

T-61.3025 Hahmontunnistuksen perusteet - Principles of pattern recognition

Tentti 20. 2. 2014 (ENGLISH on the other side)

1. On annettuna kaksi lineaarisesti erottuvaa luokkaa d -ulotteisessa hahmoavaruudessa. Esitä kolme erilaista perusteltua tapaa muodostaa luokille lineaarinen luokitin (joka siis jakaa hahmoavaruuden kahteen puoliavaruuteen sopivan hypertason avulla).
2. Muodosta yksiulotteinen Parzen-estimaatti tiheysfunktiolle p_x annetuista näytteistä

$$x^{(i)} : 2.5, 2.8, 3.4, 4.2, 4.5, 4.7, 5.2, 5.6, 7.5.$$

Käytä tasakylkisen kolmion muotoista ikkunafunktiota (kernel-funktiota). Valitse sen leveys (kannan pituus) sopivasti. Matemaattista lauseketta ei tarvita, vaan piirtämällä tehty ratkaisu kelpaa.

3. Esitä, mikä on luokittelun XOR-ongelma, ja ratkaise se jollakin epälineaarilla luokittimella.
4. Kuvan 1 koneenosassa on primitiivit a, c, e . Hyväksykö seuraava kielioppi osan:

$$G = (V_T, V_N, P, S),$$

$$V_T = \{a, c, e\}$$

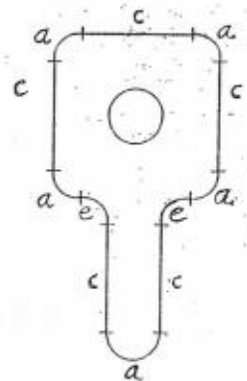
$$V_N = \{S, A, C, E, G, H, I, J, Q, R\}$$

ja produktiosäännöt P :

$$S \rightarrow GQ, S \rightarrow QG, G \rightarrow EH, H \rightarrow AIA,$$

$$I \rightarrow CJC, J \rightarrow ACA, Q \rightarrow ER,$$

$$R \rightarrow CAC, A \rightarrow a, C \rightarrow c, E \rightarrow e.$$



5. Tarkastellaan seuraavaa läheisyysmatriisia viidelle piirrevektorille:

$$P = \begin{bmatrix} 0 & 4 & 9 & 6 & 5 \\ 4 & 0 & 1 & 8 & 7 \\ 9 & 1 & 0 & 3 & 2 \\ 6 & 8 & 3 & 0 & 1 \\ 5 & 7 & 2 & 1 & 0 \end{bmatrix}$$

Käytä sekä yksinkertaista (single) että täydellistä (complete) linkkialgoritmia ja hierarkista klusterointia rakentaaksesi dendrogrammit P :stä. Kommentoi tuloksia.

ENGLISH TRANSLATIONS (FINNISH on the other side)

1. We are given two linearly separable classes in a d dimensional pattern space. Present three different well-founded ways to build a linear classifier for the classes (the classifier divides the pattern space into two half-spaces by a suitable hyperplane).
2. Construct a one-dimensional Parzen estimate for the density function p_x based on the sample

$$x^{(i)} : 2.5, 2.8, 3.4, 4.2, 4.5, 4.7, 5.2, 5.6, 7.5.$$

Use a window (kernel) function that has a shape of an isosceles triangle (two of the sides have equal lengths). Choose the width (length of the base) suitably. Mathematical expression is not needed, it is enough to plot the solution.

3. Present the XOR problem in classification, and solve it using a nonlinear classifier.
4. The machine part in Fig. 1 (on the previous page) has the primitives a, c, e . Does the following grammar accept the part:

$$G = (V_T, V_N, P, S),$$

$$V_T = \{a, c, e\}$$

$$V_N = \{S, A, C, E, G, H, I, J, Q, R\}$$

and production rules P :

$$S \rightarrow GQ, S \rightarrow QG, G \rightarrow EH, H \rightarrow AIA,$$

$$I \rightarrow CJC, J \rightarrow ACA, Q \rightarrow ER,$$

$$R \rightarrow CAC, A \rightarrow a, C \rightarrow c, E \rightarrow e.$$

5. Consider the following proximity matrix for five feature vectors:

$$P = \begin{bmatrix} 0 & 4 & 9 & 6 & 5 \\ 4 & 0 & 1 & 8 & 7 \\ 9 & 1 & 0 & 3 & 2 \\ 6 & 8 & 3 & 0 & 1 \\ 5 & 7 & 2 & 1 & 0 \end{bmatrix}$$

Apply hierarchical clustering and the single link and complete link algorithms to P and comment on the resulting dendrograms.