*Iranian Journal of Management Studies (IJMS)* *http://ijms.ut.ac.ir/*
*Vol. 13, No. 1, Winter 2020* *Print ISSN: 2008-7055*
*pp. 69-93* *Online ISSN: 2345-3745*
*Document Type:Research Paper* *DOI: 10.22059/ijms.2019.276260.673534*

# Comparing Prediction Power of Artificial Neural Networks Compound Models in Predicting Credit Default Swap Prices through Black–Scholes–Merton Model

**Asghar Beytollahi, Hadis Zeinali***
*Department of Accounting, Islamic Azad University of Kerman Branch, Iran*

(Received: February 24, 2019 ;Revised: September 2, 2019; Accepted: September 21, 2019)

## Abstract

Default risk is one of the most important types of risks, and credit default swap (CDS) is one of the most effective financial instruments to cover such risks. The lack of these instruments may reduce investment attraction, particularly for international investors, and impose potential losses on the economy of the countries lacking such financial instruments, among them, Iran. After the 2007 financial crisis, the importance of CDS has increasingly augmented because theoretically and practically, this instrument could significantly prevent catastrophes such as the mentioned crisis. The present study seeks to predict the price of CDS contracts with the Merton model as well as the compound neural network models including ANFIS, NNARX, AdaBoost, and SVM regression, and compare the predictive power of these algorithms which are among the most prestigious, intelligent models in finance. The research statistical population includes the A-rated North American and European companies which are known as the reference entities for credit default swaps. Data were collected from the Bloomberg Terminal for an eight-year period from 2008 to 2015. Contracts of 125 companies were selected as the statistical sample. The results reveal that the average predictive power of the NNARX is higher than that of other algorithms under scrutiny.

## Keywords

Derivative Financial Instruments, ANFIS, NNARX, AdaBoost, Support Vector Machine Regression.

* **Corresponding Author, Email:** hadisazeinali@gmail.com

## Introduction

Risk or uncertainty is one of the most critical issues in the finance literature and capital markets. Among the various risk classes, credit risk has drawn more attention over recent years due to the turmoil of global capital market. Credit risk refers to the probability of a loss arising from a borrower's failure to repay a debt, and subsequently, the lender may fail to receive the owned principal or interest. Such a financial event, which is one of the factors causing disruption and disorderly conditions in the capital market, is referred to as default. The increased risk of default on obligations is a kind of credit risk that may augment the risk of the capital market, reduce cash, and increase the requested interest, eventually resulting in macro-economic losses.

In an attempt to escape the adverse consequences, investors are looking for ways to overcome such risks. The use of financial derivatives such as swaps, forwards, and futures can provide investors with new risk coverage. A financial swap is a derivative contract where parties exchange financial instruments. A specific type of these contracts is known as the credit default swap (CDS), specifically designed to transfer credit risk between two or more parties. According to this contract, in the event of a default, one party is committed to compensating all or part of the losses incurred by the other party (Hull, 2009). This instrument will reduce default risk, decrease the requested interest rates, and increase capital inflows in the financial markets, consequently leading to the prosperity of the capital market and the improvement of macroeconomic variables.

The derivative market has dramatically flourished over recent years, such that the exchange value of the CDS contracts has grown to over $60 trillion by the end of 2007 (Scott, et al. 2012). This exchange size due to the investors' warm welcome reflects the growing significance and necessity of this specific type of contract, which entails further research on the various dimensions of this crucial instrument. In Iran, which has an emerging capital market, the availability of such instrument is remarkably necessary because entering a market lacking a risk coverage instrument is inconceivable for international investors, and this, per se, may cause potential investment opportunities to be easily missed.

Pricing of derivative instruments such as CDS is one of the major issues in the finance literature and capital market research. In recent years, considerable effort had been dedicated to this issue and researchers have put forth various financial and statistical models to this end. One of the most important models presented in this regard is the Merton model proposed by Robert Merton for the valuation of corporate securities (Merton, 1974). This model can prognosticate the future price of securities such as derivative instruments by taking into account their current specifications.

The previous studies on the price prediction of the CDS contracts have mainly been focused on two axes including forecasts merely based on the classical, financial models and forecasts based on a combination of classical and intelligent, financial models to compare the performance of the former with that of the latter. In this study, the results of classical models have been compared with intelligent models because, according to the findings reported by the previous research, the performance superiority of the latter over the former is strongly evident (Gündüz & Uhrig-Homburg, 2011). The present study aims to compare the performance of classical models and intelligent algorithms. To the best of the authors' knowledge, so far, there has been no research that disregards the classical financial models and compares only the predictive power of intelligent algorithms. Accordingly, this study is an attempt to fill this gap.

The purpose of the present study is to predict the price of CDS contracts based on the Merton model and the hybrid neural network (HNN) algorithms including ANFIS, NNARX, AdaBoost, and SVM regression, and compare the predictive power of these models. To this end, the data required were extracted from the Bloomberg Terminal for the eight-year period from 2008 to 2015. The statistical population of the research includes the A-rated North American and European companies which are known as the reference entities for CDS contracts. Finally, the available data associated with 125 companies were selected as the statistical sample which got analyzed with MATLAB version 4.

Credit derivatives cover the credit risk and increase the investment volume and the prosperity of the capital market. As the authors know, except some works on the theoretical description of CDSs, there has been no research on the price prediction of the CDS contracts at the time of writing, among Iranian academia. Then it seems necessary to mention there is a kind of research gap in Iran content which this study can fill and draw attention of Iranian academics and market participants to this branch of the derivatives and its potential advantages. There were some difficulties in this way like absence of required data in Iran and difficulties for accessing international data. Furthermore, this study has been done based on North American and European data and applies some limited methods like Merton financial model and mentioned algorithms. Future studies can apply different pricing models, measuring methods, time horizons and Geographical zones.

