Meta-Learning Generative Adversarial Networks for Extrapolating Nonlinear Dynamic Stochastic Systems.

Case Study: Price Forecasting of Volatile Assets for use in Algorithmic Trading.

Jashua Luna

2022

|  |  |  |
| --- | --- | --- |
|  |  |  |

**Introduction**

Financial markets, in particular cryptocurrencies which are highly volatile, have random walk properties. Furthermore, the fundamental properties which affect an asset’s price are many and attempts to measure and model them might result in changes escaping the scope of the measurement. Despite these random walk properties, the change in price of an asset observes a constant distribution so the price of an asset can be effectively modelled as a stochastic process. The famous Black-Scholes equation describes the price of an asset as one that “follows a geometric Brownian motion with constant drift and constant volatility” (Zhang 2) which satisfies this differential equation.

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

(Zhang, 2)

Where is price and is a Wiener process of Brownian motion, and and are constants denoting drift and volatility respectively. The Black-Scholes equation is known to have some limitations, however, notably its assumption of market volatility as a constant property.

In their 2000 paper, Cars H. Hommes models financial markets as a nonlinear adaptive belief system with multiple variables relating to types of investors and price vectors; stating, “evolutionary adaptive systems with heterogeneous agents using competing trading strategies [are] a natural nonlinear world full of homoclinic bifurcations and strange attractors” (Hommes 6).

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

(Hommes, 6)

Where is a nonlinear mapping, is a vector of prices, is the fraction or weight of investors of type , is a vector of parameters and and are noise terms (Hommes 6).

A problem with these solutions is the assumption of some market structure. Particularly in an adaptive belief market, a structure which may exist at one time will likely not exist in another. Abstractions can be made through which an analytical solution might be reached but this could ultimately prove fruitless in an efficient market. The model must change as quickly if not faster than the market.

|  |  |  |
| --- | --- | --- |
|  |  | (3) |
|  |  | (4) |

Where is a vectorized sample of and is a set of parameters.

Assuming the distribution of possible future states to be finite, the price of an asset can be described as a finite state system for which there is a function which precisely determines future states from information of current and past states. The function described by phi and theta can accurately model the system while the system dynamics theta trains on are constant.

Be it slippage or localized panic or euphoria, the price of a volatile asset often oscillates around a “true” price in what will be henceforth referred to as noise.

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

These high frequency undulations in price can not be profited from because of the long time it takes to enter and exit trades; primarily resulting from request-response lag and time to fill orders. This can be mitigated by placing only market orders, but this might result in slippage which in turn affects the accuracy of a model.

**Concept**

As a solution to stochastic, deterministic system with unknown inputs, a neural network is developed which can learn how the expected distribution of future data relates to the known distribution of present and past data. This way, a likely prediction can be used to implement trades. This has the added benefit of not relying on precision as it is likely that the future price of an asset would immediately diverge from the predicted price the moment a trade was enacted on the predicted information. Generative adversarial networks learn to replicate the stochasticity of data (Isola 7) and as such are appropriate for this solution (7).

|  |  |  |
| --- | --- | --- |
|  |  | (6) |
|  |  | (7) |
|  |  | (8) |

(Eq. 7 Isola 3, Eq. 8 Finn 3)

Where is a windowed set of price data samples between and (is window size) subsampled from the set containing all of the asset’s price data, is its corresponding set of parameters, is a generator which attempts to produce “real” data, and is a discriminator which trains to distinguish “real” data from that produced by the generator.

For temporal stability between predictions, the network’s previous prediction is also input. In accordance with the adaptive efficient market hypothesis, the network must be routinely actualized such that it keeps up with the changing system. This is facilitated by the meta learning model described in (8) which ensures the model can quickly and effectively train on new tasks as they arise; changing market conditions are analogous to new tasks. Finally, the proposed model takes the form of a nonlinear mapping function which maps a vector of parameters , a windowed set of price observations ,and its previous prediction to a new price prediction (6).

**Proof of Concept**

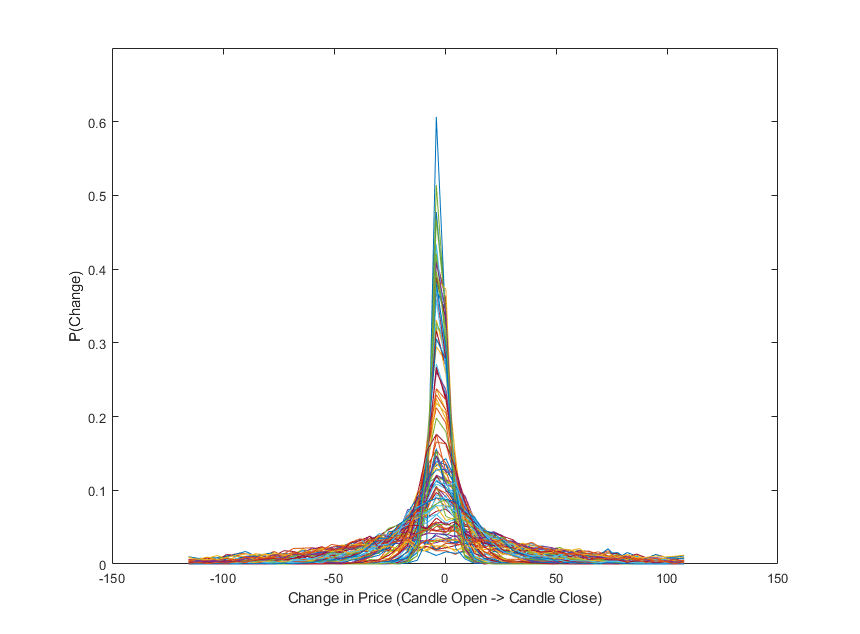
Imperative to the design of this neural network is that the change in price between candles can be reasonably described as a stochastic process, independent of market conditions. To validate this assertion, one hundred samples of two thousand candles were randomly selected from the price of Bitcoin between the years 2021 and 2022; these years were selected because they had constant and reasonable trading volume throughout. Probability distributions were then created for each of these samples using histograms, and events with a probability of occurrence below 1e-3 were discarded.

