Comparison between traditional and modern option pricing models

Alexandru Tudor
University of Twente
PO Box 217, 7500 AE Enschede
the Netherlands
a.tudor@student.utwente.nl

ABSTRACT
Today's financial markets include incredibly sophisticated mechanisms and offer a large spectrum of financial instruments. Retail investors could previously only invest in simpler products such as equities or bonds. Nowadays, derivatives are available to the general public. The lack of education made some of these financial products shift from their original aim to speculative investments since they can provide a lot of leverage.
In this research paper, we want to take a deeper look into options, one of the most commonly traded derivatives, and make a comparison between the well-known option pricing models, such as the Black-Scholes or the Binomial model, to the AI-driven pricing models that are on the horizon.
The quantitative comparison of the models will be the basis of the thesis research. We intend to do so by coding a Python tool that can retrieve the required data and compute option prices using the Black-Scholes and Binomial models. Then we will have to conduct a literature study to find pre-computed option prices using the AI-driven models. The final step will be to assess how close each estimate came to the actual market price of the option and draw the results to the research question.

Keywords
Stock market, Investing, Options, Black-Scholes model, Binomial model, AI-driven option pricing models, Data-driven option pricing models, Traditional option pricing models, Modern option pricing models, Algorithmic pricing models.

1 INTRODUCTION

The return on our investments is a very important aspect of our financial well being. Making good investment decisions that can preserve or multiplicate money over time takes a lot of research and know-how. As a result, most people delegate the responsibility of investing their money to others. You can now entrust your money to professional institutions such as mutual funds, fixed-income funds, equities funds, or hedge funds for a tiny percentage charge. The major purpose of these funds is to generate as much money as possible from the investments they make in order to keep contracting new clients. In order to do so, these institutions invest in various securities such as stocks, bonds, and commodities. But when the investments they make don't go as planned, a key aspect is to make sure they are protected from the downside risk. They do so by using financial instruments such as options. Option prices vary on a day-to-day basis therefore it is critical for a wealth manager to know what a fair price for the option contract is. The price of the option fluctuates based on multiple variables. There are a few traditional pricing models that focus on a mathematical/statistical approach such as the Black-Scholes model [2] or the Binomial model [3]. These are widely known models that have proven to be effective over time such that financial institutions use them regularly. Although these models are complex and take into account a variety of criteria, they are not designed to consider qualitative information about the underlying asset, thus pricing may be inaccurate in some instances. Therefore, we want to compare their performance with the newly implemented AI-driven data-based models that are on the rise [4, 5, 6, 7, 8].

2 PROBLEM STATEMENT AND RESEARCH QUESTION

As of today, there is research done on the comparison between multiple mathematical option pricing models but not so much on the AI-driven models. The problem the statistical models are facing is that they only take the past performance of the stock into account without looking into how the actual business is doing nowadays and how it will perform in the future. Data-based models want to fill in this gap and create algorithms that have higher predicting power by using machine learning on available data to price options with this information. This paper will analyze whether this new technology is evolved enough to outperform traditional models, or whether there is still potential for improvement. In order to try and solve the problem stated above, we need to come up with the main research question.

RQ: Which option pricing model between the statistical (Black-Scholes and Binomial) and the modern AI-driven models are predicting more accurate market option prices?

TScIT 37, July 8, 2022, Enschede, The Netherlands

This research question generates several sub-questions, the answers to which can help us further define the solution and gather more information.

SQ1: What are the factors that influence the price of an option?

SQ2: Which of the Black-Scholes and Binomial models calculate closer option prices compared to the market prices?

SQ3: According to the literature, which data-driven model that uses AI to price options has the closest estimates to market prices?

SQ4: According to statistical measures, which model is more accurate at predicting market option prices?

Answering SQ1 provides crucial knowledge for this research. Without knowing exactly what determines an option's price we cannot judge option pricing models. To answer SQ2 we need to compare the prices calculated using the traditional models with the market prices of those options. This leads us to answer SQ3 which is critical in order to make the desired comparison between the traditional and modern option pricing models. The answer to the last sub-question wraps up the whole information.

3 METHODOLOGY

To perform this research we will need to follow a specific methodology. The model that calculates the closest prices to the actual market prices will be considered the best performer. We are going to measure this based on two statistical measures "RMSE" and "MAPE". Both measure the deviation of the calculated prices from the market prices. Step 1 of the research is gathering the data of traditional option pricing models. This will be done by creating a Python implementation of the models and using it to calculate option prices for highly liquid stocks. With the gathered data, we will compute the statistical measures. Step 2 is to find the same statistical measures for the modern option pricing models. We are going to do so by performing a systematic literature review as explained in part 6 of this paper. Step 3 of the research is to compare the statistical measures of both types of models and draw the conclusions.

