A Machine Learning Approach to Investigate the Predictive Value of Pulse Pressure in ICU Mortality-Risk

Ahmed Y. A. Amer, Vranken J., Mesotten D., Vandervoort P., Storms V., Luca S., Vanrumste B., Aerts J.-M.

***Abstract*—We present a feature engineering study aiming at exploring features that may enhance the mortality-risk prediction of ICU patients. We identify pulse pressure as a potential physiological variable for ICU mortality prediction based on the classification performance of a linear hard margin support vector machines classifier. Features from the pulse pressure (PP), systolic and systolic blood pressure are used as input and the label of survival/mortality is used as an output. Complementary patients are addressed with PP which hints at using different features for different patients profiles.**

I. Introduction

Intensive monitoring of vital signs in intensive care unit (ICU) can capture clinical deterioration at an early phase and thus improve patient outcome. In the past, multiple scoring systems have been developed (e.g., APACHE, SAPS) to provide insights and even predictions regarding ICU patient mortality. However, these scoring systems are population-based and often use summarized nongranular data. This study, which is a part of a large research action, aims at exploring the use of feature engineered blood pressure variables for mortality-risk prediction of ICU patients.

II. Methods

The data is provided by intensive care unit and coronary care unit of Ziekenhuis Oost-Limburg (Genk, Belgium). The total study population consists of 447 patients with 450 admissions, of which 170 are labelled with mortality, and 280 with survival. The observations are collected with a sampling rate of 0.5-1 observation per hour. The considered number of observations per patient is defined as the last 84 observations before discharge (on average the last five days of the stay). The first 60 observations out of 84 are considered for feature extraction to predict mortality/survival 24 observations ahead at discharge (i.e., after observation 84). The used machine learning classifier is a linear hard-margin support vector machines (SVM's). This classifier is chosen because of its properties: maximizing the margin between the two classes without tolerance to error and using a linear combination between the features to predict the outcome [1].

A.Y.A Amer and B. Vanrumste are with the KU Leuven, E-MEDIA, Department of Electrical Engineering (ESAT) STADIUS, (ESAT) TC, Campus Group T, Leuven, Belgium. ahmed.youssefaliamer@kuleuven.be and bart.vanrumste@kuleuven.be

A.Y.A Amer and J.-M. Aerts are with the KU Leuven, Measure, Model & Manage Bioresponses (M3-BIORES), Department of Biosystems, Leuven, Belgium. jean-marie.aerts@kuleuven.be.

S. Luca is with Ghent University, Department of Data Analysis and Mathematical Modelling, Ghent, Belgium. stijn.luca@ugent.be.

J. Vranken, D. Mesotten, P. Vandervoort, V. Storms are with Hasselt University, Faculty of Medicine and Life Sciences, Hasselt, Belgium; Limburg Clinical Research Center / Mobile Health Unit.

J. Vranken, D. Mesotten, P. Vandervoort, V. Storms are with Ziekenhuis Oost-Limburg, Department of Anesthesiology, Department of Cardiology and Department Future Health, Genk, Belgium; Limburg Clinical Research Center / Mobile Health Unit.

III. Results and Discussion

Based on an in-depth investigation of the deceased patients and their vital signs, two behaviors reoccurred. The first behavior is that the systolic blood pressure (SBP) is descending and approaching the level of the diastolic blood pressure (DBP). The other behavior is the large difference between the SBP and DBP. For the survival patients, these behaviors are infrequent or missing. Hence, the difference between SBP and DBP, also named pulse pressure (PP) [2], can be considered as a discriminating variable. The features of *minimum, maximum, mean, median, standard deviation, variance, and energy* are extracted from SBP, DBP, PP and their first derivative. Also, *crossing the mean rate and outlier-detection* features are extracted from SBP, DBP and PP. Features are extracted with window size of 15 observations. By considering only the features extracted from SBP and DBP, the classification output is found: 72 true positives (TP), 199 true negatives (TN), 98 false negatives (FN) and 81 false positives (FP). With accuracy of 60.22%, sensitivity of 42.35%, specificity of 47.05% and F1-score of 0.446. On the other hand, considering only the extracted features from PP, the classification output is found: 45 TP, 222 TN, 125 FN and 58 FP. With accuracy of 59.33%, sensitivity of 26.5%, specificity of 43.7% and F1-score of 0.334. It is important to notice that PP features recognized 14 new true positives subject and 58 new true negatives that SBP and DBP features did not recognize. Therefore, we may conclude that different patient profiles require a different set of features and variables. Moreover, the PP variable and its extracted features can be considered in a more extensive study on feature engineering for mortality-risk prediction as the results are in line with previous medical work [2].

Acknowledgment

This research was funded by a European Union Grant through the Interreg V-A Euregio Meuse-Rhine wearIT4health project. The ICU data is provided by Ziekenhuis Oost-Limburg (ZOL) hospital in Genk/Belgium.

References

[1] Cortes C. and Vapnik V., "Support-vector networks.", *Machine learning (20)*1995, 273-297.

[2] Voors A. A., et al.," Low pulse pressureis independently related to elevated natriuretic peptides and increased mortality in advanced chronic heart failure.", *European Heart Journal*, (26) 2005, 1759-1764.