LOVELY PROFESSIONAL UNIVERSITY

INT354
Machine Learning – I

Lecture #0

Agenda, Roadmap & Expectations

Course details

- **LTP – 2 0 2**
- **Credits – 3.0**
- **Mode – BYOD**
- **Text Book**
  - **Python Machine Learning ~ Sebastian Raschka**
- **Reference Books**
  - **MACHINE LEARNING ~TOM M. MITCHELL**
  - **UNDERSTANDING-MACHINELEARNING-THEORYALGORITHMS FROM THEORY TO ALGORITHM ~ SHAI SHALEVSHAWARTZ AND SHAI BEN-DAVID**

Course Outcomes

- CO1- Explore Different types of Machine Learning and statistics used for risk minimization.
- CO2- Analyze the operations of different types of Machine Learning Classifiers.
- CO3- Examine the performance of Generative models based on Bayesian learning to solve different classification problems.
- CO4- Develop the model that predict value of continuous variable with regression analysis
- CO5- Discuss the methods for Error calculations using different Regression metrics.
- CO6- Extend the Machine Learning approach to understand the bias complexity tradeoff and algorithm independent machine learning.

PROGRAM OUTCOMES

PO1- Engineering knowledge::Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

PO2- Problem analysis::Identify, formulate, research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

PO3- Design/development of solutions::Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

PO4- Conduct investigations of complex problems::Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO5- Modern tool usage::Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

PO6- The engineer and society::Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice

PO7- Environment and sustainability::Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO8- Ethics::Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

PO9- Individual and team work::Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO10- Communication::Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO11- Project management and finance::Demonstrate knowledge and understanding of the engineering, management principles and apply the same to one's own work, as a member or a leader in a team, manage projects efficiently in respective disciplines and multidisciplinary environments after consideration of economic and financial factors.

PO 12- Life-long learning::Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PO13- Competitive Skills::Ability to compete in national and international technical events and building the competitive spirit alongwith having a good digital footprint.

MAPPING OF CO'S WITH PO'S

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO9	PO12	PO13
CO1		YES							YES	
CO2	YES	YES	YES	YES	YES				YES	YES
CO3	YES	YES	YES	YES	YES				YES	YES
CO4	YES	YES	YES	YES	YES				YES	YES
CO5		YES							YES	
CO6	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

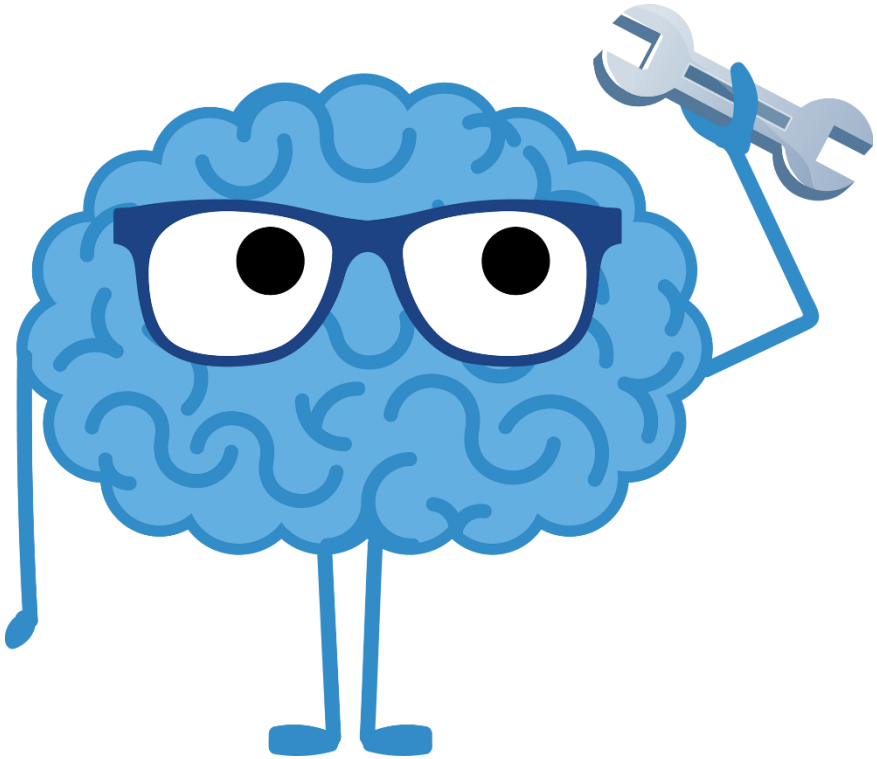
Exam and CA Details

- CA category: C020102

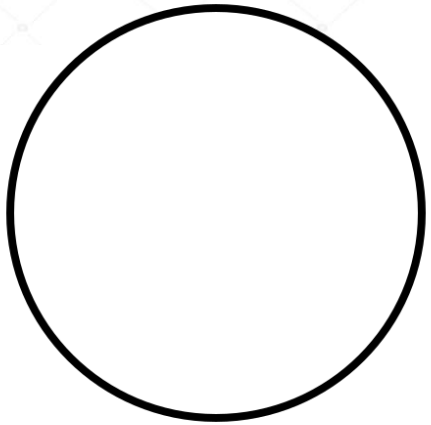
Component	Mandatory	Weightage(%)	Marks Distribution(30 marks)	Allocation week/ submission week
CBT1		30	Questions will be multiple of 5 marks or 10 marks	5/6
CBT2		30	MCQ based(1 mark each question)	12/13
PROJECT	YES	35	10 marks project execution, 10 marks presentation and viva, 10 marks report	3/12
HACKATHON	YES	35	Accuracy and efficiency 30 marks	9/10

