Reliability Models Evolution: from Survival Regression to Deep Survival

Vasiliy Krivtsov, PhD
Ford Motor Company

1. INTRODUCTION

Reliability modeling is a critical aspect of many fields, including engineering, medicine, economics, and more. Over time, the approach to reliability modeling has evolved. The shift has been from using *basic reliability models* to *survival regression models* to *deep survival* (*machine learning*) *models*. In this paper, we will discuss this methodological evolution.

1.1 RELIABILITY MODELS EVOLUTION

A. The simplest (classical) reliability models involve either a *nonparametric* (e.g., Kaplan-Meier) or *parametric* (e.g., Weibull or Lognormal lifetime distributions) approach to model the underlying random variables, such as failure times. Model parameters are estimated via respective statistical procedures, such as the *maximum likelihood* method.

B. The next, more complex, class of reliability models, in addition to describing the behavior of the underlying *survival variable*, accommodates the so-called *explanatory variables* (or *covariates*) such as the product's material properties or usage conditions. This helps to improve the prediction accuracy of future failure times, as well as identify critical variables (*risk factors*) that impact the product's reliability. Model parameter estimation is a bit more complex and involves methods, such as *partial likelihood*.

C. The most sophisticated (and most modern) class of reliability models is referred to as *deep survival models*. These models use *machine learning* (ML) methods, such as *neural networks* or *support vector machines*, to model the complex relationships between the explanatory variables and the probability of failure. The parameters of these models are estimated from large and structurally complicated data sets using special techniques, such as Bayesian networks, transfer learning, etc. The deep survival models can accommodate a wide range of explanatory variables (features), including *sensor* (not to be confused with *censored*) data.

2. APPLICATIONS

A. Numerous applications of classical reliability models can be found in engineering (e.g., [1-3]), medicine (e.g. [4]), actuarial science (e.g., [5]) and other fields.

B. Case studies involving reliability models with time-independent (and/or time-dependent) covariates can again be found in engineering (e.g. [2, 6]), drug discovery (e.g. [7]), finance (e.g. [8]), etc.

C. Finally, the applications of the deep survival models include classification problems, e.g. [9], prediction/forecasting problems, e.g. [10] and prognostic health monitoring, e.g. [11] among others.

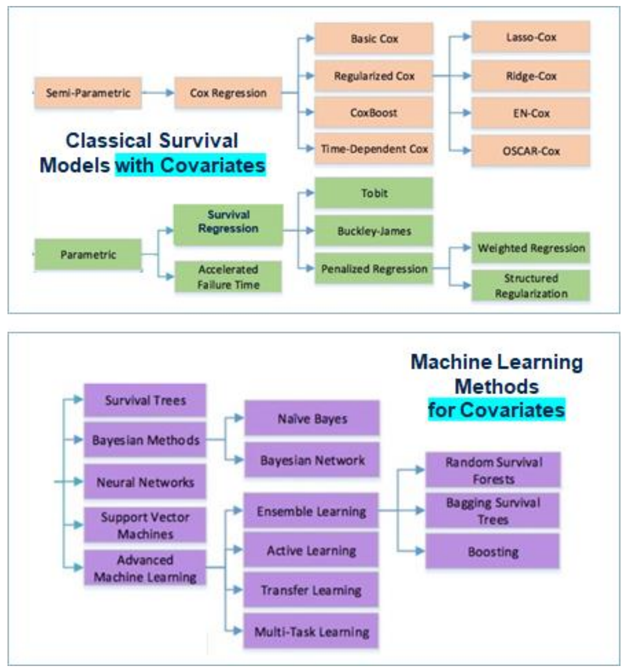

Figure 1. Machine Learning methods in survival analysis.

3. DISCUSSION

The class of reliability models with explanatory variables (described in 1.1 B) can be thought of as:

$$y(t) = f(t|\vec{\theta}, \vec{\vartheta})$$

where: t is the underlying survival variable, $y(t)$ is a probabilistic function (such as the reliability/survival function or the hazard function), $\vec{\theta}$ is the vector of parameters characterizing the probabilistic function, and $\vec{\vartheta}$ is the vector of parameters characterizing the explanatory variables (covariates).

The functional choice of $f(.)$ in the above equation is typically constrained to either the *Cox Proportional Hazards Model* or the *Survival Regression Model* — see Fig. 1. In the former case, the covariates influence the baseline hazard function, while in the latter — the location parameter of the underlying lifetime distribution.

So, the difference between the *classical survival models* and the *deep survival models* is only in the way that covariates are treated. Classical survival models typically employ a multiple linear regression structure for covariates, and parameter estimation is achieved via optimization techniques, such as the *least squares* (incl. *weighted* or *constrained least squares*) or *maximum likelihood* methods. Deep survival models accommodate covariates via multiple layers of interconnected nodes (as in *neural nets*), thus allowing to model complex, non-linear relationships and resolve complications like multicollinearity; parameter estimation here is achieved via more involved optimization techniques, such as *backpropagation*.

4. CONCLUSIONS

With the emergence of "big data", machine learning methods proved to be efficient in the fields of *image/speech reignition*, *natural language processing*, *fraud detection*, and others. The application of ML in *reliability analysis* became known as the *deep survival models*. From the standpoint of the probabilistic structure, these models are no different from classical survival models, such as the *Cox proportional hazards* model and *survival regression*. However, they become quite useful in accommodating large quantities of explanatory variables (covariates) and accounting for complex (often latent) interactions between them.

5. REFERENCES

[1] Nelson, W. (1982). *Applied Life Data Analysis*. Wiley, New York.

[2] Meeker, W.Q., Escobar L.A., and Pascual, F.G. (2021). *Statistical Methods for Reliability Data*, 2nd edition. Wiley, New York.

[3] Modarres, M., Kaminskiy M.P., and Krivtsov V.V. (2016) *Reliability Engineering and Risk Analysis*, 3rd edition, Taylor & Francis, London.

[4] Li, J. and Ma, S. (2013) *Survival Analysis in Medicine and Genetics*. Taylor & Francis, London.

[5] Macdonald, A. S., Currie, I. D., Richards, S. J. (2018). *Model ing Mortality with Actuarial Applications*. Cambridge University Press.

[6] Krivtsov, V.V., Tananko, D.E. and Davis, T.P. "Regression Approach to Tire Reliability Analysis". *Reliability Engineering & System Safety*, 2002, vol. 78, # 3, pp. 267-273.

[7] Rajula, H.S.R., Verlato, G., Manchia, M., Antonucci, N., Fanos, V. "Comparison of Conventional Statistical Methods with Machine Learning in Medicine: Diagnosis, Drug Development, and Treatment" *Medicina*, 2020, No. 56 (9), pp. 455-450.

[8] Dirick, L., Claeskens, G. and Baesens, B. Time to default in credit scoring using survival analysis: a benchmark study. *J Oper Res Soc*, 2017, 68, pp. 652–665.

[9] Li, X., Krivtsov, V.V. and Arora, K "Attention-Based Deep Survival Model for Time Series Data", *Reliability Engineering & System Safety*, 2022, Vol. 217.

[10] Chehade, A., Savargaonkar, M., and Krivtsov, V.V. "Conditional Gaussian Mixture Model for Warranty Claims Forecasting", *Reliability Engineering & System Safety*, 2022, Vol. 218.

[11] Wang W., Zhao Y., and Addepalli, S. "Remaining Useful Life Prediction using Deep Learning Approaches: A Review", *Procedia Manufacturing*, 2020, Vol. 49, pp. 81-88.