



# Algorithmic Trading

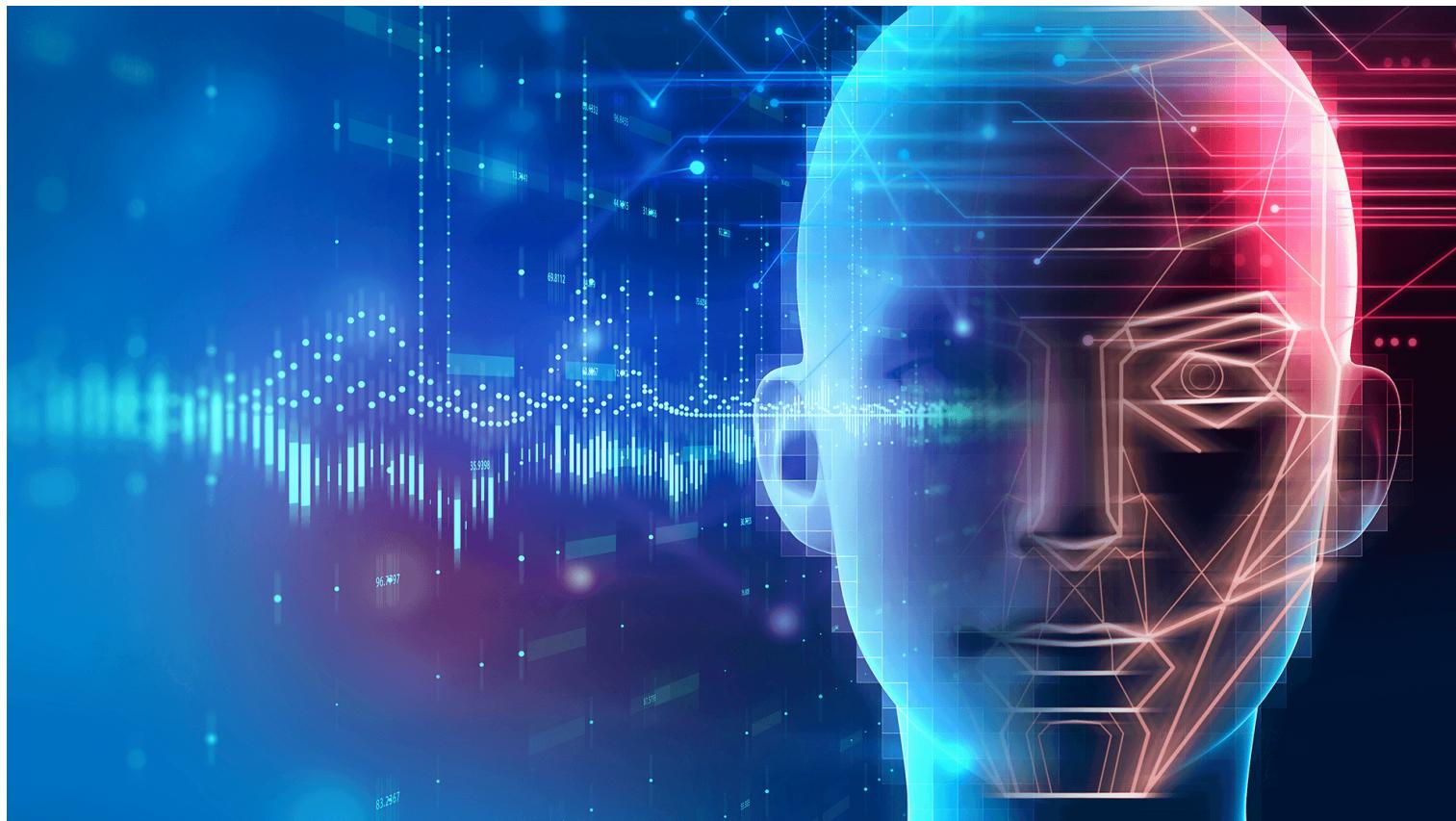
**Dr. Javier Sabio González**

*Head of Advanced Analytics & Algorithmic Trading, BBVA C&IB*

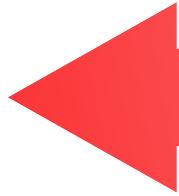
# Topics

1. Markets Macrostructure
2. Markets Microstructure
3. Algorithmic Trading Fundamentals
4. Algorithmic Execution
5. Algorithmic Market-Making
6. Algorithmic Investment
7. **The Future of Algorithmic Trading**

## 7. The Future of Algorithmic Trading



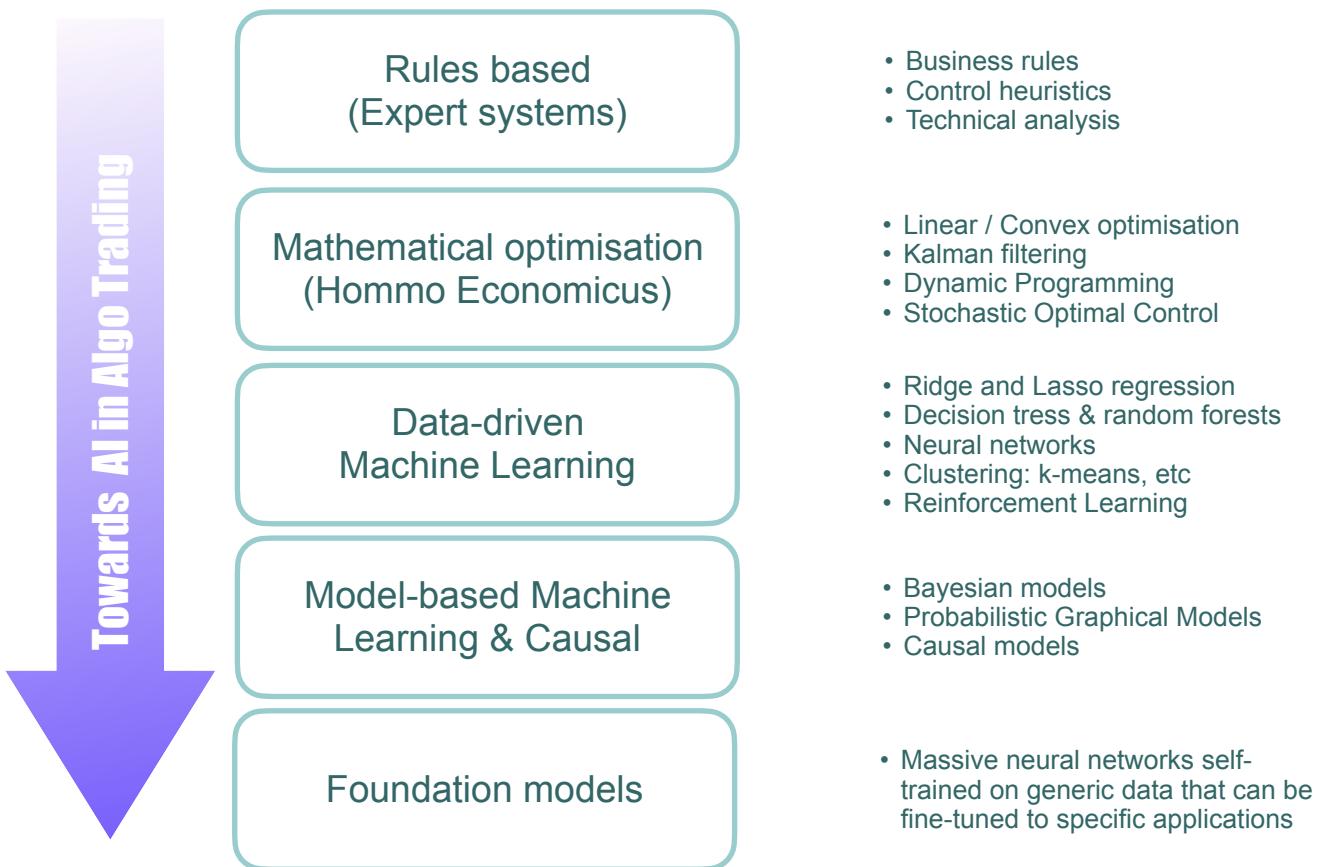
# Megatrends in Financial Markets & Algo Trading



