Applied Artificial Intelligence

Session 2: Demystifying AI, The Big Picture

Fall 2018
NC State University
Instructor: Dr. Behnam Kia
Course Website: https://appliedai.wordpress.ncsu.edu/

Review of Session 1: Course Introduction

- What type of problems are AI problems (in which scenarios you should use Artificial Intelligence)?

Review of Session 1: Course Introduction

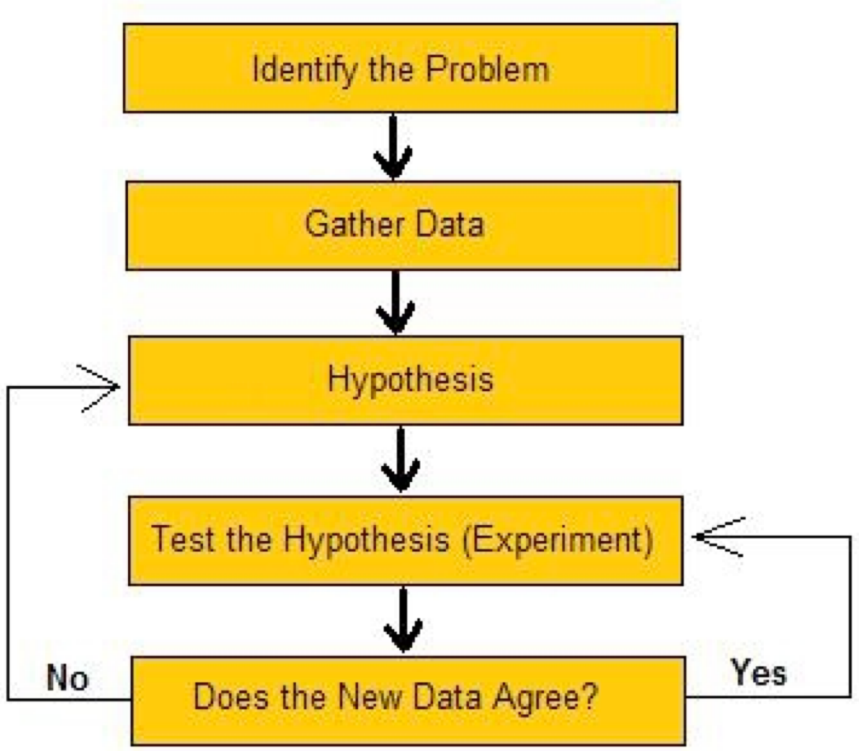
- What type of problems are AI problems (in which scenarios you should use Artificial Intelligence)?
  - Problems simple for humans to solve (intuitively), but not for computers.
  - Extracting knowledge from Big Data.
  - Solving dynamic, varying problems.

Review of Session 1: Course Introduction

- What type of problems are AI problems (in which scenarios you should use Artificial Intelligence)?
  - Problems simple for humans to solve (intuitively), but not for computers.
  - Extracting knowledge from Big Data.
  - Solving dynamic, varying problems.
- If the problem is described by a set of formal mathematical rules (coming from Math, Physics, Chemistry, Biology, etc.), and there are known methods to solve it, develop a conventional computer program and solve it. Usually this is **not** an AI problem – unless it is a hard problem in terms of complexity.

Artificial Intelligence Follows Scientific Method

- AI follows scientific method and requires observation, data collection, experimentation, etc.
- There isn't any thing magical about AI!



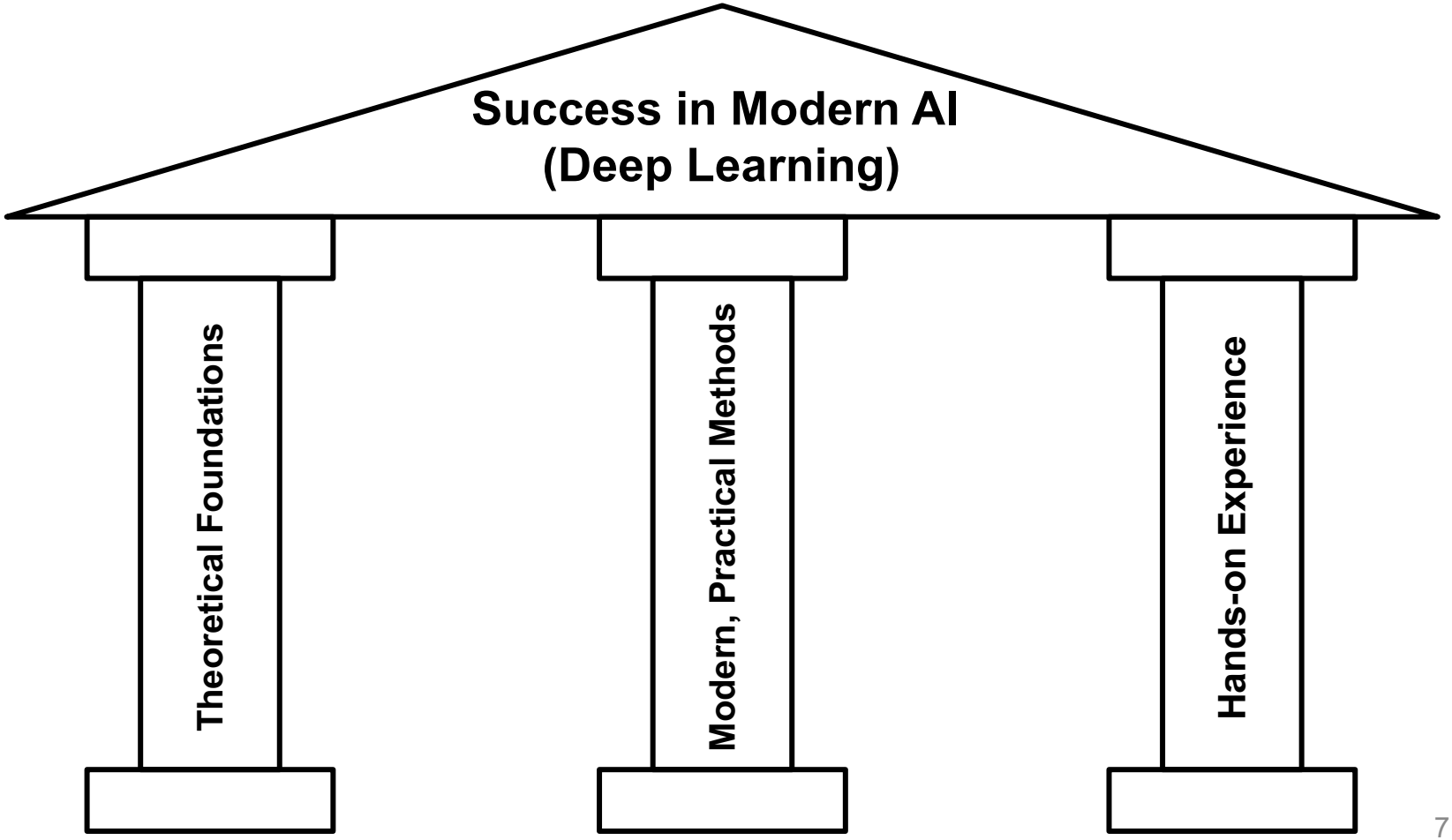
Scientific Method

What you can expect from this course.

At the end of the semester you will:

- know what an AI problem is and what is not!
- learn the basic foundations of deep learning and how to apply it to AI problems.
- gain basic hands-on experience with AI development tools and software.
- get enough experience, knowledge, and confidence to pursue on your own and learn more advanced topics.

Three Pillars to Success in Modern AI (Deep Learning)



Course Website is Online and Updated

https://appliedai.wordpress.ncsu.edu/

- Grading and course policies can be found on this website.

Session 2: Demystifying AI, The Big Picture

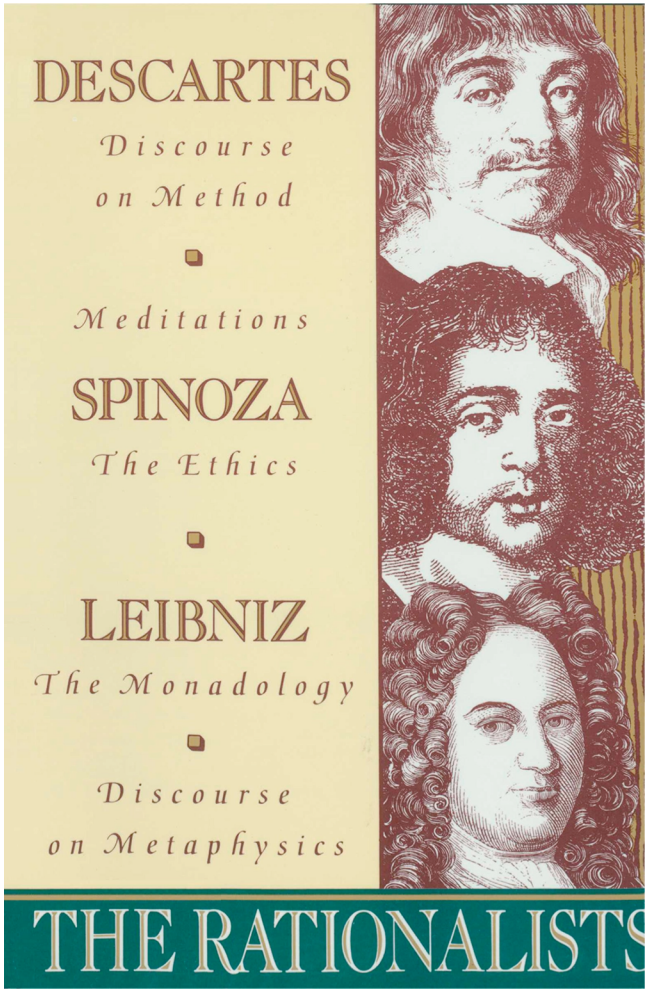
- In previous session we focused on AI problems; the problems that conventional programing fail to solve, and we need AI.
- In this session we look at the AI solutions; different AI approaches and methods to solve aforementioned problems (The Big Picture).
- These approaches and methods will be discussed at a greater detail during future sessions as we explain the course roadmap.

Some Philosophy

- In order to build Artificial Intelligence, we might ask the philosophical question of what the intelligence itself as.

- The answer to this question will show the path towards Artificial Intelligence.

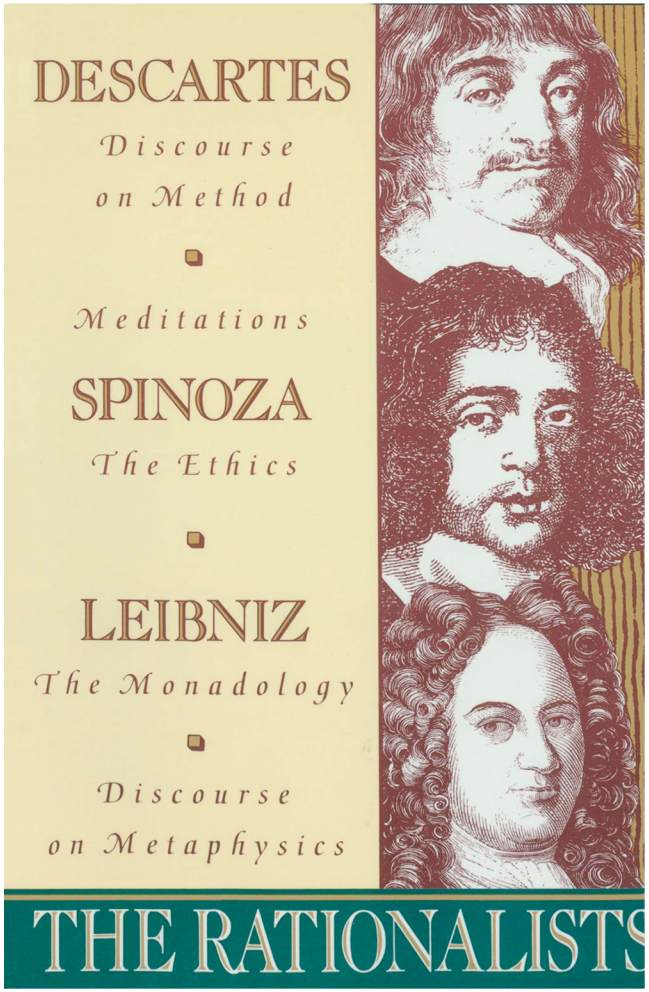
Rationalism



In philosophy, rationalism is the epistemological view that regards reason as the chief source and test of knowledge.

-Wikipedia

Rationalism



- Mind is a reasoning machine.
- It is equipped with knowledge, and with a reasoning engine it deduces new knowledge or solutions.
- So to create AI we need:
  - Knowledge representation.
  - A reasoning engine.

A famous, basic example of reasoning (logic)

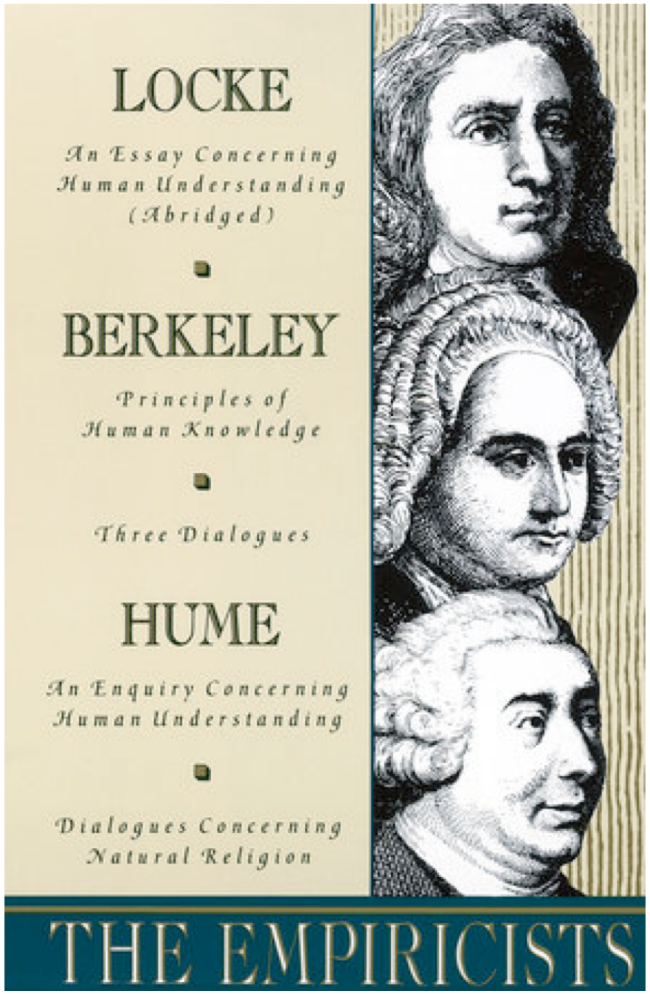
All men are mortal

Socrates is a man

} Therefore, Socrates is mortal.

