ICIC Express Letters
Part B: Applications
Volume **10**, Number **12**, December **2019**

ICIC International ©2019 ISSN 2185-2766
pp. 1047–1055

BINARY LOGISTIC REGRESSION, ARTIFICIAL NEURAL NETWORK AND SUPPORT VECTOR MACHINE TO PREDICT THE WILLINGNESS OF SAUDI RESIDENTS TOWARDS VALUE ADDED TAX

Nahid Sultana

College of Computer Science and Information Technology
Imam Abdulrahman Bin Faisal University
PO BOX 2435, Dammam 31441, Kingdom of Saudi Arabia
nszakir@iau.edu.sa

Received June 2019; accepted September 2019

Abstract. *The Kingdom of Saudi Arabia implemented VAT on January 1, 2018. Assessment of the acceptance of the residents in Saudi Arabia towards newly implemented VAT is an important area of empirical research in Business, Economics, and Social Sciences. This paper aims to analyze the acceptance of Saudi residents towards VAT by utilizing machine learning techniques. Three competing classification procedures, Binary Logistic Regression (BLR), Artificial Neural Network (ANN) and Support Vector Machine (SVM), are applied in developing the empirical models. The optimized parameters for these procedures were decided by tuning the parameters for these models and executing numerous experiments. The best optimized parameters were used to develop the final models of the proposed procedures. The experimental results showed that SVM and ANN performed better than BLR. The prediction accuracies of SVM, ANN and BLR are 99.26%, 98.52% and 85.19%, respectively, representing that the outcome of this study is highly auspicious.*

Keywords: Value added tax, Binary logistic regression, Artificial neural network, Support vector machine

1. Introduction.

In June 2016, the GCC (Gulf Cooperation Council) countries decided to execute VAT throughout the GCC region [1]. The Kingdom of Saudi Arabia and the United Arab Emirates have executed VAT, in January 1, 2018, as the major Gulf economies attempt to increase non-oil income and reduce fiscal deficits created by years of low oil prices. By introducing a 5 per cent tax on most goods and services these two countries took a crucial measure towards ending decades of tax-free living for the residents. Other Gulf countries have pushed back the implementation of VAT until 2019. The effect of VAT implementation in GCC countries specially in Saudi Arabia has rarely been studied. Machine learning techniques can be applied to measuring the level of acceptance of Saudi residents towards newly implemented VAT system.

BLR, ANN and SVM are extensively employed machine learning methods that have been recognized to perform magnificently in numerous arenas [2-4]. Logistic regression is a predictive analysis which was originated by statistician Cox in 1958 [5]. The BLR model is applied to analyzing the association between the binary dependent variable and one or more nominal, ordinal, interval or ratio-level independent variables [6,7]. ANN is a mathematical model that performs alike human neurons. The ANN has the aptitude to find out how to execute the procedure after proper training [8]. Whereas SVM is a computational procedure that learns from knowledge and samples to assign classification to objects. The key role of SVM is to classify binary dataset based on a line reaching the highest distance between the labeled data [9]. SVM has satisfactory prediction accuracy

DOI: 10.24507/icicelb.10.12.1047

for even small dataset [10-14]. This paper applied BLR, ANN and SVM classification to predicting the willingness to pay VAT of Saudi residents and compare their performance. The statistical softwares R and SPSS were used in conducting experiment to construct model. The experimental outcomes showed that SVM model performed better with a high accuracy of 99.26%, ANN had accuracy 98.52%, while BLR had accuracy 85.19%. The rest of the paper is prepared as follows. Section 2 presents the proposed predictive models using machine learning techniques, Section 3 presents the result and discussion, and the conclusion is emphasized in Section 4.

2. **Empirical Studies.** This section describes the description of the dataset and the experimental procedure for the predictive models.

