

Small Data Analyst

2. MSc Project Progress Report

Author: Sebastian Zillesen^{*}
Supervisor: Jeroen Keppens[†]
Product Owner: Isabel Sassoon[‡]

July 24, 2016

1 Introduction

This second project report summarises the progress made so far in the MSc Project "Small Data Analyst". The project is supervised by Jeroen Keppens and the requirements for the software have been closely evaluated and defined together with Isabel Sassoon, author of the underlying paper [2]. This project report builds up on the first progress report which has been handed in on June 24th.

2 Use Case Implementation

At the current situation, all initially required Use Cases have been implemented and RC4 is right now in the feedback loop. The only Use Case not implemented (which was as well marked as a *would* in the initial *MuSCoW* prioritisation scheme) relates to in-depth analysis about the algorithms performance, as this is not really relevant, as the argumentation frameworks in place are rather small, the only "big data" we have to deal with are the datasets, but this is well handled in R.

3 RC 2 & Demo

RC 2's has been successfully implemented and presented to the supervisors. The overall impression was good, the project met its requirements until that point. Major features implemented were Login/Authentication, Adding of Assumptions, R-Code execution, Initialisation of possible models after AS1.

^{*}sebastian.zillesen@kcl.ac.uk

[†]jeroen.keppens@kcl.ac.uk

[‡]isabel.sassoon@kcl.ac.uk

4 RC 3 & Demo

RC 3 focused on the implementation of the preferences between different models if certain assumptions hold [UC 10, 9-1, 19].

In addition the user management has been extended [UC 11] and adding new research question [UC 17] has been implemented.

The feedback on RC 3 was good, however we decided to make the implementation of the preferences even more intuitive by moving it from a hard-coded approach into a UI/database approach. This was then done in RC 4.

5 RC 4 & Demo

RC 4 focused on the implementation of UI features. First of all, the preferences over models have been implemented as UI solution, so a user is now able to enter preferences on the website. This involves assumption assignment and specifying the order of models in certain scenarios by drag & drop.

In addition a plotting functionality for the generated Argumentation Frameworks has been implemented [UC 15]. [1] was used as a source for the algorithm to check if an argument is acceptable w.r.t a subset of holding arguments.

Compared to the initial idea, we moved from using the argumentation based approach Sassoon proposed in [2], to the more recent approach involving different context domains and preferences expressed by relative orders in these context domains if they are applicable as Sassoon proposed in [3]. Personal preferences entered by a clinician [UC 14] have been realised by using `cancancan`¹ and personalised abili-

¹<https://github.com/CanCanCommunity/cancancan>

ties.

In addition to these planned features, the requirement for presenting plots to the user to ask him if this plot satisfies some condition was raised (this is e.g. required in the Weibul model [3]). To satisfy this need, an additional specialisation of the `QueryAssumption` has been implemented, which takes `R-Code` to generate a plot out of the dataset defined for the analysis. This plot is then presented to the user and she/he has the option to answer with yes/no. The influence of these `QueryTestAssumptions` are the same as `QueryAssumptions` the only difference is, that these assumptions contain a dynamic plot related to the dataset written in R [UC 20].

Furthermore some UI improvements have been applied. The final version was presented as playable demo to the supervisors at July 25th.

6 Project Plan Updates

There have been no significant changes to the Project Plan, as the project is running on track. In fact the completion of RC 4 was done earlier, and the additional time was used to implement another feature, called the `TestQueryAssumptions` (see above).

The coding of the web application is hereby finished (except for minor changes that might arise during the final feedback loop starting this week) and the remaining time is intended to be used for the dissertation itself.

References

- [1] DUNNE, P. E., MODGIL, S., AND BENCHCAPON, T. Computation in extended argumentation frameworks. In *in: Proc. 19th ECAI, Univ. of* (2010), Press, pp. 119–124.
- [2] SASSOON, I., KEPPENS, J., AND MCBURNEY, P. Towards argumentation for statistical model selection. In *COMMA* (2014), pp. 67–74.
- [3] SASSOON, I., KEPPENS, J., AND MCBURNEY, P. Preferences in argumentation for statistical model selection.