

Machine Learning

Lecture 1

By

Venkatesh.B IH201685013

Vikas.R IH201685100

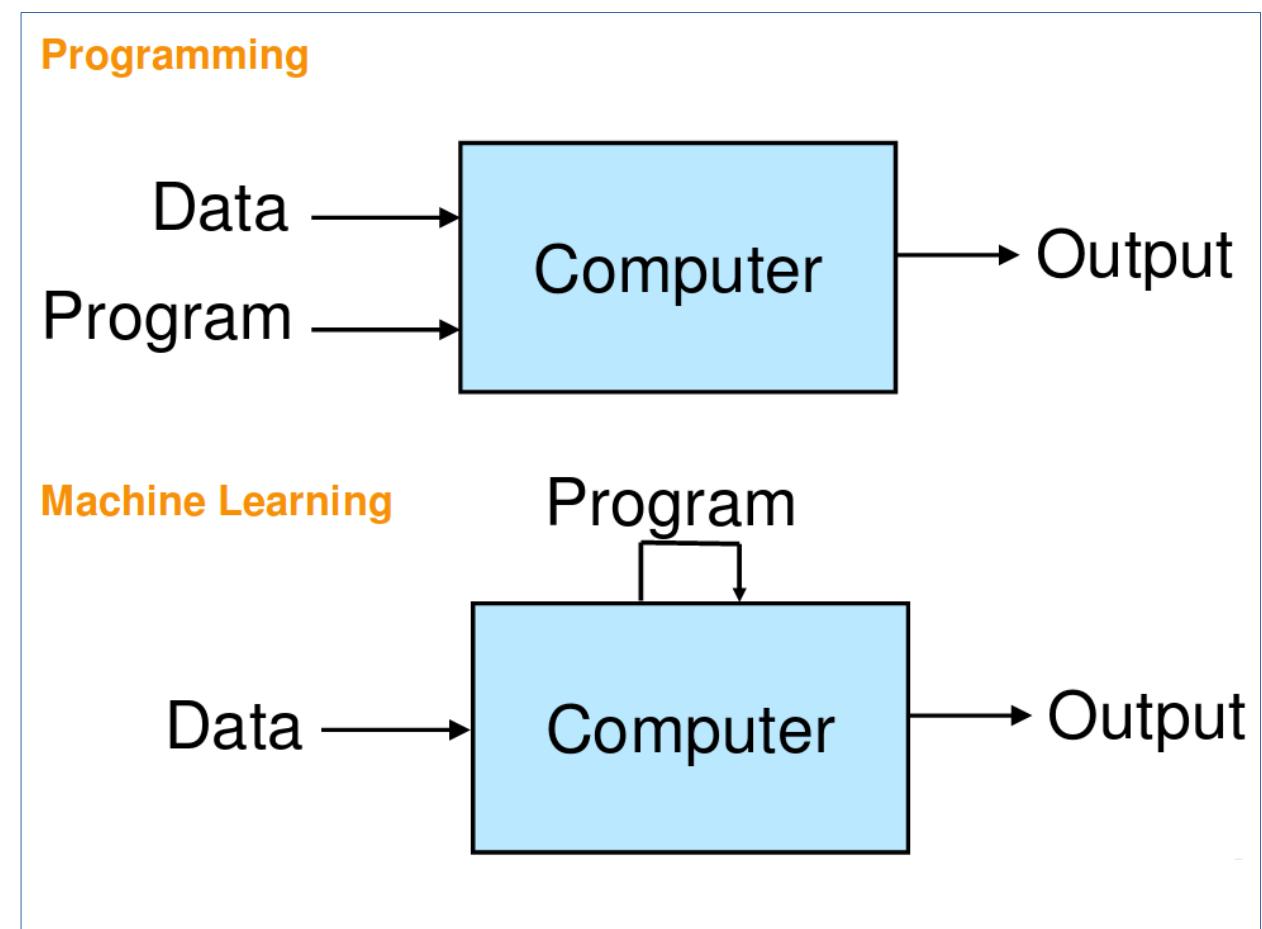
Venkatesh.S IH201685090

Generalizing from examples

- ML model's ability to perform well on new unseen data rather than just the data that it was trained on.
- It is related to the concept of overfitting. If the model is overfitted then it will not generalize well.

