

RESEARCH

Open Access



Failure prediction using personalized models and an application to heart failure prediction

Asim Roy^{1*}, Charles Bruce², Phillip Schulte², Lyle Olson² and Manasa Pola¹

* Correspondence: Asim.Roy@asu.edu

¹Arizona State University, Tempe, AZ, USA

Full list of author information is available at the end of the article

Abstract

Background: To reduce disruptions of processes and the cost of maintenance, predicting the onset of failure (or a similar event) of a physical system (or components of a physical system) has become important. Prediction of onset of failure would allow appropriate corrective actions at the right time. In this paper, we present a method to predict the “onset” of failure (the start of a degradation process or similar types of events) of a physical system that minimizes data collection and personalizes it for the physical system. The method applies to situations where one monitors the operating characteristics of the physical system at regular time intervals by means of attached sensors and other measurement instruments. It creates a model of the physical system, during normal operations, using the time-series data produced by the sensors and measurement instruments. However, it does not create or use any time-series models. It simply examines the distribution of time-series data across different time periods. It uses this model of normal operations in subsequent time periods to monitor the physical system for deviations from normality.

Results: We illustrate this method with an application to predict the “onset” of subsequent decompensated heart failures for patients already treated for a heart failure at a hospital. As part of an NIH study, these heart failure patients received two ECG patches, an accelerometer and a bio-impedance measurement device for regular monitoring for a period after their release from the hospital.

Conclusions: When dealing with non-homogenous, disparate physical systems, personalized models can be better predictors of a phenomenon compared to generalized models based on data collected from an assortment of such physical systems. In medicine such models can be a powerful addition to the set of medical diagnostic tools. And such personalized models can be built rather quickly without waiting for extensive data collection.

Keywords: Machine learning, Failure prediction, Heart failure prediction, Personalized models of physical systems



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

In general, the first step to create a model of any phenomenon is data collection. For example, to predict fraud, an organization would collect data on different fraud and non-fraud cases. And to predict breast cancer from biopsies, one would collect data on different cases where the tumors were either benign or malignant. In general, the basic idea in machine learning is to build models from a diverse set of cases so that the models can “generalize” and accurately account for the variety of cases that exemplify the phenomenon. However, the attempt to generalize from a very diverse set of cases can sometimes be problematic and can result in models that may not be very accurate in their predictions. Such diversity of cases arise often in the medical field because human bodies become very dissimilar physical systems over time. For example, one would find patients with similar medical histories exhibiting different medical conditions making it difficult to identify common medical profiles for certain medical conditions. (Creating such common profiles is the task of generalization in machine learning.) For instance, heart failure patients often have very different medical histories and, thus, makes it difficult to build highly accurate profiles for them. Hence, generalization and accuracy of prediction suffer with these kinds of phenomena.

It is often possible to redefine the problem, especially when system degradation prediction is of interest, by simply focusing on the data produced by an individual physical system. In this scenario, one simply would build models of an individual physical system that, in effect, would define its normal operating characteristics. Such models would no longer need to compare an individual system’s behavior with other similar systems. They would therefore not require extensive data collection for the purpose of “generalization.” In addition, they could be built quickly, almost instantaneously in some cases, and be able to capture the idiosyncrasies of a particular system.

Here is an example of such a situation. In one particular NIH study of decompensated heart failure (DHF) [40], further discussed later in this paper, DHF patients, after their first heart failure treatment, received a package of devices for remote patient monitoring (RPM) on their discharge from the hospital. The package included two ECG patches, an accelerometer and a bio-impedance measurement device. The NIH study collected data from individual patients with the RPM devices in order to predict the “onset” of next decompensated heart failure for such patients. However, given the diversity of the patient population in this study, any population-based predictive model, based on RPM and other medical data of the patients, would not be very accurate. In this paper, we redefine the prediction problem in such situations and show that a personalized model for each individual patient, based simply on the RPM data, would be much more accurate in its prediction of the “onset” of DHF.

In this paper, we propose a method for personalized modeling of a physical system for failure prediction (or, to be precise, predict the start of the degradation process of a system) based on time-series data produced by sensors and other measurement instruments. We then show the application of this method to predict the “onset” of subsequent decompensated heart failure of three

patients from the NIH study. Heart failures are generally a slow degradation process and are similar to slow failure processes of many other physical systems. Thus, the method can be applied to failure prediction of machinery and production processes with similar characteristics. Although the proposed method uses time-series data produced by sensors and other instruments, it does not actually construct or use any time-series models. It simply examines the distribution of time-series data across specified time cycles to make predictions about the “onset” of failure. Prediction of the “onset” of slow degradation processes is also not strictly anomaly detection but is more about trend analysis. The advantage of personalized modeling is that it does not require large amounts of data collection about other similar systems. And for situations where it is difficult to generalize from diverse population characteristics, personalized models can be far more accurate.

Literature review - remaining useful life (RUL) approach to failure prediction

Prediction of remaining useful life (RUL) of a physical system, before it is likely to require repair or replacement, is widely used for predictive maintenance. RUL methods, also called prognostic methods, are broadly classified as two types: (1) data-driven methods and (2) model-driven methods [28]. Data-driven approaches use statistical and machine learning methods on historical failure data of similar systems to build prognostic models. Traditional data-driven approaches include autoregressive (AR) and threshold autoregressive models [3, 42], projection pursuit models [12] and multivariate adaptive regression splines models [13]. More recent data-driven approaches include a variety of neural network models [2, 19, 20, 27, 37, 41]. Model-based approaches use models that describe the physics of the system [1, 7, 30].

Liu et al. [28] proposed a data-model fusion framework for system state prognostics. Nystad et al. [34] investigated the problem of estimating the remaining useful life using stochastic lifetime models and considered randomly distributed failure thresholds. Gola and Nystad [16] combined a condition monitoring system that provides reliable calculations of the actual erosion state of a choke valve during its operation with a lifetime RUL model. In effect, it adjusts a lifetime RUL model using real-time system monitoring data. Lei et al. [26] proposed a model-based method for predicting RUL that has two modules: (1) an indicator constructor which fuses mutual information from multiple features and properly correlates to the degradation processes of machinery, and (2) a RUL predictor that uses a particle filtering-based algorithm.

Casotto et al. [3] constructs a model of normal behavior of a system by fusing multiple sensor signals during its normal operation. Using this model, they then monitor the behavior of the system for degradation, expressed by drift of signals away from the normal. They fit an Autoregressive (AR) model to time series signals of multiple sensors and extract the Power Spectral Density (PSD) peaks of individual sensor readings from the roots of the corresponding AR model characteristic equation. These features, extracted during normal operations, are saved as the model of normal operating behavior. Their overall idea is very similar to ours

in the sense that they build a personalized model of a process or system and do not use any historical data. They predict failure (degradation) by observing drift in the behavior of the system. However, our signal fusion method is completely different from their method.

Qiu et al. [37] present a prognostic method that includes a wavelet filter-based method for signal de-noising, weak signature enhancement for fault identification and two Self Organizing Maps (SOMs) for performance assessment and degradation detection to detect defects at an early stage in rolling element bearings. The approach, in principle, also constructs a personalized model by training the SOMs with normal operating data of a system. We also use SOMs in our method as discussed later. However, the way they use SOMs to track system degradation is completely different from ours. Plus, we do not use any wavelet-based method for signal denoising.

