ORIGINAL PAPER

MODELING OF SURVIVAL TIME OF ORAL SQUAMOUS CELL CARCINOMAS (OSCC) IN HOSPITAL UNIVERSITI SAINS MALAYSIA USING MULTILAYER FEEDFORWARD NEURAL NETWORK

WAN MUHAMAD AMIR W AHMAD[1], MOHAMAD ARIF AWANG NAWI[1], MOHAMAD SHAFIQ MOHD IBRAHIM[1], ZALILA ALI[2] , NOR AZLIDA ALENG[3], ADAM HUSEIN[1], NURFADHLINA ABDUL HALIM[2]

Manuscript received: 06.09.2018; Accepted paper: 08.11.2018;
Published online: 30.12.2018.

Abstract. *The oral squamous cell carcinoma (OSCC) is the most common malignant neoplasm of the oral cavity with up to 50% of the mortality rate. It has been reported about 14.1 million new cancer cases and 8.2 million cancer deaths in 2012. Numbers of studies have been performed to investigate the factors that have direct and indirect or both associated with the OSCC, including their survival time. In this paper, the potential clinic pathological prognostic factors will be determined in patients who attended Hospital Universiti Sains Malaysia (USM) from 2005 to 2015 using multilayer feed-forward (MLFF) neural network. The objective of the current study is to develop a multilayer feed-forward (MLFF) neural network model of time survival of OSCC. Alcohol, tumor site, tumor size and betel quid were significant. These four variables were used to develop the best (MLFF) neural network model of OSCC.*

Keywords: *Oral Squamous Cell Carcinoma (OSCC), Clinic pathological and multilayer feed-forward (MLFF).*

1. INTRODUCTION TO ORAL SQUAMOUS CELL CARCINOMA

Majority of the population in the world are commonly occurring oral cancer diseases. According to estimates from the International Agency for Research on Cancer (IARC) in 2012, there were 14.1 million new cancer cases and 8.2 million cancer deaths worldwide [1]. As much as 3% of oral cancer includes a group of neoplasms and is the eighth most common cancer in the world. Oral cancer begins in the oral cavity which includes the entire mouth, lips, the inner lining of the lips and cheeks, the teeth gums, the portion of the mouth below the tongue, the hard palate or the bony roof of the mouth, the area behind the wisdom teeth and etc [2]. Among these, oral squamous cell carcinoma (OSCC) is the most common, comprising 95% of head and neck cancers [3]. Globally, the incidence of oral cancer holds the eighth position and show epidemiologic variability according to different geographic regions. Besides the proven causative agents like tobacco, alcohol, and human papilloma virus, there are certainly other factors that play a significant rolein the selection of treatment strategies

[1] Universiti Sains Malaysia, School of Dental Sciences, 11800 Gelugor, Penang, Malaysia.
E-mail: wmamir@usm.my.
[2] Universiti Sains Malaysia, School of Mathematical Sciences, 21030 Kuala Terengganu, Malaysia.
[3] Universiti Malaysia Terengganu, School of Informatics and Applied Mathematics, 11800 Gelugor, Penang, Malaysia.

and determination of prognosis in OSCC patients. These range from simple demographic/general physical factors to the discovery of newer molecular markers, comprising the clinical and histopathological factors. However, none of the factors can alone influence the prognosis and it has been observed that a multitude of factors/parameters needs to be included to determine the prognosis and treatment outcome of a patient [4]. The standard treatment for OSCC is a combination of surgery, radiation and chemotherapy. The 5-year survival rate of OSCC is only 50%, which has remained unchanged for a decade [5].

With the latest research developments in determining the prognosis of oral cancer lesions, researchers have gained access to a much higher data gathered. However, trade data interpretation seems less and not use the knowledge gained in the right direction to translate strategy diagnosis and better treatment. The major problem in the integration and interpretation of test data should be handled as a "network" that provides insight to further advance our understanding in addressing the underlying mechanisms of cancer in the oral cavity. Statistical survival modelingoffers an avenue for cancer network modeling. In this paper, prognostic survival cancer models are focused specifically on OSCC data [6].

2. MATERIALS AND METHODS

The archive of medical unit record of Hospital Universiti Sains Malaysia (HUSM) was reviewed and related information was extracted. A total of 16 eligible cases were selected from the list of a patient diagnosed with OSCC. The selected variables are shown in Table 1 as follows:

Table 1. Data Description.

