IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

A New Real Time Clinical Decision Support System Using Machine Learning for Critical Care Units

Noha O. El-Ganainy[1], Member, IEEE, Ilangko Balasingham[1,2], Senior Member, IEEE, Per Steinar Halvorsen[2,3] and Leiv Arne Rosseland[3,4].

[1]Norwegian University for Science and Technology (NTNU), Trondheim, Norway.
[2]Intervention Center, Division of Emergencies and Special Care, Oslo University Hospital, Oslo, Norway.
[3]Institute of Clinical Medicine, University of Oslo, Oslo, Norway.
[4]Division of Emergencies and Special Care, Oslo University Hospital, Oslo, Norway.

Corresponding author: Noha O. El-Ganainy (e-mail: noha_elganainy@yahoo.co.uk).

This work was funded in part by the European Research Consortium for informatics and mathematics (ERCIM) 'Alain Bensoussan' fellowship program.

ABSTRACT Mean arterial pressure (MAP) is an important clinical parameter to evaluate the health of critically ill patients in intensive care units. Thus, the real time clinical decision support systems detecting anomalies and deviations in MAP enable early interventions and prevent serious complications. The state-of-the-art decision support systems are based on a three-phase method that applies offline training, transfer learning, and retraining at the bedside. Their applicability in critical care units is challenging with delay and inaccuracy. In this paper, we propose a real time clinical decision support system forecasting the MAP status at the bedside using a new machine learning structure. The proposed system works in real time at the bedside without requiring the offline phase for training using large datasets. It thereby enables timely interventions and improved healthcare services. The proposed machine learning structure includes two stages. Stage I applies online learning using hierarchical temporal memory (HTM) to enable real time stream processing and provides unsupervised predictions. To the best of our knowledge, this is the first time it is applied to medical signals. Stage II is a long short-term memory (LSTM) classifier that forecasts the status of the patient's MAP ahead of time based on Stage I stream predictions. We perform a thorough performance evaluation of the proposed system and compare it with the state-of-the-art systems employing logistic regression (LR). The comparison shows the proposed system outperforms LR in terms of the classification accuracy, recall, precision, and area under the receiver operation curve (AUROC).

INDEX TERMS Clinical decision support, classification, hierarchical temporal memory (HTM), long short-term memory (LSTM), machine learning, real time prediction.

I. INTRODUCTION

In critical care units, it is essential to predict adverse events as it allows for preventive interventions and reduces clinical complications. Episodes of anomalous mean arterial pressure (MAP) values are frequent at the bedside as they are often associated with anesthesia. These episodes are linked to cardiovascular risks, multiple organ failure, and life-threatening complications, especially in critically ill patients [1-6]. As a result, evolving research activity has begun to develop real-time decision support tools to predict MAP to enable preventive preparations and reduce complications[1-6]. The use of machine learning in clinical decision support systems has become a current research hotspot[2-11]. These systems are based on offline analysis to save the challenge of processing streams in real time [3-6, 9-11]. The latter requires machine learning algorithms that are capable of adapting to high-speed streams and their fast-changing characteristics [12]. Recently, few clinical decision support systems have been proposed, working in real time using machine learning, based on stream processing [2,7,8]. These systems are mostly three phase-based, consisting of offline training on a stored data set, transfer learning, followed by bedside retraining on the monitored patient's vital signs. The problem with these systems is the generality of data and models, as the dataset does not necessarily represent a large population. Modeling must also be repeated in different clinical cases. The practicality of such systems is also questionable due to over-training, which in some cases also

TABLE I
OVERVIEW ON RELATED WORK ON CLINICAL DECISION SUPPORT SYSTEMS.

Ref	Parameter	Real Time	Three- Phase Based	Feature Selection /number of features	Data set size	Machine Learning Algorithm	Performance
[2]	Hypotension	yes	Yes	Yes/ 3022	1334	LR	AUC= 95%
[3]	Hypotension/ Kidney injury	No	offline	231	57317	LR	-
[4]	Diastolic/ Systolic	No	offline	Yes/18	285	RF	Acc= 88%
[5]	Hyper-tension	No	offline		25	RF	Acc= 60%
[6]	Hyper-tension	No	offline	No/19	100	DT, SVM, LR	Acc= 90, 86, 83 %
[7]	Acute Kidney Injury	Yes	Yes	Automatic	4490	RF	AUC=80%
[8]	Cardiotocography	Yes	Yes	No/21	2126	SVM, RF	Acc= 96, 95%
[9]	Hypovolemic shock	No	Offline	Yes/42	66	SVM	Sensitivity= 90%
[10]	Mortality	No	Offline	Yes/16	341	ANN, SVM, RF, LR	AUC= 0.84, 0.84, 0.9, 0.85.
[11]	Ectopic Pregnancy	No	Offline	Yes/33	406	MLP, DL. SVM, NB	Acc= 87, 57, 89, 68 %

leads to delayed decisions. Also, these non-personalized models have led to uncertain results [2,7,8].

This paper proposes a personalized real-time clinical decision support system that works beyond the three phase-based frameworks guidelines principle. We suggest a new machine learning prediction-and-classification system that is applied directly to the bedside in real time. The suggested framework incorporates the use for prediction and classification of online machine learning, hierarchical temporal memory (HTM), and long-term memory (LSTM), respectively, over two stages. HTM has been commonly used in financial data in real time [13-18], but it is first implemented in medical signals in our proposed system. The proposed clinical decision support system is implemented to forecast the status of MAP based on an analysis of the vital signs of the patient in real time. In our system, the HTM properties play a key role as they enable the decision support system to operate directly at the bedside, avoid retraining time delays, and provide earlier decisions. HTM forecasts the stream one-step into the future, unsupervised, without pre-modeling. These streams are fed to an LSTM classifier to forecast the MAP value one step forward, after careful selection of features. As our system is personalized, the proposed structure is applied to each patient. LSTM, the most efficient recurrent neural network (RNN) algorithm, provides a classifier with the advantage of managing the set of features, and the observation window [14, 19- 21]. It distinguishes LSTM from other neural artificial networks (ANNs) and deep learning classifiers, which are sometimes defined as inexplicable[22]. The proposed decision support system is applied to the data set of the observational study collected at the University Hospital of Oslo[30].

The proposed clinical decision support system is evaluated and compared to the state-of-the-art LR systems. Both the predictive and classification stages are carefully verified in terms of accuracy and efficiency. The proposed system outperforms LR in terms of accuracy, recall, precision, F1-score, and area under the receiver operating characteristic (AUROC). The forecast time is also considered to be a comparison metric, where it indicates the number of time steps in advance, the system can predict the event before it happens. The proposed system shows the ability to provide faster and highly accurate forecasts compared to LR. The paper is structured as follows; Section II explains the clinical motivation, related work, and decision support systems challenges. Section III is on the processes and contributions. Section IV outlines the proposed decision support system. Section V provides the results and discussion, and Section VI concludes.

