BroadSurv: A Novel Broad Learning System-based Approach for Survival Analysis

Guangheng Wu
College of Information Science and Technology
Jinan University
Guangzhou, China
wuguangheng2002@gmail.com

Junwei Duan
College of Information Science and Technology
Jinan University
Guangzhou, China
Corresponding Author: jwduan@jnu.edu.cn

Jing Wang
School of Computer Science
Guangdong polytechnic Normal University
Guangzhou, China
Corresponding Author: wj_adr@163.com

Lu Wang
Academic Affairs Office
Jinan University
Guangzhou, China
Corresponding Author: luwang1214@jnu.edu.cn

Cheng Dong
School of Intelligent System Science and Engineering
Jinan University
Zhuhai, China
donc@163.com

Changwei Lv
Shenzhen Institute of Information Technology
Shenzhen, China
lvcw@sziit.edu.cn

***Abstract*—Survival analysis (time-to-event analysis) is a set of statistic methods to analyze time-to-event data and is widely used in many fields such as economics, finance and medicine. One of the fundamental problems in survival analysis is to explore the relationship between the covariates and the survival time. Recently, with the development of deep learning-based techniques, various approaches have been proposed for survival analysis. To better handle the censoring, special cost functions or sophisticated network structures are usually designed for these methods. In this paper, a novel two-stage method is proposed to model the survival data. In the first stage, pseudo conditional probabilities are computed, which can act as the quantitative response variables in regression problems. In the second stage, with these pseudo values, a complicated survival analysis problem is transformed into a regression problem that can be effectively solved by broad learning system. The experimental results show that, with a flexible structure and a simple cost function, our proposed method has a better performance in handling the censored problems.**

***Index Terms*—survival analysis, neural network, broad learning system (BLS), pseudo survival probability**

I. Introduction

Survival analysis, also called time-to-event analysis, has a wide range of applications in medicine, finance and many other fields [1]. One of the main goals of survival analysis is to predict the survival probability by modeling the survival function based on the observed covariates. Survival data is usually more complicated because the time of interest is not fully observed for each subject in the research and this has become an inevitable problem in survival analysis. Moreover, sometimes the data may be high dimensional, such as genomics data and radiomics data. Researchers have developed various parametric and semiparametric models to handle the problem of survival analysis. However, these models are restricted to some prior assumptions, such as the proportional hazard (PH) assumption [2], which may not be satisfied in the real applications. Machine learning-based methods such as support vector machine (SVM) [3] and neural networks also have been introduced to survival analysis. Recently, a newly born methodology named broad learning system (BLS) [4] has revolutionized the artificial intelligence methods for faster training speed and higher accuracy. In this paper, a two-stage method with a discrete-time survival framework is proposed for survival analysis based on broad learning system. In the first stage, the pseudo conditional probabilities are computed for each subject based on the pseudo-observations method. In the second stage, the conditional survival probabilities are calculated by BLS with the use of these pseudo values. Compared to other PH-based survival models, our proposed model can generate the survival probability directly. What's more, with a simple architecture, our model can easily adapt to different datasets and perform efficiently.

II. Background

In this section, we first briefly describe the survival analysis and then make a description of *jackknife pseudo-observations*,

which is regarded as a way to analyze survival data. The last part is some basic concepts of classic broad learning system.

A. Survival Data

Survival data consists of three parts: observed baseline covariates X, observed survival time T, and an censoring indicator I denoting the type of event. Time T is the time since the first collection of covariates. And if an event is observed, the indicator $I = 1$, if not observed, then $I = 0$ (right-censored). In this paper, survival time is treated as discrete and the time horizon is finite.

B. Pseudo-observations

As we know, with complete data, we can simply establish a model $I(T_i > t)$ on $X_i(i = 1, ..., m)$, where $X_i = (x_{i1}, ..., x_{iq})$ denotes the q-dimensional covariates, with a generalized linear model if the survival outcome is survival probability at specific time t. When facing with the presence of censoring, pseudo-observations can help achieve the goal of transforming a survival analysis problem into a standard regression problem.

