

# Introduction to AI Algorithms

Simran Dewangan(22111056)

July 19, 2024

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Linear Regression</b>	<b>2</b>
<b>3</b>	<b>Random Forest</b>	<b>2</b>
<b>4</b>	<b>K-Nearest Neighbors (KNN)</b>	<b>3</b>
<b>5</b>	<b>K-Means Clustering</b>	<b>4</b>
<b>6</b>	<b>Naive Bayes</b>	<b>4</b>
6.1	Bayes' Theorem . . . . .	4
<b>7</b>	<b>Conclusion</b>	<b>5</b>

# 1 Introduction

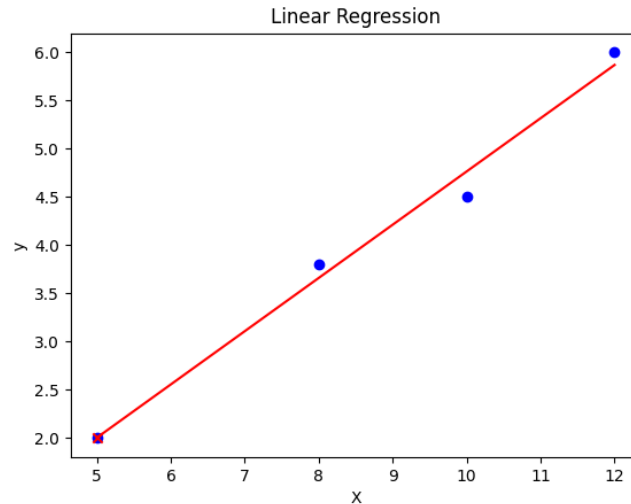
Artificial Intelligence (AI) uses different algorithms to solve various problems. In this document, we will look at some basic AI algorithms: Linear Regression, Random Forest, K-Nearest Neighbors (KNN), K-Means Clustering, and Naive Bayes.

## 2 Linear Regression

Linear Regression is a simple method for predicting a number (the dependent variable) based on one or more other numbers (independent variables). The relationship between them is a straight line, shown by the equation:

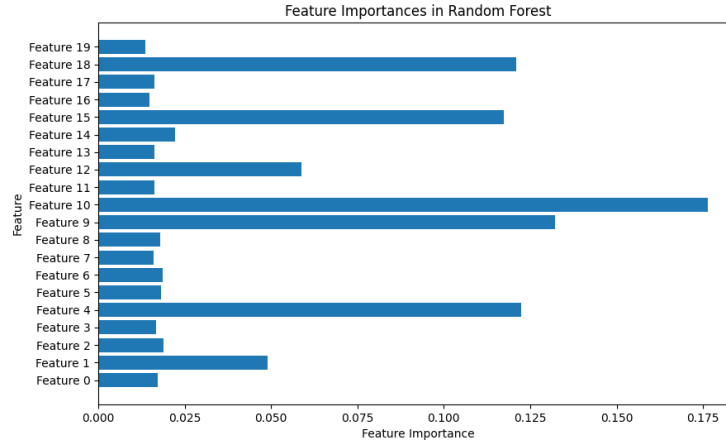
$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon \quad (1)$$

Here,  $y$  is what we want to predict,  $X$  are the input values,  $\beta$  are the coefficients that we need to find, and  $\epsilon$  is the error term. The goal is to find the line that best fits the data by minimizing the difference between the actual and predicted values.



## 3 Random Forest

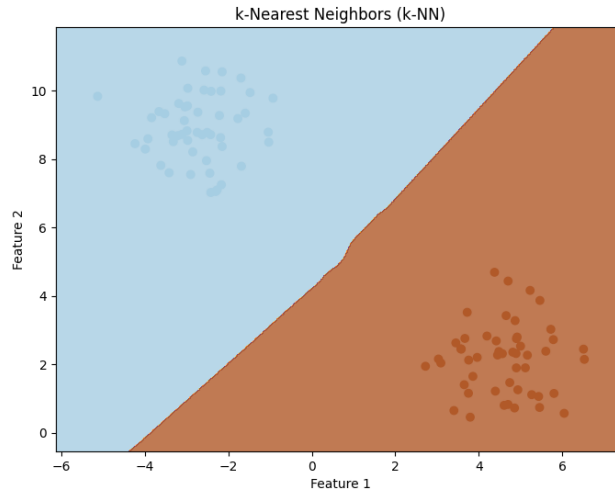
Random Forest is a technique that uses many decision trees to make a prediction. Each tree looks at a random subset of the data, and the final prediction is made by combining the results from all the trees. For classification, we use majority voting, and for regression, we use averaging.



## 4 K-Nearest Neighbors (KNN)

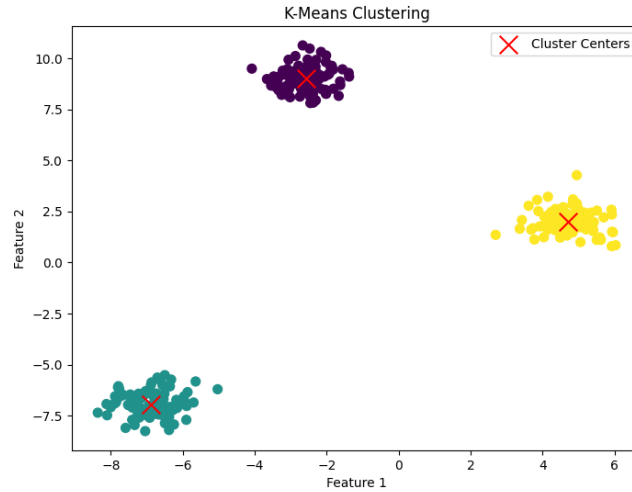
K-Nearest Neighbors (KNN) is a method used for classification and regression. To predict the value for a new data point, KNN looks at the 'k' closest data points and uses their values. The distance between data points is usually measured using Euclidean distance.

For classification, KNN assigns the most common member among the k-nearest neighbors. For regression, it calculates the average value of the k-nearest neighbors.



## 5 K-Means Clustering

K-Means Clustering is an unsupervised method that groups data into 'k' clusters. Each cluster has a centroid, and each data point belongs to the cluster with the nearest centroid.



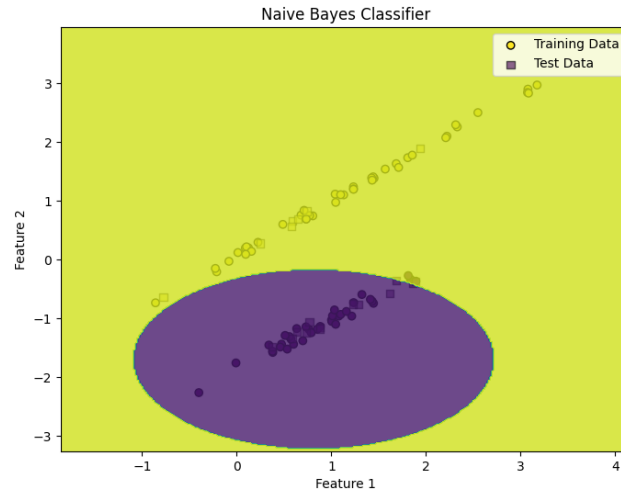
## 6 Naive Bayes

Naive Bayes is a classification method based on Bayes' Theorem. It assumes that the features are independent of each other. Despite this simple assumption, it works well in many cases.

### 6.1 Bayes' Theorem

Bayes' Theorem is written as:

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (2)$$



## 7 Conclusion

Knowing these basic AI algorithms helps in understanding more complex methods. Each algorithm has its own way of solving problems and handling data.