2021 International Conference on ICT for Smart Society (ICISS) | 978-1-6654-1697-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICISS53185.2021.9533219

A Tree-based Mortality Prediction Model of COVID-19 from Routine Blood Samples

Nunung Nurul Qomariyah
Computer Science Department,
Faculty of Computing and Media,
Bina Nusantara University,
Jakarta, Indonesia 11480
nunung.qomariyah@binus.edu

Ardimas Andi Purwita
Computer Science Department,
Faculty of Computing and Media,
Bina Nusantara University,
Jakarta, Indonesia 11480
ardimas.purwita@binus.edu

Sri Dhuny Atas Asri
Head of Functional Medical Staff
Pulmonology Department,
Pasar Minggu Regional Hospital,
South Jakarta, Indonesia, 12550
dhunyatasasri@gmail.com

Dimitar Kazakov
Computer Science Department,
University of York,
United Kingdom, YO10 5GH
dimitar.kazakov@york.ac.uk

***Abstract*—COVID-19 has been declared by The World Health Organization (WHO) a global pandemic in January, 2020. Researchers have been working on formulating the best approach and solutions to cure the disease and help to prevent such pandemics in the future. A lot of efforts have been made to develop a fast and accurate early clinical assessment of the disease. Machine Learning (ML) has proven helpful for research and applications in the health domain as a way to understand real-world phenomena through data analysis. In our experiment, we collected the retrospective blood samples data set from 1,000 COVID-19 patients in Jakarta, Indonesia for the period of March to December 2020. We report our preliminary findings on the use of common blood test biomarkers in predicting COVID-19 patient mortality. This study took advantage of explainable machine learning to examine the data set. The contribution of this paper is to explain our findings on predicting COVID-19 mortality, including the role of the top 11 biomarkers found in our dataset. These findings can be generalized, especially in Indonesia, which is now at its highest peak of the epidemic. We show that tree-based AI models performed well on predicting COVID-19 mortality, while also making it easy to interpret the findings, as they lend themselves to human scrutiny and allow clinicians to interpret them and comment on their viability.**

***Index Terms*—routine blood tests, COVID-19, mortality prediction, machine learning, classification, tree-based models**

I. Introduction

COVID-19 (Coronavirus 2019) is a disease caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). It is highly infectious and can spread easily through respiratory droplets when people sneeze or cough. The first identified case was in December, 2019 in Wuhan, China, and since the beginning of 2020, the disease has spread all over the world. On 30 January 2020, The World Health Organization (WHO) upgraded the COVID-19 outbreak to a global pandemic. Researchers from various backgrounds are making joint efforts to combat this disease and are racing to develop means of its early detection to prevent wider transmission. In Indonesia, the first two cases were confirmed on 2 March 2020 [1].

At present, more than 18 months after the first case has been identified, the number of critically ill patients with COVID-19 is still increasing, despite an active vaccination campaign taking place in several countries, including Indonesia. It is still hard to perform a differential diagnosis especially as reliable COVID-19-specific tests can be expensive, not universally accessible, and may require time to be carried out in practice. The progress of seriously ill COVID-19 patients is usually rapid and there are no clear distinctive symptoms associated with critical or severe illness. Many sudden deaths are still being reported. The importance of this research resides in identifying mortality prediction features that can be used as an early detection of COVID-19 patients whose condition is expected to deteriorate.

There are many predictors that can be used to predict the severity of the COVID-19 disease, as seen in the recent literature [2]. These include demographics, hypoxia, radiographic features (chest x-ray (CXR) and computed tomography (CT)) and laboratory biomarkers, such as D-dimer levels, C-reactive protein (CRP), Lactate dehydrogenase (LD or LDH), and high-sensitivity cardiac troponin I.

978-1-6654-1697-9/21/31.00 ©2021 IEEE

In this study, we use blood biomarkers to predict the COVID-19 patient mortality. The reason of using a blood sample is that the laboratory results can be easily collected and this analysis is routinely done for alll hospitalized patients. In addition, blood biomarkers are considered objective indicators which can be used to represent the patient condition in quantitative ways that can be easily learned by the machine learning model, which, in turn, can render the machine learning model more reliable.

