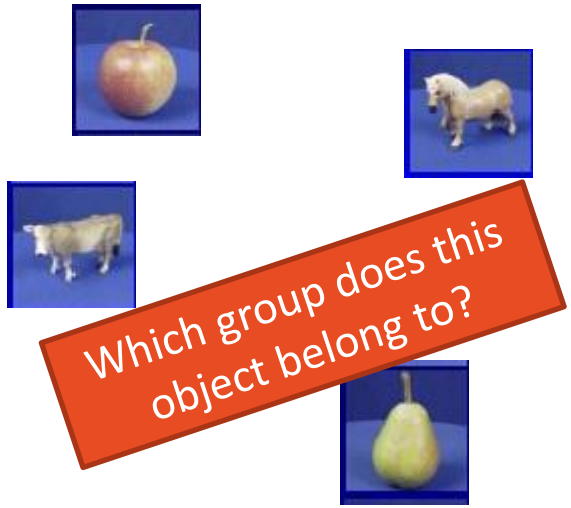
Introduction to Machine Learning

The Philosophy
Advantages
Applications
Algorithms

Venkat Reddy

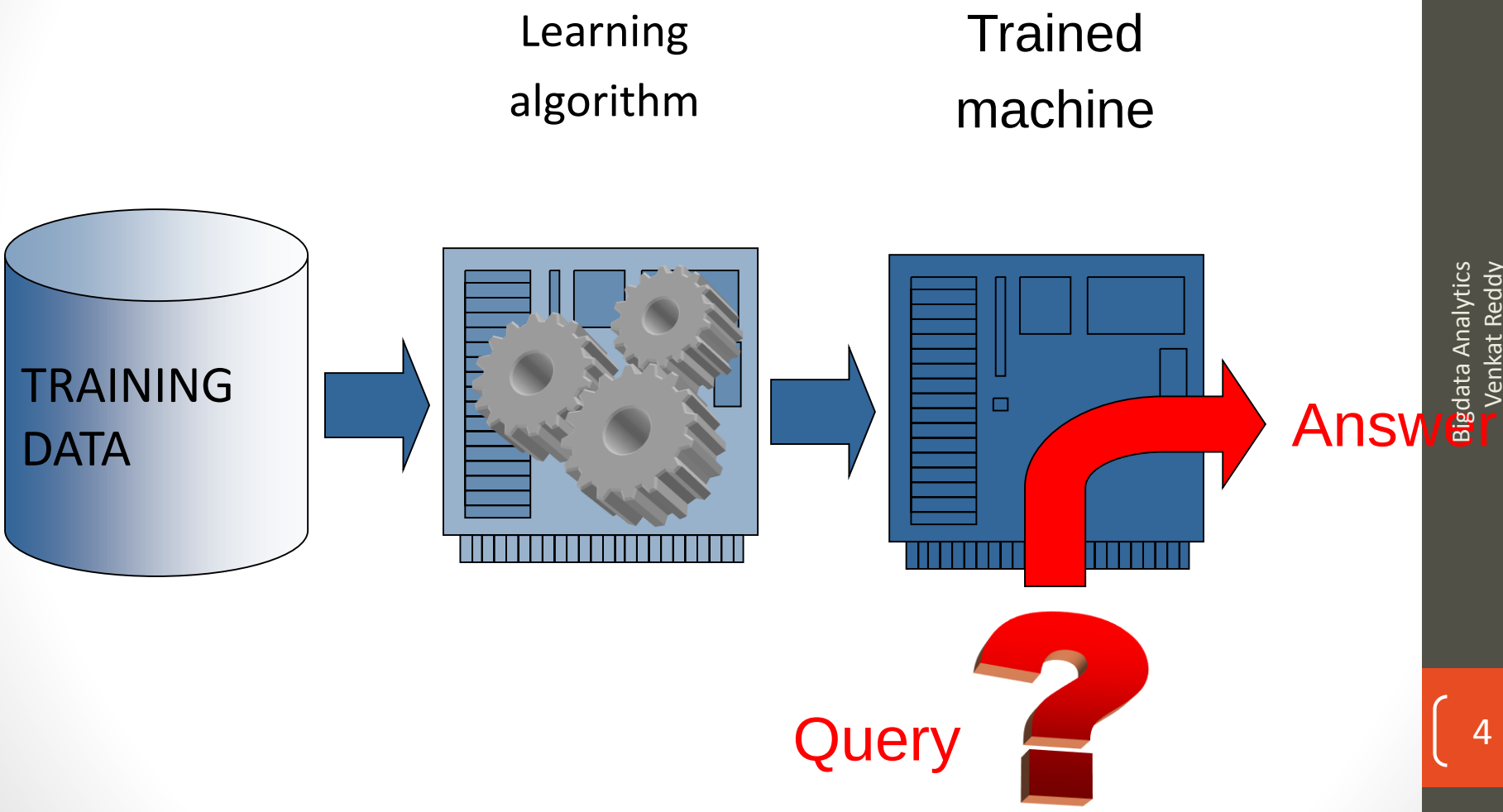
The Learning



Machine Learning

- How did our brain process the images?
- How did the grouping happen?
- Human brain processed the given images - **learning**
- After learning the brain simply looked at the new image and compared with the groups classified the image to the closest group - **Classification**
- If a machine has to perform the same operation we use Machine Learning
- We write programs for learning and then classification, this is nothing but machine learning

