

# Formulation Of A Rational Option Pricing Model using Artificial Neural Networks

Kaustubh Yadav

*School Of Computer Science and Engineering*

*(Vellore Institute Of Technology)*

Vellore, Tamil Nadu, India

kaustubh.2018@vitstudent.ac.in

**Abstract**—This paper inquires on the options pricing modeling using Artificial Neural Networks to price Apple(AAPL) European Call Options. Our model is based on the premise that Artificial Neural Networks can be used as functional approximators and can be used as an alternative to the numerical methods to some extent, for a faster and an efficient solution. This paper provides a neural network solution for two financial models, the Black-Scholes-Merton model, and the calibrated-Heston Stochastic Volatility Model, we evaluate our predictions using the existing numerical solutions for the same, the analytic solution for the Black-Scholes equation, COS-Model for Heston's Stochastic Volatility Model and Standard Heston-Quasi analytic formula. The aim of this study is to find a viable time-efficient alternative to existing quantitative models for option pricing.

**Index Terms**—Black Scholes Equation, Heston Model Calibration, Option Pricing, Stochastic Processes, Artificial Neural Networks.

## I. INTRODUCTION

Pricing financial derivatives has always been a very centric problem in the field of quantitative finance. Option contracts have seen tremendous growth after the crisis of 2008 mainly because options are considered as leverage and the party possessing can avoid risks of financial volatility by exercising these options. Options contract can also be considered as a “zero-sum game”, where if the holder of an option is benefited, that directly correlates to the loss of the other party and vice versa, this can be considered as an incentive for investing in an options contract. Pricing of Options took its first leap with the works of Black Scholes and Metron in 1973 [1], where both studies gave a model that could roughly estimate the price of a European Option, yet the Black-Scholes formula still remains one of the most used tools for pricing options. Even with the success of the Black-Scholes and Merton model, there were substantial flaws, flaws that make the model more theoretical rather than practical. The volatility for the Black-Scholes model was assumed to be constant, which in practicality is impossible, the volatility can itself be considered as a stochastic process, the Black-Scholes model also assumed that the returns on an asset are normally distributed, which implies that the stock prices follow a Geometrical Brownian Motion(GBM) ignoring the large changes in prices [2] which can also be attributed to the models' inability to price deep out-of-money options [3].

Subsequent approaches in improving on the Black-Scholes Model, included modelling volatility as a stochastic process [2], [4], [5], [6], [3]. Though these models are an improvement on the Black-Scholes Model in the sense that they use stochastic volatility to recreate practical conditions, these models still fail to capture deep out-of-money options. Also with the introduction of stochastic volatility, there are two stochastic processes which increases the computational costs, and these models don't perform exceptionally better than Black-Scholes hence Black-Scholes is still widely used. The research for finding a better solution brought non-parametric and semi-parametric models to light. These models are essentially “model-free” and take a tangent from the theory of finance and rely explicitly on data. Hence an ANN approach can be considered as a non-parametric solution. As non-parametric models rely on historical data and statistical approaches they suffer from poor-generality and as they discard the theory of finance to some extent they cannot be considered as an empirical. This paper aims to provide a semi-parametric solution that doesn't solely rely on historical data but stays true to the theory of finance and option pricing.

## II. RELATED WORK

### A. Quantitative Methods

With the introduction of the Black-Scholes equation, the academic research sought to find out new models that relax the restrictive assumptions made by the Black-Scholes Model. The first was making the volatility in the model to be a stochastic process rather than a constant, this implied the replication of the volatility smile curve as seen in the practical market conditions, which were introduced through the contributions of the Heston Model[6]. Other models, also known as Levy Models, used Levy Process to model the volatility instead of a Weiner Process, through which the volatility curve could capture the rare jumps in the underlying, these are also called jump-diffusion models [7], [8], [9].

As mentioned earlier with the increasing complexity of the model, it becomes harder to model the underlying PDE. To model such PDE, a catalog of numerical methods have been used, eg. Spectral Method, Finite Differences, Monte Carlo Simulations, Laplace-Carson transform[10], etc. Some models including Heston Models have parameters that need to be tuned with the market data, these calibrations can be done

either by historical data and or on options that are heavily traded. Hence calibration of parameters can be considered as a multidimensional optimization process, so as to reduce the distance between the market price and the model output. But calibrating over a huge amount of data increases the complexity and is computationally inefficient, hence many of the stochastic volatility models are not as mainstream.

### B. Option Pricing And Neural Networks

Artificial Neural Networks are the building blocks of Deep Learning, introduced in 1943 by McCulloch as a single neuron [11], since then ANN's have been used in many areas such as image classification, recommender systems, etc. One of the utilities of an ANN is functional approximation for non-linear equations and PDEs [12], [13]. The premise behind a functional approximator is that if the model needs to approximate a function  $y = F(x)$ , the feed-forward requires the model inputs to be multiplied with the weights of the neurons, then the weighted inputs pass through a transfer function and finally pass through an activation function, which in turn yields the first approximate, the result is then compared with the ground truth, and simultaneously the weights of the model are modified till the difference between the ground truth and models' prediction is minimal. Hence, a functional approximation is itself an optimization problem, if the model output is represented by  $y = F_{ANN}(x, w)$ , that is as a function of weights and the input the optimization can be stated as,  $\argmin_w (\sum_i ||F(x) - F_{ANN}(x_i)||)$ .

