Pattern Recognition 2024
Assignment#3

April 30, 2024

The format of your report is up to you. In general, your report should clearly show how you have obtained the results and a detailed analysis of your solutions. If you feel a bit inexperienced with writing scientific reports, have a look at the line[1]. I recommend chapter 4 of this document if (like me) English is not your mother language.

Q1. Let $\mathcal{D} = \{x_1, x_2, \cdots, x_n\}$ be a set of n independent labeled samples and let $\mathcal{D}_k(x) = \{x'_1, x'_2, \cdots, x'_k\}$ be the k nearest neighbors of x. Recall that the k-nearest-neighbor rule for classifying x is to give x the label most frequently represented in $\mathcal{D}_k(x)$. Consider a two-category problem with $P(\omega_1) = P(\omega_2) = 1/2$. Assume further that the conditional probability densities $p(x|\omega_i)$ are uniform within unit hyperspheres. Show that if k is odd the average probability of error is given by:

$$P_n(error) = \frac{1}{2^n} \sum_{j=0}^{(k-1)/2} C_n^j, \tag{1}$$

where C_n^j denotes a combination of selecting j items from a collection of n samples.

Q2. Suppose we have four normalized training samples under the two-category case: $\boldsymbol{y}_1 = (1, 4, 1)^\top, \boldsymbol{y}_2 = (1, 4, 2)^\top, \boldsymbol{y}_3 = (-1, 0, -1)^\top, \boldsymbol{y}_4 = (-1, -1, -1)^\top$. The generalized linear discriminant function $g(\boldsymbol{y}) = \boldsymbol{a}^\top \boldsymbol{y}$ is adopted to learn from the training samples and the criterion function to be minimized is set as $J_p(\boldsymbol{a}) = \Sigma_{\boldsymbol{y} \in \gamma}(-\boldsymbol{a}^\top \boldsymbol{y})$, where γ denotes the set of samples misclassified by $g(\cdot)$, i.e. $\{\boldsymbol{y}_i | \boldsymbol{a}^\top \boldsymbol{y}_i \leqslant 0, 1 \leqslant i \leqslant 4\}$.

(a) Given an initial model $\boldsymbol{a} = (-2, 0, 0)^\top$, if the **fixed-increment single-sample correction algorithm** is utilized to minimize the criterion function, what is the final resulting discriminant function with fixed learning rate $\eta = 1$?

(b) Given an initial model $\boldsymbol{a} = (-2, -1, 1)^\top$, if the **batch perceptron algorithm** is utilized to minimize the criterion function, what is the final resulting discriminant function with fixed learning rate $\eta = 0.5$ and threshold $\theta = 0.5$?

Q3. Given three samples $\boldsymbol{x}_1 = (1, 1, -2)^\top$, $\boldsymbol{x}_2 = (1, -2, 1)^\top$ and $\boldsymbol{x}_3 = (-2, 1, 1)^\top$, please reduce the original 3-dimensional samples to 1-dimensional samples using **principal component analysis** (PCA).

[1] http://www.cs.joensuu.fi/pages/whamalai/sciwri/sciwri.pdf