```
Start coding or generate with AI.
```

Linear Regression?

Linear Regression is a supervised machine learning algorithm used to predict a continuous value (number) based on input data.

It finds a straight line that best fits the data.

Linear Regression Equation

$y=mx+b$

Where:

y = Output (dependent variable)

x = Input (independent variable)

m = Slope (how fast y changes with x)

b = Intercept (value of y when x = 0)

Case Study: Predicting Student Marks

Problem Statement

A teacher wants to predict marks of a student based on the number of hours studied.

📊 Dataset

Hours Studied (x)	Marks (y)
1	35
2	40
3	50
4	60
5	65

Observation

As study hours increase, marks also increase

Relationship is approximately linear

How Linear Regression Works (Step-by-Step)

Plot data points on a graph

Draw the best-fit straight line

Minimize the error between actual and predicted values

Use the line equation to predict future values

Example Prediction

Assume the trained model gives:

Predict marks for 4 hours of study

y=7(4)+30

y=7(4)+30=58

Predicted Marks = 58

Python Peactical

[1]
✓ 3s

```
from sklearn.linear_model import LinearRegression

# Data
X = [[1], [2], [3], [4], [5]]
y = [35, 40, 50, 60, 65]

# Model
model = LinearRegression()
model.fit(X, y)

# Prediction
hours = [[4]]
predicted_marks = model.predict(hours)
print(predicted_marks)
```

```
[58.]
```

Advantages of Linear Regression

Easy to understand and implement

Works well for simple relationships

Fast computation

Useful for prediction and trend analysis

Limitations

Limitations

Assumes linear relationship

Sensitive to outliers

Not suitable for complex patterns

Real-Life Applications

House price prediction

Salary prediction

Sales forecasting

Crop yield estimation

Weather trend analysis

Logistic Regression is a supervised machine learning algorithm used for binary classification problems.

It predicts the probability that an input belongs to a particular class (usually 0 or 1) using a sigmoid (S-shaped) function.

Simple Words

Logistic Regression answers Yes/No, True/False, Pass/Fail

Output is not a number like marks, but a probability between 0 and 1

Based on this probability, the final class is decided

Mathematical Form

$$P(y = 1) = \frac{1}{1 + e^{-(wx+b)}}$$

The sigmoid function converts any value into **0–1**

If probability ≥ 0.5 → Class **1**

If probability < 0.5 → Class **0**

[ ]

```
Start coding or generate with AI.
```

Example

Predict Pass (1) or Fail (0) based on study hours

Predict Disease (Yes/No) based on symptoms

Predict Spam (Yes/No) email

Predict Spam (Yes/No) email

How Logistic Regression Works (Simple Explanation)

1. It takes the input feature (**Hours Studied**)
2. Applies a **linear equation**
3. Passes the result through a **Sigmoid function**
4. Produces a probability between **0 and 1**

Sigmoid Function:

$$P(y=1) = \frac{1}{1+e^{-z}}$$

Where:

$$z = w \cdot x + b$$

🔍 Decision Rule

- If probability ≥ 0.5 → **Pass**
- If probability < 0.5 → **Fail**

🏏 Example Prediction

Student studies for 3 hours

- Model output = **0.30**
- Prediction → ❌ Fail

Student studies for 5 hours

- Model output = **0.85**
- Prediction → ✅ Pass

Python Practical

[2] ✓ 0s

```
from sklearn.linear_model import LogisticRegression

# Training data
X = [[1], [2], [3], [4], [5], [6]]
y = [0, 0, 0, 1, 1, 1]

# Model
model = LogisticRegression()
model.fit(X, y)

# Prediction
```

```
hours = [[5]]
prediction = model.predict(hours)
print("Pass" if prediction[0] == 1 else "Fail")
```

```
Pass
```

Advantages of Logistic Regression

Easy to understand

Fast to train

Works well for binary problems

Probability-based output

Limitations

Works best when classes are linearly separable

Not suitable for complex non-linear data

A hospital wants to predict whether a patient has a disease or not based on basic medical test results.

This is a binary classification problem:

Disease = 1

No Disease = 0

Dataset (Simple Example)

Age	Blood Sugar Level	Disease
25	120	0
30	130	0
40	150	1
50	170	1
60	180	1

Logistic Regression Works (Simple Steps)

Takes patient data (Age, Sugar Level)

Applies a linear equation

Uses a sigmoid function

Uses sigmoid function

Outputs probability between 0 and 1

Classifies patient as Disease / No Disease

Q.1 predict whether a patient has a disease or not using multiple medical features with the help of Logistic Regression.

The model uses **four medical features**:

Feature	Description
Age	Patient age (years)
Blood Sugar	Sugar level (mg/dL)
Blood Pressure	BP (mmHg)
BMI	Body Mass Index

Logistic Regression?

