Introduction to Machine Learning

Module 1: Introduction

Part A: Introduction

Sudeshna Sarkar

IIT Kharagpur

Overview of Course

1. Introduction
2. Linear Regression and Decision Trees
3. Instance based learning
   Feature Selection
4. Probability and Bayes Learning
5. Support Vector Machines
6. Neural Network
7. Introduction to Computational Learning Theory
8. Clustering

Module 1

1. Introduction
   a) **Introduction**
   b) Different types of learning
   c) Hypothesis space, Inductive Bias
   d) Evaluation, Training and test set, cross-validation
2. Linear Regression and Decision Trees
3. Instance based learning
   Feature Selection
4. Probability and Bayes Learning
5. Support Vector Machines
6. Neural Network
7. Introduction to Computational Learning Theory
8. Clustering

Machine Learning History

- 1950s:
  - Samuel's checker-playing program
- 1960s:
  - Neural network: Rosenblatt's perceptron
  - Minsky & Papert prove limitations of Perceptron
- 1970s:
  - Symbolic concept induction
  - Expert systems and knowledge acquisition bottleneck
  - Quinlan’s ID3
  - Natural language processing (symbolic)

Machine Learning History

- 1980s:
  - Advanced decision tree and rule learning
  - Learning and planning and problem solving
  - Resurgence of neural network
  - Valiant's PAC learning theory
  - Focus on experimental methodology
- 90's ML and Statistics
  - Data Mining
  - Adaptive agents and web applications
  - Text learning
  - Reinforcement learning
  - Ensembles
  - Bayes Net learning

- 1994: Self-driving car road test
- 1997: Deep Blue beats Gary Kasparov

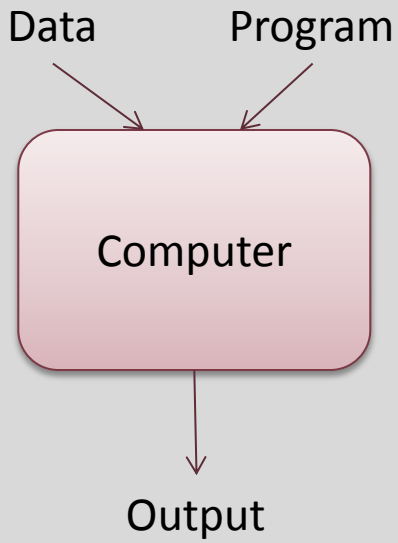
Machine Learning History

- Popularity of this field in recent time and the reasons behind that
  - New software/ algorithms
    - Neural networks
    - Deep learning
  - New hardware
    - GPU's
  - Cloud Enabled
  - Availability of Big Data

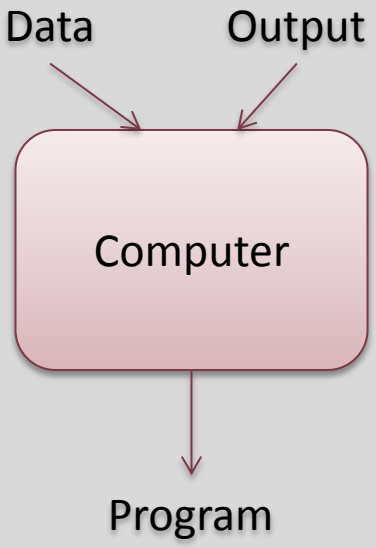
- 2009: Google builds self driving car
- 2011: Watson wins Jeopardy
- 2014: Human vision surpassed by ML systems

Programs vs learning algorithms

Algorithmic solution



Machine learning solution



Machine Learning : Definition

- Learning is the ability to improve one's behaviour based on experience.
- Build computer systems that automatically improve with experience
- What are the fundamental laws that govern all learning processes?
- Machine Learning explores algorithms that can
  - learn from data / build a model from data
  - use the model for prediction, decision making or solving some tasks

