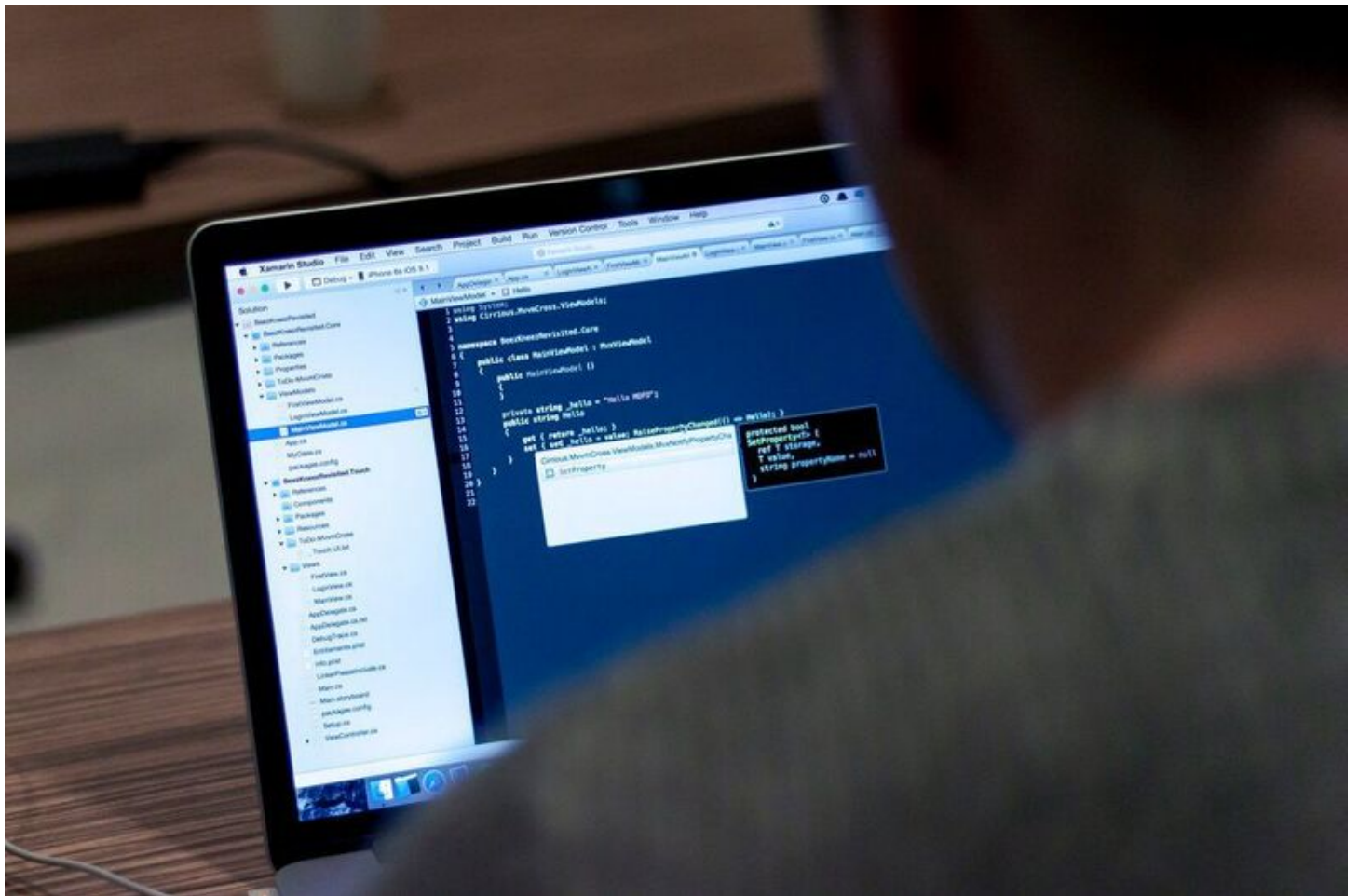


Induktive Programmierung

# Maschinelles Lernen ohne großen Datensatz ermöglichen

16.01.2020 | Autor / Redakteur: Katharina Juschkat / [Nico Litzel](#)

Die Informatikerin Ute Schmid forscht an der Methode der induktiven Programmierung – einem Ansatz, der es ermöglicht, ein System des maschinellen Lernens ohne großen Datensatz aufzubauen. Vor allem in der Medizin und der Industrie soll die Methode nützlich sein.