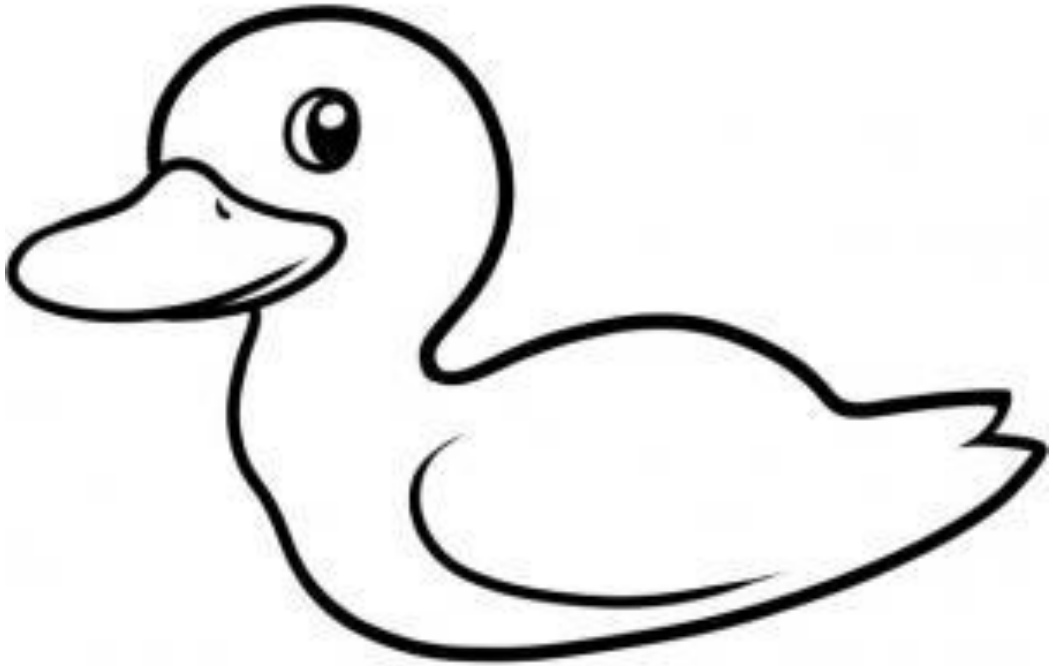
- No MTE
- ETE(25% MCQ, 75% theory)

Machine Learning



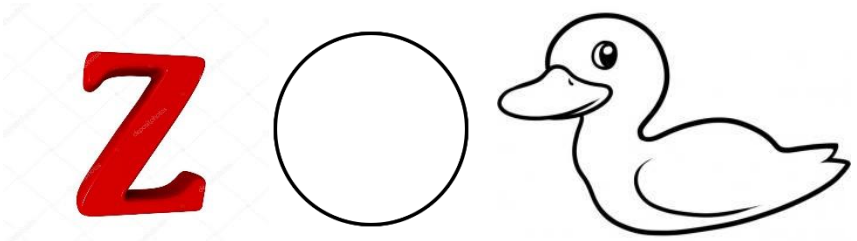
Learning – Aspect #1





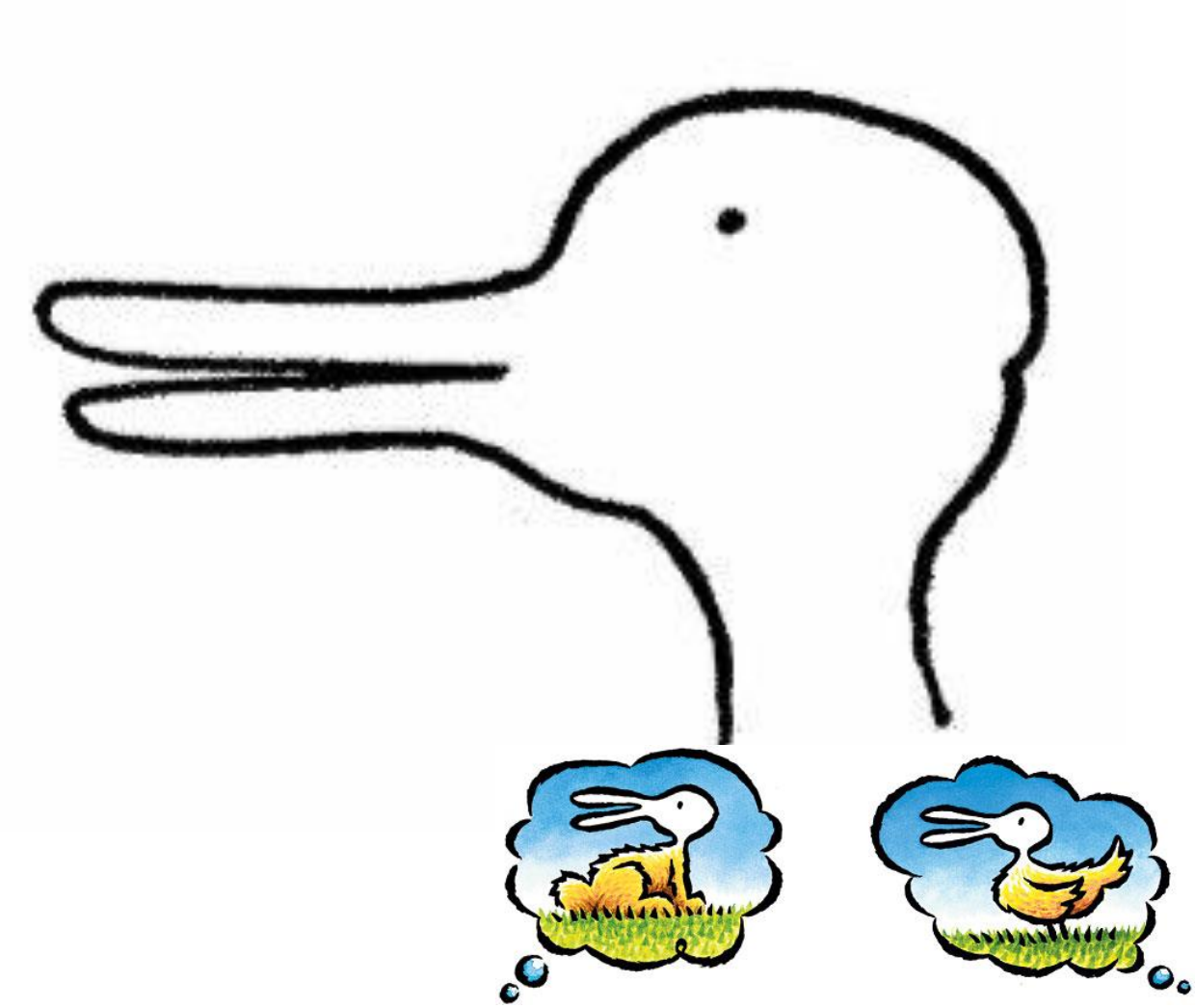
Discrete/Labeled dataset

- **You used prior learnings of -**
  - **Shapes**
  - **Patterns**
  - **Categories**
  - **Weighed them to MOST nearing entity**

