Journal of Biostatistics and Epidemiology

J Biostat Epidemiol. 2024;10(1): 53-63

Original Article

DeepWei-Cu: A Deep Weibull Network for Cure Fraction Models

Ola Abuelamayem*

Department of Statistics, Faculty of Economics and Political Science, Cairo University, Giza, Egypt.

ARTICLE INFO

Received 10.10.2023
Revised 01.11.2023
Accepted 09.02.2024
Published 15.03.2024

Key words:
Cure fraction;
Weibull distribution;
Deep learning;
Neural network;
Random censoring.

ABSTRACT

Introduction: Survival analysis including cure fraction subgroups is heavily used in different fields like economics, engineering and medicine. The main core of the analysis is to understand the relationship between the covariates and the survival function taking into consideration censoring and long-term survival. The analysis can be performed using traditional statistical models or neural networks. Recently, neural network has attracted attention in analyzing lifetime data due to its ability of efficiently estimating the survival function under the existence of complex covariates. The goal of this study is to develop a parametric neural network that can sufficiently predict survival data with cure fraction. To the best of our knowledge, this is the first time a parametric neural network is introduced to analyze mixture cure fraction models.
Methods: In this paper, we introduce a novel neural network based on mixture cure fraction Weibull loss function.
Results: Alzheimer disease dataset as long as synthetic dataset are used to study the efficiency of the model. We compared the results using goodness of fit methods in both datasets with Weibull regression.
Conclusion: The proposed neural network has the flexibility of analyzing continuous data without discretization. Also, it has the advantage of using Weibull distribution properties. For example, it can analyze data with different hazard rates (monotonically decreasing, monotonically increasing and constant). Comparing the results with Weibull regression, the proposed neural network performed better.

Introduction

Standard survival models implicitly assume the occurrence of the event of interest for all subjects in the study. Sometimes this is not the case, as part of the subjects may never experience the event of interest no matter how long the follow up time. For example, in medical field, some patients are cured and never face the recurrence of a certain disease like cancer. In economics, unemployed person may never have a job. In finance, some banks may never face bankrupt. In demography, one may never get married. (For more examples see Amico[1]).

To analyze this type of lifetime data, cure fraction models are introduced. There are two

*.Corresponding Author: ola.abuelamayem@feps.edu.eg

main types, mixture and non-mixture cure fraction models, the former is considered in our study. Boag[2] first proposed mixture cure fraction model to analyze breast cancer data. This cure model has since been applied by many using different distributions. For example, Farewell[3] applied Weibull distribution to analyze long term survivors for patients with three levels of zinc concentration. Yamaguchi[4] used the accelerated failure with generalized Gamma distribution to analyze the permanent employment in Japan. Yu et al[5] make a comparison using different distributions and found that generalized Gamma distribution is quite robust and applied the models to analyze cancer data. Kannan et al[6] applied the generalized exponential cure rate model to study relapse time for drug addicts. Martinez et al[7] analyzed gastric cancer data using cure fraction models with generalized modified Weibull distribution. Swain et al[8] and Omer et al[9] used mixture cure model with generalized Gompertz and exponentiated Weibull exponential distributions, respectively, to analyze melanoma patients.

Studying lifetime data can be done using statistical models or neural networks. Recently, the latest has attracted attention due to its flexibility in handling complex covariates. Faraggi-Simon network (Faraggi and Simon[10]) was first introduced as a nonlinear extension of Cox proportional hazard model. However, the network failed to outperform the traditional Cox model. Katzman et al[11] modified Faraggi-Simon network by applying new deep learning techniques which outperforms the traditional Cox model. Different networks were developed keeping the basic assumption of proportional hazards, see for example Zhu et al[12] and Zhu et al[13]. Pawley[14] introduced a parametric neural network (DeepWeibull) based on Weibull distribution that allows to analyze continuous data. Recently, Xie and Yu[15] introduced neural network to analyze mixture cure fraction models but preserving the assumption of proportional hazard.

