Subject Name : ML, Class Test 2, 30 Points

Q1 : [12 Points] (Python Exercise) : Fit the following data under Least Absolute Deviations regression using IRLS. The first line = "p n" where p is the number of predictors and n number of observations. The following lines are the data lines for predictor x and response variable y.
For instance : "<pred_1> ... <pred_p> y". The next line win "n" gives the number n of test cases to expect. The following lines are the test cases with predictors and expected response

```
input_str = '''2 7
0.18 0.89 109.85
1.0 0.26 155.72
0.92 0.11 137.66
0.07 0.37 76.17
0.85 0.16 139.75
0.99 0.41 162.6
0.87 0.47 151.77
4
0.49 0.18 105.22
0.57 0.83 142.68
0.56 0.64 132.94
0.76 0.18 129.71
'''

input_list = input_str.split('\n')

p,n = [ int(i) for i in input_list.pop(0).split() ]

X = empty( [n, p+1] )

X[:,0] = repeat( 1, n)

y = empty( [n, 1] )

for i in range(n):

        l = [ float(i) for i in input_list.pop(0).split() ]

        X[i, 1:] = array( l[0:p] )

        y[i] = array( l[p] )


n = [ int(i) for i in input_list.pop(0).split() ][0]
```

```
X_new = empty( [n, p+1] )
X_new[:,0] = repeat( 1, n)
y_new = empty( [n, 1] )
for i in range(n):
        l = [ float(i) for i in input_list.pop(0).split() ]
        X_new[i, 1:] = array( l[0:p] )
        y_new[i] = array( l[p] )


B = IRLS(y=y,X=X, maxiter=20)
abs_error = abs( y_new - X_new.dot(B) )
```

Implement IRLS to get abs_error for maximum iteration of 20. Choose optional parameters (if required) as per your own understanding.

HINT : Iteratively Reweighted Least Squares Numpy algorithm for IRLS in case of L1 regression. http://en.wikipedia.org/wiki/Iteratively_reweighted_least_squares

Q2 :

(6 points) We have a training set consisting of samples and their labels. All samples come from one of two classes, 0 and 1. Samples are two dimensional vectors. The input data is the form $\{X1, X2, Y\}$ where $X1$ and $X2$ are the two values for the input vector and Y is the label for this sample.

After learning the parameters of a Naïve Bayes classifier we arrived at the following table:

Table 1: Naïve Bayes conditional probabilities

	$Y = 0$	$Y = 1$
$X1$	$P(X1 = 1\|Y = 0) = 1/5$	$P(X1 = 1\|Y = 1) = 3/8$
$X2$	$P(X2 = 1\|Y = 0) = 1/3$	$P(X2 = 1\|Y = 1) = 3/4$

Denote by w_1 the probability of class 1 (that is $w_1 = P(Y = 1)$). If we know that the likelihood of the following two samples: {1,0,1},{0,1,0} given our Naïve Bayes model is 1/180, what is the value of w1? You do not need to derive an explicit value for w_1. It is enough to write a (correct ...) *equation* that has w_1 as the only unknown and that when solved would provide the value of w_1. Simplify as best as you can.

Q3 :

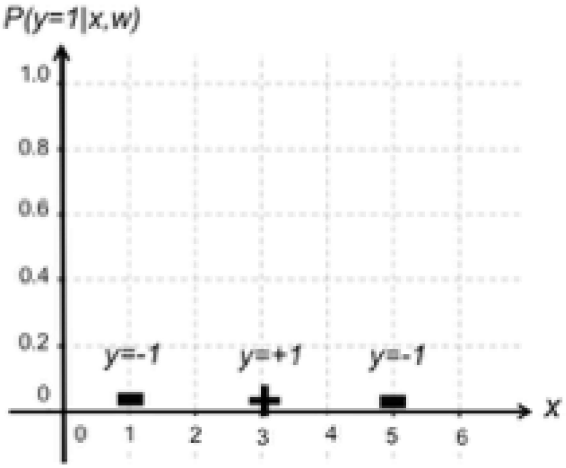
(6 points) We are given a set of two dimensional inputs and their corresponding output pair: $\{x_{i,1}, x_{i,2}, y_i\}$. We would like to use the following regression model to predict y:

$$y_i = w_1^2 x_{i,1} + w_2^2 x_{i,2}$$

Derive the optimal value for w_1 when using least squares as the target minimization function (w_2 may appear in your resulting equation). Note that there may be more than one possible value for w_1.

Q6 : [6 Points]

Assume that we have two possible conditional distributions ($P(y = 1|x, w)$) obtained by training a logistic regression on the dataset shown in the figure below:



In the first case, the value of $P(y = 1|x, w)$ is equal to 1/3 for all the data points. In the second case, $P(y = 1|x, w)$ is equal to zero for $x = 1$ and is equal to 1 for all other data points. One of these conditional distributions is obtained by finding the maximum likelihood of the parameter w. Which one is the MLE solution? Justify your answer in at most three sentences.