Technische Universität Berlin
Fakultät IV – Elektrotechnik und Informatik

Künstliche Intelligenz: Grundlagen und Anwendungen
Wintersemester 2014 / 2015

Albayrak, Fricke (AOT) – Opper, Ruttor (KI)

Schriftlicher Test – Teilklausur 2

12.02.2015

Name, Vorname: ____________________

Matrikelnummer: ____________________

Studiengang: ____________________

Hinweise:

- Überprüfen Sie bitte, ob Sie alle **10** Seiten der Klausur erhalten haben.
- Bitte versehen Sie vor Bearbeitung der Klausur alle **10** Seiten mit Ihrer Matrikelnummer.
- Bitte nicht mit einem roten oder grünen Stift schreiben.
- Bitte keinen Bleistift, keinen Tintenkiller und kein Tipp-Ex benutzen.
- Die Vorder- und Rückseiten der Klausur dürfen verwendet werden. Den Anhang (Seite 10) dürfen Sie abtrennen. Sie müssen ihn nicht abgeben.

Dieser Teil ist zur Auswertung bestimmt und soll von den Teilnehmerinnen und Teilnehmern der Klausur nicht ausgefüllt werden.

Aufgabe 1 28 Punkte	**Aufgabe 2** 28 Punkte	**Aufgabe 3** 24 Punkte	**Aufgabe 4** 20 Punkte	**Summe** 100 Punkte

Aufgabe 1 – Probabilistische Inferenz (28 Punkte)

Es existieren zwei Krankheiten, die das gleiche Symptom hervorrufen. Folgende Erkenntnisse konnten in wissenschaftlichen Studien festgestellt werden:

- Es wurde ein Gendefekt entdeckt, der bei einem Fünftel der Bevölkerung auftritt ($G = w$). Bei 10% aller Betroffenen führt er zum Ausbruch von Krankheit A ($A = w$). Ohne diesen Defekt ($G = f$) tritt Krankheit A nur in einem von 1000 Fällen auf.
- Unabhängig von den Genen tritt Krankheit B ($B = w$) bei 2% der Bevölkerung auf.
- Das Symptom tritt bei jedem 20. Menschen, der keine der beiden Krankheiten hat, auf ($S = w$). Während 50% der Patienten mit (ausschließlich) Krankheit A das Symptom haben, sind es bei Patienten mit (ausschließlich) Krankheit B nur 30%. Wenn beide Krankheiten gleichzeitig auftreten, wird das Symptom zu 90% festgestellt.
- Für Krankheit A ist ein Test vorhanden. Bei einem erkrankten Menschen fällt er zu 99% positiv aus ($T = w$), bei einem gesunden nur in 0.3% der Fälle.

(a) Zeichnen Sie ein Bayes-Netz, das zu diesem Modell passt! Die Wahrscheinlichkeitstabellen brauchen Sie hierfür nicht anzugeben. (6 Punkte)

(b) Wie groß ist die Wahrscheinlichkeit, dass der Test positiv ausfällt, wenn der Gendefekt vorliegt? (6 Punkte)

(c) Wie wahrscheinlich ist es, an Krankheit A erkrankt zu sein, wenn der Test negativ ausfällt und der Gendefekt vorliegt? (8 Punkte)

(d) Wie wahrscheinlich tritt das Symptom auf, wenn man nicht unter dem Gendefekt leidet? (8 Punkte)

Aufgabe 2 – Hidden-Markov-Modell (28 Punkte)

Wir benutzen ein Modell zur Erkennung von mathematischen Ausdrücken. Wir unterscheiden zwischen einer Folge von Zeichentypen $x_1, x_2, x_3, \ldots$ und der tatsächlich beobachteten Folge von Zeichen $y_1, y_2, y_3, \ldots$.

x_i	x_{i+1}	$P(x_{i+1}\|x_i)$
Start	Klammer(1)	1.0
Klammer(1)	Variable	0.7
Klammer(1)	Ziffer(1)	0.3
Variable	Rechensymbol	0.6
Variable	Klammer(2)	0.4
Ziffer(1)	Ziffer(2)	0.6
Ziffer(1)	Rechensymbol	0.3
Ziffer(1)	Klammer(2)	0.1
Ziffer(2)	Ziffer(2)	0.2
Ziffer(2)	Rechensymbol	0.5
Ziffer(2)	Klammer(2)	0.3
Rechensymbol	Ziffer(1)	0.8
Rechensymbol	Variable	0.2
Klammer(2)	Stop	1.0

x_i	y_i	$P(y_i\|x_i)$
Klammer(1)	(	1.0
Variable	x	0.5
Variable	y	0.3
Variable	z	0.2
Ziffer(1)	1	0.6
Ziffer(1)	2	0.3
Ziffer(1)	3	0.1
Ziffer(2)	0	0.7
Ziffer(2)	1	0.1
Ziffer(2)	2	0.1
Ziffer(2)	3	0.1
Rechensymbol	+	0.5
Rechensymbol	-	0.5
Klammer(2)	)	1.0

Die Markovkette beginnt immer mit $x_0 = \textit{Start}$ und endet mit $x_{k+1} = \textit{Stop}$. Diesen Zuständen ist keine Ausgabe zugeordnet, da sie Verbindungen zu anderen Teilen des Modells darstellen. Alle nicht angegebenen Wahrscheinlichkeiten $P(x_{i+1}|x_i)$ und $P(y_i|x_i)$ sind Null. Sie können die Zeichentypen abkürzen, um Platz zu sparen.

