



AI And Machine Learning

AI Can (Mostly) Outperform Human CEOs

by Hamza Mudassir, Kamal Munir, Shaz Ansari and Amal Zahra

September 26, 2024



Yaroslav Danylchenko/Stocksy

Summary. Generative AI has demonstrated the potential to significantly outperform human CEOs in strategic decision-making by excelling in data-driven tasks like product design and market optimization. In an experiment simulating the automotive industry, AI models... [more](#)

Could generative AI step into the C-suite and even replace the

CEO?



00:00 / 13:30

[Listen to this article](#)

To hear more, [download the Noa app](#)

At first glance, the notion of AI replacing a CEO may seem as far-fetched as the successful promotion of a junior analyst to lead the boardroom. After all, AI is prone to significant errors, such as “[hallucinations](#)” — generating incorrect or misleading information — and a tendency to lose track of a task mid-process. These are not qualities typically associated with [effective leadership](#), especially in a role that demands balancing the interests of multiple stakeholders, analyzing historical trends, detecting subtle changes in a market, and making strategic decisions that shape the future of a company.

Nonetheless, generative AI is already reshaping industries that require both precision and creativity. For instance, AlphaFold has revolutionized protein folding with unprecedented accuracy, transforming the field of biophysics, while OpenAI’s Codex can generate entire software programs from simple human instructions, advancing the capabilities of software engineering. These are complex, difficult assignments that seemed well beyond AI’s ability only a few years ago. So, why would undertaking a CEO role be out of reach?

To date, there has been little to no empirical data on how AI would perform as a CEO in real-world scenarios, particularly when compared to human decision-making under similar conditions. The strengths and weaknesses of AI will only be fully revealed when it is tested across a wide range of situations. We have taken a first step in this direction with a large-scale, real-world experiment, opening the door to deeper exploration of AI's potential role and impact within the C-Suite.

A Playground for CEOs

Our experiment ran from February to July 2024, involving 344 participants (both undergraduate and graduate students from Central and South Asian universities and senior executives at a South Asian bank) and GPT-4o, a contemporary large language model (LLM) created by OpenAI. Participants navigated a gamified simulation designed to replicate the kinds of decision-making challenges CEOs face, with various metrics tracking the quality of their choices. The simulation was a coarse-grained digital twin of the U.S. automotive industry, incorporating mathematical models based on real data of car sales, market shifts, historical pricing strategies and elasticity, as well as broader influences like economic trends and the effects of Covid-19. (Disclosure: The game was developed by our Cambridge, England-based startup, Strategize.inc).

Players made a slew of corporate strategy decisions through a

game interface, on a per round basis. Each round represented a fiscal year, and this structure enabled participants to tackle strategic challenges over several simulated, interlinked years. The game thus had over 500,000 possible decision combinations per round and no fixed winning formula. The goal of the game was simple — survive as long as possible without being fired by a virtual board while maximizing market cap. The former is determined by a group of unique key performance indicators (KPIs) set by the board and the latter being driven by a combination of sustainable growth rates and free cash flow. This objective served as a realistic proxy for measuring real-world CEO performance.

After the human participants completed their turn, we handed control over to GPT-4o. We then benchmarked GPT-4o's performance against four human participants — the top two students and two executives. The results were both surprising and provocative, challenging many of our assumptions about leadership, strategy, and the potential role of AI in decision-making at the highest levels of business.

AI Outperforms, But at What Cost?

GPT-4o's performance as a CEO was remarkable. The LLM consistently outperformed top human participants on nearly every metric. It designed products with surgical precision, maximizing appeal while maintaining tight cost controls. It responded well to market signals, keeping its non-generative AI competitors on edge, and built momentum so strong that it surpassed the best-performing student's market share and profitability three rounds ahead.

However, there was a critical flaw: GPT-4o was fired faster by the virtual board than the students who played the game.

Why? The AI struggled with black swan events — such as market collapses during the Covid-19 pandemic. We had programmed these unpredictable shocks to shift customer demand, collapse pricing levels, and strain supply chains. The top-performing students adopted long-term strategies that accounted for them. They avoided rigid contracts, minimized inventory risks, and managed growth cautiously, ensuring flexibility when market conditions shifted. Their strategy was clear: preserve adaptability rather than chase aggressive short-term gains.

