Lab Assignments List

Course Instructor: P.S.R. Patnaik.

Course Name: Elective-4 Machine Learning Year: 2019-23

COURSE OUTCOMES:

Students will be able to:
1. Interpret the need and applications of machine learning.
2. Build machine learning model for a given problem.
3. Analyze machine learning model to improve their accuracy.

Assignment 1: Differentiate between Data Mining and Machine Learning. You shall differentiate on the below mentioned points precisely.

1) What are the objectives and outcomes of data mining process
2) What are the objectives and outcomes of ML process
3) Nature of processes in DM and ML (static/offline and dynamic/online)
4) An indicative example where DM processes are more suitable than ML processes and vice-versa.

Write your answer in the text box in not more than 700 words. Cite the internet resources if you refereed any.

Assignment 2: Write a python script to implement simple linear regression using MSE method as discussed in lectures. Justify your choice of deciding learning rate and initial values for parameters used in the linear regression model. Additionally, comment on effect of learning rate and initial values of parameters on overall performance of the algorithm with help of tabulated observations of algorithm runs.

Assignment 3: Write a python script to implement logistic classifier using cross-entropy loss function as discussed in lectures. Note and comment on number of parameters to be learned and effect of learning rate on number of epochs and classifier's performance using confusion matrix.

Assignment 4: For any your chosen dataset https://archive.ics.uci.edu/ml/datasets.php build a ML model using appropriate classifier algorithm and comment on values of precision, recall and F1-score. Justify your choice of using a specific classifier algorithm out at least 3 algorithms with comparative analysis based on confusion matrix.

Assignment 5: For the classifier algorithms you have used for assignment 4 present out a detailed analysis of your experiment to find best classifier algorithm based on following points. You shall use 70% of data for training and 30% for testing the ML model.

a) Model complexity
b) Tendency to overfit

c) Variance and Bias

Assignment 6: For the classifier algorithms you have used for assignment 4 present out a detailed analysis of your experiment to find best classifier algorithm based on following points.

a) Model complexity
b) Tendency to overfit
c) Variance and Bias

You shall repeat the experiment with following train test splits with and without k fold cross validation.

a) 70:30
b) 60:40
c) 50:50

Assignment 7: For the dataset of assignment 4, using Weka apply averaging and stacking techniques for decision tree classifiers algorithms. Note your observations and comment on improvements in classifier performance w.r.t to averaging and stacking techniques applied.

Assignment 8: For the dataset of assignment 4, using scikit-learn apply averaging and stacking techniques for decision tree classifiers algorithms. Note your observations and comment on improvements in classifier performance w.r.t to averaging and stacking techniques applied.

Assignment 9: For our college collect relevant data sets from department office and by following the Machine Learning Life Cycle build a machine learning model to predict the likelihood of student to bag an offer during placement drives. Justify decisions /choices you made in various processes and stages of ML life cycle.

Assignment 10: For the ML model you built in assignment 9, explain your strategy and analysis you did to avoid overfitting, underfitting and bias in your model. Additionally detail out the methods you adopted to improve the accuracy of your ML model. Highlight the use of ML tools you have used.

Assignment 11: For the experiments you performed in assignment 6 and 7/8. Detail out your approach to improve performance of the ML model w.r.t to the following

a) Using Cross-Validation Correctly
b) Searching for the Best Hyper-Parameters
c) Testing Multiple Models

Assignment 12: Suggest applications for machine learning in the Retail Business domain and stress the need for machine learning.

Assignment 13: Suggest applications for machine learning in the Healthcare domain and stress the need for machine learning.

Assignment 14: Suggest applications for machine learning in Manufacturing domain and stress the need for machine learning.

Assignment 15: Suggest applications for machine learning in Education & Academia domain and stress the need for machine learning.

Assignment #	COs targeted	Marks
1	1	10
2	2	25
3	2	25
4	2	20
5	3	20
6	3	25
7	3	10
8	3	10
9	2	30
10	2	20
11	3	15
12	1	10
13	1	10
14	1	10
15	1	10