Interpretable active learning in support of clinical data annotation

Presented During: Quality and Safety 04
Tuesday, January 20, 2015: 12:15 PM - 1:15 PM

Poster No:

797

Submission No:

4763

Type:

Abstract

Presentation Format:

Poster Only

Authors:

Donghan Wang[1], Madalina Fiterau[2], Artur Dubrawski[2], Marilyn Hravnak[3], Gilles Clermont[4], Michael Pinsky[5]

Institutions:

[1]Carnegie Mellon University Robotics Institute, Pittsburgh, PA, [2]Carnegie Mellon University Auton Lab, Pittsburgh, PA, [3]University of Pittsburgh, Pittsburgh, PA, [4]University of Pittsburgh Medical Center, Pittsburgh, PA, [5]Univ. of Pittsburgh School of Medicine, Pittsburgh, PA

First Author / Corresponding Author:

Donghan Wang, MS - Contact Me
Carnegie Mellon University Robotics Institute
Pittsburgh, PA

Co-Author(s):

Madalina Fiterau - Contact Me
Carnegie Mellon University Auton Lab
Pittsburgh, PA

Artur Dubrawski - Contact Me
Carnegie Mellon University Auton Lab
Pittsburgh, PA

Marilyn Hravnak, RN, PhD, FCCM - Contact Me
University of Pittsburgh
Pittsburgh, PA

Gilles Clermont, MD - Contact Me
University of Pittsburgh Medical Center
Pittsburgh, PA

Michael Pinsky, MD, MCCM - Contact Me
Univ. of Pittsburgh School of Medicine
Pittsburgh, PA

Introduction:

Machine learning (ML) can be useful in applications of clinical importance, but it usually requires a supply of adjudicated data to work. Annotation of large amounts of clinical data consumes valuable time of expert clinicians. We propose to determine whether active learning (AL) can be used to reduce expert effort.

Methods:

Noninvasive vital signs monitoring data from 8 week admissions in a 24-bed step-down unit (heart rate [HR], respiratory rate [RR; bioimpedance], oscillometric blood pressure [BP], peripheral oximetry [SpO2]) recorded at 20s intervals included 585 episodes of persisting SpO2 instability (SpO2>85%). A committee of experts adjudicated the initial batch (201) of these alerts labeling them as true instabilities (133), artifacts (39) or unclear (29), used to train a Random Forest (RF) classifier to label alerts as true or artifact. AL was used to select still unlabeled episodes that, when labeled by humans, would improve the classifier the most. We used a method derived from (Fiterau, Dubrawski: Projection Retrieval for Classification, NIPS 2012) to select data that maximizes the expected information gain and presents it in a human-interpretable fashion, and compared it against a RF classifier that selects the most uncertain data. At the conclusion of each of the 3 cycles of expert adjudication (yielding 37, 43, 33 annotated cases respectively), we measured the number of still unlabeled data that cannot be confidently adjudicated by the respective models.

Results:

Informative AL method was not able to adjudicate 114 cases at the end of the 1st cycle, but at the end of 2nd it could confidently process all unlabeled data saving experts from having to label 304 episodes at that point (52% effort reduction), while RF method would allow 11% effort reduction at the end of cycle 3.

Conclusions:

Informative AL is useful in reducing data annotation efforts by expert clinicians. It could facilitate wider adoption of powerful ML methodologies in clinical informatics to enhance situational awareness of clinicians and patient outcomes.

Funding: NIH NINR R01NR013912; NSF 0911032, 1320347.

General Classification:

Clinical Research

Patient Type:

Adult

Categories - Abstracts:

Quality and Safety

Keywords:

critical care

mathematical model

multivariable modeling