xange.io

Trading. Exponentially changed
The autoregressive liquidity autoscaling ledger

XANGE

No caps on liquidity
High frequency trading, fast transaction throughput
Access to unpaired technical advantage
Flash loan contracts based on autoregressive ML strategies

X A N G E

trading. exponentially changed

A semi-permissioned ledger for decentralised exchanges over synthetic assets, with autoscaling liquidity and autoregressive functionalities

MARIO ROBERT D'AMBROSIO

University of Barcelona

Gran Via de les Corts Catalanes,
585, 08007 Barcelona, Spain

Table of Contents

Executive Summary

> Xange is a platform enabling traders to collateralize cryptocurrency holdings for regulated financial asset classes, such as forex, commodities, ETFs, CFD, complex structured products, and more so financially traded products.
>
> The layman metaphor to simplify the concept: *a social trading autoregressive Wall St*. In essence any trader is and will be capable to copy trade not based on signal copy-trading but effectively cloning AutoML models and deploying them as smart contracts, removing middleman delay, focusing on simple ROI and performance. Xange aims to address 3 main concerns for the common trader: fees, liquidity, and decision making.

The recent development in the DeFi ecosystem has brought many changes and useful applications that would substantially contribute to and diversify the general invested user and crypto-follower.
Whilst, Curve Finance, aggregators, degenerate platforms and similar upcoming never kicked off projects have vastly given a new demonstration of finance applications in the blockchain ecosystem, the door of liquidity providers enabling for reinvestment of capital through crypto aggregators remains still uncertain.
Some platforms have issued what are so-called *Synthetic assets* (Synthetix, Kwenta). Whilst has proven useful, Synthetic assets can be very beneficial in targeting a very similar logic as the CFD products. Derivatives based on price outcome can then be used as future contracts issued on-chain.
Whilst very useful in solving the economic wealth of on-chain/off-chain gateway access, synthetic assets are yet to be implemented in a way to support traditional standards of trading, pragmatic economic standards, and incentivize more liquidity provided.

The objective of the platform Xange is to enable leveraged autoregressive instruments for ANY type of CFD, this is obtained through counterparty synthetic assets bond 1:1 on real markets. Effectively Xange operates as a financial institution allowing for synthetic assets to be issued so long there is enough liquidity in the ecosystem and banks.

When liquidity is not available the strong partnership and treasury may allow for debt inheritance to position the remaining settlements.

A treasury is not used as an aggregator or yielding platform in such a case, thus not falling in the category of Decentralised Autonomous Organization treasury with governance (such as Olympus, Wonderland, Hector, and similar).

It is instead used as collateralized holding for credit lines and borrowers. Leveraged trades will only be possible for premium users with top-performing portfolios, thus mitigating liquidation risks and treasury full-scale near-total liquidation reaction (opposite to Time Wonderland).

Xange aims to then provide new strong technical indicators to solve the lag in social trading or copy trading.

Xange is an essentially highly technically driven metaverse for accessing preexisting trading instruments in a new, more resourceful pathway.

Key takeaways

Problems	**Solutions**
No autoscaling on liquidity	Parachain liquidity broker
Partial access to derivatives, if any	Institutional collateral
No copy-trading, TA, if any	Smart contract copytrading
No CFDs, complex products or CDOs	Synthetic assets over the above
No compounding over trades	Collateralize trade positions on farming protocol partners

Offering

- Xange
    - Built on Solana as fastest enabler for defi products
    - Autoscaling liquidity
    - Provider for cross chain interoperability
    - fast transaction throughput
    - possibility of further layers and possibly expansion to different ecosystems

- Xange Margin Trading Protocol
    - Leveraged trading through istitution
    - Compound yielding through farming protocols
    - Risk mitigation through in-platform portfolio management
    - Upto 1:500 leverage on flash loan

- Xange Synthetic products/derivatives
    - CDOs, CFDs and all complex structured products offerings
    - Enabling tier 3 world countries to effectively trade on Tier 1 markets
    - Bridging real asset classes with crypto asset classes, normalizing volatility
    - Enabling for custom option pricing, such as exponential (pow of 2/3 , or other), potentially enabling users to build their own synthetic assets to trade on so long liquidity is provided

- Xange autoregressors
    - Allow autoML as copytrade strategy to introduce general public to technical indicators
    - No lag due to effectively cloned smart contract, not delayed signal
    - Possibility to fine tune and stack the regressors

Introduction

What if it were possible to instantly trade, feeless, and without much investment in the decision-making process?
What if, moreover, it were possible to benchmark portfolio strategies against the odds of correct decision making and only place a leveraged trade on a quasi successful strategy ?
All this isn't yet available more so for a technical incompatibility of multiple tools and current market demands.
The cryptocurrency markets would allow to access a 'virtual' Wall St., effectively placing bets on NASDAQ assets through perpetual futures based on large pools providing liquidity

The concept behind Xange is to enable users for partially collateralised fractional shares (bitCFDs) of CFDs on a financial metaverse.

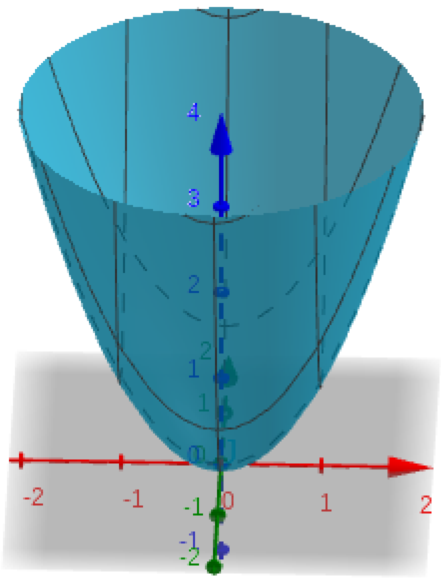


Fig Showing Gradient Descent optimisation technique in 3D plane

Objective of such offering is to achieve a higher target of economic opportunities gven the cryptoeconomic driven market accounts for over 1T USD and is supposed to quadruple by 2030, accounting for 4T and may possibly reach the highest correlating market to EUR_USD.
Moreso Xange would allow the traditional investor to benefit from new technical indicators running on the liquidity pool as smart contracts.
These would ne known as Autoregressors.
Autoregressors would be taking advantage of the nature of evolving smart contracts.

Autoregressors used would be trained on multiple neural networks. Future work would envision adapting them to user product requests. Moreover autoregressors running on the smart contracts wouldn't be impaired technically as they would just hold the 'live' model state and hyperparametrisation post

processing: no actual training would be performed on such contracts, but would all be ditributed calculus on permissioned servers.
Moreover optimisation strategies such as Gradient descent and extreme gradient Boosting, in which sattle points, or local minima and maxima are tried to avoid as bias

Problem

Hypothesis

1. The DeFi space has fallen into a rabbit hole of finding implementations not fitting real world problems
2. Users need products not new technology in the defi space, epecially new products which can fit the demand
3. The DeFi space lacks simple trading tools that can enable real world asset class trading
4. Sometimes if this is possible, liquidity is not enough to fulfill strategy or price action
5. Trading on the metaverse is somewhat still esoteric or rather exotic rendering all interest from the average user vane towards such objective
6. General public only trades crypto to crypto, not crypto to asset classes as they do not find a simple direct channel as gateway
7. There is no way to benchmark own performance towards other user in the space
8. Technical analysis is somewhat primitive and could be already replaced or *aided* by AI regressors or techniques.

Discussion

1. Whilst in nature highly speculative as assertation, it must be noticed how the latest advancements in the DeFi space have proven as very speculative, risky or without any rationale.
   We can think of the sunset of DAOs, which have appeared in great success, and vastly disappeared near after. This is because the product fit for those particular was not just less obvious, but also without clear sense.
   Moreover in the competitor matrix it appears to be clear how a lot of focus driven in the space has been given to decentralised applicatives, less attention to specific use cases enabling speculative investors, financial opportunities and technical trading.
   It is in very early beta stage as CAGR is at 12%, expected to quadruple by 2030, effectively bringing blockchain applicatives at 5Bn market cap. Most platforms offer unclear advantages over regular consolidated platforms.
   Moreover risks due to liquidation are also driving markets into less user acquisition as it is yet unclear which platforms can offer fraud guarantee protection (Snow Bank, RagnarakDAO) or cyber attack protection (Grim Finance)

2. The demand has been continuously developing towards more stable, market ready solutions.
   Better usability on the day to day is needed and requested. Moreover better suited products, structured offerings and public initiatives are needed as enabler.
   The DeFi space has had a great growth, but focused primarily on market ready technical solutions and new blockchain opportunities capitalisation, driving less attention and care to actual user interests.
   Moreover technical driven opportunities have ended to obtain very diverse results in terms of technology but very similar in terms of market supply

3. We can see this when compared to other platforms such as Kwenta, Synthetix, Frax, Curve, DeFi farming yield aggregators. Farming protocols in general do not provide trading features, and when they do it can be very limiting, as only offering treasury value locked in assets. For instance *Yearn.finance* only has liquidity bonding tools, does not enable autocompounding and trading.
   When synthetic assets are indeed offered as tradeable assets on platform the issue becomes on more complex structured products, availability of top indexes and, ovarall, scarse diversification of assets.
   On a sidenote on either of these cases simple technical indicator strategies available on the financial regulated asset market are not readily yet available.
   It would be the natural flow as the capitalization of such new asset class began to be readily available for such cryptocurrency market as well.

4. Liquidity is another big impacter of such AMM (automatic market makers), as most platforms in this space rely on sole channel of user provided liquidity as farming aggregator for bonded contracts.
   In this case the liquidity providers will be initially driven from the Solana ecosystem, in exchange for a small fee percentage. On a later stage the ecosystem will blend easily with all chain system so long the fees are relatively small and acceptable liquidity is provided.
   Liquidity will then be provided under the form of user debt and ecosystem exchanged contracts or debt obligations

5. Whilst it is true that trading is esoteric in the metaverse. This happens for a few of the aforementioned reasons and others as well. First and foremost the crypto space has opened up to trading opportunities and dexes (decentralized exchanges) only recently. The dexes and defi applications are relatively in a newborn stage, and as such, feedback user validation is still very important. Xange poses itself as a traditional brokerage platform with the advantages of the dexes powerful inheritance: high liquidity, lower fees, no caps on transactions and finally, a very fast ecosystem.

6. The idea behind a decentralized exchange is to allow users to opt for cryptocurrency in exchange trades. Essentially liquidity yield farming and autocompounding strategies are the core built from decentralised exchanges. Moreover the few available dexes (Frax, Perpetual, Kwenta) offer very limited support towards financially traditional asset classes, effectively making liquidity schemes possible only in the metaverse, not speculating over the remaining 100T (93 for only CFD in US) total market capitalizations. The idea behind a new platform as *gateway* between financial asset classes and crypto asset classes is needed to ensure capital flow in open economies and geographically agnostic market speculative, in order to drive new interest opportunities and possible investors

7. If we shall examine the current market offering in terms of decentralized exchanges, the current focus of topic discussion would resonate at a certain stage with the question around if supply is fitting demands. On this note it would be superficial to answer the question without understanding the context of such question. The markets have been behaving in a technically

driven predictable pattern, almost as if we shall say *rigged pattern*. This is vastly due to the reason markets have been copycloning itself, effectively almost progressing in a negative feedback loop. The crypto defi space has been evolving onto itself, with autocompounding strategies, more dexes, more liquidity pools, but no change to the access of flow for open economies and financial traded assets. This is why Xange would operate as brokerage firm interoperability chain for gateway synthetic trading. Moreover speculative strategies and technical contracts such as those described in point 8 will be essential to drive the platfor growth even further.

8. Finally the only technical analysis possible on such dexes or platforms is dictated by API usage from common indicators and readily available trading instruments.
   In the survey conducted, it is clear how new AI-based regressors could provide much more interesting insights than pure technical base indicators.
   Moreover, in assistance to the TA provided through autoregressing smart contracts, the possibillity of having flash loan contracts using leverage in conjunction with autocompounding yielding techniques, which are already currently available through yielding farming aggregators

Xange Metaverse

The Xange Metaverse is a diversification of a Solana based ecosystem.
Liquidity pools and providers are based on Orca, Drift, Saber and then could possibly include other parachain solutions to have higher access to liquidity, similar to the Synthetix protocol, it may rely on Curve and Kwenta for additional access to liquidity.
Xange will be then not just a platform but an Automated Market Maker, as a brokerage service, allowing for autoscaling liquidity cross-chain compatibility and arbitrage opportunities through flash loan contracts based on simple regression rules.

The risks the platform would be taking would be proportional to 2 things: **trade positions timings** and **collateralised leverage** provided from users.

