US 20190076031A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2019/0076031 A1**

Valys et al. (43) **Pub. Date: Mar. 14, 2019**

(54) **CONTINUOUS MONITORING OF A USER'S HEALTH WITH A MOBILE DEVICE**

(71) Applicant: **AliveCor, Inc.**

(72) Inventors: **Alexander Vainius Valys**, Sunnyvale, CA (US); **Frank Losasso Petterson**, Los Altos Hills, CA (US); **Conner Daniel Cross Galloway**, Sunnyvale, CA (US); **David E. Albert**, Oklahoma City, OK (US); **Ravi Gopalakrishnan**, San Francisco, CA (US); **Lev Korzinov**, San Franicsco, CA (US); **Fei Wang**, San Francisco, CA (US); **Euan Thomson**, Los Gatos, CA (US); **Nupur Srivastava**, San Francisco, CA (US); **Omar Dawood**, San Francisco, CA (US); **Iman Abuzeid**, San Francisco, CA (US)

(21) Appl. No.: **16/186,244**

(22) Filed: **Nov. 9, 2018**

Related U.S. Application Data

(63) Continuation-in-part of application No. 16/153,403, filed on Oct. 5, 2018, which is a continuation-in-part of application No. 15/393,077, filed on Dec. 28, 2016, now Pat. No. 10,159,415, which is a continuation of application No. 14/730,122, filed on Jun. 3, 2015, now Pat. No. 9,572,499, which is a continuation of application No. 14/569,513, filed on Dec. 12, 2014, now Pat. No. 9,420,956.

(60) Provisional application No. 61/915,113, filed on Dec. 12, 2013, provisional application No. 61/953,616, filed on Mar. 14, 2014, provisional application No. 61/969,019, filed on Mar. 21, 2014, provisional application No. 61/970,551, filed on Mar. 26, 2014, provisional application No. 62/014,516, filed on Jun. 19, 2014, provisional application No. 62/589,477, filed on Nov. 21, 2017.

Publication Classification

(51) **Int. Cl.**

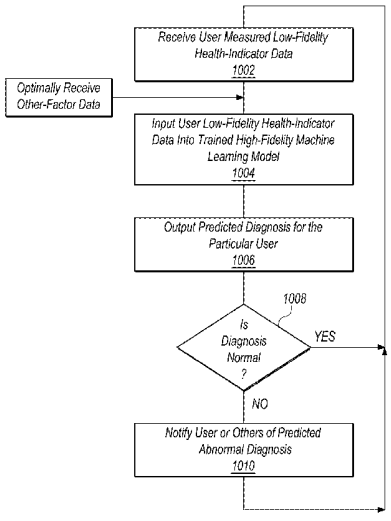
A61B 5/0205	(2006.01)
A61B 5/00	(2006.01)
G16H 40/63	(2018.01)
A61B 5/046	(2006.01)
A61B 5/024	(2006.01)
A61B 5/0245	(2006.01)
G06F 19/00	(2018.01)
A61B 5/11	(2006.01)
A61B 5/0452	(2006.01)
A61B 5/021	(2006.01)
G16H 10/60	(2018.01)
G16H 50/30	(2018.01)
G16H 15/00	(2018.01)

(52) **U.S. Cl.**

CPC ***A61B 5/02055*** (2013.01); *G16H 15/00* (2018.01); ***G16H 40/63*** (2018.01); ***A61B 5/046*** (2013.01); ***A61B 5/02405*** (2013.01); ***A61B 5/02416*** (2013.01); ***A61B 5/0245*** (2013.01); ***G06F 19/00*** (2013.01); ***A61B 5/681*** (2013.01); ***A61B 5/0022*** (2013.01); ***A61B 5/7275*** (2013.01); ***A61B 5/6898*** (2013.01); ***A61B 5/746*** (2013.01); *A61B 5/1118* (2013.01); *A61B 5/0452* (2013.01); *A61B 5/02438* (2013.01); *A61B 5/021* (2013.01); *G16H 10/60* (2018.01); *G16H 50/30* (2018.01); ***A61B 5/7264*** (2013.01)

(57) **ABSTRACT**

Disclosed herein are devices, systems, methods and platforms for continuously monitoring the health status of a user, for example the cardiac health status. The present disclosure describes systems, methods, devices, software, and platforms for continuously monitoring a user's health-indicator data (for example and without limitation PPG signals, heart rate or blood pressure) from a user-device in combination with corresponding (in time) data related to factors that may impact the health-indicator ("other-factors") to determine whether a user has normal health as judged by or compared to, for example and not by way of limitation, either (i) a group of individuals impacted by similar other-factors, or (ii) the user him/herself impacted by similar other-factors.



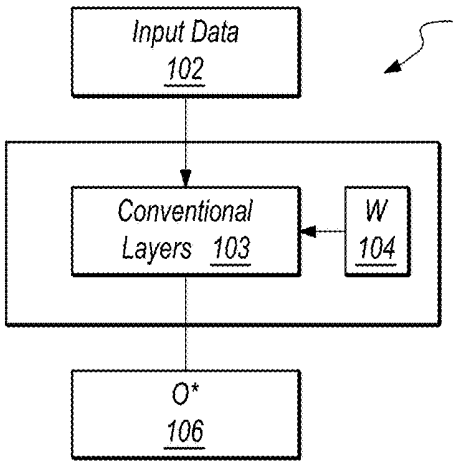


FIG. 1A
(Prior Art)

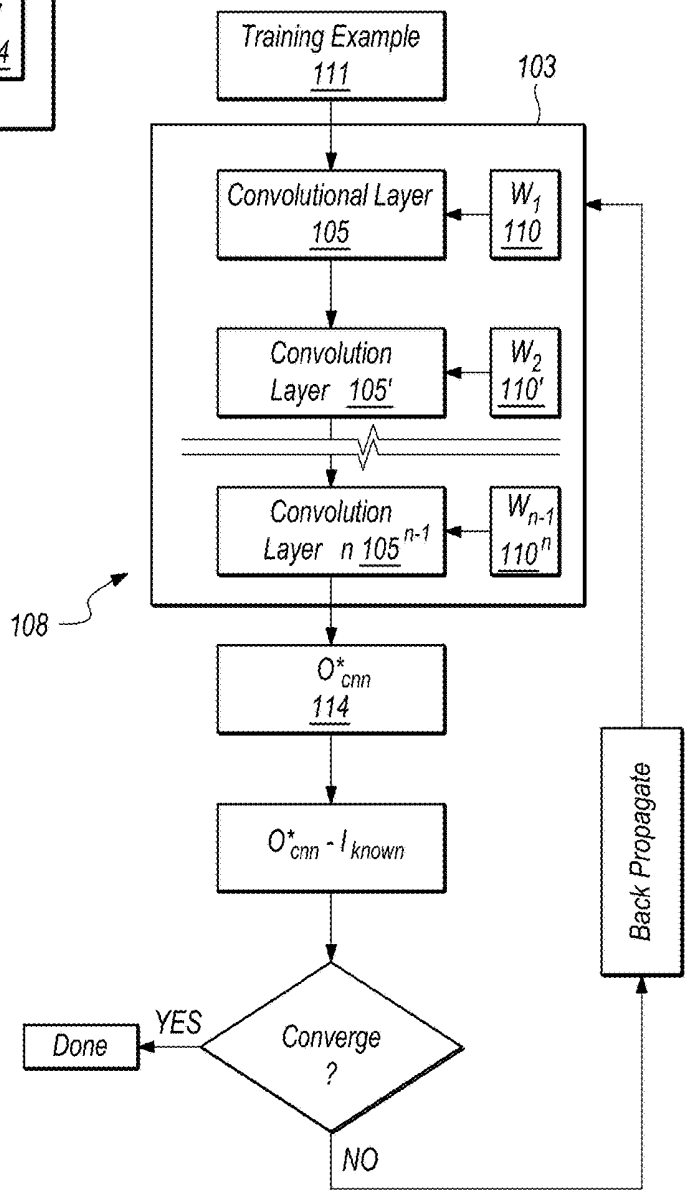


FIG. 1B
(Prior Art)

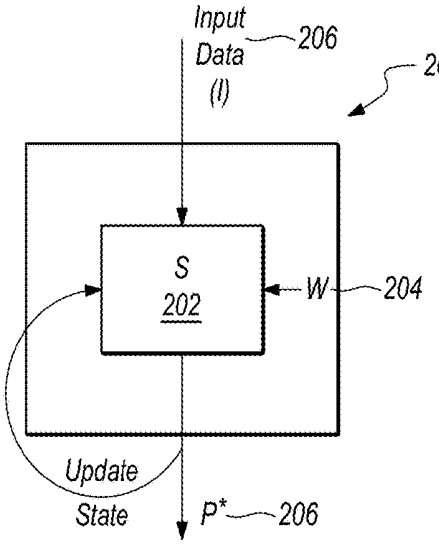


FIG. 2A
(PRIOR ART)

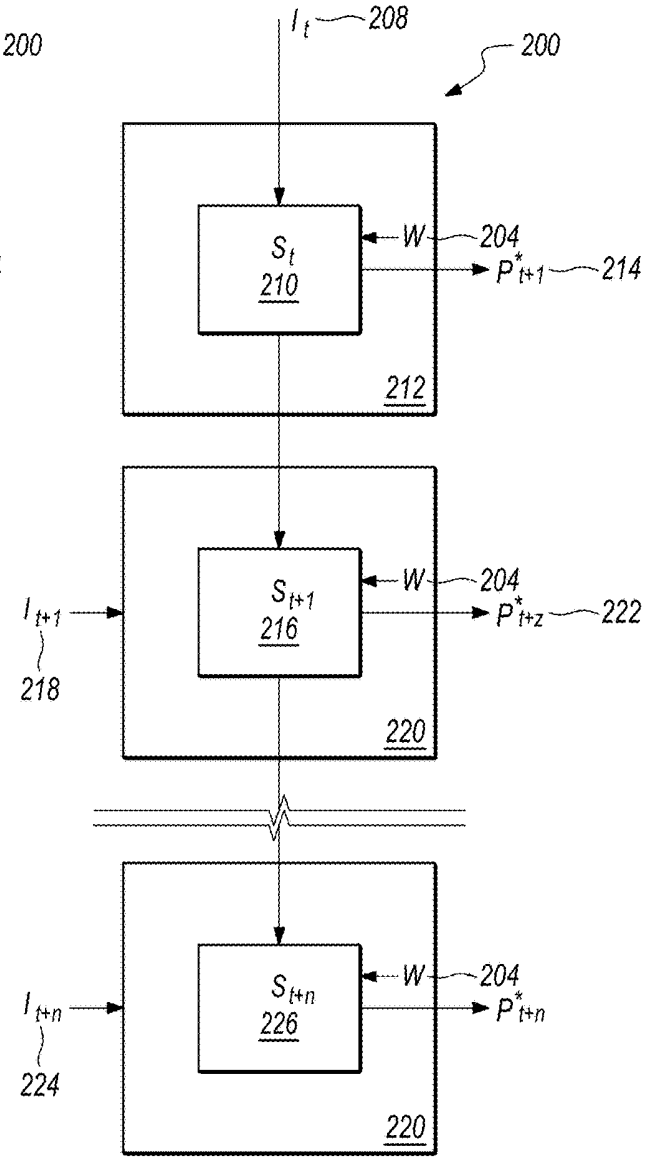


FIG. 2B
(PRIOR ART)

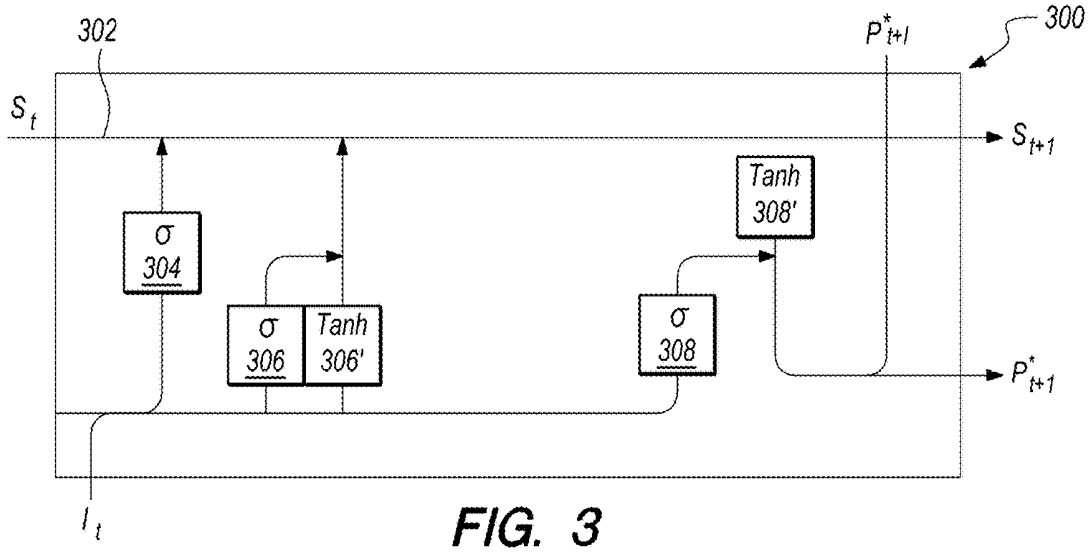


FIG. 3
(PRIOR ART)

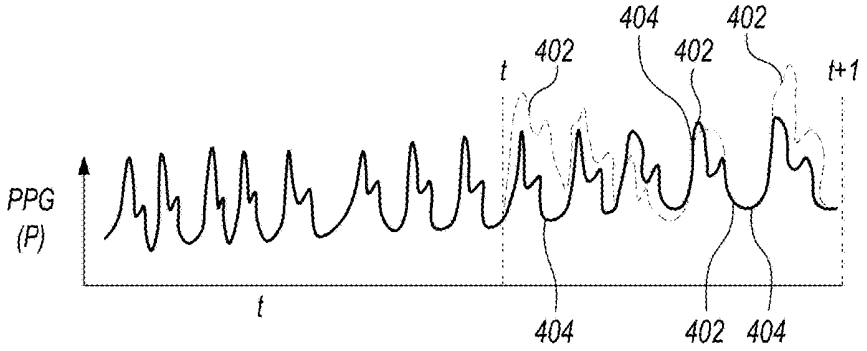


FIG. 4A

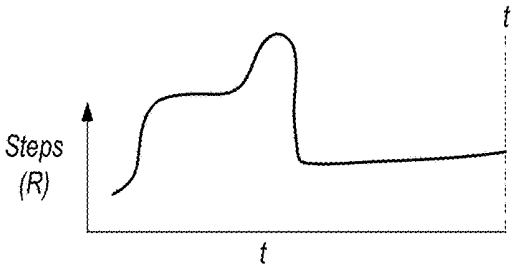


FIG. 4B

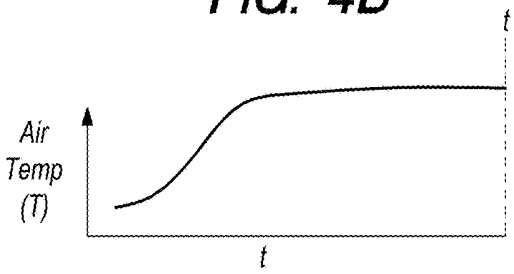


FIG. 4C

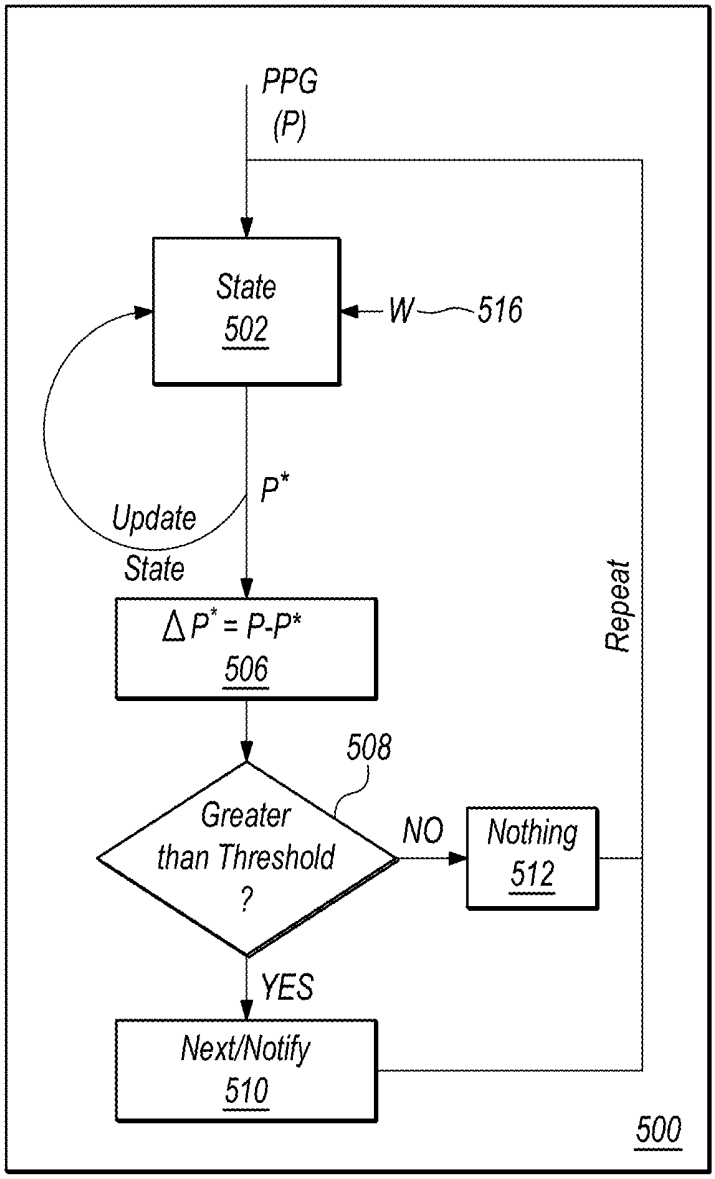

PPG
(P)
State
502
W
516
P*
Update
State
ΔP* = P-P*
506
508
Greater
than Threshold
?
NO
Nothing
512
YES
Next/Notify
510
Repeat
500

FIG. 5A

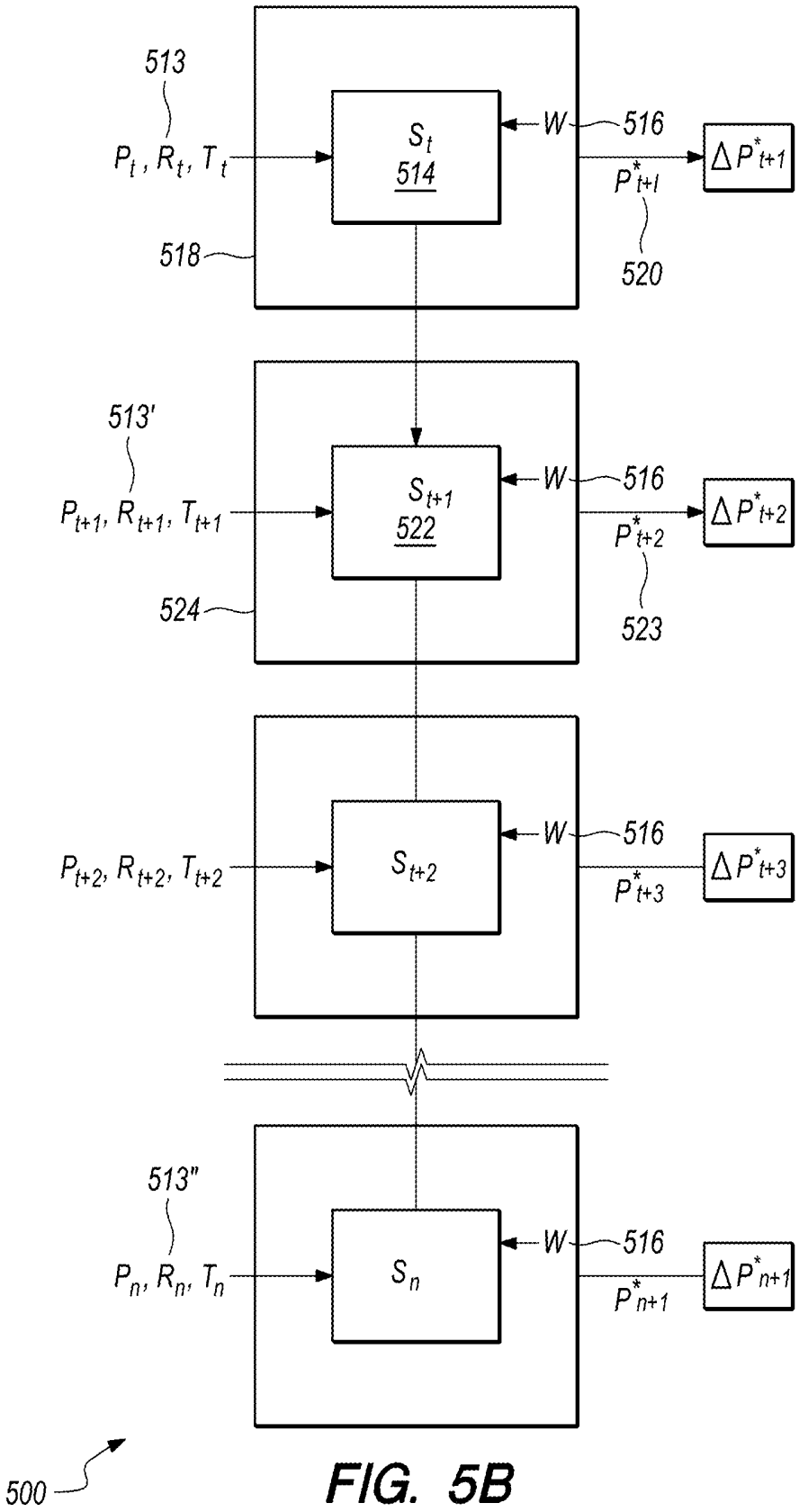

513
P_t, R_t, T_t
S_t
514
W
516
ΔP^*_{t+1}
P^*_{t+1}
518
520
513'
$P_{t+1}, R_{t+1}, T_{t+1}$
S_{t+1}
522
W
516
ΔP^*_{t+2}
P^*_{t+2}
524
523
$P_{t+2}, R_{t+2}, T_{t+2}$
S_{t+2}
W
516
ΔP^*_{t+3}
P^*_{t+3}
513"
P_n, R_n, T_n
S_n
W
516
ΔP^*_{n+1}
P^*_{n+1}
500