Output has two classes only

Can handle multiple input features

Produces probability-based decision

Widely used in healthcare diagnosis

Difference Between Linear Regression and Logistic Regression

Linear Regression	Logistic Regression
Predicts a **continuous value**	Predicts a **binary outcome**
Output is a **number**	Output is **0 or 1 (Yes/No)**

Output Type

Aspect	Linear	Logistic
Output Range	$-\infty$ to $+\infty$	0 to 1
Example	Marks = 78	Pass (1) / Fail (0)

Mathematical Model

Linear Regression

$$y = mx + b$$

Logistic Regression

$$P(y = 1) = \frac{1}{1 + e^{-(mx+b)}}$$

(Logistic regression uses a **Sigmoid function**)

- Problem Type

Linear Regression	Logistic Regression
House price prediction	Email spam detection
Salary prediction	Disease (Yes/No)
Temperature forecasting	Pass / Fail result

Graph Shape

Linear Regression ⟶ Straight line

Logistic Regression ⟶ S-shaped (Sigmoid) curve

- Learning Method

Feature	Linear	Logistic
Error Function	Mean Squared Error (MSE)	Log Loss (Cross-Entropy)
Decision Boundary	Not required	Required

Linear Regression predicts continuous values, whereas Logistic Regression predicts categorical (binary) outcomes using a sigmoid function.

Use Linear Regression ⟶ when output is a number

Use Logistic Regression ⟶ when output is Yes/No

- Design a machine learning model using Logistic Regression to predict whether a crop is at risk of disease (Yes/No) based on environmental and crop-related features.

Feature	Description
Temperature (°C)	Average weekly temperature
Humidity (%)	Relative humidity
Rainfall (mm)	Weekly rainfall
Wind Speed (km/h)	Average wind speed
Crop Age (days)	Days after sowing

Output (Target Variable)

Value	Meaning
1	Disease Risk Present
0	No Disease Risk

Temp	Humidity	Rainfall	Wind	Crop Age	Disease
28	65	20	5	18	0
30	70	25	6	22	0
32	85	80	4	30	1
33	90	95	3	35	1
31	88	70	4	28	1

Collect weather and crop data

Apply a linear combination of features

Use sigmoid function

Obtain disease risk probability

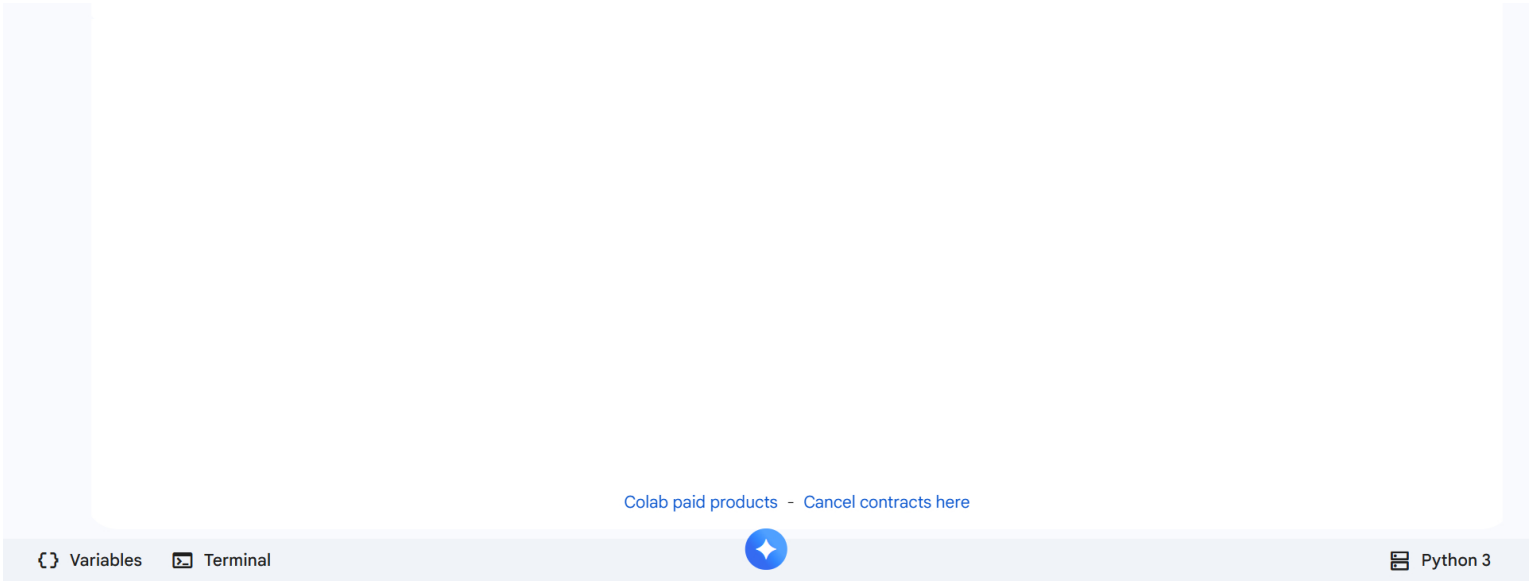
Classify crop as Risk / No Risk

Double-click (or enter) to edit

Double-click (or enter) to edit

[ ]

```
Start coding or generate with AI.
```


Colab paid products - Cancel contracts here
Variables
Terminal
Python 3