30. Juni 2006	**Prüfung aus Einführung in die Mustererkennung — SS 2006**	9.30 – 11.00
Matr.Nummer:	Nachname:	
Kennzahl:	Vorname:	

Bei der vorliegenden Prüfung können Sie eine maximale Anzahl von 30 Punkten erreichen. Bitte verwenden Sie den für die Beantwortung der Frage vorgesehenen Platz und beantworten Sie die folgenden Fragen kurz aber aussagekräftig. Sie können die Fragen auf Englisch oder Deutsch beantworten. **Keine Unterlagen sind erlaubt.**

1 Kurze Fragen

1. Ist der k-NN Klassifikator ein Überwachtes oder unüberwachtes Lernverfahren? Die nicht-passende Antwort durchstreichen. (1 Punkt)

   Überwachtes Verfahren / Unüberwachtes Verfahren

   Begründung:

2. Ist der k-NN Klassifikator ein parametrisches oder nicht-parametrisches Verfahren? Die nicht-passende Antwort durchstreichen. (1 Punkt)

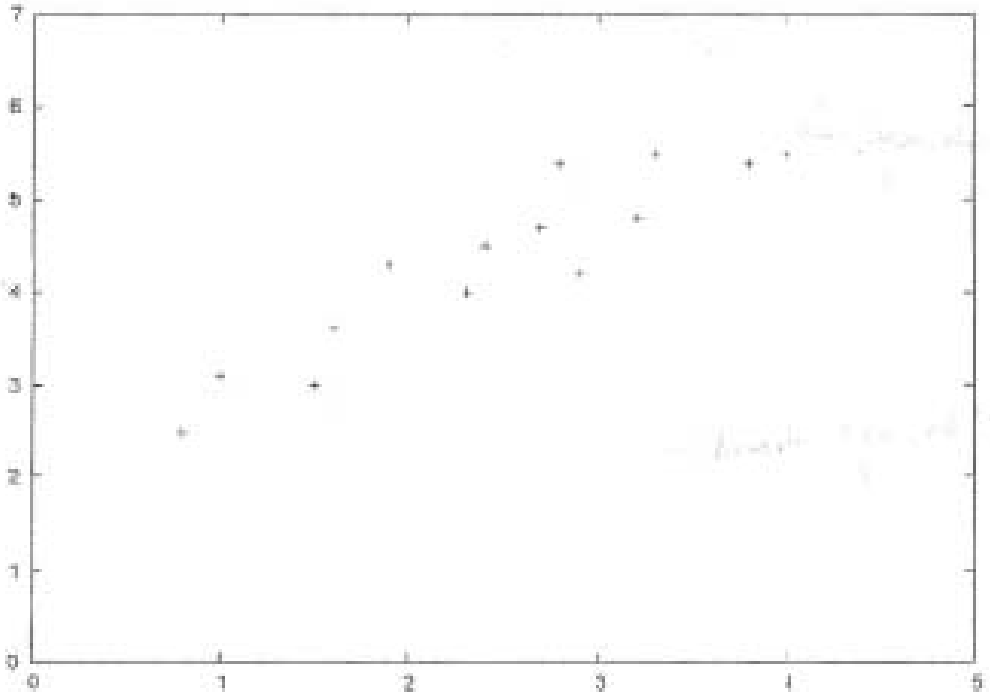
   Parametrisches Verfahren / Nicht-parametrisches Verfahren

   Begründung:

3. Unter welchen Bedinungen konvergiert der *online Perceptron Training Algorithmus* (mit fixer Lernrate)?

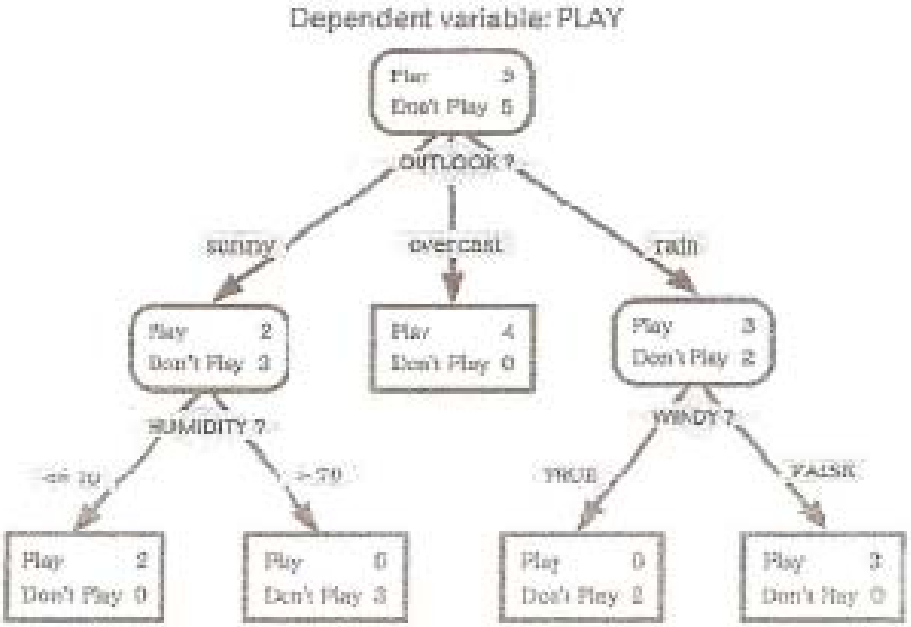
   Was macht der Algorithmus, wenn diese Bedinungen nicht erfüllt sind? (1 Punkt)

