INVESTMENTS & WEALTH MONITOR

A reprinted article from September/October 2021

Artificial Intelligence: Rage Against the Machinewashing

By Aaron Filbeck, CFA®, CAIA®, CIPM, FDP

INVESTMENTS & WEALTH INSTITUTE®

© 2021 Investments & Wealth Institute®. Reprinted with permission. All rights reserved.

Artificial Intelligence

RAGE AGAINST THE MACHINEWASHING

By Aaron Filbeck, CFA®, CAIA®, CIPM, FDP

Artificial intelligence (AI) and machine learning (ML) have the potential to revolutionize investing. These tools can empower investment managers to accelerate pattern recognition, improve forecasts, and make better decisions. But investors need some basic knowledge about AI and ML so they can detect what I call "machinewashing"—investments that make dubious claims to improvements in active management through the use of artificial intelligence.

The uses for AI may seem infinite, but most of the innovation has occurred in back- or middle-office functions (Deloitte 2019; see table 1). ML algorithms—quantitative tools that execute upon the objectives of artificial intelligence—can be as simple (e.g., training and automating a spreadsheet to do everyday tasks) or as complex (e.g., identifying relationships among seemingly disparate pieces of information) as the analyst needs them to be.

This article provides a concise overview of AI, ML, and what managers should look for when considering investments that claim to use these tools in the investment process.

UNDERSTANDING THE LANGUAGE

Artificial intelligence, machine learning, big data, and alternative data are four distinct, yet related, terms.

Artificial intelligence. Artificial intelligence is the ability of machines to solve complex problems using the problem-solving skills typically used by humans. Artificial intelligence comprises three main features: (1) the ability to acquire relevant information and obtain the rules for using it, (2) the ability to apply the rules acquired and use them to reach approximate or definite conclusions, and (3) the ability to change the process based on new information acquired.

Machine learning. Machine learning is the application of artificial intelligence through algorithms that design a sequence of actions to solve a problem. Machine learning goes outside of linear relationships and deals with optimization, prediction, and categorization but not causal inference. We can categorize machine-learning algorithms into two broad categories: supervised learning and unsupervised learning. The major difference is the starting point: theory or data. Supervised learning algorithms start with a theory. In traditional finance, we often ask, "What is the relationship between X and Y?" Supervised learning algorithms are best used for data that are labeled, such as the monthly returns of two different stocks. A supervised learning algorithm might use linear regression, measuring the linear relationship between these stocks. On the other hand, unsupervised learning algorithms start with the data and let a machine discover the relationships. When using an unsupervised learning algorithm, we ask, "Is there a relationship between any of these datapoints?" This is useful for profiling or clustering unlabeled data based on certain characteristics, but researchers may not know what problem they are trying to solve. For example, imagine a large dataset of animal features but no names of species. An unsupervised algorithm might group animals by distinct features, making the data more user-friendly and actionable.

Big data. As the name suggests, big data simply refers to the size of datasets. Big datasets may come from traditional data sources, such as financial statements or regulatory filings, and alternative data, defined below. Some big datasets may not fit on a single computer, so sophisticated techniques such as simultaneously analyzing a dataset using multiple computers may be deployed.

Alternative data. In finance, alternative data refers to data that cannot be found within traditional sources, such as regulatory filings or financial statements. Examples include satellite imagery, social media data, climate patterns, image recognition, and credit card transaction data.

Table 1

USE CASES FOR ARTIFICIAL INTELLIGENCE ALGORITHMS

Alternative Data	Using alternative data (e.g., weather forecasts, shipping movements) to gain better real-time insights
Automated Insights	Automating traditional sources of information, such as earnings transcripts
Operations Automation	Automating repetitive back-office functions
Risk Management	Monitoring suspicious transactions and triggering responses

Source: Adopted from Deloitte (2019)

© 2021 Investments & Wealth Institute. Reprinted with permission. All rights reserved.

