Hochschule Bremen
City University of Applied Sciences

BIP Machine Learning for Data Science
End2End-Project in cooperation with Fa. Emil Frey

02/10/2024

Agenda

1 Problem description
2 Procedure of the model evaluation
3 Notes on the data

Problem Description:

Task for competition:

- Determination of the expected laid up time of vehicles at the dealership location at the time of purchase.
- Target variable
  - LAID_UP_TIME [days]
- Evaluation metric
  - Root Mean Squared Error (RMSE)

Additional analytical questions:

- On which features/attributes does the laid up time in the car dealership depend? Which features have the biggest impact?
- Which vehicles (in terms of specifications) should be purchased in order to have the shortest possible laid up time at the car dealership?

Procedure of the model evaluation

- Original data
  - Fa. Emil Frey provided a dataset of the last 10-year vehicle sales in their dealerships
  - Data consist of more than 140.000 car sales and more then 100 features
- Data provided:
  - Training data set: ca. 100.000 data records including laid up time
  - Test set: ca. 40.000 data records without laid up time, which should be predicted
- Delivery
  - Model and prediction of the laid up times including columns CHASSIS_NUMBER and LAID_UP_TIME have to be uploaded
- Evaluation
  - These predictions will than compared with the true value by the teachers in order to access the performance of the models

Notes on the data

- **Date columns** in the format YYYYMMDD as integers
- Some features appear both as descriptions (with TYPE, NAME) and as codes (with ID, NUMBER) e.g. color
- Empty attributes were either not recorded or do not apply to the vehicle (e.g., leasing contract date)
- Clarification of features:
  - **SCALED_TOTAL_SALES_PRICE_BASIS**: Reference value for the vehicle's worth. Values have been mapped to the range 0-1
  - **DAY_OF_REGISTRATION:** Day of registration of specific car
  - **VEHICLE_MODEL_ID:** ID for vehicle category
  - **VEHICLE_MODEL_ID_NAME:** Vehicle category (e.g., station wagon, compact car, coupe, etc.)

Thank you for your attention!

Are there any questions?