The goal of this study is to present a parametric neural network that efficiently predicts long-term survivals without proportional hazard assumption or data discretization. Here, a Weibull mixture cure fraction neural network is presented. Weibull distribution is selected due to its flexibility in taking different shapes as well as its several applications in different fields, for example in engineering, Ng et al[16] estimated the volume of water pumping by a windmill. Velazquez et al[17] studied the reliability of electronic railway signaling system. In medical field, Yu et al[5] studied the survival function for data with testicular cancer, colon and rectum cancer, lung cancer and female breast cancer. Pascale et al[18] analyzed the survival of ischemic strokes. Lambert et al[19] studied the lifetime of cancer of the ovary and cancer of the colon patients. Achcar et al[20] estimated the model parameters for acute myelocytic leukemia and acute lymphoblastic leukemia patients. Chukwu and Folorunso[21] analyzed gastric cancer data. Yusuf and Bakar[22] analyzed acute myelogenous leukemia data. Pawley[14] used it to analyze the lifetime of breast cancer data and the survival time of seriously ill hospitalized adults.

Materials and Methods

We introduced a novel neural network based on Weibull loss function. It takes into consideration the long-term survival and efficiently analyze continuous data without discretization.

We will first explain the likelihood function under survival analysis, illustrate the basic

properties of Weibull distribution and cure fraction models. Then illustrate the derivation of the model.

Survival analysis and Weibull distribution

The main concern in survival analysis is to study the relationship between different covariates (x) and lifetime data (t_i) taking into consideration the censored observations. For n data points under right random censoring, the likelihood function can be expressed as follows

$$L_1 = \prod_{i=1}^{n} \left[f(t_i) \right]^{\delta_i} \left[S(t_i) \right]^{1-\delta_i}. \quad (1)$$

Where,

f(t): the probability density function.

S(t)= P(T> t): the survival function.

$$\delta = \begin{cases} 1, & \text{if the event is observed} \\ 0, & \text{if the event is censored} \end{cases}$$

The probability density and survival functions can take several distributions. Here, we assume that T follows Weibull distribution. The probability density function, survival function and hazard function can be expressed respectively as follows

$$f^*(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} e^{-\left(\frac{t}{\alpha}\right)^{\beta}}$$

$$S^*(t) = e^{-\left(\frac{t}{\alpha}\right)^{\beta}}$$

$$h^*(t) = \frac{f^*(t)}{S^*(t)} = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}. \quad (2)$$

Different shapes of $f(t)$ and $h(t)$ are presented in Figures 1 and 2, respectively. From Figure 1, It can be seen that Weibull distribution takes different shapes according the values of the shape parameter β. For example, when β<1 the pdf is monotonically decreasing. Also, when β =1, the pdf is monotonically decreasing but faster than that when β<1. While when β>1 the pdf increases till it reaches its peak and then monotonically decreases. The hazard rate takes three different shapes. It can be monotonically decreasing when β<1, monotonically increasing when β> 1 and constant for β=1 , as illustrated in Figure 2.