The other parts of the study are formed as follows. Section 2 presents explanation on risk, derivatives, statistic and financial models of research. Section 3 introduces a literature review on some relevant previous works. Sections 4 and 5, respectively, assign theoretical framework, hypothesis development and research methodology. Section 6 presents the research findings and the last two sections focus on conclusion and implications.

## Credit risk, Derivatives and Financial and Statistical Models for Measuring Derivatives

### 1. Risk and credit risk

Risk is defined as a series of losses which are likely to occur as a result of some events including price changes (Karen, 2016). Securities trading risk: it is defined as the possibility of a loss or drop in value and can be classified as two general categories: (1) Systemic (or market) risk which is not possible to be eliminated by diversification. (2) Nonsystematic risk which can be also covered through diversification.

Credit risk is one type of nonsystematic risk and can be defined as a financial loss risk caused by a debtor’s reduced credit quality. Credit risk has two main types: 1) Default risk in which the obligor avoids repaying a part of or total financial obligation. 2) Credit deterioration risk in which the debtor’s credit quality is reduced. In this case, the debtor’ assets value decreases and the creditor experiences a financial loss (Meissner, 2009; Chau et al., 2018).

### 2. Derivatives and hedging

Derivative is a product with a value being derived based on the value of basic variables, called bases (index, underlying asset or reference rate), in a contractual form. The underlying asset can vary in type including equity, forex, commodity, etc. (Amuthan, 2014; Marthinsen, 2018). When dealing with derivatives, a right on an asset is provided to a buyer of a derivative, which may lead the buyer to purchase or sell the asset after or during a particular time period.

As a noun, *hedging* means to protect or fend off and in financial literature is defined as the action which protects oneself against financial loss. (Šperanda & Tršinski, 2015). Financial markets are naturally marked with an extremely high rate of volatility. By using derivative products, partial or full price risks are probably transferable via locking-in asset prices through which, derivative products manage to mitigate asset price fluctuation impact on the profitability and cash flow (Amuthan, 2014).

**3. Credit Default Swap**

A credit default swap (CDS) is considered as the most popular credit derivative. CDS acts as a contract providing insurance against a default risk posed by a particular company. Here, the company is a reference entity and the company's default is a credit event. The insurance buyer is provided with the right to sell and the insurance seller agrees to purchase bonds issued by the company for their face value while occurring a credit event. The bond's total face value that can be sold is regarded as the credit default swap's notional principal. The CDS buyer makes periodic payments to the seller until the CDS's life ends or by the time of occurrence of a credit event (Hull, 2009; Aragon & Li, 2019).

**4. Financial model: Valuing the CDS Contracts by Black–Scholes–Merton Pricing Formulas**

The most famous approaches of the Black–Scholes–Merton model are the Black–Scholes–Merton formulas for European call and put options prices. These formulas are

$$c = S_0 N(d_1) - Ke^{-rt} N(d_2) \quad (1)$$

$$p = Ke^{-rt} N(-d_2) - S_0 N(-d_1) \quad (2)$$

$$d_1 = \frac{\left[\ln\left(\frac{s_0}{k}\right) + \left(r + \frac{\sigma^2}{2}\right)T\right]}{\left[\sigma\sqrt{T}\right]} \qquad d_2 = \frac{\left[\ln\left(\frac{s_0}{k}\right) + \left(r - \frac{\sigma^2}{2}\right)T\right]}{\left[\sigma\sqrt{T}\right]} = d_1 - \sigma\sqrt{T}$$

The function N(x) refers to the distribution function of cumulative probability in a standardized normal distribution. In this probability, a variable has a standard n ormal distribution, Φ (0, 1), of lower than x.

The rest of the variables need to be familiar. The variables c and p represent the European call and European put price, S0 is the stock price at the time zero, K represents the strike price, r denotes the continuously compounded risk-free rate, **σ** denotes the stock price volatility, and T represents the time to maturity of the option.

Risk-neutral valuation is regarded as an alternative approach. In a European call option, as an instance, the option value expected at maturity in a world with neutral risk is

$$\hat{E}[\max(S_t - K, 0)]$$

Where, as previously indicated, Ê represents the value expected in a world with neutral risk. According to the argument of risk neutral valuation, the expected value is the European call option price c discounted at an interest rate which is free of risk, i.e.,

$$c = e^{-rt}\hat{E}[\max(S_t - K, 0)]$$

The terms in the equation (1) can be interpreted by writing

$$\boldsymbol{C} = e^{-rt}\left[S_0 N(d_1) - Ke^{-rt}N(d_2)\right]$$

The expression N ($d_2$) shows the probability according to which the option is utilized in a world of risk-neutral, so that KN ($d_2$) is the strike price which multiplies the probability of paying the strike price. The expression $S_0 N_0\ (d_1)e^{rt}$ is the value expected in a variable risk-neutral world, being equal to $S_T$ if $S_T > K$ and zero otherwise.

Since quickly exercising an American call option has never been optimal on a non-dividend paying stock, the equation (1) equals an American call option value on a non-dividend paying stock.

Unfortunately, no exact analytical formula has been developed for an American put option value on a non-dividend paying stock.

By practically using the Black–Scholes–Merton formula, the rate of interest r is set equal to the zero-coupon risk-free rate of interest for a maturity T. This is theoretically correct if r is a common function of time. This is also correct in a theoretical viewpoint when the rate of interest is stochastic in case of the lognormal stock price at the time T and appropriate choosing of the volatility parameter. It must be mentioned that the time is normally measured as the number of trading days left in the option's life divided by the number of trading days in one year (Hull, 2009).

**5. Statistic models: Artificial Neural Networks (ANNs)**

ANNs refer to a class of models generated by biological neural systems. The concept underlying ANNs is on the basis of computing systems that are capable of learning by experience via recognizing patterns available in a data set. After identifying necessary inputs (factors), a neural network can be simply trained to form a non-linear model of the underlying system. The model is then generalized to new cases that are not part of the training data (Kumar& Walia, 2006; Halagunde Gowda 2018).