Literature review – predictive models in personalized medicine

There are many ongoing efforts to use predictive models in personalized medicine that rely more on individual patient's medical history and data. We cite a few here. Nevins et al. [32] provide a model building framework to combine multiple types of data, both genomic and clinical, of individual patients to better predict breast cancer treatment outcomes. Esteban et al. [11] model the clinical evolution of individual patients, which usually is composed of thousands of events such as ordered tests, lab results and diagnoses, to predict the future sequence of events for clinical decisions. Their specific work related to patients with kidney failure who either obtained an organ transplant or were still waiting for one. They construct neural network models in combination with embeddings in a clinical context. Che et al. [6] propose a model to personalize prediction of Parkinson's disease progression. It learns patient similarity from longitudinal and multi-modal patient records with a Recurrent Neural Network (RNN) architecture. Ng et al. [33] propose a method to build personalized predictive models to predict the onset of diabetes in patients. They show that personalized models (they use logistic regression models), built using data of similar patients, outperform global models that use data for all patients. Jiang et al. [21] propose a new method to calibrate a predictive model for individual patients; it uses a similar group of patients for calibration. Lee et al. [24] used a cosine-similarity based patient similarity metric to identify similar ICU patients from an ICU database and to dynamically build 30-day mortality prediction models for individual patients. Their model outperformed the general, one-size-fits-all model for such predictions. Lee [25] uses the same idea of finding similar patients but uses a random forest model instead of cosine similarity, to build a customized model for an individual patient. Many of these methods, particularly the ones that use a similar group of patients, use notions from the collaborative filtering method used in personalized recommendation systems in e-commerce. In summary, to build predictive models for personalized medicine, these are the kinds of approaches being used.

Literature review - predicting heart failure

Ross et al. [39] performed a systematic review of studies evaluating patient characteristics associated with hospital readmission for heart failure (HF). Rahimi et al. [38] reviewed the literature for risk prediction models for patients with

Table 1 Different types of biosensors and the biosignals generated by them. From Pantelopoulos & Bourbakis [35] and with permission of IEEE

Type of Bio-signal	Type of Sensor	Description of measured data
Electrocardiogram (ECG)	Skin/Chest electrodes	Electrical activity of the heart (continuous waveform showing the contraction and relaxation phases of the cardiac cycles)
Blood pressure (systolic & diastolic)	Arm cuff-based monitor	Refers to the force exerted by circulating blood on the walls of blood vessels, especially the arteries
Body and/or skin temperature	Temperature probe or skin patch	A measure of the body's ability to generate and get rid of heat
Respiration rate	Piezoelectric/piezoresistive sensor	Number of movements indicative of inspiration and expiration per unit time (breathing rate)
Oxygen saturation	Pulse Oximeter	Indicates the oxygenation or the amount of oxygen that is being "carried" in a patient's blood
Heart rate	Pulse Oximeter/skin electrodes	Frequency of the cardiac cycle
Perspiration (sweating) or skin conductivity	Galvanic Skin Response	Electrical conductance of the skin is associated with the activity of the sweat glands
Heart sounds	Phonocardiograph	A record of heart sounds, produced by a properly placed on the chest microphone (stethoscope)
Blood glucose	Strip-base glucose meters	Measurement of the amount of glucose (main type/source of sugar/energy) in blood
Electromyogram (EMG)	Skin electrodes	Electrical activity of the skeletal muscles (characterizes the neuromuscular system)
Electroencephalogram (EEG)	Scalp-placed electrodes	Measurement of electrical spontaneous brain activity and other brain potentials
Body Movements	Accelerometer	Measurement of acceleration forces in the 3D space

heart failure and identified the most consistently reported independent predictors of risk across models. Kansagara et al. [22] assessed the performance of the prevailing set of models that predict hospital readmission. Mortazavi et al. [31] used data from Telemonitoring to Improve Heart Failure Outcomes trial [5] to compare the effectiveness of various machine learning methods to predict 30- and 180-day all-cause readmissions and readmissions because of heart failure. Choi et al. [8] used a recurrent neural network (RNN) model to predict the initial diagnosis of heart failure. The RNN model exploited temporal relations among events in electronic health records (EHRs) and used 3884 HF records of primary care patients. They also compared it with other models such as logistic regression, neural network, support vector machine and K-nearest neighbor classifiers. Wang et al. [44] used structured and unstructured data from electronic health records to predict the onset of heart failure and varied the prediction window from 60 to 720 days before heart failure diagnosis. They used a total of 1684 heart failure records of primary care patients. Dai et al. [9] used five machine learning models to predict heart-related hospitalizations. They used EHR data of patients with heart disease from a large urban hospital in Boston.

The rest of the paper is organized as follows. The "Method" section provides an overview of the nature of the proposed personalized model and the data collected and used to predict the "onset" of heart failure using such a model. The "Results" section has the detailed steps of the proposed method. The

Table 2 Slicing across time-series to create each data point for clustering

Data Point	Time t	X(1)	X(2)	X(3)	X(4)
D1	t = 1	1	2	3	4
D2	t = 2	5	6	7	8
D3	t = 3	0	2	4	6
D4	t = 4	3	5	7	9
D5	t = 5	2	4	6	8
D6	t = 6	1	3	5	7
D7	t = 7	5	3	1	0
D8	t = 8	2	4	6	8
D9	t = 9	5	4	3	2

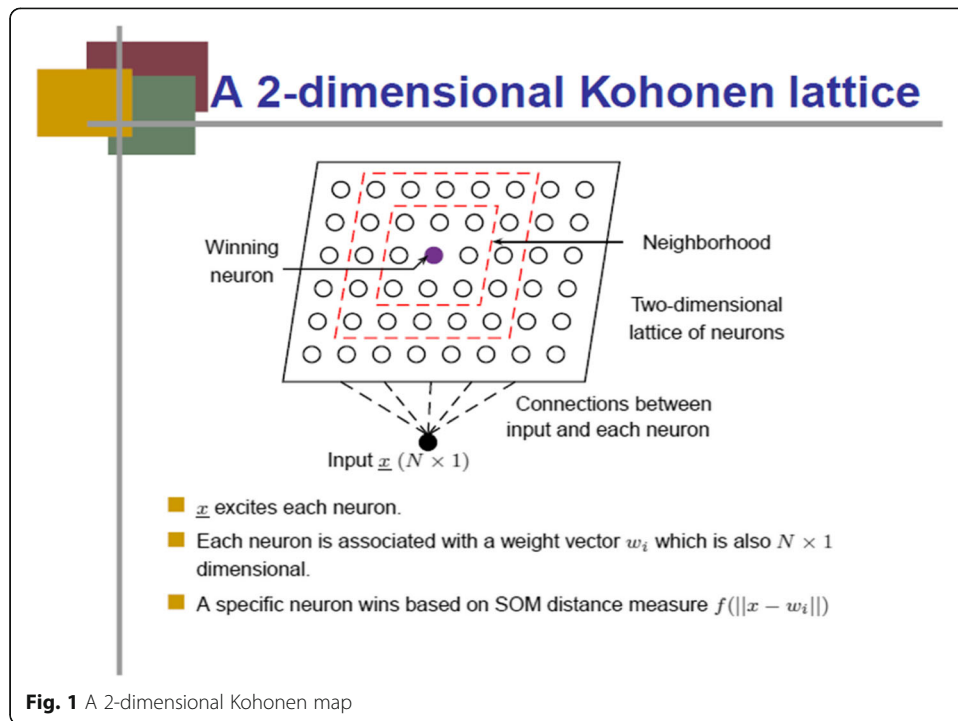


Fig. 1 A 2-dimensional Kohonen map

“Discussion” section shows the application of the method to predicting the “onset” of decompensated heart failure for three patients in the NIH study. The “Conclusions” section provides some concluding remarks.

