Automating Index Estimation for Efficient Options Trading Using Artificial Intelligence



Vivek Shukla, Rohit Sharma, and Raghuraj Singh

1 Introduction

In options trading, the contract's value depends on the underlying (stock or commodity) value. The present discussion is centred on options of stocks. There are two types of option trades: call option and put option. In the call option, buyer of the contract can exercise the right or drop the purchasing of stocks on the maturity date at a fixed rate. On the other side, there is put option. In the put option, the buyer of the contract can exercise the right to sell the stocks at a fixed rate or drop the contract. The other party involved is called writer or seller of contract. Contracts are usually made of different lot sizes that may vary from few hundred to thousand. For example, in a call option for some "ABC" share, the buyer may purchase the right of buying 100 shares at a rate of 30\$ on 15 January 2020 by paying the cost of 700\$. On maturity date, the buyer has two options: in the first option, the buyer drops the idea of purchasing because shares of underlying stock are trading at lower than 30\$. Therefore, in this case, buyer loses the money (700\$). In the second option, buyer may exercise the right to purchase the shares at a fixed rate because the market price of underlying stock is trading higher than 30\$ say 39\$. Therefore, buyer makes the profit 100 * (39 - 30) - 700 = 200\$.

In options trading [23], any contract may be a combination of various trades. Some may be call option, others may be put option. For example, there are multiple strategies like bull call spread, bear spread, etc. In these strategies, one might buy a call option and sell another call option of the same underlying asset. Many features

V. Shukla (🖂) · R. Sharma

R. Singh

Department of Computer Science & Engineering, Harcourt Butler Technical University, Kanpur, UP, India

Department of Computer Science & Engineering, Dr. Ambedkar Institute of Technology for Handicapped U.P., Kanpur, UP, India

[©] The Author(s), under exclusive license to Springer Nature Switzerland AG 2022 F. Al-Turjman et al. (eds.), *Transforming Management with AI, Big-Data, and IoT*, https://doi.org/10.1007/978-3-030-86749-2_13

of option trading [24] attract researchers to develop an efficient system that draws conclusions about the movement of index or stock so that more and more accurate decisions might be made for trading. The use of AI in finance is growing rapidly, and the recent years saw many investors begin to use it in a variety of ways, though it is unclear how many of them are using it specifically for options trading. Advances in artificial intelligence (AI) have a more significant impact even in medicine and bolster the power of medical interventions and diagnostic tools as used by authors in [20]. In the financial community, Big Data is making more of a splash than a ripple. Technology is advancing at an alarming pace, with far-reaching repercussions. Industry operations are being transformed by increasing complexity and data generation, and the finance industry is no exception. AI can efficiently sift through large quantities of Big Data to generate data predictions and cost-effective energy optimization solutions to fuel Smart City technologies such as the one used by authors in [21].

2 Constraint Programming and the CLP Scheme

Complex problems can be represented concisely in constraint solving because the properties can be represented implicitly without bindings to other variables. Constraint logic programming (CLP) provides a foundation for programming language class by combining constraint-solving and logic programming concepts. CLP(R) language is implementation of constraint logic programming approach in domain of real numbers. For example, the constraint $X \ge 1$, X < X + 2 is always solvable in CLP(R) [1, 17].

A CLP(R) clause has the form as

$$p(...):-p_1, p_2,..., p_n$$

where $n \ge 0$ and p_i represents constraints or predicates.

Constraints can have many relations $(\geq, \leq, >, =)$. CLP(R) has natural inference mechanism and declarative power of logic programming [18].

2.1 The Binomial Option Pricing Formula

In the model of binomial pricing, value is calculated for the underlying stock. In this analysis, it is assumed that the underlying stock's value can either move up or move down at the end of one-time unit. According to this assumption, value of underlying stock can be formulated as

$$C = \max \left(S \Delta + P, \ S - Q \right)$$

provided

$$uS\Delta + rP = \max (0, uS - Q)$$
$$dS\Delta + rP = \max (0, dS - Q)$$

where S represents price of stock, Δ stands for rate of change in option value with relation to change in the underlying stock, P stands for amount of currency in riskless bonds, Q stands for strike price of option contract, r is the value of interest rate +1, and u, d represent the up/down return rates +1.