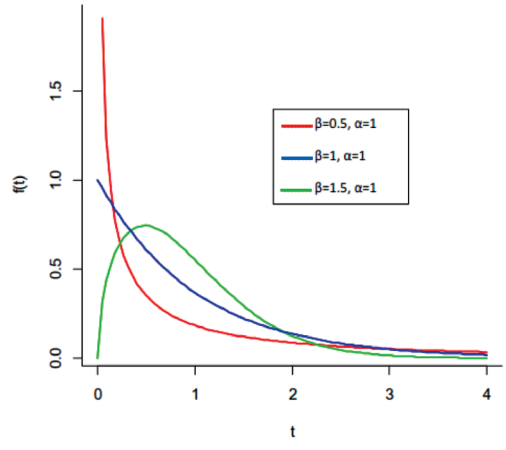

Figure 1. The probability density function for weibull distribution

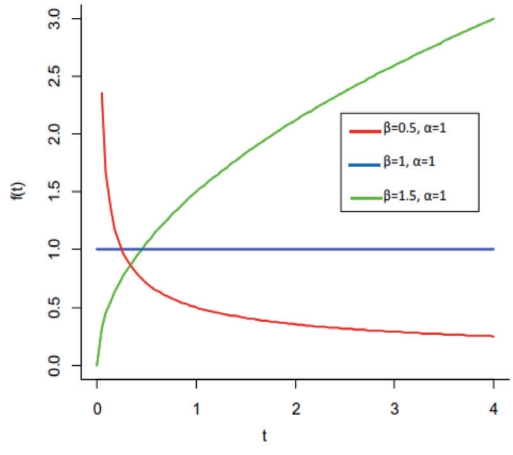

Figure 2. The hazard function for weibull distribution

Cure Fraction Models

Cure fraction models analyze a special case

of survival analysis when a portion of the population never experience the event of interest. Such subjects are referred to as cured, immune or non-susceptible. The other part of the population is those who are subject to the event of interest, they are called non-cured or susceptible. To represent the two groups of the population, we define η as follows

$$\eta = \begin{cases} 1, & \textit{if the subject is susceptible} \\ 0, & \textit{if the subject is cured} \end{cases}.$$

Let *P(η=0) = p, P(η=1) = 1 - p, F(t)* be the cumulative distribution function of the entire population and *F* (t)* is the cumulative distribution function of susceptible subjects. It is assumed that *F* (t)* is a proper cumulative distribution function, thus

$$P(T \le t \mid \eta = 1) = F^*(t)$$
$$P(T \le t \mid \eta = 0) = 0.$$

Accordingly,

$$F(t) = P(T \le t \mid \eta = 1)P(\eta = 1) + P(T \le t \mid \eta = 0)P(\eta = 0)$$
$$= (1-p)F^*(t) + 0 = (1-p)F^*(t).$$

Thus, the survival function for the whole population S(t) can be written as follows

$$S(t) = 1 - F(t) = 1 - (1-p)F^*(t) = p + (1-p)S^*(t),$$

where, *S* (t)* is the survival function of susceptible subjects.
The likelihood function in (1) can be rewritten under cure fraction model as follows

$$L_2 = \prod_{i=1}^{n} \left[(1-p_i) f^*(t_i) \right]^{\delta_i} \left[p_i + (1-p_i) S^*(t_i) \right]^{1-\delta_i}.$$

Substituting by Weibull distribution from equation (2), the likelihood function is as follows

$$L_2 = \prod_{i=1}^{n} \left[(1-p_i) \frac{\beta_i}{\alpha_i} \left(\frac{t_i}{\alpha_i} \right)^{\beta_i - 1} e^{-\left(\frac{t_i}{\alpha_i}\right)^{\beta_i}} \right]^{\delta_i} \left[p_i + (1-p_i) e^{-\left(\frac{t_i}{\alpha_i}\right)^{\beta_i}} \right]^{1-\delta_i}. \quad (3)$$

To link the covariates to the previous likelihood one can define

$$\alpha_i = \theta_1^T \boldsymbol{x}_i,$$

where, θ_1 is a vector of regression parameters and *x* is a vector of covariates.
For cure proportion, a link function must be used. The most commonly used is the sigmoid function with the following formula

$$p_i = \frac{e^{\theta_2^T \boldsymbol{x}_i}}{1 + e^{\theta_2^T \boldsymbol{x}_i}},$$

Where, θ_2 is a vector of regression parameters. This likelihood function with α_i and p_i formulas will be used in the model described in the next section.