Num.	Variables	Explanation of user variables
1.	Survival Time	Survival in months
2.	Alcohol	Alcohol taken by the patients [1= Never, 2 = Stop and 3 = Current]
3.	Tumour site	Tumour siteby the patients [1 = Gum, 2 =Tongue, 3 = Cheek, 4 = Lip]
4	Tumour size	Tumour sizeby the patients [1= Less than 2cm, 2 = Greater than 2 less than 4 cm, 3 = Greater than 4 cm]
5.	Betel quid	Betel quid taken by the patients [1 = Never , 2 = Stop, 3 = Current]

2.1. MULTILAYER FEED-FORWARD NEURAL NETWORK (MLFF)

Multilayer Feed forward Neural Network (MLFF) consists of an input layer, one or several hidden layers and an output layer. The neurons in the feed-forward neural network are generally grouped into layers. Signals flow in one direction from the input layer to the next, but not within the same layer [7]. An essential factor of successes of the neural networks depends on the training network. Among the several learning algorithms available, back-propagation (BP) has been the most popular and most widely implemented [8]. Basically, the BP training algorithm with three-layer feed-forward architecture means that the network has an input layer, one hidden layer, and an output layer. From the previous study, many

researchers using multilayer feed forward neural network (MLFF) to evaluated MSE of testing/out-sample in medical data [9-12]. In this research, the output node is fixed at one since there is only one independent variable. Thus, for the feed-forward network with N input nodes, H hidden nodes, and one output node, the values $\hat{y}$ are given by:

$$\hat{y} = g_2\left(\sum_{j=1}^{H} w_j h_j + w_0\right) \tag{1}$$

where w_j an output weight from hidden node j to the output node is w_0 the bias for the output node, g is an activation function. The values of the hidden node h_j , j=1…H are given by;

$$h_j = g_1\left(\sum_{i=1}^{N} v_{ji} x_i + v_{j0}\right),\ j = 1,...,\mathrm{H} \tag{2}$$

Here, v_{ji} the output *weight* from input node *i*to hidden node *j*, v_{j0} is the bias for hidden node *j*, $x_i$is the independent variables where *i*=1…*N* and *k* is an activation function. The architecture of the multilayer feed-forward neural network model is illustrated in Fig. 1.

