

Învățare Automată
Semiar 1: Bayes Naiv și Metoda celor mai apropiați vecini

1. Fiind dată mulțimea de exemple de antrenare $S = \{([0.5, -0.7], 1), ([0.2, 0.1], 1), ([0.4, 0.8], 2), ([-0.3, -0.8], 3), ([0.4, -0.3], 1), ([0.9, 0.3], 2), ([-0.5, -0.5], 3), ([0.2, 0.4], 2)\}$ și mulțimea de exemple de testare $T = \{([0.3, -0.1], 1), ([-0.1, -0.1], 2), ([0.2, 0.3], 2), ([-1.0, -0.5], 3)\}$:

a) Aplicați modelul Bayes naiv pentru a prezice etichetele exemplurilor de test, folosind două intervale pentru spațiul de valori.

b) Care este eroarea de clasificare a modelului Bayes naiv pentru mulțimea de testare dată?

2. Fiind dată mulțimea de exemple de antrenare $S = \{([-2, 1, 3, 1, -3, 1], 1), ([3, 1, -3, -1, 4, 0], -1), ([-1, 3, 0, 1, -2, 1], 1), ([0, 3, -5, -1, 1, 0], -1)\}$ și exemplul de testare $[0, 1, -2, -4, 2, 0]$:

a) Să se normalizeze exemplele de antrenare și testare folosind norma L_2 .

b) Determinați eticheta exemplului de test folosind un model 3-NN bazat pe distanța L_1 .