

2. MEĐUISPIT iz STROJNOG UČENJA 2010./2011.

PROBABILISTIČKI GENERATIVNI MODELI

1. BAYESOVA TEORIJA ODLUČIVANJA

- a) Napisati Bayesovu formulu, raspisati nazivnik, navesti nazive izraza koji se pojavljuju u formuli.
- b) Objasniti (pomoću skice) zašto je optimalna klasifikacijska odluka ona koja maksimizira a posteriori vjerojatnost.

2. PROCJENA PARAMETARA

- a) Procjeniteljem Θ procjenjujemo parametar populacije θ . Θ je pristran, ali ga se može korigirati tako ga se pomnoži konstantom α . Izrazite srednju kvadratnu grešku ovog procjenitelja rastavom na pristranost i varijancu (nije potreban izvod).
- b) Formalno definirajte funkciju izglednosti $L(\theta, D)$ i odgovorite koja je pretpostavka ugrađena u tu definiciju.
- c) Izračunajte log-izglednost $L(\mu=1, \sigma^2=1 | D)$ za $D=\{0.2, 0.5, 1, 2, 8, 10\}$.
- d) Napisati izvod za procjenitelja najveće izglednosti za μ i σ^2 normalne razdiobe $N(\mu, \sigma^2)$.

3. Koristimo Bayesov klasifikator za binarni problem. Primjer:

$$D = \{(6.1, y_1), (6.4, y_1), (2.1, y_1), (6.7, y_1), \\ (-6.0, y_2), (-1.8, y_2), (-2.5, y_2), (-1.1, y_2), (6.6, y_2), (-1.8, y_2)\}.$$
$$y_1 = (1, 0)^T, y_2 = (0, 1)^T. \text{ Pretpostavka je da su } x \text{ unutar razreda } C_j \text{ normalno distribuirani } N(\mu_j, \sigma_j^2).$$

- a) Odredite parametre gustoća $p(x, C_1)$, $p(x, C_2)$ i skicirajte te funkcije.
- b) Izvedite hipotezu $h_1(x | \mu_1, \sigma_1^2)$ koja inducira pouzdanost da primjer x pripada klasi C_1 .
- c) Objasnite što sve čini induktivnu pristranost dobivene hipoteze.
- d) Za neka dva primjera x_1, x_2 vrijedi $h_1(x_1) = h_2(x_2)$, $h_1(x_2) > h_2(x_2)$. U koju klasu klasificiramo x_1 i x_2 uz pretpostavku jednolikih gubitaka?

4. Navesti jednu prednost i jednu manu probabilističkih generativnih modela.

STABLA ODLUČIVANJA

5. Zadatak! Zadana je tablica s atributima. Potrebno je odabrati attribute u korijenskom čvoru pomoću algoritma ID3.
6. Razmatramo skup primjera za učenje D koji je podskup od R^2 za koje ne postoji hipoteza:
$$h_1(x_1, x_2 | \theta_0, \theta_1, \theta_2) = 1(\theta_1 x_1 + \theta_2 x_2 + \theta_0 > 0)$$
koja bi bila konzistentna s primjerima za učenje. Možemo li s ID3 za svaki takav skup D naučiti hipotezu h_2 koja jest konzistentna s primjerima za učenje. Obrazloži.
7. Usporediti induktivnu pristranost algoritma ID3 i algoritma eliminacije kandidata.
8. U skupu primjera D postoje dva označena primjera (x_1, x_2) i (y_1, y_2) takva da $x_1 = x_2$, ali $y_1 \neq y_2$. Kako to rješava ID3?
9. Koja je svrha podrezivanja stabla?

UČENJE NA TEMELJU PRIMJERA

10. Kako k utječe na VC-dimenzija modela k -NN?
11. Može li se dogoditi da hipoteza dobivena s k -NN nije konzistentna s primjerima za učenje? Navesti primjer.
12. Zašto bi netko koristio:
 - a) Klasifikaciju k -NN naspram Bayesove klasifikacije?
 - b) Bayesovu klasifikaciju naspram klasifikacije k -NN?
13. Navedite na koje se načine mogu odrediti središta radijalnih baznih funkcija u svrhu regresije?

VREDNOVANJE KLASIFIKACIJE

14. Na temelju elementarnih matrica zabune TP, FP, FN i TN definirajte mjere točnosti (Acc), preciznosti (P) i odziva (R).
15. Izračunajte mjere mikro-F1 i makro-F1 temeljem matrice zabune (bila je zadana matrica).
16. Što znači mikro-F1 > makro-F1 i je li to očekivano ? (ne mogu dešifrirati tu zadnju riječ XD)
17. Od 1000 primjera, prvi je označivač kao pozitivno označio njih 40, a drugi 80. Za 25 primjera označivači se slažu da trebaju biti klasificirani pozitivno. Izračunajte k -statistiku i odgovorite je li slaganje između označivača zadovoljavajuće.