The above formulas represent the constraints that can directly yield the CLP(R) rule as

Valuation (C, S, Delta, P, R, Q, Up, Down):

$$Up * S * Delta + R * P = \max(0, Up * S - Q),$$

Down * S * Delta + R * P = max(0, Down * S - Q),
$$C = \max(S * Delta + P, S - Q).$$

The goal can evaluate the option price for given values of S, Q, R, Up, Down. For obtaining the value of C, CLP(R) solves several linear equations and produces answers. By iterating the evaluation rule backward for a period of time, binomial pricing model is obtained. An option value tree is computed by CLP(R) clauses using above recursion [1, 19].

2.2 The Black-Scholes Model

The formula of the Black-Scholes model is limit case of formula for binomial pricing. It is based on continuous model of option trading. It represents the theoretical price of a call option in terms of strike price (k), stock price (s), the current interest rate (risk-free) r, the time to expiration t, and volatility v for the stock calculated as square v root of variance by the following equation[2]:

Theoretical value for option price = $sN(d_1) - ke^{-rt}N(d_2)$

$$d_{1} = \frac{lnln\left(\frac{s}{k}\right) + \left(r + \frac{v^{2}}{2}\right)t}{v}\sqrt{t}$$
$$d_{2} = d_{1} - v\sqrt{t}$$

where ln stands for natural log and N(y) stands for cumulative normal distribution [7, 8].

3 Complexity in Options Market

Complex adaptive system theory gives a new way to study economics. This suggests that an economic system having proactive and adaptive individuals is processbased and evolving and self-organizing rather than predictable, mechanical, and established. To optimize trading, information is exchanged, and the members in the system accumulate this knowledge [15].

The market of options [25] is very complex in terms of adaptability. It depends upon the behaviour of investors and their decision-making approach. Market and investors both are affecting each other and thus changing continuously. Following is the discussion of a simulation of agent-based model.

3.1 The Architecture and Environment of Agent-Based Model

There is a mechanism very common in the market known as Continuous Double Auction (CDA). In this model, each participant is given a role that may be of buyer or of seller. In the process of trading, the seller gives a bid for a stock. This bid can be seen by all participants of the market. An investor analyses the bid given by seller. If it might maximize the gain for the investor, then bid given by the seller is accepted, and thus a transaction is completed [16, 26].

In this model, many experimental methods are chosen for a particular trading approach. Thus, this model is very helpful in simulating a continuous double auction market. This model considers the fact that investors are not aware of any information regarding valuation and cost of entity.

3.2 Agent Strategies (Zero Intelligence Plus Model)

ZIP is machine learning-based approach that determines optimal strategy by experience. It adapts according to open auction market environment which helps to make decisions about strategies to be followed. Any agent *i* in this model has profit margin $\mu_i(t)$ and limit profit α_i . An agent can submit a bidding price $p_i(t)$ as [3]

$$p_i(t) = \alpha_i(1 - \mu_i(t))$$

and the seller can submit asking price as

$$p_i(t) = \alpha_i(1 + \mu_i(t))$$

An agent *i* is selected randomly and asking or bidding price $p_i(t)$ is the same as equal to the current price in market $q_i(t)$. If the asking or bidding price is lesser than (or more than) the current price in the market, then the transaction may happen. Whether a transaction happens or not, all agents update the profit margin $\mu_i(t)$ according to learning algorithm [9, 10].

3.2.1 Model Parameters

Every agent is allotted some parameter in this model – parameter for learning rate β_i , profit margin $\mu_i(t)$ and dynamic parameter ψ_i . Genetic algorithm is used for the optimization of transaction process. Many issues are considered whether the underlying price goes up or down, movement of option price, asking or bidding price is less or more than the current market price, etc. According to these issues, market state is analysed, and on the basis of market state, strategies are formed to either increase or maintain the profit.

