

Multiple Model Analytics for Adverse Event Prediction in Remote Health Monitoring Systems

Mohammad Pourhomayoun, Nabil Alshurafa, Bobak Mortazavi, Hassan Ghasemzadeh, Konstantinos Sideris, Bahman Sadeghi, Michael Ong, Lorraine Evangelista, Patrick Romano, Andrew Auerbach, Asher Kimchi, Majid Sarrafzadeh

Abstract— Remote health monitoring systems (RHMS) are gaining an important role in healthcare by collecting and transmitting patient vital information and providing data analysis and medical adverse event prediction (e.g. hospital readmission prediction). Reduction in the readmission rate is typically achieved by early prediction of the readmission based on the data collected from RHMS, and then applying early intervention to prevent the readmission. Given the diversity of patient populations and the continuous nature of patient monitoring, a single static predictive model is insufficient for accurately predicting adverse events. To address this issue, we propose a multiple prediction modeling technique that includes a set of accurate prediction models rather than one single universal predictor. In this paper, we propose a novel analytics framework based on the physiological data collected from RHMS, advanced clustering algorithms and multiple-model-classification. We tested our proposed method on a subset of data collected through a remote health monitoring system from 600 Heart Failure patients. Our proposed method provides significant improvements in prediction accuracy and performance over single predictive models.

I. INTRODUCTION AND BACKGROUND

New technologies such as Body Sensor Networks (BSN) and Remote Health Monitoring Systems (RHMS) allow for collecting continuous data from patients, providing clinical interventions with the goal of preventing medical adverse events (e.g., hospital readmissions, emergency visits, heart attack, and death) associated with chronic conditions. Although RHMS have shown promise in reducing healthcare costs and improving quality of care, effective analysis of the data collected by these systems and the potential benefits of utilizing such analysis is by large an open problem.

Rapid advances in many technological domains including electronics, wireless communications, internet, and sensor design has led to the development of effective RHMS that can collect varying physiological information such as

systolic and diastolic blood pressure (i.e. SBP, DBP), heart rate (HR), weight, glucose levels, EEG and ECG signals, and also physical activity from patients [1]-[8]. Recently, implantable and wearable medical devices, such as smart pills and body sensors play important roles in remote health monitoring systems by capturing, transmitting and controlling the vital information of patients [9]-[11]. RHMS provide a unique platform for physiological data collection using body area networks, information integration, data analysis, medical event prediction, and personalized education [12], [13]. As of October 2012, the US government began implementing the Hospital Readmission Reduction Program, which levies financial punishment on hospitals with high readmission rates [8]. Statistics show that nearly 20% of insured patients are readmitted to hospitals within 30 days after discharge, incurring approximately \$17 billion in charges in 2009 [8]. RHMS provide an opportunity to reduce readmission rates by early prediction of readmission based on the collected data, and then applying effective early intervention to prevent the readmission.

Conventional readmission prediction methods are usually based on a thresholding system. In these methods, the data collected from BSN are compared against pre-defined thresholds to generate appropriate alerts, predicting adverse events (such as re-hospitalization, heart attack, or death) [8]. In recent prediction techniques, analytic methods including Machine Learning algorithms are employed to build a more accurate model for adverse event prediction. The main idea of machine learning based methods is to generate an accurate predictive model based on training and mathematical algorithms, which are later used as real-time predictors by classifying the labels and predicting future adverse events. Our preliminary work from the UCLA Wireless Health Institute shows that using a machine learning engine is more accurate than relying on conventional approaches that use threshold-based predictions [8].

In this paper, we propose a new analytics framework based on the physiological data collected from RHMS, advanced clustering algorithms and multiple-model-classification which can be very effective and beneficial to build more accurate predictor models. The results demonstrate improvements in accuracy and performance of the predictor when applying the proposed method compared to typical single model classifications. In this paper, we focus on readmission prediction of patients with heart failure in the next few days following discharge. The main goal is to reduce the False Positive rate and False Negative rate of the

Mohammad Pourhomayoun, Nabil Alshurafa, Konstantinos Sideris, Bahman Sadeghi, Michael Ong, Majid Sarrafzadeh are with University of California Los Angeles (UCLA), Tel: 607-761-5754, email: mpourhoma@cs.ucla.edu.

