## **Predicting Mortality for COPD Patients Within 28 Days in ICU with XGBoost**

## Abstract

**Purpose:** This study is aimed at developing a machine learning model to predict the outcome (mortality) of patients in ICU within 28 days of admission and to compare its performance with that of three other machine learning models. On top of that we study the association of COPD with other comorbidities and impact they have on mortality.

**Methods:** Using the MIMICIII database we extracted data for all patients in ICU with COPD as primary reason for admission and then we compare the mortality impact between the patients with diabetes as a comorbidity and those without diabetes. We included adults (≥18 years) that had stayed in ICU for more than 24 hours and not more than 28 days. A machine learning model was developed to predict 28-day mortality and was compared with other 3 models.

Statistical analysis includes Chi-square test, two-tailed Student’s t-tests, univariate and multivariate logistic regression model for survival analysis. Receiver-Operating Characteristic (ROC) Analysis used to assess accuracy of predictive(scoring) models. All analyses were performed by SAS version 9.4 (SAS institute Inc., Cary, NC)

**Results**: A total of 5044 patients were enrolled but after going through the exclusion criteria the final cohort was 1358 patients. Our XGBoost model outperformed the other 3 models with AUROC of 0.841 and accuracy of 0.860. Using SHAP, the model was explained and top features for predicting mortality were discovered.

**Conclusion**: Comorbid Diabetes slightly increased the risk of 28-day mortality for patients admitted in ICU and the model (XGBoost) performed better than the other models in predicting mortality.

**Keywords**: Chronic Obstructive Pulmonary Disease, XGBoost, Intensive Care Unit

## INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD) accounts for nearly 3.2 million deaths globally in the year 2017 and ranked third among the leading causes of death worldwide. By the year 2020 COPD is believed to have caused an estimate of 6 million global deaths annually[1]. COPD is mostly associated with risk factors like smoking, exposure to fumes and smoke, occupational hazards, poor nutrition [2]. Despite having several risk factors COPD also have seral comorbidities associated with it[3][4]. In most studies they look at the prevalence of the comorbidities with diferent primary admissions. Some studies have done research on Comorbid COPD to check if it increased the risk of 28-day mortality among patients admitted to the ICU (Intensive Care Unit) for non-COPD reasons [4][15]. The prevelance of other comorbidities makes it difficult for clinical management of COPD[1]. We looked at the comorbidities in this study which are: Coronary artery disease(CAD), Hypertension(HTN), Chronic kidney disease(CKD), diabetes mellitus(DM). Recent studies have found that CAD, HTN, CKD, CAD and DM are very common comorbidities in COPD patients[1][3][5][6][12] It is known that Diabetes affects 2–37 % of patients with COPD[7], [8][13] With that in mind, researchers have strongly advised clinicians to screen patients for COPD and the other comorbidities together in order to improve patient care[5] [1]. In their study [9] found that patients with both COPD and Diabetes had 1.62 times higher risk of 2-year mortality than those with just COPD. Understanding the relationship between COPD and other comorbidities is of outmost importance in order to reduce the mortality rate among patients diagnosed with COPD.

Machine learning has been commonly used to predict disease risk, and mortality. It is important to predict mortality in good time which helps health care providers to take necessary steps to intervene in efforts to save lives. Some studies have developed prediction models for COPD in different time period after COPD diagnosis for example[10] [14] developed a model to predict mortality after 5 years of COPD diagnosis. There are several algorithms that can be used to achieve this task of predicting mortality for different diseases. In this paper, we use XGBoost to develop a predictive 28-day mortality model for COPD patients in the ICU, and to use the publicly available database MIMICIII (‘Medical Information Mart for Intensive Care’) [11][5]as a data source. In addition, the performance of the XGBoost model was compared with LR, KNN, and RF model.

## METHODS

### Data Source

Data used in this study was taken from the Medical Information Mart for Intensive Care III (MIMIC-III) database. MIMICIII is a large database with information on over 40,000 de-identified patients that who stayed in critical care units of the Beth Israel Deaconess Medical Center from the year 2001 to 2012[11]. The database contains high temporal resolution data including lab results, electronic documentation, and bedside monitor trends and waveforms. It is available freely to researchers worldwide but Access to the database has to be approved by the institutional review boards of both Beth Israel Deaconess Medical Center and Massachusetts Institute of Technology Affiliates.

### Study population

All patients who are above 18 years were considered to be enrolled as study participants but we selected only the last record was kept to avoid having duplicate patient id the last day of admission for that patient was kept mainly because it contains the actual outcome (Lived or Died). The recruitment procedure was that first we selected patients based on their length of stay. We excluded whose length of stay was less than 1 day and longer than 28 days. Lastly we removed all data with missing values more than 15%.In their research found that MIMIC-III database can have as high as 74% of missing data[12]. In their study[13] the removed all data with more than 20% missing variables. In this study we tried to keep as much as real values as possible hence removing all the records with more than 15% missing values. Figure 1, provides detailed procedure on how we recruited our participants for the study.