Machine Learning : Definition

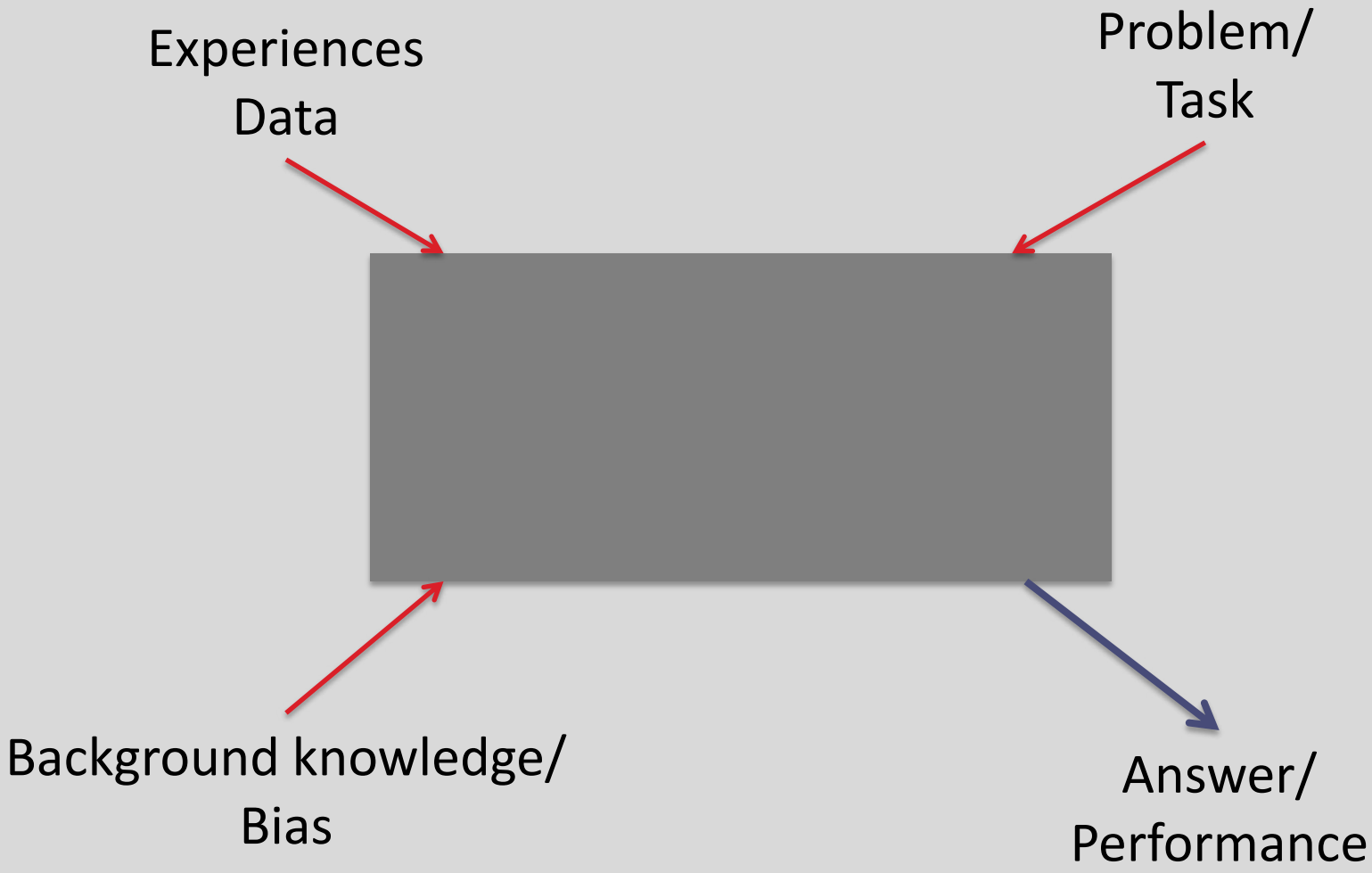
- A computer program is said to learn from *experience* E with respect to some *class of tasks* T and *performance measure* P, if its performance at tasks in T, as measured by P, improves with experience E.

