

Research Article

Two Artificial Neural Networks for Modeling Discrete Survival Time of Censored Data

Taysseer Sharaf¹ and Chris P. Tsokos²

¹Department of Mathematics and Statistics, University of South Florida, 4202 E. Fowler Avenue, CMC 342, Tampa, FL 33620, USA

²Department of Mathematics and Statistics, University of South Florida, 4202 E. Fowler Avenue, CMC 366, Tampa, FL 33620, USA

Correspondence should be addressed to Taysseer Sharaf; tsharaf@mail.usf.edu

Received 17 September 2014; Revised 17 February 2015; Accepted 23 February 2015

Academic Editor: Jun He

Copyright © 2015 T. Sharaf and C. P. Tsokos. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Artificial neural network (ANN) theory is emerging as an alternative to conventional statistical methods in modeling nonlinear functions. The popular Cox proportional hazard model falls short in modeling survival data with nonlinear behaviors. ANN is a good alternative to the Cox PH as the proportionality of the hazard assumption and model relaxations are not required. In addition, ANN possesses a powerful capability of handling complex nonlinear relations within the risk factors associated with survival time. In this study, we present a comprehensive comparison of two different approaches of utilizing ANN in modeling smooth conditional hazard probability function. We use real melanoma cancer data to illustrate the usefulness of the proposed ANN methods. We report some significant results in comparing the survival time of male and female melanoma patients.

1. Introduction

Artificial neural network (ANN) is becoming one of the most popular alternatives to conventional statistical modeling. It is actually conceived as an advanced generalized linear model. We have seen various applications of ANN utilized in different scientific subjects like engineering, economics, environment, and health, among others. For example, van Hinsbergen et al. 2009 [1] applied artificial neural networks to short-term time prediction of traffic travel time. Kingston et al. 2005 [2] proposed ANN to model water resources. In economics, Baesens et al. 2005 [3] used ANN to predict survival time of personal loan data. Baesens et al. compared the ANN model used with other survival analysis models like logistic regression and Cox PH and the results came in favor of the ANN.

In the medical sciences, most of the proposed applications of ANN were on prognostic models. For example, one of the most paramount research entities is cancer. Classifying a tumor as malignant or benign is important in cancer research. Chen et al. in 2002 used ANN to diagnose breast cancer tumors [4]. Ercal et al. in 1994 presented an ANN model to distinguish between three benign skin cancer categories

and malignant melanoma [5]. But fitting a complex nonlinear modeling such as ANN in regression problems is less prevalent. Determining the risk factors that cause cancer or modeling the survival time of a patient once he/she is diagnosed with cancer using ANN is less common.

In this present study, we are interested in utilizing ANN in survival time modeling of skin cancer (melanoma) patients. Soong et al. [6] in 2010 developed a statistical model to predict the survival time of localized melanoma patients. They used the proportional hazard model developed by Cox [7], but the assumptions of hazard function proportionality may not be applicable to a different set of data. Moreover, they did not study the effect of interaction terms. Thus, applying ANN is more applicable and efficient, especially when the data does not satisfy Cox PH assumptions. ANN does not require any assumptions that need to be justified, and it is more precise in fitting nonlinear models [8–10]. One of the basic approaches in utilizing ANN in survival analysis is by classification, whether a patient will survive over a fixed time interval or not [11]. However, the latter classification method lacks the information about the survival probability function estimates. In 1995, Lapuerta et al. proposed the use of multiple neural networks one for each time interval [12]. This model

predicts the survival probability of each time period based on a neural network trained on the observations of the same time period only. The pitfall of this approach is the large number of networks that will be trained if one studies the survival time over immense time intervals.

Other methods of ANN applied to survival time were proposed by Faraggi and Simon in 1995 [8] and by Ohno-Machado in 1996 [13]. We consider in this study the approach represented by Biganzoli et al. in 1998 [9], which was a modification of a study done by Ravdin and Clark in 1992 [14], in addition to the approach represented by Mani et al. in 1999 [10]. Thus, in this study a comparison between the two methods Biganzoli and Mani is given. Also, we study the difference between the survival time of male and female melanoma patients.

In the following section, we discuss the data used to perform our comparison, along with significant results exhibiting the differences between male and female melanoma patient survival times. In the third section we discuss briefly the two methods emphasizing the differences, advantages, and disadvantages of both. In the fourth section we present our results and identify the model that gave the best performance in estimating the survival probability function with less error.