Automatically learn programs from data

- Algorithms that can improve performance using training data.
- Typically, a large number of parameter values learned from data.
- Applicable to situations where challenging to define rules manually.



Applications

- Web Search
- Spam filters
- Recommender Systems
- Ad Placement
- Credit Scoring
- Fraud Detection
- Stock Trading
- Drug Design

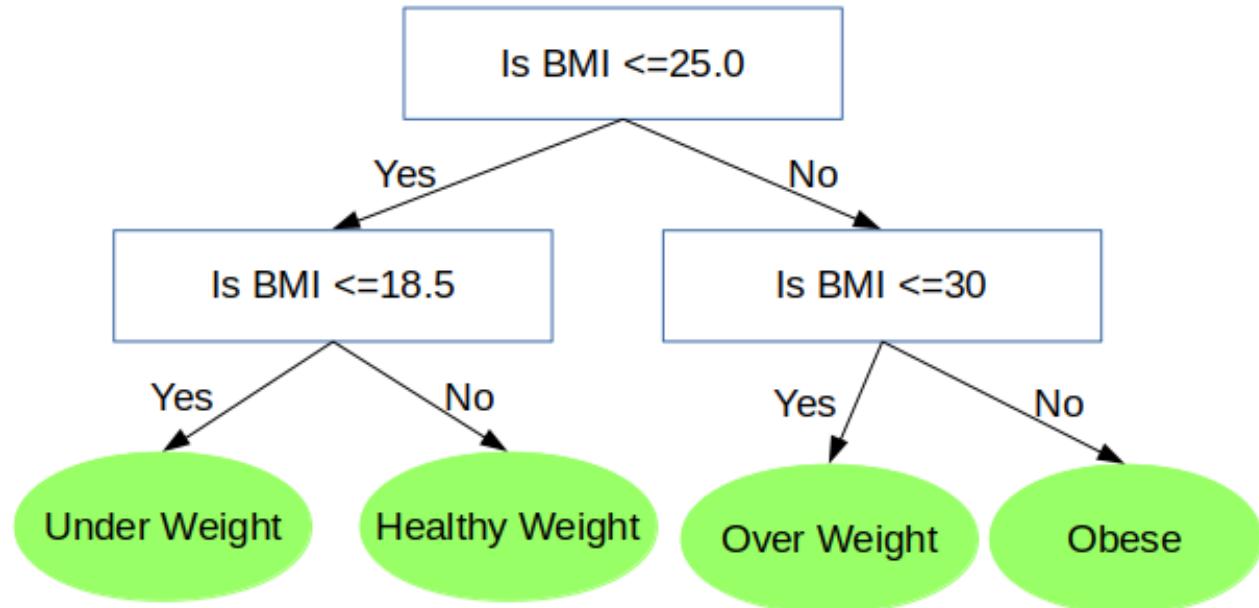
Classification

- Identifying to which of a set of categories a new observation belongs, on the basis of a training set of data containing observations whose category membership is known.
- Consider a scenario where an IIM aspirant wants to join IIM Ahmadabad, let's see whether the person gets selected or not.
- The deciding factors are CAT Score, Interview Score, Work Experience.

