<welcome to>

SIG AI

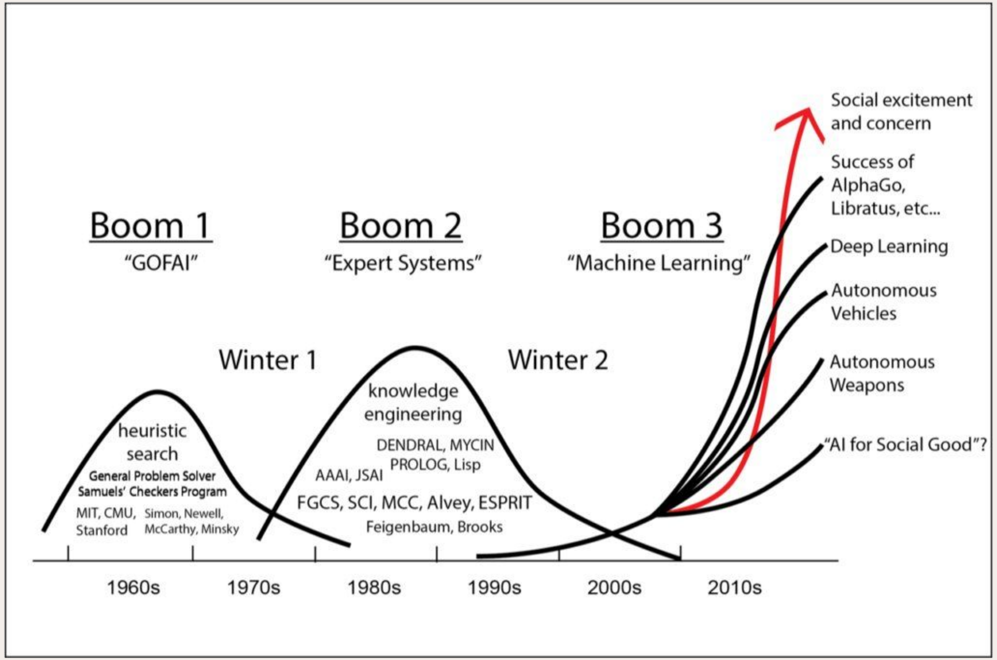
Meeting 1: 02/21/24

What is AI?

- Artificial intelligence (AI) is technology that enables computers and digital devices to do things humans do and build necessary technology to get there.
- AI Winters & Summers! (more on this later)
- A better definition by John MCCarthy: What is Artificial Intelligence
- Notion of modelling after biological systems: Boston Dynamics and the Human arm
- No turning point post-ChatGPT revolution
- In 1997, IBM Blue beats Kasparov at Chess (twice!)
- My personal favourite: AlphaGo Documentary | YouTube | Free
    - The victory is significant given the huge number of possible moves as the game progresses (over 14.5 trillion after just four moves!). **Later, Google purchased DeepMind for a reported USD 400 million.**

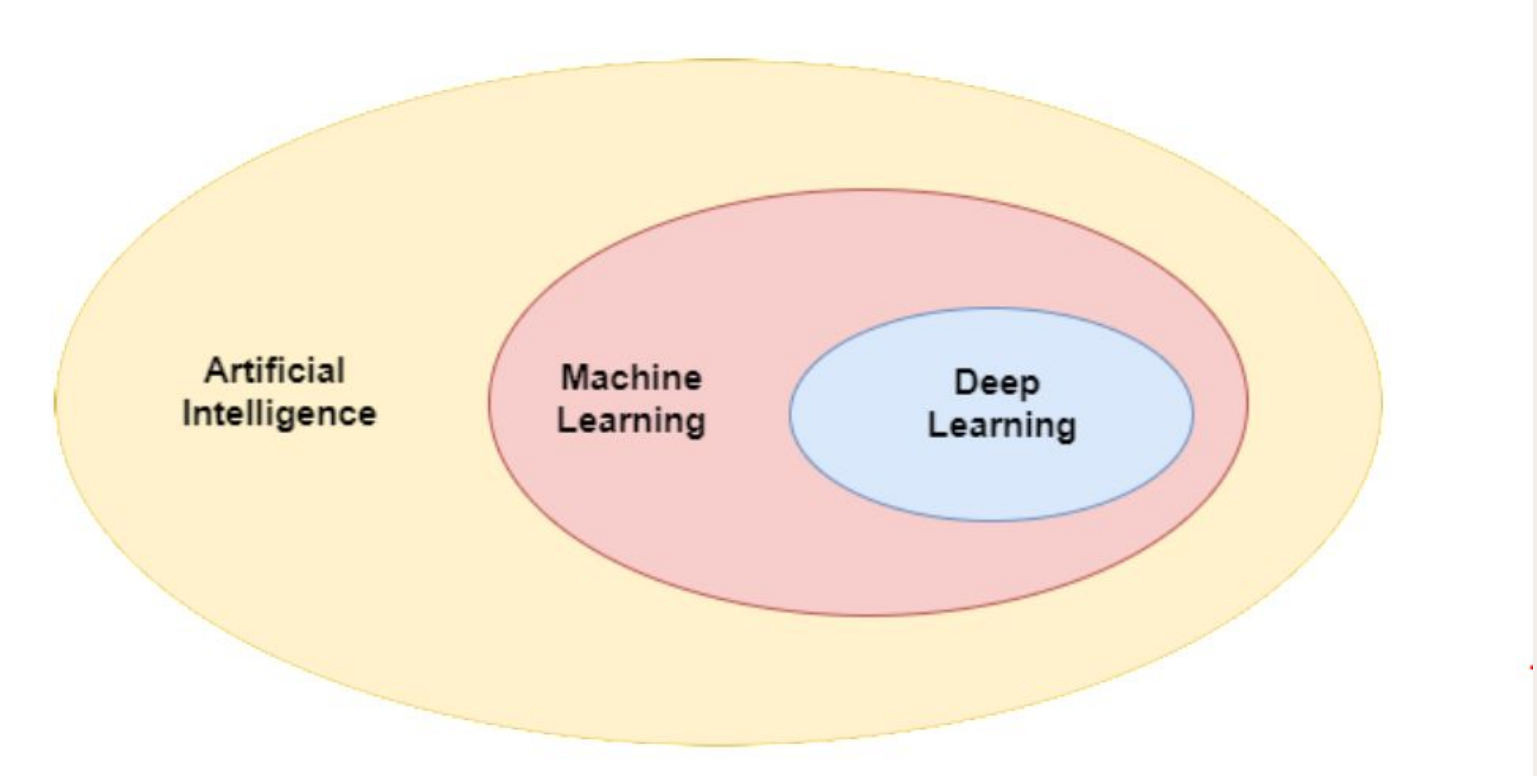
AI Winter

AI winters freeze progress
1956-1974
First wave of excitement
First neural networks and perceptrons written, first attempts at machine translation.
The U.S. Defense Advanced Research Projects Agency (DARPA) funds AI research with few requirements for delivering functioning products throughout the 1960s.
1980-1987
Renewed AI excitement
Expert systems emerge representing human decisions in if-then form. Funding picks up.
1994-present
Slow but steady progress
Computation power increases, big data provides training data, algorithms improve.
1950
1960
1970
1980
1990
2000
2010
2020
1974-1980
First AI winter
Limited applicability of AI leads to funding pullback in the U.S. and abroad.
1969: Researchers Marvin Minsky and Seymour Papert published Perceptrons, an influential book pointing out the ways early neural networks failed to live up to expectations.
1970-1974: DARPA cut its funding as enthusiasm wore thin.
1974: The Lighthill report, compiled by researcher James Lighthill for the British Science Research Council, stated: "In no part of the field [of AI] have the discoveries made so far produced the major impact that was then promised."
1987-1994
Second AI winter
Limitations of if-then reasoning become more apparent.
1987: Market for Lisp machines (specialty hardware for running AI applications) collapses.
1987: DARPA again cuts funding for AI research.
1990: Expert systems, an attempt to replicate human reasoning through a series of if-then rules, failed. The software proved hard to maintain and couldn't handle novel information, resulting in a cutback in AI development.
1991: Japanese Ministry of International Trade and Industry's Fifth Generation Computer project failed to deliver on goals of holding conversations, interpreting images and achieving humanlike reasoning.
ART: MBRUXELLE/ADOBE STOCK
©2019 TECHTARGET. ALL RIGHTS RESERVED
TechTarget

