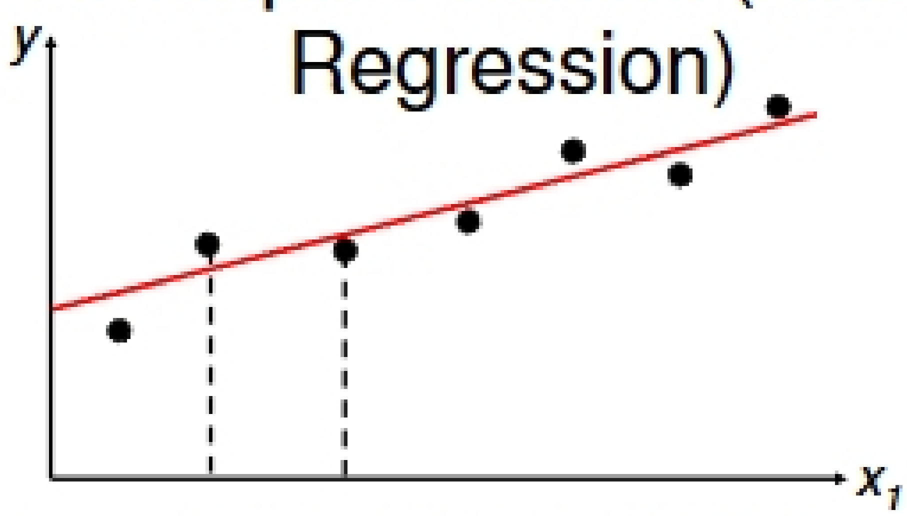
Neural Networks

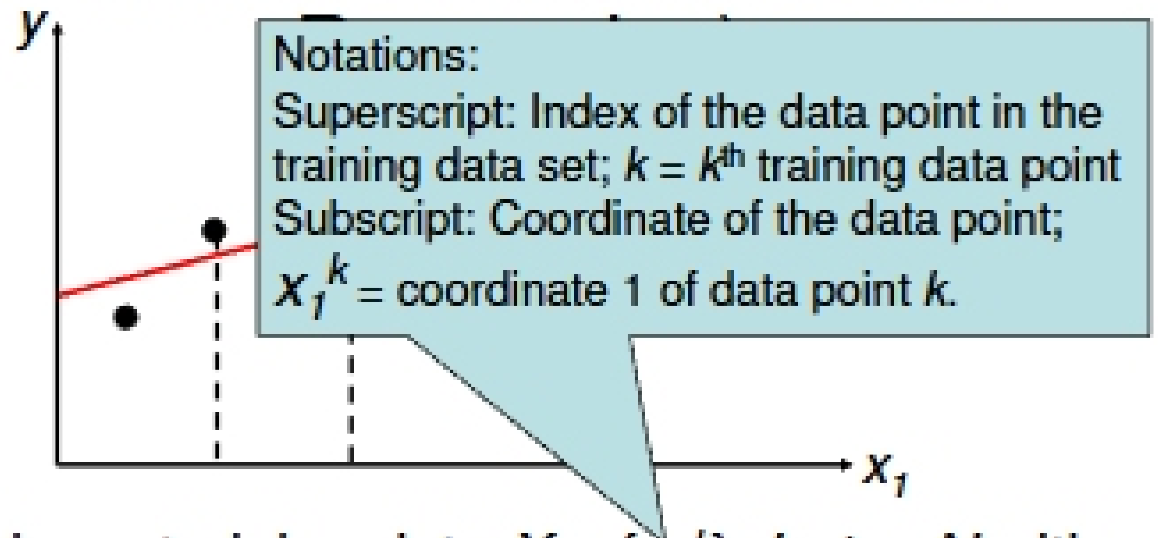
A Simple Problem (Linear Regression)



- We have training data $X = \{x_1^k\}$, i=1,..,N with corresponding output $Y = \{y^k\}$, i=1,..,N
- We want to find the parameters that predict the output Y from the data X in a linear fashion:

$$Y \approx w_o + w_1\, x_1$$

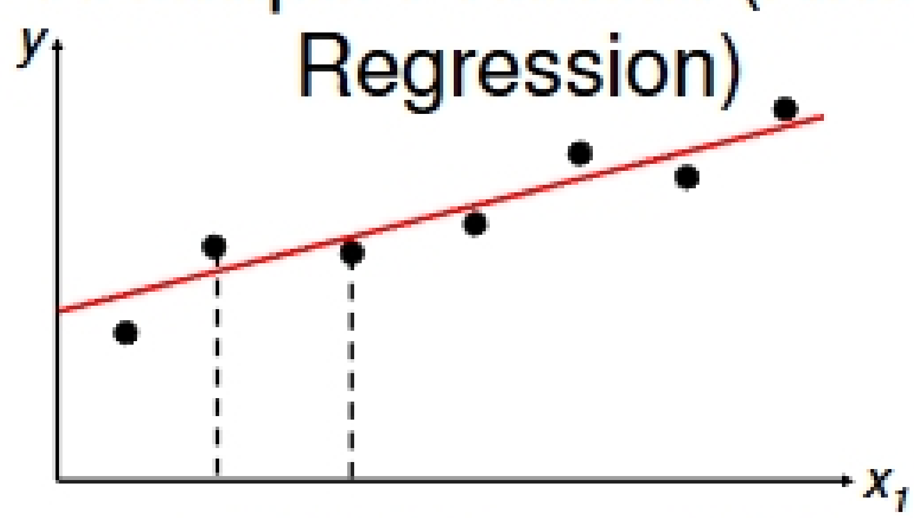
A Simple Problem (Linear



- We have training data $X = \{x_1^k\}$, $k=1,..,N$ with corresponding output $Y = \{y^k\}$, $k=1,..,N$
- We want to find the parameters that predict the output Y from the data X in a linear fashion:

$$y^k \approx w_o + w_1\, x_1^k$$

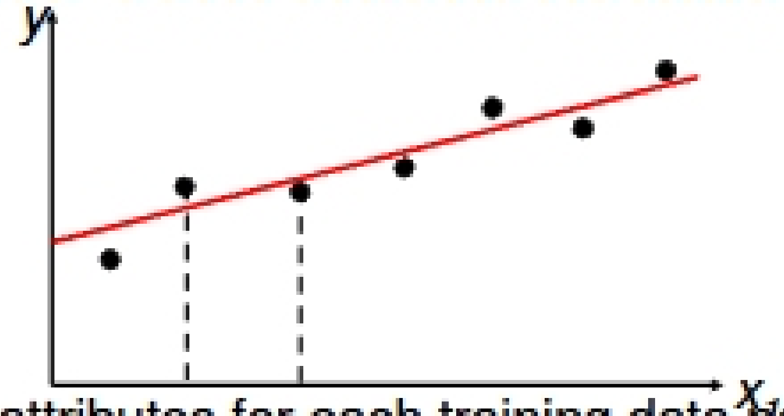
A Simple Problem (Linear Regression)



- It is convenient to define an additional "fake" attribute for the input data: $x_o = 1$
- We want to find the parameters that predict the output Y from the data X in a linear fashion:

$$y^k \approx w_o x_o^k + w_1\, x_1^k$$

More convenient notations



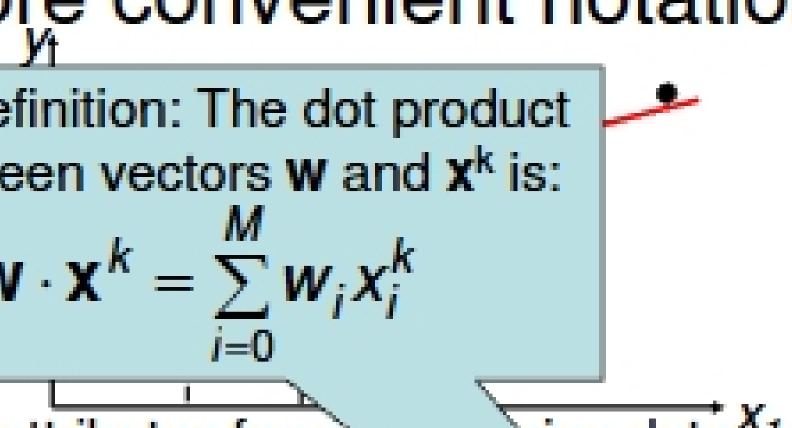
- Vector of attributes for each training data point:

$$\mathbf{x}^k = [x_o^k, .., x_M^k]$$

- We seek a vector of parameters: $\mathbf{w} = [w_o, .., w_M]$
- Such that we have a linear relation between prediction Y and attributes X:

$$y^k \approx w_o x_o^k + w_1 x_1^k + \cdots + w_M x_M^k = \sum_{i=0}^{M} w_i x_i^k = \mathbf{w} \cdot \mathbf{x}^k$$

More convenient notations



By definition: The dot product between vectors **w** and **x**^{k} is:

$$\mathbf{w} \cdot \mathbf{x}^k = \sum_{i=0}^{M} w_i x_i^k$$

- Vector of attributes for ea... ning data point:

$$\mathbf{x}^i = [x_o^i, ..$$

- We seek a vector of parameters: ... $[w_o, .., w_M]$
- Such that we have a linear relation be... een prediction Y and attributes X:

$$y^k \approx w_o x_o^k + w_1 x_1^k + \cdots + w_M x_M^k = \sum_{i=0}^{M} w_i x_i^k = \mathbf{w} \cdot \mathbf{x}^k$$