In the current study, some compound forms of ANNs are used as statistic models including AdaBoost, NNARX, ANFIS, and SVM. The following provides some explanation on the models.

**5. 1. AdaBoost**

In AdaBoost, the classification system is constantly applied to the training data. However, the focus is on different instances in this set at each application with the use of adaptive weights ($\omega_b$ (i)). This contrasts other ensembles including Bagging in which the weights are not updated. After the termination of the training process, the training set is used so that the obtained single classifiers can be combined into a final, highly accurate classifier. Therefore, the final classifier obtains a high accuracy level in the test set, as shown both empirically and theoretically by various authors.

Although the AdaBoost algorithm has several versions, Freund and Schapire's version is the most widely used version known as

AdaBoost. For purposes of simplification, only two classes are assumed without the loss of generality. A training set is given by

$$T_n = \{(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)\} \quad (3)$$

where Y obtains values of {−1, 1} . The weight $\omega_b(i)$ is assigned to each observation Xi and is initially set to 1/n. After each step, the value becomes updated. A basic classifier denoted as $C_b(X_i)$ is developed on this new training set, $T^b$, and is implemented in each training sample. The classifier's error is denoted by $\varepsilon_b$ and is calculated as

$$\varepsilon b = \sum_{i=1}^{n} \omega b(i)\ \zeta b(i) \text{ where } \zeta b(i) = \{0 \text{ if } C_b(X_i) = y_i \text{ and } 1 \text{ if } Cb(xi) \neq yi$$

The (b + 1)-th iteration will have a new weight as follows:

$$\omega_{b+1}(i) = \omega_b(i).\exp(\alpha_b \zeta_b(i)) \quad (4)$$

where $\alpha_b$ is a constant yielded by the classifier's error in the b-th iteration. More specifically, according to the above-mentioned authors,

$$\alpha_b = \ln((1 - \varepsilon_b) / \varepsilon_b$$

Afterwards, the calculated weights will be normalized and thus add up to one. Accordingly, $\varepsilon_b = 0.5 - \gamma_b$ , where $\gamma_b$ indicates the superiority of the basic classifier in the *b*-th step over the default rule in the worst case, and where the classes both have a priori probability (0.5). In this case, the observations which have been classified wrong have increased weights while the correctly classified observations have decreased weights. This, in turn, forces the single classifier to be developed in the following iteration in order to concentrate on hardest examples. In addition, differences are larger when the weights are updated in cases where the single classifier error is small. The reason is that higher importance is given to the few mentioned mistakes when achieving a high level of accuracy by the classifier. Thus, the alpha constant is interpreted as a learning rate calculated as an error function on each epoch. Moreover, this constant is utilized in the final decision rule, which attaches more significance to individual classifiers making a smaller error.

This process is iterated in any step for $b = 1,2,3,\dots,B$. Ultimately, an ensemble classifier is developed as a linear composition of single classifiers weighted by the corresponding constant $\alpha_b$.

$$C(x) = sign\ (\sum_{b=1}^{B} a_b C_b(x)) \tag{5}$$

The AdaBoost algorithm is shown as following (Freund and Schapire):

1. Start with $\omega_b(i) = \frac{1}{n}, i = 1,2,\dots,n$

2. Repeat for $b = 1,2,\dots,B$

a) Fit the classifier $C_b(x) \in \{-1,1\}$ using weights $\omega_b(i)$ on $T^b$
b) Compute: $\varepsilon_b = \sum_i^n \omega_b(i)\zeta_b(i)$ and $\alpha_b = \ln((1 - \varepsilon_b)/\varepsilon_b)$
c) Update the weights $\omega_{b+1}(i) = \omega_b(i).\ \exp\ (\alpha_b \zeta_b(i))$ and normalize them.
3. Output the final classifier $c(x) = sign\ (\sum_{b=1}^{B} \alpha_b\, C_b(x))$

Freund and Schapire revealed that by increasing the number B in the iterations, the training error level in the AdaBoost classifier tends to zero exponentially. Furthermore, they showed that the generalization or true error ($\varepsilon_R$) in the final classifier $C_F(x)$ has an upper limit, depending on the training or apparent error ($\varepsilon_A$), the size of the training set ($n$), Vapnik-Chervonenkis's dimensionality coefficient in the parametric area of basic classifiers ($d$), and the number of iterations $B$ in AdaBoost (the number of combined single classifiers).

$$\hat{\varepsilon}_R = \hat{\varepsilon}_A + O\left(\sqrt{\frac{Bd}{n}}\right) \tag{6}$$

The generalization error may decrease in the final classifier by the increased size of the training data set. However, the error increases after increasing the number of included single classifiers, which is regarded as the classifier over-fitting. An over-fitted classifier occurs when being highly, closely adjusted to the training set. In this case, the classifier loses its generalization capacity on the total population, and thus functions inaccurately in classifying previously unseen samples (Alfaro, et al., 2008).

### 5. 2. Neural network autoregressive model with exogenous inputs (NNARX)

This neural network includes the neural network autoregressive with exogenous inputs (NNARX), which has feedback connections enclosing a number of layers in the network. The linear ARX model that is generally used in time-series modeling forms the basis for the NNARX model. The defining equation for the NNARX model is presented below:

$$y_1 = f\left(y_{(t-1),} y_{(t-2),} \dots y_{(t-n),} u_{(t-1),} , u_{(t-2),}, \dots u_{(t-n),}\right)$$

Where, y(t) is the value after the output dependent signal which is regressed on previous values belonging to the output signal and also belonging to an input independent (exogenous) signal. In the feed-forward neural network, the output is fed-back to the input as a part of the standard NNARX architecture, as illustrated on the left side (Figure 1). Since the true output exists in the network training, a series-parallel architecture can be generated, where the true output is utilized rather than feeding back the estimated output, as indicated on the right side (Figure 1). The resulting network diagram with a two-layer feed-forward network used for the approximation is presented in Figure 2. This network weight type affects the network output in two different ways. Firstly, it directly affects the output as a variation in weight leads to an immediate change of the output (the first impact can be determined with the use of standard back-propagation). Secondly, it indirectly affects the output since a number of inputs to the layer, including a (t,1), are the weights' functions. (Fahimifard, et al., 2012).

