



Ventilator pressure prediction employing voting regressor with time series data of patient breaths

Health Informatics Journal
1–24

© The Author(s) 2025

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14604582241295912

journals.sagepub.com/home/jhi



Ali Raza*

Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan, Pakistan

Furqan Rustam

School of Systems and Technology, Department of Software Engineering, University of Management and Technology, Lahore, Pakistan

Hafeez Ur Rehman Siddiqui

Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan, Pakistan

Emmanuel Soriano Flores

Higher Polytechnic School, Universidad Europea Del Atlantico, Santander, Spain

Department of Project Management, Universidad Internacional Iberoamericana, Campeche, Mexico

Juan Luis Vidal Mazón

Higher Polytechnic School, Universidad Europea Del Atlantico, Santander, Spain

Project Department, Universidade Internacional Do Cuanza, Município do Kuito, Angola

Isabel de la Torre Díez

Department of Signal Theory and Communications and Telematic Engineering, University of Valladolid, Valladolid, Spain

María Asunción Vicente Ripoll

Department of Health Psychology, Universidad Miguel Hernandez Ingeniería Telemática, Alicante, Spain

Imran Ashraf

Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, South Korea

*Current Affiliation: Department of Software Engineering, The University of Lahore, Lahore, Punjab 54000, Pakistan

Corresponding authors:

Isabel de la Torre Díez, Department of Signal Theory and Communications, University of Valladolid, Paseo de Belén 15, Valladolid 47011, Spain.

Email: isator@tel.uva.es

Imran Ashraf, Department of Information and Communication Engineering, Yeungnam University, Daehak ro 280 IT Building Room 326, Gyeongsan 38541, South Korea.

Email: imranashraf@ynu.ac.kr



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further

permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Abstract

Objectives: Mechanical ventilator plays a vital role in saving millions of lives. Patients with COVID-19 symptoms need a ventilator to survive during the pandemic. Studies have reported that the mortality rates rise from 50% to 97% in those requiring mechanical ventilation during COVID-19. The pumping of air into the patient's lungs using a ventilator requires a particular air pressure. High or low ventilator pressure can result in a patient's life loss as high air pressure in the ventilator causes the patient lung damage while lower pressure provides insufficient oxygen. Consequently, precise prediction of ventilator pressure is a task of great significance in this regard. The primary aim of this study is to predict the airway pressure in the ventilator respiratory circuit during the breath. **Methods:** A novel hybrid ventilator pressure predictor (H-VPP) approach is proposed. The ventilator exploratory data analysis reveals that the high values of lung attributes R and C during initial time step values are the prominent causes of high ventilator pressure. **Results:** Experiments using the proposed approach indicate H-VPP achieves a 0.78 R^2 , mean absolute error of 0.028, and mean squared error of 0.003. These results are better than other machine learning and deep learning models employed in this study. **Conclusion:** Extensive experimentation indicates the superior performance of the proposed approach for ventilator pressure prediction with high accuracy. Furthermore, performance comparison with state-of-the-art studies corroborates the superior performance of the proposed approach.

Keywords

COVID-19, deep learning, machine learning, mechanical ventilation, ventilator pressure prediction

Introduction

A mechanical ventilator is a mechanical machine that overcomes the issue of patient breathing (ventilate) during critical illness.¹ The ventilator is utilized when a patient cannot breathe naturally and independently due to severe illness or syndrome like COVID-19.² The ventilator is connected to the patient using a hollow tube (artificial airway). The ventilator supply tube helps pump the air into the patient's lungs. The ventilator saves the patient's lives by giving them enough oxygen to the body and improving their breathing mechanism. Although the ventilator has a significant role in saving precious lives, it is also associated with many life risks. During the ventilation process in the ventilator, the air is pumped into the patient's lungs based on the air pressure.³ The low or high intensity of ventilator pressure can result in a patient's life loss. The high air pressure intensity in the ventilator causes the patient lung damage. To detect the ventilator pressure, an artificial intelligence-based system must be attached to the ventilator to alert the high or low air pressure. The air pressure in the ventilator can be controlled using the artificial intelligence-based system alert to save a patient's life risk.

Mechanical ventilation with correct air pressure is highly required for patients during COVID-19 illness to save them from respiratory distress syndrome and lung injury. According to a recent report, the reported mortality rate is 50% to 65%⁴ in patients with COVID-19 syndrome in the intensive care unit (ICU). With COVID-19 illness, 97% of patients require intensive mechanical ventilation support. The studies have shown that the pandemic mortality rate was 100% amongst patients on invasive mechanical ventilation.⁵ The patient's mortality on invasive mechanical ventilation was reported as 97% in China,⁶ 88.6% in New York,⁷ 43% in the UK,⁸ 31% in Spain,⁹ 88.8% in Australia,¹⁰ and between 40 and 60% in India.¹¹ However, with later developments in ventilators,

the fatality rate is reported to be reduced to 40%.^{12,13} Despite that, a proper ventilator with normal air pressure is a crucial need for patients.

Machine learning and deep learning-based techniques are the domains of artificial intelligence that are utilized for ventilator pressure prediction in this research study. The artificial intelligence-based systems are trained on historical data, and a predictive system is formed to predict the target values for unseen data. The main motive of artificial intelligence-based techniques is to find the patterns or relationships in data and learn from them to predict unseen data. Machine learning and deep learning-based techniques achieve high-performance accuracy results in medical data predictions.¹⁴ Nowadays, artificial intelligence-based techniques have many applications. They are most commonly used in bioinformatics for diagnosing patients, MRI-based image processing, drug discovery,¹⁵ and many more. The following are our research key contributions to ventilator pressure prediction

- A novel hybrid ventilator pressure predictor (H-VPP) approach is proposed based on a hybrid of decision tree regressor (DTR) and random forest regressor (RFR). The approach is applied to the historical data to predict the ventilator pressure prediction for COVID-19 patients.
- The ventilator exploratory data analysis (VEAA) is applied to determine the data patterns and valuable insights from the ventilator dataset. The VEAA is applied to determine the significant factors associated with the high or low pressure in the ventilator.
- For performance evaluation of the proposed approach, several machine learning and deep learning-based models are employed in this study. Such models include multilayer perceptron (MLP), linear regression (LR), DTR, RF, stochastic gradient descent regressor (SGDR), Bayesian ridge (BR), and light gradient boosting machine regressor (LGBMR). In addition, recurrent neural networks (RNN), long short-term memory (LSTM), and gated recurrent units (GRU) are also applied for experiments.
- The regression performance metrics mean absolute error (MAE), median absolute deviation (MAD), mean squared error (MSE), root mean squared error (RMSE), R-squared (R^2) score, adjusted R^2 score, and variance are used for evaluating the performance of adopted models. The regression analysis is applied based on the mapping difference between target predictions and actual values of pressure.

The remainder of this research study is divided into four sections. Section "Related work" is based on the related literature analysis and past applied state of art studies analysis. The study methodology is examined in the Section "Methodology". It also contains the details of the employed machine learning and deep learning models. The results and discussions are given in the Section "Results". Section "Conclusion" provides the conclusion of the study.

Related work

This section is based on the related literature to the current study. The past applied state-of-the-art studies are examined for ventilator pressure prediction along with the proposed techniques and performance evaluation parameters.

The prediction of ventilator weaning success based on biosignal using machine learning was proposed in¹⁶. The biosignal data of 89 patients were utilized for model building and evaluations. The RF classifier was proposed to predict ventilator weaning failure. The performance evaluation is done by bootstrapping.¹⁷ The proposed model achieved a 0.81 accuracy score of the receiver operating characteristic curve (ROC) with 95% confidence. The inspired oxygen prediction and

ventilator modes prediction for COVID-19 patients were proposed in¹⁸ which followed a deep learning approach. The proposed approach for the prediction task was the artificial neural network developed using Python programming. The model building was based on real-time patient data. The proposed model achieved a mean square error of 0.093 and an R-value of 0.81.

The authors proposed the prediction of fluid responsiveness in ventilator-induced variation in¹⁹. The patient's data was utilized for model building and evaluations. The pulse-wave transit time (PWTT) technique was utilized to predict fluid responsiveness. A 0.97 accuracy score for RoC is obtained with a p value less than 0.0001 by the proposed PWTT technique. The prediction of ventilator-associated pneumonia and respiratory distress syndrome using machine learning techniques was proposed in²⁰. The patient data from the Early versus delayed enteral nutrition (EDEN) was utilized for machine learning model building and evaluations. The multivariate binary logistic regression was the proposed approach for the prediction task. The proposed model achieved a 0.74 accuracy score for the ROC curve while the p value is 0.185.