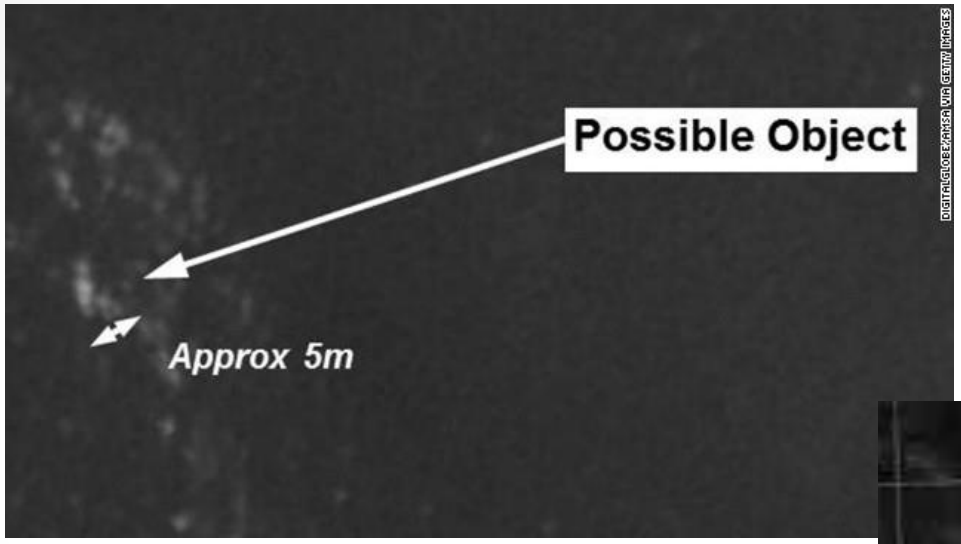
Machine Learning

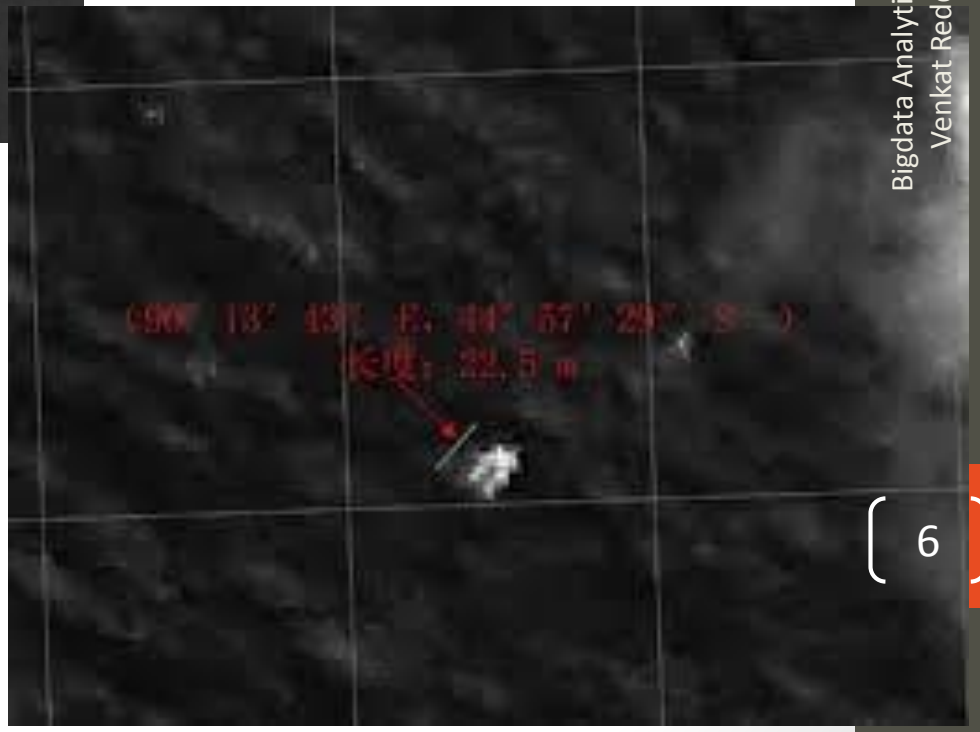


Example: Classification using ML

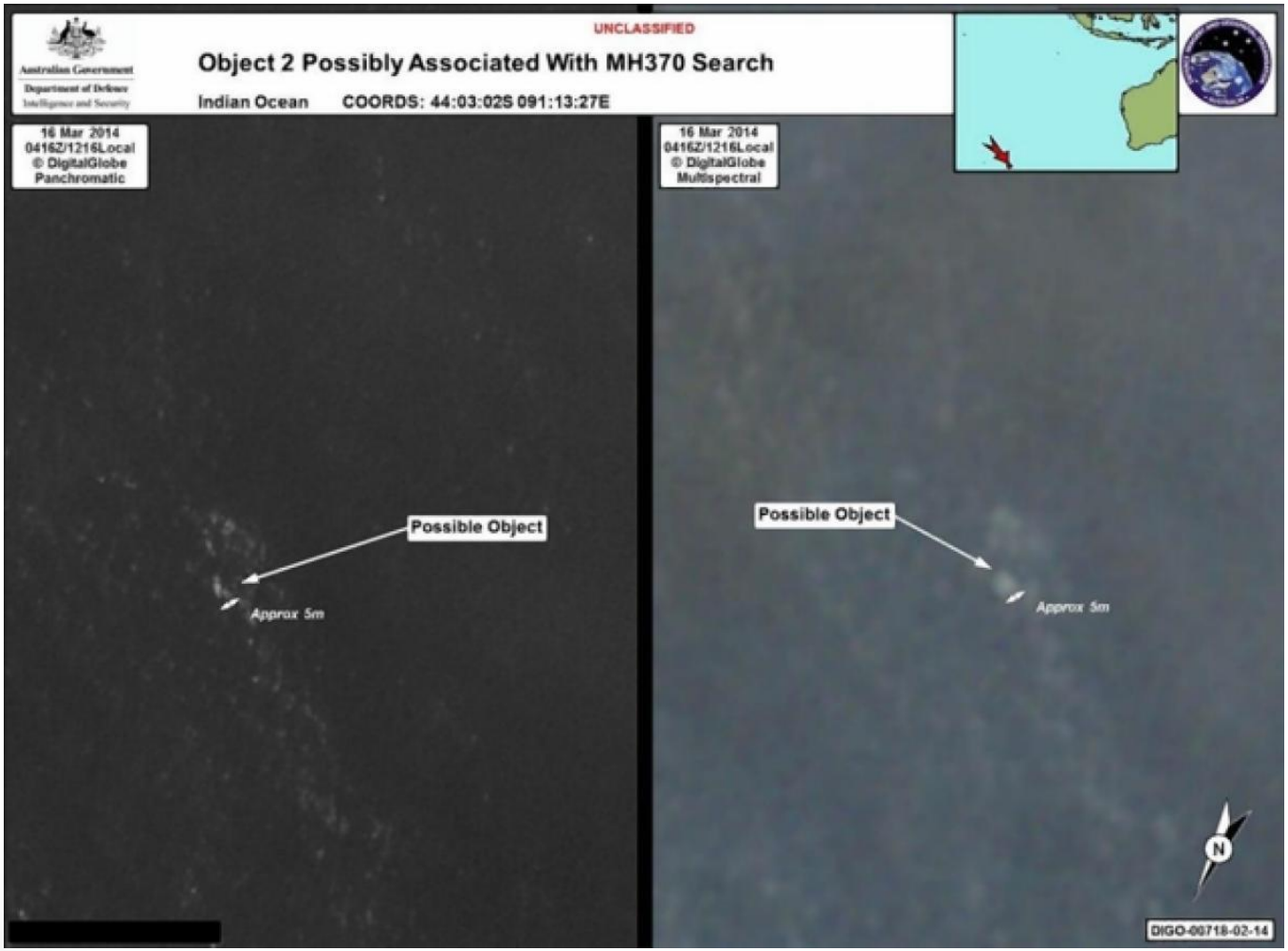
- Image processing:
  - Machine learning can be used in classification Images & objects in an image
    - Ship
    - Water
    - Rock
    - Iron object
    - Fiber Object etc.,
  - Does this really help?