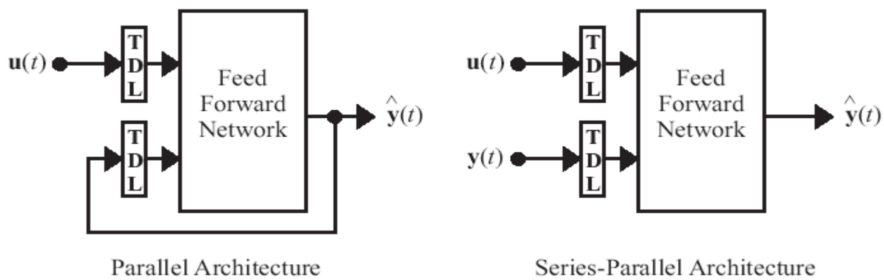



**Fig. 1. Parallel and series-parallel architectures**

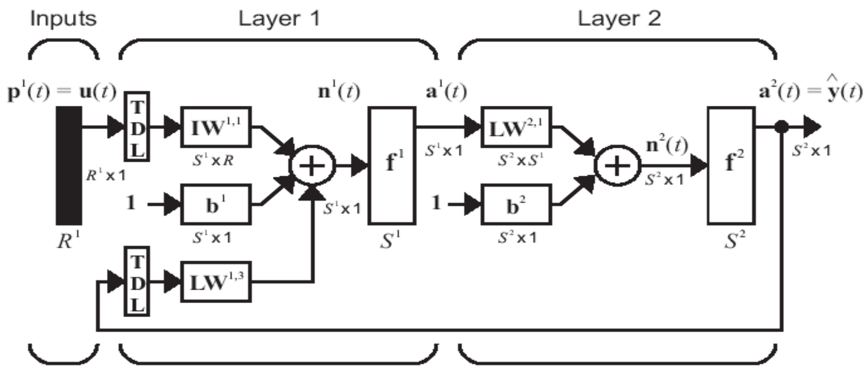



**Fig. 2. A typical neural network autoregressive with exogenous inputs (NNARX)**

## 5. 3. Support Vector Machines (SVM) Regression

Evidence shows SVM to be an appropriate alternative to conventional neural network applications; the SVM takes into account two major principles exclusively: data preprocessing specifies the performance of the SVM by using kernel functions in the first step and also using a linear learning algorithm in the second step via the kernel choice. Linear kernel function is the most basic function, which simply is the inner product in training points u and test points v:

$$\mathrm{K}(\mathrm{u},\ \mathrm{v}) = \langle\, \mathrm{u},\mathrm{v}\rangle$$

The analysis of a polynomial kernel function with degree 2 can also be considered as another alternative approach, which is commonly adopted for nonlinear modeling.

$$\mathrm{K}(\mathrm{u},\ \mathrm{v}) = (\langle\, \mathrm{u},\mathrm{v}\rangle + 1)^2$$

The third commonly used option found in the literature is the Gaussian radial basis function:

$$K(u,v) = \exp\left(\frac{\|\mathrm{u},\mathrm{v}\|^2}{2\sigma^2}\right)$$

where, **σ** is set to be 0.5 after observing the fit values of alternative parameters in the literature (Mu¨ ller et al., 1997; Gunn, 1998; Cao &

Tay, 2001). An exponential radial basis function as the last option is a similar to the Gaussian RBF:

$$K(u,v) = \exp\left(\frac{\|\mathrm{u},\mathrm{v}\|}{2\sigma^2}\right)$$

The SVM defines a loss function, known as insensitive loss function, where errors at a true value certain distance are ignored. A nonlinear regression function with one dimension and an insensitive band is shown in the following figure. The variables were used to calculate errors cost on the training points, and it was observed that the points within the insensitive band were equal to zero. Regarding the linear learning algorithm and following a thorough search for a value with the best-fitting feature, the parameter value C was set to be 10 in all runs. This parameter enables slacks in the system. This means that the samples are on the wrong side in the decision boundary (the error term penalty parameter), which is the distance in the following figure (Gündüz & Uhrig-Homburg 2011). Because of its overall good performance, this study has focused on a linear kernel for SVM analyzing.

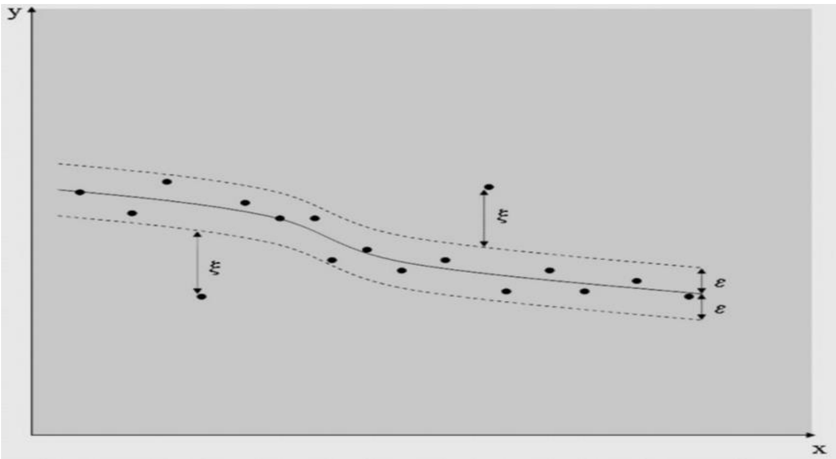



**Fig. 3. The insensitive band for a nonlinear regression function in an SVM regression setup**

**5. 4. Fundamentals of Adaptive neuro fuzzy inference system (ANFIS)**
The ANFIS complies with the multi-layer structure in ANNs with adaptive features, with a fuzzy inference system function which deals with nonlinear control through having a suitable selection and tuning the membership function of the control; the membership function can be properly tuned via pre-training and testing some input-output behaviors previously detected in the system. As a result, an adaptive structure with similar behaviors to humans is required for fuzzy logic action. This makes the ANFIS efficiently tackle with control actions, with regard to some past evidences. The ANFIS follows the Sugeno fuzzy inference system's principle. The ANFIS architecture is presented in Fig. 4. As can be observed, there exist five layers in the structure, each of which has many nodes with specific functions. The structure layer-wise functionality is described in the following:

