Supplemental Materials

The following document provides supporting information for the manuscript: "High-Dimensional Survival Analysis: Methods and Applications."

Supplemental Material: Annu. Rev. Stat. Appl. 2023. 10.
https://doi.org/10.1146/annurev-statistics-032921-022127
High-Dimensional Survival Analysis: Methods and Applications
Salerno and Li

A. Simulation Results for the Proposed Deep Neural Network for Semi-Competing Risks

We conduct simulations to illustrate the feasibility of the proposed Deep Neural Network for Semi-Competing Risks (DNN-SCR) model.

We simulate the observed data, $\mathcal{D} = \{(Y_{i1},\ \delta_{i1},\ Y_{i2},\ \delta_{i2},\ x_i);\ i = 1, \ldots, n\}$ in a fully factorial design by varying the sample size, frailty variance, log-risk function, and censoring rates, a total of 24 settings (Table A1).

Specifically, we simulate the shared frailty, γ_i, from $\Gamma(1/\theta, 1/\theta)$ with $\mathrm{Var}(\gamma_i) = \theta$ taking values of 0.5 and 2.0, corresponding to varying degrees of dependence between event times.

The baseline hazard functions, λ_{01}, λ_{02}, and λ_{03}, are taken to be Weibull distributions with the same shape and scale parameters equal to 1. We simulate two standard Normal random covariates, $X_1, X_2 \sim N(0, 1)$, which are taken to be predictive of the morbidity and mortality hazards through both a linear and non-linear log-risk function. Specifically, we first examine a linear log-risk function:

$$h_g(\boldsymbol{X}_i) = x_i^\top \boldsymbol{\beta}_g$$

with $\boldsymbol{\beta}_g = [1, 1]^\top$ for $g = 1, 2, 3$, so that the requirements for the classical model is satisfied, facilitating a fair comparison with existing methods. Then, we consider non-linear functions

$$h_g(\boldsymbol{X}_i) = \log(|\boldsymbol{X}_i|^\top \boldsymbol{\beta}_g + 1)$$

with $\boldsymbol{\beta}_g = [1, 1]^\top$ for $g = 1, 2, 3$.

Censoring times are generated from an exponential distributions to yield approximate censoring rates of 0%, 25% and 50%. We vary the number of patients as $1,000$ and $10,000$. For each parameter configuration, a total of 50 datasets are independently generated.

We compared our method to a classical MLE approach, which directly maximizes the log-likelihood function under the assumption of a semi-Markov model with Weibull baseline hazard functions. This approach assumes that the risk functions are linear combinations of the generated covariates. We compare the predictive performance of our method to the MLE approach using the average mean integrated squared error for estimating the log-risk surfaces:

$$\frac{1}{n}\sum_{i=1}^{n}[h_g(\boldsymbol{X}_i) - \hat{h}_g(\boldsymbol{X}_i)]^2; g = 1, 2, 3,$$

for each state transition hazard, separately.

As shown in Table A1, both methods accurately recover the log-risk surfaces for each state transition when the true underlying function is linear. However, in the non-linear settings, our deep neural network approach has a much lower mean integrated squared error, on average, compared to the classical MLE method, indicating a good performance of the proposed method.

Table A1: Average (SD) Mean Integrated Squared Errors for Simulated Log-Risk Surfaces for Each State Transition Hazard (i.e., $1/n \sum_{i=1}^{n}[h_g(x_i) - \hat{h}_g(x_i)]^2; g = 1, 2, 3$)

