CSE-4820/5819 Introduction to "Intro to Machine Learning"

Suining He
Department of Computer Science and Engineering
University of Connecticut
suining.he@uconn.edu

Outline

- Introduction
- Grading Scheme
- Basic Structure and Course Introduction

About Instructor

- Suining He
  - Email: suining.he@uconn.edu
- Research interests:
  - Human-centered AI
  - Location-based services, indoor localization
  - Ubiquitous sensing, data mining
  - Mobile and urban computing
- Hobbies:
  - Soccer, jogging, hiking, basketball, movies …
- Come to ITE 337, chill out (no machine learning), tea is available

Course Prerequisites

- MATH 2210Q; STAT 3025Q or 3345Q or 3375Q or MATH 3160
- Strongly recommended preparation: CSE 3500
- Basic Programming Skills are required (Python is preferred)

Recommended Textbooks

- *Paper copies are not required*
- **Pattern Recognition and Machine Learning.**
  - Christopher M. Bishop.
- **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**, 2nd Edition.
  - Aurélien Géron
- **Deep Learning with Python**, 1st Edition.
  - Francois Chollet
- **Introduction to Data Mining**, 2nd Edition.
  - Pang-Ning Tan, Michael Steinbach and Vipin Kumar
- **Convex Optimization**.
  - Boyd, Stephen, Stephen P. Boyd, and Lieven Vandenberghe.

Course Design

- Harness concepts and theories of the basic machine learning algorithms, models and mechanisms
- Hand-on practices of the basic ML algorithms and models
- Understand practical applications and design principles

Grading

- **CSE 4820:** Course performance: 100%
  - 40% for 4 written/programming assignments
  - 60% for 3 exams (1st, 2nd and Final Exams)

- **CSE 5819:** Course performance: 100%
  - 20% for 3 written/programming assignments
  - 30% for 2 exams (1st and 2nd Exams)
  - 50% for group project (2 students per team)
    - Project proposal report and presentation (15%)
    - Project mid-term report (10%)
    - Project final report and presentation (25%)

- **Bonus:** (total up to 5% for CSE4820/5819)
  - In-class/take-home quizzes/programming assignments (practice exams)
  - Total 8~10 trials (depends on course progress) sums up to 5% extra; equal weight for each

- Each student will have a total of **5** free late (calendar) days to use for homework. Each 24 hours or part thereof that an assignment is late uses up one full late day. Please note: once these late days are exhausted, no late assignments will be accepted for any reason. Students are highly encouraged to reserve your late days for unavoidable emergencies, planned travel, etc.

Overview of Course Structure

- Introduction
- Basic linear algebra and statistics
- Introduction to optimization
- Linear regression
- Logistic regression

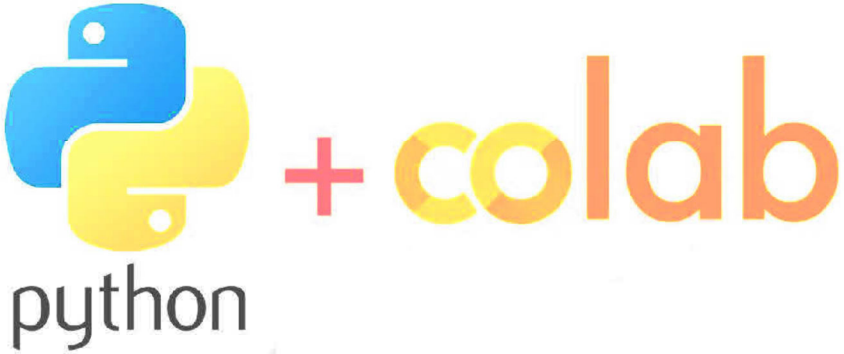
Foundations of machine learning

- Support vector machine
- Machine learning evaluation
- Clustering
- Neural network
- Convolutional neural network
- Recurrent neural network

Machine learning algorithms and models

Hands-on Machine Learning

- Google Colab for demo and test
- Tutorials and programming assignments



Assignment 0 (Optional)

- Please prepare a short description (~100 words) about yourself:

  (1) List any courses (inside/outside UConn) you have taken before related to data, machine learning and deep learning;

  (2) List any programming language you are comfortable/familiar with;

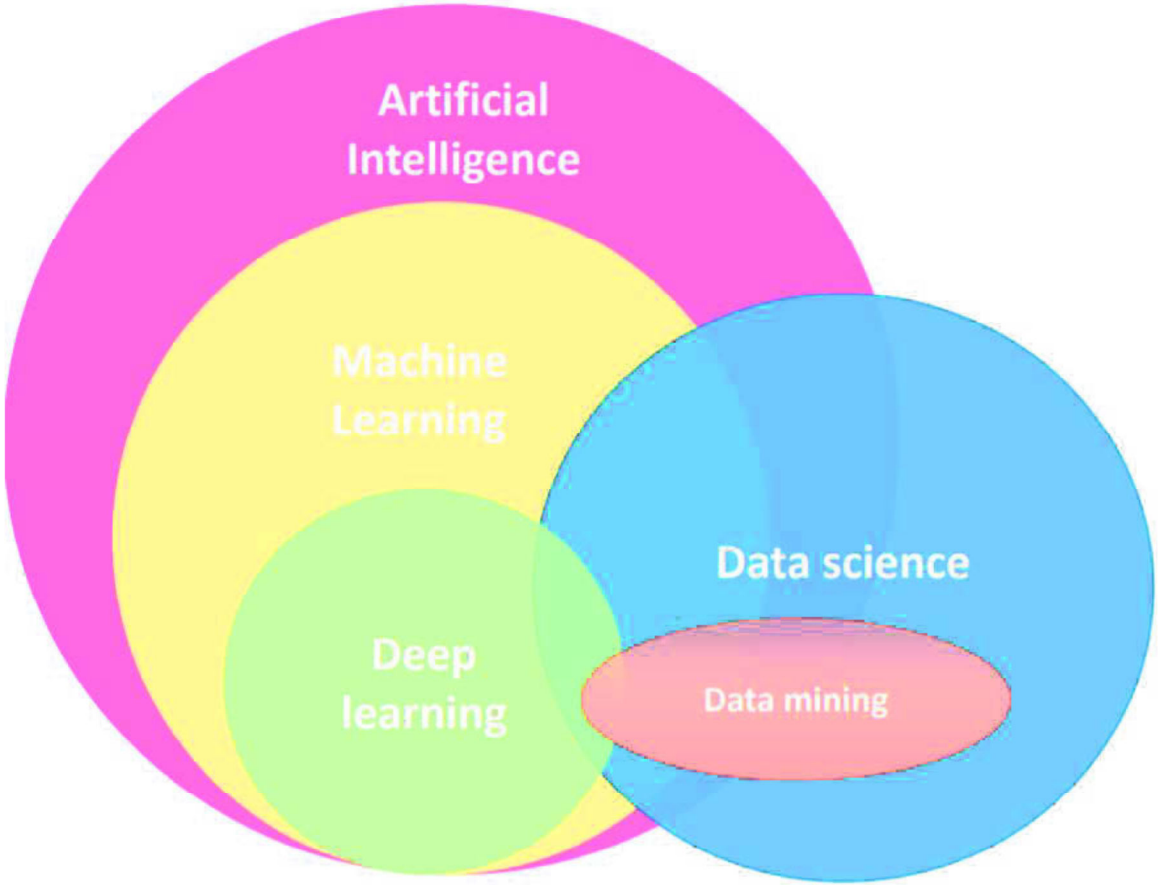
  (3) Any research interest or background?

- Due: 23:59 (end of day), August 30, through HuskyCT

What is Machine Learning

- **Machine Learning (ML)** is the science of making computer artifacts improve their performance with respect to a certain performance criterion using example data or past experience, without requiring humans to program their behavior explicitly.

Where ML Stands



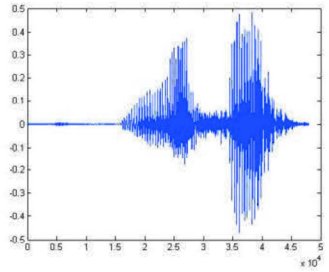
Where ML Stands

- Data mining (DM) which is also known as Knowledge Discovery Process is a field of science that is used to find out the properties of datasets
- Artificial Intelligence (AI) is a branch of science which deals with the creation of intelligent machines. These machines are called intelligent as they have their own thinking and decision-making capabilities like human beings

Why Machine Learning?

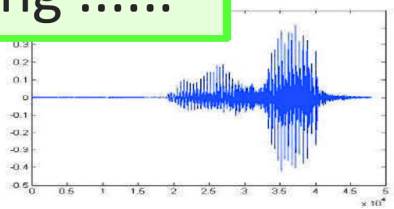
- Human expertise is too expensive
  - e.g., intrusion detection, pathology
- Human expertise does not exist
  - e.g., navigating on Mars
- Humans cannot explain their expertise
  - e.g., speech recognition, visual perception
- Problem (and hence solution) changes over time
  - e.g., network routing
- Solution needs to be adapted to particular cases
  - e.g., personalized adaptive user interface

Machine Learning

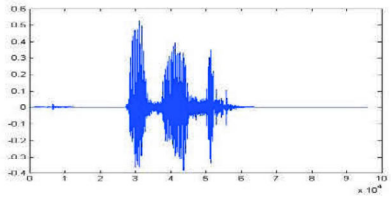


You said “Hello”

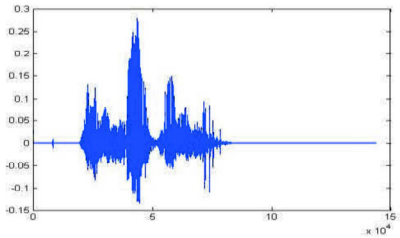
Learning ……



“Hi”



“How are you”



“Good bye”

A large amount of audio data

You write the program for learning.