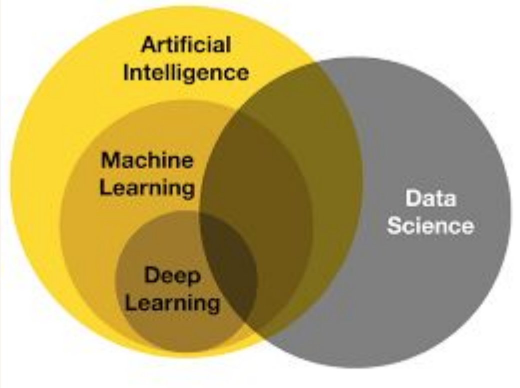


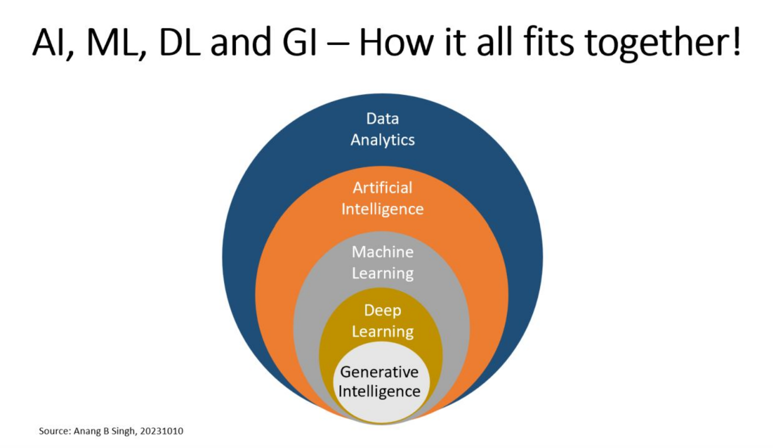
Boom 1
"GOFAI"
Boom 2
"Expert Systems"
Boom 3
"Machine Learning"
Winter 1
Winter 2
heuristic search
General Problem Solver
Samuels' Checkers Program
MIT, CMU, Stanford
Simon, Newell, McCarthy, Minsky
knowledge engineering
DENDRAL, MYCIN
PROLOG, Lisp
AAAI, JSAI
FGCS, SCI, MCC, Alvey, ESPRIT
Feigenbaum, Brooks
Social excitement and concern
Success of AlphaGo, Libratus, etc...
Deep Learning
Autonomous Vehicles
Autonomous Weapons
"AI for Social Good"?
1960s
1970s
1980s
1990s
2000s
2010s

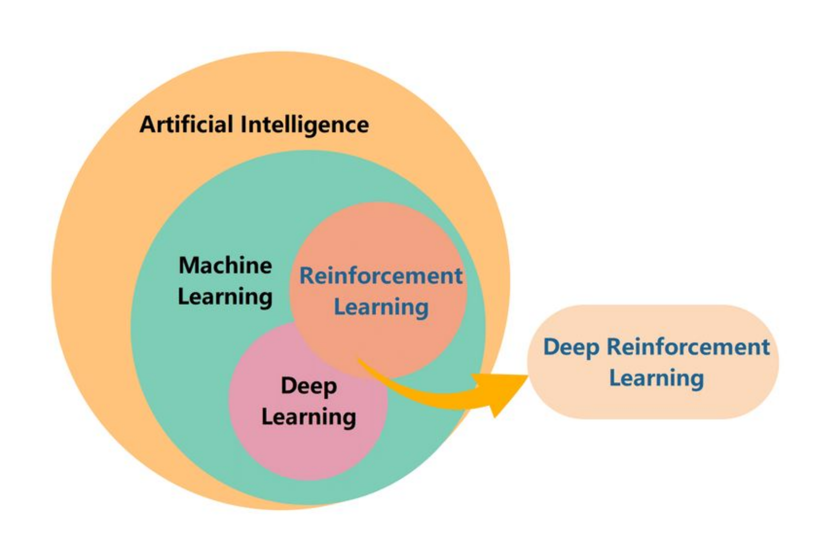
Whose their daddy?



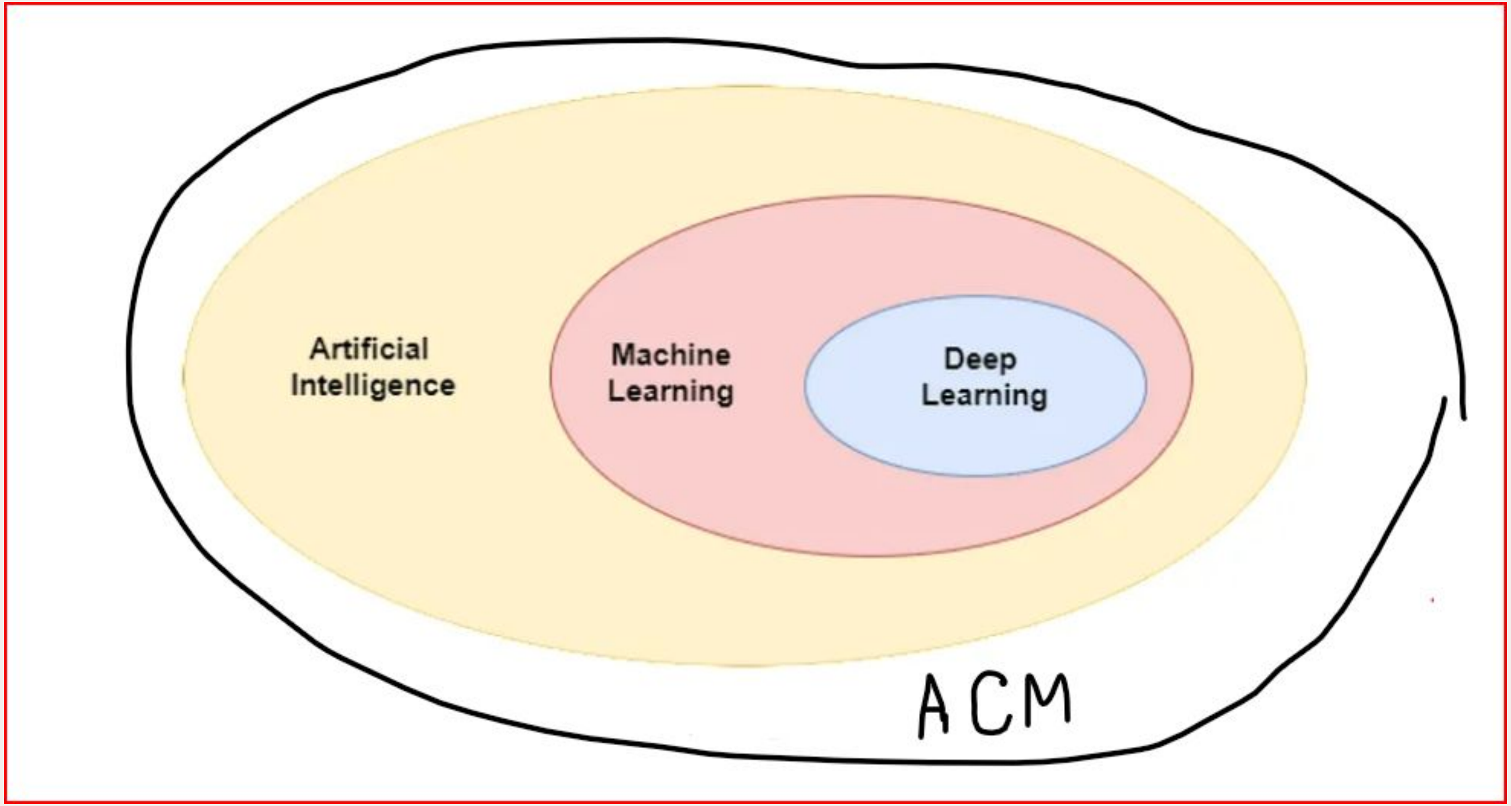
Data Scientists







Deep
Learning
Focus Groups


Artificial Intelligence
Machine Learning
Deep Learning
ACM

But, what do we agree* on?

