Recent trends in Survival analysis using Deep Learning in Medical Science: Current Perspective and Future Direction

Palak Kaushal,
Research Scholar,
Punjab Engineering College
(Deemed to be University),
Chandigarh, 160012, India

Dr. Shailendra Singh,
Professor,
Punjab Engineering College
(Deemed to be University),
Chandigarh, 160012, India

Dr. Dharam Vir,
Assistant Professor,
PGIMER,
Chandigarh, 160012, India

Email: palak.phd19cse@pec.edu.in

Abstract

Survival Analysis manifests a profound effect on healthcare services. It is the analysis of a dichotomous event, i.e., occurrence or non-occurrence of the event over a period of time. While the traditional survival analysis approaches work with some pre-assumptions and do not include time-sequential data, recently developed machine learning based systems can capture non-linearity and involve time-dependent features from multiple visits. The article highlights the ongoing research demonstrating that the machine learning-based survival analysis approaches have consistently improved compared to the state-of-art statistical survival analysis approaches.

KEYWORDS: Survival Analysis; Deep Learning; Intelligent Healthcare systems, Neural Networks

DOI Number: 10.14704/nq.2022.20.9.NQ44384 **Neuro Quantology 2022; 20(9):3330-3336**

I. INTRODUCTION

Survival Analysis is a well-developed area of statistics that signifies the probability of the event occurring beyond a certain period by estimating the survival function. It is different from the regression techniques as it includes both instances- those with the observed event of interest, i.e., uncensored events, and those with the event of interest not observed, i.e., censored events. Thus, survival analysis utilizes more information than standard models, eventually alleviating data insufficiency problems and improving performance. Survival analysis, also known as time-to-event analysis, is extensively used in medical sciences and other fields.

The event of interest can be an event of death in healthcare[1], the user clicks in customer behavior analysis[2], or device failure in engineering[3]. In this paper, we have tried to provide insights into the field of survival analysis and the current developments in deep learning based survival analysis without digging deep into the mathematical concepts.

II. CATEGORIES OF SURVIVAL ANALYSIS

The two main streams of survival analysis are traditional statistical methods and machine-learning-based methods. Figure1 shows different categories of survival analysis with some examples.

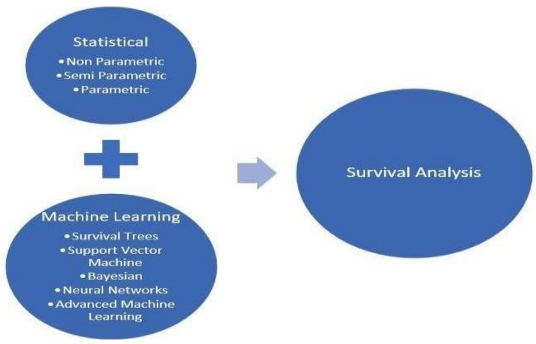

Figure-1 Different categories of Survival Analysis

A. Statistical Approaches

The statistical approaches for survival analysis can be broadly categorized into: non-parametric, semi- parametric, and parametric. The non-parametric models [4], [5] do not involve any presumptions and thus provide a rough general explanation of the probability of survival. Mostly, these describe survival probability as a function of time and give an average view of the population. But these models do not consider the individual differences in the group. The parametric model [6], on the other hand, relies heavily on the distribution of the event of interest. These models are more efficient and accurate when there is a particular distribution