The main contribution of this paper consists of: (1) reporting the 11 most prominent blood biomarkers that have been confirmed by the physicians, and (2) analyzing the discovery of three different tree-based models, and the salient features they have in common.

II. Related Work

Several studies have reported the result of the experiment with routine blood test to predict the mortality of COVID-19 patients which yielded a very high accuracy. The studies are not only shown that a machine learning is able to predict better and faster, but also suggest that in addition to the most common assessment method used to monitor the progress of common pulmonary disease such as X-rays and CT-scan images, the routine blood test can also be used as indicators of the COVID-19 severity level and predictors of the mortality.

A previous study by Yan *et al.* [3] in 2020 mentioned the three most prominent features found in the blood samples data which can predict the mortality of the COVID-19 patients, i.e., Lactic Dehydrogenase (LDH), Lymphocyte and high-sensitivity C-Reactive Protein (hs-CRP). The training experiment was conducted by using 375 patients data from Tongji hospital in Wuhan, China. The model was tested to another 110 patients data and the prediction result was claimed very accurate (over 90%). Although, this work recently received some counter statements from [4] and [5] regarding the clarity of how the blood test result were obtained and some possibilities of other complication in critically ill patients.

Another study by Habbu *et al.* [6] reported similar experiment with blood sample in India. Based on their findings, they concluded that the most correlated factors with the mortality were age, gender, and other complications such as diabetes mellitus and hypertension. Result of the experiment showed that the diabetes contributed as high as 53%, while hypertension shows as high as 33%. The other comorbidity such as cardiovascular, asthma and cerebrovascular disease were also found to be significant.

Similar study in Korea was conducted by Ko *et al.* [7]. They proposed EDRnet which was built based on deep neural network and random forest models. In their study, the model were trained on the blood test data which was obtained immediately within 24 hours after the patients being hospitalized. They claimed that their developed model can detect earlier than those which proposed earlier by Yan *et al.* [3]. The model were trained from the same data used in [3], and then tested to 106 other patient data from Korean hospitals. The accuracy result of the model reached 92%. The finding of the study were also supported by the other studies such as [8] which explain that the lower lymphocyte count were found in severe patients. This could be due to the infiltration and sequestration of CD4+/CD8+ T cells occured in patients with poor outcome. Other study which also supported the finding was conducted by Kong *et al.* [9] which mentioned that the Neutrophil-to-lymphocyte (NLR) in severe patients were found higher than the mild one. The COVID-19 disease mainly act on lymphocytes, particulary T lymphocytes. This study also suggests that patients with high NLR should be admitted to an isolation ward as early as possible. With regard to the Platelet, many studies already confirmed that the lower count will increase the severity level of the patients [10]. This specifically related to COVID-19 because the decrease of immune system may lead to inappropriate platelet activation and consumption as well as impaired megakaryopoiesis as mentioned in [10].

Moreover, Sun *et al.* [11] proposed the model based on temporal deep learning to classify the COVID-19 progression. The model was also trained on the same data published in [3]. They also proposed four COVID-19 stages definition which were never existed before. Based on their experiment, they found that low values of lymphocytes, eGFR (estimated Glomerular Filtration Rate), albumin and Serum Sodium, high values of LDH, hs-CRP, indirect bilirubin, creatinine and INR (International Normalized Ratio or also known as PT which stands for Prothrombin Time) were shown in the COVID-19 patients with critical condition. Similar result were found in the other study [12] which concluded that high-sensitivity C-reactive protein (hsCRP), aspartate aminotransferase (AST), and D-dimer were the indicators of COVID-19 mortality.

Other study [13] has collected the summary of most important biomarkers and describe their findings in critical patients. They found that C-reactive protein, Serum Amyloid A, Interleukin-6, Lactate Dehydrogenase (LDH), D-Dimer, Cardiac Troponin and Renal biomarker (Urea and Creatinin) have increased, while the White Cell Count (WCC) for NLR and Leucocyte Count (LC) have increased and decreased, respectively.