II. CLINICAL MOTIVATIONS AND CHALLENGES

MAP is a crucial clinical criterion for evaluation and examination as a predictor of adverse health conditions. A MAP anomaly can be either hypotensive or hypertensive and is defined as an interval where the values exceed thresholds specified for at least a minute [1-3]. Patients with anesthesia, surgery, and intensive care [1-6] are likely to experience anomalous levels of MAP. Clinically, it has been concluded that anomalous MAP is associated with various complications such as acute kidney injury, increased postoperative myocardial infarction levels, cardiovascular risks, as well as organ failure and death [1–6]. This explains the need for early detection of MAP in critical care units and operating rooms to allow timely intervention to reduce the potential for mortality and morbidity.

A. RELATED WORK

Various analytical tools are currently used in clinical practice to support decision-making on various clinical parameters, including MAP[2-11]. These clinical decision-making tools mostly employ machine learning techniques using algorithms such as LR, random forest (RF), supported vector machine (SVM), and deep learning. Several studies

have been performed on the early identification of events with blood pressure and supporting decisions. Edward Labs has developed an LR-based model of real-time hypotension prediction based on waveform analysis in [2]. LR was also used to calculate the dependency between hypotensive events and acute kidney injury in[3] and diastolic and systolic blood pressure in[4]. Forecasting of hypertension in pregnant women using RF has been studied in [5]. The same problem was studied and compared with classical machine learning algorithms in [6]. Acute kidney injury was predicted using RF in [7]. Cardiotocography data was analyzed using SVM and RF in [8]. SVM has been proposed to monitor cardio-cerebrovascular hemorrhage in[9]. SVM, RF, LR, and ANN have been employed and their efficacy in predicting mortality rates has been compared in[10]. While multilayer perception (MLP), SVM, deep learning, and Naives Bayes were compared to predict ectopic pregnancy in [11]. Table 1 lists the comparison between the different studies. Clinical decision support systems are mostly used offline on stored data sets [3-6, 9-11]. In 2017, Edward Labs developed a model capable of working in real time[2]. This work was considered to be a breakthrough compared to studies commonly conducted on offline analysis. This model was later developed into a commercial tool. Various frameworks followed and proposed real-time predictions, as in [7, 8].

B. CHALLENGES

Real-time stream analysis at the bedside[13-14] is the main challenge for real-time clinical decision support frameworks. Streams are usually of high speed; the data rate and volume are high. The characteristics of the stream vary constantly, which is known as the concept drift problem[12]. Such problems place many restrictions on learning algorithms because they need to be fully unsupervised without labeling or annotation. Stream prediction has been conventionally performed using linear prediction methods that are successful when signals experience high sample correlation and periodicity [24, 25]. Machine learning, however, was seen as an alternative to stream learning and predictions using statistical models, models of neural networks, and cortical algorithms[15-18]. Only the later algorithms can process streams in real time [15-18]. On the other hand, the statistical and neural network models are mostly supervised algorithms, modified to adapt to the challenges of stream predictions and real-time learning. This is achieved with time delays, buffering, and storage [12- 14, 19- 21]. But these improvements are leading to delayed forecasts and reduced performance.

Recent studies on real-time clinical decision support systems have suggested a transfer learning phase that complements the machine learning process. Applying transfer learning has become a turning point as it has strengthened the process of real time analysis of the systems and improved their performance[24]. The systems in this context avoid real-time stream processing and apply a three-phase-based framework consisting of offline training, transfer learning, and bedside retraining[2, 7, 8]. It begins with the offline training phase to train the classifier on a large stored data set. The offline procedure also includes training, testing, and validation. Transfer learning is then applied and plays a key role in these frameworks as it brings the learning achieved during the offline phase to the bedside[24]. It uses the pre-trained network offline as a basis and upgrades it by retraining the patient's data and the bedside streams in real time. After the three steps, the classifier is then able to provide decision support at the bedside. The three-phase systems are effective but lack generality and practicality [3-6, 9-11]. Deep learning models suffer from another problem; they are commonly described as "black boxes" and are not usually clinically supported. The learning and selection of features in these models are non-trackable, where the resulting alarms are considered unexplained. This problem has prompted the rise of explainable AI research to reveal more details on their automatic learning and selection of features[22].

We may conclude that the development of decision support systems rests on operating directly on real-time stream analysis outside the three-phase frames. It is expected that the next generation of decision support systems will follow an operational framework able to address these processing challenges. In this paper, we propose a decision support system modeling and processing streams in real time based on the use of cortical algorithms [15-18]. Cortical learning algorithms are a category of artificial neural networks that mimic the brain, where data is processed in real time. That means they can learn in an unsupervised manner incrementally the temporal patterns in the data. The following section explains the cortical algorithms.

III. CONTRIBUTIONS AND METHODS

A. CONTRIBUTIONS

The main contributions to this paper are as follows:

i. We are proposing a new, personalized, real-time clinical decision-making support system working at the bedside. Compared to existing systems this can be viewed as a breakthrough. As It operates beyond the three phase-based systems, eliminating both off-line training and transfer learning.

ii. The advantage of our system is based on the proposed prediction-and-classification machine learning structure using HTM-and-LSTM. HTM provides the system with the benefit of working in real time as it can model and predict streams in real time and unsupervised. For the first time, HTM has been investigated for the use of medical signals by us in[30].

iii. The proposed machine learning structure uses the LSTM classifier as a second stage to forecast a binary decision on each patient's MAP status. The system is personalized and we are developing a model for each patient. The classifier uses the vital signs of each patient as features, as we seek to investigate the inter- and intra-dependencies between the different vital signs to forecast the MAP status.

iv. Performance evaluation is provided to demonstrate the effectiveness of the proposed decision support system compared to the state-of-the-art use of LR. We apply machine learning algorithms to a data set collected at the Oslo University Hospital (OUS)[23].

B. METHODS

This subsection is dedicated to the description of HTM; it is a theoretical model that mimics different structural and algorithmic principles in the brain's neocortex [8-14]. The HTM network is trained using unsupervised Hebbian-style learning and consists of layers of cells arranged in a set of columns. The cell has three states; active, predictive / depolarized, and non-active. Each cell has two separate input regions, the proximal and the distal. The proximal region represents the feedforward input to HTM. At the same time, the distal region reflects the lateral connections within the layer that express the temporal pattern learned, and that lead to predictions. A cell is activated when it receives sufficient proximal and distal input.

