Vol.6 No.2



ISSN 2155-9562

Risk analysis: Survival models vs Machine Learning. Application to Alzheimer prediction

C. Huber-Carol, S. Gross, F. Vonta

¹University of Paris Descartes, France ²Baruch College, New York, USA ³Technical University of Athens, Greece

Abstract

 $\mathbf{W}_{ ext{e}}$ present here the statistical models that are most in use in survival data analysis. The parametric ones are based on explicit distributions, depending only on real unknown real parameters, while the preferred models are semi-parametric, like Cox model, which imply unknown functions to be estimated. Now, as big data sets are available, two types of methods are needed to deal with the resulting curse of dimensionality including non-informative factors which spoil the informative part relative to the target: on one hand, methods that reduce the dimension while maximizing the information left in the reduced data, and then applying classical stochastic models; on the other hand algorithms that apply directly to big data, i.e. artificial intelligence (AI or machine learning). Actually, those algorithms have a probabilistic interpretation. We present here several of the former methods. As for the latter methods, which comprise neural networks, support vector machines, random forests and

more (see second edition, January 2017 of Hastie, Tibshirani et al, "The elements of Statistical Learning"), we present the neural networks approach. Neural networks are known to be efficient for prediction on big data. As we analysed risk factors for Alzheimer, using a classical stochastic model, on a data set of around 5000 patients and p = 17 factors, we were interested in comparing its prediction performance with the one of a neural network on this relatively small sample size data.

Keywords—Alzheimer disease, Machine learning, Neural networks. Survival models

Biography:

C. Huber-Carol was an alumni at EcoleNormaleSupérieure (Paris France) in Mathematics and Physics, had a PHD thesis at (EigenössischeTeschnisheHochshule, ETH Zürich. Switzerland), on "Robust tests", an habilitationat Paris University in 1980, then turned, from theoretical maths and stat, to biomedical applications with a specialization in survival data analysis.

She pertains to the editorial board of Lifetime Data Analysis journal (LIDA) and is editor in chief of theISTE (Wiley) collection "Stochastic Models In Survival Analysis and Reliability". Three volumes appeared, "Reliability of Engineering Systems and Technological Risks", Vladimir Rykov, 2016, "Stochastic Risk Analysis and Management", Boris Harlamov, 2017, " Chi-squared Goodness-of-fit Tests for Censored Data", Mikhail Nikulin and Ekaterina Chimitova, 2018.

14th International Online Congress on Alzheimers & Dementia ; October 05, 2020.