  [Mitchell]

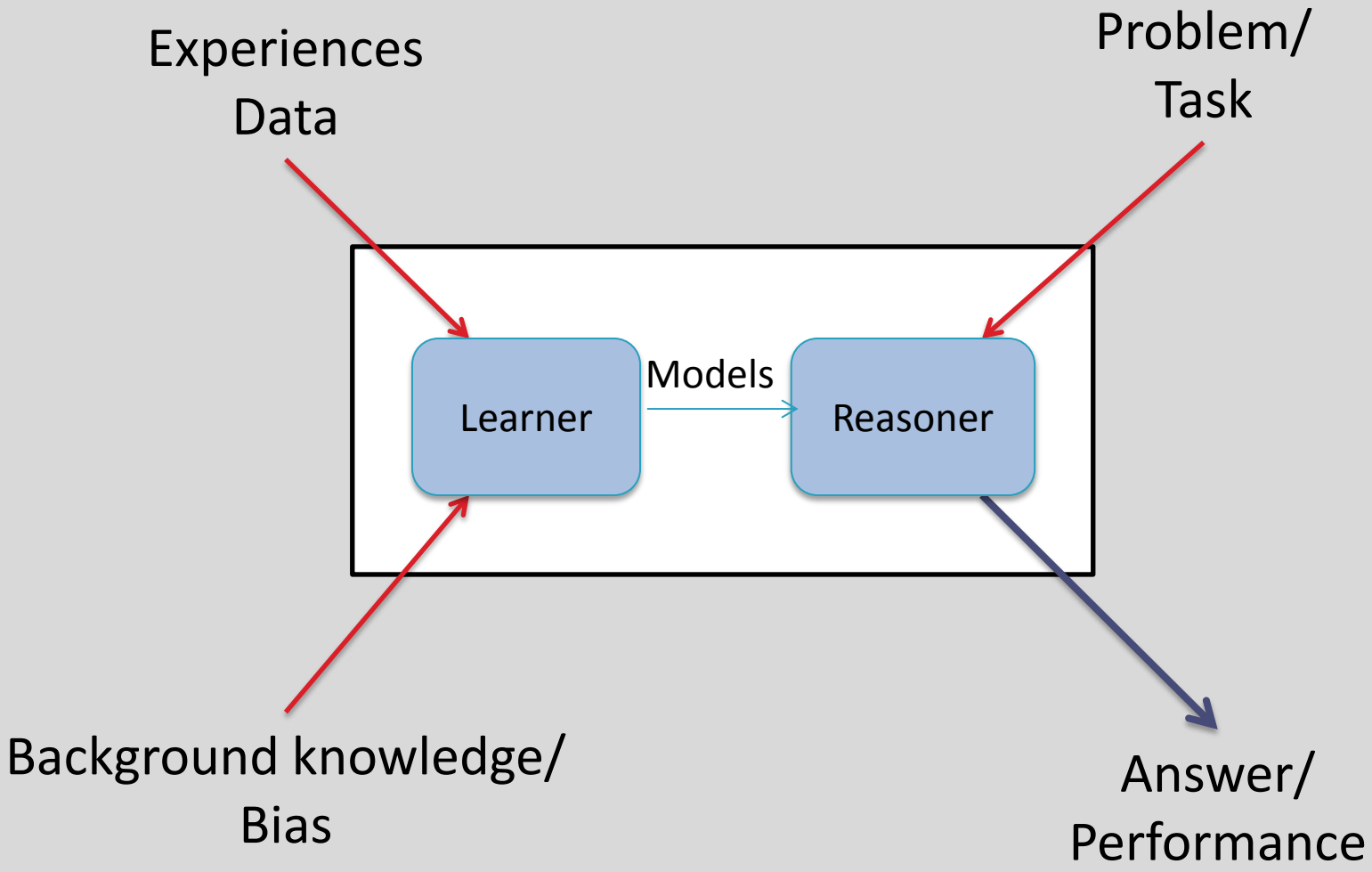
Components of a learning problem

- Task: The behaviour or task 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 :
  - For example: increasing accuracy in prediction, acquiring new, improved speed and efficiency

Black-box Learner



Learner


Experiences
Data
Problem/
Task
Learner
Models
Reasoner
Background knowledge/
Bias
Answer/
Performance

Many domains and applications

Medicine:

