Machine learning.

- Basics
- Metrics
  - Accuracy
  - confusion matrix
  - Precision
  - Recall
  - Specificity

iii

- Support vector Machine
- K- Nearest Neighbour
- Random Forest
- Neighbour based.

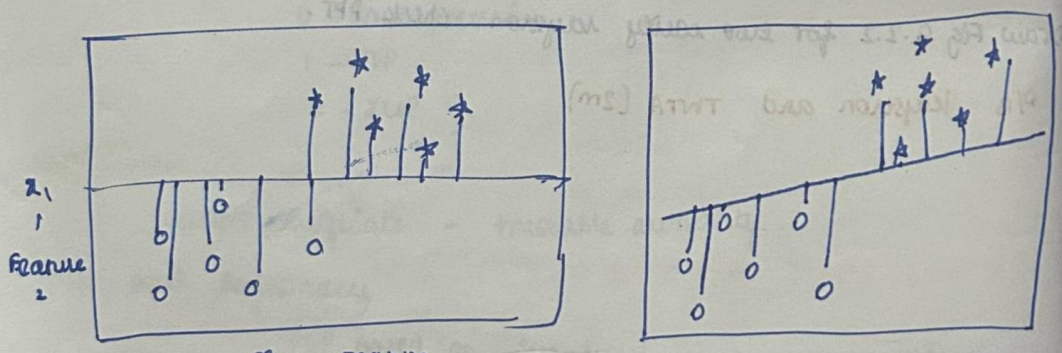
iv

- K- means clustering

v

- Neural networks.

iii unit SVM

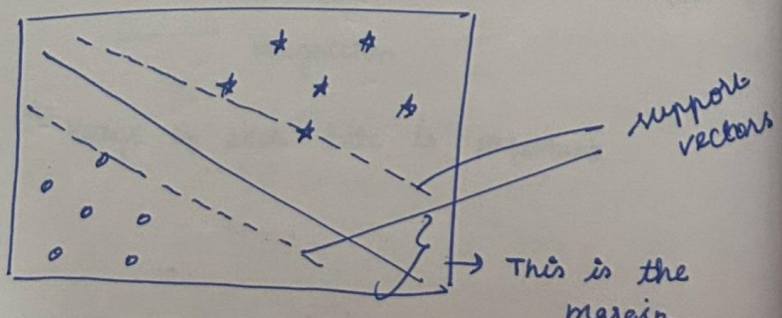


The output lies in the type.

Generate a model that can classify the given data points into 2 classification

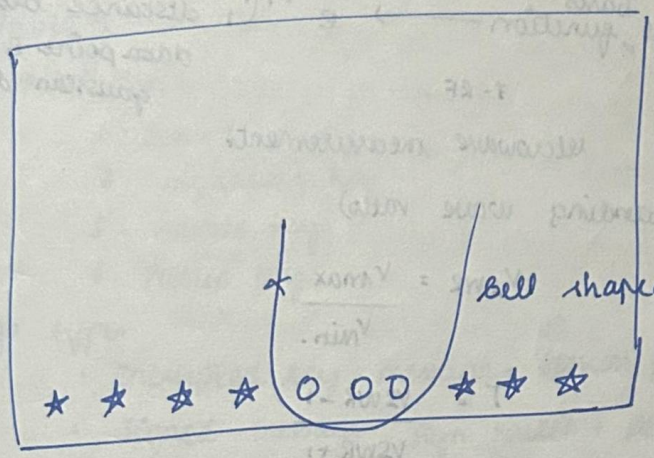
$$y = x_0 + mx_1$$

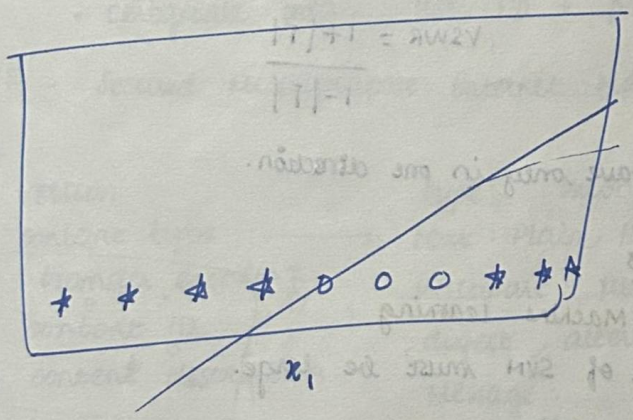
SVM forms a line such that there is a street between the two types



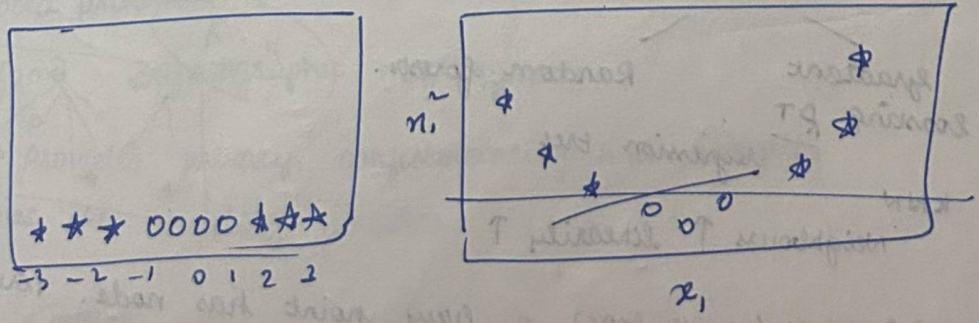
• Linearly seperable - A linear line must be able to distinguish the teppes.

• SVM is best applicable for linearly seperable data.





like there is only 1 feature. A single line cannot be drawn to seperate those. Hence we are ~~draw~~ creating another feature x_1^2 / x_1^3 / .... any polynomial.



