Form 2b – Research Master's Psychology: Peer Review Form
Research Internship Proposal

Title of research project: Bayesian Symbolic Regression

Author name: David Coba

Peer reviewer name: Maximilian Maier

Date: 25.03.2021

Write a *one page* review about the project-proposal of your fellow student. At least include the following points in your peer-review:

(1) Summary of project
This needs to be only 1-3 sentences, but it demonstrates that you understand the project-proposal and, moreover, can summarize it more concisely than the author in his abstract.

The author aims to develop a Bayesian symbolic regression, which will allow researchers to find mathematical expressions that describe the data. Unlike other machine learning techniques, this symbolic regression is highly interpretable. However, it can not yet take advantage of the benefits of Bayesian statistics (e.g., incorporating prior information). Therefore, the author's proposal seems highly relevant.

(2) Good things about the proposal (one paragraph)
This is not always necessary, especially when the review is generally favorable. However, it is strongly recommended if the review is critical. Such introductions are good psychology if you think the author needs to drastically revise the proposal.

The purpose of the proposal is clear and seems important. In addition, the proposal is very well written and, in most parts, easy to understand.

(3) Major comments
Discuss the author's assumptions, technical approach, procedure, reference, etc. Be constructive, if possible, by suggesting improvements.

No major comments.

(4) Minor comments
This section contains comments on style, grammar, etc. If any of these are especially poor and detract from the overall presentation, then they might escalate to the 'major comments' section. It is acceptable to write these comments in list (or bullet) form.

- The concept of symbolic trees requires a more detailed explanation. Are these just classification trees, or other kinds of trees?
- I am not familiar with the amount of literature on symbolic regression; therefore, this advice should be taken with skepticism, but I believe the proposal could benefit from incorporating somewhat more literature. For example, the statement "Bayesian methods are not necessarily more demanding than their evolutionary machine learning counterparts" is not supported by any literature.
- I think it would be good to have one expression set to test and try out modifications of the algorithm (like Expression 1) and another applied example to use once the algorithm is finished to avoid overfitting to Expression 1. For example, the Statistical Rethinking book has several nice mathematical models in chapter 16 that could maybe be used for this purpose.

(5) Recommendations
Provide the author with some useful recommendations that you seem fit.

I would add a more detailed explanation of symbolic trees, take another look at the literature to be sure nothing has been missed, and add a mathematical model from Statistical Rethinking chapter 16. All other aspects of the proposal are already of very high quality.