- Diagnose a disease
  - Input: symptoms, lab measurements, test results, DNA tests,
  - Output: one of set of possible diseases, or “none of the above”
- Data: historical medical records
- Learn: which future patients will respond best to which treatments

Many domains and applications

Vision:

- say what objects appear in an image
- convert hand-written digits to characters 0..9
- detect *where* objects appear in an image

Many domains and applications

Robot control:

- Design autonomous mobile robots that learn from experience to
  - Play soccer
  - Navigate from their own experience

Many domains and applications

NLP:

- detect where entities are mentioned in NL
- detect what facts are expressed in NL
- detect if a product/movie review is positive, negative, or neutral

Speech recognition

Machine translation

Many domains and applications

Financial:

- predict if a stock will rise or fall
- predict if a user will click on an ad or not

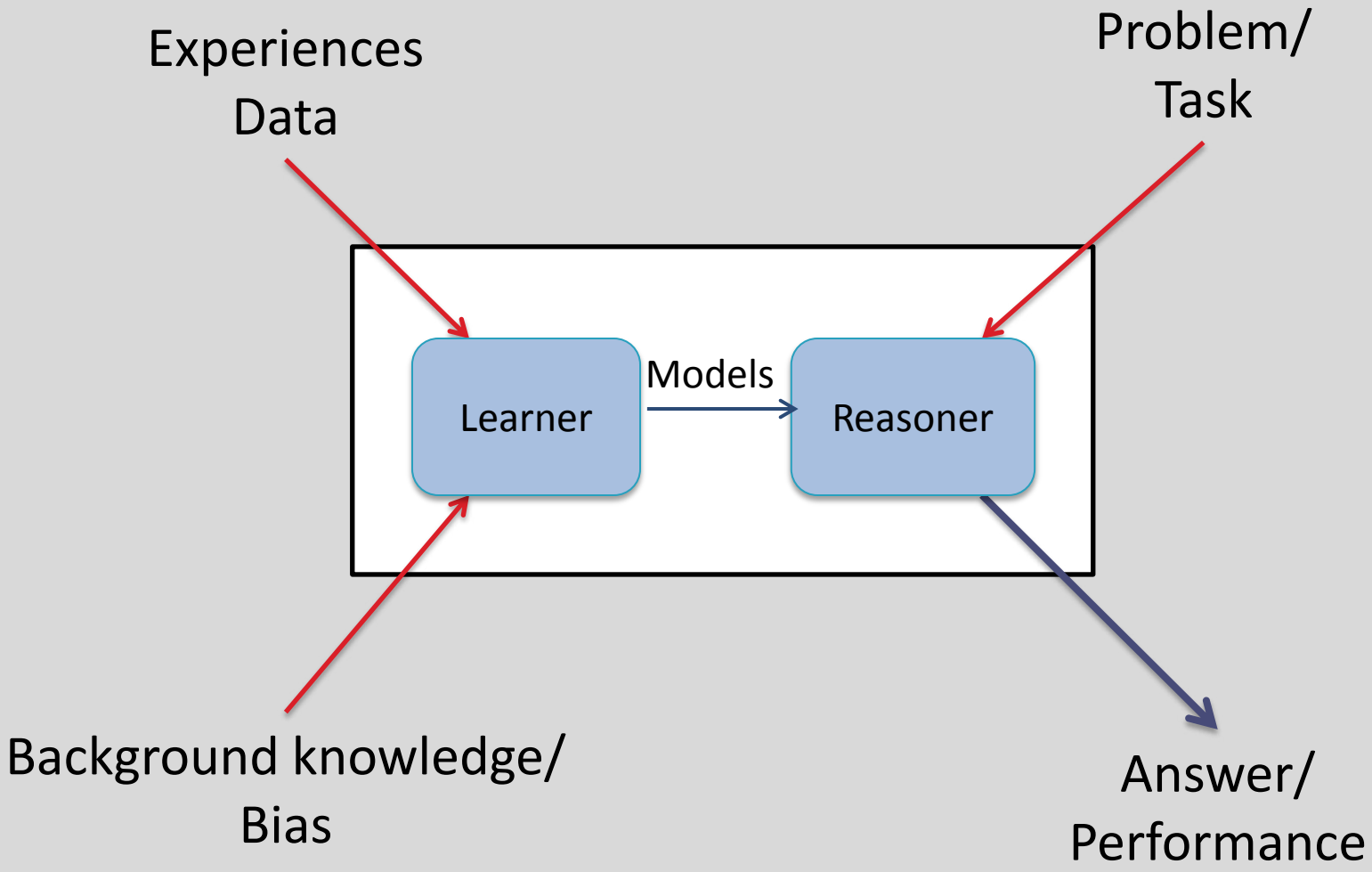