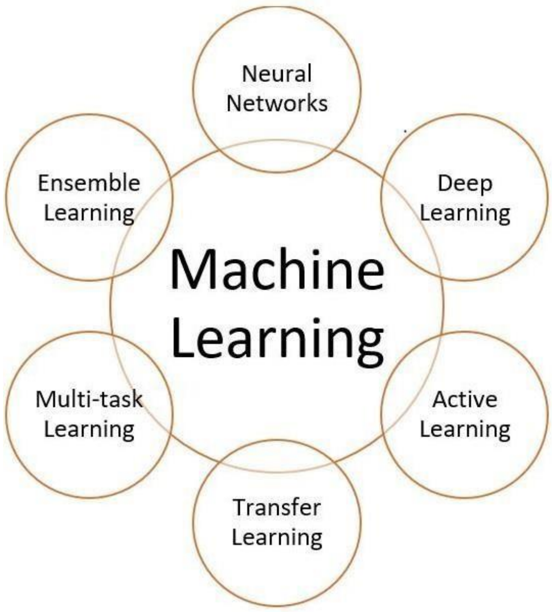

Figure-2 Different Approaches related to Machine Learning

In the event of interest, thus making them impractical in real-world problems. Eventually, the semi-parametric models[7] improve the shortcomings of both categories to a certain extent. The semi-parametric models provide personalized prediction taking into account the individual differences without considering the underlying distribution. This makes semi-parametric, more specifically the Cox Proportional Hazard model, the most frequently used statistical survival analysis model. Over time, many variants of Cox models have been studied and developed for different types of scenarios. All of the statistical approaches are primarily based on either the simple counting methodor pre-assumptions about the distribution for survival rate function and, thus, do not generalize well for real-world scenarios.

Cox proportional hazard model using Machine Learning (ML) techniques, for example, Cox-nnet [14] (an extension of Cox PH on specific datasets); DeepSurv [12]. Random Survival Forest(RSF) [15] is a fully non-parametrized continuous-time approach that uses the log-rank test to compute decision trees for right-censored data. The approach estimates the cumulative hazard rate using the Nelson-Aalen estimator. Schmid et al.[16] proposed a non-parametric RSF-based approach that computes the ensemble of discrete-time survival trees using Hellinger's distance for splitting. It functions on transformed tree-based binary classification problems of the original time-to-event problem, thus including time-varying covariates.

Another approach is developing fully-parameterized survival models that sequentially predict a distribution over time to the next event. One way is using Recurrent Neural Networks (RNN) such as RNN-SURV[11], Weibull Time- to-Event RNN, and others. The application of neural networks to time-to-event scenarios initially started with Farragi and Simon [8]. This was the first application of neural networks for survival analysis by parametrizing the Cox regression

B. Machine Learning For Survival Analysis

Over the past few years, machine learning has achieved outstanding results when combined with survival analysis with the development of technology and computational capacities. An ample amount of research is done in combining survival analysis and machine learning and showcased mixed performance[8], [9]. One reason could be that machine learning was not as advanced today.

The recent attempts of machine learning based survival analysis have been fruitful and demonstrate state-of- art performance for risk prediction[10], [11], [12], [13]. These developments have benefitted from neural network frameworks such as Keras, Theano, Tensorflow, and Pytorch, making neural networks quite simple. Figure2 lists some of the approaches related to machine learning used to solve survival analysis problems.

The application of machine learning for survival analysis can be seen in different groups. One is the extension of the model with only a single hidden layer but could not improve over the regular Cox regression model. Researchers have also attempted to include different approaches (as shown in Figure 2 like multi-task learning[17] and active learning with machine learning for survival analysis. MTLSA[17] is a multi-task learning survival approach that converts the original problem into a sequence of binary classification subproblems and calculates event probability at different times using multi-task learning. Bender et al.[18] proposed a general machine learning framework that reexpresses a survival task into a Poisson regression task, thus showcasing the application of any machine learning algorithm which can optimize Poisson log-likelihood to survival analysis task.

In contrast to machine learning, deep learning extracts the features automatically, thus improving resource utilization and time efficiency, and therefore is being extensively used nowadays by researchers. Ranganath et. al. [19] proposed a hierarchical generative approach that uses deep learning for survival analysis, specifically, Deep Exponential Family (DEF) to capture dependencies from covariates. They validated the model by stratifying patients based on the risk of developing coronary heart disease (CHD) and presented better performance than the clinically approved Framingham CHD risk score. DeepSurv[12] is a Cox proportional hazard deep multi-layer feed-forward perceptron network that predicts the event's occurrence by estimating each individual's effect on their hazard rates to the network's parametrized weights. The authors used simulated and real-world datasets for testing and validation and, lastly, a real- world medical dataset to illustrate the personalized treatment recommendations by the model. These methods still depend on the assumptions about the distributions for survival and thus suffer from the problem of generalization.

Another approach for deep-learning-based survival analysis is DeepHit [13]. The network architecture comprises of one sub-network and a group of cause- specific sub-networks. It can handle both single-underlying risk situations and multiple or competing risk situations and does not lie on assumptions about the underlying stochastic process, and handles competing risks. It also allows for the time-varying relationship between the risk(s) and covariates. The approach first parametrizes the Probability Mass Function (PMF) of the survival distribution and then combines the loglikelihood and ranking loss. It uses a time-dependent c- Index [20] for performance evaluation. Deephit has high discriminative performance but poor calibration survival estimates [21]. Dynamic- DeepHit [22] is an extension of DeepHit which handles dynamic survival analysis and works on longitudinal data. Table I lists the comparison of some most widely used deep learning models forsurvival analysis.

