Indonesian Journal of Electrical Engineering and Computer Science
Vol. 16, No. 2, November 2019, pp. 701~708
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v16.i2.pp701-708

Artificial neural network and partial least square in predicting blood hemoglobin using near-infrared spectrum

Mohd Nazrul Effendy Mohd Idrus[1], Kim Seng Chia[2]
[1,2]Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Malaysia
[2]Sensor Devices Research Laboratory, MiNT-SRC, Universiti Tun Hussein Onn Malaysia, Malaysia

Article Info

Article history:

Received Jan 26, 2019
Revised May 9, 2019
Accepted May 22, 2019

Keywords:

Artificial neural network
Hybrid model
Near infrared spectroscopy
Partial least square

ABSTRACT

Predictive models are crucial in near-infrared (NIR) spectroscopic analysis. Partial least square - artificial neural network (PLS-ANN) is a hybrid method that may improve the performance of prediction in NIR spectroscopic analysis. This study investigates the advantage of PLS-ANN over the well-known linear and non-linear modelling approaches in spectroscopy analysis that are partial least square (PLS) and artificial neural network (ANN). The results show that ANN that coupled with first order SG derivatives achieved the best prediction with root mean square error of prediction (RMSEP) of 0.3517 gd/L and coefficient of determination (R_p^2) of 0.9849 followed by PLS-ANN with RMSEP of 0.4368 gd/L and R_p^2 of 0.9787, and PLS with RMSEP of 0.4669 gd/L and R_p^2 of 0.9727. This suggests that the spectrum information may unable to be totally represented by the first few latent variables of PLS and a nonlinear model is crucial to model these nonlinear information in NIR spectroscopic analysis.

Corresponding Author:

Kim Seng Chia,
Faculty of Electrical and Electronic Engineering,
Universiti Tun Hussein Onn Malaysia, Malaysia.
Email: kschia@uthm.edu.my

1. INTRODUCTION

Nowadays, near-infrared spectroscopy (NIRS) technology showed increasing number of applications in various fields such as medical, chemical, and food analysis [1-3]. Applications that based on NIRS are developed to overcome several factors in the conventional methods which are time-consuming, destructive, and cost-effective. The intensity of the reflection and transmission of the fundamental molecular vibrations of C–H, O–H, and N–H produces absorption bands make near-infrared (NIR) useful for analyzing in the biological system [4]. By using a working range of the electromagnetic spectrum (from about 780 nm to 2500 nm), NIRS produces a spectral data which limit the usage based on several factors such as baseline drift, interference resulting in a poor signal to noise ratio, improper wavelength selection, thermal noise, and calibration issues [5]. However, these problems of nonlinearity in spectral data can be solved by using an appropriate predictive modelling and data preprocessing methods.

Hybridization or combination of two predictive modellings is frequently used in spectroscopic analysis to improve the performance of the modelling. The main purpose of these techniques used is to improve weakness that occurs within single models such as nonlinearity, redundant spectral band, and wavelength selection problems. Evolving of predictive modelling technique was a better solution to improve the accuracy of performance compared than recreating new modelling [6]. Furthermore, hybrid models were used to reduce the risk of failure by using the single model by combining several models to obtain more accurate results [7]. However, the hybrid models are not an easy process to be developed and need a deep knowledge to manipulate the models and avoid the wrong parameter to be selected.

For instance, artificial neural network (ANN) combined with multiple linear regression (MLR) has been developed to overcome the linear modelling deficiency of the traditional artificial neural networks [7]. The proposed models showed a good result in classification for both synthetic and real-life benchmark data sets and the model consistently outperforms with other predictive models such as multilayer perceptron, linear discriminant analysis, quadratic discriminant analysis, K-nearest neighbour, and support vector machines. However, MLR modelling is not a powerful linear model and can be replaced with another better linear model such as PLS. Genetic algorithm (GA) coupled with PLS showed the capability of retrieving components of interest from spectral data [8]. Furthermore, GA combined with ANN shows a reliable method in classifying an egg's freshness [9]. In addition, ANN hybrid with GA was established to predict travel agency air ticket sales revenue [10]. By using genetic operators such as reproduction, mutation and, selection, GA creates a new generation of the population which is better than the generation before. However, with a large number of spectral data, GA method for selection of variable spectral data could lead to a risk of overfitting. PLS combined with back propagation neural network (BPNN) showed effectiveness in overcoming the problems of redundant and nonlinear in spectral data [11]. Number of latent variables (LVs) generated from PLS was used as the input of BPNN to estimate the abundance of minerals on the lunar surface. Furthermore, PLS combined with BPNN achieved the best performance compared with PLS and GA-PLS. However, optimization of a number of LVs and hidden neurons used in this research is not briefly described.