Impact on Algorithmic Trading

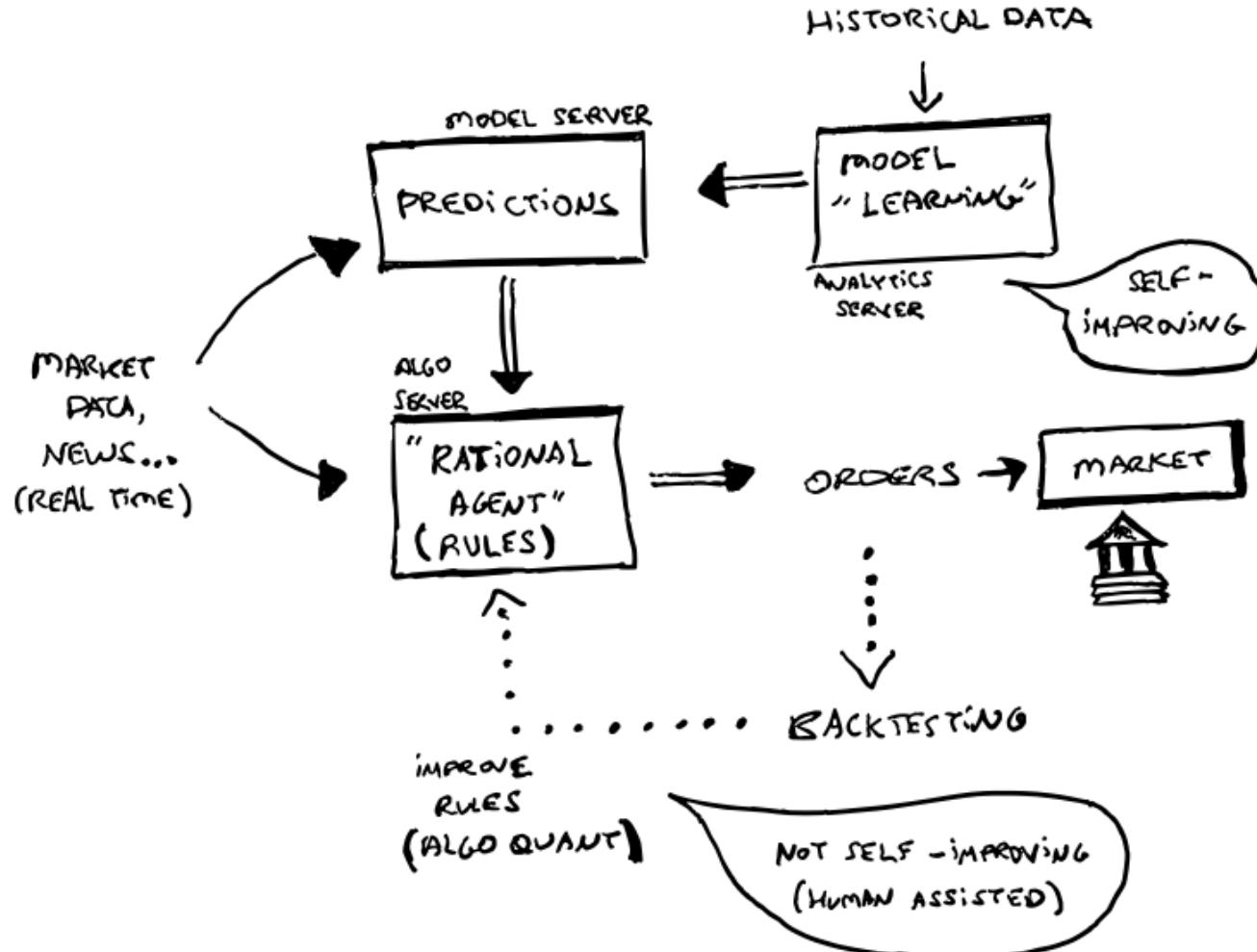
# 1. Artificial Intelligence

In [computer science](#) AI research is defined as the study of "[intelligent agents](#)": any device that perceives its environment and takes actions that maximise its chance of successfully achieving its goals.<sup>[1]</sup> Colloquially, the term "artificial intelligence" is applied when a machine mimics "cognitive" functions that humans associate with other [human minds](#), such as "learning" and "problem solving" (Wikipedia)