Decision Trees

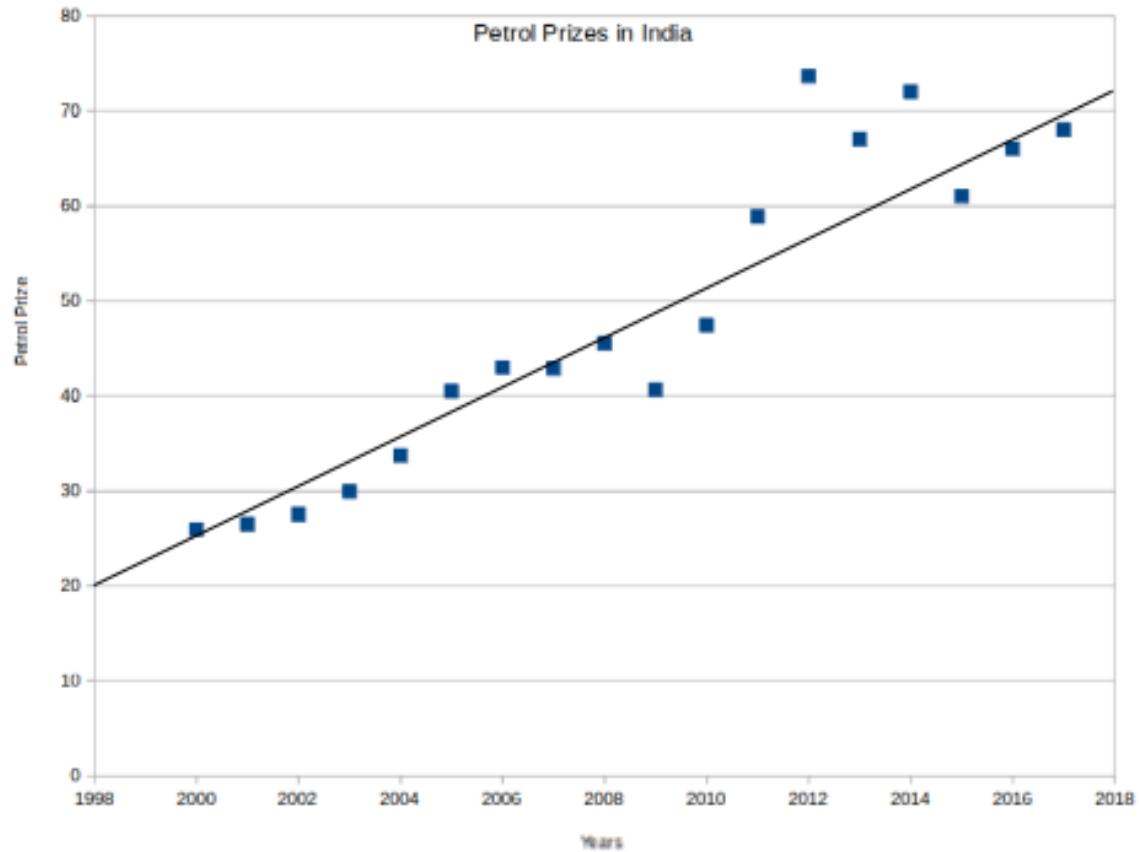
- A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility.

