Introduction to Machine Learning

Machine learning empowers computers to learn from data. It has three main types: supervised, unsupervised, and reinforcement learning. This presentation explores these categories, focusing on supervised learning with linear and logistic regression.

A by Abenezer yakob

Supervised Learning: Learning from Labeled Data

Definition

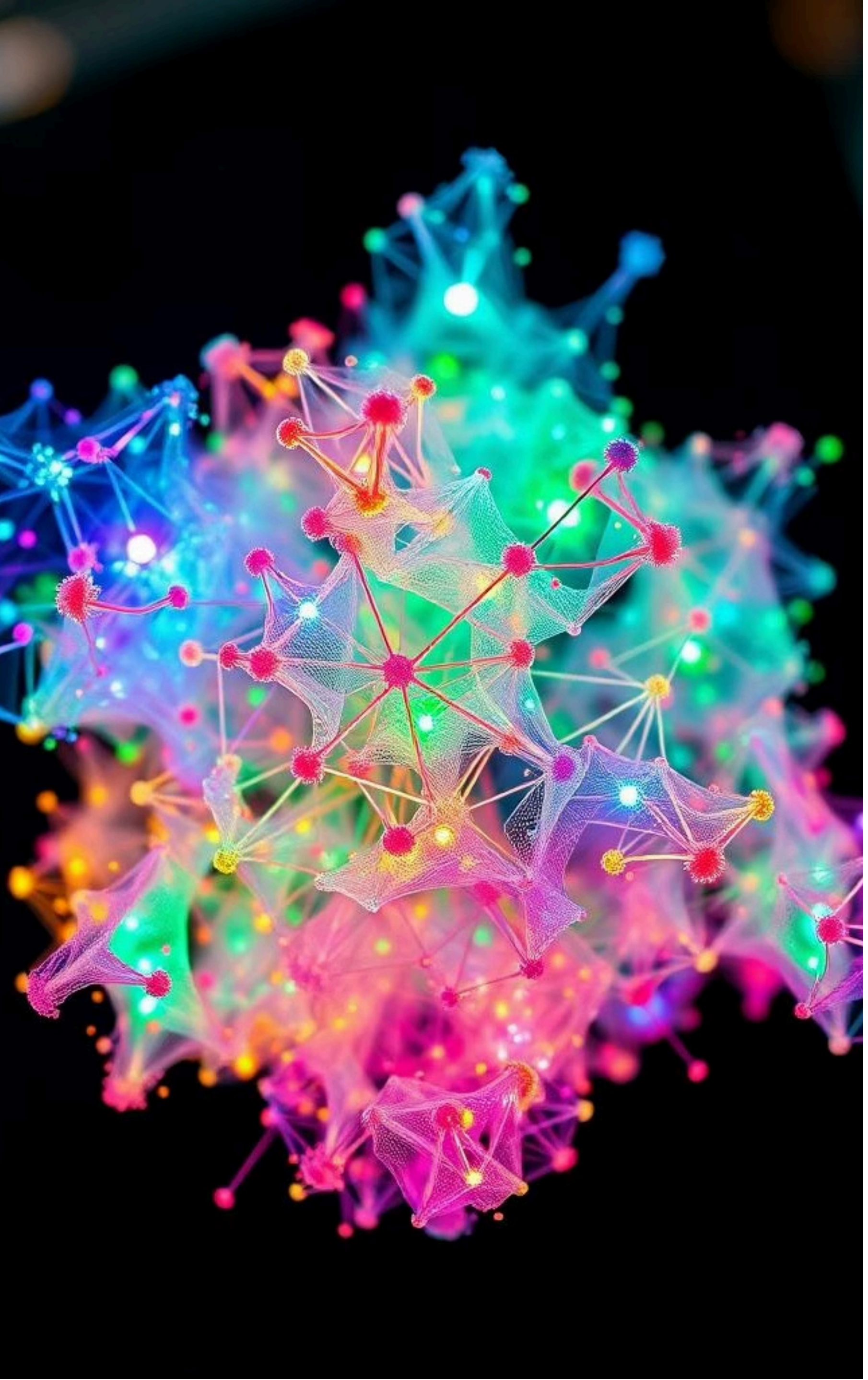
Algorithms learn from labeled data to predict or classify outcomes.