2.1. **Description of the dataset.** A quantitative research using simple random survey was conducted to address the objective of this study. The survey was conducted from February to October in 2018 in the Kingdom of Saudi Arabia. A total of 405 respondents participated in this study. There is no missing data in the dataset because the participants could not submit the online survey without completely answering all questions. The dataset consists of 14 attributes namely age, gender, nationality, social status, employment status, working sector, monthly income, VAT effect individuals spending, VAT effect monthly savings, start buying cheaper alternatives after VAT implementation, VAT will reduce social obligations, aware about VAT, VAT will have positive impact on KSA, and willing to pay VAT. Age and monthly income are the only numerical attributes and rest of the 12 attributes are nominal.

2.2. **Basic statistical analysis of the dataset.** Age of the participants ranged from 18 to 73 with a mean age of 28.34. Majority of the participants are female (67%). In terms of nationality, most of the respondents are Saudi (96%) and only 4% are other national. About 61% of the respondents are single, 35% are married and only 4% are divorced. Only 44% of the participants are employed, and about 63% of them are working in the private sector. The monthly income of the respondents ranged from 250 SAR to 150,000 SAR with average monthly income 9,537 SAR.

Spearman's correlation analysis was utilized to determine the association between individual's willingness to pay VAT with demographic factors (results shown in Table 1). As shown in Table 1, there is a significant, positive and moderately strong relationship between the respondents' willingness towards VAT and VAT will have positive impact

Table 1. Correlation of willing to pay VAT with demographic factors

Attributes and levels	Willing to pay VAT (Yes/No)	
	Correlation coefficient	Sig. (2-tailed)
Age (Numerical)	−0.096	0.053
Monthly income (Numerical)	−0.029	0.555
Employment status (Yes/No)	−0.019	0.702
VAT effect individuals spending (Yes/No)	**−0.175****	**0.000**
VAT effect monthly savings (Yes/No)	**−0.249****	**0.000**
Buy cheaper alternatives after VAT (Yes/No)	**−0.208****	**0.000**
VAT reduced social obligations (Yes/No)	**−0.123***	**0.014**
Aware about VAT (Yes/No)	**−0.270****	**0.000**
VAT will have positive impact on KSA (Yes/No)	**0.442****	**0.000**
Social status (Single/Married/Divorced)	**−0.137****	**0.006**
**Correlation is significant at the 0.01 level (2-tailed). *Correlation is significant at the 0.05 level (2-tailed).		

of VAT on KSA (p-value = 0.000 and the correlation coefficient = 0.442); the respondents' willingness towards VAT has significant but low and negative correlations with six attributes: effect of VAT on individuals spending, effect of VAT on monthly savings, start buying cheaper alternatives after VAT, VAT will reduce social obligations, awareness about VAT and social status.

2.3. **Experimental setup.** BLR, ANN and SVM techniques ware applied to conducting several experiments using the dataset described in Section 2.1 to construct best prediction model. The accuracy of the all three models was calculated and compared with each other. The optimized parameters for ANN and SVM were also determined (results shown in Table 2).

Table 2. Optimization parameters of the proposed ANN and SVM model

ANN		SVM	
Parameters	Optimum value chosen	Parameters	Optimum value chosen
Hidden layer	1 hidden layer with 8 neurons	Kernel	Polynomial kernel with degree = 3 and scale = 10
Activation function	Hyperbolic tangent activation function for the hidden layer and softmax activation function for the output layer		
Type of training	Batch training	Cost parameter	$C = 2$
Optimization algorithm	Scaled conjugate gradient with initial lambda = 0.0000005, initial sigma = 0.00005, interval center = 0 and interval offset = ±0.5		

2.3.1. *BLR model.* BLR is a statistical process for data analysis where one or more explanatory variables are used to determine a nominal dichotomous outcome (success/failure). The log of $odds = \frac{probability\ of\ success}{probability\ of\ failure}$ is used as dependent variable in BLR model that can predict the probability of a particular outcome by fitting data to a logit function, $f(T) = \frac{1}{1+e^{-T}}$, where $-\infty \leq T \leq \infty$. $f(T)$ increases monotonically from 0 to 1 as T varies from $-\infty$ to ∞. In this study BLR model was constructed using SPSS version 17. The value of the response variable considered to be one when the respondent expressed a positive willingness towards VAT payment, and zero for a negative willingness to pay VAT. The odds ratios of the independent variables and the 95% confidence intervals for odds ratios were computed. To determine the goodness-of-fit for the BLR model the Hosmer and Lemeshow test was also performed.