The theory of pseudo-observations used for survival analysis is based on the well-known jackknife method [5]. Pseudo-observations tries to handle the main problem of survival data, that is, not all subjects have appropriate responses in the study. [6]–[9] have provided an effective and direct method to explore the relationship between the observed covariates and the survival outcome when survival data is censored.

The basic idea of pseudo-observations is simple. Assume an estimator δ is available for the expectation $\delta = E[f(T)]$, for example, the Kaplan–Meier (KM) estimator [10] for $S(t) = E[I(T > t)]$ with m samples, here $S(t)$ is the survival function which represents the survival probability at a specific time t. Then the jth pseudo-observation for $f(T)$ is defined as

$$\hat{\delta}_j = m\hat{\delta} - (m-1)\hat{\delta}_{-j}, j = 1, ..., m, \tag{1}$$

where $\hat{\delta}_{-j}$ is an estimator similar to $\hat{\delta}$ applied to the observations $i \neq j$. In other words, for the jth observation, the pseudo survival probability is calculated as

$$\hat{P}_j(t) = m\hat{P}(t) - (m-1)\hat{P}_{-j}(t), j = 1, ..., m, \tag{2}$$

where $\hat{P}(t)$ is the KM estimator of $S(t)$. The pseudo values are generated once, the key point of pseudo-observations method is to replace the censored observed $f(t_j)$ by $\hat{\delta}_j$. That is to say, we could simply set up a regression model with $\hat{\delta}_j$, which is just similar to directly fitting $I(T_i > t), i = 1, ..., m$ with a regression model if these values were observed, or it may be used to calculate residuals.

C. Broad Learning System

Over the past few years, broad learning system (BLS) has raised a revolution in conventional artificial intelligence methods. With the advantages of simple structure and incremental learning algorithm, broad learning system can achieve faster training speed and higher accuracy, which can be extended to other research fields [11], [12]. BLS mainly consists of three important parts: the mapping feature nodes, the enhancement feature nodes and the desired connecting matrix. And there are two stages in the training of BLS: 1) randomly generating the connecting weights for the mapping feature nodes and the enhancement feature nodes; 2) calculating the weights. The structure of classic broad learning system is shown in Fig. 1.

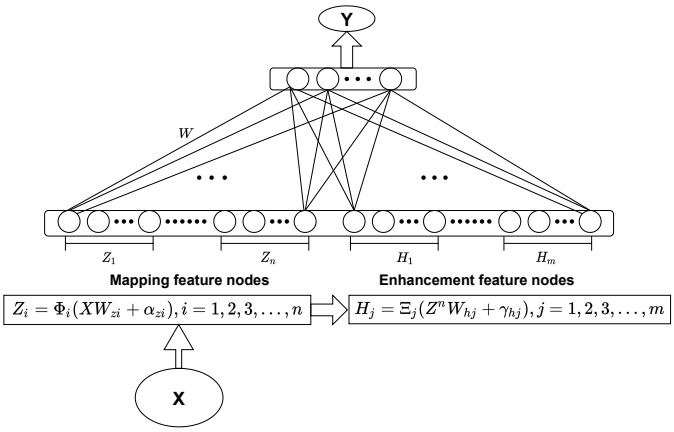

Fig. 1: Classic Broad Learning System

Suppose the training data of a supervised task has the form of $\{(X, Y)|X \in \mathbb{R}^{N\times M}, Y \in \mathbb{R}^{N\times K}\}$ from K classes, where N notes the sample size, M notes the feature dimension. For mapping feature node, it contains n groups, each consisting of k nodes, for enhancement feature node, it contains m groups and there are q nodes in a group. In this way, the ith group of mapping feature nodes can be expressed as

$$Z_i = \Phi_i(XW_{zi} + \alpha_{zi}), i = 1, 2, 3, ..., n, \tag{3}$$

where weights W_{zi} and deviation α_{zi} are randomly initialized. Next, collecting these mapping features together and denote it as $Z^n = [Z_1, Z_2, ..., Z_n]$. With the help of Z^n, BLS will expand the enhancement layer as the equation

$$H_j = \Xi_j(Z^n W_{hj} + \gamma_{hj}), j = 1, 2, 3, ..., m, \tag{4}$$

where $\Xi_j(\cdot)$ is an optional nonlinear activation function. The output of the enhancement feature nodes can be expressed as $H^m = [H_1, H_2, ..., H_m]$. Concatenating the mapping nodes and enhancement nodes together, the final form of BLS could be written as the equation

$$\begin{aligned} Y &= [Z_1, ..., Z_n|H_1, ..., H_m]W \\ &= [Z^n|H^m]W. \end{aligned} \tag{5}$$