Method

The proposed method is based on time-series data

The proposed method is about personalizing a model for a physical system to predict the “onset” of its degradation process. Such a model is solely based on

Table 3 Count of data points at each node (cluster) of a 4 × 3 (12 node) Kohonen SOM after training

Node Number	Count of data points assigned to the node
Node 1	17
Node 2	14
Node 3	87
Node 4	9
Node 5	16
Node 6	5
Node 7	26
Node 8	6
Node 9	7
Node 10	63
Node 11	23
Node 12	15

Table 4 Count of data points in the consolidated clusters of the 4 × 3 (12 node) Kohonen SOM

Nodes combined	Count of data points in the consolidated clusters
2, 3, 5, 6, 7	148
1, 4	26
8, 9, 10, 11, 12	114

data recorded at certain time intervals by a monitoring system for the physical system. It does not use or depend on any prior knowledge about such physical systems. For plants and machinery, such a monitoring system typically would consist of different types of sensors attached to them, such as the ones to measure vibration, pressure and temperature. For the heart failure case study discussed in this paper, the remote patient monitoring (RPM) system consisted of two ECG recording patches, an accelerometer to monitor the patient’s activity and a bio-impedance measuring device.

The sensors of a monitoring system can generate data at different frequencies. For example, ECG patches generate data every few milliseconds while blood pressure and weight might be recorded only a few times a day. The data generated by a sensor is essentially time-series data. When multiple sensors generate data at different frequencies, the frequencies need to be aligned for modeling purposes. There are different ways to align slow and high frequency time-series data. For example, temperature or pressure, if they are measured too frequently, can be averaged over a time interval to produce a lower frequency time-series. In the same way, if weight is measured infrequently, the same weight value can be used at subsequent time points until a new weight is recorded.

Sensor data is usually collected by external devices which then can extract additional information from them. The extracted features, in turn, define additional time-series. In general, multiple sensors and downstream devices collectively produce streaming time-series data. Thus, we can define a physical system by the characteristics of such a collection of different time-series. One way to predict the “onset” of degradation (the “onset” of failure) is to model each time-series from the data generated during normal operations of the physical system, then monitor the physical system using these time-series models and look for deviations from the normal operating mode [3]. However, we do not create any time-series models in our approach. Instead, we look for changes in the distribution of time-series data over time and then isolate one or more time-series that potentially is causing the degradation.

Table 5 Cluster sizes in consolidated form in three different Kohonen SOMs

Cluster sizes in consolidated form			Average
4 × 3 SOM	3 × 3 SOM	5 × 2 SOM	
148	133	151	144
26	33	28	29
114	122	109	115

Table 6 Four highest ranking features for three individual patients

	4 highest ranking features			
Patient A	SVC	SinTACHY	Respiration Rate	NSR
Patient B	PVC	SVC	SinTACHY	NSR
Patient C	Afib normal	Unclassified rythm	Respiration Rate	NSR + IVCD

On the decompensated heart failure study

Since we will illustrate each step of our method with a few heart failure cases, we present some background information on the NIH supported decompensated heart failure study at Mayo Clinic that provided the data for these cases. The NIH study [40] used the BodyGuardian remote monitoring system from Preventice [36]. The BodyGuardian remote health management system is an FDA 510 approved device used for remote monitoring of cardiac patients. It has a front-end that includes an adhesive snap-strip body sensor (BodyGuardian) with built-in electrodes that measure ECG signals and bio-impedance. It also has a 3-way accelerometer. Overall, the system measures heart rate, ECG, respiration rate (RR) and activity. It also communicates with off-body sensors such as a BP cuff and scale to incorporate BP and weight data. In addition, it solicits symptoms from the user thus acting as an event recorder and recording simultaneous physiologic data. It wirelessly transmits all data to a central data analysis hub.

From ECG signals, bio-impedance measurements and accelerometer data, BodyGuardian derives 56 features. It classifies activity level in the range 0 to 100, which is then binned into 10 ranges. From the activity data, it derives three basic body positions: lying, leaning and standing. We excluded activity level and body position data from our model. From ECG data, it extracts a number of features including: PVC (premature ventricular complex), SVC (supraventricular complex), NSR (normal sinus rhythm), Unclassified Rhythm, SinTachy (sinus tachycardia), SinBrady (sinus bradycardia), IVCD (interventricular conduction delay), Mobitz 1 and 2, AV Block (atrioventricular block), PJC (premature junctional complex), PAC (premature atrial contractions), SVTA (supraventricular tachyarrhythmia), AFib (atrial fibrillation - slow, normal, rapid), IVR (idioventricular rhythm), VT (ventricular tachycardia), VF (ventricular fibrillation), minimum heart rate and maximum heart rate. The data also includes blood pressure, respiration rate and weight.

We decided to have the data averaged every 5 min for modeling purposes, although the data is available on a finer time scale. In effect, we are observing the patient every 5 min. When recorded continuously during a day, one gets 288 observations. We decided to create a model using BodyGuardian data for a single day and, then, use that model to track changes in the patient’s physiological profile on subsequent days. Since the physiological measurements vary

Table 7 Cluster data distribution for patient A on subsequent days of monitoring

	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10
Cluster 1	144	282	288	252	169	119	288	288	280	158	288
Cluster 2	29	0	0	0	0	0	0	0	0	0	0
Cluster 3	115	2	0	0	0	1	0	0	1	1	0
TOTAL DATA POINTS	288	284	288	252	169	120	288	288	281	159	288

Table 8 Cluster data distribution in percentage for patient A on subsequent days of monitoring

	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10
Cluster 1	50%	99%	100%	100%	100%	99%	100%	100%	100%	99%	100%
Cluster 2	10%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Cluster 3	40%	1%	0%	0%	0%	1%	0%	0%	0%	1%	0%
TOTAL DATA POINTS	288	284	288	252	169	120	288	288	281	159	288

during the course of a day, our approach is to model the distribution of the physiological data during the day.

Since, in this study, Mayo Clinic provides a patient with the BodyGuardian device only after a heart failure treatment, in general, we create a model for a patient after a full day of recording following discharge from the hospital. One can construct models using data over several days following discharge from the hospital, but there is a risk in the sense that there could be onset of decompensation very soon after discharge. The model is meant to reflect the physiological state of the patient before the onset of a subsequent decompensation.

We do not use any clinical data of patients in our models. Nor do we use data of other patients to build each individualized patient model. This concept of creating a personalized model based predominantly on data generated by wearable biosensors is new and may have wide applicability in many situations. Table 1 (from [35]) shows some typical biosensors in use today and the biosignals generated by them. There are many factors driving the growth in usage of such wearable devices including: an aging population worldwide, the need to reduce hospital and emergency visits, and the need to monitor and manage chronic diseases remotely.

Results

Here are the steps of the proposed method:

Step 1: Select a set of time-series to use for modeling the physical system. This set may include both original sensor/device measurements (e.g. weight, blood pressure of a patient) and derived measurements (e.g. QRS complex, atrial premature complexes from an ECG).

