Early Deterioration Warning for Hospitalized Patients by Mining Clinical Data

Yi Mao, Washington University in St. Louis, USA, and Xidian University, China
Yixin Chen, Washington University in St. Louis, USA
Gregory Hackmann, Washington University in St. Louis, USA
Minmin Chen, Washington University in St. Louis, USA
Chenyang Lu, Washington University in St. Louis, USA
Marin Kollef, Washington University School of Medicine, USA
Thomas C. Bailey, Washington University School of Medicine, USA

ABSTRACT

Data mining on medical data has great potential to improve the treatment quality of hospitals and increase the survival rate of patients. Every year, 4-17% of patients undergo cardiopulmonary or respiratory arrest while in hospitals. Clinical study has found early detection and intervention to be essential for preventing clinical deterioration in patients at general hospital units. This paper proposes an early warning system (EWS) designed to identify the signs of clinical deterioration and provide early warning for serious clinical events. The EWS is designed to provide reliable early alarms for patients at the general hospital wards (GHWs). The main task of EWS is a challenging classification problem on high-dimensional stream data with irregular, multi-scale data gaps, measurement errors, outliers, and class imbalance. This paper proposes a novel data mining framework for analyzing such medical data streams. The authors assess the feasibility of the proposed EWS approach through retrospective study that includes data from 41,503 visits at a major hospital. Finally, the system is applied in a clinical trial at a major hospital and obtains promising results. This project is an example of multidisciplinary cyber-physical systems involving researchers in clinical science, data mining, and nursing staff.

Keyword: Bootstrap Aggregating, Early Warning System, EMA (Exponential Moving Average), Exploratory Undersampling, Logistic Regression

DOI: 10.4018/jkdb.2011070101
1. INTRODUCTION

Within the medical community, there has been significant research into preventing clinical deterioration among hospital patients. Data mining on electronic medical records has attracted a lot of attention but is still at an early stage in practice. Clinical study has found that 4-17% of patients undergo cardiopulmonary or respiratory arrest while in the hospital (The Joint Commission, 2008). Early detection and intervention are essential to preventing these serious, often life-threatening events. Indeed, early detection and treatment of patients with sepsis has already shown promising results, resulting in significantly lower mortality rates (Jones, Brown, Trzeciak, Shapiro, Garrett, Heffner, & Kline, 2008).

In this paper, we consider the feasibility of an Early Warning System (EWS) designed to identify at-risk patients from existing electronic medical records. Specifically, we analyzed a historical data set provided by a database from a major hospital, which cataloged 41,503 hospital visits between July 2007 and July 2011. For each visitor, the dataset contains a rich set of electronic various indicators, including demographics, vital signs (pulse, shock index, mean arterial blood pressure, temperature, and respiratory rate), and laboratory tests (albumin, bilirubin, BUN, creatinine, sodium, potassium, glucose, hemoglobin, white cell count, INR, and other routine chemistry and hematology results). All data contained in this dataset was taken from historical EMR databases and reflects the kinds of data that would realistically be available at the clinical warning system in hospitals.

Our EWS is designed to provide reliable early alarms for patients at the general hospital wards (GHWs). Unlike patients at the expensive intensive care units (ICUs), GHW patients are not under extensive electronic monitoring and nurse care. Sudden deteriorations (e.g., septic shock, cardiopulmonary or respiratory arrest) of GHW patients can often be severe and life threatening. EWS aims at automatically identifying patients at risk of clinical deterioration based on their existing electronic medical record, so that early prevention can be performed. The main task of EWS is a challenging classification problem on high-dimensional stream data with irregular, multi-scale data gaps, measurement errors, outliers, and class imbalance.

To address such challenges, in this paper, we first develop a novel framework to analyze the data stream from each patient, assigning scores to reflect the probability of intensive care unit (ICU) transfer to each patient. The framework uses a bucketing technique to handle the irregularity and multi-scaleness of measuring gaps and limit the size of feature space. Popular classification algorithms, such as logistic regression and SVM, are supported in this framework. We then introduce a novel bootstrap aggregating scheme to improve model precision and address over-fitting. Furthermore, we employ a smoothing scheme to deal with the outliers and volatility of data streams in real-time prediction.

Based on the proposed approach, our EWS predicts the patients’ outcomes (specifically, whether or not they would be transferred to the ICU) from real-time data streams. This study serves as a proof-of-concept for our vision of using data mining to identify at-risk patients and (ultimately) to perform real-time event detection.

2. RELATED WORK

Medical data mining is one of key issues to get useful clinical knowledge from medical databases. These algorithms either rely on medical knowledge or general data mining techniques.

A number of scoring systems that already exist use medical knowledge for various medical conditions. For example, the effectiveness of Several Community-Acquire Pneumonia (SCAP) and Pneumonia Severity Index (PSI) in predicting outcomes in patients with pneumonia is evaluated in Yandiola et al. (2009). Similarly, outcomes in patients with renal failures may be predicted using the Acute Physiology Score (12 physiologic variables), Chronic Health Score (organ dysfunction), and APACHE
18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:

www.igi-global.com/article/early-deterioration-warning-hospitalized-patients/63614?camid=4v1

This title is available in InfoSci-Journals, InfoSci-Journal Disciplines Medicine, Healthcare, and Life Science, InfoSci-Healthcare Administration, Clinical Practice, and Bioinformatics eJournal Collection, InfoSci-Knowledge Discovery, Information Management, and Storage eJournal Collection, InfoSci-Physical Sciences, Biological Sciences, and Engineering eJournal Collection. Recommend this product to your librarian:

www.igi-global.com/e-resources/library-recommendation/?id=2

Related Content

Discovering Protein Complexes in Protein Interaction Networks
Clara Pizzuti and Simona Ester Rombo (2009). Biological Data Mining in Protein Interaction Networks (pp. 144-168).
www.igi-global.com/chapter/discovering-protein-complexes-protein-interaction/5563?camid=4v1a

Binarization and Validation in Formal Concept Analysis
www.igi-global.com/article/binarization-validation-formal-concept-analysis/75151?camid=4v1a
Computational Modeling and Simulations in Life Sciences
[www.igi-global.com/article/computational-modeling-and-simulations-in-life-sciences/89397?camid=4v1a](www.igi-global.com/article/computational-modeling-and-simulations-in-life-sciences/89397?camid=4v1a)

Combinatorial Optimization Algorithms for Metabolic Networks Alignments and Their Applications
[www.igi-global.com/article/combinatorial-optimization-algorithms-metabolic-networks/52768?camid=4v1a](www.igi-global.com/article/combinatorial-optimization-algorithms-metabolic-networks/52768?camid=4v1a)