

2. Modellierung der Zielvariablen Bezahlmodell

Verwendete Features

- trip_eff_speed
- stations_distance
- trip_duration
- age
- starttime_hour
- starttime_month
- starttime_dayofweek
- gender

Preprocessing

- Standard scaling
- One hot encoding

Evaluierung

- Test set size: 20%
- Evaluierungsmetrik :Precision AUC (+ roc AUC) bei Grobauswahl des geeignetsten Estimators für das Problem
- Cross validation + Gridsearch, manuelle Optimierung zur finalen Parametriesierung

Verwendete Classifier

- LogisticRegression
- GradientBoostingClassifier
- RandomForestClassifier
- DecisionTreeClassifier
- MLPClassifier