

Klausur zur Lehrveranstaltung
„Maschinelles Lernen 1 – Grundverfahren“
(60 Minuten)

Nachname:	Vorname:
Matrikelnummer:	Studiengang & Semester:

Anmerkungen

- Legen Sie Ihren Studierendenausweis und ein gültiges Ausweisdokument gut sichtbar bereit.
- Tragen Sie Nachname, Vorname, Matrikelnummer, Studiengang & Semester deutlich lesbar ein und unterschreiben Sie das Klausurexemplar unten.
- Die folgenden 6 Aufgaben sind vollständig zu bearbeiten. Jede Antwort muss entweder in deutscher oder englischer Sprache formuliert sein.
- Als Hilfsmittel sind ausschließlich folgende zugelassen:
 - ein nicht programmierbarer Taschenrechner und
 - ein nicht beschriftetes Wörterbuch.
- Täuschungsversuche führen zum Ausschluss von der Klausur.
- Unleserliche oder mit Bleistift geschriebene Lösungen können von der Korrektur bzw. der Wertung ausgeschlossen werden.
- Beim Ausfüllen von Lücken gibt die Größe der Kästen keinen Aufschluss über die Länge des einzufügenden Inhaltes.
- Die Bearbeitungszeit beträgt 60 Minuten.

Ich bestätige, dass ich die Anmerkungen gelesen und mich von der Vollständigkeit dieses Klausurexemplars (Seite 1 - 14) überzeugt habe.

Unterschrift

Nur für den Prüfer

Aufgabe	1	2	3	4	5	6	Gesamt	Note
Punkte	11	7	9	9	12	12	60	
Erreicht								

**Aufgabe 1 – Induktives Lernen, Lerntheorie, Entscheidungsbäume
und Unüberwachtes Lernen****___ / 11P**

a) Definieren Sie Konsistenz und Vollständigkeit im Hypothesenraum.

(___/2P)

b) Was ist das Optimierungsziel beim Lernen von Entscheidungsbäumen bzgl. der Klassenverteilung? Welches Maß wird hierfür herangezogen? Wie wirkt sich dies auf die tiefe des gelernten Baums aus?

(___/1,5P)

c) Nennen Sie zwei Maßnahmen gegen Overfitting von Entscheidungsbäumen?

(___/1P)

d) Geben Sie die Formeln für Genauigkeit (Precision) und True-Positive-Rate (Recall) an.

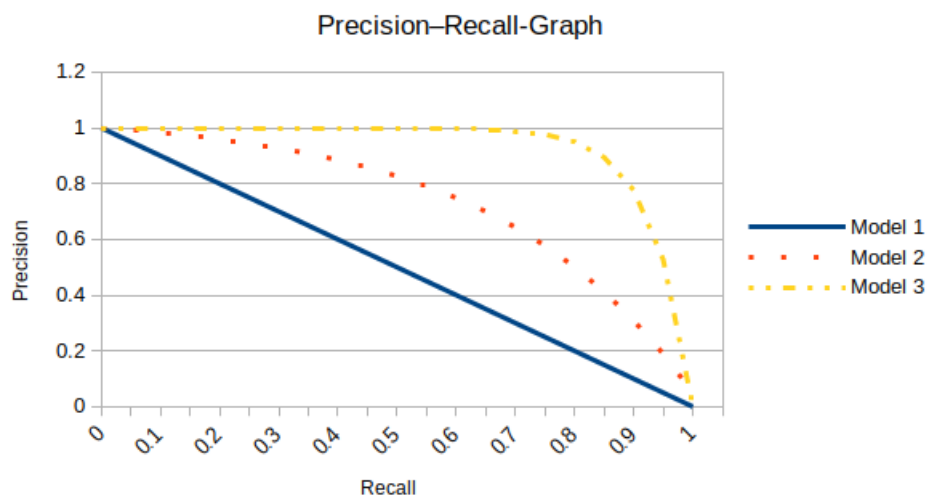
(___/2P)

- e) Nennen Sie vier positive Eigenschaften eines Random Forests (verglichen mit einem einfachen Entscheidungsbaum).

(___/2P)

- f) Welches Modell ist das Beste? Geben sie eine kurze Begründung an?