*

Unsupervised Learning

do not have labelled data, learn more about the data by inferring patterns in the dataset without reference to the known outputs

Supervised Learning

takes a known set of input dataset and its known responses to the data (output) to learn the regression/classification model.

Deep Learning

learning from examples, computer model to filter the input data through layers to predict and classify information (layers >3)

Reinforcement Learning

an agent learns from an interactive environment in a trial and error way by continuously using feedback from its previous actions and experiences

But, what do we agree on?

*

Unsupervised Learning

market segmentation and image compression

Deep Learning

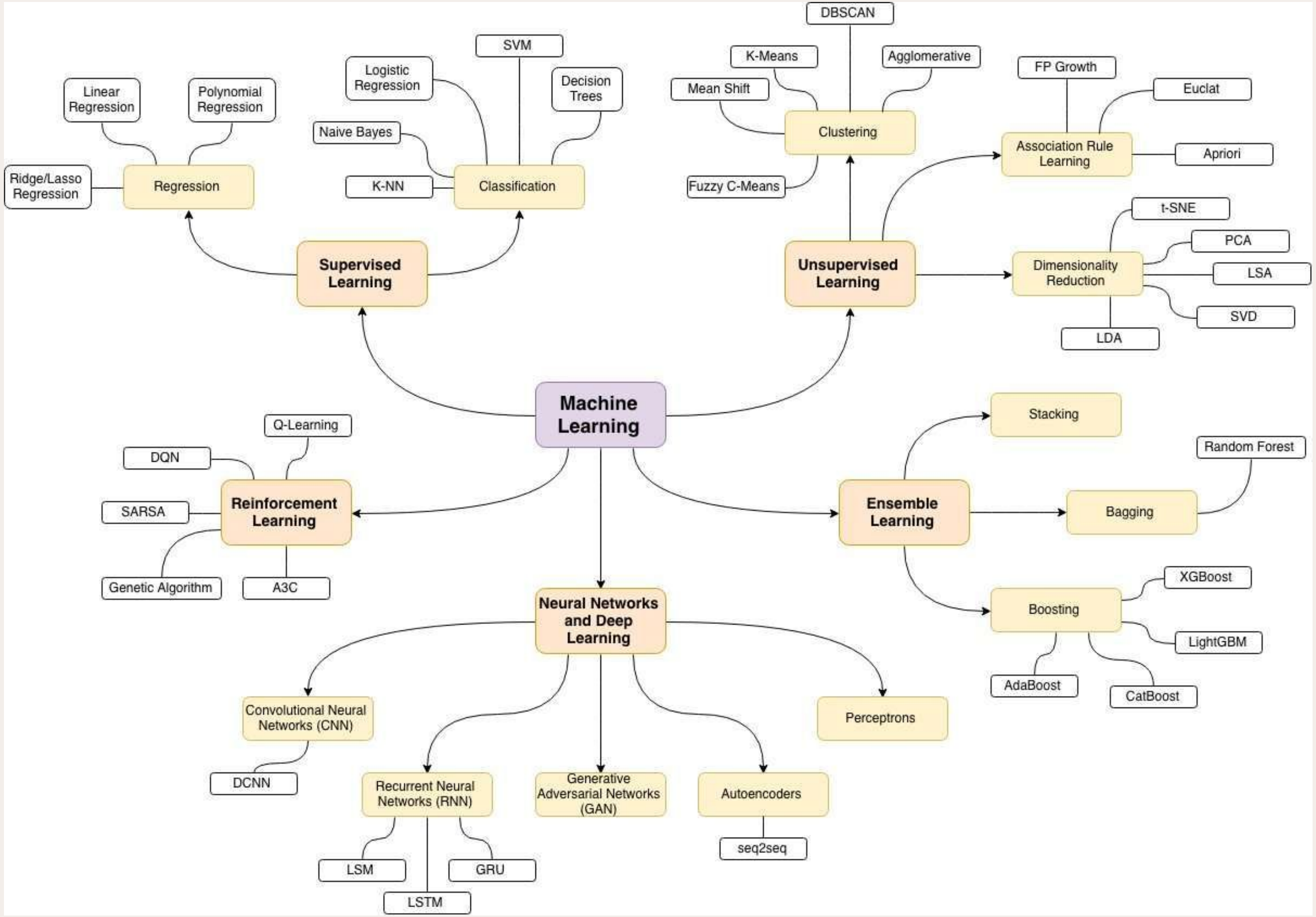
LLMs aka Large Language Models: ChatGPTTTTTTTTT

Supervised Learning

image classification and spam detection

Reinforcement Learning

game-playing AI that improves over time eg: AlphaGo


Machine Learning
Supervised Learning
Regression
Linear Regression
Polynomial Regression
Ridge/Lasso Regression
Classification
Logistic Regression
SVM
Decision Trees
Naive Bayes
K-NN
Unsupervised Learning
Clustering
DBSCAN
K-Means
Agglomerative
Mean Shift
Fuzzy C-Means
Association Rule Learning
FP Growth
Euclat
Apriori
Dimensionality Reduction
t-SNE
PCA
LSA
SVD
LDA
Reinforcement Learning
Q-Learning
DQN
SARSA
Genetic Algorithm
A3C
Ensemble Learning
Stacking
Bagging
Random Forest
Boosting
XGBoost
LightGBM
AdaBoost
CatBoost
Neural Networks and Deep Learning
Convolutional Neural Networks (CNN)
DCNN
Recurrent Neural Networks (RNN)
LSM
LSTM
GRU
Generative Adversarial Networks (GAN)
Autoencoders
seq2seq
Perceptrons

