

# Artificial intelligence and investing: it's all about the data

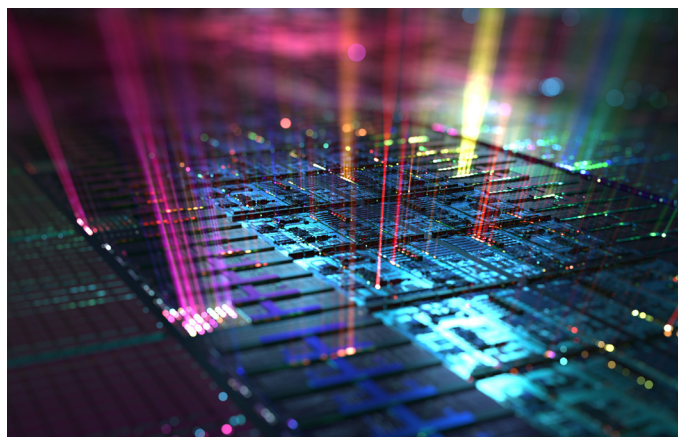
Teaching machines how to “think” has been a tantalizing goal since the early days of computing, and recent advances in generative AI have brought us closer to achieving it. However, whether AI-generated “buy/sell/hold” signals can be used for investing is partly a question of whether the value outweighs the risk. One only needs to read a few passages of AI-written content to notice that it is a simulacrum. When asked to write an essay, the program may line the words up, but the text may lack humour, nuance or deeper insight. It can struggle to solve simple math problems. Sometimes, the model will even come up with an excuse to avoid doing what you asked. Missing the mark can be harmless and amusing if nothing is at stake, but investing based on unverified AI input or output presents a real problem.

## The generative AI revolution

Of course, much of the buzz around AI stems from the release of ChatGPT in 2022. Under the hood, the original ChatGPT used a variant of a large language model called GPT-3.5, which was trained on a dataset containing hundreds of billions of words collected from the internet. One particular strength of generative large language models is their ability to summarize long blocks of text. As an example of how this could help investors, the model could parse lengthy regulatory filings for critical information that a human could overlook. A less obvious but more intriguing possibility is for the model to distill that summary into a rating or other quantitative metric.

Perhaps surprisingly, given its apparent capabilities, the model is trained only to predict the next word in a sentence as accurately as it can. Though complex in the details, the training itself amounts to having the model guess the next word, checking how big of a mistake it made, and adjusting some of its billions of model parameters (“neurons”) so it makes a smaller mistake in the future.

The simple high-level technique of applying massive amounts of data to train a very large model made of simple individual components can lead to impressive results—generally, the bigger the model, the better the result—if there is enough data.



## But...

AI systems can generate misleading or factually incorrect responses with complete conviction. Termed “hallucinations,” these responses continue to afflict large AI models in part because they are trained on vast swathes of human-generated data on the internet, itself replete with inaccuracies presented as fact. In this regard, the models are remarkably human-like. While some adjustments can be made to the models after training, it is impossible to manually check the entire training data to remove inaccurate information.

The reasoning process these AI models use to arrive at answers also is a black box. You could ask a model how it arrived at an answer, but as it is incapable of genuine introspection, the answer will still be based on a best guess of the likeliest next word. Thus, if an AI-generated signal indicates a counterintuitive trade, it will be difficult to understand the true chain of reasoning that led to it.

For systematic investors especially, backtesting can provide essential information about the effectiveness of a strategy. Generative AI models are trained on data up to the present day, and their enormous sizes mean they can memorize a great deal of historical market information. Using a signal generated from such a model risks look-ahead bias: contaminating simulated trading in a past time period with information from the future.

Generative AI holds substantial promise, but even the current generation of models faces significant challenges for use in investing. It is possible to avoid the pitfalls while reaping substantive benefits from advances in AI, however. Incorporating machine learning into investment modeling is one means, but it should be transparent and accountable. Implementing decision trees and other methodology can shine light on the rationale behind trades. The future may be AI, but transparency and using high-quality historical data are critical to avoiding bias and improving reliability.

**Federated  
Hermes**   
Limited

**The value of investments and income from them may go down as well as up, and you may not get back the original amount invested. For professional investors only.** This is a marketing communication. The views and opinions contained herein are those of the MDT team, and may not necessarily represent views expressed or reflected in other communications, strategies or products.

Issued and approved by Hermes Investment Management Limited (“HIML”) which is authorised and regulated by the Financial Conduct Authority. Registered address: Sixth Floor, 150 Cheapside, London EC2V 6ET. HIML is a registered investment adviser with the United States Securities and Exchange Commission (“SEC”). Distributed in the EU by Hermes Fund Managers Ireland Limited which is authorised and regulated by the Central Bank of Ireland. Registered address: 7/8 Upper Mount Street, Dublin 2, Ireland, D02 FT59.