UNIVERSITY OF GREENWICH
Alliance with FPT Education

AI-Powered Predictive Maintenance Training Series

Phase 1: Machine Learning Fundamentals

Instructor: Le Tran Ngoc Tran
Email: tranltn2@fe.edu.vn
Date: 14 July 2025

UNIVERSITY OF GREENWICH
Alliance with FPT Education

CONTENTS

1. About the series
2. Warm-up Games
3. Introduction to Machine Learning
4. Types of Machine Learning
5. Common Algorithms & Mechanisms
6. Python for Machine Learning
7. Supervised Learning Foundation
8. Practical Workshop

01 About the Series

Structure, Validation, Information

UNIVERSITY OF GREENWICH
Alliance with FPT Education

Series

Phase 1 ML

- Date: 14 July 2025
- Slot 1: 8h00-11h00
- Slot 2: 12h30 – 15h30
- Room: **F.309**

Phase 2 DL

- Date: 14 + 16 July 2025
- Slot 1: 8h00-11h00, 14 July
- Slot 2: 8h00 – 11h00, 16 July
- Room: **F.328**

Phase 3 NVIDIA

- Date: 22 + 23 July 2025
- Slot 1: 8h00-12h00, 14 July
- Slot 2: 8h00 – 12h00, 16 July
- Room: **IT Innovation Lab**

Attendance: **75%** + **Finish practical sessions**

02 Warm-up

Game

Warm-up Game

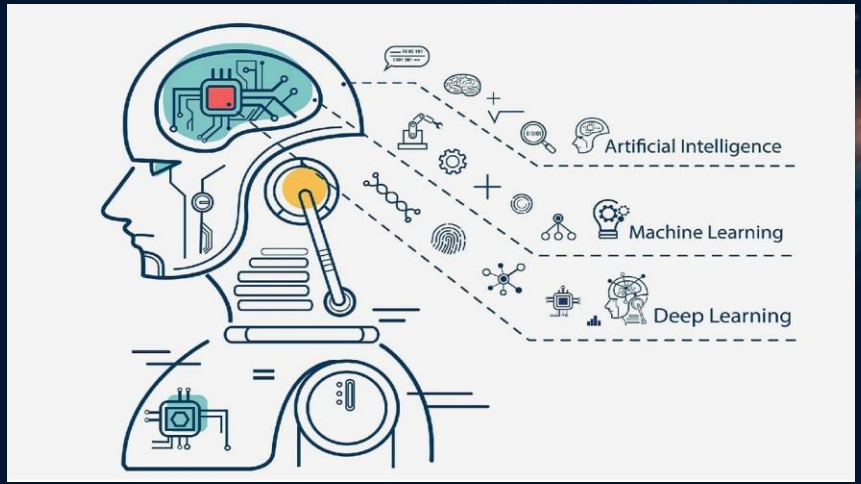
- Answer 11 questions on math, data types, probability, and statistics

03 Introduction to ML

What is ML?
Traditional Programming vs Machine Learning
Why is it important?
Applications in Industry

What is Machine Learning?

- Machine learning is "a field of study in artificial intelligence concerned with the development and study of statistical algorithms that can learn from data and generalize to unseen data, and thus perform tasks without explicit instructions” [1]
- Arthur Samuel, one of the pioneers in computer science, defined it in 1959 as "the field of study that gives computers the ability to learn without being explicitly programmed” [2]

