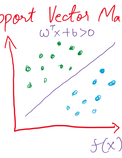
SVM, Decision Trees and Ensemble Methods

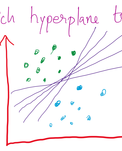
Saturday, 15 March 2025 7:27 PM

Support Vector Machines



$f(x) = sign(w^Tx+b)$

Which hyperplane to choose?



Margin

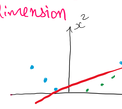
→ SVM is a maximum margin classifier

→ Issues with SVM
- → Linear Separability
- → Hard v/s soft Margin SVM

→ The Kernel trick

Not Linearly Separable Data

→ However take it to a higher dimension



High Level Idea:
Original Feature Space Can Always be Mapped to some higher dimensional space where training data is separable. Kernel Trick Ensures you do not have to explicitly take data to higher dimensions.

Decision Trees

Whether a customer will buy an apple product?

→ Decision trees are very intuitive classifiers

→ They are explainable & can work with categorical variables/features too

Points to Consider

Which feature to be used for split?
- → Gini Index
- → Entropy

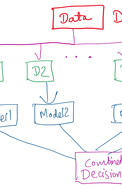
How much should be the depth of the tree
- → Stop Early
- → Prune

Ensemble Methods:-

→ Single Classifier may not be good

→ Use Multiple Classifiers like a committee of experts

Bagging: Bootstrap Aggregation



When the Model used is a decision tree, it is called a Random Forest Classifier.

Boosting:-

→ Combine "Weak" Learners to form a strong learner

→ Popular Boosting Techniques
Ada Boost Eg Grad boost