The Myth and Reality of Financial Machine Learning

Marcos López de Prado
Lawrence Berkeley National Laboratory
Computational Research Division

February 9, 2018

Key Points

- In recent years, Machine Learning (ML) has been able to master tasks that until now only a few human experts could perform.
- Financial ML is a reality:
  - Some of the most successful hedge funds in history apply ML every day.
- However, myths about Financial ML have proliferated:
  1. The Sisyphus paradigm is applicable to ML
  2. ML is a black-box
  3. Generic ML solutions work in Finance (West → East)
  4. Traditional quants know about ML (West ← East)
  5. There are many ML Portfolio Managers

- The contents of this presentation are based on my book: ***Advances in Financial Machine Learning***, Wiley (2018)

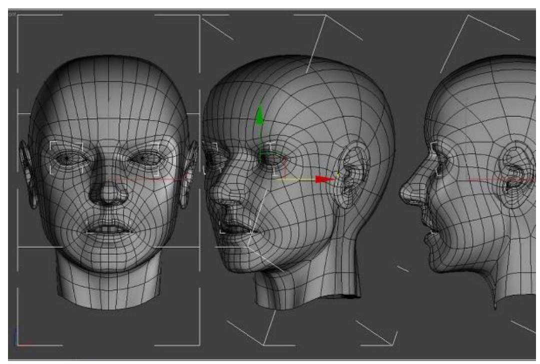
Investing in the ML Age

ML in Recent Years

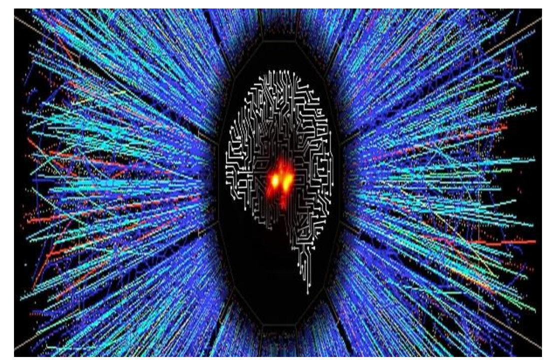
1997
Deep Blue defeats the
Chess World Champion

2011
Watson defeats 2
Jeopardy! champions

2016
AlphaGo defeats the
Go World Champion



2014
DeepFace recognizes faces
better than humans



2018
Berkeley Lab recognizes particle debris with 95% accuracy

Financial Research

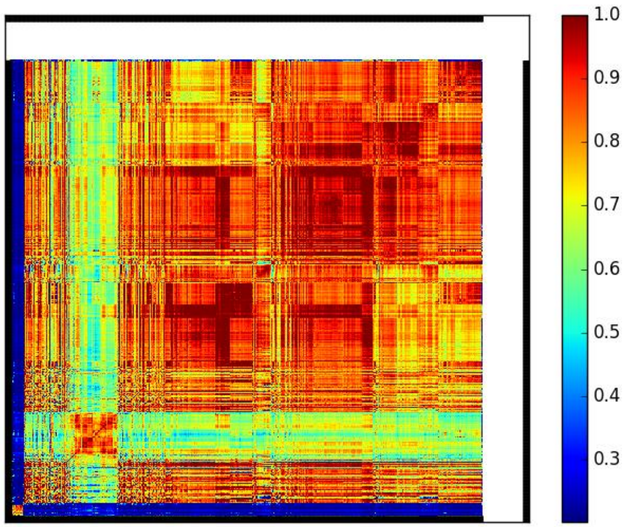
- The dismal state of 21st century financial research:
  - Story-telling prevails over objective data analysis
  - The curse of Econometrics and other 18th century mathematical tools
  - Linear theoretical models: CAPM, APT, Risk Premia, etc.
  - Multiple testing, selection bias, backtest overfitting
  - Factor investing: A few (3?) factors are well understood, however Harvey et al. [2015] show that *“most claimed research findings are likely false.”*

