

Data Science und Lineare Algebra

Didaktisch-Methodische Überlegungen

Thomas Schramm

HafenCity Universität Hamburg, Geodäsie & Geoinformatik, Überseeallee 16, 20457 Hamburg, Deutschland,
thomas.schramm@hcu-hamburg.de

Abstract. Mit dem Projekt „Linear Algebra driven by Data Science“ soll eine Selbstlerneinheit in der Form von Open Educational Resources (OER) geschaffen werden, die mit dem Anwendungsfokus „maschinelles Lernen“ die Konzepte der Linearen Algebra vermittelt.

Einführung

Lineare Algebra (LA) wird im MINT-Studium im ersten Studienjahr vermittelt und stellt neben der Differentialrechnung das Fundament der Ingenieur-Mathematik dar. Hier werden mathematische Objekte wie Vektoren und Matrizen mit ihren Operationen und Ordnungsbegriffe wie Vektorraum, Dimension, Norm, Metrik mit Konstruktionen wie Basis, Linearkombination, Abbildung, Kern etc. und deren Anwendungen eingeführt, die es ermöglichen komplexe Sachverhalte zu modellieren bzw. zu simulieren. Ein in zwei oder drei Dimensionen gewonnenes geometrisches Verständnis kann in der Abstraktion auf höhere Dimensionen und andere Objekte wie Polynome übertragen werden. Die vorherrschende Anwendung ist die Lösung linearer Gleichungssysteme und die analytische Geometrie. Sie bildet die Grundlage für numerische Anwendungen, wie maschinelles Lernen, insbesondere mit künstlichen neuronalen Netzen.

Das Projekt trägt der zunehmenden Relevanz des maschinellen Lernens insbesondere der künstlichen neuronalen Netze als nicht-symbolischen Ansatz Rechnung. Es soll versucht werden, es nicht als nachgeordnete, wenn auch wichtige Applikation zu sehen, sondern als führendes und motivierendes Beispiel, das die Einführung der aufgezählten Begriffe zwingend notwendig macht. Es soll mit einem qualitativen Verständnis für neuronale Netze beginnend, die Konzepte maschinellen Lernens und die benötigten Elemente der Linearen Algebra simultan entwickeln.

Dieses Projekt wird für die Hamburg Open Online

University (HOOU) [1] entwickelt und von dieser gefördert. Die Lerneinheit wird dann über das Portal der HOOU als OER für jedermann zur Verfügung stehen und ist damit ein zweites Modul, das neben der Differentialrechnung dem Projekt oHMint (online Higher-Mathematics for MINT-Students) [2] zugeordnet und vom sog. OMB+-Konsortium [3] beraten und evaluiert wird.

Dieses vorerst einjährige Projekt ist eine Kooperation des Autors, seiner Mitarbeiter Sören Schwenker und Kay Zobel, mit Ingenuin Gasser, Alexander Lohse der Universität Hamburg und Ruedi Seiler mit Mitarbeitern der Firma Integral Learning GmbH [4] aus Berlin als Kooperationspartnern.

1 Konzept

Die Notwendigkeit, sich mit Linearer Algebra auseinanderzusetzen, wenn maschinelles Lernen professionell betrieben werden soll, wird an vielen Stellen hervorgehoben (vergl. Z.B. [5]).

Wie betont, wollen wir aber über ein technisches Verständnis hinaus, eine kritische Grundhaltung vermitteln, die es ermöglicht, die Reichweiten der neuen Technologien des maschinellen Lernens (ML) jenseits der Hochglanzbroschüren einzuschätzen und mögliche Risiken einzuordnen. Fragen, die in diesem Zusammenhang erläutert werden sollen sind z.B.

- Wer ist für Entscheidungen verantwortlich, die auf der Basis von ML getroffen werden?
- Lassen sich ML-Entscheidungen immer verstehen?
- Wo und wie ist das gelernte Wissen repräsentiert?

Hierbei müssen die verwendeten allgemeinen Begriffe wie Data Mining, künstliche Intelligenz, maschinelles Lernen, überwachtes vs. unüberwachtes Lernen, klassifizieren (Support Vector Machines) etc. erst einmal geordnet und zumindest qualitativ eingeführt werden.

Im Weiteren beschäftigen wir uns dann mit der wesentli-

chen Methode des nicht-symbolischen maschinellen Lernens durch künstliche neuronale Netze. Mit einem qualitativen, modellhaften Verständnis für die Funktionsweise neuronaler Netze beginnend, werden deren Begrifflichkeiten, deren algorithmische Umsetzung und die genutzten mathematischen Methoden parallel entwickelt.

2 Technische Implikationen

Die Selbstlerneinheit wird auf Servern Integral Learnings auf der Basis eines vorhandenen online Linear Algebra-Kurses on-the-fly auf der speziell für Mathematik entworfenen Mumie-Plattform [6] entwickelt. In diesem Kurs stehen viele der benötigten Einheiten in der Form von Vorlesungstexten, interaktiven Beispielen und Übungen bereits zur Verfügung. Später wird diese Einheit wie der Differenzialrechnungskurs über die HOOU-Plattform verlinkt.