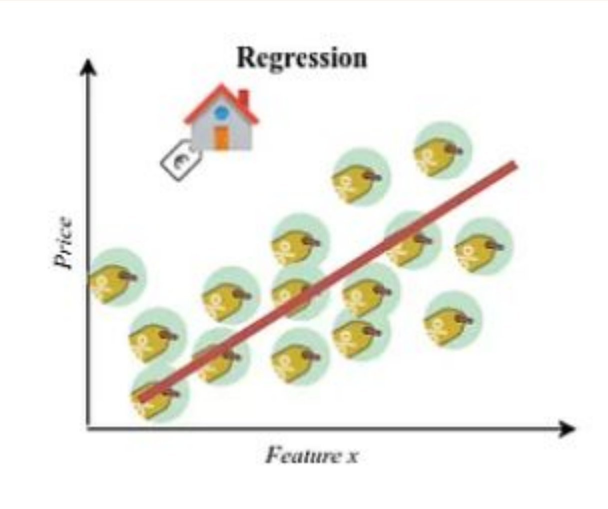
Schedule

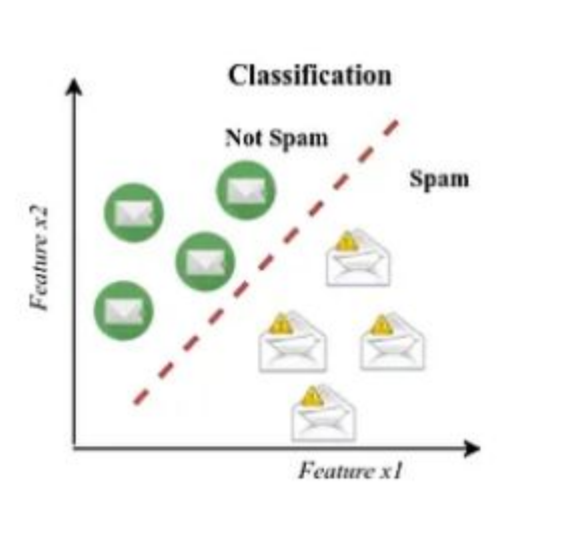
2/22	Introduction to all, basics of Supervised Learning
3/7	Supervised Learning contd.
3/21	Unsupervised Learning
4/4	Deep Learning
4/18	Reinforcement Learning

Supervised Learning

Regression

Classification





Supervised Learning

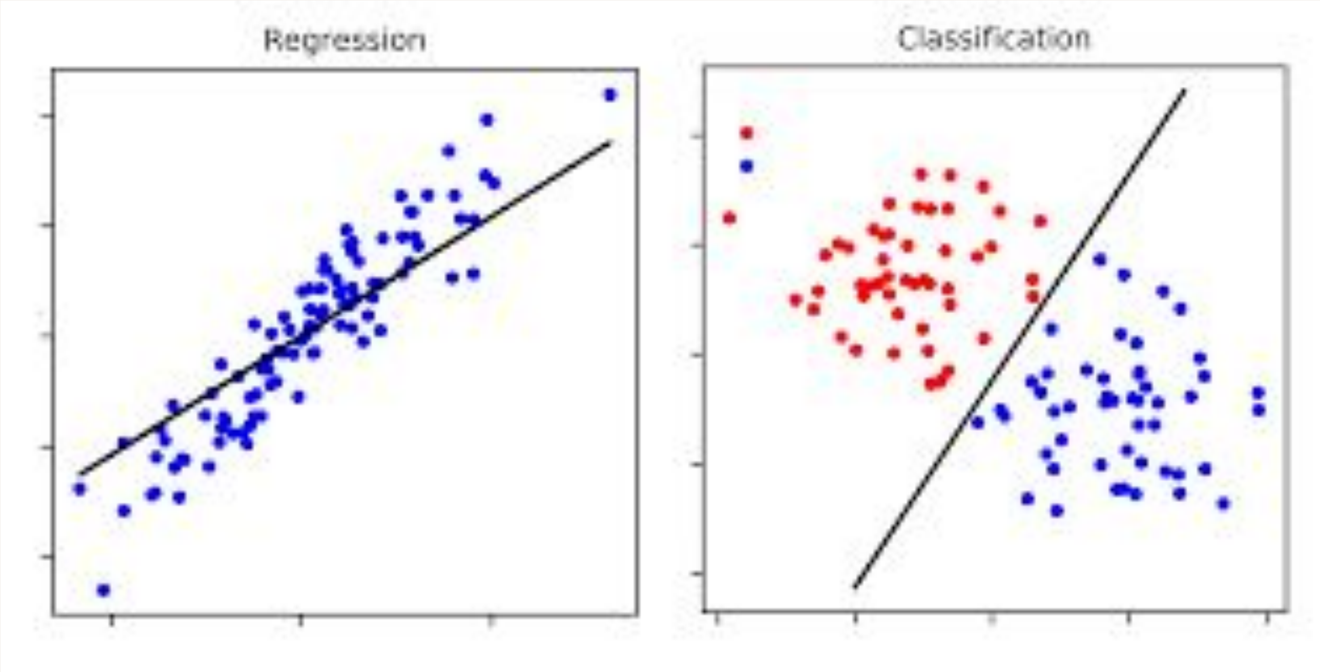
Regression

- Act on continuous target variables
- Aim for decreasing error between predicted and actual values

Classification

- build a model that can accurately assign a label or category to a new observation based on its features
- predict discrete target variables(class labels) using independent features or find a decision boundary that can separate the different classes in the target variable.
- Validated with confusion matrix, precision & recall and cross-validation

Supervised Learning



Supervised Learning

Regression

- Which of the following is NOT a regression task:
  - Predicting age of a person
  - Predicting nationality of a person
  - Predicting whether stock price of a company will increase tomorrow
  - Predicting whether a document is related to sighting of UFOs?
- Which of the following is NOT a regression task:
  - Predicting the age of a car based on its mileage
  - Predicting whether a person will buy a product or not based on their browsing history
  - Predicting the score of a student in an exam based on their study hours
  - Predicting the price of a stock next week based on past prices?

Classification

- Which of the following is NOT a classification task:
  - Predicting whether an email is spam or not based on its content
  - Predicting the breed of a dog based on its physical characteristics
  - Predicting the height of a person based on their parents' heights
  - Predicting whether a credit card transaction is fraudulent or not?
- Which of the following is NOT a classification task:
  - Predicting whether a patient has a disease or not based on their symptoms
  - Predicting the species of a flower based on measurements of its petals
  - Predicting the temperature tomorrow based on weather data
  - Predicting whether a passenger survived on the Titanic based on their age, class, and gender?

Supervised Learning

Regression

Classification

Linear Regression

On continuous data,
finding a best fit line

Logistic Regression

On categorical data, finds bet fit probability curve while doing binary classification

K-NN

Groups data depending on it’s nearest neighbours

Naive Bayes

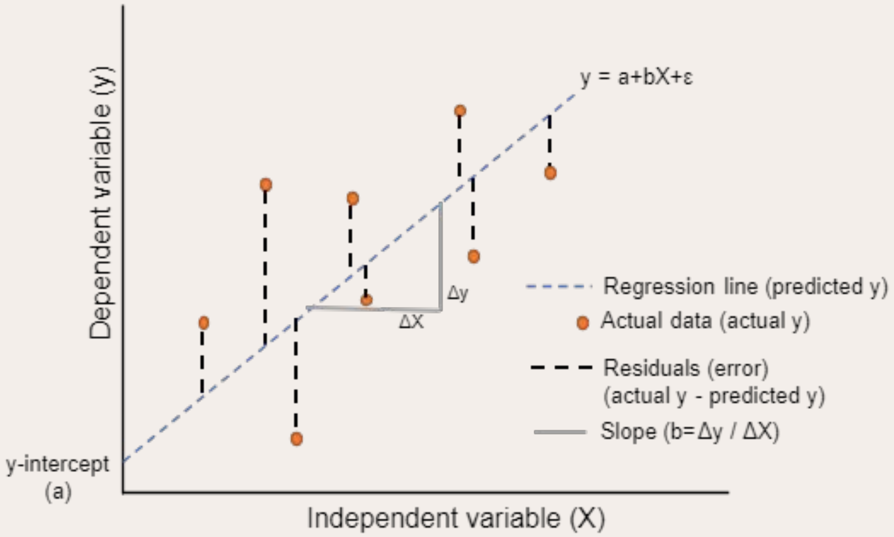
assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature all of these properties independently contribute to the probability. This method is easy to build, particularly for very large data sets.