Goal

Map input features to outputs using example data pairs.

Common Algorithms

- Linear Regression
- Logistic Regression
- Support Vector Machines (SVM)
- Decision Trees
- Random Forests
- Neural Networks



Unsupervised Learning: Discovering Hidden Patterns

Definition

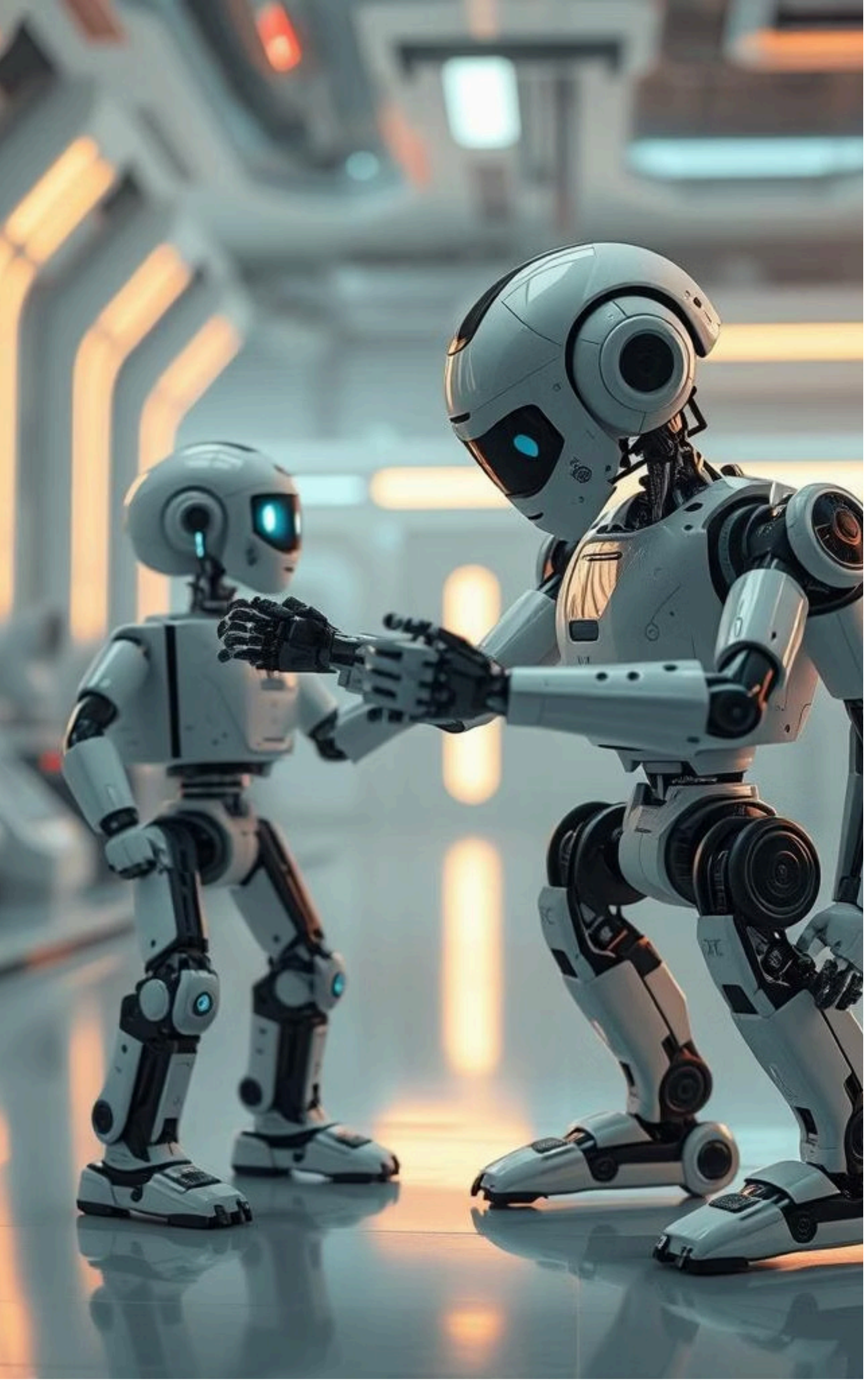
Algorithms learn from unlabeled data to identify structures and patterns.

Goal

Discover intrinsic relationships and groupings in data.