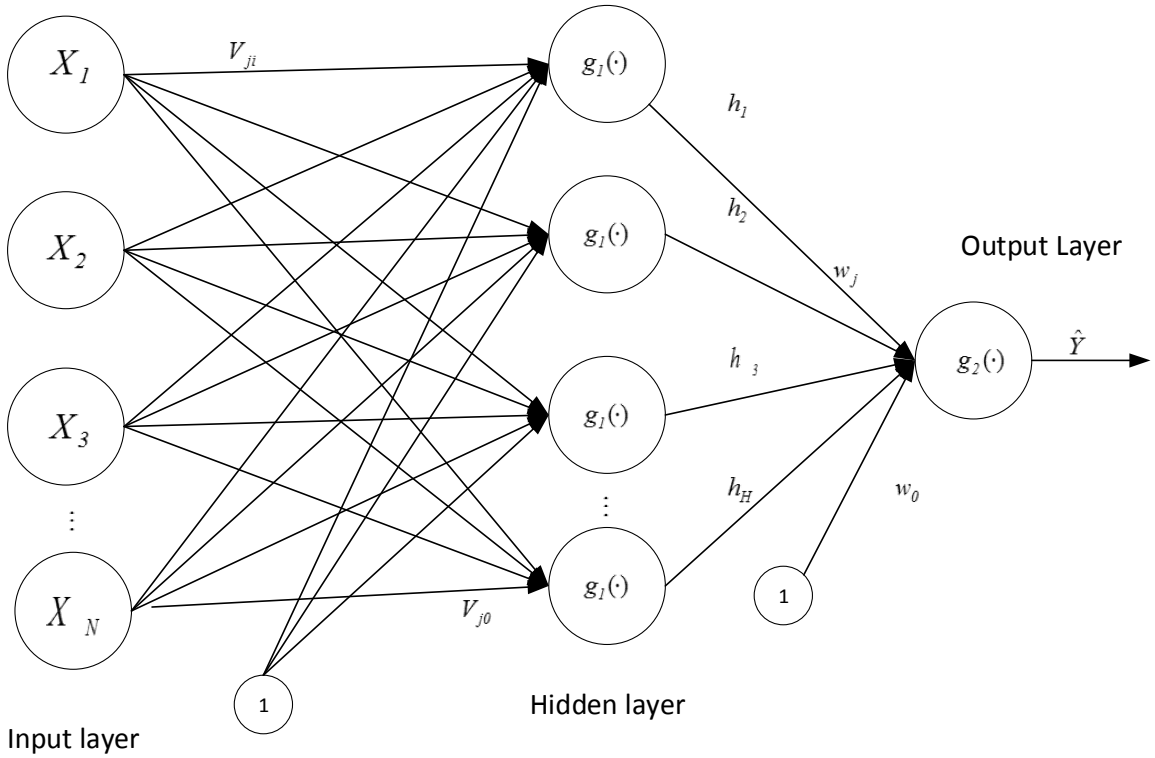

Figure 1. The architecture of the multilayer feed-forwardneural network model with one hidden layer, Ninput nodes, H hidden nodes and one output node.

The neural network architecture is composed of the number of input, hidden and output nodes. There are 5 selected variables, which were survival time, alcohol, tumor site, tumor size, and betel quid.There are four independent variables namely alcohol, tumor site, tumor size, and betel quid. The output node in this study is one node since we have one dependent variable which is survival time. The data was partitioned into two parts which are training (60%) and testing (40%).The Levenberg-Marquardt back-propagation is used as the training algorithm since it was claimed as the best training algorithm [13]. Below is the path which using SPSS Modeler to obtain the neural network analysis.

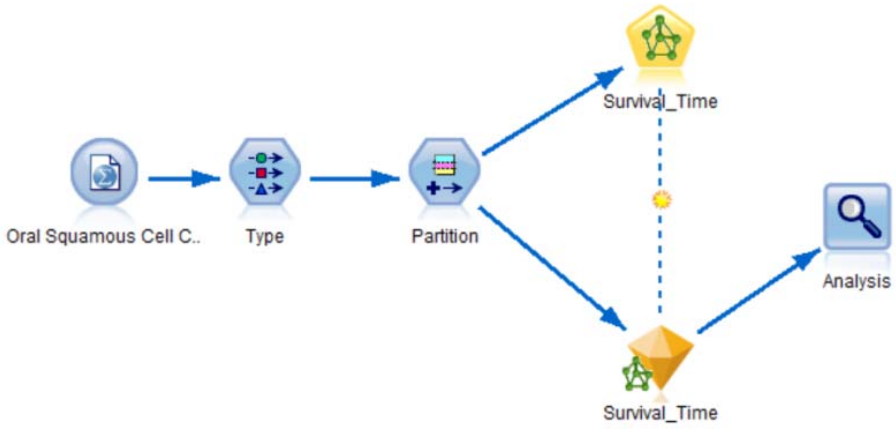

Figure 2. Neural network path in SPSS Modeller.

3. RESULTS AND DISCUSSION

Fig. 3 shows the predictor important in SPSS Modeler. Form the potential four factor predictor importance, tumour size is the most significant factor that influences the time survivor of oral squamous cell carcinoma. Betel quid is the second factor that contributing most to the time survival.

