

DEPARTMENT OF COMPUTER SCIENCE

SEMINAR

2026 SERIES

Emergent Abilities in Large Language Models: Mirage or an Elusive Predictive Frontier?

DATE & TIME

19 JAN 2026 (MON) 10:00 - 11:00 AM

VENUE

WLB104, 1/F, WING LUNG BUILDING, SHAW CAMPUS



MR. BRANDO MIRANDA

Ph.D. Student
Stanford University

ABSTRACT

This talk is presented in two parts, tracing a line of inquiry from deconstructing myths in AI evaluation to a study of some of the reasons why such myths occurred from a scientific perspective.

Mirage: First, we will address the popular notion of "emergent abilities" in LLMs. Drawing from our work, "Are Emergent Abilities of Large Language Models a Mirage?", we demonstrate that many of these sudden, unpredictable leaps in capability are often artifacts of how we measure performance—specifically, the use of non-linear or discontinuous metrics.

Elusive: The second and final part delves deeper into why predicting downstream capabilities remains so elusive, even with smooth scaling laws for pre-training. We will discuss our findings from "Why Has Predicting Downstream Capabilities of Frontier AI Models with Scale Remained Elusive?", showing that downstream evaluation metrics involve a series of transformations that progressively degrade the statistical relationship with scale. We pinpoint the core mechanism: performance depends not only on a model identifying the correct answer but also on the unpredictable probability fluctuations across a small set of incorrect choices.



SPEAKER'S
BIOGRAPHY



REGISTER NOW