Common Algorithms

- K-Means Clustering
- Hierarchical Clustering
- Principal Component Analysis (PCA)
- Association Rule Mining (Apriori)
- Anomaly Detection



Reinforcement Learning: Learning Through Interaction

Definition

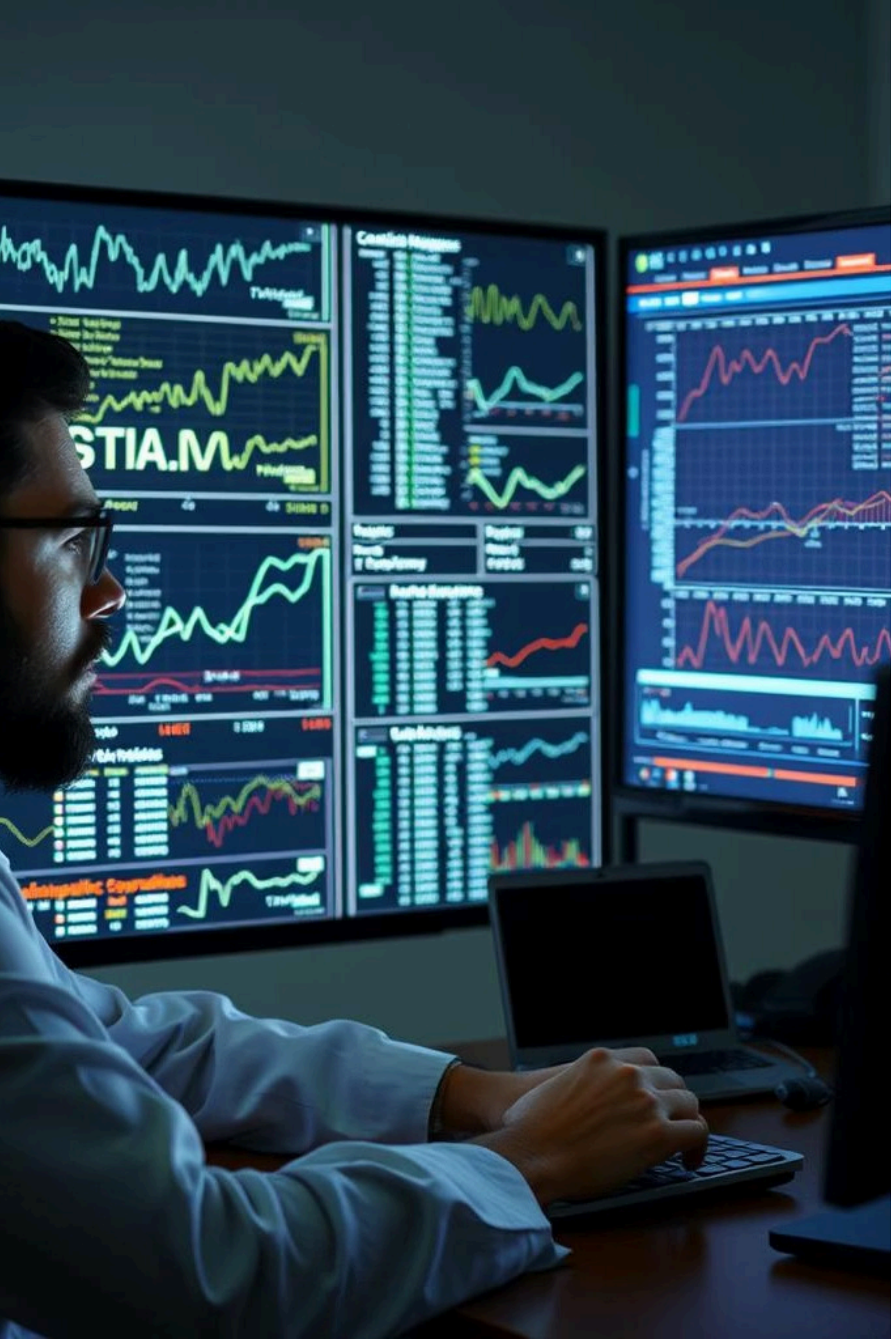
Algorithms learn by interacting with environments via trial and error.

Goal

Maximize a cumulative reward signal over time.

