

The background of the slide features a large, faint, light blue seal of Georgetown University. The seal is circular and contains an eagle with a shield on its chest, holding an olive branch and arrows. Above the eagle is a lyre. The text "MACI IN MEXICO" is at the top, "UTRAQUE UNUM" is on a banner across the eagle, and "GEORGIOPOLITANA" is at the bottom.

ANLY-601

Advanced Pattern Recognition

Assignment Project Exam Help
Spring 2018

<https://powcoder.com>

Add WeChat powcoder

L17 --- Nonparametric Classifiers

GEORGETOWN
UNIVERSITY

Non-Parametric Classifiers

Kernel or Parzen Classifiers

The obvious approach is to use a kernel density estimate for each of the class-conditional densities and proceed with a likelihood ratio test

$$\frac{\hat{p}(x|\omega_1)}{\hat{p}(x|\omega_2)} = \frac{\frac{1}{N_1} \sum_{i=1}^{N_1} \kappa_1(x - x_j^{(1)})}{\frac{1}{N_2} \sum_{i=1}^{N_2} \kappa_2(x - x_j^{(2)})} > t$$

Add WeChat powcoder

Error estimation – use leave-one-out (see x-validation) to set kernel width.
Leave-one-out is **cheap** for kernel estimators:

$$\hat{p}_L(x_k) = \frac{1}{N-1} \left[\left(\sum_{j=1}^N \kappa(x_k - x_i) \right) - \kappa(0) \right]$$

KNN Classifiers

Two paradigms:

- KNN density estimates followed by likelihood ratio test

Assignment Project Exam Help

- Voting KNN. This is what people usually mean when they refer to a KNN classifier.

<https://powcoder.com>

Add WeChat powcoder

For each test sample, select K nearest neighbors from the training data (all classes). The sample is classified according to the majority class of the k neighbors.

Asymptotics for Large N Error Rates for Voting KNN

K=1 Classify the point x as the class of its nearest neighbor

x •

$$\bullet \quad x_{NN} \in \omega_i \Rightarrow x \in \omega_i$$

Assignment Project Exam Help

The conditional risk is (classification error rate given x and x_{NN})

$$\begin{aligned} r_1 &= \text{Prob}[x \in \omega_1 \text{ and } x_{NN} \in \omega_2 \mid x, x_{NN}] + \text{Prob}[x \in \omega_2 \text{ and } x_{NN} \in \omega_1 \mid x, x_{NN}] \\ &= P(\omega_1 \mid x) P(\omega_2 \mid x_{NN}) + P(\omega_2 \mid x) P(\omega_1 \mid x_{NN}) \end{aligned}$$

<https://powcoder.com>

Add WeChat powcoder

In the limit $N \rightarrow \infty$ the data points become closer and closer together, and we can replace $P(\omega_i \mid x_{NN})$ with $P(\omega_i \mid x)$ leaving

$$r_1^* = \lim_{N \rightarrow \infty} r_1 = 2 p(\omega_1 \mid x) p(\omega_2 \mid x)$$

For Odd Number of Neighbors in General

Asymptotically, and for odd number of neighbors, the error rate at x becomes

$$r_{2k-1}^* = \lim_{N \rightarrow \infty} r_{2k-1} = \sum_{i=1}^k \frac{1}{i} \binom{2i-2}{i-1} [p(\omega_1 | x) p(\omega_2 | x)]^i + \frac{1}{2} \binom{2k}{k} [p(\omega_1 | x) p(\omega_2 | x)]^k$$

Assignment Project Exam Help

$$r_1^* = 2 p(\omega_1 | x) p(\omega_2 | x)$$

<https://powcoder.com>

$$r_3^* = p(\omega_1 | x) p(\omega_2 | x) + 4 (p(\omega_1 | x) p(\omega_2 | x))^2$$

Add WeChat powcoder

$$r_5^* = p(\omega_1 | x) p(\omega_2 | x) + (p(\omega_1 | x) p(\omega_2 | x))^2 + 12 (p(\omega_1 | x) p(\omega_2 | x))^3$$

$$r_7^* = p(\omega_1 | x) p(\omega_2 | x) + (p(\omega_1 | x) p(\omega_2 | x))^2 + 2 (p(\omega_1 | x) p(\omega_2 | x))^3 + 40 (p(\omega_1 | x) p(\omega_2 | x))^4$$

Bayes vs KNN Error

Recall that the **Bayes risk** at x is

$$r_B = \min (p(\omega_1|x), p(\omega_2|x)) = \frac{1}{2} \left(1 - \sqrt{1 - 4p(\omega_1|x)p(\omega_2|x)} \right)$$

or by Taylor series expansion

$$r_B = \sum_{i=1}^{\infty} \frac{1}{i} \binom{2i-2}{i-1} [p(\omega_1|x)p(\omega_2|x)]^i$$