The prediction of mode shifting of the ventilator for adult patients in the medical intensive care unit using machine learning technique was proposed by.²¹ The data of 1483 adult patients were utilized for model training and testing. A modified XGBoost algorithm was used for experiments. The proposed model achieved a 0.76 accuracy score for the ROC curve. The authors proposed a deep learning-based model for predicting the magnitude of asynchrony breathing in machine ventilators in²². Data of 400,000 unique asynchrony breathing patients are utilized for model building. The convolutional autoencoder model is used to carry out experiments. The proposed model achieved a 0.008 median validation error using the K-fold analysis.

The study²³ introduces a multi-task Gaussian-based neural network designed to predict the need for mechanical ventilators in COVID-19 patients. The study proposes a robust real-time prediction model to determine the likelihood of in-hospital COVID-19 patients requiring mechanical ventilation (MV). This end-to-end neural network model integrates the Multi-task Gaussian Process to manage irregular sampling rates in observational data and employs a self-attention neural network for the prediction task. The proposed model was evaluated using a large dataset of 9532 nationwide in-hospital COVID-19 patients, achieving an area under the receiver operating characteristic curve (AUROC) score of 0.79 and an area under the precision-recall curve (AUPR) score of 0.39. While the performance of the proposed method is moderate, further performance enhancement strategies are necessary.

The analysis of reviewed studies, as given in [Table 1](#), indicates that machine learning approaches are predominantly used for ventilator pressure prediction and other issues related to mechanical ventilators. Results suggest that the produced accuracy is still lower than the desired output and further research efforts are needed in this regard.

Methodology

This section describes the proposed approach in detail, along with a description of the machine learning models used for experiments and the dataset. The architecture of the proposed methodology is illustrated in [Figure 1](#). Patients' ventilator breath time series data is utilized for research model building and evaluations. The ventilator exploratory data analysis (VEEA) is applied to obtain the data patterns and insights that are the primary cause of low or high ventilator pressure. The dataset normalization is applied to transform the dataset into a unit share, achieving the best performance scores. The dataset is split with a ratio of 0.8 to 0.2, where 80% of the dataset is used for training the employed models while 20% is used for testing. We also used the GroupKFold

Table I. The ventilator pressure prediction literature summary analysis.

Ref.	Year	Approach	Dataset	Accuracy/ Error value	Research aim
¹⁶	2021	Random forest	Biosignal data of 89 patients	0.81	Prediction of ventilator weaning success based on biosignal data using machine learning
¹⁸	2022	Artificial neural network	Real-time patient data	0.093	Inspired oxygen and ventilator modes prediction using deep learning technique for COVID-19 patients
¹⁹	2020	Pulse-wave transit time (PWTT)	Patient's data	0.97	The prediction of fluid responsiveness in ventilator-induced variation using pulse-wave transit time (PWTT) was proposed
²⁰	2020	Logistic regression	Patient data from the early versus delayed enteral nutrition (EDEN)	0.74	Prediction of ventilator-associated pneumonia and respiratory distress syndrome using machine learning techniques
²¹	2022	XGBoost algorithm	Data of 1483 adult patients	0.76	Prediction of mode shifting of ventilator for adult patients in the medical intensive care unit using machine learning techniques
²²	2021	Convolutional autoencoder model	400.000 unique asynchrony breathing patients	0.008	A deep learning-based model was proposed for predicting the magnitude of asynchrony breathing in machine ventilators

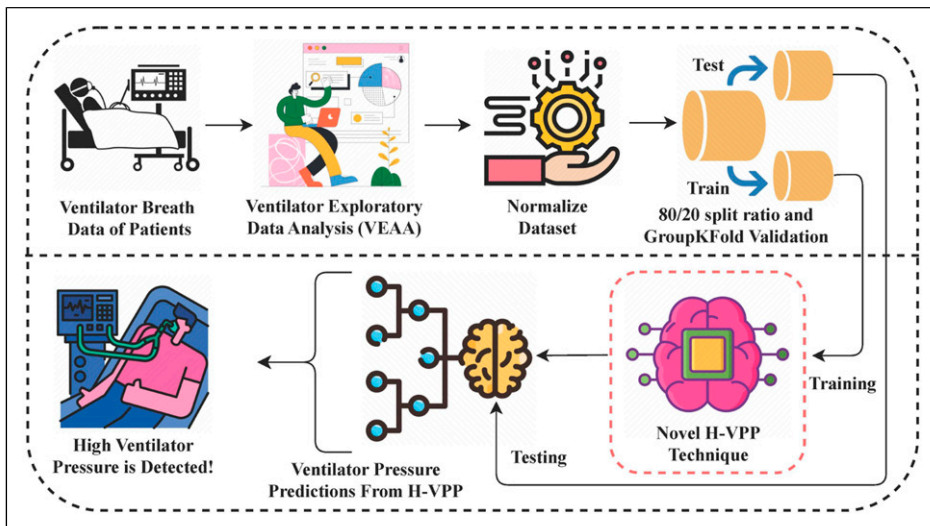


Figure 1. The architectural analysis of proposed methodology for ventilator pressure prediction.

validation split approach for results validation. The proposed predictive model is trained and tested on the ventilator dataset.

Dataset

The ventilator breaths time-series data²⁴ of patients was produced using a modified open-source ventilator²⁵ connected to an artificial bellows test lung²⁶ using a respiratory circuit. The time-series data represents an approximately three-second breath. Each row in the dataset is a time step in a breath and gives the two control signals. The control signals are relevant attributes of the lung and the resulting airway pressure. The dataset-related feature descriptive analysis is presented in [Table 2](#). It shows the details of the attributes, attribute types, and related descriptions.

- R - lung attribute indicating how restricted the airway is (in cm H₂O/L/S). Physically, this is the change in pressure per change in flow (air volume per time). Intuitively, one can imagine blowing up a balloon through a straw. We can change R by changing the diameter of the straw, with higher R being harder to blow.
- C - lung attribute indicating how compliant the lung is (in mL/cm H₂O). Physically, this is the change in volume per change in pressure. Intuitively, one can imagine the same balloon example. We can change C by changing the thickness of the balloon's latex, with higher C having thinner latex and easier to blow.
- Pressure - the airway pressure measured in the respiratory circuit, measured in cm H₂O.

Ventilator exploratory data analysis

The VEDA is a crucial process of determining hidden data patterns, hypotheses, assumptions, and cussing factors. The VEDA is based on the summary of graphical representations and statistics on data to discover patterns. The VEDA contains statistical data analysis, graphs, and charts. The VEDA helps us in our research study to find out several factors that cause high or low pressure in ventilators.

The statistical data analysis is applied to each feature present in the dataset, as shown in [Table 3](#). The analysis is based on the factors of dataset count, mean, standard deviation, minimum values, 25%, 50%, 75%, and maximum values. The analysis demonstrates that 6,036,000.0 is the total count

Table 2. The ventilator dataset descriptive features analysis.

Sr no.	Feature	Data type	Description
1	breath_id	Int64	The unique time step for patient breaths globally
2	R	Int64	Physically, this is the change in pressure per change in flow (air volume per time). The lung attribute indicates how restricted the airway is (in cm H ₂ O/L/S)
3	C	Int64	Physically, this is the change in volume per change in pressure. The lung attribute indicates the lung's compliance (in mL/cm H ₂ O)
4	time_step	float64	The actual timestamp
5	u_in	float64	The control input for the inspiratory solenoid valve. It ranges from 0 to 100
6	u_out	Int64	The control input for the exploratory solenoid valve. It can be either 0 or 1
7	Pressure	float64	The airway pressure measured in the respiratory circuit is measured in cm H ₂ O

Table 3. The statistical features analysis of the employed research dataset.

Feature	Count	Mean	Std	min	25%	50%	75%	Max
breath_id	6,036,000.0	62,838.858701	36,335.256194	1.000000	31,377.000000	62,765.500000	94,301.000000	125,749.000000
R	6,036,000.0	27.036183	19.595490	5.000000	5.000000	20.000000	50.000000	50.000000
C	6,036,000.0	26.080716	17.152314	10.000000	10.000000	20.000000	50.000000	50.000000
time_step	6,036,000.0	1.307225	0.765978	0.000000	0.642900	1.308123	1.965502	2.937238
u_in	6,036,000.0	7.321615	13.434701	0.000000	0.393662	4.386146	4.983895	100.000000
u_out	6,036,000.0	0.620449	0.485275	0.000000	0.000000	1.000000	1.000000	1.000000
Pressure	6,036,000.0	11.220408	8.109703	-1.895744	6.329607	7.032628	13.641030	64.820992

for each feature. Analyzing the target pressure feature, the mean value is 11.220,408, the minimum value is $-1.895,744$ as low ventilator pressure, and the maximum value is 64.820,992 as high ventilator pressure. This analysis represents that data variance is based on different statistical factors.