2. Materials and Methods

2.1. Data. We have 130,006 patients diagnosed with melanoma between the years 2000 and 2009 in the USA. Data accumulated from 13 registers of the Surveillance, Epidemiology, and end results program (SEER) [15]. We filter out this large dataset to contain only consummate information with respect to the patient's age at diagnosis, tumor thickness, stage of cancer, and ulceration. Soong et al. [6] in 2010 used these four variables, but their study did not consider the difference between male and female survival. We found that there exists a significant difference between the median survival time of males and females based on a 5% level of significance using the Kruskal-Wallis test. Thus, studying the effect of gender on survival time by making one model for both males and females is not statistically correct, as the survival time for male and females does not have the same distribution. Figure 1 represents a schematic diagram of the distribution of the complete data with respect to gender and cancer stage.

We train neural network model separately for males and females. Figure 1 also includes the total number of patients with complete information who were diagnosed with melanoma between the years 2000 and 2009. Patients were either alive at the end of the 10-year period (censored) or lost to follow-up during the ten-year interval (censored), in addition to patients who died because of melanoma (uncensored) (95% of male patients were censored and 98.1% of female patients were censored). We omitted patients with incomplete information. For training purposes, each of the gender category data was divided into six groups; five will be used in the cross validation technique to train and validate the neural network model, while the last group is used for prediction and for accomplishing the comparison between

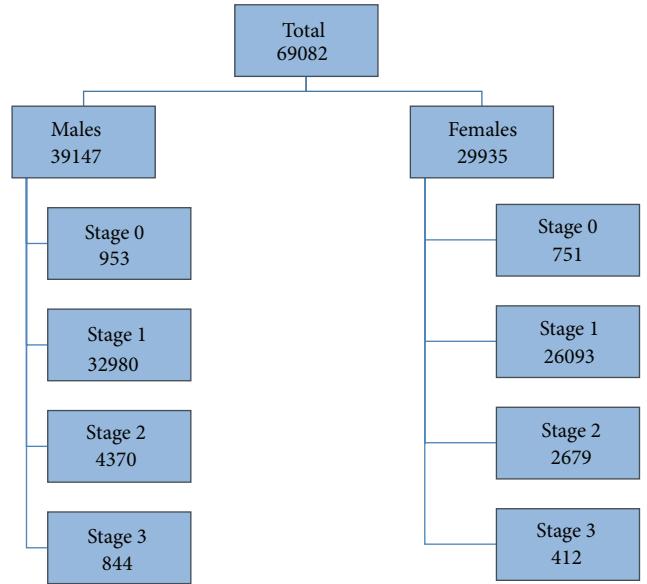


FIGURE 1: Distribution of complete information of melanoma patients.

the two modeling approaches. More details about the cross validation will follow in the next section.

In our modeling procedure, we used age at diagnosis and tumor thickness (in millimeter) as quantitative variables along with three dummy variables representing stage 1 (referring to a localized tumor), stage 2 (referring to a regional tumor), and stage 3 (referring to a distant tumor). The base level is referred to as *in situ* tumor.

2.2. Methodology. ANN was based on the model of one perceptron introduced by Rosenblatt in 1957 [16]. Today a multilayer perceptron (MLP) is known now by neural networks and consists of multiple layers of neurons. The first layer (input layer) represents the covariates or the risk factors, which are the inputs of the hidden neurons in the first hidden layer. The output of the hidden layer is the input of another hidden layer (if more than one hidden layer exists) or the output layer. This type of MLP is called feed forward artificial neural network (FFANN). We shall discuss both methods and compare them in the current study using FFANN with one hidden layer.

2.2.1. Method 1: PLANN. Partial logistic artificial neural network (PLANN) is the approach that was introduced by Biganzoli et al. [9]. PLANN is a three-layer feed forward artificial neural network with one output unit in the output layer. The activation function used in the hidden and output layer is the logistic function given by

$$f(z) = \frac{\exp(z)}{1 + \exp(z)}. \quad (1)$$

PLANN estimates the conditional hazard function that is based on the discrete survival method. The discrete survival method was introduced by Allison [17] in 1982 and then