Image processing Example



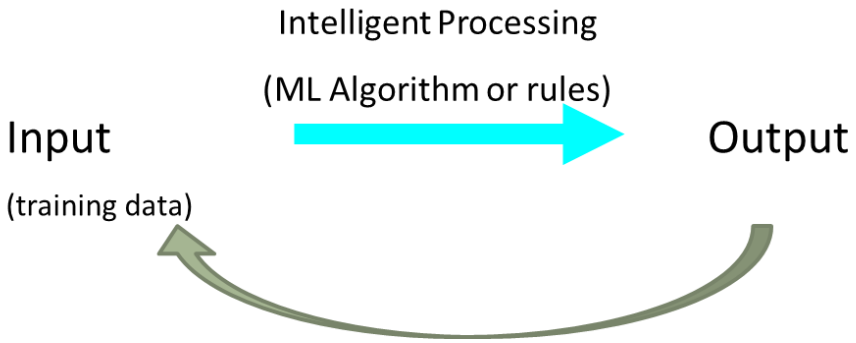


Machine learning application



The main advantage of ML

- Learning and writing an algorithm
  - Its easy for human brain but it is tough for a machine. It takes some time and good amount of training data for machine to accurately classify objects
- Implementation and automation
  - This is easy for a machine. Once learnt a machine can process one million images without any fatigue where as human brain can't
- That's why ML with bigdata is a deadly combination



Applications of Machine Learning

- Banking / Telecom / Retail
  - Identify:
    - Prospective customers
    - Dissatisfied customers
    - Good customers
    - Bad payers
  - Obtain:
    - More effective advertising
    - Less credit risk
    - Fewer fraud
    - Decreased churn rate

Applications of Machine Learning

- Biomedical / Biometrics
  - Medicine:
    - Screening
    - Diagnosis and prognosis
    - Drug discovery
  - Security:
    - Face recognition
    - Signature / fingerprint / iris verification
    - DNA fingerprinting

Applications of Machine Learning

- Computer / Internet
  - Computer interfaces:
    - Troubleshooting wizards
    - Handwriting and speech
    - Brain waves
  - Internet
    - Hit ranking
    - Spam filtering
    - Text categorization
    - Text translation
    - Recommendation

Main Parts in Machine Learning process

- Any Machine Learning algorithm has three parts
  1. The Output
  2. The Objective Function or Performance Matrix
  3. The Input

Email Spam Classification

- **The Output:** Categorize email messages as spam or legitimate.
- **Objective Function:** Percentage of email messages correctly classified.
- **The Input:** Database of emails, some with human-given labels

Hand Writing Recognition

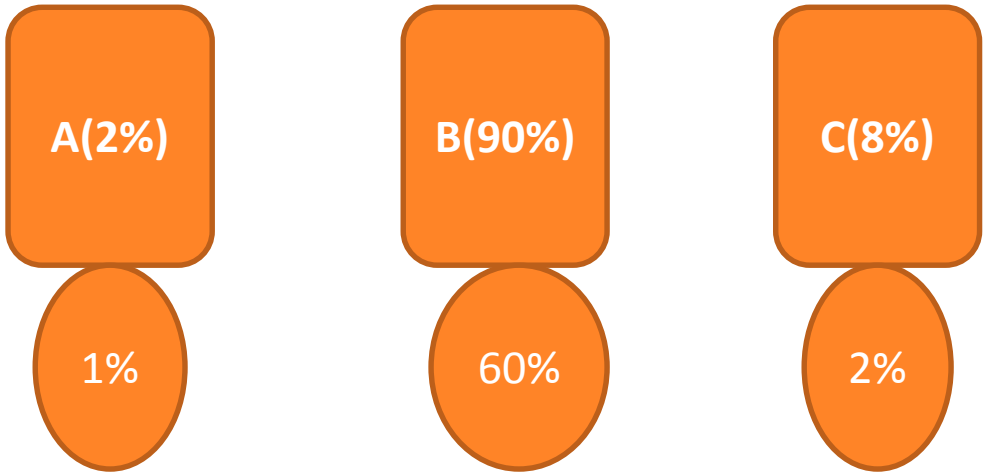
- **The Output:** Recognizing hand-written words
- **Objective Function:** Percentage of words correctly classified
- **The Input:** Database of human-labeled images of handwritten words

Some Machine Learning Methods

Everything should be made as simple as possible, but no simpler

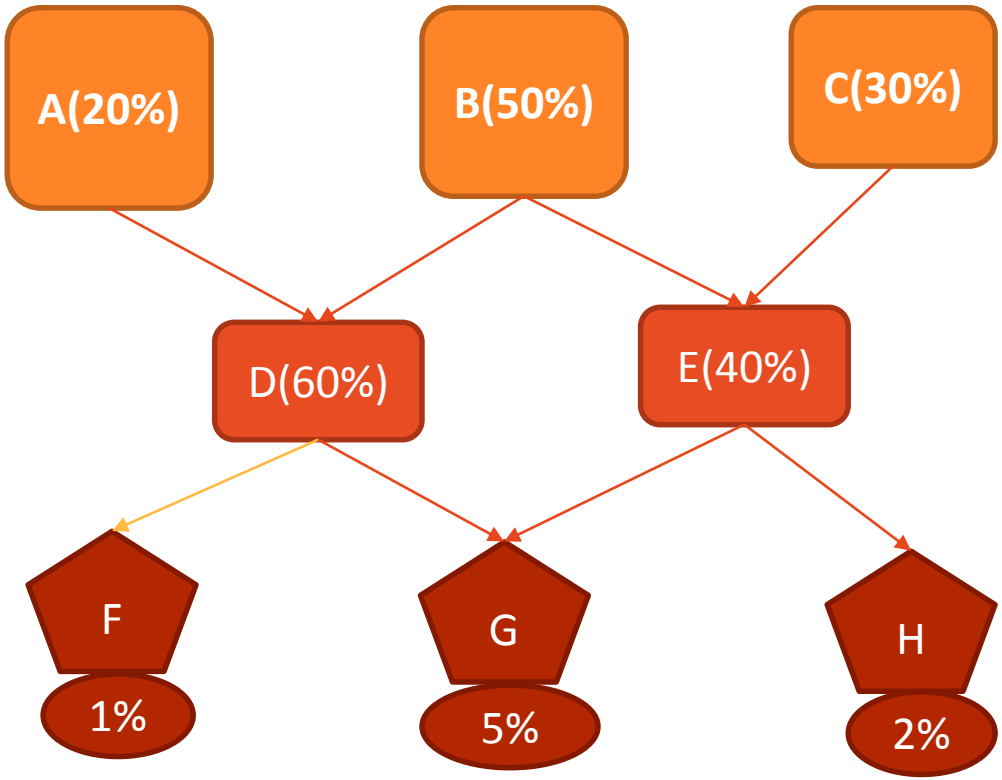
Bayes Network

The Bayes Method



- A, B, C Generate bolts, given a machine what is the probability that it will generate a defective bolt.
- Now given a defective bolt, can we tell which machine has generated it?

Several Machine in Several Layers



- Several intermediate machines
- Each has its on error chance, This problem requires a Bayesian network to solve
- Does C has anything to do if the bolt is finally coming out of F?