2.3.2. *ANN model.* An ANN is a biologically stimulated computational model constructed from hundreds of single units, known as artificial neurons, linked with coefficients/weights that create the neural structure [15-17]. In general, an ANN can be divided into the three layers: input layer, hidden layer and output layer [18-20]. The most commonly used ANN is multilayer feedforward networks. The multilayer feedforward networks use the Multi-Layer Perceptron (MLP) or the Radial Basis Function (RBF) which is a function of predictors that minimize the prediction error of target variables [18]. SPSS version 17 was used in this study to construct the MLP neural network model. Several trials were conducted to configure the best ANN model with the simplest possible structure. Before conducting the experiment, all covariates were standardized. The data were randomly assigned to training (69%) and testing (31%) datasets. The optimized neural network of this study consists of twelve input neurons, one intermediate hidden layer with eight neurons and one output layer with two neurons. The hyperbolic tangent activation function was used for the hidden layer and the softmax activation function was used for output layer. Batch, online and Mini-batch are three training type that determines how the network processes the records. The Batch training type was selected in this study because it directly reduces the total error [21]. Scaled conjugate gradient and gradient descent are two optimization algorithms used in MLP neural network to estimate the synaptic weights. The scaled conjugate gradient algorithm was selected in this study because using this algorithm to train MLP network is computationally efficient [22,23]. Initial lambda, initial sigma, interval center a_0 and interval offset a are the four parameters used in the scaled conjugate gradient algorithm [24]. The value for these four parameters was set as: initial lambda = 0.0000005, initial sigma = 0.00005, $a_0 = 0$ and $a = \pm 0.5$.

2.3.3. *SVM model.* The SVM algorithm was first developed from statistical learning theory by Vapnik and co-workers at AT&T Bell Laboratories [25-32]. In [28], for the recognition of SVM, Vapnik presented the ε-insensitive loss function defined as $|y - f(x)|_\varepsilon = \max\{0, |y - f(x)| - \varepsilon\}$ for which errors less than ε are not castigated. The SVM procedure considers training data points $(x_1, y_1), (x_2, y_2), \ldots, (x_i, y_i)$ from i samples, where $x_i \in \mathbb{R}^m$ are the input vectors of m-dimension and y_i represent the target variables. The primary goal in SVM is to construct a function $f(x)$ that has ε deviation from the expected response (y_i) based on the training data. A linear function of the form in (1) was considered to describe the mathematical basis of SVM

$$f(x) = <\boldsymbol{\omega}.\boldsymbol{x}> + b \tag{1}$$

where $<\boldsymbol{\omega}.\boldsymbol{x}>$ represents a dot product of the weight vector ω and input x in $\mathbb{R}^m$ space, and b is the bias term. To have least deviation/error (ε), the least weight vector needs to be determined by minimizing the Euclidean norm $||\omega||^2$ and by evaluating some acceptable deviations or non-negative slack variables (ξ_i and ξ_i^*) to describe the training data that lies outside the loss function. This can be expressed by the following formula (Equation (2))

$$\text{minimize } \frac{1}{2}||\omega||^2 + C\sum_{i=1}^{m}(\xi_i + \xi_i^*) \text{ subject to } \begin{cases} y_i - <\boldsymbol{\omega}.\boldsymbol{x}> - b \le \varepsilon + \xi_i \\ <\boldsymbol{\omega}.\boldsymbol{x}> + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i \text{ and } \xi_i^* \ge 0, \end{cases} \tag{2}$$

