Date :

Practical - 4

Aim: Implement linear regression using python programming

Key formulas:

1) linear equation:

$$Y = aX + b$$

- Y → dependent variable (target)
- X → independent variable (feature)
- a → slope (coefficient)
- b → intercept

2) Slope (a) calculation:

$$a = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

- $\bar{X}$ → mean of X
- $\bar{Y}$ → mean of Y

3) Intercept (b) calculation:

$$b = \bar{Y} - a\bar{X}$$

Input:
- A dataset with
  - Independent (X) → eg. Hours studied
  - Dependent variable (Y) → eg. Exam score

Expected output:
- computed slope (a) & intercept (b)
- A best fit regression line plotted over data plo
- predicted y-values for given x values.

THEORY:

Linear Regression is a foundational supervised learning algorithm that models the linear relationship between independent variable (x) & dependent variable (y). The algorithm assumes the relationship the slope intercept. the model works by minimizing the sum squared residuals (RSS) between observed & predi values, known as the 'least square method. optimizat & yeld coefficients that produce fit line through the data points.

The slope is calculated is the covariance between divided by covariance of X, while the inter derived include linearity, independy of residuals lead to biased or inefficient estimates, requiring diagnostics checks like residual diagram Leverage Analysis.

model performance is evaluated using if squared error while simple, linear regression is powerful for predictive tasks in economics, healthcare & engineering, practical implementation involves preprocessing steps like feature scaling & outlier removal to enhance accuracy

PROCEDURE :

1) Input Data
- Take user input for x & y as comma separated values.

2) Data processing
- converts inputs to numpy arrays & reshape for sklearn compatibility

3) Model training
- use LinearRegression() from sklearn.linear_model to fit the data & compute slope (a) not & intercept (b)

4) prediction
- predict y for a new x value using a trained model

5) Visualization
- plot actual data points (x & y) & regression line
- highlight the predicted & actual values.

6) Output result:
- Display the regression equation (Y = ax + b)
- Print the predicted y value for value provi

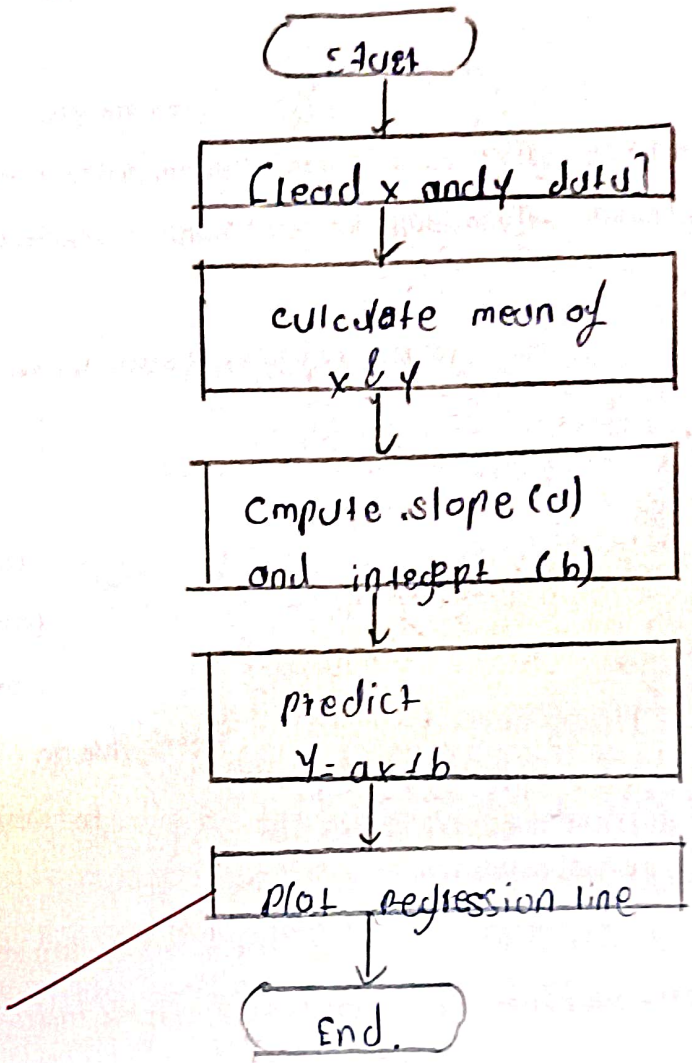
OBSERVATIONS:

1) Slope (a) : indicates the rate of change of y w.r.t. x
2) Intercept (b) : value of y when x = 0
3) The regression line minimizes the error between predicted & actual values.

Conclusion:

- Successfully implemented Linear Regression using key feature.
- The best fit line in predicting future based on training data
- The method is fundamental for prediction in machine learning.

Flowchart:


Start
Read x and y data
calculate mean of x & y
Compute slope (a) and intercept (b)
predict Y=ax+b
Plot regression line
End.