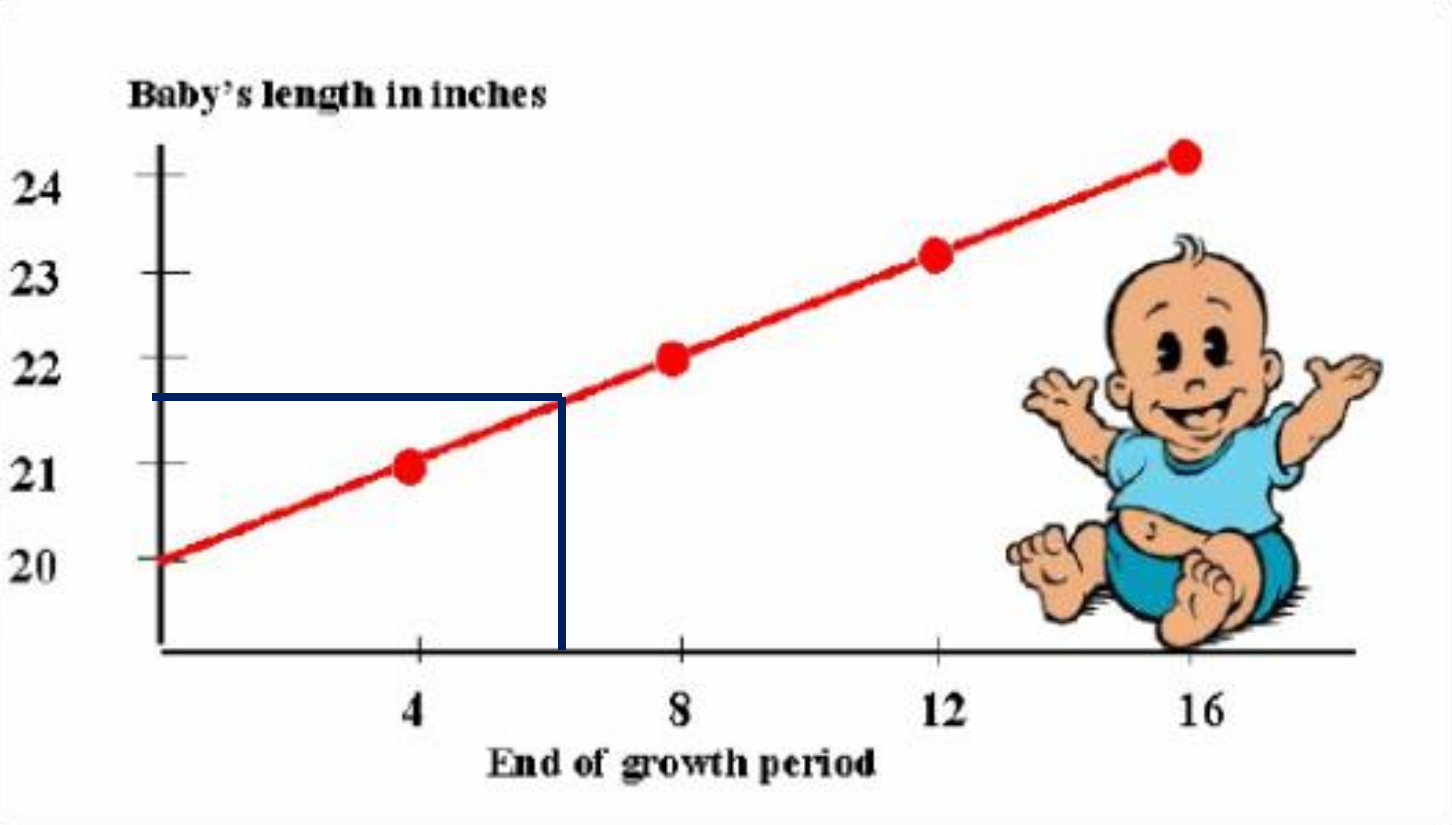


- **And finally classified it into ANSWER!!!**

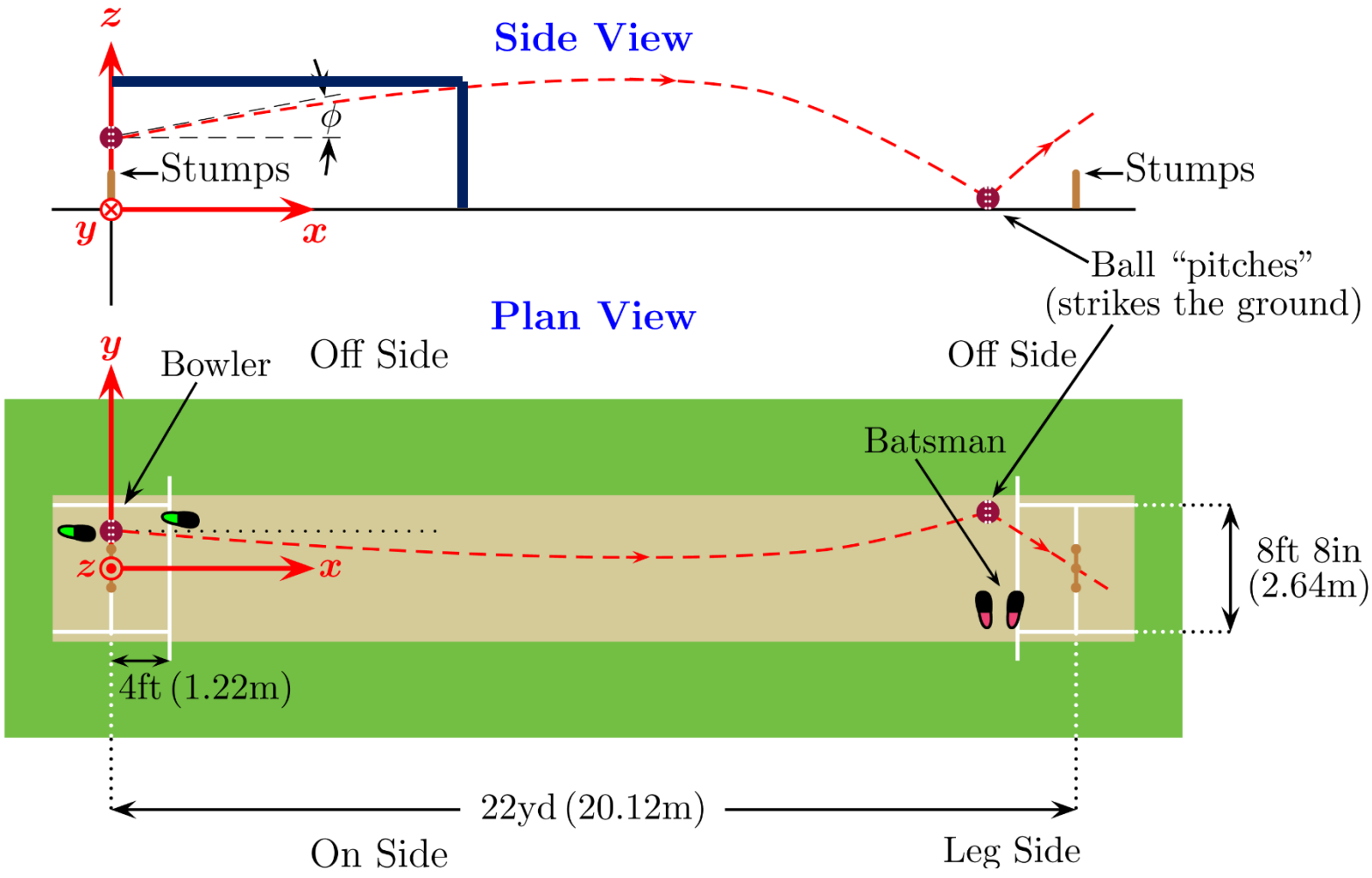
Aspect #1 Challenge



Learning – Aspect #2



Learning – Aspect #2

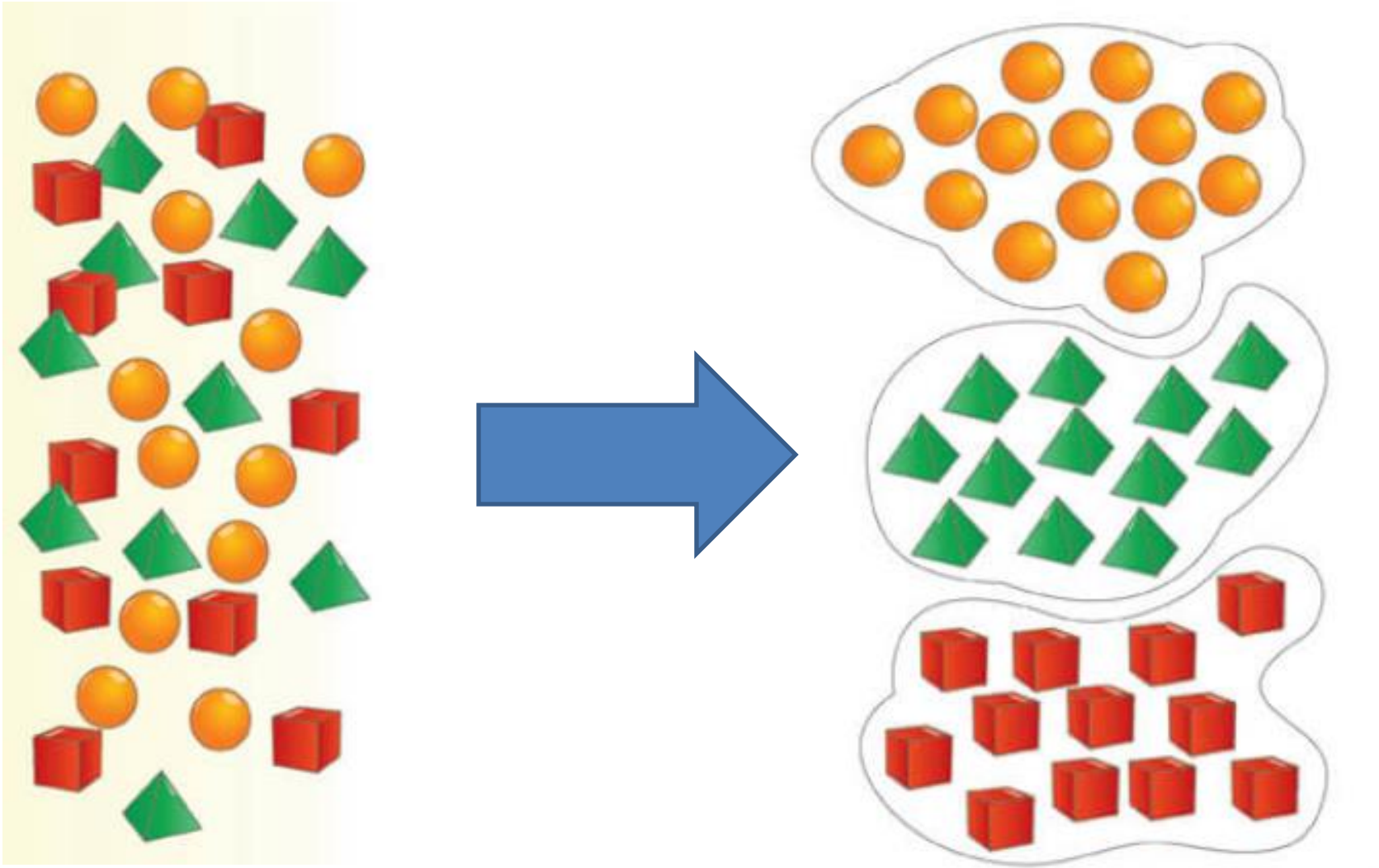


Continuous dataset

- **You used prior learnings of -**
  - **Variability**
  - **Coordinate geometry**
  - **X – Y correlation**
  - **Plot reading**

- **And finally build the ANSWER!!!**

Learning – Aspect #3



Learning – Aspect #3

Unknown dataset

- **You used prior learnings of -**
  - **Grouping objects**
  - **Extracting meaningfulness**
  - **Associating conjectures**

- **And finally build the Clusters!!!**

Learning – Aspect #4

4	3		
1	2	3	
		2	
2	1		

Learning – Aspect #4

Iterative revisit

- **You played keeping in mind**
  - **NEXT MOVE(s)**
  - **Probability of losing with a said move**
  - **Back tracking earlier opponent moves**