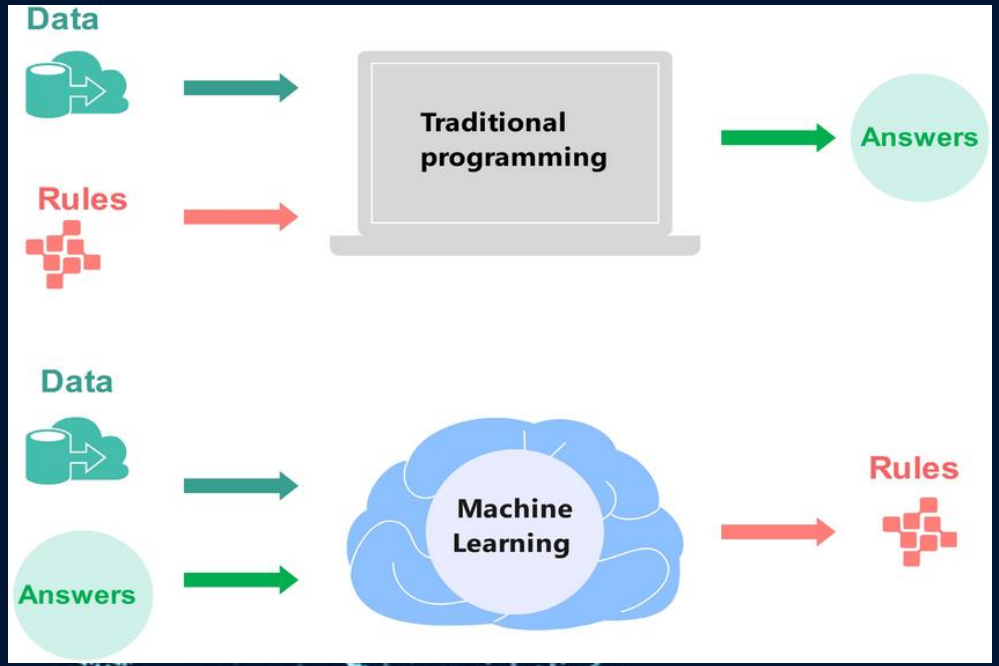


UNIVERSITY OF GREENWICH
Alliance with FPT Education

Traditional Programming vs ML

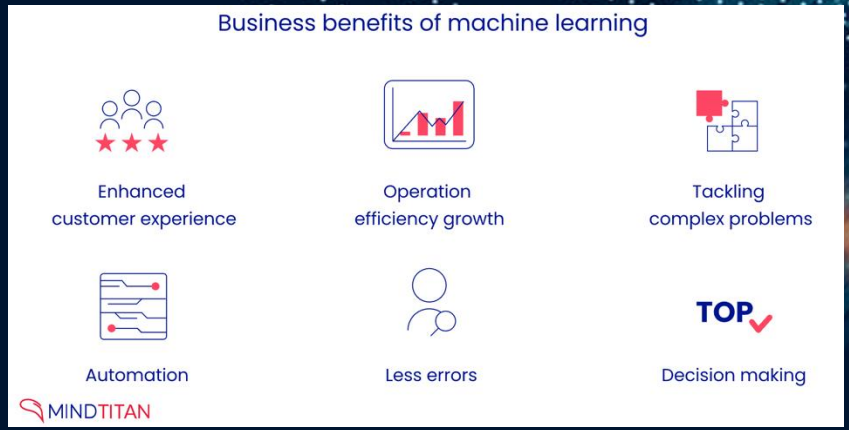
1. Traditional Programming
- Rules + Data -> Output
- Human writes explicit rules

2. Machine Learning
- Data + Output -> Rules (Model)
- Learn rules from Data



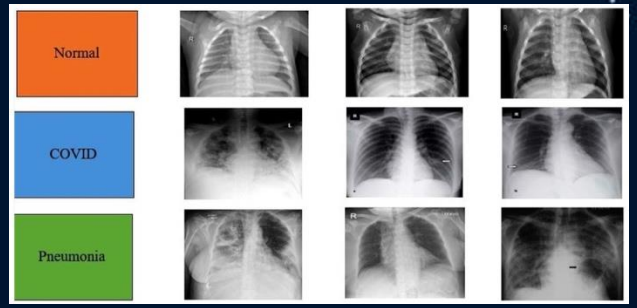
Why is Machine Learning Important?

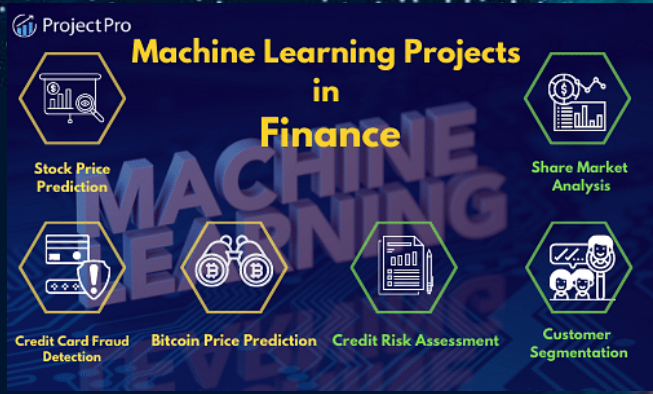
- 1. Decision Making: ML models "allow large volumes of data to be analyzed to detect trends and patterns, enabling more informed, evidence-based decisions”[3].
- 2. Automation: It "automates the development of analytical models through algorithms that interactively learn from data”[4].
- 3. Pattern Recognition: ML "can identify patterns in data that would be difficult to code manually”[5].
- 4. Scalability: It "works efficiently with big datasets and can run on powerful cloud platforms”[6].

