

Challenges Of Widespread Artificial Intelligence in Financial Institutions

Yogith Loganathan & Abhinav Choudhry

Received March 15, 2024

Accepted June 10, 2024

Electronic access July 15, 2024

The following paper covers the possible implications of the growing role of artificial intelligence (AI) and machine learning (ML) in the financial sector such as *Strategic Convergence*, *The Black Box*, *Data Collection and Embedded Bias*, *Talent Acquisition*, and *Costs and Future Financial Implications*. It acknowledges the efficiency this technology can bring, but also advances the discussion of the challenges it poses. The paper explores the need for robust regulatory frameworks, transparency, and the resources necessary to ensure responsible and ethical AI/ML advancements in the sector. The goal of this research was to synthesize the most apparent risks of AI in finance into one paper as well as offer our own unique insight into a plausible challenge.

Keywords: Black box, Talent acquisition, Embedded Bias, Big data analytics, Herding behavior

Introduction

The unprecedented volume and variety of data available in the digital age provide a fertile ground for financial institutions to harness the power of big data analytics. Surveys of firms already show that 77% of respondents believe that AI will be significant in their businesses in the coming two years¹, with companies like McKinsey estimating that the potential value of AI in banking will reach \$1 trillion².

But before exploring the role of AI/ML in finance, it's worth looking back at the origins of this technology starting in the mid-20th century. Turing's imitation game from the 1950s set the standard for measuring a machine's ability to exhibit human intelligence. During this time, AI was described as a "rule-based system," a series of if-then-else statements, incapable of learning over time³. Then, with the entry of big data and advances in storage technology, Machine Learning, a subset of AI, began to gain traction.

Machine learning uses statistical tools to learn from data to identify patterns, perform specific tasks, and provide accurate results⁴. Popular examples include search engines, malware filtering, and online shopping recommendations.

Then, in 2006, Deep Learning, a subset of ML, emerged and utilized neural networks made up of node layers. Each node behaves similarly to an artificial neuron and is linked to the rest; whenever a node's output is above a threshold value, the data is processed in the network's next layer⁴. Popular examples include large language models like ChatGPT.

The relationship between these disciplines is analogous to Russian dolls: AI makes up the outer layer, then ML, then

DL. Each layer has been or is currently utilized in finance and brings advances in productivity, accuracy, scalability, processing speeds, and efficiency⁵. Though the list of benefits seems endless, there are significant challenges that come with

the mass adoption of AI/ML models in finance. For instance, the obscurity of AI/ML algorithms may make it difficult for institutions to explain their reasoning behind certain deals—this is especially important if the client experiences major financial losses. Additionally, if ignored, the algorithms may be unfair by nature if the training data is unrepresentative of the entire population. This paper covers five potential obstacles with AI/ML in finance that we should acknowledge early so we can combat complications:

1. Strategic Convergence
2. The Black Box
3. Data Collection and Embedded Bias
4. Talent Acquisition
5. Costs and Future Financial Implications

Methodology

The literature featured in the paper were discovered via a systematic search through platforms such as google scholar as well as publications from well-recognized financial institutions. Initially, broad search topics like "AI in Finance" were used and the available literature was sifted through and chosen based on

whether they provided plausible challenges to the proliferation of AI in financial institutions. From the chosen papers, the challenges that appeared the most were synthesized into a list. Additionally, one other hypothetical challenge was added to the list based on current estimates of AI algorithms in finance. Though there is literature from four or five years in the past, they were not excluded because they were used primarily to contextualize the development of AI in finance or define certain terms.

Strategic Convergence

Start by imagining a vibrant city teeming with activity: cars navigating through traffic, cyclists weaving through lanes, and herds of joggers trying to get their steps in. Picture the path each person takes to reach their destination. Whether driving, biking, or walking, each person has their unique route, introducing variety in arrival times. Now, consider a scenario where all bankers collectively opt for a narrow alleyway shortcut to reach their offices. The once coveted 15-minute walk becomes congested, eroding the initial efficiency and destroying the variety in their arrival times.

The mass utilization of AI/ML could follow a similar pattern: the homogenization of investing strategies, like the intersection of work routes, could destroy a majority of the variety in the markets and result in a disruption of financial stability.

