

## Půlsemestrální zkouška IKR - 27. 3. 2013,

Login: .....

Podpis: .....

---

1. Popište, jaké úlohy je schopen řešit Perceptron. V jakých případech je trénovací algoritmus úspěšný a co je v takovém případě jeho řešení? V jakých případech naopak trénování selže?

- 
2. Mějme lineární klasifikátor pro dvě třídy popsany rovnicí  $y(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0$ . Popište, jak se změní rozhodovací hranice (hyperrovina oddělující třídy), když zvětšíme parametr  $w_0$

---

3. Mějme natrénované pravděpodobnostní modely (např. GMM) popisující rozložení hustoty pravděpodobnosti příznaků  $p(\mathbf{x}|C_1)$  a  $p(\mathbf{x}|C_2)$  dvou tříd  $C_1$  a  $C_2$ . Pro trénování těchto modelů jsme měli k dispozici 1000 trénovacích vzorů pro třídu  $C_1$  a 2000 trénovacích vzorů pro třídu  $C_2$ . Víme ale, že v cílové aplikaci nám bude přicházet desetkrát více vzorů z třídy  $C_1$ , než z třídy  $C_2$ . Napište, jak v cílové aplikaci budeme klasifikovat nově příchozí vzor  $x$ , tak abychom minimalizovali pravděpodobnost chyby.

---

4. Mějme dvě jednorozměrné náhodné veličiny  $x$  a  $y$  s gaussovským rozložením  $\mathcal{N}(x|\mu_x, \sigma_x^2)$  a  $\mathcal{N}(y|\mu_y, \sigma_y^2)$ , jaké bude společné (dvourozměrné) rozložení hustoty pravděpodobnosti  $p(x, y)$  za předpokladu, že jsou obě náhodné veličiny na sobě statisticky nezávislé.

---

5. Mějme kovarianční matici odhadnutou na určité množině trénovacích dat. Vlastní čísla této kovarianční matice odpovídají jistým variancím v těchto trénovacích datech. O jaké variance se jedná?

---

6. Odhad střední hodnoty gaussovského rozložení s maximální věrohodností (Maximum Likelihood) je prostý průměr trénovacích dat. Popište, jak se k tomuto výsledku dojde. Napište vzoreček objektivní funkce, kterou se snažíme maximalizovat (stačí v obecné podobě pro jakýkoli model). Naznačte postup řešení. Plné odvození bude oceněno bonusovým bodem. :-)

---

7. Popište, jak provedeme klasifikaci nového vzoru metodou k-nejbližších sousedů (K- nearest neighbours).

---

8. Co je to Detection Error Tradeoff (DET) křivka? Jakou závislost vyjadřuje? Jakou informaci z této křivky získáme navíc oproti prostému sečtení chyb, které udělá klasifikátor na testovacích datech?