Model Description

The goal of this model is to predict the parameters of cure fraction model under Weibull distribution using the neural network ability to handle complex covariates and maintaining the continuity of the data. First, a brief explanation of neural networks will be presented, then the structure of the presented network will be explained.
The main idea of neural network is to compute the output based on a functional relationship with the inputs. Neural network takes a weighted sum of the inputs with one additional term called a bias term illustrated as follows
Let $x_1, x_2, \ldots, x_n$ be a set of inputs with weights

$w_1, w_2, \ldots wn$, the neuron output is given by

$$z = b + \sum w_i x_i$$

Where: b is called a bias term

The output of the network depends on the objective of the study. So, the simple linear combination is not always the required output. Accordingly, instead of using z, a function $\sum = f(z)$ is considered and called an activation function. The choice of the activation function depends on the required range of the output. For example, if the output is a value between 0 and 1, sigmoid function could be a choice.

A neural network learns by adjusting the weights in order to minimize the observed errors. The prediction error is reflected through a function called the loss function. So, the network updates the weight by minimizing the loss function. Mean square error is a common choice for the loss function in most of the applications. However, in survival analysis one needs to take censoring information into consideration. This is done through the likelihood function. So, the loss function is the negative of log likelihood function.

The presented network is illustrated in Figure 3, DeepWei-Cu is a deep neural network that takes as input the covariates and output the estimates of cure proportion and Weibull distribution's parameters that fully characterize the survival function. The hidden layers of the network consist of a fully connected layer followed by a dropout layer. It consists of three fully connected hidden layers of widths 1, 2 and 1 of the covariates dimensions with relu activation. The output layer has a softplus activation function for Weibull distribution parameters to consist with the parameters' range $[0,\infty]$, and sigmoid activation function for cure proportion. Python is used to perform the analysis.

For training, back-propagation via the Adam optimizer is used. Each of the hidden layers have Xavier initialization and a dropout rate of 0.25. DeepWei-Cu is implemented in a TensorFlow environment with the Keras API.

The loss function is defined to be the negative log likelihood of equation 3, which can be written as follows

$$L = -\sum_{i=1}^{n} (\delta_i) \ln \left[(1-p_i) \frac{\beta_i}{\alpha_i} \left(\frac{t_i}{\alpha_i} \right)^{\beta_i - 1} e^{-\left(\frac{t_i}{\alpha_i}\right)^{\beta_i}} \right] + (1-\delta_i) \ln \left[p_i + (1-p_i) e^{-\left(\frac{t_i}{\alpha_i}\right)^{\beta_i}} \right]$$

In Figure 4, we illustrate a computational graph to compute the training loss of the proposed network: the inputs are the covariates x and the outputs are α ,β and p.

For model evaluation, both time dependent concordance index (C^{td}) and integrated brier score are considered and explained briefly as follows

Time dependent concordance index (C^{td}): Is an index to evaluate the discrimination ability of survival model taking into consideration censored observations. The main advantage of C^{td} over the usual concordance index is that there is no assumption for one-to-one correspondence between predicted survival probabilities and predicted times (i.e., no proportional hazard assumption). For a predicted survival probability ($\hat{S}$), the index is defined as follows

$$C^{td} = P(\hat{S}(T_i|X_i(t)) < \hat{S}(T_i|X_j(t)) | T_i < T_j \text{ and } \delta_i = 1)$$

For more information see Antolini et al.[23]

Brier score (*BS(t)*): Is a measure of the accuracy of probabilistic prediction. In survival analysis brier score is a measure of how well the model predicts the survival function. To account for censoring, inverse probability of censoring weighted Brier score is considered and defined as follows

$$BS(t)=\frac{1}{n}[\frac{\hat{S}_i(t)^2.I(t_i\le t,\delta_i=1)}{\hat{K}(t_i)}+\frac{\left[1-\hat{S}_i(t)^2\right].I(t_i>t)}{\hat{K}(t_i)}]$$

