AI-First Finance— How AI is Reshaping an Industry

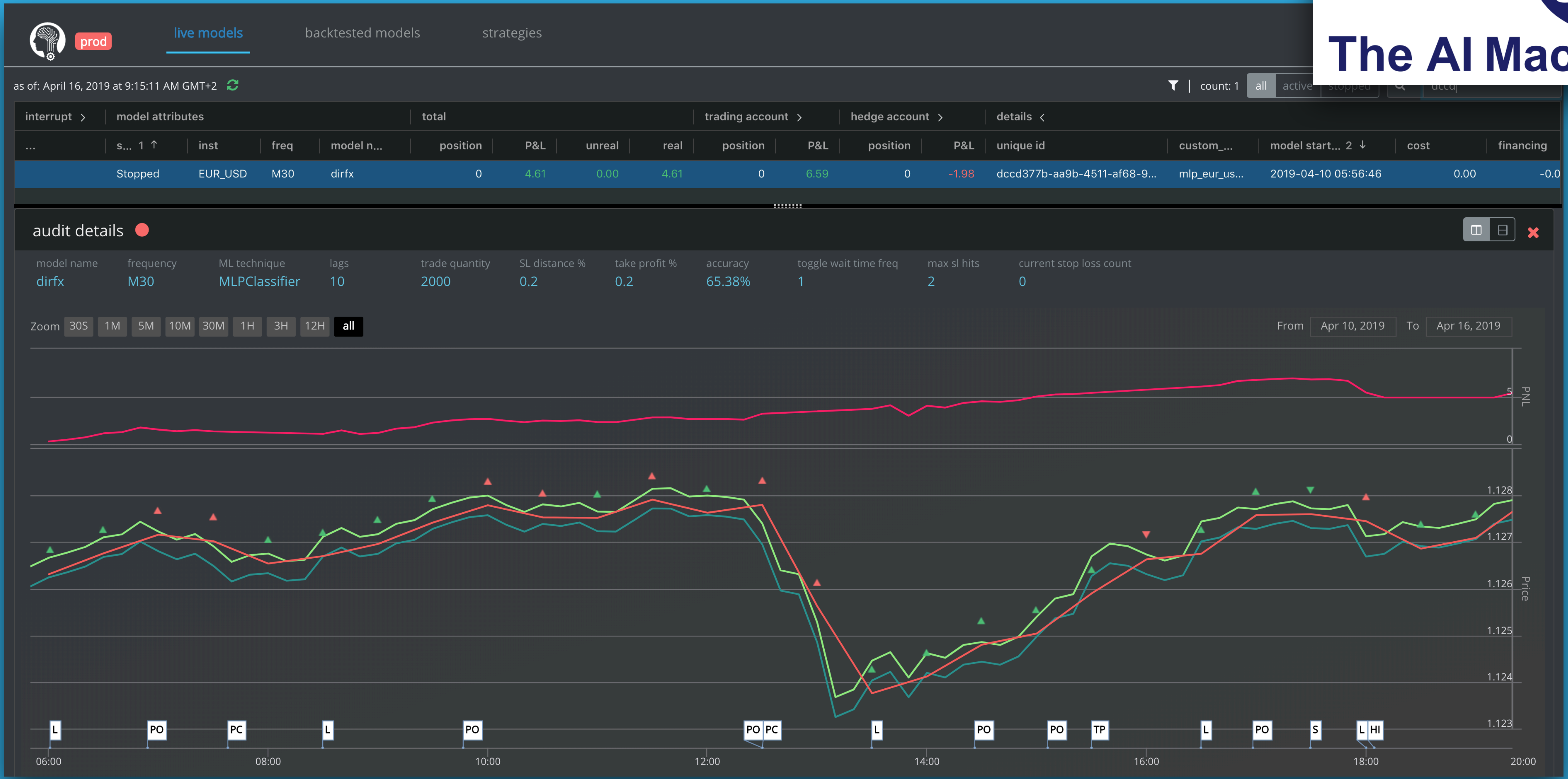
Dr. Yves J. Hilpisch

Intent International Conference
Online, 18. November 2021

Introduction

THE PYTHON QUANTS

http://tpq.io



http://aimachine.io

Certificate Progam Algo Trading


Capital Markets CIO Outlook TOP 10 ALGO TRADING SOLUTION PROVIDERS - 2019

The Python Quants

First University Certificate in Python for Algorithmic Trading

Python programming has become a key skill in the financial industry. In areas such as financial data science, computational finance or algorithmic trading, Python has established itself as the primary technological platform. At the same time, the level of Python sophistication the industry is expecting from its employees and applicants is increasing steadily. The Python Quants Group is one of the leading providers of Python for Finance training programs.

Among others, The Python Quants have tailored a comprehensive online training program leading to the first University Certificate in Python for Algorithmic Trading. Be it an ambitious student with intrigue for algorithmic trading, or a major financial institution, The Python Quants, through this systematic training program, is equipping delegates with requisite skills and tools to formulate, backtest and deploy algorithmic trading strategies based on Python.

The topics covered in the training programs offered by The Python Quants are generally not found in the typical curriculum of financial engineering or quantitative finance Master programs. Dr. Yves Hilpisch, the firm's founder and managing partner, explains, "There are courses out there that show students how to apply machine learning for the formulation and backtesting of algorithmic trading strategies. However, none of them explains the difficulties or the skills required in deploying such algorithmic trading strategies in the real world. Besides providing an introductory course that teaches Python and financial concepts from scratch, we train our delegates and clients on how best to deploy algorithmic trading strategies in automated fashion in the cloud, with, among others, real-time risk management and monitoring," explains Hilpisch, an author of three books on the topic, with "Python for Finance" (2nd ed., O'Reilly) being the standard reference in the field.

The organization's "Python for Algorithmic Trading University Certificate" consists of 200 hours of instruction, 1,200 pages of documentation and 1,000s of lines of Python code. In addition to offering both online and offline Python training, Hilpisch and his team also organize bespoke training events for financial institutions, hedge funds, banks, and asset management companies. "Most of the training is online since we have students and delegates from about 65 different countries in general. Most recently, we noticed that it's not just financial firms and students who want to deepen their algorithmic trading knowledge, but even professors of finance who want to get more involved in this popular topic," says Hilpisch.

While the Quant Platform is the most popular choice, especially for users in the financial sector who don't have access to a full-fledged, interactive, financial analytics environment, the team at The Python Quants is currently developing The AI Machine—a new platform which leverages artificial intelligence to formulate and deploy algorithmic trading strategies in a standardized manner. Hilpisch explains that it's relatively easy to write Python code for an algorithmic trading strategy, but the same can't be said about the deployment of such a strategy. "There are a few platforms out there that allow the formulation and backtesting of algorithmic trading strategies by the use of Python code. However, they usually stop exactly there. With The AI Machine, it is a single click on the 'GO LIVE' button and the strategy is deployed in real-time—without any changes to the strategy code itself," adds Hilpisch.

In 2019, The Python Quants will be introducing a new university certificate titled "Python for Computational Finance," which will focus more on original quantitative finance topics, such as option pricing, Monte Carlo simulation, and hedging. As financial institutions begin to perceive Python-based analytics as a prerequisite skill, the organization will continue to provide an "efficient and structured way of mastering all the tools and skills required in Python for Financial Data Science, Algorithmic Trading, and Computational Finance."CM

Dr. Yves Hilpisch

Dr. Yves J. Hilpisch is founder and CEO of **The Python Quants** (http://tpq.io), a group focusing on the use of open source technologies for financial data science, artificial intelligence, algorithmic trading, computational finance, and asset anagement. He is also the founder and CEO of **The AI Machine** (http://aimachine.io), a company focused on AI-powered algorithmic trading based on a proprietary strategy execution platform.

Yves has a Diploma in Business Administration, a Ph.D. in Mathematical Finance, and is Adjunct Professor for Computational Finance.

Yves is the author of six books (https://home.tpq.io/books):

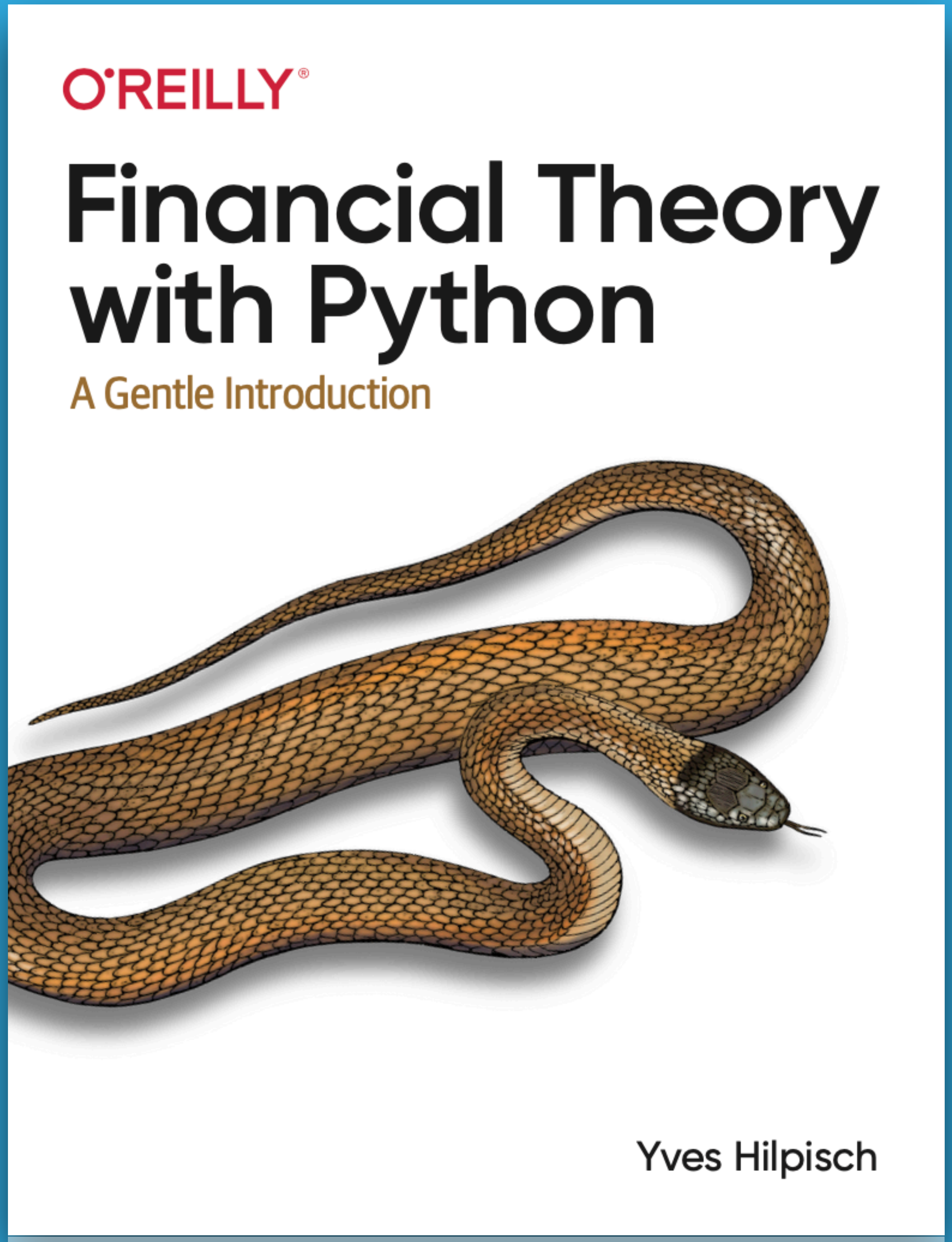
* Financial Theory with Python (2021, O'Reilly)
* Artificial Intelligence in Finance (2020, O'Reilly)
* Python for Algorithmic Trading (2020, O'Reilly)
* Python for Finance (2018, 2nd ed., O'Reilly)
* Listed Volatility and Variance Derivatives (2017, Wiley Finance)
* Derivatives Analytics with Python (2015, Wiley Finance)

Yves is the director of the first online training programs leading to **University Certificates in Python for Algorithmic Trading** (https://home.tpq.io/certificates/pyalgo), **Computational Finance** (https://home.tpq.io/certificates/compfin), and **Asset Management** (https://home.tpq.io/certificates/pyam). He also lectures on computational finance, machine learning, and algorithmic trading at the **CQF Program** (http://cqf.com).

Yves is the originator of the financial analytics library **DX Analytics** (http://dx-analytics.com) and organizes Meetup group **events, conferences, and bootcamps** about Python, artificial intelligence and algorithmic trading in London (http://pqf.tpq.io), New York (http://aifat.tpq.io), Frankfurt, Berlin, and Paris. He has given **keynote speeches** at technology conferences in the United States, Europe, and Asia.

http://hilpisch.com

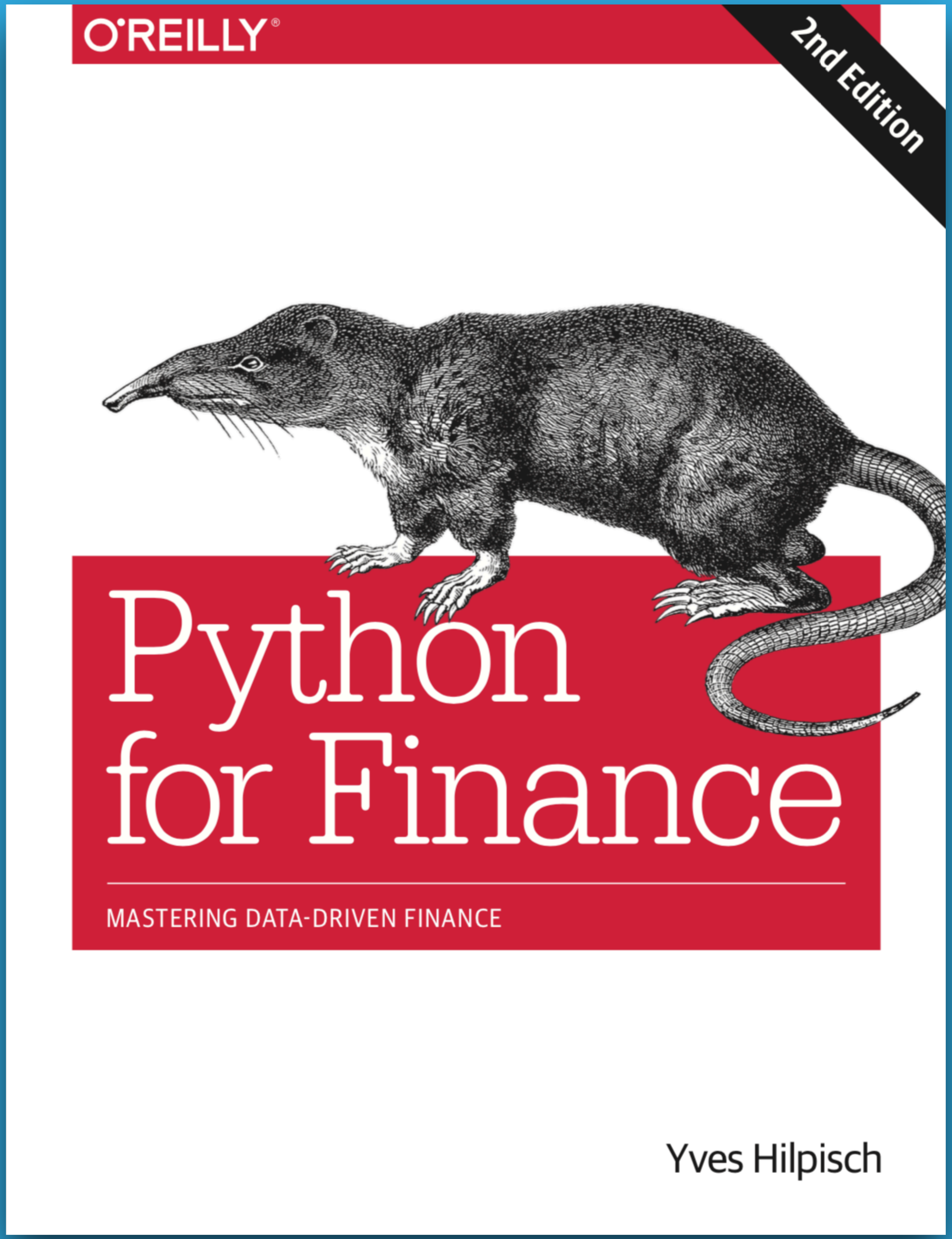
Financial Theory with Python — A Gentle Introduction

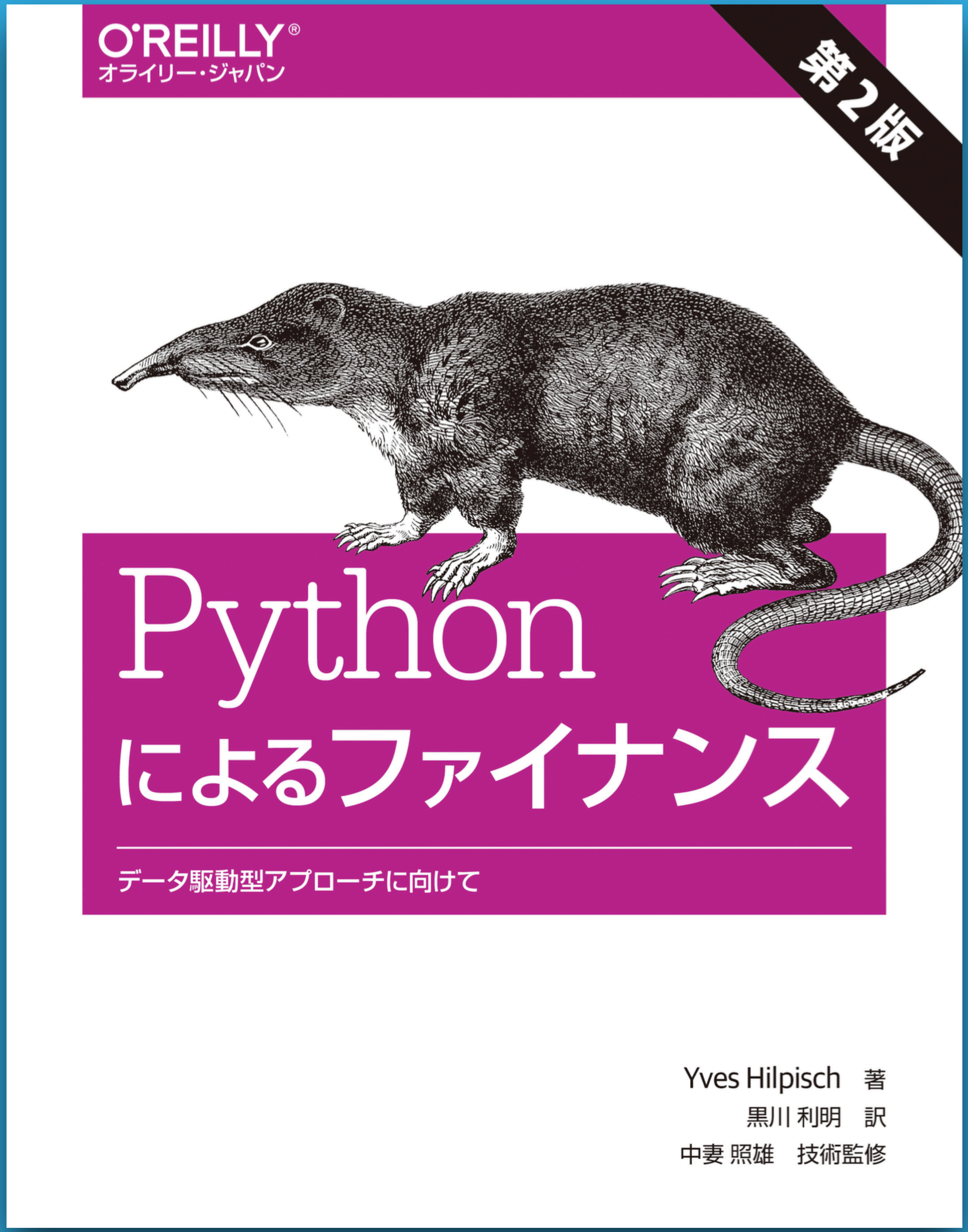


Finance with Python
Python Environments
Basic Finance Concepts and Models:
- *Risk-Return*
- *Pricing of Instruments*
- *Expected Utility Theory*
- *Mean-Variance Portfolio Theory*
- *Capital Asset Pricing Model*
- *Portfolio Optimization*

