Machine Learning

Programming Assignment

You may apply any learning algorithm (under supervised, unsupervised or reinforcement learning) to solve any real-world problem based on any real-world dataset. You may use a dataset of your own or use a publicly available dataset from a source like the following one (you can choose any other data source. Be sure to cite where you got your data from. The dataset has to be real-world data).

http://archive.ics.uci.edu/ml/index.php

https://www.kaggle.com/datasets

Based on a dataset, you may develop a machine learning algorithm (preferably using Python, but you can use any language you like) and then get the test results.

Submission

- The dataset, source code and the executable
- A **10-12** page pdf report with a description on the problem addressed, dataset used, methodology (algorithm used), results and discussion (including possible limitations/future work).
- The entire submission bundle has to be uploaded to the courseweb link (which will be available close to the deadline) as a single zip file. The name of the zip file should be your id (for example, IT15234566.zip).

The report should have an appendix that contains all the source code (added as text, not screen-shots). If the source code is not added as text in the appendix, it won't be accepted as a valid submission.

The report is the main deliverable that will be marked. However, the code should be submitted to validate what is mentioned in the report. **Both the report and the source code should be there to be accepted as a valid submission.**

All reports will be uploaded to Turnitin for plagiarism checking. If the turnitin similarity is above **30%**, **10%** of the marks will be reduced. For **50%** similarity, **50%** of the marks will be reduced. For reports with **80%** similarity, **no marks** will be given.

Deadline: End of 12th week of the lectures

Marks allocated: 30