Next, the quality of different models of PLS, ANN, and PLS-ANN to predict consumer interest rating of ready to drink green tea beverages were investigated [9]. PLS-ANN showed a better quality compared to PLS in coefficient of determination and mean square error value. However, the number of variable in the research which is 8 factors is very small to be compared with a variable in NIRS analysis. Whereas, the working range of NIRS is from 780 nm to 2500 nm have 390 to 1250 variables with a 2 nm interval. A combination of PLS and ANN also was applied in management and chemical analysis showed satisfy result [11, 12]. Even though the limitation of of PLS has been rectified using the PLS-ANN, there is no comparison among PLS, ANN, and PLS-ANN. Therefore, this study compares three types of predictive modelling i.e. linear model (PLS), nonlinear model (ANN), and a hybrid model (PLS-ANN) in predicting hemoglobin concentration using near-infrared spectral data. We also investigated the optimal number of LVs and hidden neurons to achieve optimal prediction of performance.

2. MATERIAL AND METHODS

2.1. Samples and References

The origin spectral dataset was adopted from IDRC shootout 2010 provided by Karl Norris. The Blood samples were analyzed with a NIRSystems 6500 spectrometer from 1990 to 1992. All spectral data for calibration and testing have 700 variables, were measured with a range of wavelengths that span the infrared spectrum from 1100 to 2498 nm wavelength with a 2 nm interval. The dataset contains 231 sets of calibration and 194 testing data sets to measure predictive accuracy of the modelling. Table 1 shows the summary of descriptive characteristics of the calibration and testing references data. The characteristics of calibration and testing data indicate that extrapolation samples were used for modelling. Therefore, an appropriate modelling technique needs to be implemented to predict the out of range data from the testing process. IDRC shootout 2010 has used the same extrapolation data in the tournament [13].

Table 1. Descriptive statistics of the blood hemoglobin

Data	n	Min (g/dL)	Max (g/dL)	Mean (g/dL)	Median (g/dL)	Std
Calibration	231	10.30	17.30	13.78	13.70	1.66
Testing	194	6.50	18.20	12.20	12.25	2.83
Total	425	6.50	18.20	13.06	13.40	2.40

2.2. Data Preprocessing

Savitzky-golay (SG) preprocessing method was used as a pretreatment process to remove unwanted signals such as signal to noise ratio, baseline shift effect, and slope affected from spectral data. By using a specific odd number of frame length, a single set of SG coefficient with the same number of frame length can be applied to all subsets' data to evaluate new smoothed or derivatives signal of central point data of each subset. Moreover, an optimal number of frame length should be optimized to achieve an optimum of prediction performance [13, 14]. Raw spectral data of hemoglobin was treated with a different type of SG preprocessing such as smoothing, first order, and second order sg derivatives [15]. SG coefficient can be applied to each subset data to obtain new treated data by using Equation (1).

$$Y_j = (C \otimes y)_j = \sum_{i=-\frac{m-1}{2}}^{\frac{m-1}{2}} C_i Y_{i+j} \, , \frac{m-1}{2} \leq j \leq n - \frac{m-1}{2} \quad (1)$$

where m and n is measured number frame length and total number of variables. Although, C_i is the set of sg coefficient. While, Y_{i+j} and Y_j is a related set of data before and observed value after the treatment process.

