Decoding Data: From Simple Lines to Complex Insights

An Introduction to Data Science, Regression, and Machine Learning

Jack, Collin

Data Science Seminar Series
Peak to Peak Data Science Seminar

Overview

1. Introduction to Data Science
2. Understanding Data Through Regression
3. Introduction to Machine Learning
4. Closing Remarks

Introduction to Data Science

What is Data Science?

- The science of extracting knowledge and insights from data.
- Involves a blend of statistics, computer science, and domain expertise.

Understanding Data Through Regression

Linear Regression Basics

- A statistical method for modeling the relationship between a dependent variable and one or more independent variables.
- Simple linear regression involves a single predictor and a single response variable.

Exploring Multicollinearity

- When two or more predictors in a regression model are correlated, making it difficult to distinguish their individual effects on the dependent variable.

Detecting and Handling Outliers

- Outliers can significantly affect the results of a regression analysis.
- Strategies for handling outliers include removal, transformation, or using robust regression methods.

Introduction to Machine Learning

What is Machine Learning?

- Machine Learning is a subset of artificial intelligence that involves the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

The Machine Learning Process

1. Define the problem.
2. Prepare the data.
3. Choose a model.
4. Train the model.
5. Evaluate the model.
6. Parameter tuning.
7. Make predictions.

Closing Remarks

Summary Next Steps

- Recap of the key points covered today.
- Encourage further exploration and practice with the concepts discussed.

Close

Thank you for your attention!

Any questions?

Jack, Collin

Data Science Seminar Series
Peak to Peak Data Science Seminar