Review ISZ_19

reviewers

Bartosz Sroka	400490	Points:	27/27
Norbert Podgórski	402111	Percent:	100%

The presented project fully deserves the highest possible rating. The project has been executed with great attention to detail, met all evaluation criteria, and furthermore, it can be effectively used in real life.

Problem formulation [5 | 5 pts]:

- is the problem clearly stated **[1 pt]**

The problem has been clearly defined. For someone who has never played League of Legends (LoL), the mechanisms and rules used in creating the model have been presented in a clear and understandable manner.

- what is the point of creating model, are potential use cases defined **[1 pt]**

Potential use cases has been defined - the presented solution can indeed have a real impact, such as in sports betting during tournaments.

- where do data comes from, what does it containt **[1 pt]**

The data comes from Kaggle, and the attached link provides information about the units of the variables, however, it was not described in the document. Considering further data operations, this information is not crucial.

- DAG has been drawn **[1 pt]**

It has been drawn accurately, and it includes processed data used in the project.

- confoundings (pipe, fork, collider) were described **[1 pt]**

The dependencies between the data (Pipes, Forks, Colliders) have been described in a detailed manner.

Data preprocessing [2 | 2 pts]:

- is preprocessing step clearly described **[1 pt]**

Yes, the actions taken are described in a clear manner.

- reasoning and types of actions taken on the dataset have been described **[1 pt]**

Yes, the reason for applying differential data and their standardization is provided.

Model [4 | 4 pts]

- are two different models specified **[1 pt]**

Two different models using distinct datasets have been described.

- are difference between two models explained **[1 pt]**

The differences have been explained - the second model is an extension of the first model with additional variables. Justification for the choice of these variables has been provided.

- is the difference in the models justified (e.g. does adding aditional parameter makes sense?) **[1 pt]**

Yes, the models are consistent with each other, and the second one is a rational extension of the first one.

- are models sufficiently described (what are formulas, what are parameters, what data are required) **[1 pt]**

Yes, the used data, parameters, and distributions have been described.

Priors [4 | 4 pts]

- Is it explained why particular priors for parameters were selected **[1 pt]**

Yes, it has been explained in detail.

- Have prior predictive checks been done for parameters (are parameters simulated from priors make sense) **[1 pt]**

Yes, parameters simulated from priors make sense.

- Have prior predictive checks been done for measurements (are measurements simulated from priors make sense) **[1 pt]**

Yes, measurements simulated from priors also make sense.

- How prior parameters were selected **[1 pt]**

Yes, the rationale for using specific distributions for the parameters has been provided.

Posterior analysis (model 1) [4 | 4 pts]

- were there any issues with the sampling? if there were what kind of ideas for mitigation were used **[1 pt]**

There have been issues with sampling. A solution utilizing data standardization has been presented.

- are the samples from posterior predictive distribution analyzed **[1 pt]**

Yes, they have been thoroughly analyzed.

- are the data consistent with posterior predictive samples and is it sufficiently commented (if they are not then is the justification provided) **[1 pt]**

Yes, data are consistent with posterior predictive samples and has been sufficiently analyzed.

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

Yes, they have been thoroughly analyzed.

Posterior analysis (model 2) [4 | 4 pts]

- were there any issues with the sampling? if there were what kind of ideas for mitigation were used **[1 pt]**

Due to the similarity of the model, the same issues arose and were resolved using the same approach.

- are the samples from posterior predictive distribution analyzed **[1 pt]**

The same as in the first model.

- are the data consistent with posterior predictive samples and is it sufficiently commented (if they are not then is the justification provided) **[1 pt]**

The same as in the first model.

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

The same as in the first model.

Model comparison [4 | 4 pts]

- Have models been compared using information criteria **[1 pt]**

Yes, information criteria were used to evaluate the models.

- Have result for WAIC been discussed (is there a clear winner, or is there an overlap, were there any warnings) **[1 pt]**

The WAIC criterion has been analyzed, and a superior model has been selected.

- Have result for PSIS-LOO been discussed (is there a clear winner, or is there an overlap, were there any warnings) **[1 pt]**

The PSIS-LOO criterion has been analyzed, and a superior model has been selected.

- Whas the model comparison discussed? Do authors agree with information criteria? Why in your opinion one model better than another **[1 pt]**

The difference in the metrics is small, the models operate at a similar level of accuracy, which is why the authors decided to choose a simpler model that is easier to interpret.