If the profit increases for a particular strategy, then it is believed by agent that strategy was optimal that market state and thus the probability of using that strategy again increases. On the other hand, if the profit decreases, it leads the agent to believe that strategy was not appropriate in that market state and the probability of using that strategy again is reduced. Thus, using this learning approach, the agent gets to choose strategies for making profit.

3.2.2 Optimizing Strategies Using Monte Carlo Simulation

In the market, strategies get affected by current market price and the future prices of stock. In Monte Carlo simulation, random sampling is used for state variables. The first random sample path is generated for stock price under consideration as p1, p2, p3, ..., pT over a relevant time horizon. Assuming the Brownian motion, the price for the underlying stock is [3]:

$$\frac{dS}{S} = \mu dt + \beta dz$$

where dS represents increment in stock price, μ stands for asset return, β represents the price volatility, dz stands for Wiener increment, where z is a Wiener process. Assuming the risk of neutral condition, the above formula can be rewritten as:

$$\frac{dS}{S} = \left(\mu - \frac{\beta^2}{2}\right)dt + \beta dz$$

In the above derivation of formula, assume the time interval is divided into *X* sub-intervals having length $\Delta t = T/X$. Thus, a new formula can be written as

$$\ln S_i - \ln S_{i-1} = \left(\mu - \frac{\beta^2}{2}\right) \Delta t + \beta \sqrt{\Delta t} \cdot \varepsilon_i$$

where ε_i represents standard normal distribution. If underlying price S_0 is given, then price S_i can be calculated as

$$S_i = \exp(\ln S_0 + \left(i \cdot \left[\left(\mu - \frac{\beta^2}{2}\right)\Delta t + \beta\sqrt{\Delta t} \cdot \varepsilon_i\right]\right]$$

Simulating the above formula many times and averaging it, underlying price can be calculated [13, 14].

3.2.3 Process of Strategy Formation

As it is clear from the above analysis, strong relationship exists between underlying and option price. Given the price distribution of underlying, option's life is divided into several small units. Random sampling is done on samples of distributions. With the help in simulation of price changes, movement of price's path can be predicted. Thus, value for the price of option can be calculated. By taking a random sample from the set of terminal values, the result can be analysed.

Model of equilibrium for rational expectations can be very helpful here. Over a particular period, it is assumed that $x_{i,j}$ stands for shares of underlying stock that are holed by buying agent I who is willing to accept the writer's asking price and $y_{i,j}$ stands for the writer who is willing to accept the bidding price. $\psi_{i,j}$ represents the demand for underlying stock, then the following relationship holds among them,

$$x_{i,j} = \psi_{i,j} - \psi_{i,j-1} \qquad \text{if } \psi_{i,j} \ge \psi_{i,j-1}$$

Otherwise, $x_{i,j}$ is 0

$$y_{i,j} = \psi_{i,j-1} - \psi_{i,j}$$
 if $\psi_{i,j} < \psi_{i,j-1}$

Otherwise, $y_{i,i}$ is 0

$$X_t = \sum_{N}^{i=1} x_{i,j}$$
$$Y_t = \sum_{N}^{i=1} y_{i,j}$$

where X_t denotes number of holders and Y_t stands for number of writers. The α represents the difference function between X_t and Y_t . α can be expressed as regulation speed of price as

$$\alpha (Xt - Yt)) = \{ \tan h (\alpha (Xt - Yt)), Xt \ge Yt \tan h \tan h (\alpha (Xt - Yt)), Xt < Yt \}$$

and

$$\tan h(x) = \frac{e^{x} + e^{-x}}{e^{x} - e^{-x}}$$

The price prediction for underlying price can be formulated as [3]

$$S_{t+1} = S_t \left(1 + \alpha \left(X_t - Y_t \right) \right)$$

From other paths, different random samples are taken. After repetition of this procedure hundreds of times, the set of options at a particular time is obtained. Averaging over different sample paths, calculation of option price is done. Agents tend to make a decision by comparison of market and calculated option price. This helps also in assessing the profit scenario for trader. Thus, the agents can determine whether to write or hold the options. Repetition of this process very much reflects the market situation.