III. Research Method

A. Dataset

The dataset of routine blood sample test of 1000 patients has been collected from the hospital in Jakarta, Indonesia from the period of March until December 2020. Due to the confidentiality and the permission to use the dataset, the content of the data will be kept privately by the authors but the sample structure of the dataset and column description are provided. This is done purposely so that the overall machine learning pipeline in this experiment can be clarified. Our original data consists of several blood test which were performed by the hospital during the period from the first day they were admitted to hospital until they were discharged. Each patient have been represented with several rows in the dataset. Total entries in the dataset was 10242 rows. The data also contains demographics of the patients, including age and gender, but these were excluded in our study as we want to focus on finding the most prominent biomarkers to predict the patient outcome. The dataset contained 179 biomarker features which were then reduced by 28 features only to be used in this study. The features were selected based on the most common findings in the other related work and also based on the clinicians advise. The biomarker features which were used in our experiment is shown in Table I.

B. Data Preprocessing

We performed several steps in data preprocessing. First, the missing data were imputed with the latest data of the same patient with the combination of K-Nearest Neighbour (KNN) imputer method. The KNN imputer was only used if after the all data has been carried forward to fill the missing value, there were still missing values to be filled. KNN works by checking the value of the nearest k-neighbors. Two data points are considered close if the columns that neither is missing are similar (defined as *close*). Then one latest data of each patient were taken to be processed further.

The imputation step resulted in only 984 patient data records that were eligible to be used for further processing. The dataset was imbalanced with a total number of patients who survive of 893, while the non-survivors were only 92. Machine learning algorithms tend to overfit if trained on imbalanced data. In our case, the positive class (non-survivors) is the minority class with a ratio of 1:10. We applied SMOTE [14] as one of the over-sampling method for imbalanced datasets. In this study, we take a sampling size of 500 from each class to be trained.

C. Machine Learning Algorithm

In our study, we selected tree-based model to benefit from their white-box approach advantage which allow us to see the model of the learning algorithm. With this approach, the model can be easily analyzed by the human expert. We used three algorithms in this study: (1) CART Decision Tree, (2) Random Forest, and (3) eXtreme Gradient Boosting (XGBoost).

CART stands for Classification and Regression Tree algorithm which is a term introduced by Breiman *et al.* [15]. It is a decision tree algorithm which can be used for both classification and regression problem. In CART algorithm, the data is represented in a single binary tree with the node represents the feature and the leaf represents the class decision. This algorithm is commonly used in data mining and considered simple and good enough to explain the data. Decision tree is non-parametric and can also deal with a large dataset with simplified tree-based model explanation.

Random Forest algorithm uses more than a single tree to model the data. It can be said that random forest is a collection of decision tree. It takes votes on several decisions made from more than one tree and return the majority. The more diversity of the tree is attached the better the prediction. When compared it to the classical decision tree, this algorithm performs slower but the accuracy is higher. The high accuracy in this algorithm is due to the random features which are chosen during the training process. It does not depend highly on any single or a set of features. With this method, random forest can generalize data better than decision tree.

XGBoost is short for Extreme Gradient Boosting Algorithm which is works by building up many decision trees. It uses gradient descent algorithm to optimize the search. It always tries to correct the model from previous mistake, so the next step is an improvement. The process is continued until there is no further improvement. It is a fast algorithm and can handle large data very well. XGboost can also perform well on data with imbalanced class. The main difference with the previous mentioned algorithm is that the way it builds the tree in additive, one tree at a time. This is done in a forward stage-wise process. Both algorithms also differ in the way they combining the result. Gradient boosting combine the results along the process.

We used the implementation of the three algorithms in Python languange. For the first two algorithms, we used the implementation from Scikit learn library[1], and for the last algorithm, we used the implementation package called xgboost[2]. The hyperparameter setting of each model is shown in Table II.

D. Evaluation Technique

In this study, we used standard classification evaluation technique by measuring the precision, recall, F1-score, accuracy of the model. The dataset was divided into two parts randomly with the ratio of 70% as training data and 30% as testing data. We used stratified sample when dividing the dataset to avoid overfitting. Due to the stochastic nature of the algorithms, the experiments were repeated several times and the average result were reported. The tree generated from each algorithm was selected based on the highest accuracy of the model. All the models were run on the same random state of the data splitting, so there was no bias, they were all observe the same data point. For a deeper analysis, we also reported the feature importance found by the decision tree. We used SHAP (SHapley Additive exPlanations) value for explaining random forest and xgboost