Machine Learning

This is “cat”

Learning .....

You write the program for learning.

“monkey”

“cat”

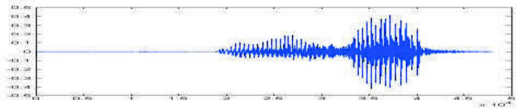
“dog”

A large number of images

ML ≈ Finding a Function $f()$

Based on Training Data

- Speech Recognition

$f($  $)=$ "How are you"

- Image Recognition

$f($  $)=$ "Cat"

- Playing Go

$f($  $)=$ "5-5" (next move)

- Dialogue System

$f($ "How are you?" $)=$ "I am fine."

(what the user said) (system response)

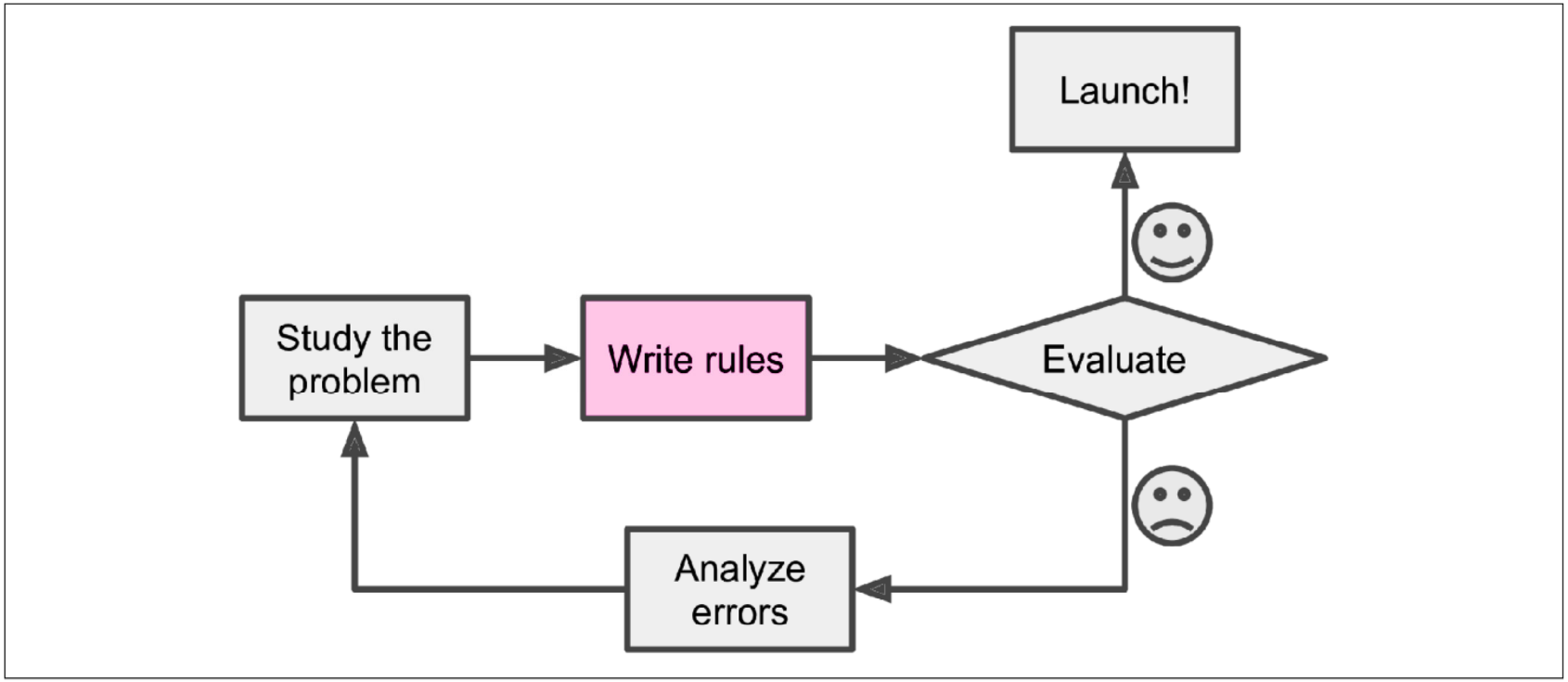
Some Features of ML

- Data is cheap and abundant (Big Data);
  - Knowledge is expensive and scarce.
  - Details of the data generation process may be unknown, but the process is not completely random.
- Learning models from data by exploiting certain patterns or regularities in the data
  - Inverting the data generation path.
- A model is often not an exact replica of the complete process, but is a good and useful approximation.
  - (George Box: “All models are wrong, but some are useful.”)
- A model may be descriptive to gain knowledge from data, or predictive to make predictions in the future, or both.
  - Almost all of science is concerned with fitting models to data

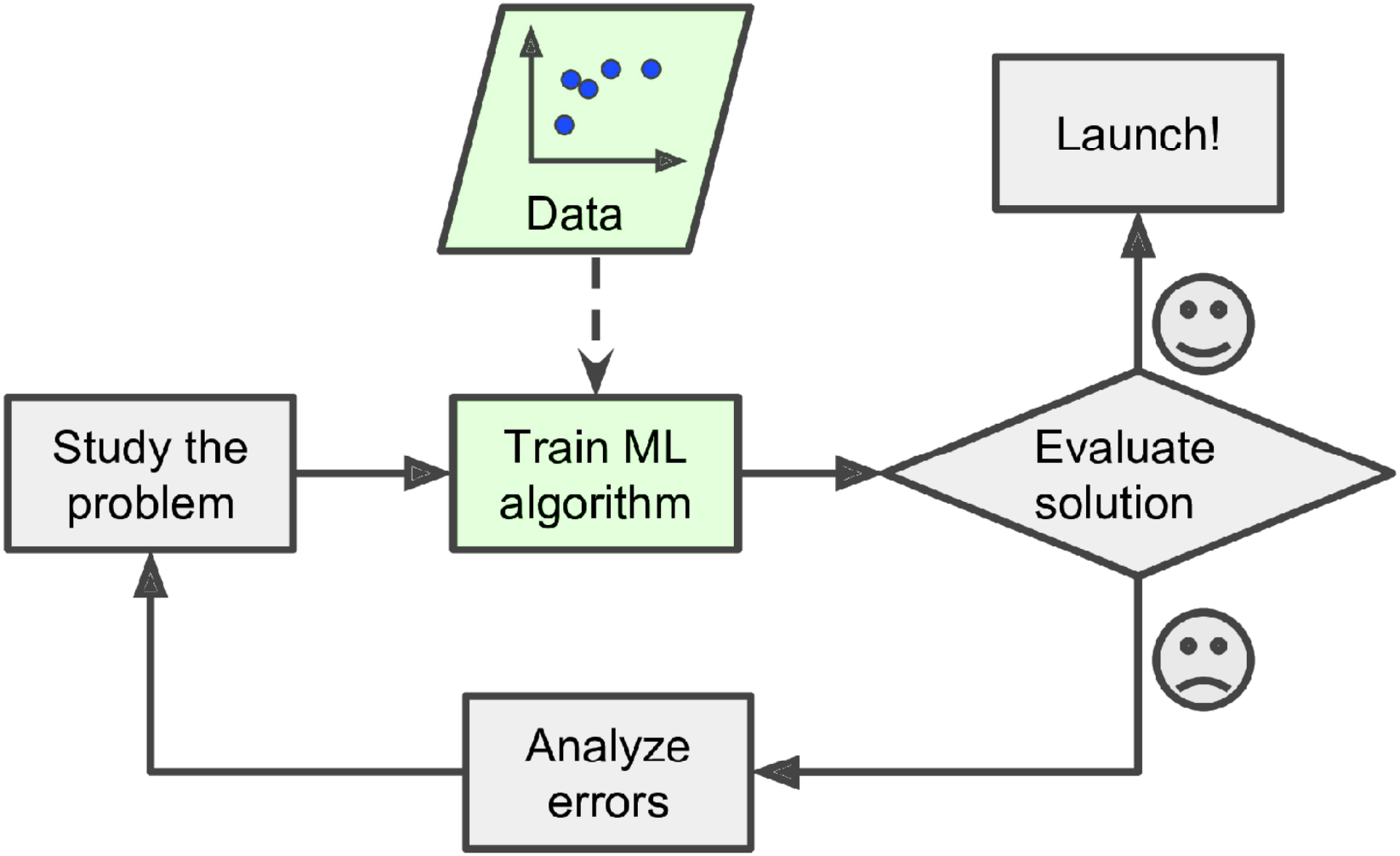
Role of Statistics and ML

- Machine learning makes extensive use of statistics in building mathematical models, because the core task is to make inference from a sample of observations.
- Role of computer science:
  - Developing efficient and accurate learning and inference algorithms
  - Common performance criteria:
    - Prediction accuracy, time complexity, space complexity.