Basic Python Concepts and Packages:
- *Major Python Idioms*
- *NumPy Package*
- *SciPy & SymPy Packages*

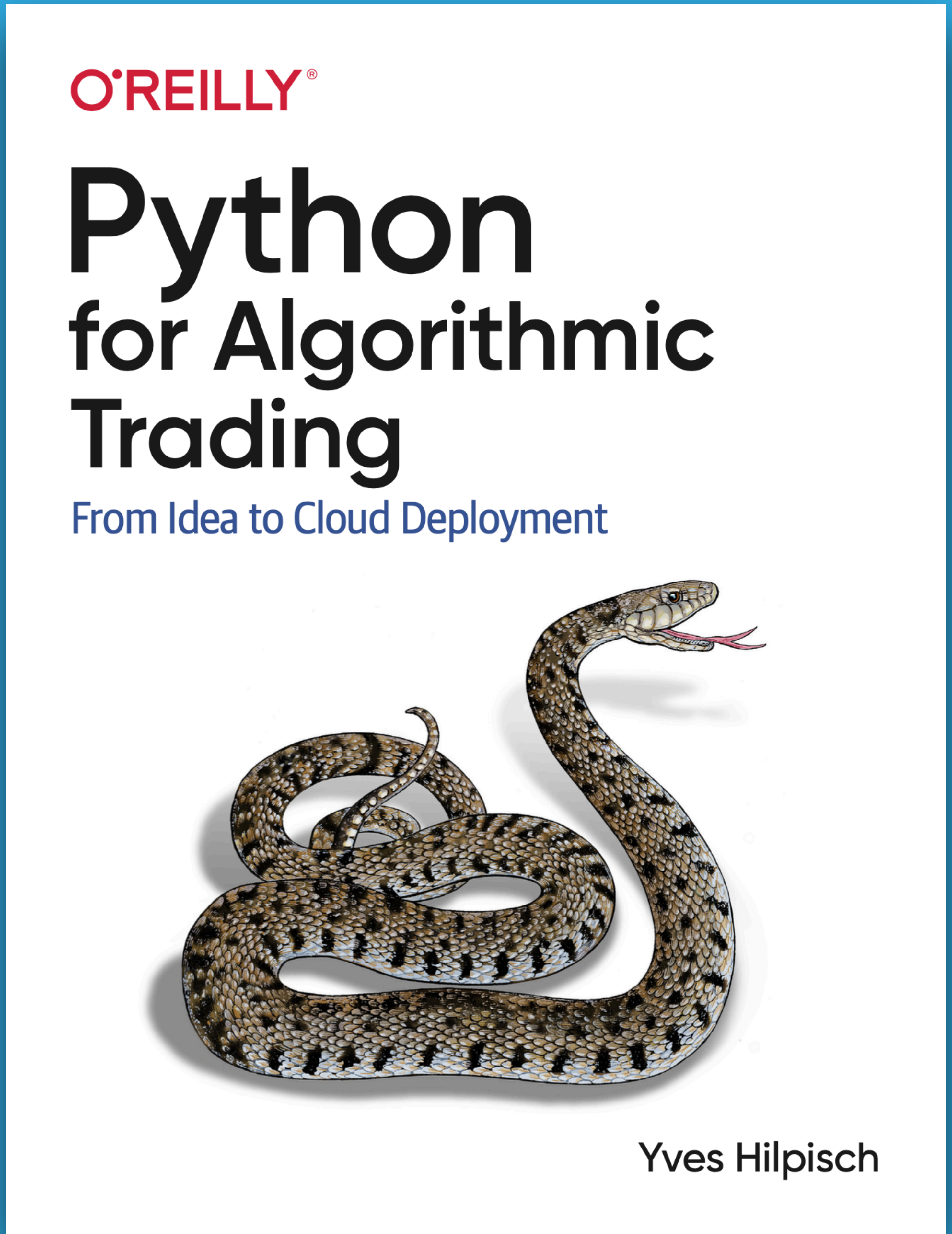
Python for Finance

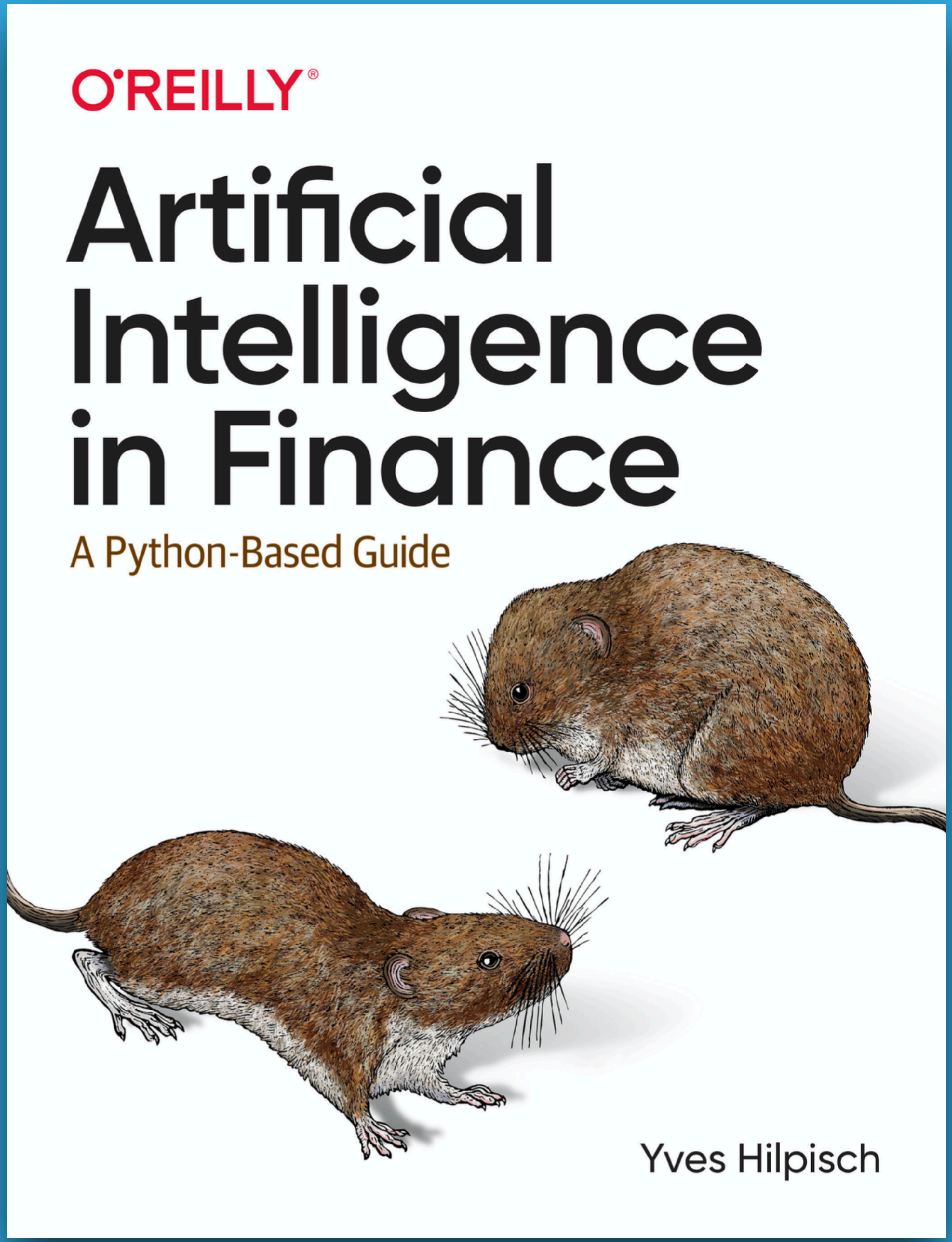




http://books.tpq.io

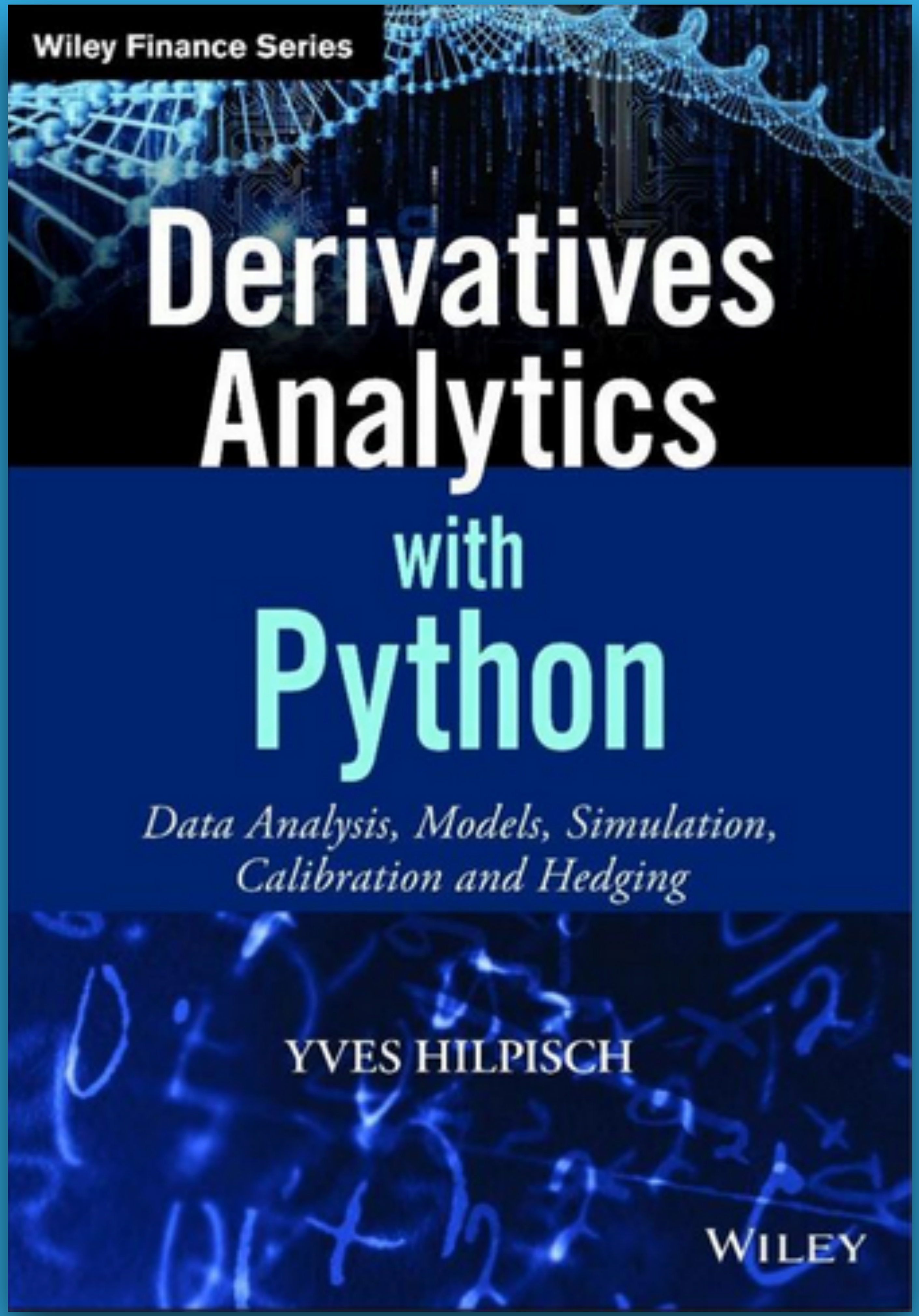
Python & AI for Finance & Trading

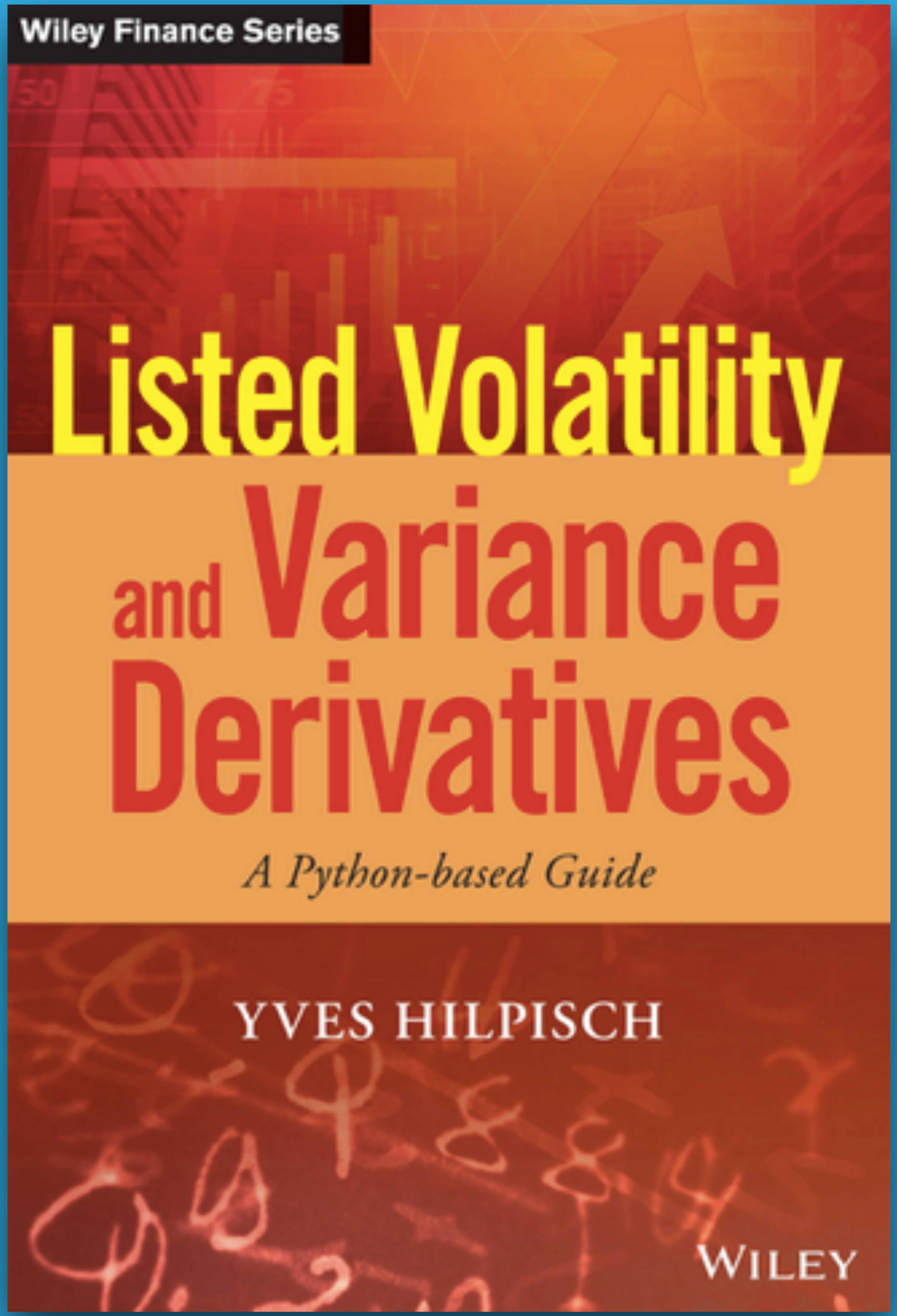




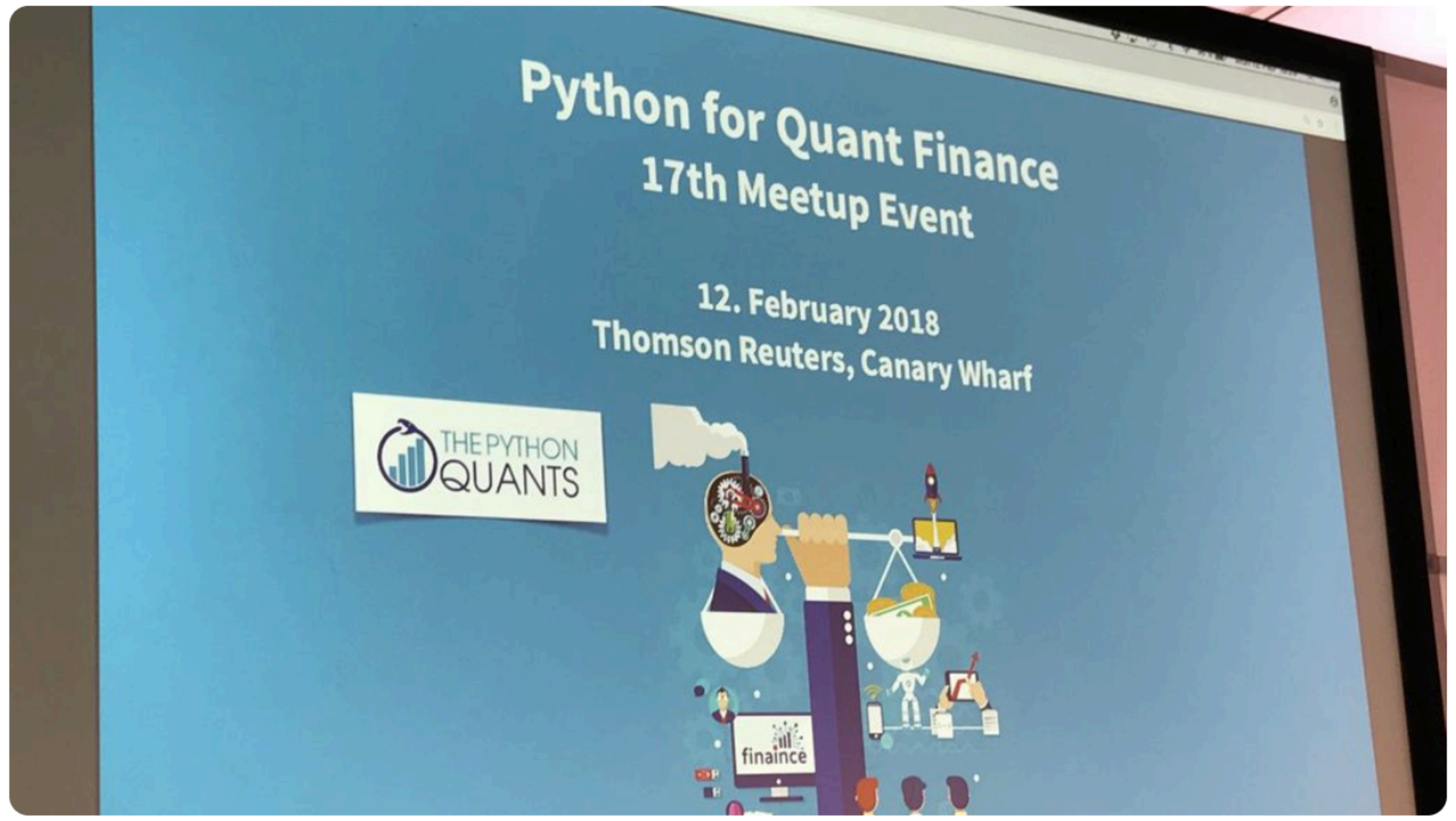
http://books.tpq.io

Quant Finance with Python





http://books.tpq.io



Python for Quant Finance

London, United Kingdom
3,416 members · Public group
Organized by **Yves H.** and **2 others**

Share:

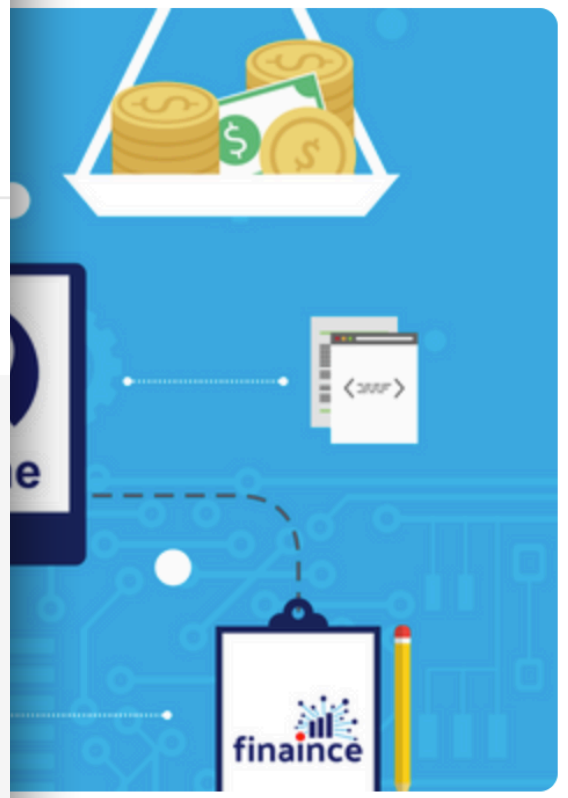
About | Events | Members | Photos | Discussions | More | Join this group

What we're about

This group is about the use of Python & AI for Quantitative Financial Applications, Algorithmic Trading and Interactive Financial Analytics.

