

Oxford Study Abroad:

Artificial Intelligence / Machine Learning Aug 2022

Situation

You are working for an astrophysics research laboratory. An advanced space programme has sent several un-manned missions to the asteroid-belt to collect rock samples from asteroids. This is the first time that such a complex mission has been achieved successfully. For the first time, samples of asteroids have been returned to Earth for scientific examination!

Geologists and Chemists have been working on the samples of rock and have tabulated their results in two data files:

1. 'China Aug 2022 Asteroid 1.csv'

and

2. 'China Aug 2022 Asteroid 2 V4.csv'

Data file (1) and (2) contain measurements from completely different groups of asteroids. No individual asteroid appears in both data files.

You are asked to complete the following two analysis tasks:

Task 1

Data file (1) ('China Aug 2022 Asteroid 1.csv') contains the results of the analysis of 4966 asteroids. Each asteroid has been given a unique identifier (ID) in the first column of data in the form "Ast_XXXX". Where the 'XXXX' is simply a unique number. There are 18 features in the data set. Each feature represents a quantity, measured in micrograms of various types of minerals found in the asteroid.

Your task is this:

- Work out if there are distinct types of asteroid, and if so, how many are there?
- If there are distinct types of asteroid, for each type of asteroid provide:
 - The typical (average) mixture of minerals for each type of asteroid, and
 - The variation in mineral mixture between asteroid types
- Provide more confidence in your results by applying a variety of different analytical methods and comparing the results

Task 2

Data file (2) ('China Aug 2022 Asteroid 2 V4.csv') contains the results of a series of experiments. The experiment is intended to understand if a particular mixture of minerals gives an asteroid that is mechanically stable. This data file contains 9 features. Each feature again represents a quantity in

micrograms of a particular mineral. The feature (column) name is given in the first row. The final column in the table provides a label, with a meaning as follows:

- 0 = “The asteroid is mechanically stable”
- 1 = “The asteroid is **not** mechanically stable”

Your task is this:

- Find a method of predicting if an asteroid is mechanically stable based on its mixture of minerals;
- Try to find multiple, different techniques to do this;
- For each technique that you use, provide a measure of its quality in terms of making good predictions
- For each technique that you use, attempt to ‘tune’ (improve) your model to give the best results possible

Presenting your Assignment

Submit your assignment in the form of a Jupyter notebook (.ipynb) file. Run every block in your file before submission so that the executable blocks show the results of the analysis as run on your computer.

Include markup (text) blocks in your notebook that describe the work you have done and the results you have obtained.

If you use any non-standard libraries or functions in your analysis, then describe those in markup blocks, including where they can be found and how they can be installed (normally a reference to a website that provides the library)