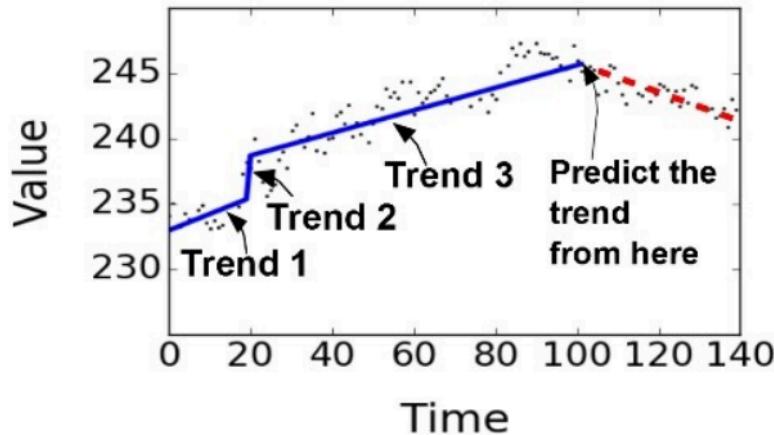


	How is it used?	Pros	Cons	Real examples (banks and fintech)
Rules based	<ul style="list-style-type: none"> <li>To code all kind of strategies (execution, market-making, investment)</li> <li>In mathematical or data-driven alots, to include corner cases &amp; exceptions not yet captured by the algorithms</li> </ul>	<ul style="list-style-type: none"> <li>Encodes business knowledge</li> <li>Easy to interpret</li> <li>Quick to build a first version</li> </ul>	<ul style="list-style-type: none"> <li>Difficult to escale up and maintain</li> <li>Not necessarily optimal</li> <li>Does not exploit data in its full potential</li> </ul>	<ul style="list-style-type: none"> <li>Suites of algos (execution, quoting) in traditional trading platforms like <a href="#">ION Trading</a> or <a href="#">Itiviti</a></li> <li>Suites of execution algos from brokers, for example <a href="#">Barclays</a>, <a href="#">Goldman Sachs</a> or <a href="#">Natixis</a></li> </ul>
Mathematical optimisation	<ul style="list-style-type: none"> <li>As the backbone for all kind of strategies</li> <li>Typically supplemented by rules</li> <li>Latest generation of algos use data-driven models inside</li> </ul>	<ul style="list-style-type: none"> <li>Easy to interpret</li> <li>Efficient to run when closed-forms are available</li> </ul>	<ul style="list-style-type: none"> <li>Model assumptions might be wrong or oversimplifying</li> <li>Difficult to evolve and adapt to practical situations</li> </ul>	<ul style="list-style-type: none"> <li>Many market-making algos in large banks are based in the <a href="#">Avellaneda-Stoikov</a> paradigm</li> <li>Popular execution algos like VWAP or Implementation Shortfall included in most suites of execution algos, see for instance <a href="#">Natixis</a></li> </ul>
Data-driven Machine Learning	<ul style="list-style-type: none"> <li>To build predictors and signals used by trading strategies</li> <li>For anomaly detection &amp; regime change</li> <li>To build the optimal strategy learning from data (Reinforcement Learning)</li> </ul>	<ul style="list-style-type: none"> <li>Very good performance in many situations</li> <li>Standard models are available in open source libraries</li> </ul>	<ul style="list-style-type: none"> <li>Black-box in many cases</li> <li>Requires large amounts of data (sometimes simulated)</li> <li>Not robust against changes in environment if not trained on purpose</li> <li>Difficult to use it for what if questions &amp; counterfactuals</li> </ul>	<ul style="list-style-type: none"> <li>ML Trading Analytics: <a href="#">Mosaic Smart Data</a> (Fintech)</li> <li>ML Market-making: <a href="#">Overbond</a> (Fintech), <a href="#">JP Morgan</a></li> <li>ML Execution: <a href="#">Borealis AI (Fintech)</a> with RBC, <a href="#">JP Morgan's LOXM</a></li> </ul>
Model-based Machine Learning	<ul style="list-style-type: none"> <li>To build predictors and signals used by wider strategies</li> <li>For anomaly detection &amp; regime change</li> <li>To build the optimal strategy learning from data</li> <li>To generate synthetic data avoiding spurious confounders</li> </ul>	<ul style="list-style-type: none"> <li>Requires less data</li> <li>Incorporates business prior knowledge easily</li> <li>Easier to interpret</li> <li>More robust against changes in environment</li> <li>Allows what if questions and counterfactuals</li> </ul>	<ul style="list-style-type: none"> <li>Requires prior knowledge</li> <li>Less developed than data-driven ML (less expertise, libraries, frameworks etc)</li> </ul>	<ul style="list-style-type: none"> <li><a href="#">Causal Lens</a> (Fintech)</li> </ul>
Foundation models	<p>Potential applications:</p> <ul style="list-style-type: none"> <li>Code implementation: copilot, translator high level — low level</li> <li>Blend unstructured alternative data into signals fed to algos</li> <li>Generative models for synthetic data including alternative data</li> <li>Human interfaces: launch algos, analyse data, RfQs, etc</li> </ul>	<ul style="list-style-type: none"> <li>Pre-trained large models</li> <li>Blending of different sources of data</li> <li>Self-learning</li> <li>Import patterns from other domains</li> <li>Simplify interfaces for business users</li> </ul>	<ul style="list-style-type: none"> <li>Costly last-mile fine-tuning</li> <li>Solutions are “approximately correct” in many cases</li> <li>False sense of confidence</li> <li>Opaque</li> </ul>	<ul style="list-style-type: none"> <li>BERT, <a href="#">Chat GPT</a> (GPT 3.5) although not yet specific applications in algo trading</li> </ul>

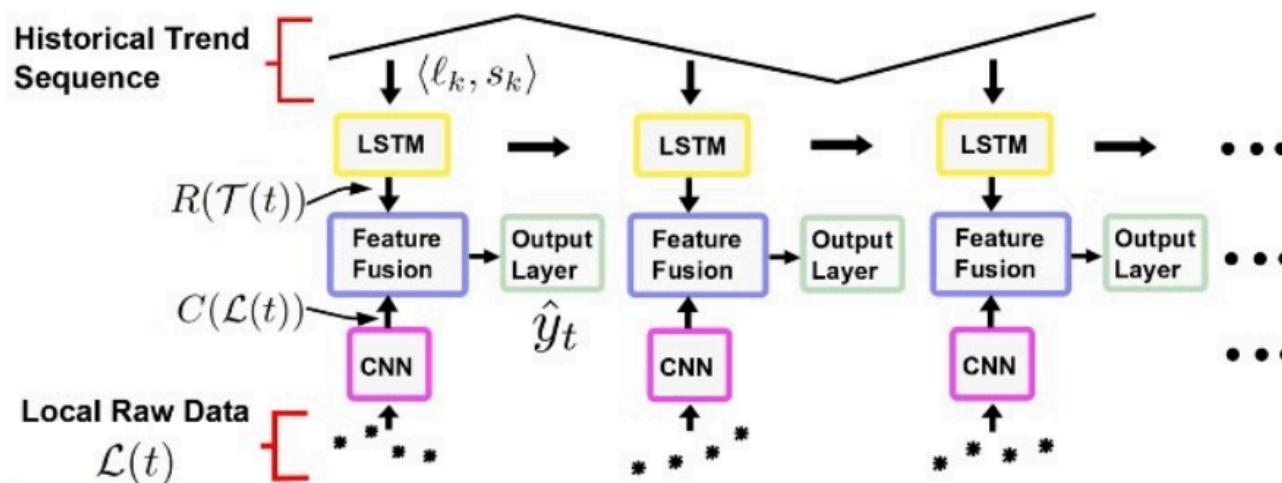
Most of productive trading algorithms today are yet based in rules / heuristics or mathematical optimisation frameworks. AI/Machine Learning is introduced via signals, predictors, and alerts:



For instance, a typical application of neural networks, or in general “discriminative” algorithms, is to use them for prediction of future market conditions: price trends, volatility, volume, demand, etc



**Example:** TreNet Hybrid Neural Network  
(CNN + LSTM) for trend prediction



There is a lot of innovation and experimentation on producing fully data-driven trading algorithms, i.e. using neural networks and reinforcement learning to build the actual strategy that solves a certain business objective: minimise transaction costs in execution, maximise P&L at a level of risk, etc

Trading decision:  
place order, change  
order, cancel order ...