Luck et al.[25] proposed a deep learning approach for survival analysis using the principles of multi- task learning and showcased that the proposed model outperformed other state-of art survival analysis models in terms of c-index and survival time prediction quality. RNN-SURV[11] is a model developed using Recurrent Neural Networks which considers the sequential nature of the survival analysis problem. It outputs the survival probability at each time interval and the corresponding risk score of each individual by transforming it into a series of binary classification problems. Nezhad et al.[26] proposed a deep active learning based survival analysis framework called Deep Active Survival Analysis (DASA) which initially uses deep learning for feature representation by producing robust features from high-dimensional data and then active learning for predicting the survival. CoxTime[21] is an extension of DeepSurv[12] in which the proportionality constraint no longer exists because of an alternative loss function that works well for both proportional and non-proportional cases.

DRSA (DEEP RECURRENT SURVIVAL ANALYSIS) [23] is a deep-learning-based survival analysis model which predicts the conditional probability of occurrence of the eventat each time, assuming its prior non-occurrence and combines both the probability density function and the cumulative distribution function of the event over time using the probability chain rule, thus forecasting survival rate at each time. The model uses RNN and appre -nds the sequential dependency pattern among neighboring time slices and therefore is efficient in backpropagating the gradient. Chi et. al. [27] proposed a semi-supervised multi-task learning approach for survival analysis which works by translating the survival analysis task into a multi-task learning problem and models survival time distribution without any assumptions.

III. DISCUSSION

Survival analysis observes the advancement towards the event of interest and plots the survival curve, indicating the probability of survival until a particular follow-up time. Some patients might not be available for follow-up after some time in medical science and are stated as censored patients. Although there is no information about what happened to them, it is sure that they were free from events until the censoring. Clinical data typically includes a number of variables, is sparse, and time-dependent, and thus applying the traditional survival analysis approach does not perform well. The labeled data is minimal, time-consuming, and challenging to collect. Especially in medical sciences, almost 50% of the data is censored, and excluding that leads to biases in thepredictions. In survival analysis, including the censored data provides the advantage of considering more samples than standard regression techniques, which helps tackle the insufficiency of data to some extent.

Most of the survival analysis models assume that the relation between the covariates and the risk is time- invariant. Also, the models do not work in the presence of competing risks or work under the assumption that the risks are independent, which is not true, especially in medical science.

DeepSurv[12] is the first Cox-PH-based neural network model which extracts non-linear representations from input data, but the proportionality assumptions confine its performance. Later, many different neural network-based approaches free from proportionality assumptions have been proposed. Most deep learning based survival models which handle competing risks have similar architecture.

The first few layers are the shared layers that learn the standard features. Then the output of the shared layer combined with original data is provided to separate arms, with each signifying a different specific risk. One of the problems with such architecture design is that as the number of competing risks increases, the number of layers also increases, thus limiting the model's scalability. Convolutional neural networks[25] and long short-term memory were used to substitute the multi-layer perceptron of DeepSurv for images and sequential data and accomplished improved performance over the traditional methods. The survey's authors discussed the application of deep learning for survival prediction for disease-specific predictive tasks in histopathology[28]. In particular, despite the fact that biomarkers and additional risk factors are repeatedly measured at different times, survival analysis is primarily done on the last available measurement. This signifies an excessive problem because the progress of biomarkers and risk factors are informative and impact the onset of disease. Therefore, it is crucial to include longitudinal measurements of covariates into the model to better understand disease progression. Rather than discarding valuable records, this allows us to make improved risk assessments of the clinical events. Thus, by including newer network architecture, better optimization techniques, and large and high-dimensional datasets, we can construct survival analysis models that will be flexible to be applied to different areas, more specifically in testing and comparing the impact of treatment or new technology. With the increasing number of publications in survival prediction in medical, it is essential to understand the recent developments, different methodologies used, and different application scenarios in which the models can be applied.

The uniqueness in the modeling approach and unfamiliarity of clinicians with the model's working and output leads to the model's unacceptability.

