Machine Learning Methods for Prediction in Intensive Care

Fabián Güiza¹, Jan Ramon¹, Daan Fierens¹, Geert Meyfroidt², Hendrik Blockeel¹, Maurice Bruynooghe¹, and Greet Van Den Berghe²

¹ Department of Computer Science, K.U.Leuven
² Department of Intensive Care Medicine, University Hospital, Leuven

Background: An average ICU patient is estimated to be described by more than 200 different parameters, making it likely that there is more information in this data than what is currently being extracted from it by humans. Machine learning methods could assist clinicians by analysing this large amount of ICU data to build models that predict the occurrence of specific clinical problems earlier than an experienced intesivist would.

Purpose: To evaluate the applicability of machine learning methods for predicting kidney dysfunction and predefined hyper-inflammatory states.

Materials and Methods: A database of 1548 patients from a randomised controlled trial studying intensive insulin therapy in a surgical intensive care unit was used for this study ³. A total of 12 prediction tasks were considered, consisting of development and recovery from hyper-inflammatory states and kidney dysfunction, with predictions ranging from 1 to 4 days in advance. Four different machine learning algorithms were used: Decision trees (DT), First Order Random Forests (FORF), Naive Bayes (NB) and Tree Augmented Naive Bayes(TAN).

Results: Criteria for discrimination and calibration were Area Under the Receiver Operator Characteristic Curve (aROC) of at least 80% and a Hosmer-Lemeshow H-statistic p-value greater than 0.05. Except for the prediction of development of inflammation, all prediction tasks regarding development satisfied the required criteria. While recovery from kidney dysfunction was predicted up to 4 days in advance, none of the predictions of recovery from hyper-inflammatory states satisfied the criteria completely. Table 1 shows results of a subset of prediction tasks.

Conclusions: For the ICU database studied and the predictive tasks considered, standard machine learning techniques, result in predictive models with good performances according to the selected criteria.

³ Van Den Berghe G, Wouters P, Weekers F, et al: Intensive Insulin Therapy in Critically Ill Patients. New England Journal of Medicine 345: 1359-1367, 2001

Task	Ν	NbEx	p(+)	DT	FORF	NB	TAN
1	1	3527	0.24	$66\% \ / \ 0.30$	70% / 0.59	70% / 0.12	73% / 0.001
1	2	2965	0.31	63% / 0.52	$66\% \ / \ 0.70$	$69\% \ / \ 0.02$	$70\% \ / \ 0.003$
2	1	6871	0.08	$82\% \ / \ 0.55$	$84\% \ / \ 0.58$	$82\% \ / \ 0.53$	83% / 0.35
2	2	5772	0.10	$79\% \ / \ 0.36$	$83\% \ / \ 0.63$	81% / 0.33	$82\% \ / \ 0.09$
2	3	5131	0.12	79%~/~0.32	$83\% \ / \ 0.63$	$81\% \ / \ 0.41$	$82\% \ / \ 0.11$
3	1	9467	0.02	$87\% \ / \ 0.86$	$91\% \ / \ 0.89$	$91\% \ / \ 0.59$	93% / 0.20
3	2	8138	0.03	$87\% \ / \ 0.57$	$91\% \ / \ 0.85$	$90\% \ / \ 0.58$	$92\% \ / \ 0.31$
3	3	7325	0.04	$85\% \ / \ 0.51$	$90\% \ / \ 0.50$	$89\% \ / \ 0.33$	$92\%\ /\ 0.24$
4	2	7084	0.05	$82\% \ / \ 0.57$	$86\% \ / \ 0.57$	86% / 0.20	88% / 0.40
4	3	6329	0.06	84% / 0.41	87% / 0.20	87% / 0.21	88% / 0.25
4	4	5790	0.07	$83\% \ / \ 0.25$	$88\% \ / \ 0.35$	$87\% \ / \ 0.04$	$87\%\ /\ 0.05$

Table 1. Subset of Results of prediction tasks concerning development of: (1)Inflammation, (2)Severe Inflammation, (3)Inflammation-Shock and (4)Kidney Dysfunction, up to N days in advance. p(+) is the percentage of positive examples from the total NbEx examples