Bobak Mortazavi, is with Yale University. Hassan Ghasemzadeh is with Washington State University, Lorraine Evangelista is with University of California Irvine, Patrick Romano is with University of California Davis, Andrew Auerbach is with University of California San Francisco, and Asher Kimchi is with Cedars-Sinai Medical Center.

predictor, which is the typical technique to evaluate the performance of a classifier. In this paper, we use the data collected through BEAT-HF study, one of the largest randomized controlled trials of telemonitoring in patients with heart failure. The primary purpose of BEAT-HF study is to compare the effect of implementing wireless remote monitoring combined with structured telephone monitoring, versus conventional care, on variation in hospital readmission. The study population includes patients with a wide range of demographic and socioeconomic characteristics [18].

The first stage is to measure and collect a patient's physiological data using body area sensors and transmit them to the analytics engine. The next stage is preprocessing the collected data and extracting the features. In the proposed method, we derive 252 statistical features from each data sample. Then, we apply various Feature Selection algorithms to select the most informative features, eliminating redundant or useless data.

In this method, we apply advanced clustering algorithms using a specific set of extracted features to first split the data samples into several clusters. Then, the most informative features of each cluster are selected as the input of individual per-cluster classifiers to label the data samples. In this method, we build more accurate predictor models to classify the data samples of each cluster individually rather than using a single universal classification model for the entire dataset. The results demonstrate significant improvements in accuracy of the analytics system in predicting medical adverse events.

The rest of the paper is organized as follows: Section II discusses the multiple modeling prediction method by providing detailed information about RHMS used to collect data, feature extraction algorithms, data clustering and group-specific feature selection, and finally multiple-model classification. Section III describes the clinical trials, dataset, results and conclusion. Finally, Section IV discusses the future work.

II. MULTIPLE MODELING PREDICTION

A. Remote Health Monitoring and Body Sensor Networks

The first stage in the RHMS for adverse event prediction is gathering the patient's physiological data using body area sensors and transmitting them to an analytics engine. In this study, 600 patients with heart failure are given wireless devices (such as a wireless weight scale and wireless blood pressure cuff) to measure weight (W), systolic and diastolic blood pressure (SBP and DBP) and heart rate (HR) on a daily basis. The devices also collect self-reported questionnaires regarding heart failure symptoms. Then, the collected data are transmitted via Bluetooth to an information gateway such as patient's smartphone. The gateway (smartphone) then transmits the data to a secure database for data storage and processing.

B. Feature Extraction

After receiving patients' physiological data from BSN, the next step is data preprocessing and feature extraction. Statistical and morphological features are the most common features used for data analytics. For example, some studies show that the weight fluctuation over a certain period of time is highly correlated to the function of the heart [15], and consequently might be a useful feature for monitoring of heart failure patients. In this study, we derive 252 total features from each data sample (i.e. each day). In order to consider the physiological status of a patient and extract the statistical features, we use three different time windows with lengths of 3, 5, and 7 days containing patients' physiological measurements (i.e. the weight, systolic and diastolic blood pressure, and heart rate in the past 3, 5 and 7 days). The statistical features contain mean, median, rms, minimum, maximum, variance, maximum fluctuation, standard deviation, and skewness of the measurements in each window (i.e. in the past 3-day, 5-day and 7-day windows).

In addition to statistical features, the threshold based alert information of the patients in the past 3, 5, and 7 days are also included as features. An alert is generated when a physiological measurement is out of the predefined acceptable range. The alerts are usually labeled as medium and high priority depending on the degree of deviation from the defined thresholds. A list of the alerts used in this study is provided in Table I. The weighted summation of these alerts in each time window is also included in the feature set. To this end, the weights of 3 and 1 are assigned to high and medium priority alerts respectively, and then the summation of the alert weights in each time window is considered as a feature.

TABLE I. THRESHOLD BASED ALERTS

Alert Label	Description	Priority
A ₁	Heart Rate 100 - 119 bpm	M
A ₂	Heart Rate > 120 bpm	H
A ₃	Heart Rate < 50 bpm	H
A ₄	Systolic BP < 80 mmHg	M
A ₅	Systolic BP > 160 mmHg	M
A ₆	Weight < 125.00 lb	M
A ₇	Weight increase of 3 lbs over 1 day	H
A ₈	Weight increase of 5 lbs over 7 days	H

We have to note that only a small portion of these 252 extracted features will be eventually used in the classification stage. In this study, after applying advanced per-cluster feature selection algorithms, only about 20 features are selected and used in the classification stage (details are available in the following sections).

C. Data Sample Clustering and Group-Specific Feature Selection

Given the diversity of patient populations and the continuous nature of patient monitoring, a single static model is insufficient for accurately predicting medical adverse events. Consequently, an effective model is more likely to be the one that includes multiple prediction models in a

hierarchical architecture. Our results show that clustering patients based on their baseline information, and performing per-cluster prediction achieves significant improvement in the accuracy of prediction compared to applying a single universal prediction model. Figure 1 demonstrates the block diagram of the proposed multiple prediction modeling.

