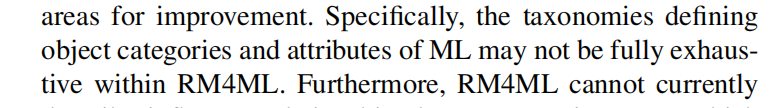
We submitted the manuscript "RM4ML: A Requirements Model for Machine Learning-enabled Software Systems". We would like to thank the referees for the careful and constructive reviews. Based on the comments of the previous referees, we have made changes to the manuscript, the details are shown as follows:

Comment1 of Reviewer 1:

Assumptions are made to define an UML extension: Are the identified "categories/objects/attributes" really extensive enough

Author reply:

Thank you for the reminder. We have made revisions to add interfaces in the meta-model that can continue to be extended and we have mentioned this possible weakness in the limitation section. As shown in the figure below.

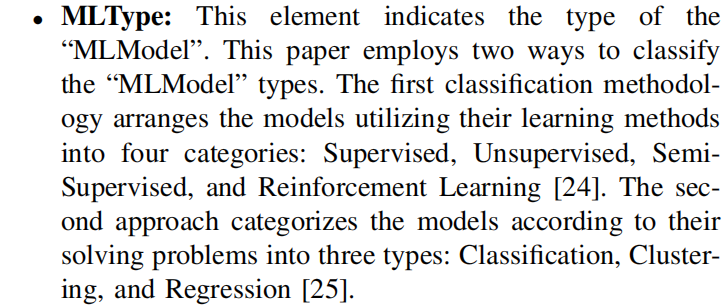


Comment2 of Reviewer 1:

Assumptions regarding the categorisation in the meta model are made. Is there a reliable source to proof this assumption?

Author reply:

Thank you for the good advice. We have made revisions to adopt commonly accepted classification criteria for classifying machine learning models and have cited two papers to illustrate the source of the classification. As shown in the figure below.

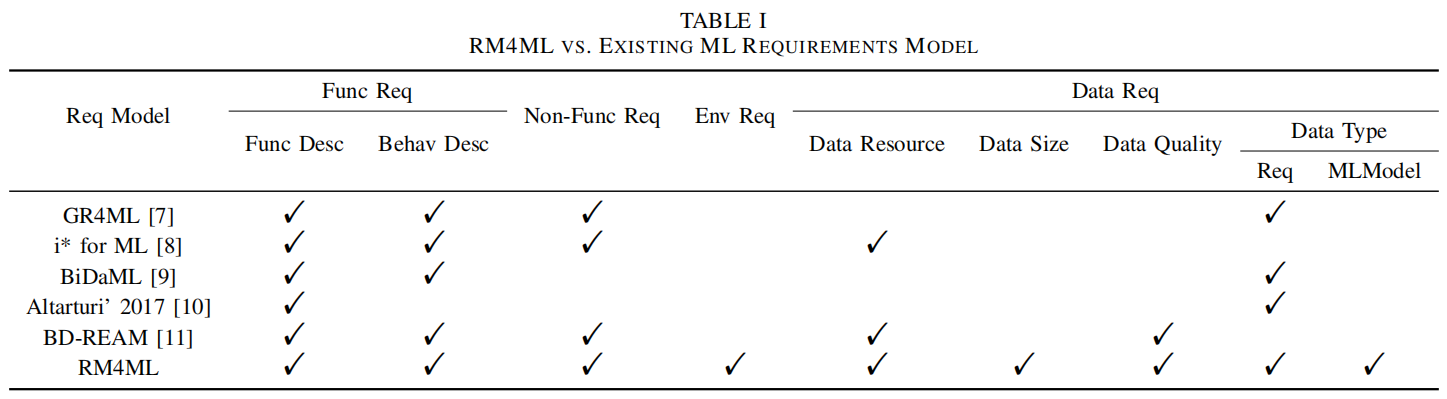


Comment3 of Reviewer 1:

more thorough evaluation on the differences between the approaches and which shortcoming their proposed approach has

Author reply:

Thank you for the reminder. We have made revisions and used tables to provide a detailed comparison of existing work in terms of functional requirements, non-functional requirements, environmental requirements, and data requirements. And the weaknesses of the existing work have been explained in the related work section. As shown in the figure below.