The HTM network at any time describes the input sequence in terms of two separate sparse representations, one at the level of the columns and one at the level of the single cells [15-17]. The sparse representation at the column level describes the current feed-forward input. The representation of the cell level denotes the learned temporal pattern from the current input and leads to predictions of the future time step. In the case of an HTM network with N columns and M cells per column, the learning rules are described using four matrices explaining both the activation rules and the network dependencies [15, 16]. The first matrix is the Activation matrix A^t, where a_{ij}^t is the activation state of the i^{th} cell in the j^{th} column at the time interval t

$$a_{ij}^t = \begin{cases} 1 & if\ j \epsilon W^t\ and\ \pi_{ij}^{t-1} = 1 \\ 1 & if\ j \epsilon W^t\ \ and\ \ \sum_i \pi_{ij}^{t-1} = 0 \\ 0 & otherwise, \end{cases} \quad (1)$$

where W^t is the matrix representing the set of active columns, which illustrates the input at time t, and π_{ij}^{t-1} denotes the predictive state of each cell determined during the previous time interval *t-1*. The activation rules work as follows, a cell is activated if and only if its column receives sufficient feedforward input, regardless of its previous state (predictive, active, or non-active) as shown in (1). However, the cell in a predictive state is privileged, more specifically it will be activated faster if its column receives enough input and will prevent other cells in the same column from being activated unless they are also in the predictive state. That is known by the intra-column inhibition. If the column receives enough input and does not have a predictive cell, it will be completely activated. The second matrix is the predictive state input is Π^t, where each element π_{ij}^t is the predictive state of the i^{th} cell in the j^{th} column for the following time interval,

$$\Pi_{ij}^t = \begin{cases} 1 & if \exists_d \left\| \widetilde{D}_{ij}^d \circ A^t \right\|_1 > \theta \\ 0 & otherwise, \end{cases} \quad (2)$$

where θ is the segment activation threshold, $\circ$ is element-wise multiplication, and $\widetilde{D}_{ij}^d$ denotes the connected synapses/connection of segment d of the i^{th} cell in the j^{h} column. A cell is said to be in a predictive state if it receives enough distal input as shown in (2), which represents the temporal pattern learned by the HTM network. Using Hebbian-style learning rules, the precision of the learning at time step t is checked at time step $t+1$. It evaluates the learning process at each time step in order to strengthen or decay the lateral connection leading to each prediction. The active lateral connection leading to successful prediction is reinforced by increasing the permanence of by the amount of p^+, and decreases the permanence of inactive connections by a small value of p^-.

$$\Delta D_{ij}^d = p^+ \dot{D}_{ij}^d \circ A^{t-1} - p^- \dot{D}_{ij}^d \circ (1 - A^{t-1}), \quad (3)$$

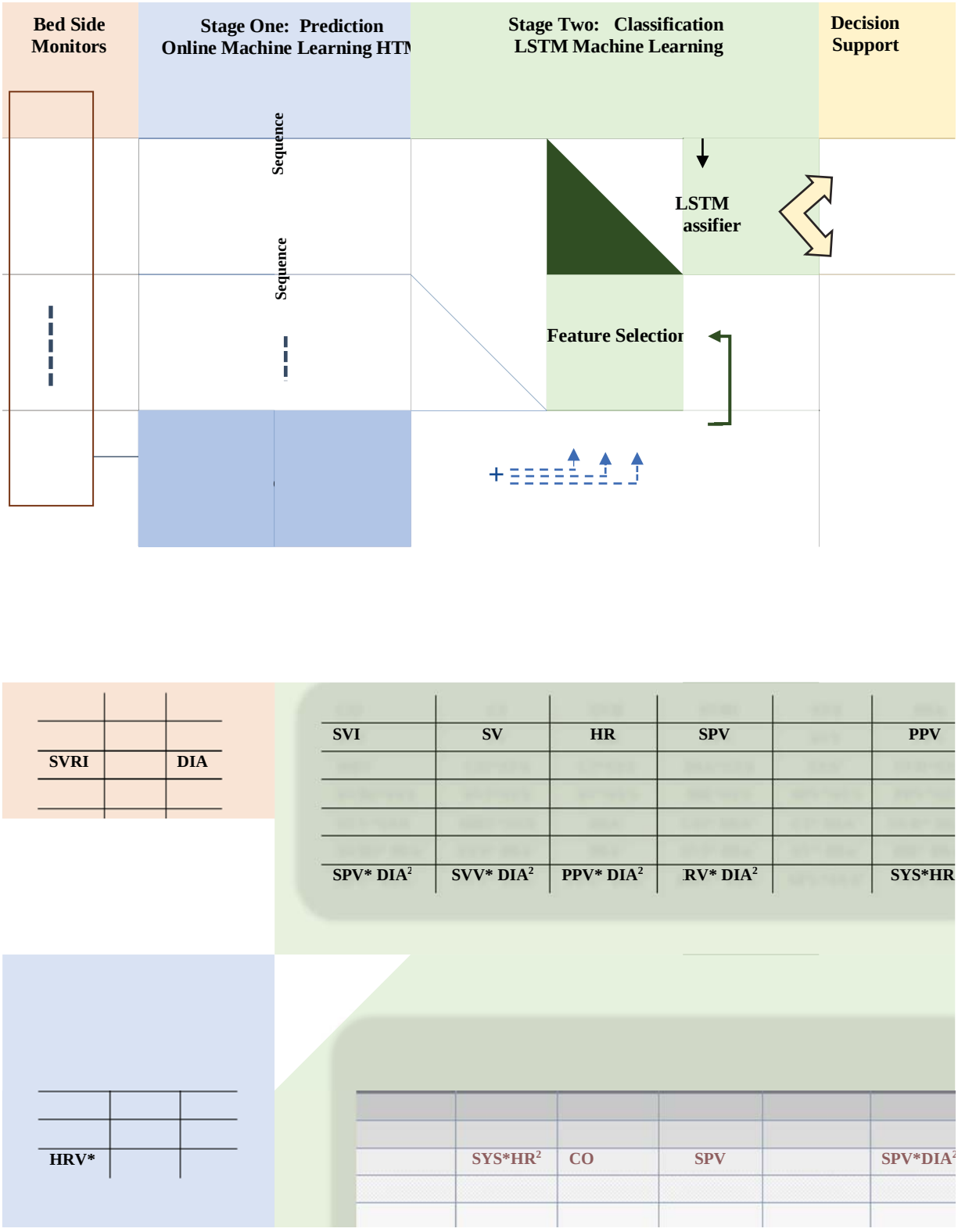
where $\widetilde{D}_{ij}^d$ denotes the connected synapses and $\dot{D}_{ij}^d$ is a binary matrix containing only the positive entries in D_{ij}^d. Next, a small decay is applied to the active segments of cells with wrong predictions; this is namely called the long-term depression rule.

$$\Delta D_{ij}^d = p^{--} \dot{D}_{ij}^d \quad (4)$$

where $a_{ij}^t = 0$ and $\left\| \widetilde{D}_{ij}^d \circ A^{t-1} \right\|_1 > \theta$, $p^{--} \ll p^-$. The fourth matrix to control the learning rules is the matrix that describes the input activity. The matrix W^t denotes the set of active columns illustrating the feedforward input. Many columns receive synaptic inputs at every time step but not all of them are activated. the number of active proximal cells for each column is calculated by applying the intercolumnar inhibitory process. Only the higher 2% of columns are selected and activated.

IV. THE PROPOSED DECISION SUPPORT SYSTEM

We propose a personalized decision support system capable to operate in real time, the system is as shown in Fig.1. At the input stage, marked in orange, the vital signs data are collected as streams directly at the bedside from the patient's monitor. In our model, we consider vital signs that are mostly monitored in clinical practice. Fig.1-A shows the complete list of these vital signs. The streams then pass to the Stage I machine learning, shown in blue, which uses HTM to provide unsupervised real-time prediction for each vital sign. The predicted streams listed in Fig.1-B are then passed to Stage II, in green. The streams undergo the feature extraction and selection then pass to the LSTM classifier. it uses the MAP ground truth and the selected features as input. At the final stage, marked in yellow, the classification decision is ready and a forecast decision on the future level of the MAP is provided.