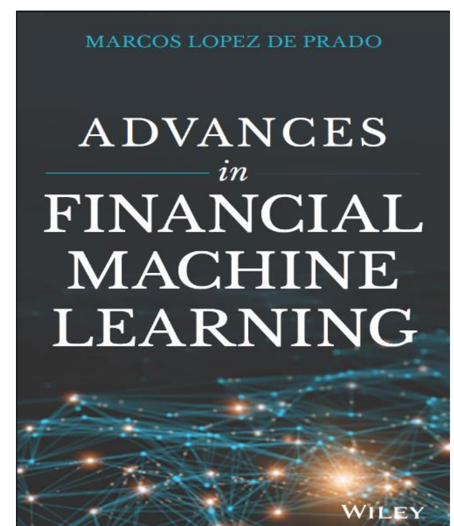
Factor Investor: “Took the risk nobody wanted. Where is my $1MM linear reward?”

Berkeley Lab’s Cori supercomputer: Scientific research requires complex techniques

Financial ML

- ML is only beginning to transform Finance:
  - 2016: Studies show that ML methods (like HRP) deliver portfolios that systematically outperform Markowitz optimization out-of-sample.
  - 2017: Four funds of Man/AHL manage $12.3 billion using AI.
  - 2017: The GIS-Liquid Strategies group manages $13 billion with 12 people.
  - 2018: KPMG's report argues that hedge funds must embrace technology or face 'treadmill to oblivion'.
  - 2018: First graduate-level textbook on ML, specifically applied to Finance.





Myth #1:
The Sisyphus Paradigm is Applicable to ML

The silo approach works for discretionary PMs

- Discretionary portfolio managers (PMs) make investment decisions that do not follow a particular theory or rigorous rationale.
- Because nobody fully understands the logic behind their bets, they can hardly work as a team and develop deeper insights beyond the initial intuition.
- If 50 PMs tried to work together, they would influence each other until eventually 49 would follow the lead of 1.

For this reason, investment firms ask discretionary PMs to work in silos.

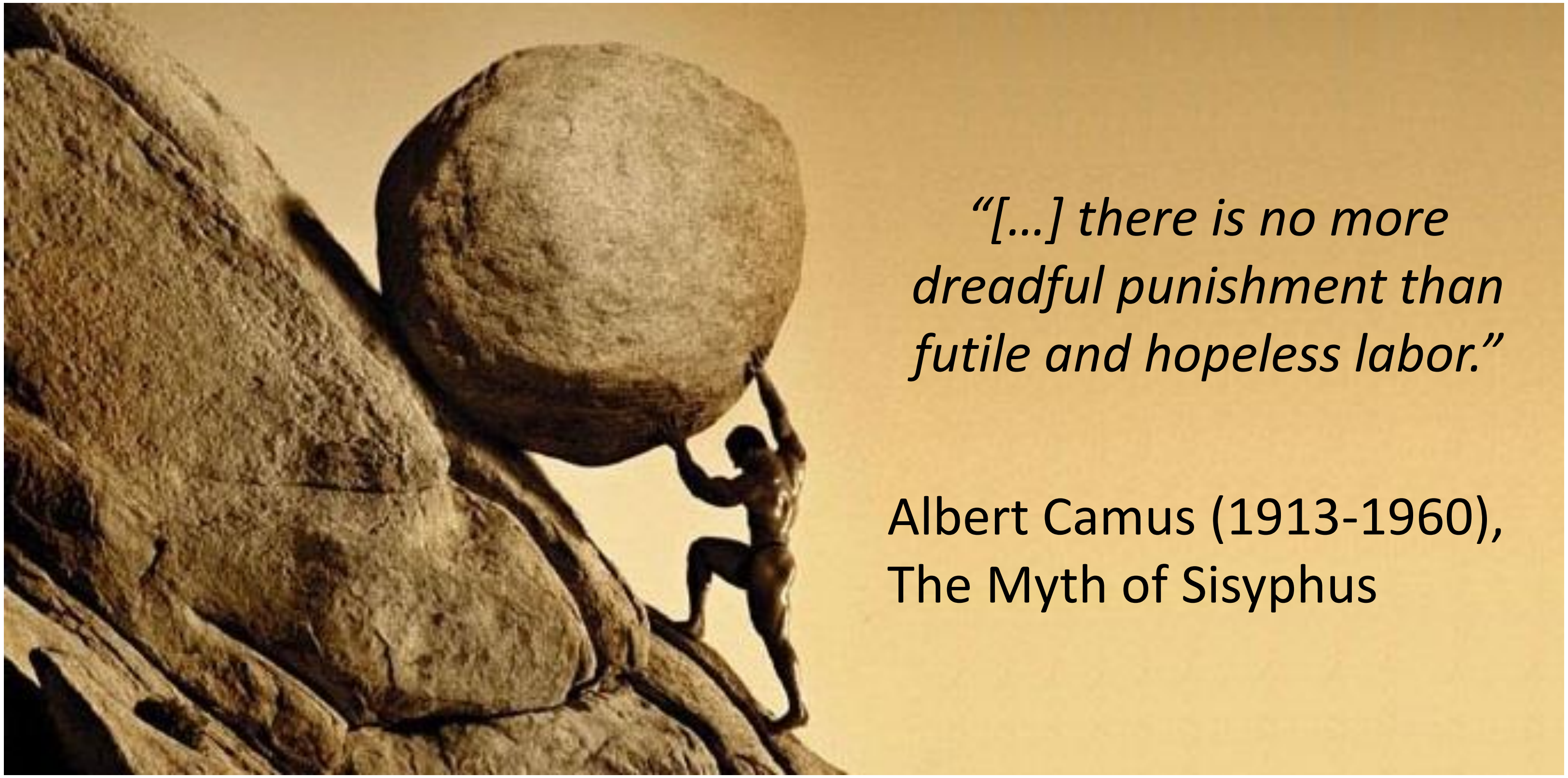
Silos prevent one PM from influencing the rest, hence protecting diversification.

