

Leren: Opdracht 1

Deadline: 30 Oktober 2014

Deze opdracht moet individueel gemaakt worden

1. Een internet winkelier gebruikt gegevens van klanten om te achterhalen wat voor klanten wat voor producten kopen. Hij zoekt klanten op op facebook en koppelt deze informatie aan de producten die ze kopen.
 - (a) Wat is in dit geval een "trainingsvoorbeeld"?
 - (b) Geef enkele voorbeelden van variabelen die het trainingsvoorbeeld beschrijven.
 - (c) Geef een voorbeeld van een "supervised" en van een "unsupervised" leertaak die interessant kunnen zijn voor de winkelier. Hoe zou het resultaat van het leren eruit kunnen zien? Geef een voorbeeld ter illustratie.
2. Given the following data:

X	5	5	3
Y	6	6	10

- (a) Manually (using only a calculator) do linear regression. First, initialize the parameters θ_0 and θ_1 such that the regression function passes through the origin (0, 0) and has an angle of 45 degrees with the x-axis. Calculate the Mean Squared Error for this hypothesis.
 - (b) Use a "learning rate" of 0.1 do one iteration of gradient descent. Give the intermediate results of your calculations, the new θ_0 and θ_1 and also compute the mean-squared error of the hypothesis after 1 iteration.
3. Can we extend the gradient descent approach to finding univariate nonlinear, quadratic functions? E.g. can we apply the same approach to finding hypotheses of the form $y = \theta_0 + \theta_1 x + \theta_2 x^2$? If so, find the gradient and update rule. If not, explain the problem.
4. Derive an equation that can be used to find the optimal value of the parameter θ_1 for linear regression from a set of data (x, y) directly, without doing gradient descent. This can be done by setting the value of the derivative of the cost for θ_1 equal to 0. Assume that the value of θ_0 is constant.