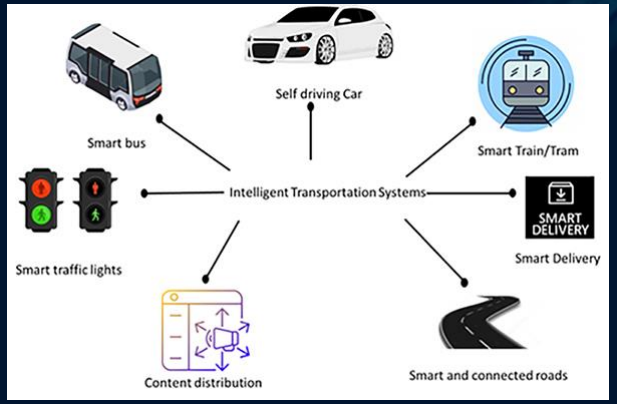


Applications across Industries

- Healthcare: Disease detection, drug discovery
- Finance: Fraud detection, credit scoring
- Retail: Recommendation engines, inventory optimization
- Transportation: Self-driving vehicles, route optimization

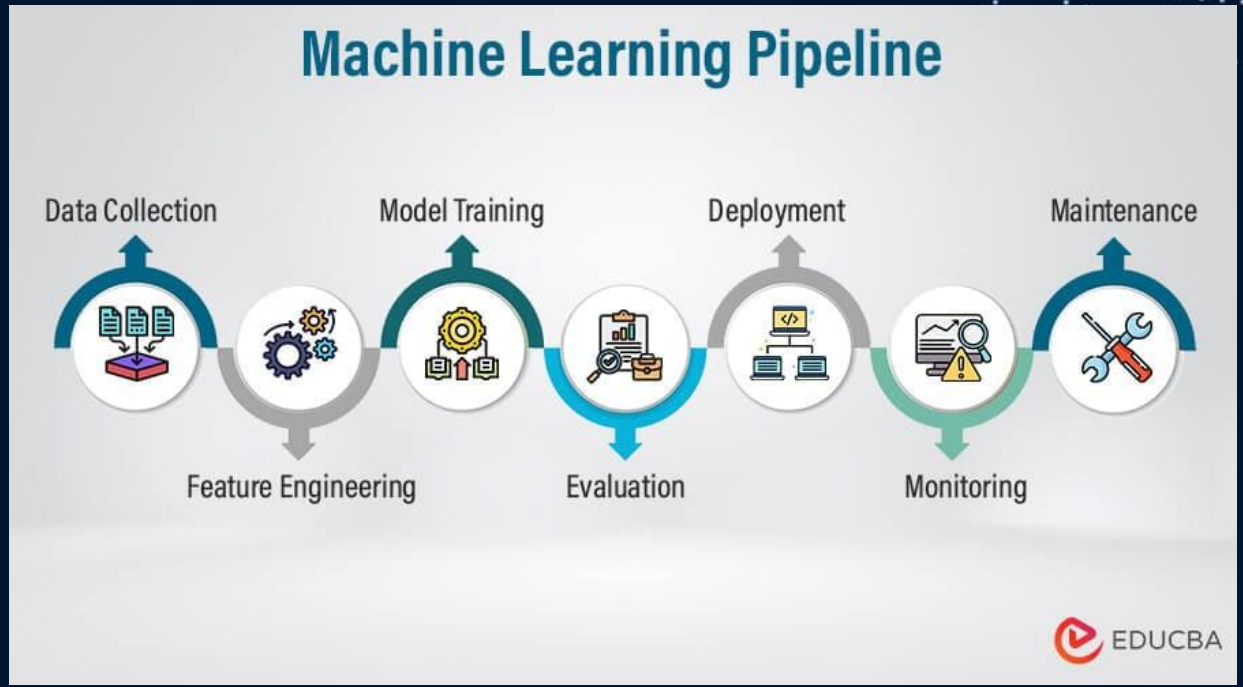






UNIVERSITY OF GREENWICH
Alliance with FPT Education

Data Science/Machine Learning Pipeline

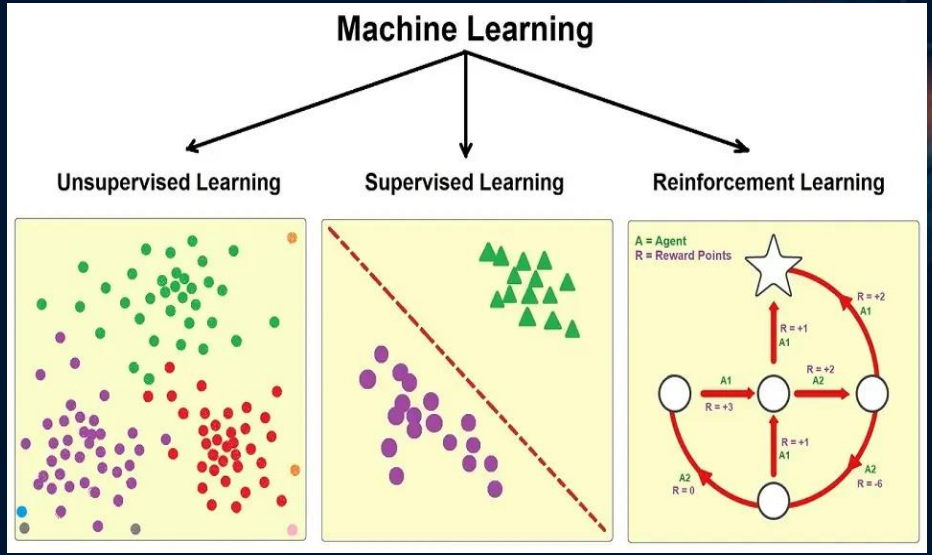


