15CSE401 MLDM

Review of Machine Learning Concepts

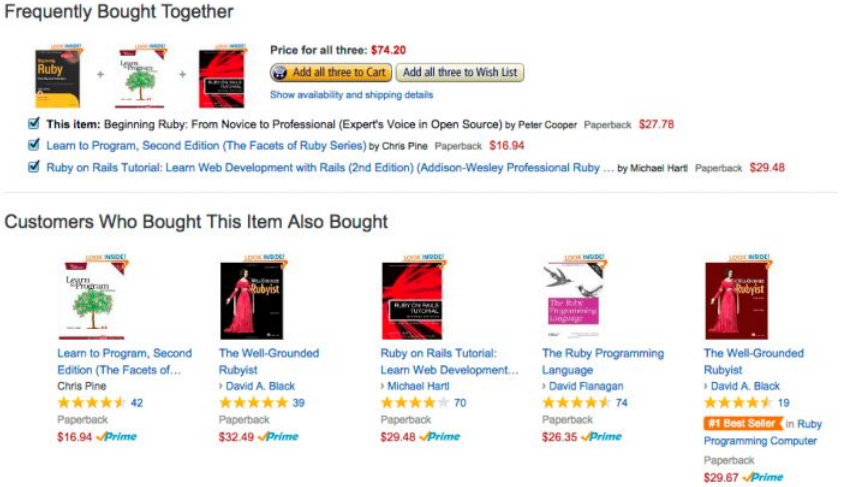
Lecture 3
21 Jul 2020

Today’s plan

- Introduction to Machine Learning (What and Why)
- Toy examples
- Real life applications
- Types of Machine Learning

What is common in all these?

What is common in all these?



- **Machine learning** replicate activities requiring
- human cognition

PC: Google Images

More examples...

NETFLIX

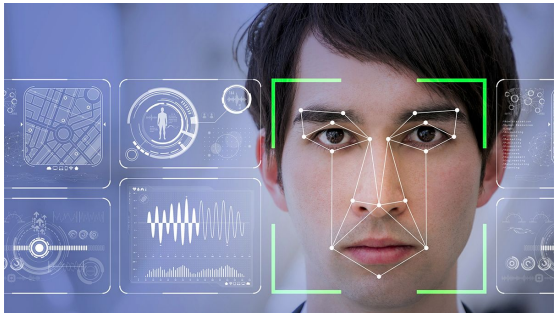
YouTube

facebook

Google

Google Search I'm Feeling Lucky

We stand in support of racial equality, and all those who search for it



PC: Google Images

Machine Learning

- Learning by Machines
- Machines -----> Computers
- Learning -----> from history

Toy example -1

-

X	Y
2	4
3	9
4	?
X	?

Toy example-1

- Y =f(x)
- What is f(x)?

X	Y
2	4
3	9
4	?
X	?

Toy example-2

- Y =f(x)
- What is f(x)?

X	Y
2	3.888
3	9.133
4	15.55
X	?

Toy example-2

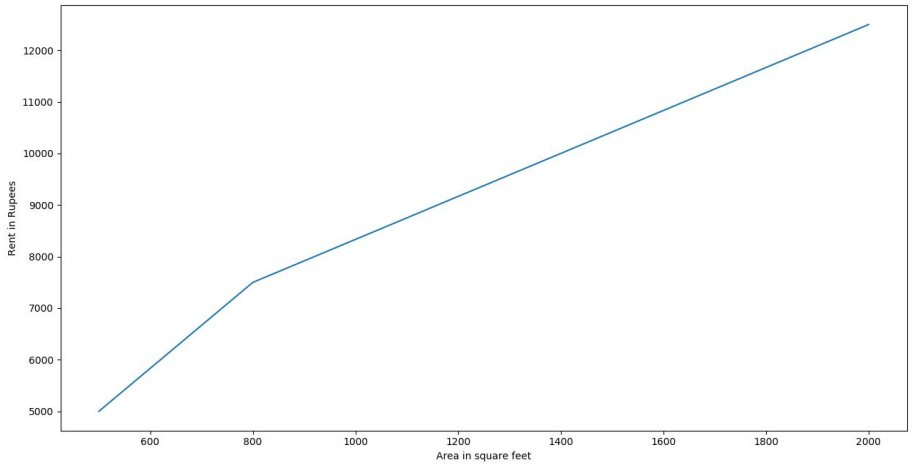
- f(x) is not explicit here
- Trial & error with many functions:
- f1(x), f2(x)..... fn(x)
- Error = Y – f(x)
- Find f(x) that minimize this error over all
- X

X	Y
2	3.888
3	9.133
4	15.55
X	?