4. Zeichnen Sie die Orientierungen des Haupteigenvektors und Zweiten Eigenvektors von einer Hauptkomponentenanalyse (*Principal Component Analysis, PCA*) für den folgenden Datensatz. Die eingezeichneten Richtungen sollen klar beschriftet sein. (1 Punkt)



5. Dieser Entscheidungsbaum entscheidet, ob man Tennis spielen gehen soll oder nicht. Die 4 Merkmale sind (*Outlook, Temperature, Humidity, Windy*). Welche Entscheidung wird für diese Merkmalsausprägungen getroffen: (sunny, 25, 60, FALSE)? Die Werte in den Knoten geben die Anzahl der Trainingsbeispiele an, die in den jeweiligen Knoten ankommen. Die nicht-passende Antwort durchstreichen. (1 Punkt)

Play / Don't Play



6. Die Aktivierungsfunktion in einem Backpropagation neuronalen Netz muss

differenzierbar / eine Signum Funktion

sein. Die nicht-passende Antwort durchstreichen. (1 Punkt)

7. Die zwei Zufallsvariablen A und B sind statistisch *unabhängig* voneinander. Welche von den folgenden Ausdrücken ist **FALSCH**? Schreiben Sie die richtige Antwort ins untere Rechteck. (1 Punkt)

   (a) $P(A|B) = P(A)$

   (b) $P(A|B) = P(B)$

   (c) $P(A, B) = P(A|B)P(B)$

   (d) $P(A, B) = P(B|A)P(A)$

   Dieser Ausdruck ist FALSCH: [ ]

2 Rechenbeispiele

1. Ein Perceptron (mit einer Signum Funktion) wird durch den folgenden Gewichtsvektor beschrieben:

$$\mathbf{w} = \begin{pmatrix} 0.3 \\ 0.8 \\ 1.7 \\ -0.8 \end{pmatrix} \tag{1}$$

Der folgende Merkmalsvektor ist mit diesem Perceptron zu klassifizieren:

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0.3 \\ -3.8 \\ 0.2 \end{pmatrix} \tag{2}$$

(a) Werden hier homogene or nicht-homogene Koordinaten verwendet? Die nicht-passende Antwort durchstreichen. (1 Punkt)

Homogene Koordinaten / Nicht-homogene Koordinaten

Begründung:

(b) Was ist der Ausgabewert von diesem Perceptron mit dem gegebenen Merkmalsvektor? Schreiben Sie diesen Wert ins Rechteck unten. (1 Punkt)

Ausgabewert vom Perceptron:

2. Sie haben zwei Klassen und wollen die Entscheidungsgrenze zwischen diesen zwei Klassen berechnen. Das Trainingsset für jede Klasse besteht aus 4 zwei-dimensionalen Vektoren. Die Trainingsvektoren für Klasse 1 sind:

$$\mathbf{x}_{11} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad \mathbf{x}_{12} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}, \quad \mathbf{x}_{13} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \mathbf{x}_{14} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \tag{3}$$

und die Trainingsvektoren für Klasse 2 sind:

$$\mathbf{x}_{21} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}, \quad \mathbf{x}_{22} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}, \quad \mathbf{x}_{23} = \begin{pmatrix} 6 \\ 2 \end{pmatrix}, \quad \mathbf{x}_{24} = \begin{pmatrix} 6 \\ 4 \end{pmatrix} \tag{4}$$

(a) Berechnen Sie die geschätzten Mittelwertvektoren $\hat{\mu}_j$ für diese zwei Klassen. Schreiben Sie diese Vektoren in die unteren Rechtecke. (1 Punkt)

Mittelwertvektor $\hat{\mu}_1 =$

Mittelwertvektor $\hat{\mu}_2 =$

(b) Berechnen Sie die geschätzte Kovarianzmatrix $\hat{\Sigma}_1$ für Klasse 1. Schreiben Sie diese Matrix ins Rechteck unten. (1 Punkt)