04 Types of ML

Supervised, Unsupervised, Reinforcement

Types of Machine Learning

- Supervised Learning: Labeled data, predict output
  (e.g., spam detection, price prediction)
- Unsupervised Learning: Unlabeled data, find structure
  (e.g., clustering, anomaly detection)
- Reinforcement Learning: Learn by trial & error
  (e.g., game AI, robotics)



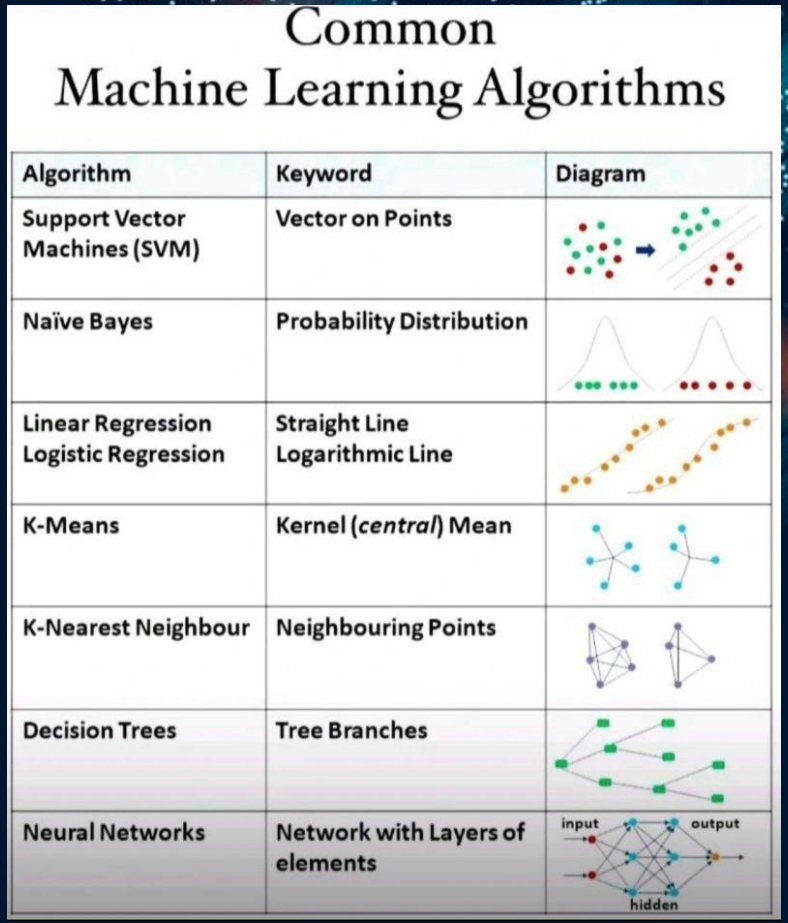
05 Common Algorithms

Regression, Classification, Clustering, ...