P&L vs benchmark  
Inventory level  
Risk level

Trading decision:



**(Deep) Reinforcement Learning**

The paradigm is similar to the one used by Google's Deep Mind to build Alpha Go. These systems learn autonomously the optimal strategy just by observing and interacting with the environment. In this example, one of these algorithms learn to play a simple Atari game and quickly has super-human performance:



Some banks are offering such Reinforcement Learning based algorithms, mostly for execution, where the objective function (minimise transaction costs) is more local and well defined. Still, they are yet the exception and mostly used for marketing purposes to showcase up the AI capabilities of the bank.

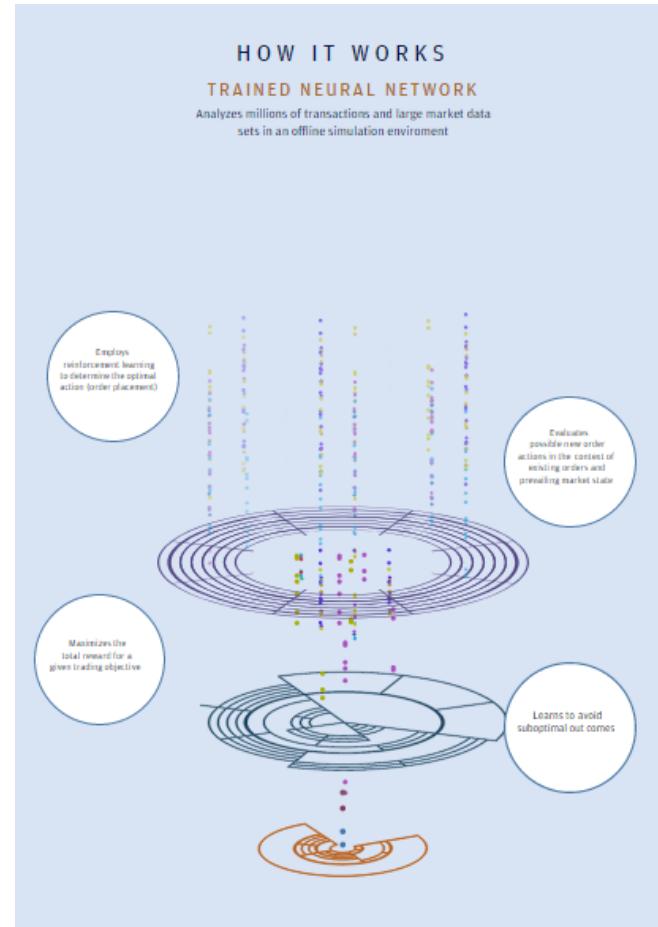
**Introducing LOXM**

Advanced equities trading powered by machine learning

LOXM, exclusively from J.P. Morgan, helps drive best execution for equities trading. Here's how.

OVERVIEW

With LOXM, we have introduced machine learning-driven execution capabilities to our proprietary equities algorithms. LOXM uses a trained neural network to optimize between liquidity demand and passive trading, based on the prevailing market state, adapting as market conditions change. Using reinforcement learning techniques, LOXM can analyze millions of simulated transactions in light of historical data to zero in on hard-to-spot optimal behaviors.



[Commodities](#)[Debt Origination](#)[\*\*Electronic Trading\*\*](#)[Equities](#)[Equity Linked Products](#)[Foreign Exchange](#)[Fixed Income](#)

Aiden is an AI-based electronic trading platform that uses the computational power of deep reinforcement learning in its pursuit of improved trading results and insights for clients.

[Request a Demo](#)

[https://www.rbccm.com/en/expertise/electronic-trading/ai-trading.page?utm\\_source=rbc&utm\\_medium=va&utm\\_campaign=aiden](https://www.rbccm.com/en/expertise/electronic-trading/ai-trading.page?utm_source=rbc&utm_medium=va&utm_campaign=aiden)

## **Why has been Deep Mind so successful with this approach and not the financial industry?**

The trading environment poses many challenges:

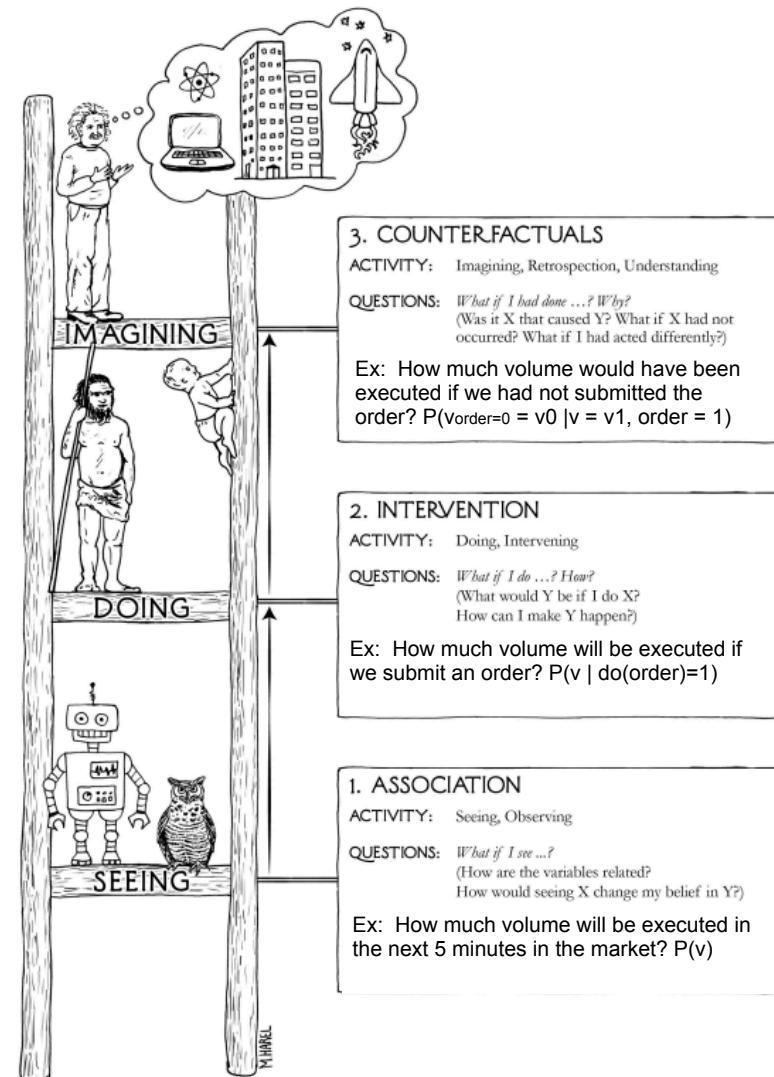
- **Stochastic environment:** the “rules” of the game are not well defined as in deterministic setups like Go, chess or even video games. The link between actions and rewards is weaker, and has to be analysed in statistical / probabilistically terms
- **Dynamic environment:** markets are continuously changing and adapting to new regimes, so past historical data might not be useful for future decisions. This is particularly harmful for techniques that require massive amounts of data
- **Training data:** pure historical market data is not well suited for training these algos, even if we consider it to be relatively stationary (which is not). These algos learn by experimentation, i.e. choosing actions and observing the reward from the environment. Past data does not have information on how the market would have reacted to these orders. Simulation can be used to address this issue, but if the simulator does not capture the real market dynamics (which is difficult), the trained algo will be good at beating the simulator but not the real market
- **Interpretability:** these techniques have the capability of choosing actions with a long term forward looking that makes them very successful in games like Go or chess. However, in a trading setup they might produce behaviours that are difficult to interpret and make the human traders responsible for the P&L and book of the algo uncomfortable

In general, one issue with the current “data-driven” AI/ML paradigm is that these techniques can solve very well specific tasks where a lot of data is available, but they lack the generalist capabilities of the human brain, which can:

- learn from few examples
- bring expertise from other domains
- ask what if & counterfactual questions.