Organizers

Yves H. and **2 others**
Message



Artificial Intelligence in Finance & Algorithmic Trading

New York, NY
345 members · Public group
Organized by **Yves Hilpisch**

Share:

About | Events | Members | Photos | Discussions | More | Manage group | Create event

What we're about

This Meetup group is concerned with data-driven and AI-first finance in general and algorithmic trading in particular. Its events cover the latest...

Organizer

Yves Hilpisch
Message

1. History of Finance
2. AI Success Stories
3. Physics Envy &
   The Beauty Myth
4. Data-Driven Finance
5. Efficient Markets
6. AI-First Finance
7. Basic Strategies
8. Conclusions

A History of Finance

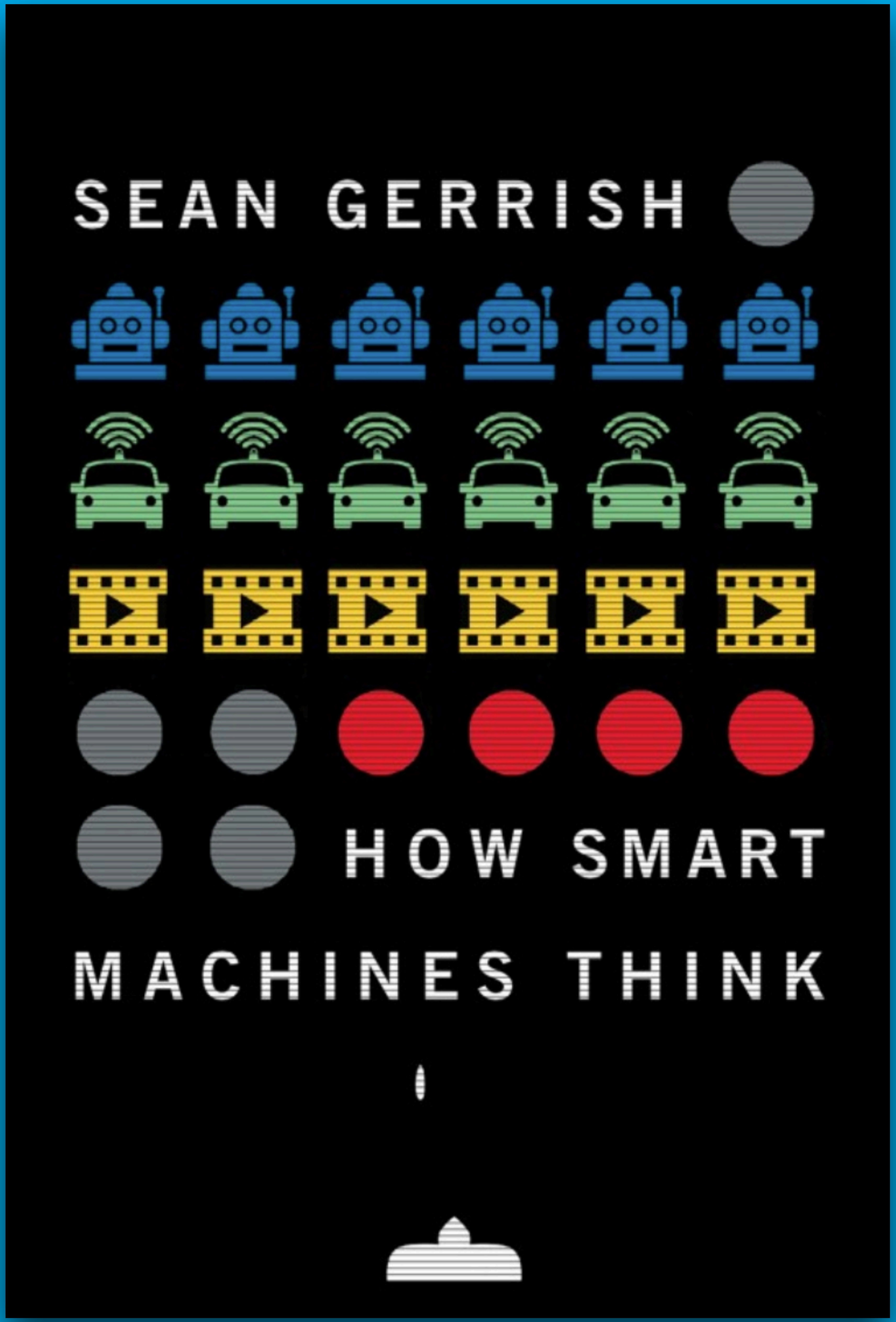
Finance has gone through multiple phases and paradigm shifts over time (1):

- **The ancient period (pre–1950):** A period mainly characterized by informal reasoning, rules of thumb, and the experience of market practitioners.
  See Rubinstein (2006): A History of the Theory of Investments, Wiley Finance.
- **The classical period (1950–1980):** A period characterized by the introduction of formal reasoning and mathematics to the field.
  See Hilpisch (2021): Financial Theory with Python, O'Reilly.
- **The modern period (1980–2000):** This period generated many advances in specific subfields of finance (for example, computational finance) and tackled, among others, important empirical phenomena in the financial markets, such as stochastic interest rates or stochastic volatility.
  See Hilpisch (2015): Derivatives Analytics with Python, Wiley Finance.

Finance has gone through multiple phases and paradigm shifts over time (2):

- **The computational period (2000–2020):** This period saw a shift from a theoretical focus in finance to a computational one, driven by advances in both hardware and software used in finance.
  See Hilpisch (2018): Python for Finance, 2nd ed., O'Reilly.
- The artificial intelligence period (post-2020): Advances in artificial intelligence (AI) and related success stories have spurred interest to make use of the capabilities of AI in the financial domain. AI-first finance describes the shift from simple, in general linear, models in finance to the use of advanced models and algorithms from AI.
  See Hilpisch (2020): Artificial Intelligence in Finance, O'Reilly.

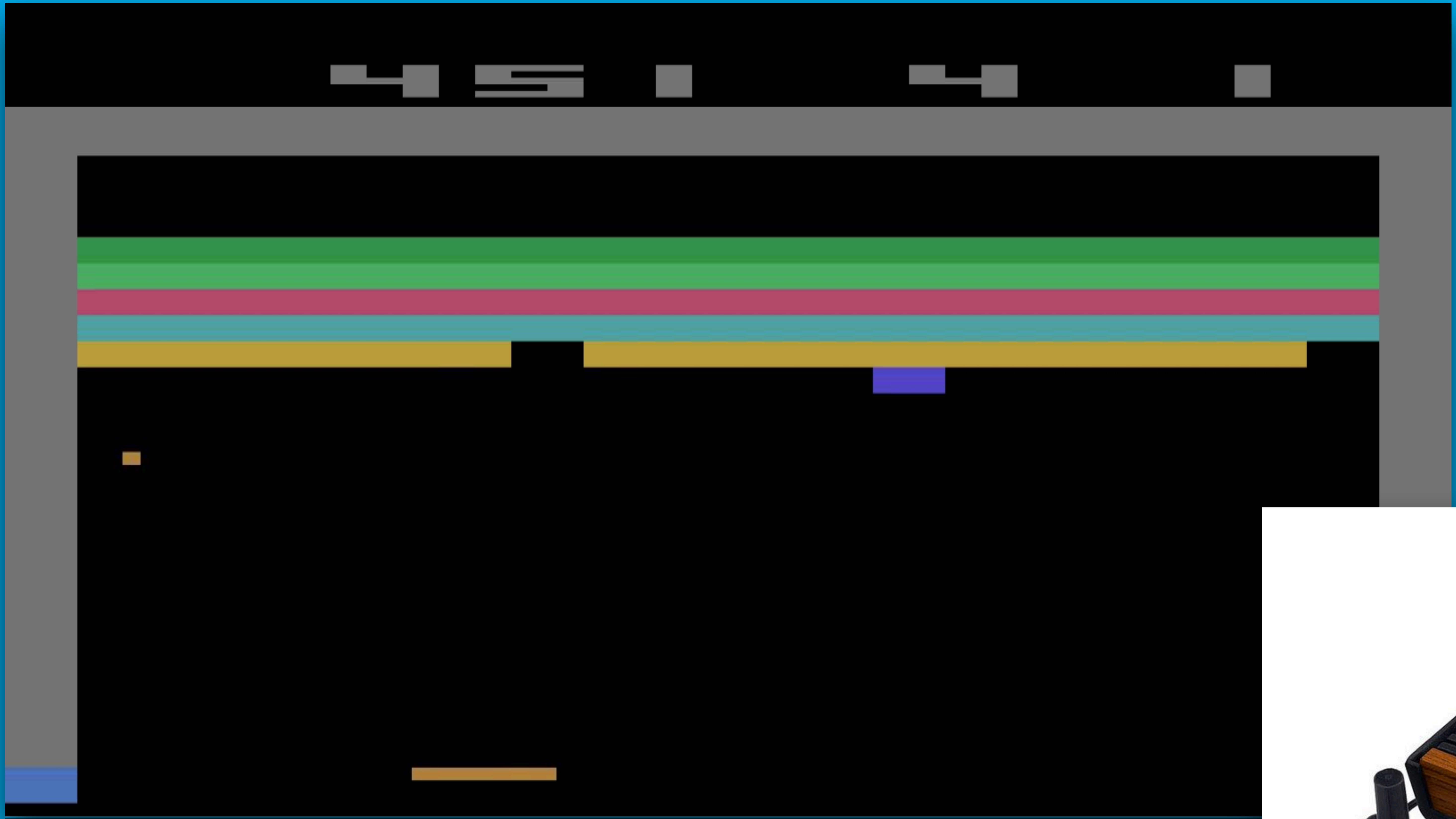
AI Success Stories



Success Stories about Deep Learning and Deep Reinforcement Learning:

- Self-Driving Cars
- Recommendation Engines
- Playing Atari Games
- Image Recognition & Classification
- Speech Recognition
- Playing the Game of Go

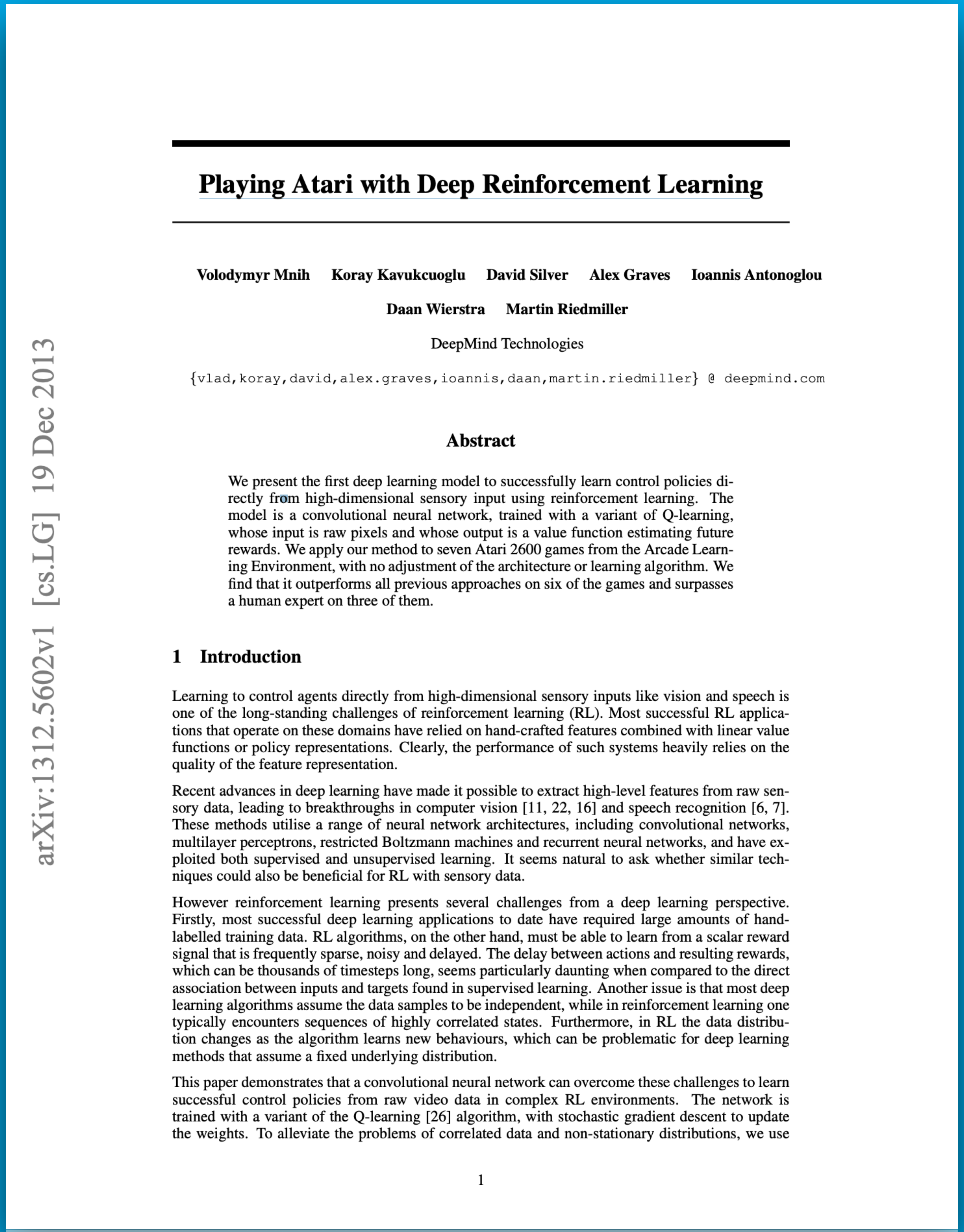
AI Success Stories —Atari Games and Reinforcement Learning


451 4 1


VIDEO COMPUTER SYSTEM™
ATARI

“We present the first deep learning model to successfully learn control policies directly from high-dimensional sensory input using reinforcement learning. The model is a convolutional neural network, trained with a variant of Q-learning, whose input is raw pixels and whose output is a value function estimating future rewards. We apply our method to seven Atari 2600 games from the Arcade Learning Environment, with no adjustment of the architecture or learning algorithm. We find that it outperforms all previous approaches on six of the games and surpasses a human expert on three of them.”

Mnih, V. (2013): “Playing Atari with Deep Reinforcement Learning”. https://arxiv.org/pdf/1312.5602v1.pdf



arXiv:1312.5602v1 [cs.LG] 19 Dec 2013

Playing Atari with Deep Reinforcement Learning

Volodymyr Mnih Koray Kavukcuoglu David Silver Alex Graves Ioannis Antonoglou
Daan Wierstra Martin Riedmiller
DeepMind Technologies
{vlad,koray,david,alex.graves,ioannis,daan,martin.riedmiller} @ deepmind.com

Abstract

We present the first deep learning model to successfully learn control policies directly from high-dimensional sensory input using reinforcement learning. The model is a convolutional neural network, trained with a variant of Q-learning, whose input is raw pixels and whose output is a value function estimating future rewards. We apply our method to seven Atari 2600 games from the Arcade Learning Environment, with no adjustment of the architecture or learning algorithm. We find that it outperforms all previous approaches on six of the games and surpasses a human expert on three of them.

1 Introduction

Learning to control agents directly from high-dimensional sensory inputs like vision and speech is one of the long-standing challenges of reinforcement learning (RL). Most successful RL applications that operate on these domains have relied on hand-crafted features combined with linear value functions or policy representations. Clearly, the performance of such systems heavily relies on the quality of the feature representation.

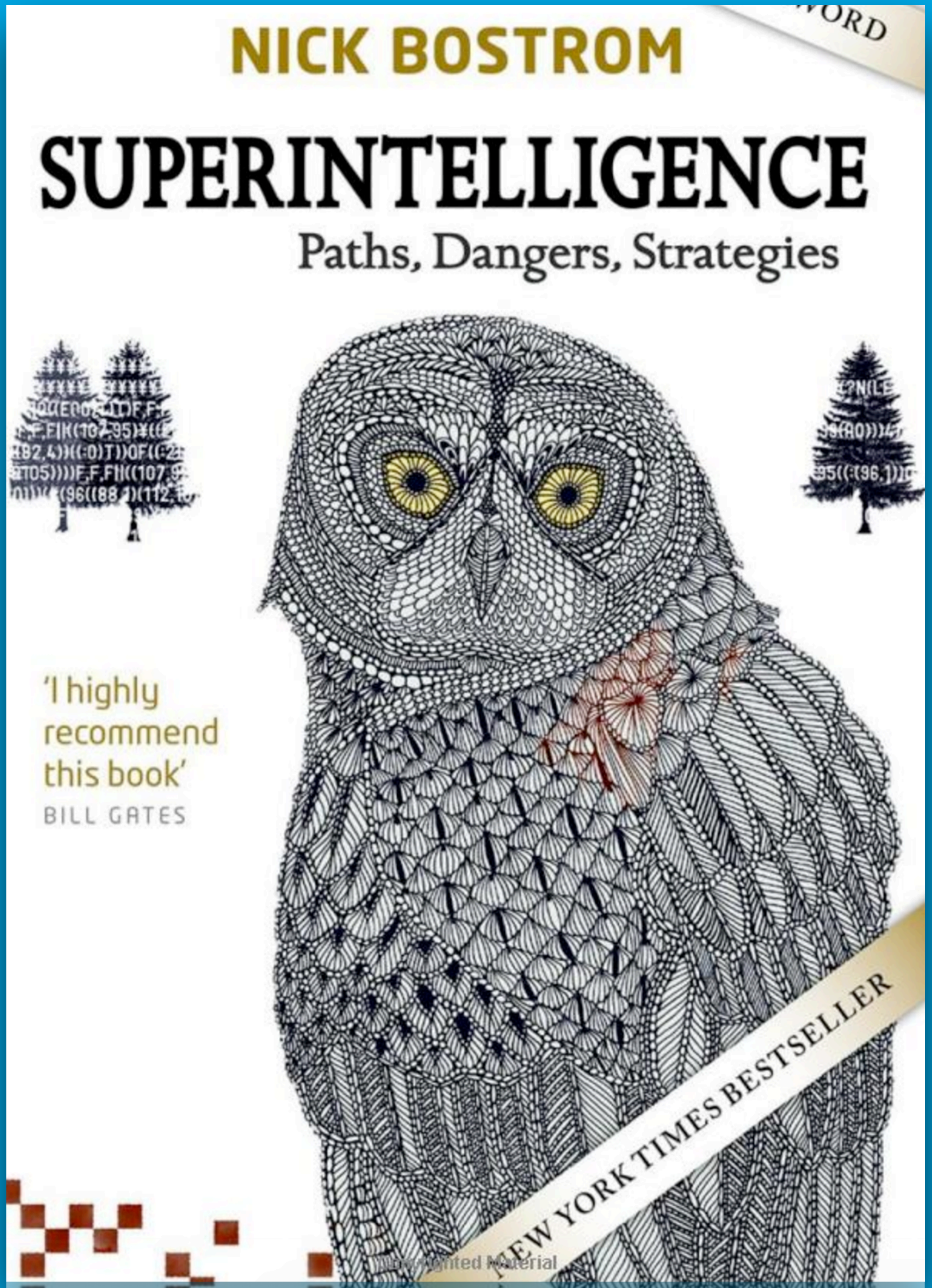
Recent advances in deep learning have made it possible to extract high-level features from raw sensory data, leading to breakthroughs in computer vision [11, 22, 16] and speech recognition [6, 7]. These methods utilise a range of neural network architectures, including convolutional networks, multilayer perceptrons, restricted Boltzmann machines and recurrent neural networks, and have exploited both supervised and unsupervised learning. It seems natural to ask whether similar techniques could also be beneficial for RL with sensory data.