(___/1,5P)



- g) Was ist für den k-Means Algorithmus mindestens festzulegen? (Hinweis: Parameter, Metriken)

(___/1P)

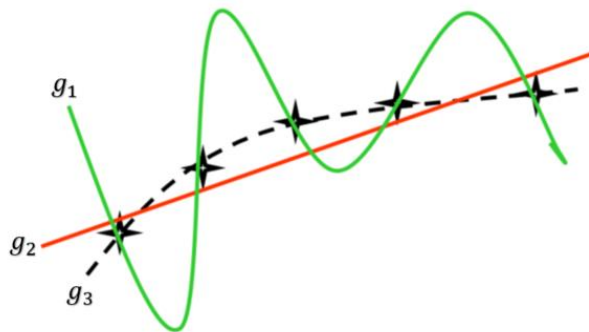
Aufgabe 2 – Neuronale Netze**___ / 7P**

- a) Geben Sie die Formel und Ableitung für die Aktivierungsfunktion ReLU an. (___/2P)
Welchen entscheidenden Nachteil hat ReLU? Welche alternative Aktivierungsfunktion versucht dies zu beheben?

- b) Was muss bei der Initialisierung der Gewichte eines Neuronalen Netzes beachtet werden? (___/2P)
Nennen und erläutern Sie kurz eine gängige Initialisierungsmethode.

- c) Die folgende Abbildung zeigt die Kurven g_1 und g_2 , die von zwei verschiedenen neuronalen Netzen an die gleichen Trainingsdaten (Sterne) angepasst wurden. Die Kurve g_3 entspricht der zu erlernenden Funktion (Grundwahrheit). (___/3P)

1. Nennen Sie die Fachbegriffe für das Fehlverhalten (zugrundeliegende Verhalten) von g_1 und g_2 und beschreiben Sie den Fehler in der Topologie des neuronalen Netzes, der zu diesem Verhalten führen kann.
2. Skizzieren Sie den typischen Verlauf des Trainings- und Testfehlers in Abhängigkeit der Trainingszyklen für g_1 und g_2 .



Aufgabe 3 – CNN**___ / 9P**

- a) Erklären Sie “Weight Sharing”. Welche Vorteile ergeben sich hierdurch gegenüber klassischen Neuronalen Netzen?

(___/1,5P)

- b) Nennen Sie zwei verschiedene Methoden zur Initialisierung von Gewichten.

(___/1P)

- c) Wie funktioniert Dropout, welche Vorteile bringt es?

(___/2P)

- d) Wie wird das Vanishing-Gradient Problem bei ResNet gelöst? Erläutern Sie dies kurz.

(___/1P)

- e) Welche Nachteile hat die Verwendung der Sigmoid Aktivierungsfunktion für tiefe neuronale Netze wie Resnet ohne „skip connections“? (___/1P)

- f) Berechnen Sie die Ausgaben der Faltungs-, sowie der Poolingsoperation, gegeben die nachfolgenden Eingangsdaten und die folgenden Operationen in sequentieller Reihenfolgen. (___/2,5P)
- Operation 1: Faltung mit 3x3, Stride 1, Padding 0
- Operation 2: Max Pooling mit 3x3, Stride 1, Padding 0

Eingangsdaten:

8	0	1	1	8
8	0	1	1	8
8	0	1	1	8
8	0	1	1	8
8	0	1	1	8

Faltungsfilter:

1	0	0
0	1	0
0	0	1

Aufgabe 4 – SVM**___ / 9P**

- a) Bezogen auf einen Support Vektor Klassifikator, beschreiben Sie das Problem welches gelöst wird, die Lösung die gefunden wird und Intuition für die Lösung.

(___/1,5P)

- b) Formulieren Sie das grundlegende Optimierungsproblem für einen Support Vector Klassifikator formal. Geben Sie außerdem die Nebenbedingungen für die Optimierung an.

(___/1,5P)

- c) Nennen Sie zwei Methoden, mit denen Support Vector Klassifikatoren auch nicht linear separierbare Daten trennen können.

(___/1P)

- d) Wie viele Stützvektoren werden mindestens benötigt, wenn der Merkmalsraum n Dimensionen hat?

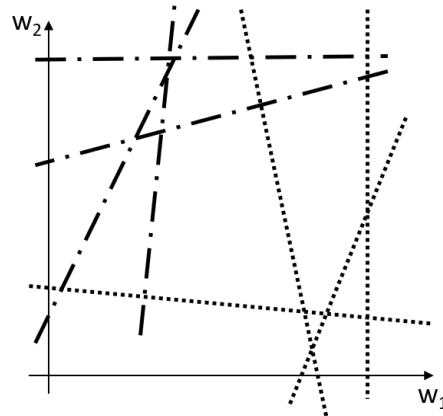
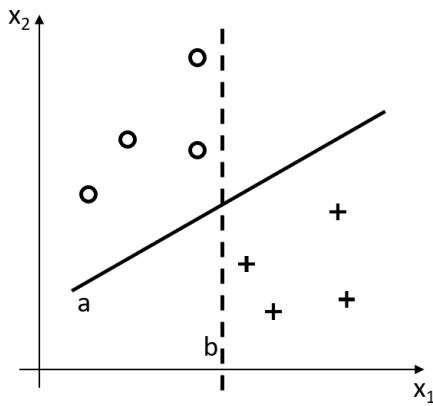
(___/1P)