4 BACKGROUND

Options are financial derivative instruments that have underlying securities such as stocks. In essence, an option is a contract between the seller (writer) and the buyer. There are two types of option contracts, call options and put options. A call option contract gives the buyer the right to buy one hundred shares of the underlying stock from the seller. These can be bought for a previously agreed price (strike price) until the date when the contract expires (maturity). A put option is very similar to the call option but instead of buying the one hundred shares, the buyer of the put option has the right to sell one hundred shares to the seller (writer). The benefit of being the writer of an option contract is that you receive the "premium" which is the price of the option. The premium is collected when the option is sold to a buyer, and it represents the maximum profit the seller can gain from the transaction [15]. In options trading, "to exercise" an option contract means to put into effect the right to buy or sell the underlying security that is specified in the options contract. There are a lot of different types of options but in this research paper, we will focus on the two most known types: European and American style options. The only difference between the two is that an American option can be exercised at any point in time while the European option can only be exercised at the expiry date. Although the owner of an option has the right to exercise it, in most cases, it is better to sell the option instead [16]. In order to comprehend why this is the case, we need to understand how options work. Therefore, we will take a look into the break-even analysis of options. Figure 1 represents the Break-Even graph of a call option. On the X-axis we have the Profit and Loss of the call buyer while the Y-axis displays the share price. Since it is a call option, the buyer profits if the underlying price is going up.

Figure 1. Call Break-Even

The initial P/L is -250$ which represents the premium paid for the option (2.5$ * 100). It is important to note that this is the maximum amount a call option holder could lose, no matter how low the share price falls. Once the share price goes past the strike price, the option starts to build value and hits the Break-Even point at 100$ per share, any further increase leads to a profit. We can observe that the Break-Even point for a call option is the strike price plus the premium paid.

$$\textit{Call B/E} = \textit{Strike price} + \textit{Premium} \tag{1}$$

The Break-Even graph of a put option is shown in Figure 2. The put buyer's profit and loss are displayed on the X-axis, while the share price is displayed on the Y-axis. Because it is a put option, the buyer will profit if the underlying price falls. The initial P/L is -250$ which represents the premium paid for the option (2.5$ * 100). It is vital to remember that no matter how high the stock price rallies, this is the maximum amount a put option holder can lose. Once the stock price falls below the strike price, the option begins to gain value and eventually reaches the Break-Even point at 97.5$ per share; any further decline in the share price results in profits. Therefore, the Break-Even point of a put option is calculated by subtracting the option premium from the strike price.

$$\textit{Put B/E} = \textit{Strike price} - \textit{Premium} \tag{2}$$

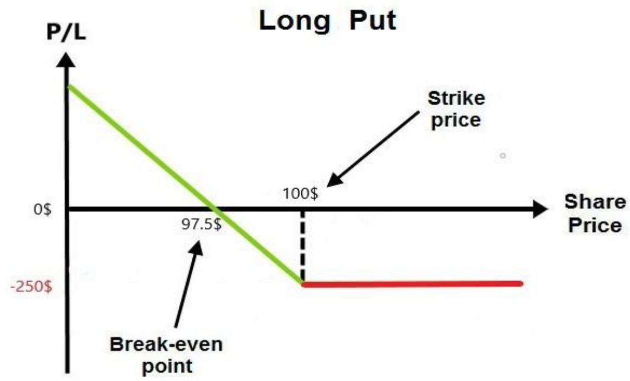

Figure 2. Put Break-Even

Table 1 shows the option chain for AAPL Calls expiring on June 3rd, 2022. The first column separates the options into three types: In-the-money (ITM), At-the-money (ATM), and Out-of-the-money (OTM). ITM options are the options that have a lower strike price than the underlying current price. ATM options have a strike price very close to the current trading price of the underlying while OTM options have a higher strike price than what the underlying is trading at. For put options, everything is the opposite since the buyer profits if the underlying price goes down.

Table 1. June 3rd, 2022 AAPL Calls[1]

AAPL Price : 145 $	***Strike Price***	***Option price***
ITM	125.00	22.35
	130.00	17.25
	135.00	13.47
	140.00	8.20
ATM	145.00	5.06
OTM	150.00	2.71
	155.00	1.27
	160.00	0.54
	165.00	0.24

We will take the ITM 140$ strike price AAPL Call option expiring on June 3rd, 2022, as an example for our illustration. Buying this call option would cost a total of 820$. Suppose that by June 3rd AAPL's price rises to 160$ (a 10.35% increase). Since it is the last day before the contract expires, we exercise the option. In this scenario, we turned the initial investment of 820$ into 2000$ (144% increase). The difference between the stock and options return perfectly shows the leverage effect that options provide. On the flip side, if by June 3rd the stock price would remain flat at 145$, the option would expire worthlessly, and the buyer would lose 820$. This is a -100% return compared to a 0% return holding the stock.

SQ1: What are the factors that influence the price of an option?