However reinforcement learning presents several challenges from a deep learning perspective. Firstly, most successful deep learning applications to date have required large amounts of hand-labelled training data. RL algorithms, on the other hand, must be able to learn from a scalar reward signal that is frequently sparse, noisy and delayed. The delay between actions and resulting rewards, which can be thousands of timesteps long, seems particularly daunting when compared to the direct association between inputs and targets found in supervised learning. Another issue is that most deep learning algorithms assume the data samples to be independent, while in reinforcement learning one typically encounters sequences of highly correlated states. Furthermore, in RL the data distribution changes as the algorithm learns new behaviours, which can be problematic for deep learning methods that assume a fixed underlying distribution.

This paper demonstrates that a convolutional neural network can overcome these challenges to learn successful control policies from raw video data in complex RL environments. The network is trained with a variant of the Q-learning [26] algorithm, with stochastic gradient descent to update the weights. To alleviate the problems of correlated data and non-stationary distributions, we use

1

AI Success Stories
—Go and AlphaGo



"Go-playing programs have been improving at a rate of about 1 dan/year in recent years. If this rate of improvement continues, they might beat the human world champion in about a decade."

Nick Bostrom (2014): Superintelligence.

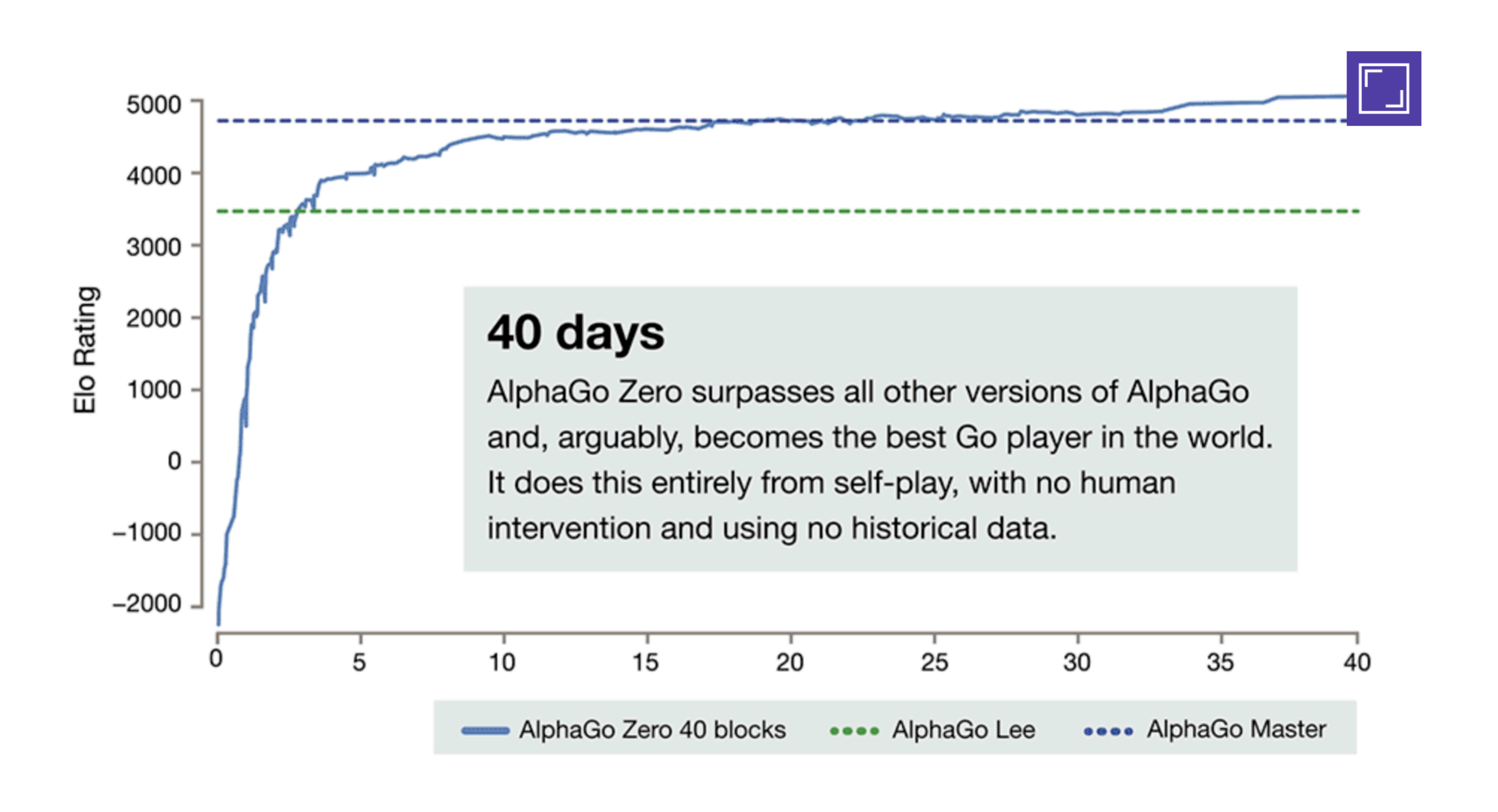
The story of AlphaGo so far

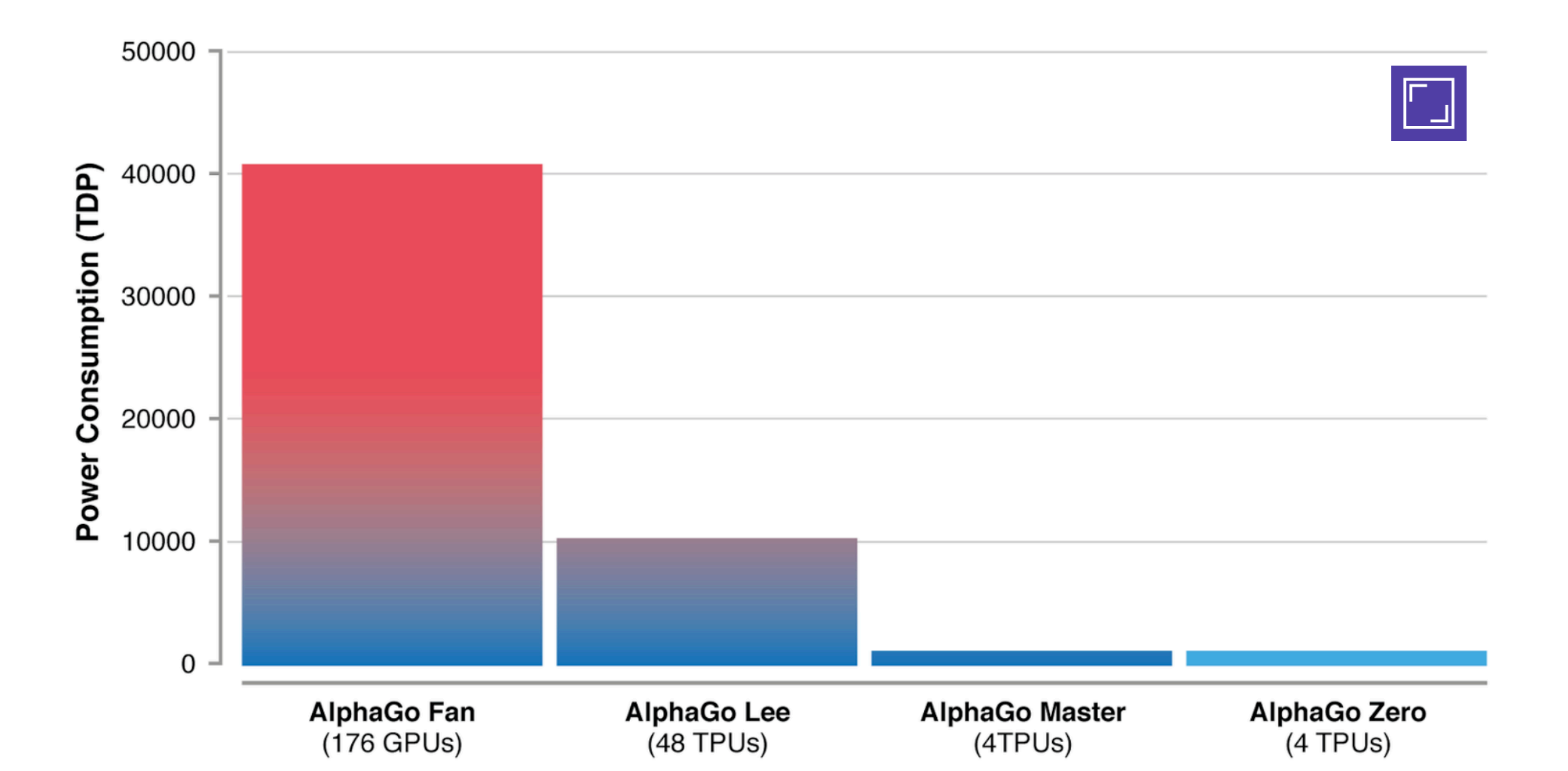
AlphaGo is the first computer program to defeat a professional human Go player, the first program to defeat a Go world champion, and arguably the strongest Go player in history.

AlphaGo's first formal match was against the reigning 3-times European Champion, Mr Fan Hui, in October 2015. Its 5-0 win was the first ever against a Go professional, and the results were published in full technical detail in the international journal, **Nature**. AlphaGo then went on to compete against legendary player Mr Lee Sedol, winner of 18 world titles and widely considered to be the greatest player of the past decade.

AlphaGo's 4-1 victory in Seoul, South Korea, in March 2016 was watched by over 200 million people worldwide. It was a landmark achievement that experts agreed was a decade ahead of its time, and earned AlphaGo a 9 dan professional ranking (the highest certification) - the first time a computer Go player had ever received the accolade.

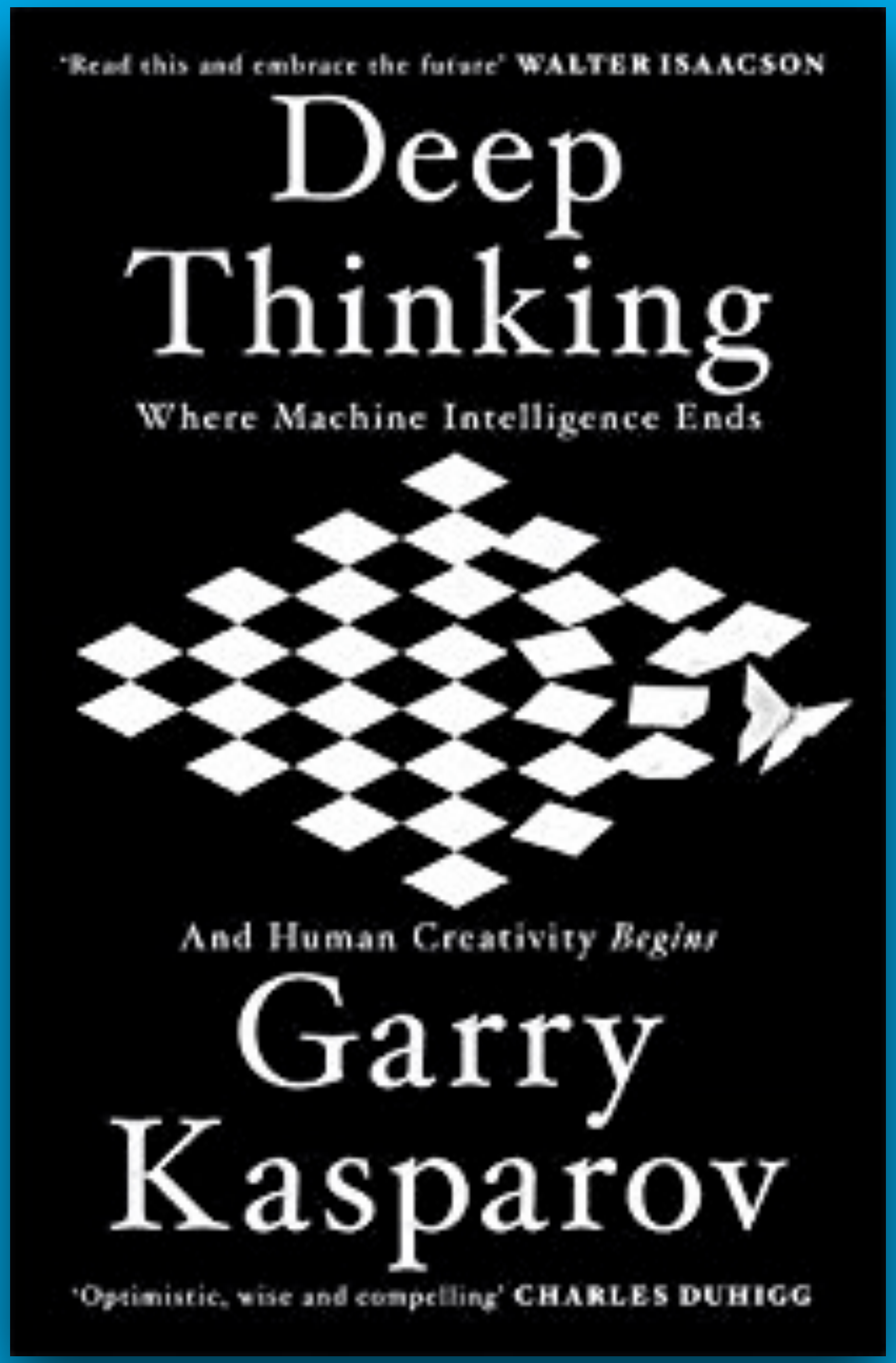
During the games, AlphaGo played a handful of **highly inventive winning moves,** several of which - including move 37 in game two - were so surprising they overturned hundreds of years of received wisdom, and have since been examined extensively by players of all levels. In the course of winning, AlphaGo somehow taught the world completely new knowledge about perhaps the most studied and contemplated game in history.





AlphaGo has become progressively more efficient thanks to hardware gains and more recently algorithmic advances

AI Success Stories
—Chess, Deep Blue & AlphaZero



"It was a pleasant day in Hamburg in June 6, 1985, … Each of my opponents, all thirty-two of them, was a computer. … it didn't come as much of a surprise, …, when I achieved

"Twelve years later I was in New York City fighting for my chess life. Against just one machine, a $10 million IBM supercomputer nicknamed 'Deep Blue'."

"Jump forward another 20 years to today, to 2017, and you can download any number of free chess apps for your phone that rival any human Grandmaster."

"Traditional chess engines – including the world computer chess champion Stockfish and IBM's ground-breaking Deep Blue – rely on **thousands of rules and heuristics handcrafted by strong human players** that try to account for every eventuality in a game. …

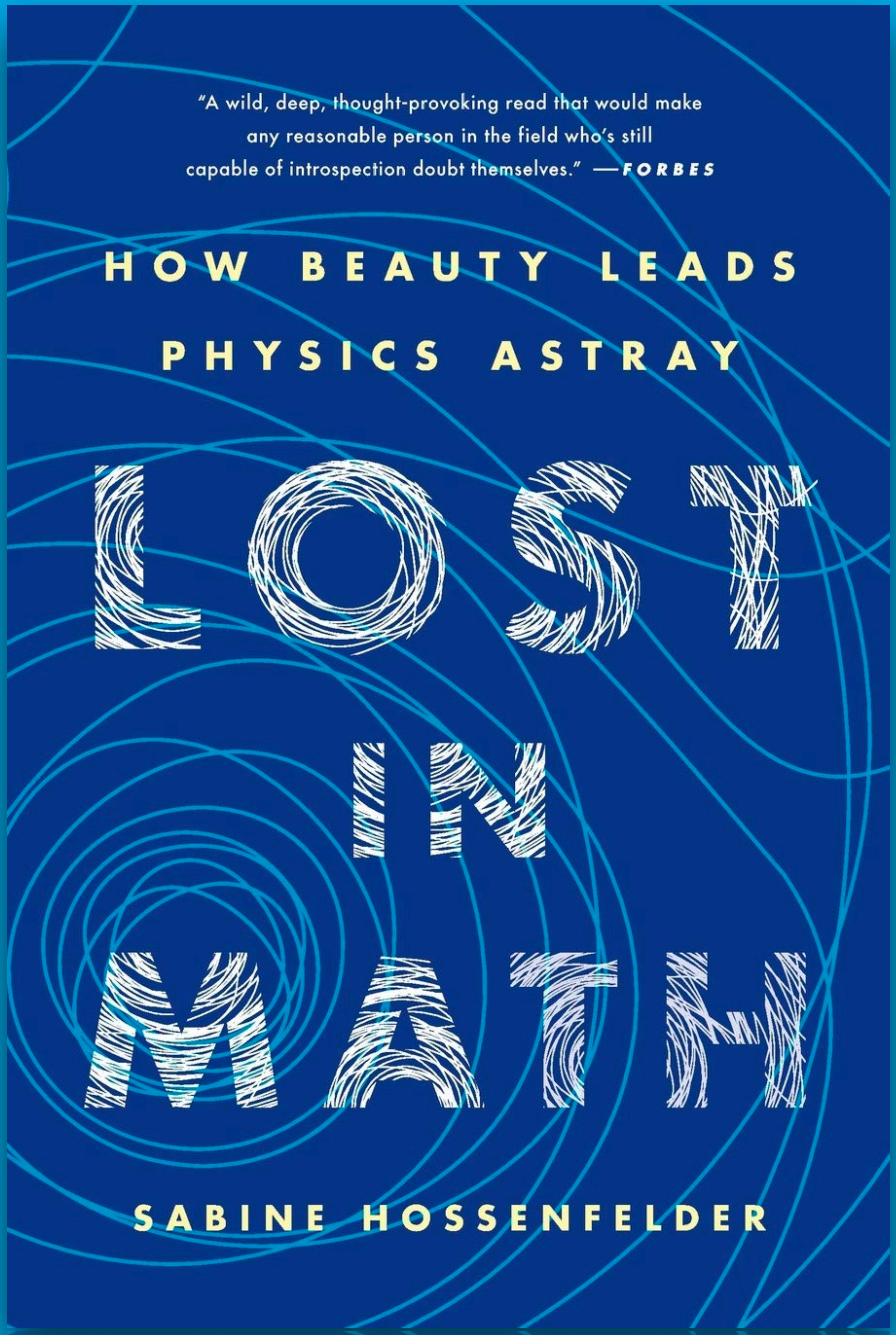
AlphaZero takes a totally different approach, replacing these hand-crafted rules with a **deep neural network** and **general purpose algorithms** that know nothing about the game beyond the basic rules."

"The amount of **training** the network needs depends on the style and complexity of the game, taking **approximately 9 hours for chess,** 12 hours for shogi, and 13 days for Go."

"In Chess, for example, it searches **only 60 thousand positions** per second in chess, compared to roughly 60 million for Stockfish."