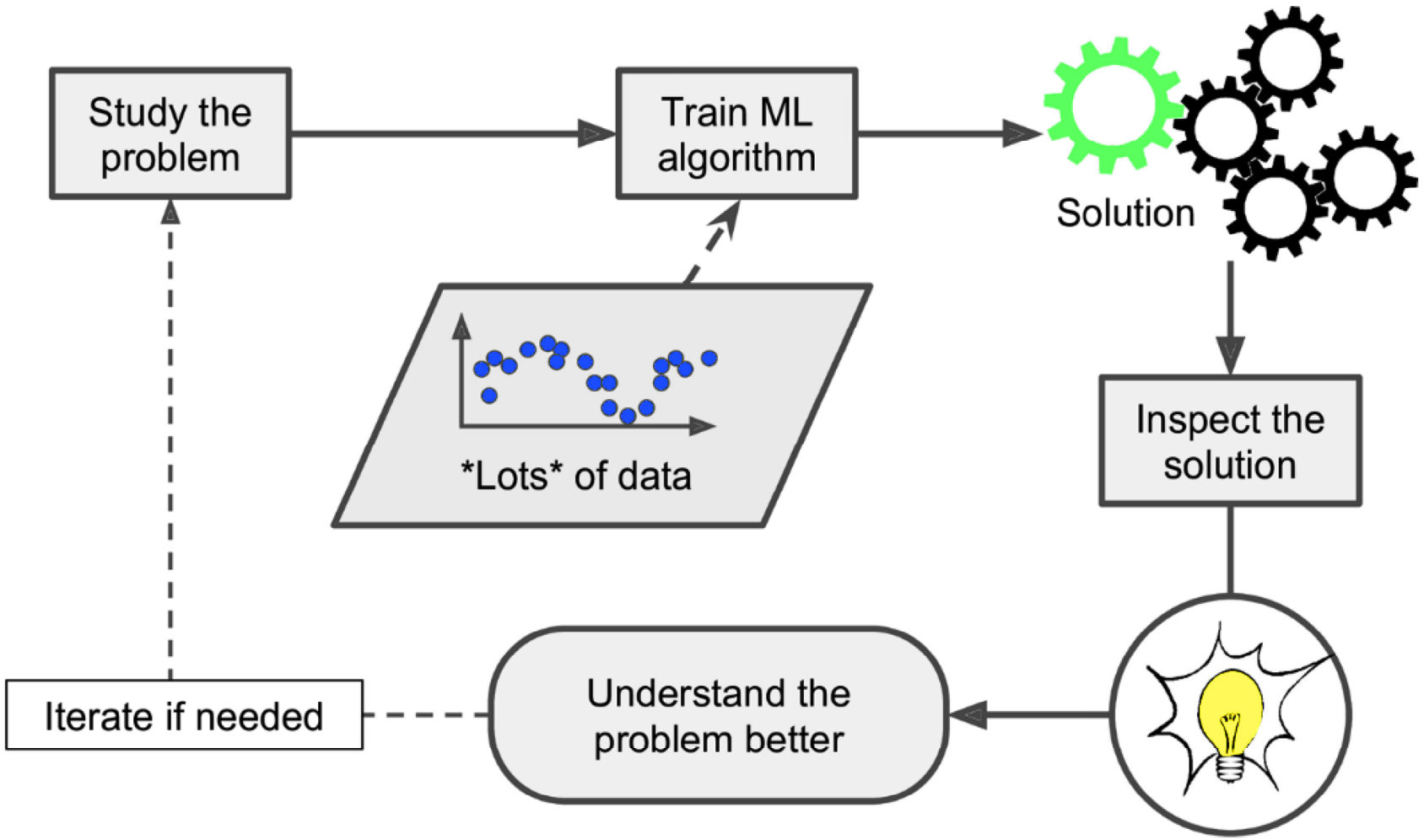
Traditional Learning Approach



Machine Learning Approach



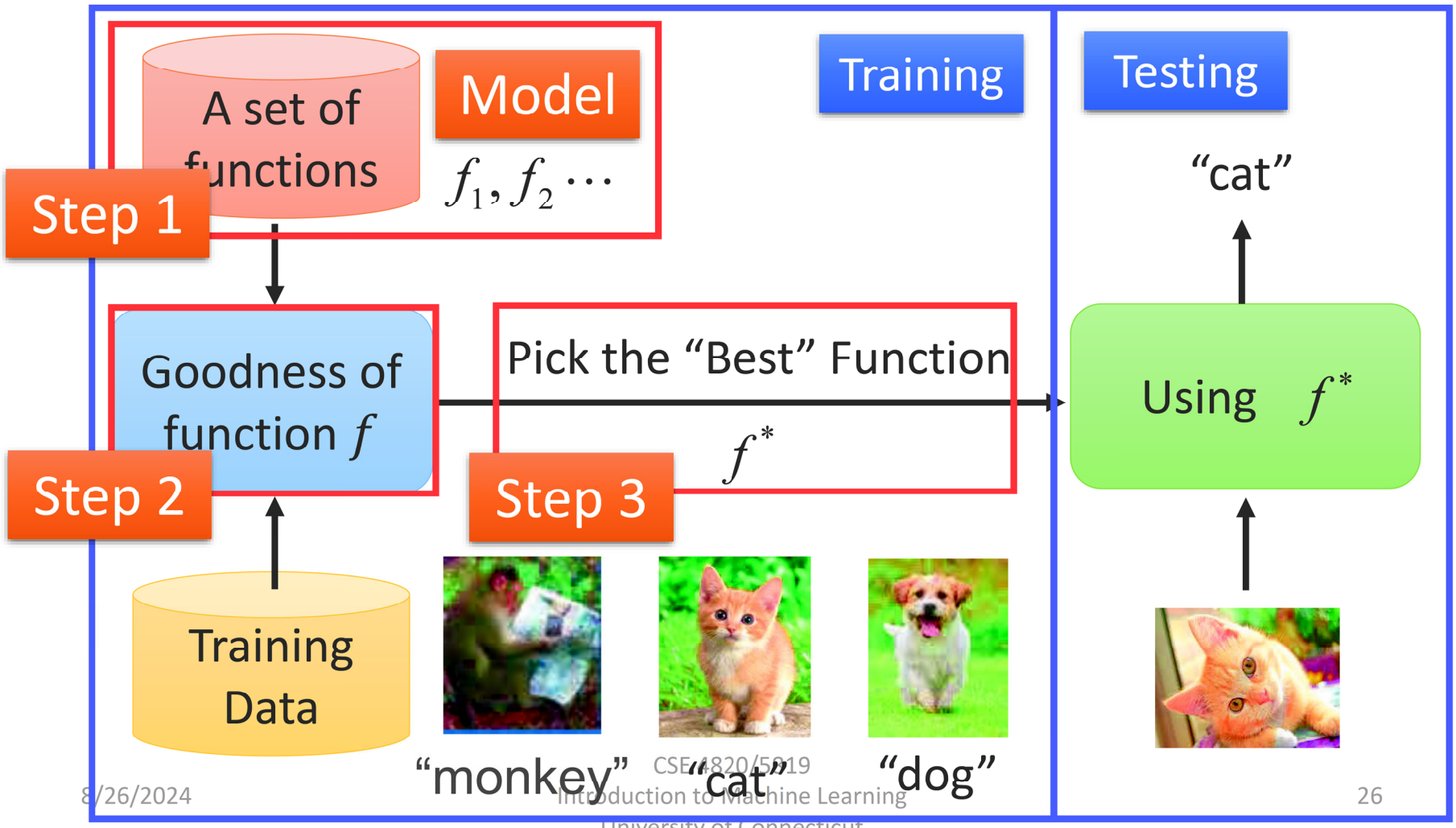
Help Humans Learn



Framework

Image Recognition:

$f($  $) =$ "cat"



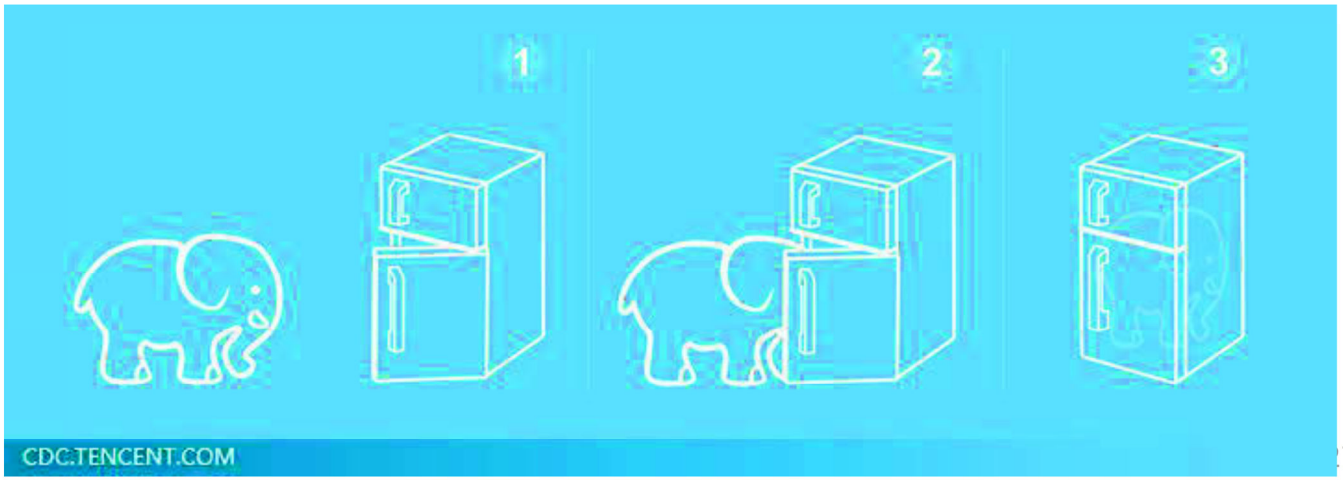
Process of ML ……

Different Tasks

Step 0: What kind of function do you want to find?