This can be applied to Option Pricing as well, by providing appropriate model parameters. Although these methods are an example of data-driven approaches in finding a solution, they do suffer from loss of generality if the training size isn't diverse, also being completely data-driven they don't conform to the empiricals of Option Pricing Theory itself. Even with these flaws, neural networks have on occasion outperformed long-standing financial models [14], [15].

The aim of this study is to provide an accurate yet time-efficient solution for pricing a European Call Option under both constant volatility (Black-Scholes) and stochastic volatility (Heston's Method) using ANN's but enforcing the rationality conditions laid out in [16] and [17] for a "rational" option pricing model which makes this study novel compared to existing models. The paper is divided into 3 sections. The first section introduces the Black-Scholes Model, Heston Model, and the rationality conditions for a pricing model, the second section is about the modeling of the ANN, description of the dataset and hyperparameter tuning, and finally training the model on APPL's European Call Options data and evaluating our predictions in the final section.

## III. BACKGROUND

An Option is a financial contract that gives the holder to buy or sell an underlying asset at a predetermined price. Options, on a high-level, can be divided into two categories: American and European. An American Option gives the holder to buy/sell an underlying on or before the expiry date, but a

European Option can only be exercised on the date of expiry. Options further are of two types, Call and Put. A call option is a contract that allows the holder to buy the underlying asset at a fixed price point on (European) or before (American) the expiry. Whereas a put option is a contract that allows the holder to sell the underlying asset at a fixed price point on (European) or before (American) the expiry. The underlying in an option contract can be stock, bond, commodity, etc but in a generic portfolio, it is usually a stock. Hence options are also called Financial Derivatives as their own value depends on an underlying asset. A stock option can be compared to a stock warrant as well, but the key difference is, an option is a contract between two investors whereas a warrant is a contract between the company that issues stock and the investor. Another key difference is that the warrants are put out by the company hence the aggregate amount of warrants in a market is positive, whereas options exist on an exchange hence the aggregate amount of options remains zero [16].

As the subject of this study is European Call Options, we shall look into the mathematical formalization of the aforementioned statements. Let the price of an underlying ( $S$ ) be \$100 we represent the present day as  $t = 0$  and the expiry of the option be in 10 days ( $T = 10$ ), hence time to maturity is ( $\tau = T - t = 10$ ). Let the exercise price for the option or the Strike Price ( $K$ ) be \$120 if the price of the underlying becomes \$150, the holder has the incentive to exercise the option as he can buy the stock for \$120 and has an opportunity to sell it at a cost of \$150, hence making a profit of \$40. But if the price of the underlying drops to \$80, the holder won't exercise the call option. Hence the payoff for a call option can be given by:

$$\max(0, S_T - K) \quad (1)$$

To quantify the worth of an option, the goal is to find the price of the option at  $t = 0$ . This can be computed by discounting if we know the price of the underlying at  $T$ ,

$$\mathbb{E}[C'(S, K, \tau)] = e^{-r\tau} \int_0^\infty \max(0, S_T - K) f(S_T | S_t, T) dS_T \quad (2)$$

Here  $C'(\cdot)$  is the price of an option at  $t = 0$ ,  $r$  is the risk-free rate or risk associated with a 3-month government treasury bond and the  $e^{-r\tau}$  is the discount term, and the function  $f(S_T | S_t, T)$  is density function that models the stochastic process  $dS_T$  and as mentioned earlier  $\max(0, S_T - K)$  is the payoff after exercising the option at time  $T$ . This is the problem of option pricing mathematically, a econometric model aims to model the stochastic process of stock movement, eg. Black-Scholes assumes that the stock movement follows a GBM and the volatility and the risk-free rates are constant, whereas the Heston Model, doesn't make the strong assumption of constant volatility.

### A. Black-Scholes Model

The Black-Scholes Model, introduced in 1973 has had a tremendous impact on the whole financial dynamic. The

model gave a methodology for pricing a non-dividend-paying European Call and Put option. The model assumes constant volatility and models the stochastic process of Stock movement as a GBM[?] hence the change in stock price is given by,

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t \quad (3)$$

Where,  $S_t$  is the price of the underlying,  $\mu$  and  $\sigma$  are constants, and  $dB_t$  is a Weiner Process for modeling the stock price. This also makes a secondary assumption that the Stock Price is log-normally distributed or the  $\log(S_t)$  follows a normal distribution. On the other hand, an Itô's Process is the one that satisfies a similar differential equation,

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t \quad (4)$$

Further an Itô's Process for a twice differentiable function  $f(X, t)$  that follows equation 4, can also be considered as a Itô's Process given by,

$$d(f(X_t, t)) = \frac{\partial f}{\partial t}(X_t, t)dt + \frac{\partial f}{\partial X_t}dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial X_t^2}dX_t^2 \quad (5)$$

where,  $dt^2$ ,  $dt dB_t$  are zero and  $dB_t^2 = dt$ . Using the same premise, we can describe a similar PDE for option pricing as well,

$$rC = \frac{\partial C}{\partial t}(S, t) + rS \frac{\partial C}{\partial S}dX_t + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}(S, t) \quad (6)$$

Here,  $C(S, t)$  represents a call option with time to expiry given by  $\tau$ , the strike price  $K$ , and the price of underlying  $S$  with a risk-free rate of  $r$ . This equation also is known as the Black-Scholes equation. Solving this PDE, with the following boundary conditions,

$$C(0, t) = 0, \forall t \quad (7)$$

$$C(S, t) \rightarrow S \text{ as } S \rightarrow \infty \quad (8)$$

$$C(S, T) = \max(0, S - K) \quad (9)$$

with the use of standard numerical methods such as the finite difference method, would yield the Black-Scholes Formula for a European Call Option.