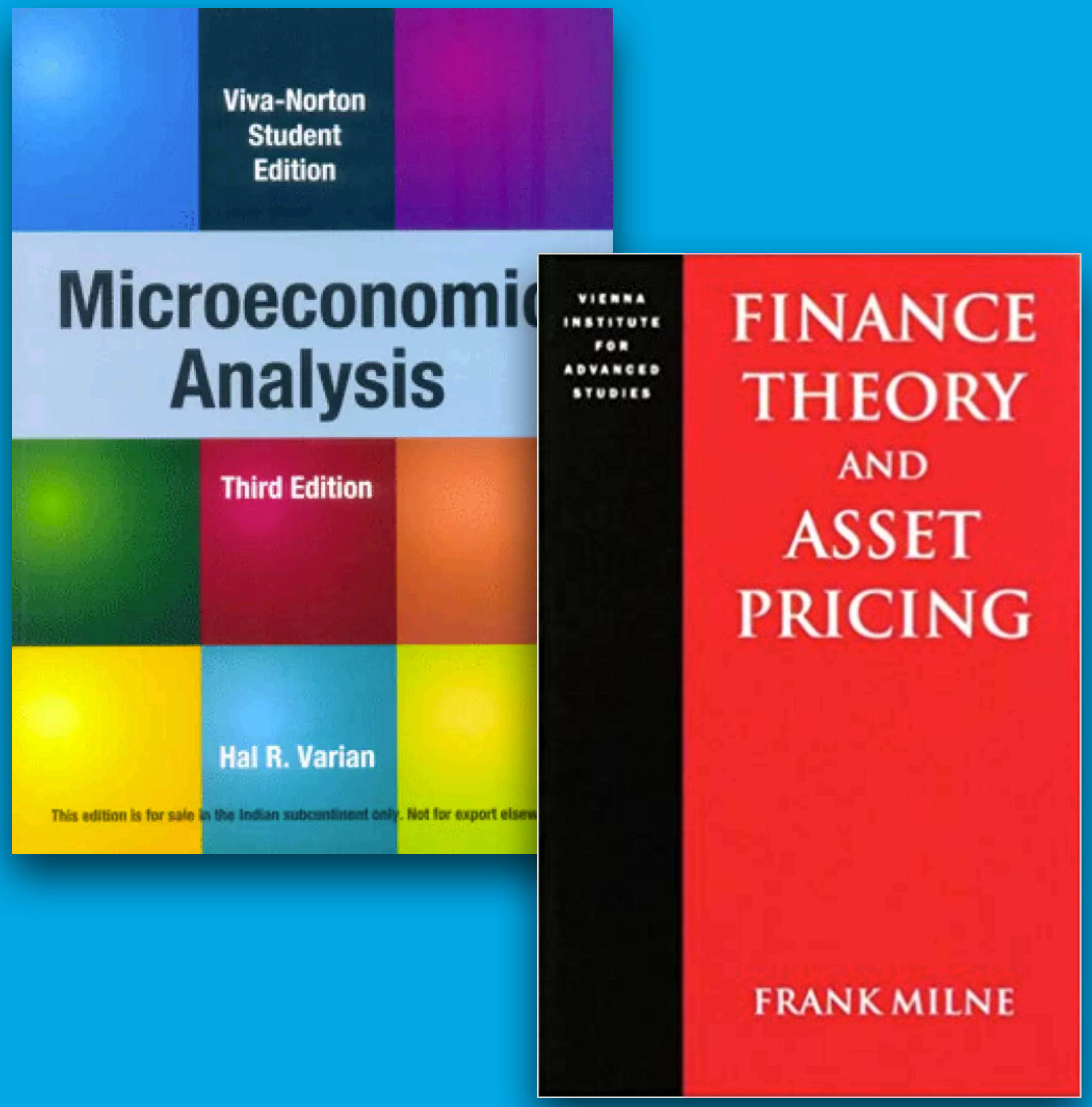
Source: http://deepmind.com

Physics Envy &
The Beauty Myth



Sabine Hossenfelder (2018): Lost in Math – How Beauty Leads Physics Astray.

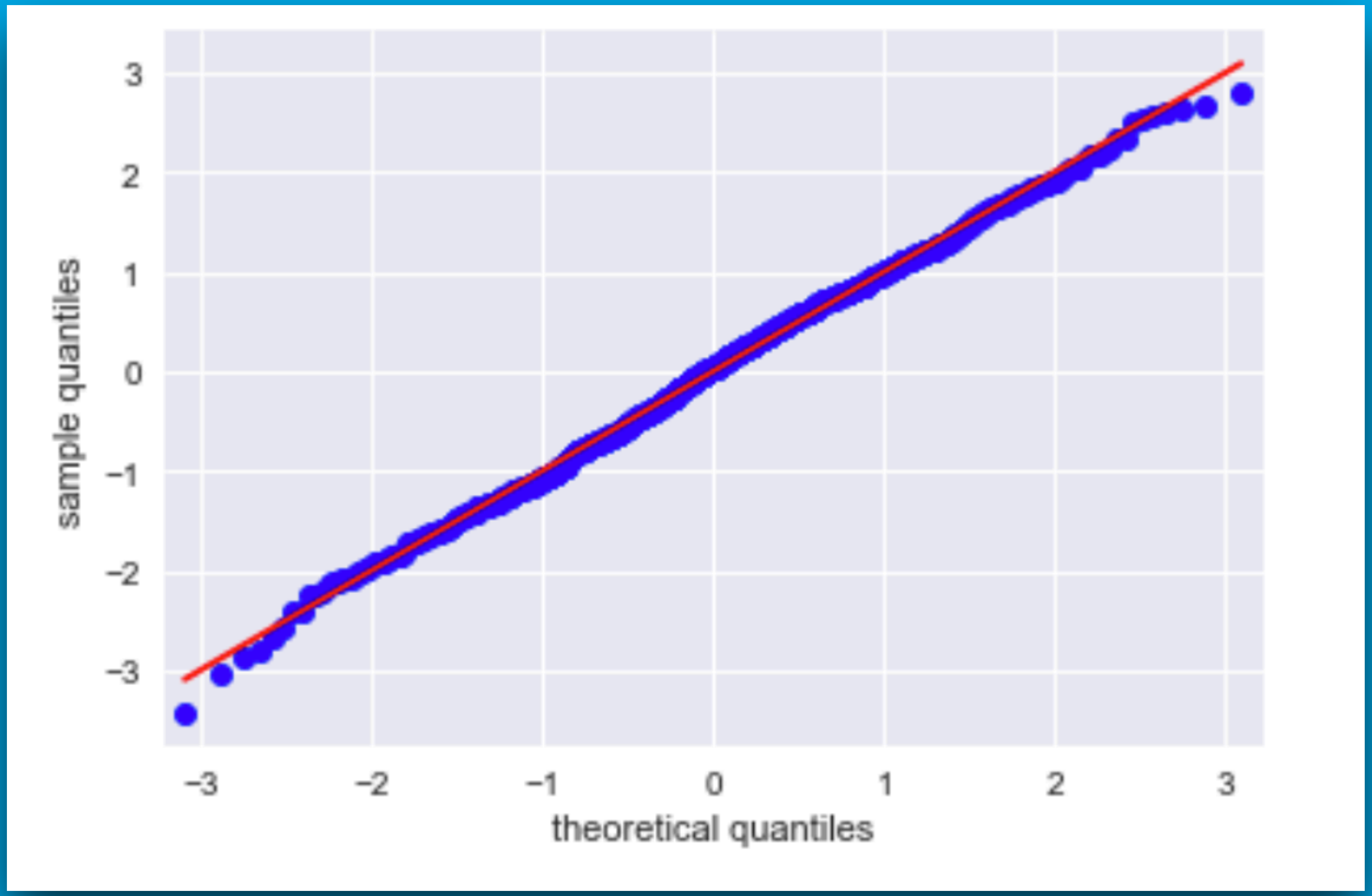
"They were so sure, they bet billions on it. For decades physicists told us they knew where the next discoveries were waiting. ... The experiments didn't reveal anything new. What failed physicists wasn't their math; it was their choice of math. They believed that Mother Nature was elegant, simple, and kind about providing clues. They thought they could hear her whisper when they were talking to themselves. Now Nature spoke, and she said nothing, loud and clear."



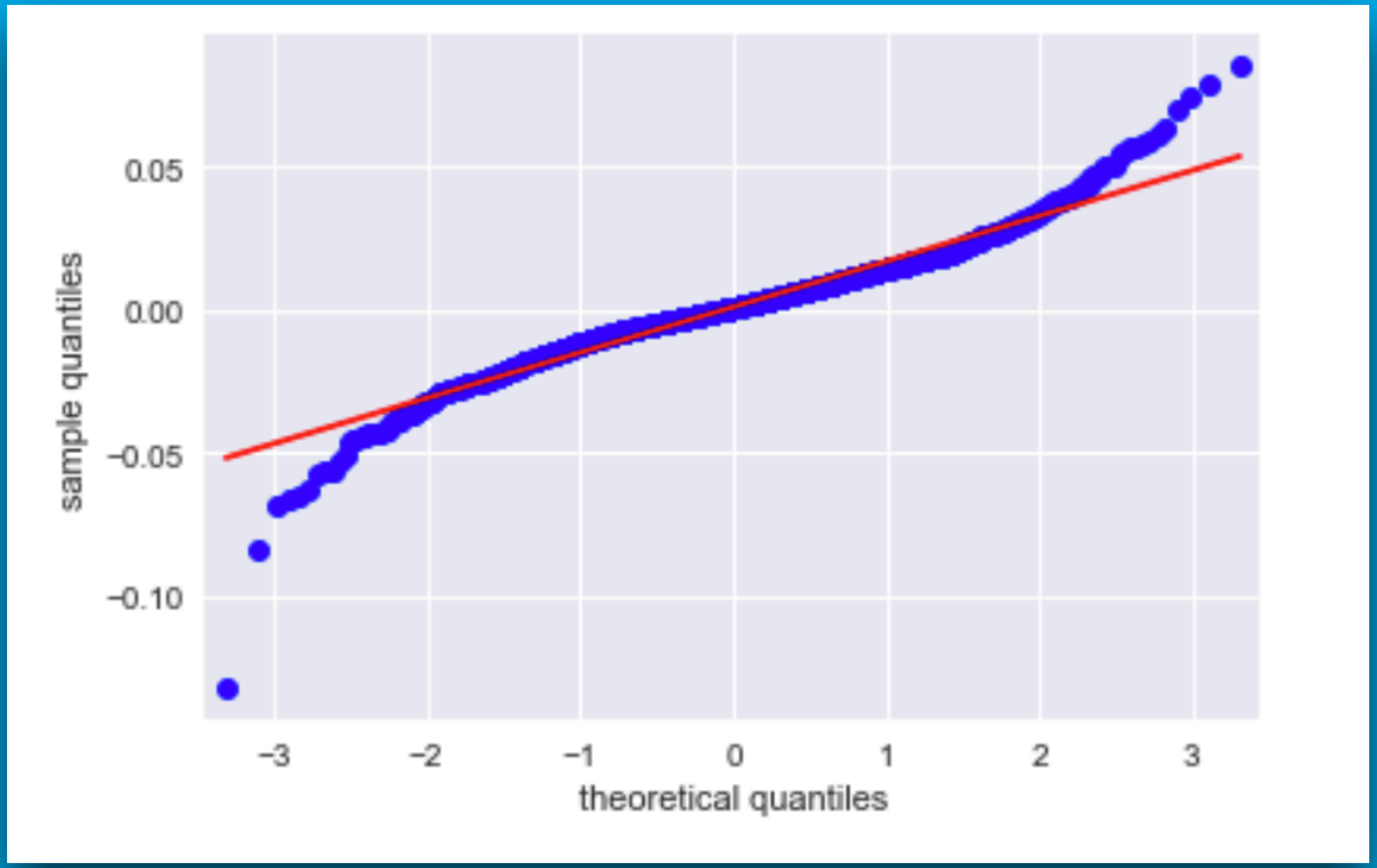
Cornerstones of Economics

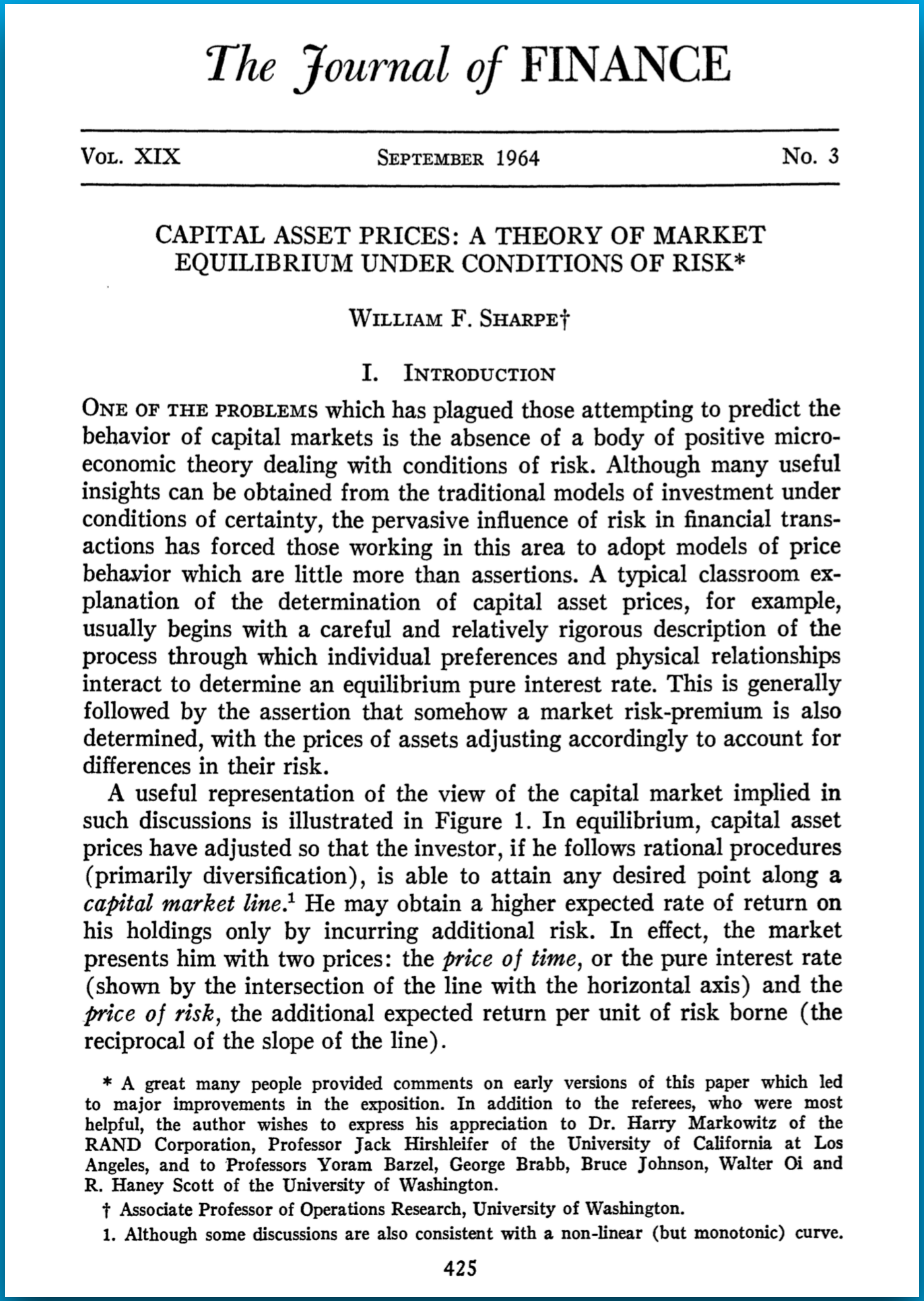
A. Expected Utility (✗)
B. Equilibrium Theory (✗)
C. Normal Distributions (✗)
D. Linear Relationships (✗)
E. Efficient Markets (✓)
F. Arbitrage Pricing (✓)

Theory



Reality





The Journal of FINANCE

VOL. XIX SEPTEMBER 1964 No. 3

CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK*

WILLIAM F. SHARPE†

I. INTRODUCTION

ONE OF THE PROBLEMS which has plagued those attempting to predict the behavior of capital markets is the absence of a body of positive microeconomic theory dealing with conditions of risk. Although many useful insights can be obtained from the traditional models of investment under conditions of certainty, the pervasive influence of risk in financial transactions has forced those working in this area to adopt models of price behavior which are little more than assertions. A typical classroom explanation of the determination of capital asset prices, for example, usually begins with a careful and relatively rigorous description of the process through which individual preferences and physical relationships interact to determine an equilibrium pure interest rate. This is generally followed by the assertion that somehow a market risk-premium is also determined, with the prices of assets adjusting accordingly to account for differences in their risk.

A useful representation of the view of the capital market implied in such discussions is illustrated in Figure 1. In equilibrium, capital asset prices have adjusted so that the investor, if he follows rational procedures (primarily diversification), is able to attain any desired point along a *capital market line*.[1] He may obtain a higher expected rate of return on his holdings only by incurring additional risk. In effect, the market presents him with two prices: the *price of time*, or the pure interest rate (shown by the intersection of the line with the horizontal axis) and the *price of risk*, the additional expected return per unit of risk borne (the reciprocal of the slope of the line).

* A great many people provided comments on early versions of this paper which led to major improvements in the exposition. In addition to the referees, who were most helpful, the author wishes to express his appreciation to Dr. Harry Markowitz of the RAND Corporation, Professor Jack Hirshleifer of the University of California at Los Angeles, and to Professors Yoram Barzel, George Brabb, Bruce Johnson, Walter Oi and R. Haney Scott of the University of Washington.

† Associate Professor of Operations Research, University of Washington.