Calculating Obesity Level using BMI



Linear Regression

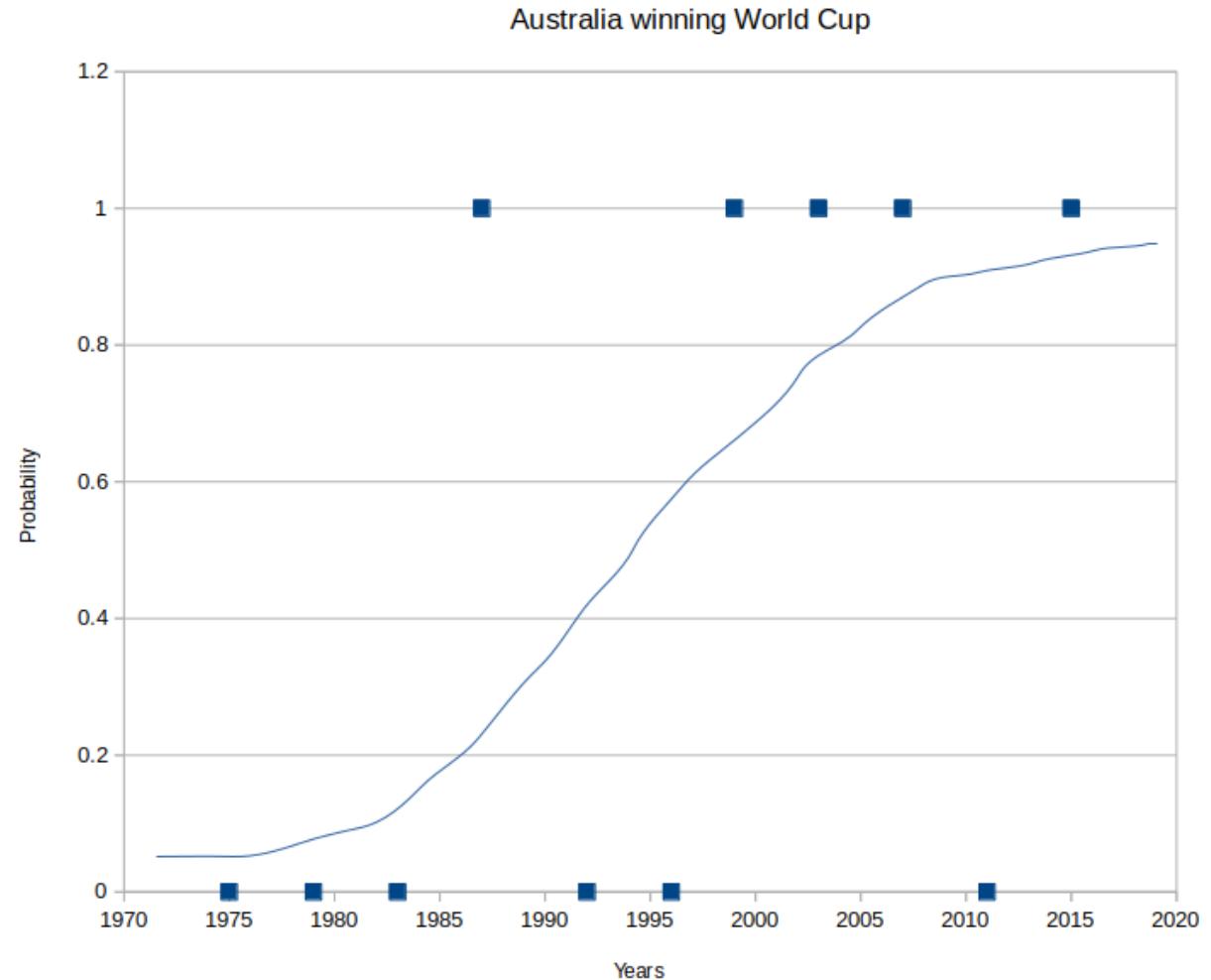
- Linear approach for modeling the relationship between a scalar dependent variable y and one or more explanatory variables denoted X .
- Consider the petrol prizes(x-axis) corresponding to each year (y-axis) in India.