$$C(S_t, t) = N(d_1)S_t - Ke^{-r\tau}N(d_2) \quad (10)$$

$$d_1 = \frac{\ln(\frac{S_t}{K}) + (r + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}} \quad (11)$$

$$d_2 = d_1 - \sigma\sqrt{\tau} \quad (12)$$

Here,  $C(S_t, t)$  represents the price of the option at time  $t$ ,  $r$  represents the risk-free rate,  $K$  represents the strike price,  $\tau$  represents the time to expiry,  $S_t$  represents the price of an underlying asset,  $\sigma$  is the standard deviation of logarithmic returns and  $N(x)$  is the cumulative normal distribution of  $x$ .

The standard deviation of logarithmic returns or  $\sigma$  is the volatility term. Hence the Black-Scholes Formula can be used to find the volatility if the Option Price is given, this type of volatility is termed implied volatility, mathematically we can represent the same with the help of the Black-Sholes Model,

$$\hat{\sigma}_{im} = \mathcal{M}_{BS}(\sigma|S, K, r, \tau, C(S_t, t)) \quad (13)$$

The following equation can only be solved by a root-finding methods such as Brent's Method as it has no analytic solution.

### B. Heston's Stochastic Volatility Model

The Heston Model was introduced in 1993 and made a significant correction to the most restrictive assumption of the Black-Scholes model, which is the assumption of constant volatility. The model considers volatility as a separate stochastic process. The first equation for the Heston model is similar to that Black Scholes formula for modeling stock price as a stochastic process. But the second equation for the Heston model, for the stochastic volatility, is derived from the mean-reverting CIR process, which is used for modeling interest rates and is given by,

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dB_t \quad (14)$$

where  $a$ ,  $b$  and  $\sigma$  are model parameters, where  $a$  represents the speed of reversion to the mean  $b$  and  $\sigma$  represents volatility, the first term in the CIR model is similar to the Vasicek model which reverts the interest rate to the mean  $b$  with a speed of  $a$ , the second term is responsible for avoiding negative interest rates given any value for  $a$  and  $b$ . The Heston Model uses the same intuition to model option pricing and can be represented with the two aforementioned stochastic processes as,

$$dS_t = rS_t dt + \sqrt{V_t}S_t dB_t^1 \quad (15)$$

$$dV_t = a(\hat{V} - V_t)dt + \eta\sqrt{V_t}dB_t^2 \quad (16)$$

$$dB_t^1 dB_t^2 = \rho dt \quad (17)$$

Here  $\hat{V}$  is the long term variance and  $V_t$  is the variance at time  $t$ , as equation 16 is analogous to equation 14,  $V_t$  reverts back to  $\hat{V}$  with a speed of  $a$ . The volatility of the variance process is given by  $\eta$  and similar to the CIR process is responsible for non-negative values for volatility. The third equation of the Heston Model explains the correlation between the stock price and the volatility stochastic process given by,  $\rho$ , it is also representative of the fact that as the stock price moves the volatility changes and higher movement in the stock price correspond to higher volatilities. Similarly, the Heston model can be explained by the following PDE,

$$\frac{\partial C}{\partial t} + \frac{\partial C}{\partial V}[a(\hat{V} - V_t) - \Delta(S, v, t)] + rS \frac{\partial C}{\partial S} + \frac{1}{2} \sigma^2 V \frac{\partial^2 C}{\partial V^2} + \rho \sigma V S \frac{\partial^2 C}{\partial S \partial V} + \frac{1}{2} S^2 V \frac{\partial^2 C}{\partial S^2} = rC \quad (18)$$

Here  $\Delta(S, v, t)$  represents the volatility risk, which is the risk of asset price movement due to a change in volatility of

the risk factors, this can also be measured with the options greek, Vega. Though there are other methods to price options with the Heston Model as well, most popular being numerical methods, such as finite difference and integration. The integration follows the Feynman-Kac Theorem, which shows that options can be discounted back if the option payoff function is known. As for this study, Heston's model is evaluated using Fourier Transform[18] and Monte Carlo Simulation provided by QuantLib's AnalyticHestonEngine.

### C. Rational Option Pricing Model

Designing an ANN option pricing model doesn't ensure the model itself follows the rules of rational option pricing. Hence it is essential to provide some empirical conditions that our model needs to satisfy, to be considered as a viable model. A similar approach was taken in [19] for a rational pricing model. The following are necessary conditions provided by the Theory Of Rational Option Pricing [16], [17].