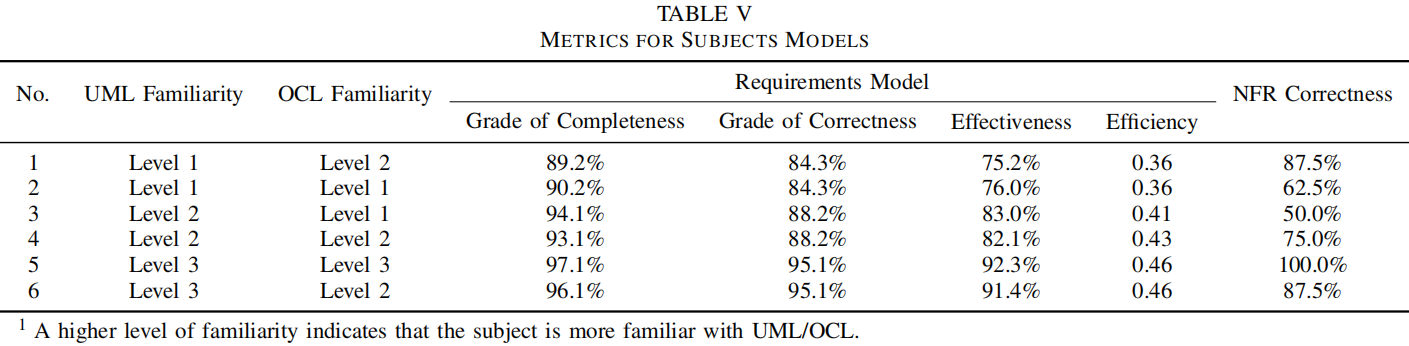


Comment4 of Reviewer 1:

The evaluation with the three different research questions, lacks comprehensive objective argumentations based on some proof (especially RQ3 is a subjective assumption of the researchers)

Author reply:

Thank you for the good advice. We now evaluated the model using manual experiments (including modeling and interviews with RM4ML), quantifying the experimental results with reference to some metrics mentioned in the literature and standards. And the experimental results have been analyzed in detail. As shown in the figure below.

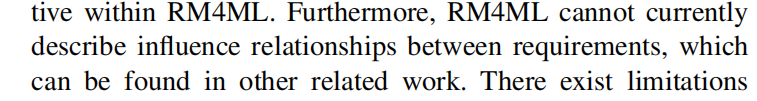


Comment5 of Reviewer 1:

Table 1 only shows the shortcomings of the other three approaches, however it would be interesting which shortcomings RM4AI has which are fulfilled by the other approaches

Author reply:

Thank you for the good advice. We present some of the shortcomings of our work relative to other work in the limitation section. As shown in the figure below.

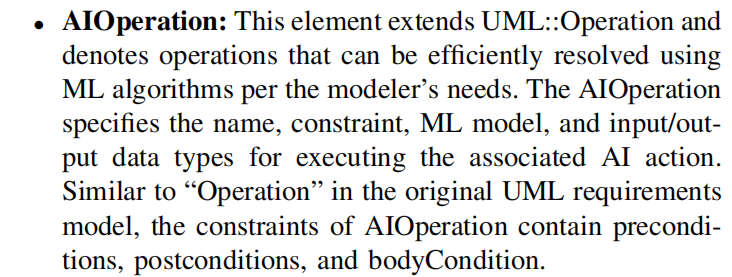


Comment6 of Reviewer 1:

On which source are the identified potential AIOperations based on in the meta-model?

Author reply:

Thank you for the reminder. We have made revisions and believe that if we can use machine learning algorithms to get a better solution for the operation is AIOperation, this element to be selected manually by the modeler. And more detailed introduction to AIOperation is given in the RM4ML section. As shown in the figure below.