Artificial Neural Networks

The Philosophy of Neural Network

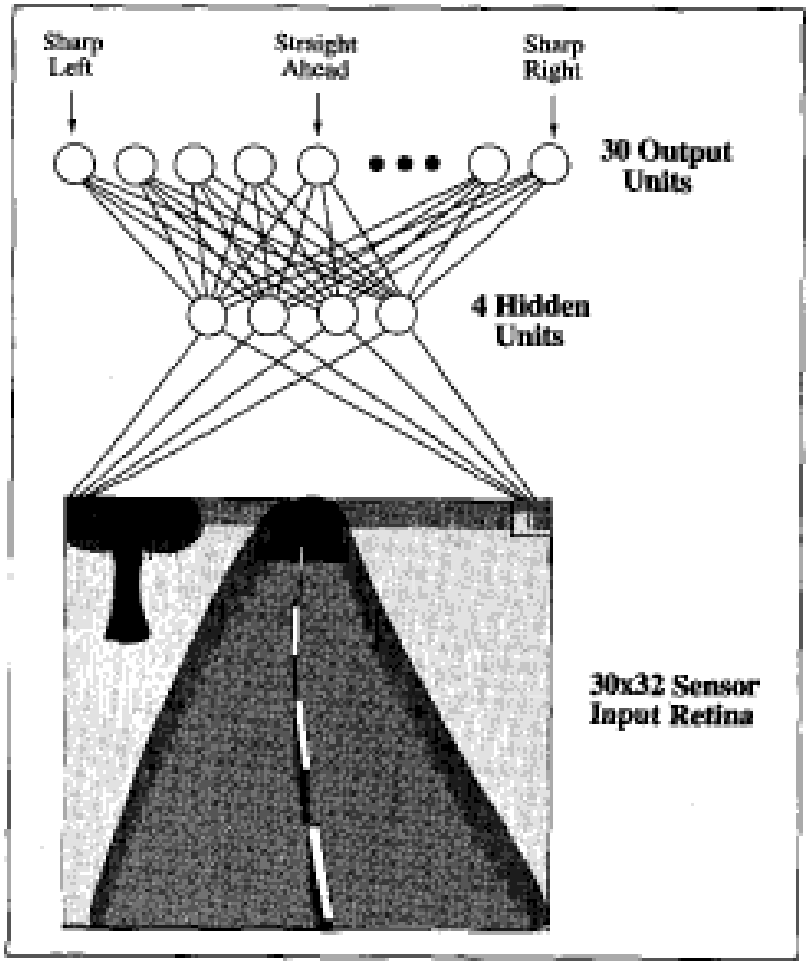
- If we have to predict whether we win in a chess game
  - Can we use logistic regression?
  - Since there can be almost infinite number of intermediate steps its almost impossible to predict the probability of winning
  - If each move is considered as an input variable then there are almost (countable)infinite moves. A coefficient for each of them is always negligible.

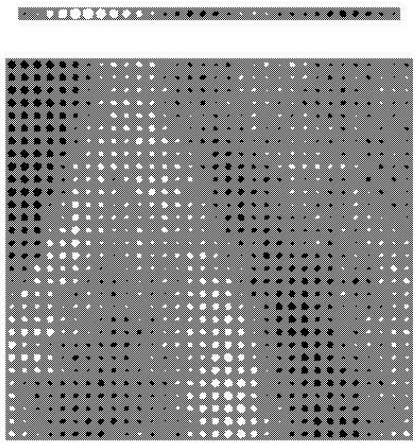
Neural network uses a simple technique, instead of predicting the final output they predict the intermediate output that lead to a win previously, this will be an input for predicting the next move

Biological Motivation for NN

- Aartificial neural networks are built out of a densely interconnected set of simple units, where each unit takes a number of real-valued inputs (possibly the outputs of other units) and produces a single real-valued output (which may become the input to many other units).
- The human brain is estimated to contain a densely interconnected network of approximately 10^{11} neurons, each connected, on average, to 10^4 others.
  - Neuron activity is typically inhibited through connections to other neurons.

ANN in Self Driving Car





PCA and FA

Look at below Cricket Team Players Data

Player	Avg Runs	Total wickets	Height	Not outs	Highest Score	Best Bowling
1	45	3	5.5	15	120	1
2	50	34	5.2	34	209	2
3	38	0	6	36	183	0
4	46	9	6.1	78	160	3
5	37	45	5.8	56	98	1
6	32	0	5.10	89	183	0
7	18	123	6	2	35	4
8	19	239	6.1	3	56	5
9	18	96	6.6	5	87	7
10	16	83	5.9	7	32	7
11	17	138	5.10	9	12	6

Describe the players

- If we have to describe or segregate the players do we really need Avg Runs, Total wickets, Height, Not outs, Highest Score, Best bowling?
- Can we simply take
  - Avg Runs+ Not outs+ Highest Score as one factor?
  - Total wickets+ Height+ Best bowling as second factor?

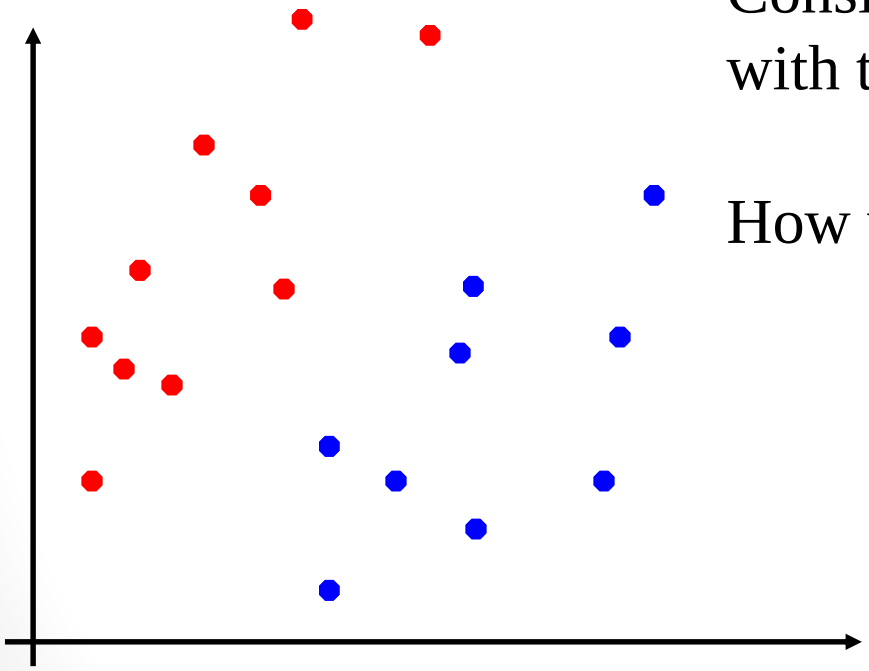
Defining these imaginary variables or a linear combination of variables to reduce the dimensions is called PCA or FA