Then in the second stage, pseudoinverse algorithm can be used to obtain the weights. In this case, the output weights of the final BLS can be calculated as

$$W = [Z^n|H^m]^+ Y. \tag{6}$$

Nevertheless, it is costly to directly compute the pseudoinverse $[Z^n|H^m]^+$ with some standard methods as the dimension may be too large in the training data. Instead, we can solve the pseudoinverse in this way:

$$arg \quad \min_W : ||BW - Y||_a^{\sigma_1} + \lambda||W||_b^{\sigma_2} \tag{7}$$

where $\sigma_1 > 0$ and $\sigma_2 > 0$, and both a and b are norm regularization. The constraint term λ is used to ensure that we can find the pseudoinverse even the generalized inverse is under ill condition. Let $B = [Z^n|H^m], \sigma_1 = \sigma_2 = a = b = 2$, the weights can be approximated as

$$W = (\lambda I + BB^T)^{-1}B^T Y. \tag{8}$$

Specially, B^+ can computed as

$$B^+ = \lim_{\lambda \to 0}(\lambda I + BB^T)^{-1}B^T. \tag{9}$$

In short, BLS adopts a flat structure, which makes it easy to train and extend. And the applications of BLS are summarized in [11]. From [11], we can get that BLS can achieve a more generalized performance compared to other conventional methods.

III. Related Works

The most common survival model in the research is the Cox proportional hazards regression (CPH) [2], which is capable of incorporating observations' covariates. CPH holds the basic assumption that the logarithm of the hazard rate is a linear function of covariates, and the hazard rate is constant, which is known as proportional hazards (PH) assumption. Similar to CPH, other existing models are more or less based on different assumptions about the potential random process and the assumed relationship between covariates and the hypothetical process parameters. For example, [13] assumes a Wiener process, while the assumption of [14] is Markov Chain. However, these models are of limited use for they are restricted to the specific strong assumptions.

Survival analysis problem has also drawn great interest in the field of machine learning. With a feed-forward neural network, Faraggi-Simon method [15] applied the neural networks to survival analysis firstly, and it lays the foundation for the nonlinear proportional hazards model. More recently, [1], [16] extend the previous general approach with more complicated network architectures or loss functions. In all, existing survival models based on neural network require the elaborate design of special loss function, sometimes even need a more complicated network structure, such as a CNN or RNN [17], [18] to handle the censored data.

Another kind of classic model is the discrete-time survival model [16], [18]–[21]. In the framework of discrete-time survival model, the observed time is divided into several intervals, and then the conditional probability is estimated for each interval. The discrete-time survival model is not subject to any assumption so that it is more flexible than the PH-based model. In a creative way, DNNSurv [22] provides another new idea for the modeling of survival data, which introduces the pseudo-observations method to transform a survival analysis problem with censored data into a standard regression problem with a discrete-time survival framework.

IV. Proposed Method

In this section, we will describe the structure of our proposed two-stage method, and give the algorithm of our method in detail.