To choose the opening prices, a method can be used which optimizes the potential profit of the makers of the market. This method can be used to offset the imbalances in public order with few limits on the allowed deviation for implied volatilities. In this method, set of implied volatilities σ over some range of possible values of implied volatilities.

Corresponding prices are examined for each element in this set. This helps in determining the side responsible for imbalances in the public order. For each imbalance in the buying side, an upward move is followed from price p_i in the set σ_i . For each price greater than p_i , product of buying imbalances is calculated, which helps determine the profit potential in the market.

A similar process is followed for the imbalances in selling side. There is a difference when followed for selling side; instead of moving forward, stepping down is followed in prices. A price can be found in each series for which profit potential can be maximum but a bound which is exchange specified for a deviation which is maximum from price p_i should also be considered.

Opening prices are picked by exploring over the sets for implied volatilities σ . The above procedure results in sets of prices that helps in maximizing the traded volume.

3.2.4 Enhancement of Model Based on Delta-Gamma Parameters

Option price can be calculated using Black-Scholes formula as

$$C_{t} \triangleq C^{(BS)} \left(S_{t}, T - t, \sigma_{t}, r_{t}, K \right)$$

= $S_{t} \phi \left(d_{+} \right) - K \exp \left(-r_{t} - (T - t) \right) \phi \left(d_{-} \right)$

where S_t is the price of underlying, *T* represents maturity time, implied volatility is represented by σ_t , r_t whose value is greater than zero represents the prevailing rate of interest, *K* represents the striking price of option, ϕ represents the cumulative distribution function and d_+ , d_- are described as

$$d_{+} / d_{-} = \frac{1}{\sigma t} \left(\ln \left(S_{t} / K \right) + r_{t} \left(T - t \right) \right)$$

Option price can be represented as

$$V_t \triangleq \left(S_t - S_0\right) / S_0$$

where V_t stands for linear return of stock price during the time (0,t) and S_0 stands for underlying price at beginning.

According to risk or position (in both cases) management, change in option price can be represented as

$$C_{t} - C_{0} = C^{(BS)}(S_{t}, T - t, \sigma_{t}, r_{t}, K) - C^{(BS)}(S_{0}, T, \sigma_{0}, r_{0}, K)$$

Some modifications can be made as

$$C_{t} - C_{0} = C^{(BS)}(S_{t}, T - t, \sigma_{t}, r_{0}, K) - C^{(BS)}(S_{0}, T, \sigma_{0}, r_{0}, K)$$

assuming that interest rate is uniform mostly.

On assuming that volatility is also constant, then above equation can be modified further as

$$C_{t} - C_{0} = C^{(BS)}(S_{t}, T - t, \sigma_{0}, r_{0}, K) - C^{(BS)}(S_{0}, T, \sigma_{0}, r_{0}, K).$$

It can be observed that option price calculation based on approximation of delta gamma parameters drops as change in underlying stock price increases for all the parameters which are considered here. This behaviour becomes more important as the value of volatility decreases (Fig. 1).

The reason for the above changes is that approximation is made at beginning while approximation becomes more accurate if it is made at t time. This modification leads to a new concept called extended delta gamma approximation.

