Abstract

Introduction: Mechanical ventilation plays a crucial role in modern medical care, as it is the primary treatment for various respiratory and breathing disorders. However, like other medical devices, mechanical ventilators are prone to generate false alarms. These can lead to alarm fatigue, a state in which medical personnel exhibit a reduced ability to notice or respond to alarms or to treatment errors when staff incorrectly assume the alarm to be valid. Because no established method exists for dealing with frequent false alarms, this work aims to identify viable algorithms for use in an autonomous system that can detect and appropriately handle false alarms resulting from equipment errors.

Methods: To identify viable candidates, I used an existing comparison study of algorithms for time series anomaly detection by Schmidl et al. [1]. I defined the requirements that the intended application imposes on a viable algorithm and developed a context-based evaluation scheme using multi-attribute decision making. I then validated the weighting of the scheme using a Monte Carlo simulation and ranked the algorithms according to their viability, separately for each machine-learning type.

Results: I identified seven key requirements, including runtime constraints, hardware demands, anomaly-detection performance and input-data characteristics. For all but one requirement, I was able to obtain relevant metrics from the available dataset and construct a scoring function using multi-attribute decision making. I selected weights for this scoring function and assessed their robustness and stability through a Monte Carlo simulation. In the simulation it became evident, that the most viable algorithms are PCC for unsupervised learning, RBForest and RobustPCA for semi-supervised learning and Normalizing Flows for supervised learning.

Discussion: All highly viable algorithms exhibited both high mean ranks, indicating high viability and low standard deviations, indicating high rank stability, across the Monte Carlo simulation. Nonetheless, the scoring function has limitations: the dataset was restricted, runtime requirements for the specific clinical use case were unavailable, one requirement could not be quantified and algorithm crashes during time series execution may have introduced bias into the dataset.

Conclusion: Future work is required to identify the most viable algorithm or ensemble for developing an autonomous alarm management system for mechanical ventilation. Such work could leverage the evaluation scheme I proposed in this work, using simulated patient data and incorporating additional context regarding available processing times. Nevertheless, given the Monte Carlo validation and the focus on algorithm rankings rather than raw scores, the resulting algorithm ranking I presented in this work provides a robust guideline for future work.