According to Hall (2000) [1] in his research paper "What you need to know about stock options.", the price of an option can be divided into two parts, Intrinsic and Extrinsic Value. One could buy the call option and immediately exercise it since the option is ITM, exercising the option would result in a positive value of 500$ (buying 100 AAPL shares at 140$ and immediately selling them at the market price). This is referred to as the option's Intrinsic Value. An OTM option has no Intrinsic Value since exercising the option after buying it would not derive any value. The rest (320$) is known as Extrinsic Value and it is derived from two main factors [2]. The days until expiration is one of them. The further the maturity date is, the bigger the extrinsic value is. This relationship is caused because a later maturity date implies more time for the option to become profitable. Therefore, a call option on AAPL with the same strike price but expiring on August 19th, 2022 is more expensive than the one in our example (8.20$ vs 14.15$[2]). The second factor that influences the Extrinsic Value of an option is Implied Volatility (IV). IV is the amount of movement the market is expecting for the specific underlying asset. Higher volatility implies higher option prices since the chances of the option becoming profitable are higher. Since volatility is a measure of movement rather than direction, no matter the type of option (call/put), the same rules apply [2]. Options are traded like stocks, on exchanges, so the price is subject to changes in offer and demand. In fact, this is what drives the implied volatility higher or lower. A market that believes a stock will rally in the upcoming days will rush to buy call options in order to profit from the upcoming movement. This creates an increase in demand and therefore an increase in option prices. Besides the time value and IV, interest rates and dividends may influence the Extrinsic Value of an option, however, these two do not change on a daily basis and the influence on the price is not so significant.

5 TRADITIONAL OPTION PRICING MODELS

Black-Scholes Model

The Black-Scholes model (BS) is one of the most important and fundamental ideas in the world of finance. It was developed in 1973 by Fischer Black and Myron Scholes [2] and still to this day, is one of the most commonly used mathematical equations for estimating the theoretical value of derivative financial instruments. The model uses the assumption that stock prices follow a "Random walk" [16], a theory that was well emphasized by Malkiel Burton in his book "A Random Walk Down Wall Street"[17]. The theory assumes that the stock prices must not follow any pattern in order for the market to be efficient. Otherwise, the stock price change could be predicted, and there would be a chance for financial gain. Besides this, the BS model also assumes that no dividends are paid until the maturity date of the option, the risk-free rate and volatility are known variables and remain constant, the option is European and there are no transaction fees associated with buying the contract. To calculate the price of an option with the Black-Scholes Model, five variables are required. The variables are the price and the volatility of the underlying asset, the time until maturity, the strike price, and the risk-free interest rates. In addition, the

[1] *https://finance.yahoo.com/quote/AAPL/options?p=AAPL&date=1660867200*

[2] *https://finance.yahoo.com/quote/AAPL/options?date=1660867200*

model predicts that asset prices will move in a geometric Brownian motion [18] with continuous drift and volatility. When applied to a stock option, the model takes into account the stock's constant price movement, the time value of money, the strike price of the option, and the period until the option expires. The mathematic equation goes as follows:

$$C = S * N(d_1) - K * e^{-r_f T} * N(d_2) \quad (3)$$

$$d_1 = \frac{\ln \frac{S}{K} + \left(r_f + \frac{\sigma^2}{2}\right) * T}{\sigma\sqrt{T}} \quad (4)$$

$$d_2 = \frac{\ln \frac{S}{K} + \left(r_f - \frac{\sigma^2}{2}\right) * T}{\sigma\sqrt{T}} \quad (5)$$

The formula for a put option is fairly similar, d_1 and d_2 are the same while the main equation is the following:

$$P = K * e^{-r_f T} * N(-d_2) - S * N(-d_1) \quad (6)$$

Where:
C = Price of the call option
P = Price of the put option
S = Current price of the underlying asset
N(d) = Cumulative normal probability density function
K = Strike price of the option
σ = The historical volatility of the underlying asset
T = Trading days until expiration of the option
r_f = Risk-free interest rate
e = Euler's number

Binomial Model

The Binomial Model was first proposed by William Sharpe in 1978 and one year later, Cox, Ross and Rubinstein formalized it and created the model that is still widely used for pricing options. Compared to other models, the Binomial model looks at the underlying price over a long period of time (using iterations) rather than focusing on a single point in time. can handle a variety of conditions and that is what sets it apart. To understand the limitations of this model better, we will take a look into the assumptions it uses. First of all, at any given moment, the price of the underlying asset can go in two directions, either up or down. The risk-free rate and the discount factor remain constant throughout the whole period. The model does not account for any transaction fees or costs, and it considers the investors indifferent to risk. The model starts with the process of building the so-called binomial tree (Appendix) which calculates the potential price of the asset over the option contract. The model begins from the current price of the underlying assets and assumes that for each period (step) of the tree, the asset price will either move up or down. The probability of the price moving up is denoted by "*p*" while the probability of it going down is "*1-p*" and it is calculated with the following formula:

$$p = \frac{e^{r_f * t/n} - d}{u - d} \quad (7)$$

$$u = e^{\sigma * \sqrt{t/n}} \quad (8)$$

$$d = e^{-\sigma * \sqrt{t/n}} \quad (9)$$

For each step of the binomial tree, the price is multiplied by the value of *u* or *d* as seen in this figure. After the binomial tree has prices computed for the underlying asset it is time to calculate the intrinsic value of the option at each node of the tree. The intrinsic value can either be positive or 0 because for a negative intrinsic value someone would simply not exercise the option. We do so by using the following formula Max[(Sn − K), 0], for a call option or Max[(K − Sn), 0], for a put option. In this manner, we calculate the intrinsic value of each option at each node. Now that we have the value of an option at each point in time (each node), the last step is to discount back to the present value. Starting from the last pairs of nodes, we multiply the specific intrinsic value with the probability associated with the stock price going in that direction and discount it with the risk-free rate as follows:

$$Option = \frac{p * OptionUp + (1 - p) * OptionDown}{(1 + r_f)^{T/n}} \quad (10)$$