[1] https://scikit-learn.org

[2] https://github.com/dmlc/xgboost

TABLE I: Biomarkers Used in Dataset

Biomarker	Feature code	Normal level (adult)	Unit
HEMATOLOGY			
Hemoglobin	HB	13.2 - 17.3	g/dL
Hematocrit	HCT	40 - 52	%
Leukocytes	LEKO	3.8 - 10.6	10^3/uL
Platelets	PLT	150 - 440	10^3/uL
Erythrocytes	ERI	4.40 - 5.90	10^6/uL
Red Cell Distribution Width	RDW	11.8 - 14.5	%
AVERAGE ERYTHROCYTE VALUE			
Mean Corpuscular Volume	MCV	80 - 100	fl
Mean Corpuscular Hemoglobin	MCH	27.5 - 33.2	pg
Mean Corpuscular Hemoglobin Concentration	MCHC	32 - 36	g/dL
COUNT TYPE			
Basophils	BASOFIL	0.0 - 1.0	%
Eosinophils	EOS	1.0 - 5.0	%
Stem Neutrophils	NEUTB	3.0 - 5.0	%
Segmented Neutrophils	SEGMEN	50 - 70	%
Lymphocytes	LIMFOSIT	25 - 50	%
Monocytes	MONOSIT	2.0 - 8.0	%
Neutrophil-Lymphocyte Ratio	NLR1	<3.12	
Erythrocyte Sedimentation Rate	LED	0 - 20	mm/hour
HEMOSTASIS			
D-Dimer	DDIMER	<0.5	ug/mL
prothrombin time	PTHSL	10.80 - 14.40	second
Activated Partial Thromboplastin Time	APTTHSL	25.00 - 35.00	second
BLOOD CHEMISTRY			
Arterial blood gas analysis			
Partial pressure of oxygen	PO2_N	71.0 - 104.0	mmHg
Oxygen saturation	O2S_N	94.0 - 100.0	%
Liver function			
Serum Glutamic Oxaloacetic Transaminase	SGOT	<50	U/L
Serum Glutamic Pyruvic Transaminase	SGPT	<50	U/L
Diabetes			
Random Plasma Glucose Test	GDSFULL	70 - 180	mg/dL
Kidney Function			
Urea	UREUM	<48	mg/dl
Creatinine	CREAT	0.70 - 1.30	mg/dL
Cardiac enzymes			
Lactate dehydrogenase	LDH	50 - 150	U/L

TABLE II: Setting of model hyper-parameters

Algorithm	Setting
decision tree	min_samples_leaf=50
	criterion=gini
random forest	max_depth=5
	n_estimators=10
	class_weight=balanced_subsample
	min_samples_split=15
xgboost	max_depth = 5
	min_child_weight = 1
	eval_metric=logloss

model. The SHAP values express how big is the contribution of each feature to the predictive power of the model.

IV. Result and Analysis

The feature importance of each model is shown in Figure 1. We show only eleven most prominent features in the dataset. From the figure we can see that those three algorithms shows similar result and this can be easily interpreted by the clinicians. The SHAP value explains how each feature contribute to the model prediction. The Top-11 features and the trend found in the dataset with regard of each class are shown in Table III. The findings is in line with the literature shown in the previous section. As mentioned in [16], the changes in lymphocytes, neutrophils, monocytes, eosinophils, and platelets are related to viral replication and hyperinflammation in COVID-19 cases. The platelet decrease, called thrombocytopenia, has been associated with severity of COVID-19. It is a common condition on patient with COVID-19. The possible cause of the platelet decrease in the blood are (1) direct infection of bone marrow cell by the virus, (2) body's immune system attacking and destroying the platelets, and (3) the aggregation of the platelet in the lungs, which caused in microthrombi and platelet consumption [17].

Arterial blood gas biomarker was also shown as important feature, as expected. The partial pressure of oxygen (PO2_N), or also known as PaO2, measuring the oxygen pressure in the arterial blood, which reflects how well the oxygen flows. We were also expecting the oxygen saturation level to be the most important features. But since we were limiting the model to focus on the showing the Top-11 features to increase the model readibility, the oxygen saturation was not captured here.