for all $i = 1, 2, \ldots, m$. The application of Lagrange multipliers (η_i, η_i^*) for effective dual space transformation extends to the function introduced in Equation (3) (for details: [10,33])

$$f(x, \eta) = \sum_{i=1}^{m}(\eta_i^* - \eta_i)K(x_i, x) + b \tag{3}$$

Here K is the kernel function that prepares non-linear problems easy to be solved linearly. Linear, polynomial and Gaussian or Radial Basis Function (RBF) are three frequently used kernel functions [34,35]. The statistical software R was used in this study to construct SVM model. The data were randomly assigned to training (70%) and testing (30%) subsets. In SVM model, the parameters kernel type and cost ware tuned, and the prediction accuracies were evaluated. The maximum accuracy 99.26% was attained with the polynomial kernel type (shown in Figure 1). For Polynomial kernel, degree = 3, scale = 10 and $C = 2$ was used for best optimal model (shown in Table 2).

3. **Result and Discussion.** This section describes the result and discussion of this study.

3.1. **BLR model.** The constructed BLR model is statistically significant, $\chi^2(11, N = 405) = 146.589$, $p < 0.001$, indicating that the model is able to distinguish between respondents who are willing to pay VAT and not willing to pay VAT [36]. The model

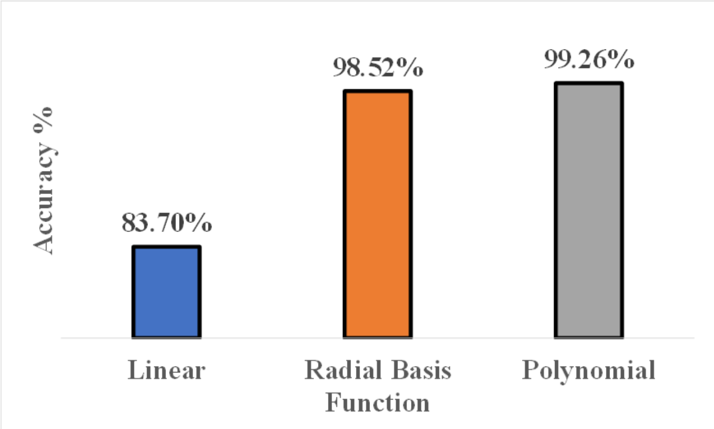

FIGURE 1. SVM kernel types and their accuracies

TABLE 3. Logistic regression predicting likelihood of willingness towards VAT

	β	S.E.	Wald	*df*	*p*-value	Odds ratio	95% C.I. for odds ratio	
							Lower	Upper
Age	.014	.030	.213	1	.644	1.014	.956	1.075
Monthly income	.000	.000	7.881	1	**.005**	1.000	1.000	1.000
Employment status (yes)	−.105	.425	.061	1	.805	.901	.391	2.073
VAT effect individuals spending (yes)	−.058	.405	.020	1	.887	.944	.427	2.089
VAT effect monthly savings (yes)	−1.371	.524	6.857	1	**.009**	.254	.091	.708
Buy cheaper alternatives after VAT (yes)	−.393	.357	1.210	1	.271	.675	.335	1.360
VAT reduced social obligations (yes)	.384	.412	.865	1	.352	1.468	.654	3.294
Aware about VAT (yes)	−1.037	.358	8.397	1	**.004**	.355	.176	.715
VAT will have positive impact on KSA (yes)	3.645	.594	37.688	1	**.000**	38.275	11.955	122.538
Social status			7.456	2	.024			
Social status (Single)	−1.653	.605	7.456	1	**.006**	.191	.058	.627
Social status (Married)	−19.140	9369.17	.000	1	.998	.000	.000	
Constant	−2.380	.861	7.641	1	.006	.093		

correctly classified 85.19% of cases. The p-value of the Hosmer and Lemeshow goodness-of-fit test is 0.089 which is larger than the significance level 0.05, suggests that there is not enough evidence to conclude that the model does not fit the data [35].