A new generation of **model based / causal inference machine learning** is expected to produce a leap forward, although it is still a work in progress

The Book of Why: The New Science of Cause and Effect – Pearl and Mackenzie



# Causal Factor Investing: Can Factor Investing Become Scientific?

75 Pages • Posted: 15 Sep 2022 • Last revised: 2 Jan 2023

Marcos Lopez de Prado

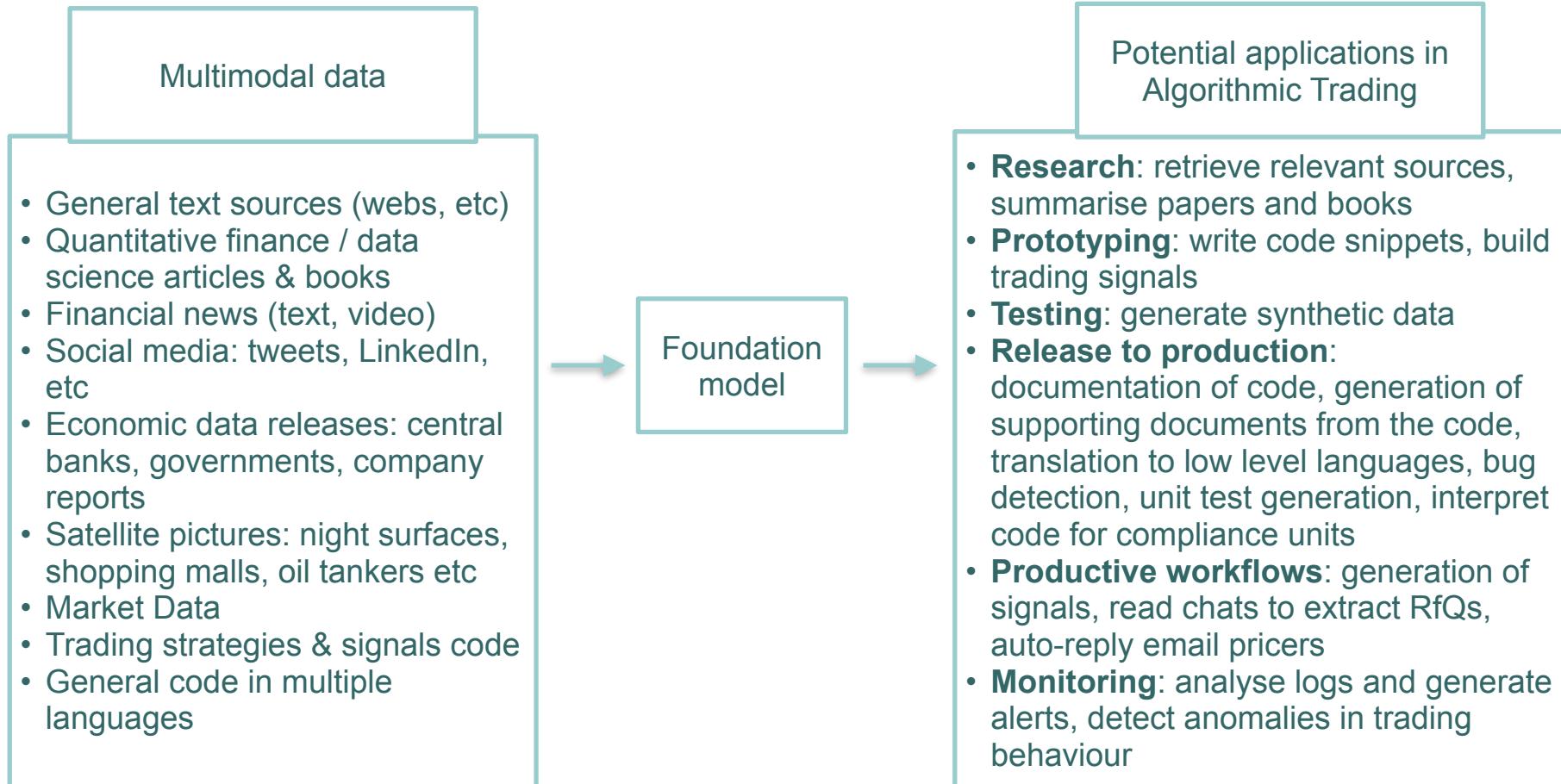
Cornell University - Operations Research & Industrial Engineering; Abu Dhabi Investment Authority; True Positive Technologies

Type	Rigor	Example
<b>Randomized controlled trials</b>	Very high	Algo-wheel experiments (e.g., section 3.3)
<b>Natural experiments</b>	High	Market-maker reaction to random spikes in order imbalance (e.g., section 3.3)
<b>Simulated interventions</b>	Medium	Estimate effect of HML using a causal graph (e.g., section 6.2.1)
<b>Econometric (observational) studies</b>	Low	Factor investing literature; backtested investment strategies (e.g., section 6.4.1)
<b>Case studies</b>	Very low	Broker report / analysis
<b>Expert opinion</b>	Anecdotal	Investment guru's prediction

Figure 12 – Hierarchy of evidence in financial research, ranked by scientific rigor

In terms of generalist capabilities, an alternative and promising route is the development of foundation models, massive neural networks pre-trained on a variety of data-sources (text, pictures, etc) in a self-supervised way, that can then be fine-tuned to specific applications.

Although there is a lot of recent interest, though, practical integrations in Algorithmic Trading workflows are still a **work in progress**



# Examples using Chat GPT

which applications of foundation models like gpt could be used for algorithmic trading?

GPT (Generative Pre-training Transformer) is a large language model that has been trained to generate human-like text. It could potentially be used in algorithmic trading in a few ways:

1. Natural Language Processing: GPT or similar models could be used to process and analyze news articles or other textual data to extract valuable information that could be used to inform trading decisions.
2. Sentiment Analysis: GPT or similar models could be used to analyze social media posts or news articles to gauge sentiment towards a particular company or industry, which could be used to inform trading decisions.
3. Generative Trading Strategies: It is possible that GPT or similar models could be used to generate trading strategies based on historical data and other inputs. However, this would likely be a very complex and risky endeavor, and it is not clear how successful it would be.

