Student Guidelines

Assessment 1

Research Study & Presentation

Due: 22 December 2019 - 11:59 pm
Total Weightage: 20%
Individual assignment

Python is one of the most frequently used programming languages in many fields, particularly in data science. It is also one of the best data science tools for the big data job.

Task

The assignment has two phases: 1) writing a report and 2) presentation of findings using Python codes.

1. Report (Weightage: 10%)

Choose data:

Choose a data from Kaggle website, https://www.kaggle.com/datasets , or a government open source data. You can also use Twitter data, which you can download using Python Tweepy package.

Analytics:

Find out what you can do with that data or what kind of decision making you can do with it. First **(Step 1)**, do an exploratory data analysis on the data that you have gathered. Exploratory data analysis is an approach for analysing data sets to summarize their main characteristics, often with visual methods. Then **(Step 2)**, Build a machine learning model on top of your data and make necessary recommendations.

Python implementation:

To be consistent with all students, implementation must be done in google Colab:

https://colab.research.google.com/notebooks/welcome.ipynb

Colab is a free notebook environment that requires no setup and runs entirely in the cloud. You need to login to google Colab and write your Python code for analysing the data. Add your google Colab account showing your name on it into your report, by clicking orange button on top-right corner and taking screenshot.

Your report should have 1500-2000 words addressing the following: information on the data and why it is important, literature review on the data and methodology you are going to work, what you are going to solve and how, plots and recommendations. The report should have at least 4-6 plots (screenshots) from your findings with explanations.

2. Presentation (Weightage: 10%)

The presentation should be a maximum of 10 minutes. It must cover the research report, research findings and visualisation and step by step discussion on how you've done this project.

Submission Guidelines

All submissions are to be submitted through turn-it-in. Drop-boxes linked to turn-it-in will be set up in the Unit of Study Moodle account. Assignments not submitted through these drop-boxes will not be considered.

Submissions must be made by the due date and time (which will be in the session detailed above) and determined by your Unit coordinator. Submissions made after the due date and time will be penalized at the rate of 10% per day (including weekend days).

The turn-it-in similarity score will be used in determining the level if any of plagiarism. Turn-it-in will check conference web-sites, Journal articles, the Web and your own class member submissions for plagiarism. You can see your turn-it-in similarity score when you submit your assignment to the appropriate drop-box. If this is a concern you will have a chance to change your assignment and re-submit. However, re-submission is only allowed prior to the submission due date and time. After the due date and time have elapsed you cannot make re-submissions and you will have to live with the similarity score as there will be no chance for changing. Thus, plan early and submit early to take advantage of this feature. You can make multiple submissions, but please remember we only see the last submission, and the date and time you submitted will be taken from that submission.

Your report should be a single word or pdf document containing your report.

Your presentation file should have a standard video format and it should not exceed 200 MB. Slides and your face should be clear in the video file. You need to submit the presentation file (not link to your video) in the provided video submission link. Please do not submit the link for your video, which will not be considered for marking.