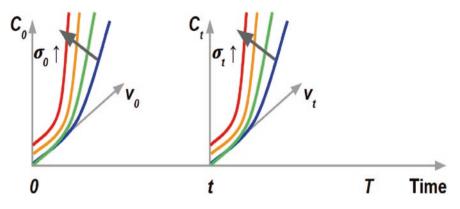


Fig. 1 Change in implied volatility and asset price [2]

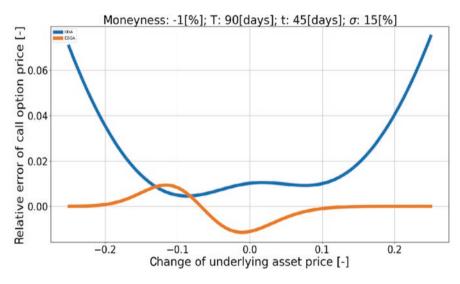


Fig. 2 Comparison of EDGA and DGA [2]

Thus, according to extended delta gamma approximation, the above equations get modified as

$$C_{t} - C_{0} = C^{(BS)} \left(S_{t}^{*}, T - t, \sigma_{t}, r_{0}, K \right) - C^{(BS)} \left(S_{t}^{*}, T, \sigma_{0}, r_{0}, K \right)$$

Using functional data analysis, change in stock prices can be predicted (Fig. 2).

The performance relies on accuracy of prediction in the changes of underlying. In this method, modelling the scalar response as linear model as

$$X = \alpha + \int_{T}^{0} X \cdot \alpha \cdot ds + \varepsilon$$

where all the variables have their usual meaning.

 α can be expanded as

$$\alpha = \sum_{K}^{\binom{k=1}{K}} b_k$$

Covariate functions can also be expanded as

$$X(s) = C\psi(s)$$

So the model can be expressed as

$$X = \beta + CJ$$

where J is expressed as,

$$J = \int \psi(s) \cdot \alpha$$

Thus, using above expressions, accurate prediction can be made for underlying asset price (Fig. 3).

Delta plays a vital role in analysing the situation of market dynamics. It represents the relationship between sensitivity of underlying price and price of option. Delta can be seen as measure of the change in the underlying price. The delta is

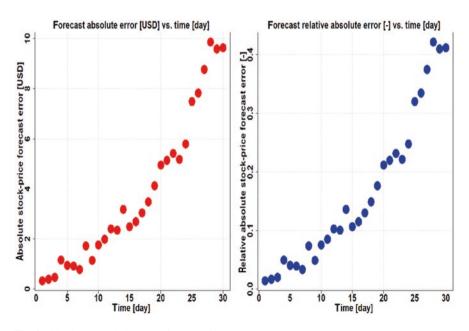


Fig. 3 Absolute vs relative error forecast [2]

calculated by change in price of option divided by change in the underlying price [11].

Concept of gamma is used to study the delta values. Gamma is used to measure the closeness of market price to the option price. As the closeness of these two prices increase, value of gamma goes higher. Gamma is also related to expiration date. As the maturity date comes closer, value of gamma increases as well.

Change in call option price can be predicted using delta and gamma approximation as

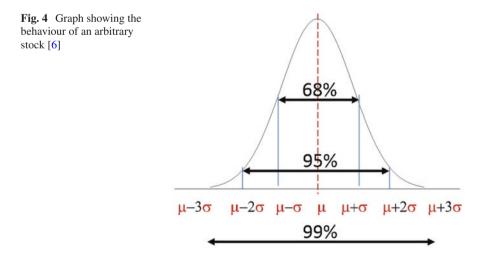
$$C_t - C_0 \approx \Delta v_t S_0 + \Gamma \left(v_t S_0 \right)^2$$

where C_t and C_0 are option prices and S_0 is underlying price, v_t is linear change in underlying asset price and Δ , Γ are parameters related to probability function and distributive function [12].

4 Estimating the Price Movement with Normal Distribution Curve of Underlying Asset

As the above calculation suggests, the price of underlying can be assessed by random sampling taken over different samples. Normal distribution curves can help to conclude the price movement over a specific time period. The Gaussian distribution curve is based on the central limit theorem. According to this theorem, random variables are identically distributed and independent when calculated average values show the same Gaussian (normal) distribution [4] (Fig. 4).

As it can be seen from the above diagram that within the first standard deviation, roughly 70% variation in underlying price occurs. If the transactions are based on a



strike price which are very close to the values in first deviation, then probability of market price ending on the price higher or closer to option value increases.

