

THE *Current*

August 20, 2025

James Badham

Google honors UCSB computer scientist for pioneering low-cost AI network models

With potential to make powerful AI systems more affordable and accessible, UC Santa Barbara computer scientist Arpit Gupta has earned two major research awards from Google to support his development of low-cost network foundation models. Gupta's work, which bridges machine learning and computer networks, aims to lower the cost of deploying large-scale AI infrastructure — an advance with broad implications for the efficiency, scalability and democratization of future technologies.

For Gupta, the awards validate his bold vision of rethinking how networks can be made “self-driving” by using a new class of machine-learning models capable of managing themselves with minimal human intervention.

The first honor, a Google Research Scholar Award, supports early-career faculty pursuing promising, high impact research. The second, the inaugural Google ML (Machine Learning) and Systems Junior Faculty Award, recognizes junior faculty worldwide conducting cutting-edge work at the interface of machine learning and systems.

Together these highly selective awards place [Gupta](#), an assistant professor of computer science, among an elite group of emerging leaders with the potential to shape the future of AI and networking.

“The (Google ML) award is going to more than 50 assistant professors in 27 U.S. universities whose research is particularly noteworthy for Google,” said Amin Vahdat, VP/GM of ML, Systems & Cloud AI. “These professors are leading the analysis, design, and implementation of efficient, scalable, secure and trustworthy computing systems. Their work crosses the technology stack, from algorithms to software and hardware, enabling machine learning and cloud computing at an increasingly massive scale.”

For decades, machine-learning problems in networking have been solved through *point solutions* — single models tailored to solve individual problems. But as networks have grown larger and more complex, maintaining separate models for each decision task has become time-consuming, computationally expensive and difficult to scale.

Gupta is pursuing a new approach he calls the *convergence principle*, his term for the idea that, instead of building and maintaining separate models for each task, it may be possible to develop a single, general-purpose *foundation model* — a large, general-purpose model pre-trained on diverse data and fine-tuned for specific tasks as needed — that can adapt to a wide variety of networking problems across different scales.

“Over time, my research has been guided by this convergence principle,” Gupta said. “We started to ask: if we want to build self-driving networks capable of making a wide range of decisions, do we really need to engineer thousands of distinct machine-learning models, or is there a smarter way to unify these capabilities in a single system?”

His work is inspired by the trajectory of the natural language processing (NLP) community, which faced similar challenges in developing task-specific models. Initially, NLP researchers built separate models for translation, sentiment analysis and question answering, each trained and maintained independently. The turning point came with the development of foundation models, such as BERT in 2018 and later GPT-3, which demonstrated that training on broad, diverse datasets could yield a single adaptable model. “They showed it was possible to move beyond point solutions by building a foundation model that could be fine-tuned for a variety of tasks,” Gupta said. “We’re exploring what it would take to bring that approach to networking.”

But networking, Gupta emphasized, poses unique challenges. “We wondered whether we could borrow what has been done in other areas and adapt it for networking, or if we would need to start from scratch and build something specific to the networking environment, a phrase referring not only to the static infrastructure but also to the dynamic, constantly changing context which includes traffic patterns, interference, competing user demands, and even malicious activity,” he says. “We realized that, while there are similarities in what makes a good model, networking has its own unique problems.”

One of Gupta’s research directions focuses on analyzing packet traces, which capture how network rules and protocols interact with that dynamic environment. By examining patterns in these packets — the smallest units of data on a network — Gupta can infer information about the state of individual hosts, subnets and even the broader network. “When I combine groups of packets sent by multiple connections on the same host, it tells me something about the host itself or the group of devices in that subnet,” he said.

Another challenge unique to networking is scale. Modern networks can transmit billions of packets every second, making it impossible to observe and process each one individually. “That problem of scale has always been there,” Gupta said, “and it means we have to build solutions that are smart enough so that we can avoid processing every packet individually.”

To address this challenge, Gupta’s group is developing what he calls a *selective representation approach*. Every packet that travels over a network carries two parts: a header and a payload. The header contains routing information, such as the source and destination, while the payload holds the actual data being sent. Deciding how much of this information to examine becomes a balancing act.

“We need to figure out how much information is enough to make an informed decision without incurring enormous computational costs,” Gupta explained. “One way to do that is by designing intelligence that minimizes the amount of data extracted from each packet while maximizing the overall insight we gain.”

In this approach, rather than trying to inspect every packet, the system identifies representative packets to act as proxies for the rest, reducing the computational

burden while accessing essential information. “A very important aspect of networking is that we can’t treat all packets the same way,” Gupta said. “What you do depends on the type of information you need. Counting packets is computationally cheap, but understanding where a packet is going and who sent it requires inspecting the headers, which is more expensive. And if you need to examine the payload — the actual content — that’s even more intensive.”

Gupta’s broader vision is to create a foundation model for networks that can leverage multi-modal data from diverse sources — including packet traces, telemetry logs and device statistics — to solve complex learning problems at different scales. This could mean making decisions every few packets, every second, or every few minutes — a flexibility not feasible with today’s task-specific models.

Currently, many networks rely on stop-gap solutions that are trained and optimized for individual use cases. While effective for narrow tasks, this approach doesn’t scale well. “Developing a task-specific model for each learning problem is not sustainable,” Gupta said. “It takes significant effort to design, train, and optimize each one. We believe a foundation model is the next step.”

Gupta is already working with [ESnet](#), a national research network, to explore how such a model might be deployed. “There’s already momentum,” he said. “Companies like Cisco and others are paying attention to what this solution is going to look like. I think in the next few years, a foundation model for networking is going to be a real thing.”

Tags

[Artificial Intelligence](#)

[Innovation & Entrepreneurship](#)

Media Contact

Debra Herrick

Associate Editorial Director

(805) 893-2191

debraherrick@ucsb.edu

About UC Santa Barbara

The University of California, Santa Barbara is a leading research institution that also provides a comprehensive liberal arts learning experience. Our academic community of faculty, students, and staff is characterized by a culture of interdisciplinary collaboration that is responsive to the needs of our multicultural and global society. All of this takes place within a living and learning environment like no other, as we draw inspiration from the beauty and resources of our extraordinary location at the edge of the Pacific Ocean.