FIG. 5B

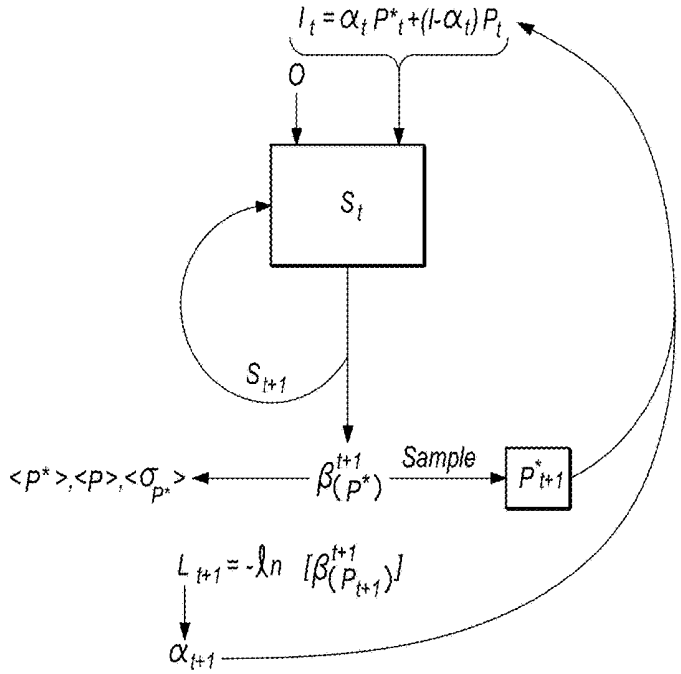


FIG. 5C

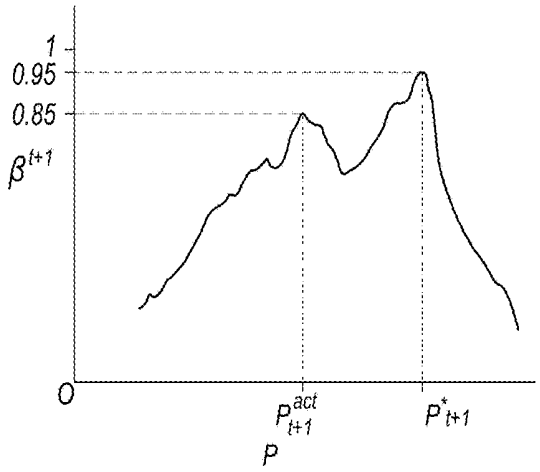


FIG. 5D

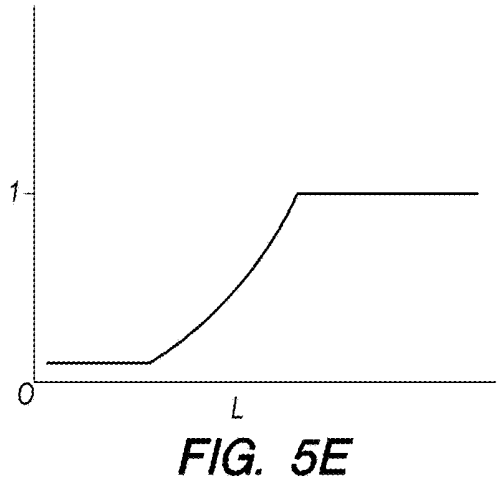


FIG. 5E

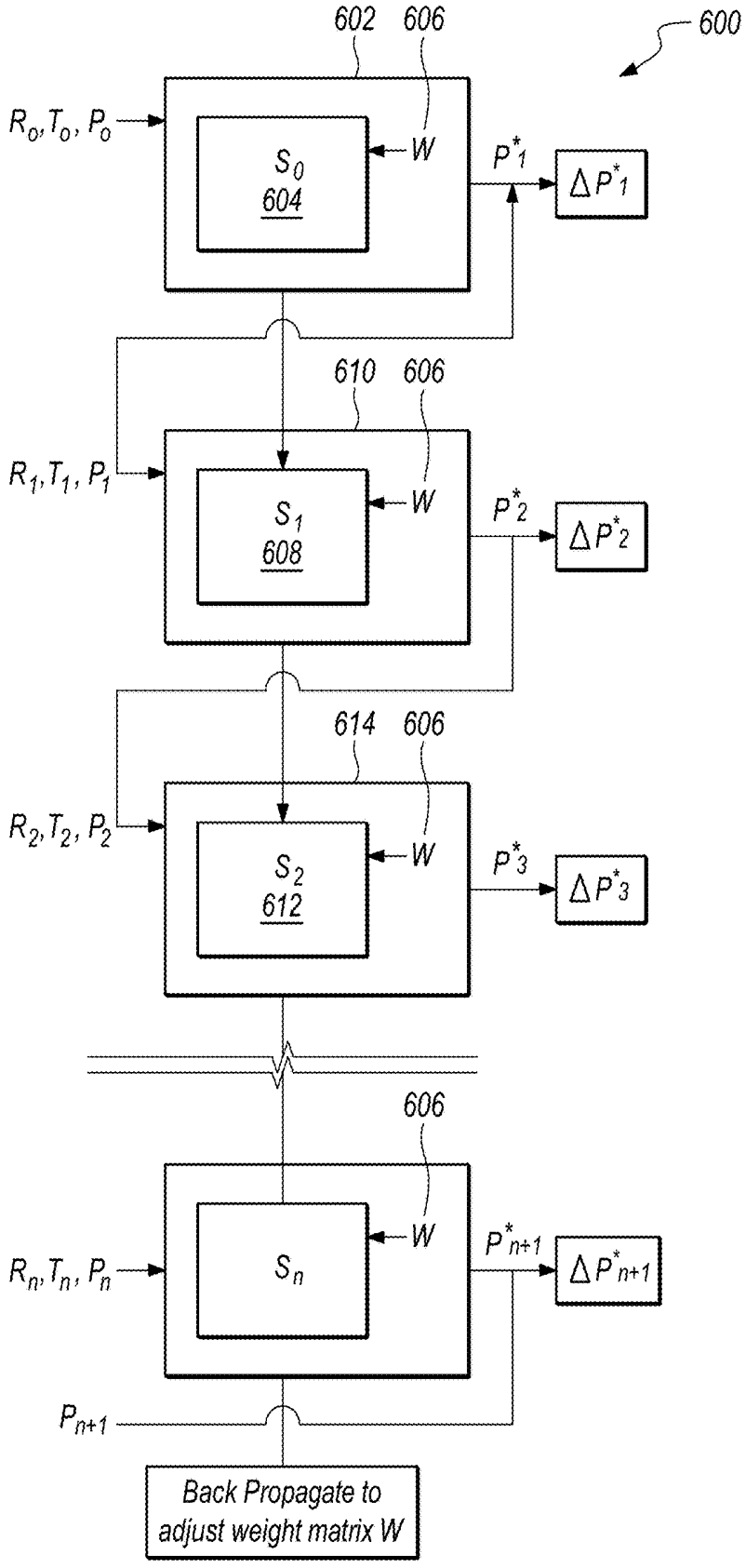

600
602
606
R_0, T_0, P_0
S_0
604
W
P^*_1
ΔP^*_1
610
606
R_1, T_1, P_1
S_1
608
W
P^*_2
ΔP^*_2
614
606
R_2, T_2, P_2
S_2
612
W
P^*_3
ΔP^*_3
606
R_n, T_n, P_n
S_n
W
P^*_{n+1}
ΔP^*_{n+1}
P_{n+1}
Back Propagate to adjust weight matrix W

FIG. 6

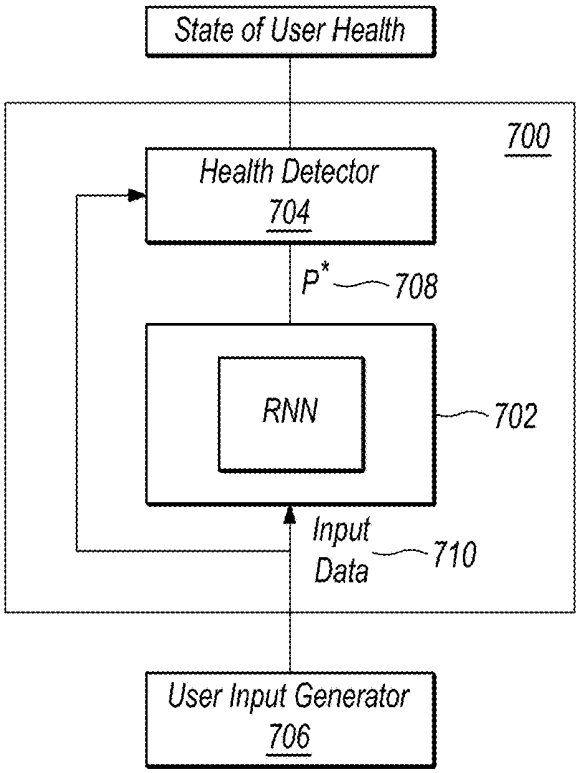

State of User Health
700
Health Detector
704
P*
708
RNN
702
Input
Data
710
User Input Generator
706

FIG. 7A

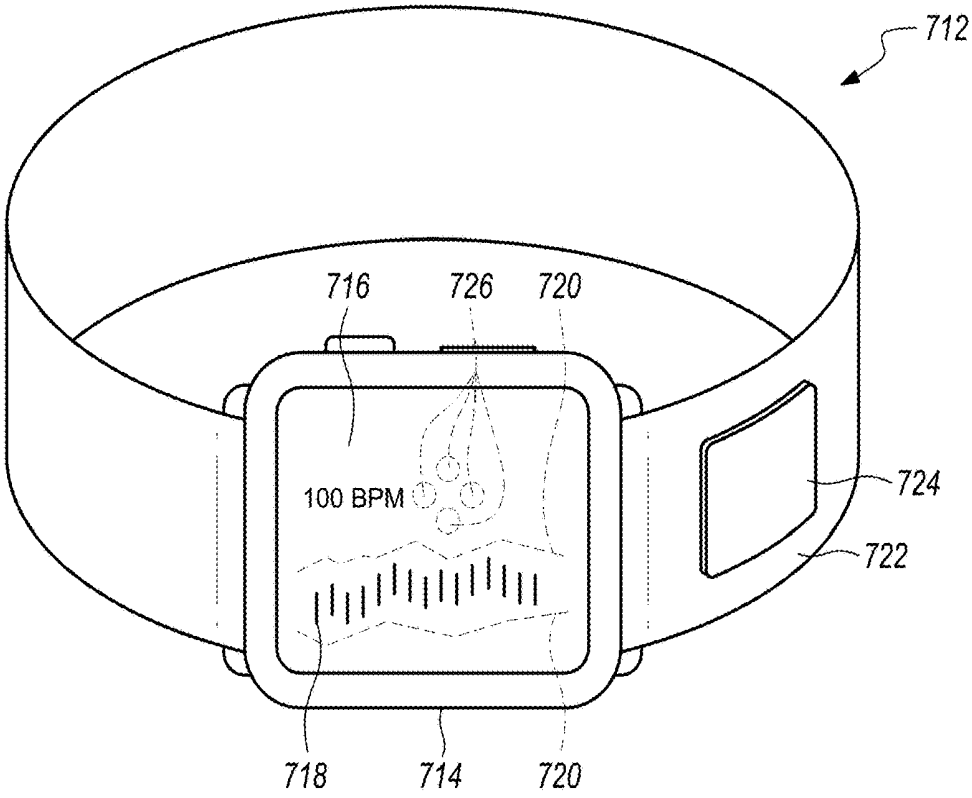

712
716
726
720
100 BPM
724
722
718
714
720

FIG. 7B

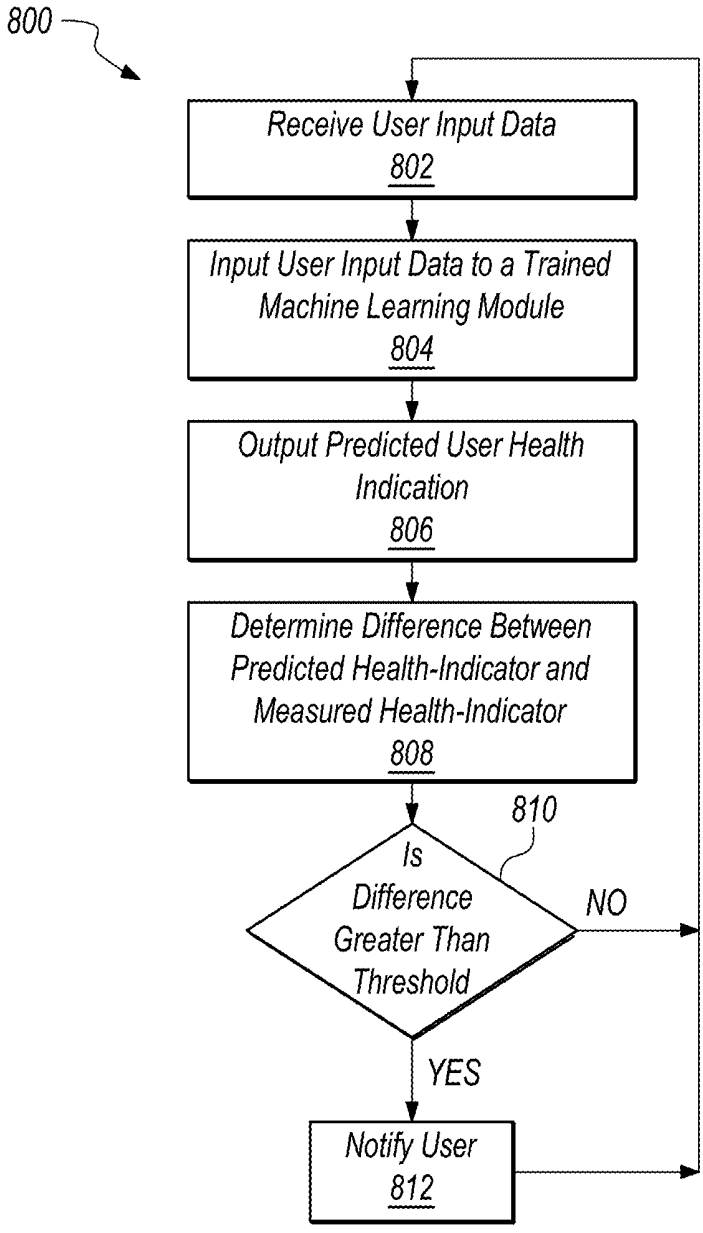

800
Receive User Input Data
802
Input User Input Data to a Trained Machine Learning Module
804
Output Predicted User Health Indication
806
Determine Difference Between Predicted Health-Indicator and Measured Health-Indicator
808
810
Is Difference Greater Than Threshold
NO
YES
Notify User
812

FIG. 8

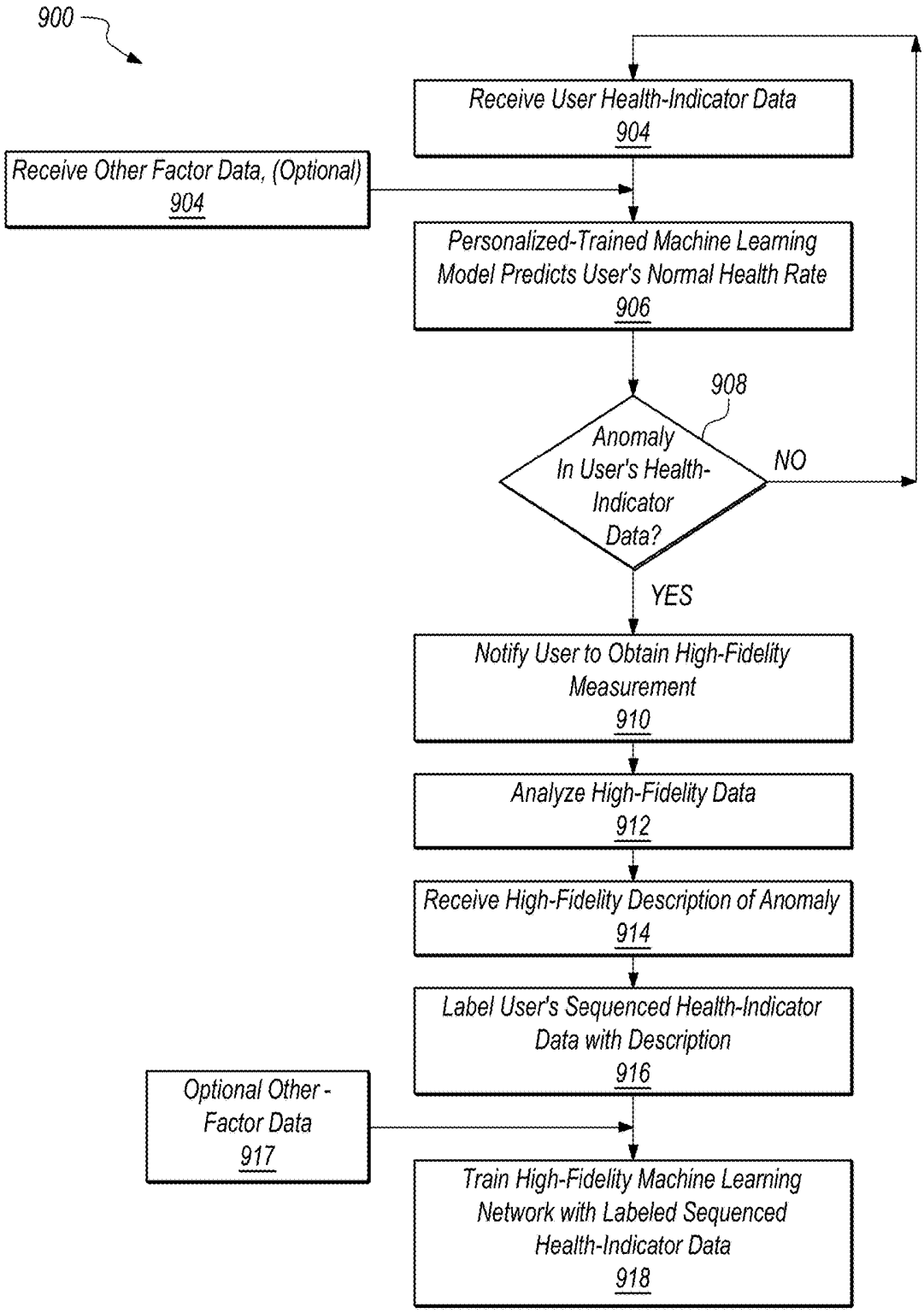

900
Receive User Health-Indicator Data
904
Receive Other Factor Data, (Optional)
904
Personalized-Trained Machine Learning Model Predicts User's Normal Health Rate
906
908
Anomaly In User's Health-Indicator Data?
NO
YES
Notify User to Obtain High-Fidelity Measurement
910
Analyze High-Fidelity Data
912
Receive High-Fidelity Description of Anomaly
914
Label User's Sequenced Health-Indicator Data with Description
916
Optional Other-Factor Data
917
Train High-Fidelity Machine Learning Network with Labeled Sequenced Health-Indicator Data
918