Where

n: Is the number of observations in the data.
Ŝ*i (t)*: is the predicted survival probability and *i=1,...,n*.
$\hat{K}(t_i)$: The estimated Kaplan-Meier survival function.
However, brier score only gives a predictive performance at a given time point t. To overcome this disadvantage. Integrated brier score is introduced to average the brier score over time interval. It has the following formula.

$$\text{üüüüü}(\quad)=\frac{1}{T_2-T_1}\int_{T_1}^{T_2}\quad(\quad)$$

Usually, T_1 is set to zero and T_2 is the maximum value of t_i. For more details, see Pawley[14] and Håvard and Ørnulf.[24]
The model is examined on two synthetic datasets and one real dataset. A brief description of the datasets is given below

Open Access Series of Imaging Studies (OASIS3)

OASIS3 dataset is an open access series of neuroimaging datasets freely available for scientific studies (see, www.oasis-brains.org). It aims to study alzheimer disease for adults ranging in age from 42-95 years. In our study, we are interested in lifetime from enrollment till first recognition of Alzheimer disease. Since not all participants will face alzheimer, we refer to this group as "cured".

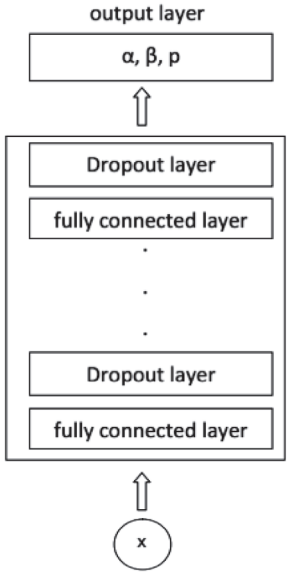

Figure 3. The architecture of Deep Wei-CU

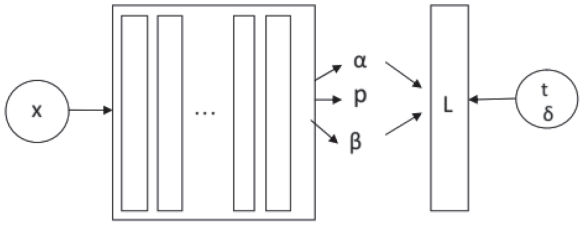

Figure 4. Computational graph to compute the training loss of Deep Wei-Cu

Clinical Dementia Rating (CDR) is commonly used to distinguish subjects with and without the disease, a value greater than zero represents demented. Hence, we restricted the sample to subjects that were non-demented at the beginning of the study. After cleaning the data of the missing, outliers and CDR > 0, our study consists of 196 participants. Table 1 gives an overview of the dataset. We used in the analysis 8 features (gender, mini-mental

state examination score, weight, height, apolipoprotein E, logarithm of geriatric depression scale, intracranial volume and subcortical gray matter volume).

To test for the applicability of cure fraction models, one usually examines the Kaplan–Meier curve. If there is a long plateau at the later part of curve, then there may be a subgroup of cured subjects. From Figure 5, It can be seen that there is a long plateau in the Kaplan-Meier curve. Hence, we can use cure fraction model to analyze OASIS3 dataset.

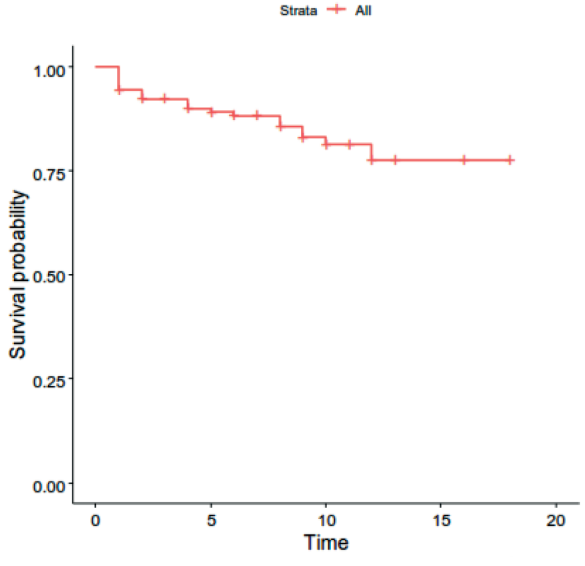

Figure 5. Kaplan-Meier Curve for OASIS3 dataset

Synthetic

Table 1 gives an overview of synthetic datasets. Four datasets are generated with the following characteristics:

Linear Weibull (LW)

Weibull model with linear relationships between the covariates and the parameters with 40,000 sample units.