It's important to note that while GPT and other foundation models have shown remarkable capabilities in a variety of tasks, they are not a panacea and their use in algorithmic trading (or any other domain) would likely require significant domain expertise and careful consideration.

write code for a trading algorithm that buys shares splitting the cost over time bins in order to achieve the vwap price over a period

Here is some pseudocode for a trading algorithm that buys shares by splitting the cost over time bins in order to achieve the Volume Weighted Average Price (VWAP) over a period:

```
Copy code
```

```
def vwap_trading_algorithm(data, shares_to_buy, start_time, end_time, bin_size):  
    # data is a list of dictionaries, each representing a tick of data  
    # each dictionary has the following keys: 'timestamp', 'price',  
    'volume'  
  
    # initialize variables  
    total_volume = 0  
    total_cost = 0  
    current_bin_start_time = start_time  
    current_bin_end_time = start_time + bin_size  
  
    # iterate through the data  
    for tick in data:  
        # check if the current time is within the current time bin
```

write code to generate a sentiment signal based on tweets

Here is some sample Python code that generates a sentiment signal based on tweets using natural language processing (NLP):

```
Copy code
```

```
import nltk  
import pandas as pd  
  
def generate_sentiment_signal(tweets):  
    """  
    Generates a sentiment signal based on the sentiment of a list of tweets.  
  
    Parameters  
    -----  
    tweets : list of str  
        The list of tweets.  
  
    Returns  
    -----  
    float  
        The sentiment signal.  
    """  
    # Preprocess the tweets  
    preprocessed_tweets = preprocess_tweets(tweets)  
  
    # Tokenize the tweets  
    tokenized_tweets = [nltk.word_tokenize(tweet) for tweet in preprocessed_tweets]  
  
    # Tag the parts of speech of the tokens
```

## 2. Cloud computing

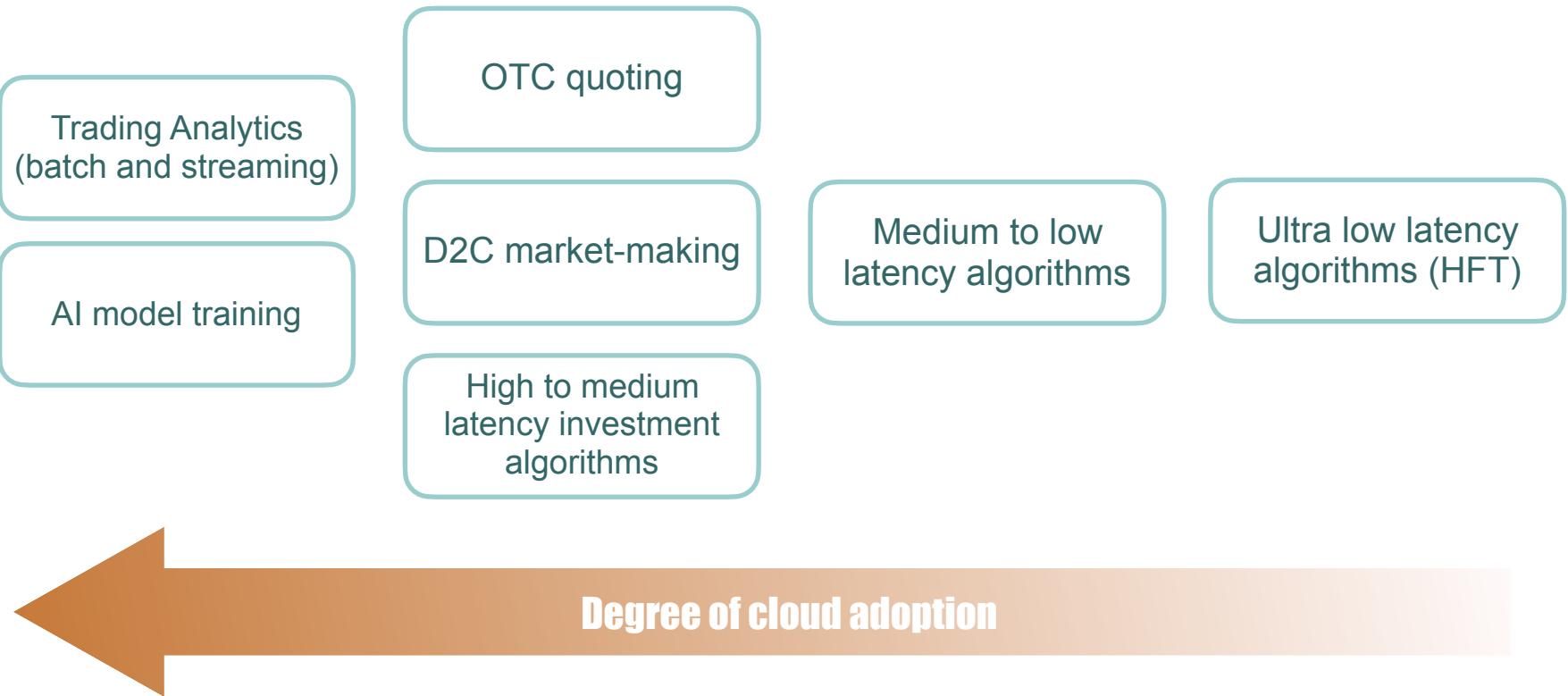
**Cloud computing**<sup>[1]</sup> is the on-demand availability of **computer system resources**, especially data storage (**cloud storage**) and **computing power**, without direct active management by the user. Large clouds often have functions **distributed** over multiple locations, each location being a data center (Wikipedia)



Trading applications, and in particular trading algorithms, are increasingly moving to cloud platforms, as a way to benefit from the latest technologies, increase scalability on demand, and get transparent costs of infrastructure and software

Although competition is intense, and the landscape evolves quickly, for trading applications it is **AWS** the most mature and extended cloud provider.

**Ultra low latency** is probably the most significant deterrent to a full-scale migration, for those use cases that require ms or sub-ms speeds and are colocated to exchanges (HFT)



# Examples

Trading analytics



Cloud-First, Real-Time  
Streaming Analytics

AI/ML training



High to medium frequency algo trading



Build on Our Platform

Build your quantitative organization from our cloud platform to leverage our 8 years of experience. With over 100,000 users we have proven we can scale to meet even the most ambitious of requirements. With code encryption and version control you can be sure your intellectual property is safe.

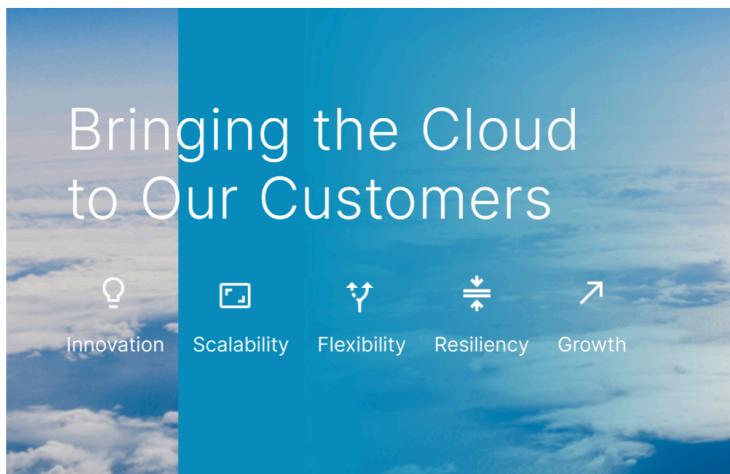
# Is the future an all in cloud trading ecosystem?