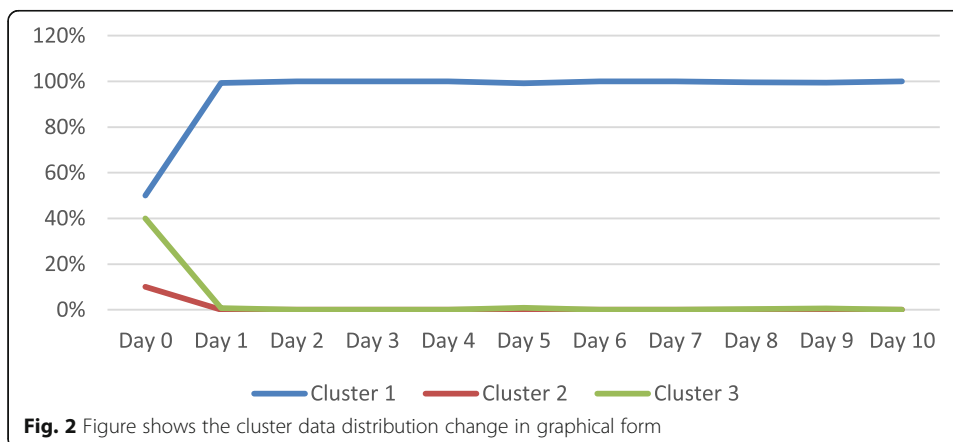


Table 9 Average feature values (or counts of feature values) for patient A on subsequent days of monitoring

	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10
SVC	1323	153	21	25	20	21	50	25	18	9	21
SinTACHY	495	86	8	21	13	22	48	18	18	15	10
NSR	375	67	3	13	7	8	13	4	1	2	1
Respiration Rate	16.20	13.00	13.00	13.00	13.00	12.58	8.53	13.38	16.27	13.51	14.79

Step 2: Determine a time interval P for recording time-series data and record the average values of the selected time-series within that time interval from streaming data. For example, the time interval P can be a millisecond, a minute or 5 min and would depend on how frequently one should observe the physical system for degradation or “onset” of failure or certain other events.

Step 3: Suppose we record the average values of N time-series in each time interval of duration P . Suppose T is the total number of observed time intervals during normal operations of the physical system, and $X(t, i)$, $i = 1 \dots N$, $t = 1 \dots T$, the average value of the i^{th} time-series at the t^{th} time interval. Thus, $X(t, i)$, $i = 1 \dots N$, is an observation of the physical system across all N time-series at time t . Each time-series i , $i = 1 \dots N$, represents a feature in the dataset.

Step 4: Using T collected observations of N time-series, create a clustering model to record the distribution of time-series data in that observation period $T * P$.

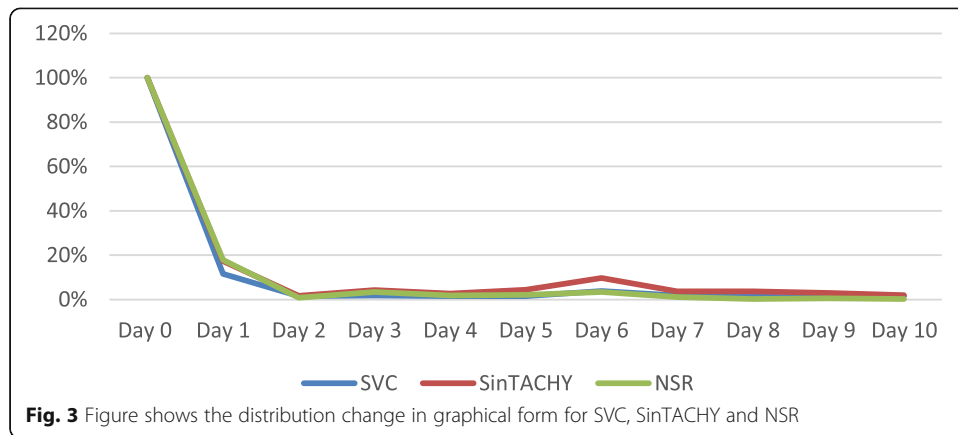
For example, one can monitor a heart failure patient every 5 min ($P = 5$ min) for a day ($T = 12 * 24 = 288$ durations of 5 min) to create a model of the patient. One can use a clustering model, such as Kohonen Self-Organizing Map (SOM) [23], to model the distribution of the time-series data. The resulting clustering model represents normal behavior of the physical system and shows the distribution of time-series values normally observed within a period of time $T * P$.

Step 5: Obtain a ranking of the features (time-series) using any unsupervised feature ranking method or from the clustering method.

Step 6: Monitor the physical system in subsequent periods of length $T * P$ using the model created in Step 4 and data from the N time-series. The monitoring tracks any changes in the distribution of time-series data. Monitoring essentially means passing the data through the clustering model and assigning each datapoint to the closest cluster. To understand the factors that cause a change from the normal behavior of the physical system, track some of the highest-ranked features (time-series) found in Step 5.

Table 10 Average feature values (or counts of feature values) as percentage of day 0 values for patient A on subsequent days of monitoring

	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10
SVC	100%	12%	2%	2%	2%	2%	4%	2%	1%	1%	2%
SinTACHY	100%	17%	2%	4%	3%	4%	10%	4%	4%	3%	2%
NSR	100%	18%	1%	3%	2%	2%	3%	1%	0%	1%	0%
Respiration Rate	100%	80%	80%	80%	80%	78%	53%	83%	100%	83%	91%



For example, one can create a model of a heart failure patient, using remote monitoring data, the day after discharge from a hospital and that model would then define “normal functioning” of the patient after the hospital treatment. One would then use that model to monitor the patient on subsequent days using the N time-series data for each day.

Step 7: If significant deviation in the distribution of the time-series data and in the trajectory of some of the highest ranked features occurs during the monitoring period, notify/alert a supervisory system of the change from normal.

Discussion

Application of the method to predict the “onset” of decompensated heart failure

The “Method” section explained Steps 1 and 2 of the method for the heart failure case study: (1) selection of time-series for modeling the patients, and (2) determination of the time interval P for time-series measurements. Step 3 sets up the data for clustering in Step 4. Here is an example of proper structuring of the data for clustering. Table 2 shows four time series – $X(1)$, $X(2)$, $X(3)$ and $X(4)$ – and 9 values for each recorded over 9 time intervals. The first data point for clustering is $D1 = [1, 2, 3, 4]$ and the last data point is $D9 = [5, 4, 3, 2]$. In essence, for the clustering model, each time-series corresponds to a feature and we slice the time-series data across time.

Step 4 of the method - clustering

In Step 4, the method creates clusters to characterize the distribution of the data during normal operations of the physical system. In our implementation, we used Kohonen’s Self Organizing Map (SOM) [23] for clustering. However, one can use other methods

Table 11 Cluster data distribution for patient B on subsequent days of monitoring

	Day 0	Day 1	Day 2	Day 3	Day 4	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15
Cluster 1	75	48	64	78	52	220	61	15	158	189	201	242	241	173
Cluster 2	31	110	69	67	60	57	69	36	125	97	83	41	42	115
Cluster 3	140	75	130	109	150	0	157	237	3	0	0	0	0	0
TOTAL DATA POINTS	246	233	263	254	262	277	287	288	286	286	284	283	283	288

Table 12 Cluster data distribution in percentage for patient B on subsequent days of monitoring

	Day 0	Day 1	Day 2	Day 3	Day 4	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15
Cluster 1	30%	21%	24%	31%	20%	79%	21%	5%	55%	66%	71%	86%	85%	60%
Cluster 2	13%	47%	26%	26%	23%	21%	24%	13%	44%	34%	29%	14%	15%	40%
Cluster 3	57%	32%	49%	43%	57%	0%	55%	82%	1%	0%	0%	0%	0%	0%
TOTAL DATA POINTS	246	233	263	254	262	277	287	288	286	286	284	283	283	288

as well such as K-means clustering [29]. Ghesmoune et al. [15] has a comprehensive survey of data stream clustering methods.