Decision Trees

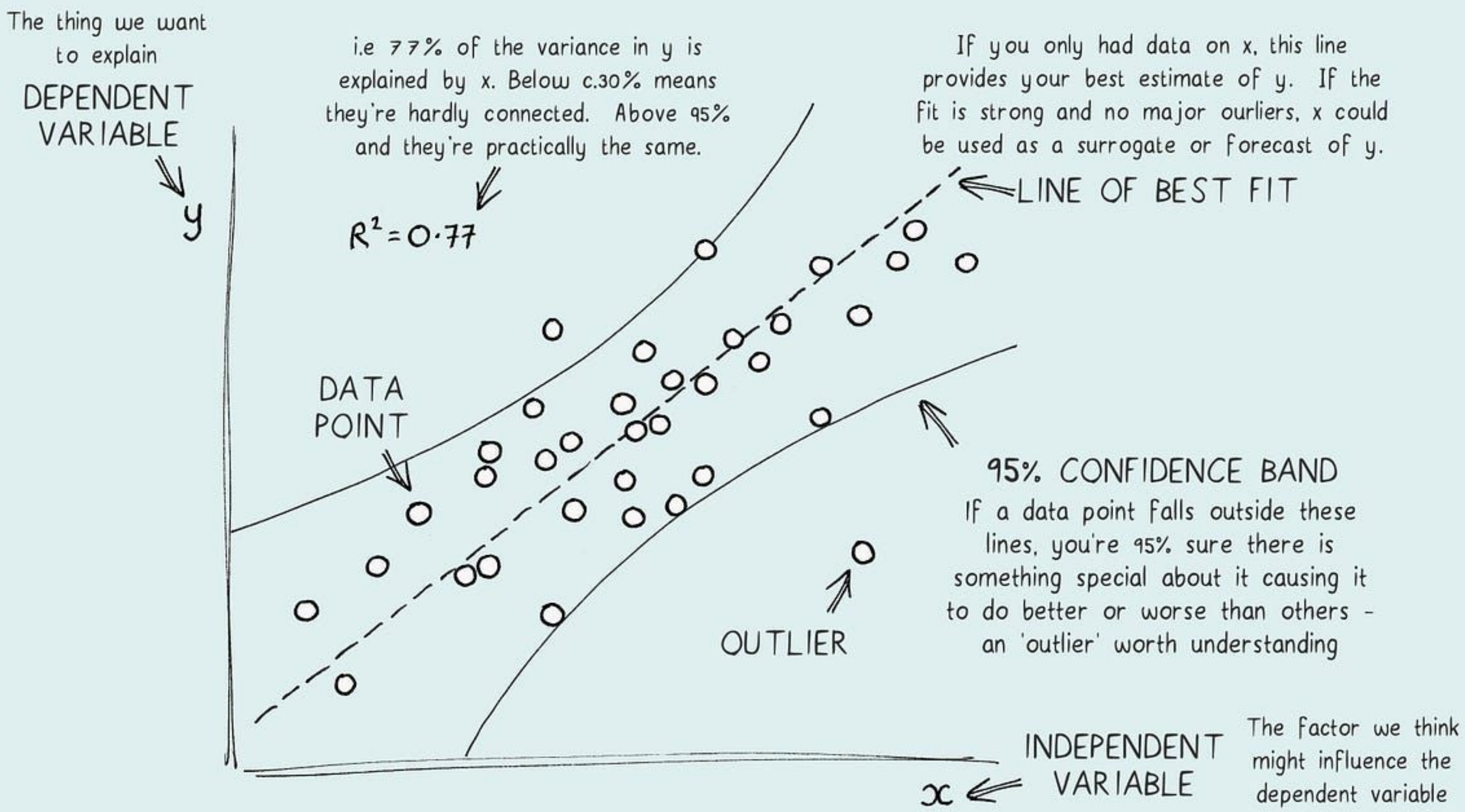
partition based on the attribute value. It partitions the tree recursively in a manner called recursive partitioning

Linear Regression

- assumes that there is a linear relationship between the independent and dependent variables, this means that the change in the dependent variable is proportional to the change in the independent variables
- y = f(x)
    - Say, f(x) = x + 7, whatever value of x is put in this equation, will decide what value y gets.
    - x is independent bc it calls the shots, y is dependent bc it's value is a result of operations on x