Step 1: define a set of functions → Step 2: goodness of function → Step 3: pick the best function

Putting An Elephant in the Fridge ……



Example of ML Models

- Learning associations
- Supervised learning:
  - Regression
  - Classification
- Unsupervised learning
  - Clustering
- Other more advanced/recent machine learning.
  - Reinforcement learning

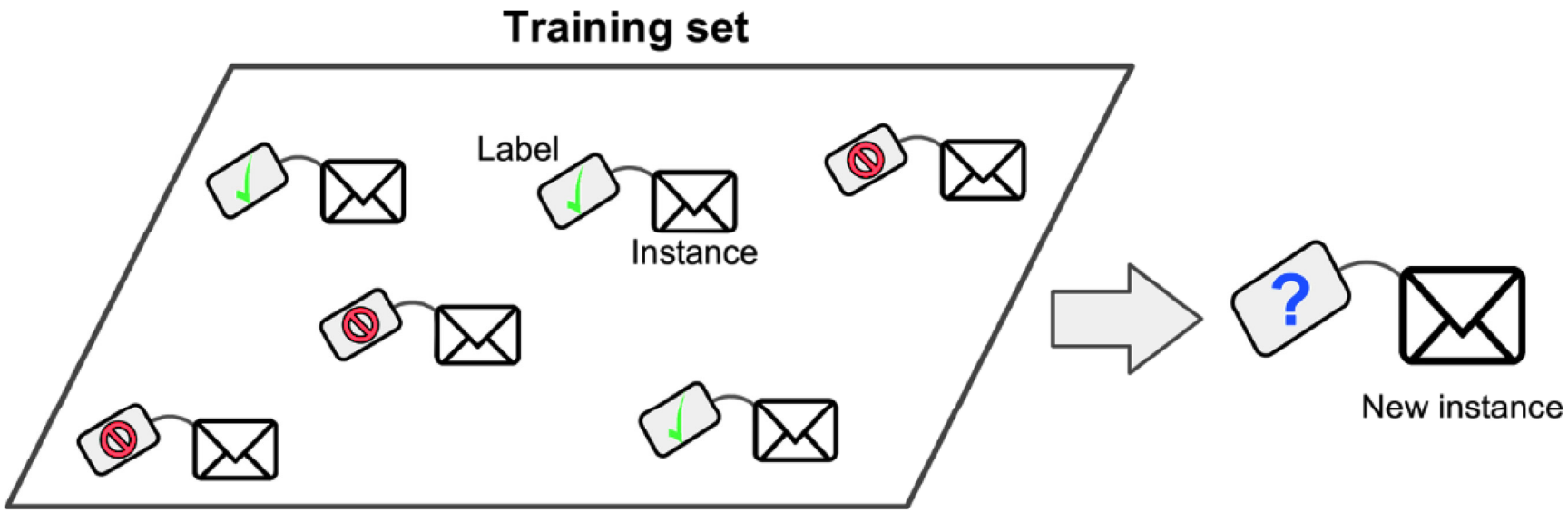
Types of Machine Learning

- It is useful to classify them in broad categories based on:
  - Whether or not they are trained with human supervision (supervised, unsupervised, semi-supervised, and reinforcement learning)
  - Whether or not they can learn incrementally on the fly (online versus batch learning)
  - Whether work by simply comparing new data points to known data points, or instead detect patterns in the training data and build a predictive model, much like scientists do (instance-based versus model-based learning)

Supervised Learning

- Starting from this lecture, we will look at different supervised learning methods
  - Model training process
- Regression
- Support Vector Machines
- Neural Networks

Supervised Learning



Regression

Regression

The output of the target function f is "scalar".

Predict PM2.5

Training Data:

Input:

9/01 PM2.5 = 63 9/02 PM2.5 = 65

Output:

9/03 PM2.5 = 100

Input:

9/12 PM2.5 = 30 9/13 PM2.5 = 25

Output:

9/14 PM2.5 = 20

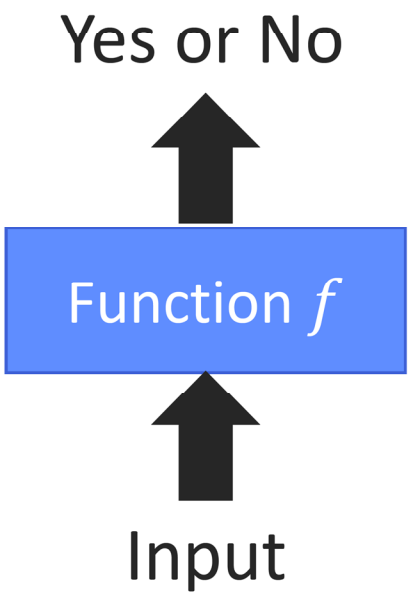
Regression

- Navigation of autonomous vehicle:
  - Angle of steering wheel
- Kinematics of robot arm:
  - Joint angles
- Recommender system:
  - Movie ratings
- Age estimation from facial image:
  - Age
- Chemical manufacturing process:
  - Yield
- Risk prediction from financial reports:
  - Risk

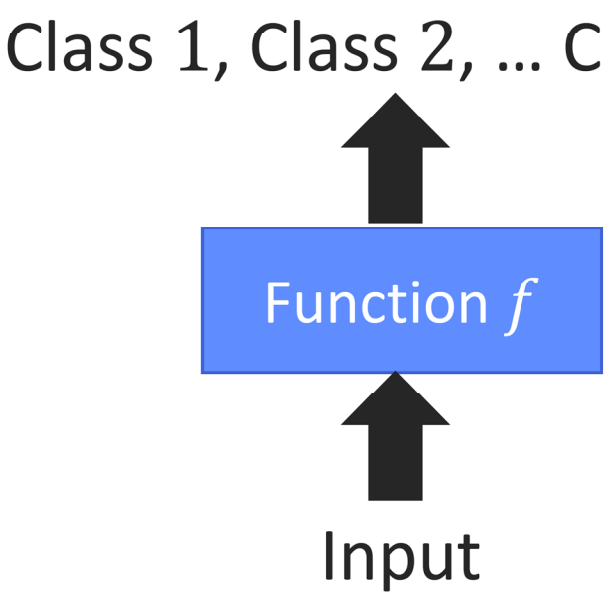
Regression

Classification

- Binary Classification

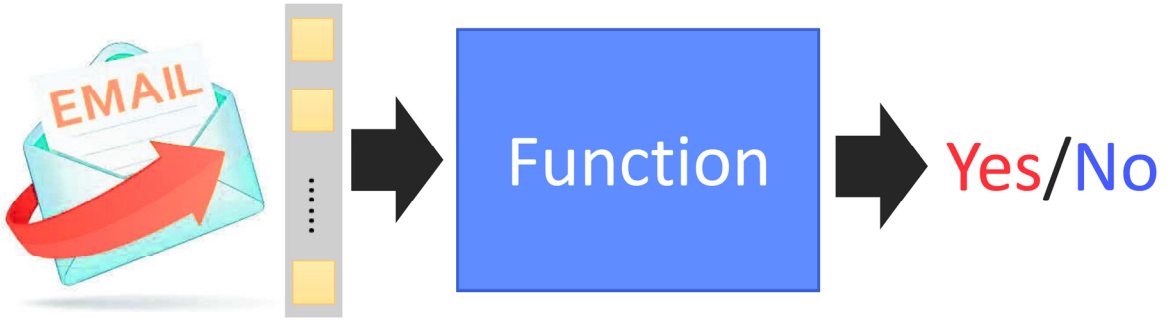


- Multi-class Classification



Binary Classification

Spam filtering

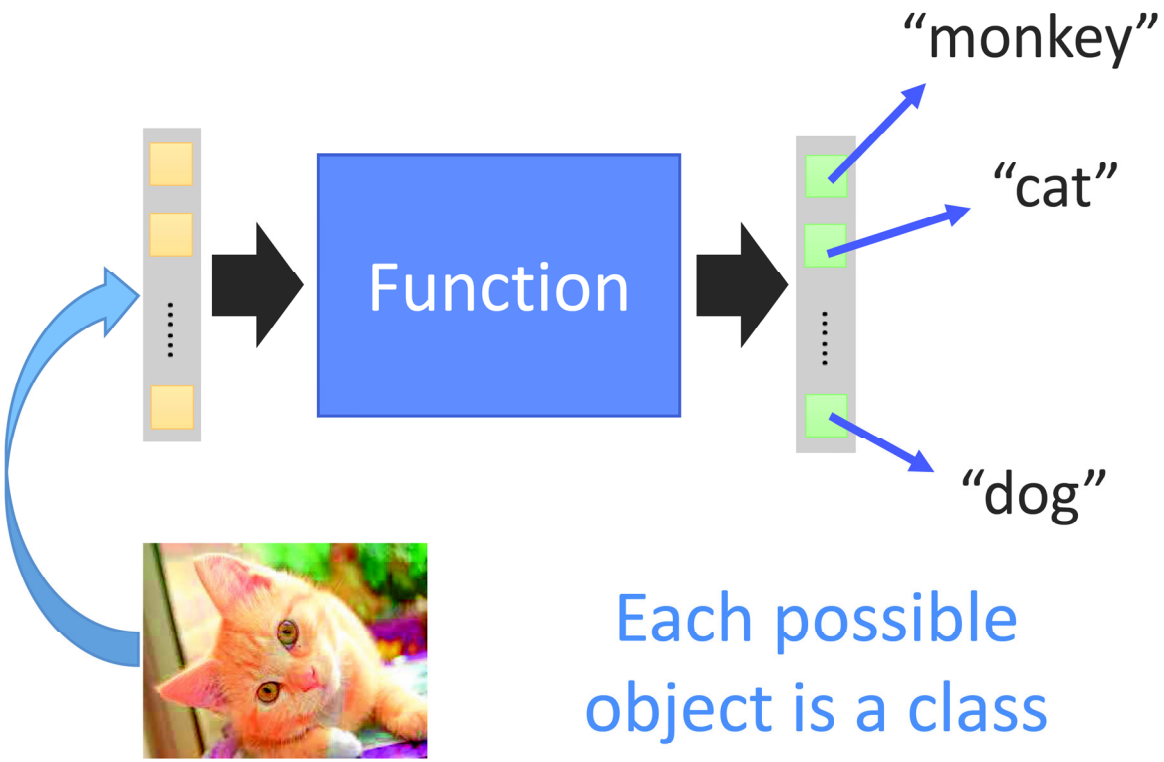


Training Data

(http://spam-filter-review.toptenreviews.com/)