Other disease indicators, such as hyperglycemia (the increase of blood sugar level), chronic renal (shown by the increase of urea level), and liver damage (shown by the increase of SGOT/SGPT level) were also captured in the Top-11 features.

The common biomarkers associated with coagulation index, including D-Dimer, prothrombin time (PT), activated partial thromboplastin time (APTT), which could sensitively reflect the blood clotting state [18], were also shown significant in the result.

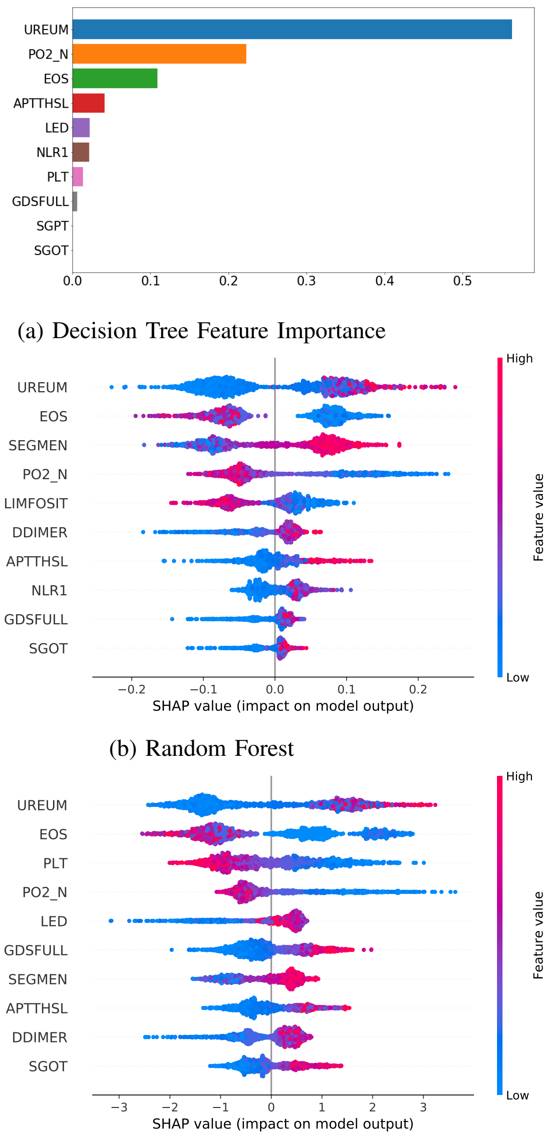
The performance of each model is shown in Table IV. It is shown that XGBoost performs better than the other two tree-based algorithms, as expected, with reasonably fast execution time. In this paper, our aim is to explore the important biomarkers in our dataset, while we can also observe the classification performance of each model.

TABLE III: Top-11 Features Trends

Feature	**Found in the class of**	
	survive	**dead**
UREUM	decrease	increase
PO2_N	increase	decrease
EOS	increase	decrease
APPTHSL	decrease	increase
LED	decrease	increase
DDIMER	decrease	increase
GDSFULL	decrease	increase
SGOT/SGPT	decrease	increase
PLT	increase	decrease
LIMFOSIT	increase	decrease
SEGMEN	decrease	increase

TABLE IV: Model Performance

Algorithm	**Class**	**Prec.**	**Recall**	**F1-score**	**Acc.**	**Exec. time (s)**
decision tree	survive	0.88	0.92	0.90	0.90	0.03
	dead	0.91	0.87	0.89		
random forest	survive	0.94	0.91	0.92	0.92	10.71
	dead	0.91	0.94	0.92		
xgboost	survive	0.99	0.96	0.98	0.98	1.03
	dead	0.96	0.99	0.98		



(a) Decision Tree Feature Importance

(b) Random Forest

(c) XGBoost

Fig. 1: Feature Importance and SHAP Summary

The tree produced by each model is shown in Figure 2, 3 and 4. Figure 2 shows the decision tree result after being trained on the dataset. One possible interpretation can be: if the D-Dimer of a patient was less than or equal to 1.568 ug/mL (the normal level < 0.5ug/mL), then the patient would be more likely to survive. If it is not the case, then we need to check on other conditions. In Figure 3, we can see that the tree produced by random forest was quite similar with the single decision tree. This is due to this method use the same approach with multiple trees. We select randomly a subtree to be shown here as an example. Figure 4 shows a subtree of xgboost model which can explain the data explicitly. One can interpret and read the subtree as if a patient found to have a high LED (Erythrocyte Sedimentation Rate) and D-Dimer was also found to be higher than normal value, while the PLT (platelet counts) lower than 404 10^3/uL (normal level 150 - 440 10^3/uL), then this patient will have higher chance to die than survive, with the probability of 0.629 ($\sigma(0.534)$).