THE BIG RISK: OVERFITTING

One risk associated with machine learning in finance is that, because markets are always changing, many models built on past information can quickly become irrelevant. Unfortunately, because none of us can see the future, we must rely on historical data and market cycles to test the viability and economic significance of new ideas. However, we also must be aware that past results are no indication of future success. Therefore, financial professionals are constantly facing a trade-off between generating a perfect backtest, i.e., a model that performs well in an historical market scenario, and creating a model that will have an acceptable level of efficacy in the future.

This is a balancing act between over- and underfitting a model to the data. An overfitted model performs well in the past because the researcher incorporates all past information into the algorithm. Therefore, the model may have too many independent variables associated with it. It looks great on past data but performs poorly in real time because many of those independent variables are irrelevant today. An underfitted model is the exact opposite. The researcher doesn't incorporate enough information into the model, which causes the model to ignore important variables and perform poorly in real time.

Neither of these situations are optimal, but there's a natural incentive to create overfitted models. After all, when presenting a backtested model to an investment committee or a prospective client, there's a natural desire to put your best foot forward.

Mathematically, we refer to this trade-off between over- and underfitting as the bias-variance trade-off, illustrated by the expression below that describes mean-squared error (MSE):

Test MSE ≈ Variance of the Function + Bias of the Function^2

However, rather than focusing on the statistical proof, let's look at this conceptually:

Test MSE is a measurement of how much a model's predictions deviate from reality. In other words, if a model perfectly predicts the market, the test MSE would be 0. It's every investor's goal to minimize test MSE. The two major contributors to test MSE are variance of the function and the square of the bias of the function.

Variance of the function is the error generated due to large changes in the underlying model when it is fed new information. A model with too many independent variables is likely to have a high function variance because each independent variable becomes sensitive to new data and information.

Bias of the function^2 is the error generated by the modeler approximating or simplifying a relationship. For example, if a modeler assumes there is a linear relationship between two stocks, but the relationship is nonlinear, the modeler has introduced bias into the model by making a simplifying assumption.

As you can probably deduce, there is a relationship between over- and underfitting and variance and bias. A model with high variance and low bias will become overfitted to the data, and a model with low variance and high bias will become underfitted to the data. The trade-off between variance and bias comes because they often conflict with each other. High variance means fewer assumptions are being made about the model, because the model is using all the historical data to create it. High bias means many assumptions are being made about the model, making it less likely to be impacted by new data or information.

CURRENT STATE OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN FINANCE

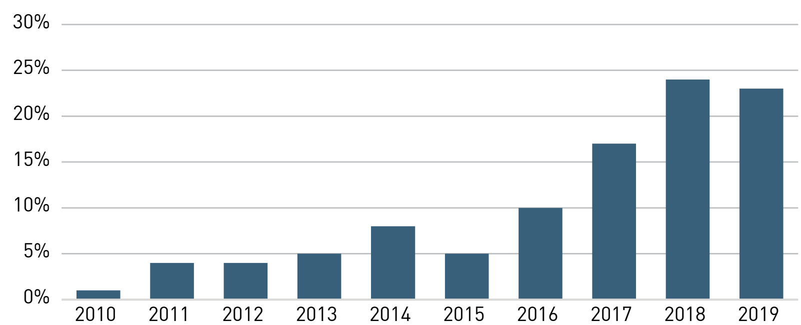
The early 19th-century German philosopher Arthur Schopenhauer wrote: "All truth passes through three stages. First, it is ridiculed. Second, it is violently opposed. Third, it is accepted as being self-evident."

I would guess that AI has reached the "self-evident" stage in the short term. Long term, however, I imagine the court of public opinion will oscillate, cycling among the three stages for some time. Indeed, topics in the financial industry often move among these stages, e.g., liquid alternatives, indexing, smart-beta factors, exchange-traded funds, and environmental, social, and governance investing (ESG).