Financial algorithms' effect on the stability of the markets can be expressed in three broad channels⁶. First, herding, the convergence of the algorithms' decisions—in our analogy, the banker's alleyway shortcut. Second, network interconnectedness, the emergence of a dependency on some central infrastructure or model. In our "teeming city," this could be all the taxis using the same GPS, and by effect the same traffic shortcuts. Third, regulatory gaps, problems in policy that lead to build-ups or seepage of systemic risk into the broader economy.

Before diving deeper into the first two channels, it's worth discussing regulatory gaps. They're often the byproduct of a quickly evolving innovation outpacing the current regulatory framework. The 2008 financial crisis is a prime example of this. Advancements in collateralized debt obligation (CDOs), credit default swaps (CDS), and subprime mortgage underwriting standards outpaced the legal constraints and blurred the risks⁶. Now, with DL models becoming gradually more autonomous it's clear to see a future where a majority of trade decisions are fully automated. When this future inevitability occurs, the volatile nature of AI/ML trades will likely make the public more vulnerable to financial losses. So, instead of continually imposing laws and restrictions as obstacles to prevent models from surpassing regulatory frameworks, we need to establish clear-cut AI regulations before a crisis unfolds.

Besides regulatory gaps, financial algorithms can affect the stability of the market via network interconnectedness and herd-

ing behaviors. Often, network interconnectedness will result in a herding behavior that will inexplicitly overfit data and result in a weak foundation for the market. Picture a world where all financial institutions have an AI/ML model based on the same available core data. If a firm buys a share of stock X (advised by the model), there's a fair probability that all the other similarly rooted models will also buy stock X⁷.

When you look around, everyone starts to buy stock X and it starts to go up. Now everyone is ecstatic, so they continue to do whatever the algorithm says and embed more trust into the model. This continues and now there's a set of oligopoly companies invested by a monopoly of firms. At this point, there are a lot of problems. As one of the critical components of the market, company X can decide to raise prices because at this point all of their competitors are already bankrupt. Additionally, if company X already runs that sector of the market, there's no shot any newcomers with innovative ideas can make it. If those aren't reasons enough, one day if one of these oligopoly companies fail or the algorithm says they're going to fail, all the firms will pull out and the market will implode.

This stresses the dangers of not only strategic convergence but also fully autonomous models. If we rely on AI to make all of our trades, the aspect of human variation will be lost, and the necessary mistakes from human error will be absent⁷. Without this, the market will be void of complexity.

However, this scenario where strategies converge and a majority of the firms depend on the same core AI/ML model has its flaws. In reality, data will probably become more safeguarded and expensive. It could be stratified by factors like quality and type, introducing a whole new level of depth into these AI models. Instead of leveling out all of the existing firms, the introduction of these models may only scale up the disparities. For instance, if firm X has historically been more successful than firm Y, these algorithms may greatly increase the productivity of firm X while firm Y functions relatively the same. This could stem from the inability of firms being able to acquire the same quality and quantity of data as their historically lucrative competitor. Ultimately, to mitigate the possibility of strategic convergence, we must create and regulate AI-related policies that will limit systemic risk and preserve healthy, fair competition.

The Black Box

The increased need, or ability, to create nuanced neural networks to synthesize data brings with it difficult technical problems. The black box, one such obstacle, describes an opaque decision system where the results have little to no explanation. The characteristics of black box models can expose organizations to vulnerabilities like biased data and decision errors which, because of the nature of the black box, will long be undetected⁸.

Opening the organization up to potential losses in capital, resources, and overall satisfaction. Ultimately, the black box raises the question of explainability and complexity, two critical factors that need to be meticulously balanced to ensure that AI/ML algorithms are as effective as possible.

There are several reasons that add to the obscure nature of AI/ML models: they are difficult to interpret, may have unknown input signals, and are often a culmination of a series of models rather than a single independent model⁹. Though there are certain scenarios where a lack of explainability can be beneficial, for instance in protecting your algorithm from outside manipulation, the drawbacks outweigh the occasional benefits. The issue of explainability, apart from the financial risks it creates also brings up serious discussions about ethics. In the case where a client loses significant capital due to a misstep from the algorithm, it will be difficult for the institution to explain how the error happened. At that level of obscurity, the trust between the clients and firms gets destroyed. So in consideration of these issues, two perspectives arise: one arguing for an accurate black box while the other settling for a slightly less complex “glass box”— a model whose inner workings are transparent and understandable to humans¹⁰.

