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

Title of research project:	Bayesian Symbolic Regression
Author name:	David Coba Castellano
Peer reviewer name:	Leonhard Volz
Date:	24th March 2021

Summary of project
The internship project concerns a Bayesian implementation of symbolic regression. For this purpose, the plan is to first implement and then adapt a previously proposed algorithm and benchmark these with other existing approaches to symbolic regression. If successful, the further aim is to implement this approach as easily usable software.

The proposal is rather concisely written, but get across the idea of symbolic regression, current and proposed approaches and their relative strengths pretty well. Moreover, the subsequent steps to be taken are presented quite clearly.

Major comments
- if I am not mistaken, the proposal should include a project summary (i.e., abstract)
- Rather remove the square brackets and phrase it within the text. Especially at the end of the project description, you could add a paragraph that expands on why this general way to specify Bayesian symbolic regression is relevant / sketch out your aims
- Generally, since you refer to very few papers as you underpinning, you could expand on what their approaches in more detail (e.g., describe the Newtonian Dynamics system if you plan to go through with that step)

Procedure
- Standard: evolutionary algorithm: the footnote implies you made the decision already, the main text however not – try to clarify
- Do you have any considerations regarding assessing overfitting of your modifications? Seems like a general issue to be aware of in symbolic regression
- I wonder about the computational feasibility – possibly explain the environment in which simulations will be conducted

Intended results:
- What you currently describe is more of a rationale that would suit the project description section more than here. I would suggest making a summary of what your tangible outcomes would be (e.g., working algorithms, performance results comparisons, possible package)

Minor comments
Project description:
- ad advantage of SR: compared to specifically fully data-driven machine learning techniques, right?

- in which sense do neural networks encode “prior information”? A bit confusing when you then go on with Bayesian alternatives later, try to make that a bit more clear

Procedure:

- What software do you plan to work with? Julia?

Work plan:

- seems rather optimistic to me, given the necessary steps within each point of the plan. Maybe expand on the individual steps necessary, might give yourself a better overview of the project as well

Recommendations

So far, the proposal is rather short, so it would be good to extend it with a bit more detail on the rationale and the background of the different approaches. Similarly, the procedure is rather vague (but that is more of an outcome given the exploratory nature of such a project). Specific recommendations or points that could be expanded upon are mentioned above.