AI and Machine Learning in Finance

Maxime Tchibozo

International FinTech Program

HEC Paris

July 2024

Agenda

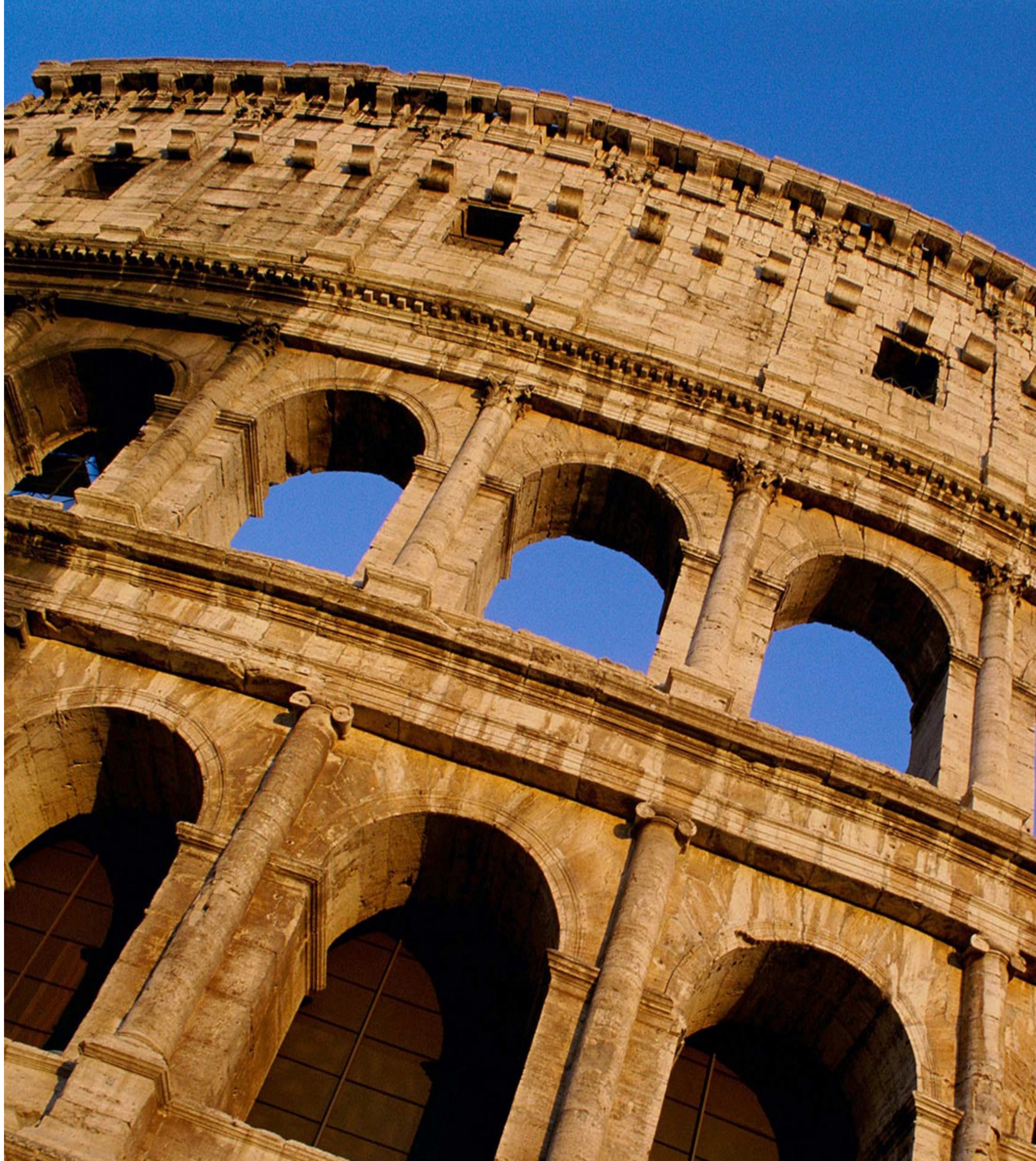
- Why are Financial Institutions interested in AI / ML?
- Overview of core AI concepts
- Principles of AI in Quant Finance
- Case Study 1: Investing
- Case Study 2: Trading
- Case Study 3: Banking
- Frontiers of AI in Finance
- Risks for AI Adoption in Finance
- Final Thoughts

This is a fully collaborative class

Please raise your hand & tell us about your prior experiences

Why are Financial Institutions interested in AI/ML?

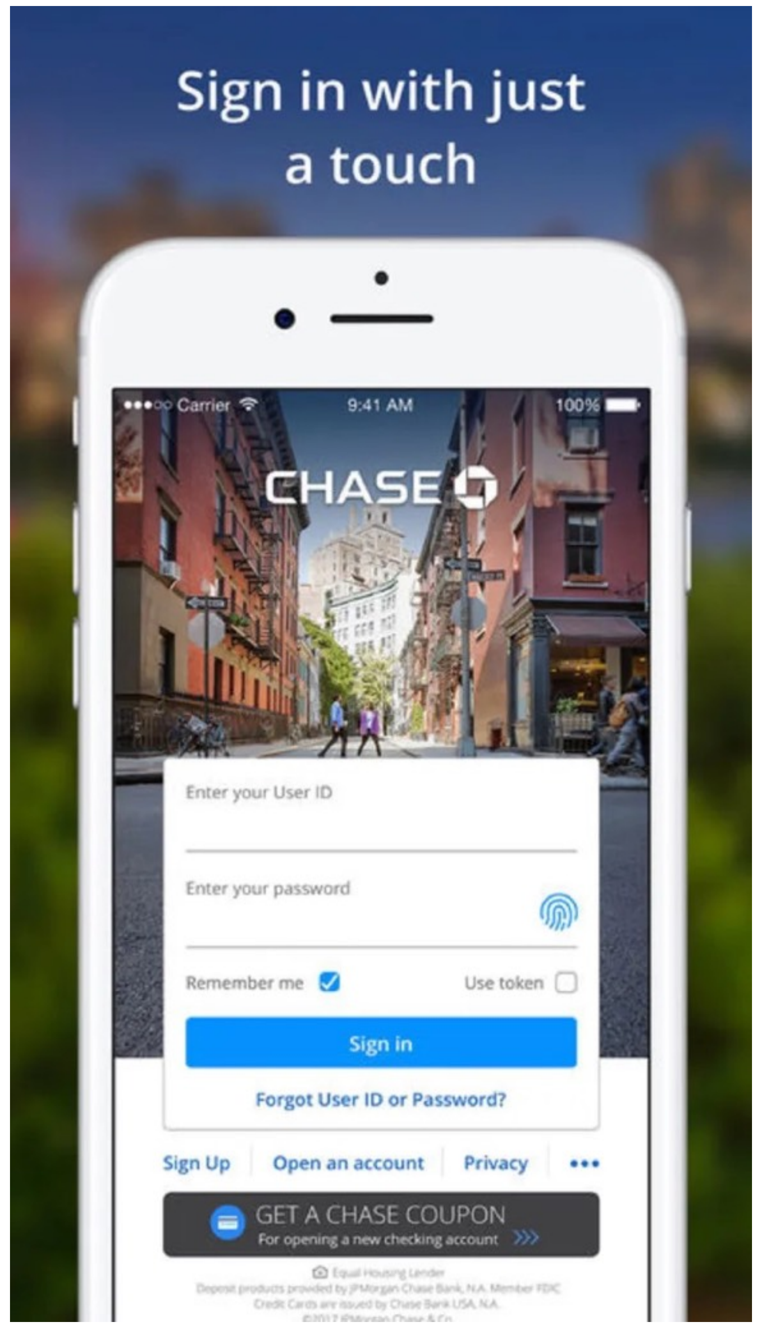
- Financial Institutions are Investing Heavily in AI
- Why Now? How the Current Paradigm of Data Emerged
- AI Methods



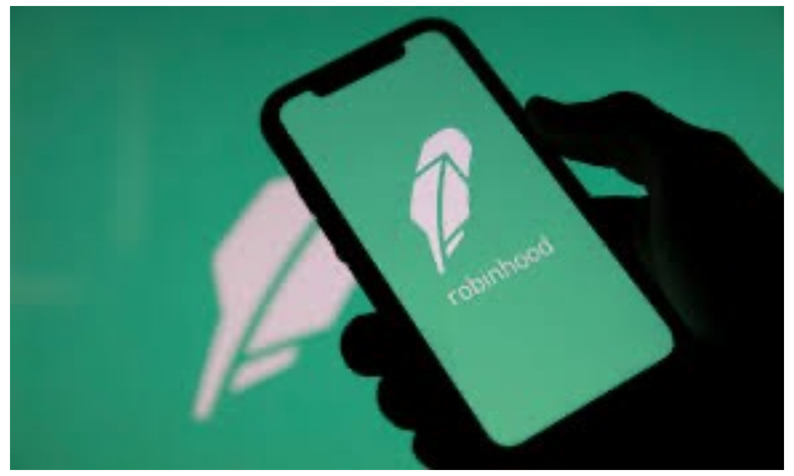
Financial Institutions are Investing Heavily in AI

CITADEL

aladdin by BLACKROCK



BANQUE DE FRANCE
EUROSYSTÈME



- Investment Banks & Asset Managers
- Retail Banks
- Hedge Funds
- Insurance Companies
- Brokerage Firms
- Government Entities

Why Now? How the Current Paradigm of Data Emerged

The AI Paradigm we see today was born from big investments made by Tech companies

Tech companies pursued the fastest possible growth, valuing scaling above all else

The Finance world has different priorities and places more emphasis on Risk Management

How AI will transform the Financial System remains an open question

90's
Emergence of the Internet

2000s
Computer Miniaturization

2010s
New Computational Methods

2020s
Cloud Computing Expands AI Reach

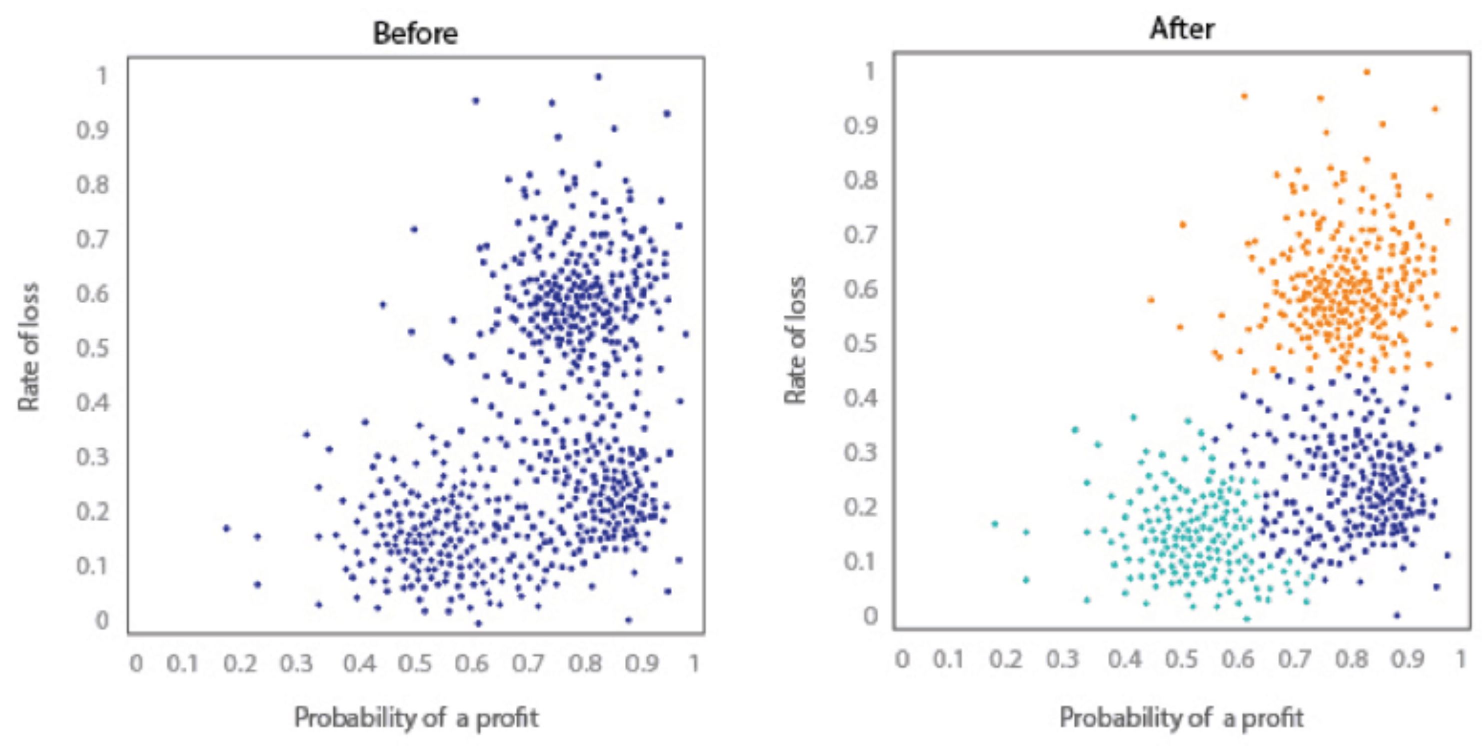
Today
Banks Seek to Capitalize on AI Opportunities.

Important to Keep in Mind: Each of these changes comes with **tradeoffs**.