As shown in Table 3, four factors (VAT effect monthly savings, aware about VAT, VAT will have positive impact on Saudi Arabia and social status) have a statistically significant contribution in the model (p-value < 0.05). The 95% confidence interval for these four factors does not contain the value of 1 indicating that these odds ratios are statistically significant [6,36,37]. VAT will have a positive impact on Saudi Arabia is the strongest predictor of willing to pay VAT (p-value < 0.000, odds ratio $= 38.275$). This indicates that a person who thinks that VAT will have a positive impact on KSA is over 38 times more likely willing to pay VAT than those who thinks VAT will not have a positive impact on KSA.

3.2. **Artificial neural network model.** The confusion matrix (shown in Table 4) and the following equations were used in this study to estimate the performance of the model.

1) $Accuracy = (TP + TN)/total = (69 + 330)/405 = 0.9852$
2) $Recall\ (Sensitivity) = TP/(TP + FN) = 69/(69 + 3) = 0.9583$
3) $Precision\ (Specificity) = TP/(TP + FP) = 69/(69 + 3) = 0.9583$
4) $F1\ Score = \frac{2*(Recall*Precision)}{(Recall+Precision)} = \frac{2*(0.9583*0.9583)}{(0.9583+0.9583)} = 0.9583$

TABLE 4. Confusion matrix for logistic regression model, ANN model and SVM model

Logistic regression				ANN				SVM (Polynomial kernel)			
Observed	Predicted			Observed	Predicted			Observed	Predicted		
	No	Yes	Percent correct		No	Yes	Percent correct		No	Yes	Percent correct
No	312	21	93.69%	No	330	3	99.09%	No	333	3	99.11%
Yes	39	33	45.83%	Yes	3	69	95.83%	Yes	0	69	100%
Overall accuracy			85.19%	Overall accuracy			98.52%	Overall accuracy			99.26%

The effect of each predictor in the ANN model based on the relative and normalized importance was analyzed and the result is shown in Figure 2. As shown in Figure 2, the monthly income of the respondents has the greatest effect on the willingness towards VAT. VAT will have positive impact on KSA, age, effect of VAT on monthly savings and social status are other four major determinants of this model.