IGURE 1. **The proposed personalized decision support system working on the bed side on real time. The input to the system is a number of monitored signs streams in real time as marked in light orange. These vital sign streams are fed to the first proposed stage for prediction, the streams are listed in the orange table . This stage is marked in blue and uses online machine learning to provide unsupervised predictions in real time. The resulting predicted streams are listed in table . They then pass to the second proposed stage, in green, employs LSTM machine learning for classification. Feature extraction and selection are applied prior to the classification. The resulting feature space after extraction and selection steps are listed in table and respectively. The later table was color coded to illustrate the importance ranking of the features. Dark shaded cells signify high impact while light shades signify low impact. Features written in black font are the ones with the least impact and are discarded. The resulting forecast decision, in yellow, is binary stating either MAP in the coming time interval is normal or experiences abnormality.**

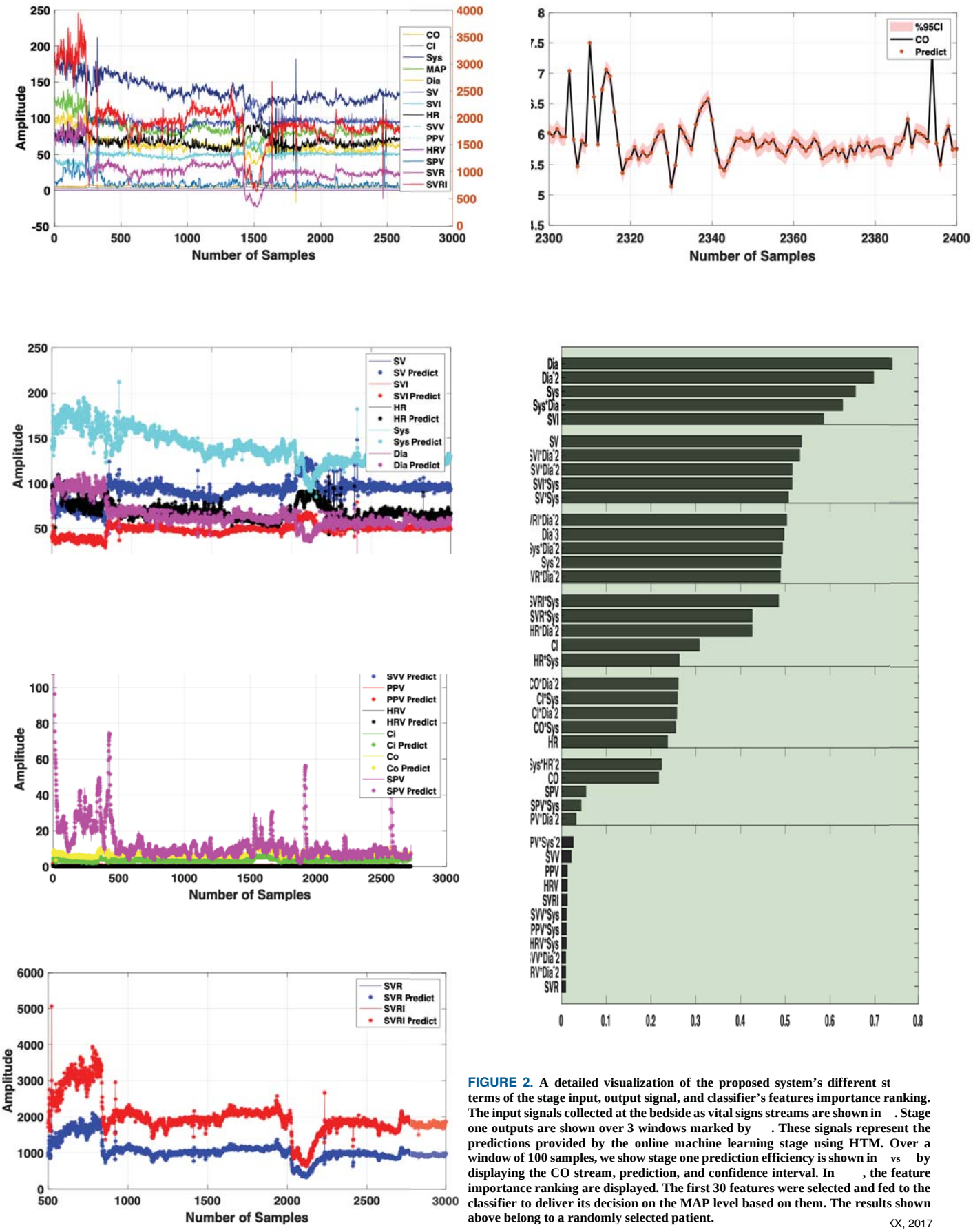


FIGURE 2. A detailed visualization of the proposed system's different st terms of the stage input, output signal, and classifier's features importance ranking. The input signals collected at the bedside as vital signs streams are shown in . Stage one outputs are shown over 3 windows marked by . These signals represent the predictions provided by the online machine learning stage using HTM. Over a window of 100 samples, we show stage one prediction efficiency is shown in vs by displaying the CO stream, prediction, and confidence interval. In , the feature importance ranking are displayed. The first 30 features were selected and fed to the classifier to deliver its decision on the MAP level based on them. The results shown above belong to a randomly selected patient.

We setup LSTM as a semi-supervised classifier and it provides a binary decision. It needs a training period until the decision on the MAP level is foreseen. In our model, we use the data collected before the spinal anesthesia for training, and the classifier is configured to deliver the decision after the spinal anesthesia. The decision-support system is personalized. This means that for each patient a complete system is trained using only its data, features, and parameters. The proposed system was applied to the data set in[23]. A detailed description of the proposed decision support system stages is given in the following subsections; the data set is defined in subsection A, stage I is clarified in subsection B, feature extraction and selection is addressed in subsection C, and an explanation of the classification stage is shown in subsection D.

A. DATA SET

The data set was collected at Oslo University Hospital following an observational study on healthy pregnant women scheduled for cesarean delivery[23]. Helsinki Declaration of ethical standards and guidelines for the Institutional Review Board (IRB) are respected. The dataset includes a continuous recording of different vital signs of 76 people, the recordings started almost 7 minutes before spinal anesthesia and until delivery. The continuously recorded vital signs are; diastolic arterial pressure (Dia), mean arterial pressure (MAP), systolic blood pressure (Sys), heart rate (HR), estimated cardiac output (CO), heart rate index (CI), stroke volume (SV), systemic vascular resistance (SVR), systemic vascular resistance index (SVRI), heart rate variability (HRV), stroke volume v. The data collected is stored in a time series, ranging from 3000 to 6000 samples per patient. -- sample reflects the median of the samples obtained over 10 seconds, with a 5-second sliding window. Patients are monitored and a forecast decision on the level of MAP is required after spinal anesthesia. The data set includes various events, such as hypotensive and hypertensive, which provide valuable benchmark data for the assessment of the proposed decision support system. For clarification, we displayed these vital signs streams for a randomly selected patient in Fig.2-A. It should be noted that the size of the data set can not be comparable to that of the medical research databases. This could, therefore, be seen as a limitation. However, the dataset has a reasonable size w.r.t to datasets collected locally in hospitals. Table 1 lists a comparison among different decision support systems in the literature including the dataset size in each study. As can be seen, there is a noticeable diversity of dataset sizes between the different studies.