Regardless, it's the stages of ridicule and violent opposition that hold the most opportunity for investors. Once someone identifies a way to outperform peers, make money, or raise assets, it becomes trickier for the customer. In the case of machine learning or artificial intelligence, how can the client or intermediary know they're getting the real deal? Are we sure this isn't a marketing

Figure 1

PROPORTION OF SYSTEMATIC HEDGE FUND LAUNCHES THAT USE ARTIFICAL INTELLIGENCE AND MACHINE LEARNING

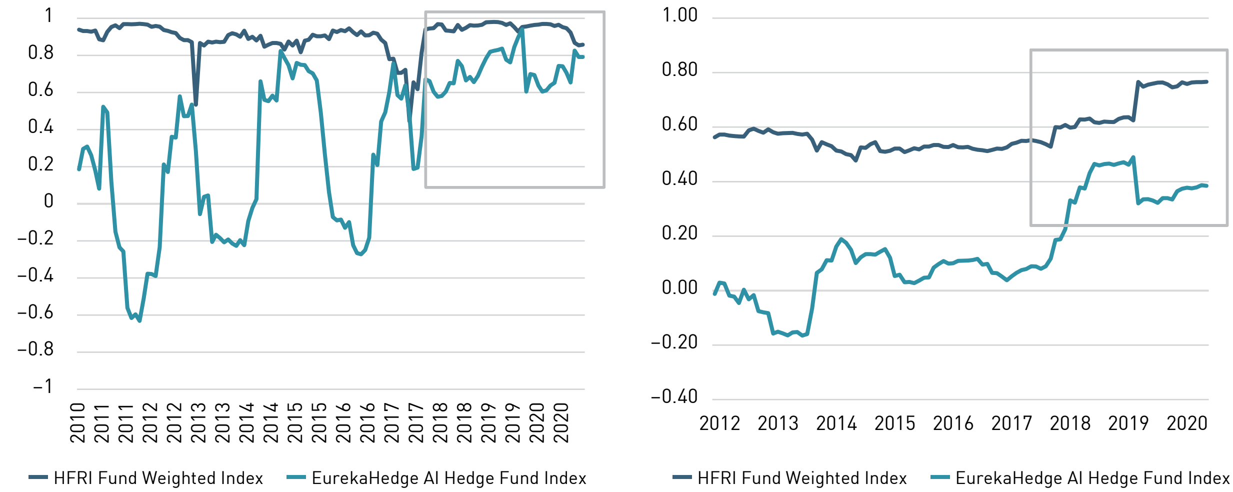


Source: Preqin 2020 Global Hedge Fund Report

© 2021 Investments & Wealth Institute. Reprinted with permission. All rights reserved.

Figure 2

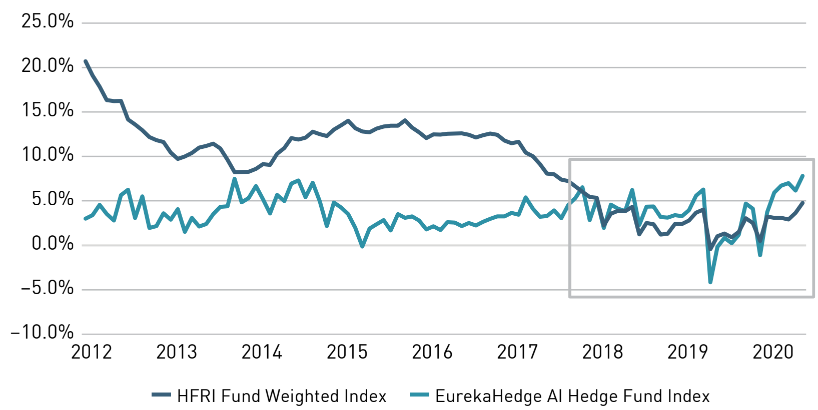
RISING CORRELATION AND BETA OF AI-DRIVEN HEDGE FUNDS TO A GLOBAL 60/40



Source: Author's calculations (monthly rolling one-year correlation against a 60-percent ACWI/40-percent Barclays Aggregate Bond), HFRI, EurekaHedge

Figure 3

CONVERGENCE OF ALPHA BETWEEN AI-DRIVEN AND BROADER HEDGE FUND COMPLEX



Source: Author's calculations (calculated using a three-year Jensen's alpha relative to a monthly-rebalanced portfolio of 60-percent MSCI ACWI/40-percent Bloomberg Barclays Aggregate Bond Index), HFRI, EurekaHedge

ploy to raise assets? In ESG, we often refer to this as "greenwashing." When it comes to data science, I call it "machinewashing."