UNIVERSITY OF GREENWICH
Alliance with FPT Education

Common ML algorithms

- Linear Regression: Fits a line to predict numeric values
- Logistic Regression: Predicts probability of class (using sigmoid)
- Decision Trees: Splits data by features to make decisions
- k-NN: Looks at closest neighbors to classify/predict
- SVM: Finds the best boundary between classes
- Naive Bayes: Uses probabilities and assumes feature independence
- K-Means: Groups data into clusters based on similarity



06 Python for ML

Python, essential libraries

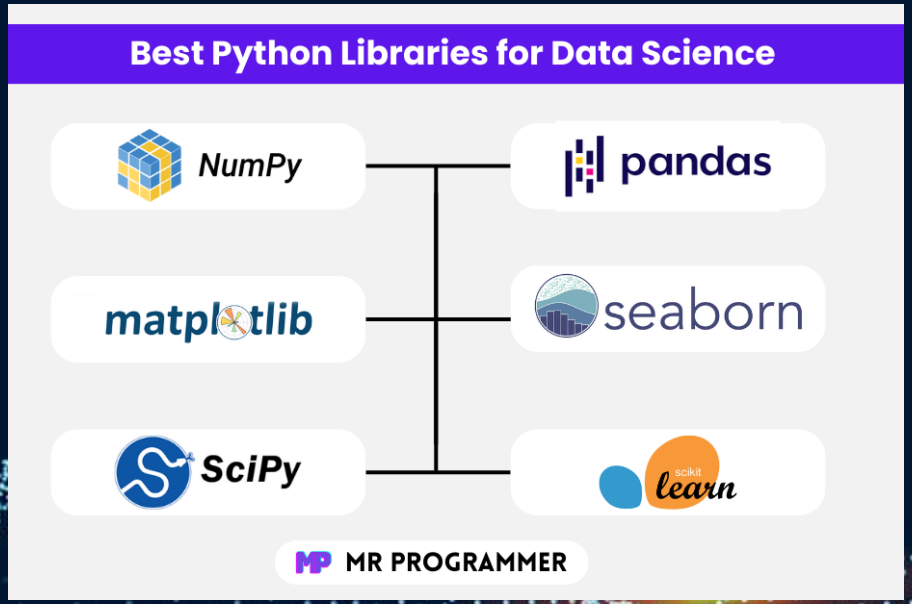
Why Python for ML?

- Why Python?

→ Easy to learn, rich ecosystem, strong ML libraries

- Key libraries:

→NumPy (arrays & math)
→ Pandas (dataframes)
→ Matplotlib (visualization)
→ Scikit-learn (ML algorithms)