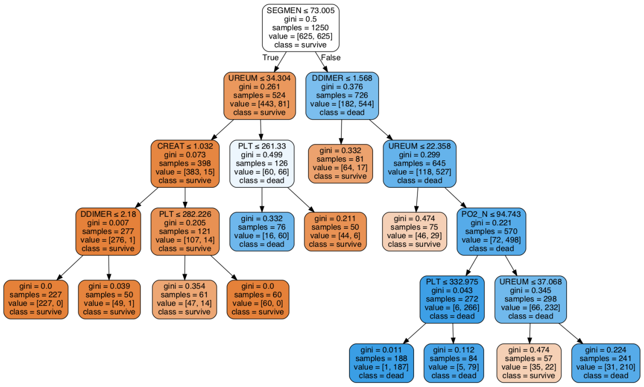


Fig. 2: Decision Tree

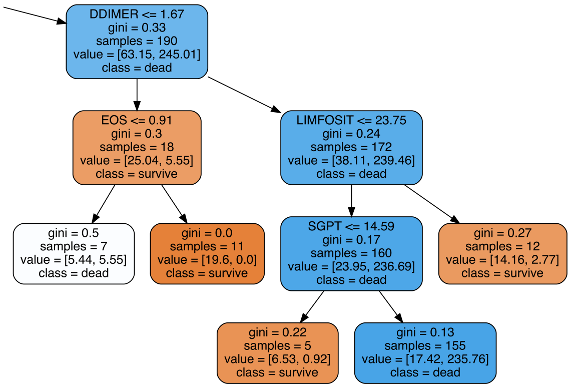


Fig. 3: Random Forest Subtree

V. Conclusions and Future Work

In this study, we have explored the biomarkers in blood sample dataset which can be the best predictors in COVID-19 patients mortality. The study was performed on 1,000 restropective patients data, in which the blood test were taken and recorded during the stay in hospital. This study used tree-based algorithms to get a better explanation of the findings in the data. Our result shows that all the algorithms, decision tree, random forest and xgboost perform well in the dataset, and yielded a valid tree to be examined by the clinicians. All the algorithms used in this study had shown the accuracy above 90% with the training execution time less than 15 seconds. In the future, we can also explore the time dimension of the data and observe whether the current finding is still valid. In addition to the blood sample data, we plan to also add clinicians observation report of the patients during the hospitalization. This report which is usually created in freetext format can be handled by employing some Natural Language Processing methods.

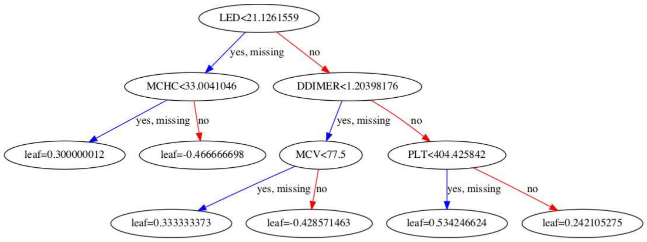


Fig. 4: XGBoost Subtree

Consent and Ethical Clearance

The patient medical records used in this study were collected by the data provider, including epidemiological, demographic, clinical, laboratory and mortality outcome information. This study has been approved by the data provider, Pasar Minggu Regional Hospital Jakarta, Ethics Committee. The requirement for patient consent was waived as this was a secondary analysis of anonymized data.

Acknowledgment

This work was supported by the Research and Technology Transfer Office of Bina Nusantara University as part of Bina Nusantara University's International Research Grant entitled "AI-Based Telemedicine for COVID-19 Patients" with the contract number of 017/VR.RTT/III/2021 and the contract date of 22 March 2021. The last author was supported by the British Council Newton Institutional Links Research grant "AI-Based Telemedicine for COVID-19 Patients".

