

Hey! What's New? 2025-6

Can AI Predict the Future?

An article in *Knowledge at Wharton*, written by Seb Murray, notes that “new research from Wharton management professor Philip Tetlock finds that combining predictions from multiple artificial intelligence (AI) systems, known as large language models (LLMs), can achieve accuracy on par with human forecasters. This breakthrough offers a cheaper, faster alternative for tasks like predicting political outcomes or economic trends.”

“What we’re seeing here is a paradigm shift: AI predictions aren’t just matching human expertise — they’re changing how we think about forecasting entirely,” Murray quotes Tetlock saying.

Dubbed as the “wisdom of the silicon crowd” by the Wharton academic and his co-authors — Philipp Schoenegger of London School of Economics, independent researcher Indre Tuminauskaite and Peter Park from Massachusetts Institute of Technology — this approach highlights how groups of AI systems can provide reliable predictions about the future.

By pooling predictions from multiple LLMs, the researchers presented a practical method for organizations to access high-quality forecasting without relying solely on expensive teams of human prognosticators. “This isn’t about replacing humans, however,” Tetlock said, “it’s about making predictions smarter, faster and more accessible.”

Individually, AI models such GPT-4, made by Microsoft-backed OpenAI, have struggled with forecasting. Previous studies revealed that their predictions were often no better than random guesses. Tetlock’s paper, “Wisdom of the Silicon Crowd: LLM Ensemble Prediction Capabilities Rival Human Crowd Accuracy,” found, however, that combining predictions from multiple models significantly boosted their accuracy.

So how does it work? According to Tetlock, the magic lies in how errors balance out. Just as human crowds average out individual biases, combining AI models cancels out inconsistencies in their predictions. Each model brings a slightly different perspective, much like human forecasters with varied expertise and experiences. “Just how human crowds balance individual biases, AI ensembles turn competing perspectives into consensus.”

His study also found that AI predictions were greatly improved — between 17% and 28% — when informed by human input, such as insights from forecasting tournaments, where people compete to predict future events accurately. These competitions provide valuable, real-time data that AI systems can incorporate into their predictions. “The best forecasts come when human intuition meets machine precision,” said Tetlock.

Interestingly, though, the researchers found the best results came from simply averaging human and AI predictions, rather than relying on the AI to synthesize them. This highlights a key takeaway: while AI is advancing, human input still plays an important role in creating the most accurate forecasts.

Tetlock and his co-authors put their methods to the test in real-world scenarios by carefully designing questions and situations that the AI models hadn’t encountered during their training. This ensured that the AI wasn’t just “cheating” by regurgitating memorized information.

The results, says Murray, were promising but revealed some challenges. “For example, the AI models struggle when there’s a significant time gap between their training data and the events they’re predicting. This lack of up-to-date knowledge can reduce accuracy. Additionally, the AI systems often exhibit overconfidence, assigning higher probabilities to outcomes that don’t align with the available evidence.”

“Resolution,” he adds, “is a technique that could be used to fix this issue by sharpening predictions to clearly distinguish between what’s likely and unlikely. The goal is to assign higher probabilities to events that actually occur and lower probabilities to those that don’t, ensuring forecasts are both confident and accurate.”

Murray believes that, “with tools in place to overcome these hurdles, the study demonstrates the practical value of AI in forecasting. In areas such as politics and economics, where big decisions depend on precise predictions, combining forecasts from LLMs is a practical, scalable, and efficient approach.”

“This is just the start. As we refine these systems, they’ll not only get more accurate but also change how we make high-stakes decisions,” Tetlock said. “The human forecasters in our comparison baseline were educated, reasonably numerate adults — but not the elite of forecasters on the public platforms (e.g., superforecasters). That is a challenge the LLMs have yet to beat.”

Learn more at [Can AI Predict the Future? - Knowledge at Wharton](#).