FIG. 9A

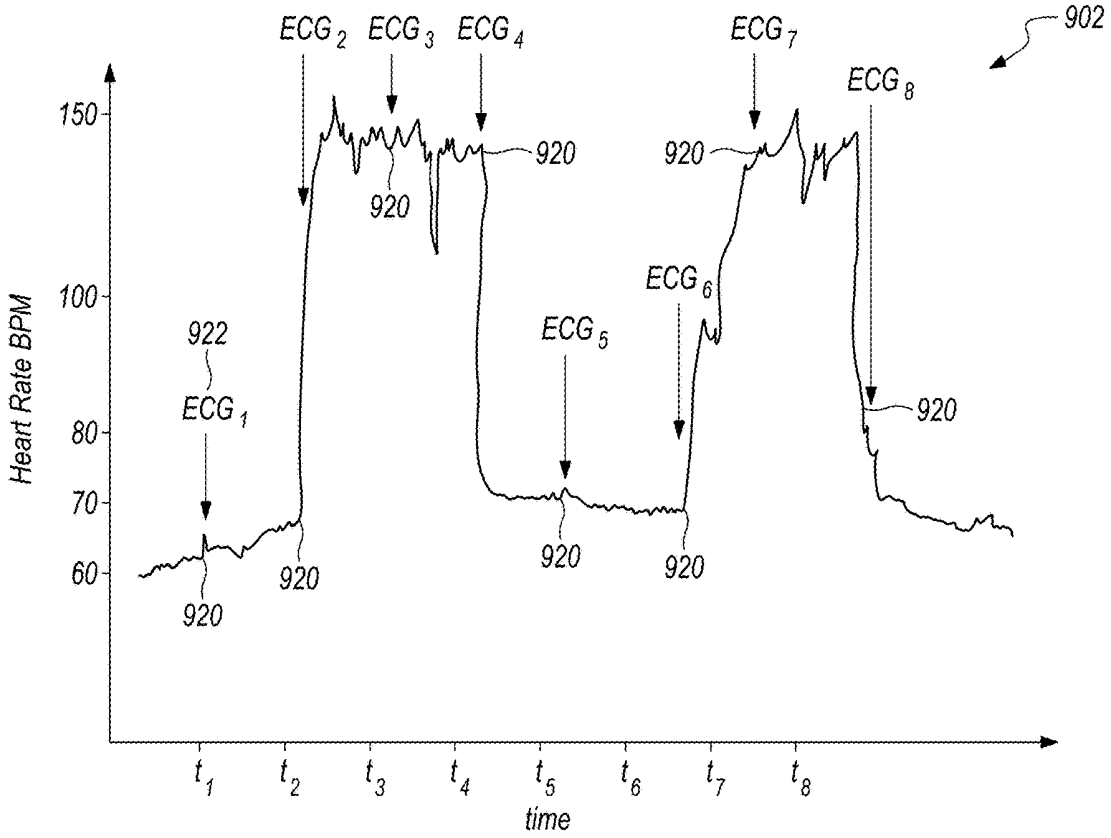

902
ECG 2
ECG 3
ECG 4
ECG 7
ECG 8
150
100
80
70
60
Heart Rate BPM
922
ECG 1
920
ECG 5
ECG 6
t 1
t 2
t 3
t 4
t 5
t 6
t 7
t 8
time

FIG. 9B

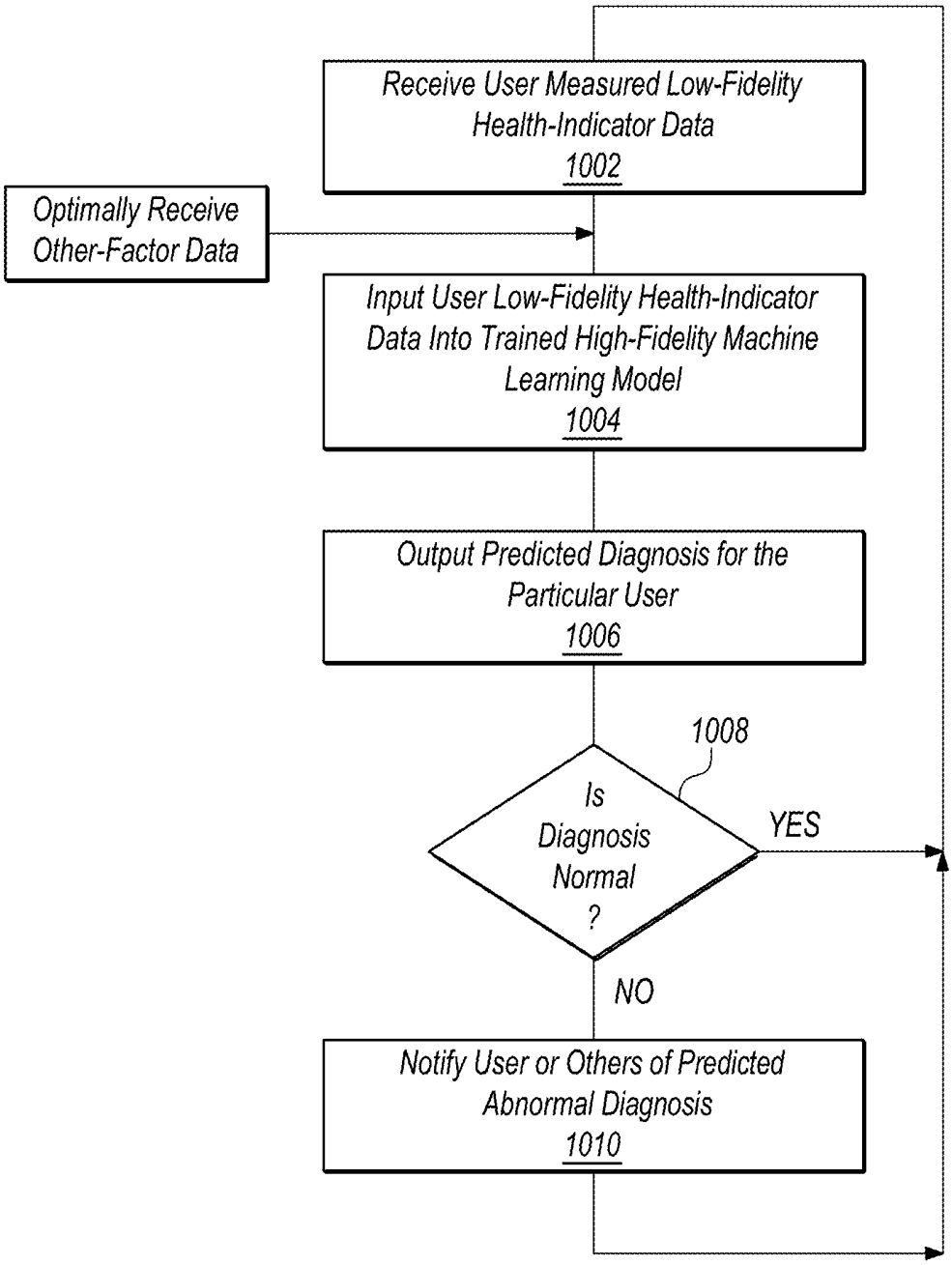

Receive User Measured Low-Fidelity Health-Indicator Data
1002
Optimally Receive Other-Factor Data
Input User Low-Fidelity Health-Indicator Data Into Trained High-Fidelity Machine Learning Model
1004
Output Predicted Diagnosis for the Particular User
1006
1008
Is Diagnosis Normal ?
YES
NO
Notify User or Others of Predicted Abnormal Diagnosis
1010

FIG. 10

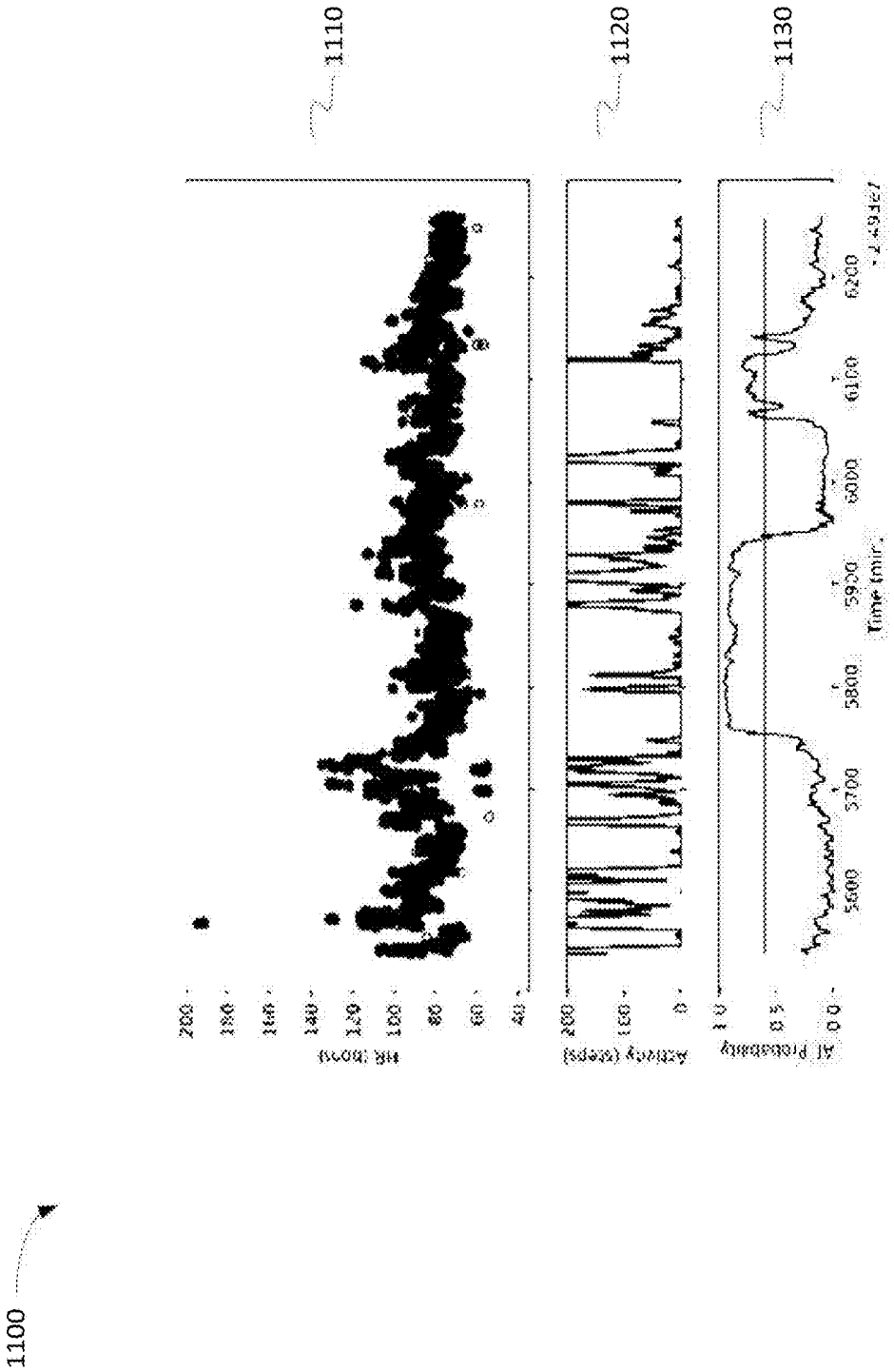

1100
1110
1120
1130
HR (BPM)
Activity (steps)
AF Probability
Time (min)

Fig. 11

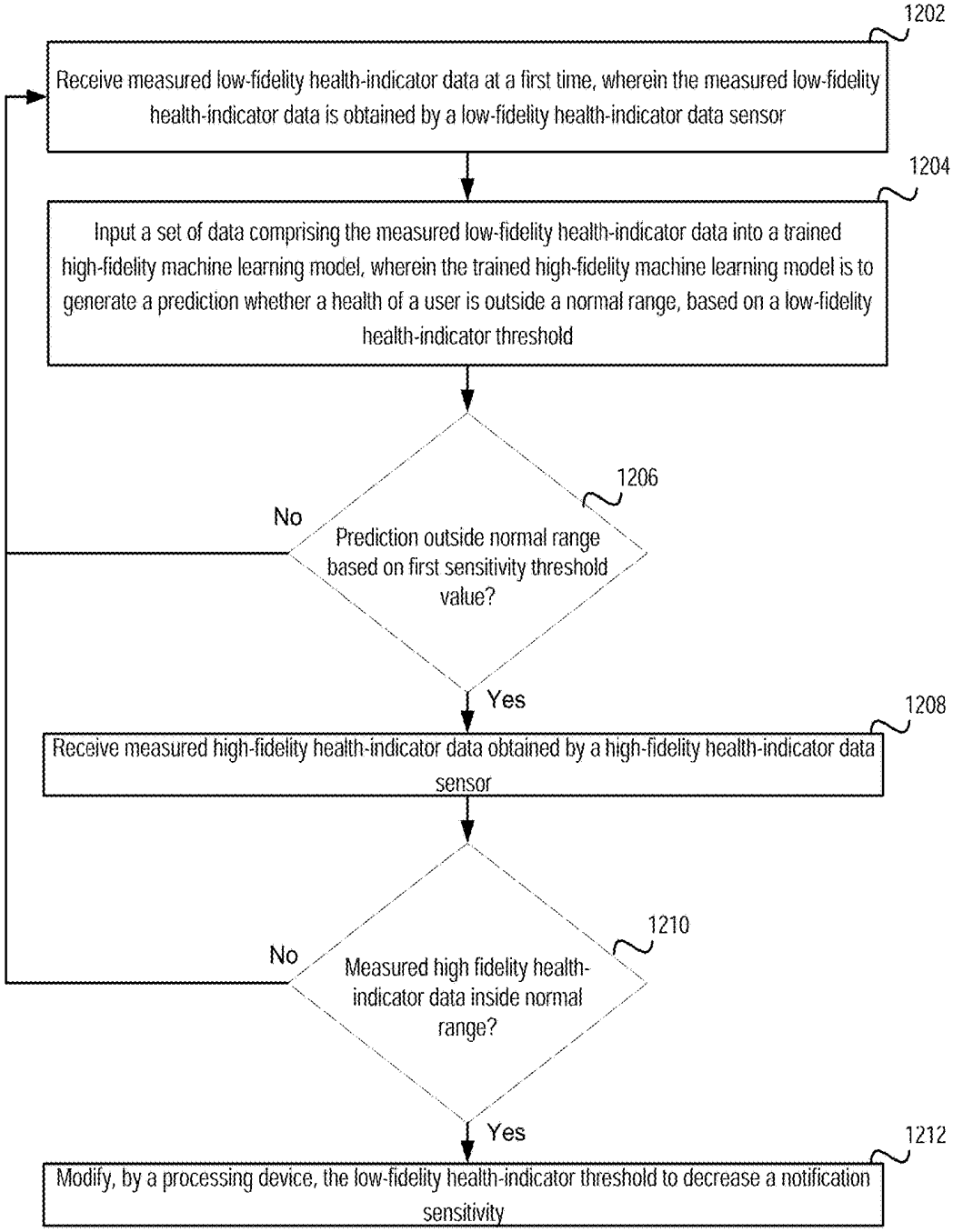

1202
Receive measured low-fidelity health-indicator data at a first time, wherein the measured low-fidelity health-indicator data is obtained by a low-fidelity health-indicator data sensor
1204
Input a set of data comprising the measured low-fidelity health-indicator data into a trained high-fidelity machine learning model, wherein the trained high-fidelity machine learning model is to generate a prediction whether a health of a user is outside a normal range, based on a low-fidelity health-indicator threshold
1206
Prediction outside normal range based on first sensitivity threshold value?
No
Yes
1208
Receive measured high-fidelity health-indicator data obtained by a high-fidelity health-indicator data sensor
1210
Measured high fidelity health-indicator data inside normal range?
No
Yes
1212
Modify, by a processing device, the low-fidelity health-indicator threshold to decrease a notification sensitivity

FIG. 12

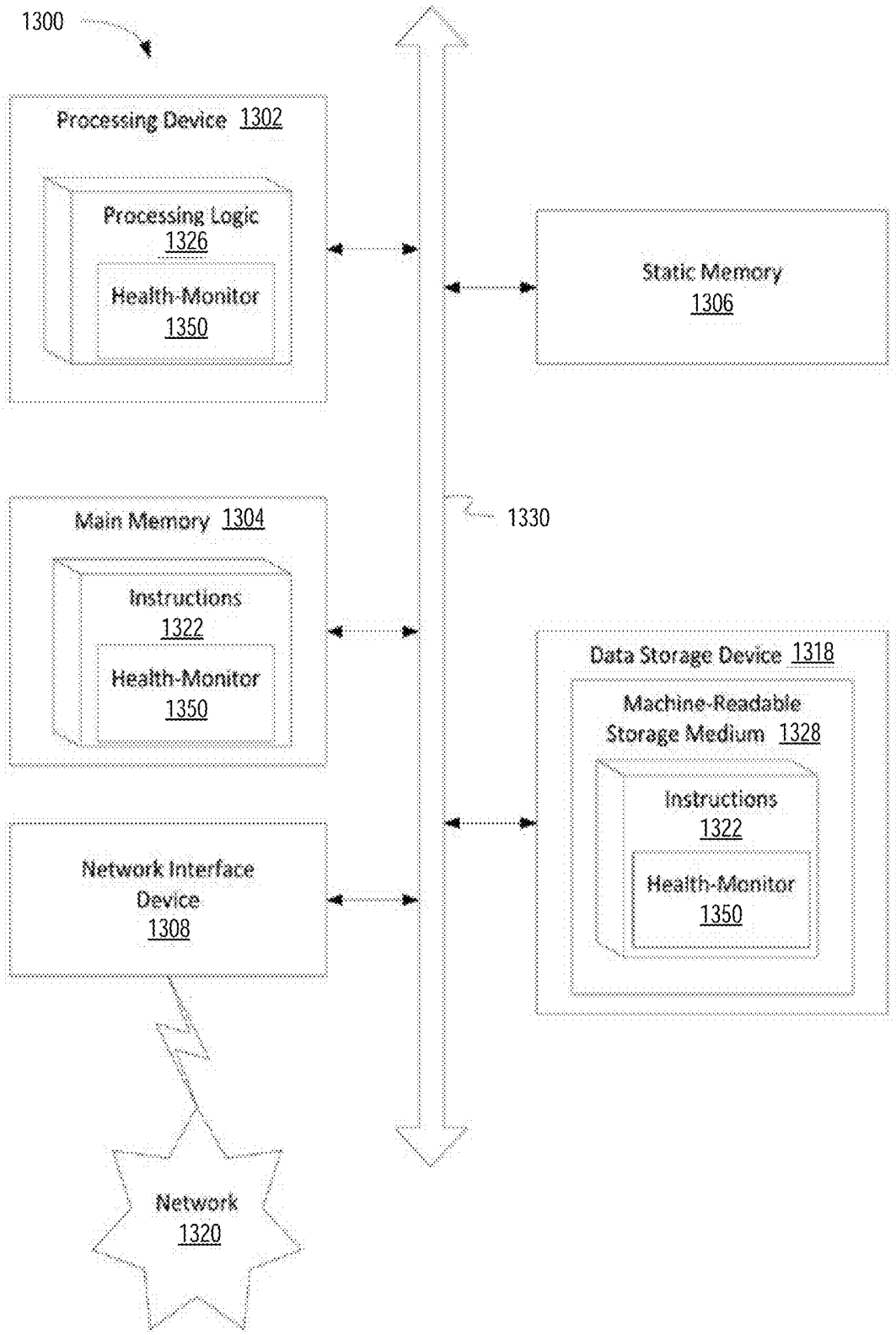

1300
Processing Device 1302
Processing Logic 1326
Health-Monitor 1350
Static Memory 1306
1330
Main Memory 1304
Instructions 1322
Health-Monitor 1350
Data Storage Device 1318
Machine-Readable Storage Medium 1328
Instructions 1322
Health-Monitor 1350
Network Interface Device 1308
Network 1320

FIG. 13

CONTINUOUS MONITORING OF A USER'S HEALTH WITH A MOBILE DEVICE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a continuation-in-part of U.S. application Ser. No. 16/153,403, filed Oct. 5, 2018, which is a continuation-in-part of U.S. application Ser. No. 15/393,077, filed Dec. 28, 2016, which is a continuation of U.S. application Ser. No. 14/730,122, filed Jun. 3, 2015, now U.S. Pat. No. 9,572,499, issued Feb. 21, 2017, which is a continuation of U.S. application Ser. No. 14/569,513, filed Dec. 12, 2014, now U.S. Pat. No. 9,420,956, issued Aug. 23, 2016, the entire contents of all of which are incorporated by reference, which claims the benefit of U.S. Provisional Application No. 61/915,113, filed Dec. 12, 2013, which application is incorporated herein by reference, U.S. Provisional Application No. 61/953,616, filed Mar. 14, 2014, U.S. Provisional Application No. 61/969,019, filed Mar. 21, 2014, U.S. Provisional Application No. 61/970,551, filed Mar. 26, 2014, which application is incorporated herein by reference, U.S. Provisional Application No. 62/014,516, filed Jun. 19, 2014, which application is incorporated herein by reference, and U.S. Provisional Application 62/569,309, filed Oct. 6, 2017, the entire contents of which are hereby incorporated by reference. This application also claims the benefit of U.S. Provisional Application 62/589,477, filed Nov. 21, 2017, the entire contents of which are hereby incorporated by reference.

BACKGROUND

[0002] Indicators of an individual's physiological health ("health-indicators")—for example and not by way of limitation: heart rate, heart rate variability, blood pressure, and ECG (electrocardiogram) to name a few—can be measured or calculated at any discrete point or points in time from data collected to measure the health-indicators. In many cases, the value of the health-indicator at a particular time, or a change over time provides information regarding the state of an individual's health. A low or high heart rate or blood pressure, or an ECG that clearly demonstrate myocardial ischemia, for example, may demonstrate the need for immediate intervention. But, readings, a series of readings, or changes to the readings over time of these indicators may provide information not recognized by the user or even a health professional as needing attention.