Maschinelles Lernen benötigt eigentlich einen großen Datensatz. Die Informatik-Professorin Ute Schmid entwickelt einen alternativen Ansatz. (Bild: gemeinfrei / Pixabay <<https://pixabay.com/de/service/terms/#license>> )

Um [maschinelles Lernen](https://www.bigdata-insider.de/was-ist-machine-learning-a-592092/) <<https://www.bigdata-insider.de/was-ist-machine-learning-a-592092/>> erfolgreich umzusetzen, braucht es vor allem massenweise Daten. Denn nur mit genügend Trainingsmaterial erkennt ein Algorithmus Gesetzmäßigkeiten und kann Regeln ableiten. Was aber, wenn es zu wenige Daten gibt?

## Maschinelles Lernen mit wenigen Daten



Die Informatik-Professorin Ute Schmid lehrt Angewandte Informatik an der Universität Bamberg und ist auch diplomierte Psychologin. Sie forscht im Bereich Künstliche Intelligenz.

(Bild: Fortiss)

Die Informatik-Professorin Ute Schmid entwickelt genau für diese Fall eine neue Methode. **Induktive Programmierung heißt die Methode**, mit der sich die Leiterin der Forschungsgruppe Kognitive Systeme der Universität Bamberg beschäftigt. Sie erforscht, wie Programme aus Ein-/Ausgabe-Beispielen oder beobachteten Eingabefolgen in einem System lernen. Die Methode ergänzt aktuelle Blackbox-Ansätze des „Deep Learning <<https://www.bigdata-insider.de/was-ist-deep-learning-a-603129/>>“, einem Teilgebiet des Maschinellen Lernens, das künstliche **neuronale Netze** <<https://www.bigdata-insider.de/was-ist-ein-neuronales-netz-a-686185/>> einsetzt und viele Trainingsdaten benötigt. Außerdem ermöglicht sie, Hintergrund- und Expertenwissen zu berücksichtigen.

## Mögliche Anwendungsfelder der induktiven Programmierung

Anwendungsfelder solcher Verfahren sind etwa die medizinische Diagnose oder die industrielle Produktion, weil es dort oft nur wenige Daten gibt, um ein Modell zu trainieren.

Ein weiterer Anwendungsbereich wäre das **sogenannte „Data Wrangling“**. Das Data Wrangling, oder die Datenaufbereitung, beschreibt den Vorgang, **Rohdaten** <<https://www.bigdata-insider.de/was-sind-rohdaten-a-920701/>> zu strukturieren, bereinigen und zu validieren, um sie für den Einsatz in Systemen des maschinellen Lernens vorzubereiten.

## Rohdaten nur mit Interaktion von Experten interpretierbar

In manchen Anwendungsbereichen, etwa der medizinischen Diagnose, gibt es oft gar keine **sogenannte „ground truth“** – also keine große Datenbasis. Das heißt, auch ein Experte kann nur nach bestem Wissen einschätzen, welche Diagnose korrekt wäre. „Es geht um die Kombination von induktiver Programmierung mit End-to-end-Ansätzen – also Deep-Learning-Ansätzen, bei denen direkt aus den Rohdaten, etwa aus Bildern, gelernt wird“, erläutert Prof. Ute Schmid. Dies setze voraus, dass semantisch interpretierbare Informationen aus den neuronalen Netzen extrahiert werden, was teilweise nur in Interaktion mit menschlichen Experten möglich ist. „Ich gehe davon aus, dass es für viele praktische Anwendungen ohnehin notwendig wird, interaktive Ansätze des maschinellen Lernens zu nutzen oder zu entwickeln, da es gar nicht möglich oder sehr teuer ist, Trainingsdaten ohne menschliche Hilfe korrekt zu kennzeichnen.“

**Methoden der induktiven Programmierung ermöglichen es, Hintergrund- und Expertenwissen beim Lernen zu berücksichtigen, was ein Lernen mit wenigen Daten ermöglichen soll.**

Prof. Schmid forscht an der induktiven Programmierung im Rahmen des **Leitprojektes „Robuste KI“** am **Forschungsinstitut Fortiss** <<https://www.fortiss.org/forschung/projekte/detail/robuste-ki>> .

Dieser Artikel stammt von unserem Partnerportal **weAutomate.de** <<https://www.weautomate.de/>> .