

**Jméno:**

1. (6 bodů). Definujte střední ztrátu (Bayesovskou ztrátu) strategie. Definujte význam všech použitých symbolů.
2. Máte  $n$  nezávislých pozorování gaussovsky rozdělené náhodné veličiny. Jaký je odhad střední hodnoty dle principu maximální věrohodnosti (2 body)? Směrodatná odchylka  $\sigma$  je známa. Dokažte pro jednorozměrný (2 bod) i vícerozměrný případ (2 body)!
3. Je dána trénovací množina  $T = \{(\mathbf{x}_i; k_i)\}, i = 1, \dots, 5, \mathbf{x}_i \in R^2, k \in \{1, -1\}$ ,  
 $T = \{(-2, 1; 1), (0, 0; -1), (0, 2; -1), (0, -3; -1), (2, 2, 1)\}$ .

Algoritmem Adaboost hledáte lineární kombinaci slabých klasifikátorů  $H(\mathbf{x}) = \sum \alpha_t h_t(\mathbf{x})$ . Máte k dispozici čtyři slabé klasifikátory a to:  $h_1(\mathbf{x}) = \delta(x_1 > 0.5)$ ,  $h_2(\mathbf{x}) = \delta(x_2 < -0.5)$ ,  $h_3(\mathbf{x}) = \delta(x_1 \leq 0.5)$ ,  $h_4(\mathbf{x}) = \delta(x_2 > 0)$ , kde  $\delta()$  nabývá hodnoty 1, je-li podmínka v závorce splněna, jinak -1.

Který z těchto klasifikátorů bude vybrán jako první (2 body)? Jakou bude mít váhu  $\alpha_1$  (2 body)? Jaké budou váhy  $w_i$  jednotlivých příkladů z trénovací množiny po prvé iteraci Adaboostu (nezapomeňte normalizovat tak, aby jejich součet byl roven jedné) (1 bod)? Který z těchto klasifikátorů bude vybrán jako druhý (1 bod)?

4. Algoritmus k-průměrů (k-means). Zapište kriteriální funkci, kterou se metoda snaží minimalizovat (1 bod)? Popište algoritmus (2 body). Dokažte (naznačte důkaz), že algoritmus konverguje do lokálního minima kriteriální funkce (2 body). Jak lze metodu k-means zobecnit (1 bod)?