



Effiziente Parallelisierung eines neuen Machine-Learning-Frameworks in Python

Maschinelles Lernen ist ein mächtiges und in der modernen Informatik häufig genutztes Werkzeug zur Realisierung komplexer Funktionen. Dank der Fähigkeit aus vorgegebenen Beispielen zu lernen ermöglichen Machine-Learning-Verfahren eine Vielzahl innovativer Anwendungen. In der Abteilung Technische Informatik wird ihr Einsatz unter anderem in Brain-Computer-Interfaces und zur biomedizinischen Datenanalyse erforscht.

Konstruktivistisches maschinelles Lernen (kML) ist ein neuer Ansatz maschinellen Lernens, mit dem gleichzeitig mehrere konkurrierende Klassifizierer für eine gegebene Fragestellung trainiert und evaluiert werden. Anders als bei klassischen Verfahren werden dabei auch zeitliche Aspekte abgebildet. Die trainierten Klassifizierer werden in einer hierarchisch geordneten Wissensbasis verwaltet und können kontinuierlich modifiziert werden.

Im Rahmen der Masterarbeit soll ein bestehender kML-Prototyp (implementiert in Bash/R) nach Python portiert und parallelisiert werden. Insbesondere soll so eine modularisierte Einbindung von Algorithmen des Machine-Learning-Packages Scikit-learn ermöglicht werden. Ziel der Arbeit ist neben der Implementierung eines Python-Packages auch dessen Evaluation sowie eine Dokumentation des Application Program Interface (API).

Die schriftliche Ausarbeitung kann auf Deutsch oder Englisch verfasst werden.

Anforderungen:

- Programmiererfahrung in Python/Scikit-learn
- Grundkenntnisse in maschinellem Lernen

Ansprechpartner:

Dr. Thomas Schmid
Paulinum, Raum P533
schmid@informatik.uni-leipzig.de