A two-dimensional rectangular grid defines a Kohonen SOM as shown in Fig. 1. For example, a grid of size 4×3 has 12 nodes.

As part of this step, since randomness is inherent in any clustering method, we train Kohonen SOMs of different sizes (e.g. 3×3 , 4×3 , 5×2) to verify the data distribution. In addition, after training the Kohonen SOMs, we combine nearby nodes (nearby clusters) of each SOM to verify the consistency of the data distribution across the SOMs. For example, suppose we train a Kohonen SOM of size 4×3 with data from 1 day of remote monitoring of a heart failure patient (which produces 288 data points) and let us suppose the resulting 12 node SOM produces the clusters shown in Table 3. We then combine the closest nodes (clusters) to produce more consolidated clusters. Table 4 shows the result of such a consolidation that produced three clusters from the 4×3 SOM clusters in Table 3. Table 5 shows the consolidated data distribution (cluster sizes) from three different Kohonen SOMs of sizes 4×3 , 3×3 and 5×2 . The consolidation process produces a data distribution that is invariant to the Kohonen SOM size.

Step 5 – feature ranking

One can get feature rankings by a variety of means [4, 10]. Table 6 shows the four highest-ranking features (out of 42 features) for three different heart failure patients. Here the features correspond to the time-series. Note that the top-ranking features

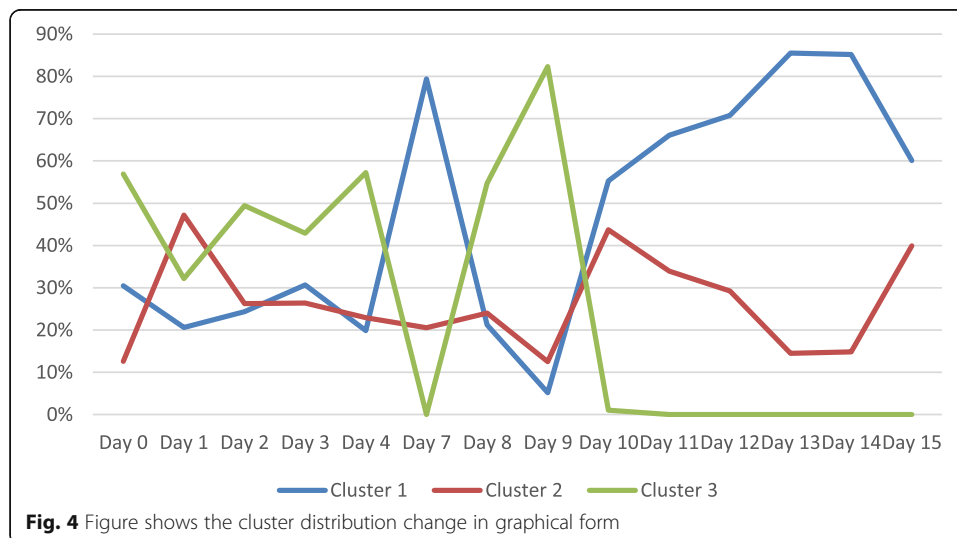


Fig. 4 Figure shows the cluster distribution change in graphical form

Table 13 Counts of feature values for patient B on subsequent days of monitoring

	Day 0	Day 1	Day 2	Day 3	Day 4	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15
PVC	514	456	285	383	713	485	838	1344	845	886	762	545	441	896
SVC	25	9	17	20	51	7	33	35	35	30	37	34	31	38
SinTACHY	0	1	1	1	1	1	5	9	4	3	1	0	1	5
NSR	173	189	173	221	142	318	315	288	444	311	121	116	204	331

(time-series) are different for the three patients and essentially characterize each patient individually along with the associated data distribution.

Step 6 – individual patient monitoring using the clustering model and the ranked features

Here we show the application of personalized clustering models and the ranked features to three different decompensated heart failure patients from the NIH supported study at Mayo Clinic. We will label the patients as A, B and C. We discuss each case separately below.

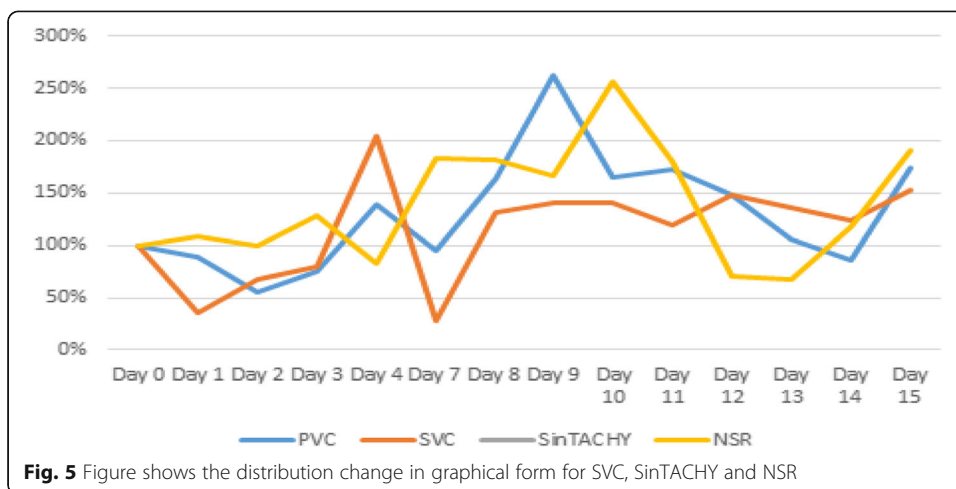
Patient A – readmitted to the hospital 17 days after hospital discharge

Patient A was a 76 years old male who had hypertension and diabetes. Remote monitoring of the patient, after the first heart failure treatment, started 10 days after discharge from the hospital. We created a personalized model on the first day of data collection (Day 0) and used it to monitor the patient on subsequent days. Assignment of data on subsequent days to the 3 Day 0 clusters (Clusters 1, 2 and 3) was on a nearest cluster basis. Table 7 shows the cluster data distribution of time-series data on subsequent days. Table 8 shows these values as a percentage of the total recordings each day. Figure 2 shows the cluster data distribution change in a graphical form. As one can observe, the cluster distribution changed drastically from Day 1. The hospital readmitted the patient for heart failure on Day 11. A fair estimate would be that the “onset” of decompensation started around Day 1 and confirmed by similar observations on subsequent days.

Table 9 shows the average values (for Respiration Rate) or counts (for SVC, SinTACHY and NSR) for each day of the four highest ranked features for patient A. Table 10 shows these values as a percentage of the values on Day 0. Figure 3 shows the distribution change in a graphical form for SVC, SinTACHY and NSR, the three highest ranked features. These physiological features provide additional information - with their upward, downward or fluctuating trends - to support a judgment about the “onset” of decompensation and about potential treatment.