Artwork from Tate Janek
http://tatejanek.com/

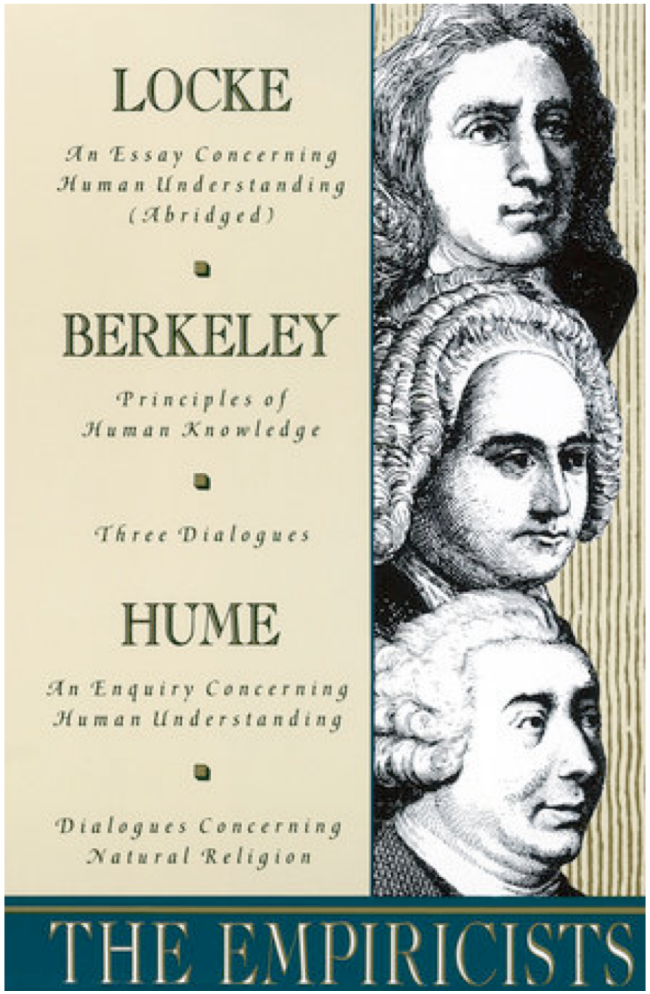
The Empiricists



The Empiricists

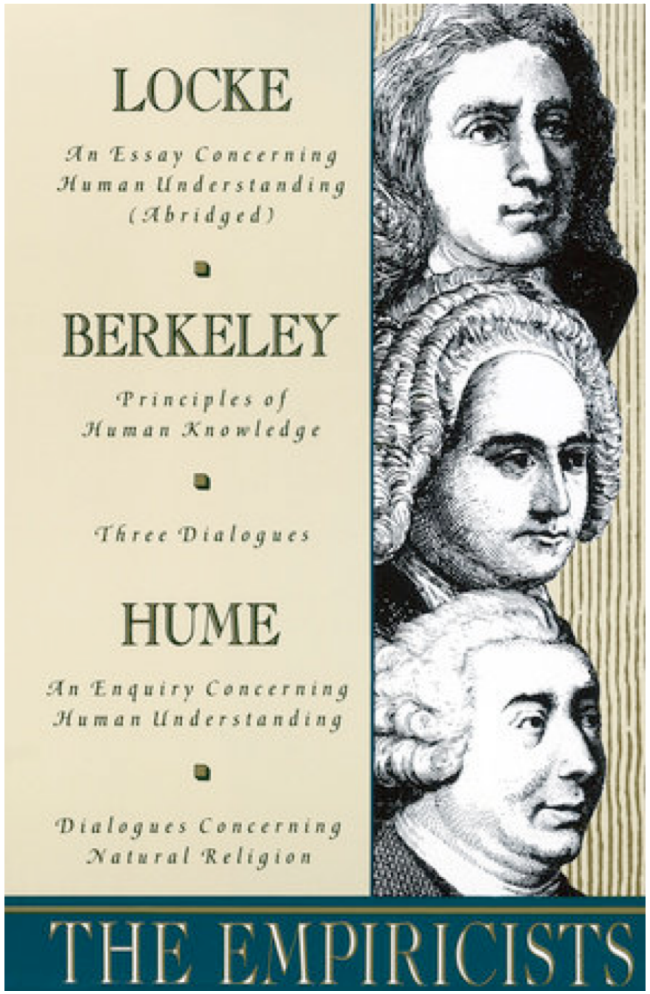
In philosophy, empiricism is a theory that states that knowledge comes only or primarily from sensory experience [observation or data].

-Wikipedia



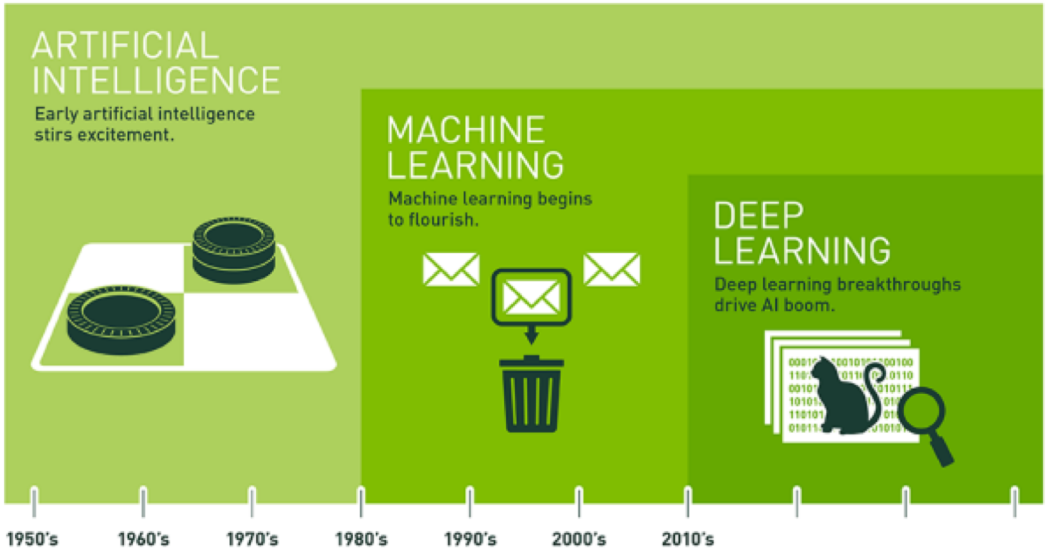
The Empiricists

- Mind is a learning machine!
- Empiricism emphasizes the role of experience, discounts the value of a priori reasoning.
- So to create AI we need:
  - Learning algorithms.
  - A lot of data.



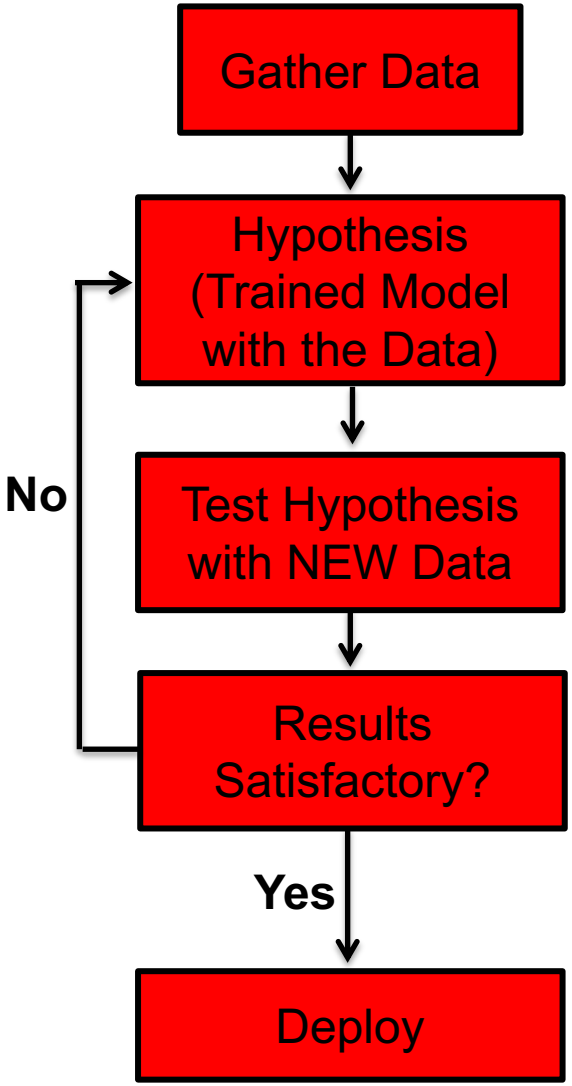
Machine Learning

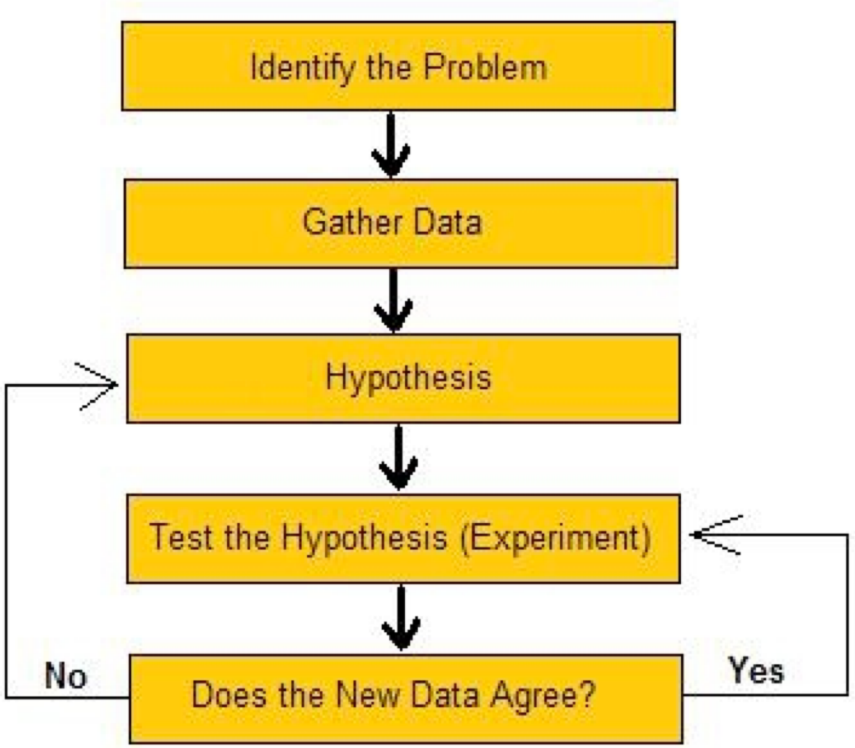
- Machine learning algorithm is an algorithm that learns from data with no need for explicit programing.



Picture from NVIDIA's deep learning institute

Machine Learning Flowchart (Which Follows Scientific Method)



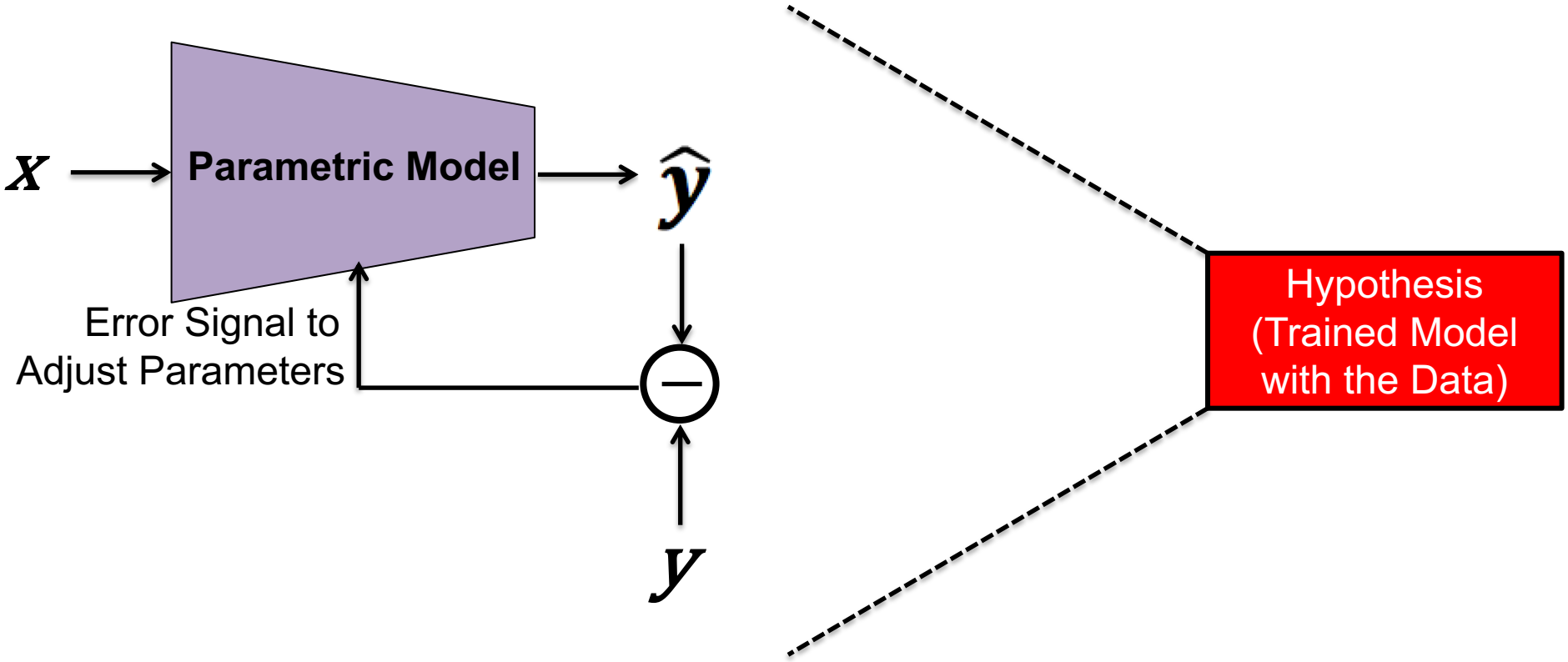


Scientific Method

Machine Learning: Training Model with Data

Training Data: (x_i, y_i), i=1,2,3,...,N

x is called features, y is label.