House Rent Prediction

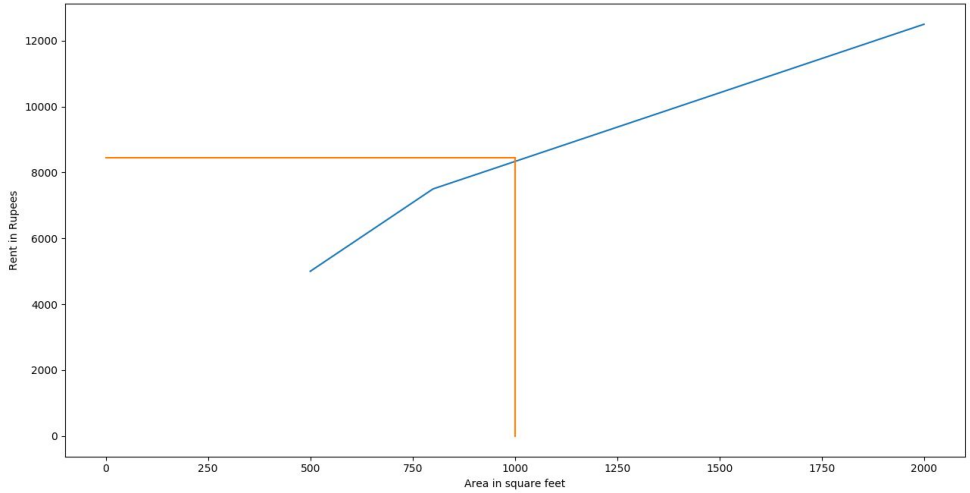
Area (Sq ft)	Rent (in Rs)
500	5000
2000	12500
800	7500
1000	?

House Rent Prediction

- 

Area (Sq ft)	Rent (in Rs)
500	5000
2000	12500
800	7500
1000	?

House Rent Prediction

- 

Area (Sq ft)	Rent (in Rs)
500	5000
2000	12500
800	7500
1000	?

Real Life Examples

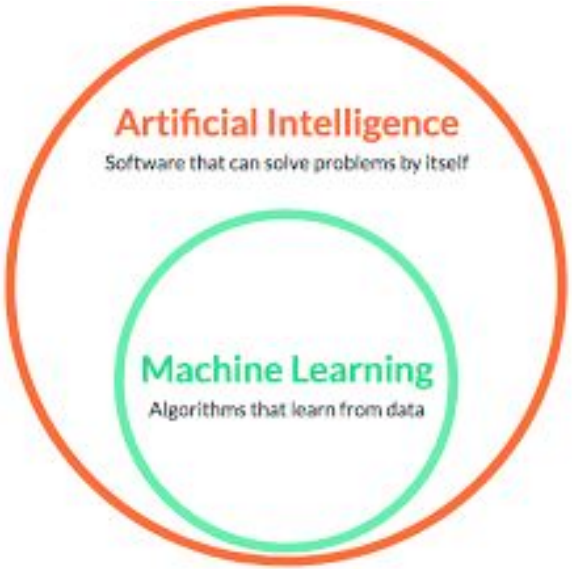
- Natural Language Processing
- Speech Recognition
- Machine Translation
- Disease prediction
- Computational drug discovery
- .....

What is Machine Learning

- “ML gives computers the ability to “**learn”** without being explicitly programmed”.[Arthur Samuel,1959]
- What is **Learning**?
- “A computer program is said to **learn** from **experience E** with respect to some **task T** and some **performance measure P**, if its performance on T, as measured by P, improves with experience E [Mitchel, 1997].
- From where get the experience

  Previous history (data)

AI and ML



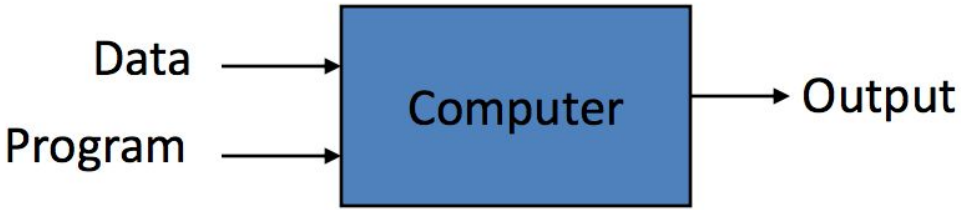
AI is the study of how to train the computers so that computers can do things which at present human can do better

ML is an application of AI that provide system the ability to automatically learn and improve from experience.

