Online Student Training for “Artificial Intelligence & Machine Learning”

(4th Feb, 2021 – 17th Mar, 2021)

LINEAR REGRESSION

Faculty Trainer

Nawvarsha Pipada

Department of Computer Science & E

Engineering College Bikane

Regression

Regression is a method of modelling a target value based on independent predictors.

This method is mostly used for forecasting and finding out cau and effect relationship between variables.

Regression techniques mostly differ based on the number of independent variables and the type of relationship between the independent and dependent variables.

Regression types

Based on type of relationship between the independent and dependent variables

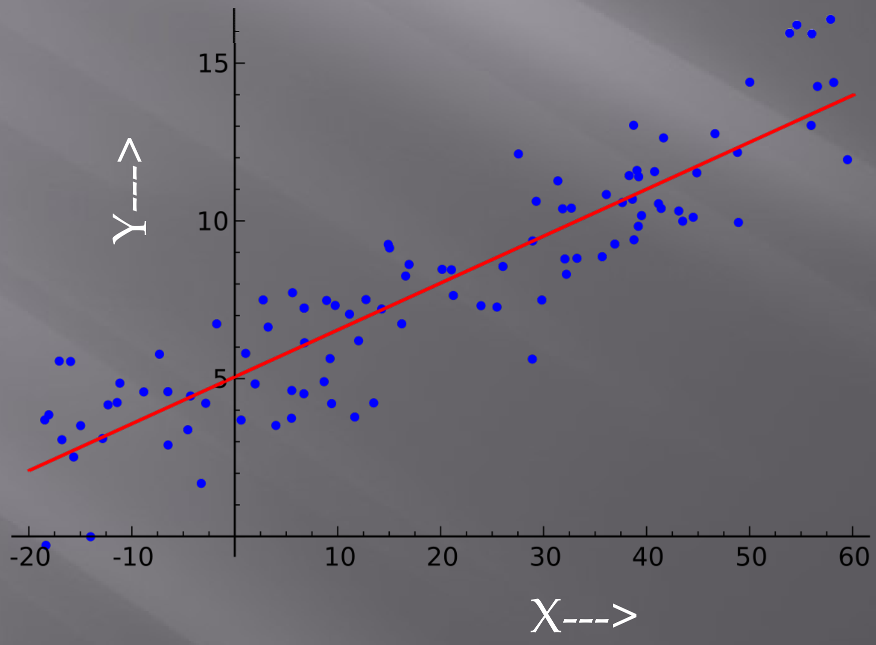
- Linear
- Non-Linear

Based on the number of independent variables (as dependent variable will be only one)

- Simple
- Multiple

Simple Linear Regression

Simple linear regression is a type of regression analysis where the number of independent variables is one and there is a linear relationship between the independent(x) and dependent(y) variable.



Simple Linear Regression

The regression line is referred to as the best fit straight line.

Based on the given data points, we try to plot a line that models the points the best.

The line can be modelled based on the linear equation shown below.

y = a_0 + a_1 * x ## Linear Equation

Two important Concepts

- Cost Function
  - A cost function is a measure of how wrong the model is in terms of its ability to estimate the relationship between X and y.
  - This is typically expressed as a difference or distance between the predicted value a the actual value
  - In our case, the cost function is actually minimization problem where we would lik minimize the error between the predicted value and the actual value. It is also know ad mean squared er