How to Adjust Parameters?

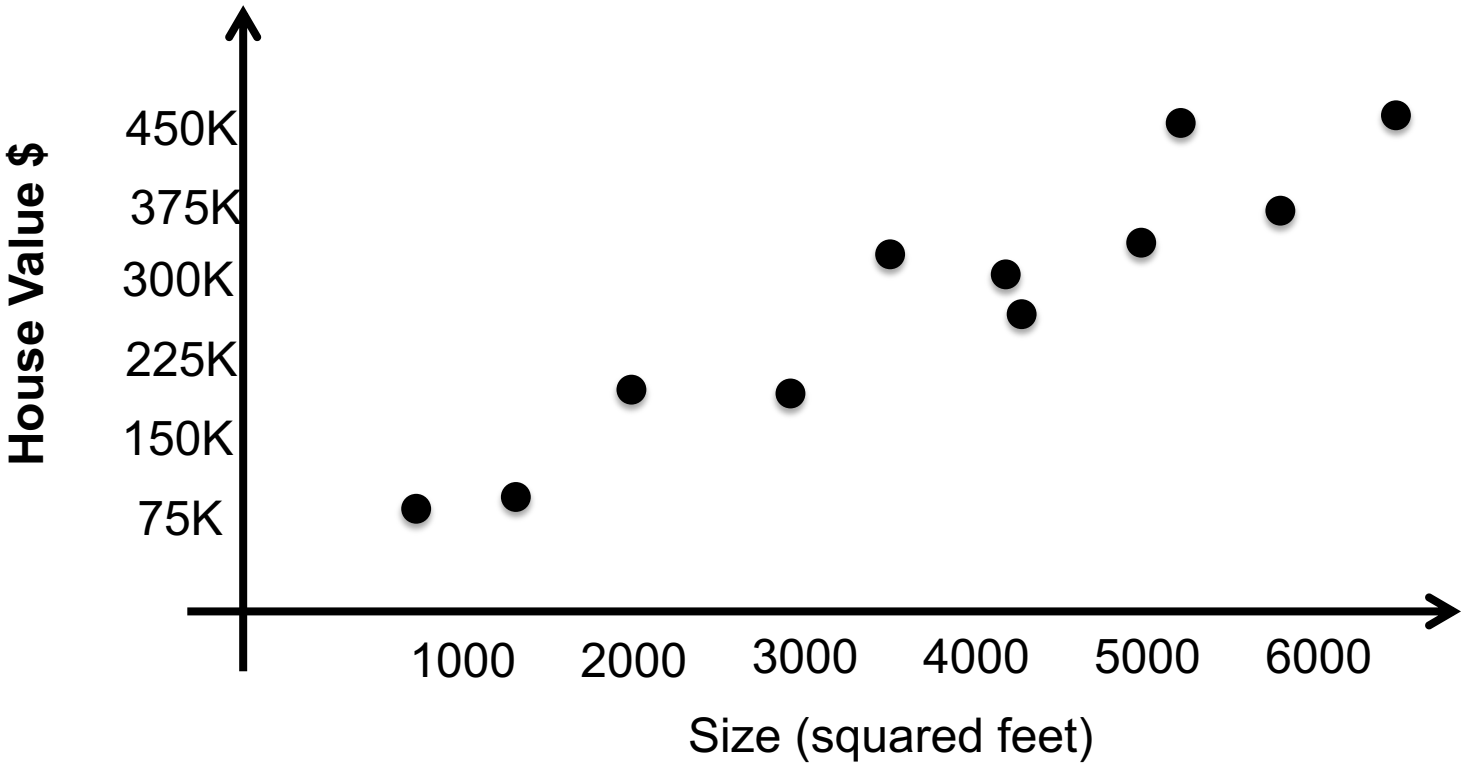
- This is an optimization problem to find parameter values, P^*, that minimizes the error.

$$P^* = \text{argmin}_P \sum_{Training set} Error(x, y, \hat{y})$$

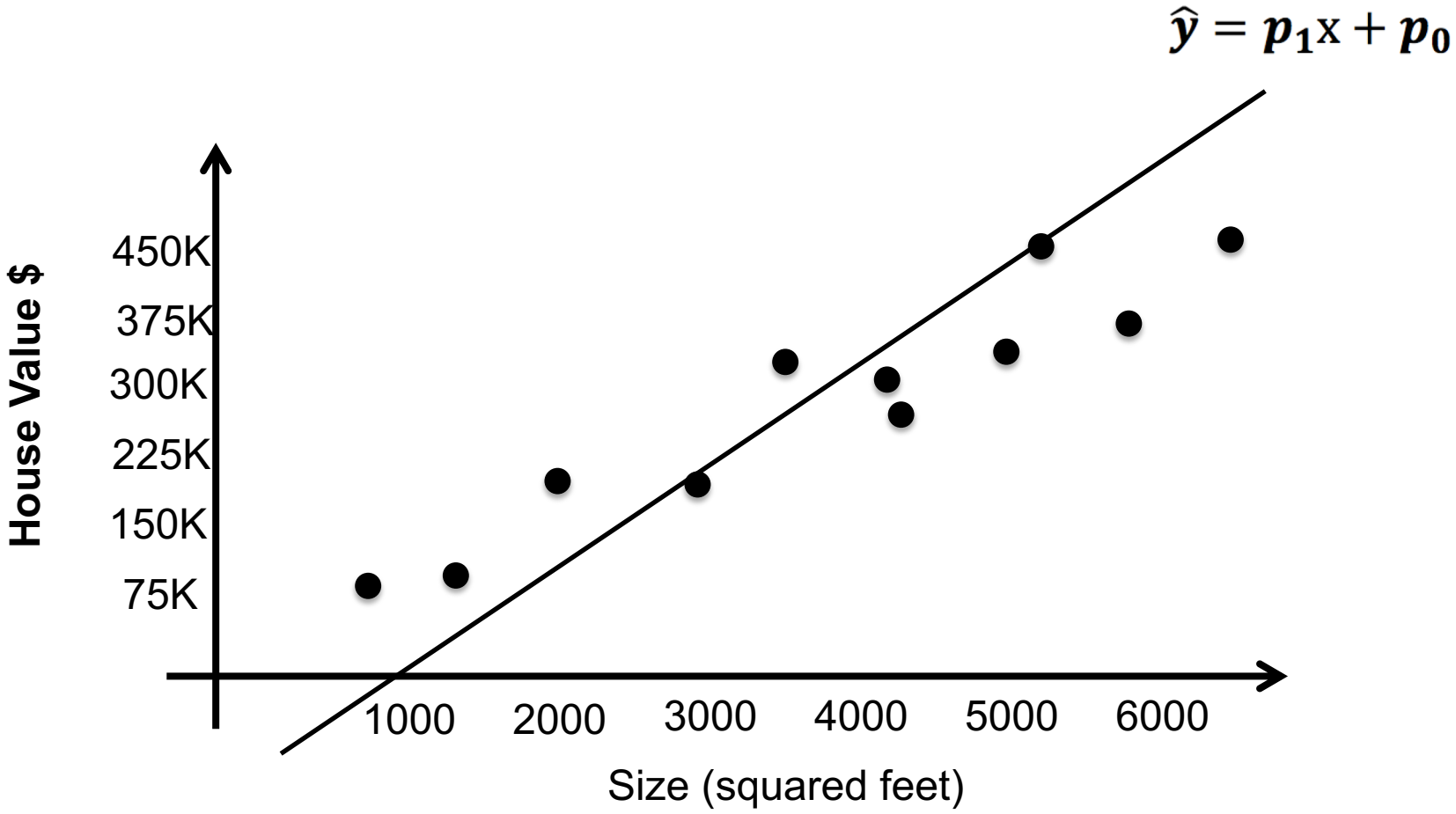
Another common way to say this is shown below, where J is cost function