Figure 1 Change in Price per Candle vs The Probability it Occurs

Figure 1 shows the probability distributions for the one hundred samples. With ten thousand samples like the ones used for figure 1, subsample mean was determined to have mean 0.0107 and standard deviation 0.7823. Subsample standard deviation was found to have mean 30.7794 with standard deviation 24.4113. The extent to which the assertion holds true is ultimately determined by the model’s performance, but from figure one, it can at least be concluded that the dataset does observe a relatively constant mean around zero with reasonable probability that change of price is less than twenty-five dollars, for any subsample.

Demonstrate:

1. System observes constant distribution of change between samples independent of market conditions.
2. Market conditions change (show the differences between wavelet transforms at different times).

**Implementation**

The price of an asset is observed at a regular frequency and samples having n observations are created. To reveal underlying trends, and encode inputs with frequency-information, samples are upscaled by the Dualtree complex wavelet transform (). Because of its complex valued wavelet (9) – orthonormal in Hilbert space – the complex wavelet transform is shift-invariant and therefore enables temporal stability between predictions.

|  |  |  |
| --- | --- | --- |
|  |  | (9) |

(Selesnick 126)

By separating the signal into its constituent frequencies, a lowpass filter can easily be applied by removing or setting to zero, the wavelet decomposition levels which correspond to high frequencies. This is important as it smooths the recovered price signal after prediction in turn facilitating trading. Additionally, high frequency properties are the hardest to learn so the lowpass filter reduces the complexity of the required architecture.

|  |  |  |
| --- | --- | --- |
|  |  | (10) |
|  |  | (11) |
|  |  | (12) |

For the network to learn system properties unique to the sample time, a set of wavelet-transformed signals is created by retaining the upscaled samples for a specified window. The meta learning model then learns features between different sample times (with varying market conditions) and determines which are relevant when faced with a “new” system state.

|  |  |  |
| --- | --- | --- |
|  |  | (12) |
|  |  | (14) |
|  |  | (15) |

Instead of the network’s previous prediction, during training the expected output plus a noise vector is input. Additionally, a noise set and noise vector is input so the adversarial network does not produce deterministic outputs (Isola 3).

**Architecture**

**Evaluation**

Results in accuracy.

Consider showing results on data generated with known distributions as well as bitcoin data.

Ie : where W is some value randomly generated from a known distribution. System can be initialized with a random point.

Show gan predictions have similar distributions to data

Introduction

About the project.

Background

About the task in hand.

Concept

How the task is divided.

Implementation

Alternatives explored for each part (why each alternative is considered and relative performance) and “final” design.

Possibly show code.

Evaluation of Design.

Implementation

Evaluation

**Appendix**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Algorithm 1

Generating training samples from price data

1. Observe price data once per minute
2. Upscale signal with complex wavelet transforms
3. Store this in dataset

Previous Version of Report

* Abstract
* Introduction
* Concept
  + Random Walk Characteristics
  + Finite States
  + Profit Maximization
* Implementation
* Input Layer
* Output Layer
  + Unideal Interface
* Prediction Layer
  + Outlier Detection and Removal
  + Spectral Densities and Wavelet Transform
  + Kalman Filter
  + Generative Adversarial Training Scheme
  + Considerations for Training
  + Network Architecture
  + Training Prediction Network
  + Wavelet Transform Layers
* Decision Layer
  + Redundant Weighted Predictions
  + Tuned Lowpass Filter
  + Binary Signal and Profit Optimization
  + Considerations for Training
  + Decision Network Architecture
* Discussion

Quotes & Other Extracts

A Quantum Like Approach to the Stock Market

*This contextual influence is responsible of the non-Kolmogorovian quantum-like behavior of the market at a statistical level. (Aerts, 1)*

*the stock price follows a geometric Brownian motion with constant drift and constant volatility. (Aerts, 2)*

*A picture containing diagram

Description automatically generated(Aerts, 2) -> Black Scholes*

*This random walk hypothesis has characterized the \_financial derivatives modelling and valuation in the banks and \_financial institutions and is strictly linked with the so-called efficient market hypothesis, that is, the assumption that financial markets are `informationally efficient' and prices of traded assets instantly change to reflect new public information. (Aerts, 3)*

*It is a basic doctrine in quantum mechanics that a property, e.g., the value of an observable, of a*

*physical entity in each state cannot be considered as a pre-existing feature of the entity (Aerts, 4)*

*the property is actualized in a measurement process and can be different if a different measurement context is considered. (Aerts, 4)*

*probability cannot be interpreted as formalizing the subjective ignorance about the actual state of the*

*entity, as it occurs for classical (Kolmogorovian) probability and statistical mechanics, but it is rather*

*caused by the presence of a lack of knowledge (fluctuations) concerning how context interacts with*

*the entity that is considered. (Aerts, 4)*

*A quantum-like potentiality structure occurs in an opinion poll because two types of questions are possible. (Aerts ,4)*

*Indeed, it is reasonable to suppose that prices of assets are not predetermined, but they depend on the simultaneous happenings of human decisions to buy and/or sell. (Aerts,4)*