

Problem 1: (35) The dataset is taken from Wikipedia’s Logistic regression [site](#), that was an interesting toy problem to solve.

A group of 20 students spends between 0 and 6 hours studying for an exam. How does the number of hours spent studying affect the probability of the student passing the exam?

Hrs (x_i)	1.0	2.0	3.0	4.0	5.0
Prob (p_i)	0.07	0.26	0.61	0.87	0.97
Pass (y_i)	0	0	0	1	1

- (15 pts) Use simple linear regression and predict the probability of pass if a student studied for 3.25 hours. Use pencil and paper to write down the detailed steps. You can use calculator or numpy for the actual calculation though.
- (20 pts) Use logistic regression to solve the above problem. Again, please provide details.

Problem 2: (30 pts) On Perceptron.

- (15 pts) Use Perceptron to implement the OR logic. Show output from each iteration (that is, the two inputs, the targeted output, and the Perceptron output) with the maximum number of iterations being 10.
- (15 pts) Use Perceptron to implement the XOR logic. Show output from each iteration (that is, the two inputs, the targeted output, and the Perceptron output) with the maximum number of iterations being 10.

Problem 3: (35) Comparison between FLD, PCA, and Perceptron.

Note that FLD and PCA are dimensionality reduction methods that only output a projection direction. Additional classification methods need to be applied to find the decision boundary. Suppose the minimum (Euclidean) distance (MD) classifier is used. On the other hand, Perceptron is a linear classifier that outputs the decision boundary directly. On the same figure, plot the four samples of the AND gate (5 pts), and compare the decision boundary from FLD+MD (10 pts), PCA+MD (10 pts), and Perceptron (10 pts). You can use whichever language that you feel comfortable (pencil & paper or Python).