Singer and Willett [18] in 1993. The discrete survival method considers grouping the continuous survival time into $k = 1, 2, \dots, K$ disjoint intervals, in which the individual records will be replicated l times, where l is the number of time period in which the event occurred. The discrete hazard probability function for time period k given a vector of covariates \mathbf{x}_i is given by

$$h_k(\mathbf{x}_i, t_k) = \frac{1}{1 + \text{EXP}\{-(\alpha_k t_k + \beta^t \mathbf{x}_i)\}}, \quad (2)$$

where t_k represents the time input for time period k . To estimate the conditional hazard in (2) PLANN uses three-layer FFANN with an activation function for the hidden and output layers given in (1). The output of the network with an H number of hidden units is given by

$$\hat{h}(\mathbf{x}_i, t_k) = f\left(b + \sum_{h=1}^H w_h f\left(a_h + \sum_{p=1}^P w_{ph} x_{ip}\right)\right), \quad (3)$$

where w_{ph} and w_h are the weights of the ANN to be estimated for the first layer and second layer, respectively, and also a_h and b are the weights for the bias connection with the hidden units and with the output unit, respectively. The target of this network is the censoring indicator c_{ik} , which is equal to 1 if the event occurred for subject i and 0 otherwise. The cost function used in PLANN is the cross entropy function which is appropriate for binary classification problems [19]. The weights of PLANN can be estimated by minimizing the cost function given by

$$E = - \sum_{i=1}^n \sum_{k=1}^{k_i} \left\{ c_{ik} \log [\hat{h}(\mathbf{x}_i, t_k)] + (1 - c_{ik}) \log [1 - \hat{h}(\mathbf{x}_i, t_k)] \right\}. \quad (4)$$

Once the network weights are estimated, the monotone survival probabilities can be easily found by converting the discrete hazard rate estimates obtained from the network output by the following equation:

$$S(t_k) = \prod_{l=1}^k (1 - h(t_l)). \quad (5)$$

The advantage of this approach is that the time dependent covariates can easily be introduced in the model, as the individual records are available for each time period. However, for large datasets or studies conducted over a long period of time this approach is inaccessible due to the immense number of replication requisites [3]. Figure 2 shows the architecture of the PLANN introduced by Biganzoli.

The first layer in Figure 2 contains the bias and one node for the time period and the rest of the nodes for the covariates. PLANN uses one input for the time to estimate smooth discrete hazard rates. However, we have used 10 nodes for the time (one for each time period) to be able to compare it with the second method.

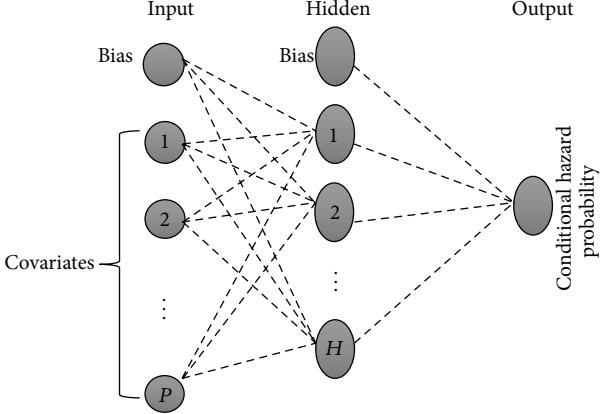


FIGURE 2: FFANN for partial logistic artificial neural network, with three layers. The input layer has p covariates and hidden layer with H hidden units and one output unit in the output layer. Activation function used in both hidden and output layer is logistic function (1).

2.2.2. Method 2. Mani et al. developed an approach which was utilized by Street [20] that predicts the survival function using a neural network with K outputs, where K is the number of time periods. He trained his network utilizing a target vector derived by Kaplan-Meier survival curves [21]. Mani used the same neural architecture as Street, but to estimate the hazard function instead. In order to estimate the hazard function, each individual or subject would have a training vector (1 by K) target of hazard probabilities h_{ik} as follows:

$$h_{ik} = \begin{cases} 0 & \text{for } 1 \leq k \leq K \\ 1 & \text{for } t \leq k \leq K \text{ and event} = 1 \\ \frac{r_k}{n_k} & \text{for } t \leq k \leq K \text{ and event} = 0. \end{cases} \quad (6)$$

