

Seminar “Probabilistic ML” im SoSe 25

Ludwig Bothmann

Inhaltsbeschreibung

Probabilistic Machine Learning

In vielen Anwendungen des maschinellen Lernens sind Unsicherheiten allgegenwärtig – sei es bei medizinischen Diagnosen, der Finanzprognose oder der Steuerung autonomer Systeme. Klassische ML-Modelle liefern oft nur Punktvorhersagen, doch wie sicher sind diese? Probabilistic Machine Learning (ProbML) bietet Methoden, um Unsicherheiten explizit zu modellieren und robuste Entscheidungen zu treffen.

Dieses Seminar widmet sich zentralen Konzepten und Techniken des probabilistischen maschinellen Lernens. Ein Schwerpunkt liegt auf der Unterscheidung zwischen aleatorischer und epistemischer Unsicherheit sowie deren Quantifizierung. Wir werden verschiedene Methoden zur Kalibrierung von Vorhersagen untersuchen und probabilistische Modelle für unterschiedliche Anwendungen kennenlernen.

Die Teilnehmer:innen werden sich eigenständig mit speziellen Bereichen dieser Themen befassen. Behandelte Methoden und Modelle umfassen unter anderem:

- Bayesianische Generalisierte Lineare Modelle
- Bayesianische Neuronale Netze und Full Bayes-Ansätze
- Gaußprozesse
- Baumbasierte Verfahren (Random Forests und Boosted Trees)
- Variational Inference
- Deep Ensembles
- Monte Carlo Dropout
- Laplace Approximation
- Conformal Prediction

Das Seminar setzt auf eine praxisnahe Auseinandersetzung mit diesen Methoden in R oder Python. Die Teilnehmer:innen sollen probabilistische Modelle selbst implementieren und ihre Eigenschaften anhand von realen Datensätzen untersuchen.

Als zentrale Referenzen für das Seminar dienen die beiden Bücher:

- “I”: Kevin P. Murphy, *Probabilistic Machine Learning: An Introduction*, MIT Press, 2022. ([Link](#))

- “II”: Kevin P. Murphy, *Probabilistic Machine Learning: Advanced Topics*, MIT Press, 2023. ([Link](#))
- “KH”: Andreas Krause, Jonas Hübötter, *Probabilistic Artificial Intelligence* (2025). ([Link](#))

Neben der inhaltlichen Komponente liegt ein besonderer Schwerpunkt dieses Seminars auf dem Aspekt des wissenschaftlichen Arbeitens. Dazu versetzen wir uns in die fiktive Situation, dass jede:r Teilnehmer:in ein wissenschaftliches Paper bei einer ML-Konferenz einreicht. Wir werden den Teilnehmer:innen die wichtigsten technischen Werkzeuge dafür vorstellen, deren Verwendung dann verpflichtend ist. Die Seminararbeit wird nach einer von uns zur Verfügung gestellten Vorlage im Format eines Papers verfasst und auch die Seminarpräsentation soll sich in der Form einem Konferenzvortrag annähern.

Das Seminar bietet somit nicht nur eine vertiefte Auseinandersetzung mit zentralen und aktuellen Themen in ML, sondern auch eine praxisorientierte Einführung in das wissenschaftliche Arbeiten und Präsentieren.

Rahmeninformationen

- Form:
 - Einführungsveranstaltung zur Organisation und evtl. ein weiterer, inhaltlicher Termin in der ersten Semesterhälfte
 - Seminarvorträge geblockt am Semesterende.
- Termin für die Vorbesprechung: TBD (April)
- Termine für die Vorträge: TBD (Semesterende)
- Abgabe der Seminararbeit: Zwei Wochen vor (!) den Vorträgen
- Dozent: Ludwig Bothmann
- Art: Bachelorseminar
- Sprache: Deutsch
- Ergänzende Bemerkungen: Das Seminar bereitet auf die Spezialisierung “Machine Learning” im Master “Statistics and Data Science” vor. Aus den Seminarthemen können auch Themen für die Bachelorarbeit entstehen.

Typischer Aufbau der Seminararbeit

- Einleitung, Hinführung zum Thema
- Related Work
- Methoden beschreiben und vergleichen
- Simulationsstudie
- Anwendung auf „echten“ Datensatz
- Discussion
- Conclusion & Outlook

Hinweise

- Recherchieren Sie eigenständig Paper zu Ihrem Thema und zitieren Sie diese in der Arbeit
- Halten Sie den Code auf jeden Fall in Git