- e) Gegeben sind Daten mit fünf unterschiedlichen Klassenzugehörigkeiten. Wie können diese mithilfe von SVMs klassifiziert werden? Wie viele SVMs werden dafür mindestens benötigt?

(___/1P)

- f) Welche der beiden Hypothesen im linken Graph ist die optimale Hypothese? Markieren Sie die entsprechenden Stützvektoren (Support Vectors) im linken Graphen (Merkmalsraum) und zeichnen Sie die Hypothese in den rechten Graphen (Hypothesenraum) ein.

(___/3P)



Aufgabe 5 – Reinforcement Learning**___ / 12P**

a) Welches Ziel verfolgt Reinforcement Learning?

(___/1P)

b) Durch welches Entscheidungsmodell lässt sich die Problemstellung beim Reinforcement Learning formal darstellen?
Welche vier Bestandteile werden für die Modellierung benötigt?

(___/3P)

c) Was besagt die Markov-Eigenschaft?

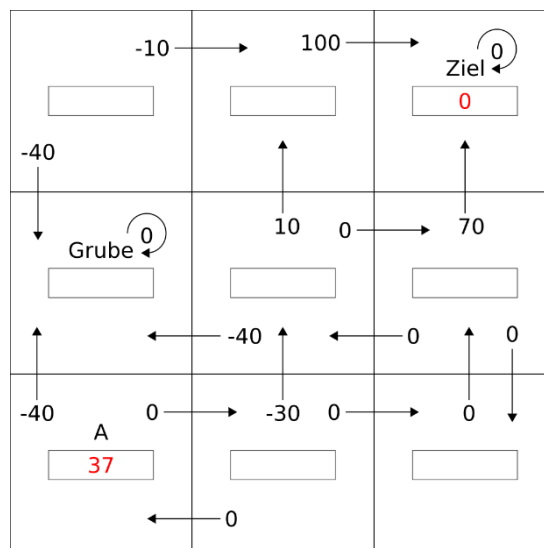
(___/1P)

d) Was ist ein wesentlicher Unterschied zwischen den Spielen „Schach“ und „Poker“ mit Bezug auf die Beschreibung der Zustände? Wodurch entsteht dieser Unterschied?

(___/1P)

- e) Gegeben ist $V^\pi(s)$: Bestimmen Sie die Optimale Strategie formal. Außerdem, (/2P)
definieren Sie die rekursive Form der Bellmann Optimalitätsgleichung in
Abhängigkeit von der optimalen Wertfunktion.

- f) Betrachten Sie die untenstehende Welt. Ein Agent kann sich mit den
angezeigten Aktionen (Pfeile) von Zelle zu Zelle bewegen. Die Belohnung für
eine Aktion entspricht der Zahl an den Pfeilen. Nehmen Sie an, dass die
optimale Strategie gelernt wurde. Tragen Sie die Zustandswerte $V^*(s)$ dieser
Strategie in die entsprechenden Kästen ein (Diskontierungsfaktor $\gamma = 0,8$).
Runden Sie ihre Ergebnisse auf ganze Zahlen. Zeichnen Sie den Pfad der
optimalen Strategie von Zelle A zum Ziel ein. (/4P)



Aufgabe 6 – Bayes, HMM und SPN**___ / 12P**

Gegeben sei das folgende Hidden Markov Model (HMM) (grafische Darstellung siehe unten) mit Zustandsraum $S = \{S_1, S_2, S_3\}$, Beobachtungsraum $V = \{V_1, V_2, V_3\}$. Die Übergangswahrscheinlichkeitsmatrix A ist gegeben durch:

(___/2P)