$$P^* = \text{argmin}_P J(P)$$

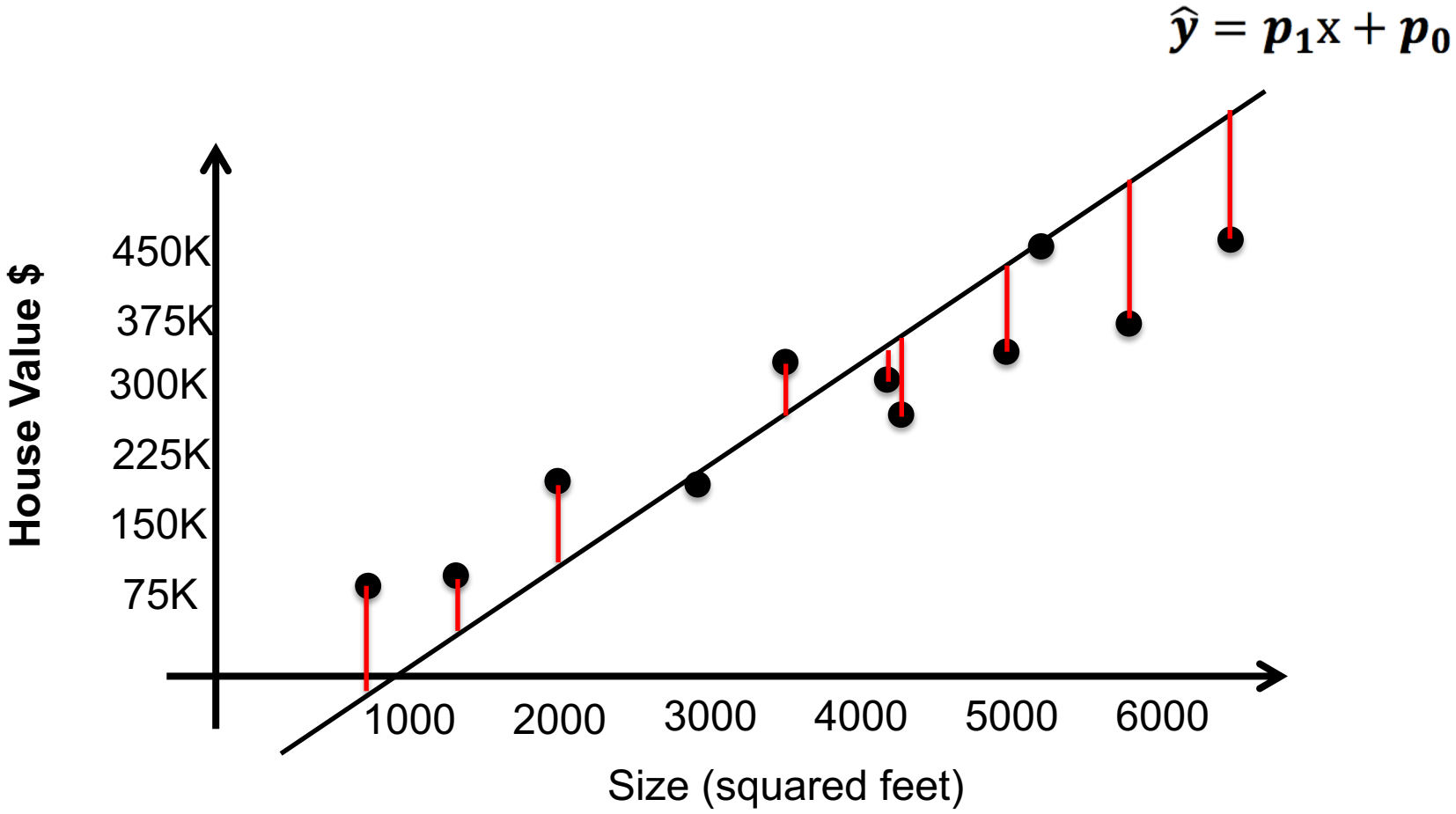
Example: House Values



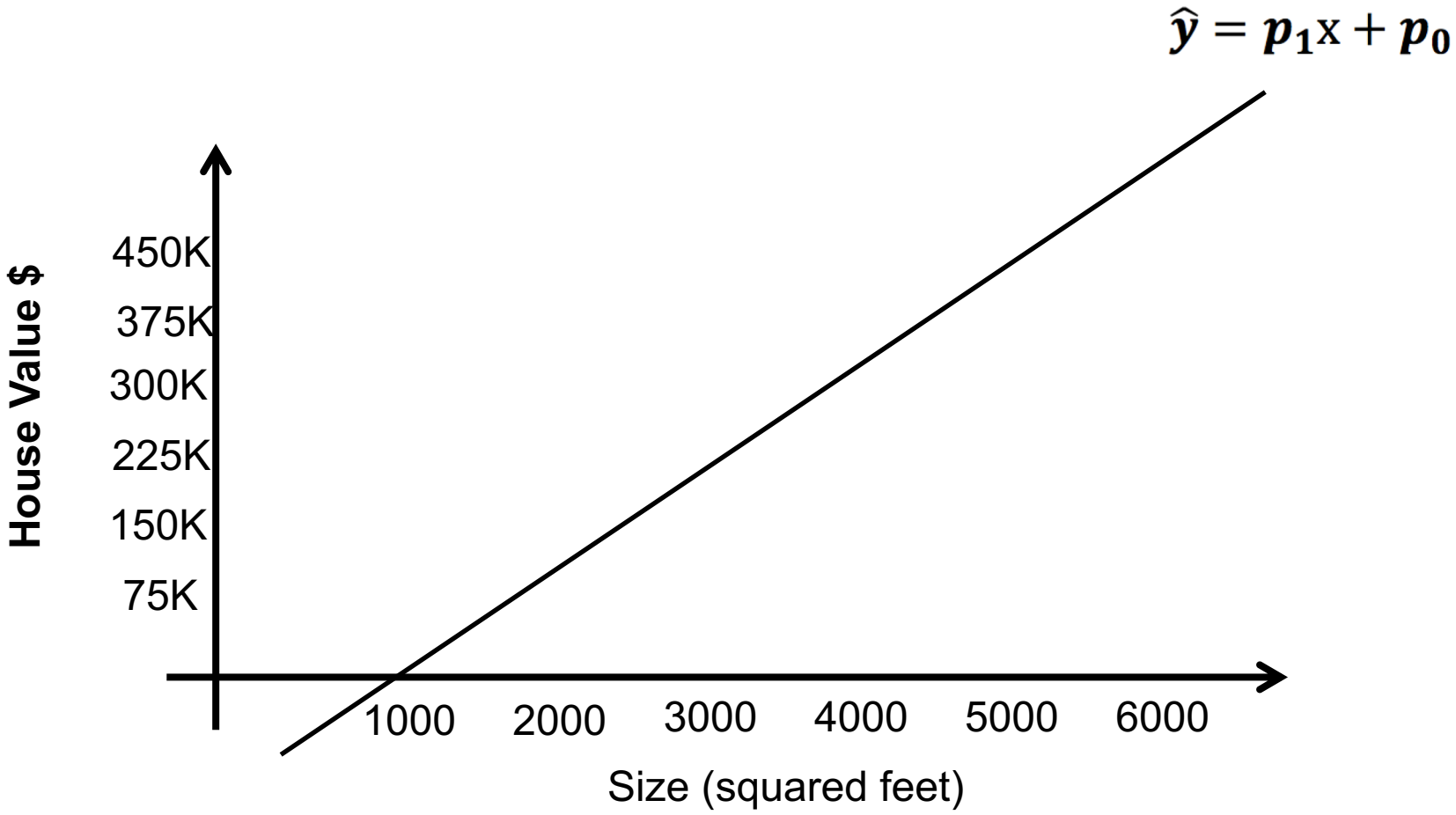
Example: House Values



Example: House Values

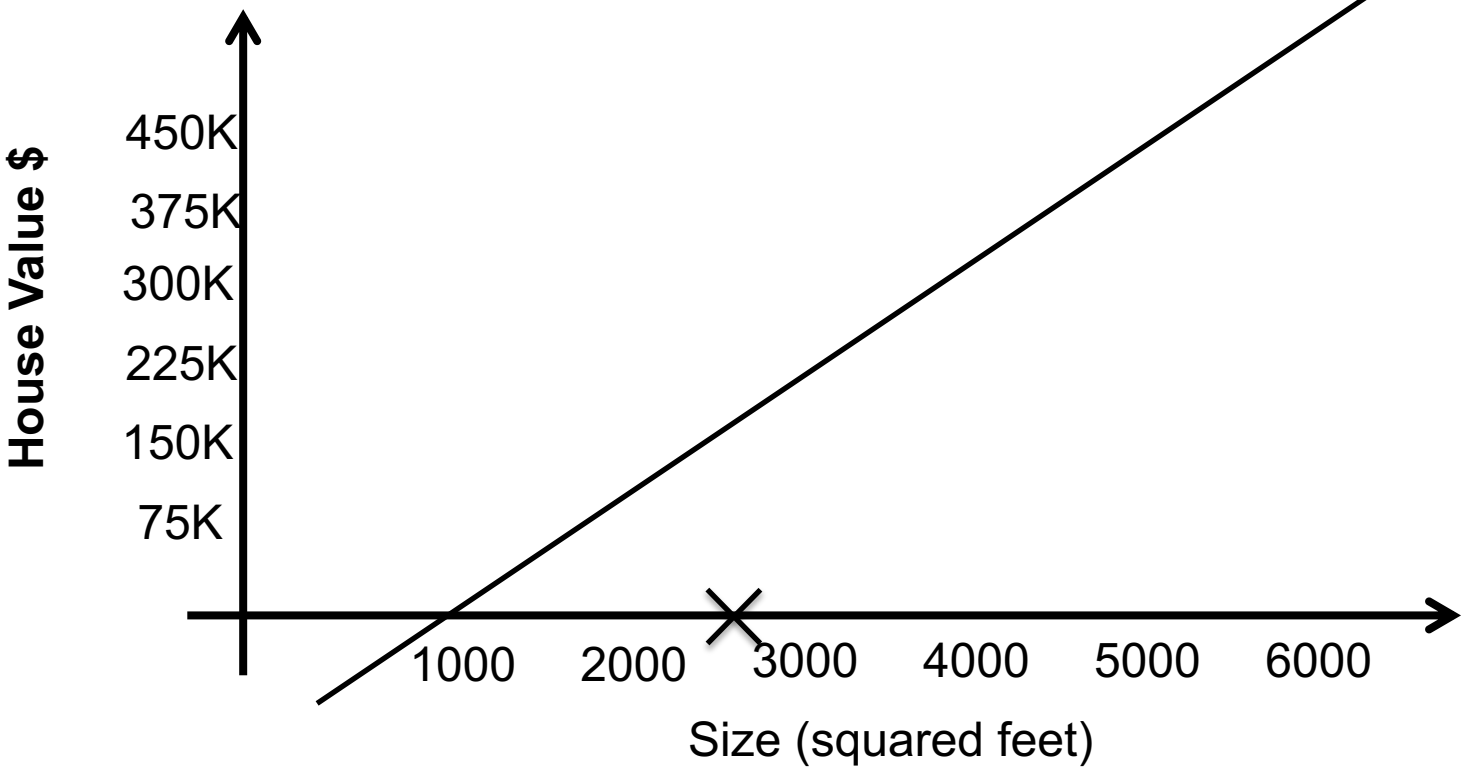


Example: House Values

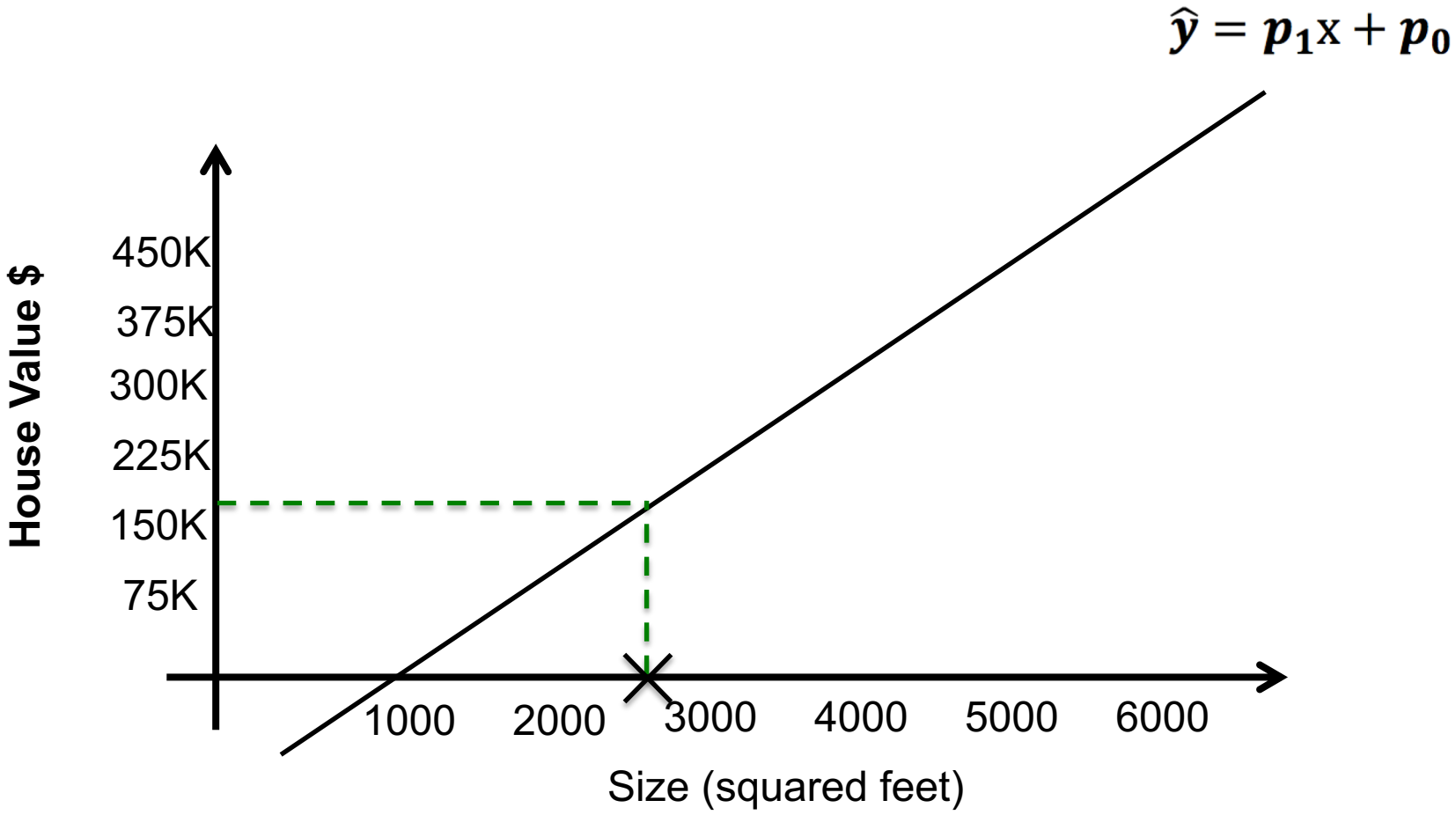


Example: House Values

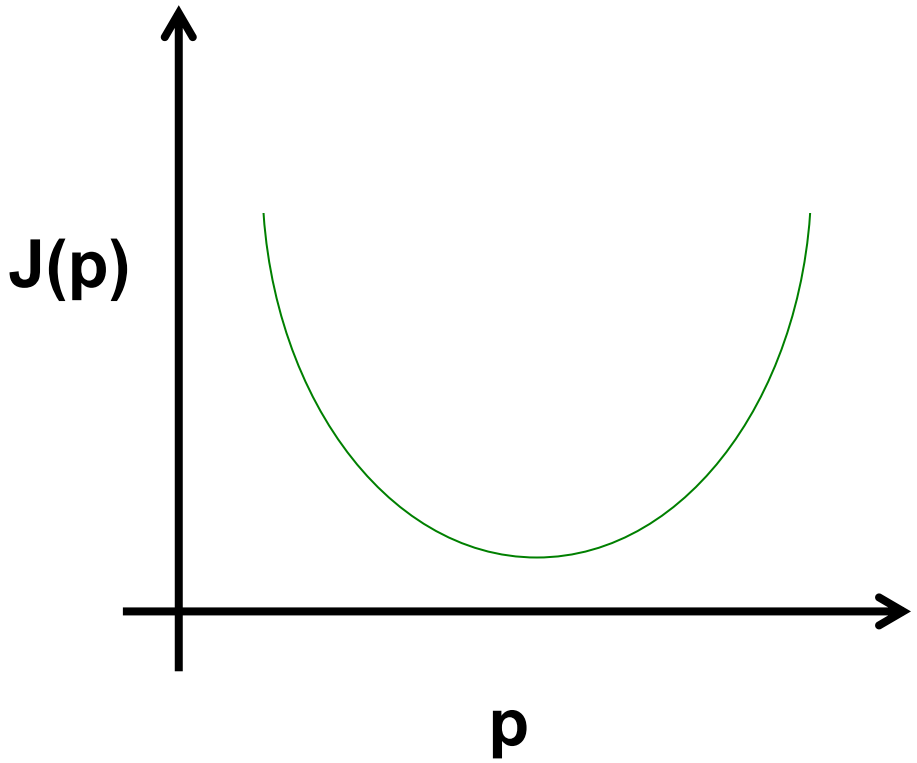
$$\hat{y} = p_1 x + p_0$$



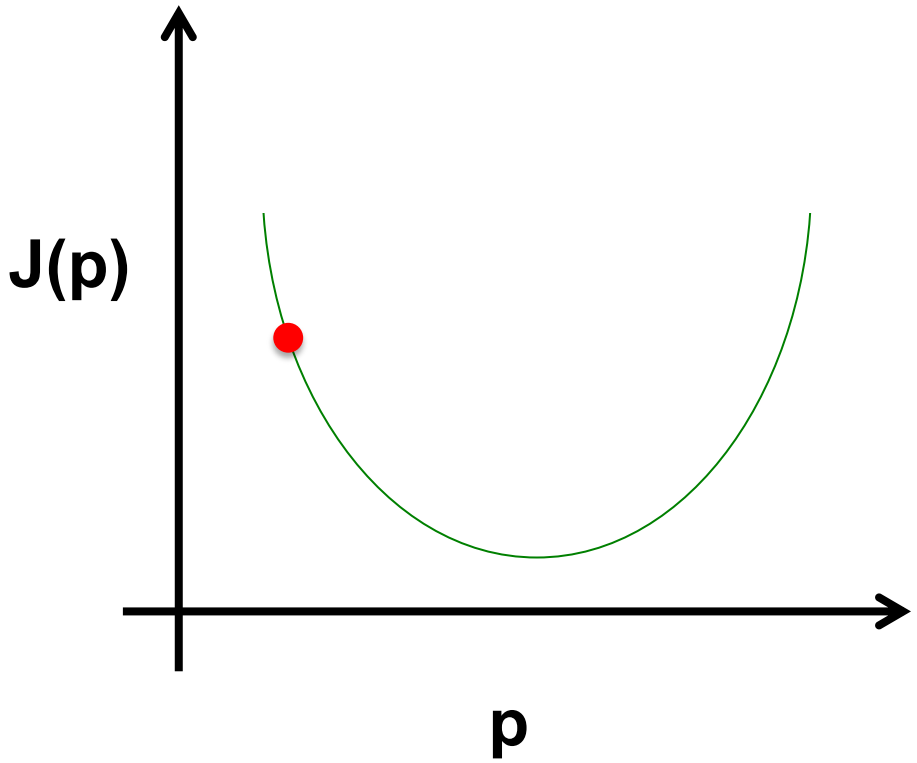
Example: House Values



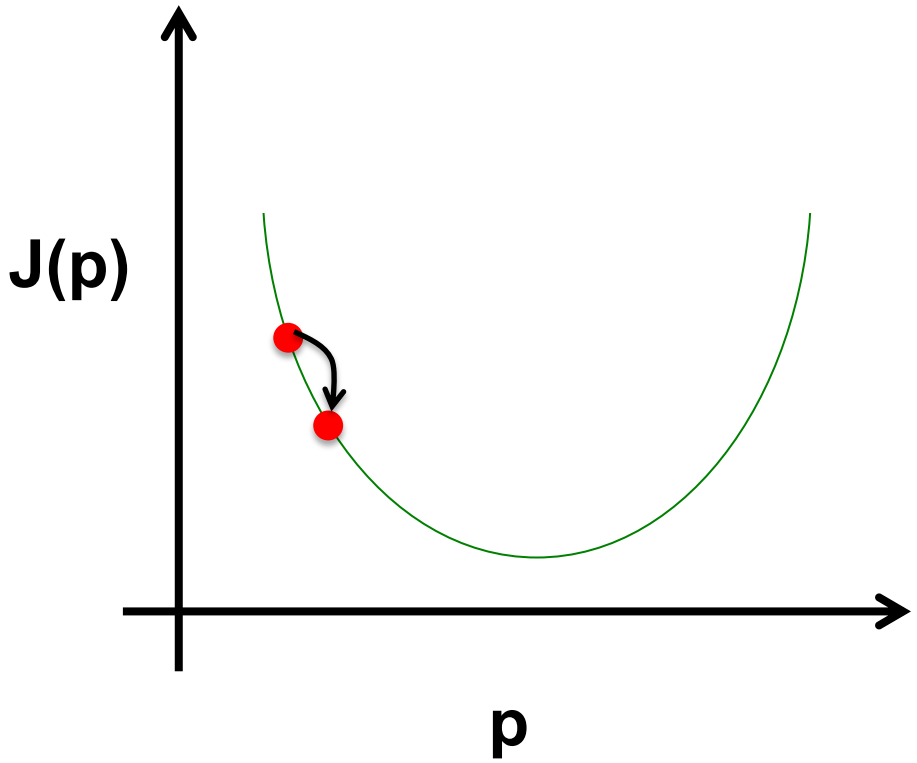
