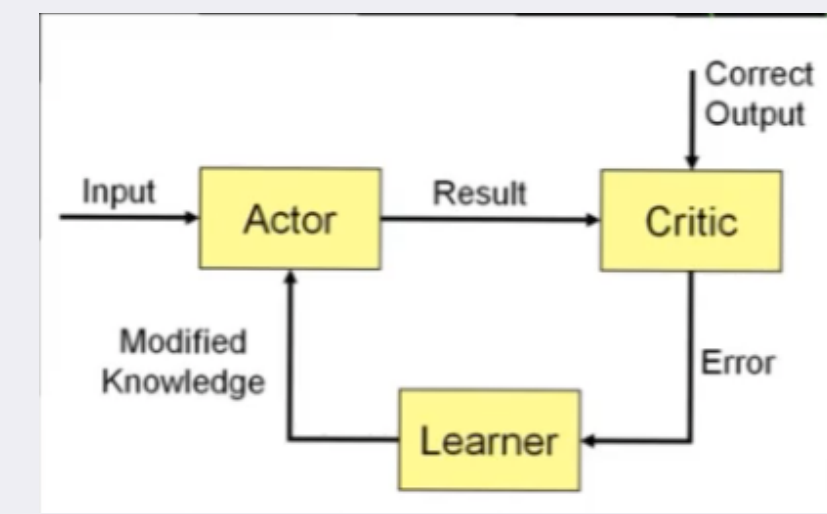
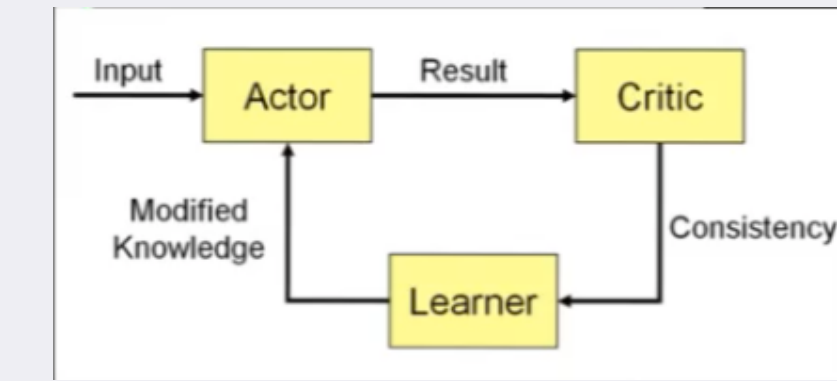


A computer program is said to learn from experience with respect to some class of tasks T and performance measure P if its performance on tasks T as measured by P improves with E.

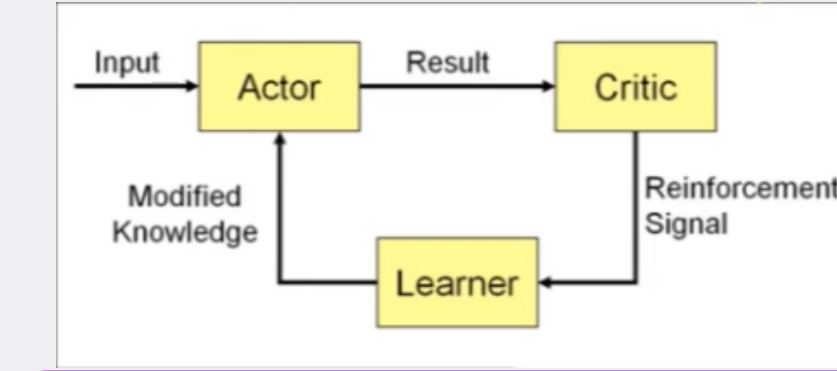


Online/offline learning



Clustering, Feature Extraction

Semi-supervised learning Self training



Sequential decision problems

Receive data instance

Make prediction

Prediction correct? 0 -> 1: Promotion 1 -> 0: Demotion

Linear discriminant function

Learning Maximizes performance relative to some performance measurement Learning is an optimization problem

set of assumptions that combined with the observed training examples, deductively entail subsequent incidence classifications made by the learner.

Inductive bias is the stuff of the learning process that enables the model to generalize.

Unbiased learner can not learn Bias-variance trade off

Category Representation bias Model itself what limitations are we imposing? How does the model type support generalization Preference bias / selection bias Performance measure Principle of Parsimony Prefer simple over complex Occam's Razor Minimize assumptions Better generalization. Overfitting

No free lunch over all possible functions that could be learned using an algorithm, we can never have one algorithm that always beats out the other algorithm. A best learning algorithm does not exist So our challenge is to match the algorithms bias to the problem

Take aways No panacea to all problems Understand the problem, match the right model Performance measure Suited to optimizing and avoid overfitting

How evaluate the performance

Performance measure Find a model can generalize Statistical learning theory Measure of loss Risk Loss function 0/1 Squared loss Linear regression Regulation

How learning algorithms function How to pick the right algorithm

Learning Models Computational Learning theory What general laws constrain our ability to learn? What type of learning problems can be solved in a reasonable amount of time / space? When can we trust the output of a learned hypothesis, and by how much can we trust it? Sample Complexity Additional information Active learning Helpful teacher model The teacher picks the examples that it knows will provide the most information sample complexity refers to the number of training examples Required to learn our target concept is some minimal level of performance. PAC: Probably Approximately Correct Concept space is PAC learnable True error bounded by epsilon approximation With probability 1-delta

What is ML

Type

Supervised Learning

Unsupervised Learning

Reinforcement learning

Example: Winnow

Winnow-2

Lecture

Why works

Learning Theory

Learning Models

M1

C1 Introduction

Examples History Related topics C2 Supervised Learning

ML

We have many problems + Data, but lack algorithms Construct a good and useful approximation Find pattern in the data Data mining: ML methods to large databases Pattern recognition

ML builds mathematical models.

Optimization Inference CS

Core task: Make inferences from a sample