BATA, Katharina & SCHMITZ, Angela
Karlsruhe, Köln

Mathematische Grundlagen im maschinellen Lernen

Theoretischer Hintergrund und Fragestellung

Durch die zunehmende Bedeutung von maschinellem Lernen (ML), beispielsweise in Empfehlungssystemen, rücken Inhalte des ML in den Fokus von Schulen und Hochschulen (**Kandelhofer et al., 2016).** Eine besondere Zielgruppe für Bildung im Bereich ML sind MINT-Studierende, welche nicht nur alltäglich mit den Ergebnissen von ML-Verfahren arbeiten, sondern diese auch selber programmieren oder weiterentwickeln. Für die Weiterentwicklung der bereits beginnenden curricularen Einbindung in MINT-Studiengänge (Shapiro et al., 2018) spielen unter anderem Anknüpfungspunkte an bestehende Lerninhalte eine Rolle (Hußmann & Prediger, 2016).

Mathematische Methoden bilden eine wesentliche Grundlage für ML (Bishop, 2006). Aus der Perspektive des Lernens von Mathematik bietet ML daher Potential zur Motivation und als Anwendungsbeispiel. Forschungsprojekte zeigen dies insbesondere für die Schule (Biehler et al., 2024), aber auch für die Grundlagenmathematik in der Hochschule (Schramm, 2020). Umgekehrt ist der Zusammenhang zwischen ML und Mathematik auch aus der Perspektive des Lernens von ML interessant. Mit der Frage „Welche mathematischen Inhalte können den Umgang mit Verfahren des ML unterstützen?“ soll am Beispiel des überwachten ML der Zusammenhang zwischen Mathematik und ML für beide Perspektiven in den Blick genommen werden.

Überwachtes maschinelles Lernen

Ausgangspunkt des überwachten ML ist ein Datensatz aus unabhängigen Merkmalen sowie einem potentiell abhängigen Merkmal (auch „Label“). Aufbauend auf dem Datensatz soll dann eine Zuordnung entwickelt werden, welche jeder beliebigen Kombination von Ausprägungen der unabhängigen Merkmale ein Label zuordnet. Das Ziel der Anwendung überwachten MLs ist damit eine abstrahierte Repräsentation der Datenstruktur in Form eines „ML-Modells“, welches nachfolgend mathematisch beschrieben wird:

In einer Menge von Datenpaaren $D = \{(x_n,\, y_n) \mid x_n \in \mathcal{X} \subseteq \mathbb{R}^{\mathrm{m}}, y_n \in \mathcal{Y} \subseteq \mathbb{R}, \mathrm{n} = 1, \dots, \mathrm{N}\}$ *sei ein unbekannter funktionaler Zusammenhang* $f\colon \mathcal{X} \to \mathbb{R}$ *der Form gegeben, dass*

$$y_{\mathrm{n}} = f(x_{\mathrm{n}}) + \mathcal{E}_{\mathrm{n}} \quad \forall\, \mathrm{n} = 1, \dots, \mathrm{N}.$$

Die $\mathcal{E}_{\mathrm{n}}$ *seien Realisierungen von Zufallsvariablen, welche potentielle Fehler aufgrund von Messungenauigkeiten oder Rauschens darstellen.*

In: L. Schick, M. Platz & A. Lambert (Hrsg.),
Beiträge zum Mathematikunterricht 2025.
58. Jahrestagung der Gesellschaft für Didaktik der Mathematik. WTM.
https://doi.org/10.37626/GA9783959873307.0

Es soll nun, ausgehend von der Menge D, der unbekannte funktionale Zusammenhang $f: \mathbb{R}^m \rightarrow \mathbb{R}$ rekonstruiert werden. Das ML-Modell ist dabei die durch das überwachte maschinelle Lernverfahren bestimmte Funktion $\hat{f}: \mathbb{R}^m \rightarrow \mathbb{R}$, welche f bezüglich eines adäquat gewählten Gütekriteriums approximiert.

Die Bestimmung der Abbildung $\hat{f}$, *also des ML-Modells*, wird üblicherweise aus der Analyse der Daten mithilfe eines mathematischen Ansatzes geschlossen. Zwei Beispiele sind in Abbildung 1 zu finden. Die Abbildung zeigt zwei Modelle, welche Datenpunkte basierend auf zwei Merkmalen (x^1 und x^2) in zwei unterschiedliche Klassen (Klasse 1: gelb, Klasse 2: blau) einteilt. Die Klassen entsprechen den Ausprägungen der y_n, die Modelle (Entscheidungsfunktionen anhand der Ausprägungen der x^i) sind durch die eingefärbten Hintergründe dargestellt. Bei der Support-Vector-Machine (SVM) wird bezüglich einer ausgewählten Fehlerfunktion eine Trennlinie zwischen den Klassen berechnet (Bishop, 2006). Bei der k-Nächste Nachbarn-Klassifikation (kNN) werden Datenpunkte durch eine Mehrheitsentscheidung benachbarter Punkte zugeordnet (ebd.).