AI Methods

Unsupervised Learning

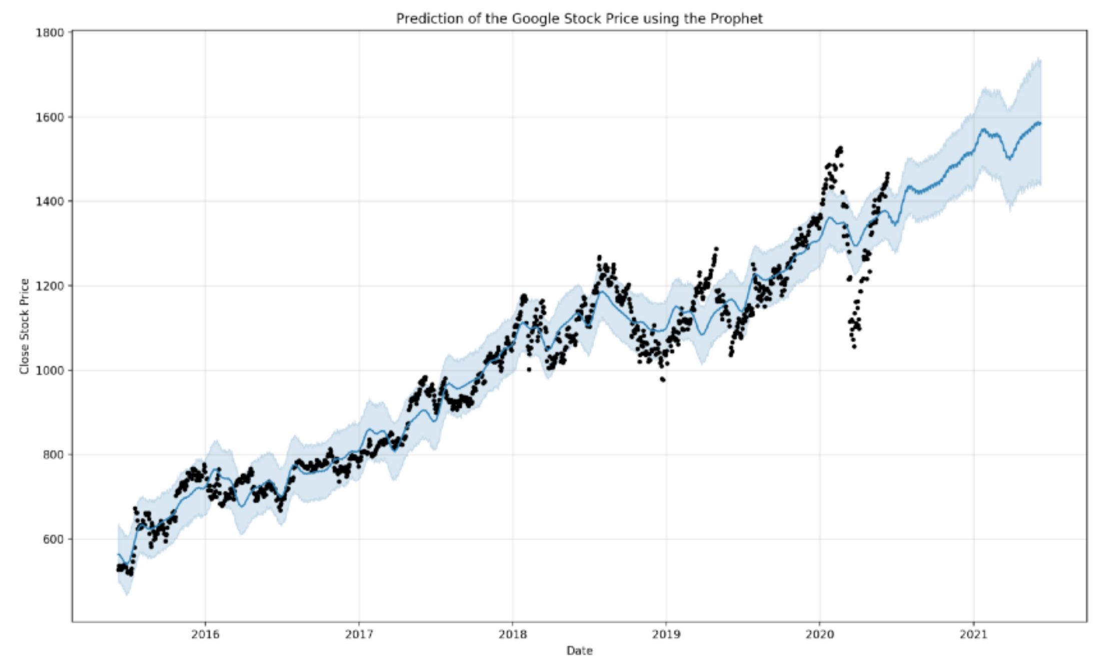
Find patterns in data without human guidance



"Here is all the data we have on past trades, can you find an interesting way to group similar trades together"
Unsupervised clustering

Supervised Learning

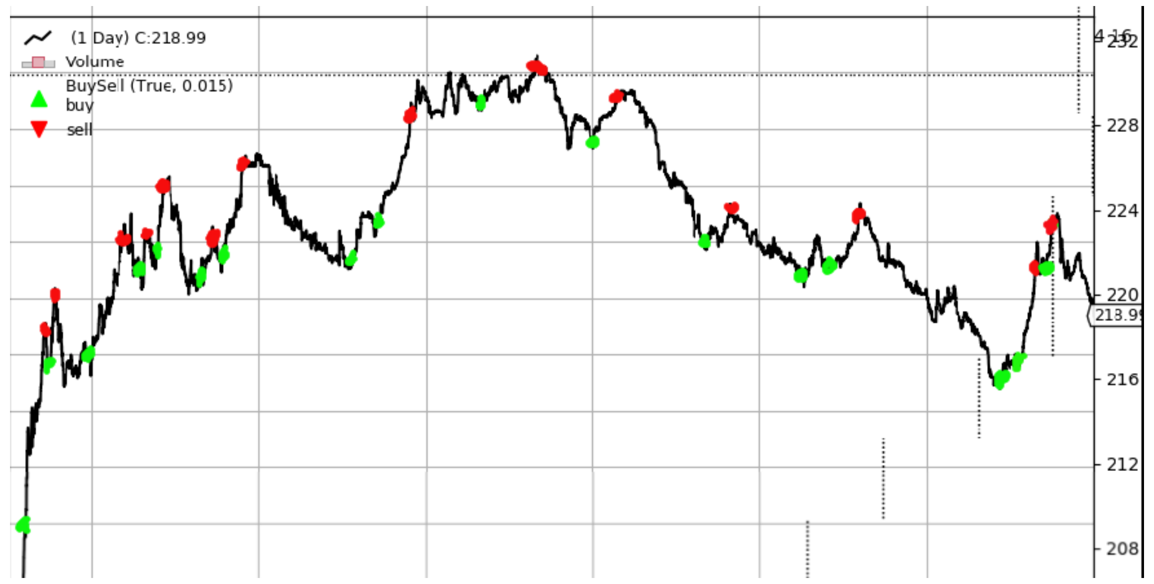
Learn patterns which best replicate a predefined outcome



"This is how the market responded at each earnings release. *Based on those past trends*, can you tell me how the share price could evolve when earnings are announced tomorrow?"
Time Series Forecasting

Reinforcement Learning

Learn to **apply** optimal **policies** in a particular environment given specific reward incentives



"Can you tell me in real-time when you think it's the right moment to buy or sell?"
Agent-Based Models

AI Methods

Unsupervised Learning

Find patterns in data without human guidance

Supervised Learning

Learn patterns which best replicate a predefined outcome

Reinforcement Learning

Learn to **apply** optimal **policies** in a particular environment given specific reward incentives

- Clustering & Segmentation
- Recommendation Systems
- Compression & Representation Learning
- Analytics & Visualization
- Data Generation & Augmentation

AI Methods

Unsupervised Learning

Find patterns in data without human guidance

Supervised Learning

Learn patterns which best replicate a predefined outcome

Reinforcement Learning

Learn to **apply** optimal **policies** in a particular environment given specific reward incentives

- Object Detection & Image Classification
- Named Entity Recognition & Text Classification
- Time Series Forecasting & Predictive Modeling
- Sentiment Analysis

AI Methods

Unsupervised Learning

Find patterns in data without human guidance

Supervised Learning

Learn patterns which best replicate a predefined outcome

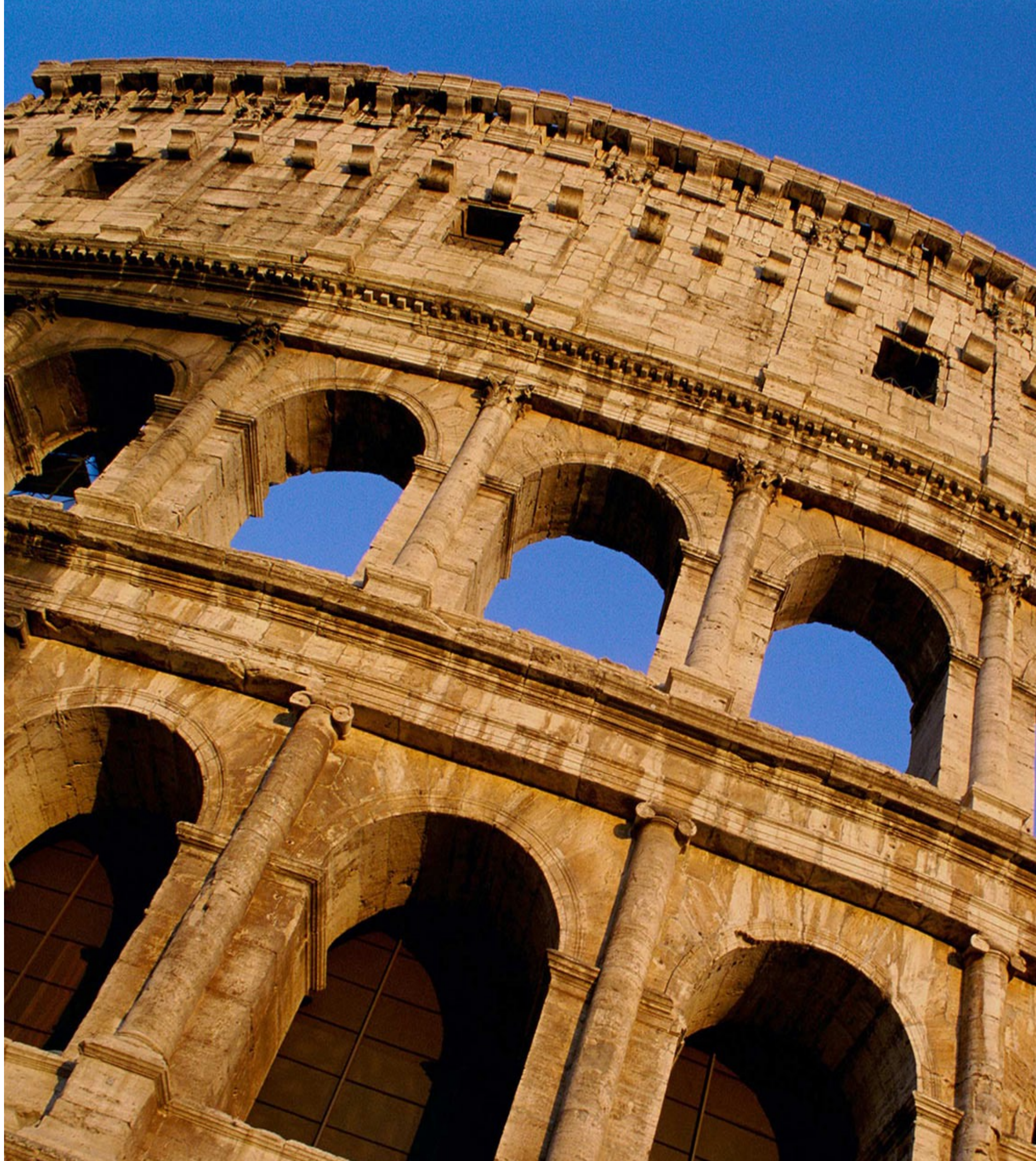
Reinforcement Learning

Learn to **apply** optimal **policies** in a particular environment given specific reward incentives

- Real-Time Decision Making
- Genetic Algorithms

Overview of core AI Concepts

- An AI Example: Setting Mortgage Interest Rates
- Identifying Main Factors
- A Data Science Approach
- Data & Representations
- Metrics & Evaluation
- Model Criticism & Update



An AI Example: Setting Mortgage Interest Rates

How should a Bank set rates for customers looking to buy a home?

A mortgage is a type of loan used to purchase real estate. The borrower agrees to pay the lender - usually a Bank - a series of recurring payments covering the principal amount of the loan and the interest.

The interest is referred to as the mortgage rate.

Which AI topic could this example fit under?

Which factors might affect how a Bank sets mortgage rates?

In France in 2023, the typical mortgage rate range was 2.5% - 4.5%.

Setting a Mortgage Interest Rate - Main Factors

Personal Factors

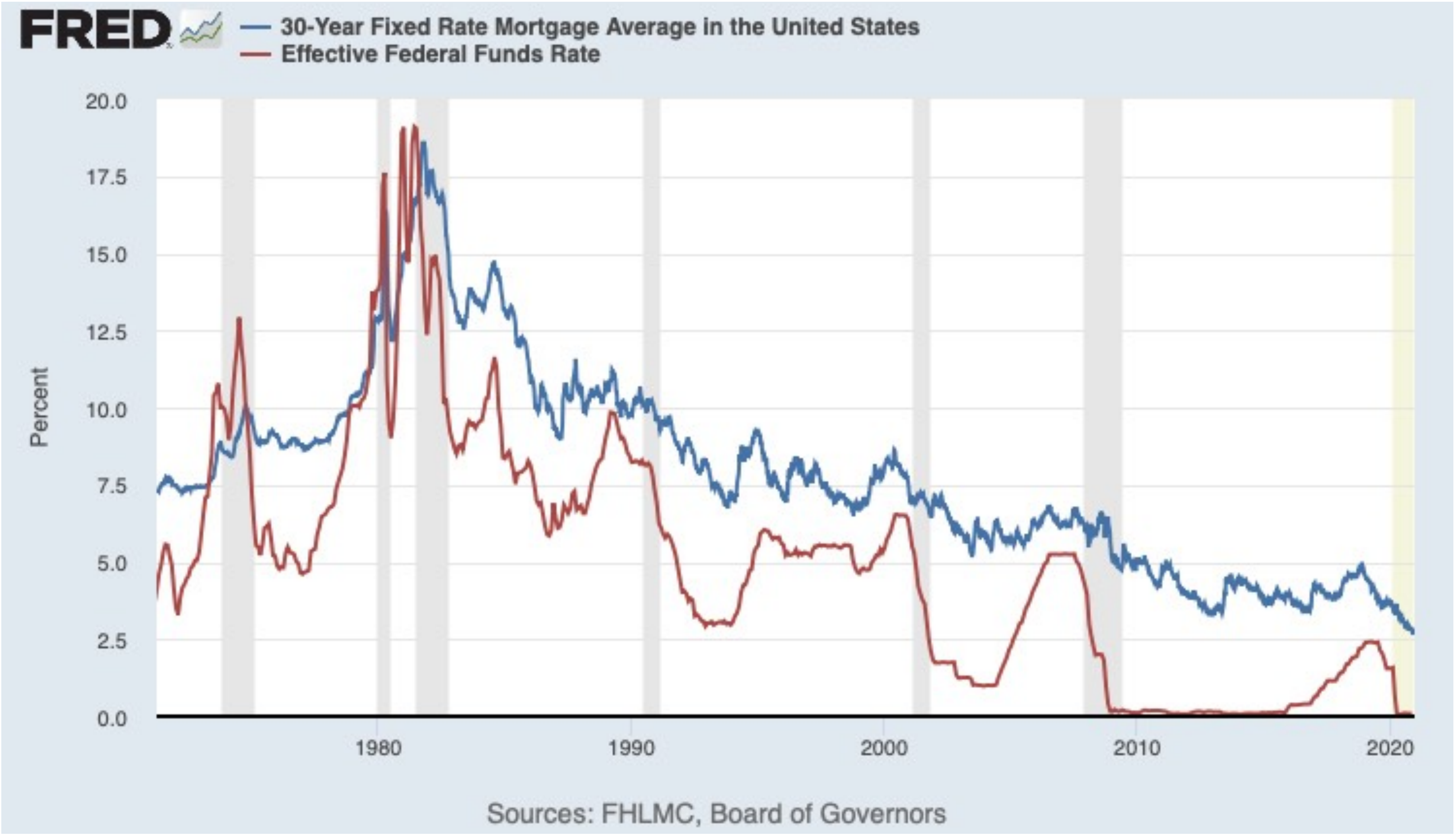
Credit score, down payment & collateral, type of property, buyer profile, …

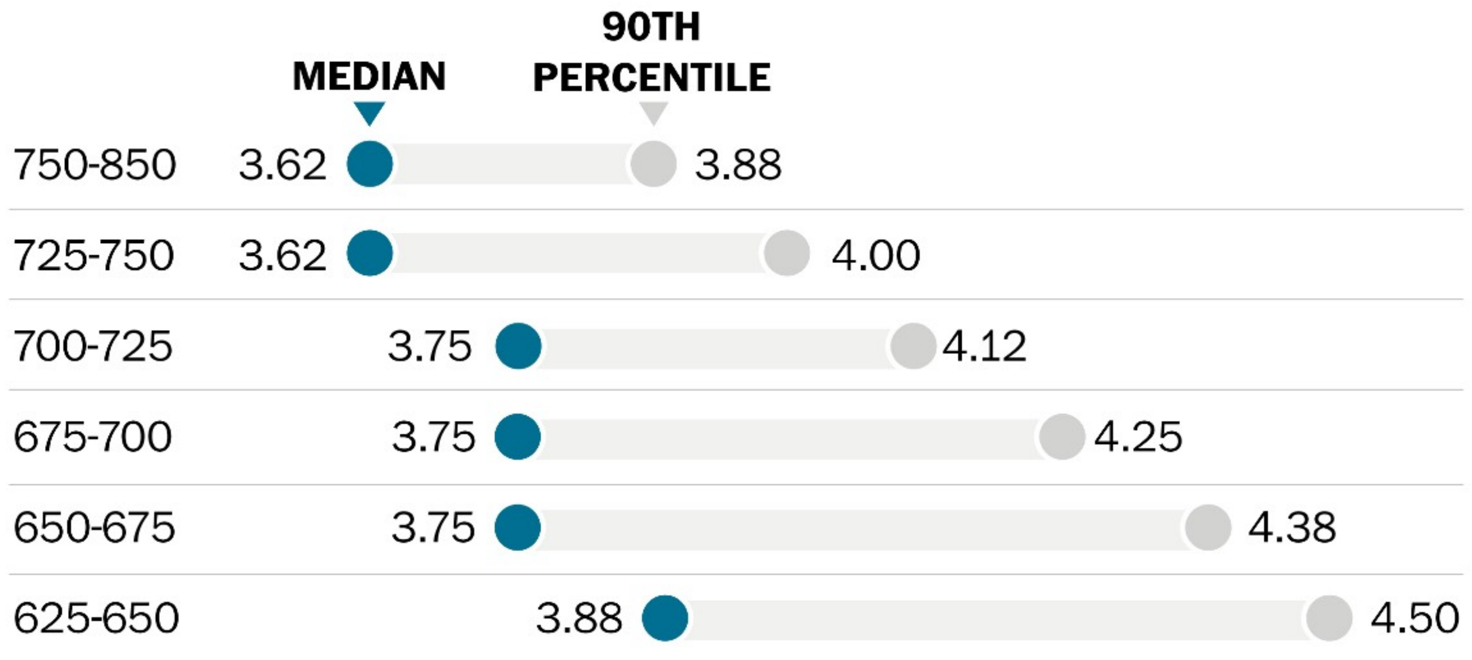
Market Factors

Central Bank interest rates, inflation, competing banks, bonds, …

Product-related Factors

Loan maturity, loan amount, …





Note: Rates are for 30-year fixed-rate mortgages processed from Sept. 1 through Nov. 7 2016.