In this study, we applied the K-Means clustering algorithm to split the data samples into 8 clusters in a two-dimensional feature space based on baseline weight and systolic blood pressure. Figure 2 shows a sample dataset clustered in this domain.

As shown in Figure 1, after clustering the data samples, we apply group-specific (per-cluster) feature selection to select the most prominent features for each cluster and reduce the redundancy. Feature selection is an important field of research in data mining and machine learning. The conventional feature selection algorithms usually focus on specific metrics to quantify the relevance and/or redundancy of each feature with the goal of finding the smallest subset of features that provides the maximum amount of useful information for prediction. Thus, the main goal of feature selection algorithms is to eliminate redundant or irrelevant features in a given feature set. Applying an effective feature selection algorithm not only decreases the computational complexity of the system by reducing the dimensionality and eliminating the redundancy, but also increases the performance of the classifier by removing irrelevant and confusing features.

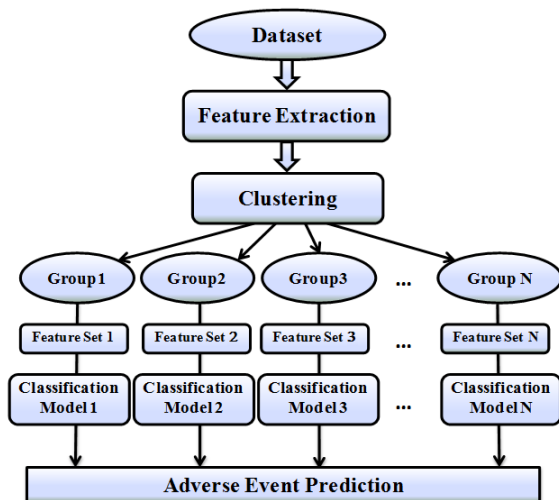


Figure 1. Multiple model classification for medical adverse event prediction.

In this paper, we tried both wrapper and filter methods; the two well-known feature selection categories. Wrapper methods usually utilize a classifier (or any other learning machine of interest) to evaluate feature subsets in an iterative manner according to their predictive power. A new feature subset is used to train a predictive model that will later be evaluated on a testing dataset to assess the relative usefulness of subsets of features [16]. Figure 3-(a) provides an illustration of the steps involved in the wrapper feature selection method.

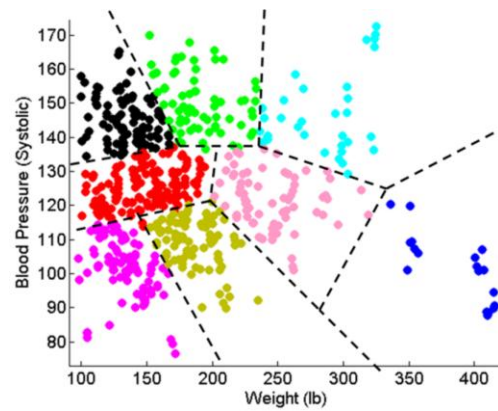


Figure 2. Clustering the data samples based on weight and blood pressure.

Filter methods use a specific metric to score each individual feature (or a subset of features together). The most popular metrics used in filter methods include correlation coefficient, mutual information, Fisher score, chi-square parameters, entropy and consistency. Filter methods are very popular (especially for large datasets) since they are usually very fast and much less computationally intensive than wrapper methods. Figure 3-(b) illustrates the steps involved in the filter feature selection method.

After selecting the best set of features for each one of the 8 clusters, these features will be used in the classification stage to build a group-specific (per-cluster) prediction model.

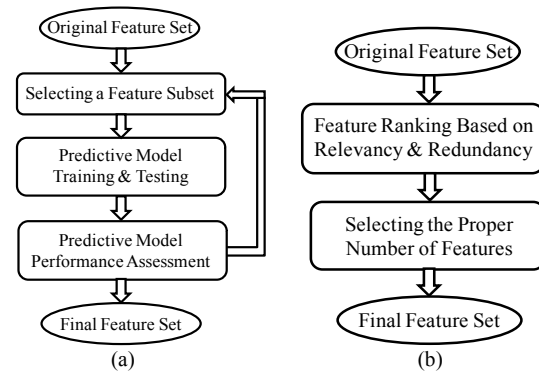


Figure 3. Feature Selection: (a) Wrapper method, (b) Filter method.

