QUESTION

[redacted]

[redacted]

$$f(x^*) = \sum_{n=1}^{N} w_n \cdot y_n$$

Now, we have the closed-form solution for weight vector as

$$\hat{w} = (X^T X)^{-1} X^T y$$

where X is the $N \times D$ input matrix, with its rows as x_n^T

The prediction $f(x^*)$ is given by

$$f(x^*) = x^{*T} \hat{w}$$

$$f(x^*) = x^{*T} (X^T X)^{-1} X^T y \tag{1}$$

$$f(x^*) = x^{*T} (X^T X)^{-1} \sum_{n=1}^{N} x_n \cdot y_n$$

Since $x^{*T}(X^T X)^{-1}$ is independent of n, it can be taken inside the summation.

$$f(x^*) = \sum_{n=1}^{N} x^{*T} (X^T X)^{-1} x_n \cdot y_n$$

$$f(x^*) = \sum_{n=1}^{N} w_n \cdot y_n$$

where $w_n = x^{*T}(X^T X)^{-1} x_n$

In KNN, w_n depends on n'th training input(w_n depends on eucledian distance of x^* from n'th training input and there is no modulation), w_n in this case depends on all the training inputs, and the similarity $x_n^T x^*$ is modulated by $X^T X$ matrix.Also, prediction in case of KNN depends only on local neighbours, but here depends on all.