Qeios

Peer Review

Review of: "Undefinable True Target Learning: Towards Learning with Democratic Supervision"

Lanzhen Yang[1]

1. College of mathematics and information science, Hebei University, China

This paper touches upon the intersection of artificial intelligence philosophy and machine learning theory, holding the potential to become the beginning of a new research field.

1. Formalization and Operationalization of the UTTL Framework: After proposing that the target is "undefinable," what becomes the learning objective? How is the optimization function defined? How can "democratic supervision" be translated into computable signals?
2. Experimental/Example Validation Section: In addition to summarizing existing works as examples, it is recommended to provide at least one thought experiment or simplified simulation. This would visually demonstrate the advantages of UTTL in handling fundamental ambiguity and multiple perspectives.

For example: Design a simulated environment where multiple "annotators" (agents) with different biases and abilities continuously label a subjective concept (e.g., "artistic value," "news fairness"). Demonstrate how the UTTL framework operates in this context and compare it with methods that pursue a "single true label."

Declarations

Potential competing interests: No potential competing interests to declare.

doi.org/10.32388/AOI1AJ