Source: Realtor.com

THE WASHINGTON POST

Key Considerations

How does each factor affect the Bank's profit (all else held equal)?

How should each of these factors be weighed?

Setting a Mortgage Interest Rate - Data Science Approach

Inputs & Data Types

Market Data

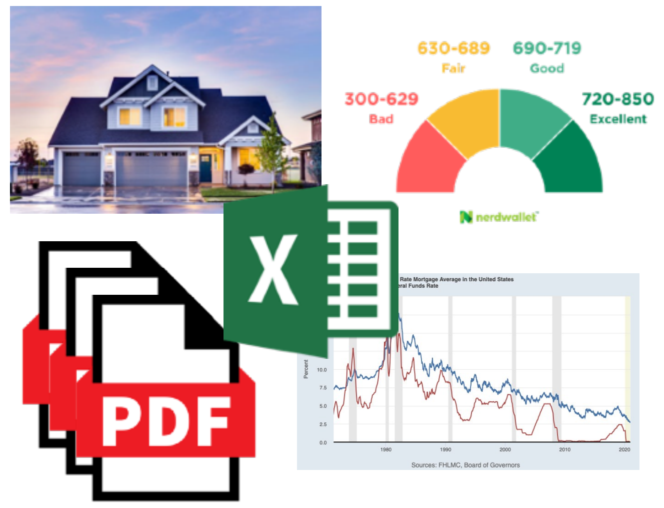
Central Bank interest rates (Spreadsheet)
Inflation figures (Spreadsheet)
Competing Banks (PDF documents)

Personal Data

Credit score (Number)
Down payment (Number)
Collateral (PDF, images)
Type of property (PDF, images, Num.)
Risk profile (PDF)

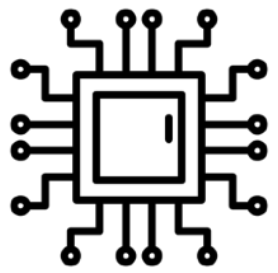
Product Data

Loan maturity (Number)
Loan amount (Number)



Representation

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ip} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{np} \end{bmatrix}}_{\mathbf{X}}$$



Embedding

Weighing

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_i \\ \vdots \\ \vdots \\ w_n \end{bmatrix}$$

Parameter Learning

Output

= 1.8%

Mortgage Interest Rate
(Number)

Regression

Setting a Mortgage Interest Rate - Data & Representations

Inputs & Data Types

Market Data

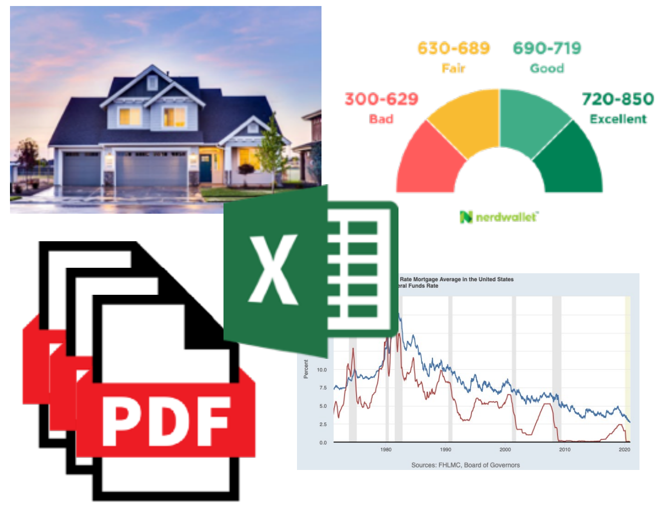
Central Bank interest rates (Spreadsheet)
Inflation figures (Spreadsheet)
Competing Banks (PDF documents)

Personal Data

Credit score (Number)
Down payment (Number)
Collateral (PDF, images)
Type of property (PDF, images, Num.)
Risk profile (PDF)

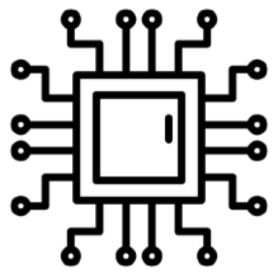
Product Data

Loan maturity (Number)
Loan amount (Number)



Representation

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ip} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{np} \end{bmatrix}}_{\mathbf{X}}$$



Embedding

All data is preprocessed and converted into a representation (a.k.a embedding).

A good embedding compresses information while preserving the most important properties of the data.

AI algorithms learn patterns from embeddings directly.

Careful consideration is required:

- Raw inputs have vastly different formats
- Numeric values might be scaled differently
- Documents might contain mostly irrelevant information
- Data may not cover all the edge cases that appear in reality

Setting a Mortgage Interest Rate - Metrics & Evaluation

AI models transform embeddings into predictions.

Based on how close predictions are to reality, an algorithm will update the weights and the embeddings.

A good AI model can predict correct values even for inputs / embeddings it has never seen before.

Careful consideration is required:

- Only unseen data should be used for evaluation
- Evaluation should be done multiple times
- Data used for evaluation should be representative of data in the real world
- The metric used for evaluation must capture underlying properties of the data
- AI decisions are not always easily interpretable
- Overfitting / Underfitting

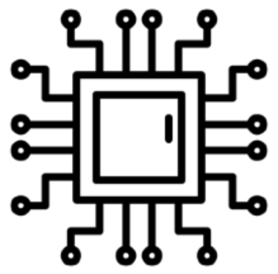
Representation | Weighing | Output | Reality

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ip} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{np} \end{bmatrix}}_{\mathbf{X}} \qquad \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_i \\ \vdots \\ \vdots \\ w_n \end{bmatrix}$$

= 1.8%

vs. 1.7%

Mortgage Interest Rate (Number)



Embedding

Parameter Learning

Regression

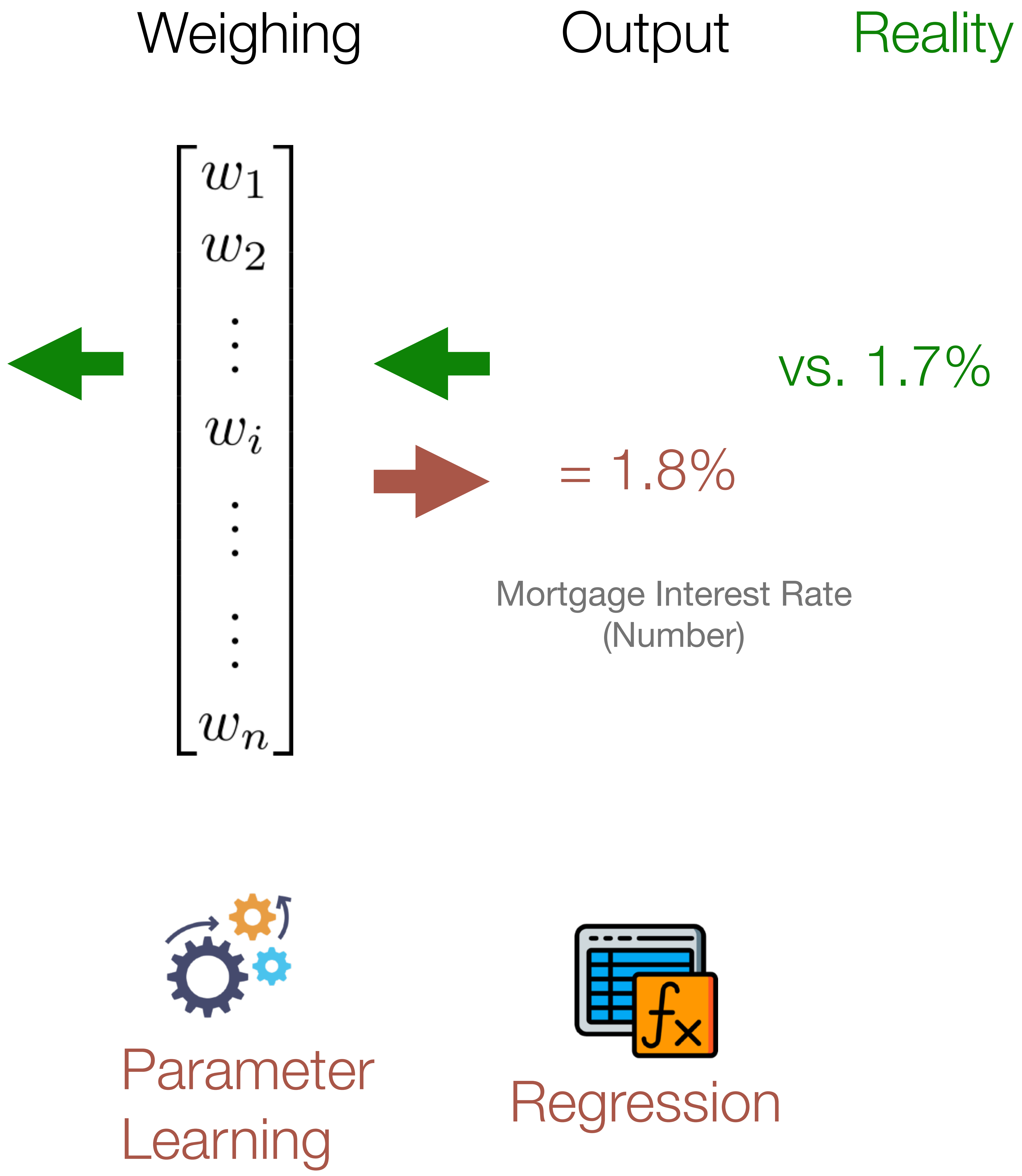
Setting a Mortgage Interest Rate - Metrics & Evaluation

Wait! The process is not yet done, even once the model has finished predicting.

We criticize the model to identify further improvements.

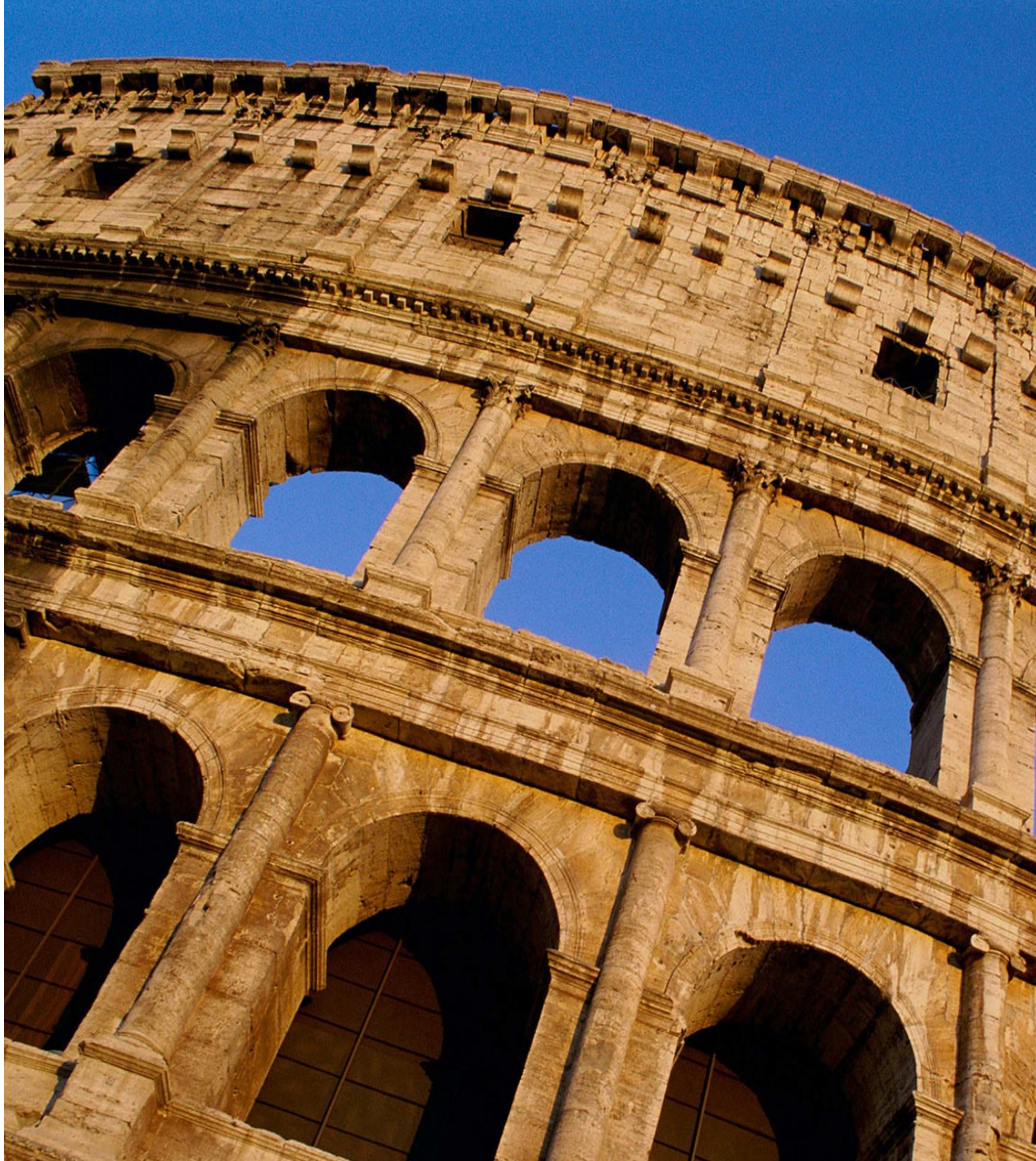
Typical model criticism questions:

- Do these predictions make common sense?
- Was all the data used informative?
- Were any important factors missed?
- Do the learnt representations capture properties of underlying data?
- Is the sample size large enough?
- Is the algorithm explainable?
- Is there a tradeoff between prediction performance and explainability of decisions?
- What kinds of responsibilities are you delegating to the algorithm?