References

[1] World Health Organization, "WHO Indonesia Situation Report - 1," 2020. [Online]. Available: https://www.who.int/docs/default-source/searo/indonesia/covid19/who-indonesia-situation-report-1.pdf

[2] B. Gallo Marin, G. Aghagoli, K. Lavine, L. Yang, E. J. Siff, S. S. Chiang, T. P. Salazar-Mather, L. Dumenco, M. C. Savaria, S. N. Aung, T. Flanigan, and I. C. Michelow, "Predictors of COVID-19 severity: A literature review," pp. 1–10, 1 2021. [Online]. Available: https://doi.org/10.1002/rmv.2146

[3] L. Yan, H.-T. Zhang, J. Goncalves, Y. Xiao, M. Wang, Y. Guo, C. Sun, X. Tang, L. Jing, M. Zhang, X. Huang, Y. Xiao, H. Cao, Y. Chen, T. Ren, F. Wang, Y. Xiao, S. Huang, X. Tan, N. Huang, B. Jiao, C. Cheng, Y. Zhang, A. Luo, L. Mombaerts, J. Jin, Z. Cao, S. Li, H. Xu, and Y. Yuan, "An interpretable mortality prediction model for COVID-19 patients," *Nature Machine Intelligence*, vol. 2, no. 5, 2020.

[4] J. L. Reeve and P. J. Twomey, "Consider laboratory aspects in developing patient prediction models," 2021.

[5] D. R. Giacobbe, "Clinical interpretation of an interpretable prognostic model for patients with COVID-19," 2021.

[6] P. Habbu, A. Kayyum shaikh, and V. Deshmukh, "An Interpretable Mortality Prediction Model for COVID -19 Patients in Solapur- Maharashtra," *International Journal of Pharmaceutical Sciences Review and Research*, vol. 66, no. 1, 2021.

[7] H. Ko, H. Chung, W. S. Kang, C. Park, D. W. Kim, S. E. Kim, C. R. Chung, R. E. Ko, H. Lee, J. H. Seo, T. Y. Choi, R. Jaimes, K. W. Kim, and J. Lee, "An artificial intelligence model to predict the mortality of COVID-19 patients at hospital admission time using routine blood samples: Development and validation of an ensemble model," *Journal of Medical Internet Research*, vol. 22, no. 12, 2020.

[8] I. Huang and R. Pranata, "Lymphopenia in severe coronavirus disease-2019 (COVID-19): Systematic review and meta-analysis," 2020.

[9] M. Kong, H. Zhang, X. Cao, X. Mao, and Z. Lu, "Higher level of Neutrophil-to-Lymphocyte is associated with severe COVID-19," *Epidemiology and Infection*, 2020.

[10] X. Zhao, K. Wang, P. Zuo, Y. Liu, M. Zhang, S. Xie, H. Zhang, X. Chen, and C. Liu, "Early decrease in blood platelet count is associated with poor prognosis in COVID-19 patients—indications for predictive, preventive, and personalized medical approach," *EPMA Journal*, vol. 11, no. 2, 2020.

[11] C. Sun, S. Hong, M. Song, H. Li, and Z. Wang, "Predicting COVID-19 disease progression and patient outcomes based on temporal deep learning," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, 2021.

[12] D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, and Z. Peng, "Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China," *JAMA - Journal of the American Medical Association*, vol. 323, no. 11, 2020.

[13] M. Kermali, R. K. Khalsa, K. Pillai, Z. Ismail, and A. Harky, "The role of biomarkers in diagnosis of COVID-19 – A systematic review," 2020.

[14] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, 2002.

[15] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification And Regression Trees*. Routledge, 10 1984.

[16] Q. Ruan, K. Yang, W. Wang, L. Jiang, and J. Song, "Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China," 2020.

[17] P. Xu, Q. Zhou, and J. Xu, "Mechanism of thrombocytopenia in COVID-19 patients," 2020.

[18] H. Long, L. Nie, X. Xiang, H. Li, X. Zhang, X. Fu, H. Ren, W. Liu, Q. Wang, and Q. Wu, "D-Dimer and Prothrombin Time Are the Significant Indicators of Severe COVID-19 and Poor Prognosis," *BioMed Research International*, vol. 2020, 2020.