[0003] Arrhythmias, for example, may occur continuously or may occur intermittently. Continuously occurring arrhythmias may be diagnosed most definitively from an electrocardiogram of an individual. Because a continuous arrhythmia is always present, ECG analysis may be applied at any time in order to diagnose the arrhythmia. An ECG may also be used to diagnose intermittent arrhythmias. However, because intermittent arrhythmias may be asymptomatic and/or are by definition intermittent , diagnosis presents challenges of applying the diagnostic technique at the time when the individual is experiencing the arrhythmia. Thus, actual diagnosis of intermittent arrhythmias is notoriously difficult. This particular difficulty is compounded with asymptomatic arrhythmias, which account for nearly 40% of arrhythmias in the US. Boriani G. and Pettorelli D., Atrial Fibrillation Burden and Atrial Fibrillation type: Clinical Significance and Impact on the Risk of Stroke and Decision Making for Long-term Anticoagulation, *Vascul Pharmacol.,* 83:26-35 (August 2016), pp. 26.

[0004] Sensors and mobile electronics technologies exist which permit frequent or continuous monitoring and recording of health-indicators. However, the capability of these sensor platforms often exceeds that of conventional medical science to interpret the data they produce. The physiological significance of health-indicator parameters, like heart rate, are frequently well defined only in specific medical contexts: for instance, heart rate is conventionally evaluated as a single scalar value out of context from other data/information that may impact the health-indicator. A resting heart rate in the range of 60-100 beats per minute (BPM) may be considered normal. A user may generally measure their resting heart rate manually once or twice per day.

[0005] A mobile sensor platform (for example: a mobile blood pressure cuff; mobile heart rate monitor; or mobile ECG device) may be capable of monitoring the health-indicator (e.g., heart rate) continuously, e.g., producing a measurement every second or every 5 seconds, while simultaneously also acquiring other data about the user such as and without limitation: activity level, body position, and environmental parameters like air temperature, barometric pressure, location, etc. In a 24-hour period, this may result in many thousands of independent health-indicator measurements. In contrast to a measurement once or twice a day, there is relatively little data or medical consensus on what a "normal" sequence of thousands of measurements looks like.

[0006] Devices presently used to continuously measure health-indicators of users/patients range from bulky, invasive, and inconvenient to simple wearable or handheld mobile devices. Presently, these devices do not provide the capability to effectively utilize the data to continuously monitor a person's heath. It is up to a user or health professional to assess the health-indicators in light of other factors that may impact these health-indicators to determine the health status of the user.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] Certain features described herein are set forth with particularity in the appended claims. A better understanding of the features and advantages disclosed embodiments will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles described herein are utilized, and the accompanying drawings of which:

[0008] FIGS. **1**A-**1**B depict a convolutional neural network that may be used accordance with some embodiments as described herein;

[0009] FIGS. **2**A-**2**B depict a recurrent neural network that may be used in accordance with some embodiments as described herein;

[0010] FIG. **3** depicts an alternative recurrent neural network that may be used in accordance with some embodiments as described herein;

[0011] FIGS. **4**A-**4**C depict hypothetical data plots to demonstrate application of some embodiments as described herein;

[0012] FIGS. **5**A-**5**E depict alternative recurrent neural networks in accordance with some embodiments as described herein and hypothetical plots used to describe some of these embodiments;

[0013] FIG. **6** depicts an unrolled recurrent neural network in accordance with some embodiments as described herein;
[0014] FIGS. **7**A-**7**B depicts systems and devices in accordance with some embodiments as described herein;
[0015] FIG. **8** depicts a method in accordance with some embodiments as described herein;
[0016] FIGS. **9**A-**9**B depicts a method in accordance with some embodiments as described herein and a hypothetical plot of heartrate versus time to demonstrate one or more embodiments;
[0017] FIG. **10** depicts a method in accordance with some embodiments as described herein.
[0018] FIG. **11** depicts hypothetical data plots to demonstrate application of some embodiments as described herein; and
[0019] FIG. **12** depicts a method in accordance with some embodiments as described herein;
[0020] FIG. **13** depicts systems and devices in accordance with some embodiments as described herein.

DETAILED DESCRIPTION

[0021] The high volume of data, complexity of interactions between health-indicators and other-factors and limited clinical guidance may limit the effectiveness of any monitoring system that attempts to detect abnormalities in continuous and/or ambulatory sensor data through specific rules based on conventional medical practice. Embodiments described herein include devices, systems, methods, and platforms that can detect abnormalities in an unsupervised fashion from time sequences of health-indicator data alone or in combination with other-factor (as defined herein) data utilizing predictive machine learning models.
[0022] Atrial fibrillation (AF or AFib) is found in 1-2% of the general population, and the presence of AF increases risk of morbidity and adverse outcomes such as stroke and heart failure. Boriani G. and Pettorelli D., Atrial Fibrillation Burden and Atrial Fibrillation type: Clinical Significance and Impact on the Risk of Stroke and Decision Making for Long-term Anticoagulation, *Vascul Pharmacol.,* 83:26-35 (August 2016), pp. 26. AFib in many people, some estimate as high as 40% of AF patients, may be asymptomatic, and these asymptomatic patients have similar risk profiles for stroke and heart failure as symptomatic patients. See, id. However, the symptomatic patients can take active measures, such as taking blood thinners or other medications, to reduce the risks of negative outcomes. Use of implantable electrical devices (CIEDs) can detect asymptomatic AF (so-called silent AF or SAF) and the duration the patient is in AF. Id. From this information, the time these patients spend in AF, or AF-burden can be determined. Id. An AF-burden of greater than 5-6 min and particularly greater than 1 hour is associated with significant increased risk of stroke and other negative health outcomes. Id. Thus, the ability to measure AF-burden in asymptomatic patients can lead to earlier interventional therapies and may reduce risks of negative health outcomes associated with AF. Id. Detection of SAF is challenging, typically requiring some form of continuous monitoring. Presently continuous monitoring for AF requires bulky, sometimes invasive, and expensive devices, where such monitoring requires a high level of medical professional oversight and review.
[0023] Many devices continuously obtain data to provide a measurement or calculation of the health-indicator data, for example and without limitation FitBit®, Apple Watch®, Polar®, smart phones, tablets among others are in the class of wearable and/or mobile devices. Other devices include permanent or semi-permanent devices on or in a user/patient (e.g., holter), and others may include larger devices in hospitals that may be mobile by virtue of being on a cart. But, little is done with this measured data other than periodically observing it on a display or establishing simple data-thresholds. Observation of the data, even by trained medical professionals, may frequently appear as normal, one primary exception being when a user has readily identifiable acute symptoms. It is tremendously difficult and practically impossible for medical professionals to continuously monitor health-indicators to observe anomalies and/or trends in data that may be indicative of something more serious.
[0024] As used herein, a platform comprises one or more customized software applications (or "applications") configured to interact with one another either locally or through a distributed network including the cloud and the Internet. Applications of a platform as described herein are configured to collect and analyze user data and may include one or more software models. In some embodiments of the platform, the platform includes one or more hardware components (e.g. one or more sensing devices, processing devices, or microprocessors). In some embodiments, a platform is configured to operate together with one or more devices and/or one or more systems. That is, a device as described herein, in some embodiments, is configured to run an application of a platform using a built-in processor, and in some embodiments, a platform is utilized by a system comprising one or more computing devices that interact with or run one or more applications of the platform.
[0025] The present disclosure describes systems, methods, devices, software, and platforms for continuously monitoring a user's data related to one or more health-indicators (for example not by way of limitation PPG signals, heart rate or blood pressure) from a user-device in combination with corresponding (in time) data related to factors that may impact the health-indicator (referred to herein as "other-factors") to determine whether a user has normal health as judged by or compared to, for example and not by way of limitation, either (i) a group of individuals impacted by similar other-factors, or (ii) the user him/herself impacted by similar other-factors. In some embodiments, measured health-indicator data alone or in combination with other-factor data is input into a trained machine learning model that determines a probability the user's measured health-indicator is considered within a healthy range, and if not to notify the user of such. The user not being in a healthy range may increase the likelihood the user may be experiencing a health event warranting high-fidelity information to confirm a diagnosis, such as an arrhythmia which may be symptomatic or asymptomatic. The notification may take the form of, for example, requesting the user to obtain an ECG. Other high-fidelity measurements may be requested, blood pressure, pulse oximeter to name two, ECG is but one example. The high-fidelity measurement, ECG in this embodiment, can be evaluated by algorithms and/or medical professionals to make a notification or diagnosis (collectively referred to herein as "diagnosis", recognizing that only a physician can make a diagnosis). In the ECG example, the diagnosis may be AFib or any other number of well-known conditions diagnosed utilizing ECGs.
[0026] In further embodiments, a diagnosis is used to label a low-fidelity data sequence (e.g., heart rate or PPG), which

may include the other-factor data sequence. This high-fidelity diagnosis-labeled low-fidelity data sequence is used to train a high-fidelity machine learning model. In these further embodiments, the training of the high-fidelity machine learning model may be trained by unsupervised learning or may be updated from time to time with new training examples. In some embodiments, a user's measured low-fidelity health-indicator data sequence and optionally a corresponding (in time) data sequence of other-factors are input into the trained high-fidelity machine learning models to determine a probability and/or prediction the user is experiencing or experienced the diagnosed condition on which the high-fidelity machine learning model was trained. This probability may include a probability of when the event begins and when it ends. Some embodiments, for example, may calculate the atrial fibrillation (AF) burden of a user, or the amount of time a user experiences AF over time. Previously AF burden could only be determined using cumbersome and expensive holter or implantable continuous ECG monitoring apparatus. Thus, some embodiments described herein can continuously monitor a user's health status and notify the user of a health status change by continuously monitoring health-indicator data (for example and not by way of limitation PPG data, blood pressure data, and heart rate data) obtained from a user worn device alone or in combination with corresponding data for other-factors. "Other-factors", as used herein, include anything that may impact the health-indicator, and/or may impact the data representing the health-indicator (e.g., PPG data). These other-factors may include a variety of factors such as by way of example not limitation: air temperature, altitude, exercise levels, weight, gender, diet, standing, sitting, falling, lying down, weather, and BMI to name a few. In some embodiments a mathematical or empirical model not a machine learning model may be used to determine when to notify a user to obtain a high-fidelity measurement, which can then be analyzed and used to train a high-fidelity machine training models as described herein.

[0027] Some embodiments described herein can detect abnormalities of a user in an unsupervised fashion by: receiving a primary time sequence of health-indicator data; optionally receiving one or more secondary time sequences of other-factor data, corresponding in time with the primary time sequence of health-indicator data, which secondary sequences may come from a sensor, or from external data sources (e.g. over a network connection, a computer API, etc.); providing the primary and secondary time sequence(s) to a pre-processor, which may perform operations on the data like filtering, caching, averaging, time alignment, buffering, upsampling and downsampling; providing the time sequences of data to a machine learning model, trained and/or configured to utilize the values of the primary and secondary time sequence(s) to predict next value(s) of the primary sequence at a future time; comparing the predicted primary time sequence values(s) generated by the machine learning module at a specific time t to the measured values of the primary time sequence at time t; and alerting or prompting the user to take an action if the difference between the predicted future time sequence and measured time sequences exceeds a threshold or criteria.

[0028] Some embodiments described herein, thus, detect when the observed behavior of the primary sequence of physiological data with respect to the passage of time and/or in response to the observed secondary sequence of data differs from what is expected given the training examples used to train the model. When the training example is gathered from normal individuals or from data that has been previously categorized as normal for a specific user, then the system can serve as an abnormality detector. If the data has simply been acquired from a specific user without any other categorization, then the system can serve as a change detector, detecting a change in the health-indicator data that the primary sequence is measuring relative to the time at which the training data was captured.

[0029] Described herein are software platforms, systems, devices, and methods for generating and using trained machine learning models to predict or determine a probability when a user's measured health-indicator data (primary sequence) under the influence of other-factor(s) (secondary sequence) is outside the bounds of normal for a healthy population (i.e., a global model) under the influence of similar other-factors, or outside the bounds of normal for that particular user (i.e., personalized model) under the influence of similar other-factors, where a notification of such is provided to the user. In some embodiments, the user may be prompted to obtain additional measured high-fidelity data that can be used to label previously acquired low-fidelity user health-indicator data to generate a different trained high-fidelity machine learning model that has the ability to predict or diagnose abnormalities or events using only low-fidelity health-indicator data, where such abnormalities are typically only identified or diagnosed using high-fidelity data.

[0030] Some embodiments described herein may include inputting a user's health-indicator data, and optionally inputting corresponding (in time) data of other-factors into a trained machine learning model, where the trained machine learning model predicts the user's health-indicator data or a probability distribution of the health-indicator data at a future time step. The prediction in some embodiments is compared with the user's measured health-indicator data at the time step of the prediction, where, if the absolute value of the difference exceeds a threshold, the user is notified that his or her health-indicator data is outside a normal range. This notification, in some embodiments, may include a diagnosis or instructions to do something, for example and not by way of limitation obtain additional measurements or contact a health professional. In some embodiments, health-indicator data and corresponding (in time) data of other-factors from a healthy population of people is used to train the machine learning model. It will be appreciated that the other-factors in training examples used to train the machine learning model may not be averages of the population, rather data for each of the other-factors corresponds in time with collection of the health-indicator data for individuals in the training examples.

[0031] Some embodiments are described as receiving discrete data points in time, predicting discrete data points at a future time from the input and then determining if a loss between discrete measured input at the future time and the predicted value at the future time exceeds a threshold. The skilled artisan will readily appreciate that the input data and output predictions may take forms other than a discrete data point or a scalar. For example, and not by way of limitation, the health-indicator data sequence (also referred to herein as primary sequence) and the other-data sequence (also referred to herein as secondary sequence) may be split into segments of time. The skilled artisan will recognize the

manner in which the data is segmented is a matter of design choice and may take many different forms.

[0032] Some embodiments partition the health-indicator data sequence (also referred to herein as primary sequence) and the other-data sequence (also referred to herein as secondary sequence) into two segments: past, representing all data before a specific time t, and future, representing all data at or after time t. These embodiments input the health-indicator data sequence for a past time segment and all other-data sequence(s) for the past time segment into a machine learning model configured to predict the most probable future segment of the health-indicator data (or distribution of probable future segments). Alternatively, these embodiments input the health-indicator data sequence for a past time segment, all other-data sequences for the past time segment and other-data sequences from the future segment into a machine learning model configured to predict the most probable future segment of the health-indicator data (or distribution of probable future segments). The predicted future segment of the health-indicator data is compared to the user's measured health-indicator data at the future segment to determine a loss and whether the loss exceeds a threshold, in which case some action is taken. The action may include for example and not by way of limitation: notifying the user to obtain additional data (e.g., ECG or blood pressure); notifying the user to contact a healthcare professional; or automatically triggering acquisition of additional data. Automatic acquisition of additional data may include, for example and not by way of limitation, ECG acquisition via a sensor operably coupled (wired or wirelessly) to a user worn computing device, or blood pressure via a mobile cuff around the user's wrist or other appropriate body part and coupled to a user worn computing device. The segments of data may include a single data point, many data points over a period of time, an average of these data points over the time period where the average may include a true average, median or mode. In some embodiments the segments may overlap in time.

[0033] These embodiments detect when the observed behavior or measurement of the health-indicator sequence of data with respect to the passage of time as impacted by corresponding (in time) other-factor sequence of data differs from what is expected from the training examples, which training examples are collected under similar other-factors. If the training examples are gathered from healthy individuals under similar other-factors or from data that has been previously categorized as healthy for a specific user under similar other-factors, then these embodiments serve as an abnormality detector from the healthy population or from the specific user, respectively. If the training examples have simply been acquired from a specific user without any other categorization, then these embodiments serve as a change detector, detecting a change in the health-indicators at the time of measurement relative to the time at which the training examples were collected for the specific user.

[0034] Some embodiments described herein utilize machine learning to continuously monitor a person's health-indicators under the impact of one or more other-factors and assess whether the person is healthy in view a population categorized as healthy under the impact of similar other factors. As the skilled artisan will readily appreciate, a number of different machine learning algorithms or models (including without limitation Bayes, Markov, Gausian processes, clustering algorithms, generative models, kernel and neural network algorithms) may be used without exceeding the scope described herein. As appreciated by the skilled artisan, typical neural networks employ, by way of example not limitation, one or more layers of nonlinear activation functions to predict an output for a received input, and may include one or more hidden layers in addition to the input and output layers. The output of each hidden layer in some of these networks is used as input to the next layer in the network. Examples of neural networks include, by way of example and not limitation, generative neural networks, convolutional neural networks and recurrent neural networks.

[0035] Some embodiments of a health monitoring system monitor heart rate and activity data of an individual as low-fidelity data (e.g., heartrate or PPG data) and detect a condition (e.g. AFib) normally detected using high-fidelity data (e.g., ECG data). For example, the heart rate of an individual may be provided by a sensor continuously or in discrete intervals (such as every five seconds). The heart rate may be determined based on PPG, pulse oximetry, or other sensors. In some embodiments, the activity data may be generated as a number of steps taken, an amount of movement sensed, or other data points indicating an activity level. The low-fidelity (e.g., heartrate) data and activity data can then be input into a machine learning system to determine a prediction of a high-fidelity outcome. For example, the machine learning system may use the low-fidelity data to predict an arrhythmia or other indication of a user's cardiac health. In some embodiments, the machine learning system may use an input of segment of data inputs to determine a prediction. For example, an hour of activity level data and heart rate data may be input to the machine learning system. The system can then use the data to generate a prediction of a condition such as atrial fibrillation. Various embodiments of the present invention are more thoroughly discussed below.

[0036] Referring to FIG. **1**A a trained convolution neural network (CNN) **100** (one example of a feed forward network), takes input data **102**, (e.g., a picture of a boat) into convolutional layers (aka hidden layers) **103**, applies a series of trained weights or filters **104** to the input data **106** in each of the convolutional layers **103**. The output of the first convolutional layer is an activation map (not shown), which is the input to the second convolution layer, to which a trained weight or filter (not shown) is applied, where the output of the subsequent convolutional layers results in activation maps that represent more and more complex features of the input data to the first layer. After each convolutional layer a non-linear layer (not shown) is applied to introduce non-linearity into the problem, which nonlinear layers may include tanh, sigmoid or ReLU. In some cases, a pooling layer (not shown) may be applied after the nonlinear layers, also referred to as a downsampling layer, which basically takes a filter and stride of the same length and applies it to the input, and outputs the maximum number in every sub-region the filter convolves around. Other options for pooling are average pooling and L2-norm pooling. The pooling layer reduces the spatial dimension of the input volume reducing computational costs and to control overfitting. The final layer(s) of the network is a fully connected layer, which takes the output of the last convolutional layer and outputs an n-dimensional output vector representing the quantity to be predicted, e.g., probabilities of image classification 20% automobile, 75% boat 5% bus

and 0% bicycle , i.e., resulting in predictive output **106** (O*), e.g. this is likely a picture of a boat. The output could be a scalar value data point being predicted by the network, a stock price for example. Trained weights **104** may be different for each of the convolutional layers **103**, as will be described more fully below. To achieve this real-world prediction/detection (e.g., it's a boat), the neural network needs to be trained on known data inputs or training examples resulting in trained CNN **100**. To train CNN **100** many different training examples (e.g., many pictures of boats) are input into the model. A skilled artisan in neural networks will fully understand the description above provides a somewhat simplistic view of CNNs to provide some context for the present discussion and will fully appreciate the application of any CNN alone or in combination with other neural networks will be equally applicable and within the scope of some embodiments described herein.