2.3. Partial Least Square

General concept idea behind of PLS modelling is to decompose both the design matrix predictor X and matrix of response Y as Equations (2) and (3)

$$X = TP^T \quad (2)$$

$$Y = UQ^T \quad (3)$$

where X is an $n \times m$ matrix of predictors, Y is an $n \times p$ matrix of response.T and U is $n \times l$ matrix that are projections of X score and Y score respectively. P and Q are $m \times l$ and $p \times l$ orthogonal loading matrices respectively. The algorithm will yield the PLS regression estimates B and B_o after estimating the factor and loading matrices T, U, P and Q for the linear regression as Equations (4)

$$Y = XB + B_o \quad (4)$$

where B and B_o is PLS regression coefficient. In this research, the coefficients of PLS regression were generated by using the MATLAB matrix routines function. Latent variable (LVs) can be extracted from PLS algorithm after the model was complete. The NIPALS algorithm is the default algorithm applied in PLS [16]. The LVs from NIPALS algorithm were further fed to ANN as the input to predict blood hemoglobin concentrations.

2.4. Artificial Neural Network

ANN as one of the nonlinear methods proved their superiority over linear method for making a calibration model [17]. In this research, multi-layer perceptron (MLP) trained with backpropagation algorithm was used as the ANN configuration [18, 19]. MLP has an input layer of source nodes, hidden layer of neurons, and an output layer of outcome network. Each neuron in hidden layer will receive input parameters of latent variables (LVs) from PLS model. Here, we have measured 1 to 50 number of LVs from PLS model and 1 to 10 number of hidden neurons in order to obtain optimal performance of predictions. Outputs of each ith neuron x_i in hidden layer were established by multiplying the corresponding weights with input parameters before passed through a transfer function to get the desired output. The process of transfering information between the input layer and hidden layers is as in Equation (5).

$$x_k = f_1(\sum_{j=1}^{n} W_{1,k,j}^{i} LVs_j + b_1) \quad (5)$$

where f_1 is the hidden layer transfer function, $W_{1,k,j}^{i}$ is the weight of the kth node in the hidden layer connected with the jth node in the input layer. While b_1 is the bias between hidden layer and input layer. In this study, we use tan-sigmoid transfer functions for hidden layer and linear transfer function for output layer to receive sums of weighed and bias input. To optimize the random initial weights, the network was trained 1000 times to achieve global prediction performance [20]. Levenberg Marquardt (LM) backpropagation algorithm was selected as training algorithm in this study [21]. The training process will stop when either the maximum number of epochs is reached, the goal performance is achieved, the performance of gradient is below minimum gradient value, the momentum update is exceeded, or the failure validation is more than the maximum amount.

2.5. Performance Validation

Training and testing performance of prediction were evaluated using root mean squared error of calibration (RMSEC) and root mean squared error of prediction (RMSEP), respectively. RMSEC and RMSEP are computed using Equation (6).

$$RMSECV \; or \; RMSEP = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \quad (6)$$

where n is the total number of samples, while $\hat{y}_i$ and y_i denote the predicted blood hemoglobin and reference blood hemoglobin, respectively. To interpret proportion of the variance in the predicted data from the reference value and to describe the relationship between blood hemoglobin and near-infrared spectrum, the coefficient of determination (R^2) was used as Equation (7).

$$R_c^2 \text{ or } R_p^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (7)$$

where $\bar{y}$ is mean of reference data, $\hat{y}_i$ and y_i denote the unseen predicted and reference blood Hb, respectively.

