

# Automatisierung von Machine Learning Workflows



## Motivation

Die Digitalisierung der Gesellschaft ist eine der wichtigsten aktuellen Entwicklungen und erzeugt einen rapide wachsenden Datenstrom. Um daraus entscheidungsrelevante Informationen zu gewinnen, werden zunehmend Verfahren des Maschinellen Lernens (ML) genutzt, da der Datenzuwachs zu dynamisch für herkömmliche hypothesengestützte Analysen ist.

Bei der Entwicklung von ML Lösungen stehen dabei eine wachsende Anzahl von Methoden und Verfahren zur Verfügung, für die Entwurfsentscheidungen getroffen und Parameter eingestellt werden müssen. ML-Spezialisten stehen deshalb vor folgenden Herausforderungen:

- No-Free-Lunch Theorem: Es gibt kein bestes Verfahren und somit auch keinen besten ML-Workflow für jede Situation
- Inductive Bias: Bei unbekannter Verteilung der Daten ist a-priori eine zuverlässige Aussage über die Performanz eines ML-Verfahrens schwer möglich.
- Wissenslücke: Das Verhalten zahlreicher etablierter ML-Verfahren ist unklar, insbesondere unter welchen Bedingungen bei gegebenen Daten eine gute Performanz zu erwarten ist.

## Lösung: CODA

Im CODA Projekt geht es um Grundlagenforschung zur Automatisierung von Algorithmenauswahl und Hyperparameteroptimierung im Entwicklungsprozess von Lösungen des Maschinellen Lernens. Die dabei gewonnenen Erkenntnisse und Lösungen stehen in Form eines Software Frameworks zur Verfügung, damit sich die Nutzer auf die verbleibenden kreativen Aspekte der Entwicklung fokussieren können.

### PROJEKT-STECKBRIEF

#### PROJEKTTITEL

CODA - Cognitive Data Analytics Framework

#### WORUM GEHT ES?

Ziel des CODA Projektes ist die Grundlagenforschung im Bereich der Algorithmenauswahl und Hyperparameteroptimierung. Auf Basis der gewonnenen Erkenntnisse entsteht ein prototypisches Software Framework zur Teilautomatisierung des Entwicklungsprozesses von Lösungen des maschinellen Lernens.

#### LAUFZEIT

von 01.01.2017 bis 31.12.2019

#### ANSPRECHPARTNER

Dr. Fikret Sivrikaya  
E-MAIL [fikret.sivrikaya@gt-arc.com](mailto:fikret.sivrikaya@gt-arc.com)  
TEL +49 (0) 30-314 74010

#### URL



<http://www.gt-arc.com>

#### SCHLAGWORTE

Maschinelles Lernen, Meta Lernen, Hyperparameter Optimierung, Modellwahl, Big Data, Smart Data, Datenanalyse

#### AUFTRAGGEBER

 Bundesministerium  
für Bildung  
und Forschung