OUTCOMES

Death = 310

Alive = 1048

ALL COPD RECORDS = 17141

Admissions(hadm\_id) = 6577

Patients(subject\_id) = 5044

RECORDS = 14773

Admissions(hadm\_id) = 5726

Patients(subject\_id) = 4511

RECORDS = 4511

Admissions(hadm\_id) = 4511

Patients(subject\_id) = 4511

RECORDS = 1358

Patients(subject\_id) = 1358

Missing values more than 15%

Length of stay

Between 1 day and 28 days

Last day of admission for all admissions

Figure 1. Cohort recruitment criteria

### Data Extraction

Data for the study was determined using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes for COPD. (COPD '490', '4910', '4911', '4912', '49120', '49121', '49122', '4918', '4919', '492', '4920', '4928', '494', '4940', '4941', '496',) [14], [15].Basically, MIMIC-III contains a lot of records, in this study, the data was extracted for all the patients admitted in the ICU with COPD as morbidity, and from this cohort we identified patients that were diagnosed with the other underlying comorbidities. MIMICIII has two databases (Metavision and Carevue)[16] in this study we focused much on the metavision database not the carevue database when getting the records.

Apart from comorbidity COPD, Medications, Laboratory tests and other clinical variables were extracted or calculated: BMI calculated from weight and height, age calculated from the date of birth and first day of hospital admission, Elixhauser Comorbidity index[17]. Other comorbidities like hypertension, Cancer were also identified from our cohort using ICD9-codes. These clinical and laboratory variables were based on first day of admission.

Outcome variables extracted include 28-day mortality (after ICU admission), length of hospital admission stay, length of ICU stay, renal replacement therapy and respiratory failure. The primary endpoint of the study was 28-day mortality. Since it is possible for a patient to have multiple ICU admission during one hospitalization ICU mortality and length of ICU stay were determined by calculating the mean ICU length of stay for each patient. For easy calculation of mortality and avoiding duplicate subject\_ids. We chose the last day because that’s when the mortality is actually known since the first few admissions might not give accurate information whether the patient is alive or dead by the time they are discharged.

All data cleaning was performed using the Python programming language (v3.7) utilizing pandas package and analysis was performed using Statistical Analysis Software **(**SAS 9.4).

### Statistical analysis

Continuous data were expressed as median (IQR) and were evaluated by two-tailed Student’s t-test wile categorical data were express as frequencies (%) and were analyzed by the Chi-square tests. The demographic information and mortality status were presented in table 1. We had the following scoring model building process. First, the univariate analyses for all variables were performed and identified the variables with p-value <0.05. Second, all potential variables identified in the first step were entered into multivariate analysis simultaneously to adjust for the potential confounding effects and only p-values <0.05 were retained in the multivariate model. (Table 2) We only retain the variable with p-values <0.05 in the multivariate model in the scoring model. The scoring model was built according to clinical stand or Receiver operating characteristic curve (ROC) detailed in appendix. The ability of the scoring model in distinguish mortality status were performed via ROC and presented in Figure 5.

