Improved AI confidence measure for autonomous vehicles

April 15 2024

Credit: Unsplash/CC0 Public Domain

A new Bar-Ilan University study addresses a fundamental question in the realm of artificial intelligence (AI): Can deep learning architectures achieve greatly above-average confidence for a significant portion of

inputs while maintaining overall average confidence?

The study's findings provide an emphatic "yes" to this question, marking a significant leap forward in AI's ability to discern and respond to varying levels of confidence in classification tasks. By leveraging insights into the confidence levels of deep architectures, the research team has opened new avenues for real-world applications, ranging from autonomous vehicles to health care.

The study was published in *Physica A: Statistical Mechanics and its Applications* by a team of researchers led by Prof. Ido Kanter from Bar-Ilan University's Department of Physics and Gonda (Goldschmied) Multidisciplinary Brain Research Center.

Ella Koresh, an undergraduate student and a contributor to the research, emphasizes the practical implications of the work. "Understanding the confidence levels of AI systems allows us to develop applications that prioritize safety and reliability," she explains.

"For instance, in the context of autonomous vehicles, when confidence in identifying a road sign is exceptionally high, the system can autonomously make decisions. However, in scenarios where confidence levels are lower, the system prompts for human intervention, ensuring cautious and informed decision-making."

Enhancing the confidence levels of AI systems holds profound implications across diverse domains, from AI-based writing and image classification to critical decision-making processes in health care and autonomous vehicles. By enabling AI systems to make more nuanced and reliable decisions when faced with uncertainty, this research sets a new standard for AI performance and safety.

More information: Yuval Meir et al, Advanced Confidence Methods in Deep Learning, *Physica A: Statistical Mechanics and its Applications* (2024). DOI: 10.1016/j.physa.2024.129758

Provided by Bar-Ilan University