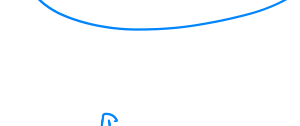
Weds

Monday

Monday: Read problem & given code



(non invertible)

(size 20)

Problem 3.3 *(Comparing k-NN performance)* Let $S := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 \le 1\}$ be a spherical ball of radius 1 in $\mathbb{R}^d$. Consider a binary classification problem in which the class labels are equally likely, i.e., $P(Y = 0) = P(Y = 1) = 0.5$, and the feature vector $x \in \mathbb{R}^d$ with the following class-conditional densities:

uniform prior $\Rightarrow$ ML?

$$p(x|y) = \text{Uniform}(\boldsymbol{\mu}_y + S), \; y = 0, 1,$$

with $\|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0\| > 4$. Let $\mathcal{D} := \{(\mathbf{X}_j, Y_j), j = 1, \ldots, n\}$ be n IID labeled training examples with joint distribution $1/2\, p(\mathbf{x}|y)$. Let $(\mathbf{X}_{\text{test}}, Y_{\text{test}})$ be a test pair which is drawn independently of $\mathcal{D}$ according to the same joint distribution $0.5p(\mathbf{x}|y)$. Let $h_{k\text{-NN}}(\mathbf{x}; \mathcal{D})$ denote the k-NN decision rule based on $\mathcal{D}$, where k is an odd positive integer. Since $\mathcal{D}$ is random, the decision rule $h_{k\text{-NN}}(\mathbf{x}; \mathcal{D})$ evaluated at any point $\mathbf{x}$ is a random variable.

(a) Compute the MAP rule and its misclassification probability.

$$p(x|y=0) = \text{Uniform}(\mu_{y=0} + S)$$
$$p(x|y=1) = \text{Uniform}(\mu_{y=1} + S)$$

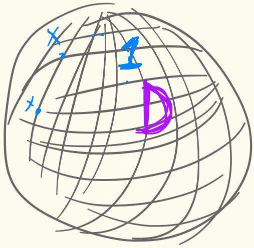
$$h_{MAP}(x) = \text{Posterior mode} = \arg\max_y p(y|x) = \arg\max_y p(x|y)\,p(y)$$

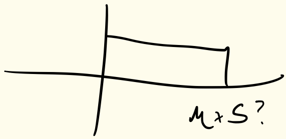
$$= \arg\max_y \tfrac{1}{2}\, p(x|y)$$

$$\arg\max_y \begin{cases} \cancel{\tfrac{1}{2}}\,\text{Uniform}(\mu_{y=0} + S) & y=0 \\ \cancel{\tfrac{1}{2}}\,\text{Uniform}(\mu_{y=1} + S) & y=1 \end{cases} \Bigg\} \; p(\cancel{y})\, p(x|y)$$

$$\text{Unif}(\mu_{y=0} + S) = \frac{1}{\mu_{y=0} + S}$$

S =





$\mu + S$?

B)

Problem 3.3 *(Comparing k-NN performance)* Let $S := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 \le 1\}$ be a spherical ball of radius 1 in $\mathbb{R}^d$. Consider a binary classification problem in which the class labels are equally likely, i.e., $P(Y = 0) = P(Y = 1) = 0.5$, and the feature vector $x \in \mathbb{R}^d$ with the following class-conditional densities:

$$p(x|y) = \text{Uniform}(\boldsymbol{\mu}_y + S),\ y = 0, 1,$$

with $\|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0\| > 4$. Let $\mathcal{D} := \{(\mathbf{X}_j, Y_j), j = 1, \ldots, n\}$ be n IID labeled training examples with joint distribution $1/2\,p(\mathbf{x}|y)$. Let $(\mathbf{X}_{\text{test}}, Y_{\text{test}})$ be a test pair which is drawn independently of $\mathcal{D}$ according to the same joint distribution $0.5p(\mathbf{x}|y)$. Let $h_{k\text{-NN}}(\mathbf{x}; \mathcal{D})$ denote the k-NN decision rule based on $\mathcal{D}$, where k is an odd positive integer. Since $\mathcal{D}$ is random, the decision rule $h_{k\text{-NN}}(\mathbf{x}; \mathcal{D})$ evaluated at any point $\mathbf{x}$ is a random variable.

(a) Compute the MAP rule and its misclassification probability.

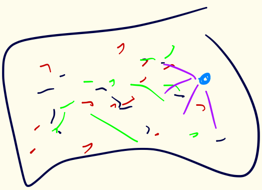
(b) Compute $P(h_{k\text{-NN}}(\mathbf{X}_{\text{test}}; \mathcal{D}) \ne Y_{\text{test}})$ for $k = 1, 3, 5, \ldots$.

(c) Compare and order the performance of the k-NN rule for $k = 1, 3, 5, \ldots$ from best to worst.

(d) Evaluate the misclassification probability of the k-NN rule for each k as $n \to \infty$.

$$P(h_{k\text{-}NN}(X_{test}; D) \ne Y_{test})$$

$K = 1$



for 1 to k NN's

hen,

$$h_{k-\text{NN}}(\mathbf{x}) = \arg\max_{y=1,\ldots,m} \underbrace{\sum_{j=1}^{k} 1(y_{(j)} = y)}_{\text{number of } k \text{ NNs of } \mathbf{x} \text{ with label } = y}$$

how many equal 1

how many equal 2

⋮

$$p_{k-\text{NN}}(y|\mathbf{x}) = \underbrace{\frac{1}{k}\sum_{j=1}^{k} 1(y_{(j)} = y)}_{\text{fraction of } k \text{ NNs of } \mathbf{x} \text{ with label } = y}$$

for $y = n$

how many k nn $= n$ / K

say $k=3$, then $P_{3nn}(y=a|x) = \frac{1}{3}\sum_{j=1}^{3}(y_{(j)} = a)$

max

$$P(h_{k\text{-}NN}(X_{test}; D) \ne Y_{test})$$

Probability of hypothesis for X_{test} being wrong

b)

$$P(h_{1NN}(X_{test}; D) \ne Y_{test}) = 1 - 1(y_1 = Y_{test})$$

$$P(h_{3NN}(X_{test}; D) \ne Y_{test}) = 1 - \frac{1}{3}\sum_{j=1}^{3} 1(y_{(j)} = Y_{test})$$

$$P(h_{5NN}(X_{test}; D) \ne Y_{test}) = 1 - \frac{1}{5}\sum_{j=1}^{5} 1(y_{(j)} = Y_{test})$$