The time series analysis based on the pressure, u_{in} , and u_{out} features using the different sample values of lung attributes $breath_id$, R , and C is analyzed in Figure 2. The analysis demonstrates that at the time steps 0 to 1, the ventilator pressure and inspiratory solenoid valve control values are high. The ventilator pressure is high when the value of R and C features is above 20. The analysis shows that after time step value 1, the pressure remains normal. This analysis shows that the ventilator pressure and inspiratory solenoid valve have high values during the initial time step values. The values of lung attributes R and C above 20 also increase the ventilator pressure.

The joint plot regression analysis based on 1000 data points is examined in Figure 3. A joint plot is based on three data plots. The first is a bivariate graph representing the regression distribution between the data of two involved variables. The second graph is placed at the top of the bivariate graph horizontally and represents the distribution of the x -axis feature. The third graph is assigned on the right side of the main bivariate graph vertically and represents the distribution of the y -axis feature. The joint plot regression analysis has the univariate and bivariate graphs together to summarize patterns in data distributions.

Figure 3(a) analyzes the bivariate regression data distribution between the u_{in} and pressure feature. The time step unit represents an approximately three-second breath. Each row in the dataset is a time step in a breath and gives the two control signals. The control signals are relevant attributes

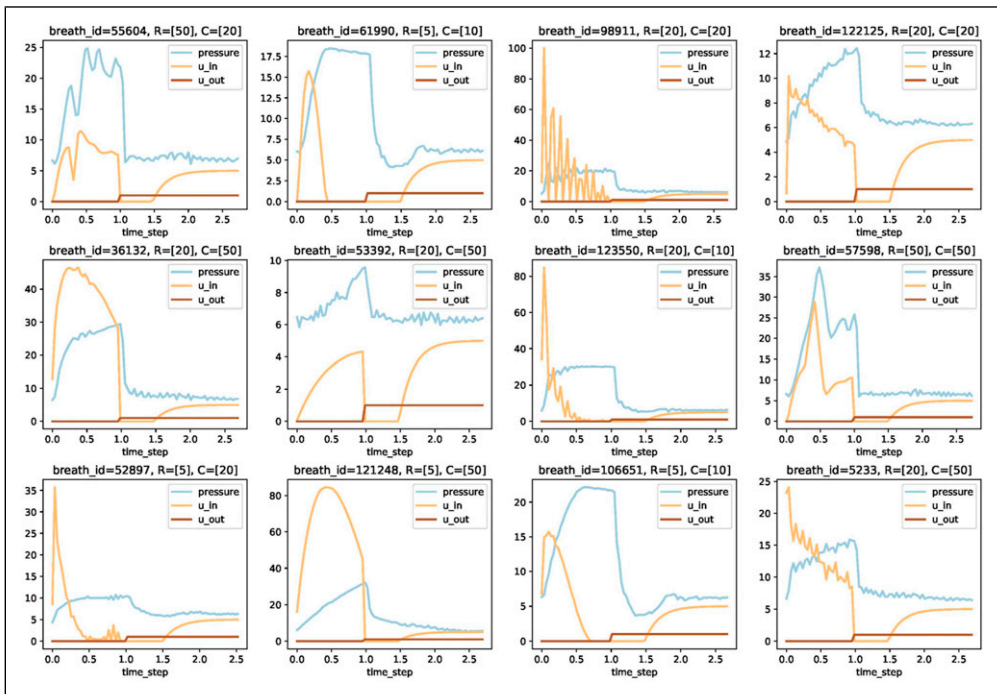


Figure 2. The effects of $breath_id$ with lung attributes R and C on ventilator pressure are analyzed in time series analysis.

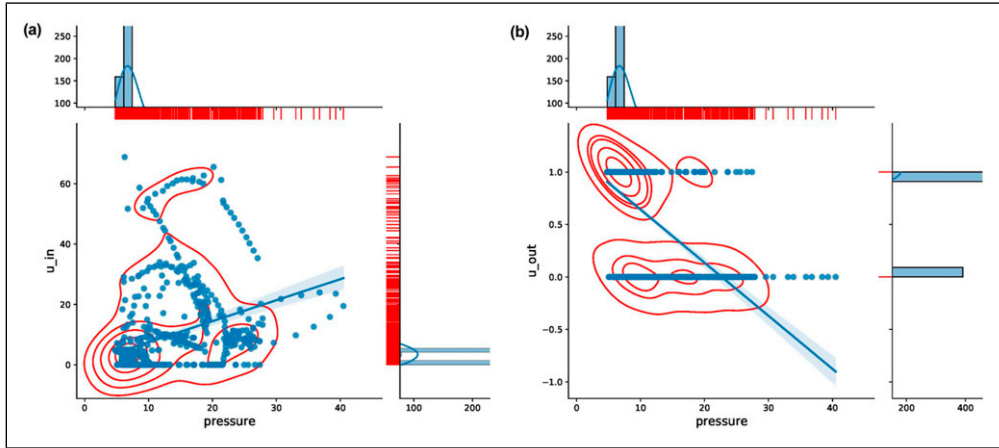


Figure 3. The joint plot regression analysis of the features u_{in} and u_{out} with the ventilator pressure feature, (a) The inspiratory solenoid valve control input analysis with pressure, and (b) The exploratory solenoid valve control input analysis with pressure.

of the lung and the resulting airway pressure. The bivariate regression analysis demonstrates the strength of correlation between features and indicates that $u_{out} - \text{pressure}$ has a 0.61 correlation while $u_{in} - \text{pressure}$ has a 0.31 correlation. The bivariate density curve shows that when the pressure values are between 0 and 30, the u_{in} has a strong distribution relationship. The regression line is also drawn among the data distributions. This analysis represents a high relationship between u_{in} and the pressure feature. The bivariate regression data distribution between the u_{out} and pressure feature is analyzed in Figure 3(b). The analysis represents that the regression line goes negative, which shows that the u_{out} and pressure features have less relationship strength. The bivariate density curve shows that when the pressure values increase, the u_{out} has the same value.

The correlation analysis between the dataset features is examined in Figure 4. The analysis demonstrates that all correlation values are positive. The high correlation between the u_{out} and time_step feature is 0.84 followed by the correlation values of 0.61 between the pressure and u_{out} feature. This analysis represents that the dataset features have good correlation values, which are best for learning techniques training for ventilator pressure predictions.

Normalize dataset

The data normalization is applied to transform all dataset feature data into a unit sphere. The min-max scaler is utilized for ventilator data scaling and normalization. The min-max scaler transforms the feature values into a unit variance scaling range. This min-max scaler scales and translates each feature individually as it is in the given range on the dataset; for example, the scale ranges between one and zero. The data normalization in this research improves the performance of employed learning techniques.

MinMaxScaler is a widely employed data preprocessing technique in machine learning. This normalization helps mitigate the influence of outliers and varying scales among different features, promoting improved model performance. MinMaxScaler preserves the relationships between data points, ensuring that the overall structure and patterns within the dataset remain intact. In summary, the advantages of MinMaxScaler include its ability to normalize data, mitigate the impact of

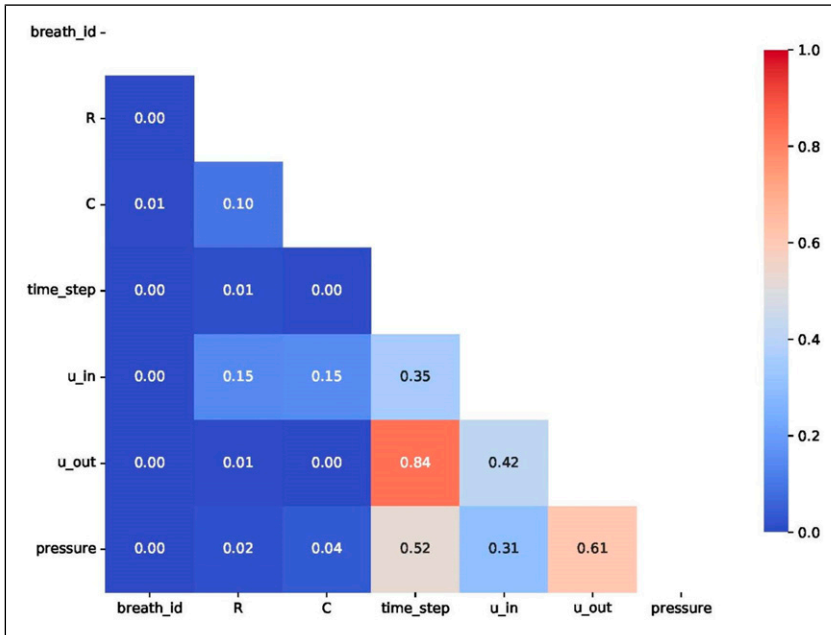


Figure 4. The correlation analysis among the ventilator dataset features used for pressure prediction.

outliers, preserve relationships between data points, and its simplicity of implementation. These attributes collectively make MinMaxScaler a valuable tool in the preprocessing toolbox, enhancing the performance and interpretability of machine learning models across diverse applications.