3. RESULTS AND DISCUSSION

3.1. Latent Variable

Figure 1 shows the value of RMSEP of PLS and PLS-ANN when a different number of latent variables (LVs) were applied with a different type of SG preprocessing. PLS-ANN without SG preprocessing tends to reach minimum 0.4368 gd/L of RMSEP when 23 number of the LVs was used. After that, the network becomes overfit when increasing more than 23 number of LVs. Meanwhile, PLS without SG preprocessing shows the further low performance prediction with 0.5307 gd/L of RMSEP when 14 number of LVs were used. However, there are not many different values of RMSEP for PLS and PLS-ANN when the number of LVs used is between 11 and 16. The effects of SG preprocessing can be seen when smoothing SG preprocessing coupled with PLS-ANN has improved its prediction performance with 0.4208 gd/L of RMSEP when 26 number of LVs used. In addition, smoothing SG coupled with PLS shows improvement with 0.5114 gd/L of RMSEP when 17 number of LVs used. However, the model became overfit when more than 17 number of LVs were used. Meanwhile, first order SG derivatives coupled with PLS-ANN shows an improvement of prediction when achieves 0.4089 gd/L of RMSEP when 27 number of LVs were used. An improvement also occurs when first order SG derivatives coupled with PLS tends to achieve 0.5024 gd/L of RMSEP when 11 number of LVs were used. PLS coupled with second order SG derivative shows increasing of prediction when 0.4669 gd/L of RMSEP when 14 number of LVs were used. However, PLS-ANN coupled with second order SG derivative shows decreasing of prediction when 0.4295 gd/L of RMSEP when 19 number of LVs were used.

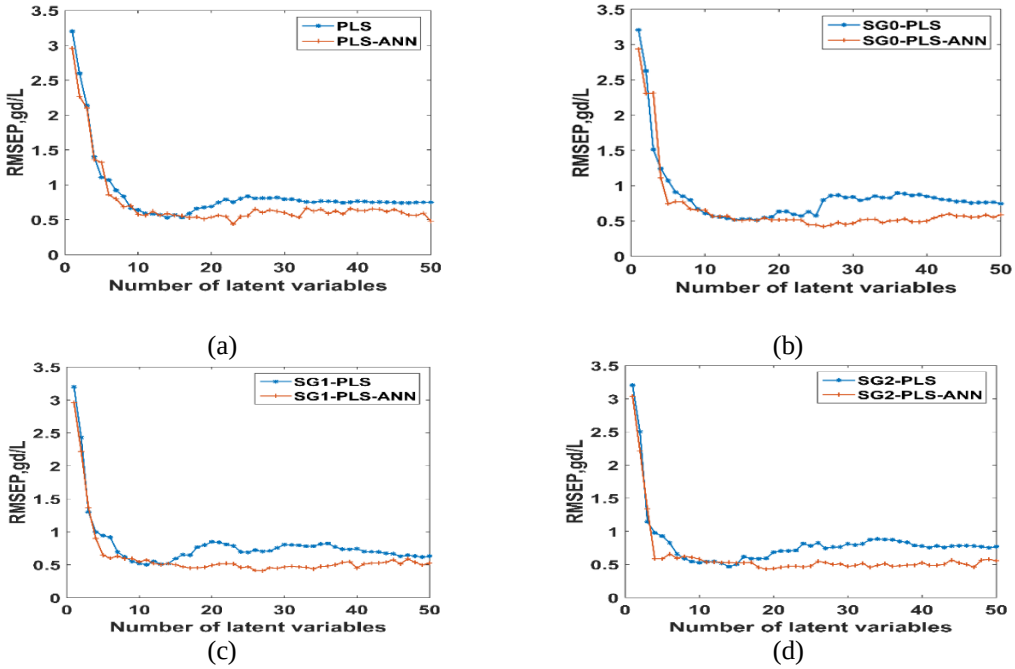

Figure 1. The root mean square error of prediction of artificial neural network (ANN) and partial least square – artificial neural network (PLS-ANN) versus the change of a number of hidden neurons with different Savitzky-golay (SG) preprocessing: (a) without SG preprocessing, (b) smoothing SG, (c) first order SG derivative, and (d) second order SG derivative

From the results, we can summarize that PLS coupled with or without SG preprocessing tends to achieve optimal prediction of performance when 11 to 17 number of LVs were used. While PLS-ANN coupled with or without SG preprocessing tends to achieve optimal prediction of performance with a high number of a LVs when 19 to 26 number of LVs were used. This indicated that the number of LVs used in PLS-ANN need to be optimized to achieve optimal prediction of performance. These results also indicated that PLS-ANN outperforms PLS performance of prediction.