Common Algorithms

- Q-Learning
- Deep Q-Networks (DQN)
- SARSA
- Policy Gradients (REINFORCE, A2C, PPO)
- Actor-Critic Methods



Diving into Supervised Learning

Prevalence

Accounts for 70% of practical machine learning applications.

Nature

Clear input-output mapping problems with labeled data.

Focus Areas

Linear Regression and Logistic Regression techniques explored.

Linear Regression: Predicting Continuous Values

1

Goal

Model linear relationships between input variables and output.

2

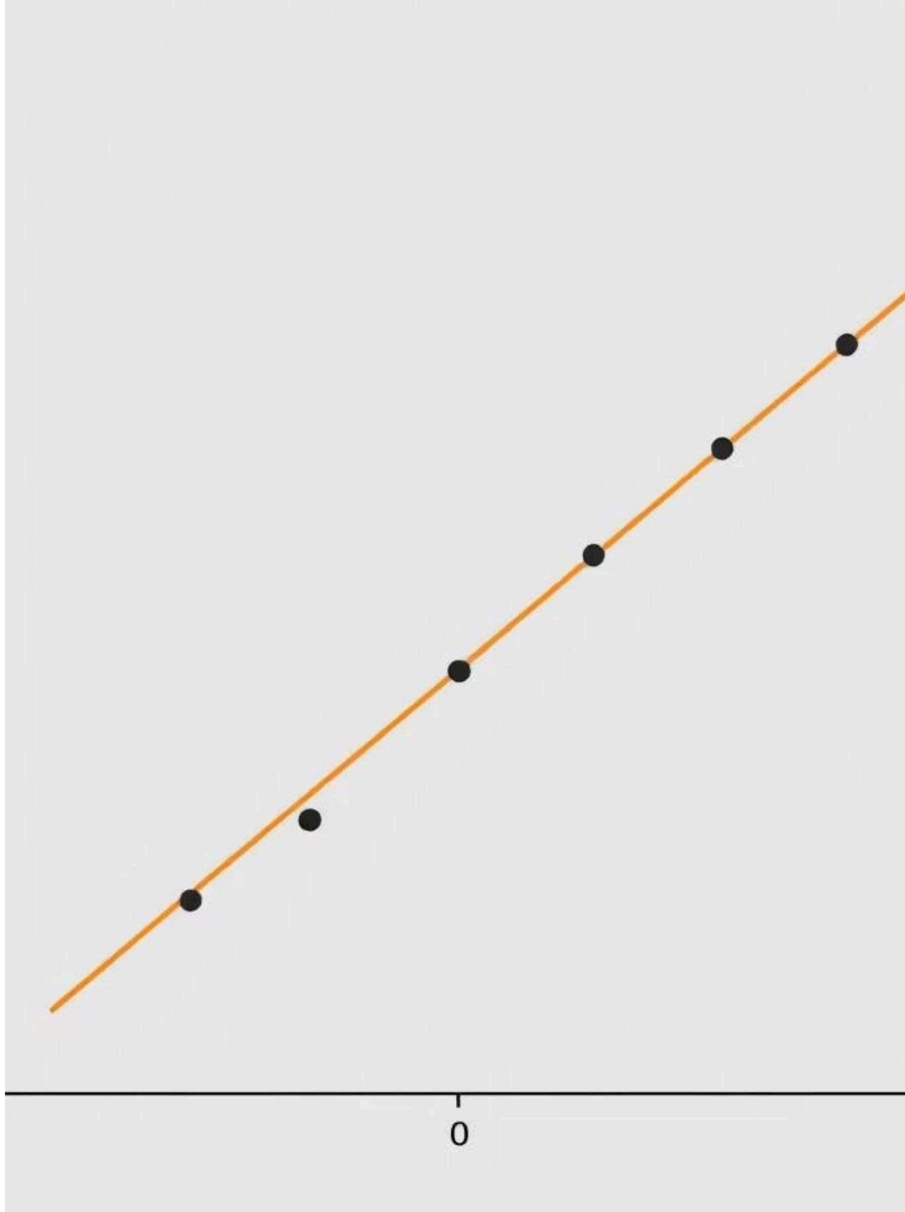
Equation

$Y = \theta_0 + \theta_1 X_1 + \theta_2 X_2 + \ldots + \varepsilon$ (predicted value formula)

3

Example

Predict house prices using size, location, and room count.



