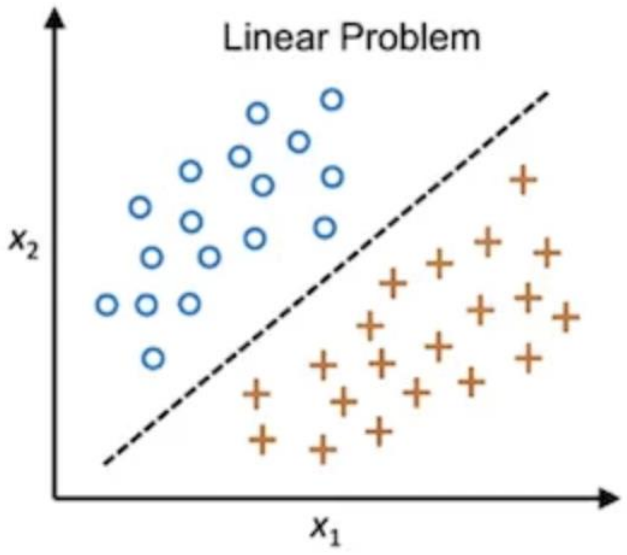
Linear vs. Non-Linear Classification

- Linear Classification refers to categorizing a set of data points to a discrete class based on a linear combination of its explanatory variables.
- On the other hand, Non-Linear Classification refers to separating those instances that are not linearly separable.
- A linear classifier is often used when classification speed is a priority as linear classifiers are known for being the fastest classifiers.
- It classifies large datasets quickly and accurately.
- On the other hand, Nonlinear classifiers are like detectives who can solve complex cases. It sorts data that doesn't follow straight patterns, like grouping things that don't fit on a simple line.

Linear Classification

- Linear Classification refers to categorizing a set of data points into a discrete class based on a linear combination of its explanatory variables.
- Some of the classifiers that use linear functions to separate classes are *Linear Discriminant Classifier, Naive Bayes, Logistic Regression, Perceptron, SVM (linear kernel).*



- Since we can tell one class apart from the other, these classes are called 'linearly-separable.
- However, an infinite number of lines can be drawn to distinguish the two classes.
- The exact location of this plane/hyperplane depends on the type of the linear classifier.

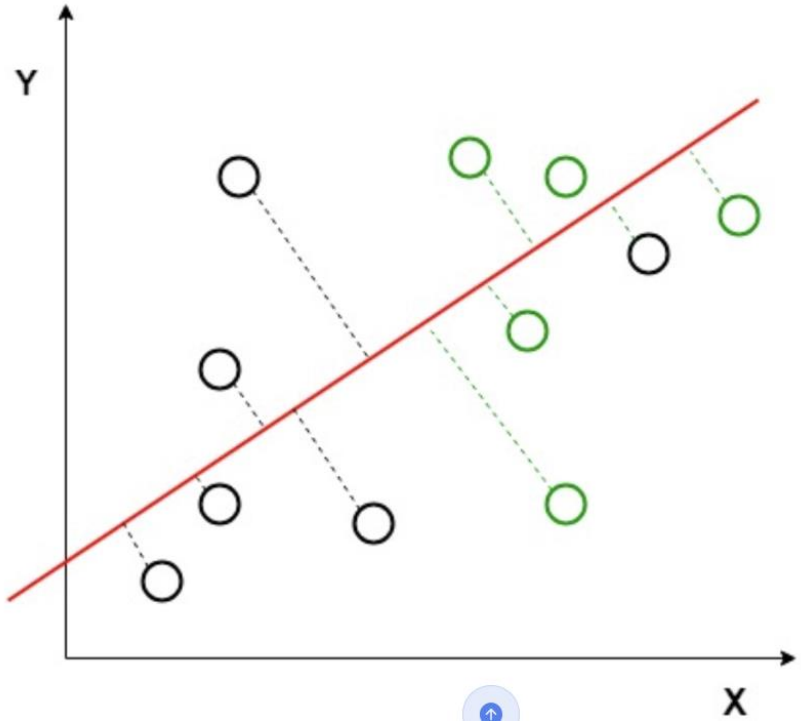
Linear Discriminant Classifier

→ It is a dimensionality reduction technique in the domain of Supervised Machine Learning.

