**Breakup of Internal Practical Marks for Machine Learning (25)**

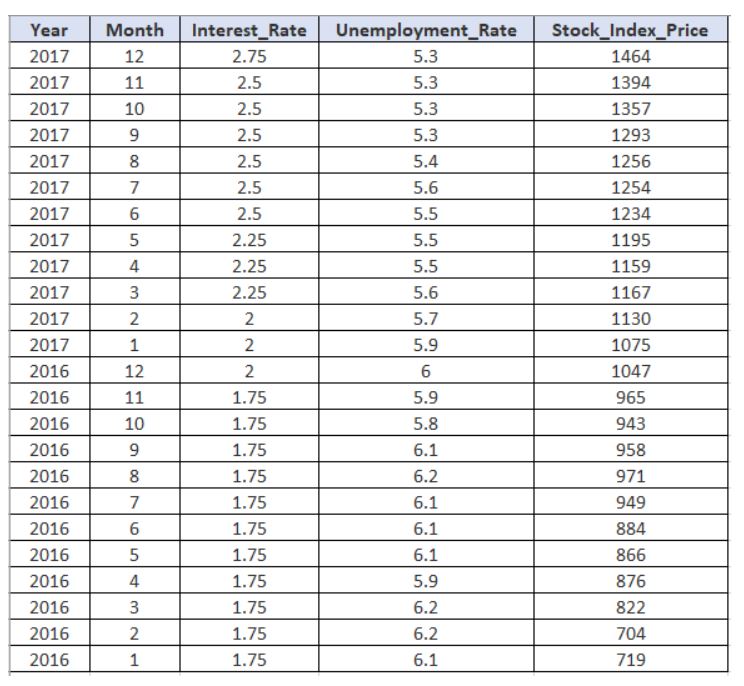
There are 20 Practical Questions of Machine Learning. Each carries 1 mark, if done and shown before Due Date. And rest 5 marks is of Class Record whose due date is 15-Apr-20.

**Due Dates**

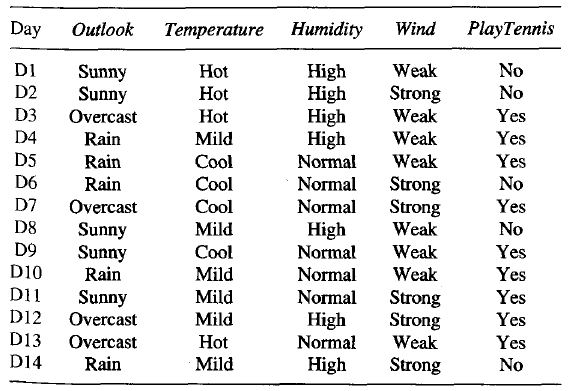
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Due Date | 12-Feb-20 | 14-Feb-20 | 19-Feb-20 | 21-Feb-20 | 26-Feb-20 | 28-Feb-20 | 4-Mar-20 |
| Q no | Q 1 to Q5 | Q 6 to 8 | Q 9 | Q 10 | Q 11 | Q 12 | Q13 |
|  |  |  |  |  |  |  |  |
| Due Date | 6-Mar-20 | 18-Mar-20 | 20-Mar-20 | 25-Mar-20 | 27-Mar-20 | 3-Apr-20 | 8-Apr-20 |
| Q no | Q14 | Q15 | Q16 | Q17 | Q18 | Q19 | Q20 |

**List of Practical Questions**

1. Write a function that returns the largest element in a vector of numbers.
2. Write a function that returns the sum of the even numbers from a vector of numbers
3. Write a function that searches a number from a vector of numbers.
4. Write a function that finds the factorial of a number.
5. Write a function that finds the mean and standard deviation of a vector of numbers.
6. Write a function to find whether the number is prime or not.
7. Write a function that returns the sum of the digits of a number.
8. Predict the stock\_index\_price based on Interest\_Rate and Unemployment\_Rate with help of formulatheta=(X’Y)-1X’X. (Matrix algebra). Use following data:



1. Using Naive Bayes and following data, predict the value of PlayTennis for an instance (Outlook=Rain, Temperature=Mild, Humidity=Normal, Wind=Weak)



1. The values of y and their corresponding values of y are shown in the table. Find the coefficients a and b for least square regression line y = ax + b. and estimate y when x = 10. Furthermore plot the scatter points and line.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| x | 0 | 1 | 2 | 3 | 4 |
| y | 3 | 4 | 5 | 4 | 6 |

1. Find the linear regression coefficients using gradient descent method for the above dataset with different learning rates (2,1,0.5,0.25,0.125) keeping initial values theta0=3 and theta1=1.
2. Let x=seq(from=-2,to=2,by=.1). Generate different subplots (2\*2 matrix) for a) y=x^3, b) y= -x^3, c) y=(2x-1)^3, d) y=2\*x^3-1. Keep same limits for x and y axis say -100 to 100 for both.
3. Scale the 2 variables of ‘cars’ dataset using mean normalization.
4. Use ‘cars’ data. It has 50 observations on 2 variables speed (mph) and dist (ft) representing speed of cars and the distances taken to stop respectively. Learn the parameters theta0 and theta1 using whole data to predict dist for the given speed. Plot the learnt model. Predict dist for 15.5 mph speed.
5. Use Boston housing data from MASS library as a dataset. Consider 70% of its data for training and rest for the testing. Predict '*median value of owner-occupied homes*' (i.e. *medv*) using *'lower status of the population*' (i.e. *lstat*) using linear regression. Generate two subplots to show: 1) Predicted values of *medv* against the original *medv* values of test dataset, 2) *medv* against *lstat*.
6. Use Boston housing data from MASS library as a dataset. Consider half of its data for training and rest for the testing. Predict 'median value of owner-occupied homes' (i.e. medv) using all the other attributes excluding 'age' attribute using multivariate linear regression and report the error percentage for the test dataset.
7. Use last 100 rows of Iris as a dataset (i.e. only versicolor and virginica but not setosa). Consider half of the data for training and rest for testing. Learn the model using logistic regression from lengths and widths of sepals and petals of training data. Use the learnt model to predict the class labels for the test dataset and report the accuracy and error.
8. Use last 100 rows of Iris as a dataset (i.e. Only versicolor and Virginia, but not sets). Consider half of the data for training and rest for testing. Learn the model using k-NN (k=5) from lengths and widths of sepals and petals of training data. Use the learnt model to predict the class labels for the test dataset and report the error.
9. Write the code for finding the weights of a perceptron using perceptron training rule for implementing OR gate. Consider all initial weights as 0 and alpha=1.
10. Use Boston housing data from MASS library as a dataset. Scale the data by any means before proceeding further. Consider half of its data for training and rest for testing. Learn the model using training data to predict *medv* from all the other attributes using neural network. Plot the learnt neural network. Predict the *medv* values for the test dataset.