[0037] FIG. **1**B demonstrates training CNN **108**. In FIG. **1**B convolutional layers **103** are shown as individual hidden convolutional layers **105**, **105'** up to convolutional layer $\mathbf{105}^{n-1}$ and the final n^{th} layer is a fully connected layer. It will be appreciated that last layers may be more than one fully connected layer. Training example **111** is input into convolutional layers **103**, a nonlinear activation function (not shown) and weights **110**, **110'** through $\mathbf{110}^{n}$ are applied to training example **111** in series, where the output of any hidden layer is input to the next layer, and so on until the final n^{th} fully connected layer $\mathbf{105}^{n}$ produces output **114**. Output or prediction **114** is compared against training example **111** (e.g., picture of a boat) resulting in difference **116** between output or prediction **114** and training example **111**. If difference or loss **116** is less than some preset loss (e.g., output or prediction **114** predicts the object is a boat), the CNN is converged and considered trained. If the CNN has not converged, using the technique of backpropagation, weights **110** and **110'** through $\mathbf{110}^{n}$ are updated in accordance with how close the prediction is to the known input. The skilled artisan will appreciate that methods other than back propagation may be used to adjust the weights. The second training example (e.g., different picture of a boat) is input and the process repeated again with the updated weights, which are then updated again and so on until the n^{th} training example (e.g., n^{th} picture of n^{th} boat) has been input. This is repeated over and over with the same n-training examples until the convolutional neural network (CNN) is trained or converges on the correct outputs for the known inputs. Once CNN **108** is trained, weights **110**, **110'** through $\mathbf{110}^{n}$ are fixed and used in trained CNN **100**, which are weights **104** as depicted in FIG. **1**A. As explained, there are different weights for each convolutional layer **103** and for each of the fully connected layers. The trained CNN **100** or model is then fed image data to determine or predict that which it is trained to predict/identify (e.g., a boat), as described above. Any trained model, CNN, RNN, etc. may be trained further, i.e., modification of the weights may be permitted, with additional training examples or with predicted data output by the model which is then used as a training example. The machine learning model can be trained "offline", e.g. trained once on a computational platform separate from the platform using/executing the trained model, and then transferred to that platform. Alternatively, embodiments described herein may periodically or continually update the machine learning model based on newly acquired training data. This updated training may occur on a separate computational platform which delivers the updated trained models to the platform using/executing the re-trained model over a network connection, or the training/re-training/update process may occur on the platform itself as new data is acquired. The skilled artisan will appreciate the CNN is applicable to data in a fixed array (e.g., a picture, character, word etc.) or a time sequence of data. For example, sequenced health-indicator data and other-factor data can be modeled using a CNN. Some embodiments utilize a feed-forward, CNN with skip connections and a Gaussian Mixture Model output to determine a probability distribution for the predicted health-indicator, e.g., heart rate, PPG, or arrhythmia.

[0038] Some embodiments can utilize other types and configurations of neural network. The number of convolutional layers can be increased or decreased, as well as the number of fully-connected layers. In general, the optimal number and proportions of convolutional vs. fully-connected layers can be set experimentally, by determining which configuration gives the best performance on a given dataset. The number of convolutional layers could be decreased to 0, leaving a fully-connected network. The number of convolutional filters and width of each filter can also be increased or decreased.

[0039] The output of the neural network may be a single, scalar value, corresponding to an exact prediction for the primary time sequence. Alternatively, the output of the neural network could be a logistic regression, in which each category corresponds to a specific range or class of primary time sequence values, are any number of alternative outputs readily appreciated by the skilled artisan.

[0040] The use of a Gaussian Mixture Model output in some embodiments is intended to constrain the network to learning well-formed probability distributions and improve generalization on limited training data. The use of a multiple elements in some embodiments in the Gaussian Mixture Model is intended to allow the model to learn multi-modal probability distributions. A machine learning model combining or aggregating the results of different neural networks could also be used, where the results could be combined.

[0041] Machine learning models that have an updatable memory or state from previous predictions to apply to subsequent predictions is another approach for modeling sequenced data. In particular some embodiments described herein utilize a recurring neural network. Referring to the example of FIG. **2**A a diagram of a trained recurrent neural network (RNN) **200** is shown. Trained RNN **200** has updatable state (S) **202** and trained weights (W) **204**. Input data **206** is input into sate **202** where weights (W) **204** are applied, and prediction **206** (P*) is output. In contrast to linear neural networks (e.g., CNN **100**), state **202** is updated based on the input data, thereby serving as memory from the previous state for the next prediction with the next data in sequence. Updating the sates gives RNNs a circular or loop feature. To better demonstrate, FIG. **2**B shows trained RNN **200** unrolled, and its applicability to sequenced data. Unrolled, the RNN appears analogous to a CNN, but in an unrolled RNN each of the apparently analogous layers appears as a single layer with an updated state, where the same weights are applied in each iteration of the loop. The skilled artisan will appreciate the single layer may itself have sub-layers, though for clarity of explanation a single layer is depicted here. Input data (I_t) **208** at time t is input into state-at-time t (S_t) **210** and trained weights **204** are applied within cell-at-time t (C_t) **212**. The output of C_t **212**

is prediction-at time step t+1 (P_{t+1}*) **214** and updated state S_{t+1} **216**. Similarly, in C_{t+1} **220** I_{t+1} **218** is input into S_{t+1} **216**, the same trained weights **204** are applied, and the output of C_{t+1} **220** is P_{t+2}* **222**. As noted above S_{t+1} is updated from S_t, therefor S_{t+1} has memory from S_t from the previous time step. For example and not by way of limitation, this memory may include previous health-indicator data or previous other-factor data from one or more previous time steps. This process continues for n-steps, where I_{t+n} **224** is input into S_{t+n} **226** and the same weights **204** are applied. The output of cell C_{t+n} is prediction P_{t+n}*. Notably, the states are updated from previous time steps giving RNNs the benefit of memory from a previous state. This characteristic makes RNNs an alternative choice to make predictions on sequenced data for some embodiments. Though, and as described above, there are other suitable machine learning techniques for performing such predictions on sequenced data, including CNNs.

[0042] RNNs, like CNNs, can handle a string of data as input, and output a predicted string of data. A simple way to explain this aspect of using an RNN is using the example of natural language prediction. Take the phrase: The sky is blue. The string of words (i.e., data) has context. So as the state is updated, the string of data is updated from one iteration to the next, which provides context to predict blue. As just described RNNs have a memory component to aid in making predictions on sequenced data. However, the memory in the updated state of an RNN may be limited in how far it can look back, akin to short-term memory. When predicting sequenced data where a longer look back, akin to long term memory, is desired, tweaks to the RNNs just described may be used to accomplish this. A sentence, where the word to be predicted is unclear from the words closely preceding or surrounding, is again a simple example to explain: Mary speaks fluent French. It is unclear from the words closely preceding that French is the correct prediction; only that some language is the correct prediction, but which language? The correct prediction may lie in the context of words separated by a larger gap than the single string of words. Long Short Term Memory (LSTM) networks are a special kind of RNN, capable of learning these long(er)-term dependencies.

[0043] As described above, RNNs have a relatively simple repeating structure, for example they may have a single layer with a nonlinear activation function (e.g., tanh or sigmoid). LSTMs similarly have a chain like structure, but (for example) have four neural network layers, not one. These additional neural network layers give LSTMs the ability to remove or add information to the state (S) by using structures called cell gates. Id. FIG. **3** shows a cell **300** for a LSTM RNN. Line **302** represents the cell state (S), and can be viewed as an information highway; it is relatively easy for information to flow along the cell state unchanged. Id. Cell gates **304**, **306**, and **308** determine how much information to allow through the state, or along the information highway. Cell gate **304** first decides how much information to remove from the cell state S_t, so-called forget-gate layer. Id. Next, cell gate **306** and **306'** determines which information will be added to the cell state, and cell gate **308** and **308'** determines what will be output from the cell state as prediction P_{t+1}*. The information highway or cell state is now updated cell state S_{t+1} for use in the next cell. LSTMs permits RNNs to have a more persistent or long(er)-term memory. LSTMs provide additional advantages to RNN based machine learning models in that output predictions take into account a context separated from the input data by longer space or time, depending on how the data is sequenced, than the simpler RNN structure.

[0044] In some embodiments utilizing an RNN, the primary and secondary time sequences may not be provided to the RNN as vectors at each time step. Instead, the RNN may be provided only the current value of the primary and secondary time sequence(s), along with the future values or aggregate functions of the secondary time sequence(s) within the prediction interval. In this manner, the RNN uses the persistent state vector to retain information about the previous values for use in making predictions

[0045] Machine learning is well suited for continuous monitoring of one or multiple criteria to identify anomalies or trends, big and small, in input data as compared to training examples used to train the model. Accordingly, some embodiments described herein input a user's health-indicator data and optionally other-factor data into a trained machine learning model that predicts what a healthy person's health-indicator data would look like at the next time step and compares the prediction with the user's measured health-indicator data at the future time step. If the absolute value of the difference (e.g., loss as described below) exceeds a threshold, the user is notified his or her health-indicator data is not in a normal or healthy range. The threshold is a number set by the designer and, in some embodiments, may be changed by the user to allow a user to adjust the notification sensitivity. The machine learning model of these embodiments may be trained on health-indicator data alone or in combination with corresponding (in time) other-factor data from a population of healthy people, or trained on other training examples to suit the design needs for the model.

[0046] Data from health-indicators, like heart rate data, are sequenced data, and more particularly time sequenced data. Heartrate, for example and not by way of limitation, can be measured in a number of different ways, e.g., measuring electric signals from a chest strap or derived from a PPG signal. Some embodiments take the derived heartrate from the device, where each data point (e.g., heart rate) is produced at approximately equal intervals (e.g., 5 s). But, in some cases and in other embodiments the derived heart rate is not provided in roughly equal time steps, for example because the data needed for the derivation is not reliable (e.g., PPG signal is unreliable because the device moved or from light pollution). The same may be said of obtaining the secondary sequence of data from motion sensors or other sensors used to collect the other-factor data.

[0047] The raw signal/data (electric signal from ECG, chest strap, or PPG signals) itself is a time sequence of data that can be used in accordance with some embodiments. For the purpose of clarity, and not by way of limitation, this description uses PPG to refer to the data representing the health-indicator. The skilled artisan will readily appreciate that either form of the data for the health-indicator, raw data, waveform or number derived from raw data or waveform, may be used in accordance with some embodiments described herein.

[0048] Machine learning models that may be used with embodiments described herein include by way of example not limitation Bayes, Markov, Gausian processes, clustering algorithms, generative models, kernel and neural network algorithms. Some embodiments utilize a machine learning

model based on a trained neural network, other embodiments utilize a recurrent neural network, and additional embodiments use LTSM RNNs. For the purpose of clarity, and not by way of limitation, recurrent neural networks will be used to describe some embodiments of the present description.

[0049] FIGS. **4**A-**4**C show hypothetical plots against time for PPG (FIG. **4**A), steps taken (FIG. **4**B) and air temperature (FIG. **4**C). PPG is an example of health-indicator data, where steps, activity level, and air temperature are examples other-factor data for other factors that may impact the health-indicator data. As will be appreciated by the skilled artisan, the other-data may be obtained from any of many known sources including without limitation accelerometer data, GPS data, a weight scale, user entry etc., and may include without limitation air temperature, activity (running, walking, sitting, cycling, falling, climbing stairs, steps etc.), BMI, weight, height, age etc. The first dotted line running vertically across all three plots represents time t at which the user data is obtained for input into a trained machined learning model (discussed below). The hashed plot lines in FIG. **4**A represent predicted or probable output data **402**, and solid lines **404** in FIG. **4**A represent measured data. FIG. **4**B is a hypothetical plot of number of a user's steps at various times, and FIG. **4**C is a hypothetical plot of air temp at various times.

[0050] FIGS. **5**A-**5**B depict a schematic for a trained recurrent neural network **500** to receive the input data depicted in FIGS. **4**A-**4**C, i.e., PPG (P), steps (R) and air temperature (T). It is again emphasized that these input data (P, R and T) are merely examples of health-indicator data and other-factor data. It will also be appreciated that data for more than one health-indicator may be input and predicted, and more or less than two other-factor data may be used, where the choice depends on for what the model is being designed. It will be further appreciated by the skilled artisan that other-factor data is collected to correspond in time with the collection or measurement of the health-indicator data. In some cases, e.g. weight, other-factor data will remain relatively constant over certain periods of time.

[0051] FIG. **5**A depicts trained neural network **500** as a loop. P, T and R are input into state **502** of RNN **500**, where weights W are applied, and RNN **500** outputs predicted PPG **504** (P*). In step **506** the difference P-P* (ΔP*) is calculated, and at step **508** it is determined if |ΔP*| is greater than a threshold. If yes, step **510** notifies/alerts the user his/her health-indicator is outside the bounds/threshold predicted as normal or predicted for a healthy person. The alert/notification/detection could be, for example and not by way of limitation, a suggestion to see/consult a doctor, a simple notification like a haptic feedback, request to take additional measurement like and ECG, or simple note without recommendation, or any combination thereof. If |ΔP*| is less than or equal to the threshold, step **512** does nothing. In both steps **510** and **512** the process is repeated with new user data at the next time step. In this embodiment, the state is updated following the output of the predicted data, and may use the predicted data in updating the state.

[0052] In another embodiment, not shown, a primary sequence of heartrate data (e.g., derived from a PPG signal) and a secondary sequence of other-factor data are provided to the trained machine learning model, which may be an RNN a CNN, other machine learning models, or a combination of models. In this embodiment, the machine learning model is configured to receive as input at reference time t:

[0053] A. A vector (V_H) of length **300** of the last **300** health-indicator samples (e.g., heart rate in beats per minute) up to and including any health-indicator data at time t;

[0054] B. At least one vector (V_O) of length **300** containing the most recent other-factor data, e.g., step count, at the approximate time of each sample in V_H;

[0055] C. A vector (V_{TD}) of length **300** where the entry at index i, $V_{DT}(i)$, contains the time difference between the timestamps of health-indicator sample $V_H(i)$ and $V_H(i-1)$; and

[0056] D. A scalar prediction interval other-factor rate O_{rate} (step rate for example and not by way of limitation) representing the mean other-factor rate (e.g., step rate) measured over the time period from t to t+τ, where τ may be, for example and not by way of limitation, 2.5 minutes and is the future prediction interval.

[0057] The output of this embodiment may be, for example, a probability distribution characterizing the predicted heart rate measured over the time period from t to t+τ. In some embodiments, the machine learning model is trained with training examples that includes continuous time sequences of health-indicator data and other-factor data sequences. In one alternative embodiment the notification system assigns a timestamp to each predicted health-indicator (e.g., heart rate) distribution of t+τ/2, thus centering the predicted distribution within the predictive interval (τ). The notification logic, in this embodiment, then considers all samples within a sliding window (W) of length $W_L=2*(\tau)$ or 5 mins in this example and calculates three parameters:

[0058] 1. Mean value of all health-indicator sequence data $\overline{H_W}$ within the time window

[0059] 2. Mean value of all model predictions of the health-indicator $\overline{H_W^*}$, which predictions timestamp falls within the time window; and

[0060] 3. Median value of the root-mean-square of each predicted health-indicator distribution within the time window (RMS_W^H); where

[0061] 4. in one embodiment if $\overline{H_W}>\overline{H_W^*}+(\psi)\times RMS_W^H$ or $\overline{H_W}<\overline{H_W^*}-(\psi)\times RMS_W^H$ where ψ is a threshold, a notification is generated.

[0062] In this embodiment, an alert is generated when the measured health-indicator is more than a certain multiple of the standard deviation away from the mean of the predicted health-indicator values within a particular window W. The window W can be applied in a sliding fashion across the sequences of measured and predicted health-indicator values, with each window overlapping the previous window in time by a designer specified fraction, e.g., 0.5 mins.

[0063] The notification may take any number of different forms. For example and not by way of limitation, it may notify the user to obtain an ECG and/or blood pressure, it may direct the computing system (e.g. wearable etc.) to automatically obtain an ECG or blood pressure (for example), it may notify the user to see a doctor, or simply inform the user the health-indicator data is not normal.

[0064] The choice of V_{DT}, in this embodiment, as input into the model is intended to allow the model to utilize information contained in the variable spacing between health-indicator data in V_H, where the variable spacing may result from algorithms deriving health-indicator data from less than consistent raw data. For example, heart rate

samples are produced by the Apple Watch algorithm only when it has sufficiently reliable raw PPG data to output a reliable heart rate value, which results in irregular time gaps between heart rate samples. In similar fashion this embodiment utilizes the vector for other-factor data (V_O) with the same length as the other vectors to handle different and irregular sample rates between the primary sequence (health-indicator) and secondary sequence (other-factor). The secondary sequence, in this embodiment, is remapped or interpolated onto the same time points as the primary time sequence.

[0065] Furthermore, in some embodiments, the configuration of data from secondary time sequences presented as input to a machine learning model from a future prediction time interval (e.g. after t) may be modified. In some embodiments, the single scalar value containing the average other-factor data rate over the prediction interval, could be modified with multiple scalar values, e.g. one for each secondary time sequence. Or, a vector of values could be used over the prediction interval. Additionally, the prediction interval may itself be adjusted. A shorter prediction interval, for example, may provide faster response to changes and improved detection of events whose fundamental timescale is short(er), but may also be more sensitive to interference from sources of noise, like motion artifacts.

[0066] Similarly, the output prediction of the machine learning model itself does not need to be a scalar. For example some embodiments may generate a time series of predictions for multiple times t within the time interval between t and t+τ, and the alerting logic may compare each of these predictions with the measured value within the same time interval.

[0067] In this preceding embodiment, the machine learning model itself may comprise, for example, a 7-layer feed-forward neural network. The first 3 layers may be convolutional layers containing 32 kernels each with a kernel width of 24 and a stride of 2. The first layer may have as input the arrays V_H, V_O, and V_{TD}, in three channels. The final 4 layers may be fully-connected layers, all utilizing hyperbolic tangent activation functions except the last layer. The output of the third layer may be flattened into one array for input into the first fully connected layer. The final layer outputs 30 values parameterizing a Gaussian Mixture Model with 10 mixtures (mean, variance, and weight for each mixture). The network uses a skip connection between the first and third fully connected layers, such that the output of layer 6 is summed with the output of layer 4 to produce the input to layer 7. Standard batch normalization may be used on all layers but the last layer, with a decay of 0.97. The use of skip connections and batch normalization can improve the ability to propagate gradients through the network.

[0068] The choice of machine learning model may affect the performance of the system. The machine learning model configuration may be separated into two types of considerations. First is the model's internal architecture, meaning the choice of model type (convolutional neural network, recurrent neural network, random forests, etc. generalized nonlinear regression), as well as the parameters that characterize the implementation of the model (generally, the number of parameters, and/or number of layers, number of decision trees, etc.). Second is the model's external architecture—the arrangement of data being fed into the model and the specific parameters of the problem the model is being asked to solve. The external architecture may be characterized in part by the dimensionality and type of data being provided as input to the model, the time range(s) spanned by that data, and the pre-or-post processing done on the data.

[0069] Generally speaking, the choice of external architecture is a balance between increasing the number of parameters and amount of information provided as input, which may increase the predictive power of the machine learning model, with the available storage and computational capacity to train and evaluate a larger model, and the availability of sufficient amounts of data to prevent overfitting.

[0070] Numerous variations of the model's external architecture discussed in some embodiments are possible. The number of input vectors, as well as the absolute length (number of elements) and time span covered, may be modified. It is not necessary that each input vector be the same length or cover the same span of time. The data does not need to be equally sampled in time—for example and not by way of limitation, one might provide a 6-hour history of heart rate data, in which data less than one hour before t is sampled at a rate of 1 Hz, data more than 1 hour before t but less than 2 hours before t is sampled at a rate of 0.5 Hz, and data older than 2 hours is sampled at a rate of 0.1 Hz, where t is the reference time.

[0071] FIG. **5**B shows trained RNN **500** unrolled. Input data **513** (P_t, R_t, and T_t) is input into state-at-time t (S_t) **514** and trained weights **516** are applied. The output of cell (C_t) **518** is prediction-at-time t+1 (P_{t+1}*) **520** and updated state S_{t+1} **522**. Similarly, in C_{t+1} **524**, input data (P_{t+1}, R_{t+1}, and T_{t+1}) **513'** is input into S_{t+1} **522** and trained weights **516** are applied and the output of C_{t+1} **524** is P_{t+2}* **523**. As noted above S_{t+1} results from updating S_t, therefor S_{t+1} has memory from S_t from the operation in cell (C_t) **518** at the previous time step. This process continues for n-steps, where input data (P_n, R_n, and T_n) **513"** is input into S_n **530** and trained weights **516** are applied. The output of cell C_t is prediction **532** P_{n+1}*. Notably, trained RNNs apply the same weights throughout, but, and importantly, the states are updated from previous time steps giving RNNs the benefit of memory from a previous time step. The skilled artisan will appreciate that the order-in-time of inputting the dependent health-indicator data may vary and would still produce the desired result. For example, the measured health-indicator data from a previous time step (e.g., P_{t-1}) and the other-factor data from the current time step (e.g., R_t and T_t) can be input into the state at the current time step (S_t), where the model predicts the health-indicator at the current time step P_t*, which is compared to the measured health-indicator data at the present time step to determine if the user's health-indicator is normal or in a healthy range, as described above.

[0072] FIG. **5**C shows an alternative embodiment of a trained RNN to determine whether a user's health-indicator sequenced data, PPG in our example, is in a band or threshold for a healthy person. The input data in this embodiment is a linear combination $I_t=\alpha_t P_t^*+(1-\alpha_t)P_t$, where P_t^* is the predicted health-indicator value at time t and P_t is the measured health-indicator at time t. In this embodiment α ranges from 0-1 nonlinearly as a function of loss (L), where the loss and α are discussed in more detail below. What is worth noting now is when α is near zero, the measured data P_t is input into the network, and when α is near one, predicted data (P_t*) is input into the network for making a prediction at the next time step. Other-factor data (O_t) at time t may optionally also be input.

