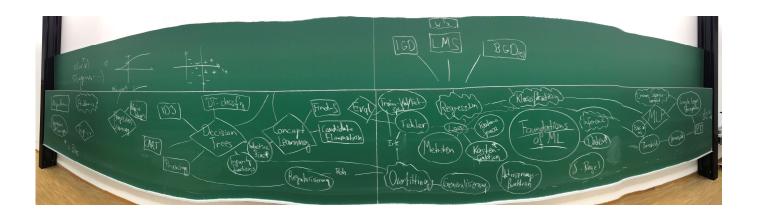
## **Foundations of Machine Learning**

Q&A

23.01.2024



## Organisatorische Fragen:

- Welche Hilfsmittel sind erlaubt? Siehe <a href="https://temir.org/teaching/machine-learning-ws23/machine-learning-ws23.html#exam">https://temir.org/teaching/machine-learning-ws23.html#exam</a>
  - Wichtig: Das einseitig eigenhändig handbeschriebene Blatt darf nicht ausgedruckt sein.
- Termin für Korrektur bereits bekannt? Nein, noch nicht klar. Studierende mit Deadline wegen Staatsexamenstermin melden sich bitte per Email bei uns. Termin für Klausureinsicht wird nach abgeschlossener Korrektur und Eintragung der Noten bekanntgegeben.
- Einschränkung klausurrelevanter Inhalte? Nein.
- Wie realistisch ist die Probeklausur? Probeklausur ähnelt in der Aufteilung der Anforderungsbereiche/Aufgabentypen und der Themenverteilung der finalen Klausur.
- Wie viele Punkte gibt es in der Klausur? Siehe Deckblatt. Grober Richtwert: 1 Punkt pro Minute, plus etwas Puffer.

## Inhaltliche Fragen:

- Regression to Classification; neuronale Netze: Wofür dient die Bias-Komponente? Verschiebung des Schwellwerts durch Änderung von  $w_0$ .
- ML:II-161: Warum ist der Cross-validation error i.d.R. kleiner als der Holdout error? –
  Dadurch, dass beim Holdout error nur ein Test set gezogen wird, kann es passieren, dass dieses nicht genau so verteilt ist wie das Training set.
- ML:VI-105: Warum ist der Hypothesenraum bei Candidate Elimination unvollständig? Dieser enthält nur Konjunktionen von Feature-Wert-Paaren, wodurch sich nicht alle erdenklichen Konzepte darstellen lassen (Beispiel: EnjoySurfing bei Sky=sunny oder cloudy, aber nicht rainy).
- Was ist Inductive Bias? Um eine Klassifikation ungesehener Examples zu ermöglichen, müssen (zusätzlich zu den Trainingsdaten) Annahmen getroffen werden. Diese bilden den Inductive Bias. Wenn man ohne Inductive Bias Aussagen über Examples treffen möchte, würde man eine Auflistung aller möglicher Examples benötigen. Wir wollen aber verallgemeinern und treffen daher Annahmen, wie eine solche Verallgemeinerung aussehen kann, ohne dass es zu Problemen kommt. Siehe ML:VI-105: Wir können komplizierte

Hypothesen zulassen, müssen dafür aber aufwendiger nach den Parametern suchen und haben ggf. Overfitting (das haben wir gesehen, als wir Punkte mit einem Polynom hohen Grades verbunden haben); oder wir nehmen einen kleineren Hypothesenraum (hier ist die Annahme konkret "das gesuchte Konzept lässt sich als Konjunktion von Feature-Wert-Paaren ausdrücken"), haben aber hier die Gefahr, dass zu sehr verallgemeinert wird, vgl. auch Folie 2-5 in <a href="https://downloads.webis.de/talks/stein\_2023b.pdf">https://downloads.webis.de/talks/stein\_2023b.pdf</a> .

 Wie ist Entropie definiert und motiviert? – Zur Definition und Motivation als Informationsgehalt siehe ML:VI-63. Für einen etwas anderen Blickwinkel auf das Thema Entropie empfehle ich dieses Video, in dem Entropie im Kontext von Wordle-Rätseln erklärt wird: <a href="https://www.youtube.com/watch?v=v68zYyaEmEA&ab\_channel=3Blue1Brown">https://www.youtube.com/watch?v=v68zYyaEmEA&ab\_channel=3Blue1Brown</a>