SIT307 Data Mining and Machine Learning



Week 2 Practical: Data and Visualisations

Overview

This week's practical will introduce students to loading data and creating visualisations using both numpy and pandas.

Practical Learning Outcomes

After completing this practical students will be able to:

PLO1 – load a range of data types using both numpy and pandas.

PLO2 – create a range of visualisations using matplotlib with numpy and pandas.

Class Activity

You will work on this activity either individually, in a small group or with the whole class as directed. This part of the practical does not need to be submitted but will provide the formative learning experience that will help you to do the Individual Problem solving task, which is required to be submitted.

Task 1: Data and Visualisations

In this task you will load a range of data file types and create visualisations using matplotlib and numpy. Secondly, students will do a similar set of tasks except using pandas for loading and storing the data.

Students should open first open "0. Data and Visualisations with numpy.ipynb" in Jupyter and work through and run each cell. Ensure you understand what each code cell is doing and how it works. Try alternatives to ensure you understand numpy and matplotlib.

Secondly open "1. Data and Visualisations with pandas.ipynb" in Jupyter and work through and run each cell. Ensure you understand what each code cell is doing and how it works. Try alternatives to ensure you understand numpy and matplotlib.

Individual Problem Solving Task

Your individually worked solution to this section should be submitted as part of your Assignment 1 Problem Solving Task Part A (Weeks 1 -3).

Task 2: Problem Solving in Python

You task this week is to create a new notebook called "[yourstudentID]_Week_2_Problem Solving Tasks.ipynb". In this you need to create cells to load the iris.txt data and generate three graphs. It is your choice which graphs you create but you should choose graphs that help provide an incite into the data you have loaded. You should also include a cell with a paragraph explaining what your exploratory analysis has told you about the data.