Logistic Regression

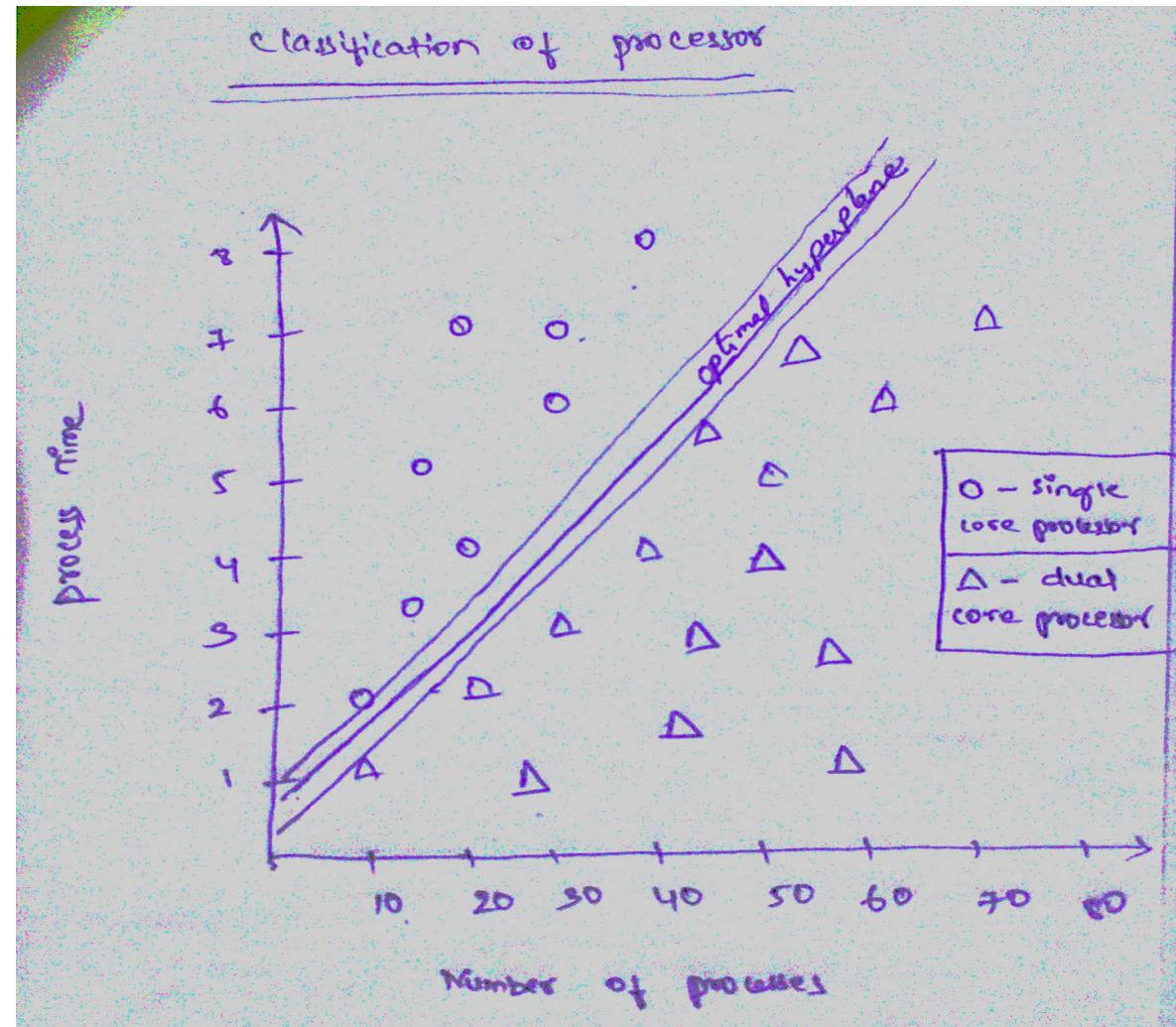
- It is a regression model where the dependent variable is categorical.
- Probability of Australia winning world Cup.
- The deciding factors are Favouring Pitches, Earlier track record- Wins,loses, Years, Players Performance.