Here, $h_{ik} = 0$ for each time period if patient i survived. $h_{ik} = 1$ from time interval t to K if patient died because of melanoma at duration t within the study time. And, for those patients who are lost to follow-up during the study of duration $t < K$, their hazards are equal to the ratio r_k/n_k , which is the Kaplan-Meier hazard estimate for time interval k . r_k is the number of patients who died because of melanoma in time period k , and n_k is the number of patients that are at risk in time interval k . For training the neural network, Mani used the logistic sigmoid function given by

$$\Phi(x) = \frac{1}{1 + e^{-x}}. \quad (7)$$

The network weights are estimated by minimizing the cost function, which is the cross entropy function. Figure 3 shows the neural network architecture utilized in Method 2. The number of units p of the input layer is equivalent to the number of independent variables or risk factors. The K output units of the output layer learn to estimate the hazard probability of each individual. Once the ANN is trained and the hazard estimates are predicted, we convert those hazard

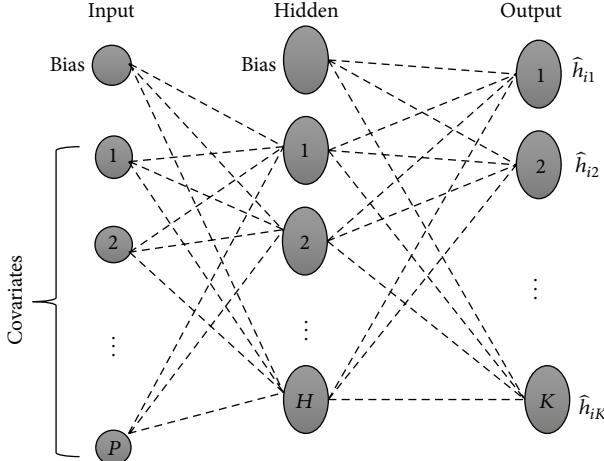


FIGURE 3: Three-layer network with K output units in the output layer where K is equal to the number of time intervals.

estimates to the survival estimates by using (5) (for each method). We have trained the weights of the ANN in both methods using the quasi-Newton algorithm.

2.3. Model Selection. Now, we are concerned with the optimal number of hidden units in the hidden layer that will give us the best neural network model. There are several methods in the literature that we can use to select the best neural network. The most popular method is the v -fold cross validation method as it does not rely on any probabilistic assumptions and helps in determining when overfitting occurs. Other statistical methods like hypothesis testing or information criteria were introduced and examined by Anders and Korn in 1999 [22] for neural network model selection, and they suggested that those statistical methods should take part in neural network modeling. However, since their proposed methods were based on certain probabilistic assumptions, it may not be always applicable in modeling real phenomena.

In order to do our comparison we took the best neural network for each method and then tested their performance on the same set of data (this set of data was removed from the training dataset). In the current study, we have used 5-fold cross validation to select the best model for each method (Methods 1 and 2). We divide the male and female datasets into six groups. Five were used in the training and validation, and the last group was used for comparing the best models from the two methods together (hold-out dataset). In addition, we use the weight decay that helps avoid overfitting and penalize large weight solutions to help in generalization. As mentioned by Ripley [23, 24] a weight decay value between $\alpha = 0.01$ and 0.1 would be more appropriate depending on the degree of fit that is expected. We have used the cross validation method along with four different values of weight decay $\alpha = \{0.025, 0.05, 0.075, 0.1\}$, to pick the best model. The same procedure of trying different weight decay values was used in [9].

The cross validation method will help us in finding the optimal number of hidden nodes. In addition, we consider

TABLE 1: Mean prediction error for the eight competing neural network models for estimating the survival time of male melanoma patients.

	Mean	St. d.v.	Count
Elia's method			
$\alpha = 0.025$	-0.3776	1.7876	85460
$\alpha = 0.050$	-0.4113	1.9388	85460
$\alpha = 0.075$	-0.308	1.4295	85460
$\alpha = 0.1$	-0.3366	1.5329	85460
Mani's method			
$\alpha = 0.025$	-0.0236	0.2029	85460
$\alpha = 0.050$	-0.0215	0.1710	85460
$\alpha = 0.075$	-0.0197	0.1539	85460
$\alpha = 0.1$	-0.0197	0.1485	85460

the model with the lowest prediction error when applied to a new data. Therefore, for each method, we picked the best model (with lowest cross validation error) and then compared its performance on the hold-out dataset. We repeated our comparison for the four values of weight decay, since for the same data two factors affect ANN performance (number of hidden units and weight decay value).