Simulation Settings					Maximum Likelihood Estimation			Deep Neural Network		
Setting	Sample Size	Frailty Variance (θ)	Log-Risk Function	Censoring Rate	h_1	h_2	h_3	h_1	h_2	h_3
1	1,000	0.50	Linear	0%	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.07 (0.05)	0.08 (0.08)	0.08 (0.05)
2	10,000	0.50	Linear	0%	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.08 (0.07)	0.08 (0.05)	0.08 (0.07)
3	1,000	2.00	Linear	0%	0.02 (0.01)	0.01 (0.01)	0.01 (0.01)	0.12 (0.07)	0.13 (0.07)	0.13 (0.09)
4	10,000	2.00	Linear	0%	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.11 (0.06)	0.11 (0.08)	0.13 (0.10)
5	1,000	0.50	Non-Linear	0%	1.80 (0.33)	1.82 (0.39)	1.85 (0.34)	0.09 (0.05)	0.09 (0.04)	0.08 (0.04)
6	10,000	0.50	Non-Linear	0%	1.80 (0.13)	1.77 (0.13)	1.78 (0.11)	0.07 (0.03)	0.08 (0.03)	0.08 (0.05)
7	1,000	2.00	Non-Linear	0%	1.92 (0.53)	1.85 (0.54)	1.96 (0.53)	0.15 (0.05)	0.15 (0.06)	0.14 (0.05)
8	10,000	2.00	Non-Linear	0%	1.82 (0.17)	1.81 (0.18)	1.83 (0.18)	0.14 (0.04)	0.12 (0.03)	0.13 (0.06)
9	1,000	0.50	Linear	25%	0.01 (0.02)	0.02 (0.01)	0.02 (0.02)	0.10 (0.06)	0.10 (0.07)	0.13 (0.12)
10	10,000	0.50	Linear	25%	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.12 (0.10)	0.12 (0.09)	0.12 (0.10)
11	1,000	2.00	Linear	25%	0.03 (0.02)	0.02 (0.02)	0.04 (0.03)	0.15 (0.10)	0.13 (0.09)	0.18 (0.12)
12	10,000	2.00	Linear	25%	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.14 (0.10)	0.12 (0.08)	0.14 (0.10)
13	1,000	0.50	Non-Linear	25%	1.96 (0.44)	2.01 (0.54)	2.24 (0.66)	0.10 (0.07)	0.10 (0.06)	0.10 (0.08)
14	10,000	0.50	Non-Linear	25%	1.95 (0.15)	1.91 (0.16)	2.16 (0.20)	0.07 (0.04)	0.09 (0.08)	0.09 (0.07)
15	1,000	2.00	Non-Linear	25%	2.06 (0.62)	1.92 (0.72)	2.25 (0.79)	0.15 (0.08)	0.15 (0.08)	0.13 (0.06)
16	10,000	2.00	Non-Linear	25%	1.88 (0.21)	1.88 (0.21)	2.04 (0.28)	0.10 (0.05)	0.11 (0.06)	0.11 (0.05)
17	1,000	0.50	Linear	50%	0.01 (0.02)	0.02 (0.02)	0.04 (0.03)	0.10 (0.07)	0.10 (0.06)	0.20 (0.15)
18	10,000	0.50	Linear	50%	0.00 (0.00)	0.00 (0.00)	0.00 (0.01)	0.10 (0.07)	0.11 (0.08)	0.17 (0.16)
19	1,000	2.00	Linear	50%	0.03 (0.03)	0.03 (0.02)	0.05 (0.05)	0.22 (0.13)	0.17 (0.13)	0.24 (0.17)
20	10,000	2.00	Linear	50%	0.00 (0.00)	0.00 (0.00)	0.01 (0.00)	0.14 (0.09)	0.14 (0.10)	0.16 (0.14)
21	1,000	0.50	Non-Linear	50%	2.06 (0.50)	2.20 (0.72)	2.61 (1.00)	0.09 (0.06)	0.13 (0.13)	0.18 (0.14)
22	10,000	0.50	Non-Linear	50%	2.03 (0.21)	2.00 (0.18)	2.36 (0.25)	0.06 (0.03)	0.09 (0.08)	0.10 (0.09)
23	1,000	2.00	Non-Linear	50%	2.16 (0.76)	2.00 (0.72)	2.41 (0.91)	0.18 (0.10)	0.18 (0.09)	0.16 (0.10)
24	10,000	2.00	Non-Linear	50%	1.92 (0.25)	1.95 (0.23)	2.22 (0.38)	0.10 (0.05)	0.11 (0.06)	0.15 (0.13)