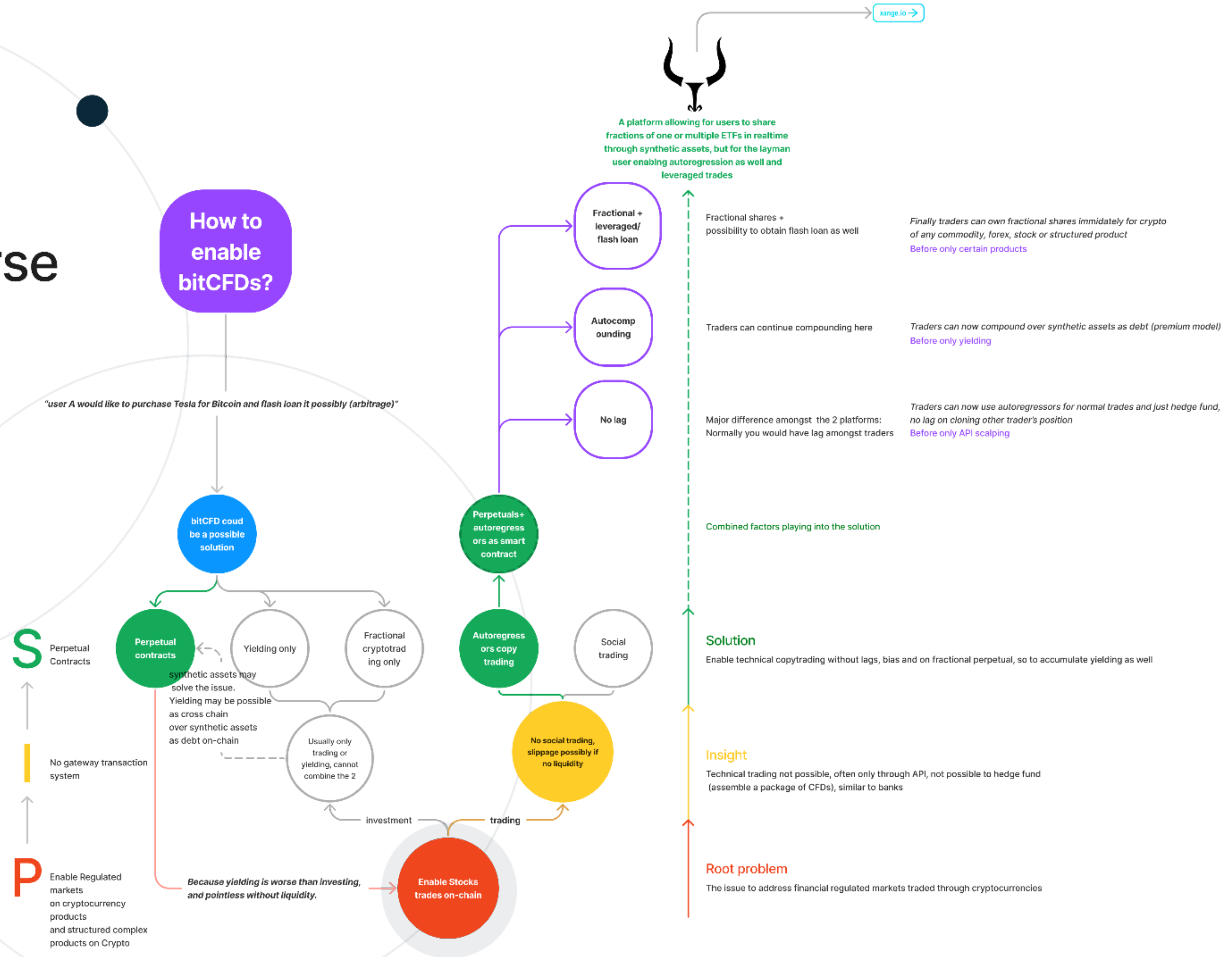
Trade Position Timings
Trade position timings will result essential to establish the leverage permissions allowed for the users to undertake and possibly to analyze and risk distribution.
Moreover trade position timing will define risk profile for the following category.
The only specific channel to use flash loan contracts will be through specific trading within short window span and through specific channel of trading (mostly HFT).
The positive thing of enabling such channel of trading on the platform is to allow for high leverage only when trades arbitrage opportunities can actually yield profits

Collateralised leverage
Collateralised leverage will be proportional to multiple factors, including: risk assessment profile (KYC), ROI history, capital involved in the settlement
Collateralised leverage is used for a trader to enrich the possibility of using even more leverage than he would envision, through a flash loan contract for instance. In such case collateralised leverage would account for a debt contract with the Xange platform.
When liquidity cannot be immediately provided Xange will factor in debt allocations from the ecosystem automatically

Xange is born.
Trading.
Exponentially
changed

Xange Metaverse Genesis

Problem → Insight → Solution



Business Formulation

After thoroughly dissecting the hypothesis it is evident how there may be multiple problems.
The problems seen as stoppers for current implementations and market offerings are to be seen with not fulfilling CFD stock offering (scarce trading offerings), not enough liquidity to being able to provide certain trading options (higher margin trades on HFT with big capital amount), and finally, no option for copy trading to allow for easiness of use and adoption.

Such problems are partly addressed, moreover, through cumbersome platforms using the most adopted but subpar solution of EVM apps (Ethereum Virtual Machine), through platform with limited trading capacity and mostly liquidity aggregators, such as dYdX, Perp (for perpetuals), Skew, Kwenta and Synthetix (the most integrated EVM applicative on the market).
Whilst proven useful for certain contexts, the interoperability amongst large financial ecosystems and the metaverse is still not possible, and as it is also not possible to foresee copytrading on autoregressing algorithms or solutions based on AI in any way.

The solution would be to allow users to trade without boundries, allow users to swap cryptocurrency for complex structured products and viceversa without perceiving timezones, copytrading lagging or similar side effects. Moreover allow to autocompound on top of smart contracts.

Xange would pose itself as a platform enabling liquidity providers as ecosystem to build enough liquidity and slippage fault tolerance to allow for synthetic assets to be issued.
Assets not enabled on the ecosystem can then be purchased over real market with collateral treasury as deposity and guarantee from a company perspective.
Moreover key aspect is the possibility to run smart contracts with autoregressing strategies that would enable the trader to run auto strategies based only on the smart contracts, without any delay or intermediaries

Early adopters would be experienced traders, whilst custom segments would be young cryptotraders and on a later stage (app phase, product map) non so crypto friendly traders.

2. Problems

- Cannot trade on major stocks from crypto directly
- Not enough liquidity to exit on leveraged trades
- Cannot copy trade strategies from cryptocurrency directly into settlements
- No autoregressor technical instrument/signal

Existing Alternatives

- DeFi Solutions (Aave, compound, Curve Finance)
- dYdX (Perpetuals)
- Perp.fi
- Skew
- Kwenta
- Synthetix

4. Solutions

- Synthetic assets through CFD Futures
- Liquidity provider aggregators
- Smart contracts running as copytraders
- ML algorithms running on the smart contracts

8. Key metrics

- Derivative trading, low slippage
- Compound aggregators, yielding on top of derivative contract
- No delay on copytrading, no lagging disadvantage
- Possibility of combining different strategies and preserving high liquidity through stablecoins
- Unlimited timezones unlike Wall Street

3. Unique Value Proposition

Allow users to trade without boundries, allow users to swap cryptocurrency for complex structured products and viceversa without perceiving timezones, copytrading lagging or similar side effects.
Moreover allow to autocompound on top of smart contract

High-level Concept

A platform enabling liquidity providers as ecosystem to build enough liquidity and slippage fault tolerance to allow for synthetic assets to be issued.
Assets not enabled on the ecosystem can then be purchased over real market with collateral treasury as deposity and guarantee from a company perspective.
Moreover key aspect is the possibility to run smart contracts with autoregressing strategies that would enable the trader to run auto strategies based only on the smart contracts, without any delay or intermediaries

9. Unfair Advantage

- No lag on social trading aspect
- No timezones
- Very high leverage due to high liquidity and collateral deposits
- Enable flash loan contracts for HFT
- Complex products can then be structured as smart contracts (indexes)
- A metaverse to allow common traders to passively compound and yied rewards based on technical indicators and real economies essentially

5. Channels

- Online channel mainy
- Offline can be as prediction only or portfolio management, operativity is kept online as the rewards are yielded through the contracts

1. Customer Segments

- Young traders, through the App will be enabled to join the 'Metaverse' explore trading opportunities and diversify their crypto baggage through the synthetic assets
- Ideally a non-crypto friendly trader who wants to join through fiat will be bound onto the aggregators using stablecoins

Early Adopters

- Small crypto friendly traders
- Very technical driven traders
- Vendor locked in ecosystem compounding users (defi space of the metaverse users, loyal users)

7. Cost structure

- OTP: One Time Payment: Marketing, deposits to API financial class providers (eToro API guarantee deposit/collateral)
- Servers for High transaction throughput (ClickHouse/InfluxDB + API infrastructure in C++)
- Load balancing VPS, scalability for upto 200B req/s, cost is around 6k$ max per month, ARC is to be estimated at 80k$ max
- Smart contract servers nodes with AutoML

6. Revenue streams

- Tier based transaction fee approach:
- Tier 0 ~10k/year: 0.05%/to
- Tier 1 10k-100k: 0.0025%/to
- Tier 2 100-250k: 0.00125%/to
- Tier 3 250k-0.5M: 0.000625%/to
- Tier 4 0.5M+/year: 0%

Subscription
- Basic: Limited to 3 Autoregressors, 50/mo + 10% profit
- Advanced: 7 Autoregressors 80/mo + 7% profit
- Premium: 20 autoregressors 120/mo + 5% profit
- Partner: Custom AR/ custom pricing

Leveraged hedging based on ecosystem trade positions. Essentially future hedging against current market based on current smart contracts, since markets will be algorithmically driven, company will be able to assess price actions and buy CFDs against the market

Business Canvas

SWOT Analysis

Strengths

First Perpetual to enable decentralised autoregressing algorithms
No lag due to effective smart contract cloning, without copy trading delays in betweenUnlimited liquidity, enabling ecosystem to merge into macroeconomic trends and bridging gap amongst defi and regulated markets, providing more liquidity into bothFirst platform to enable App usage and present itself as user friendly

Weaknesses

No option to build own ecosystem.
As liquidity must be big, it is always a cross chain solution, never will be its own ecosystem or blockchainUsers' trading activity may bring an inherent predictability and rigged concept to the markets, relying on non-randomized machine sampling.
Slight devation must be implemented to correct for artificial sampling in market making

Opportunities

Leveraged trades into regulated markets, less risk as volatility is not as high as in decentralised financial classes.
Possibility to scale as liquidity provding liquidity increase.
Cross-chain networks and compounding platforms as aggregators partnerships on top of core social trading smart contracts
Community growth, real users, not just wallers

Threats

Possibility of pre liquidation and market collapse through leveraged trades, and possibly company bankruptcy in such case.Leveraged trades must be only enabled on risk assessment case by case KYC scenario to avoid such possbility.Possibility of flash loan attacks on HFT option with leveraged trades to spike the markets and scalp the deltas produced

SWOT

ANALYSIS

STRENGTHS **WEAKNESSES**

OPPORTUNITIES **THREATS**

First Perpetual to enable decentralised autoregressing algorithms
No lag due to effective smart contract cloning, without copy trading delays in between

Unlimited liquidity, enabling ecosystem to merge into macroeconomic trends and bridging gap amongst defi and regulated markets, providing more liquidity into both

First platform to enable App usage and present itself as user friendly

No option to build own ecosystem.
As liquidity must be big, it is always a cross chain solution, never will be its own ecosystem or blockchain

Users' trading activity may bring an inherent predictability and *rigged* concept to the markets, relying on non-randomized machine sampling.
Slight devation must be implemented to correct for artificial sampling in market making

Leveraged trades into regulated markets, less risk as volatility is not as high as in decentralised financial classes.

Possibility to scale as liquidity provding liquidity increase.

Cross-chain networks and compounding platforms as aggregators partnerships on top of core social trading smart contracts

Community growth, real users, not just *wallers*

Possibility of pre liquidation and market collapse through leveraged trades, and possibly company bankruptcy in such case. Leveraged trades must be only enabled on risk assessment case by case KYC scenario to avoid such possbility.

Possibility of flash loan attacks on HFT option with leveraged trades to *spike* the markets and scalp the deltas produced

Competitor Analysis

The following is a competitor analysis based on the 3 main perpetual dexes: dYdX, Perp, Frax and Drift (Solana)

Platform	**Many Perpetuals**	**High Liquidity**	**Custom TA**
dYdX	Very limited	Yes	Some
Perp	Limited	No	No
Frax	Not many at all	No	No
Drift	Some	Yes	Possibly in future
Xange	Opportunity to create custom CFDs	Yes	Main focus

Xange would be a unqiue platform offering high autoscaling liquidity, focusing on technical autoregressors, and unlike others providing some of these functionalities, it would have the benefit of autocompounding as well in the future, so effectively yield bonding.

The protocol

Trading

Training of the Autoregressors performed is carried over on multiple window spans to fit 3 different trading timeframes: HFT (high frequency trading), swing trading (daily), and long positions (over day trading).

Trading through the platform would be performed over fractional synthetic assets (bitCFDs).
Objective would be to enable traders to place a mix of multiple positions, have more control and less manual intervention by allowing smart contracts to drive and make the trade positioning.

Another key factor for the trading platform would be to enable the social component of monitoring other people's trades and having a benchmark ROI performance chart to rely upon for cloning autoregressive strategies as well.

Finally, through the social acquisition, users which would have less experience may largely benefit from it and create performing solutions enabling a larger pool for user talent acquisition

Liquidation

Liquidation plays 2 key roles in the Xange Protocol. On one side Xange acts as corporate financial crediting institution, gathering collateral to avoid liquidation issues.
There could be 2 types of liquidations: let's identify them respectively as scenario A (provider, Xange) and scenario B (client, trader).

Scenario A.
In this scenario effectively the liquidation occurs as Xange has debt it has lost positioning from and the traders cannot pay as the liquidity pool is already in debt or dry.
In such case the financial institution's credit line is suppressed, a temporary contract is put on action with immediate holding of all company assets to repay the loan as borrower.
If the loan is not to be repaid within the agreed timespan with the crediting institutions (eToro/Fidelcrest Ltd as the beginning of partnership), bankruptcy shall be filed within 12 months and the protocol is effectively in near death situation.
This would be very similar to what happened to Time Wonderland, but with a few key considerations.

For scenario A to happen fully, many other things would have already been happening: most of the partners' debt shall be in a non crediting position, near liquidation as well, most traders on 1:500 leverage shall be in a difficult position, and even so, the platform would have suspended
leveraged trades, so effectively would have been leveraged for long time ahead liquidation from the partners as of an odd time factors.

This scenario, whilst possible, would be very rare, and could be easily be avoided by voiding the leverage crediting policy under special conditions (such as near 0 liquidity in pool).

In such case this scenario could be avoided

Scenario B.
In such scenario the user, or trader, would be liquidated due to its position not being profitable and in debt for overdue period (voiding the contract therefore) or if the liquidity of the pool were to be at risk in any near given time (as show in scenario A).

To outline how the Wonderland Time, or better yet, DAO domino effect could be here avoided as leverage would not be given to any user, it would be a performance-based unlocked feature, thus not relying on any speculative price action of a treasury, rather complex derivative trader's ability of decision making.
Thus the main difference between main treasury DAOs and Xange relies in the intrinsic non-dependant chain reaction in lqiuidating trades, as this will constitute around 5% of the company portfolio.
Other consideration for the average trader to obtain easier access to higher leverage is to benefit from flash loan contracts for HFT purposes.

Derivatives

Xange offers different products under the issuance of custom fractional fictuous shares with 1:1 bonding to real world asset classes.
Since these products will be released directly through platform and do not directly require financial regulation, the release and creation is instant, allowing for 3 types of products: CFDs, CDOs and custom created.
Thus granting the trader with custom products, own new interesting products as if the trader had an ISDA agreement with the financial insitution, whilst effectively enabling the brokerage platform to take responsibility for the structured products offering

CFDs.
Contracts for difference offered here will be the most beneficial structured product, derivative offered.
As enablers for price gambling and manipulation, they will allow for most bitCFDs, or fractional Synthetic assets to be released.

CFDs can be then combined as a whole custom derivative on a later stage to introduce new trading indexes and opportunities.

The institutional structure operating under Xange's umbrella corp will serve as brokerage platform collateralizing on real world stock market all trader's positions through API on fast exchanges such as eToro or similar, effectively confirming through the platforms only once a transaction is successful.
The way this could be seen is as creating debt over collateral trader's have been locking on the platform.
Ths could be liquidated at any time through the metaverse to use to pay buck institutions.