TABLE I COMPARISON OF DEEP LEARNING MODELS FOR SURVIVAL ANALYSIS

Model	Single Risk/ Competing Risk	TimeVarying Effects	Assumptions	Discrete Vs Continuous	Function	Github Link
Deephit[13]	Competing Risk	Yes	No	Discrete	Probability Mass Function (PMF)	https://Github.Com/Chl8856/Deephit
Dynamic Deephit[22]	Competing Risk	Yes	No	Continuous	Cumulative Incidence Function	https://Github.Com/Chl8856/DynamicDeephit
Deep Survival Machine (DSM)[19]	Competing Risk	Yes	No	Discrete	Survival Distribution	autonlab.Github.Io/Deepsurvivalmachines
Deep Recurrent Survival Analysis (DRSA)[23]	Single	Yes	No	Continuous	Survival Rate Function	https://Github.Com/Rk2900/DRSA
DeepSurv[12]	Single	No	Proportionality Hazard Assumption	Discrete	Log Hazard Ratio	https://Github.Com/Jaredleekatzman/Deepsurv
DeepCompete [24]	CompetingRisk	Yes	No	Continuous	Cumulative Hazard Function	-

IV. CONCLUSION

Thus, a tool capable of precisely predicting the condition-specific risk and supporting clinicians in making the right decisions can impact the patient's quality and longevity of life. Machine Learning based survival models have higher predictive ability than statistical models because of their capability to model non-linear relationships, high dimensional data, and complex calculations. Nevertheless, different algorithms or models developed have different strengths and weaknesses and thus can be applied for specific tasks. Also, in this era of big data, where medical data is readily available, using techniques such as deep learning will help robust data processing. More focus needs to be on developing models that predict the survival risk and infer the best treatment option to reduce the future risk at the individual level.

Thus, the application of machine learning can aid the clinical data analysis in unique ways that can help clinicians in identifying and predicting risk at an individual personalized level. Although machine learning is quite popular in commercial industries, it is still lacking in the medical field. We strongly believe that machine learning can highly revolutionize the medical field and is capable of providing better personalized and precise medicine.

REFERENCES

[1] Y. Li, L. Wang, J. Wang, J. Ye, and C. K. Reddy, “Transfer learning for survival analysis via efficient l2, 1-norm regularized cox regression,” in *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE, Conference Proceedings, pp. 231–240.

[2] N. Barbieri, F. Silvestri, and M. Lalmas, “Improving post-click user engagement on native ads via survival analysis,” in *Proceedings of the 25th International Conference on World Wide Web*, Conference Proceedings, pp. 761– 770.

[3] M. Modarres, M. P. Kaminskiy, and V. Krivtsov, *Reliability engineering and risk analysis: a practical guide*. CRC press, 2016.

[4] P. K. Andersen, O. Borgan, R. D. Gill, and N. Keiding, *Statistical models based on counting processes*. Springer Science Business Media, 2012.

[5] E. L. Kaplan and P. Meier, “Nonparametric estimation from incomplete observations,” *J Journal of the American statistical association*, vol. 53, no. 282, pp. 457–481, 1958.

[6] E. T. Lee and J. Wang, *Statistical methods for survival data analysis*. John Wiley Sons, 2003, vol. 476.

[7] D. R. Cox, “Regression models and life-tables,” *J Journal of the Royal Statistical Society: Series B*, vol. 34, no. 2, pp. 187–202, 1972.

[8] D. Faraggi and R. Simon, “A neural network model for survival data,” *Stat Med*, vol. 14, no. 1, pp. 73–82, 1995. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/7701159

[9] D. J. Sargent, “Comparison of artificial neuralnetworks with other statistical approaches:results from medical data sets,” *Cancer*, vol. 91,no. 8 Suppl, pp. 1636–42, 2001. [Online].Available: https://www.ncbi.nlm.nih.gov/pubmed/11309761

[10] T. Ching, X. Zhu, and L. X. Garmire, “Cox-nnet: An artificial neural network method for prognosis prediction of high-throughput omics data,” *PLoS Comput Biol*, vol. 14, no. 4, p. e1006076, 2018. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/29634719

[11] E. Giunchiglia, A. Nemchenko, and M. v. d. Schaar, “Rnn-surv: A deep recurrent model for survival analysis,” in *International Conference on Artificial Neural Networks*. Springer, Conference Proceedings, pp. 23–32.

[12] J. L. Katzman, U. Shaham, A. Cloninger, J. Bates, T. Jiang, and Y. Kluger, “Deepsurv: personalized treatment recommender system using a cox proportional hazards deep neural network,” *J BMC medical research methodology*, vol. 18, no. 1, pp. 1–12, 2018.