3 Exemplarische Umsetzung

Die neuentwickelten Kurslemente wurden exemplarisch in einem Lineare-Algebra-Kurs der HafenCity Universität Hamburg im Studiengang Geodäsie und Geoinformatik im zweiten Semester eingesetzt. Es war zwar nicht geplant, aber das aktuelle nur-digitale Covid-Semester zwang alle Studierenden in das Home-Office, was erfreulicher Weise zu intensiver Bearbeitung der Studieninhalte führte. Über die allgemeine Moodle-Plattform wurden die jeweiligen zu bearbeiteten Inhalte aufgezeigt, kommentiert und verlinkt. Diese wurden wöchentlich in einer Zoom-Sitzung im *flipped classroom*-Stil besprochen und Ergänzungen eingewoben. Dazu gab es ein verpflichtendes, wöchentliches formatives eAssessment auf der Plattform Möbius [7], die ebenfalls speziell für mathematisches eAssessment zur Verfügung steht. Auf dieser Plattform wird auch die finale, summative Klausur durchgeführt.

Die Aufteilung des Stoffes erfolgte nach den klassischen Mustern eines Lehrbuchs, oder einer „normalen“ Vorlesung. Hierbei wurde darauf geachtet, die essentiellen mathematischen Inhalte beizubehalten, auch wenn diese nicht zwingend für das Verständnis der künstlichen neuronalen Netze notwendig waren. Anders herum, wurden einige Elemente der Analysis (z.B. Gradientenabstieg) kurz und plausibel eingeführt, auch wenn diese nicht zwingend in eine Einführung der Linearen Algebra gehören.

Die erste Idee, ein manuell rechenbares Beispiel heranzuziehen, wurde fallen gelassen zugunsten eines Standardbeispiels zur Erkennung handschriftlicher Ziffern. Im Wesentlichen hielten wir uns hier an die Darstellung einer Folge von zitierten YouTube-Videos von 3BLUE1BROWN [8].

Dieses Beispiel wurde mit wachsender Komplexität parallel entwickelt und die Algorithmen als Python-Code vorgestellt. In einer späteren Version, sollen die Code-Schnipsel direkt auf der Mumie-Plattform ausgeführt werden. Aktuell kann der komplette Code als Jupyter-Notebook heruntergeladen, modifiziert und ausgeführt werden. Dies erfordert aktuell eine gewisse Grundkenntnis der Programmiersprache Python. Dazu wird zu Beginn auf eine kleine Lerneinheit hingewiesen. Später sollen kleine Übungen eingeführt werden, die mathematischen Konstrukte direkt auf der Mumie-Plattform in auszuführenden Code umzusetzen. Die notwendige Funktionalität dazu ist bereits vorhanden.

Um dem Anspruch gerecht zu werden, die gesellschaftlichen kritischen Aspekte der KI-Technologien zu vermitteln, wurden diverse wissenschaftliche und populäre Artikel, sowie entsprechende Videos als Zusatzmaterial angeboten und in den *flipped sessions* kritisch besprochen. Als Beispiel seien einige Youtube-Videos von Gert Scobel [9] oder zwei Kapitel aus dem aktuellen Werk Richard David Prechts [10] genannt. Die dort gestellten und oben schon genannten *großen Fragen* stießen bei den Studierenden auf großes Interesse. Besonders genannt sei die Frage nach der Verantwortlichkeit beim autonomen Fahren, der Unmöglichkeit autonomer Killerdrohnen oder die Frage, ob wir uns zu Knechten einer KI machen könnten.

4 Fazit

Wir glauben an eine sinnvolle Verbindung von Mathematik und KI im technischen und naturwissenschaftlichen Anfängerstudium und vielleicht nicht nur dort. Neben den algorithmischen Kompetenzen, die dabei erworben werden, wird die ethische Dimension und auch die damit verbundene Verantwortung der Bildungseinrichtungen deutlich. Unser erster Versuch wurde von den meisten Studierenden freundlich evaluiert, wenn auch hier und dort eine gewisse Überforderung sichtbar wurde, sich mit den verschiedenen Ebenen des Kontextes und gleichzeitig mit den auch für die Lehrenden völlig neuen Anforderungen eines digitalen Studiums auseinanderzusetzen.

5 Referenzen

- [1] Hamburg Open Online University, <https://www.hoou.de>
- [2] oHMint, www.hoou.de/materials/ohmint
- [3] OMB+, www.ombplus.de/
- [4] Integral Learning GmbH, www.mumie.net/about-il/
- [5] Donges N., Basic Linear Algebra for Deep Learning, towards data science, 2018, towardsdatascience.com/linear-algebra-for-deep-learning-f21d7e7d7f23
- [6] Mumie-Plattform, www.mumie.net/platform/

- [7] Möbius, DigitalEd, www.digitaled.com/products/assessment
- [8] Aber was *ist* nun ein neuronales Netzwerk? | Teil 1, Deep Learning, <https://youtu.be/aircAruvnKk>
- [9] Scobel, Gert, Künstliche Intelligenz philosophisch betrachtet, <https://youtu.be/oYc4p3o6IFA>
- [10] Precht, Richard David, Künstliche Intelligenz und der Sinn des Lebens, Goldmann Verlag, München 2020