Table 14 Counts of feature values in percentage of day 0 values for patient B on subsequent days of monitoring. SinTACHY was skipped because it’s Day 0 value is 0

	Day 0	Day 1	Day 2	Day 3	Day 4	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15
PVC	100%	89%	55%	75%	139%	94%	163%	261%	164%	172%	148%	106%	86%	174%
SVC	100%	36%	68%	80%	204%	28%	132%	140%	140%	120%	148%	136%	124%	152%
SinTACHY														
NSR	100%	109%	100%	128%	82%	184%	182%	166%	257%	180%	70%	67%	118%	191%



From the accelerometer data, BodyGuardian calculates whether the patient is standing, leaning or lying. Data for Patient A showed that he was standing or leaning almost 100% of the time from Day 1. Cardiologists generally know that decompensated heart failure patients cannot lie supine because of breathing problems. Thus, body position is another indicator of heart failure. In this case, multiple indicators confirm the “onset” of decompensated heart failure on Day 1.

Patient B – readmitted to the hospital 17 days after hospital discharge

Patient B was a 73 years old female who had hypertension, diabetes and other ailments. Remote monitoring of the patient, after the first heart failure treatment, started the day after discharge from the hospital. We created a personalized model on the first day of data collection (Day 0) and used it to monitor the patient on subsequent days. Assignment of data on subsequent days to the 3 Day 0 clusters (Clusters 1, 2 and 3) was on a nearest cluster basis. We dropped Days 5 and 6 because there was minimal monitoring on those days. Table 11 shows the cluster data distribution of time-series data on subsequent days. Table 12 shows these values as a percentage of the total recordings each day. Figure 4 shows the cluster distribution change in a graphical form.

As one can observe, the cluster distribution fluctuates until Day 10 and then cluster 1 size increases from Day 11 to 14. There was minimal data collection on Day 16. The hospital readmitted the patient for weakness, tiredness, abdominal pain and other problems on Day 17. From the cluster distribution in Fig. 4, one can infer that the “onset” of the medical problems most likely started on Day 11 when the cluster distribution started to change.

Table 15 Cluster data distribution for patient C on subsequent days of monitoring

	Day 0	Day 2	Day 3	Day 4	Day 12	Day 15	Day 17	Day 18
Cluster 1	89	82	100	79	142	118	103	123
Cluster 2	128	124	103	165	88	108	84	95
Cluster 3	44	76	83	44	39	21	101	70
TOTAL DATA POINTS	261	282	286	288	269	247	288	288

Table 16 Cluster data distribution in percentage for patient C on subsequent days of monitoring

	Day 0	Day 2	Day 3	Day 4	Day 12	Day 15	Day 17	Day 18
Cluster 1	34%	29%	35%	27%	53%	48%	36%	43%
Cluster 2	49%	44%	36%	57%	33%	44%	29%	33%
Cluster 3	17%	27%	29%	15%	14%	9%	35%	24%
TOTAL DATA POINTS	261	282	286	288	269	247	288	288

Table 13 shows the counts for each day of the four highest ranked features for patient B for the corresponding days in Table 11. Table 14 shows these values as a percentage of the values on Day 0. Figure 5 shows the distribution change in a graphical form for PVC, SVC and NSR, three of the four highest ranked features. These physiological features provide additional information to support a judgment about the “onset” of decompensation-related problems and about potential treatment. For this patient, accelerometer data shows that she was standing or leaning on the average 85% of the time every day from Day 10 to 15 compared to 69% average on prior days. In this case, multiple indicators confirm the “onset” of heart failure related problems on about Day 10 or 11.

Patient C – not readmitted to the hospital during the monitoring period

Patient C was a 75 years old male who had hypertension and was hospitalized for COPD or asthma. Remote patient monitoring started about 3 weeks after hospital discharge, but the patient used the devices only part of the day on many of the days. We created a personalized model on the first day of reasonably good data collection and used it to monitor the patient on subsequent days. Table 15 shows the cluster data distribution of time-series data on subsequent days where Day 0 corresponds to the day of model creation. Table 16 shows these values as a percentage of the total recordings each day. Assignment of data on subsequent days to the 3 Day 0 clusters was on a nearest cluster basis. We dropped several days because there was minimal monitoring on those days. Figure 6 shows the cluster distribution change in a graphical form. As one can observe, the data distribution remains fairly steady on the days that had good

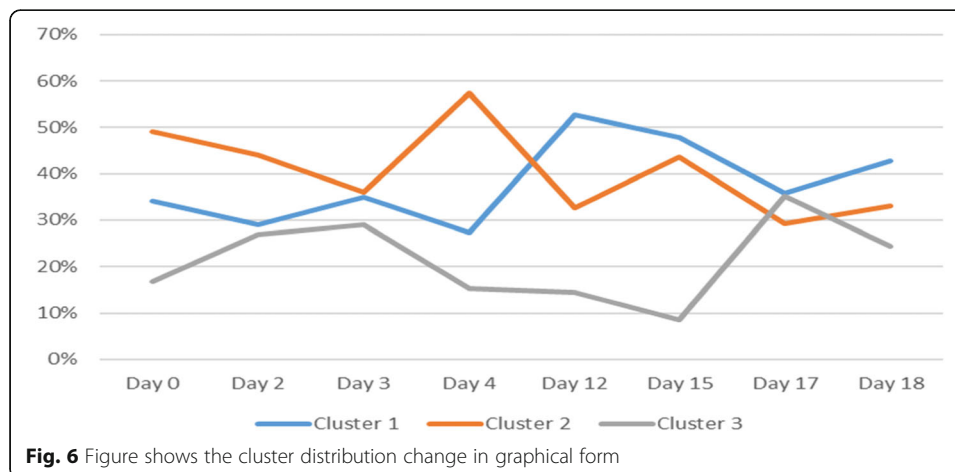


Fig. 6 Figure shows the cluster distribution change in graphical form

Table 17 Average feature values (or counts of feature values) for patient C on subsequent days of monitoring

	Day 0	Day 2	Day 3	Day 4	Day 12	Day 15	Day 17	Day 18
Afib normal	132	116	147	133	121	104	84	74
Unclassified rhythm	489	292	754	251	170	322	156	136
Respiration Rate	18.03	19.9	19.94	20.26	18.27	16.03	20.26	15.32
NSR+IVCD	45	55	45	4	6	29	42	45

monitoring. There was no re-admittance of the patient to the hospital during the monitoring period.

Table 17 shows the counts for each day of the four highest ranked features for patient C for the corresponding days in Table 15. Table 18 shows these values as a percentage of the values on Day 0. Figure 7 shows the distribution change in a graphical form for Afib normal, Unclassified rhythm, Respiration Rate and NSR + IVCD, the four highest ranked features. These physiological features provide additional information - with their upward, downward or fluctuating trends - to support a judgment about the “onset” of decompensation-related problems and about potential treatment.

For this patient, accelerometer data shows that he was standing or leaning on the average 66% of the time during a day for the days in Table 15, which means he was sleeping well. In this case, there were no significant indicators predicting the “onset” of a major medical event. In fact, the patient did not return to the hospital during the monitoring period of about a month.