3. Results

In our analysis, we used ten time intervals (12 months each) and in order to do the comparison between the PLANN and Mani's method we used 10 inputs for the ten time intervals in PLANN instead of one so that we can compare the output of the PLANN with the second method.

After training, the cross validation method resulted in choosing the networks with 52 hidden units (number of hidden nodes seems to be large but, by taking into account the number of output units, we have 10 parallel networks with five hidden nodes each) for both methods as the best model. We obtain similar results from ANN trained with the four different values of weight decay. Still, we want to examine the prediction accuracy for all eight available models to choose our best-fit model. Table 1 exhibits the comparison between the eight models for male melanoma patients and Table 2 exhibits the comparison for the female melanoma patients.

It is clear from Table 1 that using Mani's method yields better predictive neural network model than the PLANN proposed by Elia. Among the four competing models of Mani's method we have chosen the model with weight decay $\alpha = 0.1$ as the best-fit model for predicting survival times of male melanoma patients that yield smaller mean error and standard deviation.

The results in Table 2 support the decision we have made for male melanoma patients that Mani's method has a better predictive accuracy than that of the PLANN. But the best model for predicting the female melanoma patient's survival time is the model with weight decay value $\alpha = 0.075$.

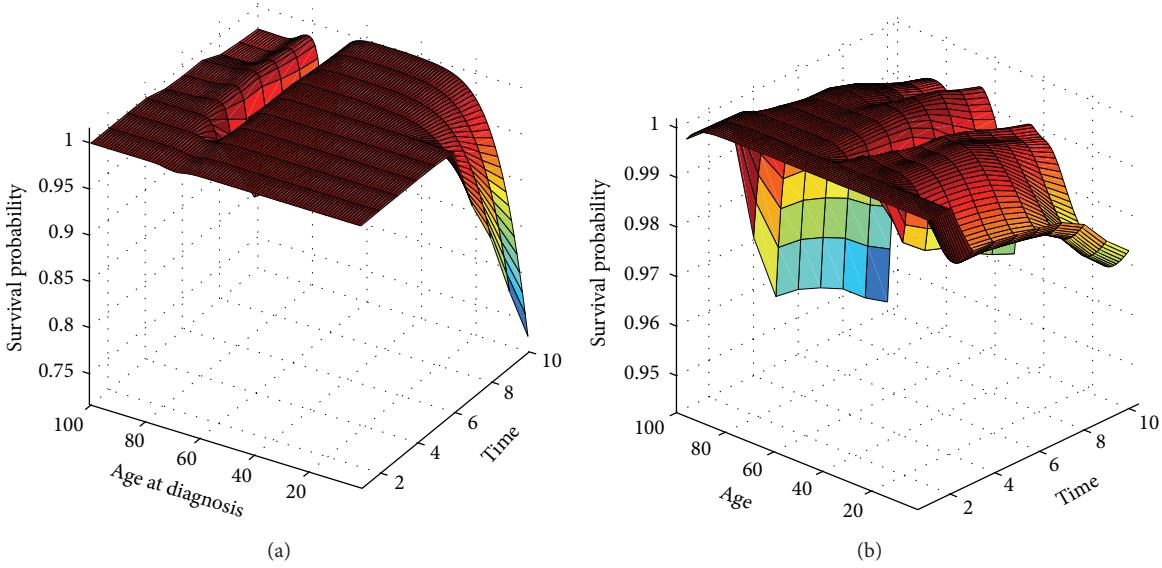


FIGURE 4: Survival probability function surface plot results for age: (a) for male melanoma patients and (b) for female melanoma patients. The survival probability is estimated as a function of time (ten-year period) and age. Other risk factors like tumor thickness are fixed at 0.58 mm with no ulceration and in the initial stage.

TABLE 2: Mean prediction error for the eight competing neural network models for estimating the survival time of female melanoma patients.