Source: https://www.geeksforgeeks.org/difference-between-machine-learning-and-artificial-intelligence/

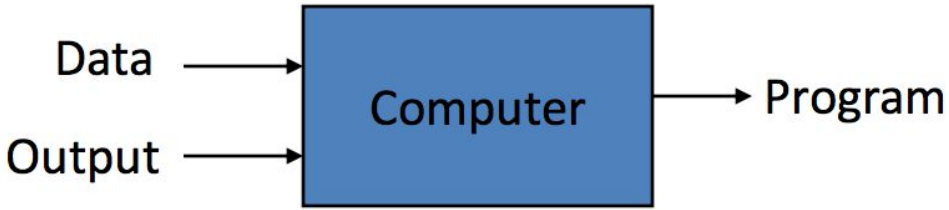
Traditional *vs* ML Programs

Traditional Programming



- Data -> Information
- Definitive cases
- Matrix multiplication

Machine Learning



- Data -> Knowledge
- Deals with uncertinity
- Next word prediction

Source: https://futurice.com/blog/differences-between-machine-learning-and-software-engineering

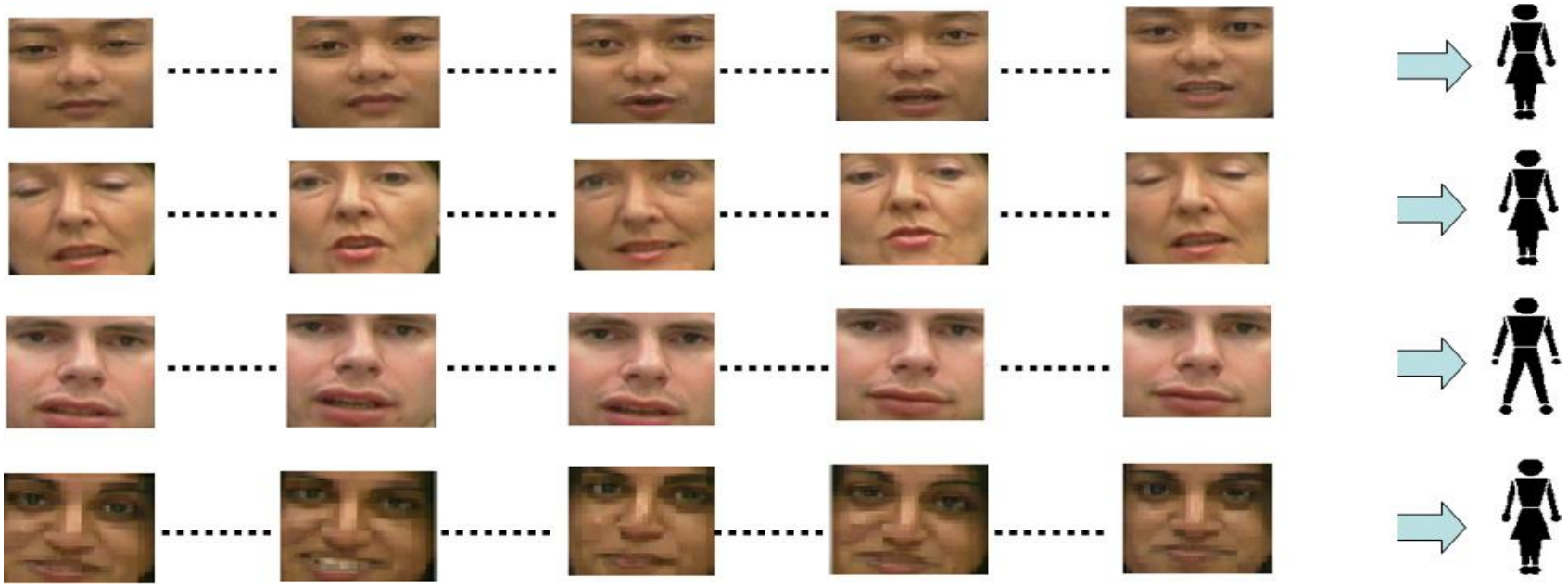
Machine Learning- Why?

