



# Supervised Learning

# Supervised Learning



- In Supervised learning, you train the machine using data which is well "labeled."
- A supervised learning algorithm learns from labeled training data, helps you to predict outcomes for unforeseen data
- For example, you want to train a machine to help you predict how long it will take you to drive home from your workplace. Here, you start by creating a set of labeled data. This data includes:
  - Weather conditions
  - Time of the day
  - Holidays
- All these details are your inputs. The output is the amount of time it took to drive back home on that specific day.

# TYPES OF SUPERVISED LEARNING

## Regression:

Regression technique predicts a single output value using training data.

**Example:** You can use regression to predict the house price from training data. The input variables will be locality, size of a house, no of rooms etc.

**Strengths:** Outputs always have a probabilistic interpretation, and the algorithm can be regularized to avoid overfitting.

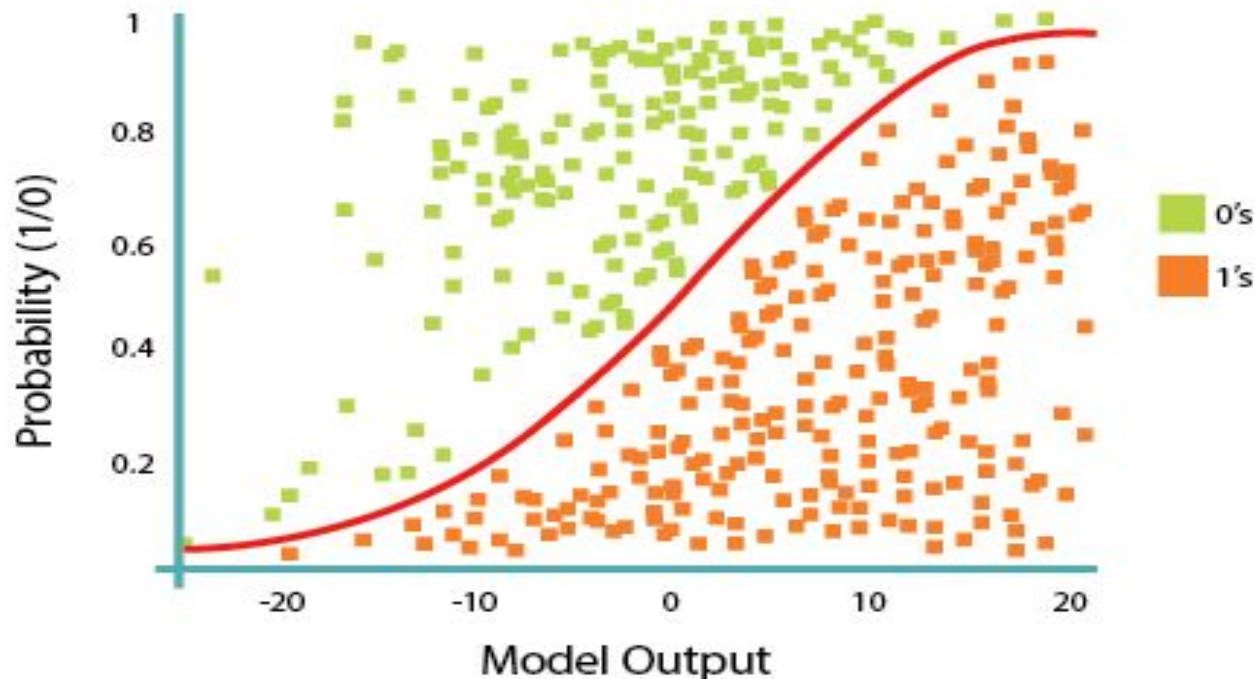
**Weaknesses:** Logistic regression may underperform when there are multiple or non-linear decision boundaries. This method is not flexible, so it does not capture more complex relationships.



# Logistic Regression



- Logistic regression method used to estimate discrete values based on given a set of independent variables.
- It helps you to predicts the probability of occurrence of an event by fitting data to a logit function. Therefore, it is also known as logistic regression.
- As it predicts the probability, its output value lies between 0 and 1.



# Naïve Bayes Classifiers



Naïve Bayesian model (NBN) is easy to build and very useful for large datasets.

- It is easy and fast to predict class of test data set. It also perform well in multi class prediction
- Naive Bayes classifier performs better compare to other models like logistic regression when you have less training data.
- It perform well in case of categorical input variables compared to numerical variable(s).