Hedge fund managers seem to have been the early adopters of these tools; hedge funds that use artificial intelligence and machine learning have continued to grow to become a meaningful proportion of the systematic hedge fund industry, according to Preqin data (see figure 1).

But recent empirical research has concluded that machines aren't the silver bullet to generating alpha. With so much noise in our financial markets, algorithms run the risk of finding meaning where there is none (Rasekhschaffe and Jones 2019). As demand increases for these strategies and competition takes hold, it will be interesting to see what happens to performance. After all, one of the many criticisms of traditional hedge fund strategies is their high and rising correlation to traditional publicly listed equities and fixed income without the high absolute returns. Will that hold true for AI hedge funds eventually? If these strategies all use the same datasets, is there a risk of crowding?

EXAMINING PERFORMANCE

The correlations of AI-driven hedge funds historically have been quite dynamic, but recently the correlations and betas of these strategies relative to a global 60/40 portfolio have converged to the broader complex (see figure 2).

Even more interesting is the convergence of the alpha generated by these strategies (see figure 3).

There are, of course, many ways to measure alpha and beta(s),[1] and I'm not trying to make any forward-looking statements about the efficacy of any individual strategy. Filbeck and Kazemi (2020) showed that any alternative investment index can be misleading due to the underlying diversification of the strategies. Figures 2 and 3 show how much more competitive the space has become in a short amount of time.

LONG-TERM VALUE PROPOSITION OF MACHINE LEARNING AND DATA

AI's value proposition isn't based entirely on its ability to generate alpha. Some of its most powerful applications focus on everything but alpha generation. AI and ML can be used to prioritize longer-term objectives such as capital preservation, operational alpha, i.e.,

© 2021 Investments & Wealth Institute. Reprinted with permission. All rights reserved.

eliminating internal inefficiencies, and enhanced communication and culture.

DUE DILIGENCE: AVOID MACHINEWASHING

Conducting due diligence on quantitative managers, and those that heavily use artificial intelligence as part of their processes, can be much more difficult than conducting due diligence on traditional managers. How do you gain insight into a black box when the content of that black box is constantly changing? In my view, qualitative analysis of a quantitative approach is likely the best approach. Sure, you need the quantitative foundations to understand the nuts and bolts, but it always comes down to asking better questions. The same way we try to avoid greenwashing, how do we avoid machinewashing our portfolios?

DETERMINE THE VALUE-ADD

The thought of AI or ML techniques used for active management conjures images of robots at the helm, making all the decisions. This is extremely rare, however, and potentially not even in existence yet. So, as an investor, you must determine (1) where these algorithms are being used and (2) whether it makes sense for algorithms to be used in the strategy.

For example, Bartram et al. (2021) find that machine learning can be applied to different steps of the portfolio management process, but each step comes with its own set of considerations. These include the following considerations described below.

Signal generation

By definition, managers want to use machine learning methods that are predictive. Although there are many issues to consider at the individual algorithm level, investors should focus on the following two big issues during the due diligence process:

Data sourcing. It is important to understand where algorithms are obtaining data in order to generate insights. For example, natural language processing can analyze transcripts of regulatory filings and earnings calls, social media, and news articles, but not everything leads to an actionable signal.

Transparency versus predictiveness. Many of the most predictive algorithms are the least transparent. This creates a trade-off between attribution of results and predictive power of the process. This can be especially challenging, and it should be questioned and analyzed, especially if the strategy requires high turnover due to short-term momentum signals.

Portfolio construction

Because machine learning techniques can identify nonlinear relationships and consume large datasets, using them can help investors create more-efficient portfolios and maximize risk-adjusted returns.