The learning algorithm can be modified in such a way to exhibit these relations of option price of assets. Figure 2 shows the behaviour of some most commonly traded stocks and indexes. It is evident from figure that variation in prices of stocks are showing the similarity between Gaussian distribution and their price variation over a period of time. Strike price vary according to moneyness of options. If the prices are chosen according to precalculated values based on standard deviation and variance, then it increases the probability of profit in the transaction (Fig. 5).

The Gaussian distribution curve for asset price can be used to predict price changes of the underlying. The kurtosis and skewness helps us to determine how different is the behaviour of the underlying asset from Gaussian distribution. Zero skewness curves are very similar to ideal distribution, and thus prediction is more accurate. Negative or less than zero skewness indicate that distribution is not even and biased toward the left area. Similarly positive or more than zero skewness value

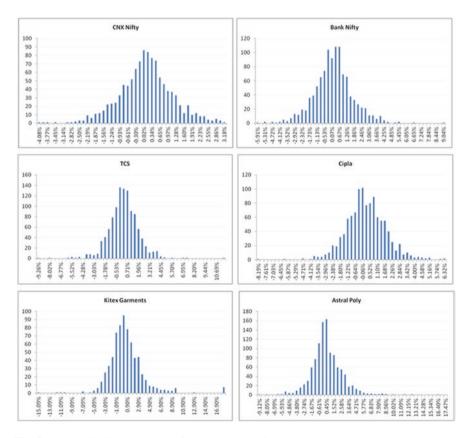


Fig. 5 Normal distribution curves of some stocks and indexes [5]

represents flatness of distribution in the right side. The more symmetric the curve, more accuracy can be achieved in predicting results.

As it can be concluded from above examples of normal distribution curves, if price estimation is between the first standard deviation, it can be used to optimize the profit because price movement is supposed to be under this range unless the market shows very strange behaviour. This calculation can be more accurate using the Monte Carlo simulation based on option pricing formula.

The planet is currently producing 2.5 quintillion bytes of data per day, which represents a once-in-a-lifetime resource for encoding, analysing, and using the data in useful ways. Machine learning and algorithms are rapidly being used in securities trading to process large amounts of data and make forecasts and judgments that humans cannot. Accurate insights into corporate decision-making models are critical in finance and trading. Humans have traditionally crunched numbers and taken choices based on inferences derived from measured risks and patterns. Computers have taken over this role today. They can compute on a large scale and pull data from a variety of sources to arrive at more precise results almost instantly. As implemented by authors in [22], even AI technologies provide promising solutions for end-to-end optimization of wireless networks. Financial analytics is no longer limited to the analysis of markets and price behaviour; it also encompasses the values that influence prices, as well as social and political patterns and the determination of degrees of support and opposition.

References

- Lassez, C., & Mcaloon, K. (1987). Constraint logic programming and option trading. *IEEE Intelligent Systems*, 2, 42–50.
- 2. Jun, J.-Y., & Rakotondratsimba, Y. (2020). *Hedging option contracts with locally weighted regression, functional data analysis, and Markov chain Monte Carlo techniques.* IEEE.
- 3. Jie, Z., Jing, C., Dongsheng, Z., & Quan, Z. (2009). An agent-based simulation model of options market. IEEE.
- 4. https://www.investopedia.com/terms/n/normaldistribution.asp
- 5. https://zerodha.com/varsity/chapter/volatility-normal-distribution
- 6. https://sphweb.bumc.bu.edu/otlt/MPH-Modules/PH717-QuantCore/PH717-Module6-RandomError/PH717-Module6-RandomError5.html
- MacBeth, J. D., & Merville, L. J. (1979). An empirical examination of the Black-Scholes call option pricing model. *The Journal of Finance*, 34(5), 1173–1186.
- 8. Raju, S. (2012). Delta gamma hedging and the Black-Scholes partial differential equation (PDE). *Journal of Economics and Finance Education*, 11(2), 51–62.
- 9. Meucci, A. (2008). Fully flexible views: Theory and practice. Fully flexible views: Theory and practice. *Risk*, *21*(10), 97–102.
- Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403), 596–610.
- Castellacci, G., & Siclari, M. (2003). The practice of delta-gamma VaR: Implementing the quadratic portfolio. *European Journal of Operational Research*, 150(3), 529–545.