We also observed that with increasing number of LVs, PLS-ANN tends to maintain the value of RMSEP compared to PLS. For instance, PLS-ANN without preprocessing overfitted when 23 number of LVs were used. However, the overfit percentage of PLS-ANN without preprocessing (increasing 49.9% from 0.4368 to 0.6547 gd/L of RMSEP) is lowest compared to PLS without preprocessing (increasing 57.7% from 0.5307 to 0.8369 gd/L of RMSEP). In addition, the overfit percentage of PLS-ANN with smoothing SG (increasing 21.7% from 0.4208 to 0.5122 gd/L of RMSEP) is lowest compared to PLS with first order SG (increasing 67.9% from 0.5114 to 0.8588 gd/L of RMSEP). Furthermore, the overfit percentage of PLS-ANN with first order SG derivatives (increasing 33.9% from 0.4089 to 0.5477 gd/L of RMSEP) is lowest compared to PLS with first order SG derivatives (increasing 69.2% from 0.5024 to 0.85 gd/L of RMSEP). Then, the overfit percentage of PLS-ANN with second order SG derivatives (increasing 27.4% from 0.4295 to 0.5473 gd/L of RMSEP) is lowest compared to PLS with second order SG derivatives (increasing 77% from 0.4669 to 0.8263 gd/L of RMSEP). These findings indicated that PLS-ANN has the minimum potential to be overfitted compared to PLS. Moreover, PLS-ANN managed to obtain a better result compared to PLS. This shows that feedforward backpropagation of ANN has had a good impact on improving prediction on PLS model. These findings agree with previous research that PLS-ANN results in a better performance than that of PLS [10, 21].

3.2. Hidden Neurons

Figure 2 shows the RMSEP of PLS and PLS-ANN when different number of hidden neurons was applied with different type of SG preprocessing. PLS-ANN without SG preprocessing tends to achieve optimum performance of prediction with 0.4368 gd/L of RMSEP when 5 number of hidden neurons were used. After that, the network is overfitting when more than 5 number of hidden neurons were used. Meanwhile, ANN without SG preprocessing is able to achieve optimum performance of prediction with 0.4607 gd/L of RMSEP when 3 number of hidden neurons were used. Smoothing SG coupled with ANN and smoothing SG coupled with PLS-ANN tends to achieve optimal prediction performance when 2 number of hidden were used with 0.4208 gd/L and 0.4499 gd/L of RMSEP respectively. However, the network is overfitted when more than 2 number of hidden neurons were used. First order SG derivatives coupled with ANN show a better result with optimum performance of prediction with 0.4089 gd/L of RMSEP when 5 number of hidden neurons was used compared with first order SG derivatives coupled with PLS-ANN.

Second order SG derivatives coupled with PLS-ANN tend to achieve optimal prediction performance 0.4295 gd/L of RMSEP when 2 number of hidden neurons were used. After that, the network is overfitted when more than 2 number of hidden neurons were used. Meanwhile, second order SG derivatives coupled with ANN achieve the lower performance of prediction with 0.4571 gd/L of RMSEP when 3 hidden neurons were used compared to second order SG derivatives coupled with PLS-ANN.

Despite with different SG preprocessing, PLS-ANN could obtain a better performance of prediction compared to ANN except first order SG derivatives preprocessing. However, first order SG derivatives coupled with ANN can be considered as the best model with the high performance of prediction 0.4089 gd/L of RMSEP when 5 number of hidden neurons were used. This finding indicates that the number of LVs was used as input of ANN is not enough to represent important information the whole spectral data itself. Therefore, ANN with directly receiving input from spectral data after treated tends to achieve a better performance of prediction compared to PLS-ANN with LVs as the inputs of ANN.

