



## VORHERSAGE VON PRÜFSTANDZEITEN FÜR VARIIERENDE AUTOMOBILKONFIGURATIONEN DURCH MACHINE LEARNING

### PROJEKTZIEL

Noch bevor Fahrzeuge die Fabriken verlassen, müssen Automobilhersteller deren Sicherheit und Zuverlässigkeit durch Prüfstandmessungen gewährleisten. Mit den unterschiedlichen Konfigurationsmöglichkeiten der Fahrzeuge gehen unterschiedliche Testzeiten einher. In der Automobilindustrie handelt es sich bei den Testständen um kritische Engpassressourcen, bei denen es Doppelbelegungen und Leerläufe zu vermeiden gilt.

Ziel des Projektes war deshalb ein Modell zu entwickeln, das präzise die Testzeit bestimmt, welche ein Fahrzeug in der jeweiligen Konfiguration auf dem Prüfstand verbringen wird, sodass die Planbarkeit der Prüfstandbelegungen optimiert werden kann.

### BEREITGESTELLTE DATEN

Es standen zwei Datensätze zur Verfügung: Der Trainingsdatensatz beinhaltete 4.200 Fahrzeugkonfigurationen mit den zugehörigen Testzeiten sowie 400 Features. Der Testdatensatz umfasste 4.200 neue Fahrzeugkonfigurationen mit den gleichen Features. Jedoch fehlten hier die entsprechenden Testzeiten.

### ANGEWANDTE METHODEN

Da der Trainingsdatensatz eine relativ hohe Anzahl von Features beinhaltete, galt es die 400 Features zu reduzieren und zu gruppieren, sodass mit weniger Features die gleiche Aussage getroffen werden konnte. Daher musste zuerst die Datenanalyse-Technik mit den besten Vorhersage-Ergebnissen gefunden werden. Aus diesem Grund wurden die folgenden beiden Ansätze genauer untersucht:

Die Multiple Correspondence Analysis (MCA) wird normalerweise für kategoriale Daten gewählt, um die Struktur im Datensatz zu identifizieren. Es gelang mit ihr ohne Informationsverlust die 400 Features auf insgesamt 14 zu reduzieren. Die Principal Component Analysis (PCA) als alternative Technik der Dimensionsreduktion wurde ebenfalls getestet. Doch bei der PCA konnte die Anzahl lediglich auf 24 wichtige Features verringert werden, sodass die MCA zur Datenreduktion für das anschließend zu entwickelnde Machine Learning-Modell verwendet wurde.

Basierend auf dem reduzierten Datensatz wurden verschiedene Machine Learning-Algorithmen verglichen. Das umfasste auch verschiedene Regressionsalgorithmen. Der Trainingsdatensatz wurde mittels Cross-Validation gesplittet. Schließlich wurde der Gradient Boosting Regressor gewählt, da dieser die besten Ergebnisse lieferte. Nach dem Parameter-Tuning des Algorithmus konnte eine Vorhersage für den ursprünglichen Testdatensatz vorgenommen werden.

## HERAUSFORDERUNGEN

Eine erste Herausforderung lag darin, dass für die große Anzahl an Features verhältnismäßig wenige Datenpunkte zur Verfügung standen und daher mittels Dimensionsreduktion zunächst die wichtigsten bestimmt werden mussten.

Die zweite Herausforderung stellte die Wahl des performantesten Regressionsalgorithmus dar, da sich anwendbare Algorithmen in ihrem Grundsatz ähneln.

## PROJEKTERGEBNIS

Mit dem entwickelten Machine Learning-Modell ist es möglich, die Prüfstandzeiten für alle Fahrzeugkonfigurationen vorherzusagen. Damit können die Testzyklen künftig effizienter ablaufen und sowohl Belegung als auch Taktung der Prüfstände noch besser optimiert werden.

Außerdem konnten die Schlüsselfaktoren identifiziert werden, die einen Einfluss auf die Testzeitlänge eines Fahrzeugs haben.