**Assumption:** If  $C(S, K, \tau)$  is the price of the option at time  $\tau$ , where  $S$  is the price of the underlying and  $K$  is the Strike, then for a call option payoff at  $\tau = 0$  is represented as

$$C(S, K, 0) = \max[0, S - K] \quad (19)$$

**Condition 1:** Two call options  $C(S, K_1, \tau)$  and  $C(S, K_2, \tau)$  with different strikes, must follow,

$$C(S, K_1, \tau) \geq C(S, K_2, \tau) \text{ iff. } K_1 > K_2 \quad (20)$$

**Condition 2:** If  $C(S, K_1, \tau)$  is the option price then  $C$  is a convex function over the strike price  $K$ ,

$$C(S, K, \tau) = \int_0^\infty \max(0, S_T - K) f(S_T | S_t, \tau) dS_T \quad (21)$$

and the first derivative is given by,

$$\frac{\partial C}{\partial K} = \int_0^K f(S_T | S_t, \tau) dS_T - 1 \quad (22)$$

**Condition 3:** As observed in equation 21 the,  $f(S_T | S_t, \tau) dS_T$  term represents the CDF of  $\mathbb{P}(S_T \leq K)$  hence,

$$\frac{\partial C}{\partial K} \leq 0 \quad (23)$$

**Condition 4:** The double derivative wrt.  $K$ , is given by  $\frac{\partial^2 C}{\partial K^2} = f(S_T | S_t, \tau)$  is the PDF hence can never be less than 0.

$$\frac{\partial^2 C}{\partial K^2} \geq 0 \quad (24)$$

**Condition 5:** If time to maturity  $\tau$  is infinity then the value of the option is the same as the underlying price,

$$C(S, K, \infty) = S \quad (25)$$

**Condition 6:** If the Strike Price,  $K$  tends to infinity the option price is 0.

$$C(S, \infty, \tau) = 0 \quad (26)$$

**Condition 7:** The lower bound on an option price can be explained with  $k$  bonds of price  $B(\tau)$  at time  $\tau$  using the following relation,

$$C(S, K, \tau) = \max[0, S - kB(\tau)] \quad (27)$$

The following conditions hold good for both American Options and European Options. The premise behind these conditions is that the option price is a differentiable function wrt.  $K$ , intuitively the derivative with respect to time of maturity, will always be greater than zero as larger the value of  $\tau$  it is more likely that Strike price is less than the underlying. These conditions will ensure that the ANN-model is fundamentally sound under the theory of rational option pricing.

## IV. METHODOLOGY

### A. Proposed System

This paper proposes 2 separate architectures for the Black-Scholes Model and Heston Model as the input parameter and the methodology differs for both.

1) *Black-Scholes ANN:* The Black-Scholes ANN is straight-forward, after data cleaning and introducing new parameters to the dataset the model inputs the normalized underlying price( $S/K$ ), the strike price ( $K$ ), time to maturity ( $\tau$ ) and, and the risk-free rates( $r$ ) and train it against the Option Prices calculated through the Black-Scholes Formula. The formulated model was devised under the rationality conditions given in the previous sections as was tested on a subset of the options data. The rationality of the model can be proved theoretically, but for sake of simplicity, we provide the practical values for the same.

2) *Calibrated Heston ANN:* As described in the previous sections the Heston Model has more of an Optimization paradigm, which increases the computation time. Heston Model requires the fine-tuning of 5 different variables making it a 5-dimensional Optimization problem. The complexity of calibrating the Heston Model can be observed by solving the Heston's Equation using a root-finding method, `scipy.optimize.broyden1` takes 2 minutes to run on 40 entries on an i5 9th Gen Core. Other methods of solving include optimizing the following function,

$$[a, \hat{V}, \eta, \rho, V_0] = \underset{[a, \hat{V}, \eta, \rho, V_0] \in X}{\operatorname{argmin}} \sum_{i=1}^n \|\mathcal{H}[a, \hat{V}, \eta, \rho, V_0] - C_T^i\|^2 \quad (28)$$

But the method used for this paper is the Quantlib implementation of the Quasi-Analytic Heston Formula, which is

based on the Levenberg-Marquardt Least-squares method[20]. As it provides the fastest solution and the trade-off between accuracy and still viable compared to Monte Carlo Simulation or solving with Optimization.

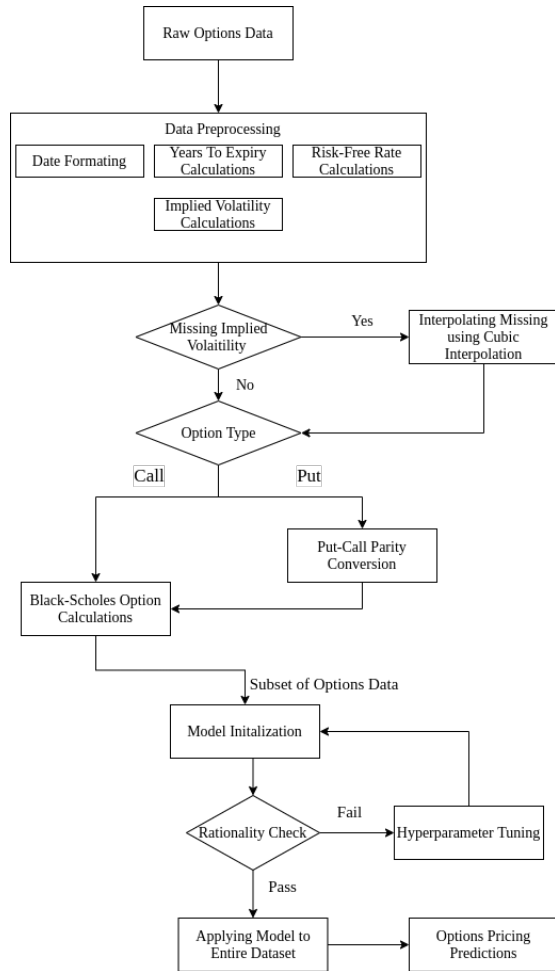


Fig. 1. Black-Scholes ANN Architecture

### B. Dataset Description and Preprocessing

The Options Data used in this study was collected from Bloomberg and CBOE. The time range of the AAPL options data is from 2012-2017 which has 1203257 option quotes. The data itself contains information on the Option such as date of expiry, underlying, strike, ask, bid, etc. To make use of the majority of the dataset we convert the Put options to Call Options using the Put-Call parity and compute some of the required columns such as years to expiry, risk-free rate, calculated by interpolating the durations(years) with synthetic rates for a 3-month bond. Next, the implied volatilities are computed using the mid-value as the option price, this approach yielded some NaN values, but later interpolated using the existing implied volatility values. For the sake of accurate predictions, we removed options with less than 10 days to expiry and deep-out of money options after performing all the above-mentioned steps the dataset had about 590240 entries.