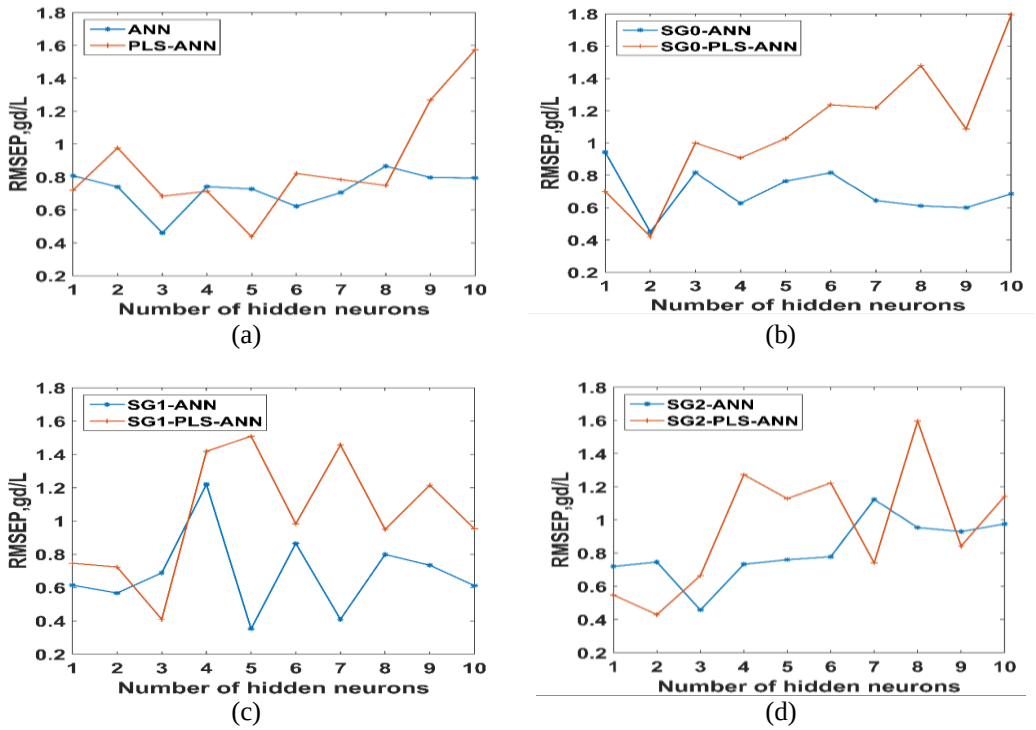

Figure 2. The root mean square error of prediction of artificial neural network (ANN) and partial least square – artificial neural network (PLS-ANN) versus the change of a number of hidden neurons with different Savitzky-golay (SG) preprocessing: (a) without SG preprocessing, (b) smoothing SG, (c) first order SG derivative, and (d) second order SG derivative

3.3. Prediction Performance

Table 2 shows the accuracy of the PLS, ANN, and PLS-ANN with different SG preprocessing methods using fixed number of frame length that is smoothing=77 nm, first order SG derivatives=27 nm, and second order SG derivatives=79 nm with the optimal number latent variables (LVs). Best prediction of performance was achieved by ANN coupled that with first order SG derivatives with RMSEP of 0.3517 gd/L and R_p^2 of 0.9849, followed by PLS-ANN that coupled with first order SG derivative with RMSEP of 0.4089 gd/L and R_p^2 of 0.9792, and then PLS that coupled second order SG derivative with RMSEP of 0.4669 gd/L and R_p^2 of 0.9727. This unusual result whereas conventional model tends to give a better result compared hybrid model is a normal finding in multivariate calibration analysis. For instance, auto regressive integrated moving average and ANN (ARIMA_ANN) are marginally showed better performance then hybrid ANN_ARIMA in Indian stock trend forecasting research [22]. Moreover, ANN showed a slightly higher performance in term of R^2 and lower RMSE compared to PLS-ANN in consumer liking scores of ready-to-drink green tea beverages predictions [9]. Furthermore, the criteria of datasets with extrapolation samples may influenced the prediction performance of different predictive modelling. However, PLS-ANN that coupled with smoothing, first order, second order SG derivative and without SG preprocessing showed satisfied result by dominating top best five of RMSEP that were 0.4208, 0.4089, 0.4295, and 0.4368 gd/L, respectively. Although ANN outperformed PLS-ANN, PLS-ANN is simpler than ANN with three hidden neurons compared to the former that needed five hidden neurons to achieve the optimal performance.