→ It is crucial in modeling differences between two groups, i.e., classes.

→ It helps project features in a high dimensions space in a lower-dimensional space.

→ Technique - Linear Discriminant Analysis (LDA) is used, which reduced the 2D graph into a 1D graph by creating a new axis. This helps to maximize the distance between the two classes for differentiation.



Logistic Regression

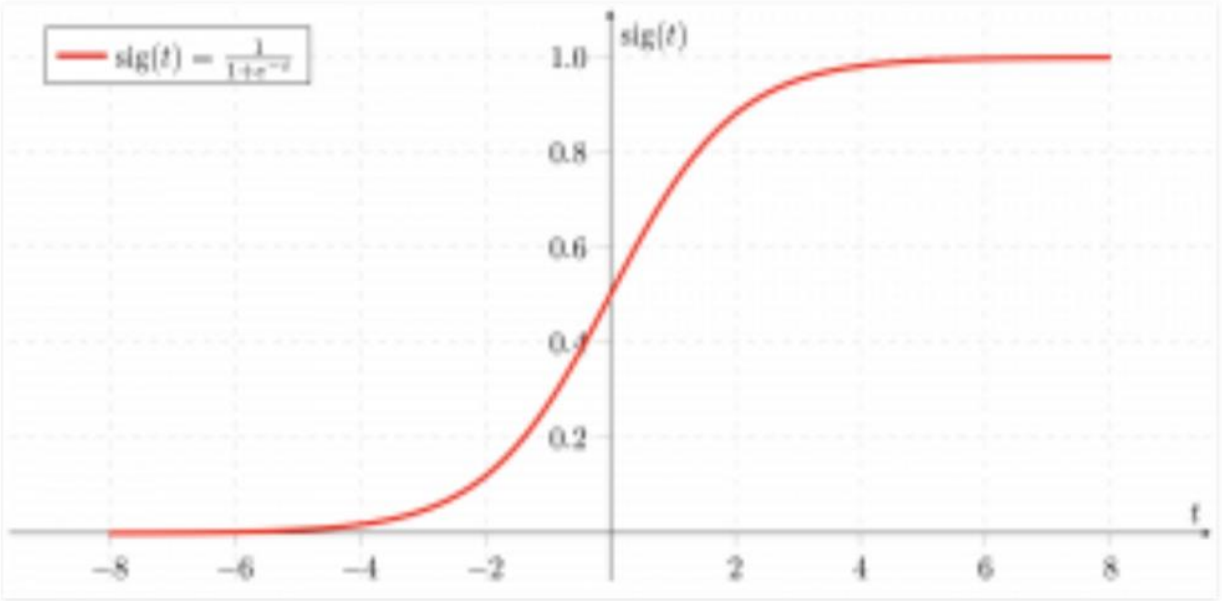
→ It is a very popular supervised machine learning algorithm.

→ The target variable can take only discrete values for a given set of features.

→ The model builds a regression model to predict the probability of a given data entry.

→ Similar to linear regression, logistic regression uses a linear function and, in addition, makes use of the 'sigmoid' function.

$$g(z) = \frac{1}{1+e^{-z}}$$



Support Vector Machine (linear kernel)

→ It is a straightforward supervised machine learning algorithm used for regression/classification.