Look at below Cricket Team Players Data

Player	Avg Runs	Total wickets	Height	Not outs	Highest Score	Best Bowling
1	45	3	5.5	15	120	1
2	50	34	5.2	34	209	2
3	38	0	6	36	183	0
4	46	9	6.1	78	160	3
5	37	45	5.8	56	98	1
6	32	0	5.10	89	183	0
7	18	123	6	2	35	4
8	19	239	6.1	3	56	5
9	18	96	6.6	5	87	7
10	16	83	5.9	7	32	7
11	17	138	5.10	9	12	6

SVM Classification

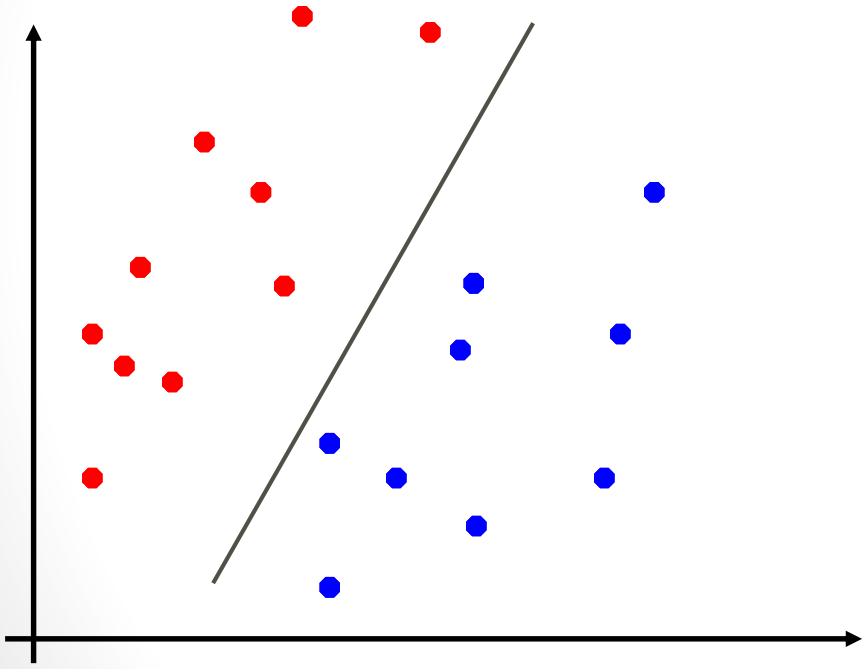
Linear Classifier

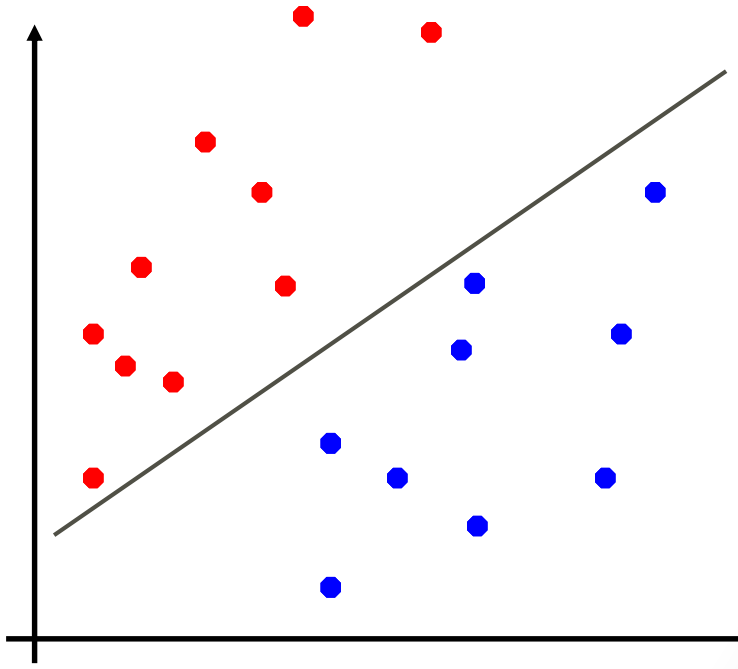


Consider a two dimensional dataset with two classes

How would we classify this dataset?

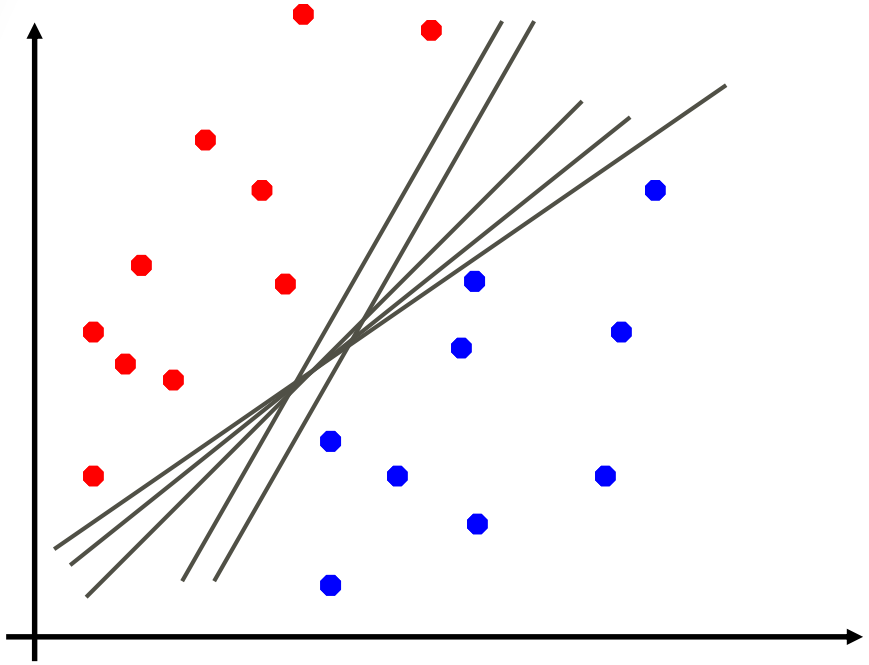
Linear Classifier





Both of the lines can be linear classifiers.