The other viable option would to ignore collateralizing trader's positions and use only the treasury to pay back for the wins or losses. Whilst interesting, liquidity could run out depending on the ecosystem's overall trades, putting at risk Xange of ending up like a Wonderland Time platform

CDOs.

Collateralized debt obligations could be seen as loans or liquidity farming structured aggregated products.

This could come under form of liquidity providing pools, collateral assets placed from traders' in the ecosystem (Synthetix, Kwento, Curve, 1Inch, Frax).

Essentially collateralized debt obligations allow for partially collateralized and in later stage fully collateralized loans in form of bonds over synthetic asset treasury lock.
This concept would be very similar to farming, the difference being no intrinsic pegging would be needed as the collateral would be supplied from the fiat bonds, not from a TVL (total treasury value locked)

Custom derivatives/products.

Custom derivatives will be the final stage of implementation, constituting effectively a baggage of *glorious* market speculative and appetible products, which may suit clients and the intrinsic price actions as well.

It is expected that the social element of the platform to be able to discover custom products will impact exponentially the adoption and *trends* amongst the platform itself.

Experienced traders may create then custom portfolio allocation with dynamic weights according to stock indexes for portfolio-managed products and use any non-polynomial function as well. This would effectively result in an unprecedented way of shorting or longing derivatives positions.

Clearly limits will be imposed on stock values, as certain high or very fractional limits will be banned, to avoid flash loan attacks or high liquidity voiding.

Example Custom product.
Suppose Alice is an expert trader, Alice loves Xange, has been using ARIMAX on AAPL stock, follows TSLA as well and DOGE, and notices how when BTC plummets in value DOGE and TSLA as well, but AAPL acquires value instead.

Alice notices how ARIMAX is very stock specific but doesn't very well interpret all the diversification and daily density numbers correctly.

So Alice decides to create her own custom product, say the *Alice* index.
Alice allocates 50% of her portfolio on a delta-neutral strategy, when the price range has been increasing for the past 1d on TSLA+DOGE+BTC (respectively 33.33% of 50% each), and the other 50% on AAPL.
When the price of any of the indicators on TSLA or BTC go down Alice would like 75% AAPL and the remaining 25% to go to TSLA+DOGE+BTC (33.33% of 25% each).

Assume a margin of 90%

Alice has just created the following function:

$$Alice(x,y,w,z) = \begin{cases} \dfrac{x+y+w}{6} + \dfrac{2z}{6} & x_{-1} + y_{-1} + w_{-1} > (x+y+w)\cdot 0.9 \\ \dfrac{x+y+w}{9} + \dfrac{6z}{9} & x_{-1} + y_{-1} + w_{-1} \leq (x+y+w)\cdot 0.9 \end{cases}$$

$$Where\ x = TSLA,\ y = DOGE,\ w = BTC,\ z = AAPL$$

$$x_{-1} = TSLA_{ytd},\ y_{-1} = DOGE_{ytd},\ w_{-1} = BTC_{ytd}$$

And there we have the Alice index, created instantly on chain without issues as a synthetic asset

Alice could then also use a custom autoregressor on this custom product and continue to yield returns on it running it as a smart contract effectively

Risks and mitigation

The major risk factor would be given by held **collaterals**, **liquidation** and market **artificial sampling**

Collaterals
One of the concerns would revolve around the idea of collaterals and deposits as such.
Collaterals as we could foresee would need to be deposited as a guarantee for debt inheritance for liquidity from the ecosystems. This would be a 2 step process as users would deposit collaterals in terms of exchange means and the platform would accept collateralization around debt inheritance to provide liquidity for the trades and capability to execute certain complex products, especially the complex custom CFDs on a later stage.

This particular step would necessitate Xange to operate as a financial broker closing positions on the real world market to settle the payments on the metaverse.
If a trader shall use leverage, thus collateralization may come into place, if such were to happen, only certain traders shall be given big leverage and for long positions, effectively mitigating the risks of liquidation and big collateral holdings.

Liquidation
By liquidation we here define the situation in which a trader's positions will be effectively put on hold on the platform and possibly voided.
Another scenario is Xange's current positions being revoked and voided on the real financial brokerage providers.
Whilst both are possible the first is to be considered the most common, as if the trader currently has created a custom CFD, placed a leveraged position and maintained it for long enough the contract shall be deemed as voided.
Xange current assets will be reverted back as sold, so effectively no liquidation shall occur on that side.
Moreover, whilst a decentralized exchange is good, a governance smart contract will have the role of policing the trades.
So effectively making Xange as semi-permissioned trading exchanged, where as trader's positions can be flagged at any time from the policy monitoring contract, and shall the risk be higher than the collateral

in any point in time (as initial collateral may be very low), the policing smart contract shall trigger a flagging system, where the positions may be liquidated.

Finally there is a possibility of witnessing Xange's positions being liquidated, this would only happen if trader's underlaying assets would account for over 50% of capitalization in being liquidated.
Considering leveraged trading will be ROI based and can be revoked at any given point in time, it is to be expected to have only 20-25% traders enabling leveraged trades, of which probably only 50% on very high risk profiles. Considering the markets can behave as random walks, the odds of success can be then simplified very aggrissively to a mere ~50% of 50% of 25% (largest risk) so effectively bringing risk profiles down to 6.25% of trader's portfolio.
The question remains open to how mcuh market capitalization and liquidation can be held in the accounts of 6.25% traders.
The policing smart contract shall ensure a no more than 3% market capitalization (so effectively less than ~40% on uniform distribution in wealth assets) to mitigate the risk of Xange liquidation and market damage.
It shall also be said that Xange may void the contracts running at any time if fraudulent or suspicious/high risk activity is to be found

Artificial sampling
By artificial sampling we here define the process during which samples of market behave in a relative predictable pattern as the trades bring more capitalization and volume to the metaverse.
This would be especially accentuated by the autoregressors smart contract running as cloned and making effectively similar trades.
On one side this could be beneficial as it would increase leveraged opportunities for the brokerage firm to speculate on and effectively purchase certain price derivative contracts based on a statistical analysis of capitalisation per smart contract type and autoregressors.

On a small percentage this would be beneficial, on a larger percentage (such as over 10%) this could damage the market spreads and potentially remove any opportunities from the markets, as they would behave in an efficient matter and be less opportunistic.
A resolution to such problem would be to let the platform allow for small randomisation of hyperparametrision on the models, and copy trade based on advertised 'best match' performance suggestions, which could rotate to allow for more strategy diversification in the markets.

On a later stage it is to be seen to have a similar percentage gain to forex EUR_USD, as this would resemble the largest manipulated market capitalization to current date of working for current buyers and sellers.

Implementation

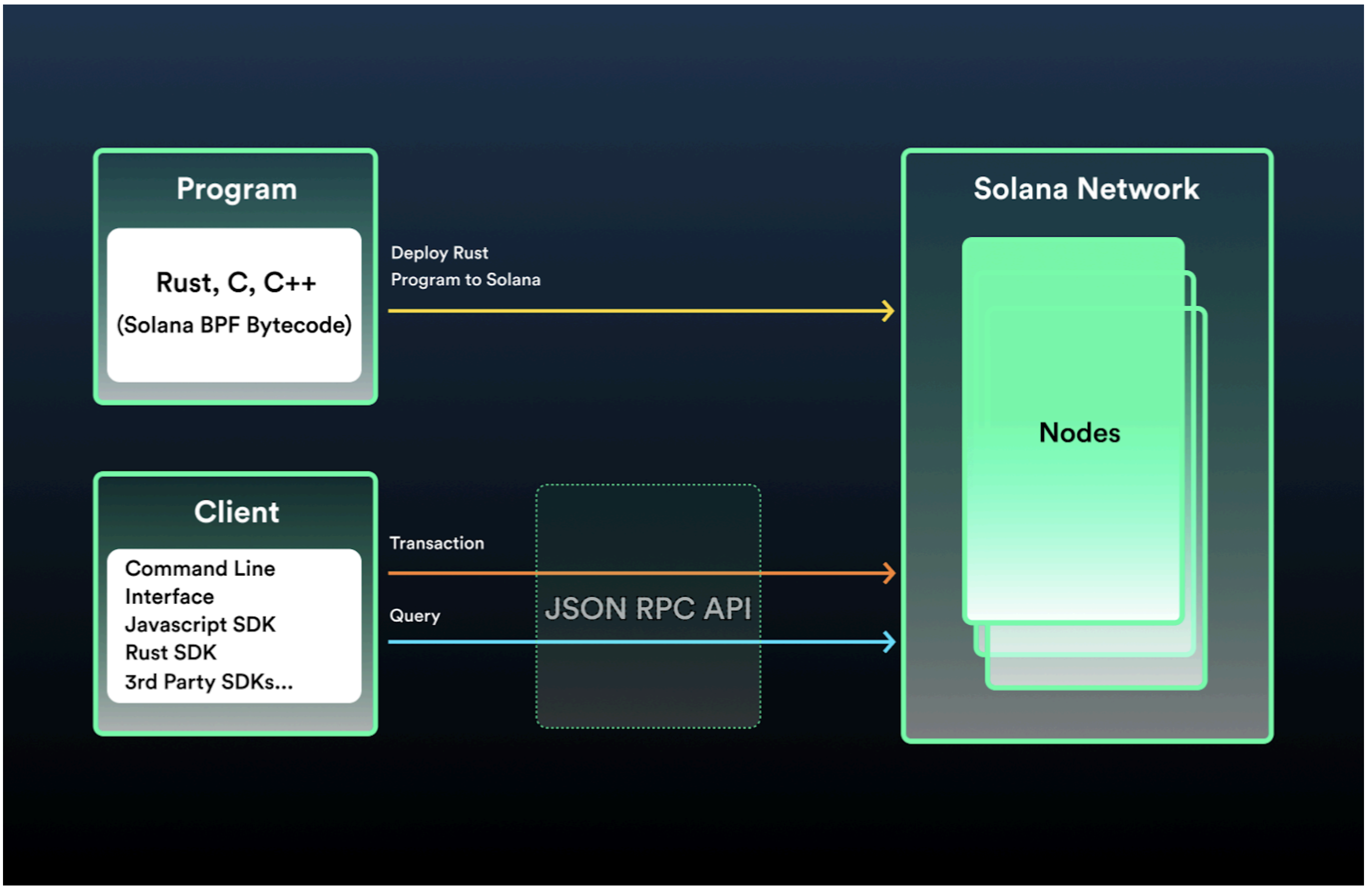
Technical implementation

Technical implementation will be carried over the Solana protocol.

Solana has been chosen as the defacto platform as enabler for most decentralized applications over the ecosystem for financial usage and specific decentralised applicaitons.

Technical implementation over Rust, C++ and python is possible for Solana smart contracts, but the ideal choice of platform is to be seen around the concept built for developers of Solana.

Solana has been selected as the platform of choice given its vast performance in terms of transactions per seconds, its wide adoption for other similar liquidity providers (Orca, Drift), and, overall, as a choice to apply for selection into the Solana Venture program.



The ecosystem - Xange Ledger

Although Solana shall be used as financial decision initially it is to be seen as the bestl

Consensus Mechanism

Consensus mechanism used for Solana, the platform of choice is proof of stake and proof of history. A custom ledger approach may require an only proof of stake with a combined possibly in the near future with a sustainability goal of proof of air, rather similar to CERN's developers suggestion for Algorand, a unique sustainability green energy provider issuance. Allthough it would be for proof of stake, unlike Algorand, it may add in compound rewards for stakers' nodes and sustainable growth.

Miners

Miners would be paid through staking rewards, not proof of work in such case. Transaction fees would be very much diminished because of this, whilst preserving the benefit of stakers involved

Smart Contracts

Smart contracts would be the core feature of the ledger, enabling autoregressors. They would be used as they would effectively enable most of the automatic market making strategies, and potentially the policing contract as well. In the future it could be possible to see custom bitCDFs given as smart contract, as they would be nothing more than mere functions running on the ledger

Solana enables smart contracts as real programs, enabling for any piece of code to run regardless of blockchain stae. The smart contracts' released on the ledger would need to be stateless, as they would interrogate the API to effectively exchange information on the hyperparametrisation and current price contract. To avoid DDoS attacks, flash loan attacks and reentrancy attacks, all smart contracts must not allow any kind of cross code injections. The fact that the smart contracts are stateless is an insurance layer of protection as audited in the recent developments over the Wormhole attack.

Features

All though Solana does not quite directly support UTXO as transaction verification system (contrary to bitcoin SV), simple payment verification is supported on chain through the SPV program. Essentially SPV Programs run as contracts deployed on the Solana network and maintain a type of public marketplace for SPV proofs that allows any party to submit both requests for proofs as well as proofs themselves for verification in response to requests.

This enables a simple and fast verification system for payments

Scaling factors and size

Scaling factor is to be expected at maximum 65000 transactions per second with Solana, for an average 400ms per block mining, thus 3 blocks per second.

Block size is at 10MB per block

This would allow for great scalability.
Probably in the near future a revision fork shall be needed to address new transaction requirements and node speed/cache clearance in stakers' servers

Public key infrastructure

Infrastructure verification would be carried over through public key BIP39-compliant seed phrases (as Solana has used).
The client wallet would need to have a seed phrase to enable public key verificatio. When comparing standards of wallet identification and seed verification, BIP39 is seen as the defacto standard, as much more performant and reliables than BIP32 or BIP44, a bit slower and less secure overall

Networking

Networking would be enabled through RPC clustered nodes. This would allow for more control, no remote code execution, provided correct security measures are taken

Roadmap

Roadmap is focused mainly on bimonthly delivery for the next 4 years.
It is divided in 3 important areas: Growth Hacking, Financial team and Developing Team
The aim of Xange is to develop a platform enabling the common trader to place market trades through autoregressive strategies over the common cryptocurrency asset classes

Development/Production Team
The developing team will be focused initially on the building of smart contract
Team will consist of 2PTE + 2FTE initially, building upto 6.5FTE.
The team will have diversified skillset over the following:

FTE
Prior to 15 months:

Senior Software Engineer (Smart Contracts), shareholder/cofounder - core team
Mid Fullstack Engineer (preferred blockchain experience)

After 15 months (in line with CFD/CDO release):
Quant developer
Blockchain engineer (smart contracts + Core)
ML Software engineer (quant developer)
Fullstack software engineer

PTE (prior 15 months)
ML Software engineer
Fullstack Engineer

Objective of the development team as shown in the figure below is to enable CDO and CFD as synthetic assets onto the platform using Solana as smart contract platform enabler.

In parallel the ML team will have already researched possible ML strategies to use and validate based on prior data for the initial markets (mainly commodities and forex).

By the time the FTE core team will have enabled the CFD and CDO development the ML team will then proceed to onchain the release of autoregressing techniques as indicators on chart and portfolio management as well.