A majority of the time there’s a trade-off between a model’s flexibility, its capacity to approximate different functions and parameters, and its explainability⁹. AI/ML algorithms are highly flexible but lack interpretability, the opposite of their linear model counterparts. For example, neural network-based algorithms, like brains, have different layers of reasoning, and in each layer, differently weighted nodes contribute to the final decision. And as the layers gradually increase, the overall explainability of the algorithm begins to decrease.

The other side argues that the accuracy–interpretability trade-off is a fallacy and more interpretable models often become more accurate, calling for interpretable linear models when possible over AI/ML models¹⁰. The Explainable Machine Learning Challenge in 2018 explored this: while teams were told to create a black box model and explain it for a given scenario, one group stood out and created an interpretable model, or a glass box, with less than a 1% error bound. Similar cases in criminal justice and medicine followed the same pattern¹⁰. So while we have the capacity to construct AI/ML algorithms, there are certain scenarios where it’s more advantageous to stick to more interpretable and traditional models.

Ultimately, the black box is a hurdle that AI/ML algorithms need to overcome as they gradually become more pervasive in finance. Until then, developing “explainable AI” or XAI models that utilize more interpretable algorithms will continue to be safer for both institutions and their clients.

Data Collection and Embedded Bias

As AI/ML algorithms become more pervasive, financial institutions increasingly rely on vast amounts of data to train these models. Consequently, firms are faced with the difficulties of finding cost-effective and quality data to ensure the algorithms are embedded with little to no bias⁵. Embedded bias, defined as an algorithm’s tendency to systematically discriminate against certain individuals or groups, is dangerous in the financial sector where public trust is essential⁹. As a result, it’s critical that companies acknowledge the underlying risks of AI before applying it; this way bias mitigation efforts can be proactive rather than reactive.

Applying effective financial algorithms requires a vast spectrum of high-quality and often expensive data, narrowing the number of firms that can create said algorithms. Additionally, the expenses for new infrastructure and data preprocessing that are a by-product of these large volumes of data might be “hidden costs” in the acquisition of AI/ML algorithms⁵. In most applications, the AI flywheel effect, the ability for AI/ML algorithms to improve themselves after they are adopted (since new data is acquired), minimizes the initial cost of acquiring data at a revenue trade-off¹¹. Essentially, companies offer their low-quality algorithm for cheaper in return for the model buyers to utilize the flywheel effect and gradually improve the model’s accuracy. During that time they experience a loss in revenue, which depending on the context may be more affordable or costlier than acquiring a large volume of data. For example, the founders of the startup Blue River Technology, a company focused on using AI to distinguish weeds from crops, established their first data manually¹². This resulted in a fairly inaccurate algorithm, but after the adoption from early users, the company had a more expansive pool of data that they used to fine-tune their model. Due to its success and contribution to pesticide optimization in farming, the company sold for over 300 million in 2017¹². However, the AI flywheel effect becomes a lot more trivial when it comes to finance. Since most firms are going to use the algorithm to generate profits off the market, using a low-performing model can increase risks associated with trading and bleed into capital.

Apart from the costs of acquiring the data, the data itself may contribute to the challenges of AI/ML algorithms. Embedded bias can occur from data collection in two ways:

1. Unrepresentative/incomplete data⁹. For example, the tendency for predictive algorithms to favor the best-represented party in the training data since there’s less associated uncertainty linked to the predictions.
2. Prejudiced data. For example, Amazon’s internal recruiting system favored men over women because of historical company hiring decisions¹³.

Both scenarios are often referred to as “garbage in, garbage out,” the concept that a flawed input will produce a flawed output. So to avoid bias stemming from data collection, it’s critical to implement measures that ensure the training data is diverse and representative of the entire population. Additionally by promoting data anonymization we can remove personally identifiable information (PII), effectively preventing the model from learning and reinforcing biases related to identity.

