Home Work 1

1.2)

1) The accuracy increases with increase in number of training instances. Gradually the rate of increase of accuracy decreases and the increase in accuracy becomes negligible with increase in number of training samples .This is my observation when i changed number of training samples for k =3.

2) The numbers (classes) 4(correct label) get confused with 9 (incorrect label) easily.

3) Initially at lower value of k the accuracy will be low because of overfitting (noise and outliers) . The accuracy increases with increase in k to one value(optimum value). After that optimal value , the accuracy decreases with increase in k because points which belong to other class increases in those k nearest points. The optimal value of k varies according to number of training samples. This is my observation when i increased value of k with limit = 500, 50000.

4) In general , smaller value of k causes overfitting(usually because of noise and outliers).

2.2)

1) k = 3

2) k = 1

3) The best k in problem 2 is not consistent with best performance k in problem.

3)

$$y = f(x) + \varepsilon, \quad E(\varepsilon) = 0, \quad Var(\varepsilon) = \sigma_\varepsilon^2$$

$$Err(x_0) = E\left((y_0 - h_S(x_0))^2 \,\middle|\, X = x_0 \right)$$

$$= E\left(\left(f(x_0) + \varepsilon - h_S(x_0) \right)^2 \right)$$

ε is independent of everything

$$= E\left[\varepsilon^2 + 2\varepsilon\left(f(x_0) - h_S(x_0)\right) + \left(f(x_0) - h_S(x_0)\right)^2 \right]$$

$$= E[\varepsilon^2] + 2E(\varepsilon)E\left(f(x_0) - h_S(x_0)\right) + E\left[\left(f(x_0) - h_S(x_0)\right)^2\right]$$

$(Var(\varepsilon) + (E(\varepsilon))^2)$

$(\because E(\varepsilon) = 0)$

$$= \sigma_\varepsilon^2 + 2 \times 0 \times E\left(f(x_0) - h_S(x_0)\right) + E\left[\left(f(x_0) - h_S(x_0)\right)^2\right]$$

$$= \sigma_\varepsilon^2 + E\left[\left(f(x_0) - h_S(x_0)\right)^2\right]$$

$(\because E(X^2) = Var(X) + (E(X))^2$

$$= \sigma_\varepsilon^2 + Var\left(f(x_0) - h_S(x_0)\right) + \left(E\left[f(x_0) - h_S(x_0)\right]\right)^2$$

$$= \sigma_\varepsilon^2 + Var\left(h_S(x_0)\right) + \left(E\left[f(x_0) - h_S(x_0)\right]\right)^2$$

$(\because f(x_0)$ is constant$)$

$Var(\text{constant}) = 0$ & $Var(A+B) = Var(A) + Var(B)$ when A, B are independent.

Finding ~~$h_S(x$~~ $Var(h_S(x_0))$

$$Var(h_S(x_0)) = Var\left(\frac{1}{k}\sum_{l=1}^{k} y(l)\right)$$

$$= \frac{1}{k^2} Var\left(\sum_{l=1}^{k} y(l)\right) \quad \left(\because Var(cx) = c^2 Var(x),\ c \text{ is constant.}\right)$$

$$= \frac{1}{k^2} Var\left(\sum_{l=1}^{k}(f(x_l) + \varepsilon)\right)$$

A, B are independent

$$\left(\because Var(A+B) = VAR(A) + Var(B)\right)$$

$$= \frac{1}{k^2}\left(\sum_{l=1}^{k}\left(Var(f(x_l)) + Var(\varepsilon)\right)\right)$$

$$= \frac{1}{k^2}\sum_{l=1}^{k} Var\, f(x_l) + \frac{1}{k^2}\sum_{l=1}^{k} Var(\varepsilon)$$

$$= \frac{1}{k^2} \times 0 + \frac{1}{k^2} \times k \times \sigma_\varepsilon^2$$

$\left(\because f(x_l)\right.$ is constant for given l, $\left. Var(f(x_l)) = 0\right)$

$$Var(h_S(x_0)) = \frac{1}{k} \times \sigma_\varepsilon^2$$

$$Err(x_0) = \sigma_\varepsilon^2 + \left(E\left(f(x_0) - h_S(x_0)\right)\right)^2 + \frac{\sigma_\varepsilon^2}{k}$$

Finding $\left(E\left(f(x_0) - h_S(x_0)\right)\right)^2$

$$\cancel{E\left(\frac{1}{k}\sum_{l=1}^{k} y(l)\right) - E}$$

$$= \left(E\left(f(x_0)\right) - E\left(h_S(x_0)\right)\right)^2$$

$$= \cancel{E\left(\frac{1}{k}\sum_{l=1}^{k}\right.}$$

$$= \left(E(f(x_0)) - E\left(\frac{1}{k}\sum_{l=1}^{k} y(l)\right)\right)^2$$

$$= \left(f(x_0) - \frac{1}{k} E\left(\sum_{l=1}^{k} (f(x_l)+\varepsilon)\right)\right)^2$$

$(\because E(f(x_0)) = f(x_0)) \Rightarrow f(x_0)$ is constant

$$= \left(f(x_0) - \frac{1}{k} \sum_{l=1}^{k} \left(E(f(x_l)) + E(\varepsilon)\right)\right)^2$$

$$= \left(f(x_0) - \frac{1}{k}\left(\sum_{l=1}^{k} E(f(x_l)) + \sum_{l=1}^{k} E(\varepsilon)\right)\right)^2$$

$$= \left(f(x_0) - \frac{1}{k}\left(\sum_{l=1}^{k} f(x_l) + \sum_{l=1}^{k} 0\right)\right)^2$$

$(\because f(x_l)$ is constant for given l $)$

$\cancel{(E(f(x_l)) = f(x_l))}$

$(E(f(x_l)) = f(x_l))$

$$= \left(f(x_0) - \frac{1}{k} \sum_{l=1}^{k} f(x_l) \right)^2$$

$$\therefore Err(x_0) = \sigma_\varepsilon^2 + \left(f(x_0) - \frac{1}{k}\sum_{l=1}^{k} f(x_l)\right)^2 + \frac{\sigma_\varepsilon^2}{k}$$