B. STAGE ONE: ONLINE PREDICTION STAGE

This stage provides unsupervised predictions of each vital sign using the HTM algorithm in real time. We have introduced HTM for the medical stream in [30]. Predictions are made in parallel for each vital sign, CO, CI, MAP, Dia, Sys, SV, SVI, SVR, SVRI, HR, SVV, PPV, SPV, and HRV. Each stream is processed separately, and each sample is fed one-by-one to an HTM algorithm that makes predictions one step into the future. Fig.1-B shows the list of the predicted vital signs streams; which are the input streams that HTM predicts one step into the future. These streams are inputs to Stage II. Fig.2-B shows the predicted streams as well as the original streams of a randomly selected patient. The streams are shown over three windows. A thorough performance evaluation of the predictive efficiency of HTM compared to the predictor LSTM-Batch is provided in our early work[30]. HTM showed better performance and is completely unsupervised.

C. STAGE TWO: FEATURE EXTRACTION AND SELECTION

Stage II includes feature extraction, feature selection, and a classifier stage. The patient's vital signs are used as features of the MAP as we aim to investigate inter-and intra-dependencies between different vital signs and MAPs for each patient. We exploit the dependencies as we examine the reflection of the MAP anomalies on the different vital signs and patterns associated with them. Classifier efficiency shows the ability to learn these dependencies and to detect early, diverse patterns of vital signs that predict MAP anomalies ahead of time before any extreme value is evident on the monitored signs.

Vital signs are used as simple and combined features through multiplication, division, and squaring The features are used in a simple and combinational form. No statistical features have been used. Although millions of features may have been included, either statistical or combination features, we have decided to use vital signs to emphasize the dependencies between vital signs and MAP. We started with 13 vital signs, which are commonly monitored parameters, CO, CI, Dia, Sys, SV, SVI, SVR, SVRI, HR, SVV, PPV, SPV, HRV. The 13 vital signs used in our system are listed in Fig.1-A. We made different combinations starting with the 13 vital signs leading to 42 features, which represent the starting space feature shown in Fig. 1-C. The size of our feature space does not necessarily need the feature selection stage. Nevertheless, we understand that such a stage is essential in practice as the parameters monitored are expanded and consequently the feature space. , we include a feature selection stage in our simulations and explore different methods to achieve a complete model that is adequate for practical implementation.

Three feature selection methods were evaluated and compared: Relieff algorithm [27], sequential forward selection (SFS)[28], and neighborhood component analysis (NCA)[29], respectively. The three methods were applied separately and their performance was evaluated in terms of accuracy and AUROC and compared to performance using the full feature space. Relieff algorithm showed the best performance and was adopted as the algorithm for selecting features in our model. This will be discussed in more detail in the following section. Relieff algorithm's selected features are listed in Fig.1-D with color code for the cell to represent

the importance of each feature. The darker the color of the cell, the more potent the feature is. The features importance ranking is provided in Fig.2-D. We can conclude that the MAP classification was highly dependent on Dia, Sys, SV, and SVI as well as on their combinations. These features are shown in the top two rows. On the other hand, as seen in 3rd to 5th rows, CI, CO, HR, and SPV along with their combinations were found to be less impactful. Whereas SVV, PPV, HRV, SVRI, SVR, and their combinations were found to have no significance in the classification process. These character names are written in black font in the bottom two rows of Fig. 1-D and Fig.2-D. These features have been discarded by the selection algorithm.

D. STAGE TWO: LSTM CLASSIFIER

LSTM is the classifier proposed to operate in a semi-supervised mood in our model. It is employed due to its efficient performance in the analysis of time series and the advantage it has of controlling the features and the observation period. As discussed in Section III, the Relieff algorithm is used in our system for the feature selection. The selected features, along with the ground truth MAP, are fed to the classifier. In the model configuration, we assign a one-minute observation period to the LSTM. The classifier divides the input into one-minute windows and learns the pattern of the dependency over this time interval. After the training phase has been completed, the classifier can predict the future range of MAP after observing patterns over a one-minute window. The classification is binary, where 0 indicates that the estimated MAP level is within the normal range, and 1 indicates that the MAP level is either anomaly due to hypotension or hypertension. The hypotensive event is defined as the MAP value < 65 mm Hg for a one-minute time window, and the hypertensive event is defined as the MAP value > 140 mm Hg for one minute [1, 2]. The classifier begins training to learn the MAP ground truth and its associated patterns on various vital signs. The training period uses data collected during patient observation time as discussed in Section IV-A, After spinal anesthesia, the classifier is assigned to provide a one-minute forecast decision on the MAP level. The performance of the proposed classifier will be evaluated and compared with the performance of the LR classifier in Section V[2, 6].

V. RESULTS AND DISCUSSION

The efficiency of the predictive stage, shown in blue in Fig.1, was investigated in our previous publication on the use of HTM for predictive medical streams[30]. In which HTM as a predictive tool has been compared to the state-of-the-art LSTM-Batch with superior performance. The predicted stream error metrics in terms of the root mean square error (RMS) and mean absolute percentage error (MAPE) improved by 50% and 60%, respectively compared to LSTM-Batch. We illustrate HTM prediction error over a window of 100 samples, the CO original signal, predicted signal, as well as the confidence interval are plot in Fig.2 A vs B. These results belong to a randomly selected patient. The prediction is shown to be accurate with a small error and confidence interval. Detailed results and discussions on the use of HTM as a predictive tool can be found in [30].

TABLE II
THE PERFORMANCE USING ALL FEATURES AND USING DIFFERENT SELECTION CRITERIA

	All Features	Relieff Selection	NCA Selection	SFS Selection
Accuracy	0.985	0.975	0.971	0.971
AUROC	0.978	0.957	0.766	0.753

NCA = Neighbor Component Analysis, SFS = Sequential Forward selection, AUROC = area under the receiver operating characteristic.

TABLE III
THE EFFECT OF HTM PREDICTIVE ACCURACY ON THE CLASSIFIER PERFORMANCE IN TERMS OF THE ACCURACY AND THE AREA UNDER THE RECEIVER CHARACTERISTIC AUROC.

	Stage Two (LSTM)	Stage One+ Two (HTM+LSTM)	Stage Two (LR)	Stage One+ Two (HTM+LR)
Accuracy	0.921	0.895	0.910	0.881
AUROC	0.940	0.886	0.800	0.719
t-test	$p < 0.05$		$p < 0.05$	

LSTM = long short-term memory, HTM = hierarchical temporal memory, LR = logistic regression.