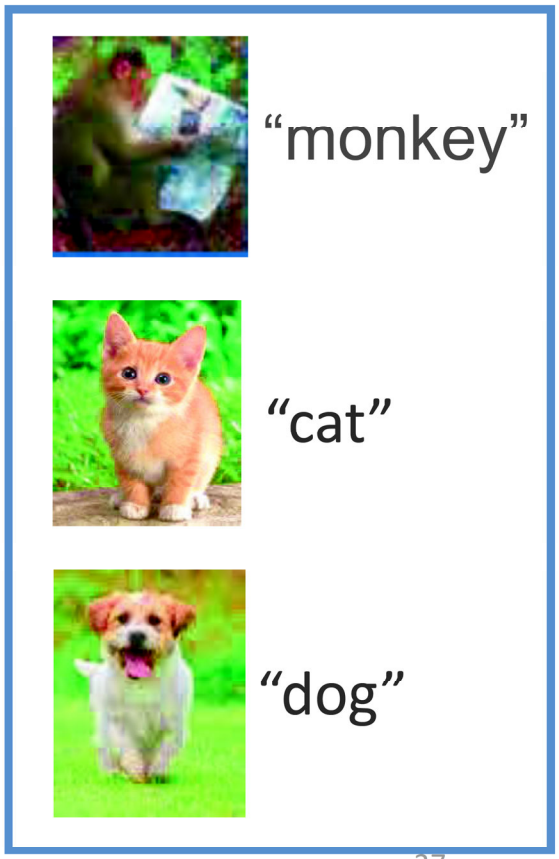
Multi-Class Classification

Image Recognition



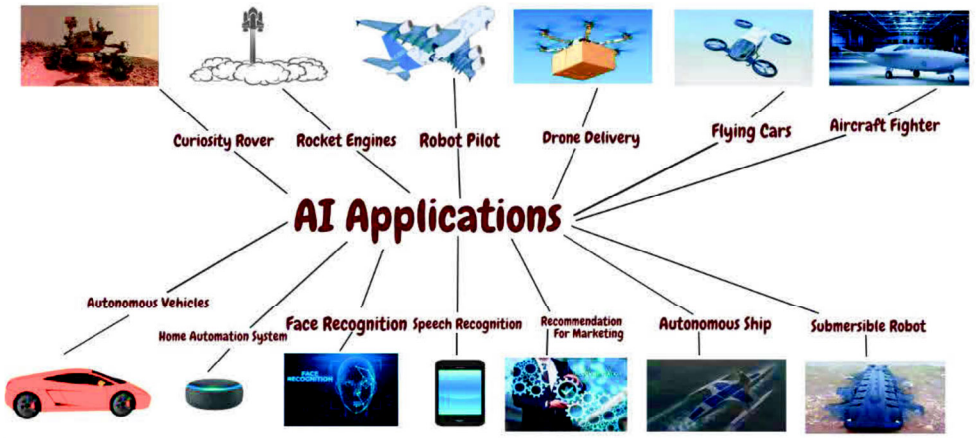
Each possible object is a class

Training Data



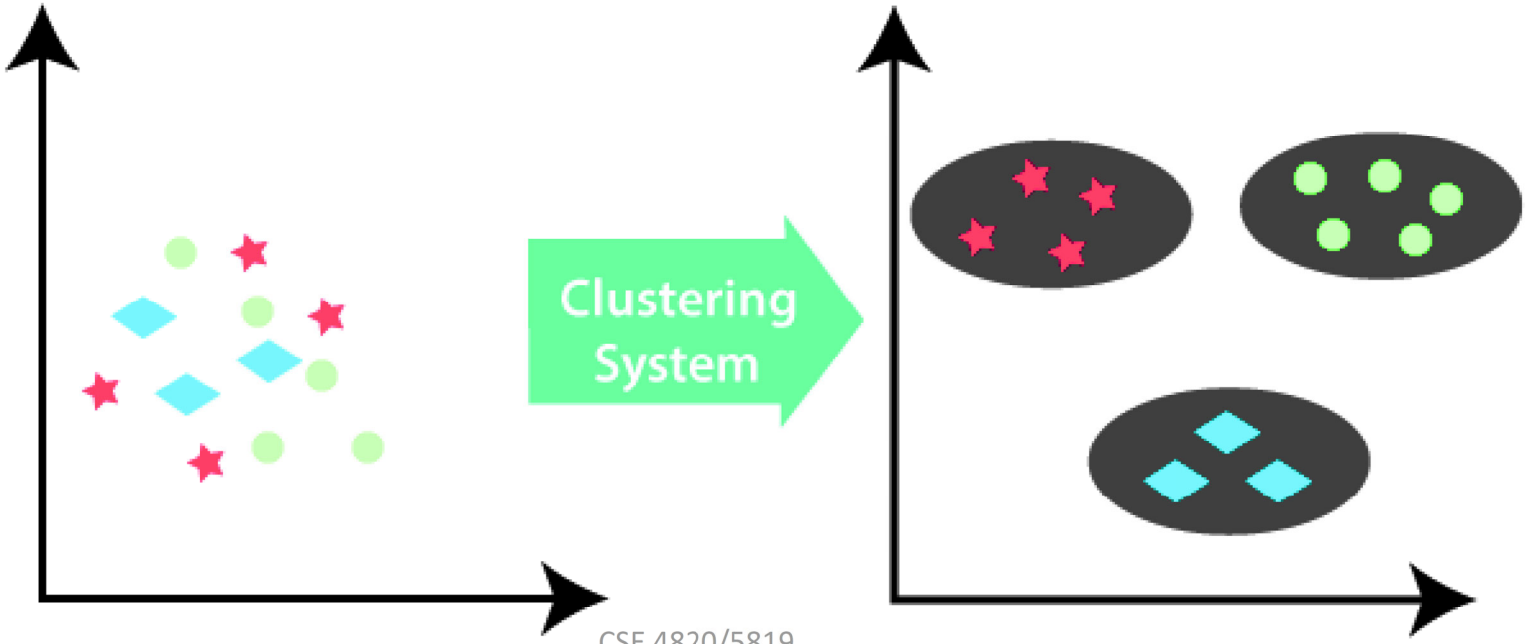
Classification

- Face recognition
- Character recognition
- Speech recognition
- Object recognition
- Biometric authentication
- Multi-touch gesture classification
- Document classification
- Spam detection
- Intrusion detection
- Terrorism detection
- Medical diagnosis
- Weather forecasting