Dataset splitting

The data splitting is required for the training and testing of employed machine learning models. The dataset splitting is needed to split the data into train and test subsets where the test subset is used to validate the models on unseen test data. The dataset splitting is performed to split the ventilator dataset for training and testing. The splitting ratio is 0.8 to 0.2 for training and testing, respectively.

Employed learning techniques

The applied machine learning and deep learning-based techniques for ventilator pressure prediction are detailed in this section. A total of ten state-of-the-art advanced machine learning and deep learning techniques are employed to predict ventilator pressure. Artificial intelligence (AI) has been widely used in the medical domain over the past decade.²⁷

As the data in the medical is increasing exponentially, the issues of extracting valuable insights from data arise. Machine learning and deep learning models are involved in the predictive process in this regard. The typical applications of AI systems are biology,²⁸ disease diagnosis, virtual nursing assistant, enhanced gene editing, proteomics, genomics, microarrays, bioinformatics^{29–32} and many more. The AI systems solve the problem of obtaining valuable insights from large biological datasets. The machine learning models utilize less computation power as compared to deep learning. Deep learning handles big data efficiently and is mainly used for vision-based applications like

image classification in MRI,³³ wireless capsule endoscopy.³⁴ This study utilizes machine learning and deep learning approaches for predicting ventilator pressure. A brief description of employed state-of-the-art machine learning and deep learning techniques is presented here.

Multilayer Perceptron MLP³⁵ is a family of feed-forward neural networks. MLP architecture contains the input, output, and hidden layers to process the input data. MLP utilizes back-propagation³⁶ techniques during training. The MLP can be used for both classification and regression problems.

Logistic Regression LR³⁷ is a supervised machine learning model mainly used to solve regression problems. The target prediction value is based on the independent data variables. The prediction is obtained by determining the linear relationship between input and output variables.

Decision Tree Regressor DTR³⁸ builds the tree-like flow chart structure for predicting the target values. The input data variables are split and placed into the tree's internal nodes. The target values are placed in the leaf of the tree.

Random Forest Regressor RFR³⁹ is based on creating a forest of multiple trees. The prediction from multiple trees is combined, and the majority prediction is selected as the final prediction value.

Stochastic Gradient Descent Regressor SGDR⁴⁰ is an efficient approach for fitting the linear regressor under loss functions such as logistic regression. The SGDR works by randomly selecting a few data samples instead of the complete dataset during every iteration in the prediction task. The SGDR determines the gradient to minimize the cost function.

Bayesian Ridge BR⁴¹ is suitable to solve the problem where the data is insufficient. The BR uses probability distribution by formulating linear regression. The target prediction values in BR are drawn from a probability distribution.

Light Gradient Boosting Machine Regressor LGBMR⁴² is an ensemble learning model used for regression and classification by constructing decision trees. The multiple involved trees in LGBMR determine the prediction values.

Recurrent Neural Network RNN⁴³ is a deep learning-based model best known for sequential data. In a traditional neural network, the outputs and inputs are independent. However, the RNN follows the looping mechanism, which is based on the working of the previous step. Outputs are input to the current step to predict the output of the layer. This way, RNNs remember the inputs because of their internal memory. The RNN works in the same behavior as the human brain's function.

Long Short-Term Memory The RNN has the problem of vanishing gradients. LSTM⁴⁴ model is an extension of RNN to overcome this issue by extending the memory. The LSTM uses three gates: input gate, output gate, and forget gate. With the help of these gates, LSTM assigns data weights. The series of gates in LSTM controls the information sequence data that enters, stores, and leaves the model network.

Gated Recurrent Unit GRU⁴⁵ is a type of RNN model with several advantages over LSTM. The GRU has less memory computation and is much faster than LSTM. Similar to LSTM, GRU utilizes gates to control input information. The GRU uses the update and reset gates to overcome the vanishing gradient problem.⁴⁶ The information passing to output is decided by these two gates.

Deep learning models are used with customized architecture and have a different number of layers, as well as, the number of neurons. Details regarding the architecture of deep learning models are given in [Table 4](#).

In a similar fashion, the performance of the machine learning models is optimized by fine-tuning several of the available hyperparameters. The hyperparameter tuning is applied to machine learning and deep learning techniques.⁴⁷ The iterative learning model training and testing process selects the best-fit hyperparameters. The best-fit hyperparameter analysis of employed learning techniques is

examined in Table 5. The hyperparameters achieve the best performance accuracy scores in predicting ventilator pressure.

Novel proposed approach

Our novel proposed H-VPP approach is based on a hybrid of DTR and RFR techniques. The architectural analysis of the H-VPP approach is analyzed in Figure 5. The whole ventilator dataset is fitted to both DTR and RFR techniques. The ventilator pressure is predicted from the DTR and RFR

Table 4. The layers stack architectural analysis of employed deep learning techniques.

Layer	Neurons unit	Output shape	Total parameter
RNN			
Recurrent neural networks	16	(None, 6, 16)	288
Dense	8	(None, 6, 8)	136
Output	1	(None, 6, 1)	9
LSTM			
Long short-term memory	16	(None, 6, 16)	1152
Dense	8	(None, 6, 8)	136
Output	1	(None, 6, 1)	9
GRU			
Gated recurrent units	16	(None, 6, 16)	912
Dense	8	(None, 6, 8)	136
Output	1	(None, 6, 1)	9

Table 5. The best-fit hyperparameters settings for employed learning techniques.

Technique	Hyper-parameters
MLP	Hidden_layer_sizes = 10, Max_iter = 10, activation = 'relu', solver = 'adam', alpha = 0.0001, learning_rate = 'constant', learning_rate_init = 0.001, tol = 1e-4, momentum = 0.9, epsilon = 1e-8, max_fun = 15000
LR	fit_intercept = True, normalize = False, copy_X = True, n_jobs = None, positive = False
DTR	criterion = "squared_error", splitter = "best", max_depth = 20, ccp_alpha = 0.0, min_samples_split = 2, min_samples_leaf = 1, max_features = None
RFR	n_estimators = 10, max_depth = 10, criterion = "squared_error", max_features = 1.0, bootstrap = True, ccp_alpha = 0.0, random_state = None
SGDR	max_iter = 10, loss = 'squared_error', penalty = 'l2', alpha = 0.0001, l1_ratio = 0.15, fit_intercept = True, tol = 1e-3, epsilon = 0.1, learning_rate = 'invscaling', eta0 = 0.01, power_t = 0.25, validation_fraction = 0.1
BR	n_iter = 10, tol = 1e-3, alpha_1 = 1e-6, alpha_2 = 1e-6, lambda_1 = 1e-6, lambda_2 = 1e-6, fit_intercept = True
LGBMR	n_estimators = 10, boosting_type = 'gbdt', num_leaves = 31, learning_rate0.1, n_estimators = 100, subsample_for_bin = 200000, min_child_weight = 1e-3, importance_type = 'split'
RNN	Optimizer = 'adam', Metrics = 'mse', Loss = 'mean_squared_error', Activation = 'linear'
LSTM	Loss = 'mean_squared_error', Optimizer = 'adam', Metrics = 'mse', Activation = 'linear'
GRU	Activation = 'linear', Optimizer = 'adam', Loss = 'mean_squared_error', Metrics = 'mse'