However, this is not the only way biases may occur. For a more complete understanding, we have to take into account human and systemic biases; though they are not isolated to AI/ML, acknowledging them can help mitigate their effects. Human bias is related to how we fill in missing information given data, such as associating race with loan approval or residency with criminal behavior¹⁴. Systemic bias is described as an institution operating in a manner that disadvantages social groups¹⁴. Both biases can play roles during data collection, algorithm training, and algorithm application, so it’s critical to be aware during most phases of constructing the model.

As AI/ML models begin to become more involved in decisions, it’s critical they’re provided with a large volume of accurate data. However, this is difficult for most companies and they may resort to low-quality data, resulting in a series of biases from the data itself. To combat this, companies should shift their bias mitigation efforts to be more proactive and there ought to be greater regulation regarding data management.

Talent Acquisition

One of the primary challenges associated with the widespread implementation of AI/ML models in finance is the shortage of skilled personnel. A recent EY survey conducted in collaboration with MIT revealed that nearly half (45%) of senior business and technology leaders believed their organizations lacked the necessary talent for AI implementation. Companies in other sectors also reported difficulties in finding suitable candidates for AI-related work: Fujitsu, a Japanese IT company that hires globally, stated that navigating the AI talent shortage was challenging to overcome¹⁵. In light of these challenges, it’s crucial to explore the specific talent that finance is missing and potential solutions that are taking place.

The inability to attract and retain top AI talent can lead to delays in project timelines and a reliance on outdated technologies. As a result, these organizations will fall behind their competitors in terms of productivity and innovation, potentially losing market share and profitability as they fail to adapt to the rapidly evolving landscape. To mitigate these disruptions, there are two types of roles that financial institutions can fill: data scientists and translators¹⁶.

Data scientists, traditionally quantitative analysts, should be comfortable with designing, training, and deploying AI/ML models¹⁶. They’ll be critical in introducing the bulk of the AI

Revolution into the company and should be prepared to maintain and calibrate the model. Currently, work in AI has a high barrier to entry where higher education is a common prerequisite. Consequently, there are fewer data scientists equipped with the necessary skills, and when they do exist, they are quickly hired into the tech sector. NYU Professor Scott Galloway estimated that as many as four out of five Ph.D. students were beginning to be recruited into the tech sector rather than entering finance or pursuing a traditional career in academia¹⁵. Possible explanations for this shift of skilled personnel stem from the better hours, compensation, and overall work-life balance that tech offers. This leaves financial institutions struggling to find experienced data scientists, despite the availability of training, unlike in the case of translators, where the necessary education has only recently made strides in the right direction.

Translators are employees who possess enough technical knowledge to understand how the model functions but are also familiar with the business and financial side of the institution¹⁶. They’ll serve as the AI Revolution’s bridge to the remainder of the employees and should be comfortable in informing high-level management on how to use the model to make knowledgeable strategies. The shortage of translators stems from the industry’s demand for a niche pool of applicants who are educated in both finance and have experience with advanced AI/ML algorithms. There are two popular ways organizations try to gain translators: they either train their existing staff in AI/ML or they recruit graduates with experience in both fields.

Amazon launched an AI education program designed to help workers develop the key skills that they would need to understand AI models⁷. Though this isn’t specifically in finance, the idea of ‘translators’ still exists in this application.

Man Group, one of the largest hedge funds with around 161.2 billion dollars under management, partnered up with Oxford to create the Oxford-Man Institute of Quantitative Finance in 2007¹⁷. Since then, the institution has graduates experienced in machine learning, financial theory, and mathematics; all critical knowledge for a finance-based translator.

To address the shortage of skilled personnel in finance, institutions must adopt a more proactive approach. They should strive to educate their current staff and also start investing in

educational institutions that aim to equip their students with the necessary multi-disciplinary skills. While this is undoubtedly challenging, especially for institutions with budget constraints, employing these methods, coupled with a quicker reaction speed, is crucial to attracting graduates away from tech into finance. However, as AI becomes more popular in finance, students may need to pursue both finance and AI to remain competitive applicants.