On a final stage the fullstack developers wil engage in building the core mobile App, as of targeting market cap of 10% of crypto capitalisation (10Bn), this will enable the average crypto enthusiast to use Xange products and services as well. It is at 5-10% market capitalisaiton if Xange shall reach such target, it will become less of just a tehcnical service niche product, but more of a general platform satisfying a metaverse of cryptotraders and institutional clients.

At this stage it shall be imperative to have the team work on a protocol in beta to manage high transactions per block throughput (TPB) and possibly expand to fiat services as creditor institution (ebanking services as gateway).

Financial team
The financial team will be needed for all the related operations for liquidity modeling, price regression modeling, but moreso as business development.

The following team will be used specifically for such an operation starting at month 6 of development:
A team consisting of 2PTE and 2FTE will be used of which:

FTE Core team:
1 Business Developer (prior experience in trading and risk assessment companies)

1 Commercial Developer (prior experience in blockchain acquisitions and VC/Grants)

PTE team:
1 Data Analyst (specializing in FTS forecasting)
1 Financial Anlyst (specializing in CFDs, structured derivatives and Risk assessment)

Financial team will have to focus on CDO development in line with development team and execute risk assessment strategies, most importantly carry over partner relationships to enable for more liquidity on platform (aligned with user/client liquidity providing as well).

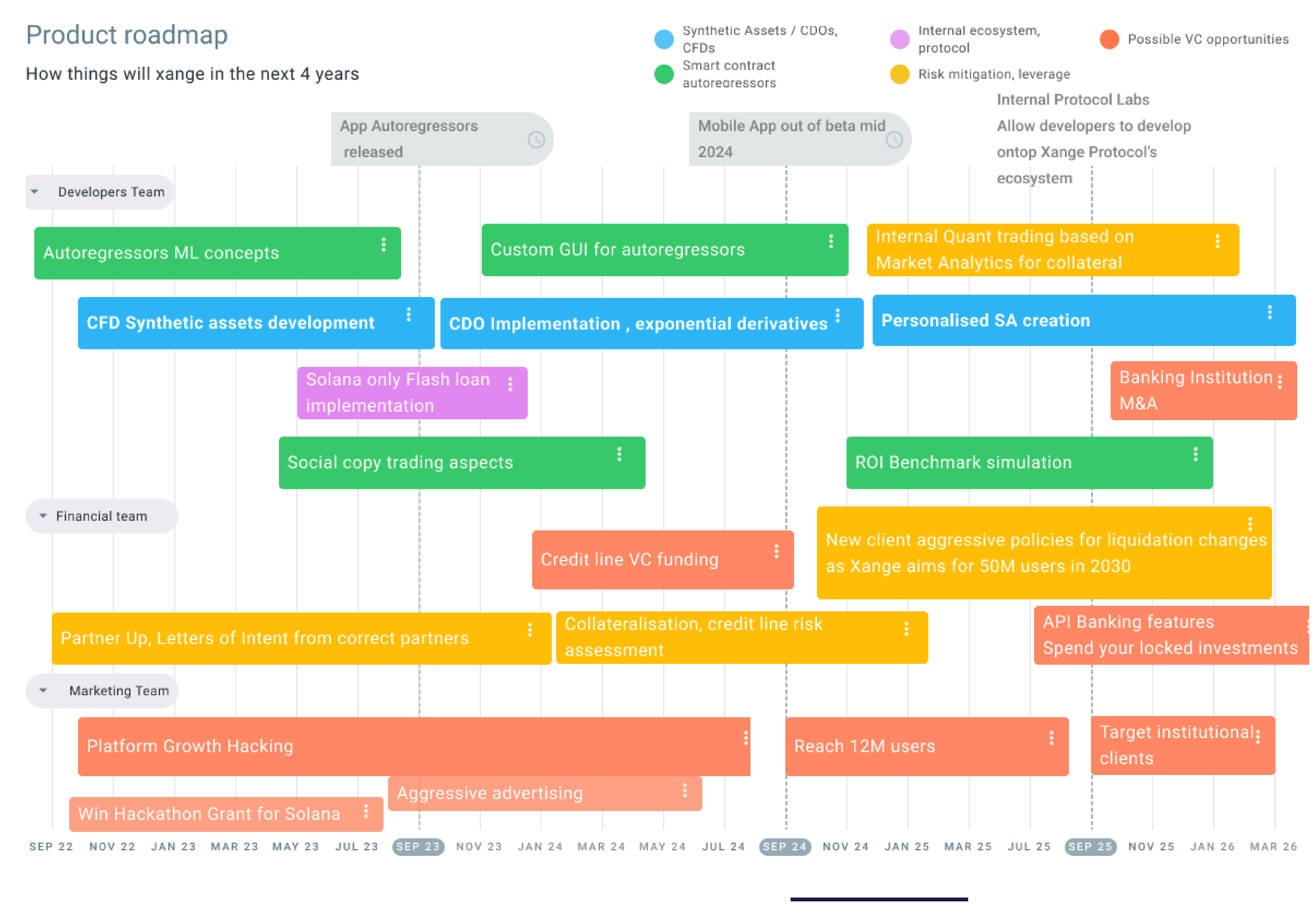
Growth Hacking team

A growth team will have the objective of scale the user targets based on the various stages of Xange: a niche product made for advanced traders, a trading paltform for everyone and finally a crypto-service financial gateway/broker for a wide audience (hedge funding/savings accounts/etc…)

Initially the growth team will be constituted of the following PTE:

PTE:
1 Marketing specialist (run targeted campaigns cross channel, ie native, display, programmatic lead and retargeting options)

The growth team specialisation will be focused mainly on executing retargeted campaigns, acquiring more users and validating beta stages of the Xange products: Trading Platform, Autoregressor smart contracts, Custom CFDs, App, Credit institution/services.



Product roadmap
How things will xange in the next 4 years
Synthetic Assets / CDOs, CFDs
Smart contract autoregressors
Internal ecosystem, protocol
Risk mitigation, leverage
Possible VC opportunities
App Autoregressors released
Mobile App out of beta mid 2024
Internal Protocol Labs
Allow developers to develop ontop Xange Protocol's ecosystem
Developers Team
Autoregressors ML concepts
Custom GUI for autoregressors
Internal Quant trading based on Market Analytics for collateral
CFD Synthetic assets development
CDO Implementation , exponential derivatives
Personalised SA creation
Solana only Flash loan implementation
Banking Institution M&A
Social copy trading aspects
ROI Benchmark simulation
Financial team
Credit line VC funding
New client aggressive policies for liquidation changes as Xange aims for 50M users in 2030
Partner Up, Letters of Intent from correct partners
Collateralisation, credit line risk assessment
API Banking features Spend your locked investments
Marketing Team
Platform Growth Hacking
Reach 12M users
Target institutional clients
Aggressive advertising
Win Hackathon Grant for Solana
SEP 22
NOV 22
JAN 23
MAR 23
MAY 23
JUL 23
SEP 23
NOV 23
JAN 24
MAR 24
MAY 24
JUL 24
SEP 24
NOV 24
JAN 25
MAR 25
JUL 25
SEP 25
NOV 25
JAN 26
MAR 26

Research validation

Reasearch was conducted in a 3 step way procedure: a first stage was done through global campaigns on discovery mode to acquire user information.

A second step involved running specific surveys and analytical data to understand target audience and user requirements.

The survey showed potential client interests in Autoregressors and possibility to autotrade in 4H span period strong interests.

After understanding possible user targeted audience and interests 3 in depth interviews have been conducted to understand possible autoregressors strategies, key takeaways to implement.

Moreover, focus of the in depth interviews was mainly to ensure product validation, main trading drivers and potential attractors.

The outcome of the in depth trading interviews had 2 main strategies outlined for the Autoregressing ARIMAX learning process.

Another potential outcome is to be seen using Harmonic patterns.

Buy/Sell support resistance strategy is outlined below

Fig showing use of buy/sell strategies

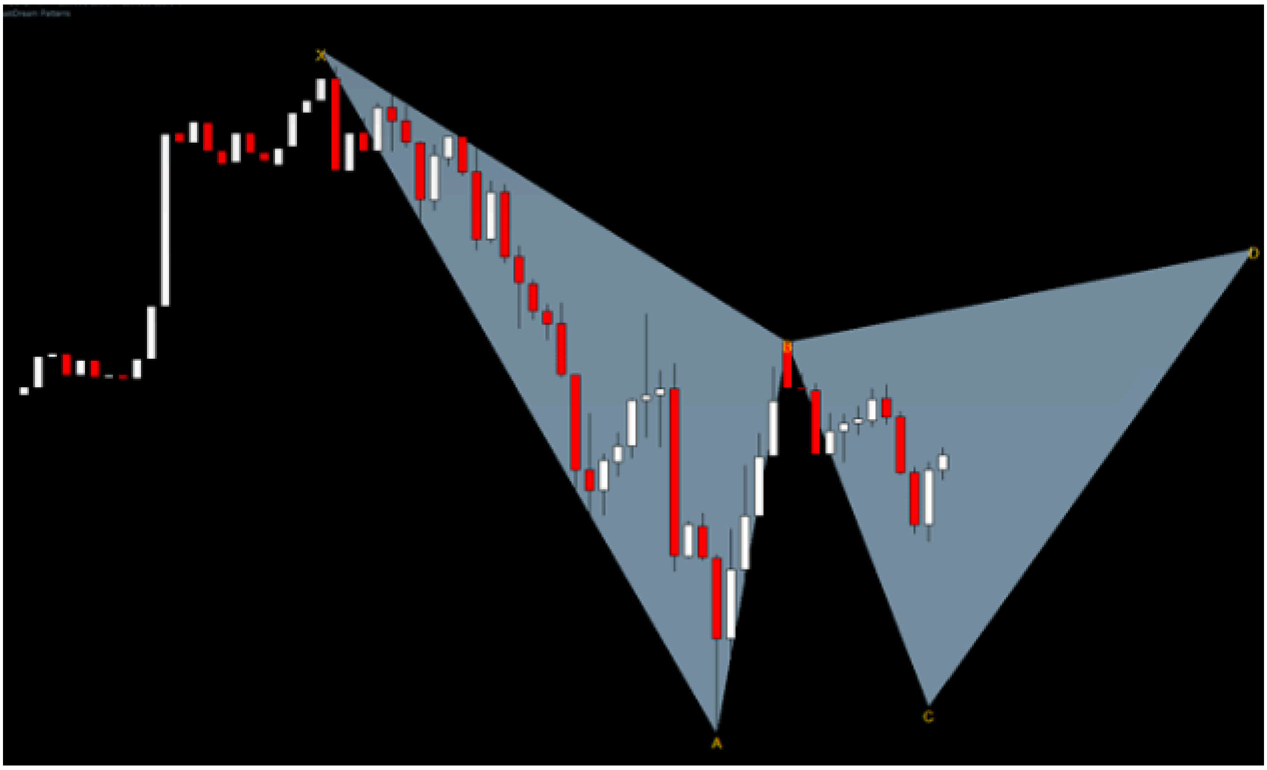


Fig showing use of harmonic patterns

Use of Confluence

Fig showing use of confluence as indicator for buying entry points in 4H span

Marketing Go-To Strategy/User acquisition

Initially the platform will be built around Solana's ecosystem.
This would allow for best growth ecosystem.
As the main interest would be initially primarily focused on user acquisition, focusing on the platform providing highest liquidity and user base.

Moreover the Solana ecosystem would offer the Solana Venture partnership program.
It would then be possible to enable synthetic contracts on such ecosystem and staff first hires through the acquired on platform grants.

On a second stage it would then be possible to partner up with diverse ecosystems to amplify liquidity provisioning and other such goals.
The main target of marketing go-to strategy shall be to increase user acquisition and thus the liquidity pool, enabling for more services and constant product validation.

Secondary target shall be to raise awareness of non decentralised-first solutions, such as Xange, merging orthodox tools to new breaucraticless fianancial streams.

Financial Breakdown

As for the Financial Breakdown it shall be important to consider costs and revenues from the intial offering and at a later stage.

For sake of simplicity, costs are estimated at 1-3yr into development and revenues as well, projected rate is then estimated for upto 5 years

Cost Breakdown

- OTP: One Time Payment: Marketing, deposits to API financial class providers (eToro API guarantee deposit/collateral)
- Servers for High transaction throughput (ClickHouse/InfluxDB + API infrastructure in C++)
- Load balancing VPS, scalability for upto 200B req/s, cost is around 6k$ max per month, ARC is to be estimated at 80k$ max
- Smart contract servers nodes with AutoML

Revenue breakdown

There are 3 income sources: fees, subscriptions, and data hedge funding.

Tier based transaction fee approach:

- Tier 0 ~10k/year: 0.05%/to
- Tier 1 10k-100k: 0.0025%/to
- Tier 2 100-250k: 0.00125%/to
- Tier 3 250k-0.5M: 0.000625%/to
- Tier 4 0.5M+/year: 0%

Subscriptions:

- Free: 1 autoregressor, 30% profit
- Basic: Limited to 3 Autoregressors, 50/mo + 10% profit
- Advanced: 7 Autoregressors 80/mo + 7% profit
- Premium: 20 autoregressors 120/mo + 5% profit
- Partner: Custom AR/ custom pricing

Hedging solution (artificial sampling limited):
Leveraged hedging based on ecosystem trade positions. Essentially future hedging against current market based on current smart contracts, since markets will be algorithmically driven, company will be able to assess price actions and buy CFDs against the market

Conclusion

Xange: current work and next steps

Based on Fisichella et al. work is possible to assume an autoregressor working on both candlestick pattern image recognition and possibly the aid of technical patterns as well, such as the harmonic channel and confluence as shown above.
It may also be possible to work with other technical indicators in the future, such as EMA 200 or the fibonacci retracement as noticed in the in depth interviews.

Current implementation of programs allowing for the autoregressors to run for synthetic assets is to be finalised.
The initial approach will be to use Solana as seen from Yahoo Finance to be the fastest solution for accounting of over 500 transactions per second (TPS) on Orca Dex.

Once the programs for the most common perpetuals on the market will be finalised, so effectively autoregressing such timeseries, the next step is to enable synthetic assets for non yet existing.

After building the ecosystem on Solana it may be necessary to have changing autoregressors to adjust for artificial sampling.
Effectively the priorities of such platform will be to grow the liquidity pool and mitigate risks.
Other challenge will be sustainable growth: the platform will want to transition to a full sustainable protocol, else user acquisition may always be partially impeded.
Moreover the technical challenges imposed by relying on an other protocol will determine technical burden and debt in case of deprecation and sunset of certain code core components.

