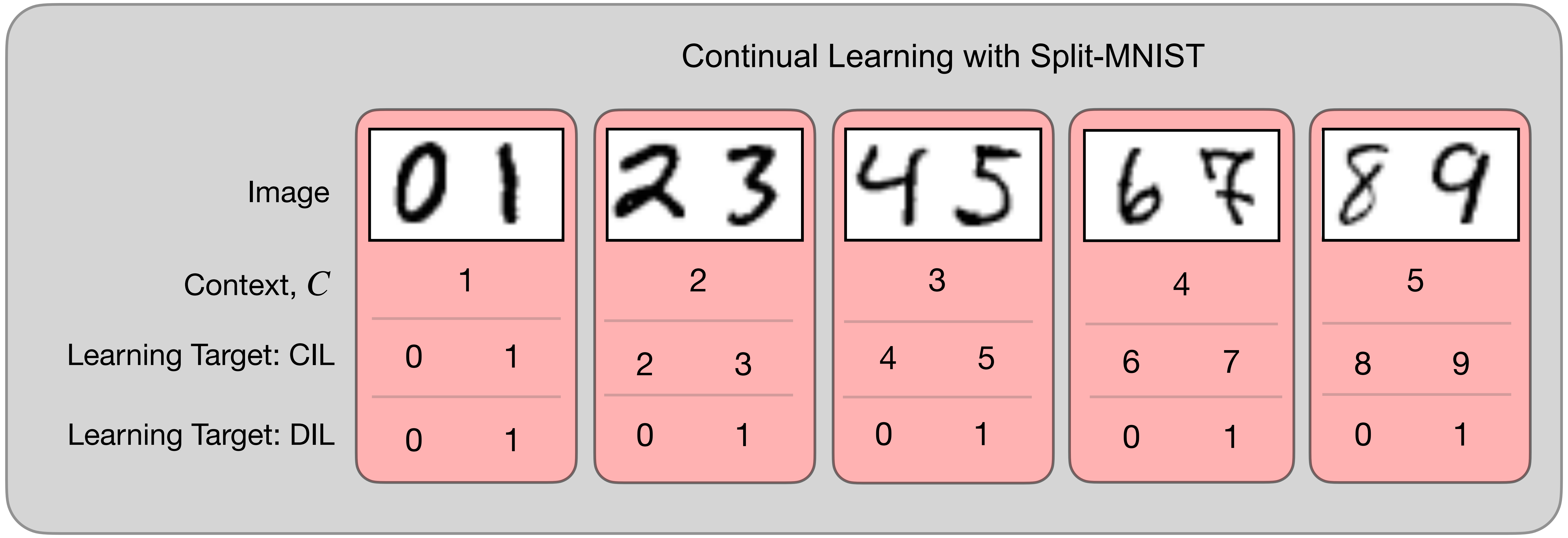
Liquid Ensemble Selection



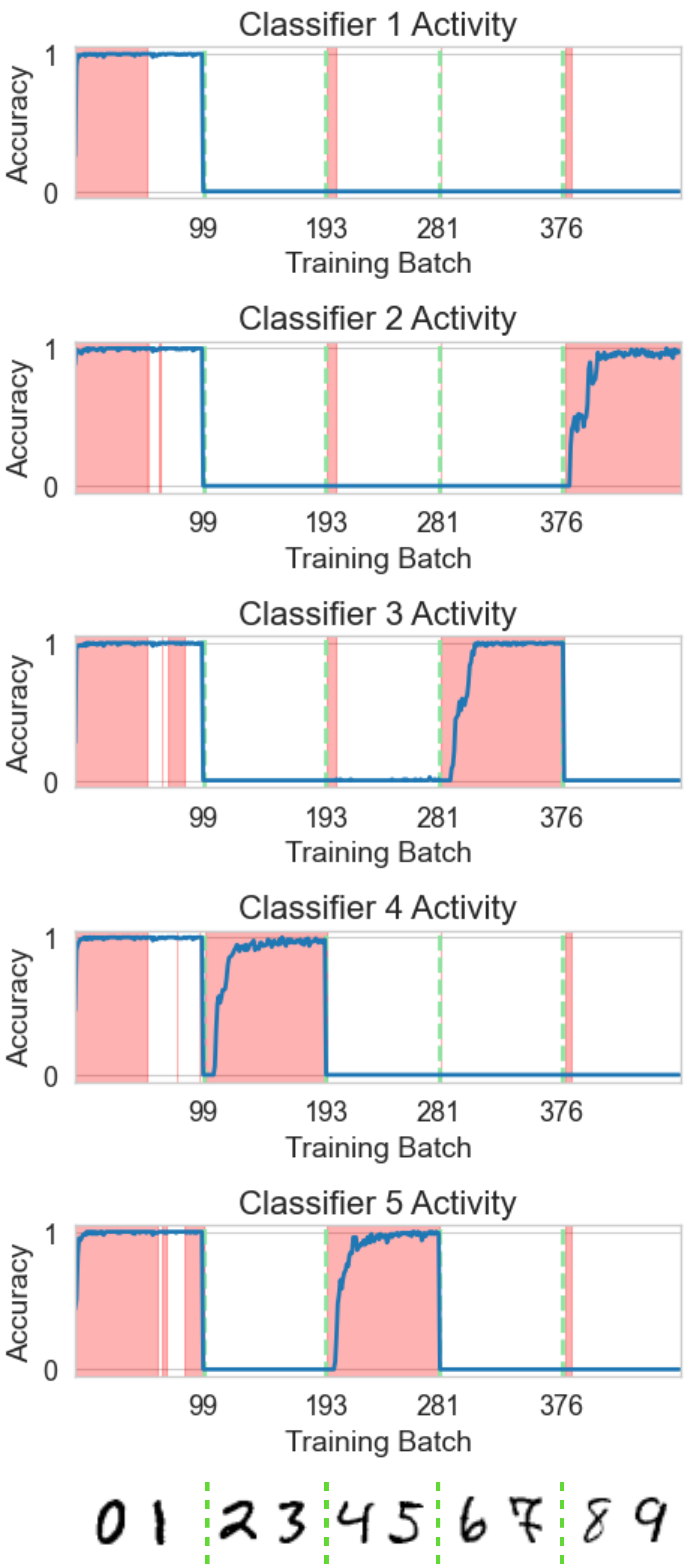
Liquid Ensemble Selection



Results



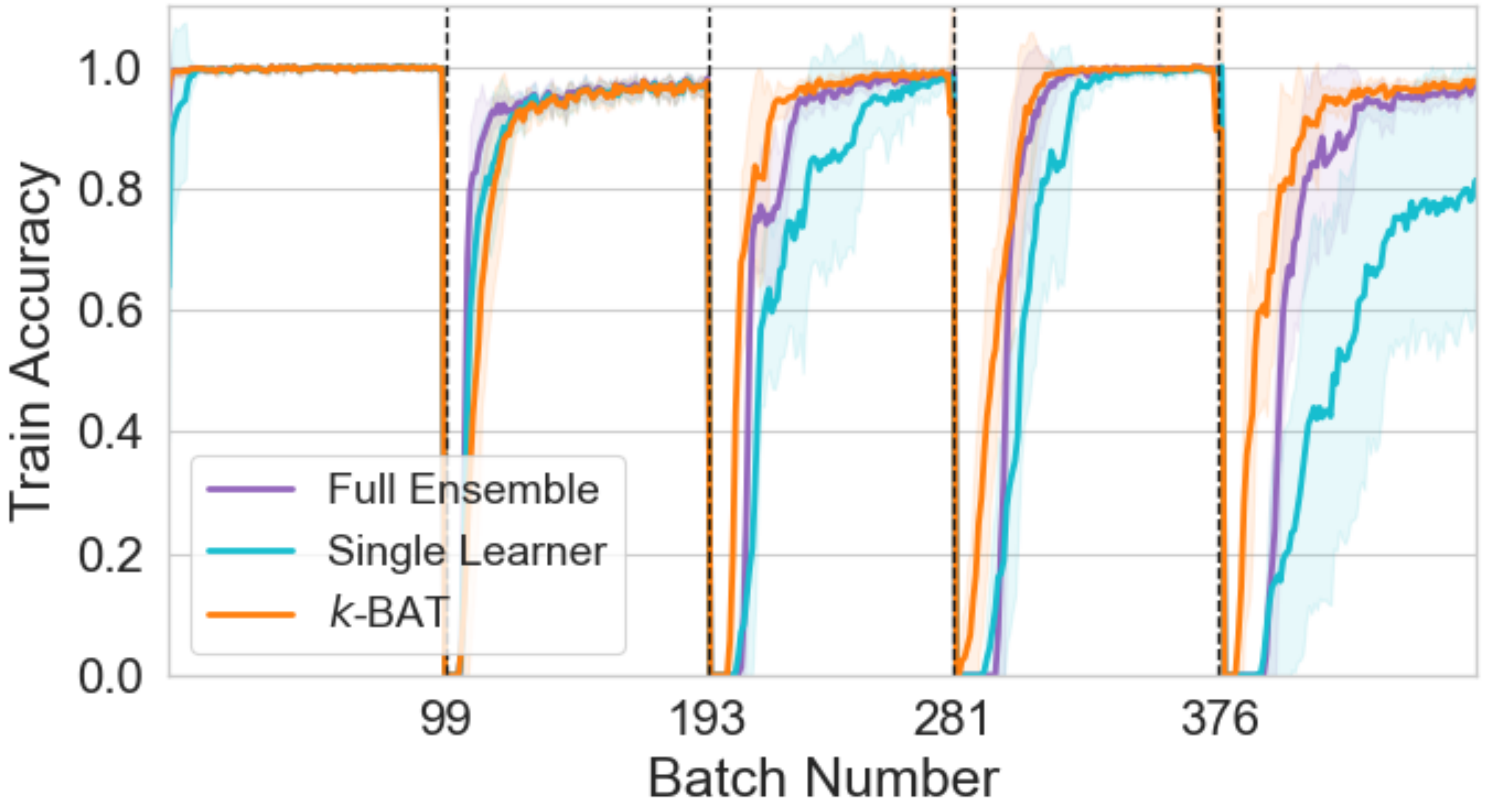
- Class Incremental Learning (CIL): Each context has a different set of classes
- Domain Incremental Learning (DIL): Each context has the same classes, but different underlying distributions mapping to those classes



