



UNIT 2

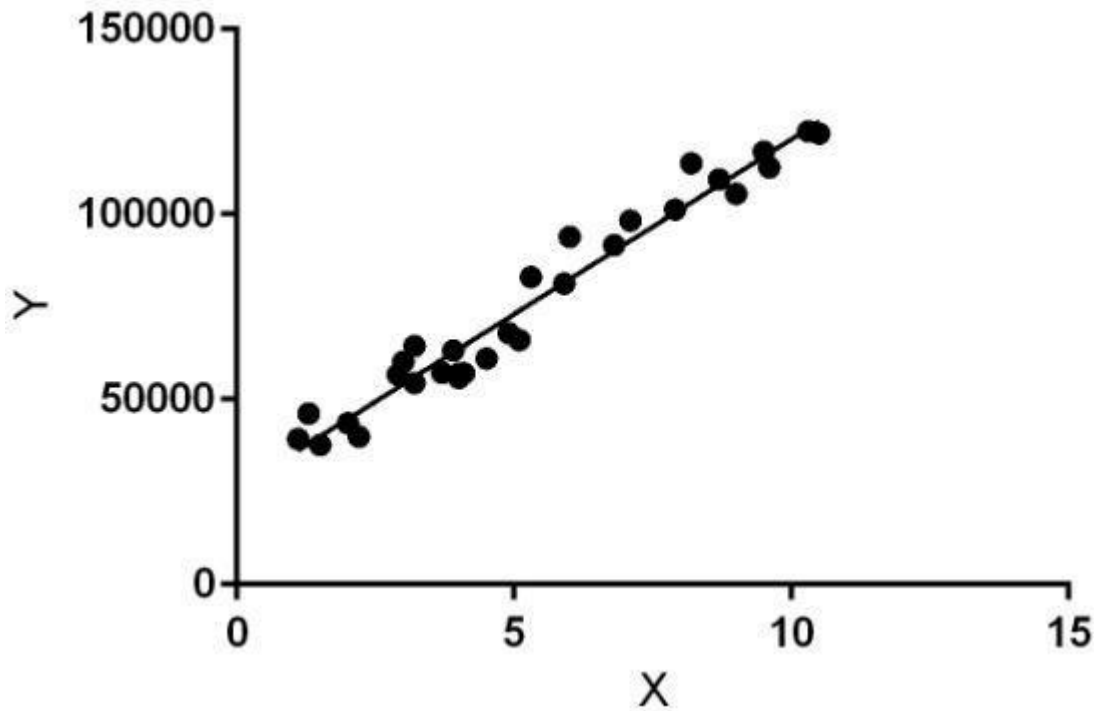
SUPERVISED LEARNING

Linear Regression

- Linear Regression is a machine learning algorithm based on supervised learning.
- It performs a regression task.
- Regression models a target prediction value based on independent variables.
- It is mostly used for finding out the relationship between variables and forecasting.
- Different regression models differ based on – the kind of relationship between dependent and independent variables they are considering, and the number of independent variables getting used.
- There are many names for a regression's dependent variable.
- It may be called an outcome variable, criterion variable, endogenous variable, or regressand.
- The independent variables can be called exogenous variables, predictor variables, or regressors.
- Linear regression is used in many different fields, including finance, economics, and psychology, to understand and predict the behaviour of a particular variable.
- For example, in finance, linear regression might be used to understand the relationship between a company's stock price and its earnings, or to predict the future value of a currency based on its past performance.
- One of the most important supervised learning tasks is regression.
- In regression set of records are present with X and Y values and these values are used to learn a function, so that if you want to predict Y from an unknown X this learned function can be used.
- In regression we have to find value of Y, So, a function is required which predicts Y given X. Y is continuous in case of regression.



- Here Y is called as criterion variable and X is called as predictor variable.
- There are many types of functions or modules which can be used for regression. Linear function is the simplest type of function.
- Here, X may be a single feature or multiple features representing the problem.



- Linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x).
- Hence, the name is Linear Regression. In the figure above, X (input) is the work experience and Y (output) is the salary of a person.
- The regression line is the best fit line for our model.

Hypothesis function for Linear Regression :

$$y = \theta_1 + \theta_2 \cdot x$$