Mean-variance optimization. Many of these algorithms don't use mean-variance optimization (MVO) as part of the process. This may end up being a better approach than MVO, which is very sensitive to inputs, but the way portfolios are constructed may not be as intuitive. Investors should ask questions about the portfolio construction process and determine if the approach is viable.

Passive or index replication. In the long-only space, indexing has become a popular approach used by investors to gain cheap market exposure. Although machine learning often is associated with active management, it also can be used for index-tracking products. Rather than purchasing every single stock in an index, machine learning could be used to build a more concentrated portfolio that provides the risk and return profile of an index but with lower transaction fees and turnover.

Trade execution

Finally, machine learning can add a lot of value through strategy and trade execution. Algorithms can learn the best times to execute trades on securities to minimize market impact that otherwise would eat into returns. Like other parts of the investment process, however, algorithms must be tested in and out of sample to ensure success. Investors should ask questions around how these managers test and implement execution algorithms and, most importantly, how the algorithms themselves evolve over time to account for dynamic markets.

These are three broad examples of where machine learning can be used in an investment process. Some strategies require only one or two of these value-adds. It's important for investors conducting due diligence to understand where and how these algorithms are being implemented. Not every problem can be solved with a computer.

> *It's important for investors conducting due diligence to understand where and how these algorithms are being implemented. Not every problem can be solved with a computer.*

FOCUS ON THE ORGANIZATIONAL CULTURE

One of the biggest differences between technology companies and traditional financial companies is culture (Harrington 2018). If stereotypes are a guide, technology companies are much more open and collaborative and financial organizations operate in siloes. To ensure a fund is successful, you must approach culture checks and qualitative due diligence factors differently. Insights are shared openly throughout the investment team, a stark contrast to the culture of finance, which prioritizes competitive edge even relative to ones' colleagues. The funds that fail are the ones that don't embrace this collaborative way of work (López de Prado 2018).

© 2021 Investments & Wealth Institute. Reprinted with permission. All rights reserved.

But investors shouldn't assume that investment management organizations will transform into technology companies overnight. And some strategies may change slower than others. Strategies that invest in exchange-traded securities may have an easier time making the transition culturally and technologically than those that operate in the private markets.

Astebro (2021) notes that, although private equity general partners have rapidly adopted artificial intelligence as part of their processes, some cultural barriers still exist. He notes that private equity and venture capital general partners traditionally do not rely on digital tools; they spend most of their time in meetings and building relationships with the founders and management of prospective portfolio companies. Because these companies aren't publicly traded, and data isn't collected during these relationship-building moments, AI and ML historically haven't been used on the front end of the process.

BE SKEPTICAL ABOUT IMPRESSIVE BACKTESTS

I've never seen a bad backtest. So it's important to find ways to test the underlying assumptions of the strategy out-of-sample and, better yet, use a process that doesn't suffer from look-ahead bias. By mitigating look-ahead bias, you can travel back in time and ignore what you already know has occurred. A good example would be a backtest that takes place during the tech bubble. In hindsight, it's easy to see that defensive value stocks would outperform expensive tech stocks, and therefore it's easy to create a backtest that simulates making the right moves. However, what if you were building an algorithm in 1997, with no view into the future? Would your algorithm buy the right securities and, more importantly, could you stick with it all the way through?

Failing to mitigate look-ahead bias will lead to a model that may not adapt to the markets in practice because it is overfitting. How many of these AI strategies were able to adapt to the COVID-19 pandemic? If you're simply relying on past relationships in capital markets, a catastrophic event such as a global pandemic might completely disrupt the algorithm altogether.

Risk management is important in investment process, and there is no exception for a process that uses algorithms. In other words, what would make this process blow up? The driver could be internal operations, or the market could break the very algorithms that are trying to take advantage of it. The quant quake of August 2007, for example, showed everyone what happens when multiple funds following similar strategies start deleveraging and unwinding positions. In early 2021, r/WallStreetBets showed everyone how strong some hedge fund risk management programs were when retail investors come together to trade in the opposite direction of the big institutions.