In Figure 3 we have an example of a Binomial tree with actual numbers. The stock is trading at 100$, the strike price is at-the-money, and the up, down, and probability are given. The purple cells represent the stock price while the yellow ones show the option price at each node. Starting from the final node we discount each value and arrive at the present value of the contract 8.52$.

STOCK	ABC	u	115%
PRICE	$100.00	d	90%
STRIKE	$100.00	p	40%
			$132.25
			$32.25
		$115.00	
		$17.18	
Stock	$100.00		$108.00
Option	$8.52		$8.00
		$90.00	
		$3.11	
			$81.00
			$0.00

Figure 3. Example Binomial Tree

SQ2: Which of the Black-Scholes and Binomial models calculate closer option prices compared to the market prices?

Now that we have a clear understanding of the traditional option pricing model, the next step is to find out how we can obtain the best performance out of them. The book "Principles of corporate finance" by Brealey, et. al. (2018) [15] outlines which strategy should be used depending on the option's type. The Black-Scholes formula is mostly suited for European-style options. This is because the formula does not allow for early exercise but rather focuses on a static point in time. The Binomial model fits best the American style option since it accounts for early exercise In the book, it is demonstrated that it is never advantageous to exercise American calls early, on stocks that do not pay dividends. This is because the option carries extrinsic value until the expiration date and an investor would prefer selling the option instead of exercising it. Therefore, it can be considered an European option and it can be valued with the Black-Scholes formula. To answer SQ2, each model performs better in certain scenarios therefore we will create a Python tool for both of them. We will use the Binomial model to price American styled options and the Black-Scholes model for pricing European options.

6 MODERN OPTION PRICING MODELS

In order to respond to the third sub-question, we will conduct a systematic literature review that will enable us to comprehend the nature of modern option pricing models, how they operate, and how efficient they are. The protocol used to find the best literature reviews was PRISMA. We started from a high number of research papers and trimmed the numbers down by adding filters until the remaining articles were fully relevant. For the research we used an all-subject database (Scopus), a finance specialized database (SSRN), and a full text database (Google Scholar). We first used SSRN for finding a literature paper on the subject. We used the search phrase: "option pricing" AND "neural networks" and received 35 results. We then searched for the word "literature" within the papers and the database returned 4 results. The most relevant one was the work of Ruf, & Wang (2019) [5] which proved to be an amazing match for our study since it summarized over one hundred papers related to the topic. Since this study has covered so many other papers, we used a lot of the references to form knowledge. We then used Google Scholar because we needed in text searches to find keywords such as "RMSE" that were crucial in the comparison. The first phrase used was: "option pricing" AND "machine learning" AND "rmse"→ 855 results. We added the year interval 2018-2022 since we want the comparison the be as relevant as possible and trimmed down to 416 results. To shrink the results even more, we added "moneyness" which assures that the comparison is done on different strike prices (ITM, ATM, OTM). The most relevant article was the one written by Ivașcu. (2021) [24]. The second search phrase was: "option pricing" AND "neural networks" AND "mape" → 465 results. After adjusting for the time period there were 127 results. We added "moneyness" to the search and shrinked the results to 46. Here we found the important articles by Cao, Liu, & Zhai. (2021) [25] and Jang, & Lee. (2019) [26]. We then used Scopus with the following search phrase: TITLE-ABS-KEY (option AND pricing AND machine AND learning) → 91 results. We used the same time interval and restricted to 56 results. Later on, we added the word "rmse" which is critical for our comparison. This limited the results to a single article by Gaspar, Lopes, & Sequeira (2020) [22] which turned out to be very useful.

The modern option pricing models are represented by AI-driven algorithms that use machine learning to price the value of options. The staple of our literature research is going to be the research of Ruf & Wang (2019) [5]. This publication is very important since the authors read over one hundred research papers on the topic of AI-driven option pricing models and compiled the whole information into a single study. Artificial intelligence (AI) has entered a new epoch. While the concept of robots capable of displaying human levels of intellect was first implemented in the 1950s, progress in constructing AI machines was limited, and AI was mostly seen as a failed venture. There has been a renaissance of AI in recent years, and based on the available studies, we will understand how it is applied in finance. There are multiple different branches of Artificial Intelligence but in option pricing, most models are based on building machine learning on Artificial Neural Networks (ANNs). ANNs attempt to mathematically replicate the way the human brain functions by creating a collection of connected nodes also known as artificial neurons. These neurons can communicate with each other through connections like the biological brain does through synapses. Besides the ethical reasons, human brains cannot be replicated because of the technological limitation we face nowadays [13], however, these models replicate a scaled-down version of a brain. As explained by Culkin, & Das (2017) [4], ANNs receive a wide range of stimuli as input and then this information is parsed through layers of neurons that learn to associate the input with output by experience. In the study, this process is associated with how children learn that touching a hot