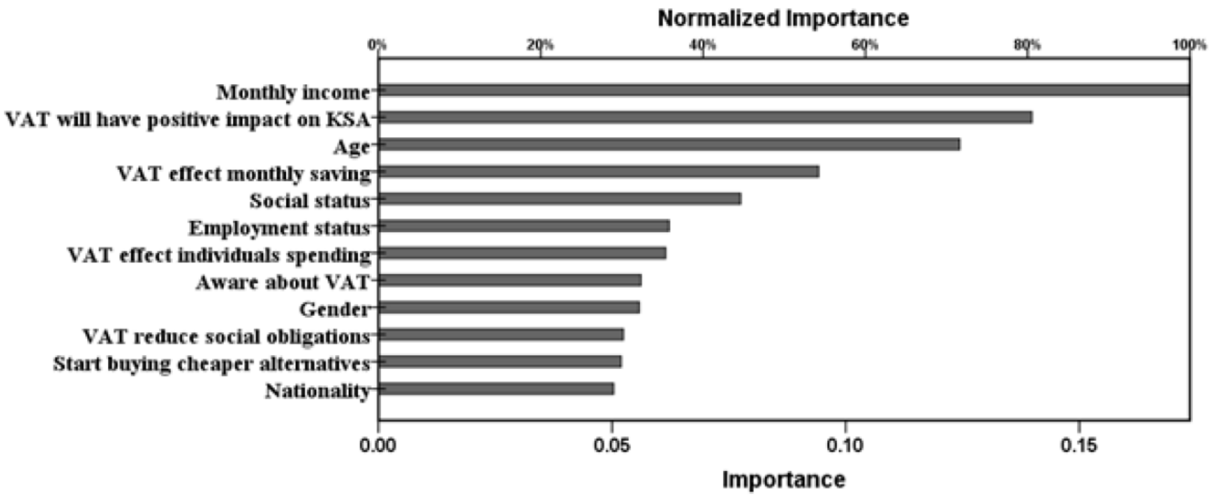

FIGURE 2. Independent variable importance chart in ANN model

3.3. **Support vector machine.** As the ANN model, the following equations and the confusion matrix were used to evaluate the performance of the SVM model.

1) $Accuracy = (TP + TN)/total = (69 + 333)/405 = 0.9926$
2) $Recall\ (Sensitivity) = TP/(TP + FN) = 69/(69 + 3) = 0.9583$
3) $Precision\ (Specificity) = TP/(TP + FP) = 69/(69 + 0) = 1$
4) $F1\ Score = \frac{2*(Recall*Precision)}{(Recall+Precision)} = \frac{2*(0.9583*1)}{(0.9583+1)} = 0.9787$

Figure 3 presents the impact of each predictor in the SVM model. As shown in Figure 3, VAT will have positive impact on KSA has the greatest effect on the willingness towards VAT as BLR model. The other four major determinants of SVM predictive model are aware about VAT, start buying cheaper alternative, age, effect of VAT on monthly savings.