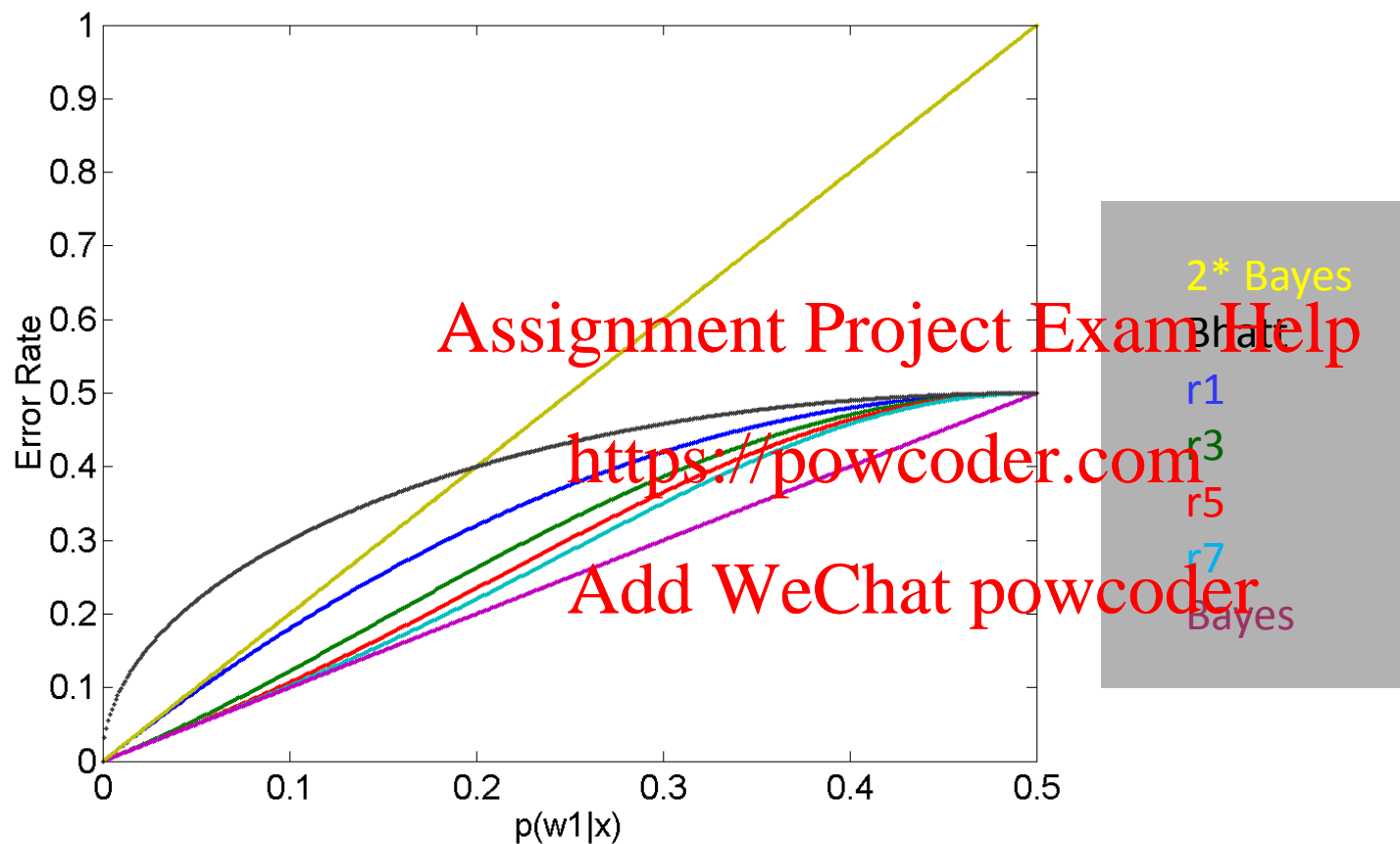
and the error rate for (odd) KNN for infinitely many training points

$$r_{2k-1}^* = \sum_{i=1}^k \frac{1}{i} \binom{2i-2}{i-1} [p(\omega_1|x)p(\omega_2|x)]^i + \frac{1}{2} \binom{2k}{k} [p(\omega_1|x)p(\omega_2|x)]^k$$

One can use these expressions to find the famous bound

$$r_B \leq \dots \leq r_5^* \leq r_3^* \leq r_1^* \leq 2 r_B$$

Bayes vs KNN Error



The KNN bound is tighter than the Bhattacharya bound $\sqrt{p(\omega_1 | x) p(\omega_2 | x)}$

Error Rate for Even Number of Neighbors

With even number of neighbors, we can have a tie (for two classes). When there's a tie, we don't make a choice. This leads to a slightly lower error rate

$$r_{2k}^* = \sum_{i=1}^k \frac{1}{i} \binom{2i-1}{i-1} [p(\omega_1|x) - p(\omega_2|x)]$$

<https://powcoder.com>

and a sequence of bounds **Add WeChat powcoder**

$$\frac{1}{2} r_B \leq r_2^* \leq r_4^* \leq r_6^* \leq \dots \leq r_B \leq \dots \leq r_5^* \leq r_3^* \leq r_1^* \leq 2 r_B$$

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder



GEORGETOWN UNIVERSITY