Xange is currently being built on the Solana ecosystem as one of the fastest transaction enabler, strong liquidity dexes to partner with (Orca, Drift).
The next step is going to be to enable non yet available fractional perpetuals on chain transactions (bitCFDs)

After Solana, whilst user acquisition continues, due to the nature of the vast change in the DeFi space, a different custom protocol may be needed

Unlike other DAOs, defi applicatives and exchanges, to mantain active user growth and potential, avoiding market cap saturation, liquidity issues or user churn high rate, it will be necessary to mantain active functionality, possibly even run a custom ledger solution.

The blockchain space, being an exponential technology, has been fastly evolving, and it is expected to continue doing so for the next years: this will pressure platforms such as Xange, reliant on technological debt to either pivot or create its own ecosystem to build upon.

We believe the best approach on the longterm solution and for longevity, cost-effectiveness of the product is and will be to mantain an active SaaS contract until the technology evolution will have reached a point or phase of maturity.

When such point in time of maturity shall be reached it should be possibel to envision a custom protocol implementation

Annex

The Autoregressor

The Autoregressor Smart Contract

Motivation

> Motivation behind the autoregressor is to enable the layman trader to aid in the decision making process without the added risk of delay due to the intrinsic nature of social copy trading.
> The smart contract cloned effectively would enable for fair opportunity, having the models running onto a centralised system and not on the smart contract itself as this would be utopic, since ML models would require high CPU/GPU computational demand.
> Thus running in cloud would mantain the costs low for the platform, and effectively the smart contract running woud query for the model with certain parametrs and sampling variance in order not to artificially develop a market driven bias, so a random sampling, queried through API and model specificity .

Foreign exchange trading accounts for over 6 trillion dollars of volume settlements daily, accounting for over 50 percent of world daily economies' settlements, considering ETFs, commodities and other asset classes [1]. Out of the over 6 trillion, 5 only go to the EUR USD pair, effectively constituting 83 percent market share, and just south of 45 percent world traded volume: for this exact purpose we find such currency pair as a perfect fit to reflect most macro and micro economic trends, seasonality and indicator of global events [2].

Machine learning interest has risen recently given the latest advancements and practical application use cases in such field. This has led to a growing number of papers and new techniques addressing real time series forecasting for non-stationary data, such as the forex market. Another growing advancement has been seen in distributed computing, through NewSQL databases and esoteric technical implementations, which have facilitated the growth of such discipline considering new database use cases facilitating the consumption and manipulation of lengthy time series.

This paper analyzes the latest advancements in real time distributed computing querying and specifically analyzes the use case of ARIMA, GRU and technical indicators with LSTM for candlestick pattern regression. Finally we provide a benchmark of the potential use cases for such algorithms and methodologies, effectively analyzing the best use case for chosen time series of the USD EUR.

With the increasing growth of machine learning applications, distributed computing advancements and neural network library availability, use cases in such space

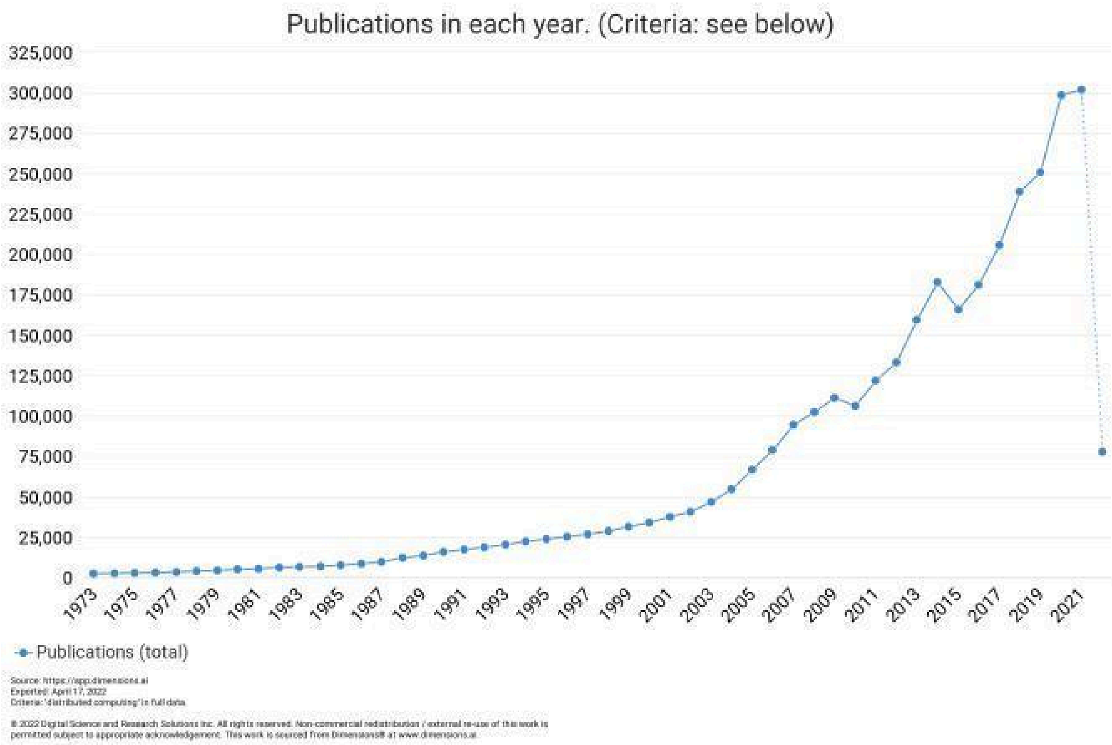


Fig. 1. Total publications for keyword 'distributed computing' on Google Scholar since 1973

have not only increased by a substantial degree, but have also seen extensive new breakthroughs in industry applicability, making choices of best practices, algorithm selections and overall Big Data infrastructure usage less obvious.

At the overwhelming annual compound growth of 39.50% CAGR [4], machine learning applications have been evolving at an unprecedented pace when compared to other exponential technologies domains.

The same applies to Big Data applications as we can run a python script computing total number of publications since 1973 for the keyword distributed computing as shown in Figure 1 And, seamlessly, the number of citations in such have also been rising, registering a steep slope of citations starting in 2013-2014, when the first release of Apache Spark software was issued, and NewSQL software had been surging in popularity, as shown in Fig. 2. Moreover the applications of machine learning in the foreign exchange market have also shown a wide interest and industry applicability.

The corresponding CAGR for forex automation is estimated at 12.5% [3] and it is found to be automated in the industry at around 70%

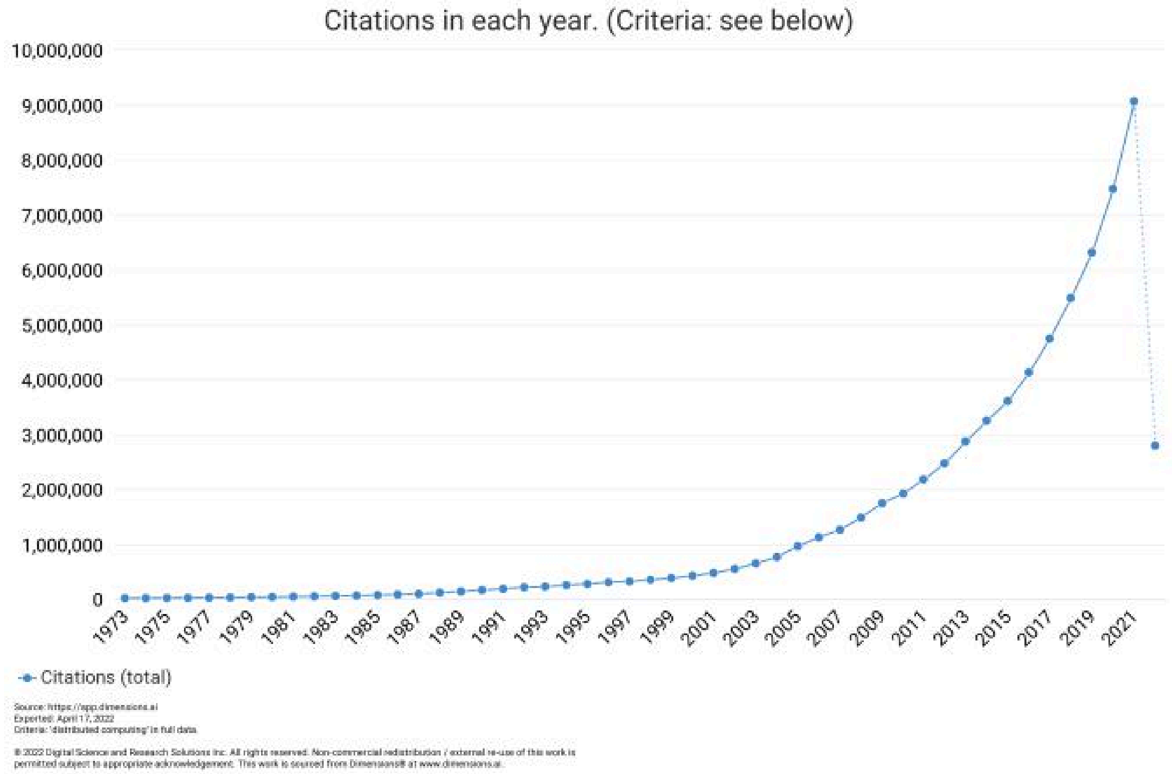


Fig. 2. Total Cit. for keyword 'distributed computing' on Google Scholar since 1973, slope spiked at 13-14

II. THE PROBLEM

With the expanding and availability of new techniques, use cases and algorithms available it becomes less trivial to decide on which methodology to apply and adopt.

The reduction of problem dimension to only one currency pair and intraday window facilitates the benchmark process in such a way: thus, this paper focuses mainly of the modern state of the art comparative research, as results may be skewed or otherwise dubious. The problem identified in such context is inherently caused and related to the latest developments in the ML field, which are driven by very narrow use-cases, further investigation thus is done to understand the relevance, applicability and causation amongst time series forecasting and modern research approaches.

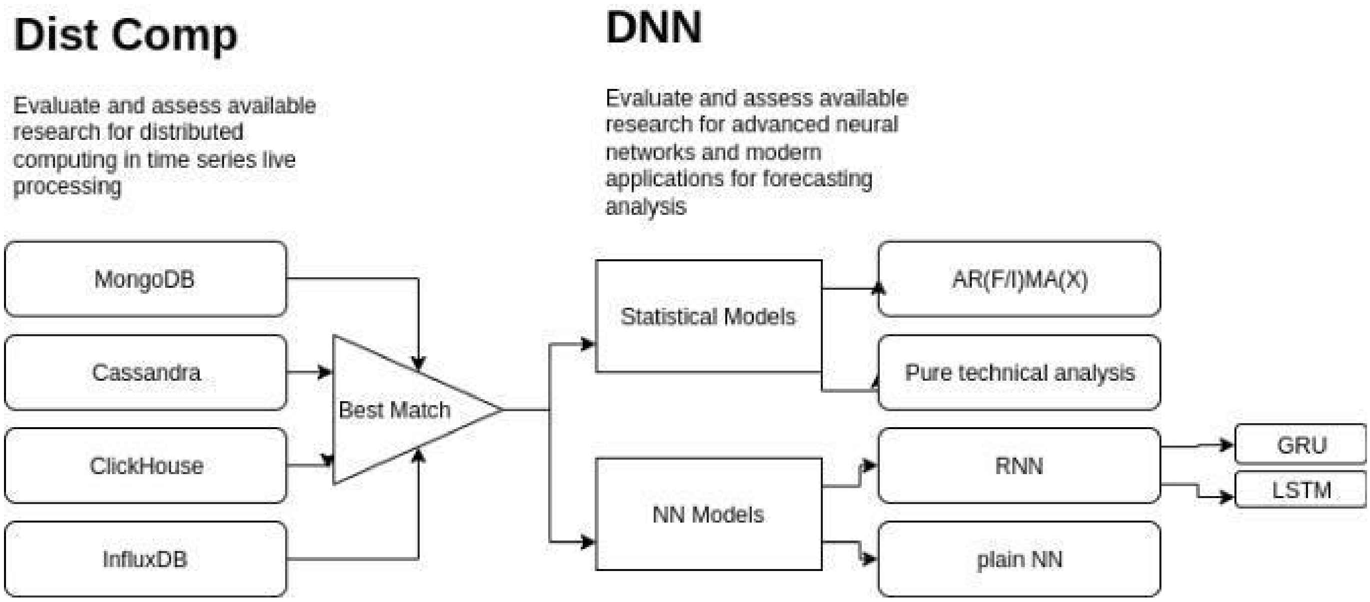
A. A common scenario

In modern literature many different algorithms have been continuously being used for regression techniques and dis-tributed computing.

We have found a growing amount of esoteric techniques in database management and aggregating clustered data systems. After the rise in adoption of technologies such as Spark, NewSQL DB have been introduced in the market.

InfluxDB, for instance, is a real time time series forecasting DB allowing for direct high performance time series storage and manipulation [6], for high performance specific scenar-ios, such as the Large Hadron Collider at CERN, mainly the ATLAS project, the exclusivity of the time and space constraints renders the comparison amongst new approaches a fundamental step. As Vasile et al. have noticed in

2020, Clickhouse as real time event observation management system for time series has also been assessed in comparison to InfluxDB for the jet propulsion recording event [5].
Other times column storage databases, such as Cassandra, or document based, such as MongoDB have been also presented with a success in literature, making the choice of technology



case dependant and less trivial. [8]

As for timeseries non-stationary data forecasting and regres-sion, many approaches have been proposed, of which mainly explore the deep neural networks techniques, ranging from auto-regressive statistical models, such as ARMA, ARIMA, ARFIMA, and recurrent neural networks, such as Long Short Term Memory models, Gated Recurrent Unit and plain neural networks.
Novel approaches using technical indicators as cross validators have also been used with promising results.
Moreover newer advancements in terms of quantum algorithms and esoteric techniques have also been reported.

B. Possible approaches

Research methodology for this paper is mainly focused on causal and descriptive available data.
Objective of such methodology is to identify a framework to use as a validator for possible distributed computing and regression applications to the specific scenario of a high volume non-stationary high frequency time series, such as the EUR/USD chosen for the purpose of intraday trading.
In Fig 3 we can see the possible decision methodology chosen when comparing all possible candidates for real time time series loading from database and forecast analysis.
Such approach allows us to thoroughly examine available present and past work and produce a result as outcome of such analysis.
Such result is the byproduct of the examined topics as in 3.

C. Objectives

Objective of the paper is to outline a methodology, based on prior findings and assumptions, which would allow to address the forecasting problem for EUR/USD forex trading, selecting a most specific and performing model and data management system for time constraints purposes.

The methodologies chosen to address the quantitative indicators of performance and Big Data systems are selected amongst traditional literature review and prior work: Mean Squared Logarithmic Error Loss (MSLE) and model accuracy for the neural networks; for the Big Data platforms loading time is used as benchmark metric.

