

### **Aufgabe 1: [DM-Motivation]**

- a) Beschreiben Sie kurz, was KDD von OLAP unterscheidet.
- b) Warum werden in der Definition von KDD Muster und Zusammenhänge in für den Menschen interpretierbarer Form angestrebt?

### **Aufgabe 2: [DM-Aufgaben]**

- a) Nennen und erklären Sie kurz die drei Eigenschaften, mit denen sich Data-Mining-Aufgaben beschreiben lassen.
- b) Was ist der Unterschied zwischen kategorischen und linearen Werten?
- c) Erläutern Sie den Unterschied der in der Vorlesung behandelten Lernarten? Ordnen Sie drei Ihnen bekannte Data-Mining-Aufgaben den verschiedenen Lernarten zu.

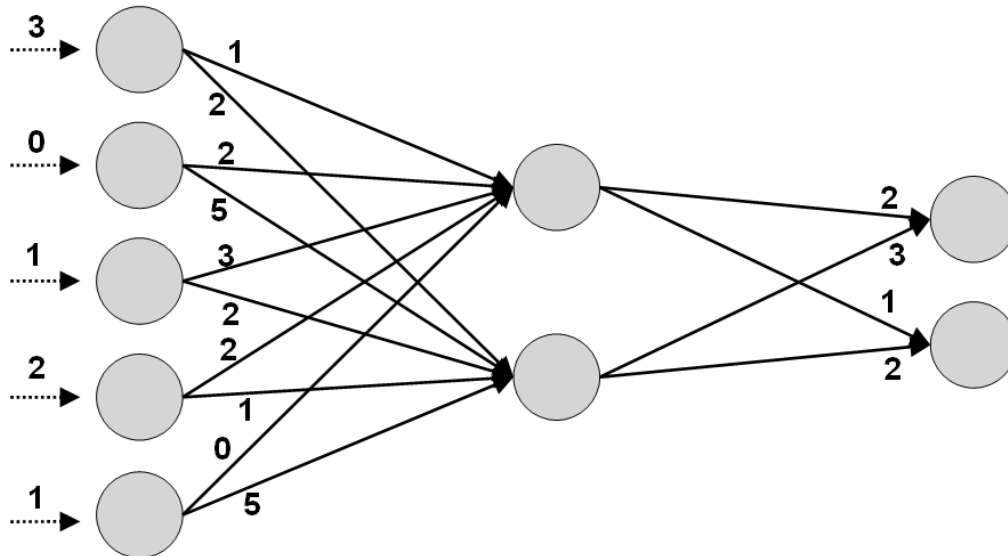
### Aufgabe 3: [DM-Verfahren: Neuronale Netze]

Schwellwertfunktion

$$s(x) = \frac{1}{5}x$$

Eingabeschicht

Ausgabeschicht



Wenden Sie die Schwellwertfunktion ausschließlich auf die versteckten Schichten an!

a) Vervollständigen Sie das gegebene Neuronale Netz bei Eingabe (3 / 0 / 1 / 2 / 1) und der Schwellwertfunktion  $s(x) = \frac{1}{5}x$ .

b) Wie muss das Netz aussehen bei Eingabe (3 / 2 / 4 / 0 / 2) und der Schwellwertfunktion  $s(x) = \frac{1}{10}x$ ?

#### **Aufgabe 4: [DM-Verfahren: Assoziationsregeln]**

TID	Itemset
1	{1,3,4}
2	{2,3,5}
3	{1,2,3,5}
4	{2,5}

- a) Die minimale Unterstützung sei 0,4. Bestimmen Sie für die gegebene Tabelle iterativ die Unterstützung für alle Itemsets und deren Kombinationen.
- b) Wieviele Scan-Vorgänge sind tatsächlich nötig (mit Begründung)?
- c) Bestimmen Sie die Konfidenz der Regeln  $3 \rightarrow 5$  und  $5 \rightarrow 3$ .
- d) Was halten Sie von der Regel  $4 \rightarrow 3$  (mit Begründung)?

#### **Aufgabe 5: [DM-KDD-Prozess]**

- a) Erläutern Sie, wieso der KDD-Prozess ein iterativer Prozess ist.
- b) Welchen Anteil hat die Phase „Data Understanding“ in einem KDD-Prozess und aus welchen Aufgaben besteht sie?