Costs and Future Financial Implications

AI/ML models have many associated costs ranging from their hardware and software components to the computational power that's required to train them. Beyond these critical costs, hidden fees from data storage and transferring also exist¹⁸. However, as these models continue to be implemented, more information about their life span and proper maintenance behaviors are discovered.

A recent study conducted by Harvard, MIT, The University of Monterrey, and Cambridge revealed that 91% AI algorithms experience temporal model degradation, or AI aging¹⁹. This challenge stems from organizations training their models to reach a specific quality but subsequently failing to retrain the models post-production to maintain that level of performance²⁰. The collegiate study was centered around four standard models and their degradation post-training:

1. Linear Regression (RV)
2. Random Forest (RF)
3. Gradient Boosting (XG)
4. Neural Network (NN)

The graph reveals that neural network and linear regression models are the most susceptible to degradation over time, with explosive degradation showing effects around the one-year mark. With this insight, companies should aim to retrain their algorithms annually to prevent complications that could hinder the model's accuracy.

To gauge the resources needed to maintain one of these models, we can perform a simplified cost analysis on an existing NN. One of the industry's largest language models, BloombergGPT, is specifically trained to support natural language processing (NLP) tasks within finance and contains 50-billion parameters²¹. According to Cerebras' AI Model Studio (Table. 1), a model around that size would likely take around 2-2.5 million dollars to train²².

Though this is less than 1% of Bloomberg's revenue from prior years, this budget may be unrealistic for smaller AI hedge funds. The costs of retraining a model annually coupled with the required hardware and software may not be affordable for many new/small hedge funds. This becomes further complicated if the model in question has yet to prove itself in the market.

Nevertheless, the costs of retraining a model annually compared to retraining it far into its degradation cycle is consistently more cost-effective. Possible methods to retrain a model well before degradation are setting alerts, creating dashboards of the model's health, and running rich model diagnostics.

Despite the current obstacles preventing AI/ML models from becoming accessible to a majority of the industry, once software and hardware costs start to be reduced and AI implementation

starts becoming easier, competition among institutions for the leading model will begin. This AI 'arms race' will have institutions continuously trying to gain an edge over each other, consequently, more and more of their budgets will be spent on improving their models.

This, in return, may cut into their profits and limit the resources they allocate to other branches of the company.

Ultimately, implementing AI/ML models comes with a range of unavoidable costs from hardware, software, computational power, and hidden fees associated with data management.

Then after deployment, costs associated with temporal model degradation become more significant, but with the current available data to the public it's difficult to estimate. However, to combat this, organizations can retrain their models regularly post-production. Though this can be fairly costly, recalibrating the algorithm to a changing environment ensures greater accuracy and mitigates the loss of profit due to error.

Conclusion

Despite the possible implications of AI/ML in finance, the technology offers unarguable value in the industry. From increasing employee productivity and reducing repetitive back-office tasks to performing big data analytics, institutions have visibly benefited from these models.

However, it's worth it to err on the side of caution; by considering the previous challenges, the industry as a whole can benefit from having appropriate regulations before there is a risk of a major financial crisis driven by automated AI agents.

Regarding obstacles like explainability and embedded bias, policymakers should take steps to prioritize client safety since decisions that are both unexplainable and based on flawed data corrupt the relationship between clients and firms. For challenges built on the premise of widespread AI/ML model implementation, institutions should aim to diverge from over-relying on their algorithm. Regulations on data distribution could also prevent certain firms from monopolizing and mitigate the possibility of systemic risk. Lastly, companies struggling to acquire the necessary talent should become more proactive and fund institutions that teach the type of employee they're looking for. Otherwise, there will continue to be a surplus of capable graduates who lack the specific requirements that firms are looking for.

Ultimately, the success of the industry in the coming years will be heavily determined by the AI policies set today. However, we should continue to value human ingenuity: just as the bustling city relied on diverse paths for varied arrival times, the world of banking benefits from a diversity of strategies and approaches.

