1vs1 classification requires KC2 linear classifiers (K = number of classes)

1vs all classification requires K linear classifiers

A linear classifier is of the form y = sign(w.x), or y = 1 if w.x > c1, 2 if c1 > w.x > c2, 3 if c2 > w.x > c3 etc

x1 > c1 and x2 > c2: this is NOT a linear classifier (it can be called a multi-linear classifier)

Logistic ridge regression = logistic loss (cross-entropy) + $\lambda * ||w||_2^2$

To solve for w: calculate the gradient with respect to w and use gradient descent (as done in case of logistic regression)

You are studying a process where the readings always lie in the range [0,1], under two possible conditions. You cannot observe the condition behind each reading, but have the prior knowledge that both conditions are equally likely. From each reading, you need to figure out the corresponding condition. Considering the reading as a random variable, you know that it follows the following probability distribution under the two conditions:

Under condition 1, the reading may be any value (uniformly) between 0 and 0.7 with probability 0.5, and any value (uniformly) between 0.7 and 1 with probability 0.5.

Under condition 2, the reading may be any value ‘x’ between 0,1 with probability proportional to x

a) Calculate the probability density functions of the distributions under both conditions

Under C1, let f1(x) be the pdf of x between 0 and 0.7, and f2(x) be the pdf of x between 0.7 and 1.

Clearly, f1(x)*0.7 = 0.5, i.e. f1(x) = 5/7 for 0 < x < 0.7

Similarly, f2(x)*0.3 = 0.5, i.e. f2(x) = 5/3 for 0.7 < x < 1

Under C2, $f(x) = kx$. Since $\int f(x)dx = 1$, we can work out that $k = 2$.

Hence the pdf $f3(x) = 2x$ for $0 < x < 1$

b) Pose the above situation as a Bayes Classifier, using the PDFs as class-conditional. What are the predictions and confidence by the classifier for different readings?

For any x, we need to estimate $p(Y=C1|x)$ and $p(Y=C2|x)$.

$p(Y=C1|x) = 1/(p(x)) * p(x|Y=C1) * p(Y=C1)$

$p(Y=C2|x) = 1/(p(x)) * p(x|Y=C2) * p(Y=C2)$

Denote $1/p(x)$ by c. $p(Y=C1) = p(Y=C2) = 1/2$

For $0 < x < 0.7$, $p(Y=C1|x) = 5c/14$ and $p(Y=C2|x) = cx$. Clearly $5c/14 + cx = 1$. Hence we get c.

For $0.7 < x < 1$, $p(Y=C1|x) = 5c/6$ and $p(Y=C2|x) = cx$. Clearly $5c/6 + cx = 1$. Hence we get c.

Note that $c = 1/p(x)$ is a function of x.

Prediction: wherever $p(Y=C1|x) > p(Y=C2|x)$, predict C1, else C2.

For $0 < x < 0.7$, predict C1 for $x < 5/14$, but predict C2 for $5/14 < x < 0.7$

For $0.7 < x < 1$, predict C1 for $x < 5/6$, but predict C2 for $5/6 < x < 1$

Confidence at x: $p(y=C1/x)$ or $p(Y=C2/x)$ depending on the prediction

c)The Bayes Error is the total probability of making a wrong prediction (irrespective of the reading value). Calculate the Bayes Error, i.e. probability of predicting the wrong condition, in case of your Bayes classifier.

At any x ($0 < x < 5/14$), probability of error $e1(x) = p(Y=C2|x) = x/(5/14+x)$

Total error probability in this region $E1 = \int e1(x)dx$ [0 to 5/14]

For $5/14 < x < 0.7$, probability of error $e2(x) = p(Y=C1|x) = 5/(5+14x)$

Total error probability in this region $E2 = \int e2(x)dx$ [5/14 to 0.7]

Similarly find error E3 for $0.7 < x < 5/6$, and E4 for $5/6 < x < 1$.

Bayes Error = E1 + E2 + E3 + E4