Convex Optimization Problem



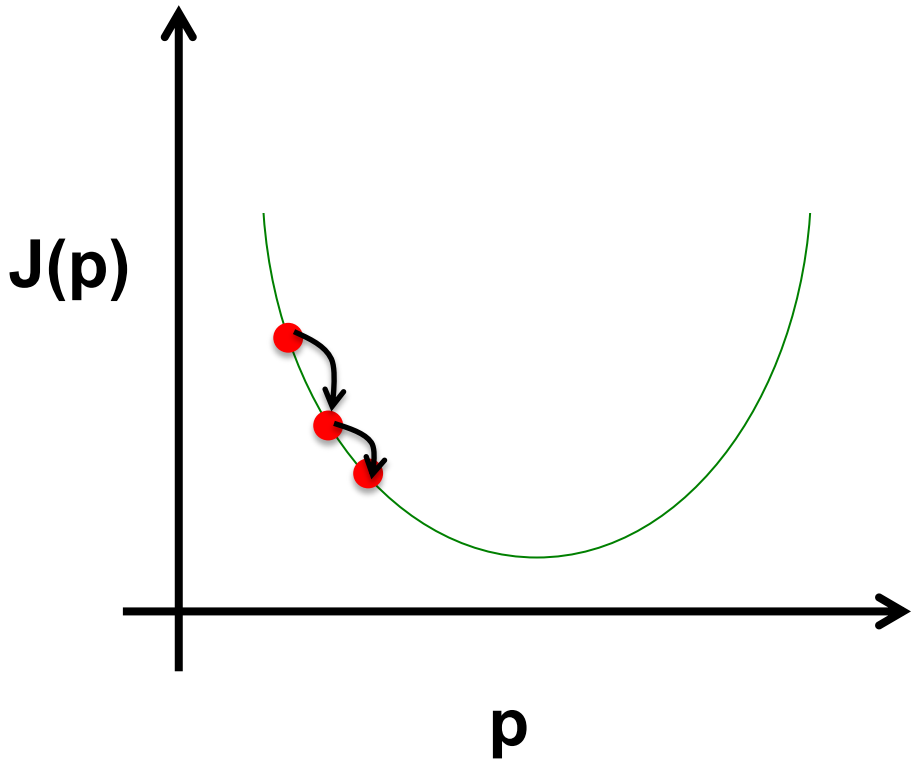
Convex Optimization Problem
Gradient Descent Method



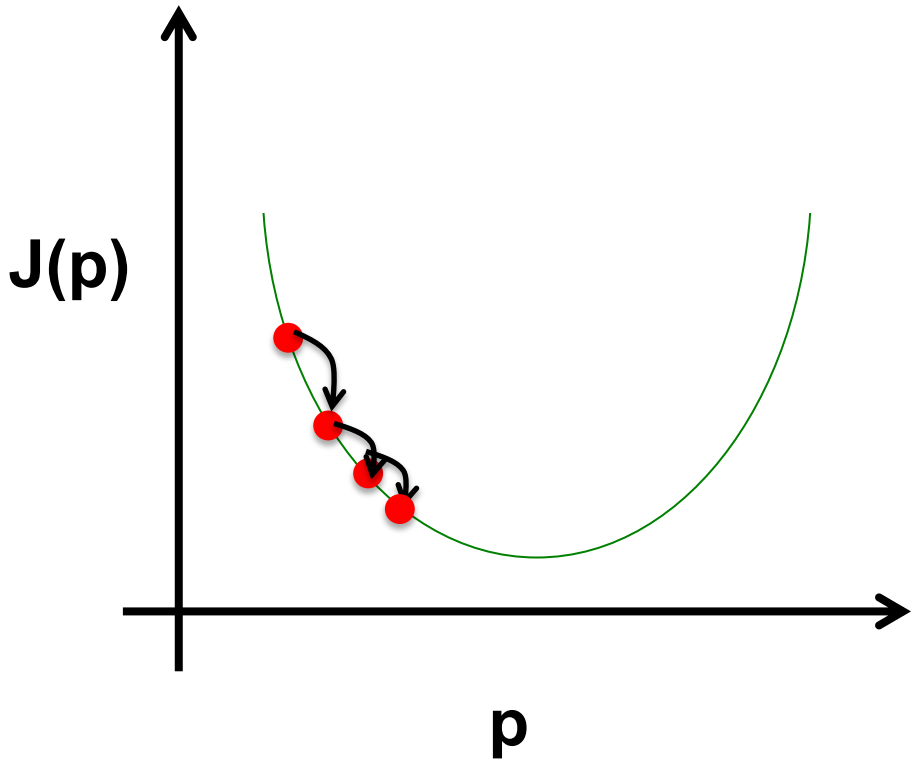
Convex Optimization Problem
Gradient Descent Method



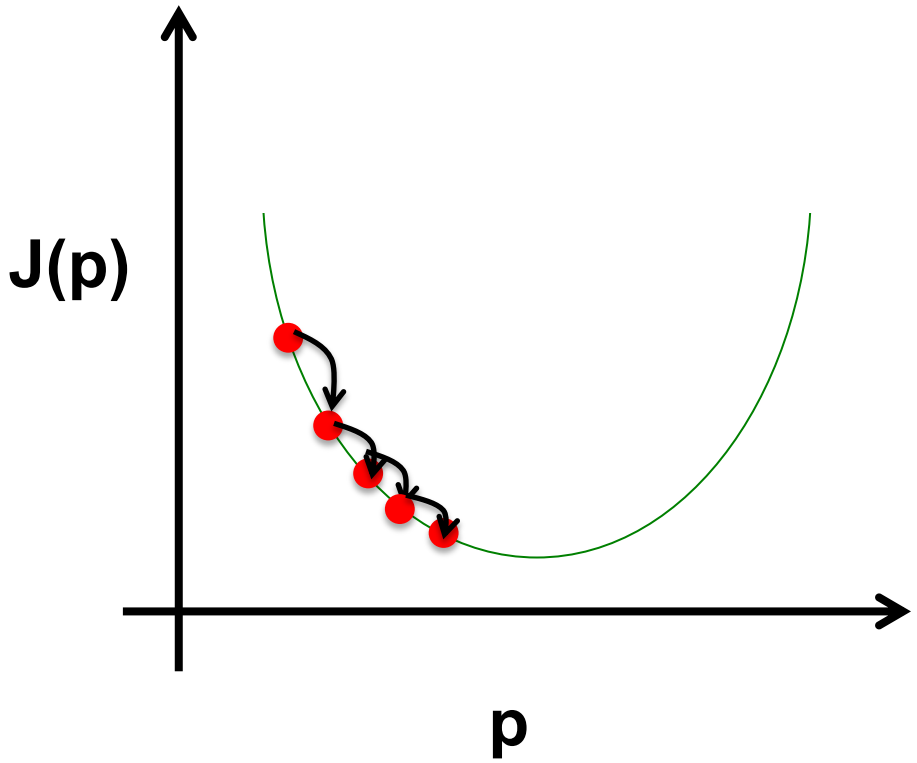
Convex Optimization Problem
Gradient Descent Method



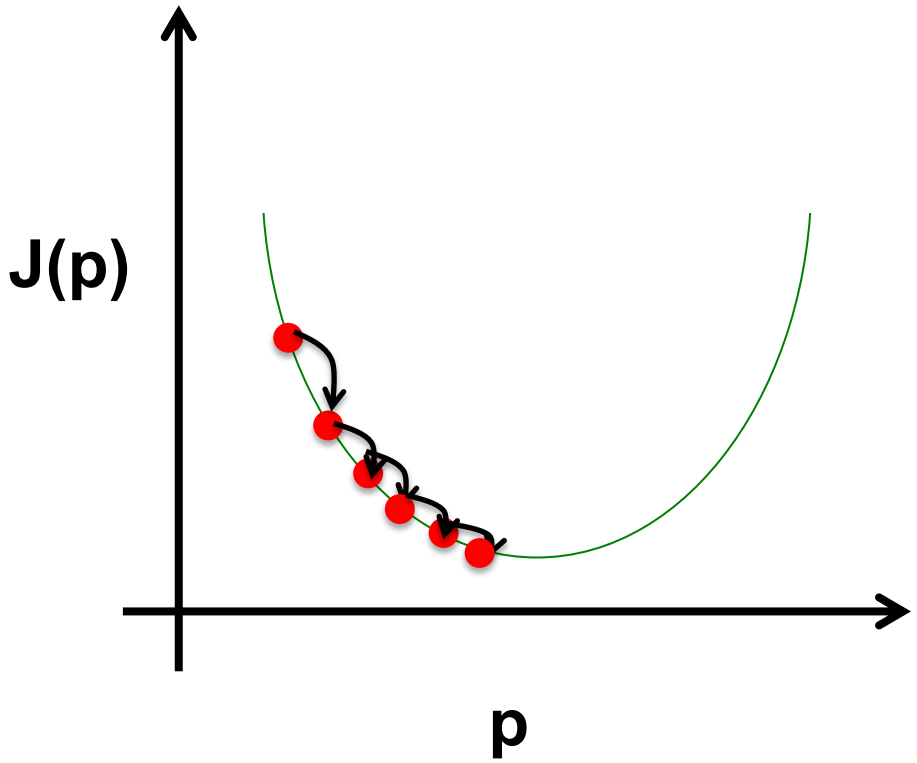
Convex Optimization Problem
Gradient Descent Method



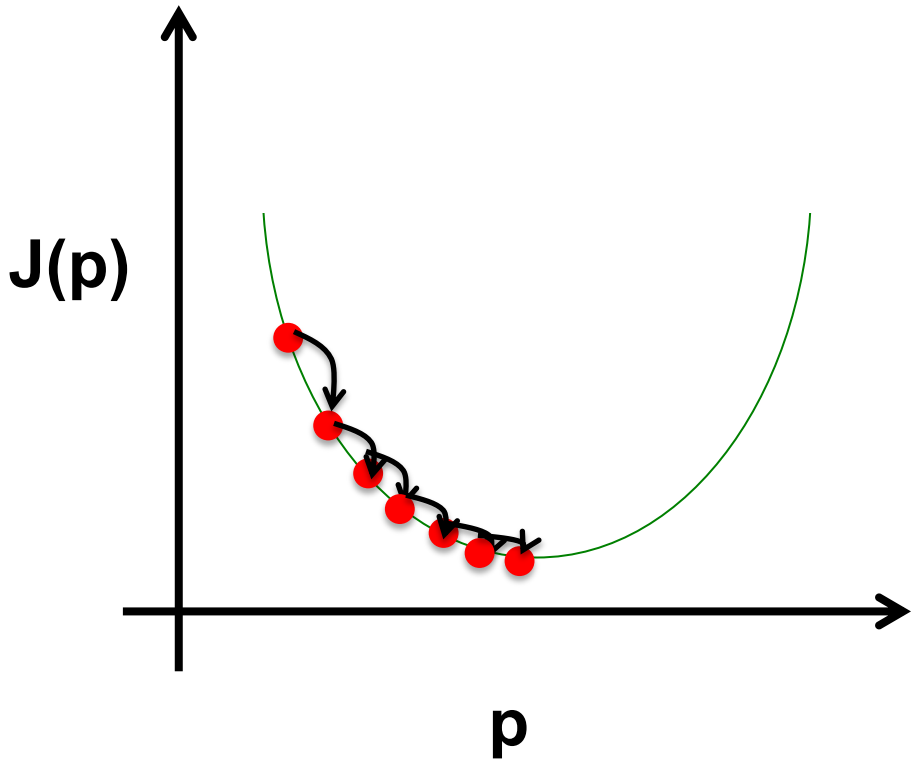
Convex Optimization Problem
Gradient Descent Method