Why Linear Regression is Important

Linear regression is a foundational and highly practical machine learning technique due to several key reasons:

- **Simplicity and Interpretability:** Its straightforward nature makes it easy to understand and interpret, serving as an excellent entry point for learning machine learning concepts.
- **Predictive Power**: It effectively predicts future outcomes based on historical data, proving valuable across diverse sectors like finance, healthcare, and marketing.
- **Foundation for Other Models:** Many advanced algorithms, including logistic regression and neural networks, build upon the fundamental principles of linear regression.
- **Computational Efficiency:** It is computationally inexpensive and performs well for problems exhibiting a clear linear relationship between variables.
- **Widespread Adoption:** It is a widely used technique in both statistics and machine learning for various regression tasks.
- **Insightful Analysis:** It provides clear insights into the relationships between variables, illustrating how much one variable influences another.

The goal of linear regression is to find a straight line that minimizes the error (the difference) between the observed data points and the predicted values. This line helps us predict the dependent variable for new, unseen data.

$$y' = \theta 0 + \theta 1x1 + \theta 2x2 + \cdots + \theta nxn$$

Where y^ is the predicted value. The goal is to minimize the error between observed and predicted values using the following cost function:

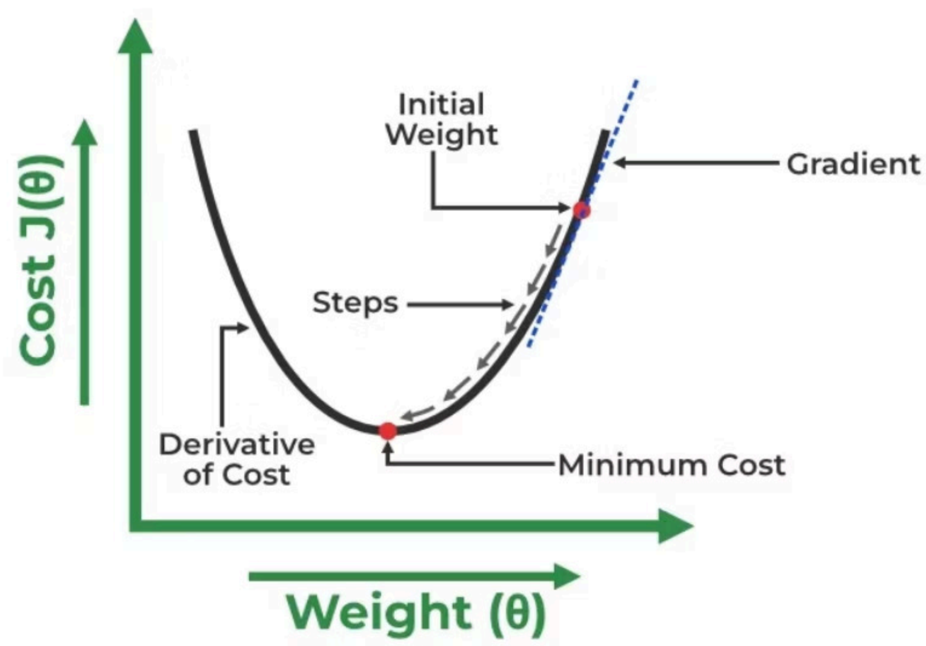
$$Costfunction(J) = n1\sum ni(yi' - yi)2$$

- $x1, x2, \ldots, xn$ are the independent variables.

The coefficients (weights) $\theta 1, \theta 2, \ldots, \theta n$ correspond to each predictor.

Gradient Descent for Linear Regression

Linear regression can be trained using **gradient descent** to minimize the **mean squared error (MSE)**. The model starts with random θ1 and θ2 values and iteratively updates them to find the best-fit line.



Gradient Descent

Logistic Regression: Predicting Categories

1

Goal

Predict probability of binary outcomes using input features.