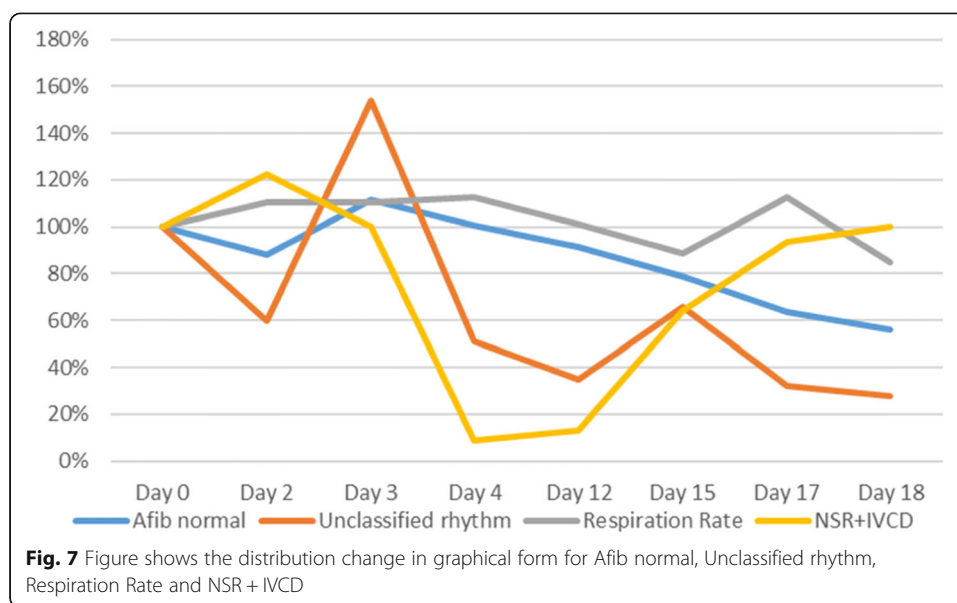
Conclusions

In medicine, where we encounter non-homogeneous, disparate physical systems, generalized predictive models based on a population of patients may be less accurate compared to personalized models, as demonstrated by Ng et al. [33]. Generalized models are also expensive to build because they require extensive data collection. And, because they require extensive data collection, there is a significant wait time for generalized models. On the other hand, a personalized model of the type proposed here can be built quickly because it needs data from just one individual patient whose model is to be built.

Note that the method proposed here does not use any information from electronic medical records (EMR) or electronic health records (EHR) of a patient or find similar patients for comparison. However, many of the other personalized predictive models, such as those of Che et al. [6], Ng et al. [33], Jiang et al. [21], Lee et al. [24] and Lee [25], use EMR/EHR data of similar patients to build their personalized predictive models.

Table 18 Average feature values (or counts of feature values) as percentage of day 0 values for patient C on subsequent days of monitoring

	Day 0	Day 2	Day 3	Day 4	Day 12	Day 15	Day 17	Day 18
Afib normal	100%	88%	111%	101%	92%	79%	64%	56%
Unclassified rhythm	100%	60%	154%	51%	35%	66%	32%	28%
Respiration Rate	100%	110%	111%	112%	101%	89%	112%	85%
NSR+IVCD	100%	122%	100%	9%	13%	64%	93%	100%



In general, in the context of personalized medicine [18], personalized models can be a powerful addition to the set of medical diagnostic tools. Personalized models offer more accurate predictions and information compared to generalized models based on a whole population of patients,

Now that ECG measurements are becoming common in many consumer and medical-grade wearable devices (such as Apple Watch Series 4 or later [14], Quardio-Core [43], AliveCor [17], Samsung S-Patch 3 [45]), personalized model building on wearable devices may not be far away.

Abbreviations

NIH: National Institutes of Health; IEEE: Institute of Electrical and Electronics Engineers; RO1: NIH Research Project Grant; ECG: Electrocardiogram; DHF: Decompensated heart failure; RPM: Remote patient monitoring; RUL: Remaining useful life; AR: Autoregressive; PSD: Power Spectral Density; SOM: Self Organizing Maps; RNN: Recurrent Neural Network; HF: Heart failure; EHR: Electronic health records; EMR: Electronic medical records; FDA: Food and Drug Administration; RR: Respiration rate; BP: Blood pressure; ICU: Intensive care unit; PVC: Premature ventricular complex; SVC: Supraventricular complex; NSR: Normal sinus rhythm; SinTachy: Sinus tachycardia; SinBrady: Sinus bradycardia; IVCD: Interventricular conduction delay; AV Block: Atrioventricular block; PJC: Premature junctional complex; PAC: Premature atrial contractions; SVTA: Supraventricular tachyarrhythmia; AFib: Atrial fibrillation - slow, normal, rapid; IVR: Idioventricular rhythm; VT: Ventricular tachycardia; VF: Ventricular fibrillation; QRS: Is the combination of three of the graphical deflections seen on a typical ECG; COPD: Chronic obstructive pulmonary disease

Acknowledgements

Not applicable.

Authors' contributions

Asim Roy was a major contributor to this manuscript. All authors read and approved the final manuscript.

Funding

This research was supported by a Mayo Clinic-ASU Alliance Fellowship awarded to Asim Roy in Summer, 2017.

Availability of data and materials

Remote patient monitoring data from the Body Guardian and associated devices was provided by Preventice Solutions (<https://www.preventicesolutions.com/>). Some patient data (de-identified) was provided by Mayo Clinic. This study was supported by a 2015 NIH RO1 grant to Dr. Charles Bruce of Mayo Clinic: [40].

Ethics approval and consent to participate

Subjects in this research study represent one arm of a two-arm randomized clinical trial [40]. Patients assigned to the active intervention and monitoring arm were included. As part of the parent trial, all subjects provided written informed consent that their data collected as part of the parent trial could be used in research. The trial was approved by the Mayo Clinic Institutional Review Board (reference# 15-001448); the IRB also approved sharing of a limited de-identified dataset with Dr. Roy (ASU) as an external partner in this research study.

Consent for publication

Patients provided written informed consent which acknowledges that study results may be made public, but individually identifying information will not be used in any publication.

Exact wording on consent document is "If the results of this study are made public, information that identifies you will not be used." And "If you are enrolled in the study, you will be identified by a study number. Only the study staff will have access to the link between your study number and your identifying information. Your identifying information will not be released outside of Mayo Clinic and any publications will exclude any kind of identifiers that could be linked with you."

Competing interests

ASU and Mayo Clinic has filed a nonprovisional patent application on this method titled "*Method and Apparatus for Event Prediction for a Physical System Based on Data Solely Collected from the Physical System*" on October 18, 2019.

Author details

¹Arizona State University, Tempe, AZ, USA. ²Mayo Clinic, Rochester, MN, USA.