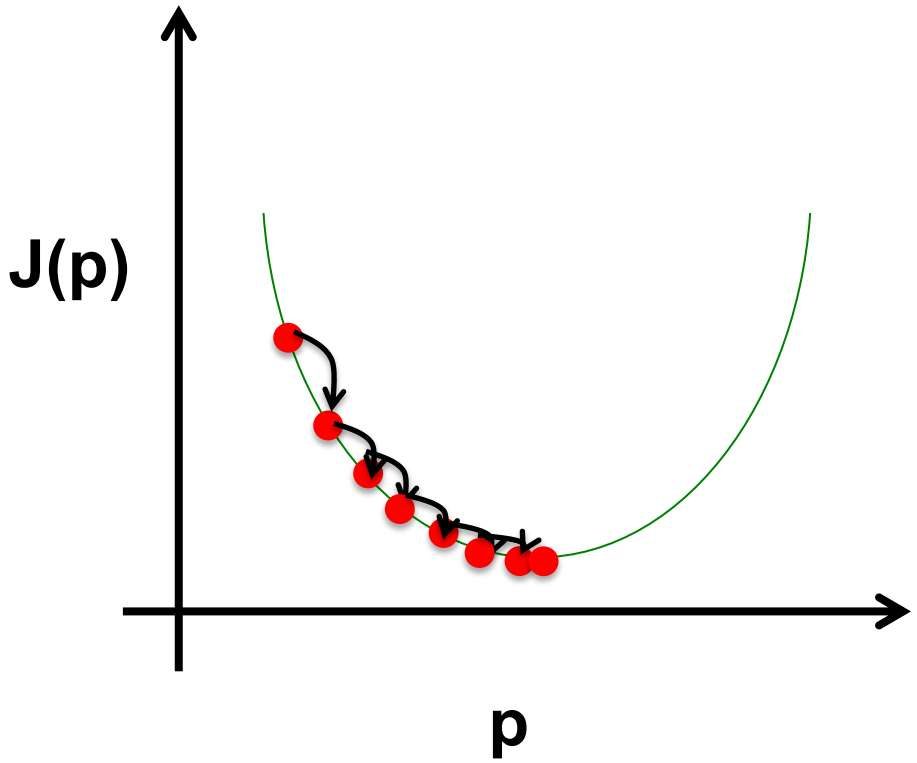
Convex Optimization Problem
Gradient Descent Method



Convex Optimization Problem
Gradient Descent Method



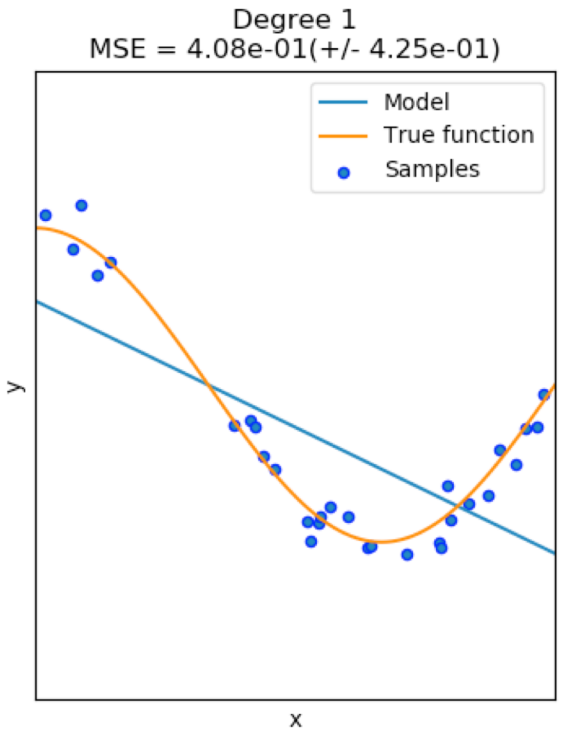
Convex Optimization Problem
Gradient Descent Method

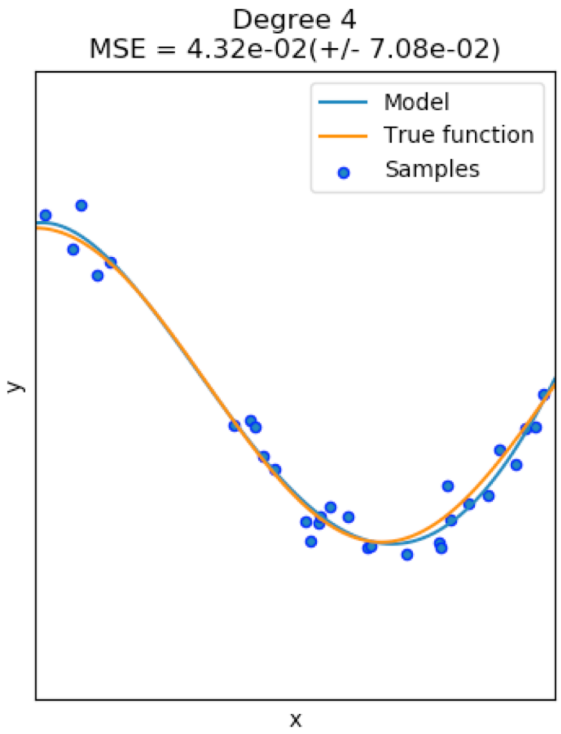


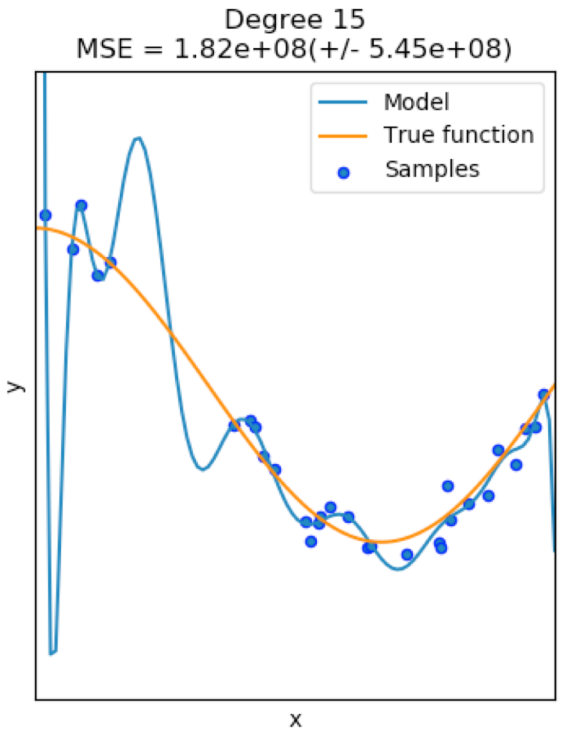
Machine Learning is More than Optimization; Generalization

- The ability to perform well on previously unobserved inputs is called generalization.

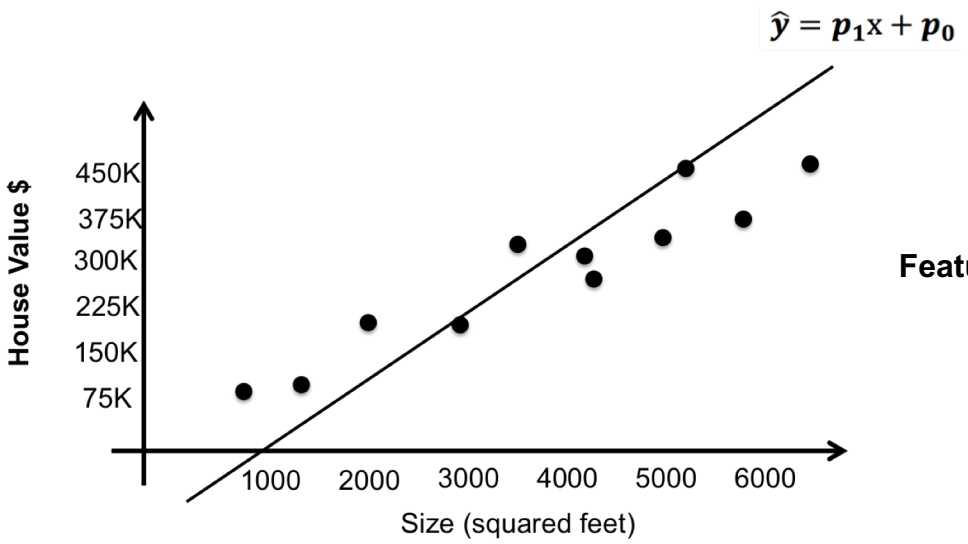
Underfitting and Overfitting

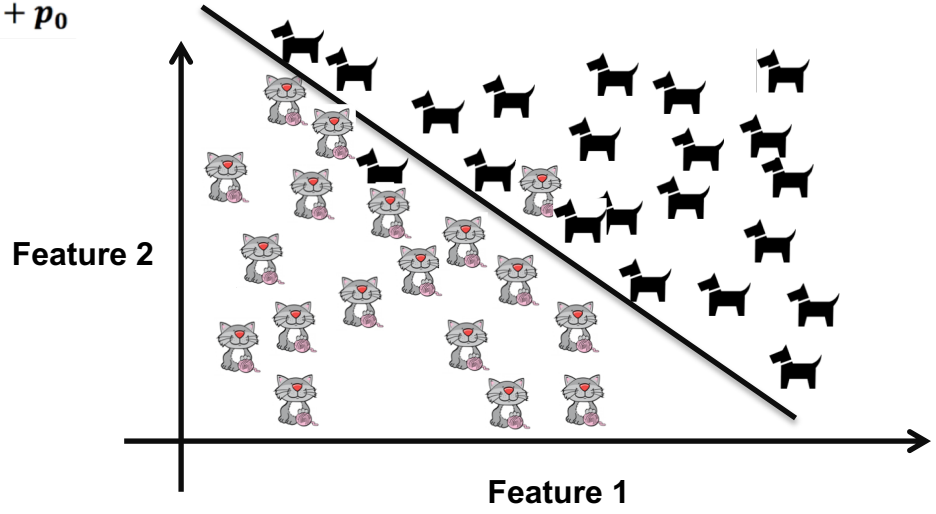






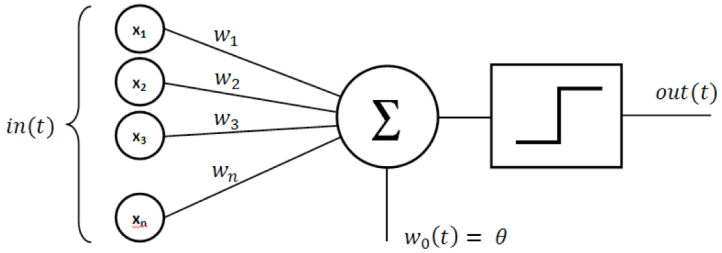
Regression vs. Classification

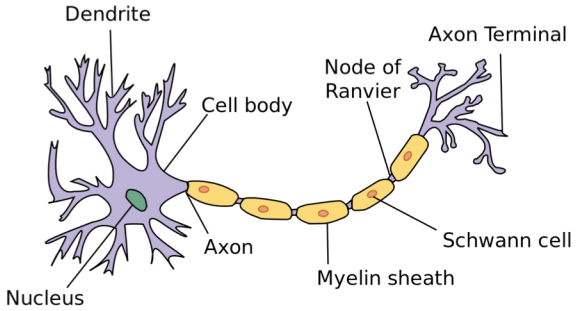




Perceptron: A Computational Neuron Model

$$y = f(WX^T) = f(\sum w_i x_i)$$



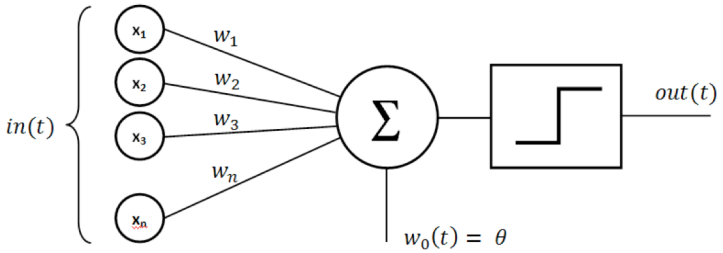


First introduced as a computational model for a nerve cell. And ever since it has carried the name of artificial neuron.

Perceptron: A Computational Neuron Model

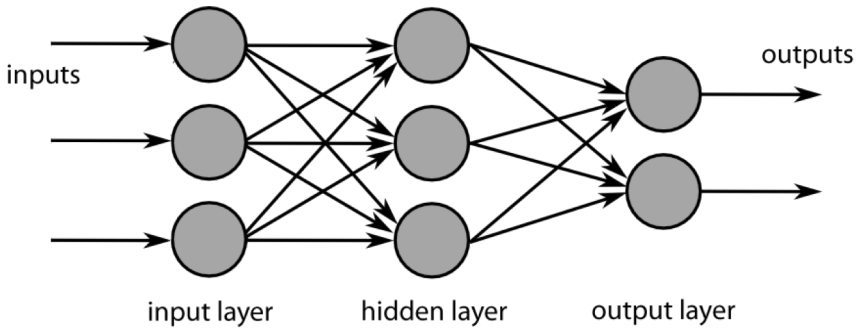
$$y = f(WX^T) = f(\sum w_i x_i)$$

- Learning parameters (cost function minimization) is a convex problem.
- But very little computational power.
- Cannot even implement a XOR gate.