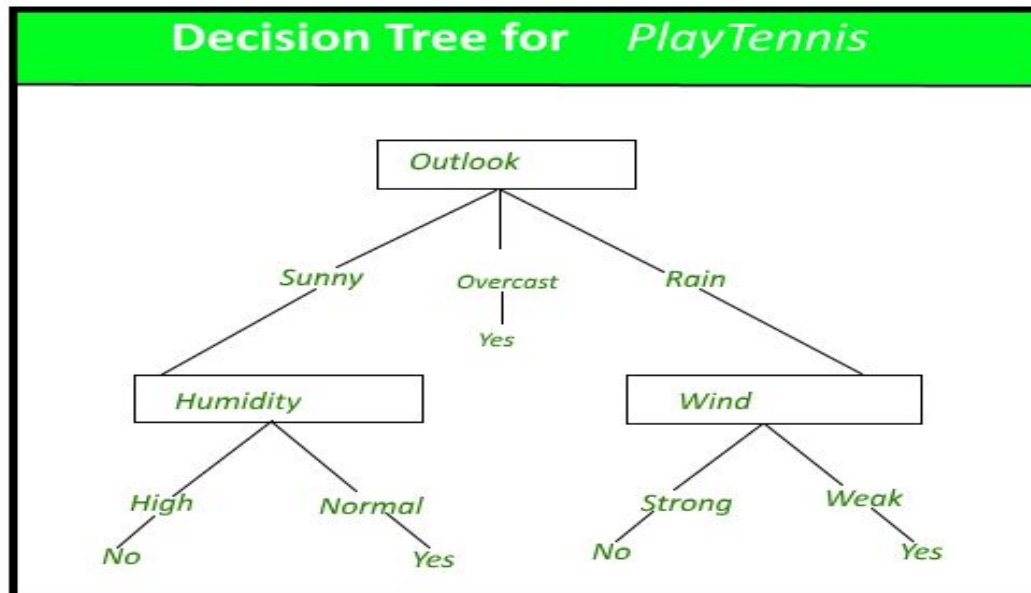
# Applications of Naive Bayes Algorithms

- **Real time Prediction:** Naive Bayes is a learning classifier and it is fast. Thus, it could be used for making predictions in real time.
- **Multi class Prediction:** This algorithm is also well known for multi class prediction feature.
- **Text classification/ Spam Filtering/ Sentiment Analysis:** Naive Bayes classifiers mostly used in text classification (due to better result in multi class problems and independence rule) have higher success rate as compared to other algorithms.
- As a result, it is widely used in Spam filtering (identify spam e-mail) and Sentiment Analysis (in social media analysis, to identify positive and negative customer sentiments)
- **Recommendation System:** Naive Bayes Classifier and Collaborative Filtering together builds a Recommendation System that uses machine learning and data mining techniques to filter unseen information and predict

# Decision Trees



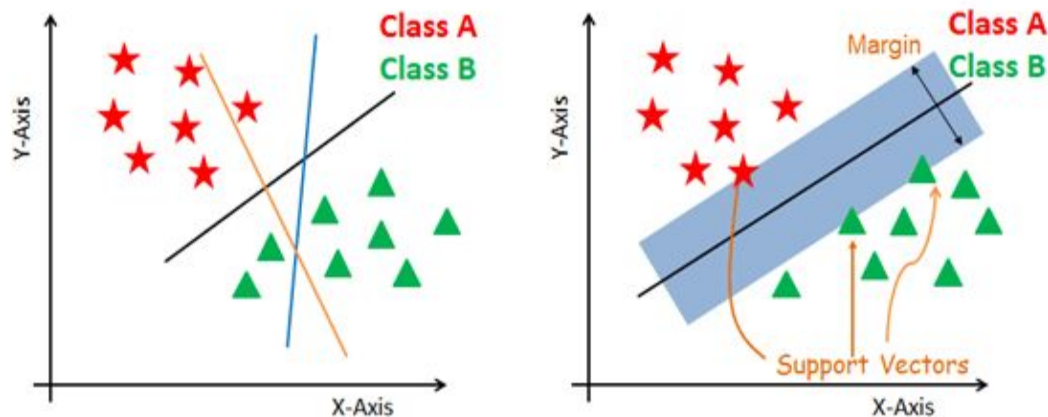
- Decisions trees classify instance by sorting them based on the feature value.
- In this method, each mode is the feature of an instance. It should be classified, and every branch represents a value which the node can assume. It is a widely used technique for classification.
- It helps you to estimate real values (cost of purchasing a car, number of calls, total monthly sales, etc.).



# Support Vector Machine



- Support vector machine (SVM) is a type of learning algorithm is based on results from statistical learning theory.
- SVM machines are also closely connected to kernel functions which is a central concept for most of the learning tasks.
- The kernel framework and SVM are used in a variety of fields. It includes multimedia information retrieval, bioinformatics, and pattern recognition.





# Supervised Learning- Summary

## Advantages

- Supervised learning allows you to collect data or produce a data output from the previous experience
- Helps you to optimize performance criteria using experience
- Supervised machine learning helps you to solve various types of real-world computation problems.

## Disadvantages

- Decision boundary might be over trained if your training set which doesn't have examples that you want to have in a class
- You need to select lots of good examples from each class while you are training the classifier.
- Classifying big data can be a real challenge.
- Training for supervised learning needs a lot of computation time.





Q.1 Explain Supervised Learning in ML?

Q.2 Explain the various types of Supervised Learning in ML?

Q.3 Explain Classification in Supervised Learning ?

Q.4 Explain Naïve Bayes Classifiers with its applications?

Q.5 Define Decision trees with examples?

Q.6 What is SVM classifier. What is its working?

Q.6 What are the advantages and disadvantages of Supervised Learning ?