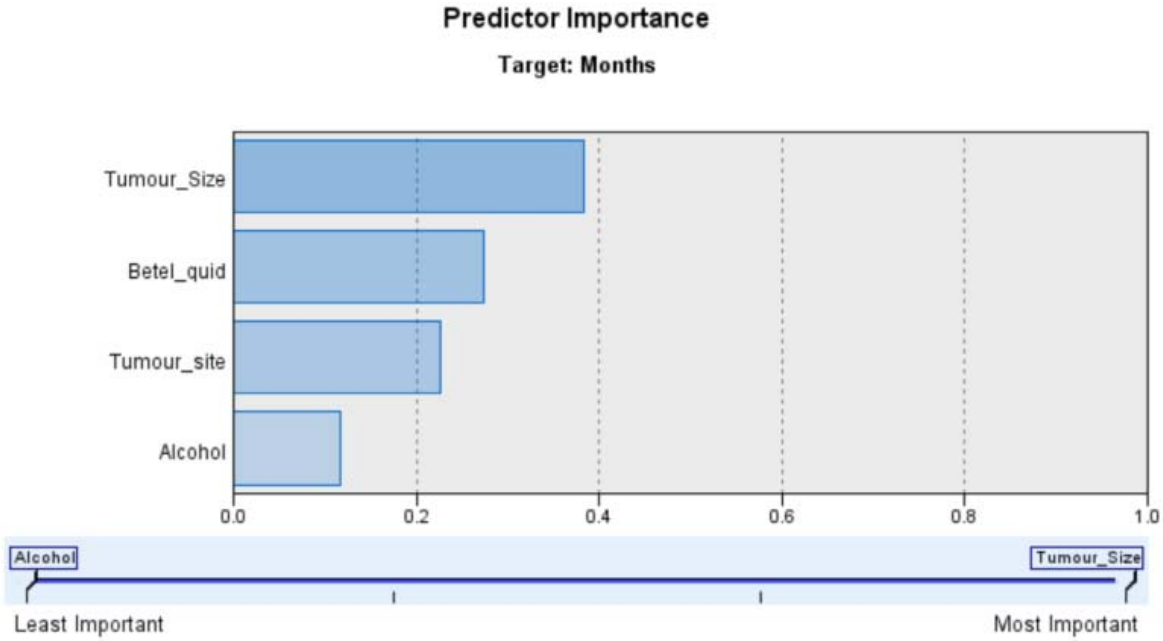

Figure 3. Predictor Importance in SPSS Modeler.

Since we consider all four independent variables (Fig. 4) as inputs for the MLFF model, then input nodes are four nodes and as survival time is considered as the output, then the output node is one. We then apply the forward procedure to find the best number of hidden nodes.

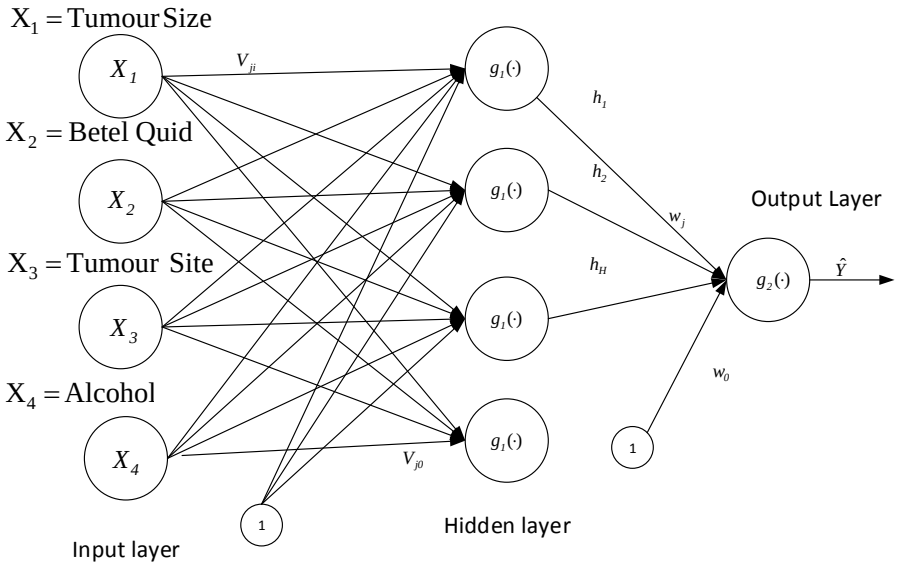

Figure 4. The Architecture of the best multilayer feed-forward neural network model with one hidden layer, four input variables, four hidden nodes,and one output node.

We found that the best number of hidden nodes for the MLFF model is four nodes. Hence, the appropriate neural network architecture which results in the best multilayer feed-forward neural network model for our case can be represented as follows:

$$\hat{y} = g_2\left(\sum_{j=1}^{4} w_j h_j + w_0\right) \tag{3}$$

where w_j is an output *weight* from hidden node *j* to the output node, w_0 is the bias for the output node and g_2 is the linear function. *h* are the values of the hidden layer nodes which can be represented as:

$$h_j = g_1\left(\sum_{i=1}^{4} v_{ji} x_i + v_{j0}\right), \; j = 1,2,3,4 \tag{4}$$

where *v* is the input weight from input node *I* to hidden node *j*, *v* is the bias for hidden node *j*and *g* is an activation function. *X* is the independent variables where X_1, X_2, X_3, X_4 are tumor size, betel quid, tumor site, and alcohol. Equation (3) and (4) can also represent as follows:

$$\hat{Y} = w_0 + w_1 h_1 + w_2 h_2 \tag{5}$$

therefore;

$$h_j = \left[1 + exp\left[-\left(v_{jo} + v_{j1}x_1 + v_{j1}x_2 + v_{j1}x_3 + v_{j1}x_4\right)\right]\right]^{-1} \quad j = 1,2, \tag{6}$$

Table 2. The results of Mean Error and Mean Absolute Error for Training and Testing.