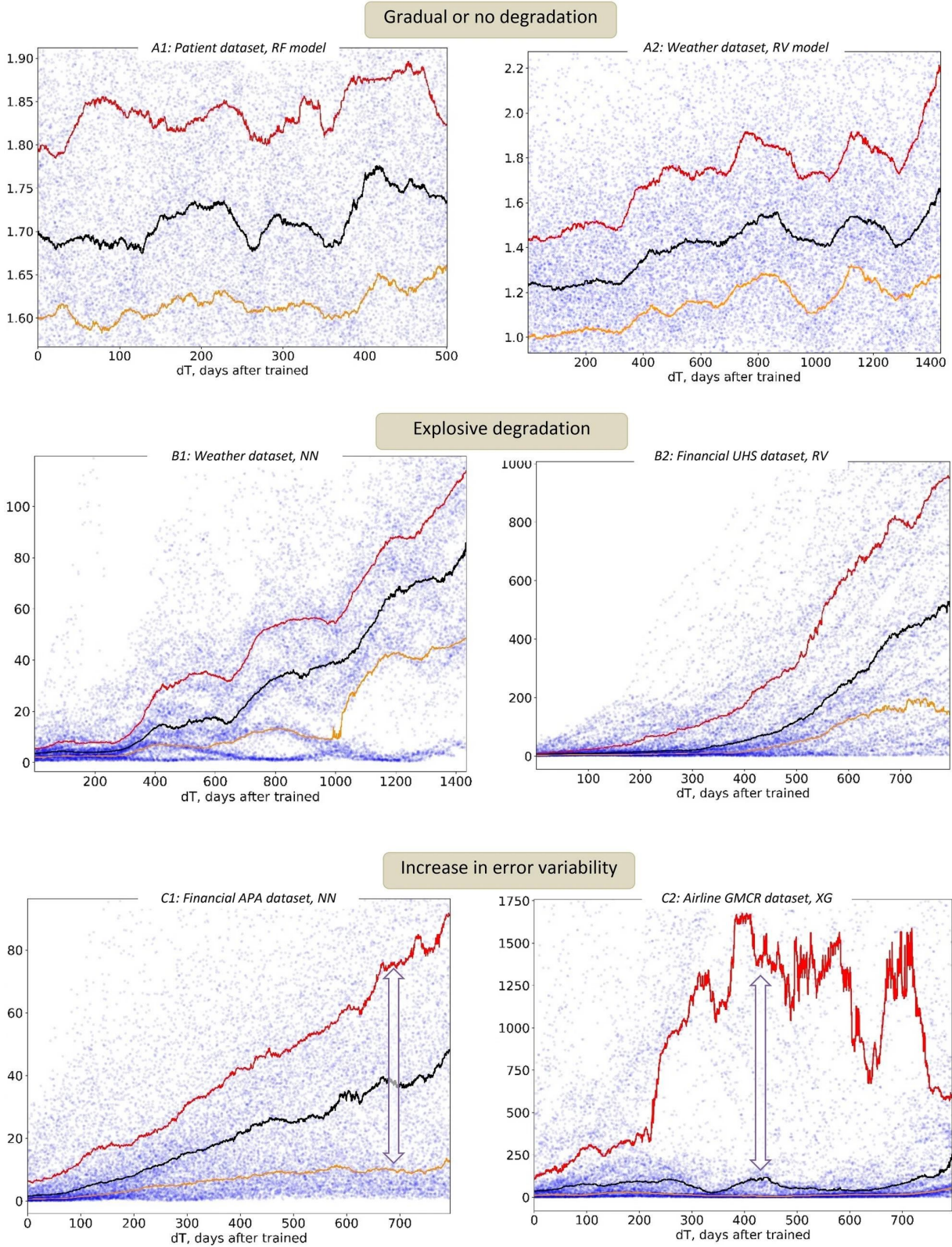


Fig. 1 Shows the relationship between degradation and the different models when they share the same dataset¹⁹.

Table 1: Shows the AI Model Studio’s cost for training a GPT model from scratch.

Model	Parameters	Tokens to Train to Chinchilla Point (B)	Cerebras Model Studio CS-2 Day to Train	Cerebras Model Studio Price to Train
GPT3-XL	1.3	26	0.4	\$2,500
GPT-J	6	120	8	\$45,000
GPT-3 6.7B	6.7	134	11	\$40,000
T-5 11B	11	34*	9	\$60,000
GPT-3 13B	13	260	39	\$150,000
GPT NeoX	20	400	47	\$525,000
GPT 70B	70	1,400	85	\$2,500,000
GPT 175B	175	3,500	Contact For Quote	Contact For Quote

* - T5 tokens to train from the original T5 paper. Chinchilla scaling laws not applicable.