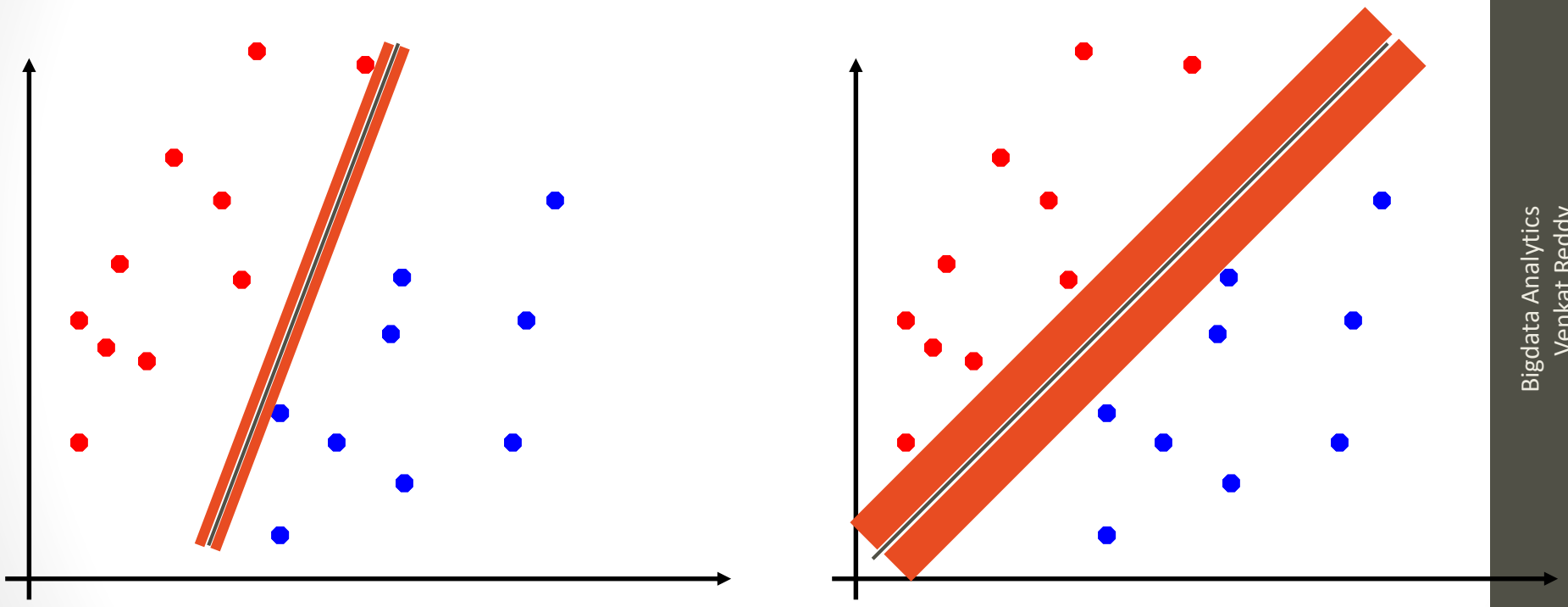
Linear Classifier



There are many lines that can be linear classifiers.

Which one is the optimal classifier.

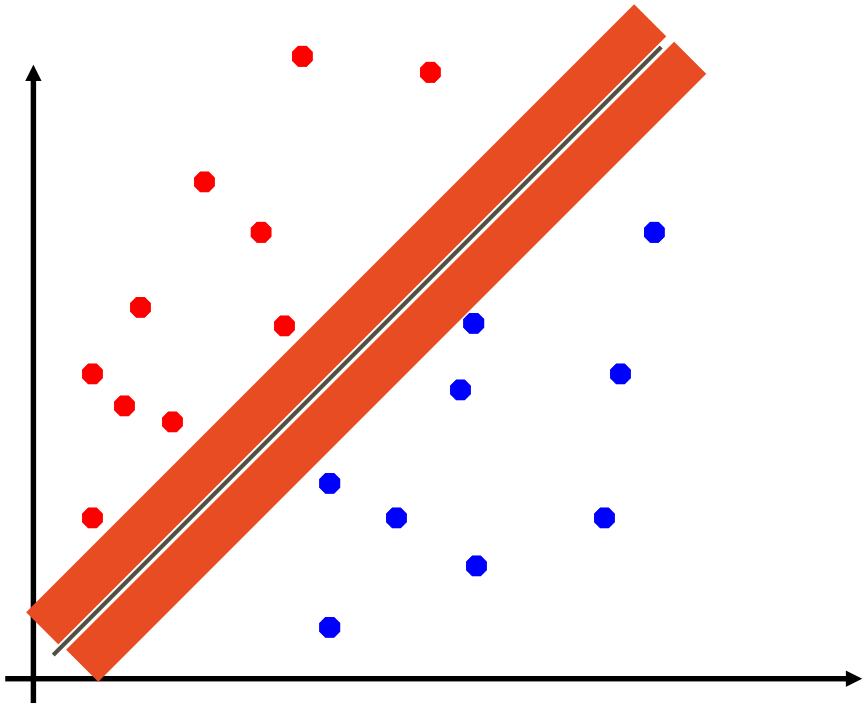
Classifier Margin



Define the **margin** of a linear classifier as the width that theboundary could be increased by before hitting a datapoint.

