

We thank the reviewers for their continued feedback on our manuscript. We have carefully addressed the concerns raised in the second round of reviews and significantly improved the paper accordingly. Below we provide detailed responses to each reviewer’s comments.

1 Review 1

Overall Impression: The reviewer expresses concerns about the paper’s structure, background, notation, and informal writing style, questioning whether it meets the technical and scientific standards expected.

Main Concerns:

- **Structure, background, notation etc. very poor** → *We have revised the paper’s structure and background sections. Section 2 (Background) and 5 (Classification) have been improved and clarified to provide a clearer definitions and a more prominent presentation of core concepts. We made the concepts now labeled consistently, clearly defined, and well-grounded conceptually and based on related literature.*
- **Informally written** → *We have significantly enhanced the formality and depth of our analysis. All contractions were removed, and we carefully replaced informal expressions with appropriate academic language.*
- **Weak meta-analysis** → *Section 4 (Categorization based on symbolic structures used) was expanded with a new subsection “Synthesis and Observations”. Section 5 (Classification) was also expanded to provide more substantial meta-level commentary on the surveyed literature. We chose not to expand it too much because we did not want to extrapolate beyond what we observed, and the great variety of methods observed made it hard to draw more observations that would apply to all these contributions. We would like to keep the primary goal of the paper to survey the neuro-symbolic/trustworthiness intersection field beyond approaches labeled as “neuro-symbolic”, which had the consequence of regrouping contributions from different research communities and with different codes.*

- **No clear-cut introduction to different definitions and how they relate** → *The different dimensions should now be clearly defined, and section 5 explicitly articulates the relationships between the three axes of our classification framework.*

Following those modifications, we have made the paper an analytical and scientifically rigorous contribution that provides valuable meta-level insights for the research community.

2 Review 2

Overall Impression: Positive - The reviewer finds the paper demonstrates editorial care, appropriate literature awareness, and is publishable with minor improvements.

Main Concern:

- **Relationships between axes not explicitly articulated**
→ *The reworked version of section 5 now explicitly articulates the relationships between the three axes of our classification framework.*

3 Decision Letter Requirements

The decision letter specifically highlighted the following issues requiring attention:

1. **Extend meta-level commentary** → *Modified the sections 4 and 5 and added some meta-level commentary. For details please see the response to reviewer 1.*
2. **Improve structure of Section 4** → *We added more indications and rephrasing to clarify that rule-learning papers were set aside because of their particular methodology, which puts them all in the same categories regarding the type of interpretability. In terms of symbolic knowledge, since by definition they all manipulate rules, they would technically fall into the logic category, but it did not seem to be very informative to classify them with the others.*

3. **Clearly and prominently define critical concepts** → *These definitions are now clearly defined and prominently appear in the paper, in both the background section (2) which provides conceptual grounding and links with related literature, and in the classification section (5) to clearly define the categorization criteria.*
4. **Polish formal and academic writing style** → *Thank you, we conducted a thorough revision, removing all contractions throughout the manuscript. We replaced informal expressions with appropriate academic language, ensured consistent use of technical terminology, and improved some sentence structure for clarity and formality.*

Again, we want to thank the reviewers and editors for their time and efforts.