Layer 1 ($L_1$): The fuzzy parameterized membership function is represented by the box in Layer 1. Every single adaptive node is described in this layer by its function as follows:

$$Q_i^1 = \mu T_i\left(x\right); i = 1, 2 \tag{7}$$

$$Q_i^1 = \mu K_{i-2}\left(y\right); i = 3, 4 \tag{8}$$

In general, '$O_i^j$' indicates the output of each node as '*i*', '*j*', and 'O' stand for node number, layer number, and output, respectively. Here, '*x*' or '*y*' demonstrates the input to the ith node while '*T*' or '*K*' represents the linguistic label given to this node. Normally, a bell-shaped function is considered as a member function, which is expressed as:

$$\mu T_i\left(x\right) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}} \tag{9}$$

In Eq. (9), the parameters of the general bell-shaped function include {a, b, c}.

Layer 2 ($L_2$): In this layer, the input signals to a node are multiplied by each other and the product result is the node output, as shown below:

$$Q_i^2 = W_i = \mu T_i(x)\mu K_i(y); i = 1,2 \tag{10}$$

The output of each node in Layer 2 is indicative of the firing strength of a rule. The structure node of Layer 2 is indicated as '$\pi$' in the following figure.

Layer 3 ($L_3$): In this layer, each node is labeled as 'N'. Normalized values of firing strengths are calculated based on the following ratio:

$$Q_i^3 = \overline{w}_i = \frac{W_i}{W_{1+}W_2}; i = 1,2 \tag{11}$$

Layer 4 (L4): In this layer, each rectangular node calculates its contribution to the net output as follows:

$$Q_i^4 = \overline{w}_i Z_i = \overline{w}_i(l_i x + m_i y + n_i); i = 1,2 \tag{12}$$

where, {l, m, n} are the parameters used in Layer 4.

Layer 5 ($L_5$): In this layer, the final output is calculated as shown below (Hsu, 2011).

$$Q_i^5 = \Sigma\overline{w}_i Z_i = \frac{\Sigma w_i Z_i}{\Sigma w_i} \tag{13}$$

(Gayen and Jana, 2017)

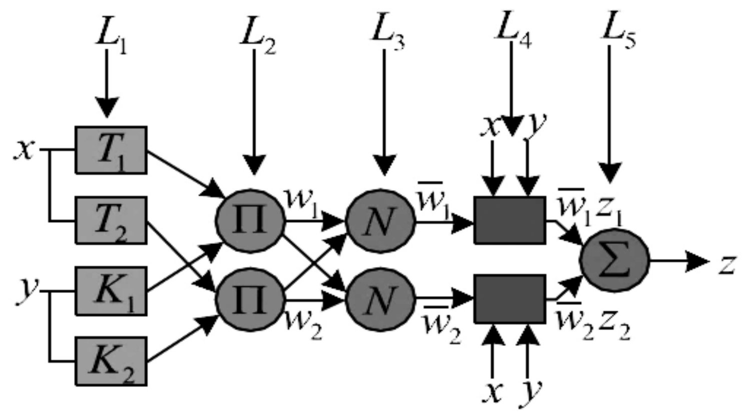



**Fig. 4. ANFIS architecture**

## Research empirical background

One of the first applications of the neural network in finance refers back to Kimoto et al. (1990), who used these networks to predict the Tokyo Stock Exchange Index. They used several neural networks trained to learn the relationship among the previous values of the various technical and economic indices in order to forecast the return on TOPIX which is the value-weighted average of all stocks listed on the Tokyo Stock Exchange.

Blaskowitz and Herwartz (2009) forecasted the EURIBOR swap term structure via the proposed adaptive models.

Gündüz & Uhrig-Homburg (2011) applied SVMs, structural models (Merton model), and reduced-form (constant intensity) models to single-name CDS bid-ask quotes from January 2001 to December 2004.

Gosh (2012) suggested a hybrid neural evolutionary methodology to predict time series, and particularly NASDAQ stock price. This is a hybrid methodology because an evolutionary computation-based optimization process has been used to design a neural network.

Anyaeche and Ighravwe (2013) dealt with the perdition of performance measure using linear regression and neural network. They used the artificial neural networks, the Back Propagation Artificial Neural Networks, as alternative predictive tools to multi-linear regression to establish an interrelationship among productivity, price recovery, and profitability as the performance indicators. Their findings revealed that ANN is the best model for establishing interconnection between the three mentioned components.

Likewise, Zahra and Seyedmohsen (2014) addressed the capabilities of neural networks and data envelopment analysis in predicting corporate productivity and, ultimately, provided decision makers with an optimal algorithm for predicting profitability.

Wang et al. (2016) intended to develop a prediction model for financial time series using an Elman recurrent neural network with a stochastic time effective function to test the predictive power of SSE, TWSE, KOSPI, and Nikkei225. The research evidence suggests that the proposed neural network outperforms some of the previous models in terms of the prediction accuracy, and its results are highly close to the actual market changes.