Input Variables*	Mean Error Training	Mean Error Testing	Mean Absolute Error Training	Mean Absolute Error Testing
X_1, X_2, X_3, X_4	0.109	1.737	13.561	19.793
X_1, X_2, X_3	0.599	7.070	14.913	23.435
X_1, X_2	1.868	3.632	14.959	20.603
X_1	-2.017	-0.939	20.668	25.323

** X_1, X_2, X_3, and X_4 represent tumor size, betel quid, tumour site and alcohol respectively.*

The purpose of the current study is to develop an MLFF neural network model of survival timefor oral cancer patients. Using the significant variables the performance of MLFF neural network model for different combinations is tested.The performance of MLFF was evaluated using Mean Error Training (MAT) and Mean Absolute Error (MAE) of

testing/out-sample. In our case, four variables were used to develop the best (MLFF) neural network model of survival time oral squamous cell carcinoma (OSCC).

4. CONCLUSIONS

The main purpose of this paper is to demonstrate such relationship and find the prognostics factor the inferences purpose. For this purpose, we developed MLFF neural network models. The data used were survival time survival of OSCC as the dependent variable and four independent variables according to their clinically important.The efficiency of MLFF was evaluated using the MAT and MAE.This paper provides a preliminary overview of the associated factor for OSCC. Patients those who have a habit of alcohol showed an increased risk of death as compared to non-alcoholic OSCC patients [6]. From the MLFF approached method, we can predict a better forecasting result in future for the decision making. These promising techniques had lead to a successful research and give the best results for the decision making especially for the decision maker.

Acknowledgments: *The authors would like to express their gratitude to Universiti Sains Malaysia (USM) for providing the research funding (RUI Grant no.1001/PPSG/8012278, School of Dental Sciences, Kampus Kesihatan).*

REFERENCES

[1] International Agency for Research on Cancer. *Personal Habits and Indoor Combustions. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* IARC Press, 100, 2012.
[2] Hermanek, P.S., *International union Against Cancer TNM classification of malignant tumors*, LH Editors, Berlin, Springer- Verlag, 4th Ed, 2nd revision, 1992.
[3] Vigneswaran, N., Williams, M. D., *Oral and maxillofacial surgery clinics of North America*, **26**(2), 123, 2014.
[4] Taghavi, N., Yazdi, I., *Arch Iran Med*, **18**(5), 314, 2015.
[5] Brunin, F., Mosseri, V., Jaulerry, C., Point, D., Cosset, J. M., Rodriguez, J., *Head Neck*, **21**, 751, 1999.
[6] Azeem, M., *Human Papillomavirus (HPV) Related Oral Squamous Cell Carcinoma In Kelantan Population*, Unpublished Msc Thesis, Universiti Sains Malaysia, Malaysia, 2018.
[7] Pham, T.D., Liu, X., *Neural networks for identification, prediction and control*, Great Britain, 1995.
[8] Darbellay, G.A., Slama, M., *International Journal of Forecasting*, **16**(1), 71, 2000.
[9] Mohamed, N., Ahmad, W.M.A.W., Aleng, N.A., Ahmad, M.H., *Applied Mathematical Sciences*, **7**(41), 2047 , 2013.
[10] Aleng, N.A., Mohamed, N., Ahmad, W.M.A.W., *Applied Mathematical Sciences*, **6**(50), 2481, 2012.
[11] Mohamed, N., Aleng, N.A., Ahmad, W.M.A.W., Ahmad, M.H., *Int. J. Contemp. Math. Sciences*, **7**(35), 1749, 2012.
[12] Aleng, N.A., Mohamed, N., Ahmad, W.M.A.W., Naing, N.N., *European Journal of Scientific Research*, **73**(1), 79, 2012.
[13] Mohamed, N., *PhD thesis - Parametric and artificial intelligence based methods for forecasting short term electricity load demand*, Universiti Malaysia Terengganu, 2011.