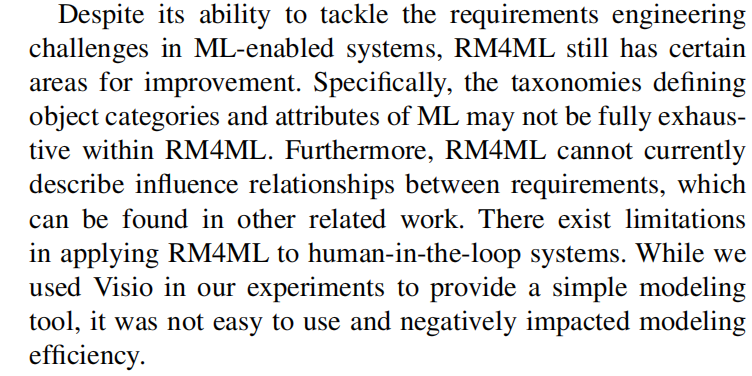


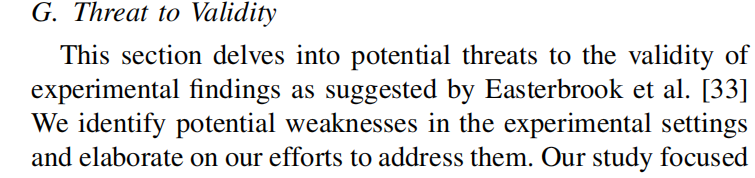
Comment7 of Reviewer 1:

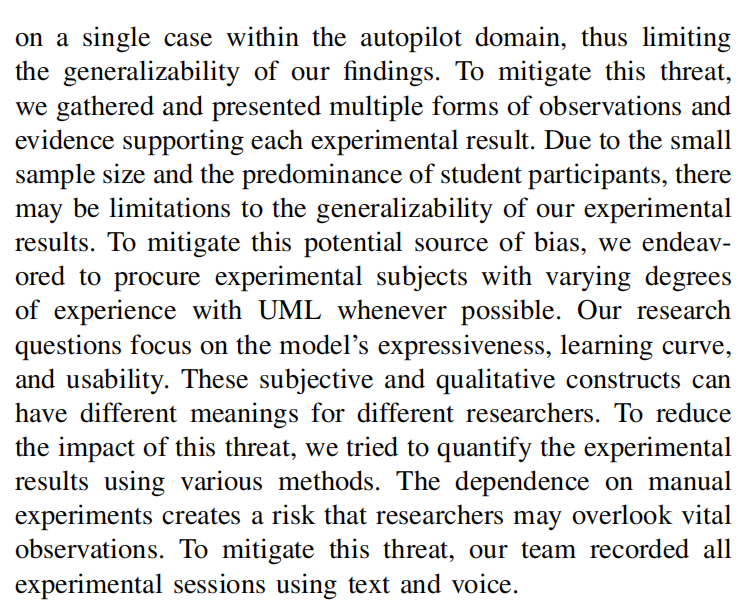
Discussion: very short and potential limitations about the evaluation are not mentioned

Author reply:

Thank you for the reminder. We have made revisions and expanded the description of the limitation section. And added some weaknesses in the experimental settings and our attempts to mitigate these weaknesses. As shown in the figures below.





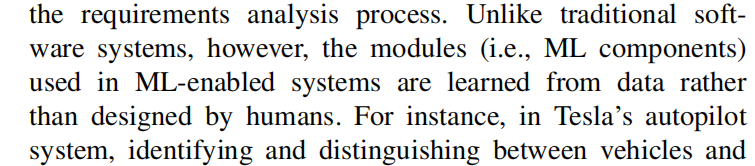


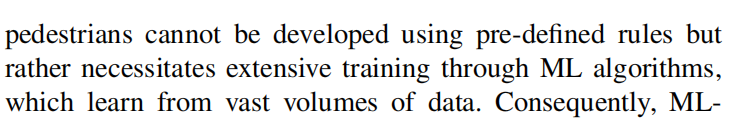
Comment8 of Reviewer 1:

"..significant differences in construction methods between AI-enhanced systems and conventional software systems" --> examples would be helpful to underline this statement

Author reply:

Thank you for the reminder. We have made revisions and illustrated the differences between ML-enabled systems and traditional software systems in the introduction section with examples from the Tesla case. As shown in the figures below.