Application in Business Intelligence

- Forecasting product sales quantities taking seasonality and trend into account.
- Identifying cross selling promotional opportunities for consumer goods.
- …

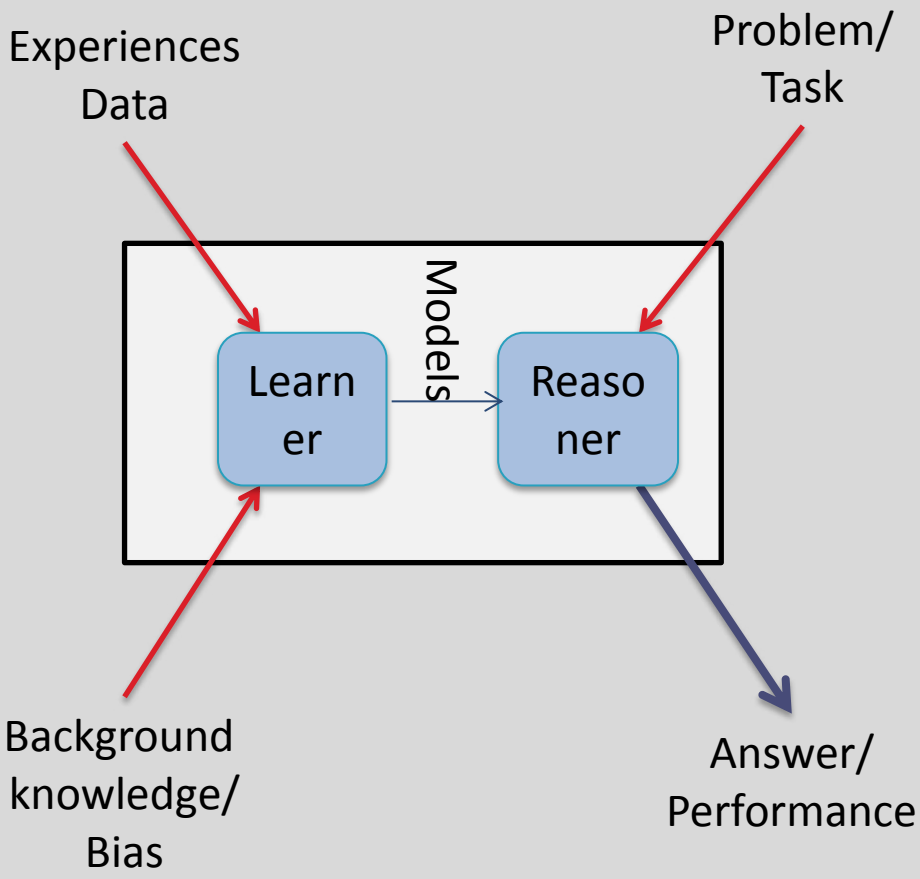
Some other applications

- Fraud detection : Credit card Providers
- determine whether or not someone will default on a home mortgage.
- Understand consumer sentiment based off of unstructured text data.
- Forecasting women's conviction rates based off external macroeconomic factors.

Learner


Experiences
Data
Problem/
Task
Learner
Models
Reasoner
Background knowledge/
Bias
Answer/
Performance

Design a Learner



1. Choose the training experience
2. Choose the target function (that is to be learned)
3. Choose how to represent the target function
4. Choose a learning algorithm to infer the target function

Choosing a Model Representation

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

- Components of Representation
  - Features
  - Function class / hypothesis language