Machine Learning Homework 1

Due on November 4, 2015

1.(20%) Rolled a die 10 times and got the outcomes as follows: $\{2, 3, 1, 5, 4, 6, 2, 1, 3, 1\}$.

(a) What is the maximum likelihood estimate for $Pr(X = 1)$ and $Pr(X = 4)$?

(b) What is the estimation of $Pr(X = 4)$ based on Laplace's law of succession ?

2.(30%) We are given a training data as follows:

Outlook	Temperature	Humidity	Windy	Play (Label)
Sunny	Hot (102°F)	High	False	-1
Sunny	Hot (105°F)	High	True	-1
Overcast	Hot (101°F)	High	False	+1
Rainy	Mild (90°F)	High	False	+1
Rainy	Cool (68°F)	Normal	False	+1
Rainy	Cool (72°F)	Normal	True	-1
Overcast	Cool (65°F)	Normal	True	+1
Sunny	Mild (94°F)	High	False	-1
Sunny	Cool (74°F)	Normal	False	+1
Rainy	Mild (86°F)	Normal	False	+1
Sunny	Mild (88°F)	Normal	True	+1
Overcast	Mild (90°F)	High	True	+1
Overcast	Hot (100°F)	Normal	False	+1
Rainy	Mild (88°F)	High	True	-1

Let

$$f_{T|P=+1}(x) \sim N(85, 16)$$

and

$$f_{T|P=-1}(x) \sim N(92, 25) \ ,$$

where $f_{T|P=y}$ stands for the conditional probability density function for the `Temperature` given `Play`=y.

(a) We have a new query x = {*Sunny, Hot, Normal, False*}. What is your prediction about `Play`? Please try two different learning algorithms that you learned in the class.

(b) Repeat (a), what is your prediction about `Play` when the query is {*Overcast,* 98°*F, Normal, True*}?

3.(50%) Let $A_+ = (0,0), (0.5,0), (0,0.5), (-0.5,0), (0,-0.5)$ and $A_- = \{(0.5,0.5), (0.5,-0.5), (-0.5,0.5), (-0.5,-0.5), (1,0), (0,1), (-1,0), (0,-1)\}$.

(a) Try to find the *hypothesis* by implementing the Perceptron algorithm in the *dual form* and replacing the inner product $\langle x^i \cdot x^j \rangle$ with $\langle x^i \cdot x^j \rangle^2$ and $R = \max_i \|x^i\|_2^2$

(b) Generate 10,000 points in the box $[-1.5, 1.5] \times [-1.5, 1.5]$ randomly as a test set. Plot these points into the hypothesis that you got in (a) and then plot the points for which $h(x) > 0$ with '+'

(c) Repeat (a) and (b) by using the training data B_+= {(0.5,0), (0,0.5), (-0.5,0), (0,-0.5)} and B_- = {(0.5,0.5), (0.5,-0.5), (-0.5,0.5), (-0.5,-0.5)}.

(d) Let the nonlinear mapping $\phi : R^2 \mapsto R^4$ defined by

$$\phi(x) = [x_1 x_2 \ , \ x_1^2 \ , \ -x_1 x_2 \ , \ x_2^2]$$

Map the training data A_+ and A_- into the feature space using this nonlinear map. find the hypothesis $f(x)$ by implementing the Perceptron algorithm in the *primal form* in the feature space.

(e) Repeat (b) by using the hypothesis that you got in (d). Please know that you need to map the points randomly generated in (b) by the nonlinear mapping ϕ first.