MSLE has been selected due to the nature of larger values that could possibly be the primary objective is to validate whether or not it is possible to build a methodology to consistently train on live data and perform model predictions for intraday trading in a large, fundamentally driven, market.

Secondary objective is to select a best effort model for real time time series forecasting of non-stationary data

III. RELATED WORK

In this paper prior research available for distributed com-puting and neural network applications has been analysed, starting from the Big Data Applications as outlined in Fig 3, following comparison of research advancements in neural networks applications and esoteric technologies. At least 10 papers per each topic were chosen and reviewed to outline current findings.

Before delving into the 2 topics of research it is also essential to mention the market hypothesis theory formulated by Eu-gene Fama [17], in which he formulates the Efficient Market Hypothesis.

The Efficient Market Hypothesis consists in stating how mar-kets tend to follow general rule of efficiency, where by effi-ciency the author defines an equilibrium state in which markets tend to be eager and, thus, efficient, attempting to satisfy all macroeconomic technical and fundamental indicators.

All though mostly efficient, Prof. Fama, has also remarked how these markets may sometimes become lazy, or rather, inefficient.

It is in such cases that technical analysis and possibly forecast-ing or modeling may have an unfair advantage, meaning the markets are theoretically under performing when compared to statistical evaluators.

To test this theory Prof. Fama has attempted three relevant information subsets [...of tests] [17].

The first form being a weak form test, in which information is just derived from historical prices; the second form a semi-strong form test, where other than price history, also fundamental information is readily available to be used at buyer/seller advantage.

Finally Prof. Fama describes the third stage of testing in which a strong from test is conducted: in such test he explicitly evaluates his theory under restricted fundamental information, supposed to be given access to only monopoly driven markets. This theory has been put under test many times and up to date is to be viewed as one of the most accepted market hypothesis: thus the key takeaway from such theory derives from the rare event of inefficiency of such markets, validating techniques such as machine learning, which can be mathematically be categorized as primitive functions of statistical states, thus again being validated only in lazy market conditions.

Another important market hypothesis lies in the random walk hypothesis formulated by Jules Regnault (1863), later then explained from MIT professor Cootner in the 1964 [20], and finally popularized by Burton Malkiel.

The random walk hypothesis consists in the theory of mod-eling any economic asset class [21], thus including stocks

and foreign currency exchanges, as the statistical model of a random walk. Thus market making can be seen as discrete fractals as modeled in probability theory, as seen also in Markov's processes, future steps are independent events of past ones.
This is also the reason random walks models can be argued as a market which algorithms cannot model or forecast, as the mathematical model is the same as the discrete fractal:

$$S_{t+1} = S_t + \mu \Delta t S_t + \sigma \sqrt{\Delta t} S_t Y_i$$ (asset modeling under RWH [21])

where

μ is a drift constant

σ is the [[standard deviation]] of the returns

Δt is the change in time

Y_i is an [[Independent and identically distributed random variables—i.i.d.]] random variable satisfying $Y_i \sim N(0, 1)$. This theory is seen by Prof. Fama as a perfect model for when markets are efficient [21].
There is still a strong debate onto whether markets behave efficiently over 50% of the time [21], effectively making ma-chine learning applications and distributed calculus applicable in such context.

A. Distributed Computing

Big data analytics can be defined as the field consisting in the discovery of patterns, correlations and inference in data through the collection, recycling and examination of large temporal, spatial multi-modal and multi-source data sources. Literature review of Big Data applications dates back to 1973, with the keyword 'Distributed Computing' first appearing ever in published journals.
After 2013, following the introduction of Apache Hadoop in 2006 and with Apache Spark technologies released in 2014, as shown in 2, such topic grew in use case, applicability and NewSQL approaches began to enter the industry market.

1) Orthodox approaches:

In common literature the traditional methodologies consist in using Apache technologies, or better yet distributed file systems (DFS), or NoSQL approaches such as column-stored value, document based databases.
Three main data managers were found when researching for DFS or NoSQL approaches.

a) Apache Hadoop:

As for the first distributed file sys-tem to be introduced in 2006, Hadoop had acquired wide usage in the industry.
Hadoop allows for distributed clusters of data available to be aggregated through the MapReduce functionality and prepro-cessed.
Hadoop's strength is to enable for high performance data availability for big clusters, when data driven applications started to incur into slowdowns, disk faults, server failures, a clustered approach became evident and needed.

Hadoop’s main strength, alongside the MapReduce function-ality, allowing for large amounts of data to be interpolated and transformed in a relatively fast way, is its resilience, allowing for server fault tolerant nodes and continuous availability.

b) **Apache Spark**:

Spark, first released in 2014, is a high performance parallel specific framework: it was specifically designed with the purpose to facilitate recursive manipulation over the same data.
Originally thought as an extension of Hadoop DFS, Spark has become increasingly used for its vast efficiency in machine learning applications.
This is due to Spark having great in-memory allocation, allow-ing for much faster machine learning applications using such. Another core aspect of Spark are the Resilient Distributed Datasets, distributed and fault tolerant, collecting subsets of elements and either producing transformations (so effectively producing other RDD) or specific actions.
Another advantage of using Spark that allows for smoother computations and less overall I/O burden is the Lazy Evalua-tion for RDD, the actions returned values will only be com-puted upon necessary need, allowing for greater performance. For clustered applications 2 main methodologies were gen-erally seen, one consisting in running Spark on Hadoop, and the other being running a Beowulf cluster with the Open Multi Platform shared memory multiprocessing software (OpenMP) which would also allow for single point aggregation of data through distributed systems (such as the Beowulf clusters).

As remarked by Reyes-Ortiz et al. in 2015 [10], although very promising and powerful after conducting a trial using the Higgs Data Set (42, 5) consisting in 11 × 106 samples of simulated signal processes, results were quite interesting, and on the contrary showed a slight superiority of Spark in-memory allocation when compared to the OpenMP platform, and also Reyes-Ortiz [10] concluded that Spark

> offer[ed] a distributed file system with failure and data replication management [..and] allow[ed] the addition of new nodes at runtime

These results were surprising as he then further concluded HPC (high performance computing) was far from reaching applicable context performance, using 2 benchmarking algo-rithms, respectively kNN, and the Pegasos variant Support Vector Machine, yet found a superiority in Spark when compared to multi processing platforms based on clustered software, such as Beowulf.
Moreover in the comparison done by Samadi et al. in 2017 it is clear that Spark is a superior product than Hadoop, but lacks in requiring a vast amount of in-memory allocation: from our results we draw the conclusion that Spark is more efficient than Hadoop to deal with a large amount of data in major cases [; however], Spark requires higher memory allocation, since it loads the data to be processed into memory and keeps them in caches for a while, just like standard databases. So the choice depends on performance level and memory constraints.

c) **MongoDB:**

MongoDB, a column-based storage system, first introduced on the market in 2009, has also seen many increasingly advantageous use cases.

In the case of machine learning applications, MongoDB is highly performing when compared to SQL traditional DB, such as MariaDB, MySQL, InnoDB engines, and overall to relational database systems.
A relational database provides the advantage of routines, triggers, procedures, stored columns and other ’manipulative functions’.
Palanisamy et al. have reviewed the major drawbacks and differences between MongoDB and Relational databases.
A relational database has the biggest limitation of having very little scalability support, which in exchange MongoDB, being a non-relational and very fast, without any connector added head preprocessing latency, has the advantage of low CPU throughput, maintaining most of the functionalities.
The reason MongoDB is not as widely adopted as MySQL is it can become challenging to create complex features and serve specific aggregated views or breakdowns of segmented data, as routine and other similar functionalities is not supported at all [13].
Moreover as Abramova et al. have compared in 2013, it appears after exhaustion of a 50/50 split test in reads and updates and 95/5 mixed between reads/modify/write cycles, MongoDB can be quite slower in speed when compared to an alternative solution offered by Apache, Cassandra, a column based SQL like non relational database: MongoDB started to reduce performance, sometimes showing poor results. Differ-ently, Cassandra just got faster while working with an increase of data. [8]

2) Esoteric Methodologies:

NewSQL:
Many new technologies have been proposed in recent years regarding the specific scenarios and specific use cases, for instance, it has been suggested for the ATLAS project at CERN to use the ClickHouse, which is a column based storage for real time jet propulsion tracking.
Moreover for time series Wang [7] has found a superior advantage for real time system monitoring and IoT applications, similar as InfluxDB, proposing the Heracles time series management system.

B. Time series forecasting

Time Series forecasting is a specific branch of Machine Learning which attempts to infer the next value(s) of a given dataset based on prior multi- or single-variate data.
Before we delve into the details of data forecasting, we hereby give a general definition of the goal and type of machine learning of interest in this particular context.
Time Series forecasting can be recursive or non recursive Data can also be stationary or nonstationary, cyclical or non cyclical and can have trend patterns as well.
Given a set of data D, classical supervised framework is defined as $D_n = \{(x_1, y_1), ..., (x_n, y_n)\}$, and sampled over an unknown distribution μ over XY where $X \in R_d$ is an input space ($x \in X$) and $Y \in R$ is an output space ($y \in Y$). In binary classification, it is solved $Y \in \pm 1$. Given a primitive F, mapping to space Y, machine learning models can

be defined as functions F : X −→ Y which best approximates μ. There exists two recognised learning methods classes: Lazy Learning (LL) and Eager Leaning (EL) [24] [23] . Whilst Eager Learning takes the whole sample, the main disadvantage is the whole dataset D_n must be loaded in-memory in order to obtain the trained model, whilst lazy models, such as kNN, do not need to store the whole dataset as they only require a sample to train on (k-mean sample)

1) ARIMA and non-stationary data:

Autoregressive moving averages have been the state of the art used for many years, as the most common and accepted models based on the MA indicator, which is at the basis of technical analysis.
ARIMA models, whilst, ARIMAX allowing for multivariate analysis, allows for specific forecasting at best with detrended stationary data.
Moreover, as usually time series in financial markets are non stationary, datasets must be detrended and also deseasonalised. The FLF-LSTM outperforms FB Prophet, ARIMA, and RNN-based models of Forex prediction. And given wavelet SVR forecasting was already benchmarked better against ARIMA for forex , it is rational to assume possible modification of LSTM derivatives could outperform Raimundo's adaptive model. Neely et al. have suggested filter, MA, and channel rules, particularly 0.5%, 1%, 2%, 3%, 4%, and 5% filter rules, as well as MA(1, 5), MA(5, 20), and MA(1, 200) rules. It was also noted in both Neely and Fisichella how a cutoff point of moving average of 150 days or lower was ideal as it would thereafter underperform [18].

2) **Other NN derivatives:**

plain NN, RNN, GRU and LSTM: Particular revised models, other than autoregressive models,
using DNN, have been tested. In 2018 Ni et al. have proposed a convolutional neural network in combination with a recurrent neural network with over 10 years of data. As the plain neural networks have more suitability over general time series forecasting, they found their model to have less of an error margin than using CNN or LSTM [30].
Gradient Recurrent Unit and plain NN have also seen a particular use case, as gradient recurrent unit allows for best non-stationary data regression, whilst plain neural networks allows for complex data, as in non-stationary and trend-driven to be regressed with better success than ARIMA.
Let x^D_t = [y_t^D, z_t^D], where y_t^D and z_t^D are correspondent volatility and price variables of log ranged timeseries D Liao et al. has found the following model, using 2 plain LSTM and an l layer DNN to better performing than a plain LSTM or 2-LSTM, essentially, combining the DNN layer to the concatenated LSTM outputs as in EQ. 1, this would allow for better volatility correction in model as factored in [28]

$$f_\Theta\left(x_t^D\right) = \mathrm{DNN}\left(\mathrm{LSTM}\left(y_t^D\right), \mathrm{LSTM}\left(z_t^D\right)\right) \qquad (1)$$

EQ 1: eq describing the 2 pair LSTM and multi scale dependence factor ([28])

3) Esoteric approaches:

a) Candlestick and technical indicators image recognition: Fisichella et al. have performed a strong literature for unorthodox regression techniques, such as the application of technical analysis candlestick pattern recognition systems, consisting in RNN image recognition pattern based on numer-ous prior literature used technical analysis

b) Genetic algorithms and quantum models: Such applications are now increasing in adoption as technological drive and understanding increases. in 2021 at Barclays a model for financial forecasting and time series prediction was trained under Sofija Dimoska [22].

As found by Prof. Dimoska the QNNs can be used effectively to model time series having, at the same time, the significant advantage of being trained significantly faster than a classical machine learning model in a quantum computer. The main drawback to such implementation derives from the imple-mentation feature engineering, requiring specific time series with small amplitude noise variations, in order to satisfy the constraints dictated by the optimisation problem which is to solve for quantum circuit training.

The quantum algorithm allows for the choice of bidirectional temporal data analysis, producing more context-driven results, especially in trends and non-stationary time series. The main model chosen is the BiLSTM, consisting of a system of 2 engineered LSTM, one propagating forwards and the other backwards, yielding more insight on the temporal data analysis.

Whilst the framework of choice and language was still driven by imperative logic (Python 3.8 and Tensorflow), nevertheless the qubit state analysis represents a non trivial problem to solve for, given the Hilbert space rule: respectively q can be measured as 0 if in state q_0, else as $q_1 = 1 - q_0$.

$$|q\rangle = \begin{bmatrix} q_0 \\ q_1 \end{bmatrix} = \begin{bmatrix} q_0 \\ 1 - q_0 \end{bmatrix} \qquad (2)$$

Whilst it is clear quantum algorithms may run and train much faster, it is yet to be seen whether time series aggressive feature engineering is recommended for such use case.

Another paper published in 2019 investigated the option of using multi-output least squares support vector regression (MLSSVR) [25]. Whilst genetic algorithms have been reviewed in a greater time span, the progressive newer models are found to be yielding better performance from swarm intelligence optimization rather than genetic evolution or fitness functions.

This is to be expected as genetic algorithms heavily rely on the nature of hardcoded functions, such as the fitness estimator, which sets the threshold for the evolutionary steps, whilst swarm intelligent algorithm be more context dependant and less evolutionary dependant.

Prof. Chou has found the nature of interval-valued time series (ITS) to be better well suited for financial forecasting, as well benchamrked in non stationary time series.
The Multi-output time series algorithm allows for better hyperparameter fine tuning and custom adjustments depending on the nature of the time series.
Main performance was seen with the USD/CAD pair (at hit rate of 53.3%) and AUD/JPY (hit rate of 51%).
Whilst the model is indeed promising as Prof. Chou himself has noticed the limitations are to be seen in the nature of the hyperparametrisation tuning, which tends to overfit and be less applicable for long-term strategies: One of its weaknesses is the need to set many parameters of the system through self-tuning. The system may not provider favorable long-term investment results. To address these limitations, future research should focus on improving predictive performance.