**Table 1: Demographic data and outcomes regarding to mortality in COPD patients**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable  Median (IQR); (n, %) | Total  (N=1,358) | Dead  (N=310) | Survive  (N=1048) | P value |
| Age | 72.0 (63.0-80.0) | 75.0 (67.0-82.0) | 70.0 (62.0-79.0) | 0.0192\* |
| Sex (Male) c | 748 (55.1) | 163 (52.6%) | 585 (55.8%) | 0.3137 |
| Comorbidity Status (Elixhausr>0) | 145 (10.7%) | 37 (11.9%) | 108 (10.3%) | 0.2405 |
| Major Comorbidity |  |  |  |  |
| CAD | 539 (39.7%) | 118 (38.1%) | 421 (40.2%) | 0.5053 |
| HTN | 755 (55.6%) | 150 (48.4%) | 605 (57.7%) | 0.0036\*\* |
| CKD | 354 (26.1%) | 109 (35.2%) | 245 (23.4%) | <0.0001\*\* |
| Cancer | 083 (06.1%) | 030 (9.7%) | 53 (5.1%) | 0.0029\*\* |
| DM | 350 (25.8%) | 80 (25.8%) | 270 (25.8%) | 0.9878 |
| Clinical Values |  |  |  |  |
| BMI | 027.2 (22.7-32.5) | 25.8 (21.1-30.8) | 27.7 (23.4-33.1) | <0.0001\*\* |
| HbA1c | 6.4 (6.4-6.4) | 6.4 (6.4-6.4) | 6.4 (6.4-6.4) | 0.1954 |
| FPG | 139.4 (119.0-139.4) | 139.4 (121.0-150.0) | 139.4 (118.0-139.4) | 0.0026\*\* |
| SBP | 116.0 (102.0-133.0) | 110.0 (97.0-126.0) | 117.0 (104.0-135.0) | <0.0001\*\* |
| PIP (N=1335) | 19.8 (16.0-24.0) | 20.0 (16.0-27.0) | 19.8 (16.0-23.0) | 0.0001\*\* |
| MAPS (N=1332) | 9.0 (8.0-11.0) | 9.5 (8.0-12.0) | 9.0 (7.0-10.0) | <0.0001\*\* |
| WBC | 10.7 (8.0-14.6) | 11.3 (8.0-16.3) | 10.6 (8.1-14.3) | 0.0424\* |
| Neutrophil (N=1,199) | 81.4 (73.4-88.0) | 84.9 (75.9-90.4) | 80.4 (73.0-87.0) | <0.0001\*\* |
| Lactate | 1.5 (1.0-2.1) | 1.7 (1.1-2.6) | 1.4 (1.0-2.0) | <0.0001\*\* |
| CREAT | 1.0 (0.7-1.4) | 1.2 (0.7-2.0) | 1.0 (0.7-1.4) | <0.0001\*\* |
| SOFA Score | 5.0 (3.0-7.0) | 5.5 (3.0-9.0) | 4.0 (3.0-6.0) | <0.0001\*\* |
| Medication |  |  |  |  |
| Norepinephrine | 463 (34.1%) | 184 (059.4%) | 279 (26.6%) | <0.0001\*\* |
| Epinephrine | 140 (10.3%) | 035 (11.3%) | 105 (10.0%) | 0.5179 |
| Vasopressin | 150 (11.1%) | 91 (29.4%) | 059 (05.6%) | <0.0001\*\* |
| Outcomes |  |  |  |  |
| ICU length of stay | 3.6 (2.0-6.5) | 5.0 (2.5-8.5) | 3.4 (2.0-6.0) | <0.0001\*\* |
| Respiratory failure | 1040 (76.6%) | 260 (83.9%) | 780 (74.4%) | 0.0006\*\* |
| Renal replacement  therapy | 69 (05.1%) | 25 (8.1%) | 44 (4.2%) | 0.0065\*\* |

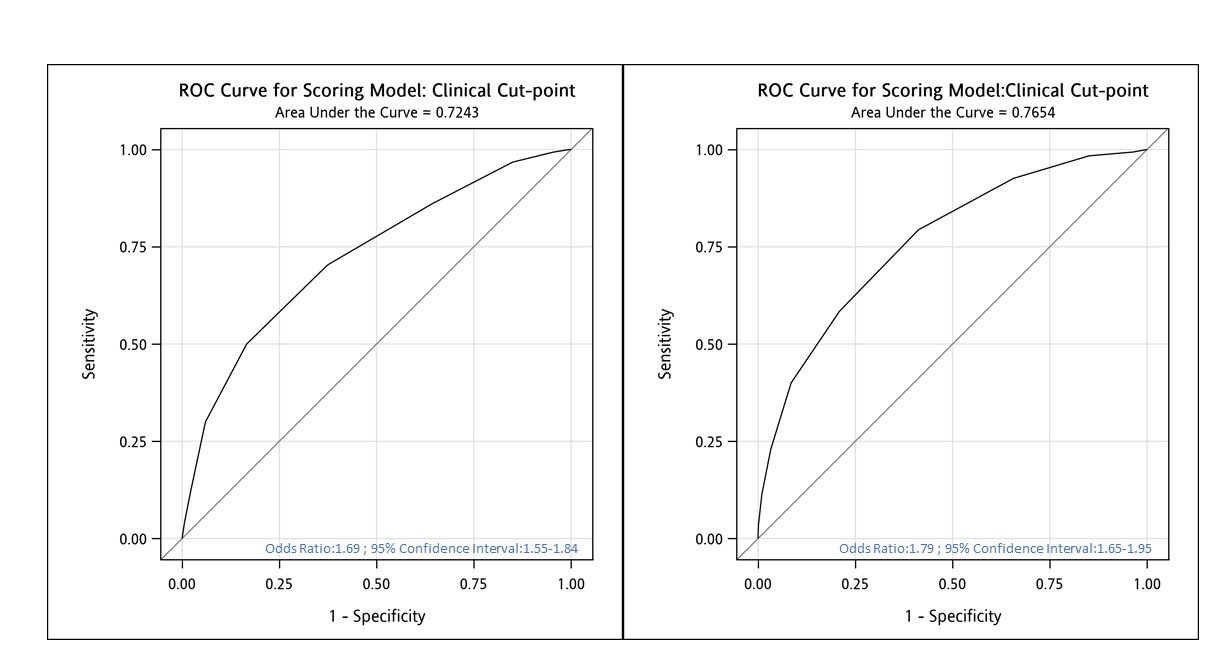
c Chi-square test. Mann-Whitney U test. \*p<0.05, \*\*p<0.01. CAD: Coronary artery disease; HTN: Hypertension; CKD: Chronic kidney disease; DM: diabetes mellitus; BMI: Body mass index; HbA1c: hemoglobin A1c; FPG: Fasting plasma glucose; SBP: Systolic blood pressure; PIP: Peak inspiratory pressure; MAPS: Mean airway pressure; WBC: White blood cell; CREAT: Creatinine; SOFA score: the sequential organ failure assessment score; ICU: intensive care unit