- **And finally moved!!!**

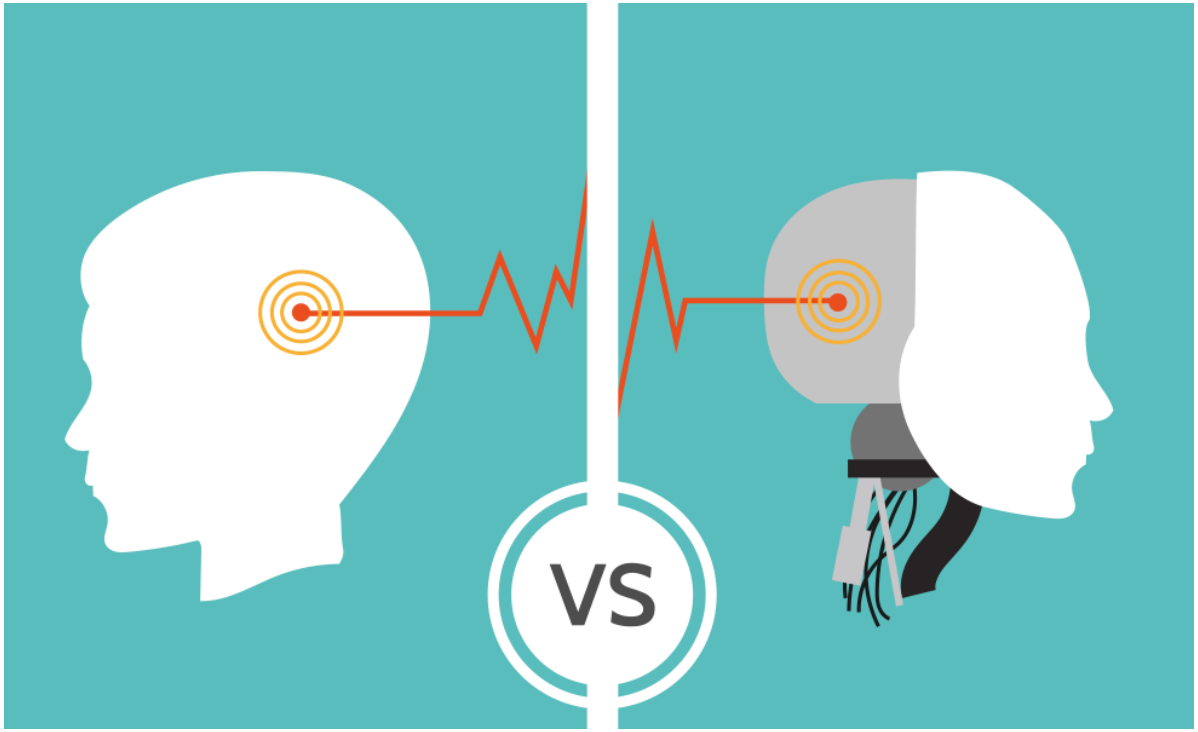
Aspect summary

- **#1 and #2**
  - **Prior information available to build answers**
- **#3**
  - **Meaningful grouping of unknown and cluttered data**
- **#4**
  - **Reiteration and loop back on pre and post information**

Aspect summary

- **#1 and #2** — Supervised
  - **Prior information available to build answers**
- **#3** — Unsupervised
  - **Meaningful grouping of unknown and cluttered data**
- **#4** — Reinforcement
  - **Reiteration and loop back on pre and post information**

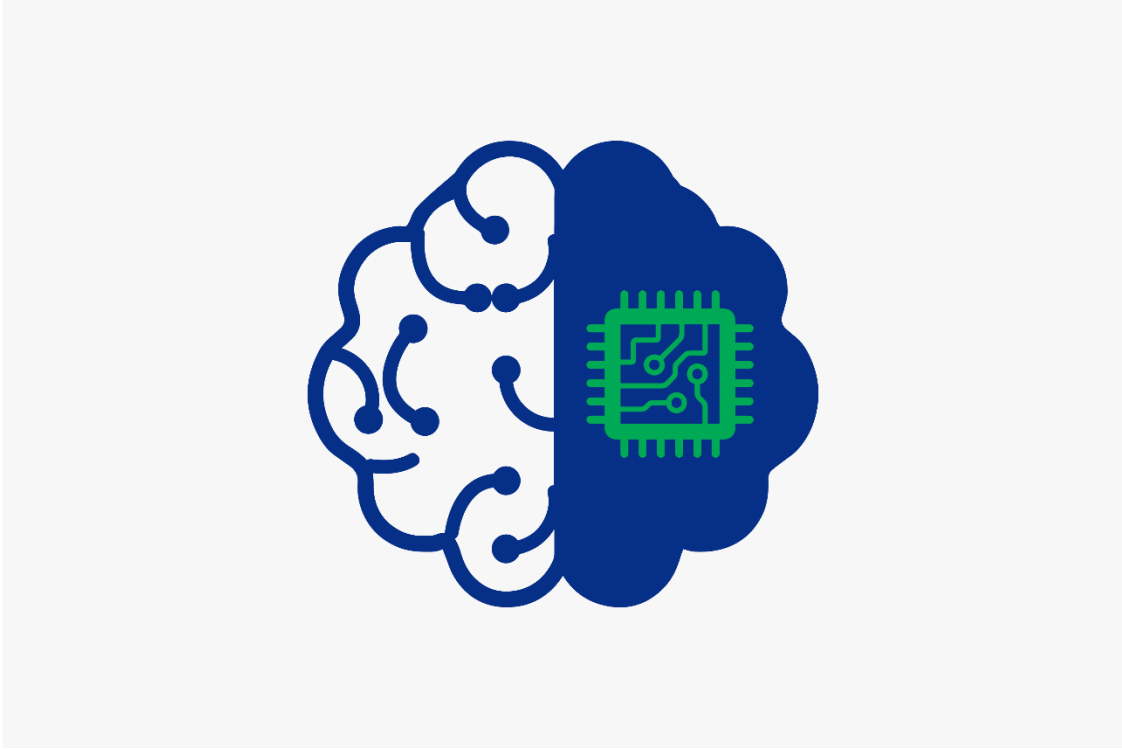
You v/s Computers



- **Everything is a PATTERN**
- **Natural ability to learn, unlearn and relearn**

- **Everything is DATA**
- **Will need code to understand data, visualize data and extract meaning**

Machine Learning



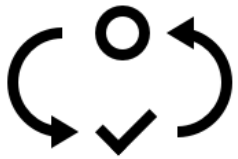
TYPES OF MACHINE LEARNING

Supervised Learning

Train an algorithm to perform classifcation and regression with a labelled data set.