A. Architecture of Model

In this paper, with the use of discrete-time survival framework and pseudo-observations, we attempt to apply broad learning system to survival analysis. In the first stage of our proposed method, we divide the observed time into s intervals which are denoted as: $(0, t_1], (t_1, t_2], ..., (t_{s-1}, t_s]$. Then calculate the pseudo conditional survival probability for each interval, where the conditional survival probability refers to the probability surviving the specific time interval under the condition that having survived in the interval before. Giving a time interval $(t_k, t_{k+1}]$, let Ω_k denote the observations who are at risk ($\Omega_0 = n$), the pseudo conditional survival probability can be calculated as

$$\hat{P}_{ik}(t_{k+1}|\Omega_k) = \Omega_k \hat{P}(t_{k+1}|\Omega_k) - (\Omega_k - 1)\hat{P}_{-i}(t_{k+1}|\Omega_k), \tag{10}$$

where $\hat{P}_{ik}(t_{k+1}|\Omega_k)$ is the KM estimator that uses the remaining survival time of observations who are in Ω_k, and so does $\hat{P}_{-i}(t_{k+1}|\Omega_k)$ but excluding the ith observation. According to the nature of pseudo observations discussed in [9], we can know that the pseudo value we obtain is 0 or 1 when there is no censoring, otherwise, the pseudo probability will be a real value. In this way, an observation's survival time was transformed into a battery of pseudo conditional survival probabilities finally. Now what we face is just a standard regression problem, which can be solved by BLS [23] easily. It is noteworthy that the number of intervals should be decided cautiously. There are two main strategies for the decision of time points, one is to set the time points on the percentage of data with equal spacing, the other is to set the time point on the research time with equal spacing. [8] points out that, in most cases, five time points evenly distributed on the research time space has already behaved quite well.

Alternatively, we can create an indicator variable for each time point, just as using $d(t)$ for time point t respectively. Suppose that $X = [X_1, ..., X_m]$ is the observed covariates, let $D = [D_1, ..., D_m]$ represent the indicator variables, where $D_m = (d_{m1}, ..., d_{ms})$ denotes the s-dimensional indicator vectors, then the input of BLS can be written as $\boldsymbol{X} = [X|D]$. Our proposed structure is just as Fig. 2 shows.

In addition, our proposed model can predict not only the conditional survival probability, but also the marginal survival probability. After obtaining the conditional survival probability $P(T > t_i|T > t_{i-1}), i = 1, ..., s$ from the model, marginal survival probability $P(T > t_i)$ can be easily computed as

$$P(T > t_i) = \prod_{k=1}^{i} P(T > t_k|T > t_{k-1}). \tag{11}$$

The algorithm of our proposed BroadSurv is shown in Algorithm 1.

V. Experiments

In this section, we will verify and evaluate our proposed method through a series of experiments on real applications.

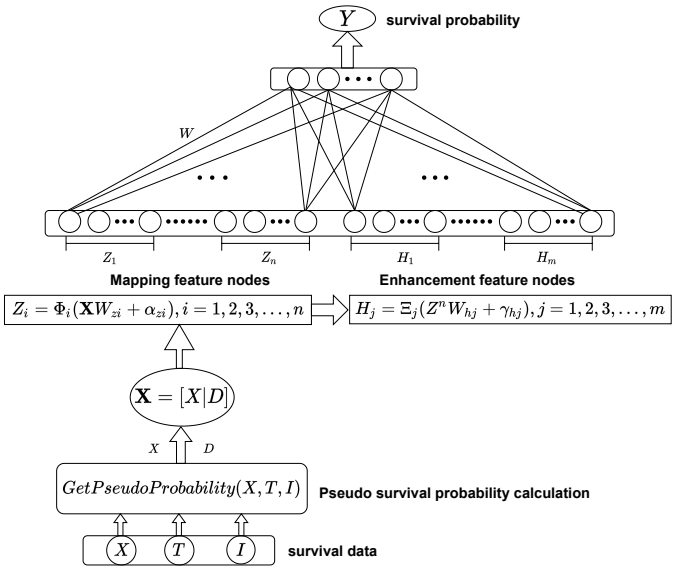

Fig. 2: Architecture of BroadSurv

Algorithm 1: BroadSurv

Input : survival data $\{X, T, I\}$;
number of groups of mapping nodes n;
number of groups of enhancement nodes m;
Output: survival probability Y.