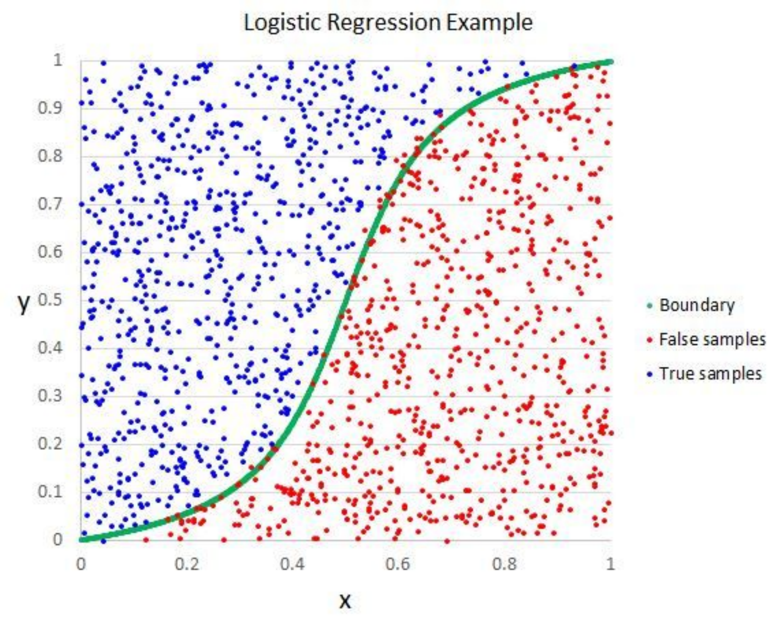


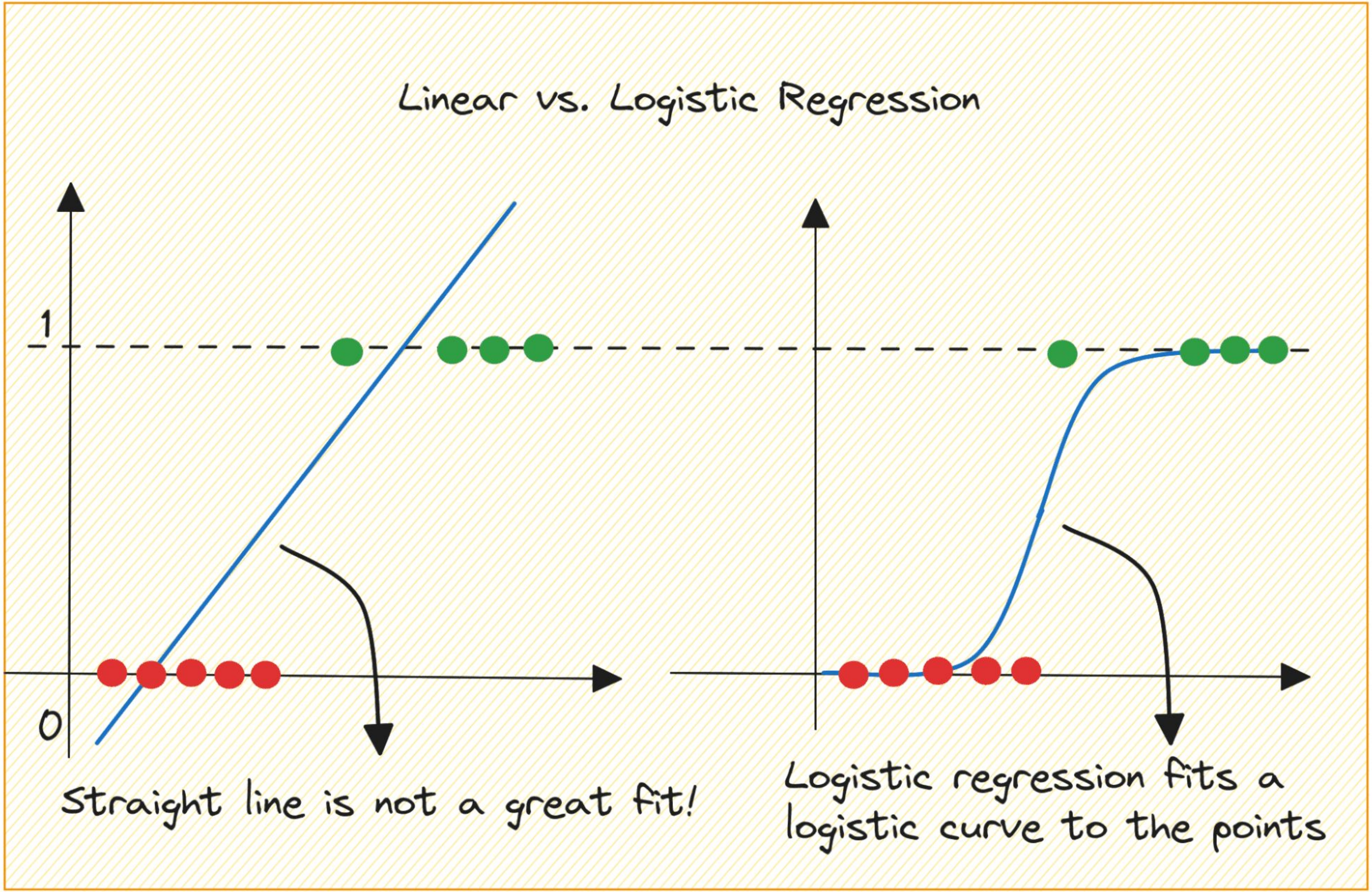
LINEAR REGRESSION



Logistic Regression

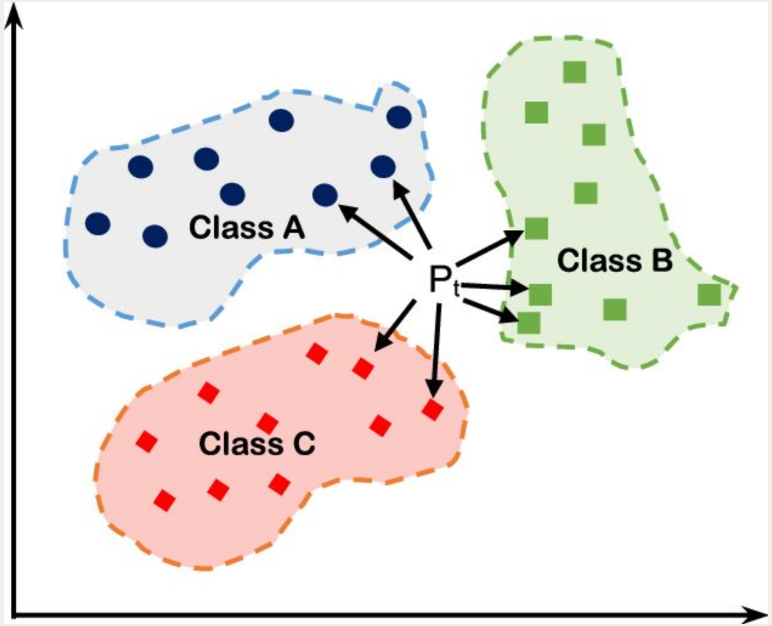
- Similar to Linear regression except 2 things:
  - out a singular value (of probability)
  - Used for categorical data
- There should not be collinearity between independent variables.



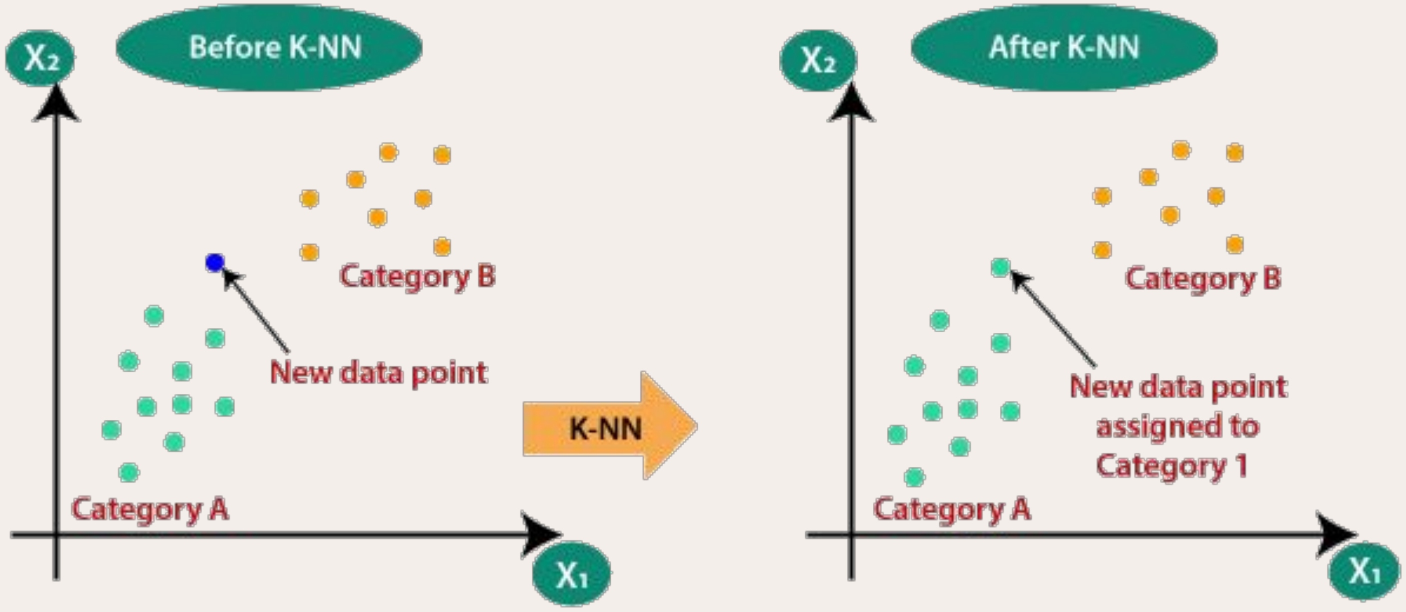

Linear vs. Logistic Regression
1
0
Straight line is not a great fit!
Logistic regression fits a
logistic curve to the points