D. Multiple Model Classification

Each patient has unique physiological reactions. Thus, in similar health risk conditions, the symptoms (and consequently the physiological data measurements and vital signs) may differ from patient to patient. In analytics systems, because the predictor model is established based on the features extracted from physiological data, designing a single model to perform the prediction for the entire dataset may fail to achieve accurate results. Thus, a major challenge in building the proposed framework is the lack of an accurate universal prediction model to support the entire dataset. To address this issue, we propose a multiple prediction modeling technique that includes a set of accurate prediction models rather than one single universal predictor for the entire dataset.

In this research, we tried various classification algorithms such as SVM, Random Forest, BayesNet, and Artificial Neural Net (ANN) as the predictor. According to our results, a Random Forest classifier with 100 trees provided fast and accurate prediction results for our dataset. Random Forest is an ensemble learning classification method comprising of a collection of decision tree predictors operating based on i.i.d random vectors. In this process, each tree casts a unit vote for the most popular class [17].

We generate Receiver Operating Characteristic (ROC) curves to demonstrate the accuracy and the ability of the classifier to correctly classify the outcomes and predict the adverse events. To compare the single prediction modeling versus multiple prediction modeling, we generate and compare ROC curves for each one of the group-specific classifiers as well as the single model prediction.

III. RESULTS AND CONCLUSION

A subset of data collected through our remote health monitoring system involving 600 patients with Heart Failure (HF) was used for this analysis. The dataset contains physiological measurements including weight, systolic and diastolic blood pressure, and heart rate collected on a daily basis from 600 patients with heart failure. Our goal was to cluster patients and perform a per-cluster adverse event prediction analysis to compare the accuracy of a single prediction model with that of a multiple model. After feature extraction (as mentioned in Section II-B), we split the dataset into 8 clusters using a K-Means clustering algorithm, and then applied individual feature selection and classification on each cluster. The clustering was performed on a two dimensional feature space including weight and blood pressure measurements. We used Random Forrest classifier and 10-fold cross validation over the data clusters and also the entire dataset to evaluate the performance.

Figure 4 shows the receiver operation characteristic (ROC) curves on a true positive rate (TPR) vs. false positive rate (FPR) plane. In this figure, the red dashed curve represents the ROC of a single model prediction trained on the entire dataset. However, other curves demonstrate the classification results for each cluster using cluster-specific feature sets and individual per-cluster classifiers. Also, Table II and III demonstrate the true positive rate for a fixed false positive rate of 30% and 20% for each one of the clusters as well as the single modeling case. As shown in Figure 4, Table II and Table III, clustering and then per-cluster feature selection and classification as an effective approach to multiple modeling can significantly improve the performance and accuracy of the adverse event prediction.

IV. FUTURE WORK

Conventional feature selection approaches do not take into consideration factors such as patient preferences, monitoring costs, and compliance level, which are critical components of a RHMS. For example, measuring weight is much more convenient for patients than measuring blood pressure in an in-home setting. Thus, in our future work, we are planning to take into account parameters such as patient preferences,

monitoring costs, and compliance levels associated with each feature in selecting the most prominent features. Also, we are planning to use contextual features and patient-reported information along with physiological features to improve the performance of the adverse event prediction.

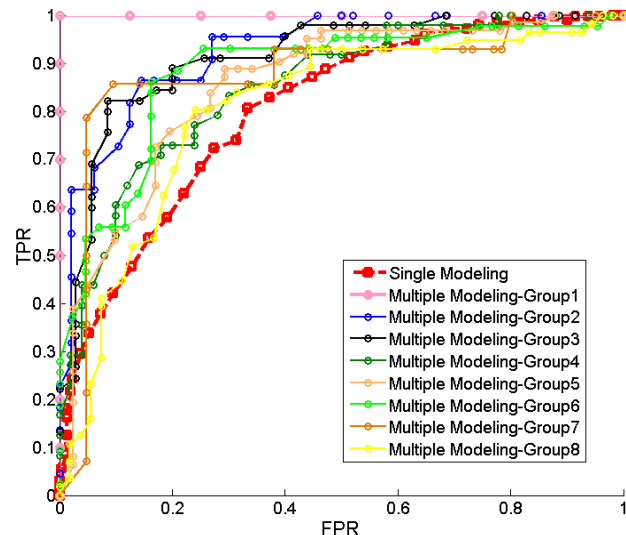


Figure 4. ROC curves for single model and cluster-specific multiple model.