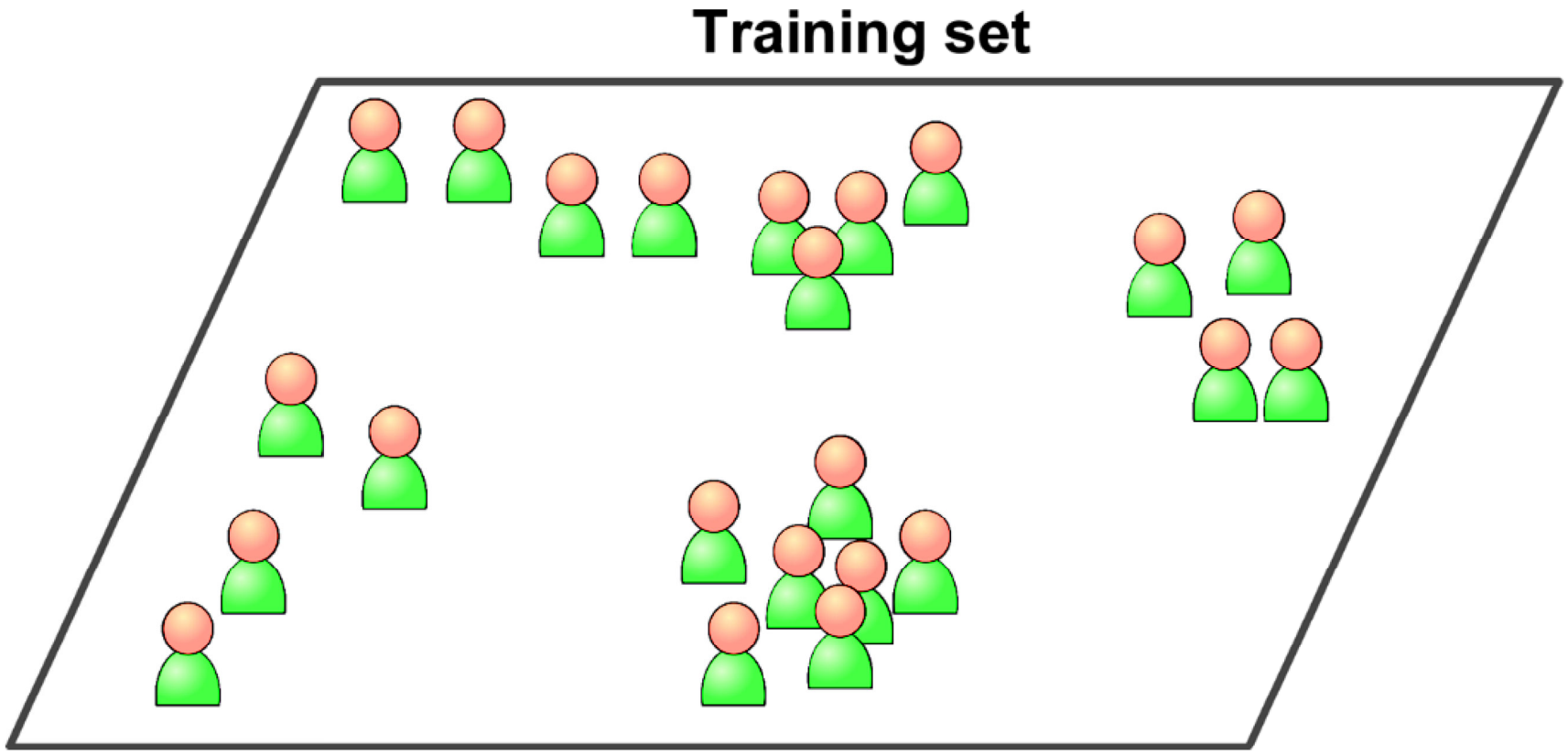


Unsupervised Learning

- Clustering
  - K-means
  - DBSCAN



Unsupervised Learning

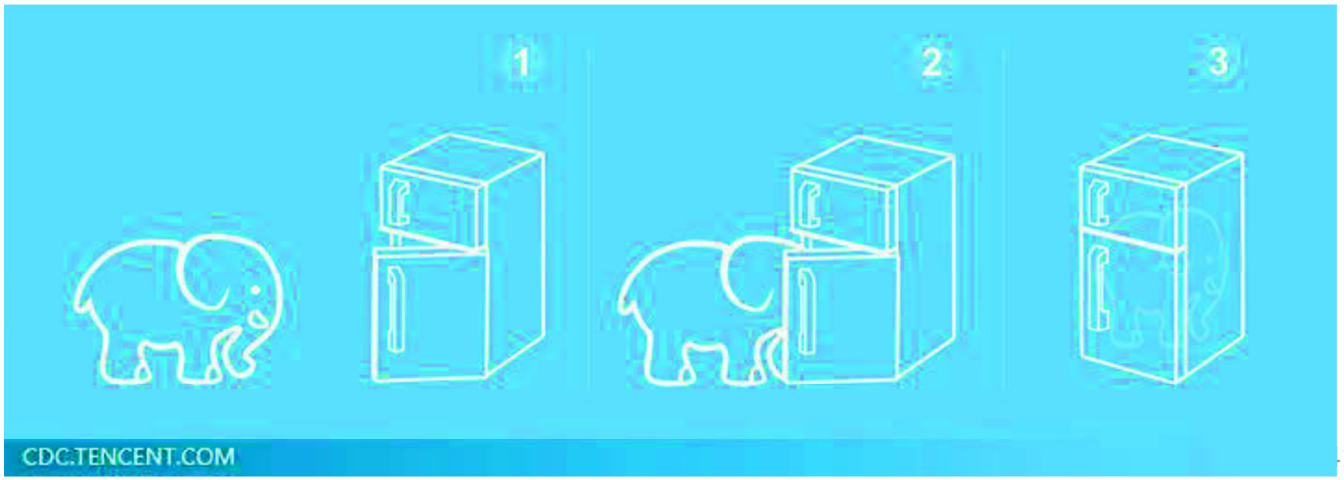


Process of ML ……

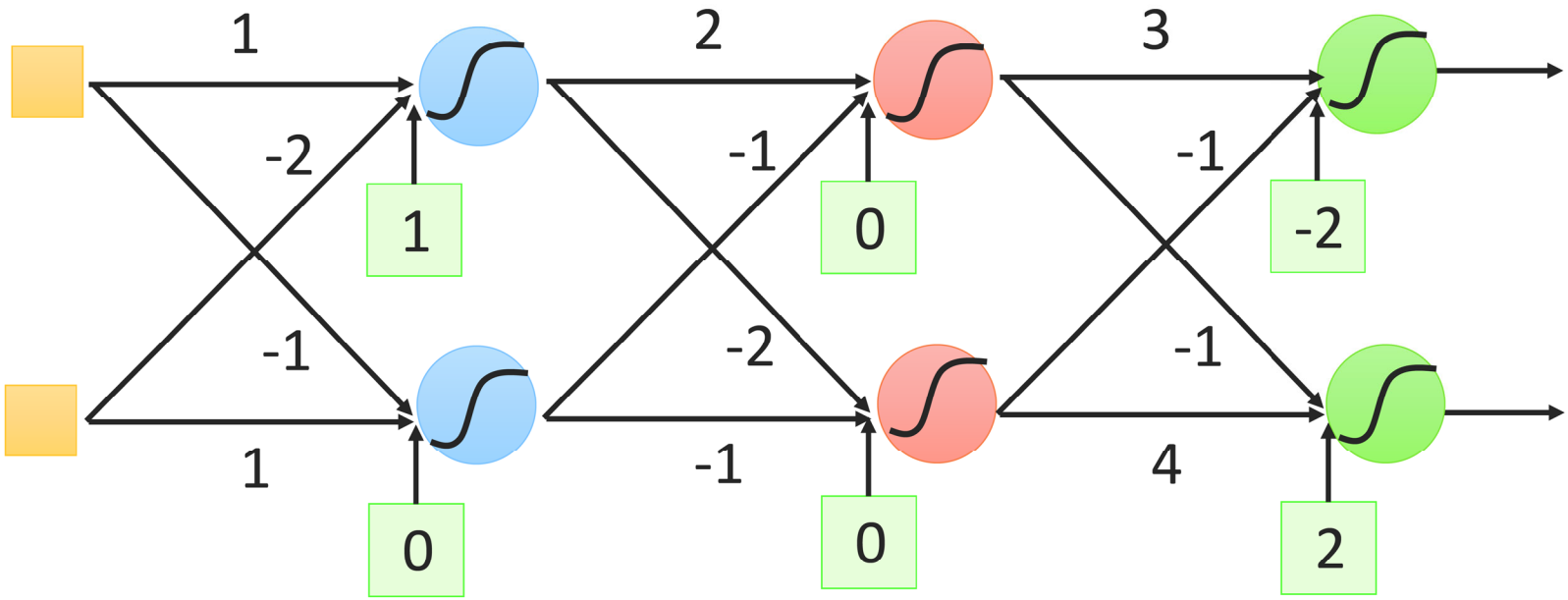
Step 0: What kind of function do you want to find?

Step 1: define a set of functions → Step 2: goodness of function → Step 3: pick the best function

Putting An Elephant in the Fridge ……



Neural Network

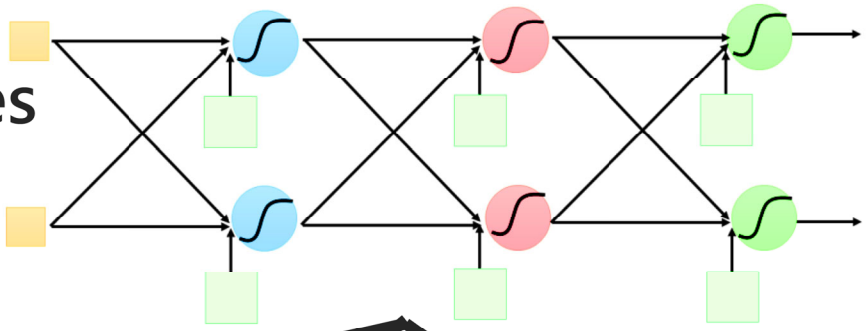


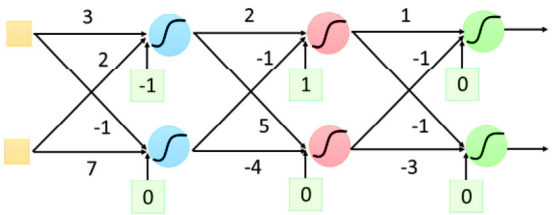
The developer only has to provide network structure.

The parameters are found automatically from data.