	Mean	St. dv.	Count
Elia's method			
$\alpha = 0.025$	-0.3384	2.1326	69130
$\alpha = 0.050$	-0.2274	14451	69130
$\alpha = 0.075$	-0.1882	1.1672	69130
$\alpha = 0.1$	-0.2011	1.2199	69130
Mani's method			
$\alpha = 0.025$	-0.0208	0.2310	69130
$\alpha = 0.050$	-0.0183	0.1849	69130
$\alpha = 0.075$	-0.0167	0.1606	69130
$\alpha = 0.1$	-0.0246	0.3022	69130

4. Discussion

In modeling survival data with ANN, it is more prevalent to utilize Method 2 (Mani's method). The prediction accuracy is much better compared to the PLANN. However, the results may change if somehow the survival data contains time varying covariates (risk factors). One can attempt to amend the PLANN model by differentiating between the individuals who survived the whole duration time and those who dropped out during the duration time, which opens another area of research in this field. With regard to learning techniques for ANN, Lisboa et al. [25] have amended the PLANN by adapting the Bayesian learning for neural networks, developed by Mackay in 1995 [26]. It is still an open problem: how the Bayesian learning will affect the performance of Mani's ANN? Is it going to change the comparison results with the PLANN? These are among other

questions that we need answers to and we open more areas of research on this type of problems.

It is clear to us that male and female melanoma patients need to be treated differently as shown by the survival plots in Figure 4, which displays the surface plot of survival probability of males (Figure 4(a)) and females (Figure 4(b)). In this figure, the survival is estimated as a function of age at diagnosis and time in years. The tumor thickness is 0.58 mm and ulceration variable is set to 0, considering that the patient was diagnosed in the initial stage. Male infant patients have less survival probability than that of female infant patients over a 10-year period, whereas a male patient at the age of 40 to 50 seems to have higher survival probabilities compared to female patients at the same age.

Figure 5 displays the surface plot of survival results for male melanoma patients for tumor thickness ranging from 0.01 mm to 9 mm. The surface plot (Figure 5(a)) is for male patient diagnosed at 20 years of age, whereas the plot (Figure 5(b)) is for male patient diagnosed at the age of 60 years. As we can see, survival estimate for young men is farther away and lower than that of older men and these findings were found similar to a recent study by Fisher and Geller in 2013 [27].

Fisher and Geller mentioned that more attention was given to older men over the past years and suggested that more awareness is needed to be addressed to young men to help in early detection of melanoma. They also mentioned the difference between young men and young women, which we can figure out by comparing the two left plots of Figures 5 and 6. The survival probability for young men (diagnosed with tumor thickness larger than 4 mm) within two years of diagnosis is too low (almost 0) compared to that of young women.

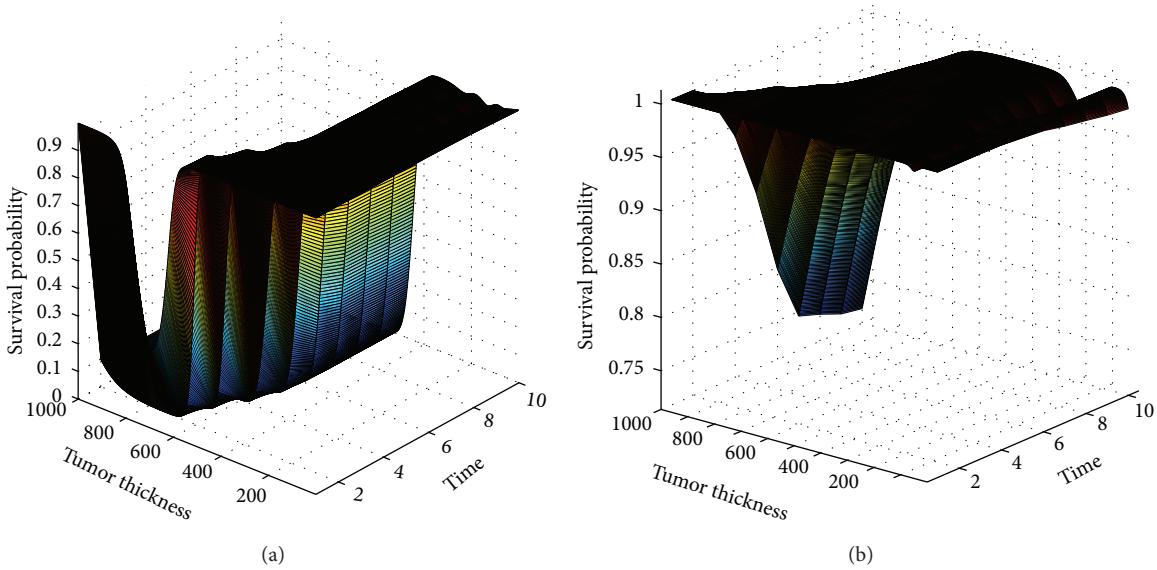


FIGURE 5: Survival probability function surface plot results for tumor thickness: (a) for the male patient diagnosed at age of 20 years old and (b) for male diagnosed at the age of 60 years.