B. Selected Methods, Citations, and Available Software

Table B1: Selected methods covered in this review, citations, and available software

Method	Citation	Available Software
Classical Survival Analysis		
Cox Proportional Hazards Model	Cox (1972)	https://cran.r-project.org/web/packages/survival/index.html
Accelerated Failure Time Models	Buckley & James (1979)	https://cran.r-project.org/web/packages/survival/index.html
Censored Quantile Regression	Portnoy (2003)	https://cran.r-project.org/web/packages/quantreg/index.html
Sub-Distribution Hazard Model	Fine & Gray (1999)	https://cran.r-project.org/web/packages/cmprsk/index.html
Illness-Death Model	Haneuse & Lee (2016)	https://cran.r-project.org/web/packages/SemiCompRisks/index.html
Regularized Cox Models		
Ridge	Verweij & Van Houwelingen (1994)	https://cran.r-project.org/web/packages/glmnet/glmnet.pdf
LASSO	Tibshirani (1997)	https://cran.r-project.org/web/packages/glmnet/glmnet.pdf
Elastic Net	Wu (2012)	https://cran.r-project.org/web/packages/glmnet/glmnet.pdf
Adaptive LASSO	Zhang & Lu (2007)	https://cran.r-project.org/web/packages/glmnet/glmnet.pdf
SCAD	Fan & Li (2002)	https://cran.r-project.org/web/packages/ncvreg/index.html
Group LASSO	Kim et al. (2012)	https://cran.r-project.org/web/packages/grpreg/index.html
Fused LASSO	Tibshirani et al. (2005)	https://cran.r-project.org/web/packages/penalized/penalized.pdf
Graphical LASSO	Friedman et al. (2008)	-
Dantzig Selector		
AFT Models	Li et al. (2014)	http://www-personal.umich.edu/ yili/adsfxns.R
Cox Model	Antoniadis et al. (2010)	-
Ultra High-Dimensional Feature Screening		
Sure Independence Screening	Fan & Lv (2008)	https://cran.r-project.org/web/packages/SIS/index.html
	Fan & Song (2010)	https://cran.r-project.org/web/packages/SIS/index.html
Principled Sure Independent Screening	Zhao & Li (2012)	http://faculty.washington.edu/acook/software.html
Buckley-James Assisted Sure Screening	Liu et al. (2020)	-
Conditional Screening	Kang et al. (2017)	https://github.com/younghhk/software/blob/master/CS.R
Forward Regression	Hong et al. (2019)	-
Inferential Methods		
Selection-Assisted Partial Regression and Smoothing	Fei & Li (2021)	https://github.com/feizhe/SPARES
Fused High-Dimensional Censored Quantile Regression	Fei et al. (2021)	https://github.com/feizhe/HDCQR_Paper
Support Vector Machines		
Rank-Based Approach	Van Belle et al. (2007)	https://cran.r-project.org/web/packages/survivalsvm/index.html
Regression Approach	Shivaswamy et al. (2007)	https://cran.r-project.org/web/packages/survivalsvm/index.html
Hybrid Approach	Van Belle et al. (2011)	https://cran.r-project.org/web/packages/survivalsvm/index.html
	Pölsterl et al. (2015)	https://cran.r-project.org/web/packages/survivalsvm/index.html
Tree-Based Methods		
Log-Rank Based	Ciampi et al. (1986)	https://cran.r-project.org/web/packages/rpart/index.html
Likelihood-Based	Ciampi et al. (1987)	https://cran.r-project.org/web/packages/rpart/index.html
Ensemble Learners		
Bootstrap Aggregation	Hothorn et al. (2004)	https://cran.r-project.org/web/packages/ipred/index.html
Gradient Boosting	Hothorn et al. (2006)	https://cran.r-project.org/web/packages/gbm/index.html
Random Survival Forests	Ishwaran et al. (2008)	https://cran.r-project.org/web/packages/randomForestSRC/index.html
	Ishwaran et al. (2011)	https://cran.r-project.org/web/packages/randomForestSRC/index.html
	Ishwaran & Lu (2019)	https://cran.r-project.org/web/packages/randomForestSRC/index.html
Censoring Unbiased Regression Trees	Steingrimsson et al. (2019)	https://cran.r-project.org/web/packages/randomForest/index.html
Deep Learning		
DeepSurv	Katzman et al. (2018)	https://cran.r-project.org/web/packages/survivalmodels/index.html
DNNSurv	Zhao & Feng (2020)	https://cran.r-project.org/web/packages/survivalmodels/index.html
Competing Risks		
DeepHit	Lee et al. (2018)	https://cran.r-project.org/web/packages/survivalmodels/index.html
Dynamic DeepHit	Lee et al. (2019)	https://github.com/chl8856/Dynamic-DeepHit
DeepCompete	Aastha & Liu (2020)	-
Hierarchical Multi-Event	Tjandra et al. (2021)	-
Semi-Competing Risks		
Penalized Estimation	Reeder et al. (2022)	-
Deep Learning	*Proposed*	-