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table 2: Univariable and Multivariable Regression Analyses of Mortality | | | | | |
| Variable | **Univariable** | | **Multivariable Model 1** | |
|  | **OR (95% CI)** | **P value** | **OR (95% CI)** | **P value** |
| Age | 1.03 (1.02-1.04) | <0.0001\* | 1.04 (1.03-1.06) | <0.0001\* |
| Sex (Male) | 0.88 (0.68-1.13) | 0.3139 | 0.76 (0.55-1.07) | 0.1164 |
| DM | 1.00 (0.75-1.34) | 0.9878 | 1.06 (0.72-1.57) | 0.7638 |
| HTN | 0.69 (0.53-0.89) | 0.0037\* | 0.73 (0.52-1.03) | 0.0722 |
| CKD | 1.78 (1.35-2.34) | <0.0001\* | 1.36 (0.89-2.09) | 0.1562 |
| Cancer | 2.01 (1.26-3.21) | 0.0034\* | 2.35 (1.28-4.30) | 0.0057\* |
| BMI | 0.97 (0.95-0.99) | 0.0002\* | 0.95 (0.93-0.97) | <0.0001\* |
| FPG | 1.01 (1.00-1.01) | 0.0002\* | 1.01 (1.00-1.01) | 0.0009\* |
| SBP | 0.98 (0.98-0.99) | <0.0001\* | 0.99 (0.99-1.00) | 0.0369\* |
| PIP | 1.05 (1.03-1.06) | <0.0001\* | 1.03 (1.01-1.05) | 0.0145\* |
| MAPS | 1.14 (1.10-1.18) | <0.0001\* | 1.10 (1.05-1.16) | 0.0002\* |
| WBC | 1.03 (1.01-1.05) | 0.0005\* | 1.02 (0.99-1.04) | 0.2249 |
| Neutrophil | 1.01 (1.00-1.02) | 0.3334 |  |  |
| LACTATE | 1.36 (1.25-1.48) | <0.0001\* | 1.20 (1.08-1.35) | 0.0012\* |
| CREAT | 1.28 (1.16-1.42) | <0.0001\* | 1.11 (0.95-1.30) | 0.1861 |
| SOFA Score | 1.14 (1.10-1.19) | <0.0001\* | 1.00 (0.95-1.06) | 0.9498 |
| Norepinephrine | 4.03 (3.09-5.25) | <0.0001\* | 2.12 (1.46-3.09) | <0.0001\* |
| Vasopressin | 6.96 (4.86-9.97) | <0.0001\* | 2.25 (1.39-3.63) | 0.0009\* |
| ICU length of stay | 1.06 (1.03-1.08) | <0.0001\* | 1.03 (1.00-1.06) | 0.1072 |
| Respiratory failure | 1.79 (1.28-2.49) | 0.0006\* | 1.66 (0.94- 2.91) | 0.0797 |
| Renal replacement | 2.00 (1.20-3.33) | 0.0074\* | 1.20 (0.58-2.47) | 0.6301 |

c Chi-square test. Mann-Whitney U test. \*p<0.05, \*\*p<0.01. CAD: Coronary artery disease; HTN: Hypertension; CKD: Chronic kidney disease; DM: diabetes mellitus; BMI: Body mass index; HbA1c: hemoglobin A1c; FPG: Fasting plasma glucose; SBP: Systolic blood pressure; SPO2: oxygen saturation; PIP: Peak inspiratory pressure; MAPS: Mean airway pressure; WBC: White blood cell; CREAT: Creatinine; ICU: intensive care unit

Model 1 adjusted for age, gender & diabetes status.