The result also indicated that the PLS model was found to have the lowest prediction performance while ANN and PLS-ANN hybrid models were found to have comparable qualities based on the values of R_p^2 and RMSEP. This result suggested that the linear PLS model cannot generalize nonlinearity of spectral data and these findings are in line with the results of the previous study [9]. While the characteristics of ANN modelling that can adapt and generalize data sets having non-linear relationships makes ANN achieve a better result in this research. In addition, PLS that coupled with second order SG derivative achieved the optimal accuracy compared to that coupled with smoothing and first order SG derivatives is in line with previous study [14]. Moreover, ANN that coupled with the first order SG derivative with a proper optimization was able to achieve a better predictive accuracy in predicting the blood hemoglobin using near-infrared spectral data with RMSEP of 0.3517 gd/L and of 0.9849, compared with the previous works [13].

Table 2. The accuracy of the PLS, ANN, and PLS-ANN with different SG preprocessing methods using the optimal tunable parameters of frame length, LVs, and hidden neurons

Modelling Method	SG preprocessing method	Frame length	Latent variable	Hidden neurons	Training		Prediction	
					RMSEC (gd/L)	R_c^2	RMSEP (gd/L)	R_p^2
PLS	Without preprocessing	N/A	14	N/A	0.3354	0.9588	0.5307	0.9648
	SG smoothing	77	17	N/A	0.3518	0.9547	0.5114	0.9673
	First order SG derivative	27	11	N/A	0.3335	0.9592	0.5024	0.9684
	Second order SG derivative	79	14	N/A	0.3287	0.9604	0.4669	0.9727
ANN	Without preprocessing	N/A	N/A	3	0.1830	0.9890	0.4607	0.9734
	SG smoothing	77	N/A	2	0.1115	0.9956	0.4499	0.9761
	First order SG derivative	27	N/A	5	0.1364	0.9938	0.3517	0.9849
	Second order SG derivative	79	N/A	3	0.1190	0.9950	0.4571	0.9746
PLS-ANN	Without preprocessing	N/A	23	5	0.1364	0.9933	0.4368	0.9787
	SG smoothing	77	26	2	0.1586	0.9908	0.4208	0.9784
	First order SG derivative	27	27	3	0.1473	0.9923	0.4089	0.9792
	Second order SG derivative	79	19	2	0.2554	0.9760	0.4295	0.9782

4. CONCLUSION

This study shows that the hybrid model (i.e. PLS-ANN) that combined linear model (i.e. PLS) and nonlinear model (i.e. ANN) achieved satisfying results in predicting the blood Hemoglobin using near-infrared spectral data. ANN that coupled with first order SG derivatives achieved the best prediction with RMSEP of 0.3517 gd/L and R_p^2 of 0.9849, followed by PLS-ANN that coupled with first order SG derivative with RMSEP of 0.4368 gd/L and R_p^2 of 0.9787, and PLS coupled second order SG derivative with RMSEP of 0.4669 gd/L and R_p^2 of 0.9727. This suggests that the related spectrum information may be excluded in the first few latent variables of PLS. Nevertheless, findings indicated that PLS-ANN was able to minimize the overfitting problem compared to PLS with fewer input variables. Furthermore, the structure of PLS-ANN is simpler than ANN, in which, the former needed three hidden neurons compared with the latter needed five hidden neurons in achieving their optimal prediction performance. Thus, a nonlinear model is crucial to model these nonlinear information in NIR spectroscopic analysis and more researches are required in understanding the potential of hybrid models in in NIR spectroscopic analysis.

ACKNOWLEDGMENTS

The author would like to acknowledge the Ministry of Education Malaysia and Research Management Centre (RMC), Universiti Tun Hussein Onn Malaysia (UTHM) for financial support GPP H312 and providing facilities for this study.

