

Ein LAMA geht online: Wie Machine-Learning-Kompetenzen durch praxisnahe Problemstellungen gefördert werden können

Gabriela Molinar¹, Marco Stang², David Lohner³, Manuela Schmidt⁴

Abstract: Dieser Artikel beschreibt die Neukonzeption einer prämierten Lehrveranstaltung zur Förderung von Machine-Learning-Kompetenzen. Ein Präsenzseminar des Labors für Angewandte Machine Learning Algorithmen wird in einen Massive Open Online Course (MOOC) umgewandelt. Dabei werden interaktive Notebooks (Jupyter Notebooks) eingesetzt, um Programmieraufgaben für eine größere Anzahl an Teilnehmenden als in Präsenz zu ermöglichen. Der Lehrveranstaltung liegt der QUA³CK-Prozess zu Grunde, der ein besonderes Augenmerk auf die Modellbildung für Maschinelles Lernen legt.

Keywords: Maschinelles Lernen, interaktive Notebooks, praxisnahe Aufgaben, MOOC

Die Vermittlung von Machine Learning (ML) Methoden wird am Karlsruher Institut für Technologie (KIT) bereits in der Lehrveranstaltung „Labor für Angewandte Machine Learning Algorithmen“ (LAMA) für Studierende der Fakultät Elektrotechnik und Informationstechnik realisiert. Nach einem Einführungsteil, auf der Basis inhaltlicher Wissensvermittlung und eng geführten Übungsaufgaben, haben Lernende die Möglichkeit, Kompetenzen durch forschendes Lernen⁵ anhand realer Szenarien zu entwickeln. Diese Projektphase wird „Into-the-Wild“ genannt, weil die Lernenden eine Problemstellung frei wählen dürfen und selbständig entscheiden, welchen Algorithmus sie für konkrete Fälle anwenden oder wie sie Daten verarbeiten. Das eigens entwickelte QUA³CK-Strukturmodell⁶ zeigt dabei die einzelnen Arbeitsschritte der Modellbildung und begleitet den gesamten Lernprozess.

¹ Karlsruher Institut für Technologie (KIT), Institut für Technik der Informationsverarbeitung (ITIV), Engesserstr. 5, 76131 Karlsruhe, gabriela.molinar@kit.edu, <https://orcid.org/0000-0003-0875-4125>

² Karlsruher Institut für Technologie (KIT), Institut für Technik der Informationsverarbeitung (ITIV), Engesserstr. 5, 76131 Karlsruhe, marco.stang@kit.edu

³ Karlsruher Institut für Technologie (KIT), Zentrum für Mediales Lernen (ZML), Karl-Friedrich-Str. 17, 76133 Karlsruhe, david.lohner@kit.edu, <https://orcid.org/0000-0003-2715-1119>

⁴ Karlsruher Institut für Technologie (KIT), Zentrum für Mediales Lernen (ZML), Karl-Friedrich-Str. 17, 76133 Karlsruhe, manuela.schmidt@kit.edu

⁵ Kerres, M.; de Witt, C.: Editorial: Theoriebildung in Mediendidaktik und Wissensmanagement. *MedienPädagogik: Zeitschrift für Theorie Und Praxis Der Medienbildung*, 6 (Theoriebildung in Mediendidaktik). 2002. <https://doi.org/10.21240/mpaed/06/2002.10.25.X>

⁶ Stock, S.; Becker, J.; Grimm, D.; Hotfilter, T.; Molinar, G.; Stang, M.; Stork, W.: QUA³CK - A Machine Learning Development Process. Proc. Artificial Intelligence for Science, Industry and Society (AISIS2019), 2020.

Der kompetente Umgang mit Machine Learning Algorithmen stellt eine wichtige Qualifikation nicht nur für jeden dar, der im Bereich der Informatik tätig ist, sondern auch für ein breiteres Publikum, z. B. für Führungskräfte verschiedener Branchen, die ein grundlegendes Verständnis der Thematik benötigen. Die mit einem Lehrpreis ausgezeichnete⁷ Präsenz-Veranstaltung, LAMA, soll nun durch die Umwandlung in den Massive Open Online Course AMALEA für die Allgemeinheit zugänglich gemacht werden.

Da der neue Online-Kurs im Vergleich zu anderen derzeit existierenden deutschsprachigen Online-Kursen im Bereich Machine Learning eine größere Anwendungsorientierung bietet, besteht die größte Herausforderung in der Skalierung der interaktiven Aufgaben – insbesondere bei der frei wählbaren Projektarbeit im „Into-the-Wild“-Teil.

Der Online-Kurs AMALEA wird auf der Lernplattform KI-Campus abgebildet. Die Kursthemen sind die Verarbeitung und Analyse von Datensätzen, Bewertung von ML-Systemen, Grundlagen des überwachten und unüberwachten Lernens, neuronale Netze, konvolutionale neuronale Netze und rekurrente neuronale Netze.

Der Inhalt der Veranstaltung ist so organisiert, dass er verschiedenen Zielgruppen ermöglicht, sich die Inhalte in unterschiedlicher Tiefe zu erarbeiten. Die ersten zwei Kapitel adressieren ein breiteres Publikum und erlauben zu verstehen, wie ML Probleme und ihre Lösungen aussehen sowie welche Schritte bis zur Identifikation eines Lösungsansatzes erfolgen müssen. Die weiterführenden Kapitel des ersten Moduls sowie das zweite Modul stehen Studierenden, Entwickler:innen und bereits vorqualifizierte Mitarbeiter:innen zur Verfügung, um die ML Methoden über die konkrete Anwendung in der Tiefe zu erlernen. Nach sechs Kurswochen Grundlagen beginnt die „Into-the-Wild“-Phase. Hierbei steht die selbständige Projektarbeit in den Jupyter Notebooks im Vordergrund. Ziel ist es, dabei das Wissen aus dem geführten Teil aktiv anzuwenden und daraus einen Lösungsansatz zu entwickeln und bewerten. Das projektbasierte Modul wird auf Basis des gewählten Lösungsansatzes und mit dem Peer-Review-Verfahren bewertet.

Die Erfahrung mit LAMA zeigt, dass die Motivation von Studierenden, sich mit ML auseinanderzusetzen, sehr hoch ist. Nun soll die Zielgruppe durch AMALEA in einem zweistufigen Modell vergrößert werden: Die ersten zwei Wochen sprechen die Allgemeinheit an, weitere vier Wochen Interessierte, die tiefer in die ML-Algorithmen einsteigen möchten. Die letzten vier Wochen des Kurses („Into-the-Wild“) richten sich an Teilnehmer:innen, die die gelernten Methoden selbständig für anwendungsnahe Problemstellungen einsetzen wollen.

⁷ https://media.bibliothek.kit.edu/world/2020/DIVA-2020-251_mp4.mp4