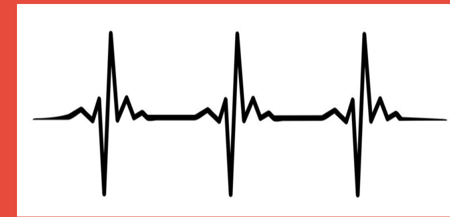
CS 668 PROJECT
Ventilator Pressure Detection

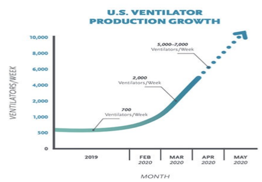
Sunanda Singareddy[1], Dr.Christelle Scharff[2]

Pace University New York – Data Science



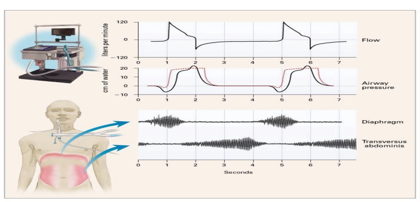
Abstract

The COVID-19 pandemic statistic count is 146,122 cases per million people and there is an incremental trend across the globe.
This project allows me to address this situation and automatically predict the right level flow of air pressure based on the actual pressure values.
As a result, predicting the pressure on ventilation before hand will help in increasing survival rates. The early estimate on **recovery rate is 97 to 99.5%**

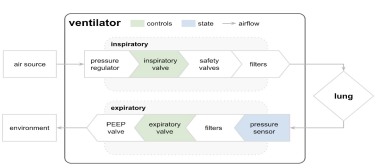


Introduction

Covid-19 is the respiratory disease which have an impact on respiratory tract including lungs.
Lungs and airways swell and become inflamed, and this infection starts in one part of lung and spreads to the other.
As the body starts to fight it, the lungs become more inflated and fill with fluid which can make it harder for patients to swap oxygen and carbon dioxide



A ventilator mechanically helps pump oxygen into body and can be set to take a certain number of breaths per minute.
It can also hold positive end-expiratory pressure (PEEP) - constant amount of low pressure to keep air sacs in lung from collapsing.
The pressure plays a vital role, and it is dependent on features such as R (resistance through with air is passed), C (compliance of volume per pressure).
Prediction of pressure must not be beyond the PIP (target pressure) and must not be below PEEP (positive end-expiratory pressure) value.



Data Preparation

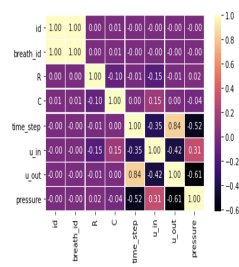
Analyzing the Dataset

Check in case there are any missing values or any null values

```
[7]:  Train_set.isna().any()

[7]:  id            False
      breath_id     False
      R             False
      C             False
      time_step     False
      u_in          False
      u_out         False
      pressure      False
      dtype: bool
```

Check the correlation in features amongst itself for feature selection



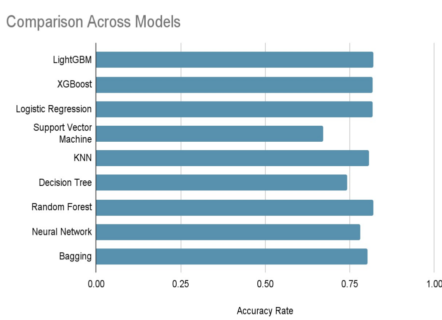
Methodology

Pressure is a continuous variable and hence for prediction **regression** is used instead of classification.

Also, **ensemble model** is better in terms of prediction as it combines the prediction from multiple models.

The bar chart is a consolidated report from the results obtained from literature review.

LightGBM and XGBoost provide the best accuracy rate as compared to other models



Experimentation

Linear Regression:
The model score given by Linear Regression is 0.38 which is not a best model for prediction.
Random Forest:
The model score given by RF is 0.74 which is still less to say a best predictor but better than Linear Regression.
XGBoost:
The model score given by XGBoost is 0.74 but the model performance if GPU is not used then it is poor. (Note: I was not able to improve as I was not able to use system GPU)
LightGBM:
This model gives the better accuracy than XGBoost and the performance is better so further analysis and predictions is done using LightGBM Ensemble Model in this project.

Model	Parameters	Train Accuracy	Test Accuracy
Linear Regression	Random state = 42	0.3841	0.3834
Random Forest	Random state = 42, Sample = 11000	0.7411	0.7389
XGBoost	Random state = 42, learning_rate = 0.9	0.7450	0.7415
LightGBM	Random state = 42, learning_rate = 0.9	0.7476	0.7472
LightGBM with Added features (Categorical Values)	**Random state = 42, learning_rate = 0.25 Feature Engineering of time series**	**0.9858**	**0.9851**

Results

LightGBM with features added:

Additional features are added using feature engineering and also converted R (5, 20, 50) and C (5, 20, 50) to numerical values. This improvised the performance, and the test and train score are nearly equivalent to 98%

```
[39]: model = lgb.LGBMRegressor(learning_rate=0.35,max_depth=27,random_state=42, num_leaves = 106)
      model.fit(x_train,y_train,eval_set=[(x_test,y_test),(x_train,y_train)], verbose=30)

      [30]    training's l2: 1.34234  valid_0's l2: 1.368
      [60]    training's l2: 1.08865  valid_0's l2: 1.12528
      [90]    training's l2: 0.970549 valid_0's l2: 1.01532

[39]: LGBMRegressor(learning_rate=0.35, max_depth=27, num_leaves=106, random_state=42)

[40]: print('Training accuracy {:.4f}'.format(model.score(x_train,y_train)))
      print('Testing accuracy {:.4f}'.format(model.score(x_test,y_test)))
      #Training accuracy 0.9655 5 depth and 0.09 and 0.33 test size
      #Testing accuracy 0.9653
      #Training accuracy 0.9711 0.10 and 7 depth and 0.35
      #Testing accuracy 0.9709
      #Training accuracy 0.9741 0.15 and 7
      #Testing accuracy 0.9739
      #Training accuracy 0.9748 11 depth 0.15
      #Testing accuracy 0.9746
      #Training accuracy 0.9793 num leaves 56
      #Testing accuracy 0.9790
      #Training accuracy 0.9840 rate 0.25 depth 13 leaves 86
      #Testing accuracy 0.9835

      Training accuracy 0.9858
      Testing accuracy 0.9851
```

Conclusion

Predicting the pressure on ventilation before hand will help in increasing survival rates. The early estimate on recovery rate is 97 to 99.5% []
To avoid hyperventilation state of patient the prediction of pressure and the flow rate must be accurate.
LightGBM is the best model in regression as compared to other models for predicting continuous variables.

Future Work

Enhance LightGBM features to improvise prediction. (My current score: 0.1502)
Use gridwise search parameters for boosting the accuracy percentage and optimize the model.
The team who won the competition with 0.0575 score, have used LSTM (Neural Network Model)

References

Yu L, Halalau A, Dalal B, Abbas AE, Ivascu F, Amin M, et al. (2021) Machine learning methods to predict mechanical ventilation and mortality in patients with COVID-19. PLoS ONE 16(4): e0249285. https://doi.org/10.1371/journal.pone.0249285

Ghazal, Sam & Sauthier, Michaël & Brossier, David & Bouachir, Wassim & Jouvet, Philippe & Noumeir, Rita. (2019). Using machine learning models to predict oxygen saturation following ventilator support adjustment in critically ill children: A single center pilot study. PLOS ONE. 14. e0198921. 10.1371/journal.pone.0198921. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0198921

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 Inform Decis Mak 21, 152 (2021). https://doi.org/10.1186/s12911-021-01506-w

T. Chen et al., "Prediction of Extubation Failure for Intensive Care Unit Patients Using Light Gradient Boosting Machine," in IEEE Access, vol. 7, pp. 150960-150968, 2019, doi: 10.1109/ACCESS.2019.2946980. https://www.researchgate.net/publication/336453098_Prediction_of_Extubation_Failure_for_Intensive_Care_Unit_Patients_Using_Light_Gradient_Boosting_Machine

Sayed, M.; Riaño, D.; Villar, J. Predicting Duration of Mechanical Ventilation in Acute Respiratory Distress Syndrome Using Supervised Machine Learning. J. Clin. Med. 2021, 10, 3824. https://doi.org/10.3390/jcm10173824

Zhu Y, Zhang J, Wang G, Yao R, Ren C, Chen G, Jin X, Guo J, Liu S, Zheng H, Chen Y, Guo Q, Li L, Du B, Xi X, Li W, Huang H, Li Y and Yu Q (2021) Machine Learning Prediction Models for Mechanically Ventilated Patients: Analyses of the MIMIC-III Database. Front. Med. 8:662340. doi: 10.3389/fmed.2021.662340 https://doi.org/10.3389/fmed.2021.662340

Zhang, Zhongheng MD1; Liu, Jingtao MD2; Xi, Jingjing MD3; Gong, Yichun MD2; Zeng, Lin PhD4; Ma, Penglin MD2 Derivation and Validation of an Ensemble Model for the Prediction of Agitation in Mechanically Ventilated Patients Maintained Under Light Sedation, Critical Care Medicine: March 2021 - Volume 49 - Issue 3 - p e279-e290 doi: 10.1097/CCM.0000000000004821 https://pubmed.ncbi.nlm.nih.gov/33470778/

Zhang, Z, Navarese, EP, Zheng, B, et al. Analytics with artificial intelligence to advance the treatment of acute respiratory distress syndrome. J Evid Based Med. 2020; 13: 301– 312. https://doi.org/10.1111/jebm.12418

Raita, Y., Camargo, C.A., Macias, C.G. et al. Machine learning-based prediction of acute severity in infants hospitalized for bronchiolitis: a multicenter prospective study. Sci Rep 10, 10979 (2020). https://doi.org/10.1038/s41598-020-67629-8

Feedback from Team

Your justifications for each model you're using are very strong - Brian
Your presentations and explanation of the topic was really good - Prachi
Presentation was well prepared and presented. I like the motivation you gave first which gave a strong point on you doing the project. - Shefali