Summary
Linear Regression is one of the simplest and most widely used algorithms in machine learning, primarily for predicting numerical outcomes. The goal of linear regression is to model the relationship between one or more independent variables and a continuous target variable by fitting a linear equation to observed data.

Highlights
📌 Definition: Linear Regression aims to find the best-fit line that minimizes the difference between the predicted and actual values.

🔍 Types:
◆ Simple Linear Regression: Uses one independent variable to predict the target.
◆ Multiple Linear Regression: Involves multiple independent variables to improve prediction accuracy.

🧮 Mathematical Foundation: The model assumes the relationship can be expressed as $Y = \beta_0 + \beta_1 X + \epsilon$, where:
◆ Y: Target variable
◆ X: Independent variable(s)
◆ β_0, β_1: Model coefficients (intercept and slope)
◆ ϵ: Error term (captures variance not explained by X)

📊 Key Metrics:
◆ R-squared: Measures the proportion of variance in the target variable explained by the model.
◆ Mean Squared Error (MSE): Calculates the average squared differences between predicted and actual values, helping gauge model accuracy.

Key Insights

📈 Interpretability: Linear Regression models are easy to interpret, making them ideal for insights and quick testing on data with a linear relationship.

⚠️ Assumptions:
◆ Linearity: There is a linear relationship between features and the target variable.
◆ Normality: The residuals (errors) are normally distributed.
◆ Homoscedasticity: Constant variance of residuals.
◆ Independence: Observations are independent of each other.

🚫 Limitations:
◆ Not effective for complex relationships that are non-linear or involve intricate interactions.
◆ Sensitive to outliers, which can heavily skew the results.

💼 Applications: Widely used in finance for trend analysis, in healthcare for predictive diagnostics, and in marketing for understanding customer behavior.