JULY 23, 2019

## Equinix to Offer Enhanced Cloud Connectivity with Amazon Web Services

Faster Connection Capacities Help Accelerate Enterprise Cloud Transformations



Bringing the Cloud to Our Customers

Innovation   Scalability   Flexibility   Resiliency   Growth

### Nasdaq and AWS Partner to Transform Capital Markets

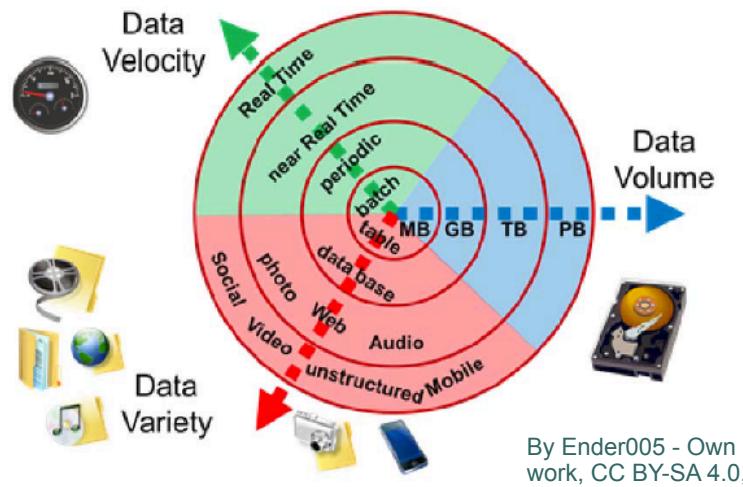
Nasdaq and AWS have announced a multi-year partnership to build the next-generation of cloud-enabled infrastructure for the world's capital markets. The partnership will enable the migration of Nasdaq's North American markets to the AWS cloud, beginning in 2022 with one of Nasdaq's U.S. options markets.

> [Read the Press Release](#)

> [View Frequently Asked Questions](#)

### 3. Big Data & Alternative Data

**Big data** is a field that treats ways to analyse, systematically extract information from, or otherwise deal with data sets that are too large or complex to be dealt with by traditional data processing application software. Big data was originally associated with three key concepts: *volume*, *variety*, and *velocity* (Wikipedia)



By Ender005 - Own work, CC BY-SA 4.0,

In practice, you may have big data if some of the following apply:

- Your data cannot be stored and analysed in a single computer
- Your application struggles to keep up with the amount of real-time data
- Your data cannot be handled with traditional relational databases like SQL

Trading algorithms have traditionally been designed and backtested with large amounts of highly structured time-series market data, from **intraday prices to limit-order-book tick data**, both in real-time (streaming) and batch. In order to handle this kind of *big data*, the industry has traditionally resorted to specialised databases and analytics engines like KDB+.

However, the majority of data stored today and considered big data is highly unstructured (pictures, movies, tweets, etc). Trading algorithms have historically not used this kind of unstructured data, but this is quickly changing with the advent of **alternative data sources**: datasets of value for investment and financial decisions that are increasingly easier to obtain in a curated way, with **multiple fin-techs** operating in this space. The analytics stack to analyse it, once the realm of big players like Google and Facebook, is now mainstream: Hadoop, Spark, Kafka, etc



Source: Quandl

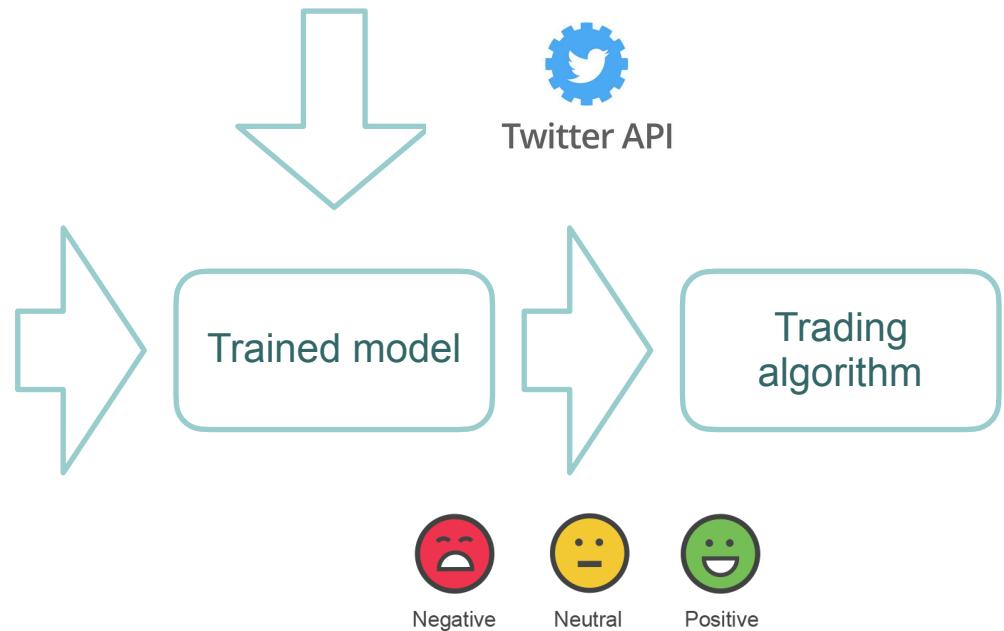


Fin-techs

# Example: sentiment analysis of Elon Musk's tweets



	tweet	subjectivity	polarity	analysis
0	👀	0.000000	0.000000	Neutral
1	: Cybertruck at Giga Texas	0.000000	0.000000	Neutral
2	🎸🎸 Austin Rocks!! 🎸🎸	0.000000	0.000000	Neutral
3	Same with Berlin	0.125000	0.000000	Neutral
4	Limited production of Model Y this year, high...	0.227619	0.029524	Positive
...	...	...	...	...
195	There will be no handles	0.000000	0.000000	Neutral
196	If there's ever a scandal about me, please c...	0.000000	0.000000	Neutral
197	_sci This comment thread is 🔥	0.000000	0.000000	Neutral
198	Don't defy DeFi	0.000000	0.000000	Neutral



## 4. Blockchain & decentralised Finance (DeFi)

**Blockchain:** decentralised ledger (book or collection of accounts in which transactions are recorded), in which every node (computer in a network) has a copy, and can add new transactions in blocks via consensus with other nodes. The history is transparent and immutable.

**Cryptocurrency:** a digital token that is cryptographically secured and transferred using blockchain technology. Leading examples are Bitcoin and Ethereum. Many types of different cryptocurrencies exist, such as stablecoins and tokens that represent digital and non-digital assets.