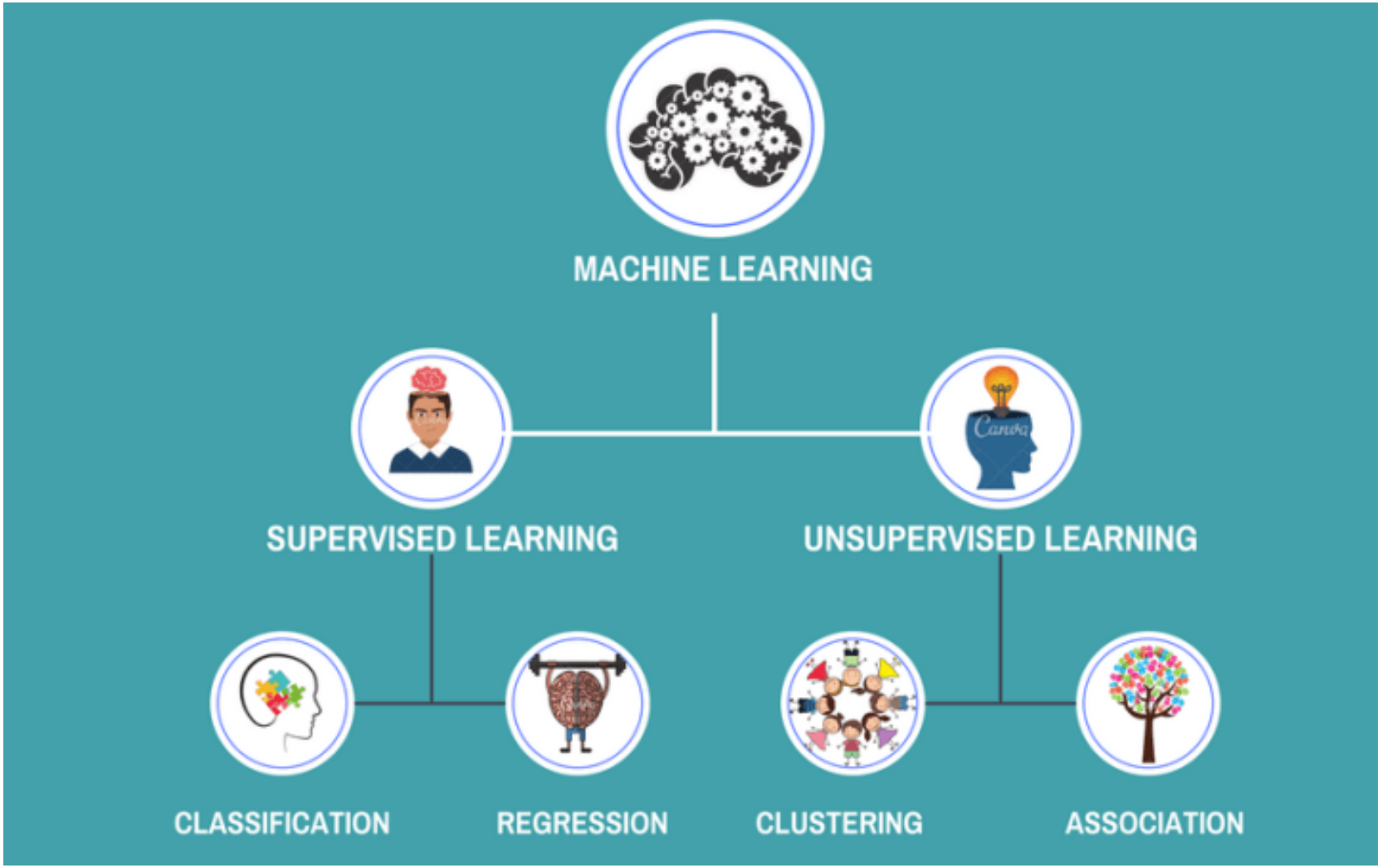
Unsupervised Learning

Train an algorithm to find clusters and associations in an unlabelled data set.



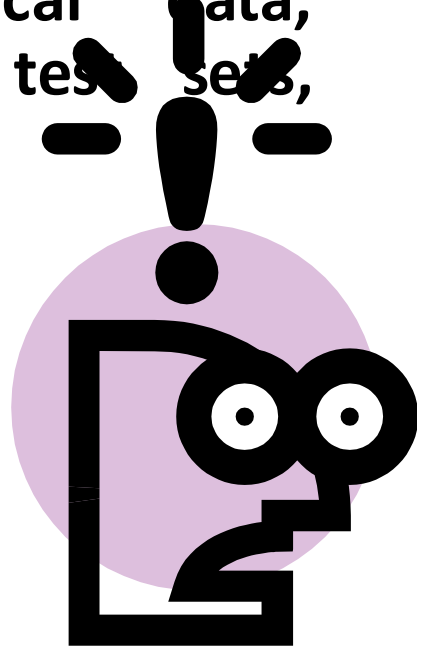
Reinforcement Learning

Train an agent to take certain actions in an environment without a data set.


MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
CLASSIFICATION
REGRESSION
CLUSTERING
ASSOCIATION

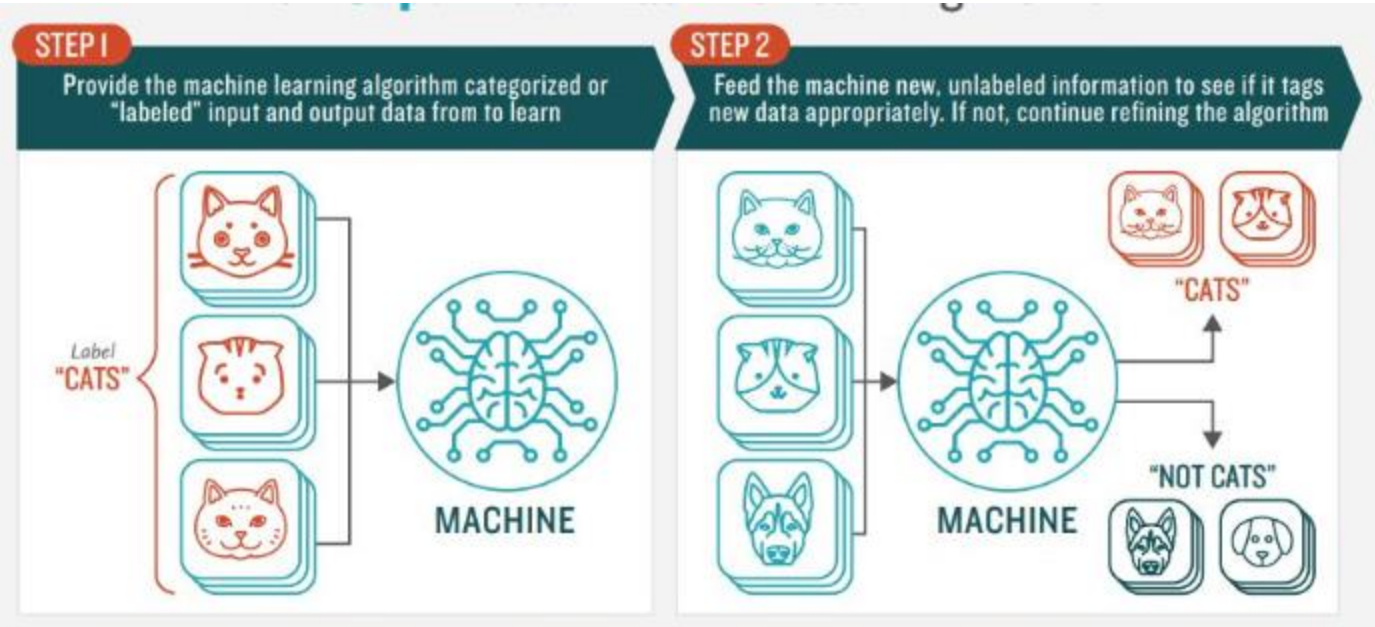
The course contents

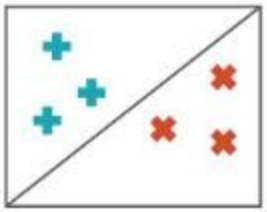
- **Introduction to machine learning : learning, need of machine learning, types of learning, well posed learning problems, designing a learning systems, statistical learning framework, empirical risk minimization, empirical risk minimization with inductive bias, PAC learning**
- **Building good training sets: data preprocessing, dealing with missing data, handling categorical data, partitioning a dataset in training and test sets, normalization, selecting meaningful features**
-



