Barbara Hammer, FB Mathematik/Informatik, Universität Osnabrück

Softcomputing (WS 2000/01)
4.Übungsblatt

Abgabe bis spätestens Donnerstag, 16.11.00, 13$\underline{^{00}}$ Uhr, Briefkasten im 4. Stock des AVZ

---

1. (7 Punkte)

   - Fete in einer nicht näher spezifizierten WG: Die 5 Personen des einen (nicht näher spezifizierten) Geschlechts trinken jeweils um die 5 (Dreieckszahl $(2, 5, 8)$) Bierchen, die 8 Personen des anderen Geschlechts jeweils um die 8 (Dreickszahl $(7, 8, 14)$) Bierchen.
   - FH-Fete: Es kommen etwa 500 Personen (Dreieckszahl $(450, 500, 550)$), die je um die 5 Bierchen (Dreieckszahl $(2, 5, 10)$) trinken.
   - Uniball: Pro Stunde gehen etwa 200 Bierchen weg ($(150, 200, 300)$), von 11 bis 1 Uhr das Doppelte. Der Ball dauert bis kurz nach 2 Uhr ($(2, 2.15, 3)$) und startet um 8.
   - Saufduell irgendwo tief in der Karibik: Kandidat 1 trinkt angenähert 20 Bierchen ($(15, 19, 20)$), Kandidat 2 trinkt etwa 30 Bierchen ($(25, 30, 35)$), wobei das mindestens $2/3$ so viel wie bei Kandidat 3 ist, und Kandidat 3 trinkt dieselbe Menge, wie beide anderen zusammen, wobei das mindestens 50 Bierchen sind. (Bedenken Sie: Fuzzy ist dazu da, auch Unmögliches zu modellieren.)

   Welche Gesamtverbrauchs-Bierchen-Fuzzymengen ergeben sich jeweils? Sofern deren Zugehörigkeitsfunktion nicht stückweise linear ist, reicht es, wenn Sie den Support und den 0.5 und 1-Schnitt angeben.

2. (5 Punkte) Wie verhalten sich die Schnitte eines Produkts von Fuzzymengen A und B gegenüber den Schnitten von A und B?

   Wie verhalten sich die Schnitte der Projektion einer Fuzzymenge A auf $X \times Y$ (Projektion nach X) gegenüber den Schnitten von A?

   Wie verhalten sich die Schnitte einer max-min-Komposition von Fuzzymengen A auf $X \times Y$ und B auf $Y \times Z$ gegenüber den Schnitten von A und B, sofern in X und Y nur endlich viele Elemente stehen?

3. (8 Punkte) Nochmal Clustering, diesmal zum Lernen einer unbekannten Funktion. Gegeben sind Trainingspunkte (x_i, y_i). Es soll die unbekannte Zuordnung $x \mapsto y$ anhand dieser Daten gelernt werden. Dazu wird auf den x_i ein Clustering durchgeführt; den Zentroiden X_j werden Funktionswerten Y_j sinnvoll zugeordnet (*); anschließend kann man Punkte x mit unbekannter Ausgabe auf die Werte $\sum A_j(x)Y_j$ abbilden, wobei $A_j(x)$ die gemäß dem fuzzy-k-means Algorithmus berechnete Zugehörigkeit von x zur Klasse j beschreibt. In (*) sind mehrere Möglichkeiten denkbar, eine davon ist folgende: $Y_j = \sum A_j(x_i)y_i / \sum A_j(x_i)$ bildet das Mittel über die dem Cluster zugeordneten Trainingspunkte.

   Implementieren Sie diese Erweiterung, wobei Sie in (*) eine beliebigen (Ihnen sinnvoll erscheinenden) Ansatz wählen können. Testen Sie, was sich dann auf den Daten aus auto1.dat (zum Training) und auto2.dat (nur zum Testen, berechnen Sie auf diesen Daten z.B. die gemittelte absolute Differenz der je aus dem Clustering prognbostizierten und tatsächlich gegebenen Funktionswerte.) ergibt. (Auto: Vorhersage des Spritverbrauchs eines Autos anhand von Gewicht, Baujahr, ..., der Verbrauch steht je Zeile an der letzten Stelle.)

   Testen Sie, was sich auf den Daten aus iris1.pat (zum Training) und iris2.pat (zum Testen) ergibt. (Iris: Die ersten vier Werte pro Zeile sind repräsentative Daten für die jeweilige Iris-Pflanze (Größe von Blüte und Stengel), der Funktionswert ist eine von drei möglichen Sorten, er ist durch den letzten Wert in der Zeile gegeben.)

P.s.: ... es kommt ab dem nächsten Zettel kein fuzzy-clustering mehr dran!