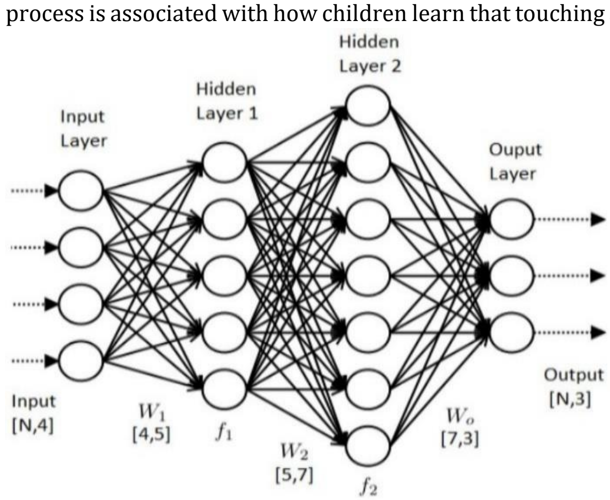

Figure 4. Artificial Neural Network[3]

stove causes pain and quickly learn not to go near one. Figure 4 illustrates how an artificial neural network looks like. First, we observe the Input Layer which represents the data introduced into the algorithm. In between the input and output, there is a hidden layer. Each hidden layer can have one or more neurons that are interconnected with each other. No matter how many nodes are in one layer, all of them are interconnected with the nodes in previous layers. Neurons receive as input, the output produced by the previous layer. For the model to work, a training data set is provided, and the model enters learning mode. Neurons apply random weights to the input and an activation function and then pass the output to the next layer. The activation function of a node in an artificial neural network determines the output of that node given an input. The result is compared to the desired output and an error is calculated. The

[3] *https://www.datasciencecentral.com/the-artificial-neural-networks-handbook-part-1/*

model then calibrates the weights in order to minimize the error. As the studies show, the number of inputs is a choice, as well as what the inputs are. Since over one hundred research papers were read, each of them having an unique ANN model, the inputs and outputs are different. However, most of the outputs are designed to calculate the option price. As for the inputs, models take a wide range of features into consideration such as strike price, stock price, interest rate, option greeks [19], volatility from calibration, historical volatility, implied volatility, GARCH-generated volatility, Kalman filter volatility, time to maturity, and, macroeconomic variables that influence volatility. The two indispensable variables fed into ANNs are the stock price and the strike price. Some models feed this data as separate variables while others use a ratio of the two, also known as moneyness. This reduces the number of inputs which makes the training of the ANN easier. Another very important feature is the volatility, which is calculated in various ways throughout the studies. Some of the studies even compare the differences between volatilities. Blynski and Faseruk (2006). [20] discovered that the used ANN has better results when using historical volatility rather than implied volatility. Another study by Y.-H. Wang (2009). [21] argues that GARCH volatility is the most efficient way to find option prices. Artificial Neural Networks present a lot of advantages when it comes to option price estimations as explained by G. Cybenko in his study [23]. The main advantage is that these models are universal approximators. Any continuous function, even ones with non-linear characteristics, can be fitted by the models. This property is demonstrated to be a product of the general architecture and training process. These proofs are extremely important because they demonstrate that any sufficiently big network may arbitrarily well approximate a given function. Only architectural choices and data quality are to blame for the constraints. However, according to Gaspar, Lopes, & Sequeira (2020) [22] several drawbacks may exist, such as the requirement for a complete and extensive dataset of historical data to train the modeling framework. This means that more complex options, particularly options not traded on public markets, cannot be priced as fairly by the ANNs as compared to other derivatives that are more widely available and have a bigger trading volume and data available. As a result, when using ANNs, we need to keep in mind that the model is calibrated to past data, which means that any changes in the future status of financial markets, such as a major financial crisis, could change the values of options, resulting in price miscalculations and the need to retrain models. Another drawback is that the models have almost no explanatory power. It is not known how much each of the attributes in the model contributes to the final result. Moreover, an attribute could bring zero contribution to the final result and the user will not know this since no explanation is provided. This is different compared to normal regressions where one can make significance tests on different parameters to see how much they contribute to the final result.