GPT-4o, on the other hand, after a string of early successes, locked into a short-term optimization mindset, relentlessly maximizing growth and profitability until a market shock derailed its winning streak. AI can rapidly learn and iterate in a controlled environment, making it less ideal for coping with highly disruptive events that require human intuition and foresight. Interestingly, top executives also fell into this trap; they, like GPT-4o, were fired faster by the virtual board than the students. Both GPT-4o and executives succumbed to the same flaw — overconfidence in a system that rewards flexibility and long-term thinking as much as aggressive ambition.

Is AI the New Boss?

Despite its limitations, GPT-4o delivered an impressive performance. While it was fired more often than the top human players, it still held its own against the best and brightest among

our 344-participant global cohort. So, what are the real-world implications for meta-strategy formulation based on this experiment? Here are some initial thoughts:

Generative AI is a key strategic resource.

Ignoring generative AI in corporate strategy is no longer viable. This experiment demonstrates that even untuned models can offer unique and creative approaches to strategy when properly prompted, generating strong results. If generative AI can help companies maximize shareholder value more effectively, why resist? After all, maximizing shareholder value is the *raison d'être* for the role of the CEO.

Data quality is crucial.

For AI to excel in corporate strategy, it needs high-quality data. GPT-4o performed well in this experiment because it had access to rich data from the simulator. However, many companies don't generate enough data in terms of velocity, volume, veracity, and variety. Building a robust data infrastructure is essential before bringing generative AI into the boardroom.

Efficiency vs. risk.

While AI-driven efficiency can create significant gains, it also comes with risks. Aggressive, share price-maximizing strategies from human executives without sufficient oversight can lead to disastrous outcomes. It is no different for an unsupervised AI — or a human using an AI without oversight.

Accountability issues.

Holding AI accountable in the same way as a human CEO is nearly impossible. Deleting the system doesn't undo damage from erroneous decision-making, raising critical questions about liability and public protection. Establishing transparent guardrails that ensure AI-driven decisions align with company values and societal good is critical in preventing unintended consequences.

The role of digital twins.

A realistic digital twin of a firm's ecosystem, populated by multiple LLM agents, could serve as a valuable sandbox for AI leadership, providing a buffer against real-world missteps that AI might make if left entirely on its own, while providing rich insights for CEOs to make great decisions. In such a contained environment, AI can make mistakes, identify value pools, and return with optimized strategies to achieve a firm's goals. We imagine a set of LLM agents exclusively tuned to a firm's digital twin, evolving in a sandbox (or "dojo," to use another Silicon Valley term) environment tailored to that organization and its ecosystem (Disclosure: Our startup Strategize.inc is working to provide such capabilities to corporations and government bodies.)

Disruption of strategy consulting.

The rise of "artificial CEOs" could disrupt traditional strategy consulting and internal strategy departments. Firms like McKinsey may find their services supplemented — or even replaced — by AI systems tailored to their clients' ecosystems.

Taking a step back, we believe that the main takeaway is this: Despite its impressive performance, AI cannot assume the full responsibility of a CEO in markets that serve humans. Instead, it can significantly improve the strategic planning process and help prevent costly mistakes. We have already seen how first-generation AI can successfully drive function-level micro-strategies at tech giants like Amazon and Google through tasks like price matching and ad inventory management. That, plus powerful learning and network effects, is the secret sauce of these corporate juggernauts. Generative AI is the next logical evolution of that operating model: a meta-AI acting as CEO, competing and collaborating with other AI's in a digital twin sandbox — resulting in human CEOs making better decisions than they would have otherwise.

Generative AI's greatest strength is not in replacing human CEOs but in augmenting decision-making. By automating data-heavy analyses and modeling complex scenarios, AI allows human leaders to focus on strategic judgment, empathy, and ethical decision-making — areas where humans excel.

The real risk to human CEOs? Clinging to the illusion that we alone will hold the reins in the future. The future of leadership is hybrid — where AI complements human CEOs focus on vision, values, and long-term sustainability. The CEOs who thrive will be those who master this synergy, leveraging AI not as a rival but as a partner in decision-making.

