Abstract 49968

| | |
|---|---|
| Type: | Scientific Abstract |
| Category: | 04.06 - Diagnostic Techniques and Monitoring (CC) |
| Presentation preference: | Either Poster or Oral |
| Status: | ACCEPTED |
| Session Code: | B94 |
| Session Title: | PREDICTIONS AND OUTCOMES IN THE ICU |
| Session Day/Date: | Monday, May 19, 2014 |
| Session Starting/Ending Times: | 2:00pm - 4:30pm |
| Session Type: | Mini Symposium |

Affiliations: School of Nursing and Department Critical Care Medicine, School of Medicine, University of Pittsburgh, and the Auton Lab, Carnegie Mellon University Pittsburgh, PA, USA.

Rationale: Bedside monitors issue single vital sign (VS) alarms (individual parameter crosses threshold), but many are due to monitoring artifacts, causing alarm fatigue. Better approaches for artifact filtering are required, but the task of classifying events as artifacts vs. real alerts is complex, requiring sophisticated algorithms acting on real-time data. Creating a bank of correctly labeled annotations of events as real or artifact to develop these algorithms is essential, but requires considerable expert input. We propose using active learning, a machine learning approach that learns a model based on the existing labeled samples, which it uses to iteratively select small numbers of VS events for expert annotation. The samples are selected in a manner which decreases the uncertainty in classifying the unlabeled samples, leading, after a limited number of annotations, to a testset performance comparable to using a large number of expert-annotated events upfront, thereby reducing expert effort.

Methods: We recruited 314 admissions to a 24-bed stepdown unit, recording 18,314 hrs of noninvasive VS monitoring data at 1/20Hz frequency for continuous peripheral pulse oximetry ($SpO_2$), heart rate, and respiratory rate; noninvasive blood pressure (BP) measured every 2h. $SpO_2$ events (<85%; n=219) were visually adjudicated and annotated by two experts (MRP, MH) as artifact (14%) or real. BP events (systolic BP<80 or >200 mmHg, diastolic BP>110 mmHg; n=96) were similarly annotated (37.5% artifact). We simulated an active learning system using expert-annotated $SpO_2$ and BP data as ground truth. First, the system builds a logistic regression model with a fraction of expert-annotated data points (e.g. 10%). It then proposes a list of events to be annotated, based on classification uncertainty. Once feedback is received (ground truth label supplied) the model is updated, and, based on this new information, the system proposes the next batch of events. We tracked model performance at each stage on hold-out data.

Results: Our active learning approach performance is reported as ability to correctly adjudicate events (by ROC curve in a 10-fold cross validation setup) vs. percentage of expert-annotated data used in the process. Near-Optimal model performance for $SpO_2$ (AUC 77%) was reached using 30% of the expert-annotated data, while optimal BP performance (AUC 90%) was achieved using 50% of the data. When only 10% of the expert-annotated data were used for $SpO_2$ or BP events, each model achieved 80% of the eventual optimal performance.

Conclusions:  An active learning method can reduce the amount of data needing human expert annotation when classifying monitoring events as artifact vs. real, although classification is more difficult for $SpO_2$. Nevertheless, further refinements of such algorithms hold promise to classify incoming monitoring data and inform clinical actions.

| Table 1. Mean Area Under the Curve (AUC) score for 10-fold cross validation at various stages of active learning to correctly annotate events as artifact or real alerts | | |
|---|---|---|
| Percentage of a priori expertly annotated events in the model | $SpO_2$ mean AUC score on test set ± SD | Blood Pressure mean AUC score on test set ± SD |
| 10% | 65% ± 15% (84% of optimum) | 72% ± 18% (80% of optimum) |
| 30% | 77% ± 11% (*optimum) | 83% ± 15% |
| 50% | 79% ± 11% | 90% ± 10% (*optimum) |
| 70% | 78% ± 10% | 90% ± 8% |
| 90% | 79% ± 10% | 91% ± 8% |
| 100% | 79% ± 10% | 91% ± 8% |