Ref:https://en.wikipedia.org/wiki/Cricket_World_Cup



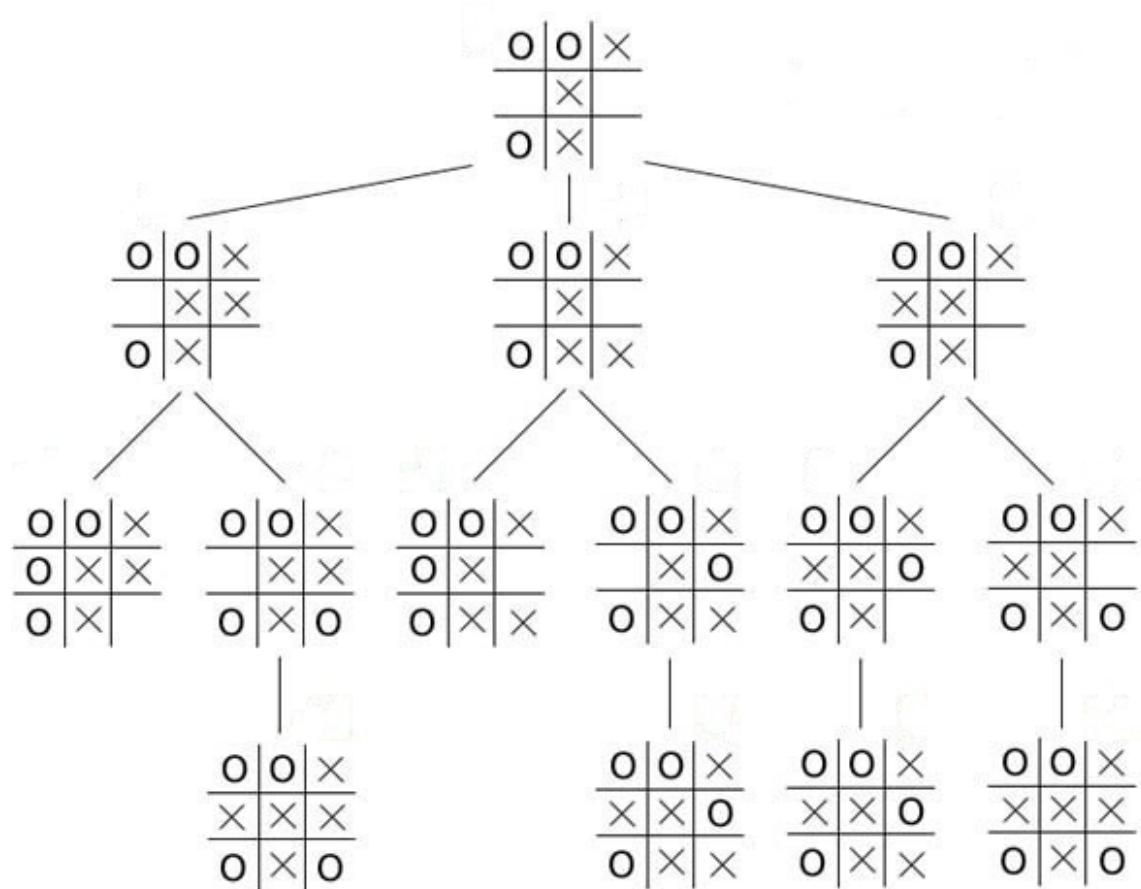
Support Vector Machine

- A discriminative classifier formally defined by a separating hyperplane. Given labeled training data the algorithm outputs an optimal hyperplane which categorizes new examples.
- Consider example of processes and its process times which classifies the core of processor.