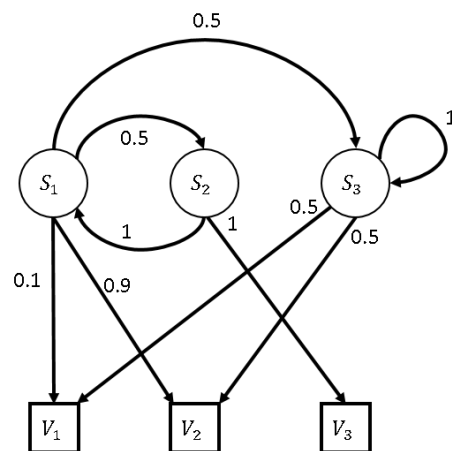
$$A = a_{(ij)} = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Die Emissions-/Beobachtungs-Wahrscheinlichkeiten $B = (b_{ik})$ sind wie folgt definiert:

$$B = b_{(ik)} = \begin{bmatrix} 0.1 & 0.9 & 0 \\ 0 & 0 & 1 \\ 0.5 & 0.5 & 0 \end{bmatrix}$$

Hinweis: q_t beschreibt den Zustand zum Zeitpunkt t , o_t die Beobachtung im Zeitpunkt t .

- a) Zu Beginn befindet sich das System in Startzustand $q_1 = S_1$.
 Sie erhalten folgende Sequenz von Messungen: $o_2 = V_3, o_3 = V_1, o_4 = V_3$
 Geben Sie die a posteriori Verteilung $P(q_4 | o_2 = V_3, o_3 = V_1, o_4 = V_3, q_1 = S_1)$ an.
Hinweis: Decodierungsproblem



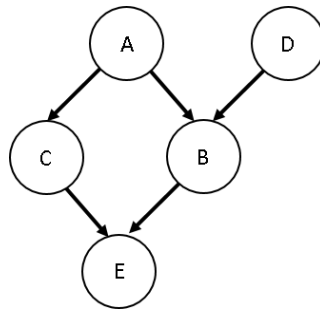
$$\begin{aligned} P(q_4 = S_1 | o_2 = V_3, o_3 = V_1, o_4 = V_3, q_1 = S_1) &= \\ P(q_4 = S_2 | o_2 = V_3, o_3 = V_1, o_4 = V_3, q_1 = S_1) &= \\ P(q_4 = S_3 | o_2 = V_3, o_3 = V_1, o_4 = V_3, q_1 = S_1) &= \end{aligned}$$

- b) Das in Aufgabe a) gegebene HMM befindet sich im Zeitpunkt t in Zustand $q_t = S_2$. Geben Sie folgende Wahrscheinlichkeiten für zukünftig mögliche Beobachtungen zum Zeitpunkt $t + 2$ bzw. $t + 3$ an: (___/3P)

$$\begin{aligned}P(o_{t+2} = V_3 | q_t = S_2) &= \\P(o_{t+2} = V_1 | q_t = S_2) &= \\P(o_{t+3} = V_3 | q_t = S_2) &= \end{aligned}$$

- c) Gegeben sei folgendes Bayes'sche Netz. Geben Sie die zugehörige Faktorisierung der Verbundwahrscheinlichkeitsdichte an: (___/2P)

$$P(A, B, C, D, E)$$

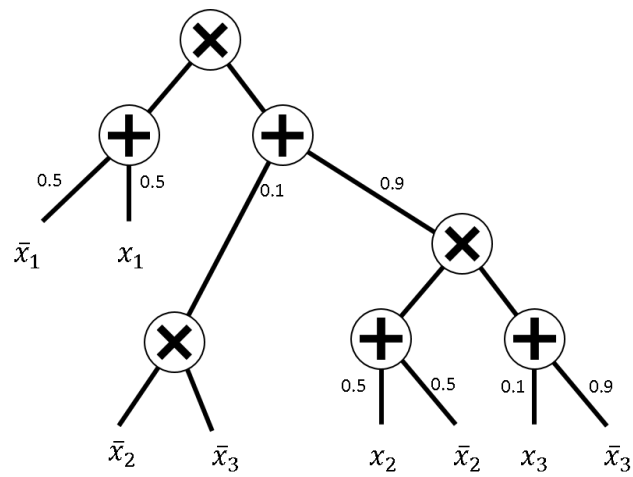


- d) Nennen Sie zwei Arten/Typen der Inferenz, die mit SPN ausgeführt werden können. (___/2P)

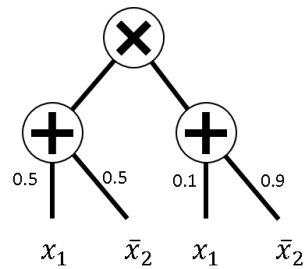
e) Geben Sie für die folgenden Sum Product Netze (SPN) jeweils an, ob diese valide sind oder nicht.

(___/3P)

A



B



C