- “We are drowning in information and starving for knowledge “— John Naisbitt
- Era of “**Big Data”**
- **Automated methods of data analysis** is a need of the hour
- **Automatically detect patterns** in data, and then use the uncovered patterns to **predict future data**, or to perform other kinds of decision making **under uncertainty**

Types of ML algorithms

- Supervised
- Unsupervised
- Semi-supervised
- Reinforcement

Supervised Learning



..... ???

Supervised Learning

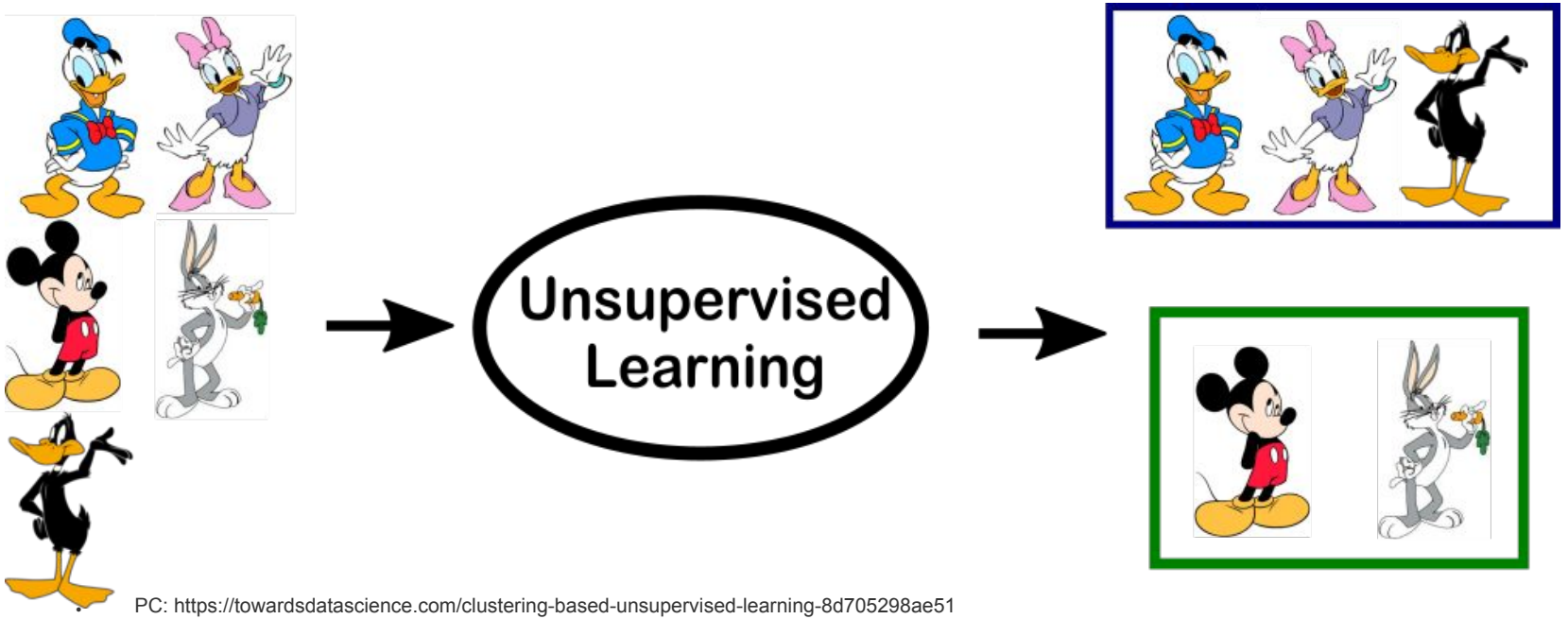
- When an algorithm learns from example data and associated target responses in order to later predict the correct response when posed with new examples

- **Classification**
  - Mapping input x to output y (where y is a class label)
  - **Binary** or **Multi-class classification**
  - **Prediction** or **Generalisation (**of unseen data)
  - P(y|X, D), X: feature vector, y: output class, D: dataset/history
- **Regression**
  - Response variable is continuous