Pang and colleagues (2018) suggested the deep long, short-term memory neural network (LSMN) with the embedded layer to predict the stock market. Applying their model to Shanghai A-shares composite index, they compared the proposed model with the baseline models. Results shed light on the superiority of the proposed model performance over its other counterparts in predicting the index under scrutiny.

## The theoretical framework and research hypotheses

Intelligent algorithms appear to be considerably more powerful than the classical models in predicting the price of a derivative instrument. Among a wide range of algorithms, different models of neural network have higher predictive power. Accordingly, four most robust hybrid neural network algorithms in the field of financial sciences and finance, namely ANFIS, NNARX, AdaBoost, and SVM, have been selected. According to the explanations provided in the previous section about the structure and characteristics of each of these algorithms, it seems that the NNARX algorithm yields higher predictive power and lower error rate in predicting the price of derivatives via the Merton financial model compared with the other three algorithms under scrutiny.

The present study chiefly intends to determine which of the algorithms under scrutiny, i.e., ANFIS, NNARX, AdaBoost, and SVM, provides the most predictive power for the Merton model. Thus, the research hypotheses are formulated as follows:

1. The AdaBoost increases the predictive power of the Merton model more than that of three other algorithms.
2. The SVM increases the predictive power of the Merton model more than that of three other algorithms.
3. The NNARX increases the predictive power of the Merton model more than that of three other algorithms.
4. The ANFIS increases the predictive power of the Merton model more than that of three other algorithms.

## Research methodology

The present study seeks to predict the price of CDS contracts. To this end, the Black–Scholes–Merton as a widely used derivative contract pricing model has been used for CDS contract pricing in this study. Data analysis and prediction have been accomplished using HNN models including NNARX, ANFIS, SVM, and AdaBoost, which have analyzed data in MATLAB software using codes optimized for financial data analysis. The statistical population of the study includes the A-rated North American and European companies considered as the reference entities for CDS contracts during the period of 2008-2015. Finally, the contracts associated with 125 companies were selected as the statistical sample by applying the following filters:

- Companies must have been active since the beginning of 2008 until the end of 2015,
- Companies must be reference entities for concluding CDS contracts,
- Companies must have a Single-A credit rating or higher,
- September 31 is considered to be the fiscal year end of companies, and
- The financial information required to investigate the company must be fully accessible.

In the first phase, predictions were made based on the existing historical data and in the second phase, the results of the first phase predictions —due to their high accuracy—were added to the historical data as the basis for the following forecasts. The latter step required historical data to be trained. Data training is a method which indoctrinates the computer the way of information processing. The training data is a primary set of data which helps a program figure out how to use technologies like neural networks to learn and yield highly accurate outcomes. Data training is one of the most vital steps in the system operation. In the current study, data training has been done through algorithms mentioned above - NNARX, ANFIS, SVM, and AdaBoost – which in the next section their structure and advantages will be briefly explained. In the following section, further explanations will be presented. Finally, in line with the research objectives and hypothesis testing, the results and outputs yielded from the study algorithms were compared to determine which algorithm has the minimum error rate and the maximum predictive power.

## Research findings

In this section, the results obtained from data processing are separately presented for four algorithms under scrutiny and two base years. As it was stated earlier, due to the availability of actual data associated with the 2008-2015 period, the forecasts for the succeeding years were carried out. Using the data available for the mentioned period, prognostication of the contract price for 2016 was carried out. Additionally, considering 2006 as the first base year, the price forecasts for 2017, 2018, and 2019 were performed (the first phase). As this prediction has relied on a large number of records related to the actual contract price and since HNN models enjoy a considerable robustness, the error rate of the algorithms for predicting the 2017 contract prices was significantly low and their prediction power was significantly high; hence, the forecasts associated with this year were viewed as 2017 actual data. Accordingly, regarding 2017 as the second base year, forecasts for 2018, 2019, and 2020 were made (the second phase). The results are presented and discussed below.

Tables 1 and 2 show the prediction results based on the data associated with the 2016 base year. As observed, according to the data related to this year, the ANFIS algorithm with 82.97% has the least error rate and is placed in the second position after NNARX. On the contrary, SVM indicates the highest error rate and the least prediction accuracy.

**Table 1. Results of prediction with AdaBoost and SVM (2016 base year)**

| Hybrid ANN Algorithms | Average Prediction accuracy | Year 3 (2019) | Year 2 (2018) | Year 1 (2017) |
|---|---|---|---|---|
| AdaBoost | 92/67 | 91/19 | 91/95 | 94/88 |
| SVM | 91/83 | 90/59 | 91/40 | 93/51 |

**Table 2. Results of prediction with NNARX and ANFIS (2016 base year)**

| Hybrid ANN Algorithms | Average Prediction accuracy | Year 3 (2019) | Year 2 (2018) | Year 1 (2017) |
|---|---|---|---|---|
| NNARX | 97/20 | 96/51 | 97/20 | 97/90 |
| ANFIS | 97/82 | 97/10 | 97/99 | 98/39 |

Tables 3 and 4 present the results obtained for the 2017 base year. In this year, the highest accuracy is obtained for the NNARX algorithm (63.96%); however, the ANFIS algorithm has also an acceptable level of precision. The least level of accuracy was achieved for AdaBoost.