1. Although some discussions are also consistent with a non-linear (but monotonic) curve.

$$\mu_i = r + \beta_i(\mu_M - r)$$

"Market Risk"

"Idiosyncratic Risk"

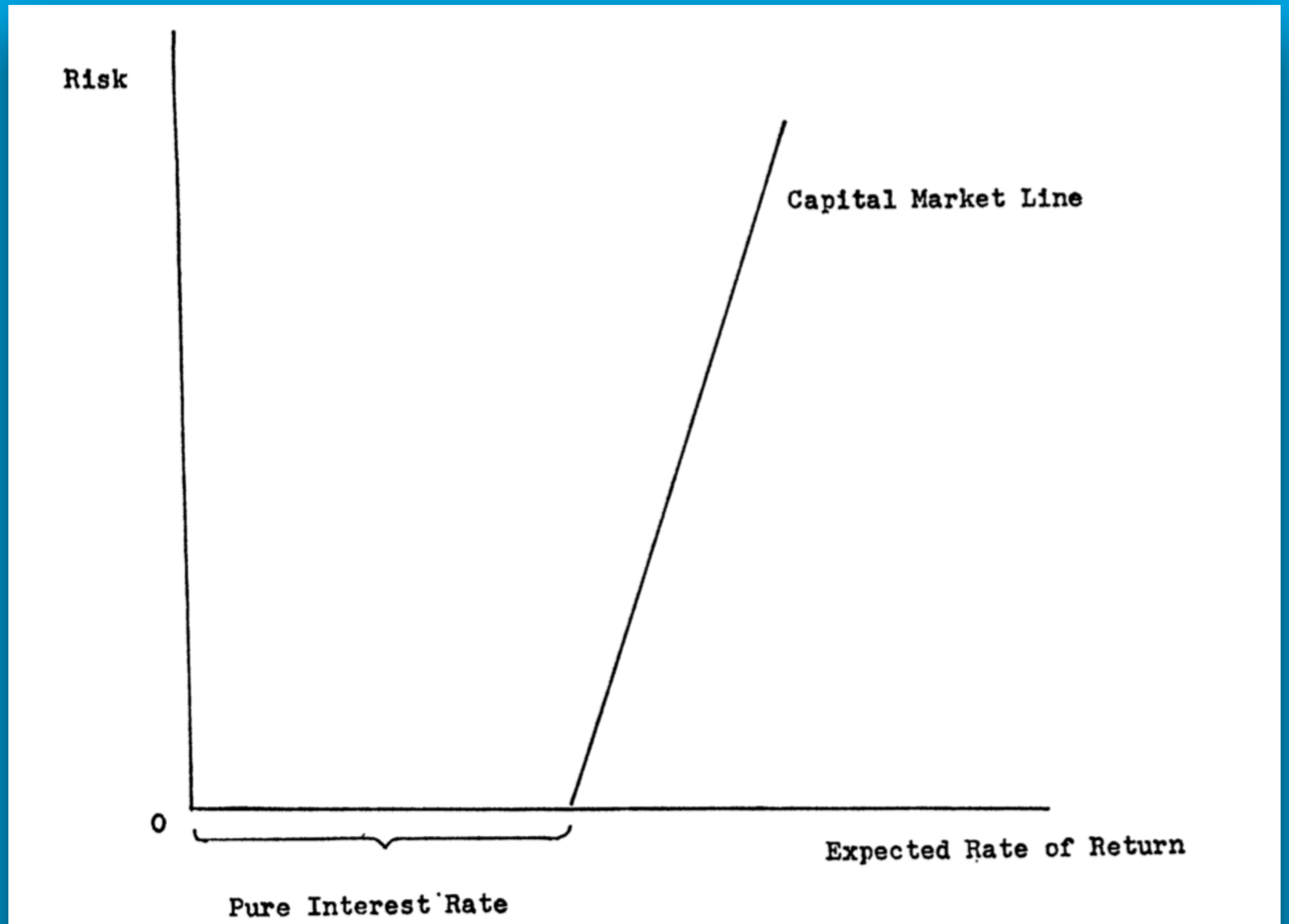
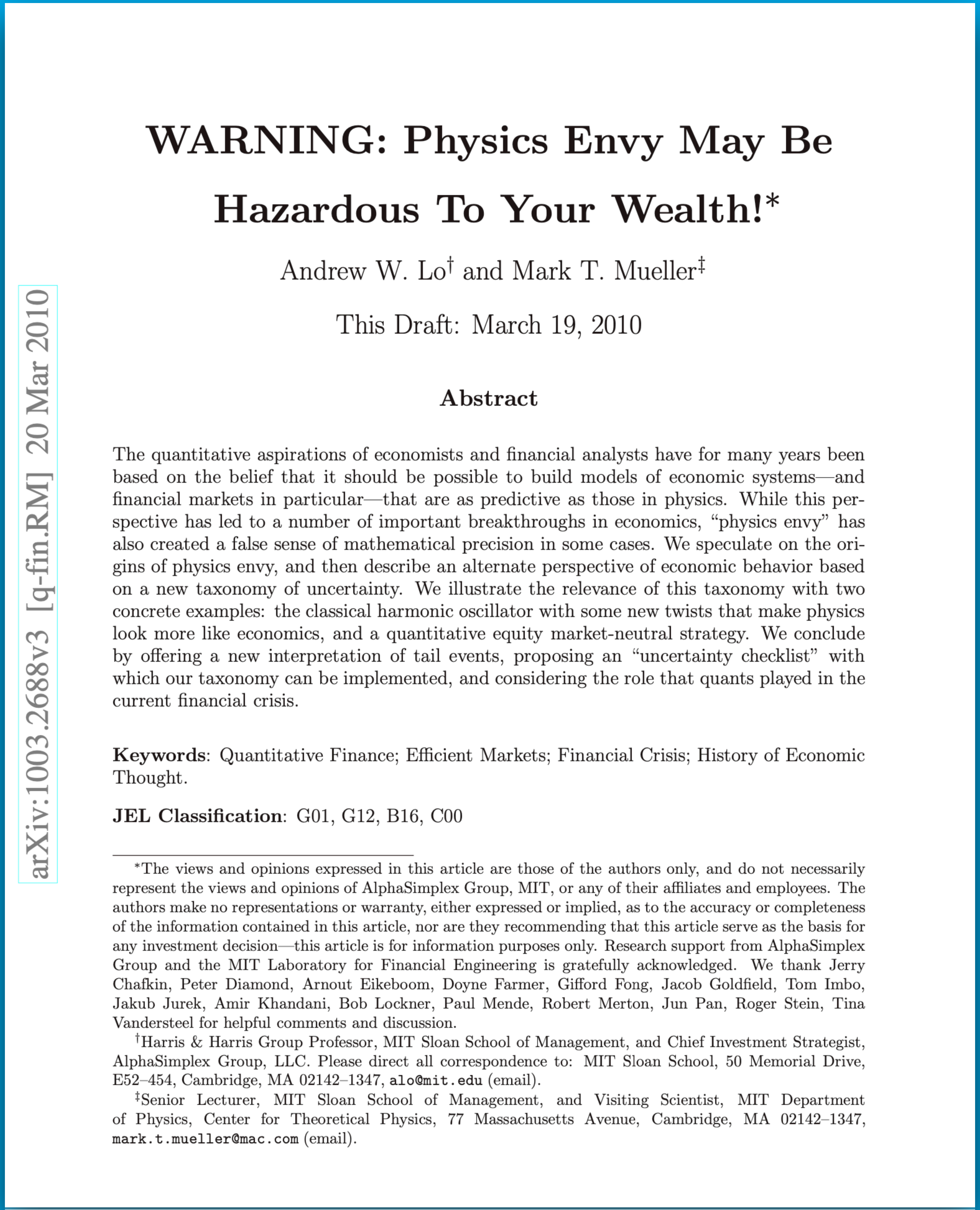


FIGURE 1



arXiv:1003.2688v3 [q-fin.RM] 20 Mar 2010

WARNING: Physics Envy May Be Hazardous To Your Wealth!*

Andrew W. Lo[†] and Mark T. Mueller[‡]

This Draft: March 19, 2010

Abstract

The quantitative aspirations of economists and financial analysts have for many years been based on the belief that it should be possible to build models of economic systems—and financial markets in particular—that are as predictive as those in physics. While this perspective has led to a number of important breakthroughs in economics, "physics envy" has also created a false sense of mathematical precision in some cases. We speculate on the origins of physics envy, and then describe an alternate perspective of economic behavior based on a new taxonomy of uncertainty. We illustrate the relevance of this taxonomy with two concrete examples: the classical harmonic oscillator with some new twists that make physics look more like economics, and a quantitative equity market-neutral strategy. We conclude by offering a new interpretation of tail events, proposing an "uncertainty checklist" with which our taxonomy can be implemented, and considering the role that quants played in the current financial crisis.

Keywords: Quantitative Finance; Efficient Markets; Financial Crisis; History of Economic Thought.

JEL Classification: G01, G12, B16, C00

*The views and opinions expressed in this article are those of the authors only, and do not necessarily represent the views and opinions of AlphaSimplex Group, MIT, or any of their affiliates and employees. The authors make no representations or warranty, either expressed or implied, as to the accuracy or completeness of the information contained in this article, nor are they recommending that this article serve as the basis for any investment decision—this article is for information purposes only. Research support from AlphaSimplex Group and the MIT Laboratory for Financial Engineering is gratefully acknowledged. We thank Jerry Chafkin, Peter Diamond, Arnout Eikeboom, Doyne Farmer, Gifford Fong, Jacob Goldfield, Tom Imbo, Jakub Jurek, Amir Khandani, Bob Lockner, Paul Mende, Robert Merton, Jun Pan, Roger Stein, Tina Vandersteel for helpful comments and discussion.

[†]Harris & Harris Group Professor, MIT Sloan School of Management, and Chief Investment Strategist, AlphaSimplex Group, LLC. Please direct all correspondence to: MIT Sloan School, 50 Memorial Drive, E52–454, Cambridge, MA 02142–1347, alo@mit.edu (email).

[‡]Senior Lecturer, MIT Sloan School of Management, and Visiting Scientist, MIT Department of Physics, Center for Theoretical Physics, 77 Massachusetts Avenue, Cambridge, MA 02142–1347, mark.t.mueller@mac.com (email).

"The quantitative aspirations of economists and financial analysts have for many years been based on the belief that it should be possible to build models of economic systems – and financial markets in particular – that are as predictive as those in physics. While this perspective has led to a number of important breakthroughs in economics, 'physics envy' has also created a false sense of mathematical precision in some cases."

Data-Driven Finance


FINANCIAL TIMES

COP clock runs down

Protest legacy

Huge Hong Kong police population keeps itself busy

Iran impasse

Long wait for liberty goes on

Briefing

Xi cements grip on China after vote puts him on par with Mao

Party says nation 'rejuvenated' • Third five-year term likely • New emphasis on socialism

Datawatch

Johnson's £4m outside earnings open him to charges of hypocrisy from MPs

De Klerk, last president of apartheid era, dies at 85

World Markets


THE WALL STREET JOURNAL.

WEDNESDAY, NOVEMBER 10, 2021 - VOL. CCLXXVIII NO. 112

WSJ.com

★★★★ $5.00

What's News

Business & Finance

World-Wide

GE Breakup Marks End of an Era

Company that symbolized U.S. manufacturing power to split into three parts

By Thomas Gryta

Over the past four decades, GE's revenue swelled and ebbed under four CEOs as it focused on different businesses, including a shift away from consumer goods and a bet on finance.

GE's revenue by segment

Tax Bill Looms as Musk Eyes Share Sale

Facebook Rife With Stolen Content

the facebook files

Latest in a series

Meta Curbs Targeted Ads

A $115,000 USC Degree Yields Low Pay, Huge Debts

Elite school hired for-profit firm to woo students to social-work master's

THE FUTURE OF EVERYTHING

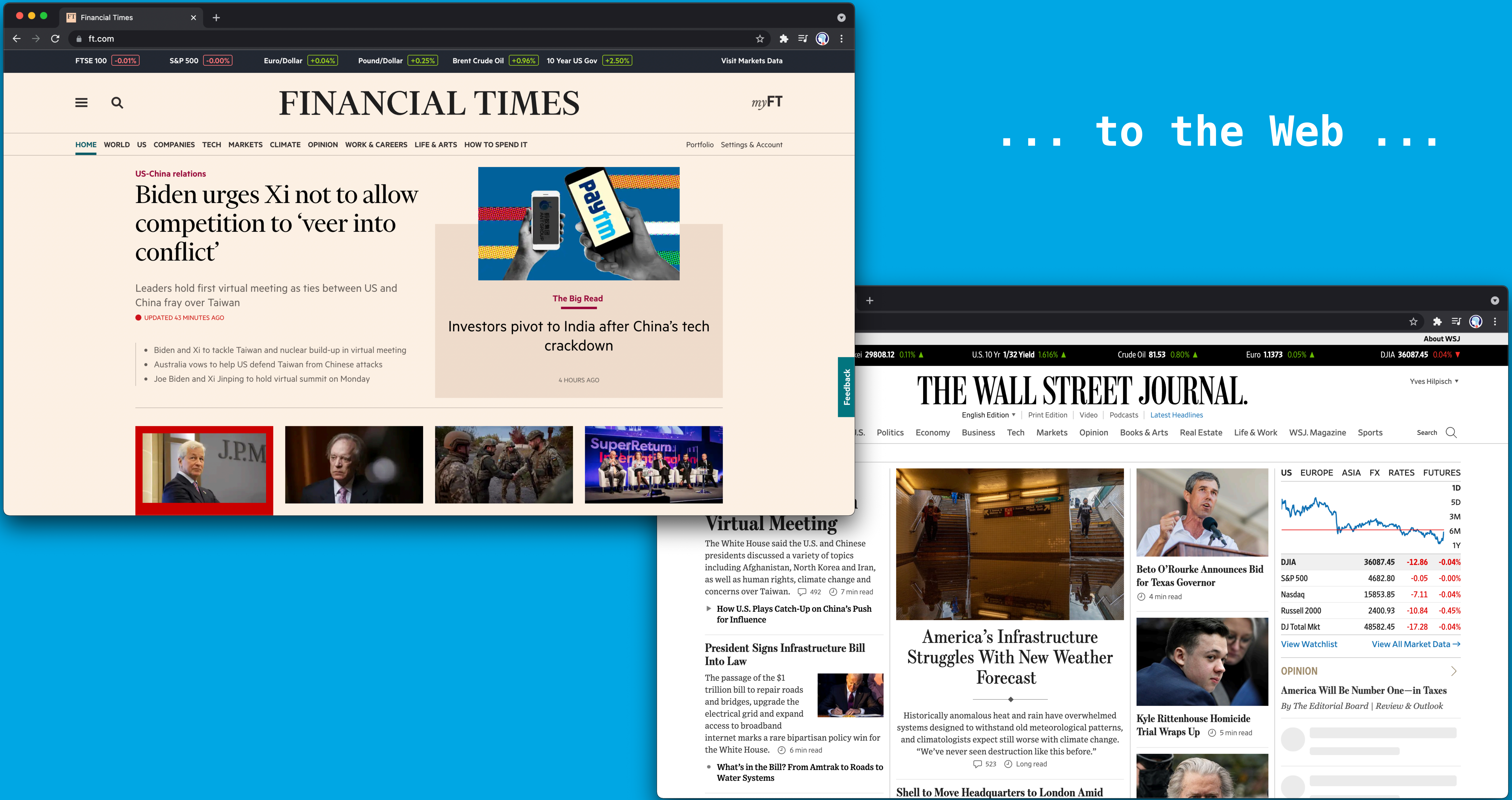
THE (ALMOST) CAR-FREE CITY

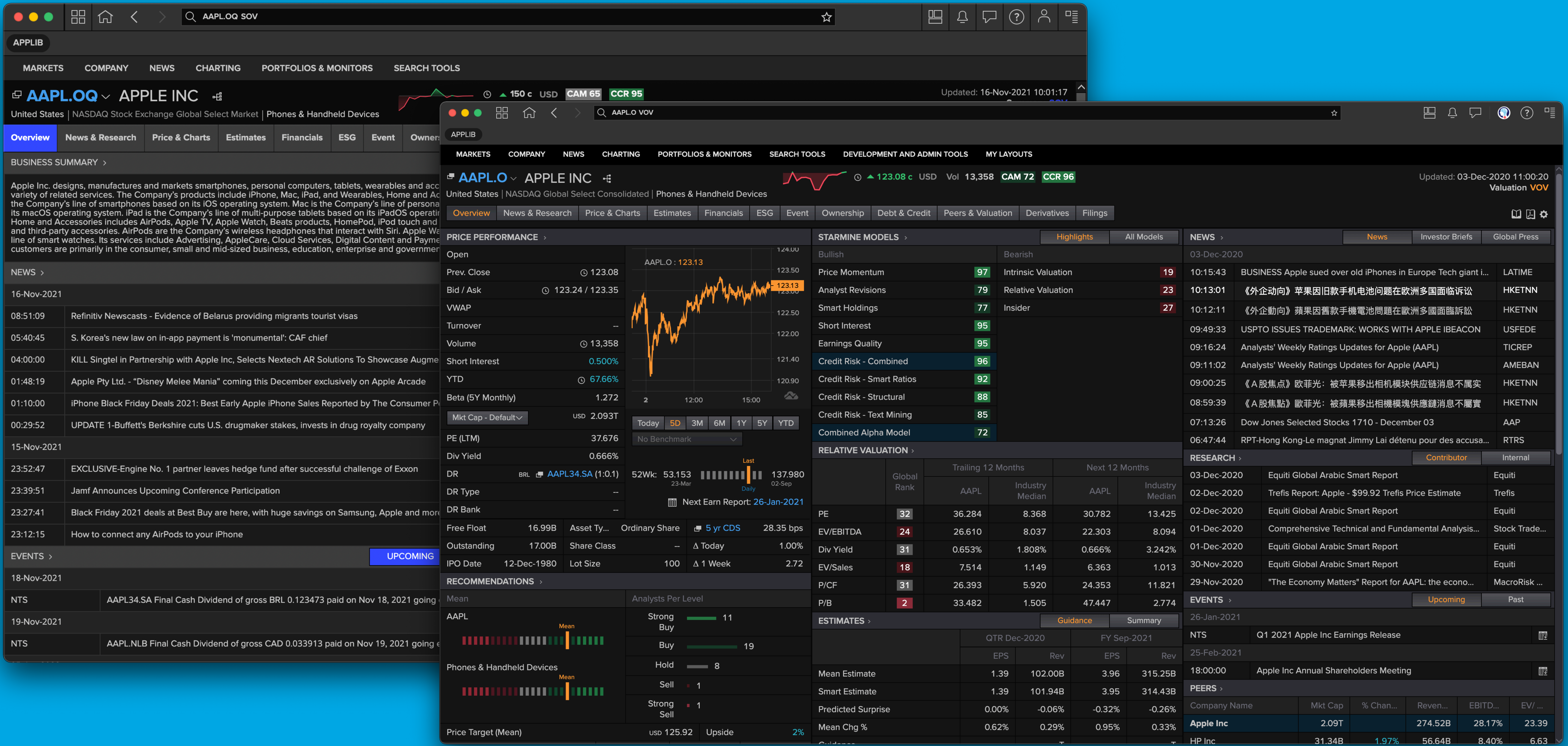
Democracy Loses Favor

Is Santa

From print …

... to the Web ...


Financial Times
ft.com
FTSE 100 -0.01%
S&P 500 -0.00%
Euro/Dollar +0.04%
Pound/Dollar +0.25%
Brent Crude Oil +0.96%
10 Year US Gov +2.50%
Visit Markets Data
FINANCIAL TIMES
myFT
HOME WORLD US COMPANIES TECH MARKETS CLIMATE OPINION WORK & CAREERS LIFE & ARTS HOW TO SPEND IT
Portfolio Settings & Account
US-China relations
Biden urges Xi not to allow competition to 'veer into conflict'
Leaders hold first virtual meeting as ties between US and China fray over Taiwan
UPDATED 43 MINUTES AGO
Biden and Xi to tackle Taiwan and nuclear build-up in virtual meeting
Australia vows to help US defend Taiwan from Chinese attacks
Joe Biden and Xi Jinping to hold virtual summit on Monday
The Big Read
Investors pivot to India after China's tech crackdown
4 HOURS AGO
Feedback
About WSJ
29808.12 0.11%
U.S. 10 Yr 1/32 Yield 1.616%
Crude Oil 81.53 0.80%
Euro 1.1373 0.05%
DJIA 36087.45 0.04%
THE WALL STREET JOURNAL.
Yves Hilpisch
English Edition | Print Edition | Video | Podcasts | Latest Headlines
Politics Economy Business Tech Markets Opinion Books & Arts Real Estate Life & Work WSJ. Magazine Sports
Search
Virtual Meeting
The White House said the U.S. and Chinese presidents discussed a variety of topics including Afghanistan, North Korea and Iran, as well as human rights, climate change and concerns over Taiwan. 492 7 min read
How U.S. Plays Catch-Up on China's Push for Influence
President Signs Infrastructure Bill Into Law
The passage of the $1 trillion bill to repair roads and bridges, upgrade the electrical grid and expand access to broadband internet marks a rare bipartisan policy win for the White House. 6 min read
What's in the Bill? From Amtrak to Roads to Water Systems
America's Infrastructure Struggles With New Weather Forecast
Historically anomalous heat and rain have overwhelmed systems designed to withstand old meteorological patterns, and climatologists expect still worse with climate change. "We've never seen destruction like this before."
523 Long read
Shell to Move Headquarters to London Amid
Beto O'Rourke Announces Bid for Texas Governor
4 min read
Kyle Rittenhouse Homicide Trial Wraps Up 5 min read
US EUROPE ASIA FX RATES FUTURES
1D 5D 3M 6M 1Y
DJIA 36087.45 -12.86 -0.04%
S&P 500 4682.80 -0.05 -0.00%
Nasdaq 15853.85 -7.11 -0.04%
Russell 2000 2400.93 -10.84 -0.45%
DJ Total Mkt 48582.45 -17.28 -0.04%
View Watchlist
View All Market Data
OPINION
America Will Be Number One—in Taxes
By The Editorial Board | Review & Outlook



... to the Terminal ...