For training the Black-Scholes ANN we split the dataset in an 80:20 ratio and for the Heston ANN, we split in a 60:40 ratio.

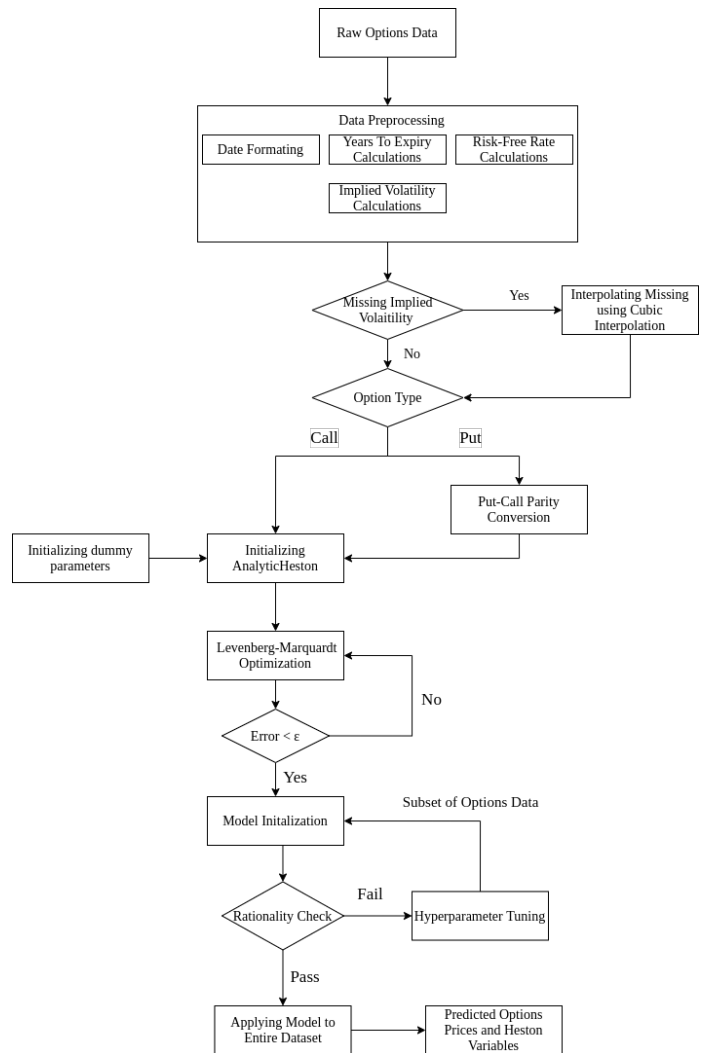


Fig. 2. HestonANN Architecture

Input Parameters	Symbol
Normalized Underlying Price	$S/K$
Time to Expiry	$\tau$
Risk-Free Rate	$r$
Implied Volatility	$\sigma$

TABLE I  
BLACK-SCHOLES ANN PARAMETERS

The output of the Black Scholes ANN is just the Option Price. For the Heston's ANN, the input parameters include  $X \in [a, \hat{V}, \eta, \rho, V_0]$  in extension with the input parameters to the Black-Scholes ANN. And the output for the Heston ANN has the Option Price and the array of parameters  $X$ , the new output is the optimized solution to the Heston's Formula as projected by the ANN. In the training process, the model learns, firstly to minimize the distance between the Heston's Price and the Mid price by tuning just the parameter of  $X$ , the

second part for the Heston ANN includes calculations of the Heston's Price given the input parameters  $X$  and the Black-Scholes inputs.

Input Parameters	Symbol
Normalized Underlying Price	$S/K$
Time to Expiry	$\tau$
Risk-Free Rate	$r$
Implied Volatility	$\sigma$
Correlation Coefficient	$\rho$
Volatility of Variance Process	$\eta$
Long Term Variance	$\hat{V}$
Mean Reversion Speed	$a$
Initial Variance	$V_0$

TABLE II  
HESTON ANN PARAMETERS

## V. RESULTS

The training process for both the models differed in many ways, the Heston ANN took a significantly large portion of time, over 3 hours with 100 epochs on an NVIDIA-GTX1650 whereas the Black-Scholes training just took over 10 minutes for 100 epochs even when the training sets differ in length significantly. For both the models, the rationality conditions were verified experimentally.

### A. Proof Of Rationality

The proof of rationality of the models is based on the work of Llagaris et al. [21], which provided a methodology for solving PDEs with a Neural Network, capitalizing on the gradient calculation using backpropagation. For the sake of simplicity  $\mathcal{N}_{HSV}$  is the neural network architecture for the Heston ANN and  $\mathcal{N}_{BS}$  be the neural network architecture for Black-Scholes ANN.

