Linear Regression Overview / Contents

www.aiquest.org



▪ Regression Concepts

▪ Types of Regression

▪ In-depth intuition of OLS

▪ Lose Functions

▪ Cost Function

▪ R Squared Values

▪ Coding with Python:

▪ Implementing Linear Regression

▪ Simple ML Project

▪ Assignment

1

Regression Concepts 

Basic Concept

• **Regression in Machine Learning:**

Regression is a technique used to predict numerical values based on input features. It models the relationship between a dependent variable (what you want to predict) and independent variables (features).

• **Example: Predicting House Prices:**

Imagine you're predicting house prices based on square footage. The regression model finds a line that best fits the data: Price = 100 \* SquareFootage + 50000. Here, 100 is the increase in price for each square foot increase, and $50,000 is the starting price estimate. This model helps estimate prices for different house sizes.

www.aiquest.org

2

Regression Concepts Examples 

1. Economics: GDP Prediction:

Using historical data, economists can predict a country's future GDP based on factors like inflation rate, unemployment rate, and consumer spending.

2. Healthcare: Patient Outcome:

Doctors can predict a patient's recovery time after surgery based on variables like age, pre-existing conditions, and the complexity of the procedure.

3. Retail: Sales Forecasting:

Retailers can use regression to forecast sales based on parameters like advertising spend, holiday season, and previous sales data.

4. Finance: Stock Price Prediction:

Traders and investors can predict stock prices by analyzing factors like trading volume, historical prices, and economic indicators.

5. Agriculture: Crop Yield Estimation:

Regression helps farmers predict crop yields based on factors like weather conditions, soil quality, and type of crop.

www.aiquest.or 3 g

Regression Concepts

Examples

6. Marketing: Customer Lifetime Value:

Marketers use regression to estimate a customer's lifetime value based on purchase history, engagement, and demographic information.

7. Education: Student Performance:

Educators can predict student performance on standardized tests using factors like attendance, study time, and past test scores.

8. Energy: Energy Consumption:

Energy companies can predict household energy consumption based on variables like weather, household size, and appliance usage.

9. Transportation: Fuel Efficiency:

Manufacturers predict a vehicle's fuel efficiency based on engine specifications, weight, and aerodynamics.

10. Real Estate: Property Valuation:

Regression helps in estimating property values based on features like location, square footage, and nearby amenities.

www.aiquest.org

4

www.aiquest.org



Types of Regression?

5

Types of Regression 

Common types

There are several types of regression techniques, each designed to handle different types of data and relationships between variables. Here are some common types of regression:

1. Linear Regression:

- Simple Linear Regression: Predicting a continuous dependent variable using a single independent variable. - Multiple Linear Regression: Predicting a dependent variable using multiple independent variables.

2. Polynomial Regression:

- Modeling nonlinear relationships by adding polynomial terms to the regression equation.

3. Ridge Regression:

- Adding a penalty term to the coefficients to prevent overfitting.

4. Lasso Regression:

- Similar to ridge regression, but with a penalty that encourages some coefficients to become exactly zero, leading to feature selection.

www.aiquest.org

6

Types of Regression 

Common types

5. Logistic Regression:

- Used for binary or multinomial classification tasks, predicting the probability of an event occurring.

6. Poisson Regression:

- Modeling count data, often used in situations where the dependent variable represents counts.

7. Time Series Regression:

- Modeling time-dependent data, considering temporal patterns and autocorrelation.

9. Nonlinear Regression:

- Fitting a nonlinear function to the data to capture complex relationships.

9. Support Vector Regression (SVR):

- Utilizes support vector machines for regression tasks, particularly suited for high-dimensional spaces.8

7

www.aiquest.org

www.aiquest.org



All you need to know about Linear Regression!

8

Linear Regression 

Basic Concepts

• Linear regression is a fundamental **supervised machine learning algorithm** used for predicting a continuous numerical value based on one or more input features.

• It models the relationship between the **dependent** variable and the **independent** variables as a linear equation.

• The goal is to find the **best-fitting line** (or hyperplane in higher dimensions) that minimizes the difference between the observed and predicted values.

• This best-fitting line represents the **linear relationship** between the input features and the target variable.

Here, 

M = Coefficient of the input feature X

C = Intercept 

X = Features

Y = Predicted Output / Label

www.aiquest.org

9

Linear Regression 

Linear Regression Representations



www.aiquest.or 10 g

Linear Regression Basic Concepts

Fig: Straight Line



X = 10, 30, 50

Y = 2\*10 + 3 =23 Y = 2\*30 + 3 = 63 Y = 2\*50 + 3 = 103

www.aiquest.or 11 g

Linear Regression

Basic Concepts

Fig: Straight Line



X = 10, 30, 50

Y = 2\*10 + 3 =23

Y = 2\*30 + 3 = 63

Y = 2\*50 + 3 = 103

| **X**  10 | **Actual**  25 | **Predicted**  23 |
| --- | --- | --- |
| 30 | 60 | 63 |
| 50 | 100 | 103 |