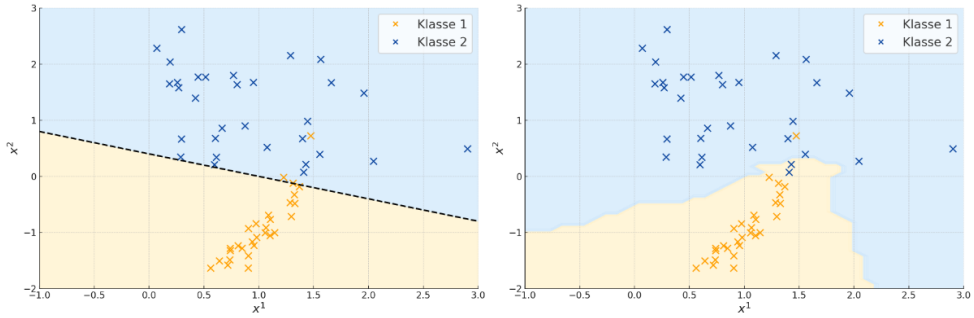

Abb. 1: Darstellung zweier mathematischer Modelle basierend auf SVM (links) bzw. kNN (rechts) (eigene Abbildung)

Methode

Zur Analyse der mathematischen Inhalte des überwachten ML wird das von Bata (im Druck) entwickelte „Modellkonzept" genutzt. Das Modellkonzept ist ein theoretisches Konstrukt, welche das konzeptuelle Wissen zu ML-Modellen darstellt und strukturiert. Im Kern des Modellkonzeptes stehen, angelehnt an die allgemeine Modelltheorie von Mahr (2008), vier „Facetten" (Bata, im Druck): *Daten, Zuordnung, Qualität* und *Nutzung*. Sie bezeichnen den konzeptuellen Fokus, mit dem das ML-Modell jeweils analysiert wird. In jeder Facette werden „Verstehenselemente" zu dem jeweiligen Fokus formuliert (ebd.). Verstehenselemente bezeichnen dabei nach Drollinger-Vetter (2011) die Teile eines Begriffs oder Konzepts, die verstanden sein müssen, um den Begriff oder das Konzept als Ganzes zu verstehen. Um die Fragestellung zu bearbeiten, werden die mathematischen Grundlagen der im

Modellkonzept formulierten Verstehenselemente spezifiziert. Der folgende Abschnitt zeigt einige Beispiele, strukturiert anhand der vier Facetten.

Ergebnisse: Mathematische Grundlagen für die Verstehenselemente

Die Datenfacette als Entsprechung zur Herstellungsperspektive bei Mahr (2008) adressiert die Datengrundlage des Modells (die Menge D). Identifizierte Verstehenselemente sind u.a. „*Ein Modell ist Daten-abhängig"* und „*Daten können Fehler und Ausreißer enthalten*". Sie zeigen auf, dass für die Auseinandersetzung mit ML *statistische Grundlagen*, wie das *Identifizieren von Ausreißern*, oder das *Argumentieren auf Grundlage von Daten* eine Rolle spielen. Insbesondere bei der Auswahl der „richtigen" Merkmale für eine Modellbildung werden Kennwerte wie *Korrelationen* oder auch Überlegungen zum *Unterschied zwischen Korrelation und Kausalität* benötigt.

Die Nutzungsfacette und die Qualitätsfacette, die die Anwendungsperspektive von Mahr (2008) abbilden, adressieren die Verwendung des Modells für eine bestimmte Fragestellung sowie die Modelleigenschaften auf technischer Ebene und im Hinblick auf die Nutzung. In der *Nutzungsfacette* werden Verstehenselemente formuliert wie beispielsweise „Das Nutzungsszenario bestimmt die Aussagekraft der ausgewählten Daten". Sie haben allerdings weniger mathematischen Charakter, sodass in dieser Facette kaum mathematische Grundlagen benötigt werden. In der *Qualitätsfacette* dagegen werden Verstehenselemente zu den mathematischen Beurteilungskriterien von Modellen formuliert, z.B. „Die Qualität eines Modells lässt sich durch die Betrachtung von Gütemaßen beurteilen". Die Gütemaße selber sind oftmals durch *Formeln* definiert, und die Auswahl des „richtigen" Gütemaßes basiert beispielsweise auf der *Unterscheidung von Fehlertypen* vergleichbar zur Unterscheidung von *α-Fehler und β-Fehler aus der Statistik*. Auch werden mit Varianz und Bias zwei Modelleigenschaften adressiert, deren mathematische Definitionen *wahrscheinlichkeitstheoretische Grundlagen* benötigen.