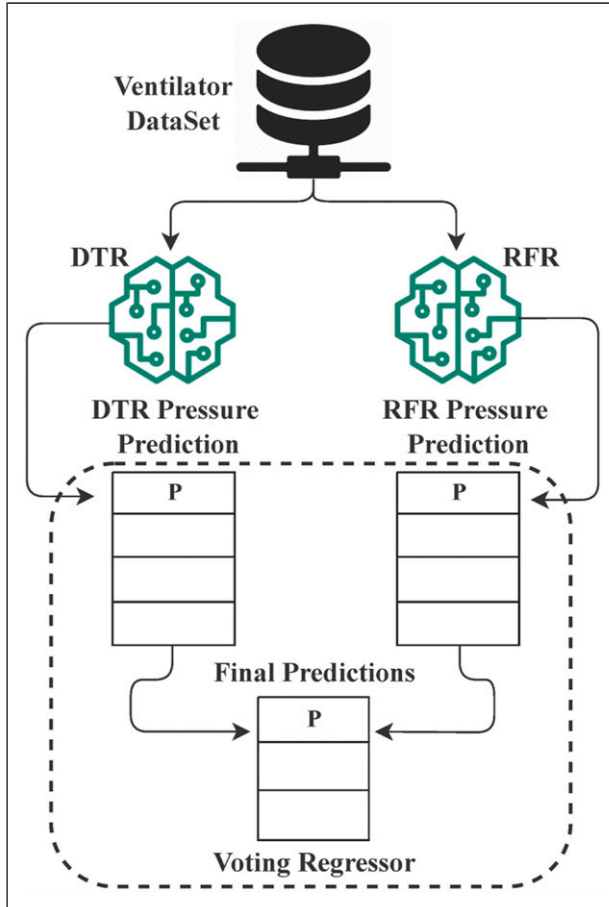


Figure 5. The architectural analysis of the proposed H-VPP approach.

techniques. The average from individual predictions is taken using a voting regressor to form a final prediction. Then a final prediction of ventilator pressure outcome with high accuracy.

The voting regressor⁴⁸ is an ensemble learning method designed for solving regression tasks. This method combines the predictions of multiple individual regressors to produce a more robust and accurate prediction. This ensemble technique falls under the category of model averaging, where diverse base regressors are trained independently, and their predictions are aggregated to form the final output. The ensemble's strength lies in its ability to reduce overfitting and enhance generalization performance by leveraging the collective wisdom of diverse models.

The `VotingRegressor` in scikit-learn can be represented mathematically as follows:

$$\hat{y}_{ensemble} = \frac{1}{N} \sum_{i=1}^N \hat{y}_{model_i} \quad (1)$$

where $\hat{y}_{ensemble}$ is the ensemble's predicted output, N is the number of individual regressors, and \hat{y}_{model_i} represents the prediction of the i -th regressor.

In the case of using `DecisionTreeRegressor` and `RandomForestRegressor` as base estimators, the ensemble prediction becomes:

$$\hat{y}_{ensemble} = \frac{1}{2}(\hat{y}_{DecisionTree} + \hat{y}_{RandomForest}) \quad (2)$$

where $\hat{y}_{DecisionTree}$ and $\hat{y}_{RandomForest}$ are the predictions of the `DecisionTreeRegressor` and `RandomForestRegressor`, respectively.

Results

Experimental results and evaluation are analyzed in this section. The comparative results of employed machine learning and deep learning techniques are validated using the performance metrics⁴⁹ MAE, MAD, MSE, RMSE, R^2 , Adjusted R^2 , and variance. These regression metrics are analyzed through mathematical notations and computation scores for employed learning techniques.

Experimental setup

Python 3.0 programming language is used for data analysis, model building, and evaluations. The Scikit-learn library module with a version of 1.0.2 is used for machine learning model building and testing. The TensorFlow library module with a version of 2.8.2 and the Keras library module with a version of 2.8.0 are utilized for building deep learning models and testing. The experiments are completed on the platform with a model Intel(R) Xeon(R), 2.20 GHz CPU, 13 GB RAM, cache size of 56,320 KB, and 12 GB NVIDIA Tesla K80 GPU.

Scientific evaluations

The scientific performance evaluations are examined using the regression metrics. This study employs seven evaluation metrics for this purpose.

MAE is the L1 loss function most commonly used for regression problems. MAE measure is the magnitude of difference between the prediction and the actual value of the dataset. It takes the average of absolute errors. It can be calculated using

$$MAE = \frac{|(y_i - y_p)|}{n} \quad (3)$$

where y_i is the actual value, y_p is the predicted value, and n is the total number of samples in the dataset.

MAD measures the model's median deviation between the actual and predicted values. MAD is best for outliers in the target variable, which is the significant reason to use MAD in combination with MAE. MAD is calculated by

$$MAD = \frac{\sum_{i=1}^n |(y_i - \hat{x}_i)|}{n} \quad (4)$$

where the \hat{X}_i is the median of the data sample, y_i is the prediction, and n is the total sample values.

MSE is determined by calculating the square average of the difference between the predicted and actual values. MSE is very sensitive to outliers in the data. The mathematical notation to calculate the MSE is

$$MSE = \frac{\sum_{i=1}^n |(y_i - \hat{x}_i)|^2}{n} \quad (5)$$

where n is the total number of samples while \hat{X}_i and y_i represent the median and prediction, respectively.

RMSE metric is the same as the MSE; however, the root of values is taken while calculating the MSE. RMSE is robust to the outliers. It is the best measure that shows how accurate the proposed model is with respect to prediction. RMSE is calculated using

$$RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^n |(y_i - \hat{x}_i)|^2} \quad (6)$$

R^2 accuracy score is a statistical performance metric utilized to determine how well an employed regression technique is in prediction for unseen data samples. R^2 accuracy score is also referred to as the coefficient of determination. It explains the proportion of variance for a target (dependent variable) by a feature (independent variable). R^2 score always lies between zero and one. R^2 is calculated using

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \quad (7)$$

Adjusted R^2 is another version of R^2 which determines the variation in the dependent variable and only the explained features have higher effects in making predictions. The adjusted R^2 is calculated using the following

$$Adjusted (R^2) = 1 - \left((1 - R^2) \frac{N - 1}{N - M - 1} \right) \quad (8)$$

where N and M are the total number of rows and columns, respectively.

The explained variance score is relatively similar to the R^2 score. The variance score explains the dispersion of errors in an input dataset. The variability of the prediction's proportions is measured by variance for a learning model. The variance score is based on the difference between the expected and the predicted values. It is calculated using

$$Variance\ score = 1 - \frac{Var(y - \hat{y})}{Var(y)} \quad (9)$$

where $Var(y - \hat{y})$ is the variance of prediction errors and $Var(y)$ is the variance of actual values.

Results of deep learning models

The training and validation performance with each epoch in employed deep learning techniques is examined in Table 6. The analysis demonstrates that the RNN model has high training loss at the first epoch and MSE values of 0.0112. The validation loss and validation MSE are also higher with a

Table 6. The training and validation performance by epochs for deep learning models.

Epoch	Training time (sec)	Training loss	MSE	Validation loss	Validation MSE
RNN					
1	365s 3 ms/step	0.0112	0.0112	0.0104	0.0104
2	335s 3 ms/step	0.0103	0.0103	0.0102	0.0102
3	331s 3 ms/step	0.0102	0.0102	0.0101	0.0101
4	333s 3 ms/step	0.0101	0.0101	0.0101	0.0101
5	329s 3 ms/step	0.0101	0.0101	0.0101	0.0101
LSTM					
1	505s 4 ms/step	0.0108	0.0108	0.0102	0.0102
2	457s 4 ms/step	0.0101	0.0101	0.0101	0.0101
3	488s 4 ms/step	0.0101	0.0101	0.0100	0.0100
4	458s 4 ms/step	0.0100	0.0100	0.0100	0.0100
5	454s 4 ms/step	0.0099	0.0099	0.0100	0.0100
GRU					
1	505s 4 ms/step	0.0107	0.0107	0.0102	0.0102
2	511s 4 ms/step	0.0101	0.0101	0.0100	0.0100
3	455s 4 ms/step	0.0099	0.0099	0.0099	0.0099
4	455s 4 ms/step	0.0098	0.0098	0.0098	0.0098
5	456s 4 ms/step	0.0098	0.0098	0.0098	0.0098

score of 0.0104. In the second epoch, the training and validation loss is decreased, which shows that the model is now fitting on data well by reducing the error. At the last epoch during training, the RNN model has the 0.0101 score value for training loss, MSE, validation loss, and validation MSE. Results show that RNN has a high error score during data training.

By analyzing the LSTM model training, during the first epoch of the model, the training loss and MSE have a value of 0.0108, which is less than the RNN model's first epoch. The validation loss and validation MSE have scores of 0.0102. The time computations for the LSTM model are high during training. During the LSTM model training, the training and validation errors are decreased as the number of epochs increases. In the last epoch, the LSTM model has a training loss and MSE value of 0.0099, much less than the first epoch. The validation loss and validation MSE have scores of 0.0100 in the last epoch. The analysis demonstrates that the LSTM model has low error scores during training.