Thus we can convert the 1D data into many dimensional data. It is called as Cornell Trick.

- Hard margin → Does not allow any data points to be within the street.
- Soft margin ↓ will allow data points to be present within the street (some).

support vectors - the points close to the margin

- Gaussian radial basis function ⟶ $e^{-\gamma|\cdot|^2}$ ↳ distance between data point & gaussian data.

7-RF

Microwave measurements

VSWR (voltage standing wave ratio)

$$V_{SWR} = \frac{V_{max}}{V_{min}}$$

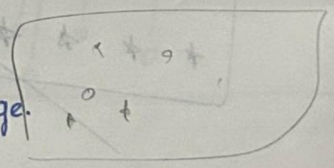
$$\Gamma = \frac{VSWR - 1}{VSWR + 1}$$

$$VSWR = \frac{1+|\Gamma|}{1-|\Gamma|}$$

Isolator - sends wave only in one direction.

Friday, Nov 3 2023

Machine Learning



NOTE: The margin of SVM must be large.

SVC - function in python

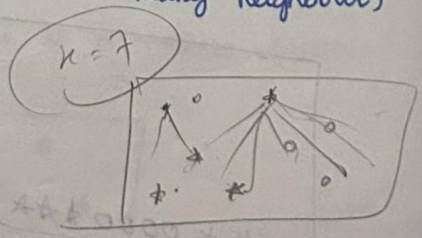
C ↑ margin ↓

When C is increased, data points can lie inside margin

① K-Nearest Neighbours ⟶ K - integer (how many neighbour)

② Decision tree

Gradient boosting DT — regression tree — Random forests.



KNN

Neighbours ↑ linearity ↑

Decision tree (miller book) → every point has node. Terminal node - leaf.

- features

Types of msg

- Signed data content → provides integrity
- Enveloped data content → provides privacy
- Digested data content
  - both original data and digest will be stored. This also provides integrity.
- Authenticated data content.
  - encryption algorithm must be applied.
  - MAC algorithm

Data $\xrightarrow[\text{algorithm}]{\text{MAC}}$ MAC function

Key management - X.509 and PGP key management. gives the trust level.

ML

Random tree

The data set is randomly picked and many trees are formed.

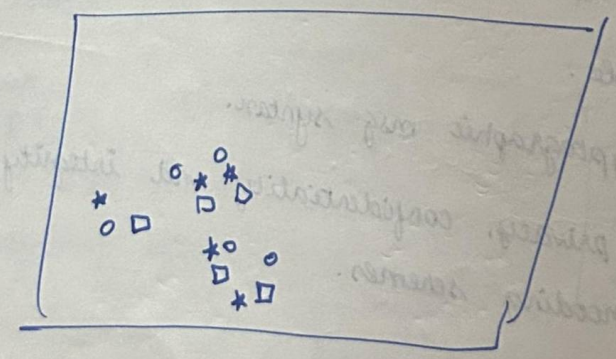
"Ensemble - collection of"

Gradient boosting - sequential tree growth
Random forest - Parallel tree growth.

K-means clusters (Hyperparameters)
- Attributes
- No. of clusters.

- Choose how many clusters are needed
- Then choose 3 random points.

Can find spherical clusters.



Read limitations

4th November 2023 Machine learning

DB Scan - Density based spatial clustering for applications with Noise

Two major parameter
- Epsilon / radius / Minimum ~~no~~ of distance
- Minimum no. of. samples

core point — min. no. of samples exist
Border point - min no. of samples doesn't exist
Noise - no samples.

Top to bottom model.

Agglomerative - bottom to top model

gaussian mixture's model (GMM)

Gaussian distribution formula / Normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left[\frac{(x-\mu)}{\sigma}\right]^2}$$

σ - variance
μ - mean.

when graph is drawn b/w x and f(x) we get a bell shaped curve.

→ when the data points are distributed elliptically we can apply this.

It is assumed that the clusters formed here are drawn using gaussian distribution model.

- clustering used for image segmentation.
- MNIST Dataset.

clustering will be helpful in preprocessing

Unit 5 - convolutional Neural

Security at the network layer: IPsec - designed by IETF. (3rd Layer)

2 modes

- Transport mode - doesn't protect IP header.
- Tunnel mode (encapsulation)

Provides security only to data travelling from transport mode to network layer layer.

Tunnel mode → protects IP header.

Two security protocol

- AH - Authentication Header
- ESP - Encapsulating Security Payload.

integrity, authentication but no privacy

integrity, authentication privacy all 3 are there.

Header has

- Next header
- Payload header
- Reserved.

IPsec has replay attack protection.

Security Association → gives logical relationship b/w 2 hosts.

There will be 2 SA

- Outbound SA
- Inbound SA

It has security association database.

- Security policy determines who to allow and who not to allow.

ISAKMP - Internet Security Association and Key Management Protocol. → 32 bits long

Nonce - random generated numbers

MD5 encryption

Machine learning

Agglomerative

- When two clusters are joining, the extreme points of both's distance must be minimized.
- Maximum distance / average distance must be minimized.
- When the variance of two clusters are similar then can be joint together.

Dendogram - visualizing heirarchical clustering

. min max scaler. → $\frac{x - \text{max value}}{[\text{max value} - \text{min value}]}$ → known as range

. Normalization :

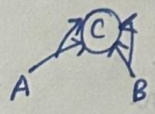
Here the numbers will be substracted from the mean of them. When the obtained ones are averaged we will get 0 and variance will get 1.

Unit 5

Neural Networks.

Buffer neural network

OR neural network



AND neural network

perceptron → will have only 2 logical unit layer.

. activation func : This determines whether neuron can will be activated / or not.