

Domain Applications of Predictive Analytics

Project Brief

Lecturer: Vikas Sahni

Introduction

This is an individual project where students will work on their own datasets.

Marks will be assigned based on effective achievement with regards to the:

1. specification of the hypothesis
2. critical assessment of applicable techniques
3. implementation of the chosen technique
4. interpretation of the results

Note: The Presentation/Demonstration/Viva is mandatory. If you do not turn up for the oral examination, the project report will NOT be assessed, and you would be awarded ZERO marks.

Project Details

You are required to choose a dataset (or related datasets) in an area of interest suitable for making predictions. Please refer to the indicative list discussed in Lecture 1 for possible areas in which the project can be done.

Some Dataset sources

1. <http://www.data.gov/>
2. <http://data.fingal.ie/>
3. <http://www.insidedm.com/search/label/Download%20datasets>
4. <https://github.com/WinVector/zmPDSwR/tree/master/KDD2009>
5. Kaggle
6. UCI Machine Learning repository

Links to pages that have links to datasets

7. <https://www.datasciencecentral.com/profiles/blogs/big-data-sets-available-for-free>
8. <https://www.kdnuggets.com/datasets/index.html>

CA1 – Project Design

Once the choice has been made on the dataset, you should produce a project design document and submit this document. The project design document should be submitted as per the CA schedule. Please use the IEEE Conference template.

Typically, the project design document should provide the following pieces of information:

1. An introduction to the project dataset (Background and Scope).
2. The goal(s) of the project.
3. Ethical concerns
4. The strategy that will be employed to analyse the dataset.
5. Preliminary visualisations
6. Applicable techniques
7. References

A project log document should be maintained throughout all stages of the project. This log should be used to accurately record:

- all actions and tasks carried out
- when actions and tasks were carried out
- the time spent on the actions and tasks

The project design document should be 6-8 pages, IEEE Conference template, and fully document your work.

CA2 – Project Report and Presentation

Typically, the project report should provide the following pieces of information:

1. Research and investigation into applicable techniques
2. The technique(s) employed
3. The findings, in terms of both quantitative results and their qualitative interpretation
4. Other relevant features of the analysis
5. References

A project log document should be maintained throughout all stages of the project. This log should be used to accurately record:

- all actions and tasks carried out
- when actions and tasks were carried out
- the time spent on the actions and tasks

The project report should be 6-8 pages, IEEE Conference template, and fully document your work.

It will also be a requirement to prepare a presentation and demonstrate the technique at the end of the semester.

Note: The Presentation/Demonstration/Viva is mandatory. If you do not turn up for the oral examination, the project report will NOT be assessed, and you will be awarded ZERO marks.