Hence  $C_{HSV}$  ie. the Heston's Option Price can be calculated as  $C_{HSV} = \mathcal{N}_{HSV}(x)$  where  $x$  is the input to the Heston ANN. The same goes for the Black Scholes ANN ie.  $C_{BS} = \mathcal{N}_{BS}(y)$  where  $y$  are the inputs to the Black-Scholes ANN.

To prove that both  $\mathcal{N}_{BS}$  and  $\mathcal{N}_{HSV}$  are rational models, we first test the models on options, just to check if they give viable results, ie. if they perform as intended. To show that validity of conditions 3 and 4, we just input the Strike price,  $K$  into our models, if our models just had a single neuron, the result due to the backpropagation(Autograd) for a single step would be  $\frac{\partial C}{\partial K}$ , this assumption can be extended to multiple layers and n-neurons in a layer. Hence we can easily compute the first and second derivatives for both models with respect to the strike. Similarly, we can extend this to time to maturity as well. The proof for the following conjecture has been beautifully theorized in [22].

As an extension, after getting the gradients, we can also solve for Black Scholes and Heston's PDE by considering the PDE itself and adding the boundary conditions given in equations 7,8,9,26 and consider the whole as a loss functions and minimizing it, but this approach is out of scope for this study.

Practically the gradients were obtained by using TensorFlow's function `tensorflow.gradients` for both the Heston and Black Scholes ANN and for both the Strike and time to maturity conditions.

Derivative	Max	Min	Average
$\frac{\partial C}{\partial K}$	-0.620087	-7.998365	-2.129823
$\frac{\partial^2 C}{\partial K^2}$	0.854321	0.004325	0.354675
$\frac{\partial C}{\partial \tau}$	85.65985	14.669981	39.655885

TABLE III  
BLACK-SCHOLES ANN DERIVATIVES

Derivative	Max	Min	Average
$\frac{\partial C}{\partial K}$	-0.956854	-13.255863	-1.125632
$\frac{\partial^2 C}{\partial K^2}$	0.965472	0.014258	0.458528
$\frac{\partial C}{\partial \tau}$	97.85759	26.55458	45.958555

TABLE IV  
HESTON ANN DERIVATIVES

### B. Black-Scholes ANN

After Rationality testing of the proposed Black-Scholes Model, it is important to validate the findings with actual options data. To train our model as mentioned AAPL Options Data was used. After 100 epochs of training, our model showed a promising result when applied to the training set and the testing set. But for actually validating the true potential of the proposed model, the same was applied to the current AAPL Options Chain. The Observed Error for the Black-Scholes ANN for both testing set and training set is given by: The

Set	MSE(%)	MAE(%)
Training Set	$4.67744 * 10^{-4}$	0.1584619
Testing Set	$3.704807 * 10^{-4}$	0.13295

TABLE V  
BLACK-SCHOLES ANN PARAMETERS

results obtained are significantly better than these following [15], [19],[23] studies even though the dataset used for this model is significantly smaller, this also reflects to the fact that a viable output is highly dependent on the model structure.

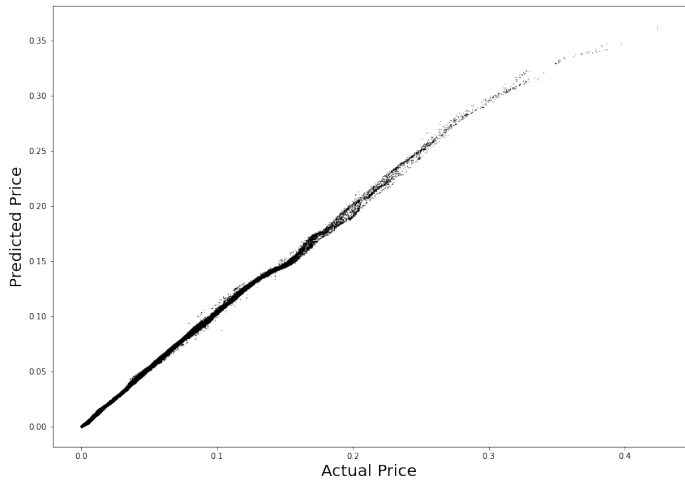


Fig. 3. Black-Scholes ANN on Testing Set  
The Error of the Black-Scholes ANN on the testing Set can be visualized as in figure 4.

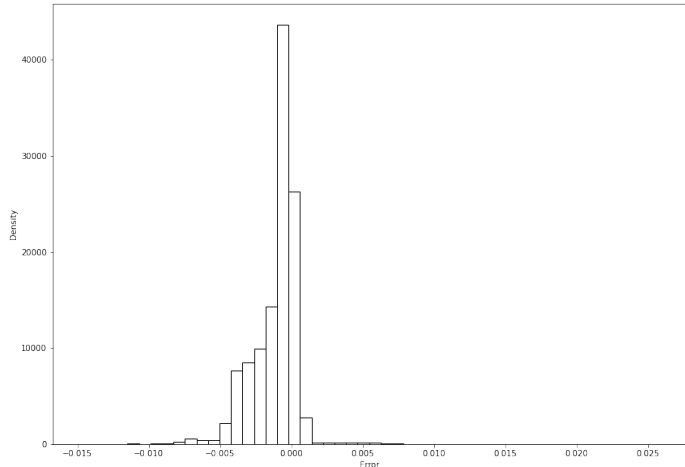


Fig. 4. Black-Scholes ANN Error on Testing Set  
And finally Black-Scholes ANN on current AAPL Options Quote Data as given by figure 5.

