COMP9322 Assignment 2

Case study description

Overview of system

The system to be developed should allow the prediction of house prices using external data and some pre-existing software. The functions of the systems should be:

- Access to a Web site to download data containing property prices.
- The application of one or more programs to the data to make some kind of prediction. These programs already exist but currently work in standalone mode and need to be launched manually.
- The presentation of the results in some form that is understandable by the customer. For example the results could be shown as a table embedded in a PDF document, or a graph.

Normally, a user needs to go to the web site, download the data, run the programs in some sequence and display the results but this has to be done several times with different parameters by different users (with typically low-level IT skills) so it is labour-intensive and error-prone.

Example

The external data source in this case is Domain.com.au and we assume we have 3 data processing programs:

- "Regression": this program takes a dataset with clean data (e.g. type "house prices cleaned") and returns results (e.g. dataset type "house price predictions").
- "Neural Network": this program takes the same inputs as the previous program but works differently. The format of the output is the same (but the results are different).
- "Preprocessing": converts data from Domain.com.au (e.g. from type "house prices domain" into dataset type "house prices cleaned").

When the system becomes operational, any user is able to:

- Import datasets: accessing Domain.com.au, supplying some user-defined parameters like the time period of interest, downloading data and creating a new dataset that can be referenced by its name (the dataset name is autocreated by the system).
- Execute models: the user is given two options of running a sequence of data processing programs on any imported dataset. In the first option, the programs "Preprocessing" followed by "Regression" are executed. In the second one, "Preprocessing" followed by "Neural Network" are executed. Executing a sequence creates a new dataset with an automatically generated name and of type "house prices cleaned".
- Downloading or visualising datasets: the user is given 2 options. Choosing amongst two programs: one that converts a dataset of type "house prices predictions" into a PDF file or one that converts it into a graph.

Deliverable

This is a group assignment. You are to use the model assigned for your group in the Jalapeno platform (shared via email) and create models to capture **only the services and information models** of the system described in the case study.

You will be evaluated based on the models you create in the Jalapeno platform under the model assigned to your group. To help the evaluators to understand and navigate your models, you are required to submit a report describing the main components of the case study you have modelled and how those components are inter-related. Name of the report should be <teamName>_Assignment2.pdf . (ONE .PDF file ONLY PLEASE)

Please include the cover sheet (word document available on WebCMS) and attach it to your assignment.

Deadline: Monday 7th May 2018
URL for report submission: https://www.dropbox.com/request/gDcPVMEEmJDvssOaYL2g

ONLY ONE SUBMISSION PER GROUP IS REQUIRED.

Note: Exercises on the lab 4 helped you to get a basic idea about the Jalapeno platform. Moreover, we have provided you with a simple solution for assignment 2 as a guideline which you can use as a starting point and extend further.