7 COMPARISON

The first step towards the comparison is to prepare the dataset for the traditional option pricing models. We created a Python tool that replicates the Black-Scholes and Binomial models. To retrieve the required data for these models, we used the Investpy API which retrieves stored historical data for publicly traded companies in the US and around the world. We used other libraries such as NumPy and Pandas to be able to manoeuvre the data and do complex calculations. Next up, we created the algorithm that calculates the prices based on the Black-Scholes and Binomial model. To get the most optimal results we chose a wide range of stocks as underlying for the options (25 European and 25 American stocks). As explained in part 5 of this study, the most efficient way to estimate option prices was to use the Binomial model for the American style options and the Black-Scholes for the European style options. The list of stocks consists of stocks with high market capitalization, high liquidity, and high trading volume since these are requirements for efficient market option prices. We created three separate data sets of option prices based on the strike price (ITM, ATM, and OTM options). The ITM and OTM options had strike prices approximately 15% outside the live stock price. For each data set, we calculated prices for call and put options on each stock. Moreover, we computed the prices for multiple expiration dates (one month, three months, six months, and one year). Therefore, for each stock, we computed a total of twenty-four different option prices resulting in a total of 1200 different option prices in all three data sets. We used Yahoo Finance[4] and the Euronext[5] website to search for live market prices. The gathering of data was a very intensive and time-consuming process because we chose to compute the prices while the markets are closed so that there are no fluctuations in prices. Hence, the whole set of data had to be calculated and retrieved from the online sources within a weekend which was very challenging but in the end rewarding. Before analyzing the results, we would like to point out some factors that could negatively influence the performance of the models. One important parameter was the historical data which was used to compute the historical volatility. For each option price, we used historical data starting from the 1st of January 2018. The reasoning behind this choice was that the last couple of years had a lot of rare events such as the pandemic and the Ukraine invasion. These resulted in a very high volatility in the stock market, so we chose a larger period in order to capture periods of the market where these events were not around. This decision turned out to be effective for the options with more distant expiration dates since the volatility is not expected to stay at high levels for periods as long as six months or one year. However, considering the recent big moves in the markets caused by rising inflation and possible future monetary policies changes, the markets are pricing in a higher volatility than the average for the last four years. This can lead to a slight undervaluation of close to expiration options by the traditional models. The last bias in our methodology might be the way we choose the market price for options. Our approach was to take the last price (meaning the

[4] *https://finance.yahoo.com/*

[5] *https://live.euronext.com/en/products/stock-options/list*

last transaction) as the market option price. However, in some cases, the bid-ask spread was completely different than the last transaction price. Often this is due to a big price change in the underlying that has not yet been corrected in the option market. In these cases, the correct approach in our opinion was to take the average between the bid and ask prices which is the most probable future settlement price. The first step after computing option prices and retrieving market option prices from the web was to calculate the difference between these two. The stocks analyzed had different prices some cost only a few dollars while others go into the thousands. This means that the option prices are not relative to each other and performing statistics on the raw difference is not relevant. Therefore, we divided the differences to the market prices to have a relative set of data and we called it an "error". Since the error could be either a positive or negative number, the best statistic measure to use was the root mean squared error (RMSE). This measure is the obvious choice in our case because we do not account for the direction of the error, we only look at its magnitude. Table 2, presented below, displays the RMSE values. These are categorized by moneyness (ITM, ATM, OTM), by option type (call or put) and by the option style (American [A] or European [E]). Each of these values were calculated from individual datasets of one hundred values each.

Table 2. Traditional Models RMSE by Option Type, Style and Moneyness.

ITM				ATM				OTM			
A CALL	A PUT	E CALL	E PUT	A CALL	A PUT	E CALL	E PUT	A CALL	A PUT	E CALL	E PUT
0.0792	0.0861	0.2608	0.1530	0.1308	0.1060	0.4500	0.5103	0.4984	0.3041	0.8595	0.3698

The ATM American calls and puts had an RMSE of 0.130 and 0.106 which classify the Binomial as a good predictor for option prices. For the European options, the error was higher (0.450 and 0.510) but still at reasonable values. The negative impact on the RMSE was made by the one-month until expiration options that were mispriced by the tool. In the case of Vivendi (Ticker: VIV), one of the options was mispriced with a value four times higher than the market price. Removing this outlier from the data would lower the RMSE score for put European options from 0.510 to 0.313. Although the improvement would have been massive, we preferred to keep the data raw. The traditional models had the best predicting power for In-the-Money options. The Binomial model had an RMSE of 0.079 and 0.086 for American calls and puts while the Black-Scholes model averaged an RMSE of 0.261 and 0.153 respectively. We had the pleasure to see calculated prices matching exactly the market prices in some cases. For the AXA company (Ticker: CS) two out of four options had an error of zero since the compared values were exactly equal. These values are very promising and show how powerful the traditional option pricing models are in optimal situations. On the other side of the spectrum, computing OTM option prices are the weak spot for the traditional models. The Black-Scholes model averaged an RMSE of 0.860 for call options which is the worse score in the whole dataset. This mainly resulted from overpricing close to maturity call options. As an example, Casino Guichard Perrachon (Ticker: CO) had a one-month OTM call price of 0.02 € while the calculated price was 0.08 €. Although the nominal difference is small (0.06 €) the difference to market price ratio is 3 which is a high value that negatively influences the overall RMSE. For put options the RMSE was 0.370, lower than for call options but still not an amazing result. For American options the narrative was similar, OTM call options averaged an RMSE of 0.498 while the put options had a slightly lower value of 0.304. The close expiration options had a negative influence on the score. For comparison, we chose the values from the work done by Gaspar, Lopes, & Sequeira (2020) [22]. In the research, they look into two ANN models. The values in Table 3 represent an average RMSE between the two ANN models and the RMSE resulted from our data. The first thing we can observe is that the ANN used performs the best with OTM options which is the exact opposite with the traditional models. The RMSE of ANNs is 0.1716, a clearly better score than the BS and Binomial model (0.6146 and 0.4012). For At-the-Money options the score is tight. ANNs average an RMSE of 0.1954. The score is beaten by the Binomial model (0.1184) but the Black-Scholes model has a higher error (0.4789). The ITM options is where the ANNs performed the worst, and the traditional models performed the best. The ANNs averaged an RMSE of 0.3804 way higher than both the Black-Scholes model (0.2069) and the Binomial model (0.0827). Ivașcu (2021) [24] performed a similar study looking into multiple option pricing models in his research paper "Option pricing using machine learning". The Artificial Neural Network tested in his study showed the same pattern as seen before. The ITM option pricing performed the worst with an RMSE of 0.59. The ATM options were priced slightly better with an RMSE score of 0.43 while the OTM options had a score of 0.13.