K-NN

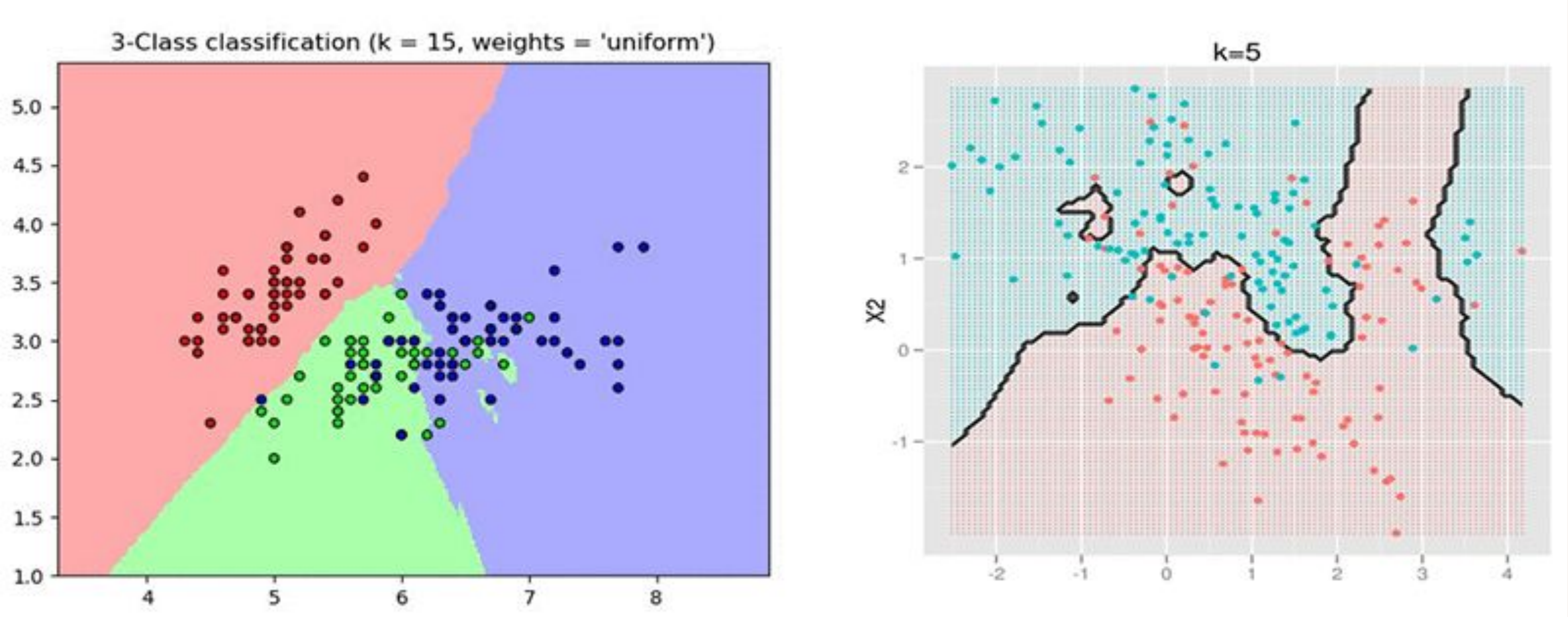
- non-parametric
- given an unclassified point, we can assign it to a group by observing what group its nearest neighbors belong to. This means a point close to a cluster of points classified as ‘Red’ has a higher probability of getting classified as ‘Red’
- Increasing k doesn’t necessarily lead to a better classification
- less sensitive to outliers compared to other algorithms



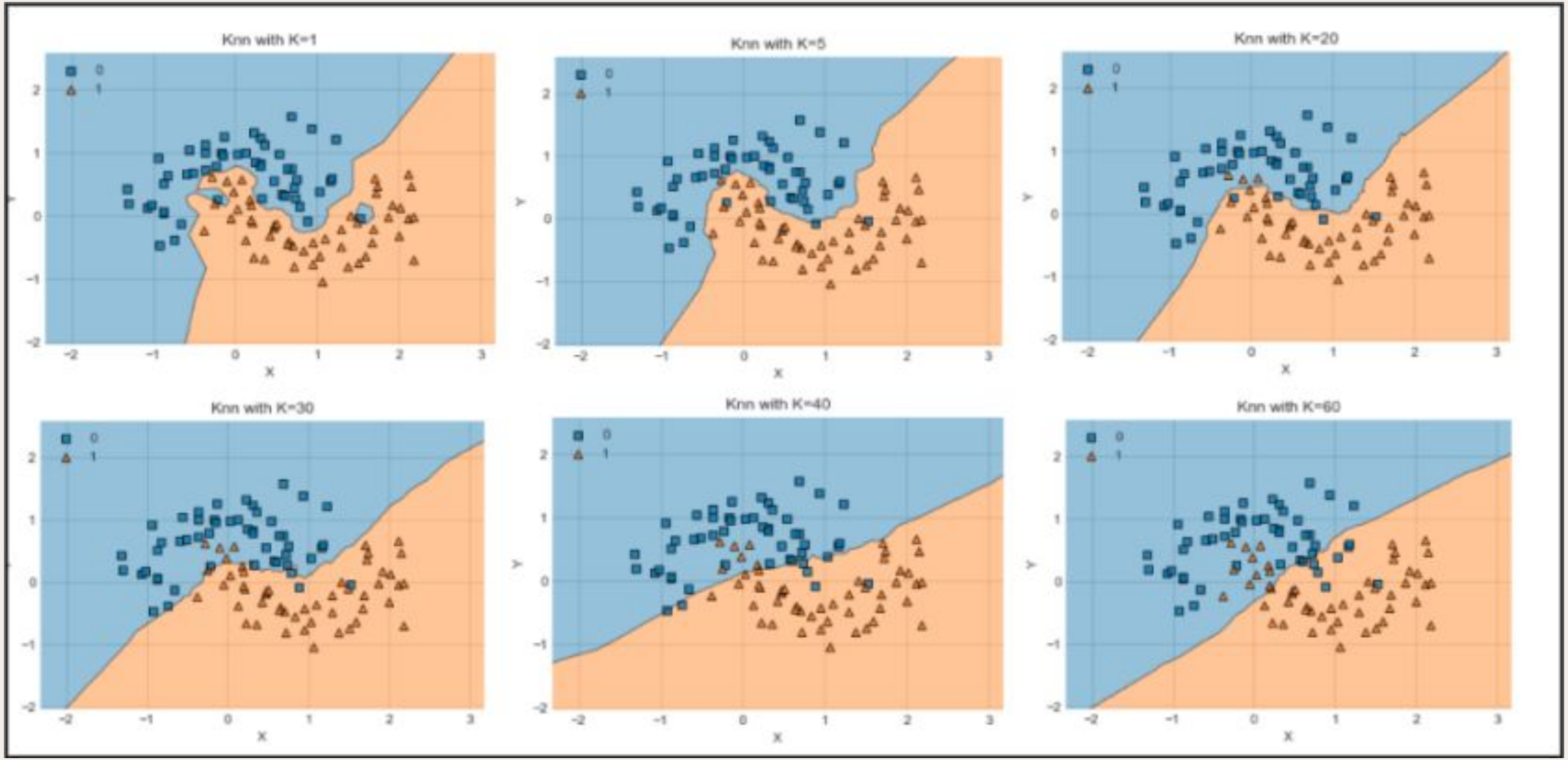
K-NN


Before K-NN
X_2
New data point
Category B
Category A
X_1
K-NN
After K-NN
X_2
New data point
assigned to
Category 1
Category B
Category A
X_1

K-NN



K-NN

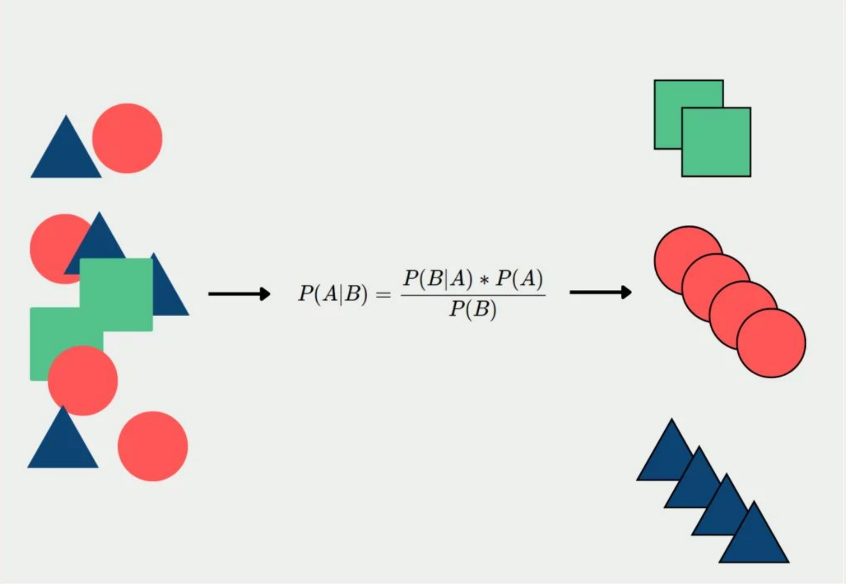


Knn with K=1
Knn with K=5
Knn with K=20
Knn with K=30
Knn with K=40
Knn with K=60