References

Aastha PH, Liu Y. 2020. Deepcompete: A deep learning approach to competing risks in continuous time domain, In *AMIA Annual Symposium Proceedings*, vol. 2020, p. 177, American Medical Informatics Association

Antoniadis A, Fryzlewicz P, Letué F. 2010. The Dantzig selector in Cox's proportional hazards model. *Scandinavian Journal of Statistics* 37(4):531–552

Buckley J, James I. 1979. Linear regression with censored data. *Biometrika* 66(3):429–436

Ciampi A, Chang CH, Hogg S, McKinney S. 1987. Recursive partition: A versatile method for exploratory-data analysis in biostatistics. In *Biostatistics*. Springer, 23–50

Ciampi A, Thiffault J, Nakache JP, Asselain B. 1986. Stratification by stepwise regression, correspondence analysis and recursive partition: a comparison of three methods of analysis for survival data with covariates. *Computational Statistics & Data Analysis* 4(3):185–204

Cox DR. 1972. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B* 34(2):187–220

Fan J, Li R. 2002. Variable selection for Cox's proportional hazards model and frailty model. *The Annals of Statistics* 30(1):74–99

Fan J, Lv J. 2008. Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society, Series B* 70(5):849–911

Fan J, Song R. 2010. Sure independence screening in generalized linear models and NP-dimensionality. *The Annals of Statistics* 38(6):3567–3604

Fei Z, Li Y. 2021. Estimation and inference for high dimensional generalized linear models: A splitting and smoothing approach. *Journal of Machine Learning Research* 22(58):1–32

Fei Z, Zheng Q, Hong HG, Li Y. 2021. Inference for high-dimensional censored quantile regression. *Journal of the American Statistical Association* :1–15

Fine JP, Gray RJ. 1999. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association* 94(446):496–509

Friedman J, Hastie T, Tibshirani R. 2008. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9(3):432–441

Haneuse S, Lee KH. 2016. Semi-competing risks data analysis: accounting for death as a competing risk when the outcome of interest is nonterminal. *Circulation: Cardiovascular Quality and Outcomes* 9(3):322–331

Hong HG, Christiani DC, Li Y. 2019. Quantile regression for survival data in modern cancer research: Expanding statistical tools for precision medicine. *Precision Clinical Medicine* 2(2):90–99

Hothorn T, Bühlmann P, Dudoit S, Molinaro A, Van Der Laan MJ. 2006. Survival ensembles. *Biostatistics* 7(3):355–373

