Day 14

Topic Covered: Curve Fitting and Regression Models in R

Summary:

Today's session introduced curve fitting techniques to model relationships between variables. Curve fitting is important for prediction, trend analysis, and understanding how variables interact. The focus was on linear and polynomial regression models, which serve as the basis for many statistical and machine learning methods. These models help describe how one variable changes with respect to another and support forecasting and analytical decision-making.

New Concepts Learned:

- **Linear Model and Polynomial Regression**
  - Used the lm() function to fit linear and polynomial regression models
  - Polynomial regression helps capture non-linear relationships by adding higher-degree terms
  - Model coefficients describe the strength and direction of relationships
- **Model Evaluation Techniques**
  - R-squared values were used to assess how well the model fits the data
  - Residual analysis helped understand model accuracy and errors
  - The predict() function was used to estimate values for new data
- **Applications of Curve Fitting**
  - Forecasting patterns in time-series data
  - Predicting outcomes based on input variables
  - Understanding behavior and trends in experimental or observational datasets

Activity:

- Created sample datasets with numerical variables
- Applied the lm() function for linear and polynomial regression
- Visualized fitted curves using ggplot2 with geom_smooth(method = "lm")
- Computed and interpreted R-squared values and residuals

- Used predict() to generate estimated values for new inputs

Challenges Faced:

Selecting the appropriate polynomial degree required careful consideration. Interpreting higher-degree polynomial coefficients was challenging. Ensuring that the model did not overfit the data was also an important concern.

Key Takeaway:

Curve fitting and regression models help quantify relationships between variables and enable predictions. Linear and polynomial regression are powerful tools for trend analysis, forecasting, and data-driven decision-making in R.