(a) Stellen Sie das Modell für die Ausdrücke in einem Übergangsdiagramm graphisch dar! Sie brauchen keine Wahrscheinlichkeiten einzutragen. (8 Punkte)

(b) Mit welcher Wahrscheinlichkeit tritt der Ausdruck „(12-y)“ in diesem Modell auf? (6 Punkte)

(c) Durch einen Übertragungsfehler kommt eine Zeichenfolge unvollständig an. Es wird „(13+x-1**E**2-y+3)“ angezeigt. **E** steht für ein fehlerhaft übertragenes Zeichen. Welches Zeichen stand dort am wahrscheinlichsten und wie sicher ist diese Vorhersage? (14 Punkte)

Aufgabe 3 – Generatives Modell (24 Punkte)

Bei einem Würfelspiel muss man eine Runde aussetzen, wenn man eine 6 würfelt. Sie haben zuerst eine 1, dann eine 3 und dann 5 mal hintereinander eine 6 gewürfelt. Sie glauben nicht an einen Zufall, sondern vermuten, dass der Würfel gezinkt ist und die Wahrscheinlichkeit für eine 6 größer als 1/6 ist.

(a) Welche Wahrscheinlichkeitsverteilung hat das Auftreten von 2 Erfolgen (1-5 gewürfelt) gefolgt von k Fehlschlägen (6 gewürfelt), wenn die Wahrscheinlichkeit eines Erfolges jedes Mal p beträgt? (4 Punkte)

(b) In einem Kasino wird ein Glücksspiel angeboten, bei dem man pro Runde 4 Euro bezahlt und die gewürfelte Augenzahl in Euro gewinnt. Was hoch ist der erwartete Gewinn pro Runde, wenn sie einen Würfel benutzen, bei dem die Wahrscheinlichkeit eine 6 zu würfeln 3/8 beträgt, während alle anderen Augenzahlen gleich wahrscheinlich sind? (4 Punkte)

(c) Wenn Sie a-priori annehmen, dass ein Würfel fair ist, würden Sie erwarten, dass im Mittel fünf von sechs Würfen ein Erfolg sind. Sie wollen diese Information als Vorwissen in Form einer Beta-Verteilung $\mathrm{Beta}(p; \alpha, \beta) = B(\alpha, \beta)\, p^{\alpha-1}(1-p)^{\beta-1}$ in ihrer Schätzung von p nutzen. Wie sollten Sie die Hyperparameter α und β wählen? Begründen Sie! (4 Punkte)

(d) Zeigen Sie, dass die Maximum-a-posteriori Hypothese für eine Folge von 2 Erfolgen gefolgt von k Fehlschlägen durch $p = (\alpha + 1)/(k + \alpha + \beta)$ gegeben ist! (8 Punkte)

(e) Welchen Wert p hat die Maximum-a-posteriori-Hypothese für $k = 5$, wenn Sie die Hyperparameter auf $\alpha = 6$ und $\beta = 2$ setzen? (4 Punkte)

Aufgabe 4 – Neuronales Netz (20 Punkte)

Betrachten Sie ein Perzeptron mit zwei reellwertigen Eingabeneuronen e_1 und e_2, zwei Gewichten w_1 und w_2 sowie einem Bias w_0. Für das lokale Feld h gilt

$$h = w_1 e_1 + w_2 e_2 - w_0$$

und zur Berechnung der Ausgabe a wird die Aktivierungsfunktion

$$a = \operatorname{sgn}(h) = \begin{cases} +1 & \text{für } h > 0 \\ -1 & \text{für } h \leq 0 \end{cases}$$

verwendet. Alle Gewichte und der Bias werden mit 1 initialisiert. Das Perzeptron soll nun die folgenden Beispiele (Soll-Ausgabe y) lernen:

e_1	$+1$	-2	$+3$	-1
e_2	$+1$	$+2$	$+2$	-3
y	$+1$	-1	-1	-1

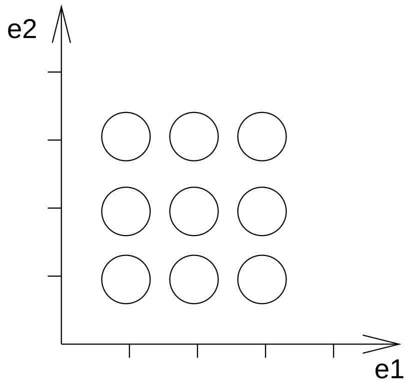
(a) Welche der vier Beispiele werden vom Perzeptron ohne eine Anpassung der Gewichte falsch klassifiziert? (4 Punkte)