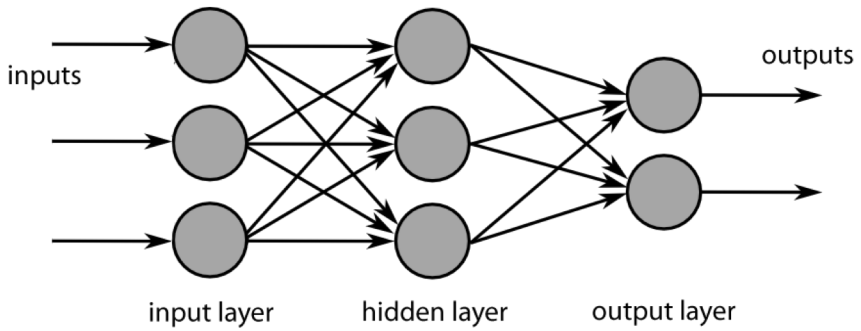
Multilayer Perceptron
Deep Feedforward Neural Network

- Arrange perceptrons (neurons) in a network. The result is a Neural Network.



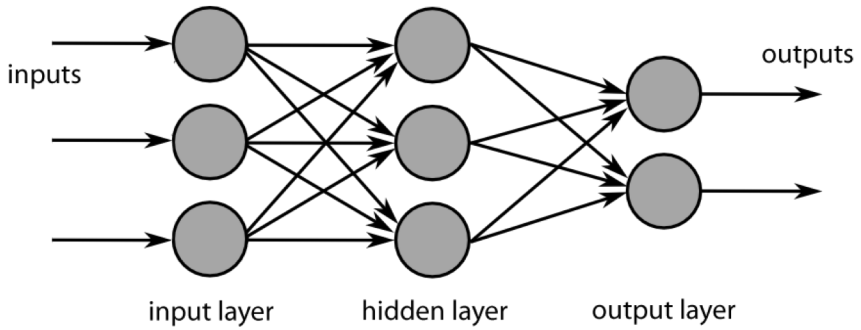
Multilayer Perceptron
Deep Feedforward Neural Network

- Multilayer Perceptron is an extremely powerful learning method.
- It is a universal function approximator (some form of a universal computing machine).



Multilayer Perceptron
Deep Feedforward Neural Network

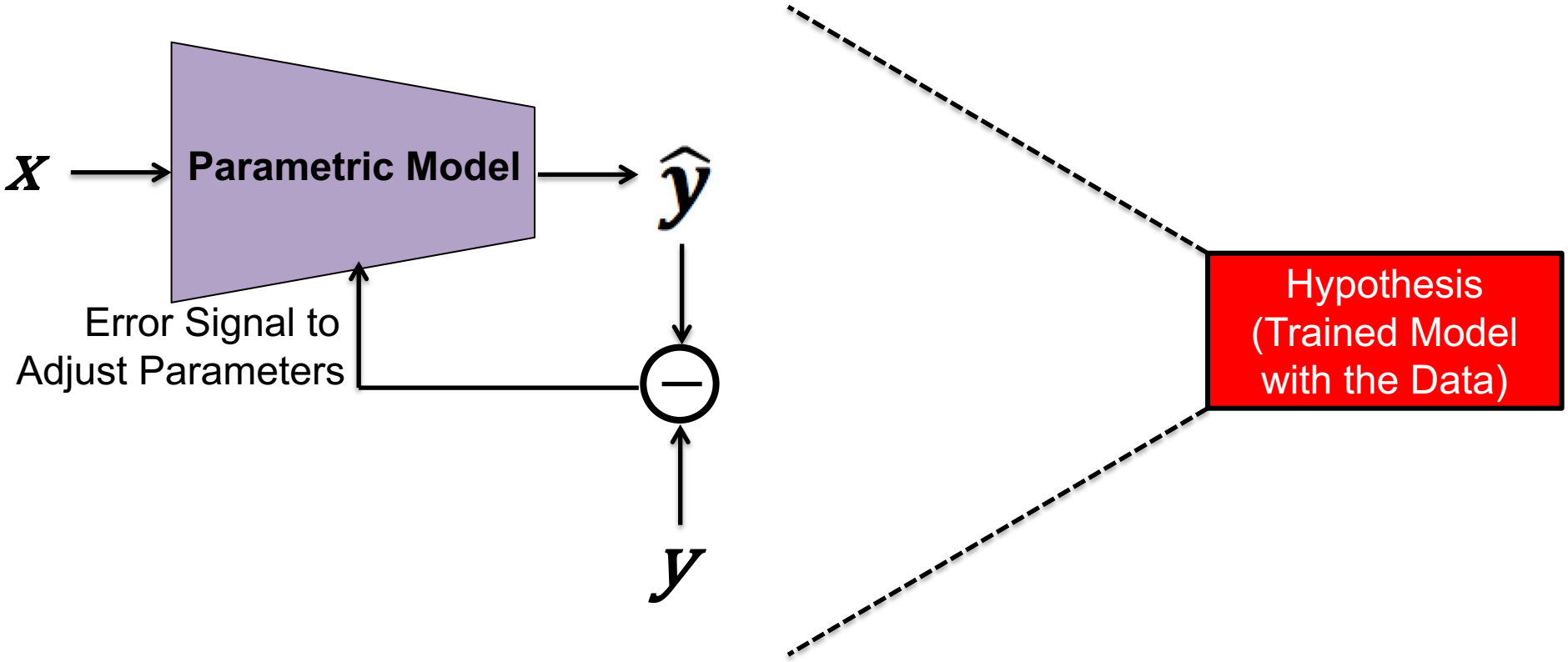
- Multilayer Perceptron is an extremely powerful learning method.
- It is a universal function approximator (some form of a universal computing machine).
- **But how to train it?**



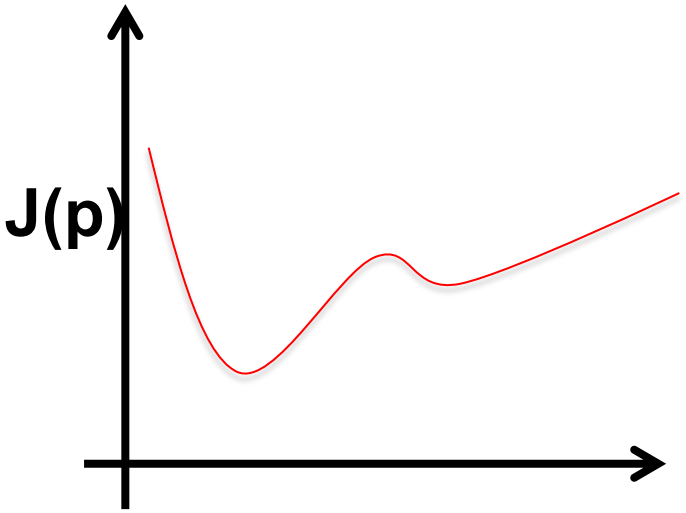
Machine Learning: Training Model with Data

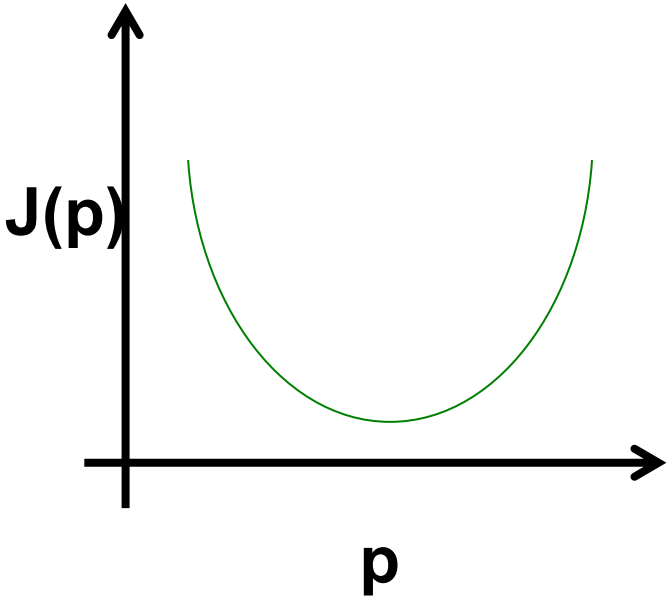
Training Data: (x_i, y_i), i=1,2,3,...,N

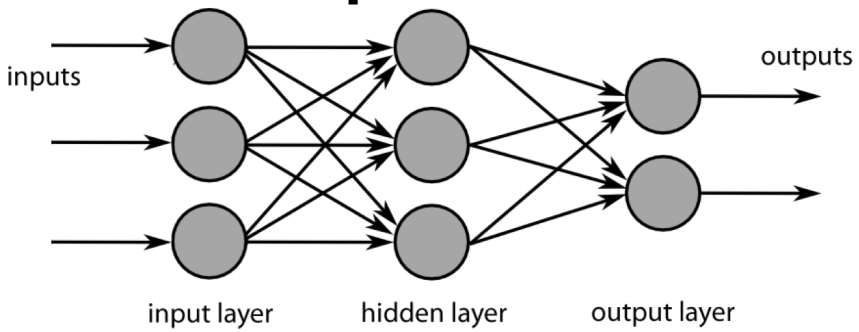
x is called features, y is label.



Multilayer Perceptron: Training
A Non-Convex Optimization Problem

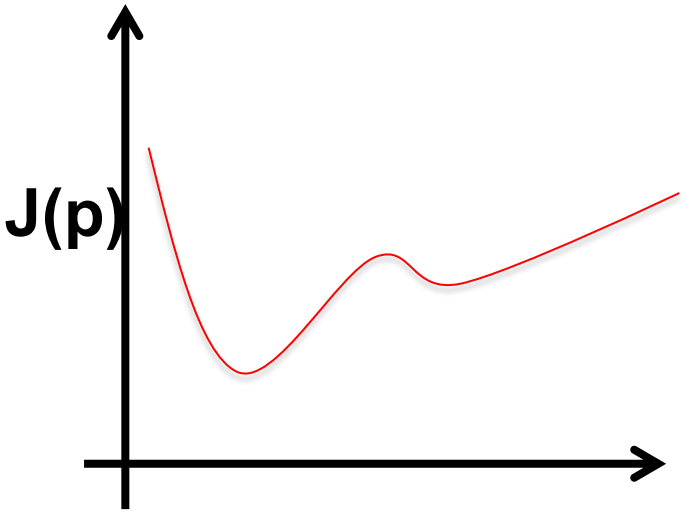


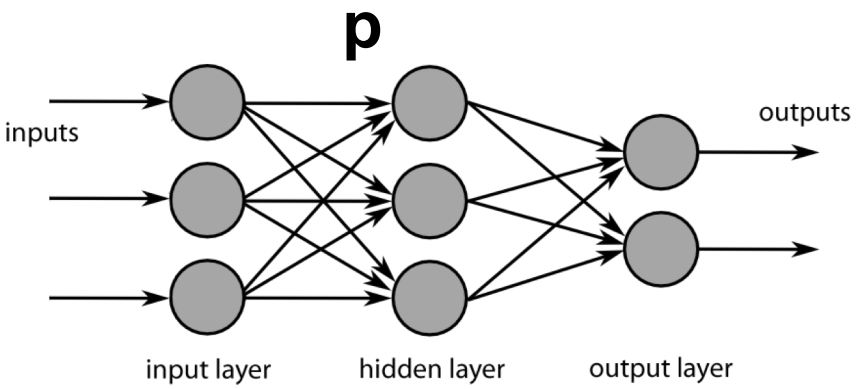




Linear methods
Perceptron
SVM
…

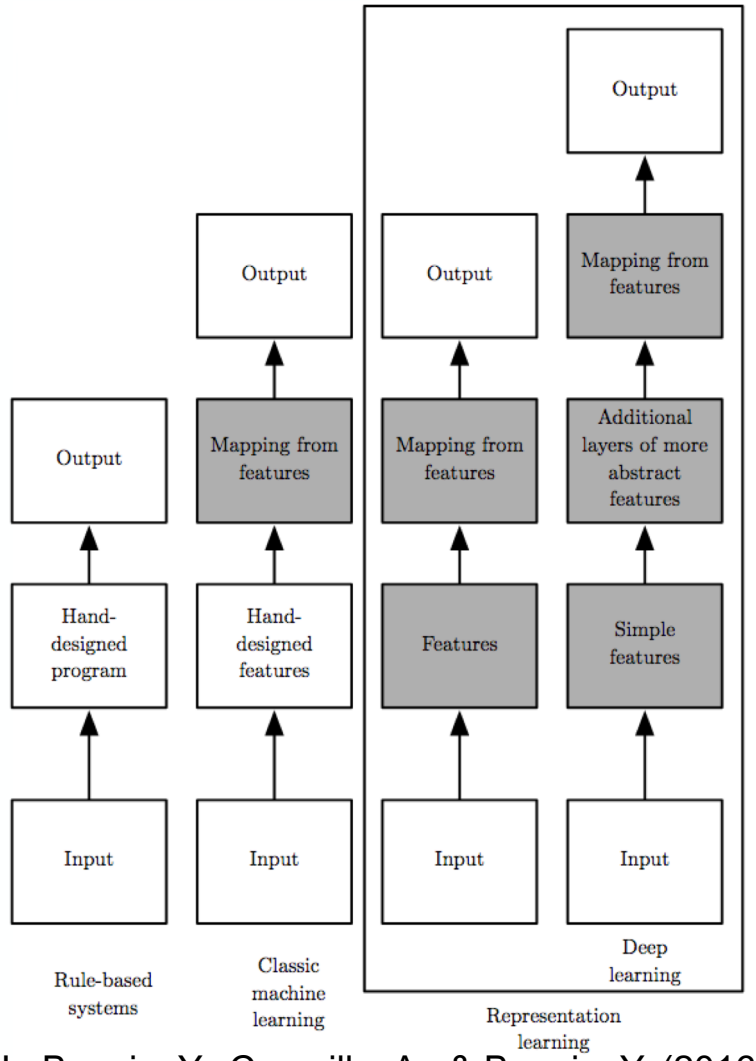
Multilayer Perceptron: Training A Non-Convex Optimization Problem





Non-convex optimization problems are usually NP-Hard!