Results

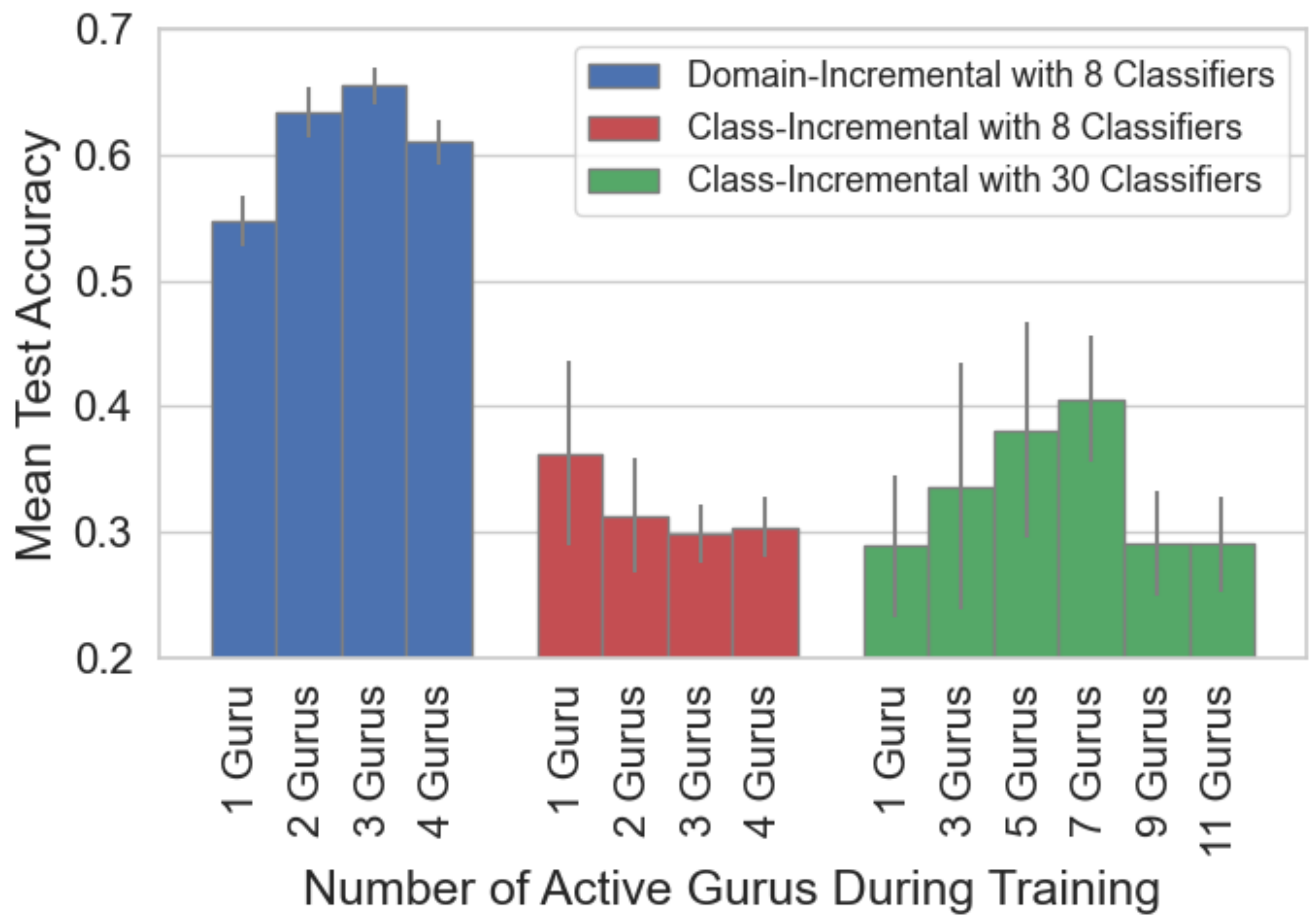
- The ensemble members do not know when the task changes
- Yet, the classifiers effectively divide and conquer

Results



In later contexts, the single model has greatly reduced plasticity

Results



Sometimes, having more ensemble members learning is helpful, and sometimes it's not

Moving Forward

- Continual learning is a relatively new field where social choice could be usefully applied
  - Online allocation of training
  - Representation of contexts
  - Aggregation of prediction

Liquid Ensemble Selection for Continual Learning

Carter Blair, Ben Armstrong, Kate Larson

cblair@uwaterloo.ca

UNIVERSITY OF WATERLOO