Table 3. RMSE on moneyness.

ITM			ATM			OTM		
BS	Binomial	ANN	BS	Binomial	ANN	BS	Binomial	ANN
0.2069	0.0827	0.3804	0.4789	0.1184	0.1954	0.6146	0.4012	0.1716

Since some studies measured the accuracy of ANNs with MAPE (mean absolute percentage error), we computed the MAPE values for our dataset. This measure is very similar to RMSE but instead of squaring the error to get rid of negative values, it takes the absolute value of the error so that it accounts for the magnitude of the error, not the direction. In the work of Jang, & Lee (2019) [26] the authors test the Bayesian neural network, and the improved version named the Generative Bayesian neural network. The improved version outperforms in pricing all types of options. The worst MAPE score was 0.2391 when pricing OTM options while the best was 0.0459 for pricing ITM options. Pricing At-the-Money options proved to be in the middle with a score of 0.1142. This research shows that there is a lot of room for improvement in the modern option pricing models since there is no exact theory behind and a lot of optimisations can be made. In Table 4 we can observe the values of the BS and Binomial models and also the results of two AI-driven models (deep neural network and Andreou et al. Neural Network) that were tested in the study paper of Cao, Liu, & Zhai (2021) [25]. According to the authors, these are one of the best performing neural networks for pricing options. The results were not what we expected considering the previous results. The modern pricing models performed the worse for OTM options, even worse than the traditional models. For pricing ATM options the results were close, the Binomial model achieved the highest score (5.75%). Both NNs had under 9% scores while the Black-Scholes model had a score of 12.35%. Lastly, the ITM options were priced better by the AI models, both scoring under 7% MAPE while the traditional models averaged 18.9%. In this research, the authors are constructing a brand new "economically meaningful" ANN model that combines the best out of the traditional and modern option pricing models. Hence, the model is called hybrid gated neural network (hGNN). The model uses the best neural architecture while considering for traditional no-arbitrage constraints that formed the base of option pricing in the past. Morover, the model has a separate neural network just for predicting the volatiliy that uses the Black-Scholes implied volatilites instead of realized volatilites which significantly improves option pricing accuracy. The results of the hybrid model are outstanding with no MAPE values higher than 0.02 and values as low as 0.0063. These values prove that the model outperforms any type of option pricing model seen before. This highlights that the most efficient option pricing model, is not one of the various studied models but a combination between the traditional and modern models.

Table 4. MAPE on moneyness.

ITM				ATM				OTM			
BS	Binomial	dNN	AnNN	BS	Binomial	dNN	AnNN	BS	Binomial	dNN	AnNN
0.2877	0.0906	0.0579	0.0691	0.1235	0.0575	0.0644	0.082	0.3799	0.2735	0.2714	0.4246

8 RECOMMENDATIONS

We developed a Python program that uses the Black-Scholes and Binomial models to calculate option prices after examining the option pricing theory and comprehending the variables that affect option prices. We simultaneously conducted research on contemporary pricing models represented by artificial neural networks and chose the most effective models for comparison. The first important aspect that we found is that the calibration of the models is often more important than the model type. A poorly designed ANN model can perform very well if the training data is adequate. At the same time, a traditional model that normally performs very well can show no predictive power if the inputs (historical data, risk-free rate) are subpar. The next important aspect is that ANNs have no exact theory or formula behind them, as the traditional models do. Because the ideal number of layers and nodes is unknown, there are many different versions of neural networks, each having advantages and disadvantages. However, the final results of our research paper show that even without an exact theory behind ANNs, they still provide incredible results. To answer the main research question, we will summarize the findings. The traditional models have performed very well when pricing ITM options and started to lose accuracy as the strike price went further Out-of-the-Money. The ANNs had different results based on the architecture of the models. Overall, the modern models proved to have real potential in pricing options, beating the traditional models in some of the cases. The best model out there seems to be the hybrid gated neural network (hGNN) developed by Cao, Liu & Zhai [25]. The model mixes classic and modern models, bringing the best of both worlds together to create an incredible model. This proves that the Artificial Intelligence involved brings massive improvements, but the fine tuning is done by the old models.