Supervised Learning

- Automatic Spam detection
- Loan Approval
- Cancer subtype prediction
- Handwritten digit/character prediction
- Stock market prediction
- Predict the age of a viewer watching a given video on YouTube

Unsupervised Learning



PC: https://towardsdatascience.com/clustering-based-unsupervised-learning-8d705298ae51

Unsupervised Learning

- Learn from the data - no labels
- Discover **“interesting structure”** in the data - **knowledge discovery**
- Does not require a human expert to label data
- Applications
  - Discovering **clusters**
  - Discovering **latent** factors
    - Topic model / Theme/ Essence of the data
    - Dimensionality reduction
  - Discovering graph structures
  - Matrix completion - image imputation
  - Association mining

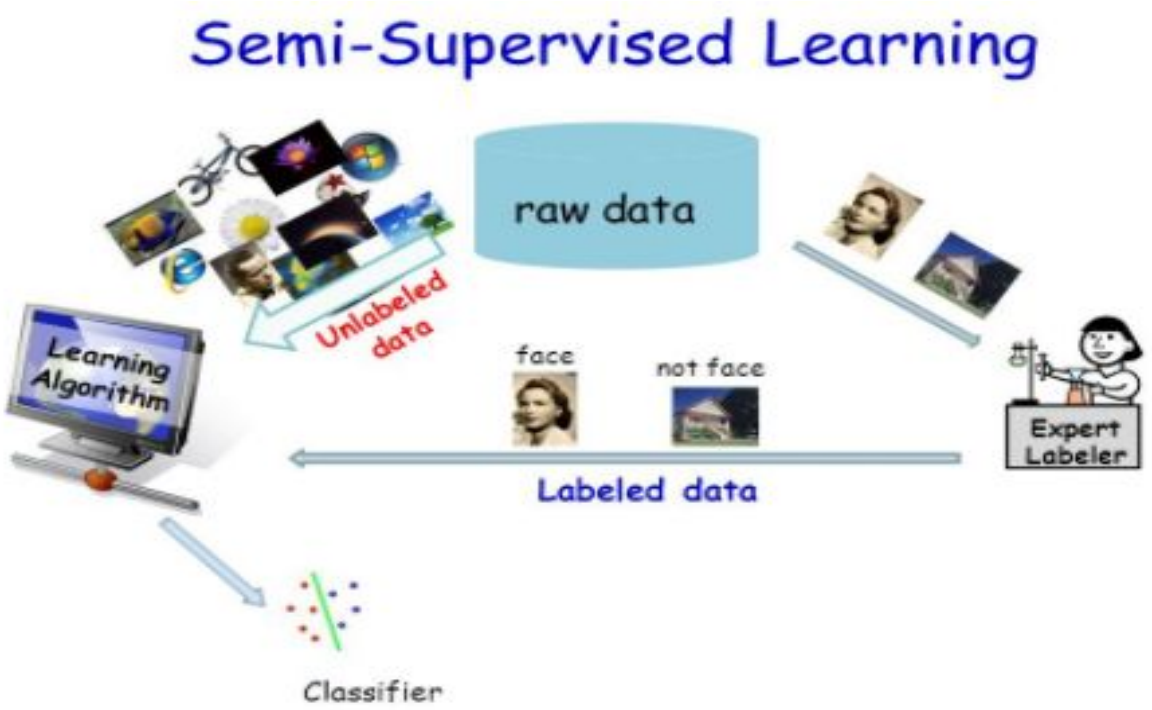
Unsupervised Learning

- Target Group identification
- News suggestion/Recommendation
- .....
- Algorithm **learns from plain examples without any associated response** (no supervision)
- **Clustering/Descriptive analysis**

Semi-supervised Learning

- Supervised learning
  - Hand labelled data
  - Costly
- Unsupervised learning
  - Limited application spectrum
- Semi-supervised Learning
  - Trained upon a combination of labeled and unlabeled data
  - First, cluster similar data using an unsupervised learning algorithm
  - Then use the existing labeled data to label the rest of the unlabeled data