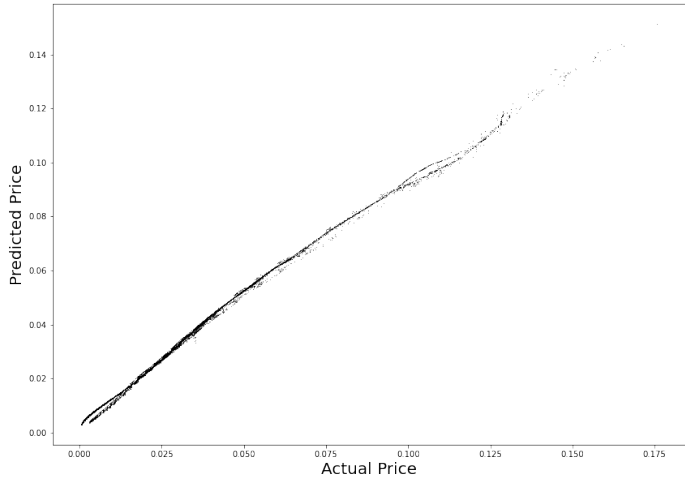


Fig. 5. Black-Scholes ANN on current AAPL Options Chain

C. Heston’s ANN

As training, the Heston Model is itself a very tedious process and requires a 2 step process, calibrating and then computing the option prices, the training sample was reduced to 60% of the entire dataset mainly due to computational time which exceeded 3 hours for the entire training set. The error rates even for the Heston ANN were comparable to other studies such as [14] and [24].

Set	MSE(%)	MAE(%)
Training Set	$1.18524 \times 10^{-3}$	0.291651
Testing Set	0.0131843	0.671406

TABLE VI  
BLACK-SCHOLES ANN PARAMETERS

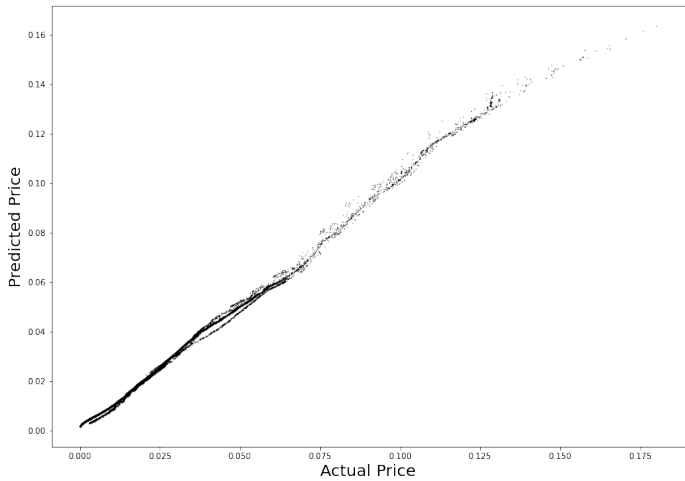


Fig. 6. Heston ANN on Testing Set

As observed, the Heston ANN doesn’t perform as well as the Black-Scholes Model this can be explained either by assuming that it is not always that Heston performs better than

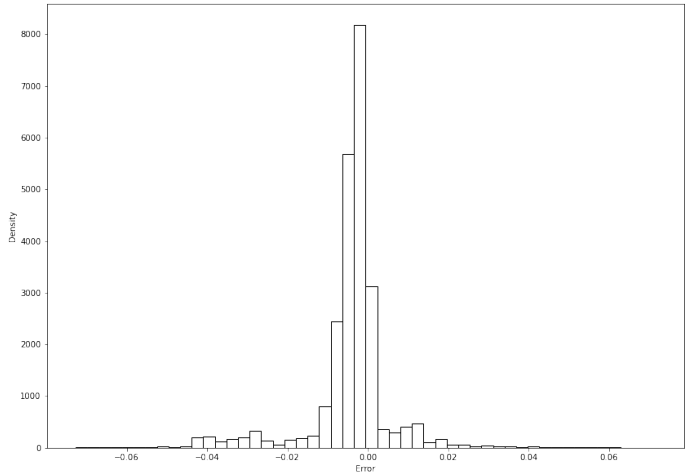


Fig. 7. Heston ANN Error on Testing Set

Black-Scholes or the other way round or attributing the latent complexities of both ANN models will never be able to exactly replicate a market. We can also make a strong comparison between the Black Scholes ANN model and the market by looking at Volatility Surface, this volatility is calculated using the inverse Black Scholes equation given by equation 13.

