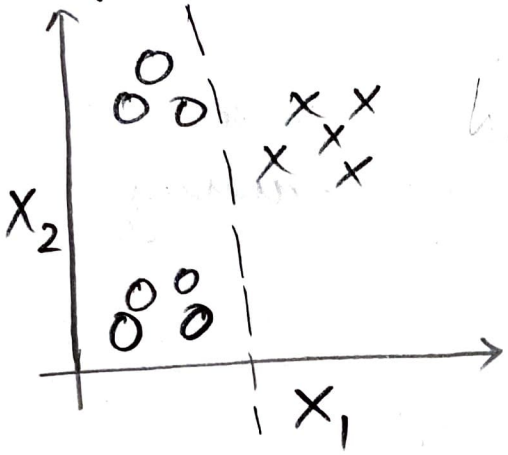
③ for Class 3



$h_\theta^{(3)}(x) = P(y=3|x;\theta)$

$h_\theta^{(i)}(x) = P(y=i|x;\theta)$

$(i = 1, 2, 3)$

One v/s all

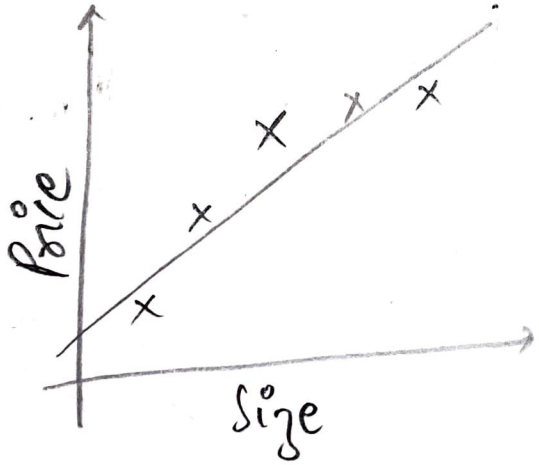
Train a logistic regression $h_\theta^{(i)}(x)$ for each class i to predict the probability that $y=i$.

On a new input x, to make a prediction, pick the class i that maximizes

$$\max_i h_\theta^{(i)}(x)$$

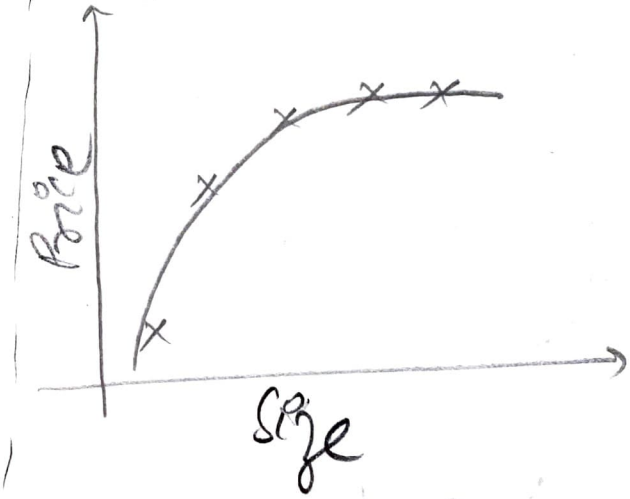
⇒ The problem of overfitting

Eg) Linear Regression (housing Prices)



$\theta_0 + \theta_1 x_1$

Underfit, High bias



$\theta_0 + \theta_1 x + \theta_2 x^2$

Just Right