- 12. Ramsay, J. O., & Silverman, B. W. (2007). *Applied functional data analysis: Methods and case studies*. Springer.
- 13. Carlo, C. M. (2004). Markov chain monte carlo and gibbs sampling. Lecture Notes for EEB, 581.
- 14. Gilks, W. R., Richardson, S., & Spiegelhalter, D. (1995). *Markov chain Monte Carlo in practice*. Chapman and Hall/CRC.
- Yadav, S. P. (2020). Vision-based detection, tracking and classification of vehicles. *IEIE Transactions on Smart Processing and Computing, SCOPUS.*, ISSN: 2287-5255, 9(6), 427–434. https://doi.org/10.5573/IEIESPC.2020.9.6.427
- 16. McMillan, L. G. (1986). "Options as a strategic investment," in New York Institute of Finance (2nd ed.). Prentice-Hall.
- Yadav, S. P., Agrawal, K. K., Bhati, B. S., et al. (2020). Blockchain-based cryptocurrency regulation: An overview. *Computational Economics*. https://doi.org/10.1007/s10614-020-10050-0
- 18. Heintze, N. C., et al. (1986). *The CLP(R) programmer's manual*. Tech. Report 73. Monash University.
- 19. Steele, G. L., Jr., & Sussman, G. J. (1979). *Constraints* (pp. 208–225). ACMSIGPLAN STAPL APL Quote Quad.
- Stephan, P., Al-Turjman, F., & Stephan, T. (2020a). Severity level classification and detection of breast cancer using computer-aided mammography techniques. Wireless Medical Sensor Networks for IoT-Based EHealth, 221–234. https://doi.org/10.1049/pbhe026e_ch13
- Chithaluru, P., Al-Turjman, F., Kumar, M., & Stephan, T. (2020). I-AREOR: An energy-balanced clustering protocol for implementing green IoT in smart cities. *Sustainable Cities and Society*, 102254. https://doi.org/10.1016/j.scs.2020.102254
- Stephan, T., Al-Turjman, F., Joseph, K. S., Balusamy, B., & Srivastava, S. (2020). Artificial intelligence inspired energy and spectrum aware cluster based routing protocol for cognitive radio sensor networks. *Journal of Parallel and Distributed Computing*. https://doi. org/10.1016/j.jpdc.2020.04.007
- Punia, S. K., Kumar, M., & Sharma, A. (2021). Intelligent data analysis with classical machine learning. In S. S. Dash, S. Das, & B. K. Panigrahi (Eds.), *Intelligent computing and applications* (Advances in intelligent systems and computing) (Vol. 1172). Springer. https://doi. org/10.1007/978-981-15-5566-4_71
- Wu, J. M. T., Wu, M. E., Hung, P. J., et al. (2020). Convert index trading to option strategies via LSTM architecture. *Neural Computing and Applications*. https://doi.org/10.1007/s00521-020-05377-6
- 25. Chen, A. P., Chen, Y. C., & Tseng, W. C. (2005). Applying extending classifier system to develop an option-operation suggestion model of intraday trading An example of Taiwan index option. In R. Khosla, R. J. Howlett, & L. C. Jain (Eds.), *Knowledge-based intelligent information and engineering systems. KES 2005* (Lecture notes in computer science) (Vol. 3681). Springer. https://doi.org/10.1007/11552413
- 26. Aggarwal, A., Alshehrii, M., Kumar, M., Alfarraj, O., Sharma, P., & Pardasani, K. R. (2020). Landslide data analysis using various time-series forecasting models. *Computers and Electrical Engineering*, 88(2020), 106858. https://doi.org/10.1016/j.compeleceng.2020.106858