**Table 3. Results of prediction with AdaBoost and SVM (2017 base year)**

| Hybrid ANN Algorithms | Average Prediction accuracy | Year 3 (2020) | Year 2 (2019) | Year 1 (2018) |
|---|---|---|---|---|
| AdaBoost | 91/89 | 90/60 | 91/01 | 94/08 |
| SVM | 93/56 | 90/49 | 90/93 | 99/27 |

**Table 4. Results of prediction with NNARX and ANFIS (2017 base year)**

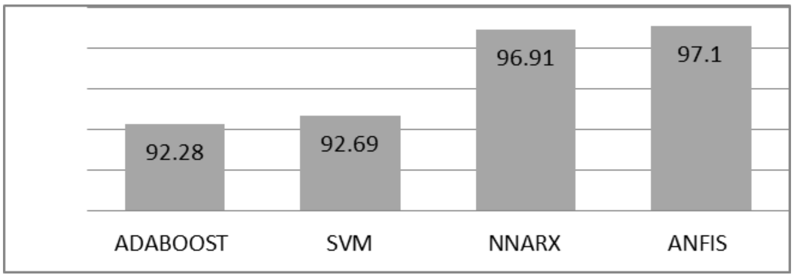
| Hybrid ANN Algorithms | Average Prediction accuracy | Year 3 (2020) | Year 2 (2019) | Year 1 (2018) |
|---|---|---|---|---|
| NNARX | 96/63 | 95/99 | 96/71 | 97/19 |
| ANFIS | 96/37 | 95/93 | 96/11 | 97/11 |

In the above tables, prediction of prices based on the Merton financial model using the four algorithms, i.e., ANFIS, NNARX, SVM, and AdaBoost, are presented as the statistical models for the two base years. As it is observed, ANFIS and NNARX algorithms indicate the highest accuracy for the base years 2016 and 2017, respectively.

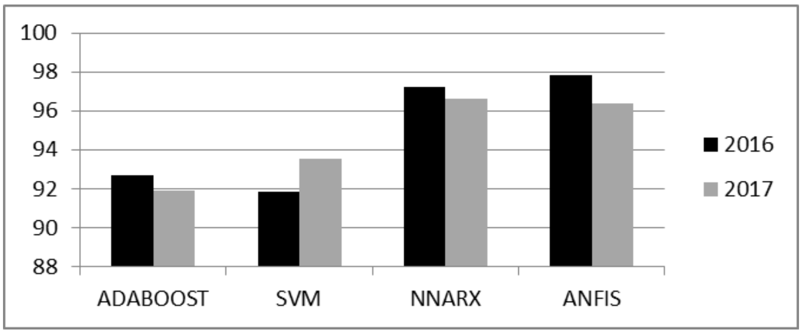
Here it is a brief explanation of how the algorithms are used in the current study. For the AdaBoost model, according to Friedman (2001), the maximum number of nodes and trees depth is respectively 10 and 4. It should be mentioned that the number of iterations is set to 800. NNARX is utilized with the *multilayer perceptron* (MLP) with hidden units including sigmoidal transfer functions used as the feedforward NN which the MLP-networks is bounded to those with single hidden layer. The results through this model are reported based on the MSE (Mean Square Error) as an evaluator of neural model, and the validity of correct predictions, regarding the direction of changes. The ANFIS architecture includes one input and one output; the model used in the current study predicts the next three days value of CDS contracts according to the previous values and forecasts the CDS prices three steps ahead. The type and number of membership functions is based on trial and error to provide the epoch number and step size which describe the model in the best way and present the lowest error. The fuzzy inference was optimized after 1000 epochs by two membership gauss shape functions with the 0.01 value for the step size. Finally in the SVM setting, the study following Gündüz & Uhrig-Homburg (2011), a linear kernel was used and the data were divided into four parts for estimation and prediction. Three quarters were used for training the SVM function. Observations of

14-day in row were used as the training input, whilst the training output was the observation on the subsequent day. The fourth quarter of the data was employed after the function training, for prediction. As test input, the observations of residual continuous 14-day, were used to predict the observation on the next day as the test output.

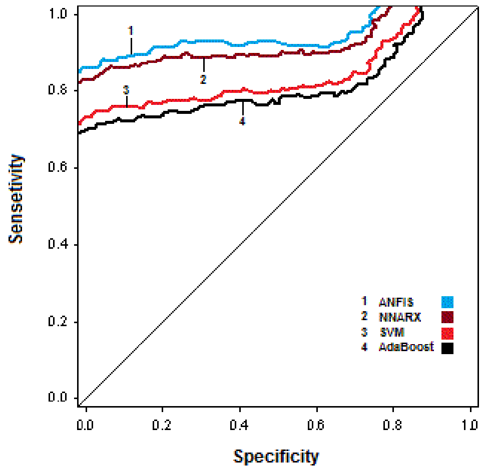
Graph (1) schematically illustrates these results. As regards the average prediction accuracy of two base years, the highest accuracy belongs to the ANFIS algorithm with the predictive power of 97.1%, as exhibited in Graph (2) which demonstrates the two-year average predictive power of models. According to the observations, the third hypothesis is confirmed, i.e., the NNARX increases the Merton model power to predict the price of CDS contracts more than three other algorithms. Graph (3) presents the Receiver Operating Characteristic curve, or ROC curve for comparing the four models visually.




**Graph 1. The comparison of the predictive power of the statistical models based on fiscal years**




**Graph 2. The two-year average predictive power of models (Percentage)**




**Graph 3. ROC curves comparing models prediction accuracy**

## Discussion and conclusions

### 1. Discussion

Classical financial models predict the price of derivatives such as options and swaps based on the existing data; consequently, they have limited power to predict the price of the futures contracts that have not hitherto been concluded. In contrast, artificial intelligence statistical models such as HNN algorithms have a significantly higher power to prognosticate the trends and perform precise predictions due to a number of strengths, e.g., application of machine learning process and capability to handle a large volume of computations. Therefore, the use of intelligent statistical models can add higher measurement precision and predictive power to the classical financial models such as the Merton model and also substantially reduce the error rate of these models.

Such computations have become increasingly widespread and drawn considerable interests in the global financial markets, such that progressive markets strive to pave the way for the application of financial robotics which can strongly reduce the permanent role of the human and negative impact of cognitive errors on the transactions. Capital market practitioners in Iran can augment the accuracy of transactions using these advanced computational methods and, hence, improve the precision in the entire capital market.

