



Erstellung einer Auto Machine Learning Pipeline

Bachelor-/Projekt-/Masterarbeit

Aufgrund der immer steigenden Computerleistung, sowie wachsenden Anzahl an Daten gerät Machine Learning immer mehr in den Fokus der Datenanalyse. Hierbei ist es besonders nützlich eine erste Baseline zu haben, anhand der man die Performance seiner Modelle bewerten kann. Hierfür eignen sich besonders die klassischen Machine Learning Methoden (e.g. Random Forest, Support Vector Machines, usw.), aufgrund ihrer einfachen Interpretierbarkeit sowie relativ geringen Rechenkapazität im Vergleich zu Neuronalen Netzen.

Das Ziel dieser Arbeit ist ein Programm zu erstellen, welches mehrere klassische Machine Learning Methoden auf einen beliebigen Datensatz anwenden kann, diese optimiert und das erhaltene Model speichert.

Ziele/Aufgaben

- Automatisierung des Feature Preprocessing, Construction und Selection
- Automatisierung der Model Evaluation und Selektion
- Automatisierung der Hyperparameteroptimierung
- Test der Pipeline anhand typischer Datensätze (e.g. Iris-setosa, MNIST, usw.)

Empfohlene Vorkenntnisse [Optional]

- Programmiersprache Python von Vorteil, aber kein muss.
- Machine Learning Grundlagen von Vorteil, aber kein muss.

Literatur

1. Automated Machine Learning: Methods, Systems, Challenges, Frank Hutter, Lars Kotthoff und Joaquin Vanschoren
2. Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques for Building Intelligent Systems, Aurélien Géron
3. https://www.alibabacloud.com/blog/6-top-automl-frameworks-for-machine-learning-applications-may-2019_595317

Betreuer Kontaktdaten

M.Sc. **Benjamin Haser**
E-mail: benjamin.haser@unibw.de
Tel: +49 89 6004 3257
Building 35/400, Raum 1401

Prof. Dr.-Ing. **Roger Förstner**
E-mail: raumfahrt@unibw.de
Tel (Sek.): +49 89 6004 3570