References

- World Economic Forum, *Transforming Paradigms a Global AI in Financial Services Survey*, World economic forum technical report, 2020.
- S. Biswas *et al.*, *AI-bank of the Future: Can Banks Meet the AI Challenge?*, 2020, <https://www.mckinsey.com/industries/financial-services/our-insights/ai-bank-of-the-future-can-banks-meet-the-ai-challenge#>.
- A. Ashta and H. Herrmann, *Strategic Change*, 2021, **30**, 211–222.
- IBM Blog, *AI Vs. Machine Learning Vs. Deep Learning Vs. Neural Networks: What’s the Difference?*, 2023, <https://www.ibm.com/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>.
- A. R. Kunduru, *From Data Entry to Intelligence: Artificial Intelligence’s Impact on Financial System Workflows*, 2023.
- G. Gensler and L. Bailey, *Generative Artificial Intelligence in Finance: Risk Considerations*, 2023, <https://doi.org/10.2139/ssrn.3723132>.
- N. Jacobson, *The Case Against Financial Algorithms - Berkeley Haas*, 2023, <https://haas.berkeley.edu/undergrad/community/blog/posts/the-case-against-financial-algorithms>.
- J. Silberg and J. Manyika, *Notes From the AI Frontier: Tackling Bias in AI (and in Humans)*, Mckinsey global institute technical report, 2019.
- E. B. Boukherouaa *et al.*, *Powering the Digital Economy: Opportunities and Risks of Artificial Intelligence in Finance*, 024, 2021.
- C. Rudin and J. Radin, *Harvard Data Science Review*, 2019.
- H. Gürkan and F. D. Véricourt, *Contracting, Pricing, and Data Collection Under the AI Flywheel Effect*, 2020, <https://doi.org/10.2139/ssrn.3566894>.
- E. Trautman, *The virtuous cycle of ai products*, 2018, <https://www.eriktrautman.com>.
- K. Hao, *This Is How AI Bias Really Happens—And Why It’s so Hard to Fix*, 2019, <https://www.technologyreview.com/s/612876/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/>.
- C. Boutin, *There’s More to AI Bias Than Biased Data, NIST Report Highlights*, 2022, <https://www.nist.gov/news-events/news/2022/03/theres-more-ai-bias-biased-data-nist-report-highlights>.
- Global Data, *AI Faces Future “Talent Crunch”, Warns Fujitsu CTO*, 2021, https://www.globaldata.com/newsletter/details/ai-faces-future-talent-crunch-warns-fujitsu-cto_326726.
- V. Brozović, *APPLICATION OF ARTIFICIAL INTELLIGENCE IN THE SECTOR OF INVESTMENT FUNDS*, 2019, <https://repositorij.efzg.unizg.hr/islandora/object/efzg:2726>.
- A. Satariano and N. Kumar, *The Massive Hedge Fund Betting on AI*, 2017, <https://www.bloomberg.com/news/features/2017-09-27/the-massive-hedge-fund-betting-on-ai>.
- A. Ng, *AI Transformation Playbook*, 2018.
- D. Vela *et al.*, *Scientific Reports*, 2022, **12**, year.
- K. He, *91% of ML models degrade over time*, 2023, <https://www.fiddler.ai/blog/91-percent-of-ml-models-degrade-over-time#:~:text=A%20recent%20study%20by%20Harvard%2C%20MIT%2C%20The%20University,organizations%20using%20ML%20models%20to%20advance%20real-life%20applications>.
- Bloomberg L.P., *Introducing BloombergGPT, Bloomberg’s 50-billion Parameter Large Language Model, Purpose-built From Scratch for Finance*, 2023, <https://www.bloomberg.com/company/press/bloomberggpt-50-billion-parameter-llm-tuned-finance>.
- T. P. Morgan *et al.*, *Counting the Cost of Training Large Language Models*, 2023, <https://www.nextplatform.com/2022/12/01/counting-the-cost-of-training-large-language-models>.