

Verteilte Systeme für mobiles Maschinelles Lernen

BACHELORARBEIT

für die Prüfung zum

Bachelor of Science

des Studiengangs Informatik / Angewandte Informatik

an der

Dualen Hochschule Baden-Württemberg Karlsruhe

von

Jens Döllmann

Abgabedatum 30. August 2021

Bearbeitungszeitraum

3 Monate

Matrikelnummer

8876462

Kurs

TINF18B4

Betreuer der Ausbildungsfirma

Florian Seiter

Gutachter der Studienakademie

Gerhard Wolf

Erklärung

Ich versichere hiermit, dass ich meine Bachelorarbeit mit dem Thema: „Verteilte Systeme für mobiles Maschinelles Lernen“ selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich versichere zudem, dass die eingereichte elektronische Fassung mit der gedruckten Fassung übereinstimmt.

Ort Datum

Unterschrift

Zusammenfassung

Abstract

Inhaltsverzeichnis

1	Einleitung	1
1.1	Was ist Maschinelles Lernen?	2

Kapitel 1

Einleitung

Hören die meisten Personen die Begriffe „Künstliche Intelligenz“ und „Maschinelles Lernen“ stellen sie sich Roboter vor: treue Diener, die all deine Aufgaben übernehmen oder eine tödliche Arme von superintelligenten Cyborgs im Kampf gegen die Menschheit. Künstliche Intelligenz (kurz KI) ist jedoch schon lange nicht mehr nur eine futuristische Fantasie und Teil von Science-Fiction-Filmen; sie wird bereits großflächig verwendet und findet Einsatz in einer Vielzahl von Anwendungen aus Bereichen in nahe zu allen Teilen der Wirtschaft. Andrew Ng, welcher unter anderem durch seine beliebten Onlinekurse zum Thema Maschinellen Lernen (ML) und Deep Learning (DL) bekannt ist,¹ vergleicht KI mit der neuen Elektrizität.

„Just as electricity transformed almost everything 100 years ago, today I actually have a hard time thinking of an industry that I don't think AI will transform in the next several years“

— Andrew Ng (Lynch 2017)

Die Anwendung, mit der ML erstmalig öffentliche Aufmerksamkeit erlangt und das tägliche Leben vieler Menschen beeinflusst, stammt aus den 90er-Jahren: Die Rede ist von den E-Mail-Spamfiltern (Géron 2017, S. 1). Es handelt sich um eine scheinbar leichte Aufgabe, welche mit traditioneller Programmierung aber dennoch nur schwer gelöst werden kann. Ein klassisches Programm besteht aus vielen statischen Regeln, im Fall des Spamfilters können diese dazu dienen, auffällige Schlüsselwörter und Satzteile zu erkennen (z. B. „Kreditkarte“, „kostenlos“, „kaufen“ usw). Ein statisches Programm ist nicht gut dafür

¹<https://www.coursera.org/instructor/andrewng>

geeignet, ein dynamisches Problem zu beschreiben. Die Absender der Spammnachrichten könnten erkennen, welche E-Mails blockiert werden und diese leicht abändern. Arbeiten sie so um das Problem herum, müssen immer wieder neue Regeln im Programm aufgenommen werden. Ein Spamfilter hingegen, welcher auf Maschinellen Lernen basiert, kann diese Regeln selbstständig erkennen. Im Training verwendet der Algorithmus hierfür Beispiel E-Mails (z. B. diese, die vom Benutzer markiert wurden), diese Daten nennt man dann den Trainingsdatensatz. Es folgt ein Programm, welches weitaus kürzer, einfacher zu verwalten und wahrscheinlich auch präziser ist.

1.1 Was ist Maschinelles Lernen?

Maschinelles Lernen ist die Wissenschaft (und Kunst) der Programmierung, die es dem Computer ermöglicht, von Daten zu lernen. Arthur Samuel und Tom Mitchell haben im Jahr 1959 und 1997 eine allgemeine und formale Definition gegeben (Géron 2017, S. 2):

„Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.“

— Arthur Samuel, 1959

„A computer program is said to learn from experience E with respect to some task T and some performance measure P , if its performance on T , as measured by P , improves with experience E .“

— Tom Mitchell, 1997

Literatur

- Géron, Aurélien (2017). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc. ISBN: 978-1-492-03264-9.
- Lynch, Shana (Mai 2017). *Andrew Ng: Why AI Is the New Electricity*. Einsichtsname: 17.05.2021. URL: <https://www.gsb.stanford.edu/insights/andrew-ng-why-ai-new-electricity>.