Stage II starts with the feature extraction and selection steps; the late-stage has been explained in Section 1V-C. Explanation of the obtained feature space of size 42 based on 13 vital signs can be seen in Fig.1 A-C. We investigate three feature selection methods such as Relieff importance ranking method, sequential forward selection method (SFS), and neighborhood component analysis (NCA) [27-29]. Each feature selection method is applied separately, and the classifier performance metrics are computed in each case. Both the accuracy and the AUROC are used to compare the three methods with the performance metrics using the entire feature space. Table II lists the complete feature space, Relieff, NCA, and SFS. Relieff shows a slight decrease in both accuracy, from 0.985 to 0.975, and AUROC, from 0.978 to 0.957, compared all feature space. The use of both NCA and SFS results in significant performance degradation of the AUROC by 20%. We, therefore, choose to use the Relieff algorithm as the feature selection method in our simulations due to its enhanced performance.

Next is the classification stage; we propose to use the LSTM classifier and compare it to the widely used and robust LR classifier. As mentioned earlier, the classifier is trained for approximately 7 minutes on the data obtained during the observation phase, then the classifier is assigned after the spinal anesthesia to provide a forecast decision on the MAP level one step into the future. Each classifier model was trained separately for each patient, as the objective of this research is to provide each patient with a personalized decision support system. The LSTM classifier is

TABLE IV
THE COMPARATIVE PERFORMANCE ANALYSIS OF THE PROPOSED SYSTEM USING LSTM CLASSIFIER AND LR USING SEVERAL METRICS.

	Proposed Decision Support (HTM+LSTM)	State-of-the-art Decision Support (LR Classifier)
Accuracy	0.895	0.881
Precision	0.762	0.576
Recall	0.912	0.506
F1-score	0.733	0.880
AUROC	0.886	0.719
SD	0.094	0.227
%95CI	0.768-1	0.437-1
Val error	0.016	0.002

LSTM = long short-term memory, HTM = hierarchical temporal memory, LR = logistic regression, AUROC= area under the receiver operating curve, SD= standard deviation, %95CI= 95% confidence interval, Val error= validation error.

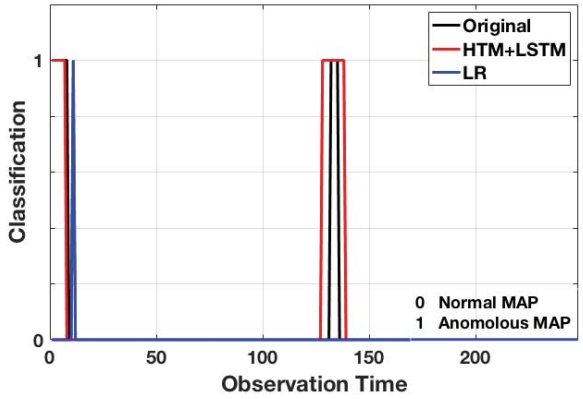


FIGURE 3. Comparison between the original classification (in black), the resulting classification using the proposed system based on LSTM (in red) and using logistic regression (in blue). Zero denotes a normal MAP status and one represents MAP disorder. The decision plot belongs to a randomly chosen patient.

TABLE V
THE PERFORMANCE EVALUATION OF DIFFERENT MODELS IN TERMS OF THE DECISIONS PRECISION AND DECISION-TO-EVENT TIME

	Precision	**Recall**	**Early decisions (sec)**	**Late Decisions (sec)**	**Missed Events**
State-of-the-art (LR)	0.576	0.506	-	200	50%
Proposed System (LSTM)	0.69	0.88	90	100	20%
Proposed System (HTM+LSTM)	0.762	0.912	150	75	10%

.

implemented using MATLAB and consists of an input layer, a bidirectional LSTM layer, a fully connected layer, a SoftMax layer, and finally a classification layer. The number of hidden units (300±5), number of epochs (400±10), and mini-batch size (30±5), which were found through an optimized search in which accuracy was observed and overfitting was prevented. The previous values refer to the mean ± percent 95 confidence interval (CI) for each parameter. The threshold was set to 1. The model was validated using nested k-fold cross-validation with an overall cross-validation value of k=10. In our simulations, the LR model is implemented using MATLAB. The sparse solver was used, where the model was validated using 10-fold cross-validation.

In our model HTM and LSTM are complementary; HTM provides unsupervised predictions in real time, and LSTM provides the classification. HTM provides the proposed framework the privilege of working unsupervised in real time, which is one step towards practical clinical decision support without excessive modeling. HTM, therefore, gives the proposed structure the advantage of early forecasts, where the predicted streams are one-time step ahead of the original sequence, which is 10 seconds in our simulations. On the other hand, given the MAP ground truth, the LSTM classifier is essential to analyze and train on predicted streams. It studies mutual dependencies and provides early MAP level forecasts. This is a role that HTM cannot do.

The proposed machine learning presents an unsupervised prediction algorithm and a semi-supervised classification algorithm. Since unsupervised learning normally leads to performance degradation, we have decided to examine the performance of each stage individually in addition to the overall/accumulative performance of the proposed structure. Furthermore, we compare each classifier performance with and without HTM as a predictive stage as described in Table III. Column one refers to LSTM without prediction, column two points to results using HTM+LSTM, column three displays results using LR, and column four highlights the results HTM+LR. When using HTM, the LSTM classification accuracy decreases from 0.92 to 0.89and from 0.94 to 0.886. Likewise, the precision of the LR classification decreases from 0.91 to 0.88 and AUROC from 0.8 to 0.71. This means the performance has slightly decreased using unsupervised HTM as a first stage. But, it is marginal compared to the advantage it offers by enabling early forecasts. This performance difference/drop was examined using a *t-test* to evaluate its significance, p value<0.05 was considered statistically significant. As pre-mentioned, HTM brings the proposed system the advantage of working in real time at the bedside, saving the retraining phase, which is expected to lead to earlier decisions compared to the state-of-the-art systems. To emphasis this comparison aspect, different parameters defining the decision-to-event time and decision success rate are listed in

Table V. It compares the decision support systems using LR, LSTM, and HTM+LSTM based on the precision, recall, and decision-to-event time. The later metrics are more reflective of the system success rate to predict different events, correct and missed events compared. Using HTM+LSTM enhanced the system precision by 20% compared to LR and by 10% compared to the use of LSTM alone. While it enhanced the recall by 40% compared to LR and 10% compared to LSTM. Which means that the false positive and negatives have decreased. We considered another parameter which is the decision-to-event time. Using HTM was shown to lead to "earlier" decisions by 50% and decreased the decision latency by 25%. The number of missed events have also decreased from 50% using LR, to 20% using LSTM, and the best rate was achieved using HTM+LSTM by only 10%.