**Figure 5 ROC curve for Scoring Models**

****



### Features

Features include, Demographic and demographic variables like Age, Gender (1 if male, 0 if female), BMI Medications: Epinephrine, Norepinephrine, Vasopressin represented by a binary 1 or 0 variable. Lab tests: HbA1c(Hemoglobin), fpg(Glucose) d1\_spo2(Oximetry) d1\_sbp(), d2\_maps(Mean Arterial Pressure), WBC(White Blood Cells Count) Neutrophils, Lactate, Creatinine, PIP(peak inspiratory pressure). A series of comorbidities for each admission represented by a binary 1 or 0 variable. Hypertension(HTN), Cancer, Chronic kidney disease (CKD), Coronary Artery Disease(CAD), Renal Failure, Renal Replacement therapy, and Elixhauser comorbidity index

### Training and Testing data

Data was spilt into Train and Test data, Training data was 80% while the 20% was used for testing. The model was interpreted by SHAP (SHapley Additive exPlanations)

### Model

A prediction model was developed to predict 28-day mortality for ICU admitted patients. For this we used Extreme Gradient Boosting (XGBoost), an ensemble machine learning method based on decision trees. The variables used predict the mortality include various demographic, laboratory, medications, a series of comorbidities and Elixhauser comorbidity index.

To confirm the effectiveness of our model (XGBoost), we employed other widely used machine learning models (LR, KNN, RF) and compared it against them

Logistic Regression (LR) is a classification method in statistical model, which was borrowed into machine learning. It is a choice in many medical data classification [18][7] used to calculate the probability of certain classes or event, and it also allows modeling and multivariate analysis of binary dependent variables. The coefficients of predictors included in the final model are estimated using the multivariate analysis and are then adjusted based on the predictors of the model. The risk estimate of the outcome is quantified by the contribution of each predictor [19][6].

K-Nearest Neighbors is a simple supervised learning machine learning algorithm that assumes similar things exists in close proximity and looks for a pattern in those occurrences. KNN can be used for both classification problems or regression problems despite being a simple algorithm it can still give competitive results. However, in the industry, KNN is widely used in classification problems.

Random Forest (RF) is an ensemble machine learning algorithm, which puts together multiple decision trees to predict the outcome based on the average probability of all the trees on each subset of data samples to obtain better predictive performance that can’t be obtained by a single algorithm. Apart from producing precise predictions, being fast and easy to implement, RF can handle large number of input variables without overfitting.

Extreme Gradient Boosting popularly knowns as XGBoost, is an improved algorithm based on the gradient boosting decision tree. It combines a set of machine learning algorithms to come up with a better machine learning algorithm as a whole. To solve many data science problems in a fast and accurate way XGBoost provides a parallel tree boosting (also known as GBDT, GBM). The boosted trees in XGBoost is divided into regression trees and classification trees and the core of the algorithm is to optimize the value of the objective function. Some of the advantages of XGBoost includes scalability in all scenarios, and being fast.

## DISCUSSION

The study selected patients who were admitted to ICU with COPD as a primary reason, and investigated the relationship between it and other comorbidities and clinical outcomes. The main comorbidity focused on was DM, to see if there is higher mortality rate on those patients that had COPD DM positive compared to those that had COPD DM negative Results of the study indicated that there was no significant impact on mortality rate for COPD with DM positive for the patients after ICU admission.

On the other hand, COPD with HTN had higher occurrences which is 48% of our total population

We attempted to compare the performance of the models to predict 28-day mortality using the data we had. Using the AUROC (area under the receiver operating characteristic) curve analysis we found that AUROC for predicting 28-day mortality in XGBoost(AUROC: 0.841 ACCURACY:0.860 F1 score: 0.91) which was better than the other machine learning models. KNN(AUROC: 0.732 ACCURACY: 0.808 F1 score: 0.89) LR(AUROC: 0.819 ACCURACY: 0.830 F1 score: 0.90) RF(AUROC: 0.813 ACCURACY: 0.808 F1 score: 0.89)

Our main focus when comparing the models was on AUROC and XGBoost performed better on AUROC. Using AUROC to measure model performance is considered to be the best way to measure a binary classification model than using accuracy. Accuracy in most cases is based on probability while AUROC is how well a model can classify the outcome. Using accuracy on imbalanced dataset is not a good practice because the model might just as well predict the majority class only while missing the minor class which in case of mortality is dangerous. Since our dataset was imbalanced we opted to use the AUROC to rank the model performance between the 4 models developed.

The following are the parameters used in XGBoost: learning\_rate =0.1, n\_estimators=100, scale\_pos\_weight=0.95, max\_depth=9, min\_child\_weight=4, max\_delta\_step =10, subsample=0.9, colsample\_bytree=1. Other pareameters were left to be default parameters.