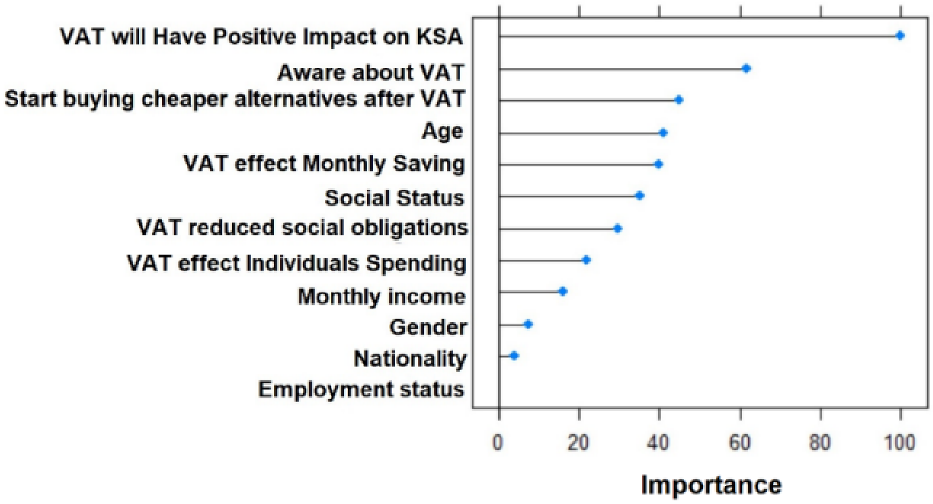

Figure 3. Independent variable importance chart in SVM model

4. **Conclusions.** BLR, ANN and SVM are three distinguished machine learning techniques. All these techniques have been recognized to perform extremely well in numerous fields. The aim of this study was to apply these machine learning techniques to exploring the acceptance of Saudi residents towards newly implemented VAT system. A random sample of participants representing the Kingdom of Saudi Arabia was studied in this paper. The BLR, ANN and SVM models were constructed to predict the willingness towards VAT and have been compared against each other to determine the best predictive model. The dataset was first preprocessed, and 10-fold cross validation was applied as the procedure for dividing the training and testing datasets. The optimized parameters for ANN and SVM were determined. Considerable experiments were conducted employing various values of the parameters for these procedures. The results indicate that the SVM model performs better with prediction accuracy 99.26%; however, the prediction accuracy for ANN and BLR are 98.52% and 85.19%, respectively. Moreover, based on the three proposed model in this study, the factor "VAT will have positive impact on Saudi Arabia" is one of the most important predictors that positively affect the willingness of the Saudi residents towards VAT. The other five significant predictors of the willingness towards VAT are VAT effect monthly savings, the awareness about VAT, monthly income, start buying cheaper alternatives after VAT implementation and social status. The implementation of these model would help to improve the acceptance towards newly implanted VAT system if the government puts more focus on the predictors. Future research works are expected to improve the VAT system in Saudi Arabia and all other GCC countries as well by expanding the scope of the survey both in terms of the sample size and the questionnaire.

REFERENCES

[1] *General Authority of Zakat and Tax (GAZT), Kingdom of Saudi Arabia*, https://www.vat.gov.sa/en, 2019.

[2] R. G. Brereton and G. R. Lloyd, Support vector machines for classification and regression, *Analyst*, vol.135, no.2, pp.230-267, 2010.

[3] T. Ahmad and M. N. Aziz, Data preprocessing and feature selection for machine learning intrusion detection systems, *ICIC Express Letters*, vol.13, no.2, pp.93-101, 2019.

[4] D. Wang, B. Zhang, W. Zhou and W. Song, Interval-valued financial time series prediction based on improved Elman neural network, *ICIC Express Letters*, vol.13, no.2, pp.159-166, 2019.

[5] D. R. Cox, The regression analysis of binary sequences (with discussion), *J. R. Stat. Soc.*, vol.Series B, pp.215-242, 1958.

[6] D. S. Moore, G. P. McCabe and B. A. Craig, *Introduction to the Practice of Statistics*, 9th Edition, W. H. Freeman, 2016.

[7] A. J. M. Sufian, *Methods and Techniques of Social Research*, The University Press Ltd., 1998.

[8] A. Prasad and J. B. Edward, Importance of artificial neural networks for location of faults in transmission systems: A survey, *The 11th International Conference on Intelligent Systems and Control (ISCO)*, pp.357-362, 2017.

[9] T. S. Furey, N. Cristianini, N. Duffy, D. W. Bednarski, M. Schummer and D. Haussler, Support vector machine classification and validation of cancer tissue samples using microarray expression data, *Bioinformatics*, vol.16, no.10, pp.906-914, 2000.

[10] T. O. Owolabi, K. O. Akande and S. O. Olatunji, Estimation of surface energies of hexagonal close packed metals using computational intelligence technique, *Appl. Soft Comput.*, vol.31, pp.360-368, 2015.

[11] K. O. Akande, S. O. Olatunji, T. O. Owolabi and A. AbdulRaheem, Feature selection-based ANN for improved characterization of carbonate reservoir, *SPE Saudi Arabia Section Annual Technical Symposium and Exhibition*, 2015.

[12] C. Cortes and V. Vapnik, Support-vector networks, *Machine Learning*, vol.20, no.3, pp.273-297, 1995.

[13] F. E. Gunawan, Improving the reliability of F-statistic method by using linear support vector machine for structural health monitoring, *ICIC Express Letters*, vol.12, no.12, pp.1183-1193, 2018.

[14] Z. Wang, Q. Zeng and M. Lv, Modeling the external truck arrivals in container terminals based on DBN and SVM, *ICIC Express Letters*, vol.12, no.10, pp.1033-1040, 2018.

[15] J. M. Zurada, *Introduction to Artificial Neural Systems*, West Publishing Co. St. Paul, MN, USA, 1992.

[16] J. Gasteiger, X. Li, V. Simon, M. Novič and J. Zupan, Neural nets for mass and vibrational spectra, *J. Mol. Struct.*, vol.292, pp.141-159, 1993.

