

Data Analytics – Assessment

Project developer:

Jakub Majcher

Rafał Skrzypek

Project reviewer:

Jakub Grzesiak

Rafał Kośła

Dorota Kowalczyk

Problem formulation [0-5 pts]:

is the problem clearly stated [1 pt]	1 pt
what is the point of creating model, are potential use cases defined [1 pt]	1 pt
where do data comes from, what does it contain [1 pt]	1 pt
DAG has been drawn [1 pt]	1 pt
confoundings (pipe, fork, collider) were described [1 pt]	1 pt

Data preprocessing [0-2 pts]:

is preprocessing step clearly described [1 pt]	1 pt
reasoning and types of actions taken on the dataset have been described [1 pt]	0.5 pt

Model [0-4 pts]

are two different models specified [1 pt]	1 pt
are difference between two models explained [1 pt]	1 pt
is the difference in the models justified (e.g. does adding additional parameter makes sense?) [1 pt]	0 pt
are models sufficiently described (what are formulas, what are parameters, what data are required) [1 pt]	1 pt

Priors [0-4 pts]

Is it explained why particular priors for parameters were selected [1 pt]	1 pt
Have prior predictive checks been done for parameters (are parameters simulated from priors make sense) [1 pt]	1 pt
Have prior predictive checks been done for measurements (are measurements simulated from priors make sense) [1 pt]	1 pt
How prior parameters were selected [1 pt]	1 pt

Posterior analysis (model 1) [0-4 pts]

were there any issues with the sampling? if there were what kind of ideas for mitigation were used [1 pt]	1 pt
are the samples from posterior predictive distribution analyzed [1 pt]	1 pt
are the data consistent with posterior predictive samples and is it sufficiently commented (if they are not then is the justification provided)	1 pt

have parameter marginal distributions been analyzed (histograms of individual parameters plus summaries, are they diffuse or concentrated, what can we say about values) [1 pt]	1 pt
---	------

Posterior analysis (model 2) [0-4 pts]

were there any issues with the sampling? if there were what kind of ideas for mitigation were used [1 pt]	1 pt
are the samples from posterior predictive distribution analyzed [1 pt]	1 pt
are the data consistent with posterior predictive samples and is it sufficiently commented (if they are not then is the justification provided)	1 pt
have parameter marginal distributions been analyzed (histograms of individual parameters plus summaries, are they diffuse or concentrated, what can we say about values) [1 pt]	0 pt

Model comparison [0-4 pts]

Have models been compared using information criteria [1 pt]	1 pt
Have result for WAIC been discussed (is there a clear winner, or is there an overlap, were there any warnings) [1 pt]	0 pt
Have result for PSIS-LOO been discussed (is there a clear winner, or is there an overlap, were there any warnings) [1 pt]	1 pt
Was the model comparison discussed? Do authors agree with information criteria? Why in your opinion one model better than another [1 pt]	1 pt

Summary: 23.5pkt/27pkt