

2. Workshop "Verfahren zur Interpretierbarkeit von neuronalen Netzen" - Dienstag, 21. Juni 2022, 9:30 bis 11:30 Uhr

Termin: Dienstag, 21. Juni 2022, 9:30 bis 11:30 Uhr

Das Zustandekommen der Ergebnisse von neuronalen Netzen ist häufig selbst für Expert*innen nur schwer nachvollziehbar. Für eine sichere und verantwortungsvolle Anwendung von neuronalen Netzen in sicherheitskritischen Bereichen ist eine Erklärung jedoch essenziell notwendig. Deswegen sind Verfahren zur Interpretierbarkeit von neuronalen Netzen ein aktives Forschungsfeld. Ziel ist es neben dem Erklären von Ergebnissen auch falsches Modellverhalten aufdecken zu können. Dazu sollen in diesem Workshop relevante Methoden vorgestellt und konkrete Werkzeuge zur Umsetzung von mehr Interpretierbarkeit diskutiert werden.



Interessierte Teilnehmer registrieren sich bitte für die Veranstaltung unter: [ANMELDUNG](#) (Log-in Code: neuro-net2)

9:30 - 9:35 Uhr	Begrüßung und Rückblick auf den letzten Workshop	
9:35 - 10:20 Uhr	Regellernen zur Schwachstellenanalyse und Erklärbarkeit von Black-Box Modellen	Dr. Daniel Becker, Fraunhofer IAIS
10:20 - 10:50 Uhr	Advantages, properties and limitations of model agnostic explainability techniques in Machine Learning	Dr. Antoine Gautier, QuantPi
10:50 - 10:55 Uhr	<i>Pause</i>	
10:55 - 11:25 Uhr	Diskussionsrunde	
11:25 - 11:30 Uhr	Verabschiedung / Ausblick zu kommenden Workshops	

Bei Fragen stehen wir Ihnen sehr gerne zur Verfügung.

Dr. Maximilian Poretschkin

Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS
E-Mail: maximilian.poretschkin@iais.fraunhofer.de
Telefon: +49 2241 14-1984

Rebekka Göрге

Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS
E-Mail: rebekka.goerge@iais.fraunhofer.de

Alexander Zimmermann

DIN
E-Mail: alexander.zimmermann@din.de
Telefon: +49 30 2601-2153

gefördert durch:

Gefördert durch:

Ministerium für Wirtschaft,
Industrie, Klimaschutz und Energie
des Landes Nordrhein-Westfalen