9 LIMITATIONS

There were several restrictions because this research study was completed in a constrained amount of time. The assessment of the risk-free rate for European stocks was one of them. The yield on treasury bonds for the eurozone was used as the risk-free rate, which was set at 2%. One could argue that using the risk-free rate of the country in which the company has the largest operating activity would have been the most accurate way to carry out the data collection. Even more exact would be a weighted average of risk-free yields depending on the revenue from each nation the company operates in. However, this would have required much more investigation outside the subject of study. Another limitation we faced was finding an open-source ANN software that we could use to gather data on the modern models. We were unable to build an ANN from scratch because to time constraints, but we were able to solve this issue by conducting a literature review, which gave us the information we needed.

10 FUTURE WORK

We covered a lot during this research and achieved very great results. However, there are ideas that can be further worked upon. An interesting approach would be using an ANN on the same exact data as for the traditional models. This could not be done here due to the lack of time. Working on merging both conventional and contemporary models is also required in order to advance option pricing models even further. More effort in this direction will result in new, more effective models.

REFERENCES

[1] Hall, B. J. (2000). What you need to know about stock options. Harvard Business Review, 78(2), 121-121.

[2] Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. Journal of Political Economy, 81(3), 637–654.

[3] Conroy, R. M. (2009). Binomial option pricing.

[4] Culkin, R., & Das, S. R. (2017). Machine learning in finance: the case of deep learning for option pricing. Journal of Investment Management, 15(4), 92-100.

[5] Ruf, J., & Wang, W. (2019). Neural networks for option pricing and hedging: a literature review. arXiv preprint arXiv:1911.05620.

[6] Ferguson, R., & Green, A. (2018). Deeply learning derivatives. arXiv preprint arXiv:1809.02233.

[7] Palmer, S. (2019). Evolutionary algorithms and computational methods for derivatives pricing (Doctoral dissertation, UCL (University College London)).

[8] Hirsa, A., Karatas, T., & Oskoui, A. (2019). Supervised deep neural networks (DNNs) for pricing/calibration of vanilla/exotic options under various different processes. arXiv preprint arXiv:1902.05810.

[9] Malliarakis, N. (2020). Profitability comparison between option pricing models for American stock options after Subprime crisis (Doctoral dissertation, Theseus).

[10] Dar, A. A., & Anuradha, N. (2018). Comparison: binomial model and Black Scholes model. Quantitative finance and Economics, 2(1), 230-245.

[11] Lee, C. F., Chen, Y., & Lee, J. (2016). Alternative methods to derive option pricing models: review and comparison. Review of Quantitative Finance and Accounting, 47(2), 417-451

[12] Jang, H., & Lee, J. (2019). Generative Bayesian neural network model for risk-neutral pricing of American index options. Quantitative Finance, 19(4), 587-603.

[13] Hahn, J. T. (2013). Option pricing using artificial neural networks: an Australian perspective. Bond University.

[14] Hull, J., Treepongkaruna, S., Colwell, D., Heaney, R., & Pitt, D. (2013). Fundamentals of futures and options markets. Pearson Higher Education AU.

[15] Brealey, R. A., Myers, S. C., Allen, F., & Mohanty, P. (2020). Principles of corporate finance. Tata McGraw-Hill Education.

[16] Fama, E. F. (1965). Random Walks in Stock Market Prices. Financial Analysts Journal, 21(5), 55–59.

[17] Malkiel, Burton Gordon. (2003). A random walk down Wall Street : the time-tested strategy for successful investing. New York :W.W. Norton.

[18] Osborne, M. F. M. (1959). Brownian Motion in the Stock Market. Operations Research, 7(2), 145–173.

[19] Paunović, J. (2014). Options, Greeks, and risk management. Singidunum Journal of Applied Sciences, 11(1), 74-83.

[20] L. Blynski and A. Faseruk. Comparison of the effectiveness of option price forecasting: Black–Scholes vs. simple and hybrid neural networks. Journal of Financial Management & Analysis, 19(2):46–58, 2006.

[21] Y.-H. Wang. Nonlinear neural network forecasting model for stock index option price: hybrid GJR–GARCH approach. Expert Systems with Applications, 36(1):564–570, 2009a.

[22] Gaspar, R. M., Lopes, S. D., & Sequeira, B. (2020). Neural network pricing of american put options. Risks, 8(3), 73.

[23] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. Mathematics of control, signals and systems, 2(4), 303-314.

[24] Ivașcu, C. F. (2021). Option pricing using machine learning. Expert Systems with Applications, 163, 113799.

[25] Cao, Y., Liu, X., & Zhai, J. (2021). Option valuation under no-arbitrage constraints with neural networks. European Journal of Operational Research, 293(1), 361-374.

[26] Jang, H., & Lee, J. (2019). Generative Bayesian neural network model for risk-neutral pricing of American index options. Quantitative Finance, 19(4), 587-603.

APPENDIX

A.1 Binomial Tree