Received: 25 February 2020 Accepted: 23 June 2020

Published online: 08 July 2020

References

1. Adams DE. Nonlinear damage models for diagnosis and prognosis in structural dynamic systems. In: Component and systems diagnostics, prognostics, and health management II, vol. 4733: International Society for Optics and Photonics; 2002. p. 180–91.
2. Atiya AF, El-Shoura SM, Shaheen SI, El-Sherif MS. A comparison between neural-network forecasting techniques-case study: river flow forecasting. *IEEE Trans Neural Netw.* 1999;10(2):402–9.
3. Casoetto N, Djurdjanovic D, Mayor R, Ni J, Lee J. Multisensor process performance assessment through use of autoregressive modeling and feature maps. *J Manuf Syst.* 2003;22(1):64.
4. Chandrashekar G, Sahin F. A survey on feature selection methods. *Comput Electr Eng.* 2014;40(1):16–28.
5. Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, et al. Telemonitoring in patients with heart failure. *N Engl J Med.* 2010;363(24):2301–9.
6. Che C, Xiao C, Liang J, Jin B, Zhu J, Wang F. An rnn architecture with dynamic temporal matching for personalized predictions of parkinson's disease. In: Proceedings of the 2017 SIAM international conference on data mining: Society for Industrial and Applied Mathematics; 2017. p. 198–206.
7. Chelidze D, Cusumano JP. A dynamical systems approach to failure prognosis. *J Vib Acoust.* 2004;126(1):2–8.
8. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inform Assoc.* 2016;24(2):361–70.
9. Dai W, Brisimi TS, Adams WG, Mela T, Saligrama V, Paschalidis IC. Prediction of hospitalization due to heart diseases by supervised learning methods. *Int J Med Inform.* 2015;84(3):189–97.
10. Dy JG. Unsupervised feature selection. In: Computational methods of feature selection; 2008. p. 19–39.
11. Esteban C, Schmidt D, Krompaß D, Tresp V. Predicting sequences of clinical events by using a personalized temporal latent embedding model. In: 2015 international conference on healthcare informatics: IEEE; 2015. p. 130–9.
12. Friedman JH, Stuetzle W. Projection pursuit regression. *J Am Stat Assoc.* 1981;76(376):817–23.
13. Friedman JH. Multivariate adaptive regression splines. *Ann Stat.* 1991;19(1):1–67.
14. Frisch DR. A novel technique to expand the electrocardiographic recording capability from an Apple watch. *Am J Med.* 2019;132(8):940–1.
15. Gheshmoune M, Lebbah M, Azzag H. State-of-the-art on clustering data streams. *Big Data Anal.* 2016;1(1):13.
16. Gola G, Nystad BH. From measurement collection to remaining useful life estimation: defining a diagnostic-prognostic frame for optimal maintenance scheduling of choke valves undergoing erosion. In: Annual conference of the prognostics and health management society; 2011. p. 26–9.
17. Halcox JP, Wareham K, Cardew A, Gilmore M, Barry JP, Phillips C, Gravenor MB. Assessment of remote heart rhythm sampling using the AliveCor heart monitor to screen for atrial fibrillation: the REHEARSE-AF study. *Circulation.* 2017; 136(19):1784–94.
18. Hamburg MA, Collins FS. The path to personalized medicine. *N Engl J Med.* 2010;363(4):301–4.
19. Heimes FO. Recurrent neural networks for remaining useful life estimation. In: 2008 international conference on prognostics and health management: IEEE; 2008. p. 1–6.
20. Husmeier D. Neural networks for conditional probability estimation: forecasting beyond point predictions: Springer Science & Business Media; 2012.
21. Jiang X, Osl M, Kim J, Ohno-Machado L. Calibrating predictive model estimates to support personalized medicine. *J Am Med Inform Assoc.* 2012;19(2):263–74.
22. Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, Kripalani S. Risk prediction models for hospital readmission: a systematic review. *JAMA.* 2011;306(15):1688–98.
23. Kohonen T. The self-organizing map. *Proc IEEE.* 1990;78(9):1464–80.
24. Lee J, Maslove DM, Dubin JA. Personalized mortality prediction driven by electronic medical data and a patient similarity metric. *PLoS One.* 2015;10(5).
25. Lee J. Patient-specific predictive modeling using random forests: an observational study for the critically ill. *JMIR Med Inform.* 2017;5(1):e3.
26. Lei Y, Li N, Gontarz S, Lin J, Radkowski S, Dybala J. A model-based method for remaining useful life prediction of machinery. *IEEE Trans Reliab.* 2016;65(3):1314–26.
27. Liang Y, Liang X. Improving signal prediction performance of neural networks through multiresolution learning approach. *IEEE Trans Syst Man Cyber B Cybern.* 2006;36(2):341–52.

28. Liu J, Wang W, Ma F, Yang YB, Yang CS. A data-model-fusion prognostic framework for dynamic system state forecasting. *Eng Appl Artif Intell*. 2012;25(4):814–23.
29. Lloyd S. Least squares quantization in PCM. *IEEE Trans Inf Theory*. 1982;28(2):129–37.
30. Luo J, Bixby A, Pattipati K, Qiao L, Kawamoto M, Chigusa S. An interacting multiple model approach to model-based prognostics. In: SMC'03 conference proceedings. 2003 IEEE international conference on systems, man and cybernetics. Conference theme-system security and assurance (Cat. no. 03CH37483), vol. 1: IEEE; 2003. p. 189–94.
31. Mortazavi BJ, Downing NS, Bucholz EM, Dharmarajan K, Manhapra A, Li SX, et al. Analysis of machine learning techniques for heart failure readmissions. *Circ Cardiovasc Qual Outcomes*. 2016;9(6):629–40.
32. Nevins JR, Huang ES, Dressman H, Pittman J, Huang AT, West M. Towards integrated clinico-genomic models for personalized medicine: combining gene expression signatures and clinical factors in breast cancer outcomes prediction. *Hum Mol Genet*. 2003;12(suppl_2):R153–7.
33. Ng K, Sun J, Hu J, Wang F. Personalized predictive modeling and risk factor identification using patient similarity. In: AMIA summits on translational science proceedings, 2015; 2015. p. 132.
34. Nystad BH, Gola G, Hulsund JE. Lifetime models for remaining useful life estimation with randomly distributed failure thresholds. In: First european conference of the prognostics and health management society, vol. 3; 2012.
35. Pantelopoulou A, Bourbakis NG. A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Trans Syst Man Cybern Part C Appl Rev*. 2009;40(1):1–12.
36. Preventice. BodyGuardian Remote Monitoring System.™ <https://www.preventicesolutions.com/patients/body-guardian-heart.html>. Accessed 10/3/2019.
37. Qiu H, Lee J, Lin J, Yu G. Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. *Adv Eng Inform*. 2003;17(3–4):127–40.
38. Rahimi K, Bennett D, Conrad N, Williams TM, Basu J, Dwight J, et al. Risk prediction in patients with heart failure: a systematic review and analysis. *JACC Heart Fail*. 2014;2(5):440–6.
39. Ross JS, Mulvey GK, Stauffer B, Patlolla V, Bernheim SM, Keenan PS, Krumholz HM. Statistical models and patient predictors of readmission for heart failure: a systematic review. *Arch Intern Med*. 2008;168(13):1371–86.
40. Technologies for Maintenance of Independent Living in Heart Failure Patients, NIH RO1 2015, Bruce, Charles J, Mayo Clinic Jacksonville, Jacksonville, FL, United States.
41. Tian Z. An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring. *J Intell Manuf*. 2012;23(2):227–37.
42. Tong H, Lim KS. Threshold autoregression, limit cycles and cyclical data. In: Exploration of a nonlinear world: an appreciation of the Howell Tong's contributions to statistics; 2009. p. 9–56.
43. Walker AL, Muhlestein JB. Smartphone electrocardiogram monitoring: current perspectives. *Adv Healthc Technol*. 2018;4:15–24.
44. Wang Y, Ng K, Byrd RJ, Hu J, Ebadollahi S, Daar Z, et al. Early detection of heart failure with varying prediction windows by structured and unstructured data in electronic health records. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC): IEEE; 2015. p. 2530–3.
45. Wong SP, Jacobsen K, Signal N, Chan C, Voss J. Validation of the Samsung S-patch-a novel ambulatory ECG monitoring device. *Circulation*. 2018;138(Suppl_1):A15566.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