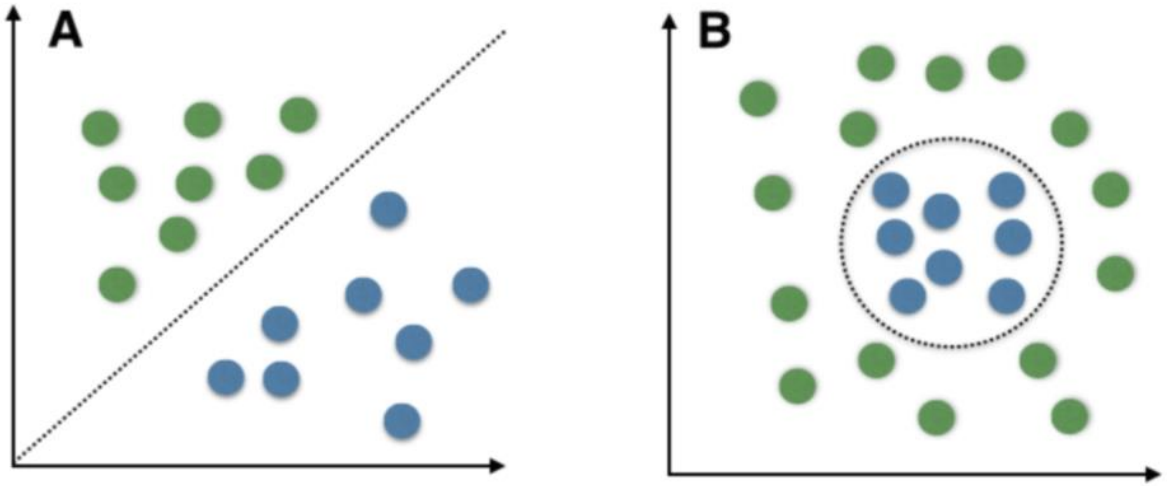
→ This model finds a hyper-plane that creates a boundary between the various data types.

→ It can be used for binary Classification as well as multinomial classification problems.

→ A binary classifier can be created for each class to perform multi-class Classification.

→ In the case of SVM, the classifier with the highest score is chosen as the output of the SVM.

→ SVM works very well with linearly separable data but can work for non-linearly separable data as well.

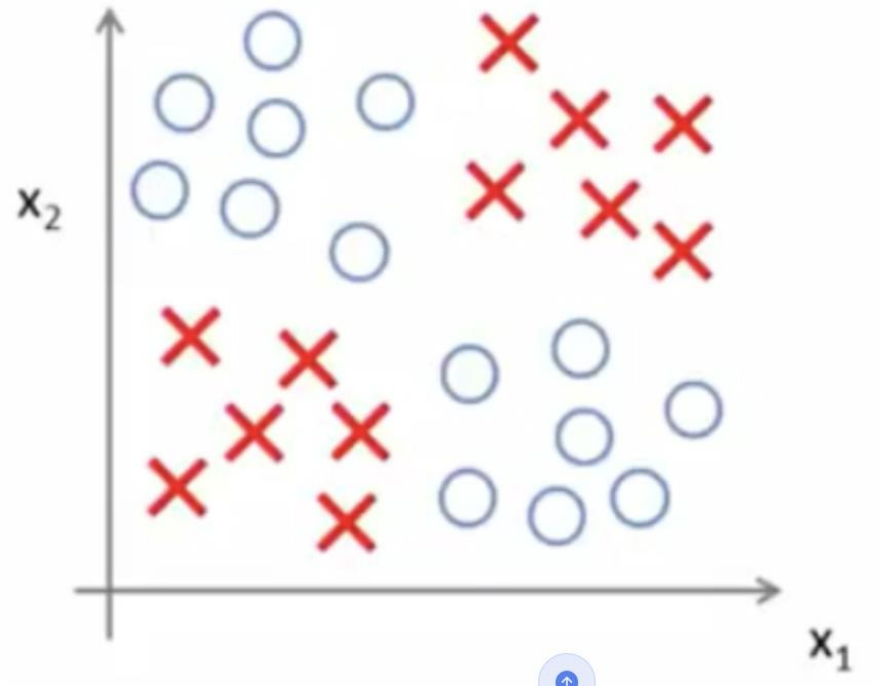


A: Linearly Separable Data B: Non-Linearly Separable Data

What is Non-Linear Classification

→ Non-linear classification refers to categorizing those instances that are not linearly separable.

→ Some of the classifiers that use non-linear functions to separate classes are Quadratic Discriminant Classifier, Multi-Layer Perceptron (MLP), Decision Trees, Random Forest, and K-Nearest Neighbours (KNN).

