



Co-funded by
the European Union

*Alliance for Fostering
Business and Education Innovation
through Digital Supply Chains*

Data Analytics

Interpretation of the results and extraction of
useful insights





Evaluating Model Performance

- **Accuracy:** Measure of total correct predictions as a proportion of total cases.
- **Precision:** Ratio of true positives to total predicted positives, important in minimizing false positives.
- **Recall:** Ratio of true positives to actual positives, critical where missing a positive is costly.
- **F1 Score:** Harmonic mean of precision and recall, balances both metrics.
- **MSE and MAE:** Measures for quantifying errors in regression models, illustrate average model prediction error.



Feature Importance Analysis



- **Purpose:** Determine the contribution of each feature to the prediction accuracy of the model.
- **Tools and Techniques:** Methods like Gini importance, permutation importance, and model-specific metrics.
- **Actionable Insight:** Identifying key drivers that can be targeted to improve outcomes or reduce risks.



Understanding Model Predictions

- **Explanation Tools:** Use of LIME for local interpretations, SHAP for global insights.
- **Model Transparency:** Explaining individual predictions to stakeholders.
- **Practical Example:** Case study on applying explainability to improve customer satisfaction.



Visualizing Model Behavior



- **Scatter Plots:** Illustrate relationships and distribution between two variables.
- **Heatmaps:** Show correlation among multiple variables at once.
- **Decision Boundaries:** Visualize how model classifications are separated in the feature space.
- **Clustering Analysis:** Reveal natural groupings and patterns that inform model behavior.

Cross-Validation and Generalization



Co-funded by
the European Union



- **Cross-Validation:** Technique to evaluate how the results of a statistical analysis generalize to an independent data set.
- **Model Robustness:** Ensuring the model performs consistently across different datasets.
- **Tuning Complexity:** Adjusting parameters to find the optimal balance between bias and variance.



Analyzing Misclassifications

- **Error Types:** Distinguishing between type I and type II errors.
- **Impact Assessment:** How these errors affect business outcomes.
- **Strategic Corrections:** Adjusting model thresholds or retraining to reduce errors.



Integrating Domain Knowledge



- **Expert Input:** Leveraging industry expertise to interpret model outputs.
- **Contextual Relevance:** Ensuring predictions align with domain-specific expectations.
- **Case Studies:** Examples of successful collaborations that enhanced model utility.



Identifying Patterns and Trends



- **Pattern Recognition:** Techniques to identify recurring themes or anomalies in data.
- **Trend Analysis:** Longitudinal study of data to predict future movements.
- **Strategic Business Decisions:** Utilizing insights from patterns and trends to guide business strategies.

Outlier Detection and Management



Co-funded by
the European Union



- **Detection Techniques:** Statistical tests and visual methods to identify outliers.
- **Influence on Models:** How outliers can skew results and lead to poor model performance.
- **Handling Strategies:** Options include ignoring, transforming, or using robust models.



Iterative Process of ML Analysis

- **Feedback Loops:** Incorporating new data and feedback to refine models.
- **Continuous Learning:** Adapting models in response to evolving data patterns.
- **Innovation:** Encouraging ongoing research and development to stay ahead of market trends.