UNIVERSITY OF GREENWICH
Alliance with FPT Education

Practical Session

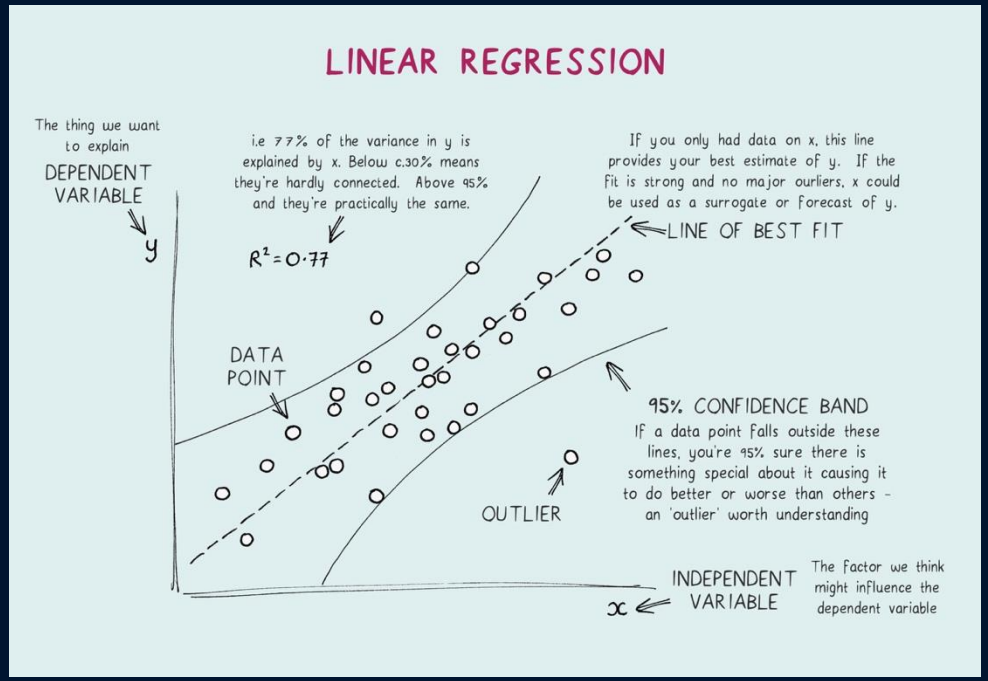
- Access this link: https://drive.google.com/file/d/1nIsyxk1tS60mqnSa2WVze2Z1nSqGwCEx/view?usp=sharing

07 Supervised Learning

Regression, Classification

UNIVERSITY OF GREENWICH
Alliance with FPT Education

Linear Regression



- Definition: Linear regression is a supervised learning algorithm used to predict a continuous target variable based on one or more input features.
- Purpose: Finds the best-fit straight line (regression line) through the data.
- Applications: House price prediction, sales forecasting, risk assessment.

UNIVERSITY OF GREENWICH
Alliance with FPT Education

Model

- $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon$

Where:

- Y is the dependent variable.
- β_0 is the intercept (the value of Y when all X's are zero).
- $\beta_1, \beta_2, \ldots, \beta_p$ are the coefficients for the independent variables $X_1, X_2, \ldots, X_p$.
- ε is the error term.
- $\beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon$

How does it work?

- Goal: Minimize the difference between actual and predicted values
- Loss Function: Mean Squared Error (MSE)

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

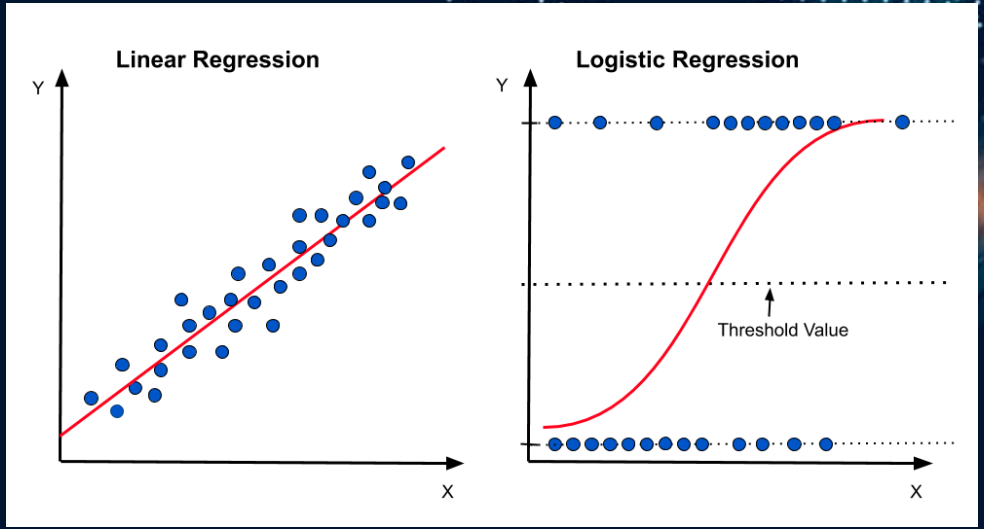
- Optimization: Uses methods like Ordinary Least Squares (OLS) or Gradient Descent to find the best parameters.

