

Case ID:M23-184P

Published: 3/5/2024

Inventors

Paulo Shakarian

Noel Ngu

Lakshmivihari Mareedu

Abhinav Koyyalamudi

Contact

Physical Sciences Team

Predicting Accuracy of a Machine Learning Model

In the realm of natural language processing, large language models (LLMs) have emerged as powerful tools capable of understanding and generating human-like text. In terms of accuracy, LLMs are known to have performance issues, specifically when reasoning tasks are involved. And despite their advancements, LLMs often face challenges when tasked with solving complex math word problems (MWP). These challenges stem from the need for multi-step reasoning and comprehension of mathematical concepts, presenting a barrier to reliable performance.

Existing solutions (e.g., meta-learning and introspection attempts to predict when an LLM will succeed for a given input) primarily rely on the output of LLMs without providing insights into the reliability of their responses. Consequently, there is a growing demand for algorithmic tools that can predict the performance of LLMs, offering users a means to assess the accuracy of the model's answers.

Researchers at Arizona State University have developed an algorithm that predicts the performance of a large language model (LLM). Leveraging machine learning introspection techniques, the algorithm analyzes features extracted from the symbolic representations of math word problems to estimate the accuracy of LLM-generated responses.

By examining the structural characteristics of math word problems, the algorithm generates a quantitative score indicating the reliability of LLM outputs. This predictive model enhances the transparency and trustworthiness of LLM-generated solutions, allowing users to make informed decisions based on the model's predicted performance. The algorithm could become part of an app, website, or desktop software used for interfacing with an LLM. It can be used as a form of middleware running on the same system as the LLM and/or as a form of software library and fully integrated with LLM software (e.g., as a training process for the LLM).

Related publication: [An Independent Evaluation of ChatGPT on Mathematical Word Problems \(MWP\)](#)

Potential Applications:

- Educational Technology
- Automated Assessment Systems
- Chatbot and Virtual Assistant Development

Benefits and Advantages:

- Improved Accuracy

- Transparency and Trust
- Efficient Learning