Calculate the pseudo conditional survival probabilities $X, D = GetPseudoProbability(X, T, I)$;
Set the input of BLS $\mathbf{X} = [X|D]$;
Initialize $i = j = 1$;
while $i \leq n$ **do**
  Randomly initialize W_{zi}, β_{zi};
  Calculate $Z_i = \Phi_i(\mathbf{X}W_{zi} + \beta_{zi})$;
Get the mapping groups $Z^n = [Z_1, Z_2, ..., Z_n]$;
while $j \leq m$ **do**
  Randomly initialize W_{hj}, β_{hj};
  Calculate $H_j = \Xi_j(\mathbf{X}W_{hj} + \beta_{hj})$;
Get the enhancement groups $H^m = [H_1, H_2, ..., H_m]$;
Set $B = [Z^n|H^m]$;
Calculate B^+ with Eq. (9);
Calculate W with Eq. (8);
Calculate survival probability $Y = \mathbf{X}W$.

A. Metrics

In order to verify the performance of our model, we used the criteria concordance index (C-index) [14] to evaluate the performance of survival model in the follow-up experiments. C-index represents the scores of all observation pairs with correct order of survival prediction pairs, that is, the probability that the predicted survival time of two subjects is the same as the actual survival time order. C-index is quite useful to evaluate proportional hazards models for it only depends on the order of predictions. See [14] for more details. What's more, C-index is a real value between 0 and 1, where 0.5 is the result of random predictions (if C-index < 0.5, it indicates a bad performance worse than a random guess), 1 is best concordance and 0 is complete anti-concordance.

B. Datasets

We apply our proposed model to the following four real world, publicly available survival datasets, which are introduced in [24]. The brief descriptions of these datasets are as followed:

METABRIC The Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) dataset includes gene expression profiles and clinical features used to identify subgroups of breast cancer. There are 1980 patients in total, 58% of deaths were observed, with a median survival of 116 months. [25].

SUPPORT The Study to Understand Prognoses Preferences Outcomes and Risks of Treatment (SUPPORT) is a larger dataset from study designed to learn about the survival of late patients who accept life-support [26]. There are totally 9,105 patients and 14 covariates in the dataset.

Rot. & GBSG The Rotterdam tumor bank dataset(Rot.& GBSG) includes 1,546 patients who suffer from node-positive breast cancer, of whom approximately 90% have an observed time of death [27].

FLCHAIN The Assay of Serum Free Light Chain (FLCHAIN) [28] has a total number of 7,874 subjects with 8 observed covariates.

After preprocessing, such as selecting the numerical covariates and removing the missing values, a summary of the data we used is provided in Table I.

TABLE I: The summary of adopted datasets

Datasets	Training Size	Testing Size	Covariates	Censoring Rate
METABRIC	1,523	381	9	0.42
SUPPORT	7,098	1,775	14	0.32
Rot.& GBSG	1,546	686	7	0.43
FLCHAIN	4,567	1,957	8	0.70

C. Results

We will compare our proposed method with classic Cox proportional hazards regression (CPH) [2], DeepSurv [1] and DNNSurv [22]. Both CPH and DeepSurv are PH-based method, and DeepSurv is a model based on deep learning method. DNNSurv also adopts a discrete-time survival framework. For our proposed BroadSurv and DNNSurv, we first choose 5 time points evenly distributed on the event time space to evaluate their performances. Table II shows the C-index of these models. As is shown in the table, our proposed BroadSurv outperforms other survival models in the datasets METABRIC, Rot.&GBSG and FlCHAIN. Compared to the PH-based models like CoxPH and DeepSurv, BroadSurv can actually achieve a better performance. In addition, since BroadSurv and DNNSurv both use the similar discrete-time survival framework, we also evaluate the performances of BroadSurv and DNNSurv with the number of intervals ranging from 2 to 10. The result is as Fig. 3 shows. Compared to the

TABLE II: C-index of experiments

Methods	METABRIC	SUPPORT	Rot.&GBSG	FLCHAIN
CoxPH	0.628	0.589	0.666	0.790
DeepSurv	0.636	0.611	0.674	0.790
DNNSurv	0.631	**0.608**	0.687	0.802
BroadSurv	**0.643**	0.601	**0.696**	**0.802**

similar discrete-time survival model DNNSurv, our proposed BroadSurv also behaves well.