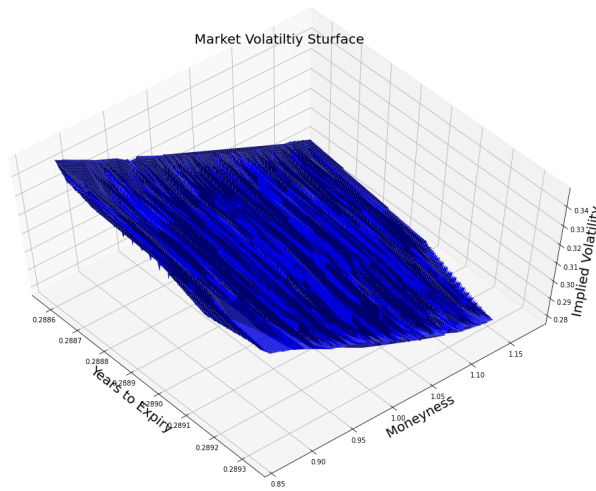


Fig. 8. Market Volatility Surface(Mid) for Expiry on 2013-01-19

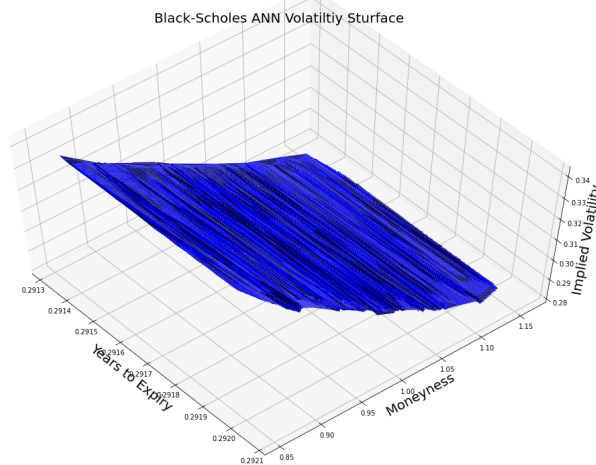


Fig. 9. Black Scholes Volatility Curve for Expiry on 2013-01-19

## VI. CONCLUSION

The aim of this study was to formulate deep learning models that replicate financial models, not just learning through historical data by also having an “intuition” of the theory of finance and rational option pricing as well. The results achieved by both models have a significant weight in this field of research, but there is always scope for a better and more precise model but the computational complexity will still be a resistance. But with the PDE solver by [21] the process of solving financial PDEs can be accelerated to an extent. As for future work, we shall look more into solving the entire PDE by a Neural Network eliminating the process of training and testing and reducing computation time significantly.

## REFERENCES

- [1] F. Black and M. Scholes, “The pricing of options and corporate liabilities,” *Journal of political economy*, vol. 81, no. 3, pp. 637–654, 1973.
- [2] G. Bakshi, C. Cao, and Z. Chen, “Empirical performance of alternative option pricing models,” *The Journal of finance*, vol. 52, no. 5, pp. 2003–2049, 1997.
- [3] C. J. Corrado and T. Su, “Implied volatility skews and stock index skewness and kurtosis implied by s&p 500 index option prices,” *Journal of Derivatives*, vol. 4, no. 4, pp. 8–19, 1997.
- [4] D. S. Bates, “Post-’87 crash fears in the s&p 500 futures option market,” *Journal of econometrics*, vol. 94, no. 1-2, pp. 181–238, 2000.
- [5] S. R. Das and R. K. Sundaram, “Of smiles and smirks: A term structure perspective,” *Journal of financial and quantitative analysis*, pp. 211–239, 1999.
- [6] S. L. Heston, “A closed-form solution for options with stochastic volatility with applications to bond and currency options,” *The review of financial studies*, vol. 6, no. 2, pp. 327–343, 1993.
- [7] S. G. Kou, “A jump-diffusion model for option pricing,” *Management science*, vol. 48, no. 8, pp. 1086–1101, 2002.
- [8] R. C. Merton, “Option pricing when underlying stock returns are discontinuous,” *Journal of financial economics*, vol. 3, no. 1-2, pp. 125–144, 1976.
- [9] A. Cartea and M. G. Figueroa, “Pricing in electricity markets: a mean reverting jump diffusion model with seasonality,” *Applied Mathematical Finance*, vol. 12, no. 4, pp. 313–335, 2005.
- [10] Z. Zhou and X. Gao, “Numerical methods for pricing american options with time-fractional pde models,” *Mathematical Problems in Engineering*, vol. 2016, 2016.
- [11] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.
- [12] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989.
- [13] K. Hornik, “Approximation capabilities of multilayer feedforward networks,” *Neural networks*, vol. 4, no. 2, pp. 251–257, 1991.
- [14] S. Liu, C. W. Oosterlee, and S. M. Bohte, “Pricing options and computing implied volatilities using neural networks,” *Risks*, vol. 7, no. 1, p. 16, 2019.
- [15] N. Gradojevic, R. Gençay, and D. Kukulj, “Option pricing with modular neural networks,” *IEEE transactions on neural networks*, vol. 20, no. 4, pp. 626–637, 2009.
- [16] R. C. Merton, “Theory of rational option pricing,” *The Bell Journal of economics and management science*, pp. 141–183, 1973.
- [17] H. Föllmer and A. Schied, *Stochastic finance: an introduction in discrete time*. Walter de Gruyter, 2011.
- [18] P. Carr and D. Madan, “Option valuation using the fast fourier transform,” *Journal of computational finance*, vol. 2, no. 4, pp. 61–73, 1999.
- [19] Y. Yang, Y. Zheng, and T. M. Hospedales, “Gated neural networks for option pricing: Rationality by design,” *arXiv preprint arXiv:1609.07472*, 2016.
- [20] D. W. Marquardt, “An algorithm for least-squares estimation of non-linear parameters,” *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.
- [21] I. E. Lagaris, A. Likas, and D. I. Fotiadis, “Artificial neural networks for solving ordinary and partial differential equations,” *IEEE transactions on neural networks*, vol. 9, no. 5, pp. 987–1000, 1998.
- [22] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.
- [23] C. Dugas, Y. Bengio, F. Bélisle, C. Nadeau, and R. Garcia, “Incorporating second-order functional knowledge for better option pricing,” in *Advances in neural information processing systems*, 2001, pp. 472–478.
- [24] O. Pironneau, “Calibration of heston model with keras,” 2019.