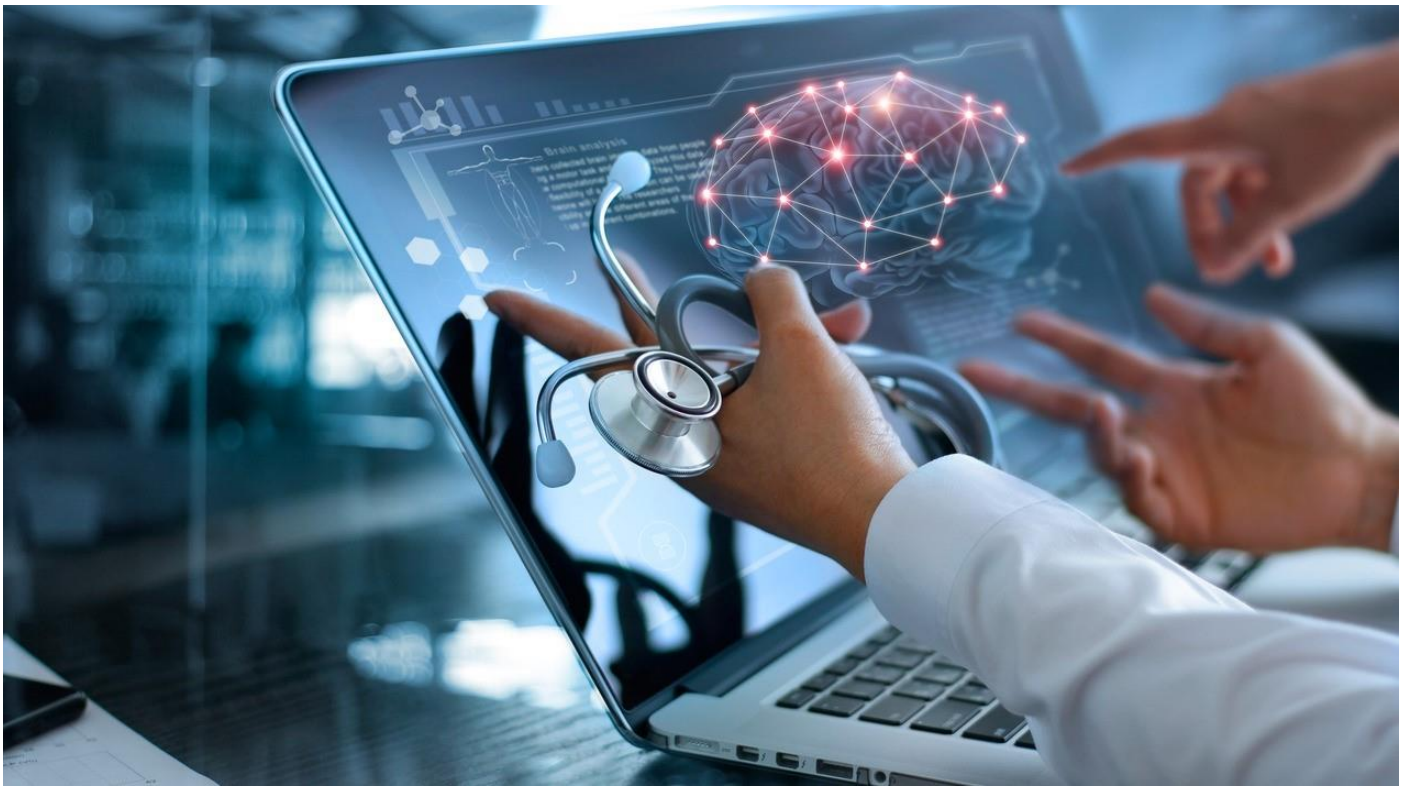
# NTU Navigating Complexity to Enhance AI Models

Alita Sharon | February 7, 2024



Understanding how the brain combines previously acquired skills to tackle new challenges remains a complex puzzle in neuroscience. Nanyang Assistant Professor Hiroshi Makino, from NTU's Lee Kong Chian School of Medicine, sheds light on this intriguing process through a study involving mice and theoretical analyses with artificial intelligence (AI) models. The research not only unveils the brain's ability to compose new skills but also holds implications for enhancing AI models.

Mice were trained in behavioural experiments involving tasks requiring the manipulation of a joystick to move an object towards a destination. Successful completion of the task resulted in a water reward. Subsequently, the mice were trained to associate licking a waterspout with receiving water. The complexity increased when the mice were challenged with a combined task – using the joystick to move the waterspout to a specific location and then licking it to obtain their water reward.

Assistant Professor Makino delved into the neural activity of both mice and AI models. The focus was on understanding how the mice integrated their learned skills to accomplish the composite task. The findings revealed a mechanism where the brain combines representations of pre-learned action values from constituent subtasks. This insight into the brain's learning process holds potential for improving our understanding of cognitive functions.

The study incorporated theoretical predictions from the field of deep reinforcement learning, where agents learn to solve composite tasks by combining representations of pre-learned action values from simpler subtasks. AI models provided a theoretical framework that helped in understanding the learning process observed in mice. The theoretical predictions were validated through empirical testing on the mice, showcasing a convergence between artificial and biological systems.

Assistant Professor Makino believes that this research not only enhances our understanding of how the brain learns but also has implications for improving AI models in the future. The ability to compose new skills from a pre-acquired repertoire is a crucial aspect of biological intelligence, and the study offers valuable insights into this fundamental cognitive process.

The study draws parallels between deep RL algorithms, which leverage policy entropy to express stochastic policies, and the initial high variability observed in behaviour during pretraining. This algorithmic convergence between artificial and biological systems prompts further exploration into the mechanisms that promote exploration for future learning.

This groundbreaking study not only unravels the mysteries of how the brain combines learned skills but also establishes a crucial connection between artificial intelligence models and biological systems. As we delve deeper into the brain's ability to tackle new challenges through the integration of existing knowledge, the study opens avenues for future research in neuroscience and AI, with potential implications for advancements in both fields.

NTU Singapore, under the leadership of President Professor Ho Teck Hua, has unveiled a new initiative to address global opportunities and challenges posed by artificial intelligence (AI). The university aims to boost AI education, research, and innovation, introducing new programmes such as the Bachelor of Science in AI and Society and the Turing AI Scholars Programme. These efforts align with Singapore's national AI strategy and solidify NTU's position as a leading global university.

Singapore is keenly focused on deploying Artificial Intelligence across various fields to enhance efficiency and drive innovation. Recognising the transformative potential of AI, the Singaporean government has been actively fostering an ecosystem conducive to the development and application of AI technologies.

Singapore proactively tackles safety and ethical concerns tied to AI technologies by investing in research and development to formulate guidelines and standards prioritising safety in AI deployment. This approach seeks to foster public trust in AI while minimising associated risks linked to its extensive adoption.