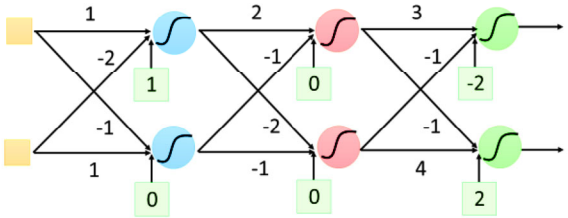
function set

Neural network structures

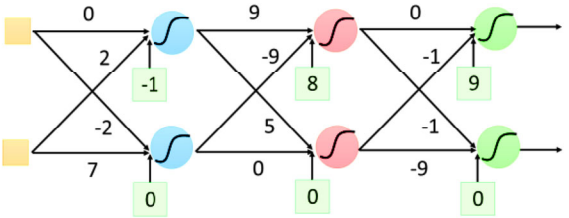




f_1



f_2



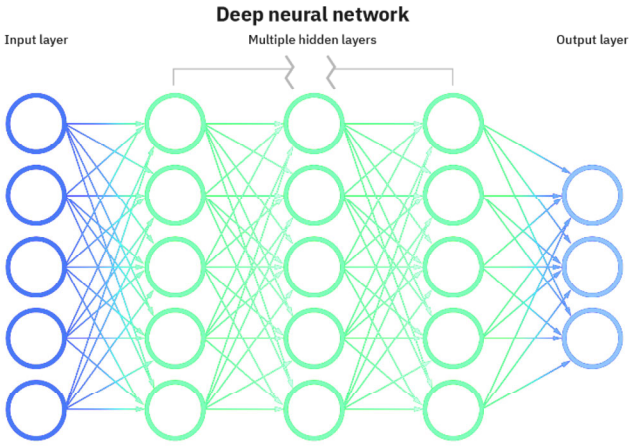
f_3

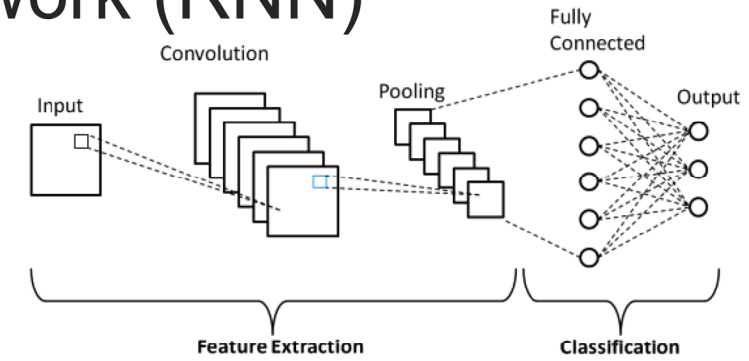
Neural network parameters learned from different tasks

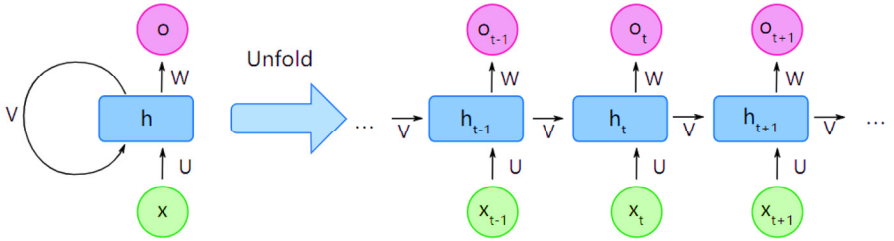
Machine learning!

Neural Network

- Introduction to neural network
- Convolutional neural network (CNN)
- Recurrent neural network (RNN)

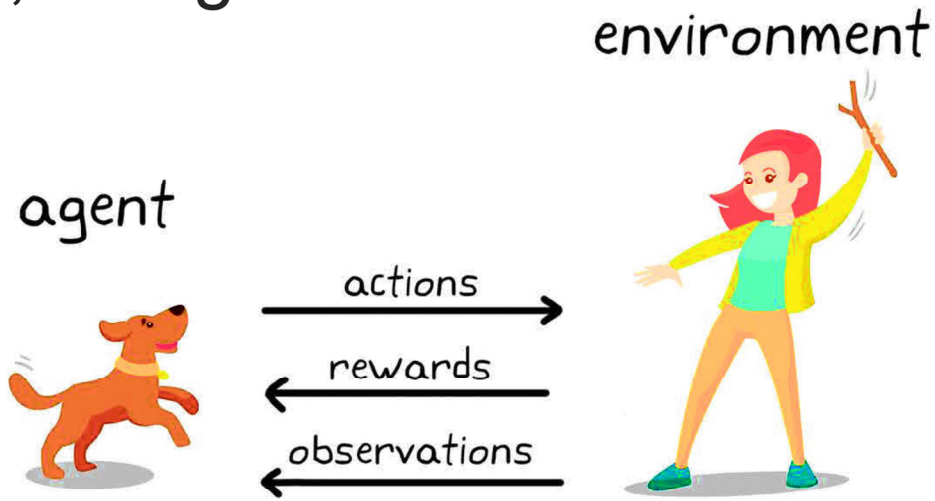






Reinforcement Learning

- Reinforcement Learning is a very different BEAST.
- The learning system, called an ***agent*** in this context, can observe the ***environment***, select and perform ***actions***, and get ***reward*** in return

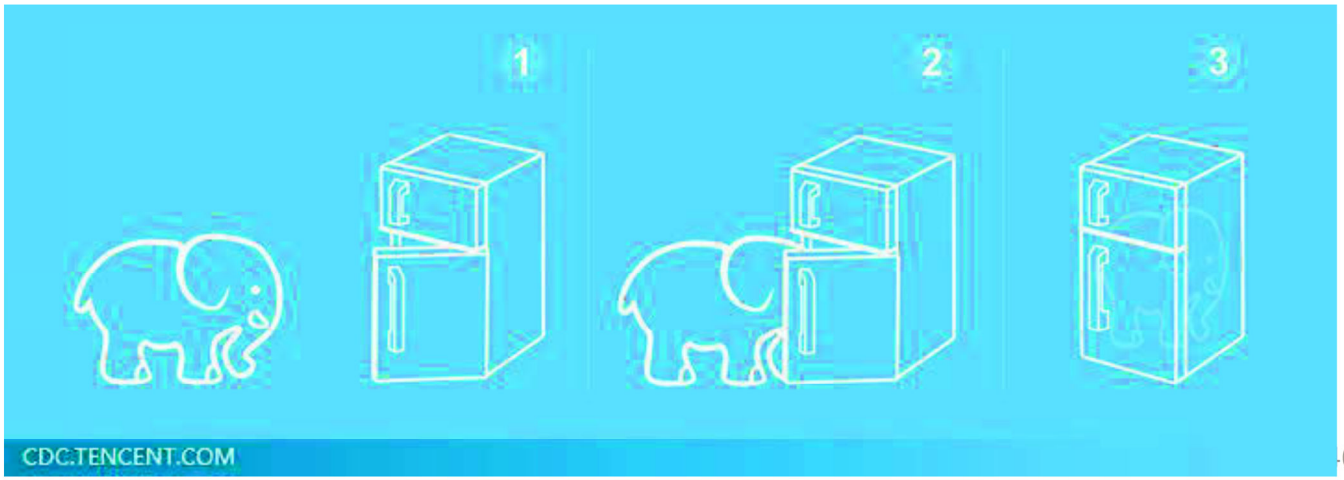


Process of ML ……

Step 0: What kind of function do you want to find?

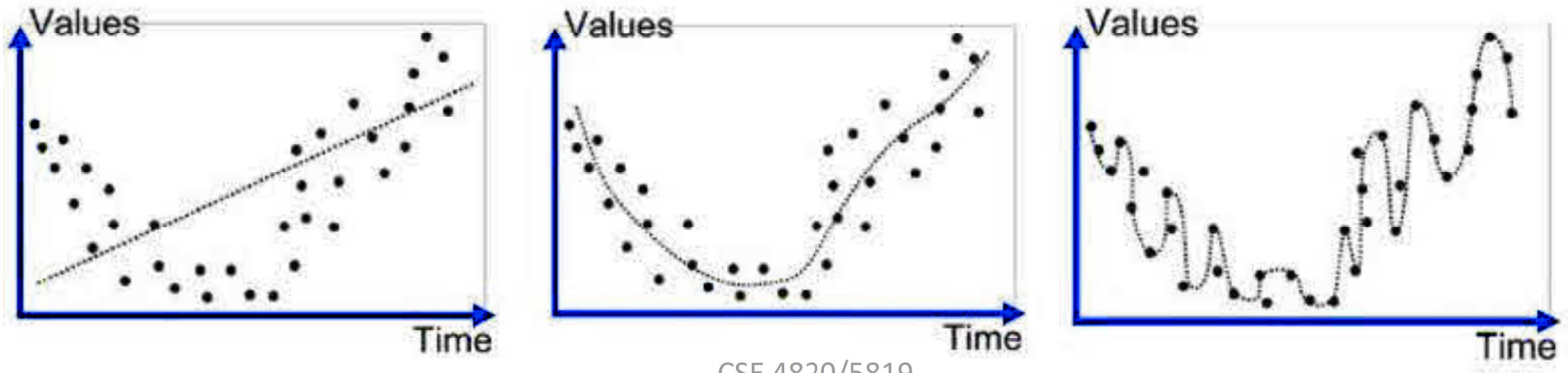
Step 1: define a set of function → Step 2: goodness of function → Step 3: pick the best function

Putting An Elephant in the Fridge ……



Goodness of Function

- Objective function
  - Minimizing the “difference” between the values of training data and the predicted ones
  - **Loss function**
- Determines the ways or performance of conducting machine learning

