

Peer Review

Reviewers:

Daniel Jurkowski 407200

Adam Pękala 405380

Authors:

“Diabetes Prediction based on medical records”

Konrad Prokop

Kaia Rupniak

Section	Criterion	Points	Description
Problem formulation	is the problem clearly stated	1/1 points	The problem is clearly stated in an informative way.
	what is the point of creating a model, are potential use cases defined?	1/1 points	Potential use cases are defined.
	where do data come from, and what does it contain	1/1 points	Data comes from The Pima Indians Diabetes Database and is fully described.
	DAG has been drawn	1/1 points	DAG has been drawn properly.
	confoundings (pipe, fork, collider) were described	1/1 points	Confoundings are specified.
Section sum:	5/5 points		

Section	Criterion	Points	Description
Data processing	is the preprocessing step clearly described	1/1 points	Preprocessing is clearly described. Additionally, a visualization is made.
	reasoning and types	1/1 points	Actions and

	of actions taken on the dataset have been described		reasonings are described.
Section sum:	2/2 points		

Section	Criterion	Points	Description
Model	are two different models specified	1/1 points	Two models are specified.
	are difference between two models explained	1/1 points	The difference is explained.
	is the difference in the models justified (e.g. does adding additional parameter makes sense?)	1/1 points	The authors have added additional predictors, which make sense in this case.
	are models sufficiently described (what are formulas, what are parameters, what data are required)	0.75/1 points	Models are adequately described, and stan code is provided. However, it would be nice if formulas were displayed in a “mathematical” way.
Section sum:	3.75/4 points		

Section	Criterion	Points	Description
Priors	Is it explained why particular priors for parameters were selected	1/1 points	Yes, the authors provided their reason for selecting priors.
	Have prior predictive checks been done for parameters (are parameters simulated from priors make sense)	1/1 points	Parameters simulated from priors make sense, authors compared it to real observed data.

	Have prior predictive checks been done for measurements (are measurements simulated from priors make sense)	1/1 points	The measurements simulated from the priors were compared with the actual data to ensure they align well.
	How prior parameters were selected	1/1 points	The prior parameters were selected based on the analysis of the dataset and the visualizations that highlight the significant features influencing diabetes.
Section sum:	4/4 points		

Section	Criterion	Points	Description
Posterior analysis (model 1)	were there any issues with the sampling? if there were what kind of ideas for mitigation were used	1/1 points	There were no significant issues observed during the sampling process. T
	are the samples from posterior predictive distribution analyzed	1/1 points	The results were compared with the actual outcomes to evaluate the model's performance.
	are the data consistent with posterior predictive samples and is it sufficiently commented (if they are not then is the justification provided)	1/1 points	The consistency of the data with posterior predictive samples was checked
	have parameter marginal distributions been analyzed	1/1 points	The marginal distributions of the parameters were

	(histograms of individual parameters plus summaries, are they diffuse or concentrated, what can we say about values)		analyzed.
Section sum:	4/4 points		

Section	Criterion	Points	Description
Posterior analysis (model 2)	were there any issues with the sampling? if there were what kind of ideas for mitigation were used	1/1 points	There were no significant issues observed during the sampling process.
	are the samples from posterior predictive distribution analyzed	1/1 points	The results were compared with the actual outcomes to evaluate the model's performance.
	are the data consistent with posterior predictive samples and is it sufficiently commented (if they are not then is the justification provided)	1/1 points	The consistency of the data with posterior predictive samples was checked.
	have parameter marginal distributions been analyzed (histograms of individual parameters plus summaries, are they diffuse or concentrated, what can we say about values)	1/1 points	The marginal distributions of the parameters were analyzed.
Section sum:	4/4 points		

Section	Criterion	Points	Description
Model comparison	Have models been compared using information criteria	1/1 points	The models were compared using the WAIC and LOO criterion.
	Have result for WAIC been discussed (is there a clear winner, or is there an overlap, were there any warnings)	1/1 points	The WAIC values have been discussed. No warnings are present.
	Have result for PSIS-LOO been discussed (is there a clear winner, or is there an overlap, were there any warnings)	1/1 points	The PSIS-LOO values have been discussed. No warnings are present.
	Was the model comparison discussed? Do authors agree with information criteria? Why in your opinion one model better than another	1/1 points	The model comparison was discussed.
Section sum:	4/4 points		

Total points: 26.75/27

Total percentage: 99%