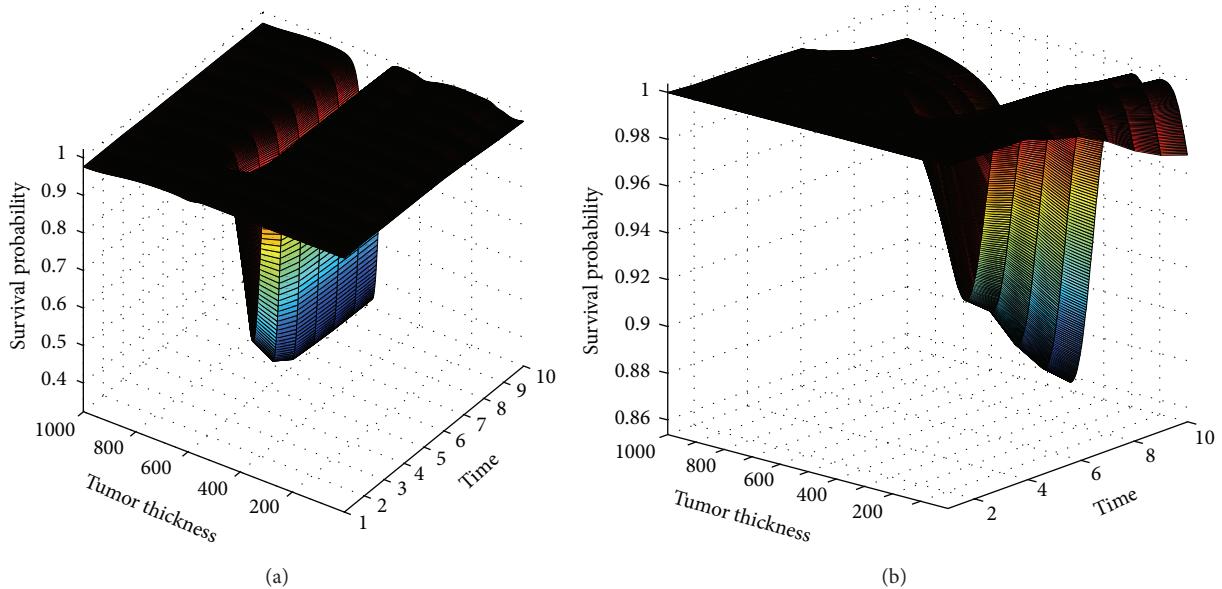


FIGURE 6: Survival probability function surface plot results for tumor thickness: (a) for the female patient diagnosed at the age of 20 years and (b) for female diagnosed at the age of 60 years.

Some of our significant findings were found to be similar to those found in another study by Gamba et al. in 2013 [28]. However, more investigation and statistical data analysis are required to better understand the causes of the differences between young males and females and to plan new strategies to fight the major pernicious form of skin cancer (melanoma) [29].

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] C. P. I. van Hinsbergen, J. W. C. van Lint, and H. J. van Zuylen, "Bayesian committee of neural networks to predict travel times with confidence intervals," *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 5, pp. 498–509, 2009.
- [2] G. B. Kingston, M. F. Lambert, and H. R. Maier, "Bayesian training of artificial neural networks used for water resources modeling," *Water Resources Research*, vol. 41, no. 12, Article ID W12409, pp. 1–11, 2005.
- [3] B. Baesens, T. van Gestel, M. Stepanova, D. van den Poel, and J. Vanthienen, "Neural network survival analysis for personal