Non-linear Weibull (NLW)

Weibull model with cubic relationships between the covariates and the parameters with 40,000 sample units.

The data is generated as follows

1) For each unit (i= 1,...,n):

a) generate two covariates from uniform(-1,1) or standard normal

b) Set

$\alpha_i = 20 + 10x_{1i} - 10x_{2i}$, for linear data.

$\alpha_i = 80 + 40x_{1i}^3 - 30x_2^{i2} - 5x_{1i}$, for non-linear data.

$$p_i = \frac{e^{3x_{1i}-2x_{2i}}}{1+e^{3x_{1i}-2x_{2i}}}$$

c) Generate t_i from Weibull with parameters α_i and β=1 and 1.1 for linear and nonlinear, respectively.

2) Randomly select 60% of the units to be censored.

3) For censored data, generate a right censored event time from uniform $(0,t_i)$.

4) Set δ_i=0 for censored data, otherwise δ_i=1.

5) If $t_i > 500$, set t_i= 500. This is done to keep the harmony of the data.

Results

To study the performance of DeepWei-Cu, we

Table 1. Descriptive statistics of OASIS3 and synthetic datasets.

	No. of uncensored	No. of censored	No. of features
OASIS3	26 (13.3%)	170 (86.7%)	8
synthetic	16000 (40%)	24000 (60%)	2

Table 2. Concordance Index (integrated brier score) for linear Weibull (NL) and nonlinear Weibull (NL) synthetic datasets under uniform covariates.

		Oracle	DeepWei-Cu	RW
Datasets	LW	0.6405 (0.3632)	0.5819 (0.2610)	0.5738 (0.4186)
	NLW	0.5021 (0.3248)	0.5067 (0.2891)	0.5007 (0.3941)

Table 3. Concordance Index (integrated brier score) for linear Weibull (NL) and nonlinear Weibull (NL) synthetic datasets under standard normal covariates.

		Oracle	DeepWei-Cu	RW
Datasets	LW	0.5326 (0.3827)	0.5197 (0.1855)	0.5118 (0.2623)
	NLW	0.5529 (0.3927)	0.5512 (0.2543)	0.5350 (0.2557)

Table 4. Concordance Index (integrated brier score) for OASIS3

DeepWei-Cu	RW
0.8986 (0.0929)	0.8620 (0.0964)

compare the results with the usual parametric statistical cure fraction model. The results illustrate that DeepWei-Cu outperforms Weibull regression.

For evaluation, we split the data into three randomly train/test datasets, with 80% of the units in each training set and the remaining 20% in the test set. We reserved 20% of the training set as a validation set. All datasets are generated with 60% censoring. To test the performance of the model, we used time dependent concordance index (C^{td}) and integrated brier score.

The concordance index and the integrated brier score for synthetic and real datasets are illustrated in Tables 2 till 4, respectively. High C^td and low integrated brier score indicate better model performance in learning the patients' survival distribution.

In synthetic datasets we can obtain the oracle C^{td} and oracle integrated brier score using the true survival distribution. This helps in providing a benchmark against which to compare the model. If the model approaches near the oracle, then it has likely learned the true distribution. Sometimes due to random chance, a model may beat the oracle metric (Pawley [14]). From Table 2 and 3, It can be seen that the concordance index and the integrated brier score for DeepWei-Cu are very close to the oracle one, which indicate the model's ability to learn the true survival distribution. Also, it is clear that DeepWei-Cu has higher C^{td} and lower integrated brier score than that of the regression Weibull (RW) in both linear and nonlinear cases.