**Smart contract:** programs stored on a blockchain that run when predetermined conditions are met. They typically are used to automate the execution of an agreement so that all the participants can be immediately certain of the outcome. Users interact with smart contract by sending transactions that trigger the execution of the function that defines the smart contract. They are one of the salient features of the Ethereum blockchain vs Bitcoin (as well as many new ones like Solana)

**Decentralised Finance (DeFi):** a financial infrastructure that does not rely on centralised institutions like banks. Exchange, lending, borrowing and trading are conducted on a peer-to-peer basis using blockchain technology and smart contracts

**Automated Market Maker (AMM, within DeFi):** a smart contract that holds assets on both sides of a trading pair and continuously quotes a price for buying and for selling. Based on executed purchases and sales, the contract updates the asset size behind both the bid and the ask and uses this ratio to define a pricing function

# Crypto vs DeFi

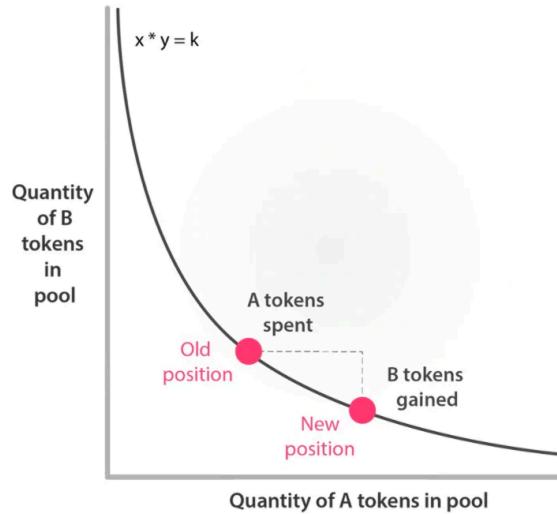
- When using algos to trade **cryptocurrencies** in exchanges like [Coinbase](#), the fundamentals are the same as per any other asset class. These exchanges typically use Limit Order Book structures and most of the trading algorithms used in other assets can be easily adapted.
- **DeFI**, on the other hand, represents a new playground for trading algorithms, with new rules that potentially require deep adaptations of previous algorithms or the design of new ones from scratch. To mention some (see [link](#) for a detailed discussion):
  - **Block time**: transactions are only confirmed when the current block is confirmed in the blockchain, which for Ethereum takes on average 10-15s.
  - **Trade transparency and front running**: pending transactions sit in pending transaction pool (mempool) where contents are visible to everyone. Moreover, the order of the transactions within a block is controlled by miners, who might have economic incentives to order them in a way that benefits them, for instance if they are both miners and arbitrageurs (so called [MEV](#), Maximal Extractable Value). Such a front running is therefore a very present risk in the DeFi ecosystem, a practice that is not illegal in these applications, given the full transparency of the system and the lack of regulatory supervision so far.
  - **Costs**: transactions in blockchains like Ethereum have a cost (called “gas”), which is determined in auctions (Priority Gas Auctions, PGAs). This makes transaction cost estimation challenging for trading algorithms operating in the DeFi ecosystem.

# Algorithmic Trading in DeFI: Automated Market Makers (AMM)

- **Automated Market Makers (AMM)** like [UniSwap](#), are trading algorithms that are used to create liquidity in a pair of tokens (e.g. cryptocurrencies) by incentivising agents (liquidity providers) to contribute their tokens into a pool from which other agents (liquidity takers) can exchange their tokens in a proportion given by a mathematical formula. For instance, a typical one is the “constant function market - maker”:

$$x \times y = k$$

TOKEN X QUANTITY      TOKEN Y QUANTITY      CONSTANT

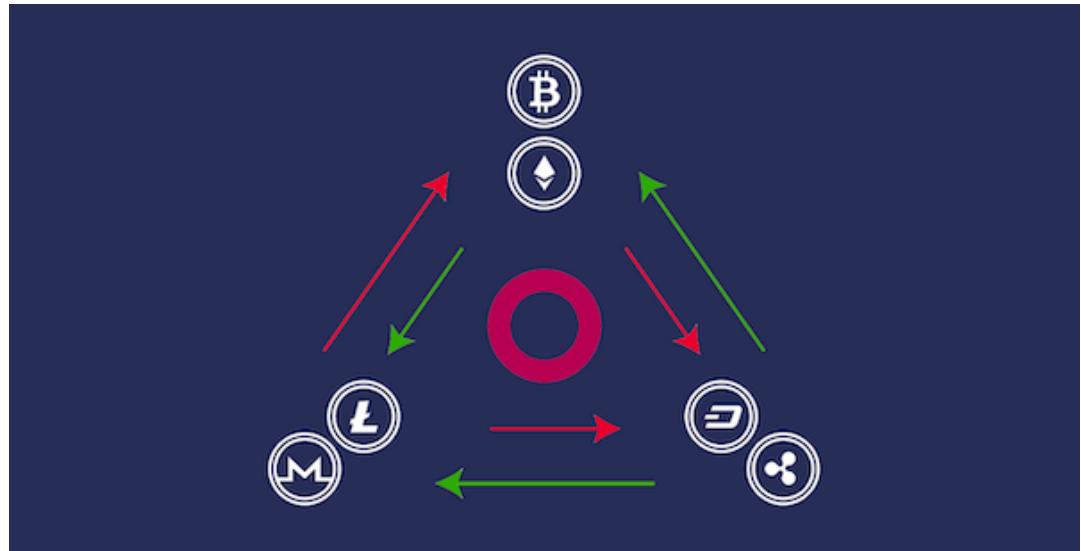


Source: Coinbase

- The rationale for this pricing function is to incentivise arbitrageurs to **replenish liquidity** into the pool when prices deviate between pools or exchanges. They do not address though common problems in market - making like inventory risk or information asymmetry, so we expect evolution in this direction.
- First versions of Uniswap tied liquidity providers to a **common AMM protocol**. However, the latest release of Uniswap, [v3](#), moves closer to order-driven types of markets, allowing for liquidity providers to customise their pricing strategies. This opens up the possibility of a new breed of market-making algorithms, that could potentially import ideas from the market-makers used in order-driven markets

# Algorithmic Trading in DeFI: Arbitrageurs

- Being DeFI a decentralised ecosystem, the role of **arbitrageurs** is critical in many smart contracts. For instance, as mentioned in the previous slide, current versions of AMMs require external actors to keep the prices in their asset pools in line with other pools or exchanges. Trading algorithms can automate the exploitation of these arbitrages, potentially using similar ideas to the ones used in other assets but including new ingredients from the DeFI ecosystem, like for example **flash loans**
  - A flash loan is a non-collateralised loan with zero counterpart risk and zero duration. It has no counterpart risk because in a single transaction a) the loan is created, b) all buying and selling using the loan is completed, c) the loan is paid in full.
  - It is therefore very suitable to exploit arbitrage opportunities.



# Algorithmic Trading in DeFi: Fintechs

- Algorithmic Trading in DeFi is still in its infancy. The tools available to build trading algorithms in a simple way for traditional markets (including crypto) are still lacking due to the complexities of coding in the blockchain
- Some interesting ideas are starting to emerge from the Fintech world, for example [Mercor](#)

