Module 2 Summary and Highlights

Congratulations! You have completed this lesson. At this point in the course, you know:

- Regression models relationships between a continuous target variable and explanatory features, covering simple and multiple regression types.
- Simple regression uses a single independent variable to estimate a dependent variable, while multiple regression involves more than one independent variable.
- Regression is widely applicable, from forecasting sales and estimating maintenance costs to predicting rainfall and disease spread.
- In simple linear regression, a best-fit line minimizes errors, measured by Mean Squared Error (MSE); this approach is known as Ordinary Least Squares (OLS).
- OLS regression is easy to interpret but sensitive to outliers, which can impact accuracy.
- Multiple linear regression extends simple linear regression by using multiple variables to predict outcomes and analyze variable relationships.
- Adding too many variables can lead to overfitting, so careful variable selection is necessary to build a balanced model.
- Nonlinear regression models complex relationships using polynomial, exponential, or logarithmic functions when data does not fit a straight line.
- Polynomial regression can fit data but mayoverfit by capturing random noise rather than underlying patterns.
- Logistic regression is a probability predictor and binary classifier, suitable for binary targets and assessing feature impact.
- Logistic regression minimizes errors using log-loss and optimizes with gradient descent or stochastic gradient descent for efficiency.
- Gradient descent is an iterative process to minimize the cost function, which is crucial for training logistic regression models.

Mark as completed