The overall performance of the prediction-and-classification system using HTM-LSTM is compared with the state-of-the-art LR in terms of precision, accuracy, recall, F1-score, AUROC, SD, and 95 % CI. The validation error is also taken into account. Table IV lists all the results reflecting the average of each metric over the entire data set. The comparison shows that the proposed system outperforms LR in terms of the specified metrics. We have considered another comparative measure, which is the time margin between the forecast and actual event. As a sample, the resulting decision on the MAP level of a randomly selected patient is observed in Fig. 2. The original classification is shown in black, the decision using the proposed system is shown in red and the classification using the LR is shown in blue. The proposed system was able to predict both events correctly. The second "event" was predicted 4-time windows ahead of time, equivalent to 200 seconds. The average prediction time per patient was 150 seconds and the delayed predictions were 75 seconds on average. The prediction of the "normal" MAP status is also shown to encounter a delay of 100 seconds. On the other hand, the LR classifier is unable to predict both events. This is also consistent with the performance evaluation listed in Table IV and Table V, which indicates a higher recall than the precision values. These results show the advantage of using HTM for earlier and successful decisions on the MAP level compared to the state-of-the-art systems.

VI. CONCLUSION

This paper proposes a new personalized real-time decision support system implementing a new hybrid prediction-and-classification machine learning system. The proposed system is considered to be a breakthrough compared to state-of-the-art real-time systems using three phase-based models as it operates directly at the bedside. It saves extensive offline modeling, uncertainty, and delays associated with conventional systems. The proposed new machine learning framework consists of two stages: hierarchical temporal memory (HTM) predictor stage, first applied to medical signals, and long short-term memory (LSTM) classifier.

The proposed systems have the privilege of working at the bedside in real time, supported by the capabilities of HTM. The framework investigates and learns from the dependencies between the vital signs and the MAP, where the vital signs are used as features of the MAP. Careful selection and extraction of the features were performed using the Relieff algorithm. The LSTM classifier was then configured to predict the MAP level for each patient by detecting the associated pattern on the patient's vital signs.

The proposed system performance is assessed against the state-of-the-art logistic regression (LR). Accuracy, the area under the receiver operating characteristics (AUROC), precision, recall, and F1-score are used as evaluation criteria. The proposed system has shown improved performance, a 20% improvement in AUROC, a 20% improvement in precision, and a 35% increase in recall. The proposed system also lead to a 50% increase in early decisions, decreased latency by 25%, and decreased missed events by 50% resulting in an average decision-to-event time of 150 sec.

REFERENCES

[1] B. Pratt, L. Roteliuk, F. Hatib, J. Frazier, RD Wallen, "Calculating arterial pressure-based cardiac output using a novel measurement and analysis method," Biomedical Instrument Technology, vol. 41, pp. 403-411, 2007.

[2] F. Hatib, Z. Jian, S. Buddi, C. Lee, J. Settels, K. Sibert, J. Rinehart, and M. Cannesson," machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis," Anesthesiology vol. 129, pp. 663–674, 2018.

[3] V. Salmasi, K. Maheshwari , D. Yang, EJ. Mascha, A. Singh, DI. Sessler, A. Kurz," Relationship between intraoperative hypotension, defined by either reduction from baseline or absolute thresholds, and acute kidney and myocardial injury after noncardiac surgery: A retrospective cohort analysis," Anesthesiology vol. 126, pp. 47–65, 2017.

[4] R. He, Z. Huang, L. Ji, J. Wu, H. Li, Z. Zhang, "Beat-to-Beat ambulatory blood pressure estimation based on random forest," 13th IEEE International Conference on Wearable and Implantable Body Sensor Networks *BSN*, California, USA, 2016.

[5] M. Moreira, J. Rodrigues, A. Oliveira, K. Saleem, and A. Venancio Veto, "Predicting Hypertensive Disorders in High-risk Pregnancy Using the Random Forest Approach," IEEE International Conference on Communications ICC, 2017.

[6] A. Hiwale, P. Talele, and R. Phalnikar," Prediction of Pregnancy-Induced Hypertension Levels Using Machine Learning Algorithms", Computing in Engineering and Technology, vol.1025, pp.597-608, 2020.

[7] M. Flechet, F. Guiza, M. Schetz, P. Wouters, I. Vanhorebeek, I. Derese, J. Gunst, I. Spriet, M. Casaer, G. Van den Berghe, and G. Meyfroidt," AKIpredictor, an online prognostic calculator for acute kidney injury in adult critically ill patients: development, validation, and comparison to serum neutrophil gelatinase-associated lipocalin," Intensive Care Medicine, vol. 43, pp. 764-773, 2017.

[8] V. Nagendra, H. Gude, D. Sampath, S. Corns, S. Long, "Evaluation of SVM and RF Classifiers in a real-time fetal monitoring system based on cardiotocography data" IEEE Conference on Computational intelligence in Bioinformatics and Computational Biology, Manchester, UK, August 2017.

[9] B. Van Der Ster, F. Bennis, T. Delhaas, B. Westerhof, W. Stok, and J. Van Lieshout, "Support Vector Machine Based Monitoring of Cardio-Cerebrovascular Reserve during Simulated Hemorrhage," Frontiers in Physiology, vol. 8, pp. 1-10, 2018.

[10] M. Hsieh, M. Hsieh, C. Chen, C. Hsieh, C. Chao, and C. Lai, "Comparison of machine learning models for the prediction of mortality of patients with unplanned extubating in intensive care units," *Scientific Reports*, vol. 8, pp. 1-7, November 2018.

[11] A. Fernandez, D. Fernandez, and M. Sanchez, "A Decision Support system for predicting the treatment of ectopic pregnancies"

International Journal of Medical Informatics, vol.129, pp.198-204, 2019.

[12] A. L'Heureux, K. Grolinger, H. El-Yamany, and M. Capretz, "Machine learning with big data: challenges and approaches," IEEE Access, vol.5, pp. 7776-7797, April 2017.

[13] C. Wang, Z. Zhao, L. Gong, L. Zhu, Z. Liu, and X. Cheng, "A distributed Anomaly Detection System for In-Vehicle Network Using HTM," IEEE Access, vol.6, pp. 9091-9098, March 2018.

[14] J. Mackenzie, J. F. Roddick, and R. Zito, "An Evaluation of HTM and LSTM for short-term arterial traffic flow prediction," IEEE Transactions on Intelligent Transportation Systems, pp. 1-11, August 2018.

[15] S. Ahmed, A. Lavin, S. Purdy, and Z. Agha, "Unsupervised Real time Anomaly Detection for Streaming Data," Neurocomputing, Vo.262, pp. 134-147, November 2017.

[16] Y. Cui, S. Ahmad, and J. Hawkins, "Continuous online sequence learning with an unsupervised neural network model," Neural Computation, Vol.28, Issue.11, pp. 2474-2504, November 2016.

[17] J. Hawkins and S. Ahmad, ''Why neurons have thousands of synapses, a theory of sequence memory in the neocortex,'' *Frontiers Neural Circuits*, vol. 10, no. 23, pp. 1–13, March 2016.

[18] S. Ahmad and J. Hawkins. (2015). Properties of sparse distributed representations and their application to hierarchical temporal memory. [Online]. Available: http://arxiv.org/abs/1503.07469

[19] A. Pulver and S. Lyu, " LSTM with working memory," IEEE joint conference on neural networks *IJCNN*, 2017.