Introduction

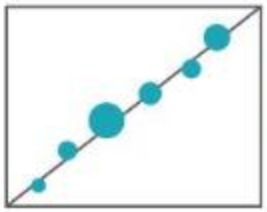
Unit 1





CLASSIFICATION

Sorting items into categories

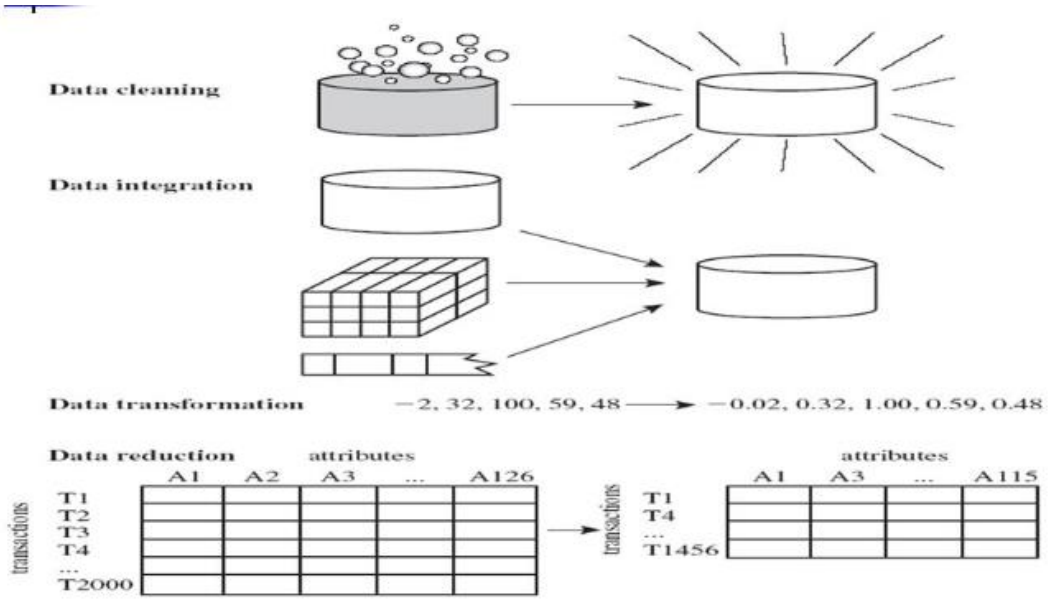


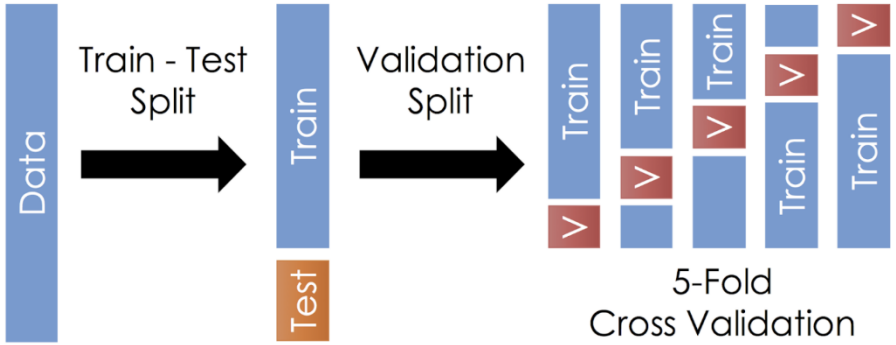
REGRESSION

Identifying real values (dollars, weight, etc.)