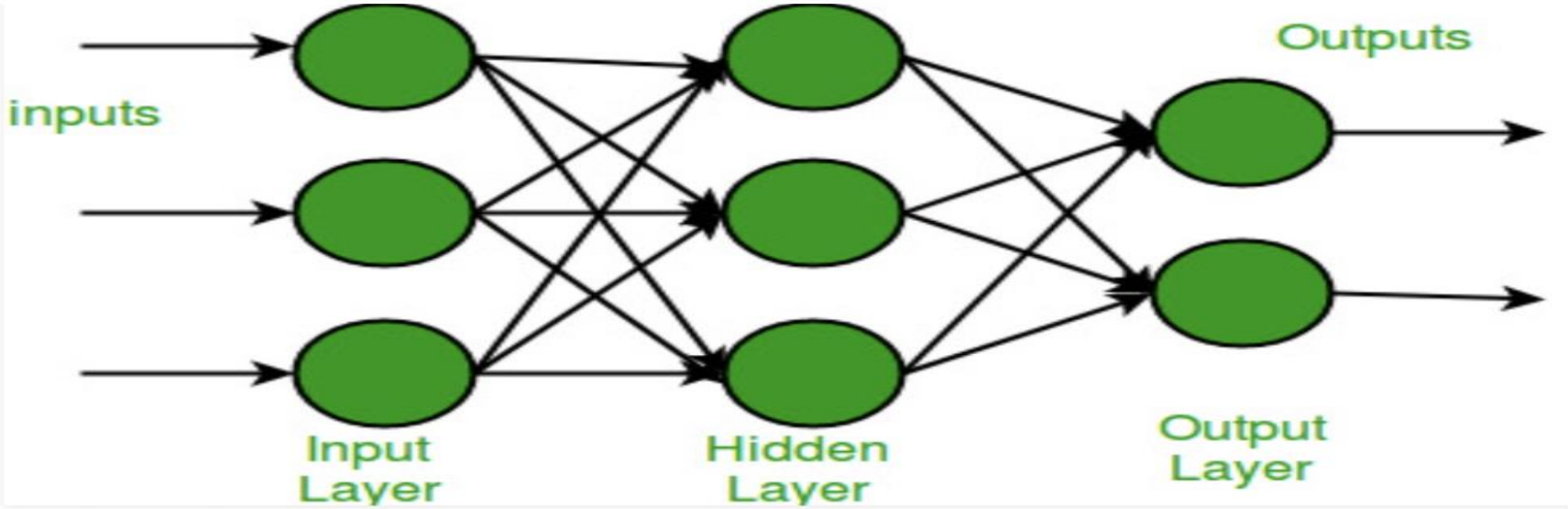


Multi-Layer Perceptron (MLP)

→ This is nothing but a collection of fully connected dense layers. These help transform any given input dimension into the desired dimension.

→ It is nothing but simply a neural network.

→ MLP consists of one input layer(one node belonging to each input), one output layer (one node belonging to each output), and a few hidden layers (>= one node belonging to each hidden layer).



→ In the above diagram, we notice three inputs, resulting in 3 nodes belonging to each input.

→ There is one hidden layer consisting of 3 nodes.

→ There is an output layer consisting of 2 nodes, indicating two outputs.

→ Overall, the nodes belonging to the input layer forward their outputs to the nodes present in the hidden layer. Once this is done, the hidden layer processes the information passed on to it and then further passes it on to the output layer.