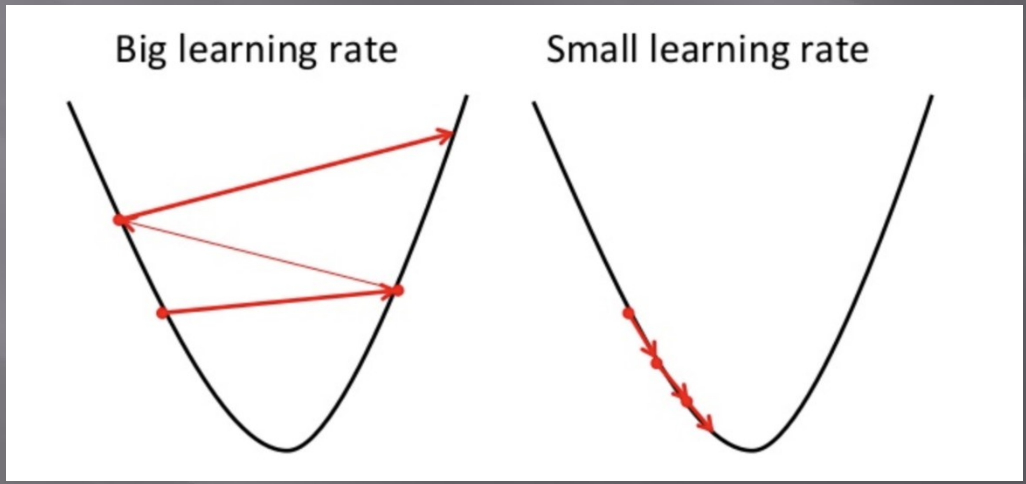
$$minimize \frac{1}{n} \sum_{i=1}^{n} (pred_i - y_i)^2$$

$$J = \frac{1}{n} \sum_{i=1}^{n} (pred_i - y_i)^2$$

Two important Concepts

Gradient Descent

- Gradient descent is a method of updating a_0 and a_1 to reduce the cost function(MS
- The idea is that we start with some values for a_0 and a_1 and then we change these values iteratively to reduce the cost.
- Gradient descent helps us on how to change the values.



To find the gradients

$$J = \frac{1}{n}\sum_{i=1}^{n}(pred_i - y_i)^2$$

$$J = \frac{1}{n}\sum_{i=1}^{n}(a_0 + a_1 \cdot x_i - y_i)^2$$

$$\left.\begin{aligned}
\frac{\partial J}{\partial a_0} &= \frac{2}{n}\sum_{i=1}^{n}(a_0 + a_1 \cdot x_i - y_i) \implies \frac{\partial J}{\partial a_0} = \frac{2}{n}\sum_{i=1}^{n}(pred_i - y_i) \\
\frac{\partial J}{\partial a_1} &= \frac{2}{n}\sum_{i=1}^{n}(a_0 + a_1 \cdot x_i - y_i) \cdot x_i \implies \frac{\partial J}{\partial a_1} = \frac{2}{n}\sum_{i=1}^{n}(pred_i - y_i) \cdot x_i
\end{aligned}\right\} \text{Grad}$$

Updating a_0 and a_1

Alpha is the learning rate which is a hyperparameter that you must speci

A smaller learning rate could get you closer to the minima but takes more time to reach the minima, a larger learning rate converges sooner but ther is a chance that you could overshoot the minima.

$$a_0 = a_0 - \alpha \cdot \frac{2}{n} \sum_{i=1}^{n} (pred_i - y_i)$$

$$a_1 = a_1 - \alpha \cdot \frac{2}{n} \sum_{i=1}^{n} (pred_i - y_i) \cdot x_i$$

Goodness-of-fit

R-squared is a statistical measure of how close the data are to th fitted regression line.

It is also known as the coefficient of determination, or the coefficient of multiple determination for multiple regression.

In general, the higher the R-squared, the better the model fits your data

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \overline{y})^2}$$

Steps for Implementing Linear Regression

1. Check if there exist linear relationship between independent and dependent variable.
2. Load training and test data from dataset files.
3. Initialize the values of learning rate, slope, intercept number of runs (update slope and intercept) .
4. Predict the value of target for training data and calculate the residual error.
5. Update the values of slope and intercept using gradients to minimize residual errors.
6. After getting, final values of slope and intercept, predict the targe values for testing data.
7. Evaluate the fitness of model using R2-Score.

THANK YOU!!!!!