Another aspect dealt with in this study is the use of financial derivatives, especially the credit derivative instrument like CDS as the most widely used type. As mentioned in the earlier sections, the use of a credit derivative instrument will cover credit risk and, therefore, increase not only the investment volume but also the prosperity of the capital market. Far as the authors know, there was no research, at the time of writing, on the pricing and price prediction of the CDS contracts in the Iranian academia, and some limited research has been dedicated to a sheer description of this critical issue. Since the derivative financial instrument is not traded in Iran, there is no information on its exchanges in the Iranian market. Therefore, this study used international data in an attempt to pave the way for the familiarity of the Iranian researchers and capital market practitioners with this crucial category of the financial instruments. This, per se, can open up new avenues for the entry of this very instrument into the Iranian market and make the capital market take advantage of its benefits.

## 2. Conclusion

As it was expected and also as observed in the numerical reports, the findings reveal that the use of ANN algorithms and classical financial models such as the Merton model will provide high accuracy for the prediction of derivative pricing. Among the four algorithms used in this study, NNARX and ANFIS algorithms yield higher prediction accuracy. In explaining this supremacy, the efficiency resulted from the neuro-fuzzy system and the feedback function can be mentioned as the most crucial strengths of the ANFIS and NNARX algorithms, respectively, compared with other HNN models. According to the numerical results obtained from analyses, the highest one-year predictive power is related to the ANFIS algorithm for the base year 2016, while the highest two-year predictive power is related to the NNARX algorithm. Accordingly, the third hypothesis is also approved, i.e., the NNARX algorithm yields the most predictive power and the least error rate for the Merton financial model compared with the other three algorithms.

The researcher has also confronted with inevitable limitations that overshadowed the overall trend of the work, e.g., lack of data availability in the Iranian market and difficulty in accessing the international data. As the statistical population of the present study was chosen from the North American and European countries, future studies are recommended to address other geographical areas. In this study, only the Merton model and four intelligent algorithms were used to analyze data; hence, future works are suggested to apply other financial and statistical models to that end.

## Scientific and practical applications

As it was mentioned, so far, there has been no study that merely compares the power of intelligent algorithms to predict the price of CDS contracts; therefore, this study tries to fill this research gap and pave the way for future research in the Iranian academic environment. Investors, academic researchers, investment companies, and other capital market practitioners are potential users of such research. This study also prepares the ground for the familiarity of the Iranian capital market practitioners and researchers with credit derivative instruments and opens up new avenues for the entry of such instruments into the Iranian market, prosperity of the capital market, and positive macroeconomic outcomes. Notably, using these instruments, investors can cover credit risk and make more confident investment decisions. Investment companies can also improve their profitability and attract more funds by taking advantage of the benefits of this instrument. The benefits of credit risk coverage can also help enterprises managers to achieve a higher level of efficiency in their economic activities.

## References

Alfaro, E., García, N., Gámez, M., & Elizondo, D. (2008). Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks. *Decision Support Systems, 45*(1), 110-122.

Amuthan, R. (2014). *Financial Derivatives*. Himalaya Publishing House.

Anyaeche, C. O., & Ighravwe, D. E. (2013). Predicting performance measures using linear regression and neural network: A comparison. *African Journal of Engineering Research, 1*(3), 84-89.

Aragon, G. O., & Li, L. (2019). The use of credit default swaps by bond mutual funds: Liquidity provision and counterparty risk. *Journal of Financial Economics, 131*(1), 168-185.

Blaskowitz, O., & Herwartz, H. (2009). Adaptive forecasting of the EURIBOR swap term structure. *Journal of Forecasting*, *28*(7), 575-594.

Chau, F., Han, C., & Shi, S. (2018). Dynamics and determinants of credit risk discovery: Evidence from CDS and stock markets. *International Review of Financial Analysis*, 55, 156-169.

Fahimifard, S. M., Mohaddes, S. A., Mohammadi, H., & Keshtkar, R. (2012). Application of NNARX to agriculture sector value added forecasting: A case of Irans agriculture sector. *African Journal of Agricultural Research*, *7*(27), 3918-3924.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: An event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, *25*(4), 355-373.

Gayen, P. K., & Jana, A. (2017). An ANFIS based improved control action for single phase utility or micro-grid connected battery energy storage system. *Journal of Cleaner Production*, *164*, 1034-1049.

Gündüz, Y., & Uhrig-Homburg, M. (2011). Predicting credit default swap prices with financial and pure data-driven approaches. *Quantitative Finance*, *11*(12), 1709-1727. Halagunde Gowda, G. R. (2018). *Classificatory statistical modelling on adoption of drough coping mechanisms*. (Doctoral dissertation, DEPARTMENT OF FARM ENGINEERING

INSTITUTE OF AGRICULTURAL SCIENCES BANARAS HINDU UNIVERSITY VARANASI).).

Hull, J. (2009). *Options, futures and other derivatives*/John C. Hull.

Kimoto, T., Asakawa, K., Yoda, M., & Takeoka, M. (1990, June). Stock market prediction system with modular neural networks. In 1990 IJCNN *international joint conference on neural networks* (pp. 1-6). IEEE.

Marthinsen, J. E. (2018). *Risk takers: Uses and abuses of financial derivatives*. Walter de Gruyter GmbH & Co KG.

Meissner, G. (2009). *Credit derivatives: application, pricing, and risk management*. John Wiley & Sons.

Pang, X., Zhou, Y., Wang, P., Lin, W., & Chang, V. (2018). An innovative neural network approach for stock market prediction. *The Journal of Supercomputing*, 1-21.

Scott, K. E., & Taylor, J. B. (Eds.). (2012). *Bankruptcy Not Bailout: A Special* (No. 625). Hoover Press.

Wang, J., Fang, W., & Niu, H. (2016). Financial time series prediction using elman recurrent random neural networks. *Computational intelligence and neuroscience*, Volume 2016, Article ID 4742515, 14 pages.