The silo approach fails with quant PMs

- The boardroom's mentality is, let us do with quants what has worked with discretionary PMs.
- Let us hire 50 PhDs, and demand from each of them to produce an investment strategy within 6 months.
- This approach typically backfires, because each of these PhDs will frantically search for investment opportunities and eventually settle for:
  - A false positive that looks great in an overfit backtest; or
  - A standard factor model, which is an overcrowded strategy with low Sharpe ratio, but at least has academic support.
- Both outcomes will disappoint the investment board, and the project will be cancelled.
- Even if 5 of those 50 PhDs found something, they would quit.

Sisyphean Quants

- Firms directing quants to work in silos, or to develop individual strategies, are asking the impossible.
- Identifying new strategies requires large teams working together.



The Meta-Strategy Paradigm (1/3)

- The complexities involved in developing a true investment strategy are overwhelming:
  - Data collection, curation, processing, structuring,
  - HPC infrastructure,
  - software development,
  - feature analysis,
  - execution simulators,
  - backtesting, etc.
- Even if the firm provides you with shared services in those areas, you are like a worker at a BMW factory who has been asked to build the entire car alone, by using all the workshops around you.
  - One week you need to be a master welder, another week an electrician, another week a mechanical engineer, another week a painter, ... try, fail and circle back to welding. It is a futile endeavor.

The Meta-Strategy Paradigm (2/3)

- It takes almost as much effort to produce one true investment strategy as to produce a hundred.
- Every successful quantitative firm I am aware of applies the meta-strategy paradigm.
- Your firm must set up a research factory
  - where tasks of the assembly line are clearly divided into subtasks.
  - where quality is independently measured and monitored for each subtask.
  - where the role of each quant is to specialize in a particular subtask, to become the best there is at it, while having a holistic view of the entire process.
- This is how Berkeley Lab and other U.S. National laboratories routinely make scientific discoveries, such as adding 16 elements to the periodic table, or laying out the groundwork for MRIs and PET scans: https://youtu.be/G5nK3B5uuY8