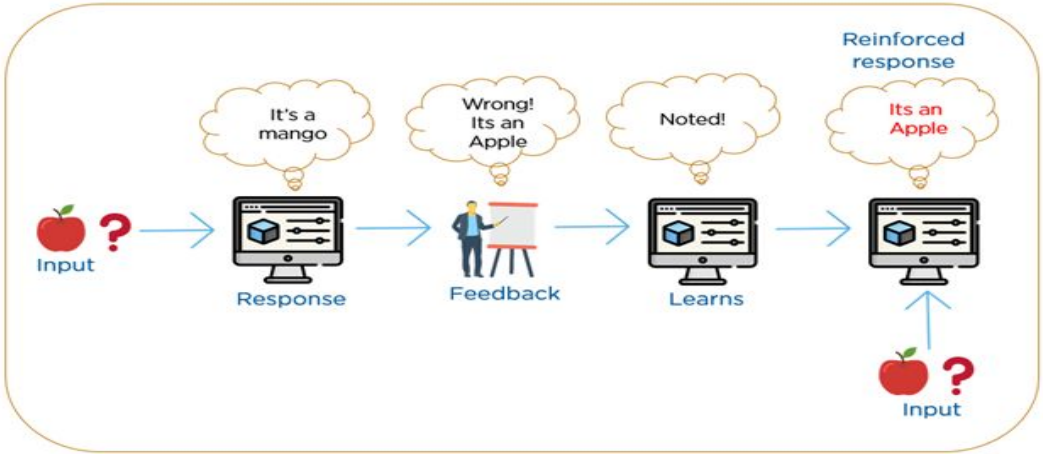
Semi-supervised Learning



- Speech Recognition
- Internet content classification
- Bioinformatics applications

Reinforcement Learning

- Accompany an example with **positive or negative feedback** according to the solution the algorithm proposes
- **Learning by trial and error**
- **“how to act or behave when given occasional reward or punishment signals”**



Machine Learning- Basic concepts

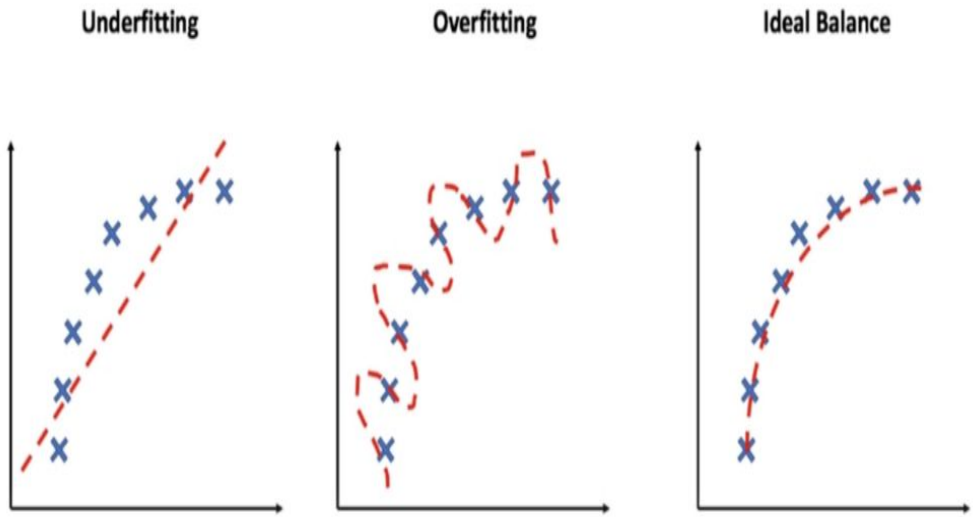
Parametric vs Non-parametric models

- **Parametric models**
  - Summarizes data with a set of **parameters of fixed size** (independent of the number of training examples) is called a **parametric model**.
  - Assume fixed form for mapping function
  - Eg: Naive Bayes, Simple Neural Network, Logistic regression etc

- **Non-parametric models**
  - **Non-parametric methods** are good when you have a lot of data and no prior knowledge
  - Do not make strong assumptions about the form of the mapping function
  - Eg: KNN, SVM, Decision Tree

Overfitting

- **Modeling error** that occurs when a function is too closely fit to a limited set of data points.
- Happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data



Model Selection

- How should we pick the right one?
- Compute the misclassification (error) rate of each method, select with minimum error
  - Training error
  - Generalization error (test error)
  - Validation error

- Cross Validation
  - K-Fold CV
  - Leave One Out Cross Validation (LOOCV)

Next

- Core concepts in Data Mining
  - Measuring similarity and dissimilarity

Reading Assignment

- **Chapter 1 and 2**
- Kevin P. Murphey, “Machine Learning, a probabilistic perspective”, The MITPress, 2012.

Thank you