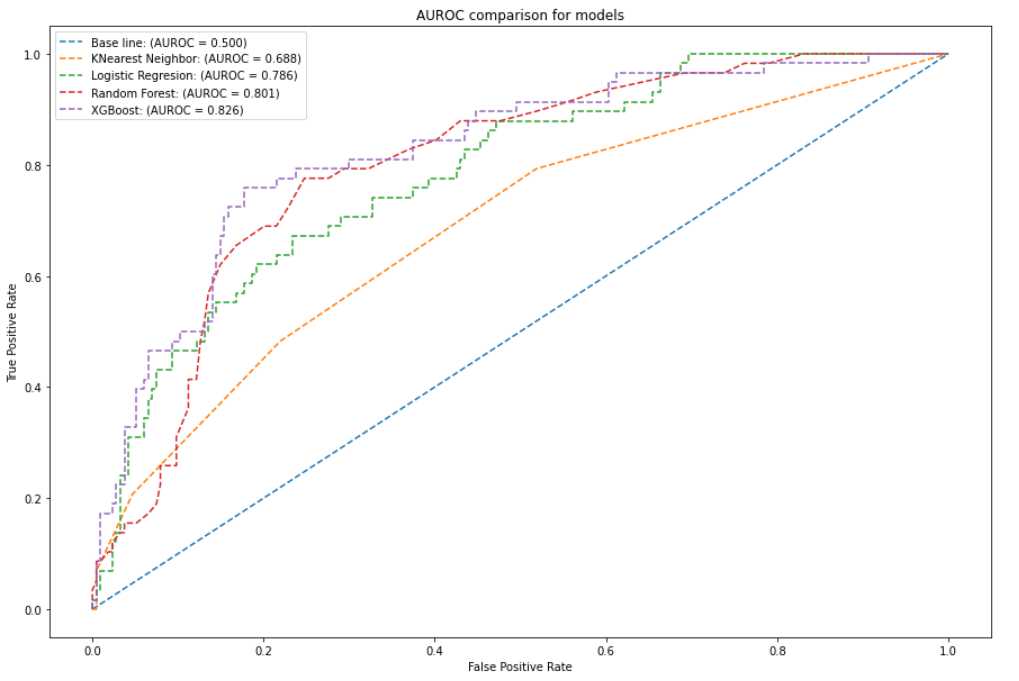


Figure 2. Area Under the Receiver Operating Characteristic curves showing the performance of the XGBoost, KNN, LR, and RF for predicting 28-day mortality in critically ill COPD patients.

On top of this we performed cross validation on all the models, still our method come out with the best mean accuracy in comparison with the other 3 models. We applied 5 K-fold cross validation so that we could determine which model will come out to be the best. KNN had an Accuracy: 78.13% and standard deviation of (1.55%), LR Accuracy: 82.62% standard deviation (1.75%), RF Accuracy: 82.11% standard deviation (1.94%) and XGBoost scored an Accuracy: 83.51% standard deviation (1.53%)

### Model explanation

Most complex machine learning models cannot be easily interpreted by just looking at the model itself as the best way to understand it. We used SHAP a set of tools to interpret our black box model[20] thus XGBoost. This helps in understanding the global model structure by combining several local explanations for each prediction [21][11]. SHAP values interpret the impact of having a certain value for a given feature in comparison to the prediction we'd make if that feature took some baseline value. It helps in breaking down how the model works for an individual prediction as well as a summary of all the features importance.

Figure 3. shows the impact each feature has on the model output. This chart shows the magnitude of each feature over all samples. It also shows the distribution of impact each feature has. The x axis shows how the value helps the model towards positive outcome or negative outcome, in our case red means towards 1 (death) and Blue towards the negative which is 0 (Alive) The color represents the feature value. Red indicating Higher and Blue indicating Low for example in the figure, on pip, the higher the pip the higher the chances of not surviving. Figure 4 Shows top 20 ranking variables used to train XGBoost Model. X axes show the average impact of model output magnitude, expressed by SHAP values.