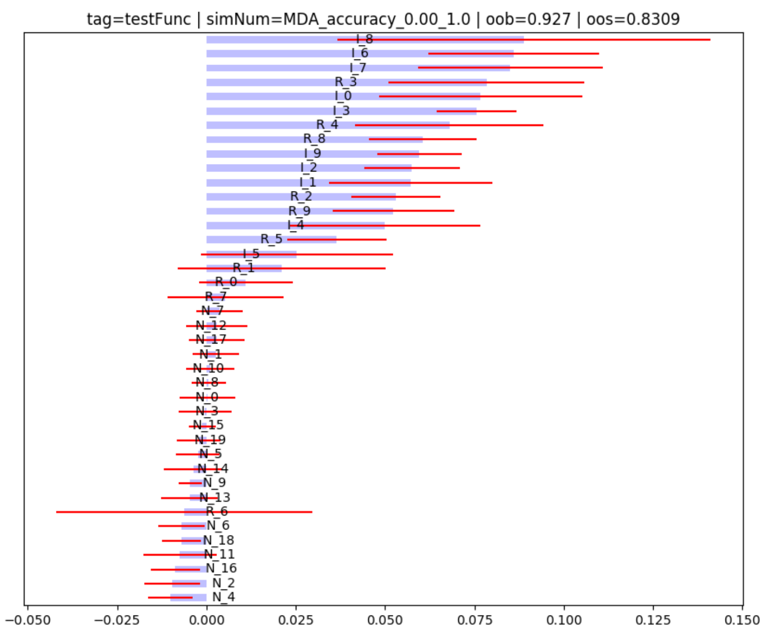
The Meta-Strategy Paradigm (3/3)

Practical Application	Classic approach	Quantitative Meta-Strategy
Selection & Hiring **(Example 1)**	Interview candidates with SR (or any other performance statistic) and track record length above a given threshold. Pros: Trivial to implement. Cons: Unknown (possibly high) probability of hiring unskilled PMs.	Design an interview process that recognizes the variables that affect the probability of making the wrong hire: • False positive rate. • False negative rate. • Skill-to-unskilled odds ratio. • Number of independent trials. • Sampling mechanism. Pros: It is objective and can be improved over time, based on measurable outcomes. Cons: More laborious.
Oversight **(Example 2)**	Allocate capital as if PMs were asset classes. Pros: Trivial to implement. Cons: Correlations are unstable, meaningless. Risks are likely to be concentrated.	Recognize that PMs styles evolve over time, as they adapt to a changing environment. Pros: It provides an early signal while the style is still emerging. Allocations can be revised before it is too late. Cons: Allocation revisions may be needed on an irregular calendar frequency.
Stop-Out **(Example 3)**	Stop-out a PM once a certain loss limit has been exceeded. Pros: Trivial to implement. Cons: It allows preventable problems to grow until it is too late.	For any drawdown, large or small, determine the expected time underwater and monitor every recovery. Even if a loss is small, a failure to recover within the expected timeframe indicates a latent problem. Pros: Proactive. Address problems before they force a stop-out. Cons: PMs may feel under tighter scrutiny.

Myth #2:
ML is a Black Box

Feature Importance

- ML algorithms identify patterns in a high dimensional space.
- These patterns associate features with outcomes.
- The nature of the relationship can be extremely complex, however we can always study what features were more important.
  - E.g., even if a ML algorithm may not derive an analytical formula for Newton's Gravitational Law, it will tell us that *mass* and *distance* were the key features.



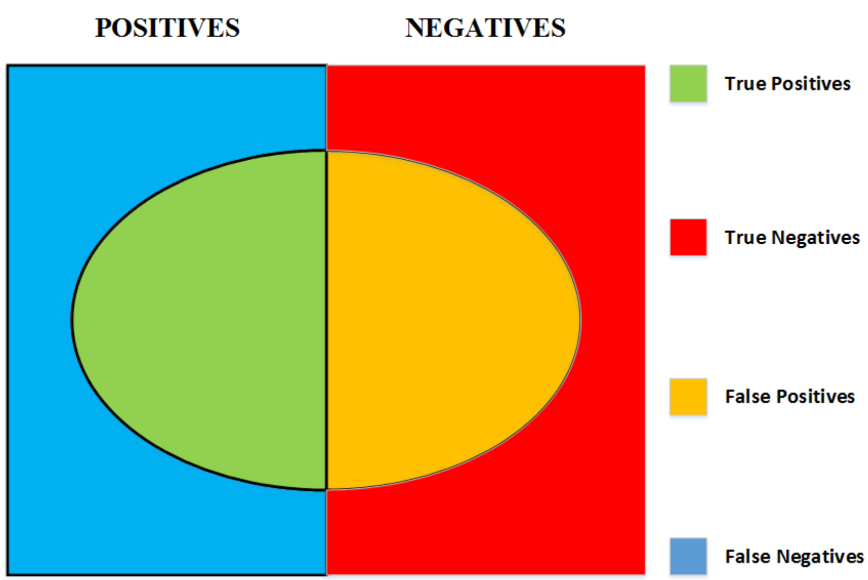
In traditional statistical analysis, key features are often missed as a result of the model's misspecification.