[13] C. Lee, W. Zame, J. Yoon, and M. van der Schaar, “Deephit: A deep learning approach to survival analysis withcompeting risks,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/vie w/11842

[14] T. Ching, X. Zhu, and L. X. Garmire, “Cox-nnet: An artificial neural network method for prognosis prediction of high-throughput omics data,” *PLoS Comput Biol*, vol. 14, no. 4, p. e1006076, 2018. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/29634719

[15] H. Ishwaran, U. B. Kogalur, E. H. Blackstone, and M. S. Lauer, “Random survival forests,” *J The annals of applied statistics*, vol. 2, no. 3, pp.841–860, 2008.

[16] M. Schmid, T. Welchowski, M. N. Wright,and M. Berger, “Discretetime survival forests with hellinger distance decision trees,” *Data Mining and Knowledge Discovery*, vol. 34, no. 3, pp. 812–832, 2020. [Online]. Available: https://doi.org/10.1007/s10618-020-00682-z

[17] Y. J. R. C. Li Y, Wang J, “A multi-task learning formulation for survival analysis,” pp. 1715–1724, 2016. [Online]. Available: https://doi.org/10.1145/2939672.2939857

[18] A. Bender, D. Rugamer, F. Scheipl, and B. Bischl, “A general machine¨ learning

framework for survival analysis," ser. Machine Learningand Knowledge Discovery in Databases. Springer International Publishing, Conference Proceedings, pp. 158–173.

[19] E. N. B. D. Ranganath R, Perotte A, "Deep survival analysis," pp. 101–114, 2016. [Online].Available: https://proceedings.mlr.press/v56/Ranganath16. html

[20] L. Antolini, P. Boracchi, and E. Biganzoli, "Atimedependent discrimination index for survivaldata," *Stat Med*, vol. 24, no. 24, pp. 3927–44,2005. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16320281

[21] B. Havard Kvamme and I. Scheel, "Time-to-event prediction with neural˚ networks and cox regression," *Journal of Machine Learning Research*, vol. 129, 2019.

[22] C. Lee, J. Yoon, and M. V. Schaar, "Dynamic-deephit: A deep learning approach for dynamicsurvival analysis with competing risks based onlongitudinal data," *IEEE Trans Biomed Eng*, vol.67, no. 1, pp.122–133, 2020. [Online].Available: https://www.ncbi.nlm.nih.gov/pubmed/30951460

[23] K. Ren, J. Qin, L. Zheng, Z. Yang, W. Zhang, L. Qiu, and Y. Yu, "Deep recurrent survival analysis," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, Conference Proceedings, pp. 4798–4805.

[24] L. Y. Aastha, Huang P, "Deepcompete : A deep learning approach to competing risks in continuous time domain," *AMIA Annu Symp Proc*, pp. 177–86, 2021.

[25] M. Luck, T. Sylvain, H. Cardinal, A. Lodi, and Y. Bengio, "Deep learning for patient-specific kidney graft survival analysis," 2017.

[26] M. Z. Nezhad, N. Sadati, K. Yang, and D. Zhu, "A deep active survival analysis approach forprecision treatment recommendations: Application of prostate cancer," *Expert Systemswith Applications*, vol. 115, pp. 16–26, 2019.[Online]. Available: https://www.sciencedirect.com/science/article/p ii/S0957417418304949

[27] S. Chi, Y. Tian, F. Wang, Y.Wang, M. Chen, and J. Li, "Deep semisupervised multitask learning model and its interpretability for survival analysis," *IEEE J Biomed Health Inform*, vol. 25,no. 8, pp. 3185–3196, 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/33687852

[28] C. L. Srinidhi, O. Ciga, and A. L. Martel, "Deep neural network models for computational histopathology: A survey," *Med Image Anal*, vol. 67, p. 101813, 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/33049577

[29] C. Lee, J. Yoon, and M. V. Schaar, "Dynamic-deephit: A deep learning approach for dynamicsurvival analysis with competing risks based onlongitudinal data," *IEEE Trans Biomed Eng*, vol.67, no. 1, pp.122–133, 2020. [Online].Available: https://www.ncbi.nlm.nih.gov/pubmed/30951460

ACKNOWLEDGMENT

The authors would like to thank all the related authors and their respective institutions for providingtheir support for the research.