A Predictive Analytics Methodology to Assess and Optimize Readmission Risk in Heart Failure Patients

Pavitra Krishnaswamy1+*, Savitha Ramasamy1*, Abdullah Al-Dujaili2, Srinath Sridharan1, Geraldine Goh3, Tong Shao Chuen3, Khin Chaw Yu Aung3, Gerard Leong Kui Toh3, Michael Ross Macdonald3, Sheldon Lee Shao Guang3, Cao Yan3, Suresh Sundaram2, Chow Wai Leng3+

1Institute for Infocomm Research, Singapore
2Nanyang Technological University, Singapore
3Changi General Hospital, Eastern Health Alliance, Singapore
*pavitrak@i2r.a-star.edu.sg, wai.leng.chow@easternhealth.sg

Abstract
We describe a predictive modeling and decision support framework that integrates machine learning and optimization for personalized clinical decision support. We pilot the approach on data from a congestive heart failure patient cohort, and demonstrate the ability to predict and optimize readmission risk in a clinically meaningful manner.

Introduction
The accumulation of vast medical datasets within hospitals offers opportunities to leverage a rich combination of clinical history, physiological readings and administrative data accumulated during routine clinical care for personalized clinical decision support. Yet, most clinical decision support systems that are deployed in practice primarily perform data visualization and do not incorporate advanced predictive analytics or data mining methodologies.

Historically, research on clinical data mining tools has focused on: (a) time series analysis or physiological waveforms, (b) bio-statistical methods for epidemiological analysis of trends in diagnostic results and/or patient outcomes, and (c) word search tools that employ semantic processing of clinician queries for retrieving relevant patient treatment guidelines (Iavindrasana et al. 2009). These data mining tools provide objective outputs (e.g., quantitative trends in diagnostic results, or a list of treatments provided for similar patients) that readily fit within the clinical workflow.

Recently, several studies have investigated the use of machine learning tools for predicting clinical outcomes and risk profiles for adverse events using retrospective clinical datasets (Kansagara, et al. 2011). However, as these predictions are based on highly complex datasets and models, they are often disconnected from the clinical actions or decisions that might improve the predicted outcomes. This gap poses a significant challenge for adoption of machine learning based predictive tools within clinical practice.

To surmount this translation gap, there is need for data-driven approaches that can not only predict outcomes, but also derive plausible care delivery decisions that may optimize these outcomes for a given patient. Such capabilities would enable clinicians to interpret machine learning predictions and evaluate plausible decisions/parameters to target for optimal outcomes – and move machine learning predictive tools to the realm of clinical decision support.

We propose a new three-step framework to integrate machine learning and optimization techniques for personalized clinical decision support. First, use a retrospective database of multimodal clinical data to train machine learning models and proactively predict clinical outcomes of interest. Second, use the machine learning models to perform black-box optimization for deriving optimal decisions that can facilitate improved care. Third, incorporate constraints based on past similar patients to personalize the decisions for the current patient under consideration. To illustrate the approach, we consider the use case of managing readmission risk in patients admitted for congestive heart failure.

Clinical Use Case
Congestive heart failure (CHF) is a leading cause of hospitalizations, particularly in the rising elderly demographic. Hospital visits account for ~70% of the overall cost of CHF, and a large proportion of hospital visits are readmissions due to exacerbation. As such, there is interest in data-driven tools to predict readmission risk, proactively assess
Questions of interest include: Is readmission risk low enough to discharge a patient? If not, what is the likely cause of readmission? What parameters are optimal targets to reduce risk? Tools to address these questions would be valuable for planning inpatient or transitional care, and targeting resources to those at highest risk.

Proof-of-Concept Evaluation

Data: We develop the proposed predictive analytics approach for this use case, as illustrated above, using retrospective data from an observational study conducted at the Changi General Hospital, Singapore. This study collected extensive clinical data for 530 CHF patients over 688 index and readmission episodes during Nov 2014 – Oct 2015. The dataset comprises over 160 features including demographic information, admitting clinical attributes, in-patient management logistics, prescription data, discharge clinical indicators, post-discharge plan, and best practice element tracking. We consider each episode independently, and preprocess the data by coding text and subjective features into standardized named entities, filling in missing data with medians of available values, and eliminating redundancies to reduce feature dimensionality. We label readmission risk as high if there is any readmission within 1 year following a given episode, and low otherwise. We label readmission cause as heart failure if the readmission is due to CHF-exacerbations and non-heart failure otherwise.

Methods: We cast the discharge planning problem as a series of classification and black box optimization problems. First, we develop neural network based classifiers for 1-year readmission risk and readmission cause respectively. In particular, we trained Radial Basis Function (RBF) classifiers, so as to quantize probability of the predicted classes. We use data from 10 randomly selected subsets of 80% of the patients to train the classifier. Next, we optimize the readmission risk outcomes predicted by the RBF classifier. As the explicit functional form of the readmission risk is unavailable, we employ a black-box optimization technique. In particular, given the complex multivariate and ill-conditioned nature of the objective function, we employ a Covariance Matrix Adaptation optimizer which is based on an evolutionary strategy defined by the correlations within the data (Hansen and Ostermeier 2001). Amongst the features used to train the RBF classifiers, only the clinically actionable features are specified as decision variables for optimization. All decision variables are bounded by admissible clinical ranges. The optimizer outputs the best decisions for the patient under evaluation, with the importance ranking to enable clinicians to prioritize accordingly.

Results: The classifiers predict readmission risk and cause accurately with testing AUC of 79% and 69% respectively. Evaluating the optimizer results in view of the predicted causes, we find that the optimal decisions are clinically meaningful. For episodes followed by heart failure readmissions, the optimizer prioritizes actions to regulate indicators of cardiac function. For episodes followed by non-cardiac related readmissions, the optimizer prioritizes actions to regulate non-cardiac clinical indicators. For example, the optimal decisions for cases followed by chronic kidney disease related readmissions tend to prioritize regulation of renal function. For episodes with no subsequent readmissions, the optimizer prioritizes actions to reduce length of hospital stay. Finally, the optimal decisions and their priority order are implicitly personalized to individual patients, therefore adding unique value beyond generic therapy guidelines formulated at systems/cohort levels.

Future Directions

We put forth a novel framework and preliminary system for optimal personalized clinical decision support. Future efforts will focus on techniques to propagate uncertainty of machine learning predictions using stochastic optimization techniques. Applications stratifying patients based on resource utilization or risk profiles, and setting care management targets to improve outcomes can be anticipated.

References