Project (100%)

Predictive Analytics

Introduction

Using a dataset selected from the list of sites supplied on Moodle, or a dataset of your own chosing, you are required identify problems that can be solved or questions that can be answered by means of predictive analytics and then documenting it as a report.

While it is recommended that you use RapidMiner, you may use any other relevant tools to analyse a specific problem or issue with data that requires the use of some of the algorithms and techniques that were covered in the programme.

You should present your findings in a written report of between four and six pages, assuming a 12pt font is used.

Your report should capture the following aspects that are relevant to your project:

- Executive Summary (10%)
  - What problem did you address, what did you find, and (very briefly) how?
- Introduction (10%)
  - Provide some information on the dataset and the topic it pertains to.
  - What motivated you to investigate the problem or question - what are the challenges and/or implications?
- Methodology (40%)
  - What methods are you applying? E.g. are you comparing different techniques, combining techniques, etc.
  - Evaluation criteria/requirements – how do/will you know that your predictive analytics solution is good/valid?
  - Why do you think this proposed approach is suitable?
- Evaluation (40%)
  - Present the results of your models and say how they demonstrate that you are (partially) solving the problem at hand.
  - Discussion of results, their interpretation(s) and implications

Submission

You should submit your report via Moodle. You should also submit a zipped archive of your RapidMiner process (or code for any other tool you choose to use) via the ancillary files link.

Marking

Your submission will be marked according to the areas above and their respective weightings. Any submission marked at 40% and under 60% will be awarded a pass. Any submission marked at 60% and under 80% wil be awarded a merit. Submissions marked at 80% or over will be awarded a distinction.

Assistance

If you need assistance with any part of this, please ask. This includes identifying useful or interest analyses, choosing appropriate models, interpreting results, or creating a zipped archive of RapidMiner processes.