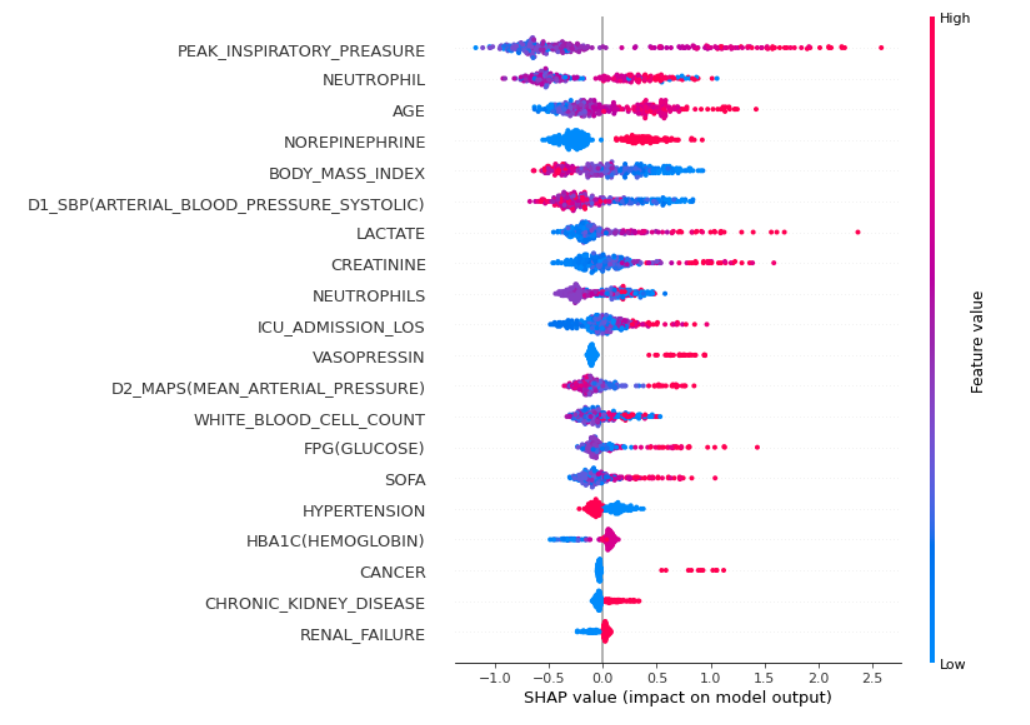


Figure 3. Feature impact on the model output.

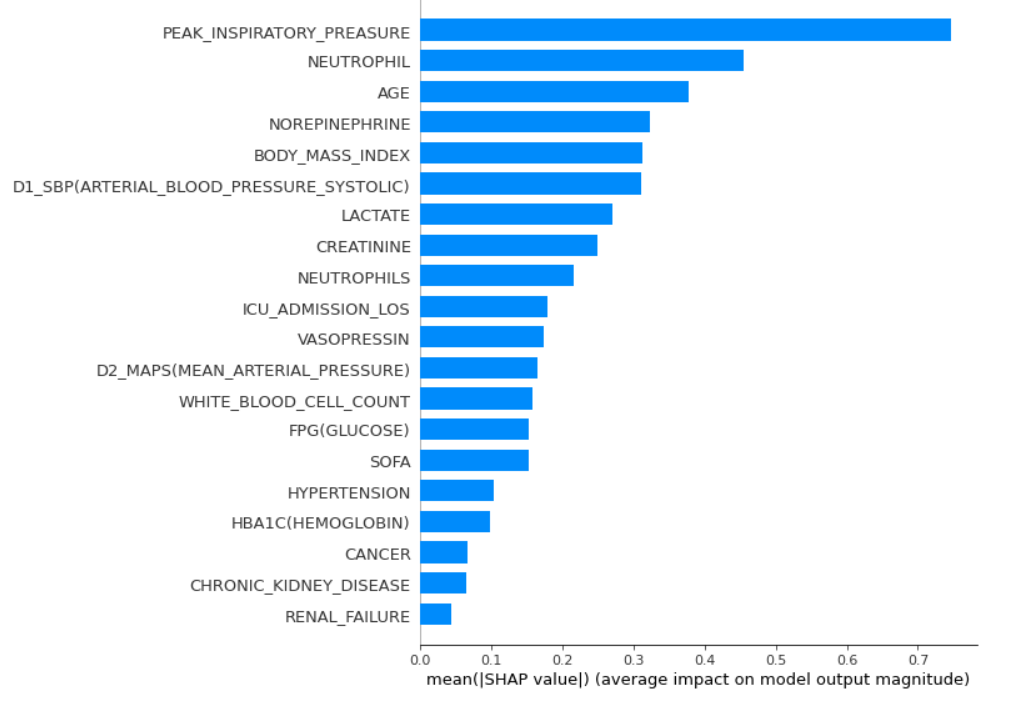


Figure 4 Average impact of features (top 20 ranking variables) on model output.

## CONCLUSION

Four models for predictions of 28-day mortality in ICU for patients with COPD were developed XGBoost model with other machine learning algorithms performed better in both AUROC curve and Model Accuracy with the AUROC of 0.841. important features were discovered that would help in alerting health service providers in time to save lives. The prediction of mortality is important in healthcare and proper care of the patient can be done soon after the prediction is made in good time.

The dataset contained a lot of missing values made us drop a lot of rows which was more than 50% of the main cohort.

Future work, we will continue to look at other possible features that can be used to predict mortality in COPD patients in addition to the ones we have already studied.

### References

[1] A. Cavaillès *et al.*, “Comorbidities of COPD,” *Eur. Respir. Rev.*, vol. 22, no. 130, pp. 454–475, Dec. 2013, doi: 10.1183/09059180.00008612.

[2] WHO, “Chronic obstructive pulmonary disease (COPD),” Jun. 21, 2021. https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd) (accessed Jun. 29, 2021).

[3] G. Hillas, F. Perlikos, I. Tsiligianni, and N. Tzanakis, “Managing comorbidities in COPD,” *Int. J. Chron. Obstruct. Pulmon. Dis.*, vol. 10, pp. 95–109, 2015, doi: 10.2147/COPD.S54473.

