Assignment Description

Design and Development of a Data Analytics Service

- This is a group assignment. The max size of a group is 5 (no fewer than 4, no more than 5)
- You need to have two meetings with your mentor during Week 11 and 12
- Demo Dates: To be organised with mentors during Week 13
- Late penalty: 20 percent of the final mark per day

Summary

In this assignment, you are asked to develop a Data Analytics Service . Besides a few requirements that you must meet, the specification of the service you'd build is **deliberately left open** . You are expected to develop a plan and execute it with your group members all throughout this assignment period.

Rational

The idea here is to develop a REST API that provide analytical services for the consumer in a certain field. The scenario is that you need to obtain data from one or more data sources, pre-process the data so it is usable to train a machine learning model, and accordingly provide to the consumer the ability to do prediction in certain context.

Some examples of analytical service :

- Clarfa i which is an Image recognition service (https://clarifai.com/models)
- Sentiment Analysis for Movie reviews (https://towardsdatascience.com/deploying-a-machine...)
- Proprty Price prediction using Machine learning (https://yalantis.com/blog/predictive-algorithm-for...)

Available Data Sources

Explore and investigate the following data sources. You should choose a few data sets that could give you an interesting scenario for your service . You can choose from the following sources or any real-world dataset (i.e. synthetic or datasets used for learning exercises are not allowed) .

- Kaggle Datasets
- Explore NSW Government Public Datasets
- Explore Australian Federal Government Public Datasets
- Explore United Nation Pubic Datasets
- Explore Worldbank's Datasets

- CIA (USA) The World Factbook

You can also mix the data sources with some other existing APIs if that could make your service more interesting! For example, you can choose from the following:

- http://openweathermap.org
- Map APIs (e.g., Google Maps)
- Visualisation APIs (e.g., Google Charts, Charts.js)
- Photo APIs (e.g., Pixabay API)
- Any other API that you think you'd be useful for the assignment (only after a specific approval from your mentor).

One note: I'd NOT use F acebook or twitter APIs (or similar social-network type APIs) as a source to get data as I find their APIs very limited in terms of extracting interesting information that is suitable for the assignment. However, you could use them to post messages (e.g., twitting a new event).

In Summary you Need to:

1. Com e up with a scenario
2. Select the Data sources that would help fulfil your scenario
3. Perform Data Integration and pre-processing if needed .
4. Building a machine learning model to fulfil the scenario (using the dataset that you have prepared for training and evaluation)
5. Designing a R EST ful API to allow the consumption of your service (you need to consider an authentication scheme for the consumers of your service)
6. Designing a Simple Client with GUI (you are free to use whatever you like whether it is a simple HTML- java-script , ASP, php, JSP, JSF, or even window-based interfaces)

Group Work

Must establish a shared repository and show evidence of work built over the assignment period (not just the last 2 days!).
Must establish a role and tasks in each group. The actual work of the member in that role/tasks must be linked with the evidence of work e.g., email/messaging, repository commits , etc.).

Meetings with Mentor

You can arrange meetings with the tutors from Week 10. Before the final demo, you are required to have two meetings. Each meeting will be assessed. During the meetings, your mentor (tutor) will give feedback on your work in progress. So utilise the time as much as you can.

- Meeting One - more or so complete design documentation (10% of the total mark).
- Meeting Two - an early implementation of the service , demo of work in progress (10% of the total mark).

Marking Scheme

The following shows how the marks will be distributed. Detailed marking schemes will be released later.

- Meeting One - 10%
- Meeting Two - 10%

- Week 13 - 70% (demo), 10% (group work)