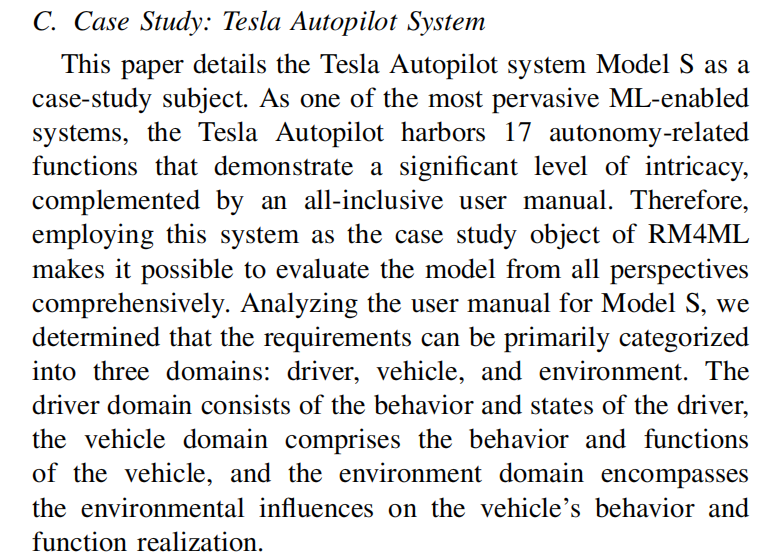


Comment9 of Reviewer 1:

however if the evaluation is based on this example, it is difficult to generally answer the proposed three RQ (How can you generally identify that the proposed model specify the elements in AI-enhanced systems with only one short example and no real use case?)

Author reply:

Thank you for the reminder. We have made revisions and used the case of Tesla Autopilot, which is highly complex and widely used. And this case is described in detail in the case study section. As shown in the figure below.

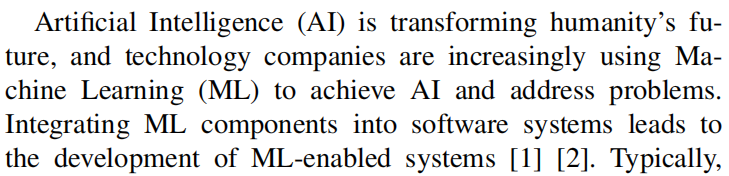


Comment1 of Reviewer 2:

Unclear definition of AI-enhanced systems

Author reply:

Thank you for the reminder. We have made revisions and changed the RM4ML applicable system to ML-enabled system. And the ML-enabled system is defined in the introduction section. As shown in the figure below.

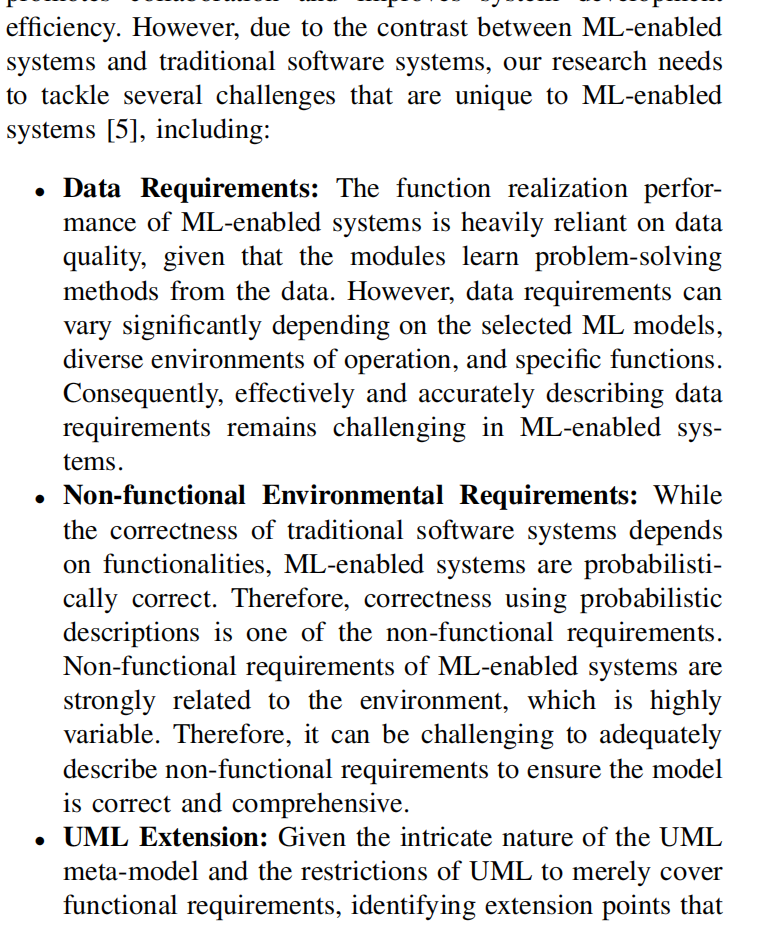


Comment2 of Reviewer 2:

The problem that the paper aims to study and resolve is not clearly stated. Unclear definition of the challenges that the paper aims to tackle (gap in the literature? Or specific need of RE for AI-enhanced systems?)

Author reply:

Thank you for the reminder. We have made revisions and added three challenges of RE for ML-enabled systems: " Definition of data requirements", " Description of non-functional requirements and environment" and " Extension of UML ". These three challenges are described in detail in the introduction section. As shown in the figure below.

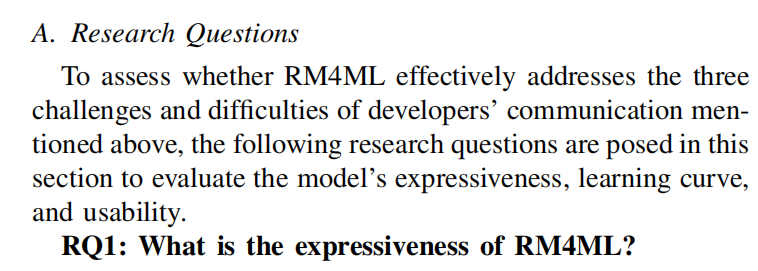


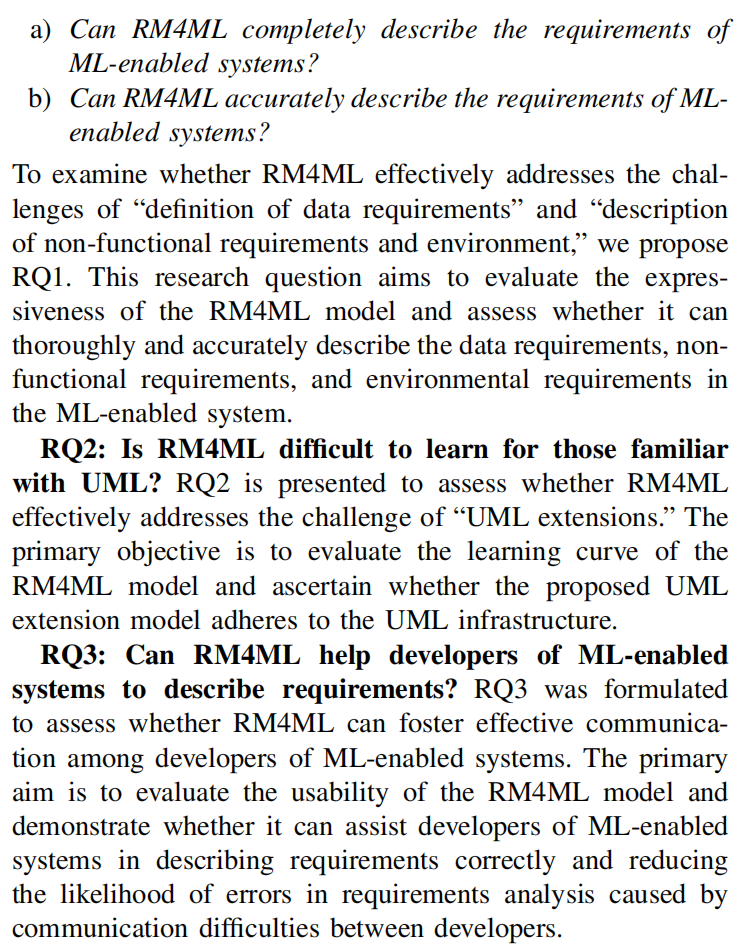
Comment3 of Reviewer 2:

The research questions are formulated in a way that the goal is only to evaluate the RM4AI model and not to answer specific needs when modeling requirements for AI-enhanced systems.