Naive Bayes

- a collection of classification algorithms based on Bayes' Theorem. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e. every pair of features being classified is independent of each other
- Why 'Naive'? – classifier assumes that the features used to describe an observation are conditionally independent

$$\boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)}}$$



*

Day	outlook	temp	humidity	windy	play
1	sunny	hot	high	FALSE	no
2	sunny	hot	high	TRUE	no
3	overcast	hot	high	FALSE	yes
4	rainy	mild	high	FALSE	yes
5	rainy	cool	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
7	overcast	cool	normal	TRUE	yes
8	sunny	mild	high	FALSE	no
9	sunny	cool	normal	FALSE	yes
10	rainy	mild	normal	FALSE	yes
11	sunny	mild	normal	TRUE	yes
12	overcast	mild	high	TRUE	yes
13	overcast	hot	normal	FALSE	yes
14	rainy	mild	high	TRUE	no

X (feature vector) = (Sunny, Hot, Normal, False)
y = No

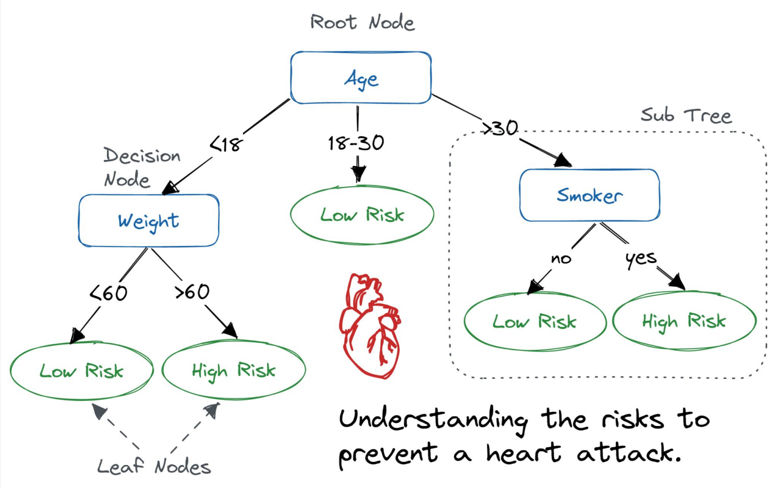
$$P(No|today) = \frac{P(SunnyOutlook|No)P(HotTemperature|No)P(NormalHumidity|No)P(NoWind|No)P(No)}{P(today)}$$

$$P(No|today) \propto \frac{3}{5}\cdot\frac{2}{5}\cdot\frac{1}{5}\cdot\frac{2}{5}\cdot\frac{5}{14} \approx 0.0068$$

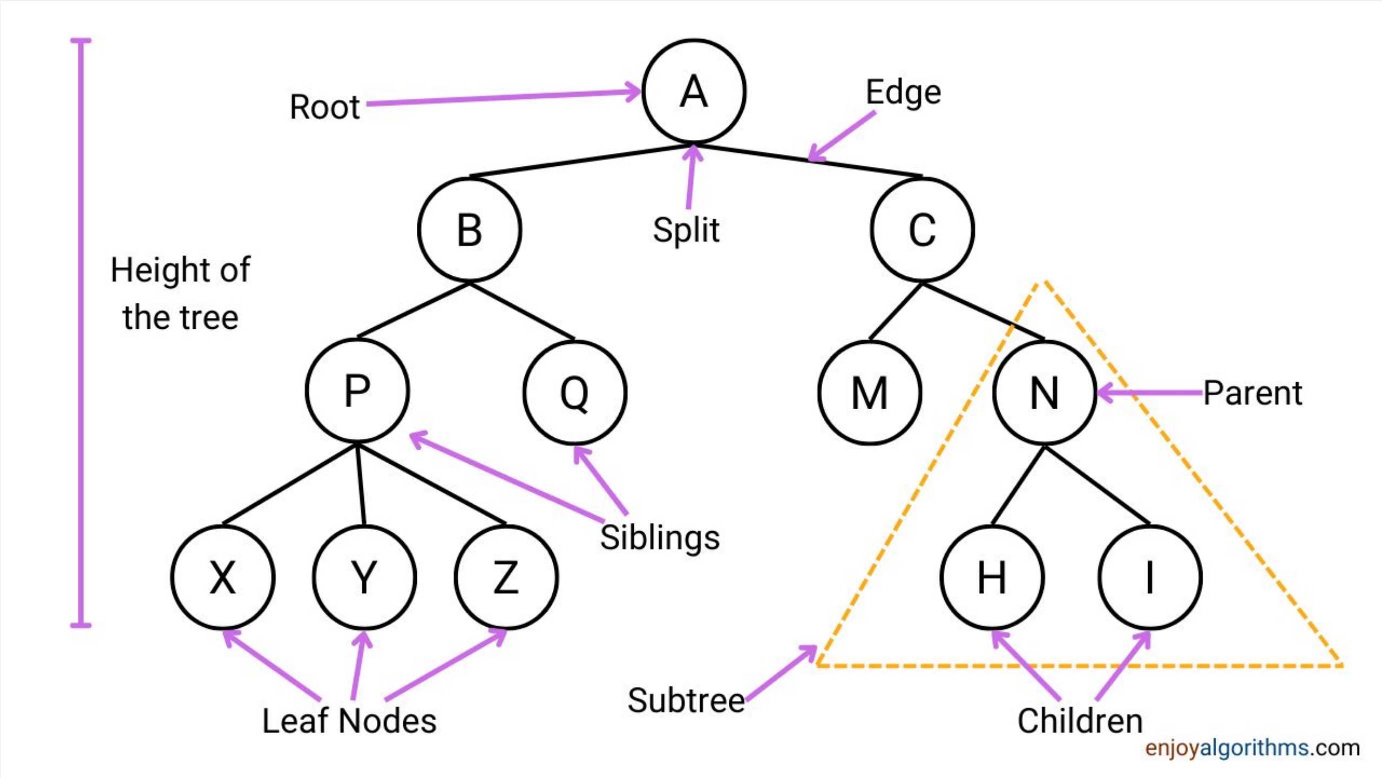
So basically, P(y|X) here means, the probability of "Not playing golf" given that the weather conditions are "Sunny outlook", "Temperature is hot", "normal humidity" and "no wind" = 0.0068
< more in STAT 381 or IE 342 >

Decision Trees

- tree can be "learned" by splitting the source set into subsets based on Attribute Selection Measures, which is a criterion used to evaluate the usefulness of different attributes for splitting a dataset
- less requirement of data cleaning compared to other algorithms
- useful for solving decision-related problems



Decision Trees



Day	outlook	temp	humidity	windy	play
1	sunny	hot	high	FALSE	no
2	sunny	hot	high	TRUE	no
3	overcast	hot	high	FALSE	yes
4	rainy	mild	high	FALSE	yes
5	rainy	cool	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
7	overcast	cool	normal	TRUE	yes
8	sunny	mild	high	FALSE	no
9	sunny	cool	normal	FALSE	yes
10	rainy	mild	normal	FALSE	yes
11	sunny	mild	normal	TRUE	yes
12	overcast	mild	high	TRUE	yes
13	overcast	hot	normal	FALSE	yes
14	rainy	mild	high	TRUE	no