In ML analysis, we give up closed-form specifications in exchange for identifying what variables are truly relevant.

Once we know *what* are the factors at play, we can develop a theory of *how*.

Meta-labeling

- Suppose that you have a model for setting the side of the bet (long or short):
  - You just need to learn the size of that bet, which includes the possibility of no bet at all (zero size).
  - This is a situation that practitioners face regularly. We often know whether we want to buy or sell a product, and the only remaining question is how much money we should risk in such bet.



Meta-labeling builds a secondary ML model that learns how to use a primary exogenous model.

The secondary model does not learn the side. It learns the size.

How to use Meta-labeling

- Meta-labeling is particularly helpful when you want to achieve higher F1-scores:
  - First, we build a model that achieves high recall, even if the precision is not particularly high.
  - Second, we correct for the low precision by applying meta-labeling to the positives identified by the primary model.
- Meta-labeling is a very powerful tool in your arsenal, for three additional reasons:
  - ML algorithms are often criticized as black boxes. Meta-labeling allows you to build a ML system on a white box.
  - The effects of overfitting are limited when you apply meta-labeling, because ML will not decide the side of your bet, only the size.
  - Achieving high accuracy on small bets and low accuracy in large bets will ruin you. As important as identifying good opportunities is to size them properly, so it makes sense to develop a ML algorithm solely focused on getting that critical decision (sizing) right.

Meta-labeling for "Quantamental" Firms

- You can always add a meta-labeling layer to any primary model, whether that is an ML algorithm, a econometric equation, a technical trading rule, a fundamental analysis…
- That includes forecasts generated by a human, solely based on his intuition.
- In that case, meta-labeling will help us figure out when we should pursue or dismiss a discretionary PM's call.
- The features used by such meta-labeling ML algorithm could range from market information to biometric statistics to psychological assessments.
- Meta-labeling should become an essential ML technique for every discretionary hedge fund. In the near future, every discretionary hedge fund will become a quantamental firm, and meta-labeling offers them a clear path to make that transition.

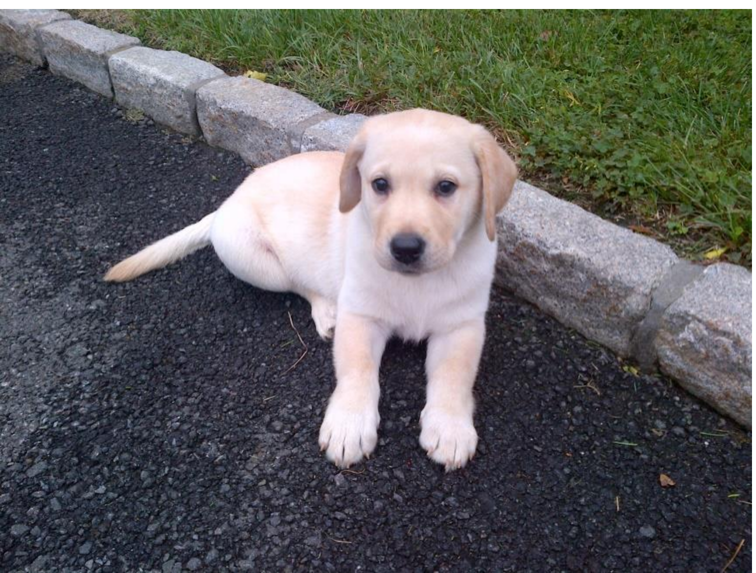
Myth #3:
Generic Solutions Work In Finance
(West → East)

Example: The "spilled samples" problem (1/2)

- Most non-financial ML researchers can assume that observations are drawn from IID processes. For example, you can obtain blood samples from a large number of patients, and measure their cholesterol.
- Of course, various underlying common factors will shift the mean and standard deviation of the cholesterol distribution, but the samples are still independent: There is one observation per subject.
- Suppose you take those blood samples, and someone in your laboratory spills blood from each tube to the following 9 tubes to their right.
  - That is, tube 10 contains blood for patient 10, but also blood from patients 1 to 9. Tube 11 contains blood from patient 11, but also blood from patients 2 to 10, and so on.