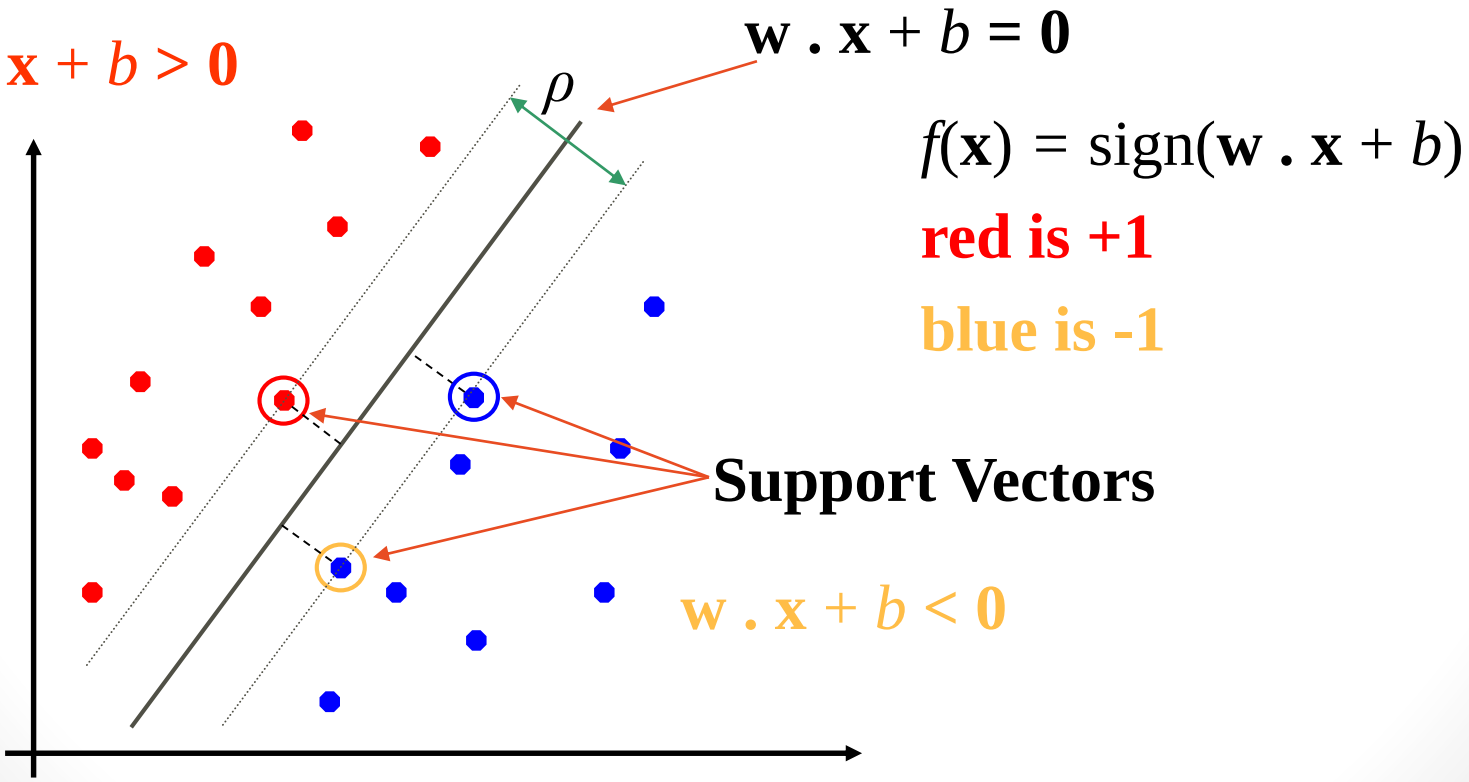
Maximum Margin

- The **maximum margin linear classifier** is the linear classifier
- with the maximum margin.
- This is the simplest kind of SVM (Called Linear SVM)



Support Vectors

- Examples closest to the hyper plane are ***support vectors***.
- ***Margin*** ρ of the separator is the distance between support vectors.

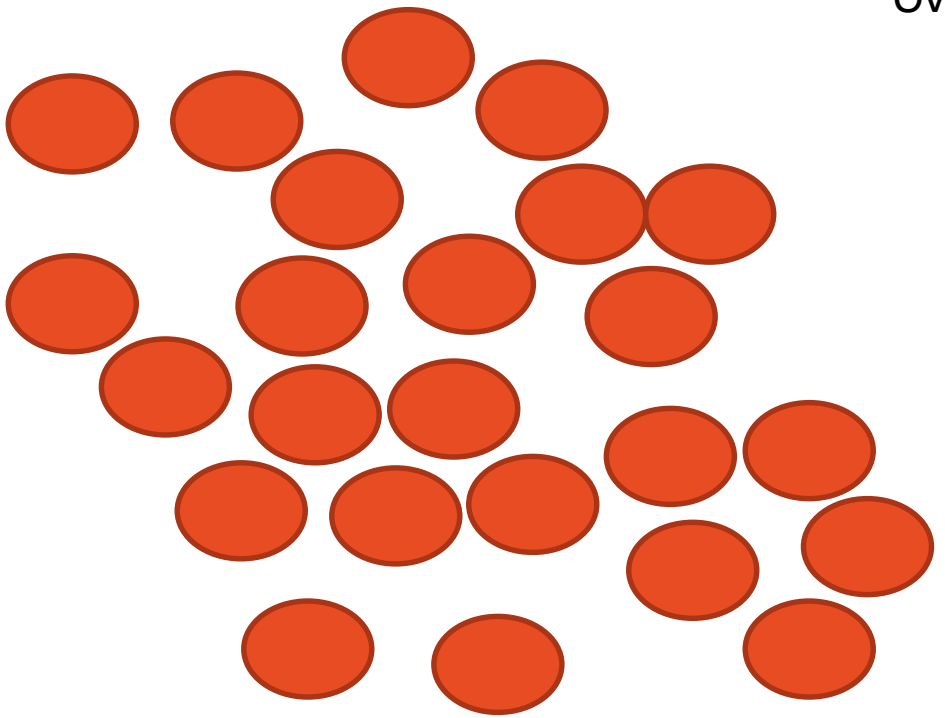


Difference Traditional Models & Machine Learning

- ML is more heuristic
- Focused on improving performance of a learning agent
- Also looks at real-time self learning and robotics – areas not part of data mining
- Some algorithms are too heuristic that there is no one right or wrong answer
- Lets take K-Means example