Different Models of AI Systems Over Time



Deep learning, Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Cambridge: MIT press.

Pioneers and Heroes of Deep Learning

Geoffrey E. Hinton

Yann LeCun

Jürgen Schmidhuber

Yoshua Bengio

- The freedom to express the learning problem as a non-convex optimization problem gives immense modeling power to the algorithm designer, but often such problems are NP-hard to solve.

Pioneers and Heroes of Deep Learning

Geoffrey E. Hinton

Yann LeCun

Jürgen Schmidhuber

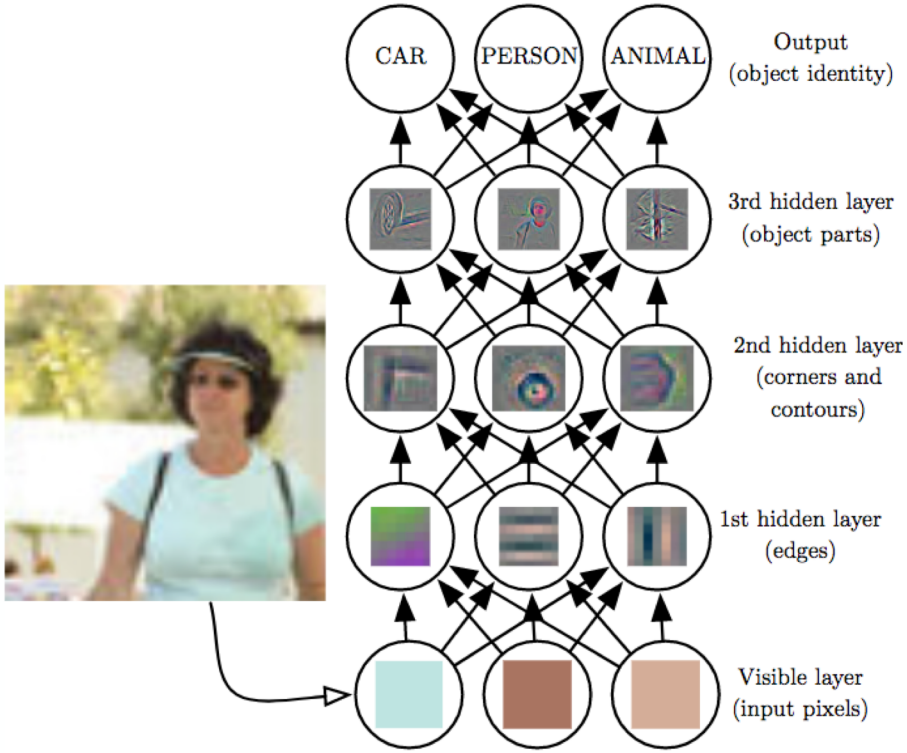
Yoshua Bengio

Deep Learning

Deep Learning, a Machine Learning technique that learns to represent the world as a deep nested hierarchy of concepts.

Deep Learning

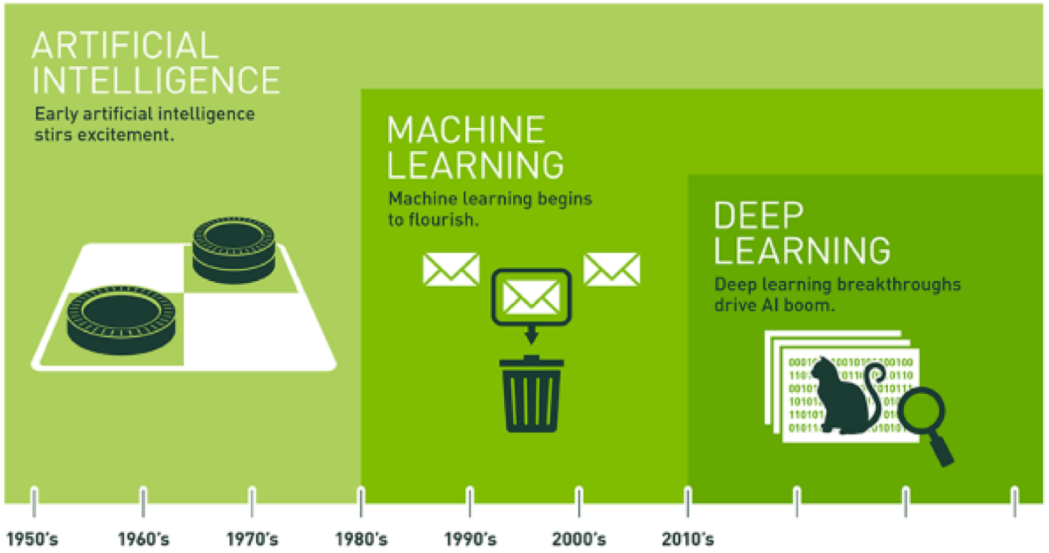
- Give the entire data to the network, and it itself picks a chooses the features and patterns.
- Requires many layers of network, therefore such networks are called deep networks.



Deep learning, Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Cambridge: MIT press.

Machine Learning

- Machine learning algorithm is an algorithm that learns from data with no need for explicit programing.



Picture from NVIDIA's deep learning institute

Pioneers and Heroes of Deep Learning

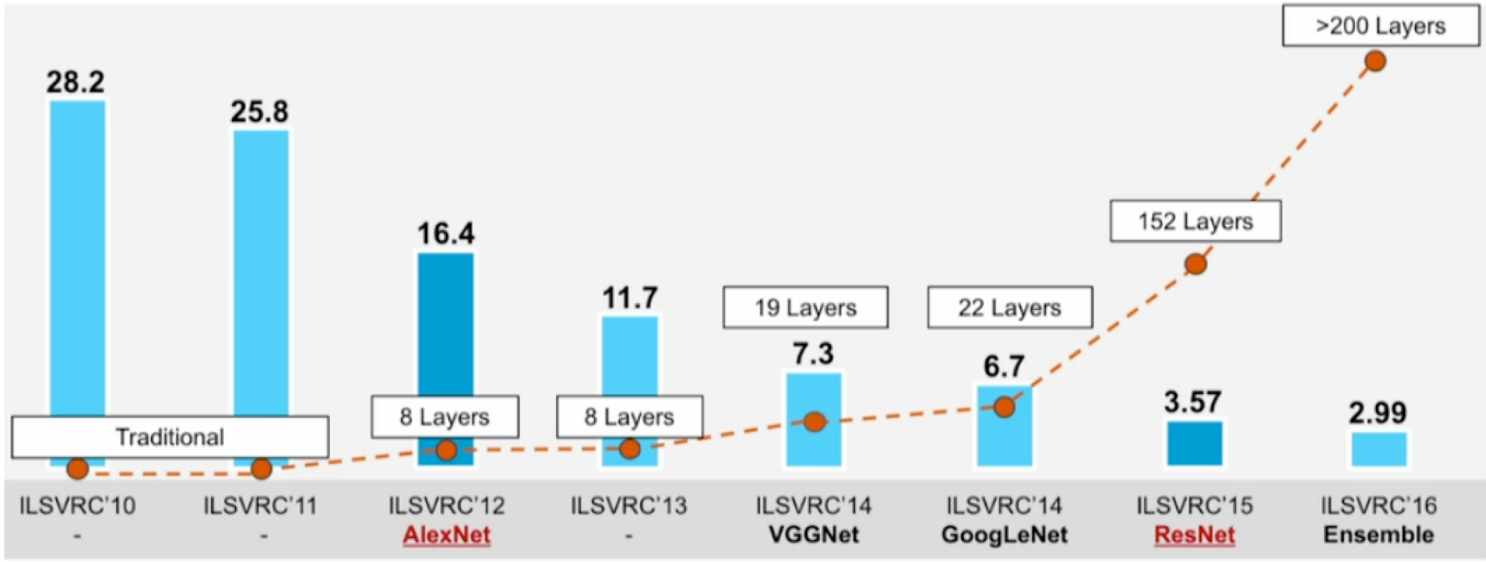
Geoffrey E. Hinton

Yann LeCun

Jürgen Schmidhuber

Yoshua Bengio

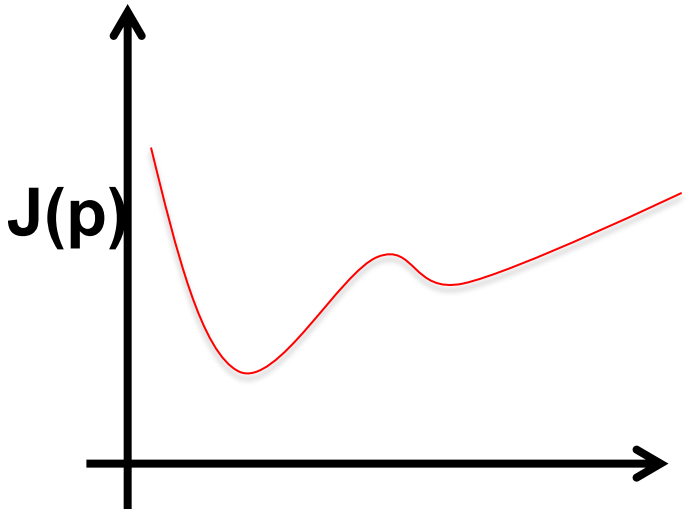
2012 ImageNet Large Scale Visual Recognition Challenge (ILSVRC2012)

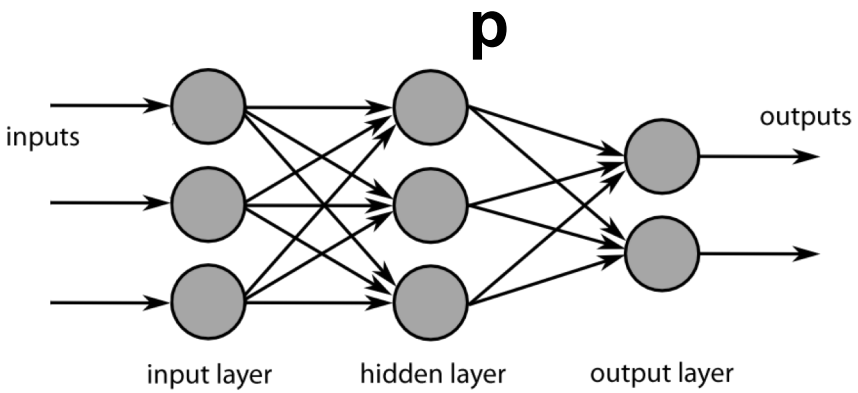


Picture from http://sqlml.azurewebsites.net/

Deep Learning
So how did they solve the training problem after all?

- Back propagation method, known for many years
- Stochastic Gradient Descent
- A lot of Data
- A lot of Compute power (GPU)



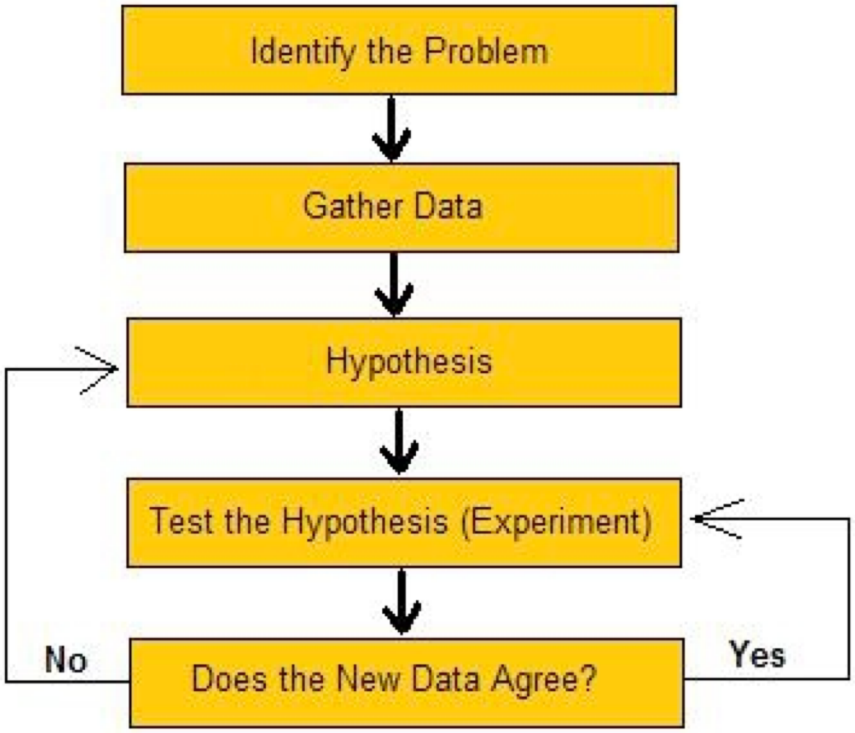


Parametric Models:

- Linear Models
- Perceptron
- Multilayer Perceptron
- Convolutional Network
- Recurrent Neural Network
  - LSTM
  - GRU
- .
- .
- .

Searching as an AI Mechanism

- AI: solving problems that don't have a formal solution.
- Explore and exploit the solution state and try them on the problem to see it solves it or not.



Scientific Method

Next Session

Programing And Development Tools

(Or How Easy It Is to Develop AI Solutions)

Reading Assignment: Deep Learning, Chapter 1 Introduction (pages 1-26)