Die wohl reichhaltigste Facette im Hinblick auf die mathematischen Grundlagen ist die *Zuordnungsfacette*. Sie entspricht dem Modellobjekt nach Mahr (2008) und bezieht sich auf die durch das ML entstehende Zuordnung von Merkmalen zum Label. Durch Verstehenselemente wie „Ein Modell ordnet zu" oder „Die Zuordnung unterliegt einer Vorschrift" werden die Abbildung $\hat{f}$ sowie der zu ihr führende mathematische Ansatz in den Blick genommen. Für die Formulierung und Interpretation von $\hat{f}$ muss zunächst einmal der Umgang mit *reellwertigen Funktionen mehrerer Veränderlicher* möglich sein. Zusätzlich führt der Blick auf die mathematischen Ansätze wie die oben genannte SVM oder kNN zu Grundlagen wie *Optimierung*, *Vektor- und Matrizenrechnung* sowie *Abstandsberechnung*.

Diskussion

Exemplarisch wurden mathematische Grundlagen des überwachten ML dargestellt, indem sie mittels einer Analyse anhand der Verstehenselemente des Modellkonzeptes (Bata, im Druck) systematisch erfasst und strukturiert wurden. Die mathematischen Grundlagen können beispielsweise als sinnvoller Schritt hin zur Lehre von ML genutzt werden, um die notwendigen mathematischen Voraussetzungen der Studierenden zu beurteilen und Anknüpfungspunkte an das Vorwissen zu finden (Hußmann & Prediger, 2016). Auch kann man sie zur Beurteilung der Inhalte mathematischer Grundlagenmodule verwenden. Zudem kann ML für alle identifizierten mathematischen Grundlagen als Anwendungsbeispiel und Motivation dienen.

Die untersuchte Fragestellung kann mit einem vergleichbaren Ansatz auf weitere Themen des ML, wie z.B. bestärkendes Lernen, übertragen werden.

Im Vortrag wird ein konkreter ML-Modellerstellungsprozesses exemplarisch mit unterschiedlichen mathematischen Ansätzen dargestellt. Dabei werden die genannten und weitere Verstehenselemente sowie zugehörige mathematische Grundlagen exemplarisch identifiziert.

Literatur

Bata (im Druck). *Maschinelles Lernen lernen - Entwicklung und Erforschung einer Lehr-Lernumgebung in den Ingenieurswissenschaften*. Springer.

Biehler, R., Schönbrodt, S., & Frank, M. (2024). KI als Thema für den Mathematikunterricht. *mathematik lehren*, 244, 2–7. Friedrich-Verlag.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Drollinger-Vetter, B. (2011). *Verstehenselemente und strukturelle Klarheit: Fachdidaktische Qualität der Anleitung von mathematischen Verstehensprozessen im Unterricht*. WAXMANN.

Kandlhofer, M., Hirschmugl-Gaisch, S., Huber, P., & Steinbauer, G. (2016). Artificial intelligence and computer science in education: From kindergarten to university. In *Frontiers in Education Conference (FIE)*. IEEE.

Hußmann, S., & Prediger, S. (2016). Specifying and Structuring Mathematical Topics. *Journal für Mathematik-Didaktik*, 37 (S1), 33–67. https://doi.org/10.1007/s13138-016-0102-8

Mahr, B. (2008). Ein Modell des Modellseins. Ein Beitrag zur Aufklärung des Modellbegriffs. In U. Dirks & E. Knobloch (Hrsg.), *Modelle* (S. 187–218). Peter Lang Verlag.

Schramm, T. (2020). Data Science und Lineare Algebra: Didaktisch-methodische Überlegungen. *ARGESIM Report*, 59, 477–479. https://doi.org/10.11128/arep.59.a59067

Shapiro, R. B., Fiebrink, R., & Norvig, P. (2018). How Machine Larning Impacts the Undergraduate Computing Curriculum. *Communications of the ACM, 61* (11), 27–29. https://doi.org/10.1145/3277567