Tick Data

[4]:
```
%%time
data = ek.get_timeseries('AAPL.O',
                         start_date='2021-11-15 15:00:00',
                         end_date='2021-11-15 15:30:00',
                         interval='tick',
                         fields=['*'])
```

```
CPU times: user 120 ms, sys: 14.5 ms, total: 135 ms
Wall time: 2.7 s
```

[5]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 41213 entries, 2021-11-15 15:00:00.004000 to 2021-11-15 15:29:59.936000
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   VALUE   41145 non-null  Float64
 1   VOLUME  41213 non-null  Int64
dtypes: Float64(1), Int64(1)
memory usage: 1.0 MB
```

[14]:
```
data.tail()
```

[14]:

AAPL.O	VALUE	VOLUME
Date		
2021-11-15 15:29:59.134	150.4456	10
2021-11-15 15:29:59.313	150.445	1
2021-11-15 15:29:59.588	150.4409	150
2021-11-15 15:29:59.745	150.445	1
2021-11-15 15:29:59.936	150.4488	5

... to powerful APIs.

Apple Event on 18. October 2021 (https://www.apple.com/de/apple-events/october-2021/).

```
[9]: headlines = ek.get_news_headlines(query='R:AAPL.O macbook',
                                       count=5,
                                       date_from='2021-10-18',
                                       date_to='2021-10-19')
```

```
[10]: headlines
```

[10]:

	versionCreated	text	storyId	sourceCode
2021-10-18 23:30:18.401	2021-10-18 23:30:18.401000+00:00	Apple is finally fixing the things people hate...	urn:newsml:reuters.com:20211018:nNRAh2psl1:1	NS:WASHPO
2021-10-18 23:10:18.012	2021-10-18 23:10:18.012000+00:00	Apple event – live: Macbook Pro and other new ...	urn:newsml:reuters.com:20211018:nNRAh2kj3a:1	NS:INDEPE
2021-10-18 21:41:19.927	2021-10-18 21:41:19.927000+00:00	New MacBook Pro features ultra-fast chips, ret...	urn:newsml:reuters.com:20211018:nNRAh2u38b:1	NS:EFEING
2021-10-18 21:33:50.860	2021-10-18 21:33:50.860000+00:00	Apple Event: MacBook Pro 2021 alleged pictures...	urn:newsml:reuters.com:20211018:nNRAh2u1wj:1	NS:TIMIND
2021-10-18 21:33:50.623	2021-10-18 21:33:50.623000+00:00	Apple launches new MacBook Pro: Price, specs a...	urn:newsml:reuters.com:20211018:nNRAh2u1vv:1	NS:TIMIND

... to powerful APIs.

```
[11]: story = headlines.iloc[0]
```

```
[12]: story
```

```
[12]: versionCreated                    2021-10-18 23:30:18.401000+00:00
      text            Apple is finally fixing the things people hate...
      storyId               urn:newsml:reuters.com:20211018:nNRAh2psl1:1
      sourceCode                                              NS:WASHPO
      Name: 2021-10-18 23:30:18.401000, dtype: object
```