The GRU model has almost similar error rates as compared to LSTM. During the first epoch of the GRU model, the training loss and MSE have a value of 0.0107, which is less than the LSTM model's first epoch. The validation loss and validation MSE also have scores of 0.0102, the same as the LSTM model. The GRU model has fewer time computations during training as compared to the LSTM model. The error rates decrease as the GRU model epochs are increased. In the last epoch, the GRU training loss and MSE have a value of 0.0098, much less than the first epoch. The validation loss and validation MSE also have scores of 0.0098 in the last epoch. The analysis demonstrates that the GRU model has lower error scores during training than the LSTM model.

Prediction results of models

The comparative performance analysis of employed machine learning and deep learning techniques on unseen test data is given in [Table 7](#). The analysis demonstrates that the novel proposed approach

Table 7. Comparative performance of employed machine learning and deep learning models.

Technique	MAE	MAD	MSE	RMSE	R^2 score	Adjusted R^2 score	Variance
MLP	0.041	0.014	0.005	0.072	0.64	0.64	0.640
LR	0.059	0.009	0.023	0.095	0.38	0.38	0.383
DTR	0.029	0.007	0.003	0.059	0.76	0.76	0.760
RFR	0.031	0.008	0.003	0.061	0.74	0.74	0.747
SGDR	0.059	0.023	0.009	0.095	0.38	0.38	0.383
BR	0.059	0.023	0.009	0.095	0.38	0.38	0.383
LGBMR	0.049	0.029	0.005	0.076	0.60	0.60	0.608
RNN	0.082	0.077	0.008	0.094	0.39	0.39	0.603
LSTM	0.078	0.072	0.008	0.091	0.43	0.43	0.629
GRU	0.079	0.074	0.008	0.091	0.43	0.43	0.648
Proposed	0.028	0.007	0.003	0.056	0.78	0.78	0.782

outperforms all the used machine learning and deep learning models. By analyzing the MAE, the lowest score value is 0.028, which is achieved by the proposed approach. The highest MAE score of 0.082 is by the RNN model. The MAD analysis shows that the lowest score value of 0.007 is commonly achieved by the proposed approach and DTR model while the highest MAD of 0.077 is by the RNN model. Similarly, the best value for MSE, that is, 0.003 is obtained by three models including DTR, RFR, and the proposed approach. On the other hand, the highest MSE of 0.023 is achieved by the LR technique.

For the RMSE, 0.056 is the lowest score which is obtained by the proposed hybrid approach while the highest RMSE of 0.095 is achieved by LR, BR, and SGDR techniques. Scores of MAE, MAD, MSE, and RMSE show the error between the predicted and actual values, and higher values show the inability of the model to predict a correct output. Obtained scores from the proposed approach indicate that the proposed hybrid model performs much better than other employed approaches. R^2 and adjusted R^2 accuracy scores have the same values in this analysis. The highest R^2 score of 0.78 is achieved by the proposed approach whereas the minimum R^2 score of the study is 0.38 which is obtained by LR, BR, and SGDR techniques. R^2 values for the proposed model suggest that it outperforms other models regarding the R^2 score. This analysis demonstrates that the novel proposed approach has less error rate than other approaches, leading to high accuracy for ventilator pressure prediction.

The bar chart shows the comparative analysis of the R^2 accuracy score for all employed machine learning and deep learning techniques, as given in Figure 6. The analysis demonstrates that a high accuracy score of 78% is achieved by the proposed approach in comparison with other techniques. The DTR and RFR techniques also achieved good accuracy scores. The lowest R^2 score of 38% is achieved by LR, SGDR, and BR models.

The regression analysis is demonstrated based on the mapping of actual ventilator pressure values with the predicted pressure values. The 10,000 data points are taken under consideration for regression analysis. In Figure 7(a), the DTR technique is analyzed for ventilator pressure prediction. The DTR prediction regression analysis shows that the model poorly predicts the pressure values from 0.5 or above. In Figure 7(b), the RFR model for ventilator pressure prediction is analyzed. The RFR prediction regression analysis shows that the model poorly predicts the pressure values from 0.4 or above. The RFR model has a higher error rate than the DTR model. In Figure 7(d), the proposed technique is examined for ventilator pressure prediction. The proposed H-VPP prediction

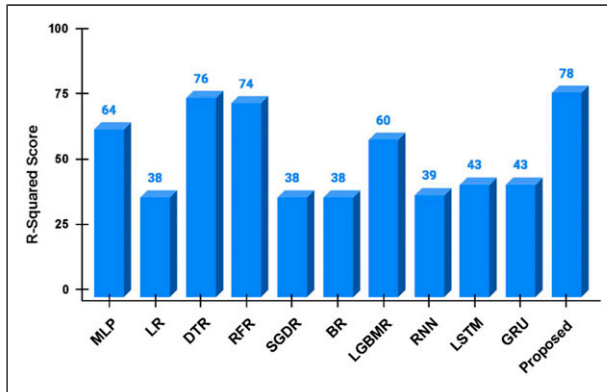


Figure 6. The R^2 score performance analysis for employed models.

regression analysis demonstrates that the pressure values are almost correctly predicted and have high accuracy compared to DTR and RFR. Only a few data points have an error in predicting the pressure values of 0.6 or above predicted by the proposed model.

Performance comparison with existing approaches

The comparative performance analysis of the proposed approach with other state-of-the-art studies is conducted in Table 8. For this purpose, the models from the selected studies are built and used with the dataset used in this study for a fair comparison. The state-of-the-art models XGBoost, light GBM, RNN, LR, DT, and CNN are applied for comparison. The comparative performance evaluations are based on the MAE, MSE, R^2 score, and RMSE metrics. The analysis demonstrates that the proposed approach outperforms with state of the art models with a 0.78 R^2 score. Similarly, the error metrics show that the proposed approach has lower error scores as compared to existing models.

Discussions

Mechanical ventilator plays a vital role in saving millions of lives. Patients with COVID-19 symptoms need a ventilator to survive during the pandemic. The pumping of air into the patient's lungs using a ventilator requires a particular air pressure. High or low ventilator pressure can result in a patient's life loss as high air pressure in the ventilator causes the patient lung damage while lower pressure provides insufficient oxygen. Consequently, precisely predicting ventilator pressure is a task of great significance in this regard. The analysis of previous studies indicates that machine learning approaches are predominantly used for ventilator pressure prediction, and results suggest that the produced accuracy is still lower than the desired output, and further research efforts are needed in this regard.

This study presents a novel H-VPP approach for precise and accurate ventilator pressure prediction. The proposed H-VPP approach is based on the hybrid of DTR and RFR, where the predictions from these models are regressed to predict the final output. Extensive experiments are performed involving seven machine learning and three deep learning models to investigate the performance of the proposed approach regarding MAE, MAD, MSE, RMSE, R^2 score, adjusted R^2