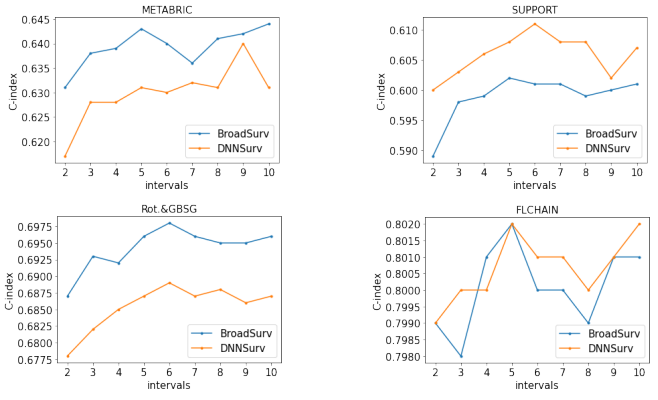

Fig. 3: C-index of BroadSurv and DNNSurv

VI. Conclusion

In this paper, a novel two-stage broad learning system-based method is proposed to make prediction of survival probability in survival analysis. In the first stage, we calculate pseudo survival probabilities with a discrete-time survival framework, and then we make risk predictions based on broad learning system in the second stage. With pseudo values and discrete-time survival framework, the analysis of censored survival data can be first simplified to a standard regression problem and then effectively addressed by broad learning system. To our knowledge, it is the first time broad learning system is applied to survival analysis with censored data. The experimental results demonstrates that our proposed BroadSurv outperforms the PH-based methods in the real survival datasets when the PH assumption is not satisfied. Unlike the hazard ratio and prognostic index in the PH-based models, BroadSurv can directly compute the survival probability. Compared to discrete-time survival framework like DNNSurv, BroadSurv can also achieve superior or comparable performance.

Acknowledgment

This work was supported in part by the Research Platforms and Program of the Education Department of Guangdong Province in China under Grant 2017KTSCX113, in part by the Fundamental Research Funds for the Central Universities, Jinan University under Grant 21619412, in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2021A1515011999, in part by the National Key R&D Program of China under Grant 2018YFC2002500, in part by the Research Platform and Projects of Guangdong Education Department under Grant 2018GkQNCX028, and in part of Stable Support Plan for Colleges and Universities of Shenzhen under Grant SZIITWDZC2021B03.

References

[1] Jared L Katzman, Uri Shaham, Alexander Cloninger, Jonathan Bates, Tingting Jiang, and Yuval Kluger. Deepsurv: personalized treatment recommender system using a cox proportional hazards deep neural network. *BMC medical research methodology*, 18(1):1–12, 2018.

[2] David R Cox. Regression models and life-tables. breakthroughs in statistics. *New York*, 372:527–41, 1992.

[3] Vanya Van Belle, Kristiaan Pelckmans, Sabine Van Huffel, and Johan AK Suykens. Support vector methods for survival analysis: a comparison between ranking and regression approaches. *Artificial intelligence in medicine*, 53(2):107–118, 2011.

[4] CL Philip Chen and Zhulin Liu. Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. *IEEE transactions on neural networks and learning systems*, 29(1):10–24, 2017.

[5] Rupert G Miller. The jackknife-a review. *Biometrika*, 61(1):1–15, 1974.

[6] Per Kragh Andersen, John P Klein, and Susanne Rosthøj. Generalised linear models for correlated pseudo-observations, with applications to multi-state models. *Biometrika*, 90(1):15–27, 2003.

[7] John P Klein and Per Kragh Andersen. Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. *Biometrics*, 61(1):223–229, 2005.

[8] Per K Andersen and John P Klein. Regression analysis for multistate models based on a pseudo-value approach, with applications to bone marrow transplantation studies. *Scandinavian Journal of Statistics*, 34(1):3–16, 2007.

[9] Per Kragh Andersen and Maja Pohar Perme. Pseudo-observations in survival analysis. *Statistical methods in medical research*, 19(1):71–99, 2010.

[10] Edward L Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.