REFERENCES

[1] V. Gomes, et al., "Characterization of neural network generalization in the determination of pH and anthocyanin content of wine grape in new vintages and varieties," *Food Chemistry*, vol. 218, pp. 40-46, 2017.
[2] Y. Zhang, et al., "Comparision of Several Preprocessing Algorithms Based on Near Infrared Spectroscopic Measurement of Glucose in Aqueous Glucose Solutions," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 12, pp. 2683-2689, 2014.
[3] H. A. G. Al-kaf, et al., "A comparison between single layer and multilayer artificial neural networks in predicting diesel fuel properties using near infrared spectrum," *Petroleum Science and Technology*, vol. 6466, pp. 1-8, 2018.
[4] M. Pojić, et al., "The Application of Near Infrared Spectroscopy in Wheat Quality Control," *Infrared Spectroscopy - Life and Biomedical Sciences*, vol. 1, pp. 167-184, 2012.
[5] J. Yadav, et al., "Prospects and limitations of non-invasive blood glucose monitoring using near-infrared spectroscopy," *Biomedical Signal Processing and Control*, vol. 18, pp. 214-227, 2015.
[6] K. D. Chito, et al., "Incremental model learning for spectroscopy-based food analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 167, pp. 123-131, 2017.
[7] M. Khashei, et al., "A novel hybrid classification model of artificial neural networks and multiple linear regression models," *Expert Systems with Applications*, vol. 39, pp. 2606-2620, 2012.
[8] K. Song, et al., "Remote sensing of chlorophyll-a concentration for drinking water source using genetic algorithms (GA)-partial least square (PLS) modeling," *Ecological Informatics*, vol. 10, pp. 25-36, 2012.
[9] P. Yu, et al., "Development of a partial least squares-artificial neural network (PLS-ANN) hybrid model for the prediction of consumer liking scores of ready-to-drink green tea beverages," *Food Research International*, vol. 103, pp. 68-75, 2018.
[10] H. Huang, "A Hybrid Neural Network Prediction Model of Air Ticket," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 11, pp. 6413-6419, 2013.

[11] S. Li, et al., "Hybridization of partial least squares and neural network models for quantifying lunar surface minerals," *Icarus*, vol. 221, pp. 208-225, 2012.
[12] S. Hosseinpour, et al., "Exact estimation of biodiesel cetane number (CN) from its fatty acid methyl esters (FAMEs) profile using partial least square (PLS) adapted by artificial neural network (ANN)," *Energy Conversion and Management*, vol. 24, pp. 389-398, 2016.
[13] B. Igne, et al., "The 2010 IDRC software shoot-out at a glance," *NIR news*, vol. 21, pp. 14, 2010.
[14] K. S. Chia, et al., "Evaluation of common pre-processing approaches for visible (VIS) and shortwave near infrared (SWNIR) spectroscopy in soluble solids content (SSC) assessment," *Biosystems Engineering*, vol. 115, pp. 82-88, 2013.
[15] J. Luo, et al., "Properties of Savitzky-Golay digital differentiators," *Digital Signal Processing: A Review Journal*, vol. 15, pp. 122-136, 2005.
[16] S. Wold, et al., "PLS-regression: A basic tool of chemometrics," *Chemometrics and Intelligent Laboratory Systems*, vol. 58, pp. 109-130, 2001.
[17] R. M. Balabin, et al., "Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction," *Chemometrics and Intelligent Laboratory Systems*, vol. 88, pp. 183-188, 2007.
[18] D. Svozil, et al., "Introduction to multi-layer feed-forward neural networks," *Chemometrics and Intelligent Laboratory Systems*, vol. 39, pp. 43-62, 1997.
[19] S. Haykin, "Feedforward Neural Networks: An Introduction," *Convergence*, pp. 1-16, 1998.
[20] G. Binetti, et al., "Cultivar classification of Apulian olive oils: Use of artificial neural networks for comparing NMR, NIR and merceological data," *Food Chemistry*, vol. 219, pp. 131-138, 2017.
[21] H. Karim, et al., "Comparison of Neural Network Training Algorithms for Classification of Heart Diseases," *IAES International Journal of Artificial Intelligence*, vol. 7, pp. 185-189, 2018.
[22] N. Merh, et al., "A Comparison Between Hybrid Approaches of Ann and Arima for Indian Stock Trend Forecasting," *Business Intelligence Journal*, pp. 23-43, 2010.