[0073] I_t and O_t are input into state S_t which, in some embodiments, outputs a probability distribution (β) of the predicted health-indicator data (P_{t+1}*) at time step t+1 ($\beta_{(P^*_{t+1})}^{t+1}$), where $\beta_{(P^*)}$ is the probability distribution function of predicted health-indicator (P*). In some embodiments, the probability distribution function is sampled to select a predicted health-indicator value at t+1 (P_{t+1}*). As appreciated by the skilled artisan $\beta_{(P^*)}$ may be sampled using different methods depending on the goals of the network designer, which methods may include taking the mean value, max value or a random sampling of the probability distribution. Evaluating β^{t+1} using the measured data at time t+1 provides the probability the state S_{t+1} would have predicted for the measured data.

[0074] To illustrate this concept, FIG. **5**D shows a hypothetical probability distribution for a range of hypothetical health-indicator data at time t+1. This function is sampled, for example at maximum probability 0.95, to determine a predicted health-indicator at time t+1 (P_{t+1}*). The probability distribution (β^{t+1}) is also evaluated using the measured or actual health-indicator data ($P_{t+1}^{act.}$), and a probability is determined that the model would have predicted if the actual data had been input into the model. In this example $\beta_{(P_{t+1}^{act.})}^{t+1}$ is 0.85.

[0075] A loss may be defined to help determine whether to notify a user his or her health status is not in a normal range as predicted by the trained machine learning model. The loss is chosen to model how close the predicted data is to the actual or measured data. The skilled artisan will appreciate many ways to define loss. In other embodiments described herein, for example, the absolute value of the difference between the predicted data and the actual data ($|\Delta P^*|$) is a loss. In some embodiments, the loss (L) may be $L=-1n[\beta_{(P)}]$, where $L_{t+1}=-1n[\beta_{(P_{t+1}^{act.})}^{t+1}]$. L is a measure of how close the predicted data is to the measured or actual data. $\beta_{(P)}$ ranges from 0 to 1, where 1 means the predicted value and measured value are the same. Therefore, a low loss indicates the predicted value is probably the same as or close to the measured value; in this context it means the measured data looks like it comes from a healthy/normal person. In some embodiments, thresholds for L are set, e.g., L>5, where the user is notified the health-indicator data is outside the range considered healthy. Other embodiments may take an average of losses over a period of time and compare the average to a threshold. In some embodiments, the threshold itself may be a function of a statistical calculation of the predicted values or an average of the predicted values. In some embodiments, the following equation may be used to notify the user the health-indicator is not in a healthy range:

$$|\langle P_{range}\rangle - \langle P_{range}^*\rangle| > f^{(\sigma_{(\langle P^*_{range}\rangle)})}$$

[0076] $\langle P_{range}\rangle$ is determined by a method of averaging the measured health-indicator data over a time range

[0077] $\langle P_{range}^*\rangle$ is determined by a method of averaging predicted health-indicator data over the same time range;

[0078] $(\sigma_{(\langle P^*_{range}\rangle)})$ is the median of the sequence of standard deviations derived from the network over the same time range; and

[0079] $f^{(\sigma_{(\langle P^*_{range}\rangle)})}$ is a function of the standard deviation evaluated at P_{range}* and may serve as the threshold.

[0080] The methods of averaging that may be used include, by way of example not limitation, average, arithmetic mean, median and mode. In some embodiments, outliers are removed so as not to skew the calculated number.

[0081] Referring back to the input data ($I_t=\alpha_t P_t^*+(1-\alpha_t)P_t$) for the embodiment depicted in FIG. **5**C, α_t is defined as a function of L and ranges from 0 to 1. For example, $\alpha(L)$ may be a linear function, or a non-linear function, or may be linear over some range of L and non-linear over a separate range of L. In one example, as shown in FIG. **5**E, the function $\alpha(L)$ is linear for L between 0 and 3, quadratic for L between 3 and 13, and 1 for L greater than 13. For this embodiment, when L is between 0 and 3 (i.e., when the predicted health-indicator data and measured health-indicator data nearly match), the input data I_{t+1} will be approximately the measured data P_{t+1}, as $\alpha-1$ will be near zero. When L is large, e.g., greater than 13, $\alpha(L)$ is 1, which makes the input data $I_{t+1}=P_{t+1}$*, the predicted health-indicator at time t+1. When L is between 1 and 13, $\alpha(L)$ varies quadratically, and the relative contributions of predicted and measured health-indicator data to the input data will also vary. The linear combination of predicted health-indicator data and measured health-indicator data weighted by $\alpha(L)$ permits, in this embodiment, weighting the input data between predicted and measured data at any particular time step. In all these examples the input data may also include the other-factor data (O_t). This is only one example of self-sampling, where some combination of predicted data and measured data are used as input to the trained network. The skilled artisan will appreciate many others may be used.

[0082] Machine learning models in embodiments use a trained machine learning model. In some embodiments, the machine learning models use a recurrent neural network, which requires a trained RNN. As an example, and not by way of limitation, FIG. **6** depicts an unrolled RNN to demonstrate training a RNN in accordance with some embodiments. Cell **602** has initial state S_0 **604** and weight matrix W **606**. Step-rate data R_0, air temperature data T_0 and initial PPG data P_0 at the time step zero are input into state S_0, weight W is applied, and a predicted PPG (P_1*) at the first time step is output from cell **602**, and ΔP_1* is calculated using PPG obtained at time step 1 (P_1). Cell **602** also outputs updated state at time step 1 **608** (S_1), which goes into cell **610**. Step rate data R_1, air temperature data T_1 and PPG data P_1 at time step 1 are input into S_1, weight **606** W is applied, and a predicted PPG (P_2*) at the time step 2 is output from cell **610**, and ΔP_2* is calculated using PPG (P_2) obtained at time step 2. Cell **610** also outputs updated state at time step 2 **612** (S_2), which goes into cell **614**. Step rate data R_3, air temperature data T_3 and PPG data at time step 3 (P_3) are input into S_2, weight **606** W is applied, and a predicted PPG (P_3*) at time step 3 is output from cell **614**, and ΔP_3* is calculated using PPG obtained at time step 3 (P_3). This is continued until state at time-step-n **616** is output and ΔP_{n+1}* is calculated. The ΔP^*'s are used in back propagation to adjust the weight matrix, similar to the training of convolutional neural networks. However, unlike convolutional networks, the same weight matrix in recurrent neural networks is applied at each iteration; it is only modified in back propagation during training. Many training examples with health-indicator data and corresponding other-factor data are input into RNN **600** over and over until it converges. As discussed previously, LTSM RNNs may be used in some embodiments where the states of such networks provide a longer term contextual analysis of input data, which may

provide better prediction when the network learns long(er)-term correlations. As also mentioned and the skilled artisan will readily appreciate other machine learning models will fall within the scope of embodiments described herein, and may include by way of example not limitation CNN or other feed-forward networks.

[0083] FIG. **7**A depicts a system **700** that predicts whether a user's measured health-indicators are within or outside a threshold of normal for that of a healthy person under similar other-factors. System **700** has machine learning model **702** and health detector **704**. Embodiments for machine learning model **702** include a trained machine learning model, a trained RNN, CNN or other feed forward network for example (and not by way of limitation). The trained RNN, other network or combination of networks may be trained on training examples from a population of healthy people from whom health-indicator data and corresponding (in time) other-factor data has been collected. Alternatively, the trained RNN, other network or combination of networks may be trained on training examples from a particular user, making it a personalized trained machine learning model. The skilled artisan will appreciate training examples from different populations may be selected depending on the use or design for the trained network and system in general. The skilled artisan will also readily appreciate that the health-indicator data in this and other embodiments may be one or more health-indicators. For example and not by way of limitation, one or more of PPG data, heartrate data, blood pressure data, body temperature data, blood oxygen concentration data and the like could be used to train the models and to predict the health of a user. Health detector **704** uses prediction **708** from machine learning model **702** and input data **710** to determine whether a loss, or other metric determined by analyzing the predicted output with the measured data, exceeds a threshold considered normal and thus unhealthy. System **700** then outputs a notification or the state of a user's health. This notification may take many forms as discussed herein. Input generator **706** continuously obtains data with a sensor (not shown) from a user wearing or in contact with the sensor, where the data represents one or more health-indicators of the user. Corresponding (in time) other-factor data may be collected by another sensor or acquired through other means as described herein or as readily apparent to the skilled artisan.

[0084] Input generator **706** may also collect data to determine/calculate other-factor data. Input generator, for example and not by way of limitation, may include a smart watch, wearable or mobile device (e.g., Apple Watch® or FitBit® smart phone, tablet or laptop computer), a combination of smart watch and mobile device, a surgically implanted device with the ability to transmit data to a mobile device or other portable computing device, or a device on a cart in a medical care facility. Preferably user input generator **706** has a sensor (e.g., PPG sensor, electrode sensor) to measure data related to one or more health-indicators. The smart watch, tablet, mobile phone or laptop computer of some embodiments may carry the sensor or the sensor may be remotely placed (surgically embedded, contacted to the body remote from the mobile device, or some separate device) where, in all these cases, the mobile device communicates with the sensor in order to gather health-indicator data. In some embodiments, system **700** may be provided on the mobile devices alone, in combination with other mobile devices, or in combination with other computing systems via communication through a network through which these devices may communicate. For example and not by way of limitation, system **700** may be a smart watch or wearable with machine learning model **702** and health detector **704** located on the device, e.g., the memory of the watch or firmware on the watch. The watch may have user input generator **706** and communicate with other computing devices (e.g. mobile phone, tablet, lap top computer or desk top computer) via direct communication, wireless communication (e.g., WiFi, sound, Bluetooth, etc) or through a network (e.g., internet, intranet, extranet etc.) or a combination thereof, where trained machine learning model **702** and health detector **704** may be located on the other computing devices. The skilled artisan will appreciate that any number of configurations of system **700** may be utilized without exceeding the scope of embodiments described herein.

[0085] Referring to FIG. **7**B smart watch **712**, in accordance with an embodiment, is depicted. Smart watch **712** includes watch **714** which contains all the circuitry and microprocessors, and processing devices (not shown) known to the skilled artisan. Watch **714** also includes display **716**, on which a user's health-indicator data **718** may be displayed, in this example heart rate data. Also displayed on display **716** may be the predicted health-indicator band **720** for the normal or the healthy population. In FIG. **7**B the user's measured heart rate data does not exceed the predicted healthy band, so in this particular example no notification would be made. Watch **714** may also include watch band **722**, and high-fidelity sensor **724**, for example an ECG sensor. Alternatively, watch band **722** may be an expandable cuff to measure blood pressure. Low-fidelity sensors **726** (shown in shadow) are provided on the back of watch **714** to collect user health-indicator data, such as PPG data, which can be used to derive heart rate data or other data like blood pressure, for example. Alternatively, as will be appreciated by the skilled artisan, a fitness band may be used in some embodiments, such as FitBit or Polar, where the fitness bands have similar processing power and other-factor measurement devices (e.g., ppg and accelerometers).

[0086] FIG. **8** depicts an embodiment of a method **800** for continuously monitoring a user's health status. Step **802** receives the user input data, which may include data for one or more health-indicators (aka primary sequence of data) and corresponding (in time) data for other-factors (aka secondary sequence of data). Step **804** inputs the user data into a trained machine learning model, which may include a trained RNN, CNN, other feed-forward network as described herein or other neural network known to the skilled artisan. In some embodiments, the health-indicator input data may be one or a combination of predicted health-indicator data and measured health-indicator data, e.g., a linear combination, as described in some embodiments herein. Step **806** outputs data for one or more predicted health-indicators at a time step, which outputs may include, by way of example not limitation, a single predicted value, a probability distribution as a function of predicted values. Step **808** determines a loss based on the predicted health-indicator, where, for example and not by way of limitation, the loss may be a simple difference between predicted and measured health-indicators, or some other appropriately selected loss function (e.g. negative log of a probability distribution evaluated at the value for the measured health-indicator). Step **810** determines if the loss

exceeds a threshold considered normal or unhealthy, where the threshold may be, for example and not by way of limitation, a simple number picked by the designer, or a more complex function of some parameter related to the prediction. If greater than the threshold, step **812** notifies the user that his or her health indicator exceeds a threshold considered normal or healthy. The notification, as described herein, may take many forms. In some embodiments, this information may be visualized to the user. For example and not by way of limitation the information can be displayed on a user interface such as a graph that shows (i) measured health-indicator data (e.g., heart rate) and other-factor data (e.g., step count) as a function of time, (ii) a distribution of predicted health-indicator data (e.g., predicted heart rate values) generated by the machine learning model. In this way, the user can visually compare the measured data points to the predicted data points and determine by visual inspection whether their heart rate, for example, falls into the range expected by the machine learning model.

[0087] Some embodiments described herein have mentioned using a threshold to determine whether to notify a user or not. In one or more of these embodiments, the user may change the threshold to adjust or tune the system or method to more closely match the user's personal health knowledge. For example, if the physiological indicator used is blood pressure and the user has higher blood pressure, then embodiments may frequently alert/notify the user that his health-indicator is outside normal or healthy range from a model trained on a healthy population. Thus, certain embodiments permit the user to increase the threshold value so the user is not notified so frequently that his/her health-indicator data exceeds what is considered normal or healthy.

[0088] Some embodiments preferably use the raw data for the health-indicators. If the raw data is processed to derive a specific measurement, e.g., heart rate, this derived data may be used in accordance with embodiments. In some situations, the provider of a health monitoring apparatus does not have control of the raw data, rather what is received is processed data in the form of a calculated health-indicator, e.g., heart rate or blood pressure. As will be appreciated by the skilled artisan, the form of the data used to train a machine learning model should match the form of the data collected from the user and input into the trained model, otherwise the predictions could prove erroneous. For example, the Apple Watch gives heart rate measurement data at unequal time steps, and does not provide raw PPG data. In this example, a user wears an Apple Watch that outputs heart rate data in accordance with Apple's PPG processing algorithm with heart rate data at unequal time steps. The model is trained on this data. Apple deciding to change its algorithm for providing the heart rate data may render the model trained on data from the previous algorithm obsolete to use on data input from the new algorithm. To account for this potential issue, some embodiments resample the irregularly spaced data (heart rate, blood pressure data, or ECG data etc.) onto a regularly spaced grid and sample from regularly spaced grid when collecting data to train the model. If Apple, or other supplier of data, changes its algorithm, the model needs only to be retrained on newly collected training examples, but the model does not need to be reconstructed to account for the algorithm change.

[0089] In a further embodiment, the trained machine learning model may be trained on the user's data, resulting in a personalized trained machine learning model. This trained personalized machine learning model can be used in place of or in combination with the machine learning models trained on a healthy population of people described herein. If used by itself, a user's data is input into the personalized trained machine learning model, which would output a prediction of that individual's health-indicator in the next time step that is normal for that user, which is then compared with the actual/measured data from the next time step in a manner consistent with embodiments described herein to determine whether the user's health-indicators had differed by some threshold from what is predicted normal for that user. In addition, this personalized machine learning model could be used in combination with the machine learning model trained on training examples from a population of healthy people to generate predictions and associated notifications as related to both what is predicted normal for that individual user and predicted normal for the healthy population of people.

[0090] FIG. **9**A depicts a method **900** in accordance with another embodiment, and FIG. **9**B shows a hypothetical plot **902** of heart rate (by way of example not limitation) as a function of time for the purpose of explanation. Step **904** (FIG. **9**A) receives user heart rate data (or other health-indicator data) and, optionally, corresponding (in time) other-factor data, and inputs this data into a personalized-trained machine learning model. In some embodiments, the personalized-trained model is trained on the user's individual health-indicator data and, optionally, corresponding (in time) other-data as described herein. Thus, in step **906** the personalized-trained machine learning model predicts normal heart rate data for that individual user under conditions of the other-factor(s), and step **908** identifies aberrations or anomalies in the user's health-indicator data as compared to what is predicted as normal for that particular user. Some embodiments receive the user's health-indicator data from a wearable device (e.g., Apple Watch, smart watch, FitBit®, etc.) on the user, or from another mobile device (e.g., tablet, computer, etc.) in communication with a sensor on the user (e.g., Polar® strap, PPG sensor etc.), which is discussed throughout this description.

[0091] A loss may be defined to help determine whether to notify a user, in step **908**, that the user's measured data is anomalous to what is predicted as normal for that particular user. The loss is chosen to model how close the prediction is to the actual or measured data. The skilled artisan will appreciate many ways to define loss. In other embodiments described herein and equally applicable here, for example, the absolute value of the difference between the predicted value and the absolute value $|\Delta P^*|$ is a form of a loss. In some embodiments, the loss (L) may be $L=-1n[\beta_{(P)}]$, where $L_{t+1}=1n[\beta_{(P_{t+1}{}^{act})}{}^{t+1}]$. L, generally, is a measure of how close the predicted data is to the measured data. $\beta_{(P)}$, the probability distribution in this example, ranges from 0 to 1, where 1 means the predicted data and measured data are the same. Therefore, a low loss, in some embodiments, indicates the predicted data are probably the same as or close to the measured data. In some embodiments, thresholds for L are set, e.g., L>5, where the user is notified an anomalous condition exists from that predicted for that particular user. This notification may take many forms, as described elsewhere herein. As also described elsewhere herein, other embodiments may take an average of losses over a period of time and compare the average to a threshold. In some embodiments, as described in more detail elsewhere herein,

the threshold itself may be a function of a statistical calculation of the predicted data or an average of the predicted data. Loss has been described in more detail elsewhere herein, and for the sake of brevity will not be discussed further here. The skilled artisan will also appreciate the input and predicted data may be scalar values, or segments of data over a time period. For example and not by way of limitation, a system designer may be interested in 5 minute data segments, and would input all the data prior to time t and all other-data for t+5 min, predict the health-indicator data for t+5 mins and determine a loss between measured health-indicator data for the t+5 min segment against the predicted health-indicator data for the t+5 min segment.

[0092] Step **908** determines if an anomaly is present or not. As discussed this may be determined if the loss exceeds a threshold. As previously described, the threshold is set by choice of the designer and based on the purpose of the system being designed. In some embodiments the threshold may be modified by the user, but preferably not so in this embodiment. If an anomaly is not present, the process is repeated at step **904**. If an anomaly is present, step **910** notifies or alerts the user to obtain a high-fidelity measurement, an ECG or blood pressure measurement for example and not by way of limitation. In step **912**, the high-fidelity data is analyzed by an algorithm, a health professional or both and is described as normal or not normal, and if not normal some diagnosis may be assigned, e.g., AFib, tachycardia, bradycardia, atrial flutter, or high/low blood pressure depending on the high-fidelity measurement obtained. It is noted for clarity, that notification to record high-fidelity data is equally applicable and possible in other embodiments, and in particular embodiments using general models described above. The high-fidelity measurement, in some embodiments, may be obtained directly by the user using a mobile monitoring system, such as ECG or blood pressure systems, which may be associated with the wearable device in some embodiments. Alternatively, the notification step **910** causes automatic acquisition of the high-fidelity measurement. For example, the wearable device may communicate with a sensor (hard-wired or via wireless communication) and obtain ECG data, or it may communicate with a blood pressure cuff-system (e.g., wrist band of a wearable or an armband cuff) to automatically obtain a blood pressure measurement, or it may communicate with an implanted device such as a pace maker or ECG electrodes. Systems for remotely obtaining an ECG are provided, for example, by AliveCor, Inc., such systems include (without limitation) one or more sensors contacting the user in two or more locations, where the sensor collects electrical cardiac data that is transmitted, either wired or wirelessly, to a mobile computing device, where an app generates an ECG strip from the data, which can be analyzed by algorithms, a medical professional or both. Alternatively, the sensor may be a blood pressure monitor, where the blood pressure data are transmitted, either wired or wirelessly, to the mobile computing device. The wearable itself may be a blood pressure system having a cuff with ability to measure health-indicator data and optionally with an ECG sensor similar to that described above. The ECG sensor may also include an ECG sensor such as that described in co-owned U.S. Provisional Application No. 61/872,555, the contents of which is incorporated herein by reference. The mobile computing device may be, for example and not by way of limitation, a computer tablet (e.g., iPad), smart phone (e.g., iPhone®), wearable (e.g., Apple Watch) or a device (maybe mounted on a cart) in a healthcare facility. The mobile computing device could be, in some embodiments, a laptop computer or a computer in communication with some other mobile device. The skilled artisan will appreciate that a wearable or smartwatch will also be considered mobile computing devices in terms of the capabilities provided in the context of embodiments described herein. In the case of a wearable, the sensor may be placed on the band of the wearable where the sensor may transmit the data wirelessly or by wire to the computing device/wearable, or the band may also be a blood pressure monitoring cuff, or both as previously described. In the case of a mobile phone, the sensor may be pads attached to or remote from the phone, where the pads sense electrical cardiac signals and wirelessly or by hardwire communicate the data to the wearable or other mobile computing device. More detailed descriptions for some of these systems are provided in one or more of U.S. Pat. Nos. 9,420,956; 9,572,499; 9,351,654; 9,247,911; 9,254,095; and 8,509,882 and one or more of US Patent Application Publication Numbers 2015/0018660; 2015/0297134; and 2015/0320328, all of which are incorporated herein in their entirety and for all purposes. Step **912** analyzes the high-fidelity data and provides a description or diagnosis, as previously described.