- loan data," *Journal of the Operational Research Society*, vol. 56, no. 9, pp. 1089–1098, 2005.
- [4] D.-R. Chen, R.-F. Chang, W.-J. Kuo, M.-C. Chen, and Y.-L. Huang, "Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks," *Ultrasound in Medicine and Biology*, vol. 28, no. 10, pp. 1301–1310, 2002.
- [5] F. Ercal, A. Chawla, W. V. Stoecker, H.-C. Lee, and R. H. Moss, "Neural network diagnosis of malignant melanoma from color images," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 9, pp. 837–845, 1994.
- [6] S.-J. Soong, S. Ding, D. Coit et al., "Predicting survival outcome of localized melanoma: an electronic prediction tool based on the AJCC melanoma database," *Annals of Surgical Oncology*, vol. 17, no. 8, pp. 2006–2014, 2010.
- [7] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society, Series B: Methodological*, vol. 34, no. 2, pp. 187–220, 1972.
- [8] D. Faraggi and R. Simon, "A neural network model for survival data," *Statistics in Medicine*, vol. 14, no. 1, pp. 73–82, 1995.
- [9] E. Biganzoli, P. Boracchi, L. Mariani, and E. Marubini, "Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach," *Statistics in Medicine*, vol. 17, no. 10, pp. 1169–1186, 1998.
- [10] D. R. Mani, J. Drew, A. Betz, and P. Datta, "Statistics and data mining techniques for lifetime value modeling," in *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '99)*, pp. 94–103, San Diego, Calif, USA, August 1999.
- [11] L. Bottacci, P. J. Drew, J. E. Hartley et al., "Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions," *The Lancet*, vol. 350, no. 9076, pp. 469–472, 1997.
- [12] P. Lapuerta, S. P. Azen, and L. LaBree, "Use of neural networks in predicting the risk of coronary artery disease," *Computers and Biomedical Research*, vol. 28, no. 1, pp. 38–52, 1995.
- [13] L. Ohno-Machado, "Sequential use of neural networks for survival prediction in AIDS," in *Proceedings of the AMIA Annual Symposium Proceedings Archive*, pp. 170–174, 1996.
- [14] P. M. Ravdin and G. M. Clark, "A practical application of neural network analysis for predicting outcome of individual breast cancer patients," *Breast Cancer Research and Treatment*, vol. 22, no. 3, pp. 285–293, 1992.
- [15] Surveillance Epidemiology and End Results (SEER) Program, *Research Data (1973–2009)*, Division of Cancer Control and Population Sciences, National Cancer Institute, Surveillance Research Program, Surveillance Systems Branch, 2012, <http://www.seer.cancer.gov/>.
- [16] F. Rosenblatt, *The Perceptron: A Perceiving and Recognizing Automaton*, Cornell Aeronautical Laboratory, 1957.
- [17] P. D. Allison, "Discrete-time methods for the analysis of event histories," *Sociological Methodology*, vol. 13, pp. 61–98, 1982.
- [18] J. D. Singer and J. B. Willett, "It's about time: using discrete-time survival analysis to study duration and the timing of events," *Journal of Educational and Behavioral Statistics*, vol. 18, no. 2, pp. 155–195, 1993.
- [19] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, New York, NY, USA, 1995.
- [20] W. N. Street, "A neural network model for prognostic prediction," in *Proceeding of the 15th International Conference on Machine Learning*, San Francisco, Calif, USA, 1998.
- [21] E. L. Kaplan and P. Meier, "Nonparametric estimation from incomplete observations," *Journal of the American Statistical Association*, vol. 53, pp. 457–481, 1958.
- [22] U. Anders and O. Korn, "Model selection in neural networks," *Neural Networks*, vol. 12, no. 2, pp. 309–323, 1999.
- [23] B. D. Ripley, "Neural networks and flexible regression and discrimination," *Journal of Applied Statistics*, vol. 21, no. 1-2, pp. 39–57, 1994.
- [24] B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge, Cambridge, UK, 1996.
- [25] P. J. G. Lisboa, H. Wong, P. Harris, and R. Swindell, "A Bayesian neural network approach for modelling censored data with an application to prognosis after surgery for breast cancer," *Artificial Intelligence in Medicine*, vol. 28, no. 1, pp. 1–25, 2003.
- [26] D. J. C. MacKay, "Probable networks and plausible predictions—a review of practical Bayesian methods for supervised neural networks," *Network: Computation in Neural Systems*, vol. 6, no. 3, pp. 469–505, 1995.
- [27] D. E. Fisher and A. C. Geller, "Disproportionate burden of melanoma mortality in young US men: the possible role of biology and behavior," *JAMA Dermatology*, vol. 149, no. 8, pp. 903–904, 2013.
- [28] C. S. Gamba, C. A. Clarke, T. H. M. Keegan, L. Tao, and S. M. Swetter, "Melanoma survival disadvantage in young, non-hispanic white males compared with females," *JAMA Dermatology*, vol. 149, no. 8, pp. 912–920, 2013.
- [29] A. C. Society, "Cancer Facts & Figures 2013," 2013, <http://www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-036845.pdf>.