In real world, we can't guarantee that the data is generated from Weibull distribution. To compare the model and test its applicability, we apply it on OASIS3 real dataset. From table 4, it can be seen that, DeepWei-Cu has higher C^{td} and lower integrated brier score than that of RW, which indicates that Deep-WEi-Cu has a better performance. Also, the concordance index has high value (> 0.8) indicating a strong model.

Discussion

Survival analysis models have a pre-assumption

of the occurrence of the event of interest for all observations. However, in some studies like Alzheimer, diabetic and cancer, we face the situation where part of the population may never face the event of interest. In this case, we need to take into consideration the effect of long-term survivors to guarantee accurate analysis.

In our study, we introduced a novel cure fraction Weibull neural network. It has the advantage of analyzing continuous data without the need of discretization. Also, we took into consideration the effect of long-term survivals in the loss function. Besides, benefiting from the flexibility and properties of Weibull distribution by not restricting to proportional hazard assumption.

To evaluate the network performance, both discriminative ability and accuracy are tested using time dependent concordance index and integrated brier score, respectively. The analysis was performed on both synthetic and real datasets. Two synthetic datasets were considered using linear and nonlinear covariates. Comparing the results of the proposed network with Weibull regression, the former performed better in both covariates cases and model performance matrices.

A straight forward generalization to this work can be done by changing the used distribution. Also, in this study we considered mixture cure fraction models, one can study non-mixture ones. Moreover, we only considered the case of structured covariates, a generalization can be made to handle both structured and unstructured ones. Furthermore, we used backpropagation technique in training the network, other methods could be used. Finally, it is suggested for further studies to take into consideration any pre-existing information about the parameters. This can be done using Bayesian technique.

Conclusion

This paper presents a novel model (DeepWei-Cu) to be used in the mixture cure fraction survival analysis. It trains a neural network to predict the parameters of Weibull distribution and cure proportion with respect to the covariates under right random censoring. To the best of our knowledge, this is the first time a parametric neural network is used in analyzing lifetime data with cure subgroup.

DeepWei-Cu has the flexibility of analyzing continuous data without discretization. Also, it has the advantage of using Weibull distribution properties. For example, it can analyze data with different hazard rates (monotonically decreasing, monotonically increasing and constant). As a test, we compared the performance of DeepWei-Cu with the performance of previous statistical models. The numerical results showed that, DeepWei-Cu performed better.

Conflicts of interest

The authors declare that they have no conflicts of interest.

References

1. Amico, M. and Keilegom, I., V. Cure models in survival analysis. Annual Review of Statistics and Its Application. 2018; 5: 18.1-18.32.

2. Boag, J., W. Maximum likelihood estimates of patients cured by cancer therapy. Journal of the Royal Statistical Society. Series B. 1949; 11(1): 15-53.

3. Farewell, V., T. The use of mixture models for the analysis of survival data with long-term survivors. Biometrics. 1982; 38: 1041-1046.

4. Yamaguchi, K. Accelerated failure-time regression models with a regression model of surviving fraction: an application to the analysis of "permanent employment" in Japan. Journal of the American Statistical Association. 1992; 87(418): 284-292.

5. Yu, B., Tiwari, R., C., Cronin, K., A. and Feuer, E. J. Cure fraction estimation from the mixture cure models for grouped survival data. Statistics in Medicine. 2004; 23: 1733-1747.

6. Kannan, N., Kundu, D., Nair, P. and Tripathi, R., C. The generalized exponential cure rate model with covariates. Journal of Applied Statistics. 2010; 37(10): 1625-1636.

7. Martinez, E., Z., Achcar, J., A., Jacome, A., A. and Santos, J., S. Mixture and non-mixture cure fraction models based on the generalized modified Weibull distribution with an application on gastric cancer data. Computer Methods and Programs In Biomedicine. 2013, 112(3); 343-355.

