



Abschlussarbeiten im Bereich „Transparenz intelligenter Systeme“

Zielsetzung der Arbeit

Am Fachgebiet Wirtschaftsinformatik beschäftigen wir uns u.a. mit den Auswirkungen innovativer digitaler Technologien auf Unternehmen und Gesellschaft. Da Künstliche Intelligenz vielerorts als „Basistechnologie des 21. Jahrhunderts“ bezeichnet wird, findet sie zunehmend Einzug in Informationssysteme aller Art. Haupttreiber der aktuellen Entwicklungen ist dabei die zunehmende Verbreitung statistischer Verfahren des Maschinellen Lernens (ML). Oftmals sind diese Modelle jedoch von sehr komplexer Natur und stellen für Menschen sogenannte Black-Box-Systeme dar. Dies bedeutet, dass wichtige Merkmale intelligenter Systeme, z.B. verwendete Daten und Gewichtungen für getroffene Vorhersagen, für Anwender nicht nachzuvollziehen sind.

In der Information Systems (IS) Forschung interessiert uns vor allem, wie sich die Transparenz intelligenter Systeme auf die Akzeptanz und Interaktion mit intelligenten Systemen auswirkt. Dafür soll im Rahmen eines Forschungsprojekts ein System entwickelt werden, das moderne Verfahren für die Interpretation von ML-Modellen für Anwender bereitstellt. Ein besonderer Fokus soll hierbei auf der Benutzererfahrung liegen.

Zu vergebende Abschlussarbeiten

- Thema 1: Interpretationsverfahren für ML-Anwendungen im Bereich Textklassifizierung
- Thema 2: Interpretationsverfahren für ML-Anwendungen im Bereich Bildklassifizierung
- Thema 3: Interpretationsverfahren für ML-Anwendungen basierend auf strukturierten Daten

Methodik / Vorgehensweise

- Literaturrecherche hinsichtlich Interpretationsverfahren für den jeweiligen Anwendungsfall
- Prototypische Entwicklung eines oder mehrerer Interpretationsverfahren

Empfohlenes Vorwissen

- Mathematische Grundlagen im Bereich Maschinelles Lernen und/oder Künstlicher Intelligenz
- Erfahrung im Umgang mit gängigen ML-Frameworks wie *TensorFlow*, *PyTorch* (für Themen 1 & 2) oder *Scikit-learn* (Thema 3)

Beginn/Betreuer

Beginn ab sofort.

Bei Interesse bitte mit konkreter Themenpräferenz, Lebenslauf und Leistungsspiegel melden bei:

Felix Peters

(peters@is.tu-darmstadt.de)