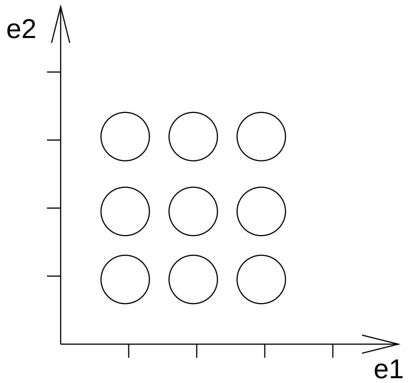
(b) Wie ändern sich die Gewichte, wenn das neuronale Netz gemäß der Perzeptron-Lernregel mit jedem Beispiel einmal trainiert wird? Verwenden Sie $\lambda = 0.4$ als Lernrate und passen Sie auch den Bias w_0 an. (4 Punkte)

(c) Stellen Sie das Klassifikationsproblem in der e_1-e_2-Ebene graphisch dar und lösen Sie es durch möglichst wenige lineare Entscheidungsgrenzen! (6 Punkte)

(d) Kann ein Perzeptron die Beispiele bei genügend langem Training exakt lernen? Begründen Sie Ihre Antwort! (2 Punkte)

(e) Was ist die kleinste Anzahl linearer Entscheidungsgrenzen, die ausreicht, um jede beliebige Zuordnung der unten abgebildeten Beispiele in die Klassen +1 und −1 korrekt zu klassifizieren? Zeichnen sie diese Entscheidungsgrenzen ein und geben Sie eine Zuordnung an, die mit weniger Grenzen nicht klassifizierbar ist. (4 Punkte)





Anhang – darf abgetrennt werden

Aufgabe 1: Annahmen des Modells

- Es wurde ein Gendefekt entdeckt, der bei einem Fünftel der Bevölkerung auftritt ($G = w$). Bei 10% aller Betroffenen führt er zum Ausbruch von Krankheit A ($A = w$). Ohne diesen Defekt ($G = f$) tritt Krankheit A nur in einem von 1000 Fällen auf.
- Unabhängig von den Genen tritt Krankheit B ($B = w$) bei 2% der Bevölkerung auf.
- Das Symptom tritt bei jedem 20. Menschen, der keine der beiden Krankheiten hat, auf ($S = w$). Während 50% der Patienten mit (ausschließlich) Krankheit A das Symptom haben, sind es bei Patienten mit (ausschließlich) Krankheit B nur 30%. Wenn beide Krankheiten gleichzeitig auftreten, wird das Symptom zu 90% festgestellt.
- Für Krankheit A ist ein Test vorhanden. Bei einem erkrankten Menschen fällt er zu 99% positiv aus ($T = w$), bei einem gesunden nur in 0.3% der Fälle.

Aufgabe 2: Nachrichtenmodell

x_i	x_{i+1}	$P(x_{i+1}\|x_i)$
Start	Klammer(1)	1.0
Klammer(1)	Variable	0.7
Klammer(1)	Ziffer(1)	0.3
Variable	Rechensymbol	0.6
Variable	Klammer(2)	0.4
Ziffer(1)	Ziffer(2)	0.6
Ziffer(1)	Rechensymbol	0.3
Ziffer(1)	Klammer(2)	0.1
Ziffer(2)	Ziffer(2)	0.2
Ziffer(2)	Rechensymbol	0.5
Ziffer(2)	Klammer(2)	0.3
Rechensymbol	Ziffer(1)	0.8
Rechensymbol	Variable	0.2
Klammer(2)	Stop	1.0

x_i	y_i	$P(y_i\|x_i)$
Klammer(1)	(	1.0
Variable	x	0.5
Variable	y	0.3
Variable	z	0.2
Ziffer(1)	1	0.6
Ziffer(1)	2	0.3
Ziffer(1)	3	0.1
Ziffer(2)	0	0.7
Ziffer(2)	1	0.1
Ziffer(2)	2	0.1
Ziffer(2)	3	0.1
Rechensymbol	+	0.5
Rechensymbol	-	0.5
Klammer(2)	)	1.0

Aufgabe 3: Priorverteilung

$$\mathrm{Beta}(p;\alpha,\beta) = B(\alpha,\beta)\, p^{\alpha-1}(1-p)^{\beta-1}$$

Aufgabe 4: Beispiele

e_1	$+1$	-2	$+3$	-1
e_2	$+1$	$+2$	$+2$	-3
y	$+1$	-1	-1	-1