C. Optimization techniques

IV. REVIEW METHODOLOGY

The literature review was carried as summary for selected papers, in regards to Cooper [15]. Publications cited and used are provided as means to pursue the aforementioned research objectives. All research, whilst not inclusive of all keyword search terms has been selected for specificity and to address topics needed for the problem investigation. Whilst we examined more citations than hereby presented, only the purposeful most applicable have been hereby cited.
Moreover the review has also been carried over in accordance to Kitchenham and Charters outlined guidance in Guidelines for Performing Systematic Literature Reviews in Software Engineering [16].

A. Review Search Terms

When coming to selection sources and databases used, mul-tiple sources were used, cross referenced and, inspecting the arguments and findings provided in comparison to available prior cited research.
Moreover the data sources used were mainly validated through cited prior work, if the paper appeared to have relevancy and pertinence but failed to have citations and, the cause of such did not correlate to a temporal factor, this paper does not provide a direct citation to such research due to peer review ambiguous examination.
Search terms used were concatenated with keywords such as 'AND', 'OR' and were chosen to match specificity to the topics within the forex trading context.
For Distributed computing the keywords in the set ['forex', 'trading'] were used in conjunction with [('neural network', 'ARIMA', 'LSTM', 'comparative machine learning'), ('big data', 'spark', 'hadoop', 'influxdb')] using as interpolators ['AND', 'OR', ' ', 'VS'].
Data sources used can be found as shown in table I

TABLE I
TABLE SHOWING DATA SOURCES CHOSEN AND ANALYZED FOR RESEARCH SELECTION

Database sources	URL
ArXiv	https://arxiv.org/
Springer	https://link.springer.com
ACM Digital Library	https://dl.acm.org
IEEE Xplore	https://ieeexplore.ieee.org
Wiley Online Library	https://onlinelibrary.wiley.com

B. Review Selection Criteria

Following search criteria, after examination, papers were selected into possible candidates, depending on suitability given multiple excluding factors as detailed.
The papers selected were chosen on a basis of specificity, applicability, pertinence, credibility. The papers not meeting such criteria were not included a priori, as not meeting the review guidelines nor possible field pertinence.
The following selection criteria were used explicitly.
Inclusion criteria:

A The paper is relevant to the topic of Big Data or Neural Networks, particularly for forecasting time series
B The paper describes a comparative approach when using neural network techniques for forecasting
C The paper describes a comparative approach when using distributed computation for datasets
C The paper addresses forecasting optimization problems E The paper includes forex technical indicators or regression problems

Exclusion criteria:

A The paper exhausts multiple big data techniques, which none can be pertinent to the application context of time series analysis
B The paper exhausts multiple neural networks techniques, which none is applicable to the context of time series analysis
C The paper has never been officially published, cited yet or is pending publication
D The paper provides techniques previously presented through other papers, which included the authors finding as well and are relevant to the topic but not the application context defined in the research problem

V. PRELIMINARIES

The Forex market is driven by 2 essential approaches when it comes to performing and comparing investment strategies: fundamental analysis and technical analysis.
We refer to fundamental analysis as the decision process of investment based on market acquired insight knowledge, this in the form of major news events, trends, political speculation and overall behavioural sciences.
All though out of the scope of the paper, it is worth mentioning trading strategies using sentiment analysis have also been tested with an interesting profitability outcome: in a

comprehensive study in 2021 Ma et al. have found that there have been successful use cases of NN models and LSTM when performing sentiment analysis, particularly for specific markets, using news feeds such as through Twitter data, on, specific segments, such as the Saudi market [26].

. Ma et al. have performed a comprehensive study and research in this regard, comparing the NN models to support vector machine, concluding performance of RF and SVMs was better than NB; the highest performance can be found in maximum entropy and linear SVM.

The scope of this paper is rather to outline a strategy for technical analysis.

Technical analysis, first introduced in the 1800s through Charles Dow (The Dow theory), stating a market's performance and trend is directly linked to its Average indicators, thus quantifying an uptrend when the MA (moving average) is to advance a previous importance high.

This theory, along the context of efficient market making, and investment automating strategies growth, economic boost, has led to a great increase and interest in advancing technical understanding of averages indicators: the concepts of Bollinger Bands, Moving Average, Average True Range, Stoller Bands, Exponential Moving Average, Donchian channels and derivatives.

A. Trading Candlesticks

For the purpose of analysing stock trends and technical econometrics, a specific framework has been outlined: the candlestick. A candlestick is an object describing, shape, time window, market low, market high, and related open close range.

It is through such objects that a recent advance of technical analysis consists in applying technical indicators to candlesticks charts for recognizing visual hint/clues for possible trend or reversal signals.

In Fig. 4 we outline a basic candlestick structure: the shadow, the short or long bottom tail of price demand and offer not closed, and same for upper shadow. It is also possible to identify patterns based on specific candlestick sequences or patterns, such as the doji or the popularised flag, especially trending in highly volatile markets, such as cryptocurrencies. Color of the candlestick is red for closing lower than opening, else green for uptrend.

It is possible to understand volatility and price confirmation (or support) through the long tails, and possibly high volatility and consolidation in opposite case, as reversal generally would occur if a long tail/head is identified before a settled candlestick with shorter tail.

In the foreign exchange market the trading hours can differ based on the market, there are globally 4 hour different timezones, the market making no settlements on weekends.

We will focus on the NYSE/NASDAQ timezone for the scope of this paper.

Moreover a 4H candlestick window span has been chosen, as generally accepted for intraday trading [18]. There are

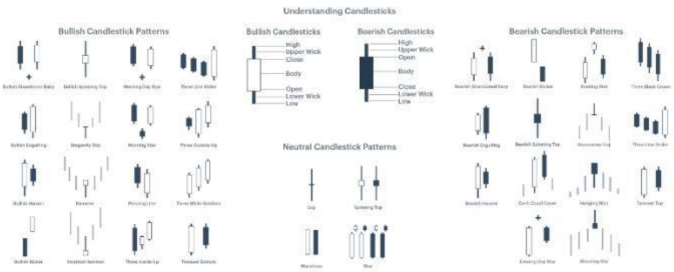


Fig. 4. Figure showing candlestick patterns and structure

three categories for forex exchanged currency pairs: Majors, Minors and Exotic.
We have chosen the largest, most traded major pair for the scope of this paper, namely EUR/USD. A market short position is to be defined when a base currency is to be sold for the quote currency, vice versa for buying the base through the quote it is to be defined as long, the base currency being EUR, the quote being the USD dollar.
In trading terminology an increasing market base price settlement is to be referred as bullish, viceversa a quote increase over a stable or diminishing base is to be referred as to bearish.

B. Technical indicators strategies

The first technical indicators to appear in history were correlated to the price opening, price closing and volume of settlements, thus defining the average movements and trends of settlements. As technical analysis became more used and accepted as a supporting investment methodology, new technical indicators have been added to the field, ranging to the traditional support and resistance bands, essentially describing the price range movements and breakouts or consolidations, retracements, an evolution of support and resistance bands, such as the fibonacci retracements, which describes in a recursive fashion the support and retracement bands under exponential curve volumes ($\Theta(\log n)$). Finally more advanced strategies have been highlighted, such as the Bollinger Bands, donchian channels, and exponential moving averages.

C. Moving Average

The simple moving average describes the unweighted mean of k samples in the window span used of data sample. Normally the same number of points is taken from the central value of reference in a time series. In our case being the 10th value $p_1$0 for k data points, with 10 previous p points and consecutive 10 for a 20 window time span.

D. Exponential Moving Average

An improvement on moving averages has been performed using the exponential moving average model design. The EMA model can be seen as a series limit towards infinity

$$S_t = \begin{cases} Y_0, & t = 0 \\ \alpha Y_t + (1 - \alpha) \cdot S_{t-1}, & t > 0 \end{cases} \qquad (4)$$

- The coefficient α represents the degree of weighting decrease, a constant smoothing factor between 0 and 1
  . A higher α discounts older observations faster.
- Y_t is the value at a time period t.
- S_t is the value of the EMA at any time period t.

Exponential smoothing can be used as a model based on EMA, quite similar to statistical autoregressors, it is best used for non stationary data. As SES (simple exponential smoothing) can be quite generalised and not so characteristics for level shifts, spikes and specific drops in timeseries, new models have been tested in technical analysis and machine learning.

Technical analysis has proposed the recursive approach, the fibonacci recursion as a classic standard for support and re-sistance bands at 70% 30% regressions or inflation. Moreover Monfared et al. have tested with better success a variation of SES, mainly the revised SES, RSES, to be more specific for taking into consideration price short timed high percentage oscillations.

EQ 4: exponential moving average: a recursive approach

E. Donchian channels

The Donchian channel is one of the earliest technical instruments to be defined. They highlight correlation between the current price and trading ranges over selected periods. The typical donchian channel strategy is to buy when the prior low of falls under the $\min(x | \forall x \in [Price_{20d}])$, and to sell when the price rises over $\max(x | \forall x \in [Price_{20d}])$. UC = Highest High in Last N Periods

Middle Channel = ((U C + LC)/2) LC = Lowest Low in Last N periods where:

U C = Upper channel

N = Number of minutes, hours, days, weeks, months

Period =Minutes, hours, days, weeks, months

LC = Lower channel

Whilst Bollinger Bands plot a simple moving average (SMA), donchian channels plot the extreme values for the price anal-ysis.

$$BOLU = MA(TP, n) + m * \sigma[TP, n]$$
$$BOLD = MA(TP, n) - m * \sigma[TP, n]$$

where:

BOLU = Upper Bollinger Band

BOLD = Lower Bollinger Band

MA = Moving average

TP(typical price) = (High + Low + Close) ÷ 3

n = Number of days in smoothing period (typically 20)

m = Number of standard deviations (typically 2)

$\sigma[TP, n]$ = Standard Deviation over last n periods of TP

(5)

EQ 5: system describing 20 day period rule with $\pm 2\sigma$ in John Bollinger's model

VI. EVALUATION

We here evaluated the selected InfluxDB and Cassandra as for data preprocessing, after reviewing the possible overall advantages for large data and storage speed in distributed computing.

Results were conclusive of InfluxDB superiority as it per-formed the query in under 200ms of over 20 years of daily historical forex data used (provided by ECB for the EUR-USD pair).

A general query selector has been used to fetch all possible records and matches.

As for the models used in literature, we found as shown in ta-ble ?? an historically accepted large usage of ARIMA models, LSTM models and plain NN as to be the most effective in for case specific pairs. All though this paper highlights the related consistency and quantity of work available for the mentioned models, such as NN, ARIMA and LSTM, it must also be noted how esoteric approaches, whilst still experimental, could be considered higher performing, when studied in depth in the course of research. As Ma et. al have noticed it should be said sentiment analysis using RNN and entropy measurements algorithms through news feeds, and more so, combined with newer approaches, such as the MQL5 trading strategy using an image neural network classification or recognition system using genetic algorithms [18], could potentially lead to a much better ROI (200% found in Fisichella et al. using GA).

TABLE II

TABLE SHOWING FINDINGS ACROSS DIFFERENT MODELS

Model	Findings
plain NN	• Very easy to train • Doesn't capture seasonality nor trends • Multivariate analysis can be incorporated as technical indicators with more neurons
RNN	• C_RNN better than plain RNN or CNN, and LSTM for error precision (RMSE) • Better suitability for autocorrelations and less specificity
GRU	• Hodrick and Prescott filter (HP filter) used with GRU decomposes better trends and volatility as factors, improving plain RNN or GRU (Xiong et al. [31])
LSTM	• Best for intraday trading, must be retrained completely every 2 weeks • Lag value can be adjusted for long or short term. Not best when used against target volatility
AR(FI)MA(X)	• ARIMA better than ARMA, ARFIMA best suited • Not ideal for all TS, data should be deseasonalised • ARIMAX is superior as allows for multivariate with technical indicators
Esoteric	• QA bidirectional LSTM is promising as faster than ordinary models (through qubit entanglement), can be better suited for HFT approaches ideally capturing specific trends not caught by human trader (as bidirectional) • 2p LSTM Capture both inter/intra-day autocorrelation. • 2p LSTM lagging in the prediction against target volatility. • 2p LSTM capture cross currencies correlations • Fisichella's GA Image approach is very promising for select pairs (200% ROI)

When comparing the models ARIMA and plain NN have been chosen as most literature had listed such findings, all though as expected our results carried over for over 20 years of forex data, lead to a clear higher accuracy rate and overall better score on the NN model with 4 neurons and using as activation function the 'relu' function.

Results have been carried over after MSE and loss function for ARIMA as evaluation, after also benchmarking the accuracy of the models.

CHALLENGES AND FUTURE SCOPE

The main challenge highlighted in the evaluation and as-sessment phase is to be able to identify the specific validation framework to enable for best selection model. Moreover another challenge is posed by the intention of model selection. This paper has seen a specific performance of LSTM on rela-tively short time ranges, better overall performance using plain neural network as a non-pair specific and general approach, but less specificity for intraday trading and no particular trend insight.

The common literature reviewed has focused mainly on the error selection (RMSE) and accuracy of a model in the re-gression problem: all though partly correct, due to the efficient market hypothesis [17] it is also evident how a deterministic and probabilistic approach in just forecasting may not be always ideal.