Data Pre-processing

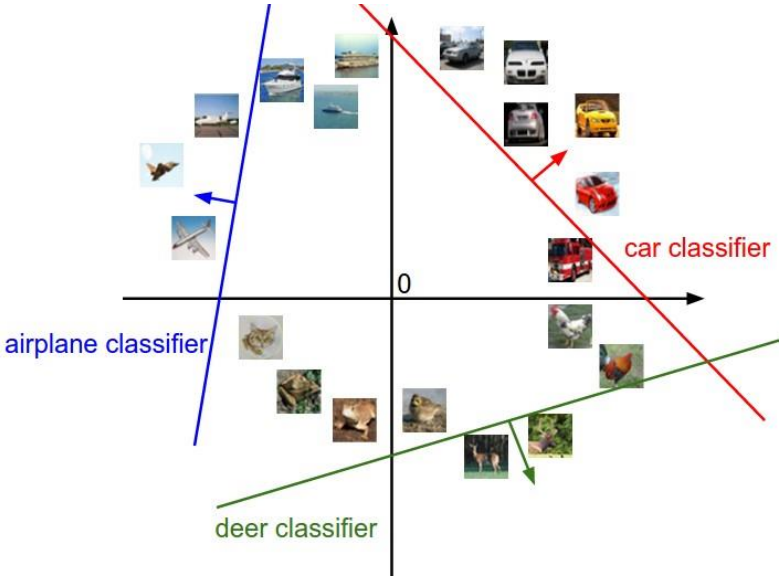
Unit 1

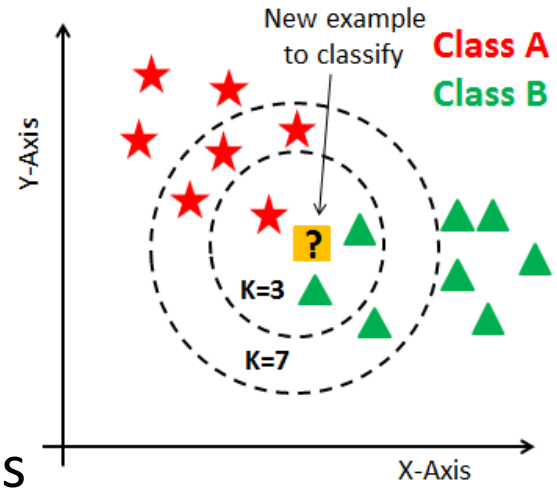




Scikit-learn classifiers

Unit 2





SVM, K-NN, Bayesian Algorithms

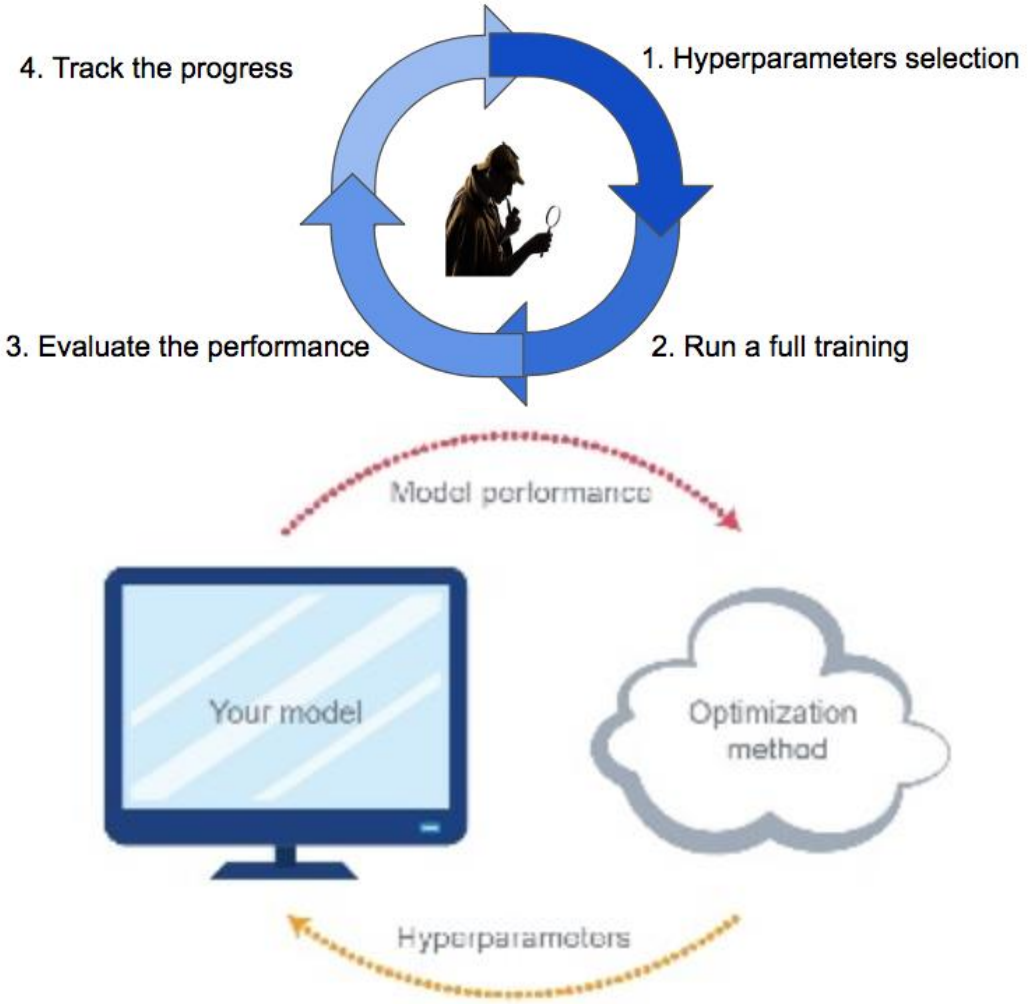
Generative Models

Unit 3

- **Maximum likelihood estimator**
- **Bayesian Learning**
- **Bayes Theorem**
- **Brute-force concept learning**
- **Bayes optimal classifier**
- **Gibs algorithm**
- **Naïve bayes classifier**
- **EM algorithm**

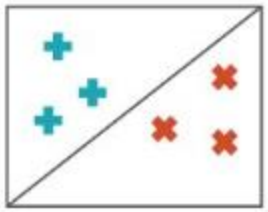
Model Evaluation

Unit 3



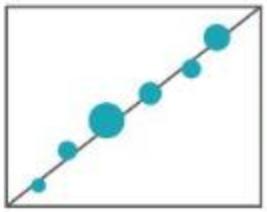
Regression Analysis

Unit 4



CLASSIFICATION

Sorting items into categories



REGRESSION

Identifying real values (dollars, weight, etc.)

Regression Metrics

Unit 5

- R^2 Score
- mean absolute error
- mean squared error
- mean squared logarithmic error
- mean absolute percentage error
- explained variance score
- D^2 score visual evaluation of regression models

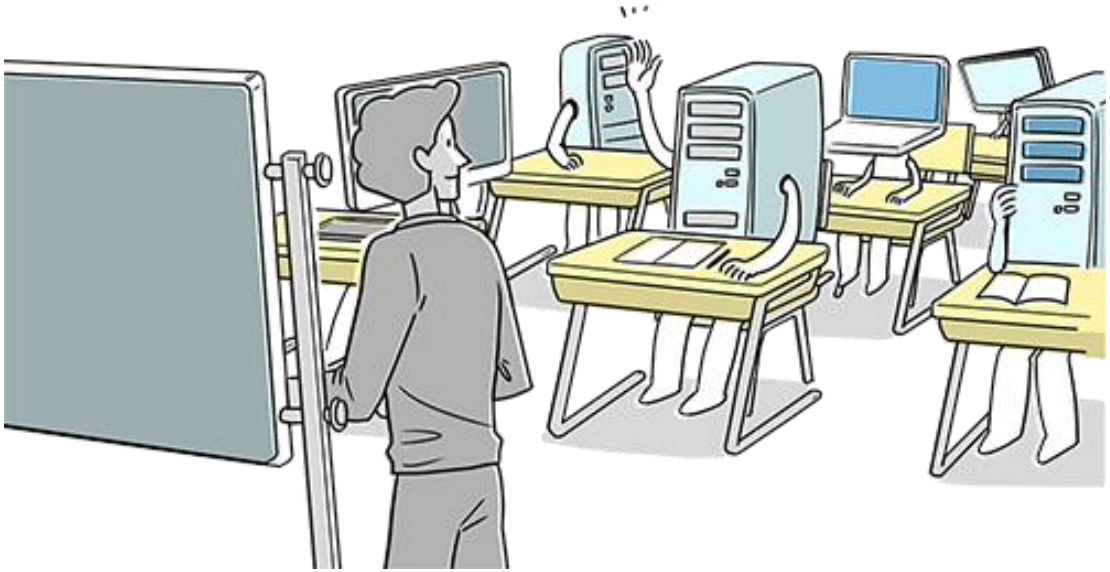
Unit 6

The bias-complexity tradeoff : no free lunch theorem, error decomposition, the VC-Dimension, the Rademacher complexity, the Natarajan dimension

Algorithm-Independent machine Learning : combining classifiers,majority voting classifier, re-sampling for estimating statistics, lack of inherent superiority of classifier, bagging and boosting classifier, random forest classifier and regressor, support vector classifier and regressor

Get Set Go!!!

Get trained to TRAIN systems…



Build futuristic solutions…

Assessment expectations & example