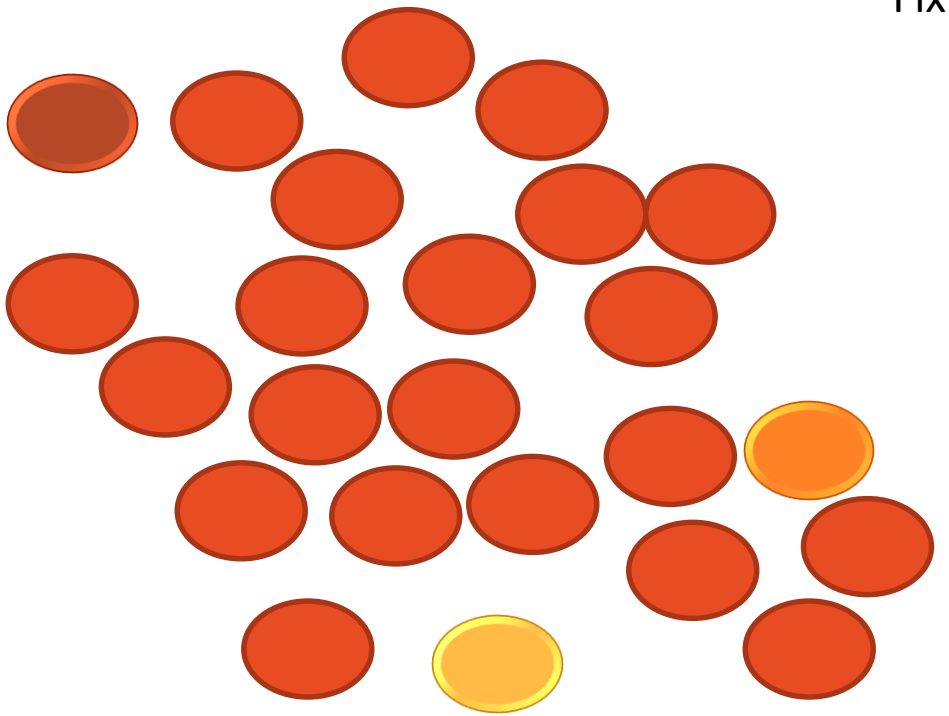
K-Means clustering

Overall population

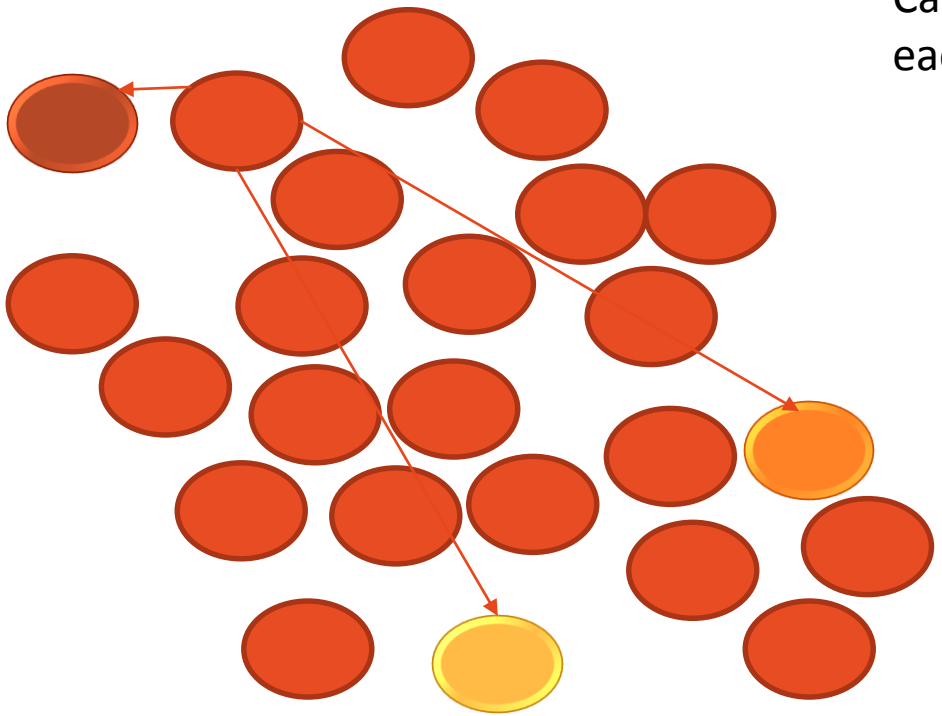


K-Means clustering

Fix the number of clusters

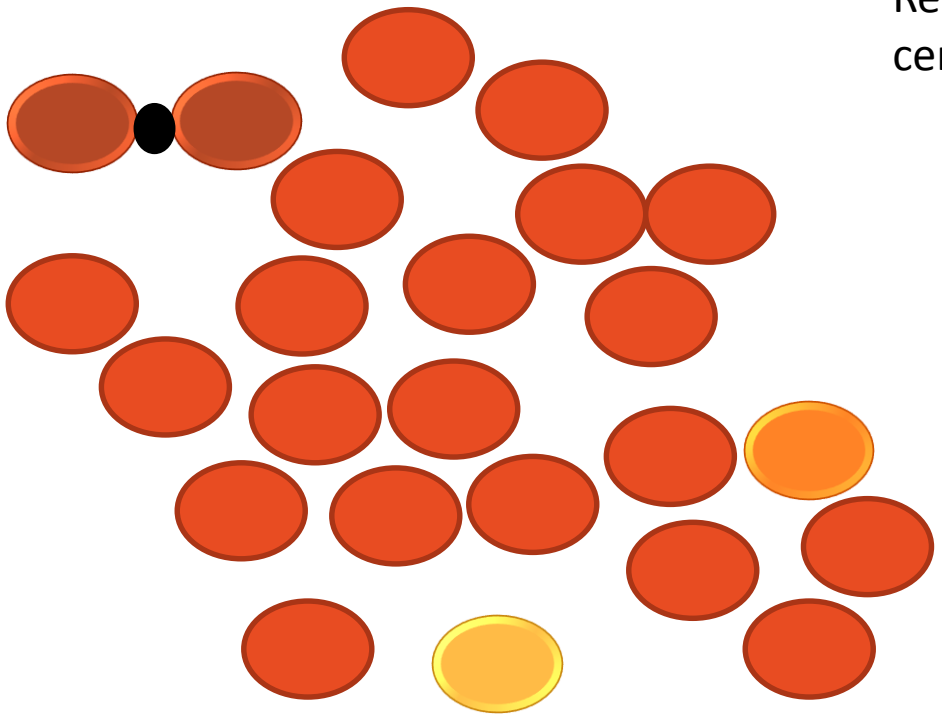


K-Means clustering



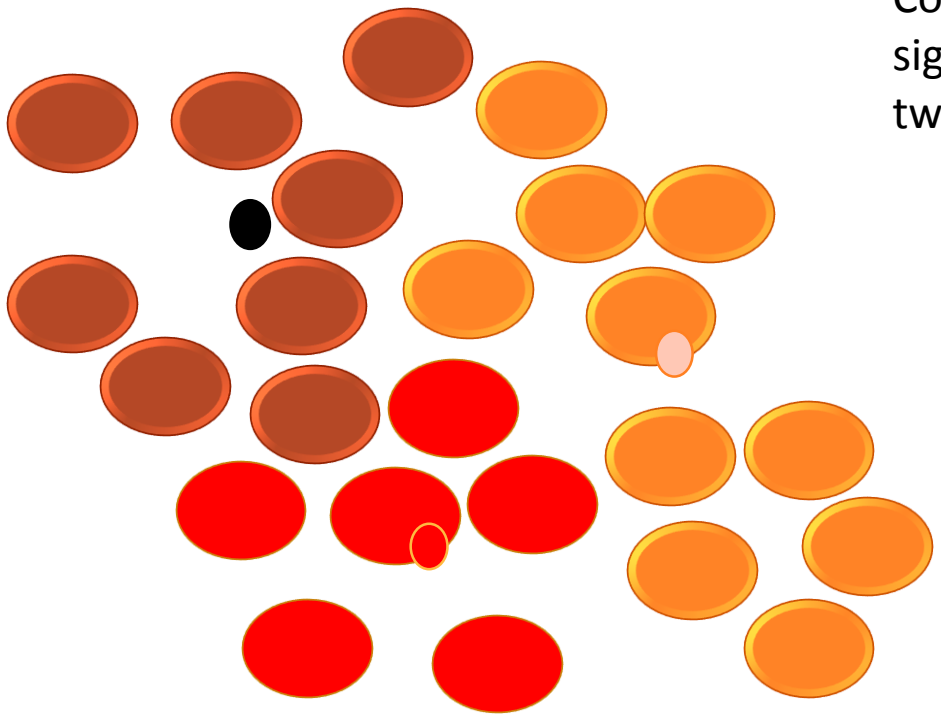
Calculate the distance of each case from all clusters

K-Means clustering



Re calculate the cluster centers

K-Means clustering



Continue till there is no significant change between two iterations

Calculating the distance

	Weight
Cust1	68
Cust2	72
Cust3	100

Which of the two customers are similar?

	Weight	Age
Cust1	68	25
Cust2	72	70
Cust3	100	28

Which of the two customers are similar now?

	Weight	Age	Income
Cust1	68	25	60,000
Cust2	72	70	9,000
Cust3	100	28	62,000

Which two of the customers are similar in this case?

Distance Measures

- Euclidean distance
- City-block (Manhattan) distance
- Chebychev similarity
- Minkowski distance
- Mahalanobis distance
- Maximum distance
- Cosine similarity
- Simple correlation between observations
- Minimum distance
- Weighted distance
- Venkat Reddy distance

- Not sure all these measures will result in same clusters in the above example.
- So, same ML algorithms with different results. Generally this is not the case with conventional methods

Thank you

More ..

http://www.slideshare.net/21_venkat/presentations

https://www.facebook.com/groups/SASanalysts/