[11] Xinrong Gong, Tong Zhang, CL Philip Chen, and Zhulin Liu. Research review for broad learning system: Algorithms, theory, and applications. *IEEE Transactions on Cybernetics*, 2021.

[12] Junwei Jin, Yanting Li, Tiejun Yang, Liang Zhao, Junwei Duan, and C.L. Philip Chen. Discriminative group-sparsity constrained broad learning system for visual recognition. *Information Sciences*, 576:800–818, 2021.

[13] Mei-Ling Ting Lee and GA Whitmore. Proportional hazards and threshold regression: their theoretical and practical connections. *Lifetime data analysis*, 16(2):196–214, 2010.

[14] Ira M Longini Jr, W Scott Clark, Robert H Byers, John W Ward, William W Darrow, George F Lemp, and Herbert W Hethcote. Statistical analysis of the stages of hiv infection using a markov model. *Statistics in medicine*, 8(7):831–843, 1989.

[15] David Faraggi and Richard Simon. A neural network model for survival data. *Statistics in medicine*, 14(1):73–82, 1995.

[16] Margaux Luck, Tristan Sylvain, Héloïse Cardinal, Andrea Lodi, and Yoshua Bengio. Deep learning for patient-specific kidney graft survival analysis. *arXiv preprint arXiv:1705.10245*, 2017.

[17] Xinliang Zhu, Jiawen Yao, and Junzhou Huang. Deep convolutional neural network for survival analysis with pathological images. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 544–547. IEEE, 2016.

[18] Eleonora Giunchiglia, Anton Nemchenko, and Mihaela van der Schaar. Rnn-surv: A deep recurrent model for survival analysis. In *International Conference on Artificial Neural Networks*, pages 23–32. Springer, 2018.

[19] Michael F Gensheimer and Balasubramanian Narasimhan. A scalable discrete-time survival model for neural networks. *PeerJ*, 7:e6257, 2019.

[20] Stephane Fotso. Deep neural networks for survival analysis based on a multi-task framework. *arXiv preprint arXiv:1801.05512*, 2018.

[21] Changhee Lee, William R Zame, Jinsung Yoon, and Mihaela van der Schaar. Deephit: A deep learning approach to survival analysis with competing risks. In *Thirty-second AAAI conference on artificial intelligence*, 2018.

[22] Lili Zhao and Dai Feng. Dnnsurv: Deep neural networks for survival analysis using pseudo values. *arXiv preprint arXiv:1908.02337*, 2019.

[23] CL Philip Chen, Zhulin Liu, and Shuang Feng. Universal approximation capability of broad learning system and its structural variations. *IEEE transactions on neural networks and learning systems*, 30(4):1191–1204, 2018.

[24] Håvard Kvamme, Ørnulf Borgan, and Ida Scheel. Time-to-event prediction with neural networks and cox regression. *arXiv preprint arXiv:1907.00825*, 2019.

[25] Christina Curtis, Sohrab P Shah, Suet-Feung Chin, Gulisa Turashvili, Oscar M Rueda, Mark J Dunning, Doug Speed, Andy G Lynch, Shamith Samarajiwa, Yinyin Yuan, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, 486(7403):346–352, 2012.

[26] William A Knaus, Frank E Harrell, Joanne Lynn, Lee Goldman, Russell S Phillips, Alfred F Connors, Neal V Dawson, William J Fulkerson, Robert M Califf, Norman Desbiens, et al. The support prognostic model: Objective estimates of survival for seriously ill hospitalized adults. *Annals of internal medicine*, 122(3):191–203, 1995.

[27] M Schumacher, G Bastert, H Bojar, K Hübner, M Olschewski, W Sauerbrei, C Schmoor, C Beyerle, RL Neumann, and HF Rauschecker. Randomized 2 x 2 trial evaluating hormonal treatment and the duration of chemotherapy in node-positive breast cancer patients. german breast cancer study group. *Journal of Clinical Oncology*, 12(10):2086–2093, 1994.

[28] Terry M Therneau and Thomas Lumley. Package 'survival'. *Survival analysis Published on CRAN*, 2(3):119, 2014.