Moreover non-stationary data, even when deseasonalized and

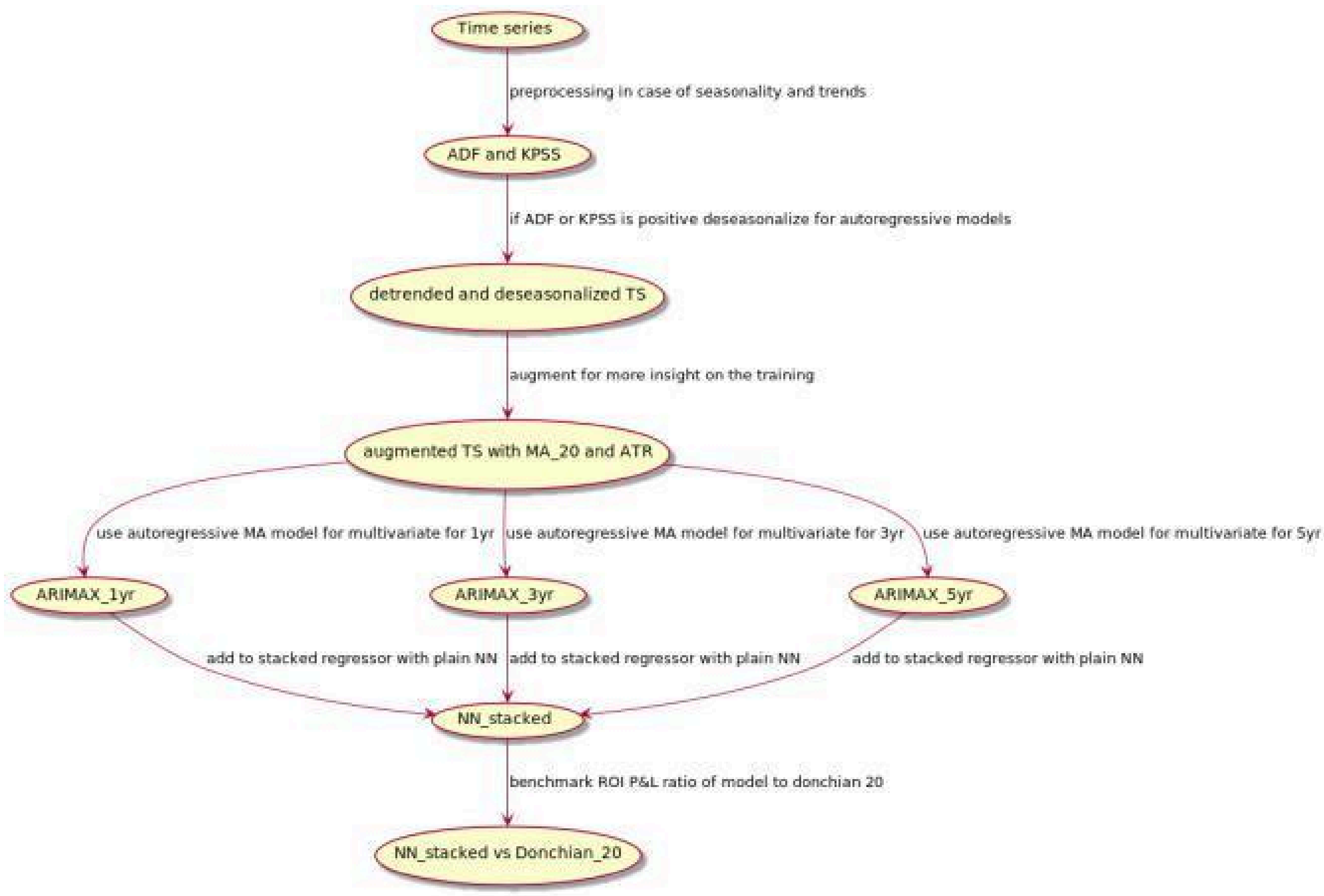


Fig. 5. Evaluation framework proposed for future literature benchmarking

fed into autoregressor models, is not always performing. In a normal trading context usually the trader's performance will be assessed through its ROI metrics.

In fig 5 we propose a framework for validating models based on profitable and not profitable classification over nor-mal model selection for regression, moreover the literature reviewed is not sufficient in multivariate time series forecasting to suggest such an approach could be underperforming when compared to single variate. We suggest using the Average True Range, MA(20) as multivariate inputs in conjunction with price range and comparing the strategy to the Donchian channel 20 day long/short rule.

CONCLUSION

Results support strong findings indicating superiority of InfluxDB when compared to other like services, such as Cassandra. ARIMA, although superior to other models tested such as ARMA, it is still not as performming as plain Neural Network regression. Moreover as ROI performance is clearly related to shorting and longing specific timeframe, which is something not commonly analysed in relevant work, a stacked model could be used with multiple yearly trained models based on a plain NN layer, and more so a bidirectional LSTM may even be researched, split into the 2 strategies, 1 for shorting and the other for longing positions. It is clear as the trading world is increasingly being automated with newer, revised versions of technical indicators, multivariate TS forecasting may better mimic the decisions of an experienced trader, where technical signals are also factored as part of the regression, for the validation, as outlined in 5 it would be optimal to validate models based on ROI performance, rather than solely relying on RMSE.

Although these findings do indicate the relevance and edge of plain Neural Networks when it comes to performance, it is also evident how the specificity of the time series influences such findings.

DATA AVAILABILITY

Data used in this paper has been provided to us by the . Data used for market daily open, close, high and low has been audited and collected by statistical public services institutions under GDPR compliance. Moreover data by the European Central Bank is publicly made available on their website (1999-Present EUR USD Data).

References

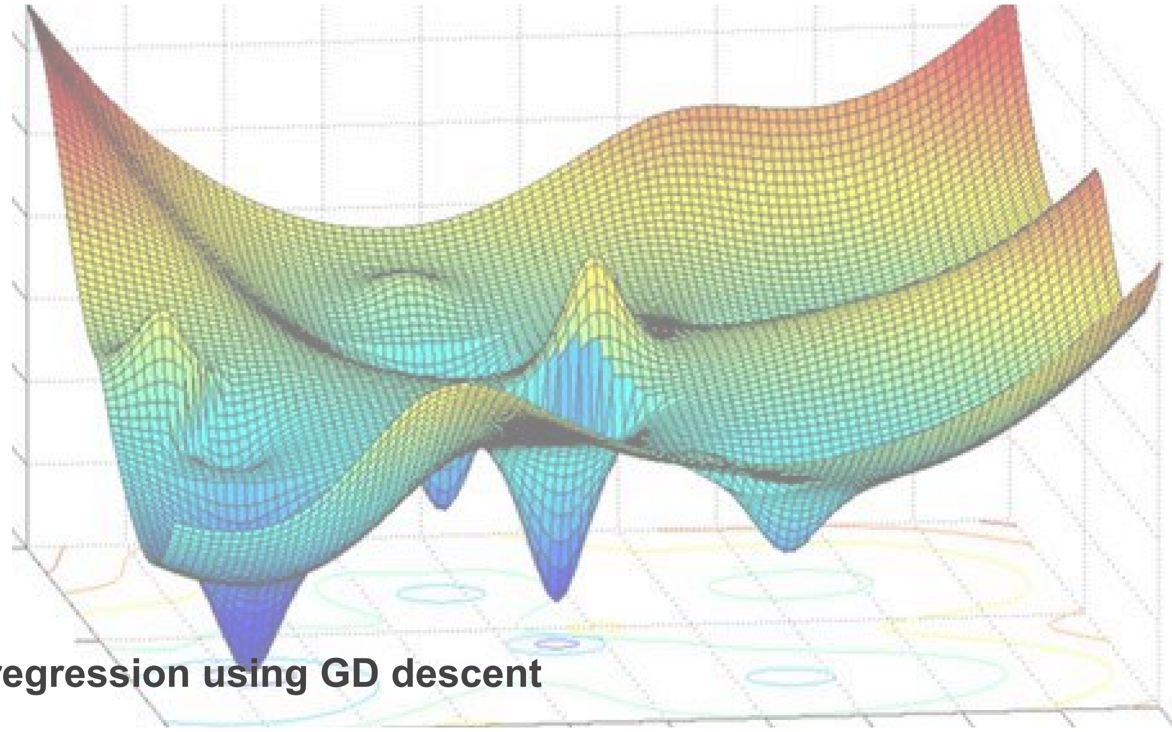
1. Cls-group.com. (2022). FX Trade Volume report.
2. 1 Nasdaqtrader.com. (2022). Full Market Share Summary.
3. and, R. (2021). Foreign Exchange Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2021-2026.
4. Vantage Market Research,https://www.vantagemarketresearch.com (2022). Machine learning as a service Market Global Size, Share, Forecast 2022 - 2028 | Vantage Market Research.
5. Vasile, M., Avolio, G. and I. Soloviev (2020). Evaluating InfluxDB and ClickHouse database technologies for improvements of the ATLAS operational monitoring data archiving.
6. Beermann, T., A. Vartapetian, Vokac, P., J. Elmsheuser, Barberis, D., S. Crépé-Renaudin, Svatos, M., Wolters, H., A. Alekseev and I. Glushkov (2020). Implementation of ATLAS Distributed Computing monitoring dashboards using InfluxDB and Grafana.
7. Wang, Z., Xue, J. and Shao, Z. (2021). Heracles: An Efficient Storage Model And Data Flushing For Performance Monitoring Timeseries.
8. Abramova, Veronika and Jorge Bernardino. "NoSQL databases: MongoDB vs cassandra." C3S2E ¢13 (2013).
9. Europa.eu. (2022).
10. Reyes-Ortiz, J.L., Oneto, L. and D. Anguita (2015). Big Data Analytics in the Cloud: Spark on Hadoop vs MPI/OpenMP on Beowulf
11. Samadi, Y., Zbakh, M. and Tadonki, C. (2017). Performance comparison between Hadoop and Spark frameworks using HiBench benchmarks. Concurrency and Computation: Practice and Experience, [online] 30(12), p.e4367.
12. Anderson, M.J., Smith, S., Sundaram, N., M. Capotă, Zhao, Z., Dulloor, S.R., Satish, N. and Willke, T.L. (2017). Bridging the Gap between HPC and Big Data frameworks.
13. Sowndarya Palanisamy and P. Suvithavani (2020). A survey on RDBMS and NoSQL Databases MySQL vs MongoDB.
14. ACM Conferences. (2015). SMiLer | Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data.
15. H. M. Cooper, "Organizing knowledge syntheses: A taxonomy of literature reviews," Knowl. Soc., vol. 1, no. 1, pp. 104–126, Mar. 1988.
16. B. A. Kitchenham and S. M. Charters, Guidelines for Performing Systematic Literature Reviews in Software Engineering. Princeton, NJ, USA: Citeseer, 2007.
17. Jstor.org. (2022). Efficient Capital Markets: A Review of Theory and Empirical Work on JSTOR
18. Fisichella, M. and Garolla, F. (2021). Can Deep Learning Improve Technical Analysis of Forex Data to Predict Future Price Movements? IEEE Access, [online] 9, pp.153083–153101.
19. Fischer, T. and Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, [online] 270(2), pp.654–669.
20. The MIT Press (2016). The Random Character of Stock Market Prices. [online] The MIT Press.
21. Jstor.org. (2022). The Random Walk Hypothesis and Technical Analysis on JSTOR.
22. Emmanoulopoulos, D. and Dimoska, S. (2022). Quantum Machine Learning in Finance: Time Series Forecasting.
23. I. Hendrickx and A. Van Den Bosch. Hybrid algorithms with instance-based classification. In Machine Learning: European Conference on Machine Learning, pages 158–169, 2005.
24. T. M. Mitchell. Machine learning. WCB. McGraw-Hill Boston, 1997.
25. Chou, J.-S., Truong, D. and Le, T.-L. (2020). Interval Forecasting of Financial Time Series by Accelerated Particle Swarm-Optimized Multi-Output Machine Learning System.
26. Ma, H., Ma, J., Wang, H., Li, P. and Du, W.-C. (2021). A Comprehensive Review of Investor Sentiment Analysis in Stock Price Forecasting.

27. Hu, Z., Zhao, Y. and Khushi, M. (2021). A Survey of Forex and Stock Price Prediction Using Deep Learning. Applied System Innovation,
28. Liao, S., Chen, J. and Ni, H. (2021). Forex Trading Volatility Prediction using Neural Network Models. arXiv.org.
29. Monfared, M.A.S., Ghandali, R. and Esmaeili, M. (2014). A new adaptive exponential smoothing method for non-stationary time series with level shifts. Journal of Industrial Engineering International
30. Ni, L., Li, Y., Wang, X., Zhang, J., Yu, J. and Qi, C. (2019). Forecasting of Forex Time Series Data Based on Deep Learning. Procedia Computer Science, [online] 147, pp.647–652
31. ACM Other conferences. (2021). GRU Stock Price Forecasting Method Based on HP Filter | 2021 International Conference on Aviation Safety and Information Technology.
32. Allied Market Research. (2021). Allied Market Research. [online] Available at: https://www.alliedmarketresearch.com/crypto-currency-market#:~:text=The%20global%20cryptocurrency%20market%20size,12.8%25%20from%202021%20to%202030.
33. GitHub. (n.d.). aave/aave-protocol. [online] Available at: https://github.com/aave/aave-protocol/blob/master/docs/Aave_Protocol_Whitepaper_v1_0.pdf.
34. Team, S. (2022). English - Synthetix System Documentation. [online] Synthetix.io. Available at: https://docs.synthetix.io/litepaper/ [Accessed 13 May 2022].
35. Juliano, A. (2017). dYdX: A Standard for Decentralized Margin Trading and Derivatives. [online] Available at: https://whitepaper.dydx.exchange/.
36. Gitbook.io. (2021). Margin System - Zeta. [online] Available at: https://zetamarkets.gitbook.io/zeta/zeta-protocol/margin-system
37. 01.xyz. (2022). Introduction - 01 - Decentralized Derivatives Exchange. [online] Available at: https://docs.01.xyz/
38. Liu, J.A. (2018). Atomic Swaptions: Cryptocurrency Derivatives. arXiv.org. [online] doi:10.48550/arXiv.1807.08644.
39. Arusoaie, A. (2020). Certifying Findel Derivatives for Blockchain. arXiv.org. [online] doi:10.48550/arXiv.2005.13602.
40. Härdle, W.K., Harvey, C.R. and Reule, R.C.G. (2019). Understanding Cryptocurrencies. SSRN Electronic Journal. doi:10.2139/ssrn.3360304.
41. Wonderland.money. (2021). Introduction - Wonderland. [online] Available at: https://docs.wonderland.money/
42. Yahoo.com. (2018). Dragonfly Research Says Solana's Speed Surpasses ETH & other Chains.
43. Sayeed, S., Marco-Gisbert, H. and Caira, T. (2020). Smart Contract: Attacks and Protections. IEEE Access, [online] 8, pp.24416–24427. doi:10.1109/access.2020.2970495.

Declaration

I, Mario Robert D'Ambrosio, declare that the Submitted Research Paper is my original work and no part of it has been published any where else in the
past. I take full responsibility, that if in future, the paper is found invalid according to basic
rules, the last decision will be of the Authorities concerned. Any form of plagiarism will lead
to disqualification of the paper

2022.05.13
Dublin, Ireland



ML autoregression using GD descent

This paper aims to outline the business proposition of an exponentially newly adapted trading product, throught for modern demands, such as technically driven indicators (autoregressors), sought as future contracts to avoid delay, interchain operability built on an ecosystem.
Such as to provide autoscaling liquidity, as debt shall be allocated to provide more liquidity in case of missing direct access to assets.
Moreover the platform shall be seen as a regulated effective financial institution as to enable many synthetic assets through high liquidity and leveraged positions, a hybrid model shall be thought.
We hereby propose a hybrid regulatory metaverse which can be accessed through both fiat and non fiat asset classes, to play at timezone agnostic market price auction, with the aid of automated rule decision making.

Fig above shows the Gradient descent optimisation for regression problems in the typical machine learning context of multivariate plane regressions.
It is also possible to view the sattle points, points in which the ML autoregressors do not perform as much.
Objective of this paper is thus to conclude the findings advising for a possible autoregressive strategy on the most traded price auction, the forex EURUSD pair.

mario.dambrosio@aerospacereasearch.net
xange.io