Hothorn T, Lausen B, Benner A, Radespiel-Tröger M. 2004. Bagging survival trees. *Statistics in Medicine* 23(1):77–91

Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. 2008. Random survival forests. *The Annals of Applied Statistics* 2(3):841–860

Ishwaran H, Kogalur UB, Chen X, Minn AJ. 2011. Random survival forests for high-dimensional data. *Statistical Analysis and Data Mining: The ASA Data Science Journal* 4(1):115–132

Ishwaran H, Lu M. 2019. Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival. *Statistics in Medicine* 38(4):558–582

Kang J, Hong HG, Li Y. 2017. Partition-based ultrahigh-dimensional variable screening. *Biometrika* 104(4):785–800

Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. 2018. Deepsurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Medical Research Methodology* 18(1):1–12

Kim J, Sohn I, Jung SH, Kim S, Park C. 2012. Analysis of survival data with group lasso. *Communications in Statistics-Simulation and Computation* 41(9):1593–1605

Lee C, Yoon J, Van Der Schaar M. 2019. Dynamic-deephit: A deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. *IEEE Transactions on Biomedical Engineering* 67(1):122–133

Lee C, Zame WR, Yoon J, van der Schaar M. 2018. Deephit: A deep learning approach to survival analysis with competing risks, In *Thirty-second AAAI conference on artificial intelligence*

Li Y, Dicker L, Zhao SD. 2014. The Dantzig selector for censored linear regression models. *Statistica Sinica* 24(1):251

Liu Y, Chen X, Li G. 2020. A new joint screening method for right-censored time-to-event data with ultra-high dimensional covariates. *Statistical Methods in Medical Research* 29(6):1499–1513

Pölsterl S, Navab N, Katouzian A. 2015. Fast training of support vector machines for survival analysis, In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 243–259, Springer

Portnoy S. 2003. Censored regression quantiles. *Journal of the American Statistical Association* 98(464):1001–1012

Reeder HT, Lu J, Haneuse S. 2022. Penalized estimation of frailty-based illness-death models for semi-competing risks. *arXiv preprint arXiv:2202.00618*

Shivaswamy PK, Chu W, Jansche M. 2007. A support vector approach to censored targets, In *Seventh IEEE international conference on data mining (ICDM 2007)*, pp. 655–660, IEEE

Steingrimsson JA, Diao L, Strawderman RL. 2019. Censoring unbiased regression trees and ensembles. *Journal of the American Statistical Association* 114(525):370–383

Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K. 2005. Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67(1):91–108

Tibshirani RJ. 1997. The lasso method for variable selection in the Cox model. *Statistics in Medicine* 16:385–395

Tjandra D, He Y, Wiens J. 2021. A hierarchical approach to multi-event survival analysis, In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, pp. 591–599

Van Belle V, Pelckmans K, Suykens J, Van Huffel S. 2007. Support vector machines for survival analysis, In *Proceedings of the Third International Conference on Computational Intelligence in Medicine and Healthcare (CIMED2007)*, pp. 1–8

Van Belle V, Pelckmans K, Van Huffel S, Suykens JA. 2011. Support vector methods for survival analysis: a comparison between ranking and regression approaches. *Artificial Intelligence in Medicine* 53(2):107–118

Verweij PJ, Van Houwelingen HC. 1994. Penalized likelihood in Cox regression. *Statistics in Medicine* 13(23-24):2427–2436

Wu Y. 2012. Elastic net for Cox's proportional hazards model with a solution path algorithm. *Statistica Sinica* 22:27

Zhang HH, Lu W. 2007. Adaptive lasso for Cox's proportional hazards model. *Biometrika* 94(3):691–703

Zhao L, Feng D. 2020. Deep neural networks for survival analysis using pseudo values. *IEEE Journal of Biomedical and Health Informatics* 24(11):3308–3314

Zhao SD, Li Y. 2012. Principled sure independence screening for Cox models with ultra-high-dimensional covariates. *Journal of Multivariate Analysis* 105(1):397–411