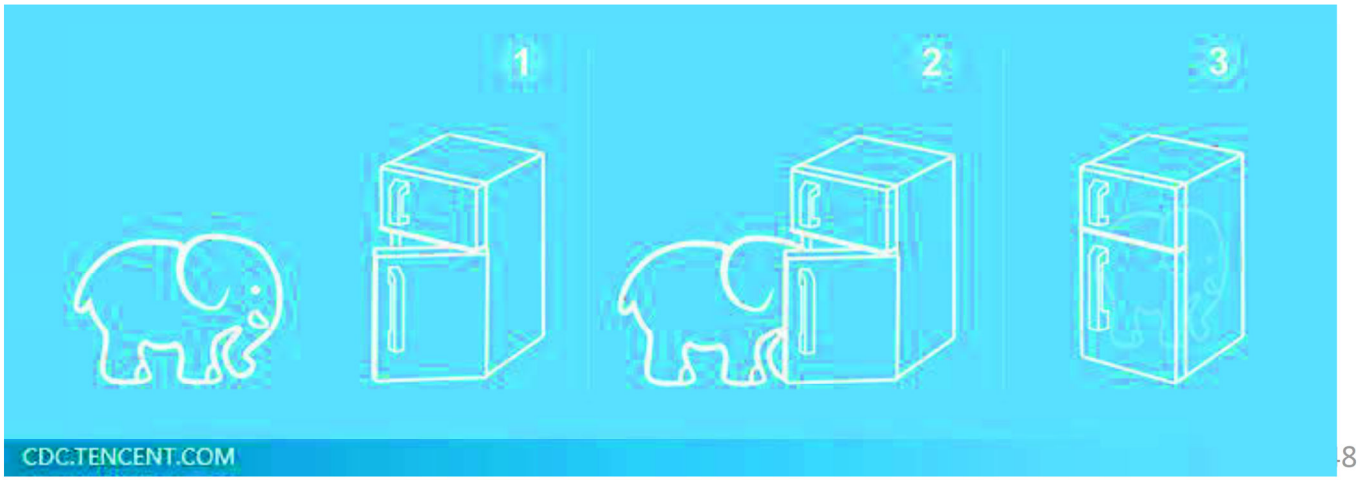


Process of ML ……

Step 0: What kind of function do you want to find?

Step 1: define a set of functions → Step 2: goodness of function → Step 3: pick the best function

Putting An Elephant in the Fridge ……



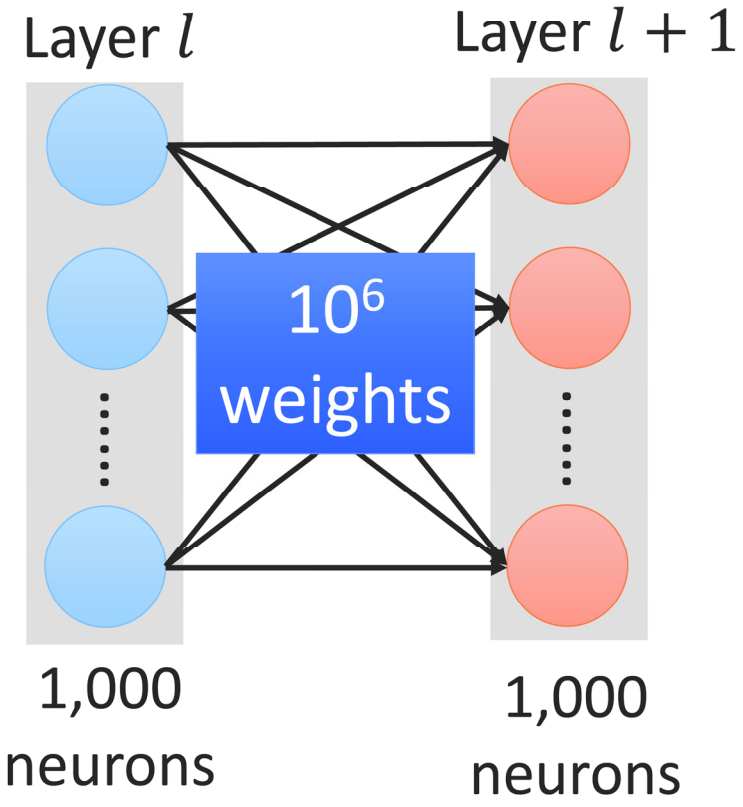
How to Pick the Best Function?

Enumerate all possible values

Network parameters $\theta = \{w_1, w_2, w_3, \cdots, b_1, b_2, b_3, \cdots\}$

Millions of parameters

Today a network can have more than 100M parameters.



Gradient Descent

PYTORCH

theano

Caffe

Microsoft
CNTK

Deep Learning library produced by Amazon
DSSTNE

Process of ML ……

Step 0: What kind of function do you want to find?

Regression, Classification, ……

Step 1: define a set of functions

Neural Network

……

Step 2: goodness of function

Supervised Learning

……

Step 3: pick the best function

Gradient Descent

……

Why we need to learn ML?

http://www.express.co.uk/news/science/651202/First-step-towards-The-Terminator-becoming-reality-AI-beats-champ-of-world-s-oldest-game

AI will replace some jobs? New Jobs： AI Trainers

AI Trainer

Machines train themselves.
Why do we need AI trainers??

Pokemon fight themselves.
Why do we need Pokemon trainers??

AI Trainers

Step 1: define a set of function → Step 2: goodness of function → Step 3: pick the best function

- Pokemon Trainers
  - Pick different pokemons for the fights
    - Different properties and types
- AI Trainers
  - Pick models, loss functions
    - Different models and loss function solves different problems

Defending type

Attacking type	NORMAL	FIGHT	FLYING	POISON	GROUND	ROCK	BUG	GHOST	STEEL	FIRE	WATER	GRASS	ELECTR	PSYCHC	ICE	DRAGON	DARK	FAIRY
NORMAL	1×	1×	1×	1×	1×	½×	1×	0×	½×	1×	1×	1×	1×	1×	1×	1×	1×	1×
FIGHT	2×	1×	½×	½×	1×	2×	½×	0×	2×	1×	1×	1×	1×	½×	2×	1×	2×	½×
FLYING	1×	2×	1×	1×	1×	½×	2×	1×	½×	1×	1×	2×	½×	1×	1×	1×	1×	1×
POISON	1×	1×	1×	½×	½×	½×	1×	½×	0×	1×	1×	2×	1×	1×	1×	1×	1×	2×
GROUND	1×	1×	0×	2×	1×	2×	½×	1×	2×	2×	1×	½×	2×	1×	1×	1×	1×	1×
ROCK	1×	½×	2×	1×	½×	1×	2×	1×	½×	2×	1×	1×	1×	1×	2×	1×	1×	1×
BUG	1×	½×	½×	½×	1×	1×	1×	½×	½×	½×	1×	2×	1×	2×	1×	1×	2×	½×
GHOST	0×	1×	1×	1×	1×	1×	1×	2×	1×	1×	1×	1×	1×	2×	1×	1×	½×	1×
STEEL	1×	1×	1×	1×	1×	2×	1×	1×	½×	½×	½×	1×	½×	1×	2×	1×	1×	2×
FIRE	1×	1×	1×	1×	1×	½×	2×	1×	2×	½×	½×	2×	1×	1×	2×	½×	1×	1×
WATER	1×	1×	1×	1×	2×	2×	1×	1×	1×	2×	½×	½×	1×	1×	1×	½×	1×	1×
GRASS	1×	1×	½×	½×	2×	2×	½×	1×	½×	½×	2×	½×	1×	1×	1×	½×	1×	1×
ELECTR	1×	1×	2×	1×	0×	1×	1×	1×	1×	1×	2×	½×	½×	1×	1×	½×	1×	1×
PSYCHC	1×	2×	1×	2×	1×	1×	1×	1×	½×	1×	1×	1×	1×	½×	1×	1×	0×	1×
ICE	1×	1×	2×	1×	2×	1×	1×	1×	½×	½×	½×	2×	1×	1×	½×	2×	1×	1×
DRAGON	1×	1×	1×	1×	1×	1×	1×	1×	½×	1×	1×	1×	1×	1×	1×	2×	1×	0×
DARK	1×	½×	1×	1×	1×	1×	1×	2×	1×	1×	1×	1×	1×	2×	1×	1×	½×	½×
FAIRY	1×	2×	1×	½×	1×	1×	1×	1×	½×	½×	1×	1×	1×	1×	1×	2×	2×	1×

These matchups are suitable for Generation VI onward.

AI Trainers

Step 1: define a set of function → Step 2: goodness of function → Step 3: pick the best function

- Pokemon Trainers
  - Pick different pokemons for the fights
    - Different properties and types
  - Sometimes the pokemons might not work well with the trainers
    - Need experienced trainers!

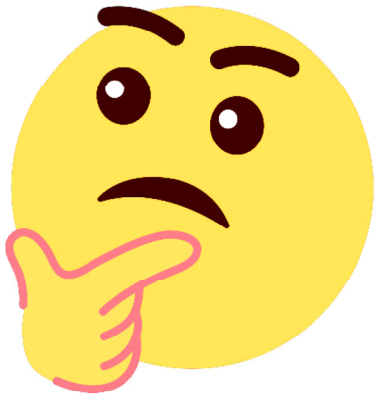
- AI Trainers
  - Pick models, loss functions
    - Different models and loss function solves different problems
  - Best function is hard to find
    - e.g., Deep Learning
    - Need experienced trainer!

AI Trainers

- Good trainers, good fighters
- Good trainers, good AIs

Last but not the Least …

- You come here not to learn machine learning…
  - You come here to learn how to learn machine learning
  - You come here to learn how to tame machine learning
  - You come here to learn how to avoid (some) machine learning



Assignment 0

- Please prepare a short description (~100 words) about yourself:
  1) List any courses (inside/outside UConn) you have taken before related to data, machine learning and deep learning;
  2) List any programming language you are comfortable/familiar with;
  3) Any research interest or background?

Next Lecture

- Warm-Up
- Basic linear algebra and statistics
- Optimization