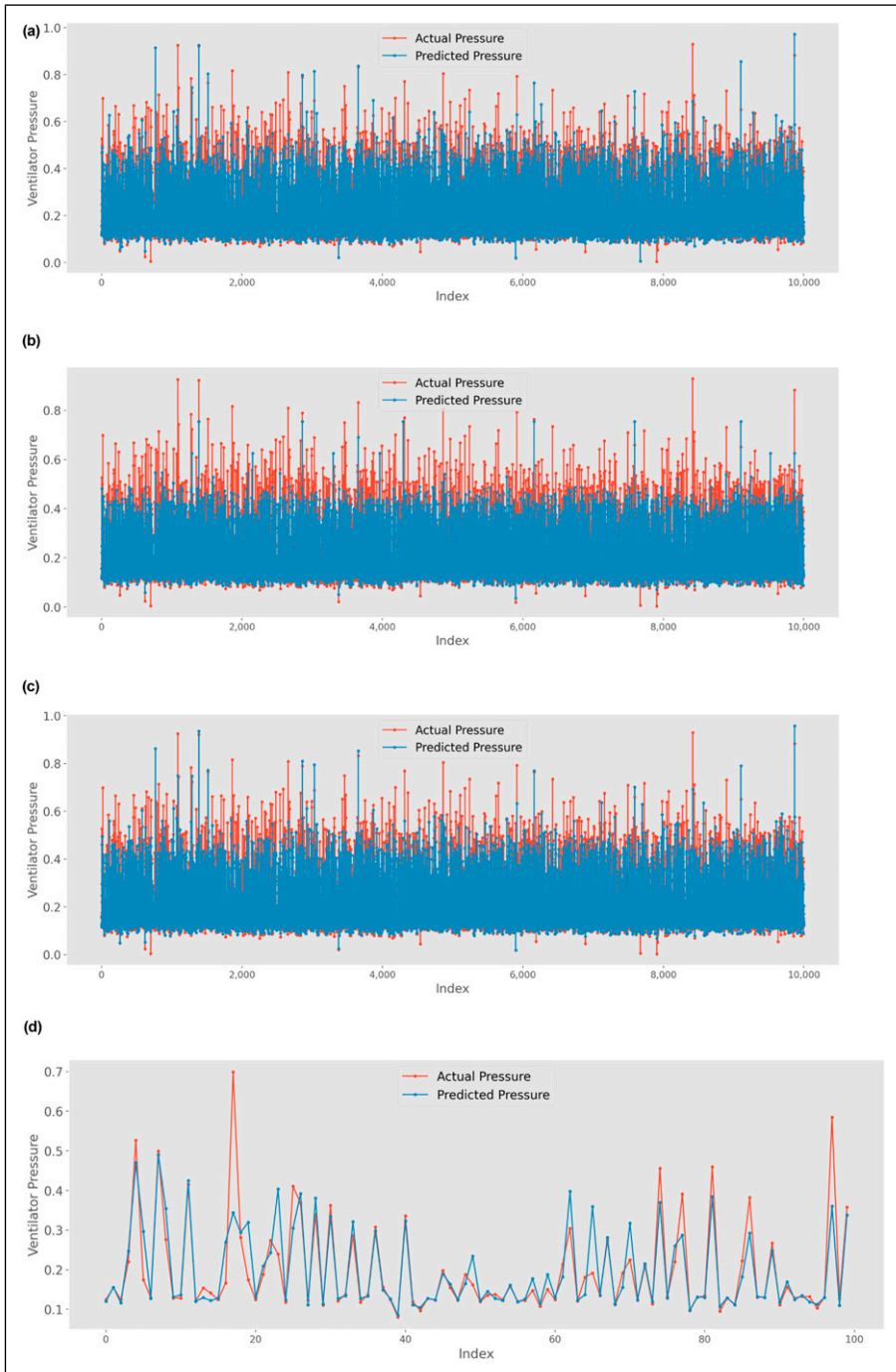


Figure 7. Regression analysis of, (a) DTR technique for ventilator pressure prediction, (b) RFR approach for ventilator pressure prediction (c) Proposed H-VPP approach for ventilator pressure prediction, (d) 100 sample points of Proposed H-VPP approach for ventilator pressure prediction.

Table 8. Performanc comparison with state-of-the-art approaches.

Ref.	Year	Learning type	Technique	MAE	MSE	R^2 score	RMSE
50	2021	Machine learning	XGBoost	0.127	0.018	0.24	0.135
51	2021	Machine learning	LightGBM	0.049	0.005	0.60	0.076
52	2021	Deep learning	Recurrent neural network	0.082	0.008	0.39	0.094
21	2022	Machine learning	XGBoost	0.127	0.018	0.24	0.135
53	2021	Machine learning	XGBoost	0.127	0.018	0.24	0.135
54	2022	Machine learning	Linear regression	0.059	0.009	0.38	0.095
55	2021	Machine learning	Decision tree	0.029	0.003	0.76	0.059
56	2021	Deep learning	Convolutional neural networks	0.059	0.009	0.38	0.095
Proposed	2022	Machine learning	Novel H-VPP	0.028	0.003	0.78	0.056

score, and variance. Results suggest that the proposed approach outperforms all the employed models with a 0.78 R^2 score. The applied VEDA reveals that the prominent cause of high ventilator pressure is the high values of lung attributes R and C during initial time step values.

Limitations

In this research study, we have proposed a novel H-VPP approach for precise and accurate ventilator pressure prediction. However, the proposed research has some limitations. The R^2 score of our approach could be enhanced by minimizing the Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) rates. Additionally, the applied deep learning models are computationally expensive and can be optimized by reducing their layered architectures. Regarding study limitations and future work, a transfer learning-based model will be developed for ventilator pressure predictions.

Conclusion

Mechanical ventilators have become important considering the recent surge in the COVID-19 pandemic and play an integral part in saving countless lives. However, high or low ventilation pressure can cause lung damage to patients, and pressure prediction is important to avoid such complications. This study presents a novel H-VPP approach for precise and accurate ventilator pressure prediction. The proposed H-VPP approach is based on the hybrid of DTR and RFR where the predictions from these models are regressed to predict the final output. Extensive experiments are performed involving seven machine learning and three deep learning models to investigate the performance of the proposed approach regarding MAE, MAD, MSE, RMSE, R^2 score, adjusted R^2 score, and variance. Results suggest that the proposed approach outperforms all the employed models with a 0.78 R^2 score. Error metrics also show the superior performance of the proposed approach. Performance comparison with existing state-of-the-art corroborates the superior performance of the proposed model. The applied VEDA reveals that the prominent cause of high ventilator pressure is the high values of lung attributes R and C during initial time step values.

Author contributions

Conceptualization, Ali Raza and Furqan Rustam; Data curation, Furqan Rustam and Hafeez Ur Rehman Siddiqui; Formal analysis, Ali Raza, Hafeez Ur Rehman Siddiqui and Maria Asuncio'n Vicente Ripoll;

Funding acquisition, Emmanuel Soriano Flores; Investigation, Juan Luis Vidal Mazon and Isabel de la Torre Diez; Methodology, Hafeez Ur Rehman Siddiqui and Emmanuel Soriano Flores; Project administration, Maria Asuncion and Vicente Ripoll; Resources, Emmanuel Soriano Flores; Software, Emmanuel Soriano Flores and Juan Luis Vidal Mazon; Supervision, Imran Ashraf; Validation, Isabel de la Torre Diez, Maria Asuncion Vicente Ripoll and Imran. Ashraf; Visualization, Juan Luis Vidal Mazon and Isabel de la Torre Diez; Writing – original draft, Ali Raza and Furqan Rustam; Writing – review & editing, Imran Ashraf.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study is funded by the European University of Atlantic.

ORCID iD

Imran Ashraf  <https://orcid.org/0000-0002-8271-6496>

Data availability statement

The data used in this study can be requested from the corresponding authors.

References

1. Zuckerberg J, Shaik M, Widmeier K, et al. A lung for all: novel mechanical ventilator for emergency and low-resource settings. *Life Sci* 2020; 257: 118113.
2. Tobin MJ, Laghi F and Jubran A. Caution about early intubation and mechanical ventilation in COVID-19. *Ann Intensive Care* 2020; 10(1): 78.
3. Mehedi IM, Shah HS, Al-Saggaf UM, et al. Adaptive fuzzy sliding mode control of a pressure-controlled artificial ventilator. *J Healthc Eng* 2021; 2021: 1926711.
4. Oliveira E, Parikh A, Lopez-Ruiz A, et al. ICU outcomes and survival in patients with severe COVID-19 in the largest health care system in central Florida. *PLoS One* 2021; 16(3): e0249038.
5. Yang X, Yu Y, Xu J, et al. Clinical course and outcomes of critically ill patients with sars-cov-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 2020; 8(5): 475–481.
6. Wang Y, Lu X, Li Y, et al. Clinical course and outcomes of 344 intensive care patients with COVID-19. *Am J Respir Crit Care Med* 2020; 201(11): 1430–1434.
7. Richardson S, Hirsch JS, Narasimhan M, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York city area. *Jama* 2020; 323(20): 2052–2059.
8. ICNARC I. Icnarc report on covid-19 in critical care. *Icnarc COVID-19 study case mix program database*. London: Intensive Care National Audit & Research Centre, 2020.
9. Barrasa H, Rello J, Tejada S, et al. Sars-cov-2 in Spanish intensive care units: early experience with 15-day survival in Vitoria. *Anaesth Crit Care Pain Med* 2020; 39(5): 553–561.
10. Le Grande C and Dow A. *Best doctors and nursesTM: Australia leads world in COVID-19 ICU survival rates*. Melbourne, Australia: The Age. Retrieved from: <https://www.theage.com.au/national/best-doctors-and-nurses-australia-leads-world-in-covid-19-icu-survival-rates-20200714-p555byq.html>.2020