12 www.aiquest.org

Linear Regression 

Basic Concepts

F(x) = mx + c = y = predicted output

13 www.aiquest.org

Linear Regression 

Basic Concepts

F(x) = mx + c = y = predicted output





14 www.aiquest.org

Linear Regression 

Basic Concepts

**Formula 01:** Slope,



15

www.aiquest.org

Linear Regression 

Maths

| **Given Dataset**  **x y** | |
| --- | --- |
| 1 | 2 |
| 2 | 3 |
| 3 | 5 |
| 4 | 4 |



www.aiquest.org

16

Linear Regression 

Formula

**Formula 02:** Slope,



17

www.aiquest.org

Linear Regression

Maths

www.aiquest.org



| **Given Dataset**  **x y** | |
| --- | --- |
| 1 | 2 |
| 2 | 3 |
| 3 | 5 |
| 4 | 4 |

18

Linear Regression

Maths

**Note:** Both formulas are mathematically **equivalent** and will yield the same result when applied correctly. So, slope ��=0.8 (Both Formula). You are allowed to apply any of them.

www.aiquest.org



|  |
| --- |

19

Linear Regression Let’s See Another Math From Scratch

20

www.aiquest.org

Linear Regression

Raw Calculation

Slope, m = Σ((x - x̄) \* (y - ȳ)) / Σ((x - x̄)^2) Intercept, c = ȳ - m \* x̄

Where:

x is a data point on the independent variable (x-axis). y is the corresponding dependent variable (y-axis). x̄is the mean of the independent variable.

ȳ is the mean of the dependent variable.



| **Feature (X)**  5 | **Class (Y)**  50 |
| --- | --- |
| 7 | 65 |
| 4 | 42 |
| 8 | 76 |
| 2 | 23 |
| 10 | 105 |
| 7 | ? |

21 www.aiquest.org

Linear Regression 

Raw Calculation

Prediction

|  |
| --- |

22 www.aiquest.org

Linear Regression 

Raw Calculation

Final Goal

|  |
| --- |

www.aiquest.or 23 g

Linear Regression 

Raw Calculation

24 www.aiquest.org

Linear Regression 

Raw Calculation

25 www.aiquest.org

Linear Regression Python Implementation

26

www.aiquest.org

Linear Regression

Raw Implementation with Python 

**Step 1: Calculate the means of X and Y**

mean\_x = df['X'].mean()

mean\_y = df['Y'].mean()

**Step 2: Calculate the slope (m) using the formula**

numerator\_m = ((df['X'] - mean\_x) \* (df['Y'] - mean\_y)).sum()

denominator\_m = ((df['X'] - mean\_x)\*\*2).sum()

slope\_m = numerator\_m / denominator\_m

**Step 3: Calculate the y-intercept (c) using the formula**

intercept\_c = mean\_y - (slope\_m \* mean\_x)

**Step 4: Predict the value for X = 7 using the regression line equation Y = mX + c** x\_predict = 7

y\_predict = (slope\_m \* x\_predict) + intercept\_c

**Show the calculated slope, intercept, and the predicted value of Y**

slope\_m, intercept\_c, y\_predict

27 www.aiquest.org

Linear Regression

Best Fit Line with Sklearn 

Using Sklearn Library 

Data Set Value of M & C 

Line

**Ref:** Click the Link

28 www.aiquest.org

Linear Regression

Visual Prediction

Data Set Value of M & C **Ref:** Click the Link

Line

www.aiquest.or 29 g

Linear Regression Lose/Cost Function

30

www.aiquest.org

Splitting Dataset 

100% 

70% 15% 15%

31 www.aiquest.org

Linear Regression

Cost Function**The cost function** is a function, which is associates a cost with a **decision.**

50 min

40 min

Map

30 min20 min

Cost Function 55 min

Dhanmondi

Dhanmondi Dhanmondi Dhanmondi Dhanmondi

32 www.aiquest.org

Linear Regression

Loss Vs. Cost Function

**1. Loss (or Error) for a Single Sample:**

• When you calculate the difference between the actual value and the predicted value for a single data point, it's generally referred to as a "loss" or "error" for that specific data point.

• This term is used to describe the discrepancy between the prediction and the true value for a single instance.

**2. Cost (or Loss) for the Entire Dataset:**

• When you calculate the average or total of these losses/errors across the entire dataset, it's often referred to as the "cost" or "loss" for the dataset.

• The term "cost" or "loss" is used to describe the overall quality of the model's predictions for the entire dataset.

33 www.aiquest.org

Linear Regression

Residuals 

**Residual** = Observed Value - Predicted Value



**Predicted Result:**

• Pred = 0.8 \* X + 1.5

• Y\_pred = (0.8\*1)+ 1.5 = 2.3

• Y\_pred = (0.8\*2)+ 1.5 = 3.1

• Y\_pred = (0.8\*3)+ 1.5 = 3.9

• Y\_pred = (0.8\*4)+ 1.5 = 4.7

www.aiquest.or 34 g

Linear Regression

L1 Loss & L2 Loss



35 www.aiquest.org

Linear Regression

MAE: L1 Loss

36 www.aiquest.org

Linear Regression

MSE: L2 Loss

37 www.aiquest.org

Linear Regression 

RMSE: L2 Loss

**NOTE:** To verify the results, please read the Python lecture notes.

38 www.aiquest.org



Thanks for your patience!

Let’s Implement with Python

www.aiquest.or 39 g