Principles of AI in Quant Finance

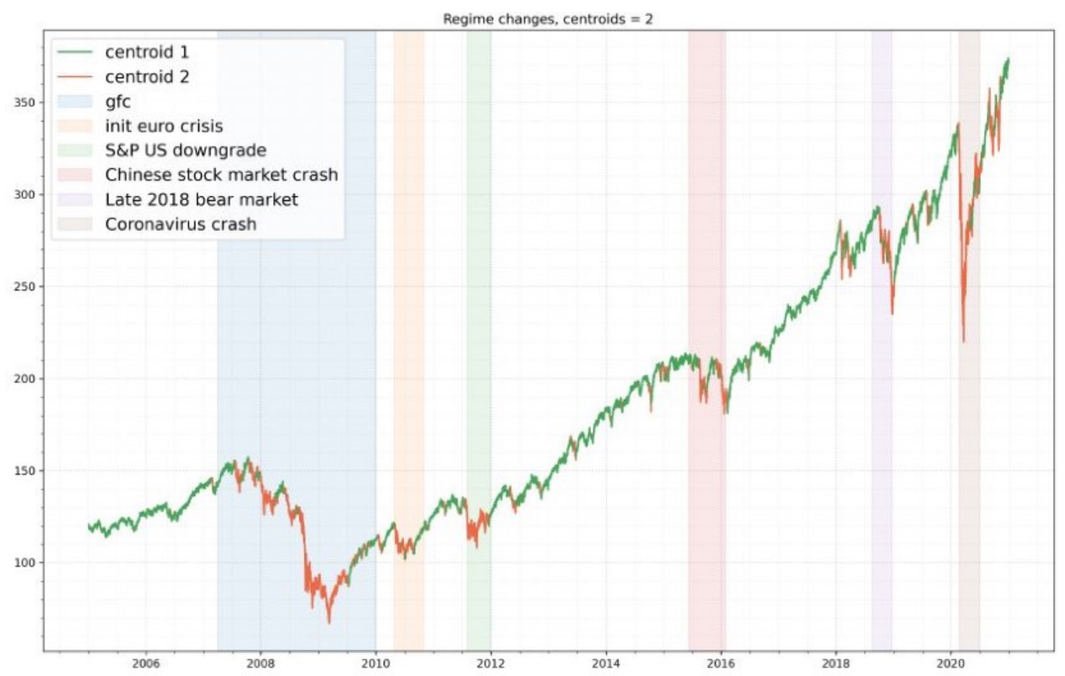
- Applications of AI in Quant Finance
- Case Study: Thesis-Driven Investing (Cat-Bonds)
- Thesis-Driven Investing: A Word of Caution



Applications of AI in Quant Finance

Price and Volatility Forecasting

Identify predictive patterns from time series data



Real-time detection of market regime changes
[Horvath, Issa and Muguruza, 2021]

Portfolio Optimization

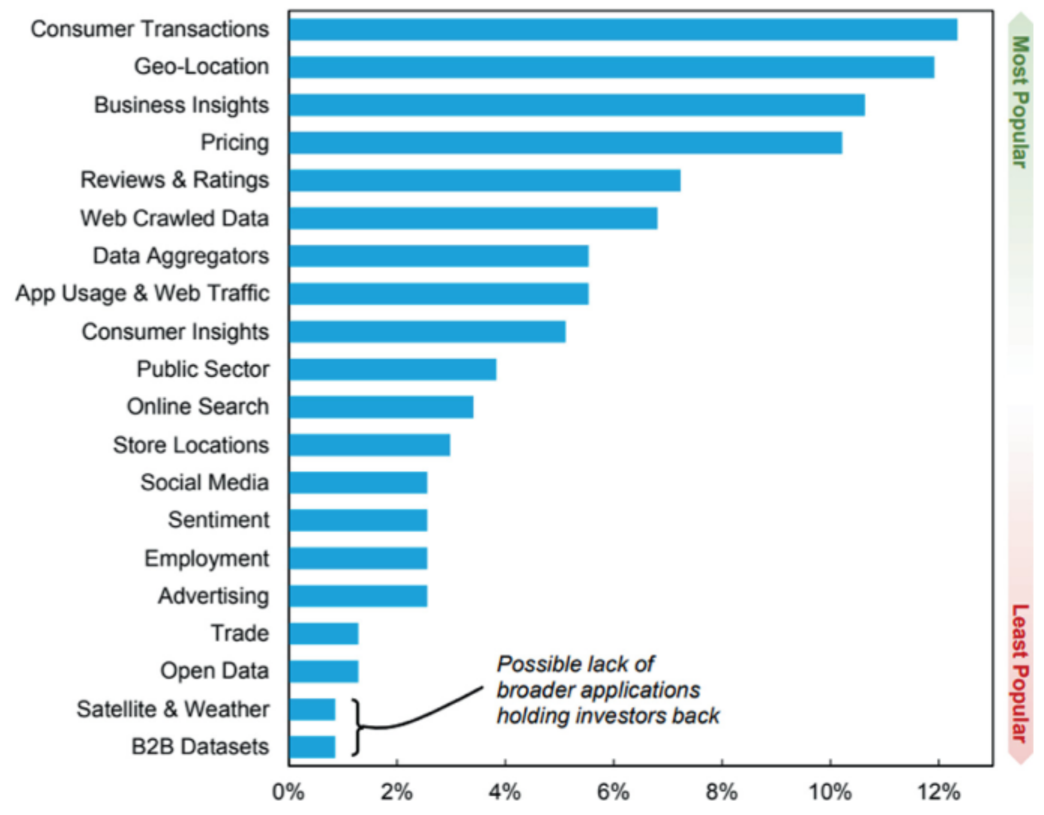
Detect asset correlations and adjust asset allocations to balance risk and reward

Key		SPY	IJH	IWM	EFA	SCZ	EEM	BLV	CSJ	IGSB	DBC	IYR	RWX	GLD
U.S. Large Cap Stocks	SPY	1.00	0.92	0.91	0.88	0.84	0.65	-0.18	0.40	-0.25	0.39	0.43	0.61	-0.01
U.S. Mid-Cap Stocks	IJH	0.92	1.00	0.97	0.82	0.83	0.72	-0.17	0.59	-0.30	0.60	0.57	0.69	0.22
U.S. Small-Cap Stocks	IWM	0.91	0.97	1.00	0.79	0.78	0.59	-0.25	0.44	-0.46	0.58	0.41	0.60	0.10
International Developed Market Large-Cap Stocks	EFA	0.88	0.82	0.79	1.00	0.97	0.80	-0.30	0.51	-0.25	0.49	0.37	0.79	0.15
International Developed Market Small-Cap Stocks	SCZ	0.84	0.83	0.78	0.97	1.00	0.81	-0.34	0.56	-0.27	0.47	0.43	0.78	0.21
Emerging Market Stocks	EEM	0.65	0.72	0.59	0.80	0.81	1.00	-0.04	0.81	0.05	0.67	0.65	0.82	0.61
Long-term Taxable Bonds	BLV	-0.18	-0.17	-0.25	-0.30	-0.34	-0.04	1.00	-0.08	0.75	0.05	0.41	-0.07	0.47
Short-term Taxable Bonds	IGSB	0.40	0.59	0.44	0.51	0.56	0.81	-0.08	1.00	0.18	0.64	0.62	0.69	0.62
Intermediate-term Municipal Bonds	MUB	-0.25	-0.30	-0.46	-0.25	-0.27	0.05	0.75	0.18	1.00	-0.16	0.35	-0.04	0.33
Commodities - Broad	DBC	0.39	0.60	0.58	0.49	0.47	0.67	0.05	0.64	-0.16	1.00	0.30	0.49	0.66
U.S Real Estate	IYR	0.43	0.57	0.41	0.37	0.43	0.65	0.41	0.62	0.35	0.30	1.00	0.72	0.64
International Real Estate	RWX	0.61	0.69	0.60	0.79	0.78	0.82	-0.07	0.69	-0.04	0.49	0.72	1.00	0.47
Gold	GLD	-0.01	0.22	0.10	0.15	0.21	0.61	0.47	0.62	0.33	0.66	0.64	0.47	1.00

Clustering for real-time portfolio risk management
[Lopez do Prado, 2020]

Thesis-Driven Investing

Develop differentiated investment strategies by leveraging alternative data sources

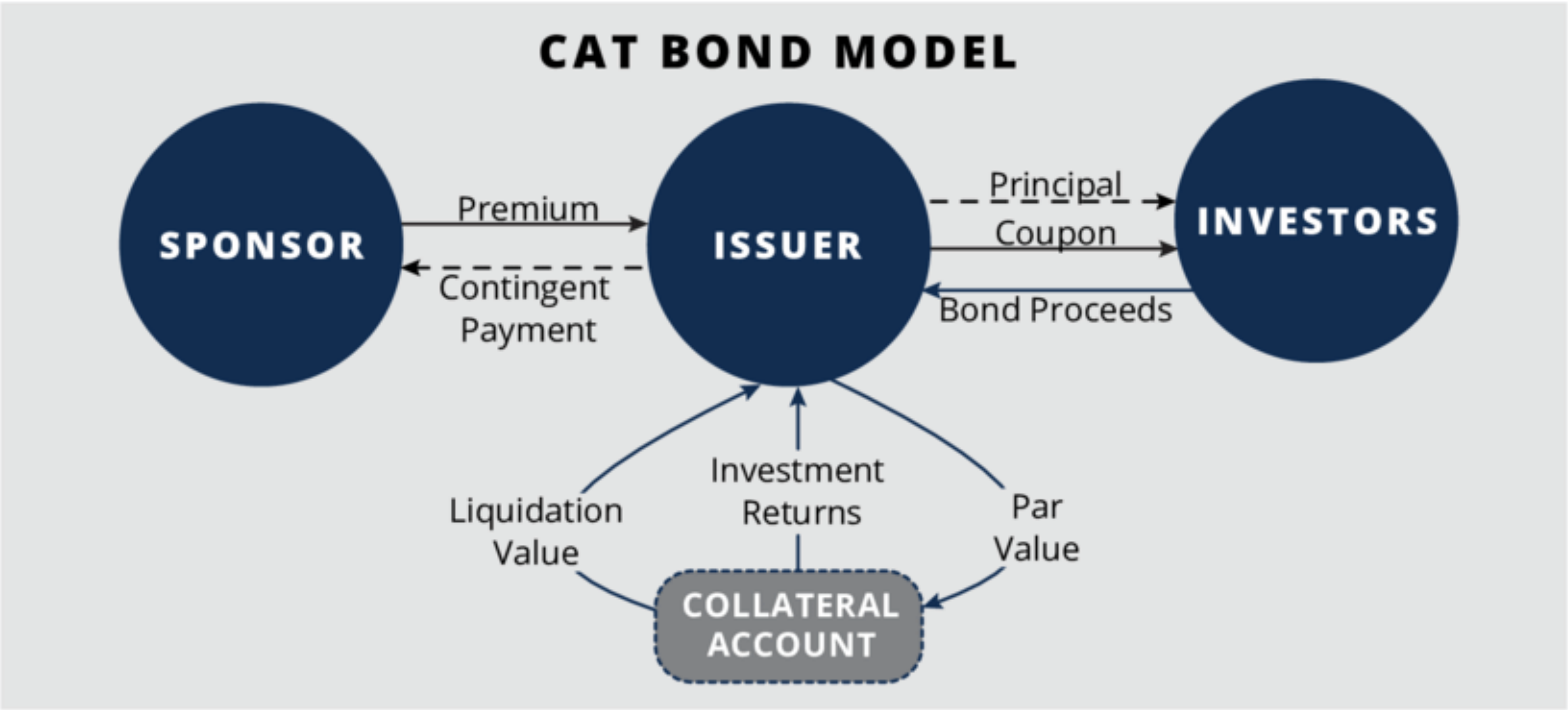


Alternative data for quantitative trading (social media, election polls, environmental data, satellite data)
[Charoenwong and Kwan, 2021]

Case Study: Thesis-Driven Investing (Cat Bonds)

Catastrophe bonds are financial instruments that transfer risks of catastrophic events (e.g. hurricanes, earthquakes) from insurers to investors.

Investors purchase bonds and receive regular coupon payments from insurers. If catastrophes occur, the principal for the bond is used to cover insurer losses and investors lose money.



Rhodes and Vajjhala (2015) - Leveraging Catastrophe Bonds as a Mechanism for Resilient Infrastructure Project Finance

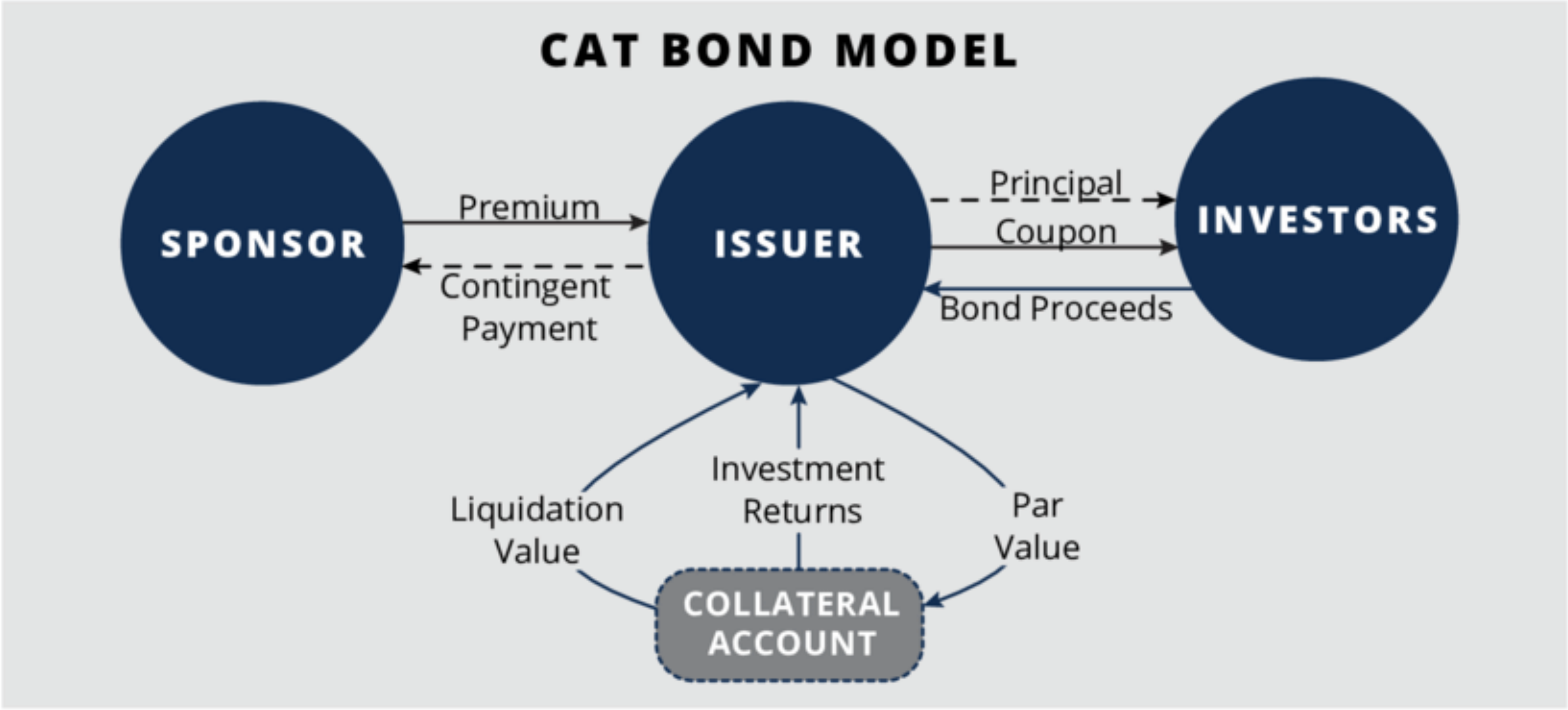
Investment Thesis

As catastrophe prediction technologies improve, legacy insurance models are likely to misprice catastrophic risks, creating opportunities for investors

Case Study: Thesis-Driven Investing (Cat Bonds)

Catastrophe bonds are financial instruments that transfer risks of catastrophic events (e.g. hurricanes, earthquakes) from insurers to investors.