11. COVID F. *Deaths in ICU suggest hospital care improving: study*. India: NDTV Convergence Limited, 2020, p. 19.
12. Rahim F, Amin S, Noor M, et al. Mortality of patients with severe COVID-19 in the intensive care unit: an observational study from a major COVID-19 receiving hospital. *Cureus* 2020; 12(10): e10906.
13. Armstrong R, Kane A and Cook T. Outcomes from intensive care in patients with COVID-19: a systematic review and meta-analysis of observational studies. *Anaesthesia* 2020; 75(10): 1340–1349.
14. Masud M, Sikder N, Nahid AA, et al. A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework. *Sensors* 2021; 21(3): 748.
15. Gupta R, Srivastava D, Sahu M, et al. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol Divers* 2021; 25(3): 1315–1360.
16. Park JE, Kim TY, Jung YJ, et al. Biosignal-based digital biomarkers for prediction of ventilator weaning success. *Int J Environ Res Publ Health* 2021; 18(17): 9229.
17. Egbert J and Plonsky L. Bootstrapping techniques. *A Practical Handbook of Corpus Linguistics*. New York: Springer, 2020, pp. 593–610.
18. Radhakrishnan S, Nair SG and Isaac J. Multilayer perceptron neural network model development for mechanical ventilator parameters prediction by real time system learning. *Biomed Signal Process Control* 2022; 71: 103170.
19. Yamashita K. Pulse-wave transit time with ventilator-induced variation for the prediction of fluid responsiveness. *Acute Med Surg* 2020; 7(1): e484.
20. Wu Z, Liu Y, Xu J, et al. A ventilator-associated pneumonia prediction model in patients with acute respiratory distress syndrome. *Clin Infect Dis* 2020; 71(Supplement_4): S400–S408.
21. Cheng KH, Tan MC, Chang YJ, et al. The feasibility of a machine learning approach in predicting successful ventilator mode shifting for adult patients in the medical intensive care unit. *Medicina* 2022; 58(3): 360.
22. Loo N, Chiew YS, Tan CP, et al. A machine learning approach to assess magnitude of asynchrony breathing. *Biomed Signal Process Control* 2021; 66: 102505.
23. Zhang K, Karanth S, Patel B, et al. A multi-task Gaussian process self-attention neural network for real-time prediction of the need for mechanical ventilators in COVID-19 patients. *J Biomed Inf* 2022; 130: 104079.
24. Brain G. *Google brain - ventilator pressure prediction*. California: Kaggle, 2021. <https://www.kaggle.com/competitions/ventilator-pressure-prediction/data>.
25. The People’s Ventilator Project. *A fully-open supply-chain resilient pressure-control ventilator for the people*. California, USA: Plos One, 2020. <https://www.peoplesvent.org/en/latest/>.
26. IngMar Medical. *QuickLung products*. USA: Ingmar Medical, 2020. <https://www.ingarmed.com/product/quicklung/>.
27. Dwivedi YK, Hughes L, Ismagilova E, et al. Artificial intelligence (AI): multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *Int J Inf Manag* 2021; 57: 101994.
28. Helmy M, Smith D and Selvarajoo K. Systems biology approaches integrated with artificial intelligence for optimized metabolic engineering. *Metab Eng Commun* 2020; 11: e00149.
29. Shastry KA and Sanjay H. Machine learning for bioinformatics. *Statistical modelling and machine learning principles for bioinformatics techniques, tools, and applications*. New York: Springer, 2020, pp. 25–39.
30. Rupapara V, Rustam F, Aljedaani W, et al. Blood cancer prediction using leukemia microarray gene data and hybrid logistic vector trees model. *Sci Rep* 2022; 12(1): 1000.
31. Rustam F, Reshi AA, Ashraf I, et al. Sensor-based human activity recognition using deep stacked multilayered perceptron model. *IEEE Access* 2020; 8: 218898–218910.

32. Imran A, Alnumay WS, Rashid A, et al. Prediction models for covid-19 integrating age groups, gender, and underlying conditions. *Comput Mater Continua (CMC)* 2021; 67(3): 3009–3044.
33. Roca P, Attye A, Colas L, et al. Artificial intelligence to predict clinical disability in patients with multiple sclerosis using flair mri. *Diagn Interv Imaging* 2020; 101(12): 795–802.
34. Rustam F, Siddique MA, Siddiqui HUR, et al. Wireless capsule endoscopy bleeding images classification using cnn based model. *IEEE Access* 2021; 9: 33675–33688.
35. Yulita I, Abdullah A, Helen A, et al. Comparison multi-layer perceptron and linear regression for time series prediction of novel coronavirus COVID-19 data in west java. In: *Journal of Physics: Conference Series*. IOP Publishing, 2021, vol 1722, p. 012021.
36. Singh A, Kushwaha S, Alarfaj M, et al. Comprehensive overview of backpropagation algorithm for digital image denoising. *Electronics* 2022; 11(10): 1590.
37. Rath S, Tripathy A and Tripathy AR. Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model. *Diabetes Metabol Syndr* 2020; 14(5): 1467–1474.
38. Taufiqurrahman A, Putrada AG and Dawani F. Decision tree regression with adaboost ensemble learning for water temperature forecasting in aquaponic ecosystem. In: *2020 6th International Conference on Interactive Digital Media (ICIDM)*. IEEE, 2020, pp. 1–5.
39. Yang S, Zhang L, Fan J, et al. Experimental study on erosion behavior of fracturing pipeline involving tensile stress and erosion prediction using random forest regression. *J Nat Gas Sci Eng* 2021; 87: 103760.
40. Soni K and Singh S. Bitcoin price prediction-an analysis of various regression methods. In: *2022 IEEE 12th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*. IEEE, 2022, pp. 271–276.
41. Saqib M. Forecasting COVID-19 outbreak progression using hybrid polynomial-bayesian ridge regression model. *Appl Intell* 2021; 51(5): 2703–2713.
42. Gong M, Bai Y, Qin J, et al. Gradient boosting machine for predicting return temperature of district heating system: a case study for residential buildings in tianjin. *J Build Eng* 2020; 27: 100950.
43. Bai Y, Xie J, Liu C, et al. Regression modeling for enterprise electricity consumption: a comparison of recurrent neural network and its variants. *Int J Electr Power Energy Syst* 2021; 126: 106612.
44. Ji Y, Liew AWC and Yang L. A novel improved particle swarm optimization with long-short term memory hybrid model for stock indices forecast. *IEEE Access* 2021; 9: 23660–23671.
45. Li W, Wu H, Zhu N, et al. Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit (gru). *Information Processing in Agriculture* 2021; 8(1): 185–193.
46. Hu Z, Zhang J and Ge Y. Handling vanishing gradient problem using artificial derivative. *IEEE Access* 2021; 9: 22371–22377.
47. Cooney C, Korik A, Folli R, et al. Evaluation of hyperparameter optimization in machine and deep learning methods for decoding imagined speech eeg. *Sensors* 2020; 20(16): 4629.
48. Yulisa A, Park SH, Choi S, et al. Enhancement of voting regressor algorithm on predicting total ammonia nitrogen concentration in fish waste anaerobiosis. *Waste Biomass Valorization* 2023; 14(2): 461–478.
49. Chicco D, Warrens MJ and Jurman G. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *PeerJ Comput Sci* 2021; 7: e623.
50. Yu L, Halalau A, Dalal B, et al. Machine learning methods to predict mechanical ventilation and mortality in patients with COVID-19. *PLoS One* 2021; 16(4): e0249285.
51. Otaguro T, Tanaka H, Igarashi Y, et al. Machine learning for the prediction of successful extubation among patients with mechanical ventilation in the intensive care unit: a retrospective observational study. *J Nippon Med Sch* 2021; 88(5): 408–417. JNMS–2021_88.
52. Mamandipoor B, Frutos-Vivar F, Peñuelas O, et al. Machine learning predicts mortality based on analysis of ventilation parameters of critically ill patients: multi-centre validation. *BMC Med Inf Decis Making* 2021; 21(1): 152.

53. Lin MY, Li CC, Lin PH, et al. Explainable machine learning to predict successful weaning among patients requiring prolonged mechanical ventilation: a retrospective cohort study in central Taiwan. *Front Med* 2021; 8: 663739.
54. Chen WT, Huang HL, Ko PS, et al. A simple algorithm using ventilator parameters to predict successfully rapid weaning program in cardiac intensive care unit patients. *J Personalized Med* 2022; 12(3): 501.
55. Abujaber A, Fadlalla A, Gammoh D, et al. Machine learning model to predict ventilator associated pneumonia in patients with traumatic brain injury: the c. 5 decision tree approach. *Brain Inj* 2021; 35(9): 1095–1102.
56. Jia Y, Kaul C, Lawton T, et al. Prediction of weaning from mechanical ventilation using convolutional neural networks. *Artif Intell Med* 2021; 117: 102087.