

Performing exploratory
data analysis
, Feature engineering
and Predictive
modeling.

Title Slide:

- Title: Performing Exploratory Data Analysis, Feature Engineering, and Predictive Modeling
- Your Name

• Date **Slide 1: Introduction**

Brief overview of the presentation.

Slide 2:

Exploratory

Data Analysis

(EDA)
Definition and importance of EDA.

Key EDA tasks:

Data visualization

Slide 3: Data Visualization

Title: Data Visualization in
EDA

Types of EDA plots:

Scatter plots

Histograms

Box plots

Heatmaps

Benefits of data visualization.

Slide 4: Data Summary

Statistics

Title: Summary Statistics in
EDA

Mean, median, and mode

Variance and standard
deviation

Skewness and kurtosis

Slide 5: Data Cleaning

- Title: Data Cleaning in EDA
- Handling missing data
- Outlier detection and treatment
- Data transformation (e.g., scaling)

Slide 6: Feature

Engineering

Definition and importance of feature engineering.

Techniques for creating new features:

Feature extraction

Feature transformation

Slide 7: Feature Extraction

- Title: Feature Extraction
- Principal Component Analysis (PCA)
- Text data vectorization (e.g., TF-IDF)
- Time series feature extraction (e.g., lag features)

Slide 8: Feature

Transformation

Transformation

Standardization and
normalization

Log transformation

Binning and discretization

Slide 9: Feature Selection

- Title: Feature Selection
- Methods for feature selection:
- Recursive Feature Elimination (RFE)
- Feature importance from tree-based models
- Univariate feature selection

Slide 10: Predictive

Modeling

Title: Predictive Modeling
Introduction to predictive modeling.

Steps in building a predictive model:

Data splitting

Model selection

Slide 11: Data Splitting

- Title: Data Splitting
- Importance of training and testing data sets.
- Cross-validation techniques.

Slide 12: Model Selection

Title: Model Selection

Types of models:

Linear regression

Decision trees

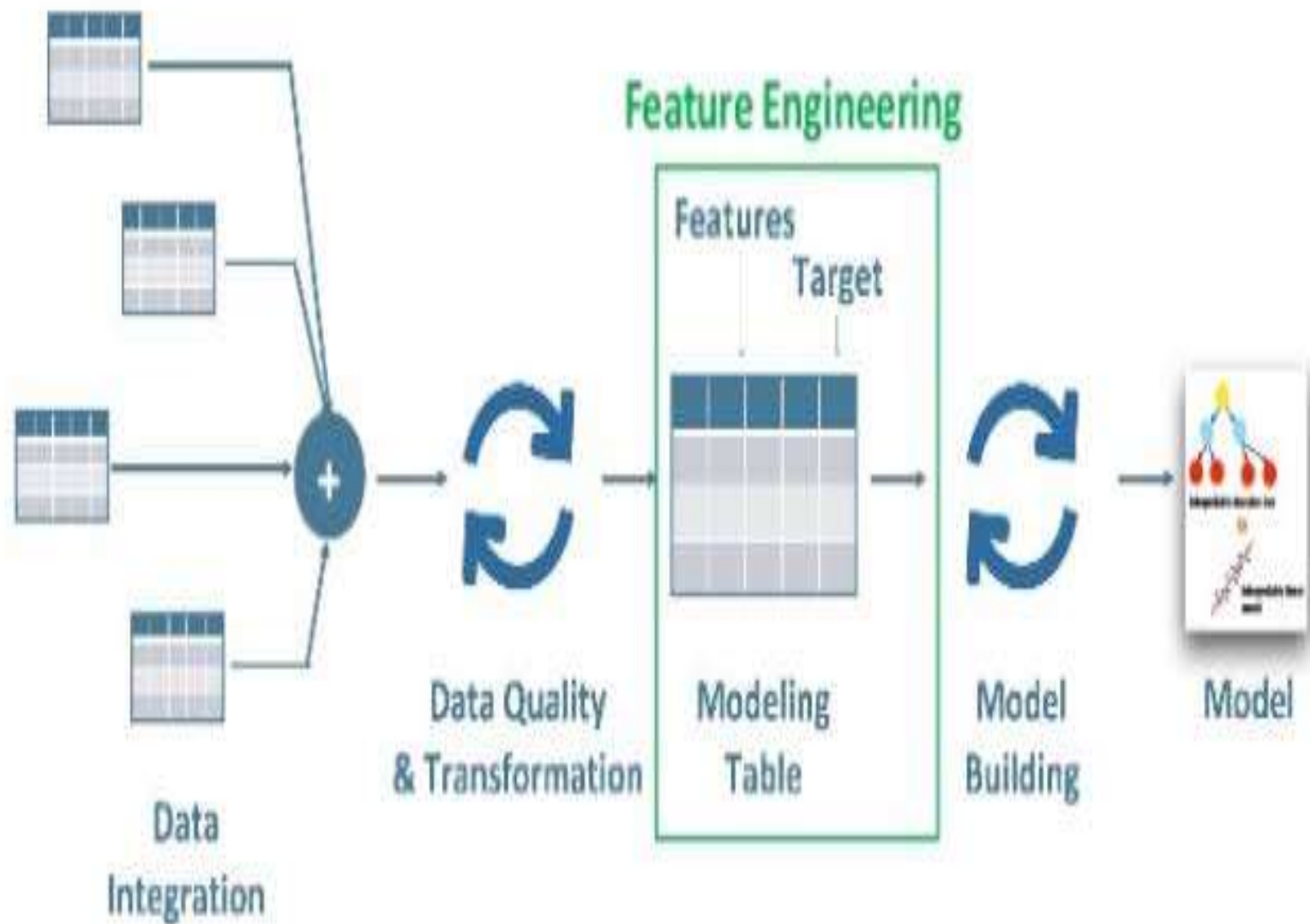
Random forests

Neural networks

Model selection criteria (e.g.,
accuracy, F1 score)

Slide 13: Model Training and Evaluation

- Title: Model Training and Evaluation
 - Training the selected model.
 - Model evaluation metrics:
 - Mean Absolute Error (MAE)
 - Mean Squared Error (MSE)
 - ROC curve (for classification)



Slide 14: Conclusion

- Summarize the key points from each section.
- Highlight the importance of EDA, feature engineering, and predictive modeling in extracting insights from data.