Investors purchase bonds and receive regular coupon payments from insurers. If catastrophes occur, the principal for the bond is used to cover insurer losses and investors lose money.



Rhodes and Vajjhala (2015) - Leveraging Catastrophe Bonds as a Mechanism for Resilient Infrastructure Project Finance

Investment Thesis

As catastrophe prediction technologies improve, legacy insurance models are likely to misprice catastrophic risks, creating opportunities for investors

Implementation

- **Model:** Leverage climate and seismology models to estimate catastrophic event probabilities and price potential catastrophe impacts
- **Data:** weather data, seismic activity data, historical disaster records
- **Investment Strategy:** Investors purchase bonds when they believe insurers are overestimating catastrophe risks

According to Swiss Re, Cat Bonds returned ~20% returns for hedge funds in 2023

Break

Thesis-Driven Investing - A Word of Caution

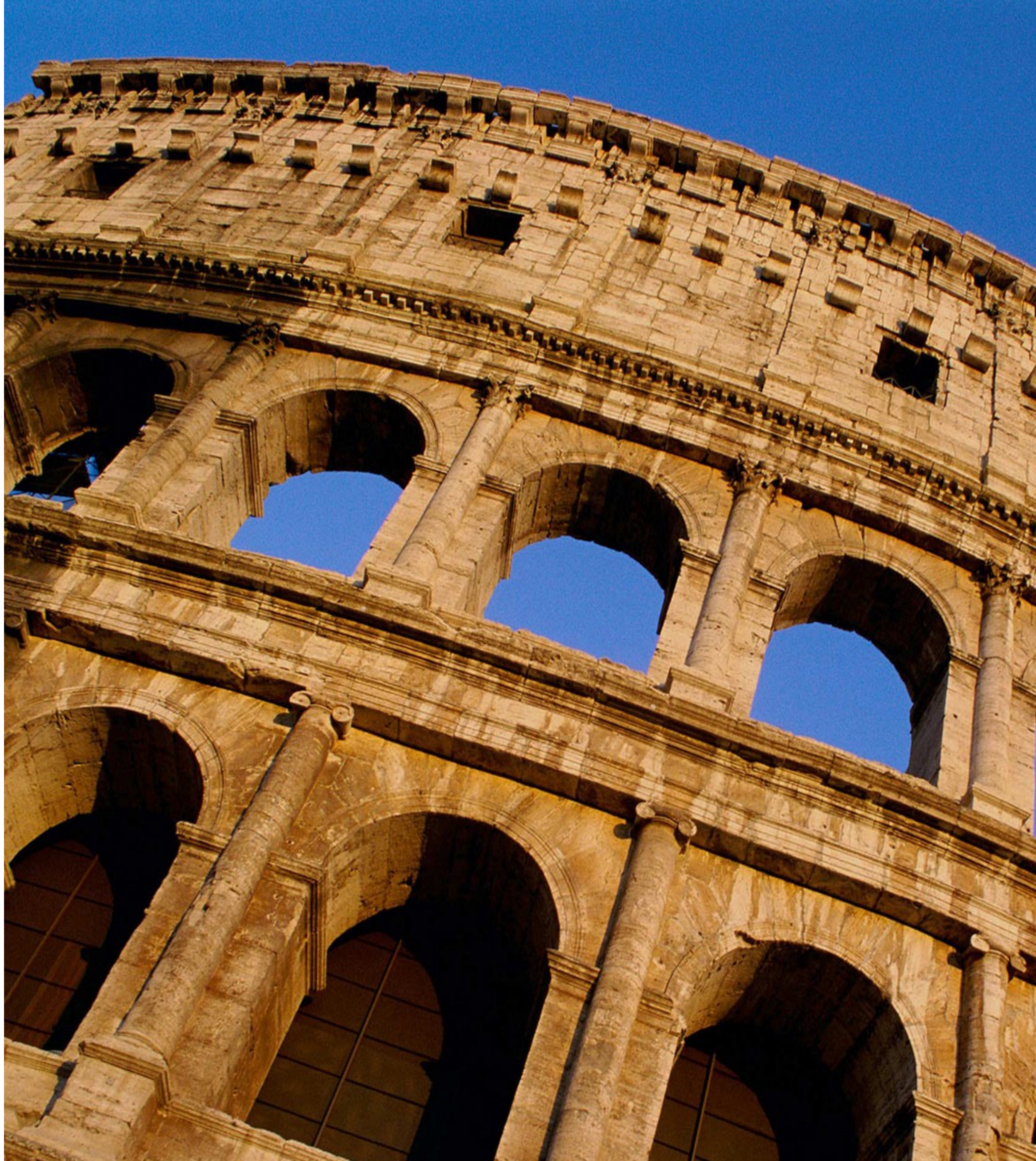
- Efficient Markets: **Being first matters**. A valid model can lose money if someone else has already exploited the opportunity
- Correlated Returns: Features derived from alternative data don't always correlate with returns in intuitive ways
- Alternative data can provide useful sanity checks to test simpler investment theses
- Data is more valuable than you think. Finding quality data is a full time job

Due next class

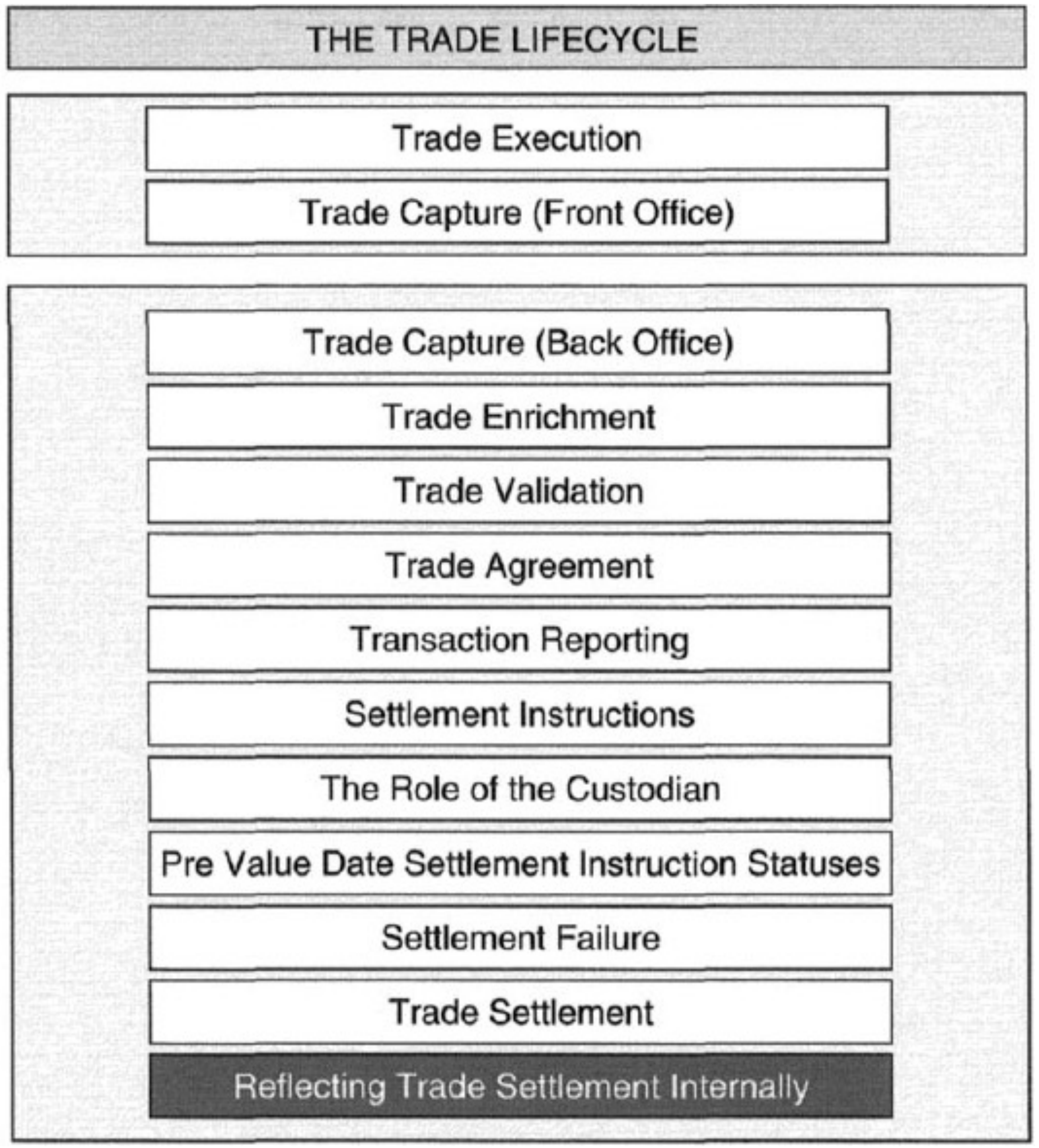
Develop an investment thesis from a topic you know well; identify alternative data sources that can validate the thesis; Derive a simple investment strategy

Case Study #2: Trading

- Trade Execution
- Trade Validation
- Know Your Customer



Trading Case Study: Lifecycle of a Trade



A trader initiates a trade with the purchase or sale of a security, commodity, option. The buyer and seller enter into an agreement, with a set price, volume, etc.

Operational teams then specify the logistics of the trade between parties.

- We compare traditional and AI-driven approaches to the trading lifecycle.

Key Questions to Keep in Mind

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

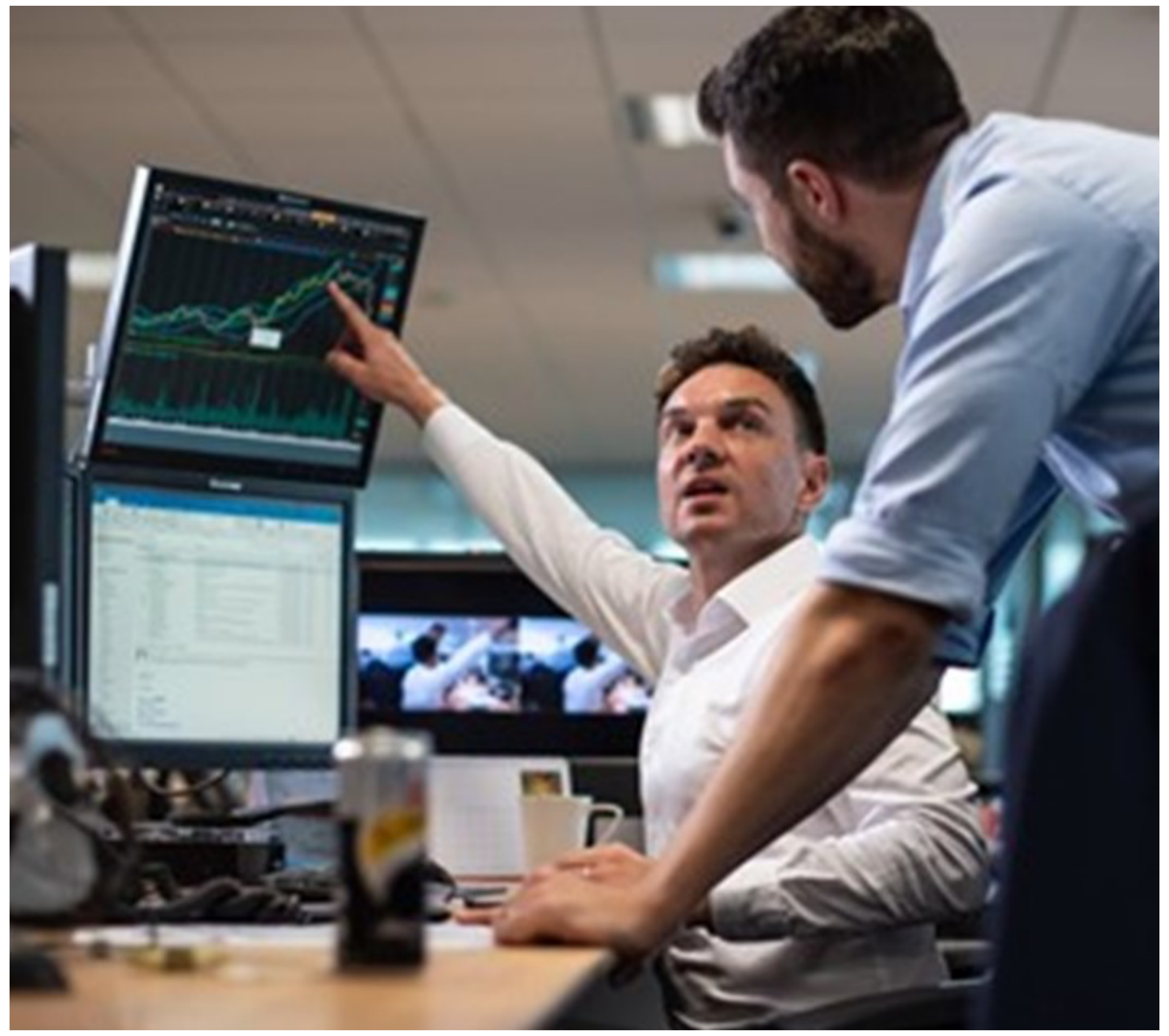
Which parts of the traditional process would benefit from being **data-driven**?

Trading Case Study: Lifecycle of a Trade - Trade Execution

The trader is a human expert in anticipating price or volatility fluctuations for a given market.

They have experience, know the market and get pricing assistance from quants.

They use their knowledge and intuition to time the trade and get the best possible price.



Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

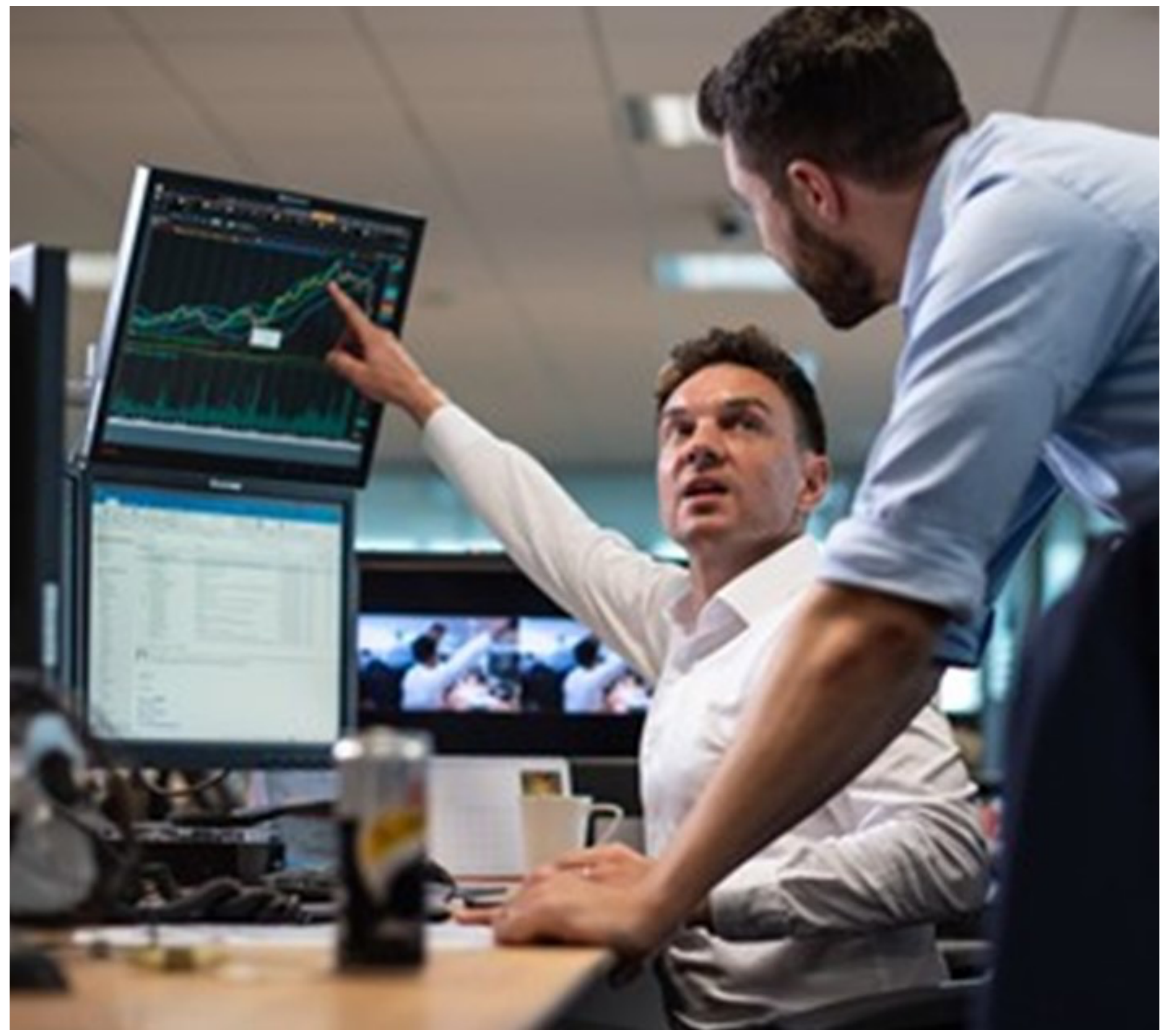
Which parts of the traditional process would benefit from being **data-driven**?

Trading Case Study: Lifecycle of a Trade - Trade Execution

The trader is a human expert in anticipating price or volatility fluctuations for a given market.

They have experience, know the market and get pricing assistance from quants.

They use their knowledge and intuition to time the trade and get the best possible price.



Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Bottlenecks (non-exhaustive)

Attention, Speed, Risk Limits, Latency, Emotion, …

How can AI help?

- **News Aggregation Methods**: Real-time Big Data summarization of the web (Bloomberg, Twitter, Reddit, Government Data).
- **Machine Learning-Based Pricing**: Combine Quant probabilistic pricing methods with Machine Learning-based estimators.
- **Identifying Trading Opportunities**: Automatically compare market patterns to historically similar trading charts.

Trading Case Study: Lifecycle of a Trade - Trade Validation

Operational teams reconcile information from the trade with the counterpart. Buyer and seller send each other documentation with the agreed upon details.

Ops. teams verify that both parties have the exact same information. In some cases, they have two days to validate the details and disclose the trade to government entities.

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Trading Case Study: Lifecycle of a Trade - Trade Validation

Operational teams reconcile information from the trade with the counterpart. Buyer and seller send each other documentation with the agreed upon details.

Ops. teams verify that both parties have the exact same information. In some cases, they have two days to validate the details and disclose the trade to government entities.

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Bottlenecks (non-exhaustive)

Time, Complexity, Documentation Errors, …

How can AI help?

- **Automated Trade Detail Extraction**: Document Understanding algorithms extract prices, amounts, and dates from PDF & images.
- **Document Management**: Classification algorithms automatically categorize document types (invoice, contract, bill of lading, …).
- **Anomaly Detection**: Algorithms can automatically flag documents with missing or incorrect information.

Trading Case Study: Lifecycle of a Trade - Know Your Customer

Every trade transiting through a Bank is screened. Brokers found guilty of approving transactions involving sanctioned entities can be fined.

Banks must collect identifiable details such as personal information. These identifiers are crossed with public watchlists and flagged if they involve sanctioned individuals.

This is a key topic in crypto.

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Trading Case Study: Lifecycle of a Trade - Know Your Customer

Every trade transiting through a Bank is screened. Brokers found guilty of approving transactions involving sanctioned entities can be fined.

Banks must collect identifiable details such as personal information. These identifiers are crossed with public watchlists and flagged if they involve sanctioned individuals.

This is a key topic in crypto.

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Bottlenecks (non-exhaustive)

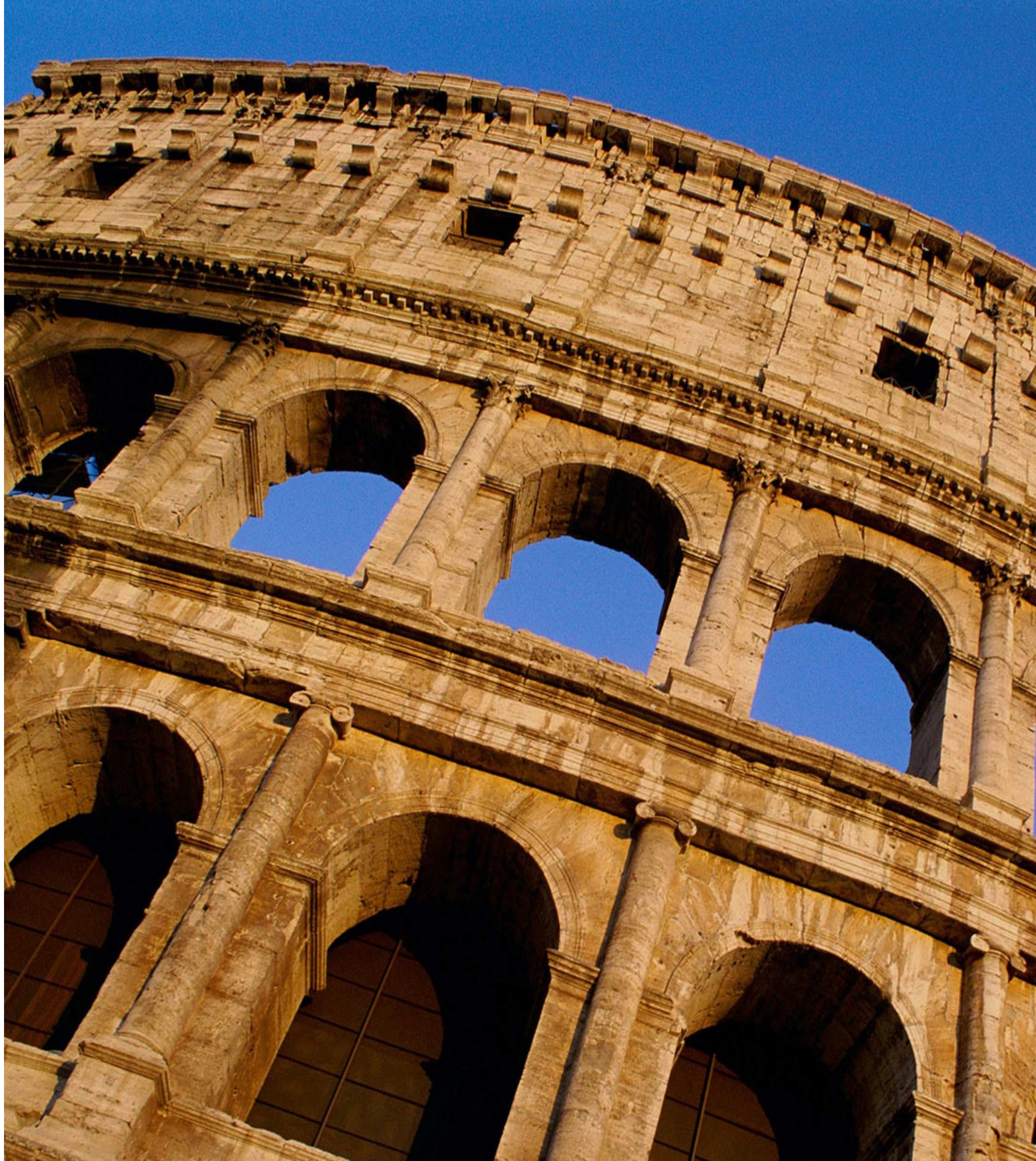
Data Complexity, Volume, Public Data, Regulatory Changes

How can AI help?

- **Entity Recognition**: Matching algorithms can disambiguate companies and persons appearing under multiple names.
- **Public Watchlist Monitoring**: Automatic aggregation of data from public sources can help quickly adapt to geopolitical events.
- **Anomaly Detection**: Algorithms reduce operational workload by automatically reviewing and labelling transactions flagged by algorithms.

Case Study #3: Banking

- Client Onboarding
- Due Diligence
- Portfolio / Loan Management



Investment Banking Case Study: Banking Lifecycle

FRONT OFFICE	
Investment Banking Division (IBD)	
Product Groups	Industry Groups
Mergers and Acquisitions (M&A) **Financing / Global Capital Markets** Debt Capital Market Equity Capital Market Leveraged Finance Capital Market	Healthcare Real Estate Transporatation Utility and Energy Natural Resources Industrials Consumer Retails Media and Telecommunication

MIDDLE OFFICE
Risk Management
Treasury
Others

Corporate and Investment Banks are banks for companies (and governments).

Companies raise money to grow, either organically or through acquisitions. Banks provide companies with debt, assist them in raising equity, or advise on potential M&A deals.

- We compare traditional and AI-driven approaches to the banking lifecycle.

Key Questions to Keep in Mind

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Investment Banking Case Study: Client Onboarding

The investment banker works on behalf of a company looking to sell part or all of its business. They cover companies in a specific sector and are experts in matching companies from this sector with buyers.

In the deal preparation phase, the banker orchestrates the sale/acquisition process, coordinates with potential buyers, and negotiates the best terms.

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Investment Banking Case Study: Client Onboarding

The investment banker works on behalf of a company looking to sell part or all of its business. They cover companies in a specific sector and are experts in matching companies from this sector with buyers.

In the deal preparation phase, the banker orchestrates the sale/acquisition process, coordinates with potential buyers, and negotiates the best terms.

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Bottlenecks (non-exhaustive)

Speed, High Demand, Market Appetite, Marketing, …

How can AI help?

- **Recommendation Systems**: Given historical data of deals involving similar buyers and sellers, how should we orchestrate this deal? What is the likelihood of this deal going through?
- **News Aggregation & Sentiment Analysis**: Which companies are likely to have an appetite for deals based on information they shared publicly?
- **Anomaly Detection**: Which potential buyers can be instantly rejected based on informative metrics?

Investment Banking Case Study: Due Diligence

Once the banker has identified a suitable buyer, they grant them exclusive access to the client's financial and operational information. All relevant information is quickly and thoroughly checked to minimize deal uncertainty.

Bankers and lawyers draft a legal contract for the purchase and agree on a deal valuation.

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Investment Banking Case Study: Due Diligence

Once the banker has identified a suitable buyer, they grant them exclusive access to the client's financial and operational information. All relevant information is quickly and thoroughly checked to minimize deal uncertainty.

Bankers and lawyers draft a legal contract for the purchase and agree on a deal valuation.

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Bottlenecks (non-exhaustive)

Discretion, Modeling Issues, Data Quality, Market Risk, …

How can AI help?

- **Contract Understanding**: Algorithms can scan contracts to avoid clauses which shift the balance of the deal.
- **Compliance & Anomaly Detection**: Client information may contain accounting issues which affect the deal price.
- **Pricing**: Machine Learning can help generate realistic scenarios to stress-test financial models.

Investment Banking Case Study: Portfolio / Loan Management

Aside from M&A, banks lend money to corporate clients to help them execute their strategy. Loan monitoring teams keep track of all the loans the Bank makes and fees it receives.

Multiple banks can participate in a loan, in which case they receive information on all the other banks' participation.

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Investment Banking Case Study: Portfolio / Loan Management

Aside from M&A, banks lend money to corporate clients to help them execute their strategy. Loan monitoring teams keep track of all the loans the Bank makes and fees it receives.

Multiple banks can participate in a loan, in which case they receive information on all the other banks' participation.

Where are the **bottlenecks** in the traditional process?

What is stopping the Bank from **scaling** its action?

Which parts of the traditional process would benefit from being **data-driven**?