2

Logistic Function (p)

The probability 'p' of a positive outcome is given by the sigmoid function:

$$p = 1/(1 + e^{-z})$$

3

Linear Combination (z)

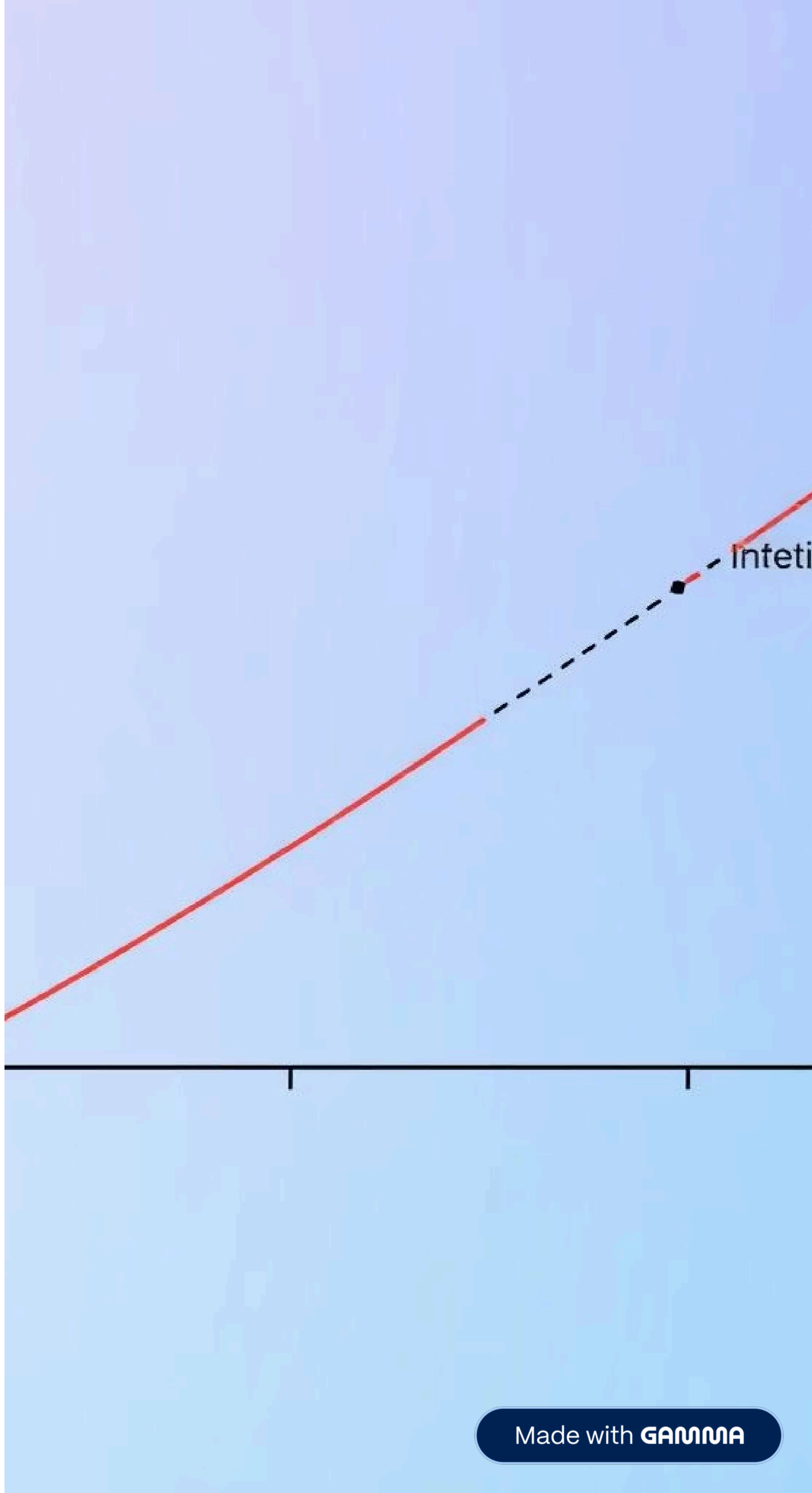
The input 'z' to the sigmoid function is a linear combination of the input features (X_i) and their corresponding weights (θ_i), similar to the output of linear regression:

$$z = \theta_0 + \theta_1 X_1 + \theta_2 X_2 + \cdots + \theta_n X_n$$

4

Example

Predict customer churn based on demographics and usage.



Conclusion: Supervised Learning Applications

Linear Regression

Used for sales forecasting with accuracy within 5%.

Logistic Regression

Effective in fraud detection achieving 90% accuracy.

Choosing Algorithms

Match algorithm to your data type and problem goals.