Decision Tree

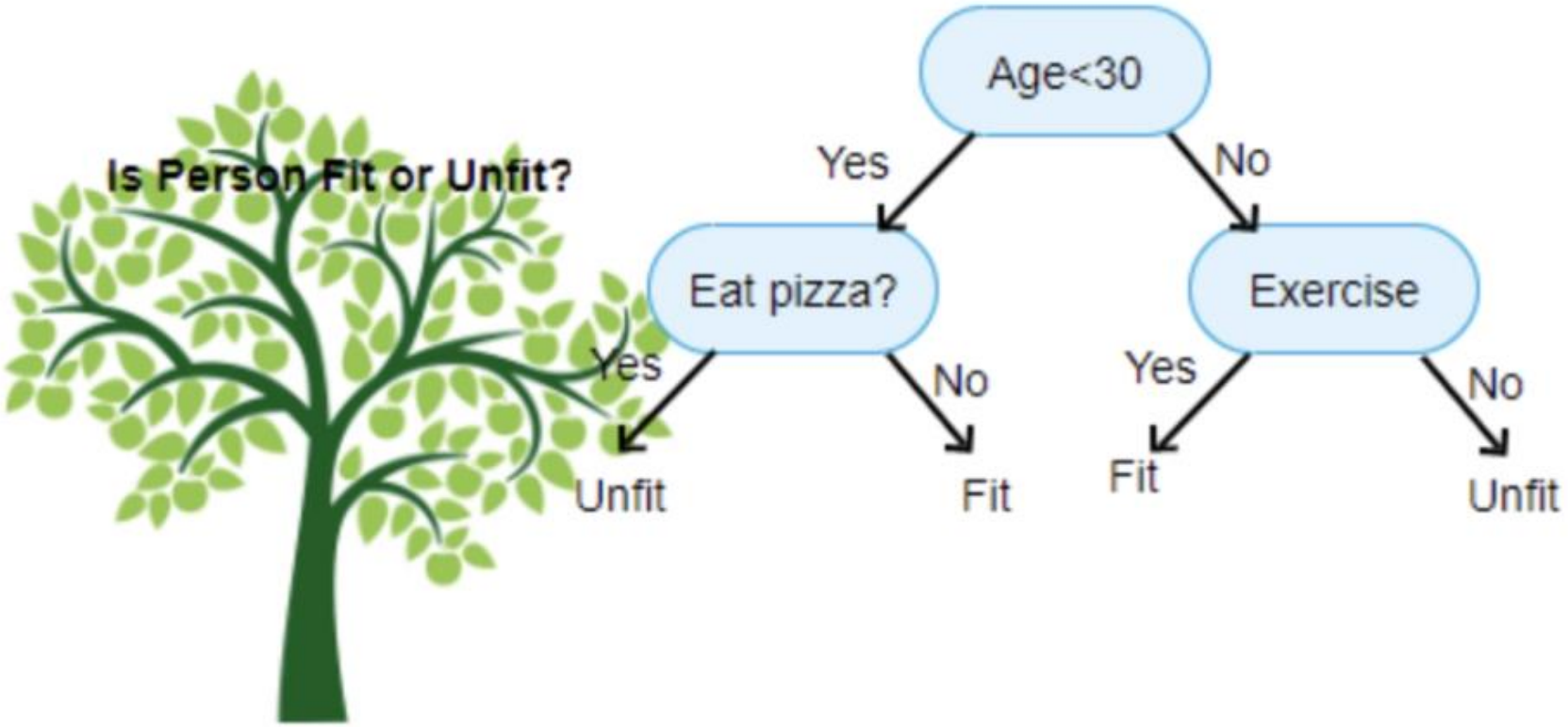
→ It is considered to be one of the most valuable and robust models.

→ Instances are classified by sorting them down from the root to some leaf node.

→ An instance is classified by starting at the tree's root node, testing the attribute specified by this node, then moving down the tree branch corresponding to the attribute's value, as shown in the above figure.

→ The process is repeated based on each derived subset in a recursive *partitioning* manner.

→ For a better understanding, see the diagram below.


Is Person Fit or Unfit?
Age<30
Yes
No
Eat pizza?
Exercise
Yes
No
Yes
No
Unfit
Fit
Fit
Unfit

Linear vs Non-Linear Classification

Now, we will briefly sum up all that we've learned and try to compare and contrast Linear Classification and Non-Linear Classification.

S.No	Linear Classification	Non-Linear Classification
1.	Linear Classification refers to categorizing a set of data points into a discrete class based on a linear combination of its explanatory variables.	Non-Linear Classification refers to categorizing those instances that are not linearly separable.
2.	It is possible to classify data with a straight line.	It is not easy to classify data with a straight line.
3.	Data is classified with the help of a hyperplane.	The utilization of kernels is made to transform non-separable data into separable data.