[20] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Stwunebrink, and J. Schmidhuber, "LSTM: a search space odyssey," IEEE Trans on neural networks and learning systems, vol. 28, no. 10, pp. 2222-2232. October 2017.

[21] G. Ercolano, D. Riccio, and S. Rossi, "Two Deep approaches for ADL recognition: A multi-scale LSTM and a CNN-LSTM with 3D matrix skeleton representation

[22] I. Garcia-Magarino, R. Muttukrishnan, and J. Lloret "Human-Centric AI for Trustworthy IoT Systems with Explainable Multilayer Perceptron", IEEE Access, Vol.7, 2019.

[23] M. Erango, A. Frigessi. L. A. Rosseland. (2018, July). A three minutes supine position test reveals higher risks of spinal anesthesia induced hypotension during cesarean delivery. An observational study. F1000 Research. DOI: 10.12688/f1000research.15142.1

[24] D. Kearney, S. McLoone, and T. Ward, "Investigating the Application of Transfer Learning to Neural Time Series Classification"30th Irish Signals and Systems Conference, June 2019.

[25] J. Makhoul, "Linear prediction: A tutorial review," Proceeding of IEEE, vol. 63, pp. 561-580, April 1975.

[26] P. Prandoni and M. Vetterli, "An FIR cascade structure for adaptive linear prediction," IEEE Trans on signal process, vol.46, pp 2566-2571, Sept 1998.

[27] M. Robnik-Sikonja, and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Machine Learning*, vol. 53, pp. 23–69, 2003.

[28] R. Kohavi and G. H. John, "Wrappers for features subset selection," *Artificial Intelligence, vol.97. pp-273-324, 1997.*

[29] W. Yang, K. Wang, and W. Zuo, "Neighborhood Component Feature Selection for High-Dimensional Data," Journal of Computers, vol. 7, pp. 161-168, January 2012.

[30] N. El-Ganainy, I. Balasingham, P. S. Halvorsen, L. A. Rosseland, "On the performance of hierarchical temporal memory predictions of medical streams in real time," in 13th International Symposium on Medical Information Communication Technology *ISMICT*, Oslo, Norway, 2019.

[31] M. Taylor *et al.*, (Feb. 2016). *NuPIC: 0.5.0*. [Online]. Available: https://doi.org/10.5281/zenodo.46074

NOHA O. EL-GANAINY received her Ph.D. in Electrical Engineering and electro Physics from the Faculty of Engineering, Alexandria University, Egypt in 2010. She is currently affiliated as a postdoctoral fellow at Norwegian university for science and technology NTNU funded by a European research consortium for informatics and mathematics (ERCIM).

From 2010 to 2017, she worked as an assistant professor at the Arab Academy for science, technology, and maritime transport (AASTMT). She has teaching experience in different institutions in both Egypt and Norway. Her current research work addresses machine learning with a special interest in algorithms learning and modeling different types of data in real time. Her previous research projects exploited the use of different signal processing techniques in wireless communication systems on the MAC layer including signal detection and estimation, cooperative and cognitive transmission/receiving, channel estimation, and noise cancellation. She has also worked on image processing watermarking, stenography, change, and motion detection.

Dr. El-Ganainy has published several scientific papers and articles in international journals and conferences and received the young scientist's award from the union radio scientific international (URSI) in 2011. She has served as a reviewer, TPC, and track chair for many international journals and conferences.

ILANGKO BALASINGHAM (SM'12) received the M.Sc. and Ph.D. degrees from the Department of Electronic Systems, Norwegian University of Science and Technology (NTNU), Trondheim, Norway in 1993 and 1998, respectively, both in signal processing. He performed his Master's degree thesis at the Department of Electrical and Computer Engineering, University of California Santa Barbara, USA. From 1998 to 2002, he worked as Research Engineer developing image and video streaming solutions for mobile handheld devices at Fast Search & Transfer ASA, Oslo. Fast was a startup and was acquired by Microsoft Inc. and renamed as Microsoft Development Center Norway.

Since 2002 he has been with the Intervention Centre, Oslo University Hospital, where he is Head of Section for Medical ICT Research and Head of Wireless Biomedical Sensor Network Research Group. He is a Professor of Medical Signal Processing and Communications at NTNU since 2006. His research interests include super robust short-range communications for both in-body and on-body sensors, body area sensor network, microwave short range sensing of vital signs, short range localization and tracking mobile sensors, signal and image processing, and molecular and Nano communication networks.

Prof. Balasingham has authored or co-authored more than 90 journal papers, 170 full conference papers, 8 book chapters, 42 abstracts, and 16 articles in popular press and holds 6 issued patents and 10 disclosure of inventions. He has supervised 21 Postdocs, 21 PhDs, and 30 Masters. He was the General Chair of the 2012 Body Area Networks (BODYNETS) and 2019 IEEE International Symposium of Medical Information and Communication Technology (ISMICT) and TPC Chair of the ACM NANOCOM 2015. He serves as Area Editor of Elsevier Nano Communication Networks and Steering Committee Member of ACM NANOCOM.

PER STEINAR HALVORSEN received a medical degree from the University of Oslo, Norway, in 1990. He was an authorized specialist in anesthesiology at the Oslo University Hospital in 2005 and received a Ph.D. degree in 2010 and M.H.A in 2014 at the University of Oslo, Norway. In 2019 he was appointed as a Professor in anesthesiology at the University of Oslo and currently also holds a clinical position as a senior consultant and section manager at The Intervention Centre at Oslo University Hospital."

LEIV ARNE ROSSELAND received his medical degree from the University of Oslo in 1991 and graduated as a Specialist in Anesthesiology and Intensive care in 1998. He is a Professor of Anesthesiology at the Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. He is also Head of the Department of Research and Development, Division of Emergencies and Critical Care, Oslo University Hospital, Oslo, Norway. He has worked as a Consultant Anesthesiologist at Oslo University Hospital since 2000 and in the same year started his Ph.D., which he obtained in 2005. He has been a Professor at the Institute of Clinical Medicine since 2011 and became the Head of the Department of Research and Development in 2014.

His research interest is in clinical anesthesiologic research, especially pain, inflammation, and intensive care medicine. In particular, his research includes hemodynamic changes in healthy and pathologic pregnancies and the relationship between acute pain and inflammation. His group also studies pain and its impact on cardiovascular homeostasis, pain and symptom burden in intensive care unit (ICU) patients and family caregivers, as well as evaluating patient pathways after major surgery and ICU admission.

He has received prizes from the Norwegian Association of Anesthesiologists annual conference in 2013 and was named Researcher of the Month in January 2018 by the South-Eastern Norway Regional Health Authority. He is the recipient of a Fast Track to Innovation Grant (project number: 691262) from the European Research Council, which is a collaboration between Oslo University Hospital and two industry partners to develop the INTUI-VIEW Ultrasound Tracking System, an intelligent needle tracking system that aims to use ultrasound imaging for improved minimally invasive interventions.

Prof. Rosseland has published 73 original scientific research papers, two book chapters, and numerous national and international conference presentations and reports.