[17] S. Agatonovic-Kustrin and R. Beresford, Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research, *J. Pharm. Biomed. Anal.*, vol.22, no.5, pp.717-727, 2000.

[18] I. N. da Silva, D. H. Spatti, R. A. Flauzino, L. H. B. Liboni and S. F. dos Reis Alves, *Artificial Neural Networks*, Springer International Publishing, Cham, 2017.

[19] B. D. Ripley, *Pattern Recognition via Neural Networks*, Cambridge University Press, 1996.

[20] D. A. Asfani, A. F. Ilman, N. W. A. Sanjaya, I M. Y. Negara, D. Fahmi, D. R. Sawitri, M. Wahyudi and H. L. Al-Azmi, The modelling of low voltage arc flash based on artificial neural network, *International Journal of Innovative Computing, Information and Control*, vol.14, no.4, pp.1389-1405, 2018.

[21] M. F. Møller, A scaled conjugate gradient algorithm for fast supervised learning, *Neural Networks*, vol.6, no.4, pp.525-533, 1993.

[22] T. Marwala, Introduction to finite-element-model updating, in *Finite-Element-Model Updating Using Computional Intelligence Techniques*, London, Springer London, 2010.

[23] N. Z. Zacharis, Predicting student academic performance in blended learning using artificial neural networks, *Int. J. Artif. Intell. Appl.*, vol.7, no.5, 2016.

[24] R. Rojas, *A Systematic Introduction*, Springer-Verlag, 1996.

[25] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, 1995.

[26] J. Li, B. Zhang, J. Shi, J. Li, B. Zhang and J. Shi, Combining a genetic algorithm and support vector machine to study the factors influencing CO_2 emissions in Beijing with scenario analysis, *Energies*, vol.10, no.10, p.1520, 2017.

[27] A. A. Adewumi, T. O. Owolabi, I. O. Alade and S. O. Olatunji, Estimation of physical, mechanical and hydrological properties of permeable concrete using computational intelligence approach, *Appl. Soft Comput.*, vol.42, pp.342-350, 2016.

[28] I. O. Alade, T. A. Oyehan, I. K. Popoola, S. O. Olatunji and A. Bagudu, Modeling thermal conductivity enhancement of metal and metallic oxide nanofluids using support vector regression, *Adv. Powder Technol.*, vol.29, no.1, pp.157-167, 2018.

[29] Z. Yuan and B. Huang, Prediction of protein accessible surface areas by support vector regression, *Proteins Struct. Funct. Bioinforma.*, vol.57, no.3, pp.558-564, 2004.

[30] N. S. Raghavendra and P. C. Deka, Support vector machine applications in the field of hydrology: A review, *Appl. Soft Comput.*, vol.19, pp.372-386, 2014.

[31] R. Collobert and S. Bengio, SVMTorch: Support vector machines for large-scale regression problems, *J. Mach. Learn. Res.*, vol.1, no.2, pp.143-160, 2001.

[32] M. O. Stitson, A. Gammerman et al., Support vector regression with ANOVA decomposition kernels, in *Advances in Kernel Methods*, MA, USA, MIT Press Cambridge, 1999.

[33] T. O. Owolabi, K. O. Akande and S. O. Olatunji, Development and validation of surface energies estimator (SEE) using computational intelligence technique, *Comput. Mater. Sci.*, vol.101, pp.143-151, 2015.

[34] T. O. Owolabi and M. Gondal, Novel techniques for enhancing the performance of support vector regression chemo-metric in quantitative analysis of LIBS spectra, *J. Anal. At. Spectrom.*, 2017.

[35] J. B. Gao, S. R. Gunn, C. J. Harris and M. Brown, A probabilistic framework for SVM regression and error bar estimation, *Machine Learning*, vol.46, nos.1-3, pp.71-89, 2002.

[36] P. Julie, *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS*, 4th Edition, Open University Press, 2010.

[37] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, 7th Edition, Pearson Education, 2013.