[0093] In step **914**, diagnosis or categorization of the high-fidelity measurement is received by a computing system, which may be in some embodiments the mobile or wearable computing system used to collect the user's heart rate data (or other health-indicator data), and in step **916** the low-fidelity health-indicator data sequence (heart rate data in this example) is labeled with the diagnosis. In step **918**, the labeled user's low-fidelity data sequence is used to train a high-fidelity machine learning model, and optionally other-factor data sequence is also provided to train the model. The trained high-fidelity machine learning model, in some embodiments, has the capability to receive measured low-fidelity health-indicator data sequence (e.g., heart rate data or PPG data) and optionally other-factor data and give a probability or predict or diagnose or detect when a user is experiencing an event typically diagnosed or detected using high-fidelity data. The trained high-fidelity machine learning model is able to do this because it has been trained on user's health-indicator data (and optionally other-factor data) labeled with diagnoses of the high-fidelity data. Thus, the trained model has the ability to predict when a user is having an event associated with one or more of the labels (e.g., Afib, high blood pressure etc.) solely based on measured low-fidelity health-indicator input data sequence, e.g. heart rate or ppg data (and optionally other-factor data). As the skilled artisan will appreciate, the training of the high-fidelity model can take place on the user's mobile device, remote from the user's mobile device, a combination of the two, or in a distributed network. For example and not by way of limitation, the user's health-indicator data could be stored in a cloud system, and this data can be labeled in the cloud using the diagnosis from step **914**. The skilled artisan will readily appreciate any number of ways and manners to store, label and access this information. Alternatively, a global trained high-fidelity model could be used, which would be trained on labeled training examples from a population of people experiencing these conditions typically diagnosed or detected with high-fidelity measurements. These global training examples would provide low-fidelity data

sequences (e.g., heart rate) labeled with conditions diagnosed using a high-fidelity measurement (e.g., Afib called from a ECG by a medical professional or an algorithm).

[0094] Referring now to FIG. **9**B, plot **902** shows a schematic of heart rate plotted as a function of time. Aberrations **920** from the user's normal heart rate data occurred at times t_1, t_2, t_3, t_4 t_5, t_6, t_7, t_8. Normal, as described above, means that the predicted data for this particular user was within a threshold of the measured data, where the aberrations are outside the threshold. At aberrations from normal some embodiments prompt the user to obtain a more definitive or high-fidelity reading, by way of example not limitation an ECG reading, identified as ECG_1, ECG_2, ECG_3, ECG_4, ECG_5, ECG_6, ECG_7, ECG_8. As described above the high-fidelity reading could be automatically obtained, the user may obtain it, and it could be things other than an ECG, e.g., blood pressure. High-fidelity readings are analyzed by algorithm, health professional or both to identify the high-fidelity data as normal/abnormal and to further identify/diagnose abnormal, AFib for example and not by way of limitation. This information is used to label the health-indicator data (e.g., heart rate or PPG data) at the point(s) of anomaly **920** in the user's sequenced data.

[0095] The distinction between high-fidelity and low-fidelity data is one where high-fidelity data or measurements are typically used to make a determination, detection or diagnosis, where low-fidelity data cannot readily be used for such. For example, an ECG scan may be used to identify, detect or diagnose arrhythmias, whereas heart rate or PPG data do not typically provide this capability. As the skilled artisan will appreciate, the description herein relating to machine learning algorithms (e.g., Bayes, Markov, Gausian processes, clustering algorithms, generative models, kernel and neural network algorithms) apply equally to all embodiments described herein.

[0096] In some situations users remain asymptomatic despite that issues may be present, and even if symptoms present it may be impractical to obtain the high-fidelity measurement necessary to make a diagnosis or detection. For example and not by way of limitation, arrhythmias particularly AF may not present and even when symptoms do present it is notoriously difficult to record an ECG at that moment, and without expensive, bulky and sometimes invasive monitoring devices it is incredibly difficult to continuously monitor the user. As discussed elsewhere herein, it is important to understand when a user experiences AF because AF, at a minimum, may be a causal factor in stroke among other serious conditions. Similarly, and as discussed elsewhere, AF burden may have similar import. Some embodiments allow for continuous monitoring of arrhythmias (e.g., AF) or other serious conditions using only the continuous monitoring of low-fidelity health-indicator data, such as heart rate or ppg along with optional other-factor data.

[0097] FIG. **10** depicts a method **1000** in accordance with some embodiments of health monitoring systems and methods. Step **1002** receives measured or actual user low-fidelity health-indicator data (e.g., heart rate or PPG data from a sensor on a wearable), and optionally receives corresponding (in time) other-factor data, which may impact the health-indicator data as described herein. As discussed elsewhere herein the low-fidelity health-indicator data may be measured by a mobile computing device, such as a smart watch, other wearable, or computer tablet. In step **1004**, the user's low-fidelity health-indicator data (and optionally the other-factor data) is input into a trained high-fidelity machine learning model, which, in step **1006**, outputs a predicted identification or diagnosis for the user based on the measured low-fidelity health-indicator data (and optionally corresponding (in time) other-factor data). Step **1008** asks if the identification or diagnosis is normal, which, if yes, the process starts over. If the identification or diagnosis is not normal, step **1010** notifies the user of the problem or detection. Optionally, the system, method or platform may be set up to notify any combination of the user, family, friends, healthcare professionals, emergency 911, or the like. Which of these people are notified may depend on the identification, detection or diagnosis. If the identification, detection or diagnosis is life threatening, then certain people may be contacted or notified that may not be notified if the diagnosis is not life threatening. In addition, in some embodiments, the measured health-indicator data sequence is input into the trained high-fidelity machine learning model and the amount of time a user is experiencing an abnormal event (e.g., difference between onset and cessation of the predicted abnormal event) is calculated, permitting a better understanding of the abnormal burden on the user. In particular, AF burden may be highly important to understand in preventing stroke and other serious conditions. Thus, some embodiments allow continuous monitoring of abnormal events with a mobile computing device, a wearable computing device or other portable device capable of only acquiring low-fidelity health-factor data, and optionally other-factor data.

[0098] FIG. **11** depicts example data **1100** analyzed based on low-fidelity data to generate a high-fidelity output prediction or detection, according to some embodiments as described herein. While described with reference to detection of atrial fibrillation, similar data may be generated for additional predictions of high-fidelity diagnosis based on low-fidelity measurements. The first chart **1110** shows heart rate calculations over time for a user. The heart rate may be determined based on PPG data or other heart rate sensors. The second chart **1120** shows activity data for a user during the same time period. For example, the activity data may be determined based on step count, or other measurements of movement of the user. The third chart **1130** shows a classifier output from a machine learning model and a horizontal threshold for when a notification is generated. A machine learning model may generate the prediction based on an input of low-fidelity measurements. For example, the data in the first chart **1110** and the second chart **1120** may be analyzed by a machine learning system as described further above. The result of the machine learning system analysis may be provided as the atrial fibrillation probability shown in chart **1130**. When the probability is over a threshold value, shown in this case as above 0.6 confidence, a health monitoring system can trigger a notification or other alert for the user, a physician, or other users associated with the user.

[0099] In some embodiments, the data in charts **1110** and **1120** may be provided as continuous measurements to a machine learning system. For example, the heart rate and activity levels may be generated as measurements every 5 seconds in order an accurate measurement. A segment of time with multiple measurements can then be input to a machine learning model. For example, the previous hour of data can be used as an input to the machine learning model. In some embodiments, shorter or longer periods of time may

be provided rather than one hour. As shown in FIG. **11**, the output chart **1130** provides an indication of periods of time in which a user is undergoing an abnormal health event. For example, the periods when the prediction is over a certain confidence level may be used by a health monitoring system to determine atrial fibrillation. This value can then be used to determine an atrial fibrillation burden on the user during the measured time period.

[0100] In some embodiments, a machine learning model to generate the predicted output in chart **1130** may be trained based on labeled user data. For example, the labeled user data may be provided based on high-fidelity data (such as an ECG reading) taken at a time period when low-fidelity data (e.g., PPG, heart rate) and other data (e.g., activity level or steps) is also available. In some embodiments, the machine learning model is designed to determine if there was likely atrial fibrillation during a preceding time period. For example, the machine learning model may take an hour of low-fidelity data as an input and provide a likelihood there was an event. Accordingly, training data may include hours of recorded data for a population of individuals. The data can be health-event-labeled-times when a condition was diagnosed based on high-fidelity data. Accordingly, if there was a health-event labeled time based on high-fidelity data, the machine learning model may determine that any one hour window of low-fidelity data with that event that is input into the untrained machine learning model should provide a prediction of the health-event. The untrained machine learning model can then be updated based on comparing the prediction with the label. After repeating for a number iterations and determining that the machine learning model has converged, it may be used by a health monitoring system to monitor for atrial fibrillation of users based on low-fidelity data. In various embodiments, other conditions than atrial fibrillation may be detected using low-fidelity data.

[0101] FIG. **12** depicts a method **1200** in accordance with some embodiments as described herein The method **1200** may be performed by processing logic that comprises hardware (e.g., circuitry, dedicated logic, programmable logic, microcode, etc.), software (e.g., instructions run on a processing device to perform hardware simulation), or a combination thereof. In one embodiment, processing logic is executed by a kernel of an operating system associated with the hardware described herein. It should be noted that the operations of the methods described herein may be performed in any order and combination.

[0102] In one embodiment, at step **1202**, processing logic receives measured or actual user low-fidelity health-indicator data (e.g., heart rate or PPG data from a sensor on a wearable) at a first time, and optionally receives corresponding (e.g., in time) other-factor data, which may impact the health-indicator data, as described herein. In one embodiment, an example of other-factor data may include an activity level of the user.

[0103] As discussed herein, the low-fidelity health-indicator data may be collected and measured by a low-fidelity health-indicator data sensor (e.g., a PPG) and a corresponding mobile computing device, such as a smartwatch, other wearable (e.g., a fitness band), computer tablet, a laptop computer, etc. In one embodiment, the low-fidelity health-indictor data sensor is located on the mobile computing device. In other embodiments, the low-fidelity health-indictor data sensor is remotely located from the mobile computing device, and operatively coupled via a communication network (e.g., Bluetooth, WiFi, Near-field communication (NFC), etc.) or wires.

[0104] At step **1204**, the user's low-fidelity health-indicator data (and optionally the other-factor data) is input, by processing logic, into a trained high-fidelity machine learning model, which outputs a predicted identification or diagnosis for the health (e.g., cardiac) of the user based on the measured low-fidelity health-indicator data (and optionally corresponding (e.g., in time) other-factor data). For example, processing logic may input a set of data including the measured low-fidelity health-indicator data, and optionally including the other-factor data, into a trained high-fidelity machine learning model, to generate a prediction whether the health of the user outside a normal range, based on a low-fidelity health-indicator threshold. In various embodiments, the trained high-fidelity machine learning model may include one or more of: a generative neural network, a recurrent neural network (RNN), a feed-forward neural network, etc.

[0105] At step **1206**, processing logic determines whether the prediction is outside the normal range. In one embodiment, the prediction is based on a comparison of a probability being outside the normal range to a threshold value (e.g., a sensitivity threshold) over a period of time (e.g., a time threshold). For example, to determine that the prediction is outside the normal range, processing logic may determine that a probability that the prediction is outside the normal range is above a first sensitivity threshold for a first time threshold. In various embodiments, it is contemplated that variations of the sensitivity threshold value, such as an average value over a period of time, may be evaluated over the first time threshold. In such cases, suitable variations of the sensitivity threshold value, in addition to the raw sensitivity threshold value, are considered to be represented by the term "sensitivity threshold" described herein. Furthermore, for the purposes of the present disclosure, it is contemplated that the low-fidelity health-indicator threshold includes one or more of: a sensitivity threshold or a time threshold.

[0106] If it is determined at step **1206** that the prediction (e.g., the predicted identification or diagnosis) is inside the normal range, the process starts over at step **1202**. If the prediction is outside the normal range (e.g., the prediction includes an indication of atrial fibrillation, for example), at step **1208** processing logic automatically receives measured high-fidelity health-indicator data obtained by a high-fidelity health-indicator data sensor (e.g., an ECG sensor). Alternatively, the processer may alert the user to obtain a measurement from the high-fidelity data sensor, which measured high-fidelity health-indicator data is received by the processing logic as described above. Processing logic may optionally automatically send a request for the measurement and receipt of such high-fidelity health-indicator data (e.g., to the high-fidelity health-indicator data sensor) upon determining that the prediction is outside the normal range.

[0107] At step **1210**, processing logic determines whether the measured high-fidelity health- indicator data is normal or abnormal, i.e., whether the measured high-fidelity health-indicator data indicates that the health of the user is normal or abnormal. If the measured high-fidelity health-indicator data is abnormal, processing logic may notify the user of the abnormality and continue to step **1202**. Optionally, processing logic may be set up to notify any combination of the user,

family, friends, healthcare professionals, emergency 911, or the like, of the abnormality as well as any additional suitable information, such as the location of the user. Which of these people are notified, and what information is sent, may depend on the specific identification, detection or diagnosis. If the identification, detection or diagnosis is life threatening, then certain people may be contacted or notified that may not be notified if the diagnosis is not life threatening, for example. In addition, in some embodiments, the measured health-indicator data sequence is input into the trained high-fidelity machine learning model and the amount of time a user is experiencing an abnormal event (e.g., difference between onset and cessation of the predicted abnormal event) is calculated, permitting a better understanding of the abnormal burden on the user. In particular, AF burden may be important to understand in preventing stroke and other serious conditions. Thus, some embodiments may allow continuous monitoring of abnormal events with a mobile computing device, a wearable computing device or other portable device capable of acquiring low-fidelity health-factor data, and optionally other-factor data.

[0108] If the measured high-fidelity health-indicator data is normal at step **1210**, processing logic modifies, by a processing device, the low-fidelity health-indicator threshold at step **1212**. In one embodiment, the low-fidelity health-indicator threshold corresponds to the first sensitivity threshold, and to modify the low-fidelity health-indicator threshold processing logic modifies the first sensitivity threshold to a second sensitivity threshold (e.g., a first value of the first sensitivity threshold to a second value) to decrease a notification sensitivity (e.g., that determines when a user is notified of a possible abnormality). For example, in one embodiment, processing logic may increase the first sensitivity threshold to a higher second sensitivity threshold, and thus decrease the notification sensitivity so that the user receives fewer notifications of abnormalities, which notification may be more indicative of when a user is actually experiencing an abnormality. In one embodiment, the modification to the low-fidelity health-indicator threshold may be temporary, and upon the expiration of a time interval (e.g., one day, one week, one month, etc.), the value corresponding to the low-fidelity health-indicator threshold may be modified back to the original value, or some other value. For example, processing logic may modify the second sensitivity threshold back to the first sensitivity threshold in response to the expiration of a time interval.

[0109] In another embodiment, the low-fidelity health-indicator threshold corresponds to the first time threshold, and to modify the low-fidelity health-indicator threshold processing logic may modify the first time threshold to a second time threshold. It should be noted that there are a multitude of methods by which a new low-fidelity health-indicator threshold may be determined, and the present disclosure provides some non-limiting examples of such methods. For example, the new low-fidelity health-indicator threshold may be determined randomly, as a percentage of a previous low-fidelity health-indicator threshold, based on historical data, based on a selection from predetermined low-fidelity health-indicator thresholds, based on a machine learning model, based on user-specific data, based on general-population data, etc. The system and methods described herein are equally capable of utilizing additional, equally suitable methods for determining a new low-fidelity health-indicator threshold.

[0110] In one embodiment, modification of the sensitivity threshold may not done in real-time on the wearable device, but instead done asynchronously on some other processing system (e.g. in the cloud) and the new sensitivity thresholds may be downloaded to the device. Similarly, on the device itself, when determining if and by how much to modify the sensitivity threshold, processing logic may analyze not just the immediately prior notification and high-fidelity measurement, but might also consider multiple prior notifications and measurements (e.g. each notification and ECG recorded in the past week, for example), and modify the sensitivity threshold by applying a formula that considers all, or some portion, of that data.

[0111] It should be noted that each of the operations described with respect to method **1200** may be performed automatically (e.g., without human intervention). Specifically, the modification of the low-fidelity health-indicator threshold may be done automatically, upon one or more determinations that particular conditions have been met, as described herein. In one non-limiting example of detecting atrial fibrillation from a PPG, a globally-trained high-fidelity machine learning model operating on PPG data may respond differently to different individuals due to biological variation, skin color, etc. Advantageously, this mechanism allows the model to initially be set at high sensitivity so as to ensure abnormal conditions are detected, and to dynamically adjust its sensitivity downwards (e.g., increasing specificity) in order to adjust to the specific requirements and/or particularities of each individual. Advantageously, this model may lead to a higher population sensitivity/specificity than a global model without such adjustments.

[0112] FIG. **13** illustrates a diagrammatic representation of a machine in the example form of a computer system **1300** within which a set of instructions, for causing the machine to perform any one or more of the methodologies discussed herein, may be executed. In alternative embodiments, the machine may be connected (e.g., networked) to other machines in a local area network (LAN), an intranet, an extranet, or the Internet. The machine may operate in the capacity of a server or a client machine in a client-server network environment, or as a peer machine in a peer-to-peer (or distributed) network environment. The machine may be a personal computer (PC), a tablet PC, a set-top box (STB), a Personal Digital Assistant (PDA), a cellular telephone, a web appliance, a server, a network router, a switch or bridge, a hub, an access point, a network access control device, or any machine capable of executing a set of instructions (sequential or otherwise) that specify actions to be taken by that machine. Further, while only a single machine is illustrated, the term "machine" shall also be taken to include any collection of machines that individually or jointly execute a set (or multiple sets) of instructions to perform any one or more of the methodologies discussed herein. In one embodiment, computer system **1300** may be representative of a server, mobile computing device, wearable, or the like configured to perform health monitoring as described herein.

[0113] The exemplary computer system **1300** includes a processing device **1302**, a main memory **1304** (e.g., read-only memory (ROM), flash memory, dynamic random access memory (DRAM)), a static memory **1306** (e.g., flash memory, static random access memory (SRAM), etc.), and a data storage device **1318**, which communicate with each other via a bus **1330**. Any of the signals provided over various buses described herein may be time multiplexed

with other signals and provided over one or more common buses. Additionally, the interconnection between circuit components or blocks may be shown as buses or as single signal lines. Each of the buses may alternatively be one or more single signal lines and each of the single signal lines may alternatively be buses.

[0114] Processing device **1302** represents one or more general-purpose processing devices such as a microprocessor, central processing unit, or other processing device. More particularly, the processing device may be complex instruction set computing (CISC) microprocessor, reduced instruction set computer (RISC) microprocessor, very long instruction word (VLIW) microprocessor, or processor implementing other instruction sets, or processors implementing a combination of instruction sets. Processing device **1302** may also be one or more special-purpose processing devices such as an application specific integrated circuit (ASIC), a field programmable gate array (FPGA), a digital signal processor (DSP), network processor, or the like. The processing device **1302** is configured to execute processing logic **1326**, which may be one example of a health-monitor **1350** and related systems for performing the operations and steps discussed herein.

[0115] The data storage device **1318** may include a machine-readable storage medium **1328**, on which is stored one or more set of instructions **1322** (e.g., software) embodying any one or more of the methodologies of functions described herein, including instructions to cause the processing device **1302** to execute a health-monitor **1350** and related processes as described herein. The instructions **1322** may also reside, completely or at least partially, within the main memory **1304** or within the processing device **1302** during execution thereof by the computer system **1300**; the main memory **1304** and the processing device **1302** also constituting machine-readable storage media. The instructions **1322** may further be transmitted or received over a network **1320** via the network interface device **1308**.

[0116] The machine-readable storage medium **1328** may also be used to store instructions to perform a method for monitoring user health, as described herein. While the machine-readable storage medium **1328** is shown in an exemplary embodiment to be a single medium, the term "machine-readable storage medium" should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, or associated caches and servers) that store the one or more sets of instructions. A machine-readable medium includes any mechanism for storing information in a form (e.g., software, processing application) readable by a machine (e.g., a computer). The machine-readable medium may include, but is not limited to, magnetic storage medium (e.g., floppy diskette); optical storage medium (e.g., CD-ROM); magneto-optical storage medium; read-only memory (ROM); random-access memory (RAM); erasable programmable memory (e.g., EPROM and EEPROM); flash memory; or another type of medium suitable for storing electronic instructions.

[0117] The preceding description sets forth numerous specific details such as examples of specific systems, components, methods, and so forth, in order to provide a good understanding of several embodiments of the present disclosure. It will be apparent to one skilled in the art, however, that at least some embodiments of the present disclosure may be practiced without these specific details. In other instances, well-known components or methods are not described in detail or are presented in simple block diagram format in order to avoid unnecessarily obscuring the present disclosure. Thus, the specific details set forth are merely exemplary. Particular embodiments may vary from these exemplary details and still be contemplated to be within the scope of the present disclosure.

[0118] Additionally, some embodiments may be practiced in distributed computing environments where the machine-readable medium is stored on and or executed by more than one computer system. In addition, the information transferred between computer systems may either be pulled or pushed across the communication medium connecting the computer systems.

[0119] Embodiments of the claimed subject matter include, but are not limited to, various operations described herein. These operations may be performed by hardware components, software, firmware, or a combination thereof.