Kovarianzmatrix $\hat{\Sigma}_1 =$

(c) Sie wollen Entscheidungsfunktionen, die auf dem Bayes Theorem basieren, verwenden. Wie berechnen sich gemäß Bayes-Theorem die *a posteriori* Wahrscheinlichkeiten $p(\omega_j|\mathbf{x})$ (1 Punkt)

(d) Sie modellieren die Klassen durch zwei-dimensionale Gauß'sche Verteilungen:

$$p(\mathbf{x}|\omega_j) = \frac{1}{(2\pi)^{\frac{a}{2}} \left|\hat{\Sigma}_j\right|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x}-\hat{\mu}_j)^T \hat{\Sigma}_j^{-1}(\mathbf{x}-\hat{\mu}_j)\right] \tag{5}$$

wobei $a = 2$. Schreiben Sie die Diskriminanten-Funktion $g_j(\mathbf{x}) = \ln p(\omega_j|\mathbf{x})$ hier aus (die Gauß'sche Verteilung für $p(\mathbf{x}|\omega_j)$ soll auch in diesen Ausdruck hineingenommen werden). N.B. Der Ausdruck soll eine Funktion von $\hat{\mu}_j$, $\hat{\Sigma}_j$, ω_j, $P(\omega_j)$, $p(\mathbf{x})$ und $\mathbf{x}$ sein. Klassen-spezifischen Werten sollen nicht verwendet werden. (1 Punkt)

(e) Warum kann man die Terme: $-\ln 2\pi - \ln p(\mathbf{x})$ auslassen, ohne dass die Entscheidungsregionen und Entscheidungsgrenzen sich ändern. (1 Punkt)

(f) Sie nehmen an, dass $P(\omega_1) = P(\omega_2) = \frac{1}{2}$. Die Diskriminanten-Funktionen für die zwei Klassen schauen so aus:

$$g_1(\mathrm{x}) = -\frac{3}{4}x_1^2 + 3x_1 - \frac{3}{4}x_2^2 + \frac{9}{2}x_2 - \frac{39}{4} - \frac{1}{2}\ln\left(\frac{4}{9}\right) + \ln\left(\frac{1}{2}\right) \tag{6}$$

$$g_2(\mathrm{x}) = -\frac{3}{4}x_1^2 + 9x_1 - \frac{3}{4}x_2^2 + \frac{9}{2}x_2 - \frac{135}{4} - \frac{1}{2}\ln\left(\frac{4}{9}\right) + \ln\left(\frac{1}{2}\right) \tag{7}$$

wobei x_1 und x_2 die zwei Komponenten vom einem belibiegen Vektor x sind. Berechnen Sie den Ausdruck für die Entscheidungsgrenze. (1 Punkt)

(g) Begründen Sie, warum die von Ihnen berechnete Entscheidungsgrenze richtig ist. (1 Punkt)

(h) Schreiben Sie den Ausdruck für die Mahalanobis-Distanz d zwischen einem beliebigen Vektor $\mathbf{x}$ und der Klasse 1 (mit Mittelwertvektor $\hat{\mu}_1$ und Kovarianzmatrix $\hat{\Sigma}_1$). (1 Punkt)

(i) Berechnen Sie die Mahalanobis Distanz zwischen Klasse 1 und dem Vektor

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{8}$$

Schreiben Sie die berechnete Distanz ins untere Rechteck. (1 Punkt)
TIP: Die Inverse von einer 2×2 Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{9}$$

ist

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{10}$$

Mahalanobis Distanz $d =$ [ ]

3 Detaillierte Fragen

1. Beschreiben Sie kurz das k-NN Verfahren. Welche Vor- bzw. Nachteile besitzt es gegenüber Verfahren wie dem Perceptron oder dem Mahalanobis Distanz Klassifikator? (2 Punkte)

2. Beschreiben Sie das *Single Linkage Hierarchical Clustering* Verfahren. Wie funktioniert der Algorithmus? Erklären Sie mittels einem 2-dimensionalen Beispiel. (2 Punkte)

3. Sie haben die Klassenwahrscheinlichkeiten $P(\omega_1|\mathrm{x})$ und $P(\omega_2|\mathrm{x})$ für ein Problem mit zwei Klassen berechnet, wobei x der Merkmalsvektor ist. Für welche Klasse soll, laut *Bayes Decision Rule* (Bayes Entscheidungsregel), entschieden werden? Beweisen Sie, dass die Bayes-Entscheidungsregel die Fehlerwahrscheinlichkeit $P(error)$ im Fall von 2 Klassen immer minimiert. (2 Punkte)

4 Aufsatz

Wählen Sie eines der folgenden Themen und schreiben Sie 100–200 Wörter darüber (6 Punkte). Die Themen sind:

1. PCA in der Gesichtserkennung.
2. Der Aufbau von Entscheidungsbäumen.
3. Bias, Varianz und Generalisierungsfähigkeit.