8. Swain, P., K., Grover, G. and Goel, K. Mixture and non mixture cure fraction models based on generalized gompertz distribution under Bayesian approach. Tatra Mountains Mathematical Publication. 2016; 66: 11-135.

9. Omer, M., Abu Bakar, M., Adam, M. and Mustafa, M. Cure models with exponentiated Weibull exponential distribution for the analysis of melanoma patients. Mathematics. 2020; 8(11), 1926; https://doi.org/10.3390/math8111926.

10. Faraggi, D., and Simon, R. A neural network model for survival data. Statistics in Medicine. 1995; 14:73-82.

11. Katzman, J., Shaham, U., Cloninger, A., Bates, J., Jiang, T. and Kluger,'Y. Deep Survival: A Deep Cox Proportional Hazards Network. Stat. 2016, 1050 (2), 1–10; https://doi.org/10.1186/s12874-018-0482-1.

12. Zhu, X., Yao, J and Huang, J. Deep Convolutional Neural Network for Survival Analysis with Pathological Images. IEEE International Conference on Bioinformatics and Biomedicine. 2016: DOI: 10.1109/BIBM.2016.7822579.

13. Zhu,X., Yao, J., Zhu, F. and Huang, H. WSISA: Making Survival Prediction from Whole Slide Histopathological Images. IEEE Conference on Computer Vision and Pattern Recognition. 2017: DOI: 10.1109/CVPR.2017.725.

14. Pawley, M. DeepWeibull: a deep learning approach to parametric survival analysis. M.Sc. Thesis, Departement of Mathematics, Imperial College London. 2020.

15. Xie, Y. and Yu, Z. Mixture cure rate models with neural network estimated nonparametric components. Computational Statistics. 2021; 36, 2467-2489.

16. Ng, S,. K., Mclachlan, G., J., Yau, K., W., Lee, A., H. Modelling the distribution of

ischaemic stroke-specific survival time using an EM-based mixture approach with random effects adjustment. Statistics in Medicine. 2004; 23: 2729-2744.

17. Velazquez, M., Rodriquez, J., Carmen, M., Murrieta, F. and Eslava, G. Application of the Weibull Distribution to Estimate the Volume of Water Pumping by a Windmill. Journal of Power and Energy Engineering. 2016; 4(9): 36-51.

18. Pascale, E., Freneaux,T., Sista, R., Sannino, P., Marmo, P., Bouillaut, L. Application of the Weibull distribution for the optimization of maintenance policies of an electronic railway signaling system. European Safety and Reliability Conference. 2017; 8p: hal-01521640.

19. Lambert, P., C., Thompson, J., R., Weston, C., L. and Dickman, P., W. Estimating and modeling the cure fraction in population-based cancer survival analysis. Biostatistics. 2007; 8(3): 576-594.

20. Achcar, J., A., Coelho-Barros, E., A. and Mazucheli, J. Cure fraction models using mixture and non- mixture models. Tatra Mountains Mathematical Publications. 2012; 51: 1-9.

21. Chukwu, A., U. and Folorunso, S., A. Determinant of flexible parametric estimation of mixture cure fraction model: an application of gastric cancer data. West African Journal of Industrial and Academic Research. 2015; 15(1): 139-156.

22. Yusuf, M., U. and Bakar, M., R. Cure models based on Weibull distribution with and without covariates using right censored data. Indian Journal of Science and Technology. 2016; 9(28): 1-12.

23. Antolini, L., P. Boracchi, and E. Biganzoli. A time-dependent discrimination index for survival data. Statistics in Medicine. 2005; 24(24): 3827- 3944.

24. Håvard Kvamme and Ørnulf Borgan (2023). The Brier Score under Administrative Censoring: Problems and Solutions. Journal of Machine Learning Research.2023; 24(2): 1-26.