Example: The "spilled samples" problem (2/2)

- Now you need to determine the features predictive of high cholesterol (diet, exercise, age, etc.), without knowing for sure the cholesterol level of each patient.
- That is the equivalent challenge that we face in financial ML.
  - Labels are decided by outcomes.
  - Outcomes are decided over multiple observations.
  - Because labels overlap in time, we cannot be certain about what observed features caused an effect.



My friend Luna can recognize faces, like Google or FaceBook.

Luna is not so good at investing, and Google's ML would probably fail miserably if applied to financial markets.

Finance is not a plug-and-play subject as it relates to ML

Myth #4:
Traditional Quants Know About ML
(West ← East)

Traditional Quants: “P” and “Q”

- “Quants” only exist in the Press.
  - What does a finance professor have in common with a physicist?
- In reality, the “quant” world is populated by two very distinct (and often adversarial!) species: *P-quants* and *Q-quants*.
- **P-quants**:
  - Primary education: Economics, finance, statistics.
  - Main goal: Risk and portfolio management.
  - Main tools: Statistics, linear algebra, calculus, econometrics.
  - Main language: MatLab, R.
- **Q-quants**:
  - Primary education: Physics, engineering, mathematics.
  - Main goal: Derivatives pricing.
  - Main tools: Stochastic calculus, PDE, Martingale pricing.
  - Main language: C, C++.

Data Scientists

- In contrast, ML practitioners (or Data Scientists):
  - Primary education: Mathematics, statistics, computer science.
  - Main goal: Data analysis for prediction.
  - Main tools: Discrete math, kernel methods, experimental math.
  - Main language: Python, R, C, ...
- Unlike P-quants, data scientists search for patterns.
  - They are not beholden to any particular theory regarding its origin or cause.
  - This creates the risk of overfitting, which must be managed carefully.
- Unlike Q-quants, data scientists do not place emphasis on exactness.
  - The theorems cover the classifiers, not the patterns.
  - Data scientists excel at modelling hierarchical relations.
  - Heuristics are common.

Myth #5:
There Are Many ML Portfolio Managers

The Shortage of (*True*) ML Funds

- There is only a handful of ML funds or ML portfolio managers with a track record. The hype is mostly based on paper trading!!
- The conundrum:
  - ML experts have very little experience with finance.
  - Finance experts have very little experience with ML.
- It will take many years (perhaps a decade) to form a workforce with joint expertise and experience in both areas (Finance + ML).
- Until then, our best hope is to re-train some of the existing “P” and “Q” quants.
  - This presents the risk that those retrained ML quants will make common mistakes. It also explains why so many (pseudo) ML funds have failed.
- Signs that an operation does *not* use true ML:
  - Follows the Sisyphean paradigm.
  - Researchers require less than 100 TFLOPS per headcount (as of 2018).

THANKS FOR YOUR ATTENTION!

Bio

Marcos Lopez de Prado manages several multibillion-dollar funds for institutional investors using machine learning algorithms. He is also a research fellow at *Lawrence Berkeley National Laboratory* (U.S. Department of Energy, Office of Science).

One of the top-10 most read authors in finance (SSRN's rankings), Marcos has published dozens of scientific articles on machine learning in the leading academic journals, and he holds multiple international patent applications on algorithmic trading.

Marcos earned a PhD in Financial Economics (2003), a second PhD in Mathematical Finance (2011) from *Universidad Complutense de Madrid*, and is a recipient of Spain's National Award for Academic Excellence (1999). He completed his post-doctoral research at *Harvard University* and *Cornell University*, where he teaches a Financial Machine Learning course at the School of Engineering. Marcos has an Erdös #2 and an Einstein #4 according to the *American Mathematical Society*.

Disclaimer

- The views expressed in this document are the authors’ and do not necessarily reflect those of the organizations he is affiliated with.
- No investment decision or particular course of action is recommended by this presentation.
-