Overview

- The components and the architecture of a learning problem
- Why a learner needs a bias
- Identify the sources of error for a prediction

Learning

Learning is the ability to improve one's behavior based on experience.

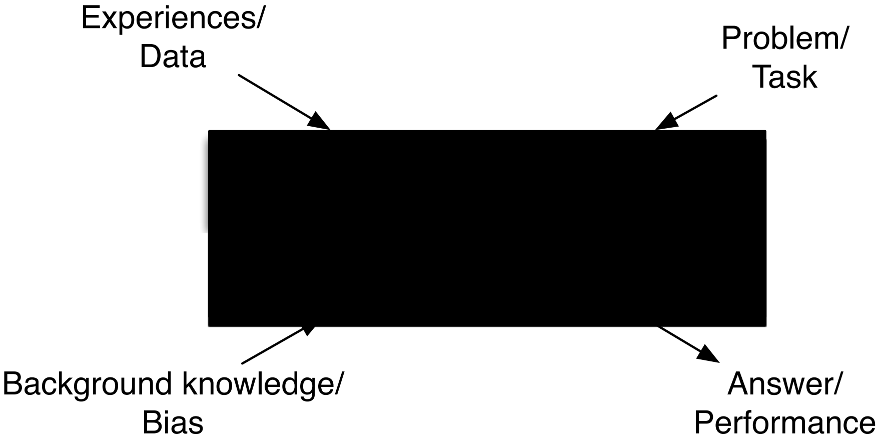
- The range of behaviors is expanded: the agent can do more.
- The accuracy on tasks is improved: the agent can do things better.
- The speed is improved: the agent can do things faster.

Components of a learning problem

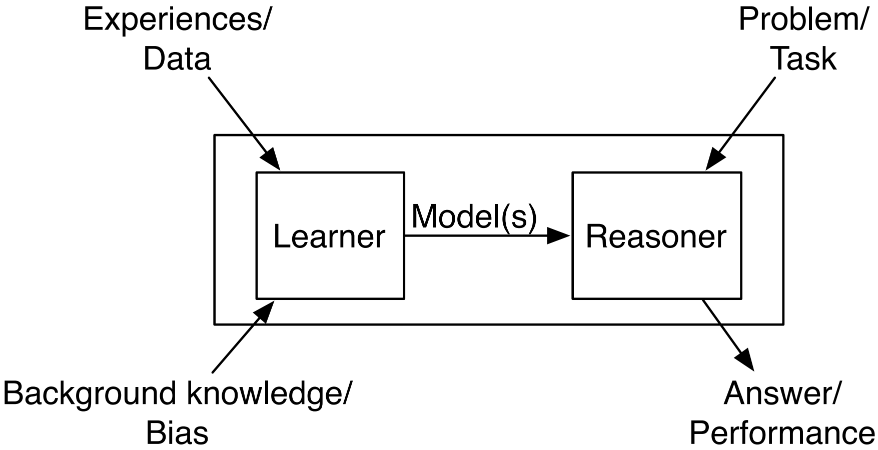
The following components are part of any learning problem:

- **task** The behavior or task that's being improved.
  For example: classification, acting in an environment
- **data** The experiences that are being used to improve performance in the task.
- **measure of improvement** How can the improvement be measured?
  For example: increasing accuracy in prediction, new skills that were not present initially, improved speed.

Black-box Learner



Learning architecture



Choosing a representation for models

- The richer the representation, the more useful it is for subsequent problem solving.
- The richer the representation, the more difficult it is to learn.

Common Learning Tasks

- Supervised classification Given a set of pre-classified training examples, classify a new instance.
- Unsupervised learning Find natural classes for examples.
- Reinforcement learning Determine what to do based on rewards and punishments.
- Analytic learning Reason faster using experience.
- Inductive logic programming Build richer models in terms of logic programs.
- Statistical relational learning learning relational representations that also deal with uncertainty.

Example Classification Data

Training Examples:

	Action	Author	Thread	Length	Where
e1	skips	known	new	long	home
e2	reads	unknown	new	short	work
e3	skips	unknown	old	long	work
e4	skips	known	old	long	home
e5	reads	known	new	short	home
e6	skips	known	old	long	work

New Examples:

e7	???	known	new	short	work
e8	???	unknown	new	short	work

We want to classify new examples on feature *Action* based on the examples' *Author*, *Thread*, *Length*, and *Where*.

Feedback

Learning tasks can be characterized by the feedback given to the learner.

- Supervised learning What has to be learned is specified for each example.
- Unsupervised learning No classifications are given; the learner has to discover categories and regularities in the data.
- Reinforcement learning Feedback occurs after a sequence of actions.

Measuring Success

- The measure of success is not how well the agent performs on the training examples, but how well the agent performs for new examples.
- Consider two agents:
  - P claims the negative examples seen are the only negative examples. Every other instance is positive.
  - N claims the positive examples seen are the only positive examples. Every other instance is negative.

Measuring Success

- The measure of success is not how well the agent performs on the training examples, but how well the agent performs for new examples.
- Consider two agents:
  - *P* claims the negative examples seen are the only negative examples. Every other instance is positive.
  - *N* claims the positive examples seen are the only positive examples. Every other instance is negative.
- Both agents correctly classify every training example, but disagree on every other example.

Bias

- The tendency to prefer one hypothesis over another is called a bias.
- Saying a hypothesis is better than N's or P's hypothesis isn't something that's obtained from the data.
- To have any inductive process make predictions on unseen data, an agent needs a bias.
- What constitutes a good bias is an empirical question about which biases work best in practice.

Learning as search

- Given a representation, data, and a bias, the problem of learning can be reduced to one of search.
- Learning is search through the space of possible representations looking for the representation or representations that best fits the data, given the bias.
- These search spaces are typically prohibitively large for systematic search.
- A learning algorithm is made of a search space, an evaluation function, and a search method.

Data

- Data isn't perfect:
  - the features given are inadequate to predict the classification
  - there are examples with missing features
  - some of the features are assigned the wrong value
- overfitting occurs when distinctions appear in the training data, but not in the unseen examples.

Overfitting

- In statistics, overfitting is the production of an analysis that corresponds too closely or exactly to a particular set of data, and may therefore fail to fit additional data or predict future observations reliably.
- An overfitted model is a statistical model that contains more parameters than can be justified by the data.
- Underfitting occurs when a statistical model cannot adequately capture the underlying structure of the data. An underfitted model is a model where some parameters or terms that would appear in a correctly specified model are missing.
- Underfitting would occur, for example, when fitting a linear model to non-linear data. Such a model will tend to have poor predictive performance.

Overfitting

- As an extreme example, if the number of parameters is the same as or greater than the number of observations, then a model can perfectly predict the training data simply by memorizing the data in its entirety.
  Such a model, though, will typically fail severely when making predictions.
- To lessen the chance of, or amount of, overfitting, several techniques are available.
  The basis of some techniques is:
  - either to explicitly penalize overly complex models,
  - or to test the model's ability to generalize by evaluating its performance on a set of data not used for training, which is assumed to approximate the typical unseen data that a model will encounter.

Errors in learning

Errors in learning are caused by:

- Limited representation (representation bias)
- Limited search (search bias)
- Limited data (variance)
- Limited features (noise)

Characterizations of Learning

- Find the best representation given the data.
- Delineate the class of consistent representations given the data.
- Find a probability distribution of the representations given the data.