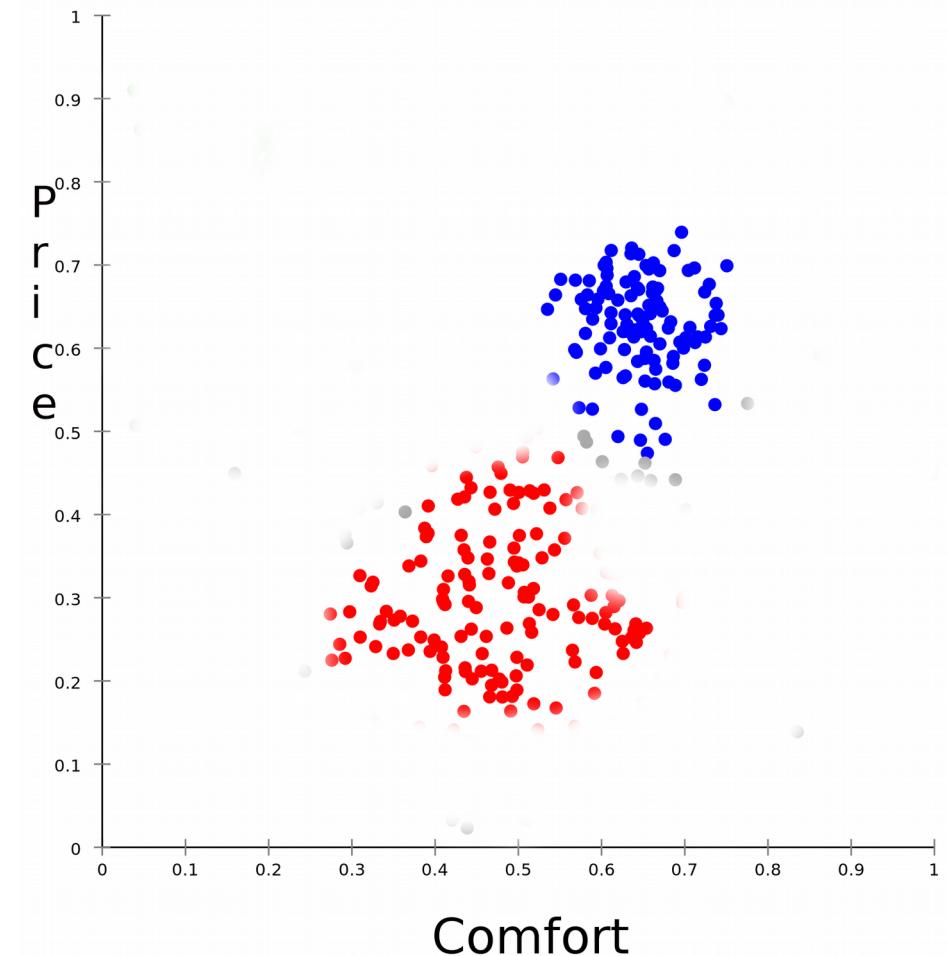
Neural Network

- A computing system made up of a number of interconnected processing elements, which process information by their dynamic state response to external inputs.
- Consider the example of Tic Tac Toe at a state in a game and its possible outcomes.



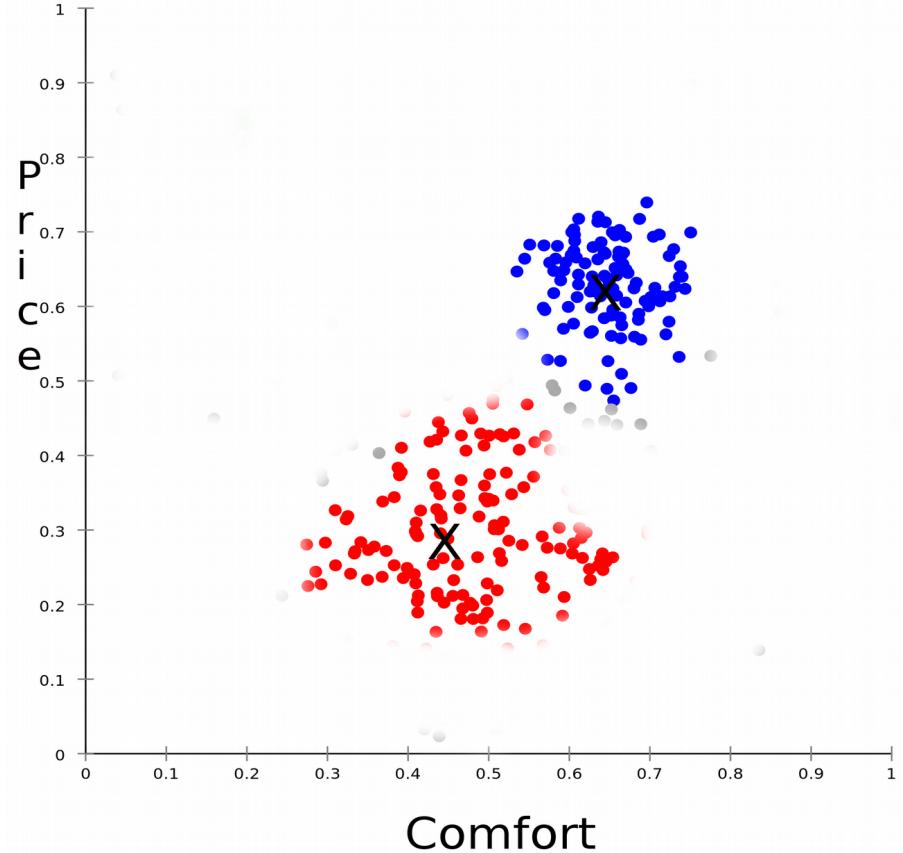
Clustering

- Task of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups.
- Consider an example of identifying Car types depending on the factors comfort(in scale 0 to 1 on x axis) and Price(in millions on y axis).



K-Means

- Starts with k initial clusters as specified. At each iteration, the records are assigned to the cluster with the closest centroid, or center.
- Consider the above stated example of identifying Car types.



Hierarchical Clustering

- It is a cluster analysis which seeks to build a hierarchy of clusters.
- Consider the example of Calculating population at each level by means of Election constituencies (x axis) and their respective Population(in hundred millions on y axis) of India.