Evaluating Model Performance

- R^2 Score (Coefficient of Determination): Proportion of variance explained by the model
- Mean Squared Error (MSE): Average squared difference between actual and predicted values
- Mean Absolute Error (MAE): Average absolute difference between actual and predicted values

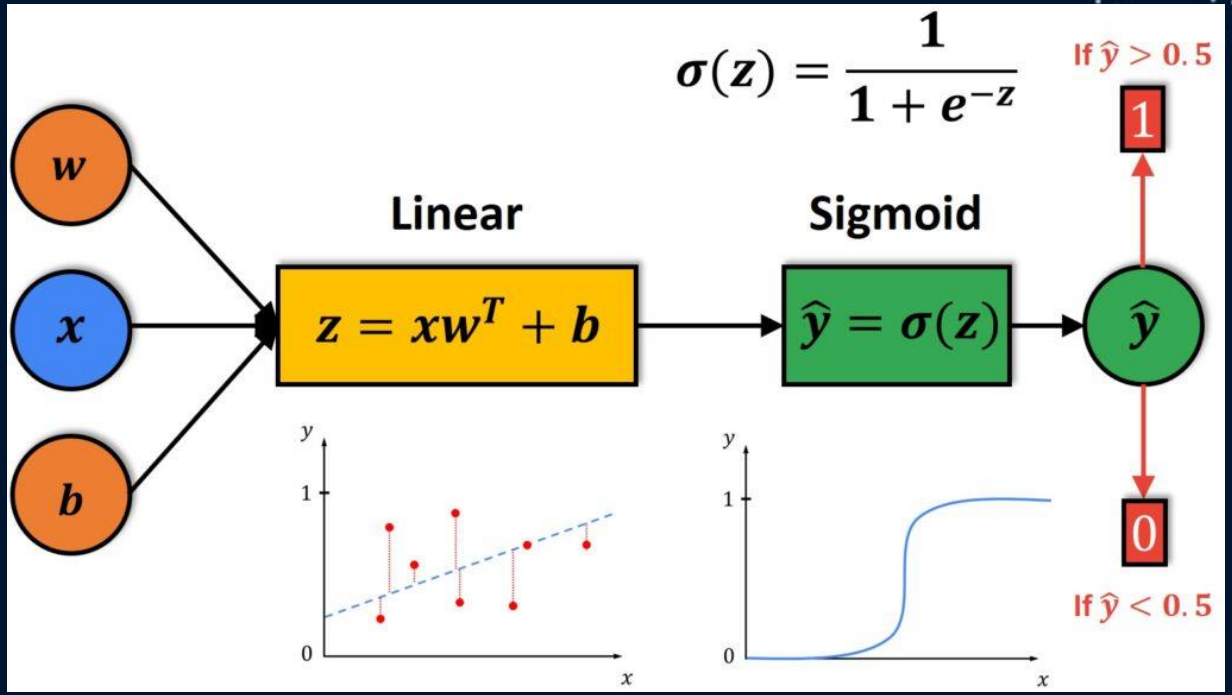
Logistic Regression

- Definition: Logistic regression is a supervised learning algorithm used for classification tasks, predicting the probability that an input belongs to a certain class.
- Purpose: Converts linear combinations of inputs into probabilities using the sigmoid function.
- Applications: Email spam detection, disease diagnosis, customer churn prediction.



UNIVERSITY OF GREENWICH
Alliance with FPT Education

Model



How does it work?

- The process of training a logistic regression model involves finding the optimal values for the coefficients (b0, b1, ... bn) that best fit the training data.
- This is typically done using maximum likelihood estimation (MLE).
- MLE finds the coefficients that maximize the likelihood of observing the given data.

Evaluating Model Performance

- Accuracy: Proportion of correct predictions
- Precision, Recall, F1-score: For imbalanced classes
- Confusion Matrix: True/False Positives/Negatives
- ROC Curve & AUC: Probability thresholds and model discrimination

UNIVERSITY OF GREENWICH
Alliance with FPT Education

References

- [1] https://en.wikipedia.org/wiki/Machine_learning
- [2] https://ischool.syracuse.edu/what-is-machine-learning/
- [3] https://londoncg.com/en/blog/machine-learning-uses-applications-and-importance-of-ml
- [4] https://www.sciencedirect.com/topics/computer-science/machine-learning
- [5] https://www.peerbits.com/blog/machine-learning-vs-traditional-programming.html
- [6] https://www.geeksforgeeks.org/blogs/machine-learning-pipeline/
- [7]

UNIVERSITY OF GREENWICH

Alliance with FPT Education

Thank You !