```
[13]: news_text = ek.get_news_story(story['storyId'])
```

```
[14]: from IPython.display import HTML
```

```
[15]: HTML(news_text)
```

[15]: The demise of MagSafe charging. An inelegant Touch Bar. Limited selection of ports. The laundry list of complaints about Apple's laptops has steadily grown over the past five years. Now, Apple is finally walking back those changes.

On Monday, the Cupertino, Calif., company unveiled a pair of new MacBook Pro laptops, powered by its latest homegrown processors and free of the many limitations that plagued earlier models. It also showed off a set of updated AirPods and colorful HomePod mini smart speakers. Riding high from record Mac sales last year, Apple made sure to make its new MacBooks the star of its virtual event Monday.

Still, computers that run Apple's MacOS software account for only a fraction of the overall PC landscape — just over 7 percent as of the end of the second quarter, according to market research firm IDC. Its market share has slipped from 8 percent in the first quarter and 7.6 percent a year earlier, IDC data showed. The changes on display Monday seem to be geared more toward

... to powerful APIs.

EXPERT OPINION

Contact Editor: **Brian Brannon**, bbrannon@computer.org

The Unreasonable Effectiveness of Data

Alon Halevy, Peter Norvig, and Fernando Pereira, *Google*

Eugene Wigner's article "The Unreasonable Effectiveness of Mathematics in the Natural Sciences"[1] examines why so much of physics can be neatly explained with simple mathematical formulas such as $f = ma$ or $e = mc^2$. Meanwhile, sciences that involve human beings rather than elementary particles have proven more resistant to elegant mathematics. Economists suffer from physics envy over their inability to neatly model human behavior. An informal, incomplete grammar of the English language runs over 1,700 pages.[2] Perhaps when it comes to natural language processing and related fields, we're doomed to complex theories that will never have the elegance of physics equations. But if that's so, we should stop acting as if our goal is to author extremely elegant theories, and instead embrace complexity and make use of the best ally we have: the unreasonable effectiveness of data.

One of us, as an undergraduate at Brown University, remembers the excitement of having access to the Brown Corpus, containing one million English words.[3] Since then, our field has seen several notable corpora that are about 100 times larger, and in 2006, Google released a trillion-word corpus with frequency counts for all sequences up to five words long.[4] In some ways this corpus is a step backwards from the Brown Corpus: it's taken from unfiltered Web pages and thus contains incomplete sentences, spelling errors, grammatical errors, and all sorts of other errors. It's not annotated with carefully hand-corrected part-of-speech tags. But the fact that it's a million times larger than the Brown Corpus outweighs these drawbacks. A trillion-word corpus—along with other Web-derived corpora of millions, billions, or trillions of links, videos, images, tables, and user interactions—captures even very rare aspects of human behavior. So, this corpus could serve as the basis of a complete model for certain tasks—if only we knew how to extract the model from the data.

Learning from Text at Web Scale

The biggest successes in natural-language-related machine learning have been statistical speech recognition and statistical machine translation. The reason for these successes is not that these tasks are easier than other tasks; they are in fact much harder than tasks such as document classification that extract just a few bits of information from each document. The reason is that translation is a natural task routinely done every day for a real human need (think of the operations of the European Union or of news agencies). The same is true of speech transcription (think of closed-caption broadcasts). In other words, a large training set of the input-output behavior that we seek to automate is available to us *in the wild*. In contrast, traditional natural language processing problems such as document classification, part-of-speech tagging, named-entity recognition, or parsing are not routine tasks, so they have no large corpus available in the wild. Instead, a corpus for these tasks requires skilled human annotation. Such annotation is not only slow and expensive to acquire but also difficult for experts to agree on, being bedeviled by many of the difficulties we discuss later in relation to the Semantic Web. The first lesson of Web-scale learning is to use available large-scale data rather than hoping for annotated data that isn't available. For instance, we find that useful semantic relationships can be automatically learned from the statistics of search queries and the corresponding results[5] or from the accumulated evidence of Web-based text patterns and formatted tables,[6] in both cases without needing any manually annotated data.

8 1541-1672/09/$25.00 © 2009 IEEE *Published by the IEEE Computer Society* IEEE INTELLIGENT SYSTEMS

Eugene Wigner's article "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" examines why so much of physics can be neatly explained with simple mathematical formulas such as f = ma or e = mc^2. Meanwhile, sciences that involve human beings rather than elementary particles have proven more resistant to elegant mathematics. Economists suffer from physics envy over their inability to neatly [and successfully] model human behavior. An informal, incomplete grammar of the English language runs over 1,700 pages. Perhaps when it comes to natural language processing and related fields, we're doomed to complex theories that will never have the elegance of physics equations. But if that's so, we should stop acting as if our goal is to author extremely elegant theories, and instead embrace complexity and make use of the best ally we have: the unreasonable effectiveness of data.

IP[y]:
IPython

jupyter

NumPy

matplotlib

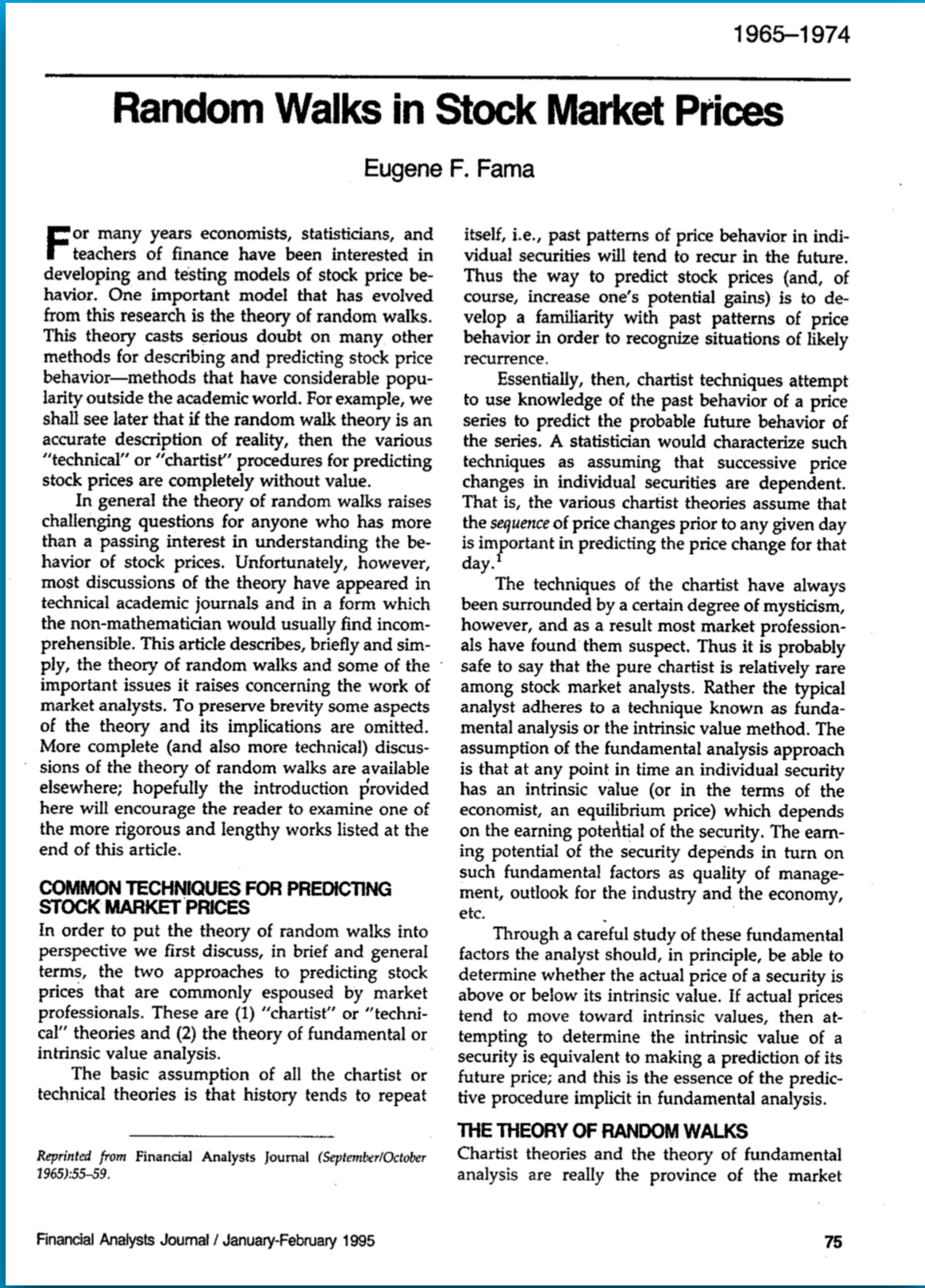
pandas

$y_it = \beta' x_{it} + \mu_i + \epsilon_{it}$

xlwings

X

Efficient Markets



Random Walks in Stock Market Prices

Eugene F. Fama

For many years economists, statisticians, and teachers of finance have been interested in developing and testing models of stock price behavior. One important model that has evolved from this research is the theory of random walks. This theory casts serious doubt on many other methods for describing and predicting stock price behavior—methods that have considerable popularity outside the academic world. For example, we shall see later that if the random walk theory is an accurate description of reality, then the various "technical" or "chartist" procedures for predicting stock prices are completely without value.

In general the theory of random walks raises challenging questions for anyone who has more than a passing interest in understanding the behavior of stock prices. Unfortunately, however, most discussions of the theory have appeared in technical academic journals and in a form which the non-mathematician would usually find incomprehensible. This article describes, briefly and simply, the theory of random walks and some of the important issues it raises concerning the work of market analysts. To preserve brevity some aspects of the theory and its implications are omitted. More complete (and also more technical) discussions of the theory of random walks are available elsewhere; hopefully the introduction provided here will encourage the reader to examine one of the more rigorous and lengthy works listed at the end of this article.

COMMON TECHNIQUES FOR PREDICTING STOCK MARKET PRICES

In order to put the theory of random walks into perspective we first discuss, in brief and general terms, the two approaches to predicting stock prices that are commonly espoused by market professionals. These are (1) "chartist" or "technical" theories and (2) the theory of fundamental or intrinsic value analysis.

The basic assumption of all the chartist or technical theories is that history tends to repeat itself, i.e., past patterns of price behavior in individual securities will tend to recur in the future. Thus the way to predict stock prices (and, of course, increase one's potential gains) is to develop a familiarity with past patterns of price behavior in order to recognize situations of likely recurrence.

Essentially, then, chartist techniques attempt to use knowledge of the past behavior of a price series to predict the probable future behavior of the series. A statistician would characterize such techniques as assuming that successive price changes in individual securities are dependent. That is, the various chartist theories assume that the *sequence* of price changes prior to any given day is important in predicting the price change for that day.[1]

The techniques of the chartist have always been surrounded by a certain degree of mysticism, however, and as a result most market professionals have found them suspect. Thus it is probably safe to say that the pure chartist is relatively rare among stock market analysts. Rather the typical analyst adheres to a technique known as fundamental analysis or the intrinsic value method. The assumption of the fundamental analysis approach is that at any point in time an individual security has an intrinsic value (or in the terms of the economist, an equilibrium price) which depends on the earning potential of the security. The earning potential of the security depends in turn on such fundamental factors as quality of management, outlook for the industry and the economy, etc.

Through a careful study of these fundamental factors the analyst should, in principle, be able to determine whether the actual price of a security is above or below its intrinsic value. If actual prices tend to move toward intrinsic values, then attempting to determine the intrinsic value of a security is equivalent to making a prediction of its future price; and this is the essence of the predictive procedure implicit in fundamental analysis.

THE THEORY OF RANDOM WALKS

Chartist theories and the theory of fundamental analysis are really the province of the market

Reprinted from Financial Analysts Journal *(September/October 1965):55–59.*

Eugene F. Fama (1965):

"For many years, economists, statisticians, and teachers of finance have been interested in developing and testing models of stock price behavior. One important model that has evolved from this research is the theory of random walks. This theory casts serious doubt on many other methods for describing and predicting stock price behavior—methods that have considerable popularity outside the academic world. For example, we shall see later that, if the random-walk theory is an accurate description of reality, then the various "technical" or "chartist" procedures for predicting stock prices are completely without value."—Eugene F. Fama (1965): "Random Walks in Stock Market Prices"

Michael Jensen (1978): “Some Anomalous Evidence Regarding Market Efficiency”:

“A market is efficient with respect to an information set S if it is impossible to make economic profits by trading on the basis of information set S.”

If a stock price follows a (simple) random walk (no drift & normally distributed returns), then it rises and falls with the same probability of 50% (“toss of a coin”).

In such a case, the best predictor of tomorrow’s stock price – in a least-squares sense – is today’s stock price.

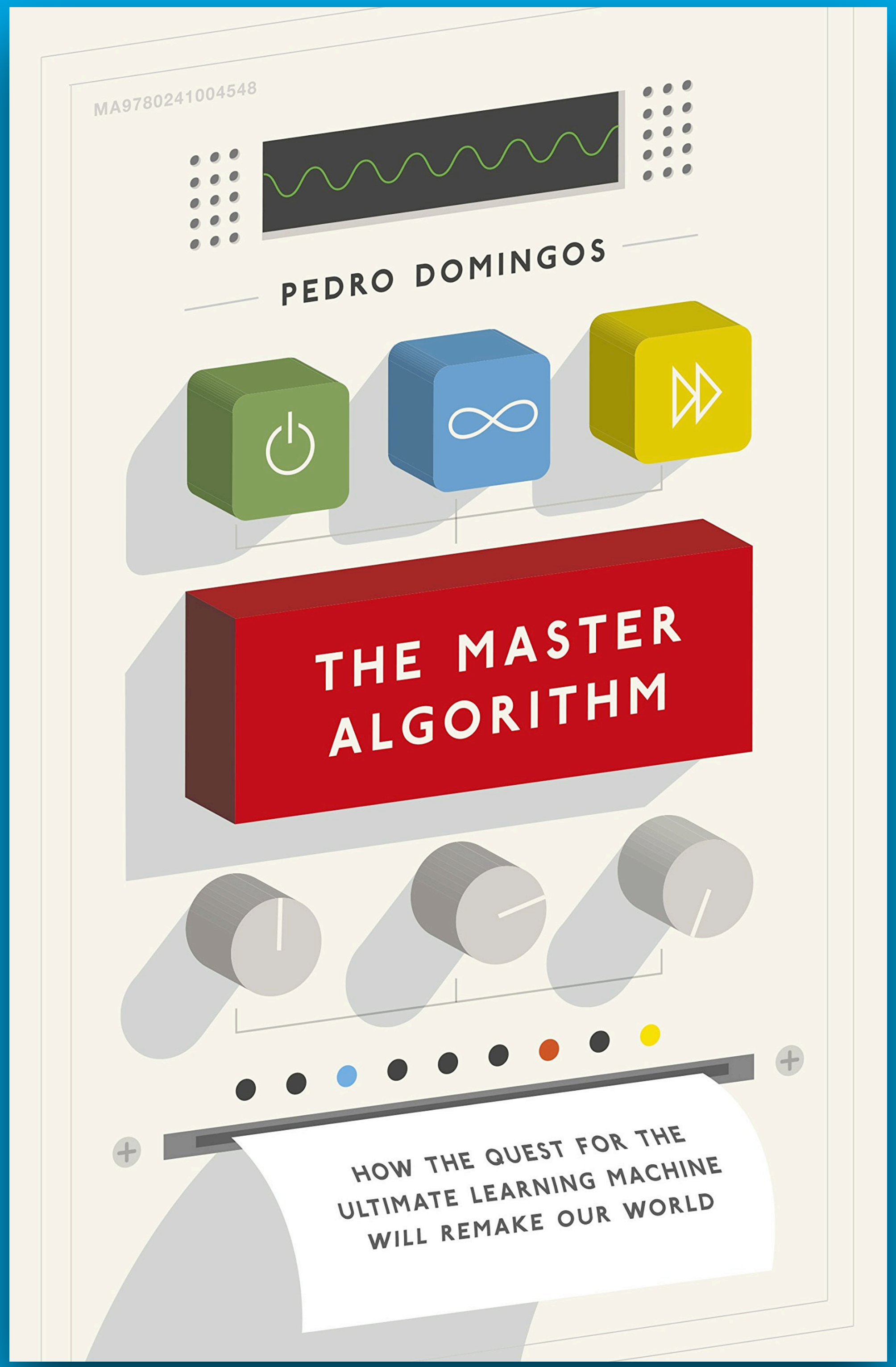
AI-First Finance

scientific method

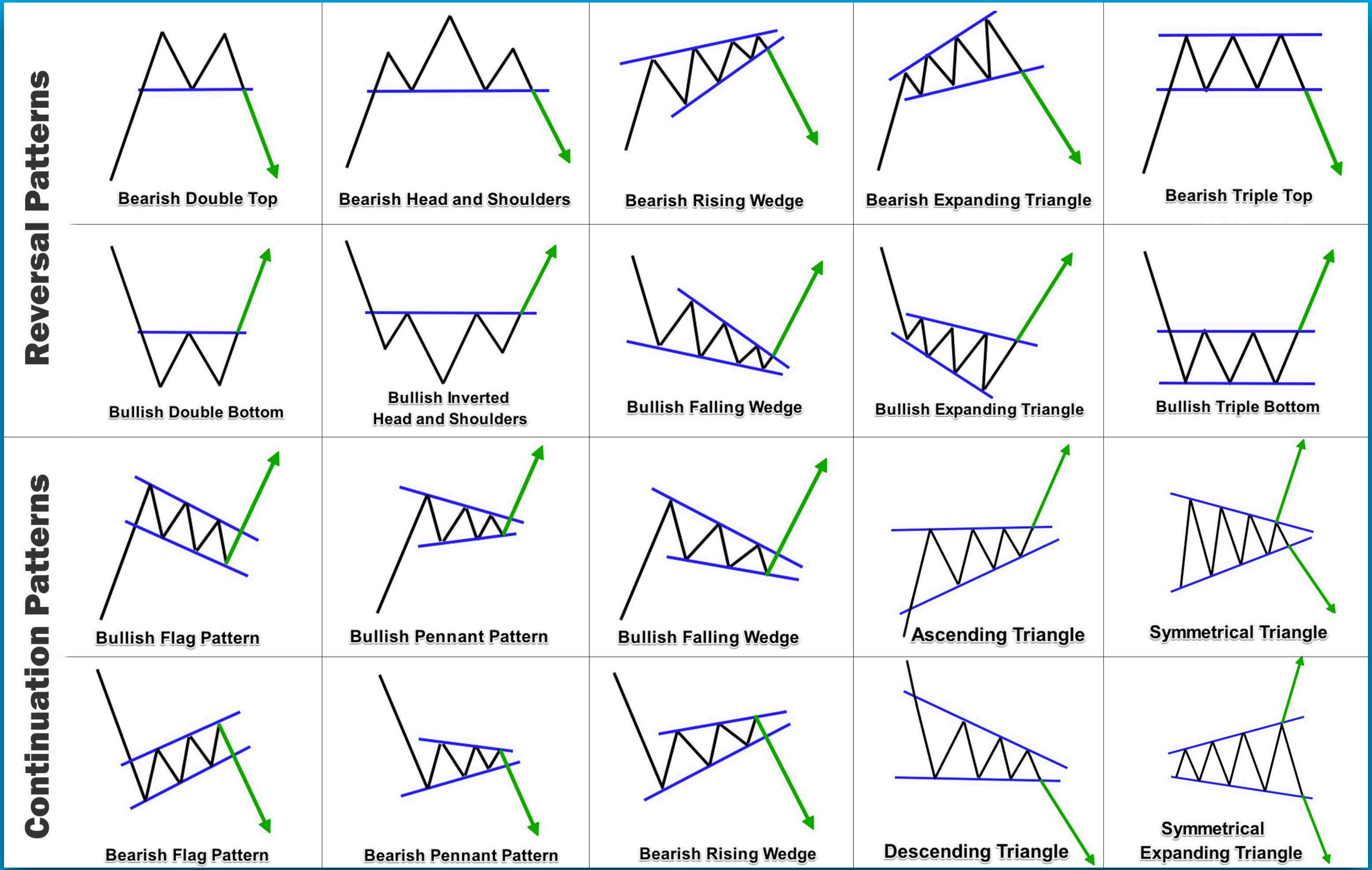
noun

a method of procedure that has characterized natural science since the 17th century, consisting in systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses.

"criticism is the backbone of **the scientific method**"

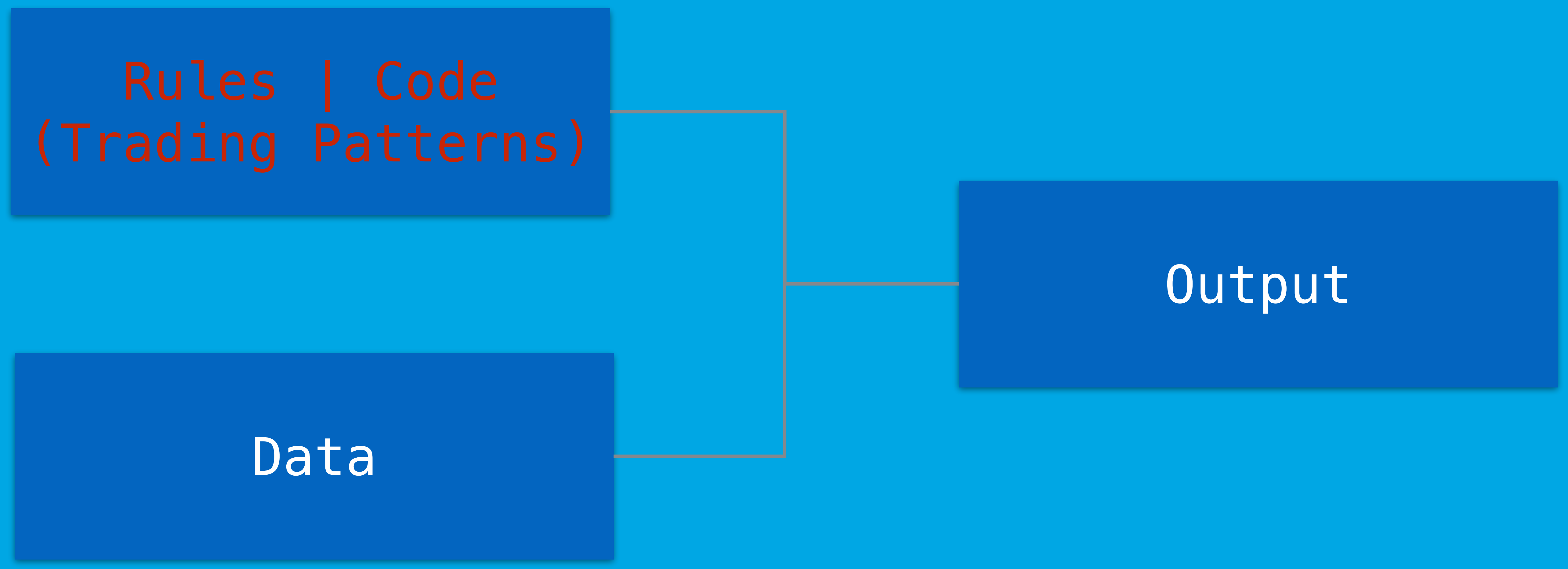


"Machine learning is the scientific method on steroids. It follows the same process of generating, testing, and discarding or refining hypotheses. But while a scientist may spend his or her whole life coming up with and testing a few hundred hypotheses, a machine-learning system can do the same in a second. Machine learning automates discovery. It's no surprise, then that it's revolutionizing science as much as it's revolutionizing business."

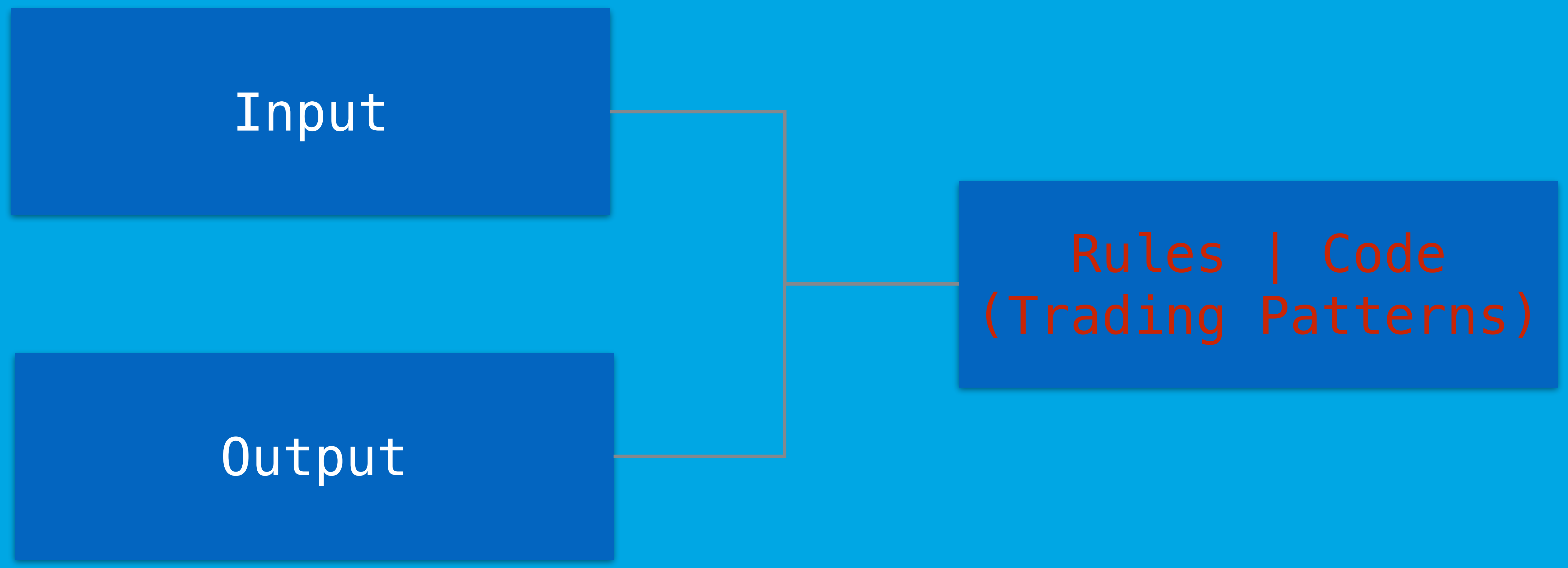


Reversal Patterns
Bearish Double Top
Bearish Head and Shoulders
Bearish Rising Wedge
Bearish Expanding Triangle
Bearish Triple Top
Bullish Double Bottom
Bullish Inverted Head and Shoulders
Bullish Falling Wedge
Bullish Expanding Triangle
Bullish Triple Bottom
Continuation Patterns
Bullish Flag Pattern
Bullish Pennant Pattern
Bullish Falling Wedge
Ascending Triangle
Symmetrical Triangle
Bearish Flag Pattern
Bearish Pennant Pattern
Bearish Rising Wedge
Descending Triangle
Symmetrical Expanding Triangle

Programming.



Machine Learning.



Financial Markets

x

y

"non-linear, complex, changing"

Finance History

"normative economics = assumptions,

(too) "simple and elegant theories"

$f(\bullet)$

"hardly any supporting

$f(x) \neq y$

"brain-driven & beauty myth"

AI in Finance = finaince

"positive economics = data, relationships,

x

"general, parametrizable,

$m(\bullet, a, b)$

"might show good performance, but

$m(x, a*, b*) \approx y$

"data-driven & AI-first"

Keras

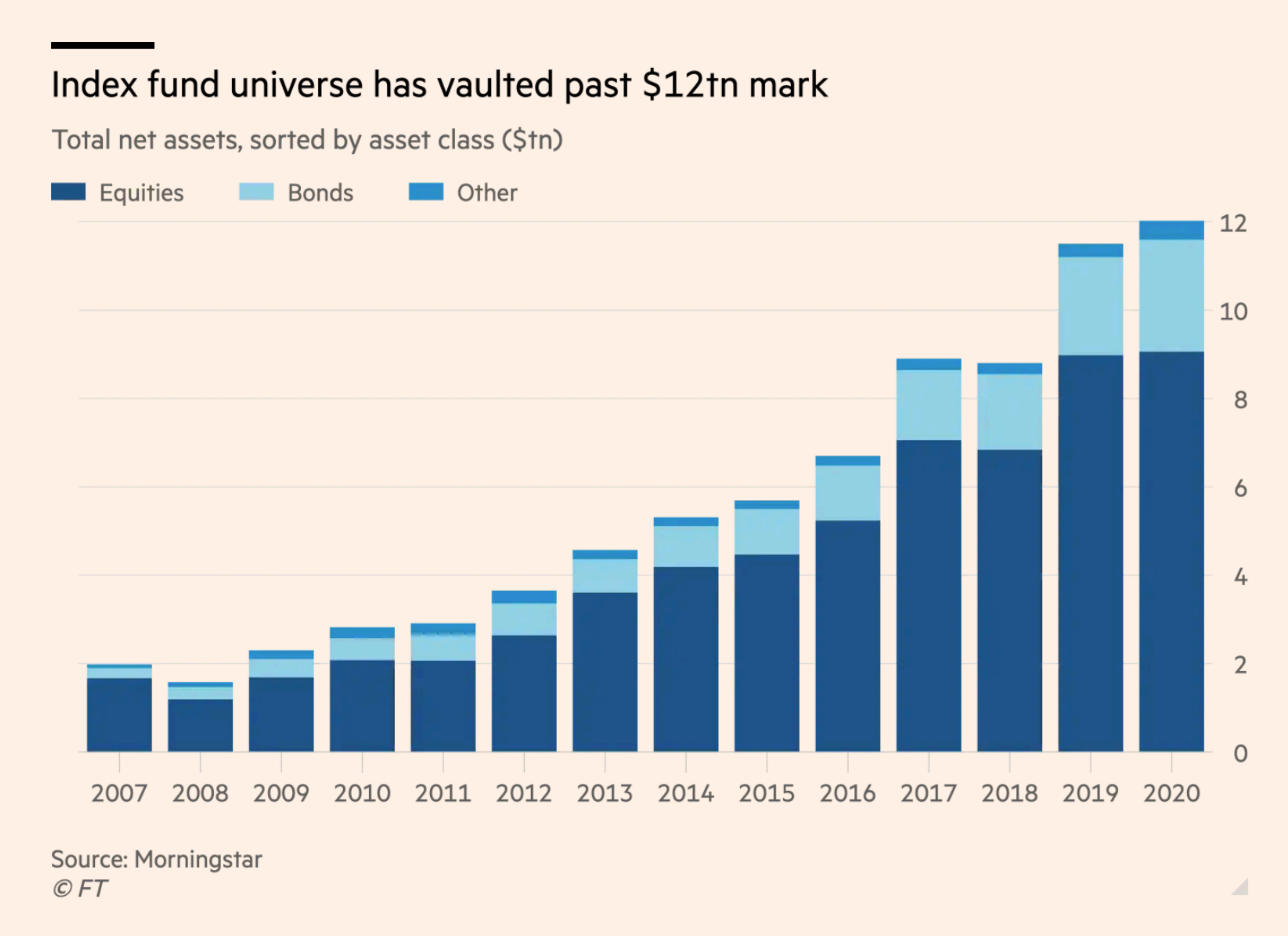
scikit learn

TensorFlow

NVIDIA®

Basic Strategies

Basic Strategy 1: Going all passive.



“It is true that the indices that passive funds track have over time morphed from being supposedly neutral snapshots of markets into something that actually exerts power over them, thanks to the growth of passive investing.

Mr Green argues that this helps explain why active managers are actually seeing their performance worsen as passive investing grows. The more money index funds garner, the better their holdings do in exact proportion to their weighting, and the harder it is for traditional discretionary investors to keep up.”

—Robin Wigglesworth: "A theory of (almost) everything for financial markets." Financial Times, 29. December 2020.

Basic Strategy 2: Going all in on data & AI.

Many Foundations of Finance are Flawed

A. Expected Utility (✘)
B. Equilibrium Theory (✘)
C. Normal Distributions (✘)
D. Linear Relationships (✘)
E. Efficient Markets (✓)
F. Arbitrage Pricing (✓)

Data-Driven & AI-Based Approaches

A. Data Science
B. Machine Learning
C. Deep Learning
D. Reinforcement Learning

Basic Strategy 2: Going all in on data & AI.

An Overview Of Artificial Neural Networks for Mathematicians

Leonardo Ferreira Guilhoto

Abstract

This expository paper first defines what an Artificial Neural Network is and describes some of the key ideas behind them such as weights, biases, activation functions (mainly sigmoids and the ReLU function), backpropagation, etc. We then focus on interesting properties of the expressive power of feedforward neural networks, presenting several theorems relating to the types of functions that can be approximated by specific types of networks. Finally, in order to help build intuition, a case study of effectiveness in the MNIST database of handwritten digits is carried out, examining how parameters such as learning rate, width, and depth of a network affects its accuracy. This work focuses mainly on theoretical aspects of feedforward neural networks rather than providing a step-by-step guide for programmers.

Contents

“In the mathematical theory of artificial neural networks, the universal approximation theorem states that a feed-forward network with a single hidden layer containing a finite number of neurons can approximate continuous functions on compact subsets of R^n, under mild assumptions on the activation function. The theorem thus states that simple neural networks can represent a wide variety of interesting functions when given appropriate parameters; however, it does not touch upon the algorithmic learnability of those parameters.”

–https://en.wikipedia.org/wiki/Universal_approximation_theorem

Basic Strategy 2: Going all in on data & AI.

DEEP ORDER FLOW IMBALANCE: EXTRACTING ALPHA AT MULTIPLE HORIZONS FROM THE LIMIT ORDER BOOK

PETTER N. KOLM, JEREMY TURIEL AND NICHOLAS WESTRAY

Petter N. Kolm is Clinical Professor and Director of the Mathematics in Finance Master's Program at NYU's Courant Institute of Mathematical Sciences, New York, NY 10012
petter.kolm@nyu.edu

Jeremy D. Turiel is Ph.D. Candidate at University College London, Department of Computer Science, 66-72 Gower Street, London WC1E 6EA
jeremy.turiel.18@ucl.ac.uk

Nicholas Westray is Visiting Researcher in Financial Machine Learning at NYU's Courant Institute of Mathematical Sciences, New York, NY 10012
nicholas.westray@nyu.edu

Abstract. We employ deep learning in forecasting high-frequency returns at multiple horizons for 115 stocks traded on Nasdaq using order book information at the most granular level. While raw order book states can be used as input to the forecasting models, we achieve state-of-the-art predictive accuracy ... training simpler "off-the-shelf" artificial neural networks on stationary in- ...derived from the order book. Specifically, models trained on order flow ...antly outperform most models trained directly on order books. Using ...ectional regressions we link the forecasting performance of a long short- ...memory network to stock characteristics at the market microstructure ...vel, suggesting that "information-rich" stocks can be predicted more accurately. Finally, we demonstrate that the effective horizon of stock specific forecasts is approximately two average price changes.

1. Introduction

In this article we employ deep learning (DL) in forecasting high-frequency returns at multiple horizons for 115 stocks traded on Nasdaq using order book information at the most granular level. In the last decade, DL has experienced enormous success, outperforming more traditional approaches in areas such as image classification, computer vision and natural language processing (Krizhevsky et al., 2012; LeCun et al., 2015; Schmidhuber, 2015; Goodfellow et al., 2016; Devlin et al.,

Date: August 6, 2021.
Key words and phrases. Artificial neural networks; Deep learning; Financial machine learning; High-frequency trading; Limit order books; Market microstructure; Multiple horizons; Order flow; Return predictability.

Data used for study: 10 TB.

"We employ deep learning in forecasting high-frequency returns at multiple horizons for 115 stocks traded on Nasdaq using order book information at the most granular level. While raw order book states can be used as input to the forecasting models, we achieve state-of-the-art predictive accuracy by training simpler \`off-the-shelf\` artificial neural networks on stationary inputs derived from the order book."

"Finally, our proposed approach is similar to that of a real-world production setting where the models are updated on a rolling basis. Based on our experience from training the deep learning models in this article across a large number of stocks, we conclude that deploying these models at large scale in practice is fully feasible and no longer a pipe-dream."

Lightening Talk by Petter Kolm.

Conclusions

1. Finance has long been driven by the “**beauty myth**” – elegant but too simplistic models, equations and approaches.
2. The availability of **big financial data** (historical–streaming, structured–unstructured) gives rise to ***data-driven finance***.
3. It can be assumed that the **“unreasonable effectiveness of big data”** holds true in the financial domain as well.
4. Due to the availability of big data (e.g. billions of hours of virtual car driving, billions of self-played games), **Artificial Intelligence** (AI) is changing almost every area of our lives.
5. It is therefore to be assumed that in the same way the **combination of *data-driven and AI-first finance*** will influence and change finance, investing, and trading for good.

The Python Quants GmbH

Dr. Yves J. Hilpisch
+49 3212 112 9194
tpq.io | ai@tpq.io | @dyjh