To mitigate the risk of overfitting a strategy and to increase the probability of out-of-sample success, Rasekhschaffe and Jones (2019) suggest diversifying and combining multiple models together, such as the following:

- Using more than one algorithm to forecast the same security
- Using multiple time periods to test models out-of-sample
- Using unique forecast styles for different horizons, e.g., technical analysis for short-term signals versus fundamental analysis for long-term signals

Using a combination of models allows the manager to diversify away from a single algorithm that drives decisions. However, investors performing due diligence on managers may want to ask questions about the individual algorithms and how those algorithms fit together. Multiple models can help diversify the strategy's drivers of performance, but investors will want to be wary of too many meaningless signals and algorithms to the point that, in practice, the strategy doesn't execute properly.

PUTTING IT ALL TOGETHER

We say past performance is no indication of future results, and this is especially true when trying to analyze a strategy that is quite literally learning as time goes on. The confusion, and lack of education,[2] can cause us as investment professionals to be afraid to ask the necessary questions to determine the efficacy of a strategy. The world is messy, but let's not machinewash it.

Aaron Filbeck, CFA®, CAIA®, CIPM, FDP, is director of global content development at CAIA Association, where he provides vision, leadership, and strategic direction for content agenda and member education programs. He chairs CAIA Association's global editorial board and sits on the organization's content and exams leadership team. He earned a BS with distinction in finance and a master of finance from Penn State University. Contact him at afilbeck@caia.org.

ENDNOTES

1. Because this is for illustrative purposes, I did not use style factors.
2. Financial Data Professional (FDP) Institute, https://fdpinstitute.org.

REFERENCES

Astebro, Thomas. 2021. An Inside Peak at AI Use in Private Equity. *Journal of Financial Data Science* 3, no. 3 (summer): 97–107.

Bartram, Söhnke M., Jürgen Branke, Giuliano De Rossi, and Mehrshad Motahari. 2021. Machine Learning for Active Portfolio Management. *Journal of Financial Data Science* 3, no. 3 (summer): 9–30.

Deloitte. 2019. Artificial Intelligence: The Next Frontier for Investment Management Firms. https://www2.deloitte.com/us/en/pages/financial-services/articles/ai-next-frontier-in-investment-management.html.

Filbeck, Aaron, and Hossein Kazemi. 2020. An Alternative View of Manager Selection Risk. CAIA Association (January 16). https://caia.org/blog/2020/01/16/an-alternative-view-of-manager-selection-risk.

Harrington, Cynthia. 2018. For Fintech Career Switchers, the Hardest Skills to Master May Be the Soft Ones. CFA Institute (August 8). https://www.cfainstitute.org/en/research/cfa-magazine/2018/for-fintech-career-switchers-the-hardest-skills-to-master-may-be-the-soft-ones.

López de Prado, Marcos. 2018. The 10 Reasons Most Machine Learning Funds Fail. *Journal of Portfolio Management* 44, no. 6: 120–133.

Rasekhschaffe, Keywan, and Robert Jones. 2019. Machine Learning for Stock Selection. *Financial Analysts Journal* 75, no. 3 (third quarter): 70–88.

© 2021 Investments & Wealth Institute. Reprinted with permission. All rights reserved.

5619 DTC Parkway, Suite 500
Greenwood Village, CO 80111
Phone: +1 303-770-3377
Fax: +1 303-770-1812
www.investmentsandwealth.org

© 2021 Investments & Wealth Institute®. Reprinted with permission. All rights reserved.

INVESTMENTS & WEALTH INSTITUTE® is a registered mark of Investment Management Consultants Association Inc. doing business as Investments & Wealth Institute. CIMA®, CERTIFIED INVESTMENT MANAGEMENT ANALYST®, CIMC®, CPWA®, CERTIFIED PRIVATE WEALTH ADVISOR®, RMA®, and RETIREMENT MANAGEMENT ADVISOR® are registered certification marks of Investment Management Consultants Association Inc. doing business as Investments & Wealth Institute.