

## IDSR

Integration von Domänenwissen in Symbolische Regression

|                                 |   |                        |            |
|---------------------------------|---|------------------------|------------|
| <b>Programm / Ausschreibung</b> | Dissertationen FH OÖ, Dissertationsprogramm FH OÖ, Dissertationsprogramm der FH OÖ 2023 | <b>Status</b>          | laufend    |
| <b>Projektstart</b>             | 01.10.2023  | <b>Projektende</b>     | 31.10.2025 |
| <b>Zeitraum</b>                 | 2023 - 2025   | <b>Projektlaufzeit</b> | 25 Monate  |
| <b>Keywords</b>                 | Scientific Machine Learning, Symbol Regression, Physics-inspired Machine Learning       |                        |            |

### Projektbeschreibung

Ausgangssituation und Motivation:

Auf Grund ihrer hohen Komplexität, lässt sich das Verhalten von Systemen und Prozessen in kritischen Anwendungsbereichen wie etwa der Medizintechnik, dem Finanzsektor oder dem Energiesektor, nicht allein aus theoretischen Überlegungen und physikalischen Gegebenheiten modellieren. Um mit der steigenden Anforderung nach Effizienz und Genauigkeit sowie mit dem zunehmenden Umfang an Simulationen und Experimenten Schritt zu halten, ist Daten-basierte Modellierung inzwischen etabliert. Dabei werden Algorithmen für maschinelles Lernen eingesetzt, um Modelle zu trainieren, welche aus Eingabedaten genaue Vorhersagen generieren. Jedoch werden diese Modelle oft durch hochkomplexe Funktionsformen (z.B. neuronale Netze) dargestellt, deren innere Strukturen und Funktionsweisen nicht leicht zu verstehen sind. Ein noch offenes Thema von hohem Interesse ist die Entwicklung von Daten-basierten Modellen, welche den zugrunde liegenden funktionalen Prinzipien folgen (physikalische Gesetze abbilden) und ein gutes Extrapolationsverhalten aufweisen. Aus diesem Grund konzentrieren sich viele aktuelle Arbeiten auf neue Möglichkeiten zur Inklusion von Domänenwissen in die Modellierung. In diesem Zusammenhang gewinnt das Thema der symbolischen Regression zunehmend an Bedeutung, da auf Grund der Modelldarstellung als kurze Formel die Möglichkeit zur Integration von physikalischen Gesetzen einfach möglich ist.

Ziele und Innovationsgehalt:

Das angestrebte Ziel dieser Arbeit ist es Algorithmen und Methoden zu entwickeln, welche vorhandenes Domänenwissen in den Daten-basierten Modellierungsprozess miteinbeziehen, um plausiblere, robustere und extrapolationsfähigere Modelle zu finden. In diesem Kontext spielt die Unsicherheit der Modelle eine wesentliche Rolle, deshalb sollen die entwickelten Methoden zusätzlich Unsicherheitsintervalle für die Prognosen liefern.

Angestrebte Ergebnisse:

Neben den entwickelten Algorithmen und theoretischen Analysen sollen als Ergebnis dieser Arbeit, Algorithmen implementiert und als open source Software veröffentlicht werden. Die Dissertation wird fertig gestellt und es werden zwei Konferenzbeiträge und ein Journalartikel veröffentlicht.

## **Projektpartner**

- FH OÖ Forschungs & Entwicklungs GmbH