Author reply:

Thank you for the reminder. We have made revisions and the three research questions are currently proposed to evaluate the expressiveness, learning curve, and usability of the model, and to verify whether our proposed model can meet the three challenges presented. We introduce the three research questions in the Research Question section. As shown in the figures below.





Comment4 of Reviewer 2:

The evaluation does not provide specific arguments. There is no definition of a validation criterion or metric. Metrics are used to quantify the complexity of RM4AI, but no information is given to evaluate that complexity. Are there some thresholds to specify the level of complexity? When is the complexity too high? Moreover, how can we know that, when comparing the values of the metrics for RM4AI and UML, the difference in the metrics values are significant or not? The authors state that the model is “easy to learn”. Yet, they do not provide the arguments needed to infer that conclusion from the different metrics values.

Author reply:

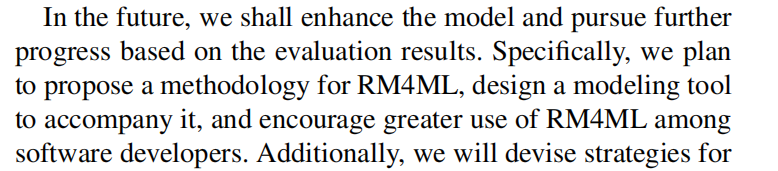
Thank you for the reminder. We have made revisions and analyzed the modeling process of the expert model to derive thresholds, compared the quantified values in the subject models to the thresholds, and demonstrated that the models are "easy to learn" through interviews with the subjects. The experimental results are discussed in the Experimental Results section.

Comment5 of Reviewer 2:

Improvement of the evaluation process is mentioned in the conclusion, but without many details. Specifying what “extensive” means in “through extensive experiments” (cf. conclusion) would allow the reader to understand what other types of experiments can be done to further assess the RM4AI model. Giving some examples of those other types of systems could help to do that.

Author reply:

Thank you for the good advice. We have made revisions and add work on evaluation to our future work, and we will invite more software developers to use this model to model the requirements for their systems. As shown in the figure below.

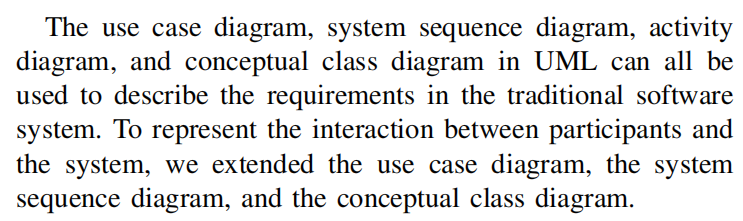


Comment1 of Reviewer 3:

The presentation, in general, is ok to follow. However, in some places, it lacks sufficient justifications of certain choices, e.g., why extending use case, sequence and class diagrams, not others.

Author reply:

Thank you for the reminder. We have made revisions to add descriptions of why use case diagrams, system sequence diagrams, and concept class diagrams were selected for expansion in the RM4ML section. As shown in the figure below.

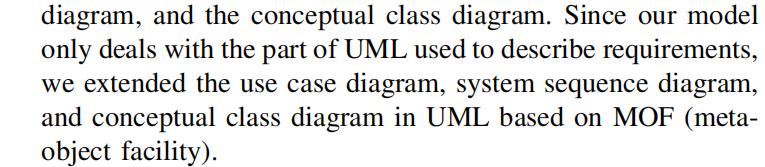


Comment2 of Reviewer 3:

Since it is an extension to UML, it is expected to propose a profile, where stereotypes are proposed. The paper however does not present such a profile and it is not clear how the extension is implemented from the tooling aspect.

Author reply:

Thank you for the reminder. Since we only need to deal with the parts of the UML that are available to represent requirements, we are using a MOF-based approach to extend the UML. And we introduce it in the RM4ML section. As shown in the figure below.



Comment3 of Reviewer 3:

The NFR extension part is not specific to AI-enhanced systems. In the literature, there are ways to specifying/modeling NFRs. It is hard to see why a new extension is needed.

Author reply:

Thank you for the reminder. The model description section of our article explains why we want to extend the non-functional requirements element. As shown in the figure below.