[4] W. Huang, R. Xie, Y. Hong, and Q. Chen, “Association Between Comorbid Chronic Obstructive Pulmonary Disease and Prognosis of Patients Admitted to the Intensive Care Unit for Non-COPD Reasons: A Retrospective Cohort Study,” *Int. J. COPD*, vol. 15, pp. 279–287, Feb. 2020, doi: 10.2147/COPD.S244020.

[5] Patricia Inachio PhD, “With Type 2 Diabetes and COPD Increasing, Researchers Look for Links,” Apr. 30, 2019. https://copdnewstoday.com/2019/04/30/type-2-diabetes-copd-increasing-researchers-look-links/ (accessed Jun. 29, 2021).

[6] S. Gläser, S. Krüger, M. Merkel, P. Bramlage, and F. J. F. Herth, “Chronic Obstructive Pulmonary Disease and Diabetes Mellitus: A Systematic Review of the Literature,” *Respiration*, vol. 89, no. 3, pp. 253–264, 2015, doi: 10.1159/000369863.

[7] J. S. Rana *et al.*, “Chronic obstructive pulmonary disease, asthma, and risk of type 2 diabetes in women,” *Diabetes Care*, vol. 27, no. 10, pp. 2478–2484, Oct. 2004, doi: 10.2337/diacare.27.10.2478.

[8] P. Rogliani, G. Lucà, and D. Lauro, “Chronic obstructive pulmonary disease and diabetes,” *COPD Res. Pract.*, vol. 1, no. 1, p. 3, Aug. 2015, doi: 10.1186/s40749-015-0005-y.

[9] T.-W. Ho, C.-T. Huang, Y.-J. Tsai, A. S.-Y. Lien, F. Lai, and C.-J. Yu, “Metformin use mitigates the adverse prognostic effect of diabetes mellitus in chronic obstructive pulmonary disease,” *Respir. Res.*, vol. 20, Apr. 2019, doi: 10.1186/s12931-019-1035-9.

[10] S. J. Kiddle, H. R. Whittaker, S. R. Seaman, and J. K. Quint, “Prediction of five-year mortality after COPD diagnosis using primary care records,” *PLOS ONE*, vol. 15, no. 7, p. e0236011, Jul. 2020, doi: 10.1371/journal.pone.0236011.

[11] A. E. W. Johnson *et al.*, “MIMIC-III, a freely accessible critical care database,” *Sci. Data*, vol. 3, no. 1, p. 160035, May 2016, doi: 10.1038/sdata.2016.35.

[12] H. Zhang, M.-F. Balcan, and D. P. Woodruff, “Medical Missing Data Imputation by Stackelberg GAN,” p. 13.

[13] N. Hou *et al.*, “Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost,” *J. Transl. Med.*, vol. 18, no. 1, p. 462, Dec. 2020, doi: 10.1186/s12967-020-02620-5.

[14] R. A. Deyo, D. C. Cherkin, and M. A. Ciol, “Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases,” *J. Clin. Epidemiol.*, vol. 45, no. 6, pp. 613–619, Jun. 1992, doi: 10.1016/0895-4356(92)90133-8.

[15] H. Quan *et al.*, “Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data,” *Med. Care*, vol. 43, no. 11, pp. 1130–1139, Nov. 2005, doi: 10.1097/01.mlr.0000182534.19832.83.

[16] Johnson, Alistair, Pollard, Tom, and Mark, Roger, “MIMIC-III Clinical Database.” PhysioNet, 2015. doi: 10.13026/C2XW26.

[17] “Elixhauser Comorbidity Index - an overview | ScienceDirect Topics.” https://www.sciencedirect.com/topics/medicine-and-dentistry/elixhauser-comorbidity-index (accessed Jun. 29, 2021).

[18] C.-A. Hu *et al.*, “Using a machine learning approach to predict mortality in critically ill influenza patients: a cross-sectional retrospective multicentre study in Taiwan,” *BMJ Open*, vol. 10, no. 2, p. e033898, Feb. 2020, doi: 10.1136/bmjopen-2019-033898.

[19] J. Liu, J. Wu, S. Liu, M. Li, K. Hu, and K. Li, “Predicting mortality of patients with acute kidney injury in the ICU using XGBoost model,” *PLOS ONE*, vol. 16, no. 2, p. e0246306, Feb. 2021, doi: 10.1371/journal.pone.0246306.

[20] Na Serere, “Black Box Model Using Explainable AI with Practical Example,” *Analytics Vidhya*, Oct. 24, 2020. https://www.analyticsvidhya.com/blog/2020/10/unveiling-the-black-box-model-using-explainable-ai-lime-shap-industry-use-case/ (accessed Jun. 29, 2021).

[21] S. M. Lundberg *et al.*, “From local explanations to global understanding with explainable AI for trees,” *Nat. Mach. Intell.*, vol. 2, no. 1, pp. 56–67, Jan. 2020, doi: 10.1038/s42256-019-0138-9.