Bottlenecks (non-exhaustive)

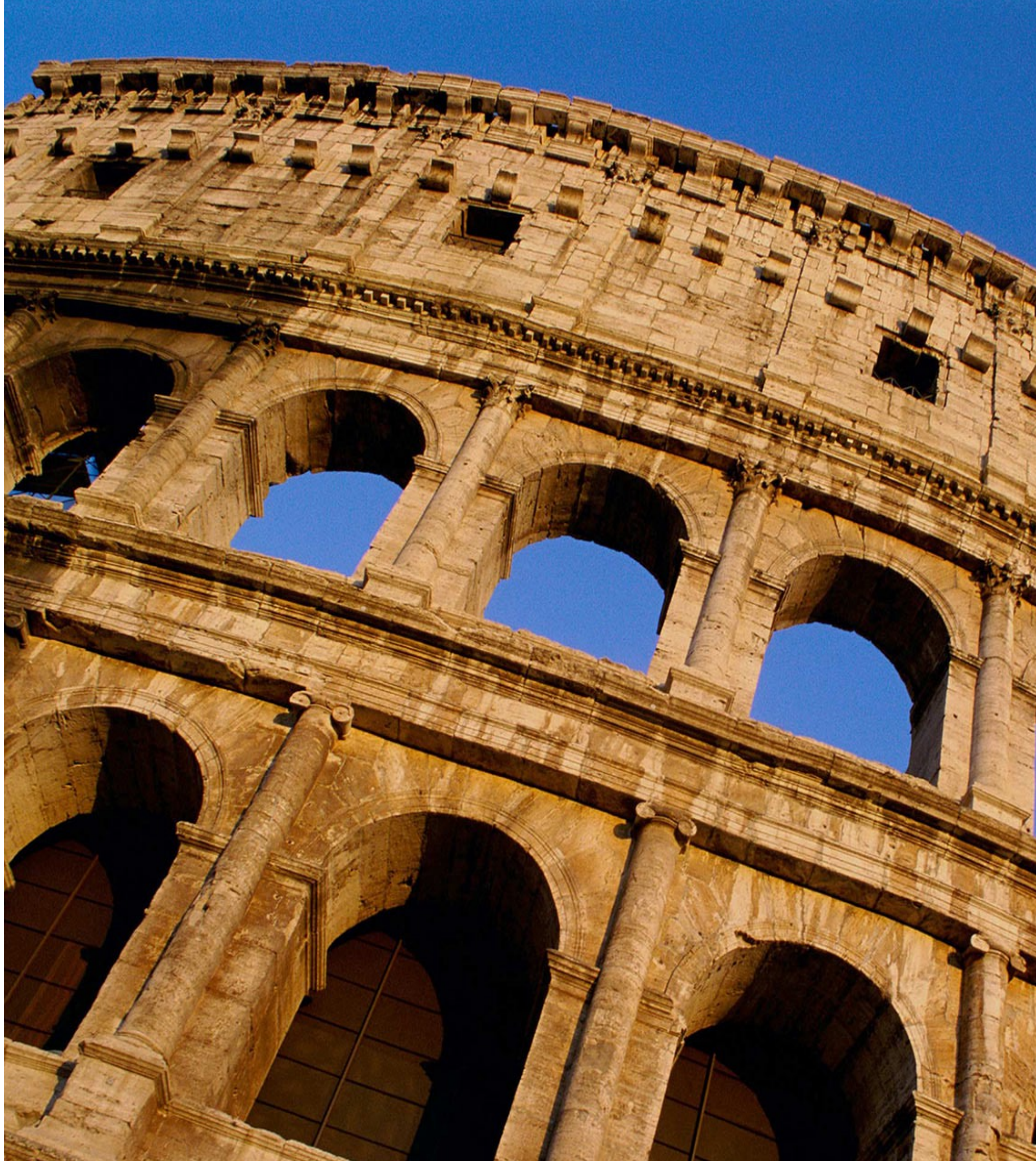
Data Quality, Competition, Yield, Attention, …

How can AI help?

- **Document Understanding**: Loans involving many banks generate large amounts of documentation.
- **Market Intelligence**: Valuable intel on competing banks can be extracted from information companies disclose.
- **Analytics**: Automatically identify low-fee / low-yield loans to reallocate capital to more profitable clients.

Frontiers of AI in Finance

- LegalTech & Contract Understanding
- Document Management & Search Engines
- Causality & Interpretable Machine Learning

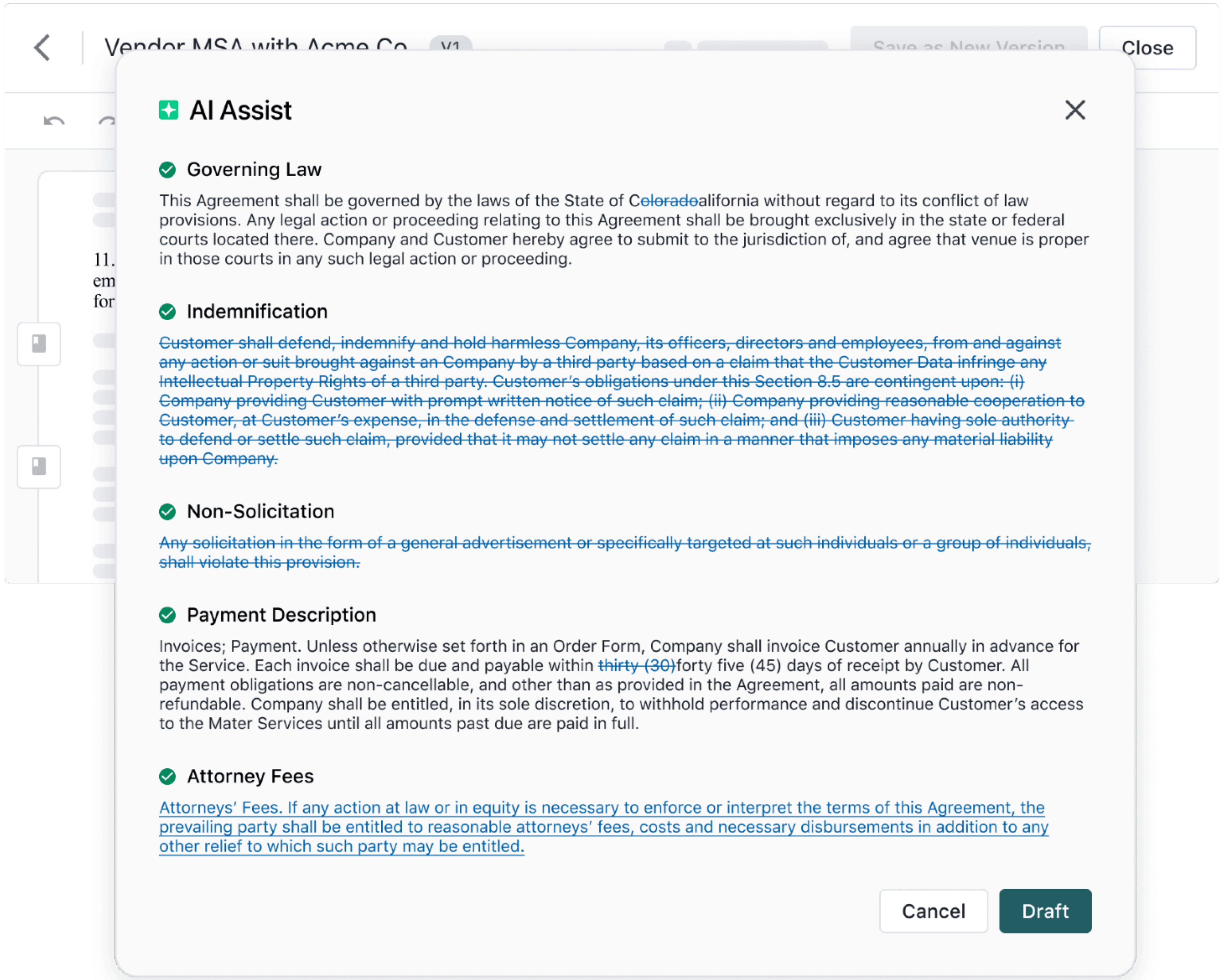


LegalTech & Contract Understanding

Agreements between buyers and sellers are made legally binding with contracts.

- In Banking, large deals or large loans come with equally long contracts.
- In Trading, contracts related to complex products (options, swaps, derivatives) cause many headaches for Trade Validation and Risk teams.

The contract is perhaps the most important document, but can hide unfair clauses. If a bad deal is signed, it can become valid!



Clause extraction software can identify provisions in contracts and summarize them into simple English (ironclad)

Example Financial use-cases

Under which conditions does this contract allow me to sell my participation in a loan?

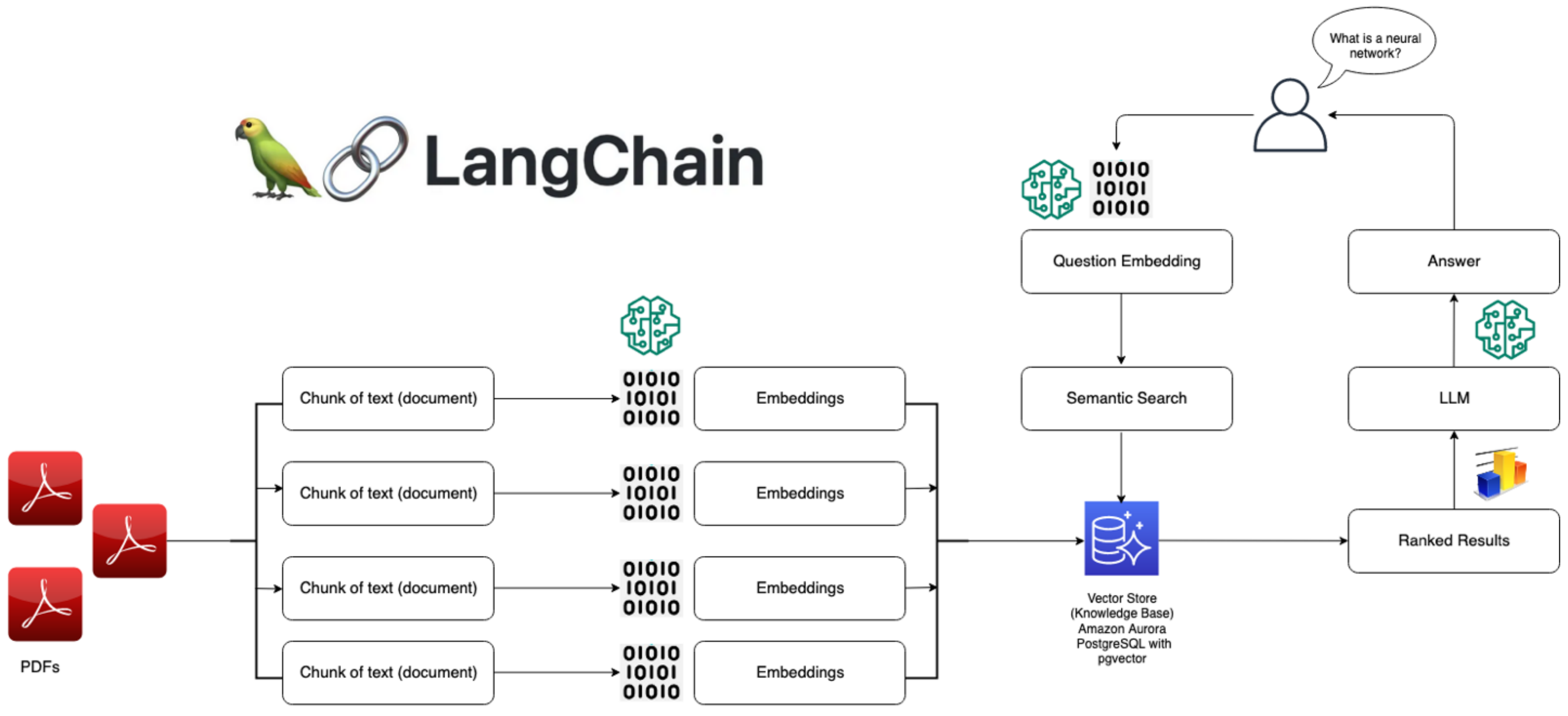
What does this company mean by "Entry into a Material Agreement"?

Document Management & Search Engines

Advances in Natural Language Processing make it possible to directly query documents.

Combined with a recommendation system, this can effectively create a search engine across all files in a database.

This allows users to query information from private files without sharing data externally.



Example Financial use-cases

Which loans on our portfolio could potentially be affected by the Credit Suisse blow-up?

What was the sentiment of analysts in this earnings call?

What differences exist between these two versions of the same file?

Causality & Interpretable Machine Learning

Causal Inference is a framework used in Statistics to quantify the effect variables have on outcomes.

Given the importance of audits in Finance, causality could help bridge the gap between opaque AI models and explainable financial models.

Causal Inference plays a large role in Economics, but is challenging to implement in practice and requires extreme care to avoid misinterpreting results.

Example Financial use-cases

What is the effect of an M&A announcement on the share price of the target?

Why was this transaction flagged as fraudulent?

What is the effect of an increase in interest rates on depositor withdrawal rates?

(1) Economists use Causal Inference…

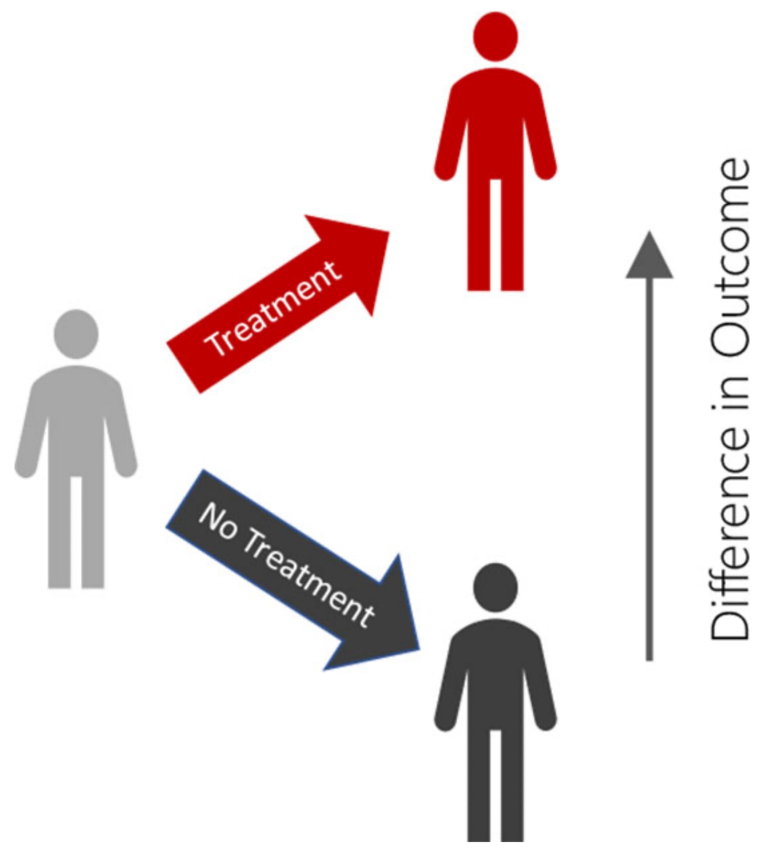
…to evaluate the effect of policies on labor markets.