TABLE II. TRUE POSITIVE RATE FOR FIXED FALSE POSITIVE RATE = 30%

Dataset	TPR for FPR=30%
Cluster 1	100%
Cluster 2	95%
Cluster 3	91%
Cluster 4	84%
Cluster 5	89%
Cluster 6	93%
Cluster 7	85%
Cluster 8	83%
Single Modeling	75%

TABLE III. TRUE POSITIVE RATE FOR FIXED FALSE POSITIVE RATE = 20%

Dataset	TPR for FPR=20%
Cluster 1	100%
Cluster 2	86%
Cluster 3	89%
Cluster 4	68%
Cluster 5	76%
Cluster 6	88%
Cluster 7	86%
Cluster 8	73%
Single Modeling	60%

ACKNOWLEDGMENT

The authors would like to appreciate the BEAT-HF team at UCLA, UC Davis, UC Irvine, UC San Diego, UC San Francisco and Cedars-Sinai Health System, and especially the BEAT-HF publication committee for their valuable comments and feedback.

REFERENCES

- [1] M. Lan, N. Alshurafa, H. Ghasemzadeh and M. Sarrafzadeh, "WANDA: An End-to-End Remote Health Monitoring and Analytics System for Heart Failure Patients," *Wireless Health Conf.*, 2012b.
- [2] M. K. Suh, et al., "Dynamic Self-adaptive Remote Health Monitoring System for Diabetics," *EMBC*, 2012.
- [3] M. K. Suh, et al., "A Remote Patient Monitoring System for Congestive Heart Failure," *JOMS*, June 2011a.
- [4] N. Alshurafa, et al. "Anti-Cheating: Detecting Self-Inflicted and Impersonator Cheaters for Remote Health Monitoring Systems with Wearable Sensors," *BSN* 2014.
- [5] M. K. Suh, et al., "WANDA B: Weight and Activity with Blood Pressure Monitoring System for Heart Failure Patients," in *IEEE Workshop on IREHSS*, 2010b.
- [6] N. Alshurafa, et al. "Battery optimization in smartphones for remote health monitoring systems to enhance user adherence." *Int. Conf. on Pervasive Technologies Related to Assistive Environments*. 2014.
- [7] N. Alshurafa, et al. "Improving Compliance in a Remote Health Monitoring System through Smartphone Battery Optimization". *IEEE J Biomed Health Inform*, 2014.
- [8] S. I. Lee, H. Ghasemzadeh, B. J. Mortazavi, M. Lan, N. Alshurafa, M. Ong, M. Sarrafzadeh, *Remote Health Monitoring Systems: What Impact Can Data Analytics Have on Cost?*, in *Wireless Health*, 2013.
- [9] M. Pourhomayoun, Z. Jin and M.L. Fowler, "Accurate Localization of In-Body Medical Implants Based on Spatial Sparsity," *IEEE Transactions on Biomedical Engineering*, 2013.
- [10] M. Pourhomayoun, M. Fowler, Z. Jin, "A Novel Method for Medical Implant In-Body Localization," *Int. Conf. of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2012.
- [11] M. Pourhomayoun, Z. Jin and M.L. Fowler, "Indoor Localization, Tracking and Fall Detection for Assistive Healthcare Based on Spatial Sparsity and Wireless Sensor Network," *International Journal of Monitoring and Surveillance Technologies Research*, 2013.
- [12] K. Hung and Z. Yuan-Ting, Implementation of a WAP-based telemedicine system for patient monitoring, *Information Technology in Biomedicine*, *IEEE Trans. on*, vol. 7, no. 2, pp. 101-107, 2003.
- [13] A. G. Logan, et al., "Mobile Phone-Based Remote Patient Monitoring System for Management of Hypertension in Diabetic Patients," *American Journal of Hypertension*, 2007.
- [14] A. Smith, *Pew Internet*. Nearly half of American adults are smartphone owners, <http://pewinternet.org/Reports/2012/Smartphone-Update-2012/Findings.aspx>, March 2012.
- [15] S. I. Chaudhry, Y. Wang, J. Concato, T. M. Gill and H. M. Krumholz, "Patterns of weight change preceding hospitalization for heart failure," *Circulation*, vol. 116, pp. 1549-1554, 2007.
- [16] Isabelle Guyon, Andre Elisseeff, "An Introduction to Variable and Feature Selection", *Journal of Machine Learning Research* 3 (2003).
- [17] Breiman, Leo. "Random Forests". *Machine Learning* 45 (1): 5–32, (2001).
- [18] Black JT, Romano PS, Sadeghi B, Auerbach AD, Ganiats TG, Greenfield S, Kaplan SH, Ong MK; BEAT-HF Research Group. A remote monitoring and telephone nurse coaching intervention to reduce readmissions among patients with heart failure: study protocol for the Better Effectiveness After Transition - Heart Failure (BEAT-HF) randomized controlled trial. *Trials*. 2014 Apr 13;15(1):124. doi: 10.1186/1745-6215-15-124.