[0120] Although the operations of the methods herein are shown and described in a particular order, the order of the operations of each method may be altered so that certain operations may be performed in an inverse order or so that certain operation may be performed, at least in part, concurrently with other operations. In another embodiment, instructions or sub-operations of distinct operations may be in an intermittent or alternating manner.

[0121] The above description of illustrated implementations of the invention, including what is described in the Abstract, is not intended to be exhaustive or to limit the invention to the precise forms disclosed. While specific implementations of, and examples for, the invention are described herein for illustrative purposes, various equivalent modifications are possible within the scope of the invention, as those skilled in the relevant art will recognize. The words "example" or "exemplary" are used herein to mean serving as an example, instance, or illustration. Any aspect or design described herein as "example" or "exemplary" is not necessarily to be construed as preferred or advantageous over other aspects or designs. Rather, use of the words "example" or "exemplary" is intended to present concepts in a concrete fashion. As used in this application, the term "or" is intended to mean an inclusive "or" rather than an exclusive "or". That is, unless specified otherwise, or clear from context, "X includes A or B" is intended to mean any of the natural inclusive permutations. That is, if X includes A; X includes B; or X includes both A and B, then "X includes A or B" is satisfied under any of the foregoing instances. In addition, the articles "a" and "an" as used in this application and the appended claims should generally be construed to mean "one or more" unless specified otherwise or clear from context to be directed to a singular form. Moreover, use of the term "an embodiment" or "one embodiment" or "an implementation" or "one implementation" throughout is not intended to mean the same embodiment or implementation unless described as such. Furthermore, the terms "first," "second," "third," "fourth," etc. as used herein are meant as labels to distinguish among different elements and may not necessarily have an ordinal meaning according to their numerical designation.

[0122] It will be appreciated that variants of the above-disclosed and other features and functions, or alternatives thereof, may be combined into may other different systems or applications. Various presently unforeseen or unanticipated alternatives, modifications, variations, or improvements therein may be subsequently made by those skilled in

the art which are also intended to be encompassed by the following claims. The claims may encompass embodiments in hardware, software, or a combination thereof.

[0123] In addition to the embodiments described above, the present disclosure includes, without limitation, the following example implementations.

[0124] Some example implementations provide a method of monitoring a user's cardiac health. The method can include, receiving measured health-indicator data and other-factor data of a user at a first time, inputting, by a processing device, the health-indicator data and other-factor data into a machine learning model, wherein the machine learning model generates predicted health-indicator data at the next time step, receiving the user's data at the next time step, determining, by the processing device, a loss at the next time step, wherein the loss is a measure between the predicted health-indicator data at the next time step and the user's measured health-indicator data at the next time step, determining that the loss exceeds a threshold, and outputting, in response to determining that the loss exceeds a threshold, a notification to the user.

[0125] In some example implementations of the method of any example implementations the trained machine learning model is a trained generative neural network. In some example implementations of the method of any example implementations the trained machine learning model is a feed-forward network. In some example implementations of the method of any example implementations the trained machine learning model is a RNN. In some example implementations of the method of any example implementations the trained machine learning model is a CNN.

[0126] In some example implementations of the method of any example implementations the trained machine learning model is trained on training examples from one or more of: a healthy population, a population with heart disease, and the user.

[0127] In some example implementations of the method of any example implementations the loss at the next time step is the absolute value of the difference between the predicted health-indicator data at the next time step and the user's measured health-indicator at the next time step.

[0128] In some example implementations of the method of any example implementations the predicted health-indicator data is a probability distribution, and wherein the predicted health-indicator data at the next time step is sampled from the probability distribution.

[0129] In some example implementations of the method of any example implementations the predicted health-indicator data at the next time step is sampled according to a sampling technique selected from the group consisting of: the predicted health-indicator data at maximum probability; and random sampling the predicted health-indicator data from the probability distribution.

[0130] In some example implementations of the method of any example implementations the predicted health-indicator data is a probability distribution (β), and wherein the loss is determined based on a negative logarithm of the probability distribution at the next time step evaluated with the user's measured health-indicator at the next time step. In some example implementations of the method of any example implementations the method further includes self-sampling of the probability distribution.

[0131] In some example implementations of the method of any example implementations the method further includes averaging the predicted health-indicator data over a period of time steps, averaging the user's measured health-indicator data over the period of time steps, and determining the loss based on an absolute value difference between the predicted health-indicator data and the measured health-indicator data.

[0132] In some example implementations of the method of any example implementations the measured health-indicator data comprises PPG data. In some example implementations of the method of any example implementations the measured health-indicator data comprises heart rate data.

[0133] In some example implementations of the method of any example implementations the method further includes resampling irregularly spaced heart rate data onto a regularly spaced grid, wherein the heart rate data is sampled from the regularly spaced grid.

[0134] In some example implementations of the method of any example implementations the measured health-indicator data is one or more health-indicator data selected from the group consisting of: PPG data, heart rate data, pulse oximeter data, ECG data, and blood pressure data.

[0135] Some example limitations provide an apparatus comprising a mobile computing device comprising a processing device, a display, a heath-indicator data sensor, and a memory having instructions stored thereon that, when executed by the processing device, cause the processing device to: receive measured health-indicator data from the health-indicator data sensor at time and other-factor data at a first time, input health-indicator data and other-factor data, into a trained machine learning model, and wherein the trained machine learning model generates predicted health-indicator data at a next time step, receive measured health-indicator data and other-factor data at the next time step, determine a loss at the next time step, wherein the loss is a measure between the predicted health-indicator data at the next time step and the measured health-indicator data at the next time step, and output a notification if the loss at the next time step exceeds a threshold.

[0136] In some example implementations of any example apparatus the trained machine learning model comprises a trained generative neural network. In some example implementations of any example apparatus the trained machine learning model comprises a feed-forward network. In some example implementations of any example apparatus the trained machine learning model is a RNN. In some example implementations of the method of any example implementations the trained machine learning model is a CNN.

[0137] In some example implementations of any example apparatus the trained machine learning model is trained on training examples from one of the group consisting of: a healthy population, a population with heart disease and the user.

[0138] In some example implementations of any example apparatus the predicted health-indicator data is a point prediction of the user's health-indicator the next time step, and wherein the loss is the absolute value of the difference between the predicted health-indicator data and the measured health-indicator data at the next time step.

[0139] In some example implementations of any example apparatus the predicted health-indicator data is sampled from a probability distribution generated from the machine learning model.

[0140] In some example implementations of any example apparatus the predicted health-indicator data is sampled according to a sampling technique selected from the group

consisting of: a maximum probability; and random sampling from the probability distribution.

[0141] In some example implementations of any example apparatus the predicted health-indicator data is a probability distribution (β), and wherein the loss is determined based on a negative logarithm of β evaluated with the user's measured health-indicator at the next time step.

[0142] In some example implementations of any example apparatus the processing device is further to define a function α ranging from 0 to 1, wherein I_t comprises a linear combination the user's measured health-indicator data and the predicted health-indicator data as a function of α.

[0143] In some example implementations of any example apparatus the processing device is further to perform self-sampling of the probability distribution.

[0144] In some example implementations of any example apparatus the processing device is further to: average, using an averaging method, the predicted health-indicator data sampled from the probability distribution over a period of time steps, average, using the averaging method, the user's measured health-indicator data over the period of time steps, defining the loss the absolute value of the averaged predicted health-indicator data and the measured health-indicator data.

[0145] In some example implementations of any example apparatus the averaging method comprises one or more methods selected from the group consisting of: calculating an average, calculating an arithmetic mean, calculating a median and calculating a mode.

[0146] In some example implementations of any example apparatus the measured health-indicator data comprises PPG data from a PPG signal. In some example implementations of any example apparatus the measured health-indicator data is heart rate data. In some example implementations of any example apparatus the heart rate data is collected by resampling irregularly spaced heart rate data onto a regularly spaced grid, and the heart rate data is sampled from the regularly spaced grid. In some example implementations of any example apparatus the measured health-indicator data is one or more health-indicator data selected from the group consisting of: PPG data, heart rate data, pulse oximeter data, ECG data, and blood pressure data.

[0147] In some example implementations of any example apparatus the mobile device is selected from the group consisting of: a smart watch; a fitness band; a computer tablet; and a laptop computer.

[0148] In some example implementations of any example apparatus the mobile device further comprises a user high-fidelity sensor, wherein the notification requests the user to obtain high-fidelity measurement data, and wherein the processing device is further to: receive an analysis of the high-fidelity measurement data; label the user measured health-indicator data with the analysis to generate labeled user health-indicator data; and use labeled user health-indicator data as a training example to train a trained personalized high-fidelity machine learning model.

[0149] In some example implementations of any example apparatus the trained machine learning model is stored on the memory. In some example implementations of any example apparatus the trained machine learning model is stored on a remote memory, wherein the remote memory is separate from the computing device and wherein the mobile computing device is a wearable computing device. In some example implementations of any example apparatus the trained personalized high-fidelity machine learning model is stored on the memory. In some example implementations of any example apparatus the trained personalized high-fidelity machine learning model is stored on a remote memory, wherein the remote memory is separate from the computing device and wherein the mobile computing device is a wearable computing device.

[0150] In some example implementations of any example apparatus the processing device is further to predict that the user is experiencing atrial fibrillation and determine an atrial fibrillation burden of the user.

[0151] Some example implementations provide a method of monitoring a user's cardiac health. The method can include receiving measured low-fidelity user health-indicator data and other-factor data at a first time, inputting data comprising the user health-indicator data and other-factor data at the first time, into a personalized high-fidelity trained machine learning model, wherein the personalized high-fidelity trained machine learning model makes a prediction if the user's health-indicator data is abnormal, and if the prediction is abnormal, sending a notification that the user's health is abnormal.

[0152] In some example implementations of the method of any example implementations the trained personalized high-fidelity machine learning model is trained on measured low-fidelity user health-indicator data labeled with an analysis of high-fidelity measurement data.

[0153] In some example implementations of the method of any example implementations the analysis of high-fidelity measurement data is based on user specific high-fidelity measurement data.

[0154] In some example implementations of the method of any example implementations the personalized high-fidelity machine learning model outputs a probability distribution, wherein the prediction is sampled from the probability distribution.

[0155] In some example implementations of the method of any example implementations the prediction is sampled according to a sampling technique selected from the group consisting of the prediction at a maximum probability and random sampling the prediction from the probability distribution.

[0156] In some example implementations of the method of any example implementations an averaged prediction is determined by averaging, using an averaging method, the prediction over a period of time steps, and wherein the averaged prediction is used to determine if the user's health-indicator data is normal or abnormal.

[0157] In some example implementations of the method of any example implementations the averaging method comprises one or more methods selected from the group consisting of: calculating an average, calculating an arithmetic mean, calculating a median and calculating a mode.

[0158] In some example implementations of the method of any example implementations the personalized high-fidelity trained machine learning model is stored in a memory of a user wearable device. In some example implementations of the method of any example implementations the measured health-indicator data and other-factor data are time segments of data over a time period.

[0159] In some example implementations of the method of any example implementations the personalized high-fidelity trained machine learning model is stored in a remote memory, wherein the remote memory is located remotely from a user wearable computing device.

[0160] In some example implementation a health monitoring apparatus may include a mobile computing device comprising a microprocessor, a display, a user heath-indicator data sensor, and a memory having instructions stored thereon that, when executed by the microprocessor, cause the processing device to: receive measured low-fidelity health-indicator data and other-factor data at a first time, wherein measured health-indicator data is obtained by the user health-indicator data sensor, input data comprising the health-indicator data and other-factor data at the first time, into a trained high-fidelity machine learning model, wherein the trained high-fidelity machine learning model makes a prediction if the user's health-indicator data is normal or abnormal; and in response to the prediction being abnormal, send a notification to at least the user that the user's health is abnormal.

[0161] In some example implementations of health monitoring apparatus of any example implementation the trained high-fidelity machine learning model is a trained high-fidelity generative neural network. In some example implementations of health monitoring apparatus of any example implementation wherein the trained high-fidelity machine learning model is a trained recurrent neural network (RNN). In some example implementations of health monitoring apparatus of any example implementation the trained high-fidelity machine learning model is a trained feed-forward neural network. In some example implementations of health monitoring apparatus of any example implementation the trained high-fidelity machine learning model is a CNN.

[0162] In some example implementations of health monitoring apparatus of any example implementation the trained high-fidelity machine learning model is trained on measured user health-indicator data labeled with based on user specific high-fidelity measurement data.

[0163] In some example implementations of health monitoring apparatus of any example implementation the trained high-fidelity machine learning model is trained on low-fidelity health-indicator data labeled based on high-fidelity measurement data, wherein the low-fidelity health-indicator data and the high-fidelity measurement data is from a population of subjects.

[0164] In some example implementations of health monitoring apparatus of any example implementation the high-fidelity machine learning model outputs a probability distribution, wherein the prediction is sampled from the probability distribution.

[0165] In some example implementations of health monitoring apparatus of any example implementation the prediction is sampled according to a sampling technique selected from the group consisting of: the prediction at a maximum probability; and random sampling the prediction from the probability distribution.

[0166] In some example implementations of health monitoring apparatus of any example implementation an averaged prediction is determined by averaging, using an averaging method, the prediction over a period of time steps, and wherein the averaged prediction is used to determine if the user's health-indicator data is normal or abnormal.

[0167] In some example implementations of health monitoring apparatus of any example implementation the measured health-indicator data and other-factor data are time segments of data over a time period.

[0168] In some example implementations of health monitoring apparatus of any example implementation the averaging method comprises one or more methods selected from the group consisting of: calculating an average, calculating an arithmetic mean, calculating a median and calculating a mode.

[0169] In some example implementations of health monitoring apparatus of any example implementation the personalized high-fidelity trained machine learning model is stored in the memory. In some example implementations of health monitoring apparatus of any example implementation the personalized high-fidelity trained machine learning model is stored in a remote memory, wherein the remote memory is located remotely from the wearable computing device. In some example implementations of health monitoring apparatus of any example implementation the mobile device is selected from the group consisting of: a smart watch; a fitness band; a computer tablet; and a laptop computer.

What is claimed is:

1. An apparatus, comprising:

a processing device;

a low-fidelity heath-indicator data sensor operatively coupled to the processing device;

a high-fidelity health-indicator data sensor operatively coupled to the processing device; and

a memory having instructions stored thereon that, when executed by the processing device, cause the processing device to:

receive measured low-fidelity health-indicator data at a first time, wherein the measured low-fidelity health-indicator data is obtained by the low-fidelity health-indicator data sensor;

input a set of data comprising the measured low-fidelity health-indicator data into a trained high-fidelity machine learning model, wherein the trained high-fidelity machine learning model is to generate a prediction whether a health of a user is outside a normal range, based on a low-fidelity health-indicator threshold;

in response to a determination that the prediction is outside the normal range: receive measured high-fidelity health-indicator data obtained by the high-fidelity health-indicator data sensor; and

in response to a determination that the measured high-fidelity health-indicator data is inside the normal range: modify the low-fidelity health-indicator threshold to decrease a notification sensitivity.

2. The apparatus of claim **1**, wherein the low-fidelity health-indicator threshold corresponds to a first sensitivity threshold, and wherein to modify the low-fidelity health-indicator threshold the processing device is to modify the first sensitivity threshold to a second sensitivity threshold.

3. The apparatus of claim **2**, the processing device further to: modify the second sensitivity threshold to the first sensitivity threshold in response to an expiration of a time interval.

4. The apparatus of claim **2**, wherein to determine that the prediction is outside the normal range, the processing device is to: determine that a probability that the prediction is outside the normal range is above the first sensitivity threshold for a first time threshold.

5. The apparatus of claim **4**, wherein the low-fidelity health-indicator threshold corresponds to the first time threshold, and wherein to modify the low-fidelity health-

indicator threshold the processing device is to modify the first time threshold to a second time threshold.

6. The apparatus of claim **1**, wherein the high-fidelity health-indicator data sensor comprises an electrocardiogram (ECG) sensor.

7. The apparatus of claim **1**, wherein the low-fidelity health-indicator data sensor comprises a photoplethysmography (PPG) sensor.

8. The apparatus of claim **1**, wherein the apparatus is one of: a smartwatch, a fitness band, a computer tablet, or a laptop computer.

9. The apparatus of claim **1**, wherein the trained high-fidelity machine learning model comprises one or more of: a generative neural network, a recurrent neural network (RNN), or a feed-forward neural network.

10. The apparatus of claim **1**, wherein the set of data further comprises a record of activity level of the user.

11. A method, comprising:

receiving measured low-fidelity health-indicator data at a first time, wherein the measured low-fidelity health-indicator data is obtained by a low-fidelity health-indicator data sensor;

inputting a set of data comprising the measured low-fidelity health-indicator data into a trained high-fidelity machine learning model, wherein the trained high-fidelity machine learning model is to generate a prediction whether a health of a user is outside a normal range, based on a low-fidelity health-indicator threshold;

in response to determining that the prediction is outside the normal range: receiving measured high-fidelity health-indicator data obtained by a high-fidelity health-indicator data sensor; and

in response to determining that the measured high-fidelity health-indicator data is inside the normal range: modifying, by a processing device, the low-fidelity health-indicator threshold to decrease a notification sensitivity.

12. The method of claim **11**, wherein the low-fidelity health-indicator threshold corresponds to a first sensitivity threshold, and wherein to modify the low-fidelity health-indicator threshold the method further comprises: modifying the first sensitivity threshold to a second sensitivity threshold.

13. The method of claim **12**, further comprising: modifying the second sensitivity threshold to the first sensitivity threshold in response to an expiration of a time interval.

14. The method of claim **12**, wherein to determine that the prediction is outside the normal range, the method further comprises: determining that a probability that the prediction is outside the normal range is above the first sensitivity threshold for a first time threshold.

15. The method of claim **14**, wherein the low-fidelity health-indicator threshold corresponds to the first time threshold, and wherein to modify the low-fidelity health-indicator threshold the method further comprises: modifying the first time threshold to a second time threshold.

16. The method of claim **11**, wherein the high-fidelity health-indicator data sensor comprises an electrocardiogram (ECG) sensor.

17. The method of claim **11**, wherein the low-fidelity health-indicator data sensor comprises a photoplethysmography (PPG) sensor.

18. The method of claim **11**, wherein the processing device corresponds to one of: a smartwatch, a fitness band, a computer tablet, or a laptop computer.

19. The method of claim **11**, wherein the trained high-fidelity machine learning model comprises one or more of: a generative neural network, a recurrent neural network (RNN), or a feed-forward neural network.

20. The method of claim **11**, wherein the set of data further comprises a record of activity level of the user.

* * * * *

专利名称(译)	使用移动设备持续监控用户的健康状况		
公开(公告)号	US20190076031A1	公开(公告)日	2019-03-14
申请号	US16/186244	申请日	2018-11-09
[标]申请(专利权)人(译)	阿利弗克公司		
申请(专利权)人(译)	ALIVECOR INC.		
当前申请(专利权)人(译)	ALIVECOR INC.		
[标]发明人	VALYS ALEXANDER VAINIUS PETTERSON FRANK LOSASSO GALLOWAY CONNER DANIEL CROSS ALBERT DAVID E GOPALAKRISHNAN RAVI KORZINOV LEV WANG FEI THOMSON EUAN SRIVASTAVA NUPUR DAWOOD OMAR ABUZEID IMAN		
发明人	VALYS, ALEXANDER VAINIUS PETTERSON, FRANK LOSASSO GALLOWAY, CONNER DANIEL CROSS ALBERT, DAVID E. GOPALAKRISHNAN, RAVI KORZINOV, LEV WANG, FEI THOMSON, EUAN SRIVASTAVA, NUPUR DAWOOD, OMAR ABUZEID, IMAN		
IPC分类号	A61B5/0205 A61B5/00 G16H40/63 A61B5/046 A61B5/024 A61B5/0245 G06F19/00 A61B5/021 A61B5/0452 A61B5/11 G16H10/60 G16H15/00 G16H50/30		
CPC分类号	A61B5/681 G16H15/00 A61B5/046 A61B5/02405 A61B5/02416 A61B5/0245 G06F19/00 G16H50/30 A61B5/746 A61B5/6898 A61B5/7275 A61B5/02055 A61B5/0022 A61B5/1118 A61B5/0452 A61B5/02438 A61B5/021 G16H10/60 A61B5/7264 G16H40/63		
优先权	61/915113 2013-12-12 US 61/953616 2014-03-14 US 61/969019 2014-03-21 US 61/970551 2014-03-26 US 62/014516 2014-06-19 US 62/589477 2017-11-21 US		
外部链接	Espacenet USPTO		

摘要(译)

本文公开了用于连续监测用户的健康状态的设备，系统，方法和平台，例如心脏健康状态。本公开描述了用于连续监视来自用户设备的用户的健康指示数据（例如但不限于PPG信号，心率或血压）的系统，方法，设备，软件和平台，以及相应的（及时）与可能影响健康指标的因素（“其他因素”）相关的数据，以确定用户是否具有正常健康状况，例如但不限于（i）一组受类似其他因素影响的个人，或（ii）受类似其他因素影响的用户。