(2) Pharma companies use Causal Inference…

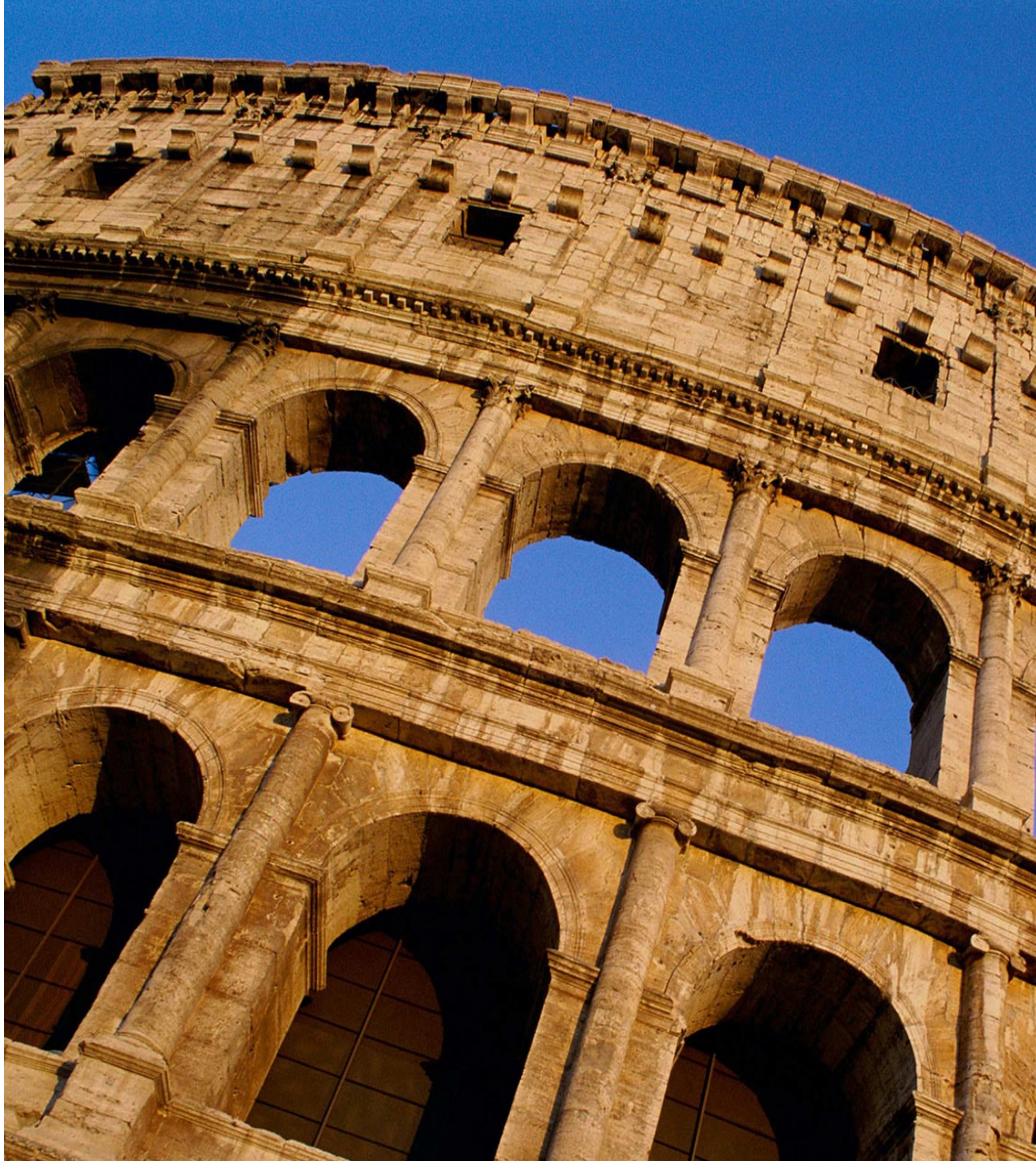
…to measure the effect of a treatment (vaccine, antibiotic, etc.) on a patient's medical condition.

(3) Social media companies use Causal Inference…

…to measure the effect of a UI change on user behavior.

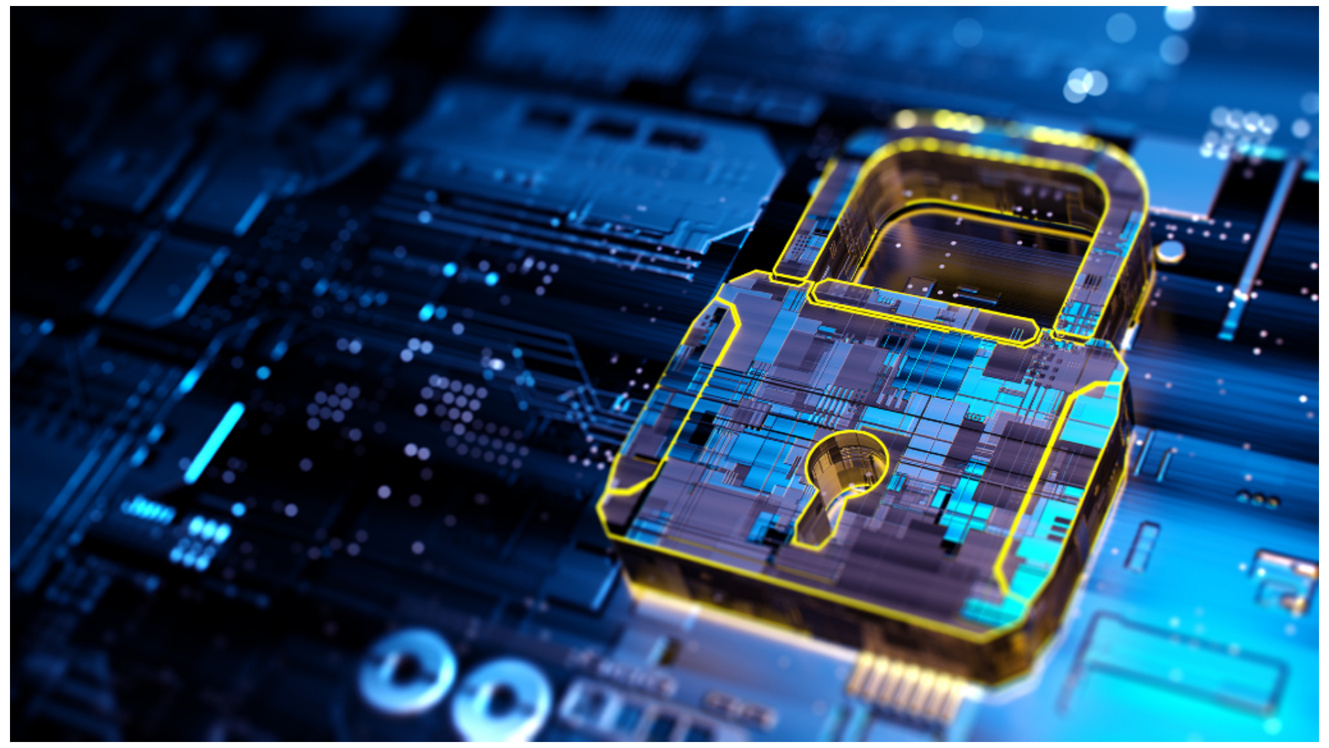


Risks for AI Adoption in Finance

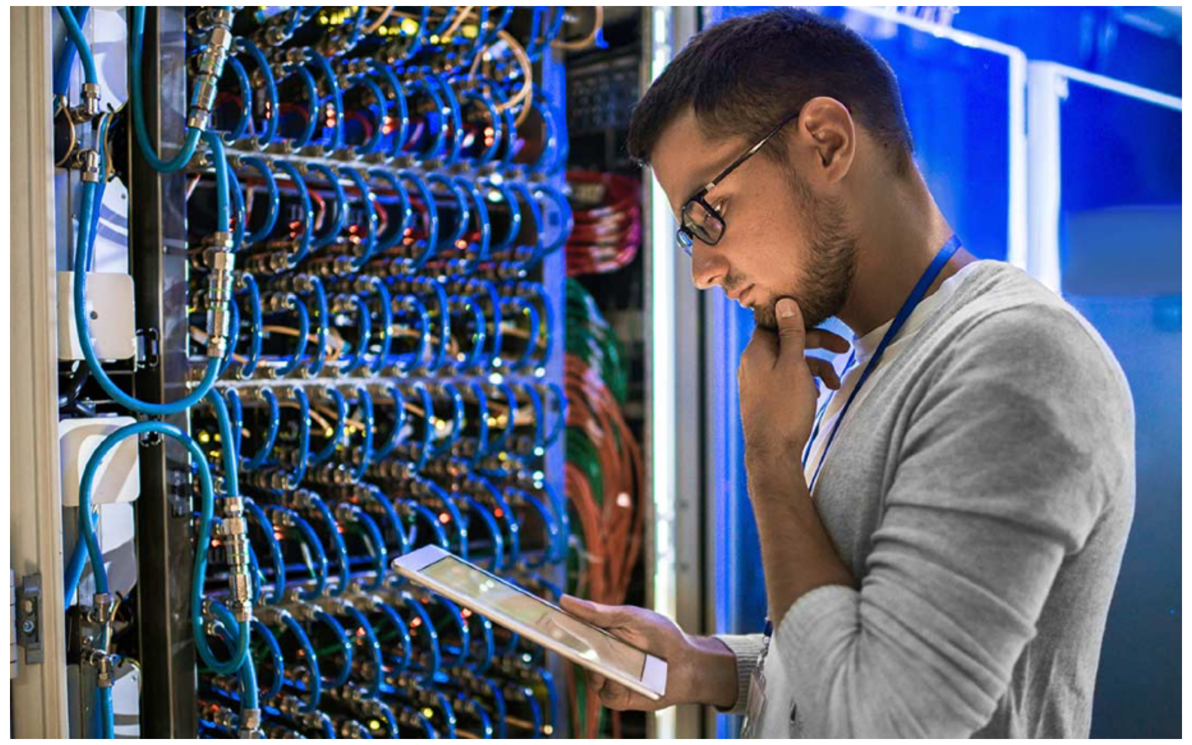


Risks From AI in Finance

Explainability

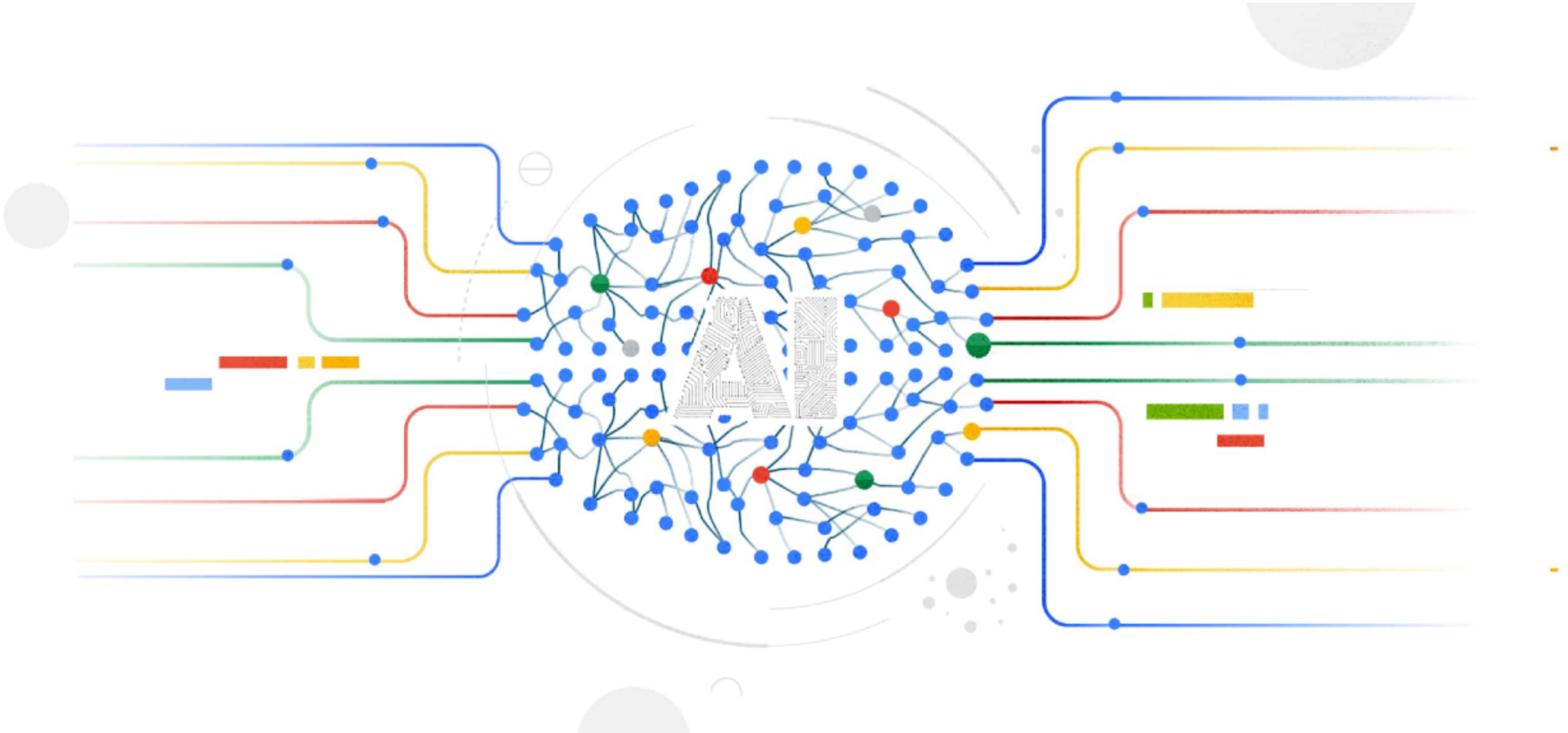


Cybersecurity



Infrastructure

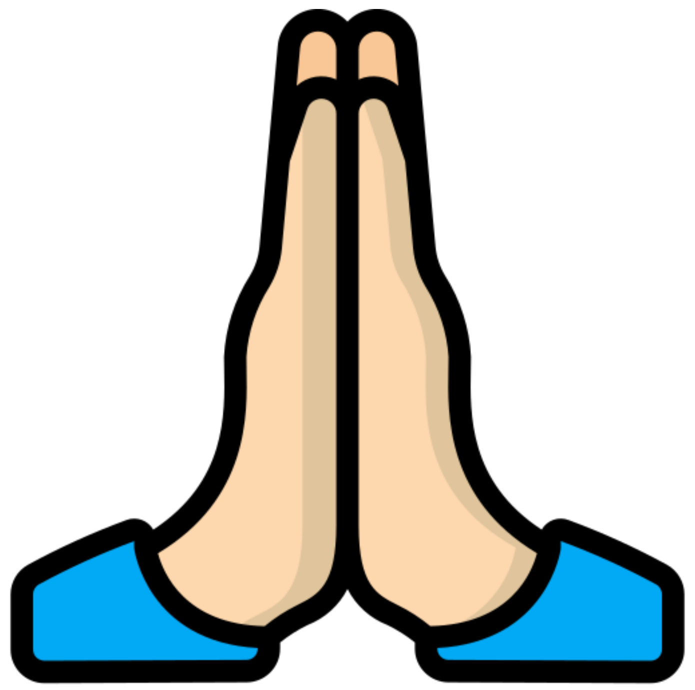
Data Privacy



Model Errors

Final Thoughts

- Many AI in Finance topics we didn’t cover (Asset Management, Private Equity, Equity Research, etc.)
- All feedback welcome (positive and negative)
- Slides available at github.com/mtchibozo/hec-international-fintech
- Many of these challenges will be up to you to solve, **start coding!**
- If you can’t beat them, use them ; algorithms easily replace simple time-consuming tasks, but human supervision will still be required for complex, risky, and capital-intensive tasks



Thank You!