

Indian Journal of Modern Research and Reviews

This Journal is a member of the 'Committee on Publication Ethics'

Online ISSN:2584-184X



Research Article

Modeling The Factors Affecting of Outcome COVID-19 Patients, Based on Artificial Intelligence Networks

Rusul F. Abedi

College of Health and Medical Techniques, Kufa, Al-Furat Al-Awsat Technical University (ATU)

Corresponding Author: * Rusul F. Abedi

DOI: <https://doi.org/10.5281/zenodo.23337845>

Abstract

Background: The coronavirus epidemic, defined as (COVID-19, firstly appeared in China (Wuhan city) in Dec. month 201 and rapidly spread worldwide. Therefore, accurate information about all clinical and non-clinical features of this epidemic has become important for societies and countries. The goal of this research is to investigate clinical, demographic, and laboratory variables of COVID-19 and their influence on patient response (death/recovery) by using robust statistical methods to solve the multicollinearity problem among independent variables. **Methods:** This study was based on a randomly selected sample of 200 patients using the RAND () function in Python, selected from the Zenodo database for two New York hospitals in 2022. We proposed the use of Neural Network Logistic Regression compared with Ridge Logistic Regression to solve problems in the data, such as multicollinearity. **Results:** In assessing the effects of variables on the clinical outcome of COVID-19 patients (death or discharge), it was found that the factors (Age, EOS PCT, LYM PCT, MONO PCT, WBC, MCHC, LUC ABS, Glucose Random, Protein Total, BASO PCT, K Bld, NEU ABS) exhibited statistically significant effects with $0.0001 < p\text{-value} < 0.05$, whereas the remaining factors showed no statistically significant effects ($p\text{-value} > 0.05$). **Conclusion:** The proposed method demonstrated high efficiency in predicting patient outcomes, achieving an accuracy of 0.85 and an area under the curve (AUC) of 0.95. In comparison, the logistic regression method achieved an accuracy of 0.825 and an AUC of 0.817.

Manuscript Information

- ISSN No: 2584-184X
- Received: 07-06-2026
- Accepted: 10-07-2026
- Published: 30-09-2026
- MRR:4(9); 2026: 334-341
- ©2026, All Rights Reserved
- Plagiarism Checked: Yes
- Peer Review Process: Yes

How to Cite this Article

Abedi R.F. Modeling the Factors Affecting of Outcome COVID-19 Patients, Based on Artificial Intelligence Networks. Indian J Mod Res Rev. 2026;4(9):334-341.

Access this Article Online



www.mrrjournal.in

Keywords: COVID-19, Ridge Logistic Regression, ROC Curve, Neural Network Analysis

1. INTRODUCTION

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), presents with a wide spectrum of outcomes, from asymptomatic infection to acute respiratory distress syndrome, multi-organ failure, and death. This variability reflects differences in how individual host physiological systems respond to the virus, particularly the immune, metabolic and electrolyte-regulating systems, and not only differences in viral exposure.

Age is the most consistent determinant of outcome. In hospitalized adults, older age was independently associated with in-hospital mortality [1]. Physiologically, aging is accompanied by immune senescence, a chronic low-grade inflammatory state, and reduced functional reserve of the lungs, heart, and kidneys, all of which limit the ability to tolerate the systemic stress of severe infection.

The circulating leukocyte profile provides a direct window into the host immune response. Patients with severe COVID-19 tend to show an elevated total white blood cell (WBC) count and absolute neutrophil count (NEU ABS), together with a reduced lymphocyte percentage (LYM PCT), a pattern that reflects an exaggerated innate response and impaired adaptive immunity [2].

Metabolic and biochemical variables also influence outcome. Elevated glucose has been shown to favor SARS-CoV-2 replication and to promote monocyte inflammatory responses through a glycolysis-dependent mechanism [3], so random blood glucose may reflect both pre-existing metabolic disease and stress-induced hyperglycemia. Total serum protein, which is mainly albumin and globulin reflects nutritional status, hepatic synthetic function.

Artificial neural network can model such complex interactions among demographic, hematological and biochemical variables. Zhu JS (2020) studied deep-learning artificial intelligence analysis of clinical variables to predict mortality in COVID-19 patients. The study consisted of 181 hospitalized patients with confirmed COVID-19 infection in Wuhan, China. The data were split into 85% training and 15% testing sets. Prediction performance was compared with that of severity scores. The deep-learning prediction model and the resultant risk stratification score may prove useful in clinical decision-making under time-sensitive conditions [4].

Lin JK. et al. (2021) used an artificial neural network (ANN) to predict the mortality of COVID-19 patients using routine blood samples at hospital admission in Tunisia and compared ANN and CNN in terms of model performance at admission time. The results showed that ANN yielded high efficiency with an

AUC of 0.96 compared with 0.91 from CNN based on testing data [5]. Devi SKC et al. (2024) proposed a ridge logistic regression model to obtain robust predictions of COVID-19 mortality in India and to solve the multicollinearity problem among explanatory variables. They relied on criteria such as MSE, RMSE, and the R^2 score and concluded that poor feature selection led to low model generalization [6].

Saeed Reza M. et al. (2025) conducted a systematic review of 36 research papers involving artificial intelligence techniques for diagnosing COVID-19 prognosis using CT or chest X-ray images. The models showed sensitivity values of 71% for mortality, 88% for severity assessment, and 67% for oxygen requirement prediction [7].

This research differs from previous studies that focused on determining whether a person was infected with COVID-19 (positive/negative), especially at the beginning of the pandemic in December 2019. Our research focuses on the outcomes of infected patients (recovered or died), considering a range of explanatory variables expected to influence mortality.

The goal of this research was to develop a deep-learning artificial intelligence model (Neural Networks) with Ridge regularization to study the factors influencing patient response and to solve the multicollinearity problem among independent variables to obtain reliable determinants of demographic, clinical, and laboratory features affecting response (death/discharge)

2. Study Design

2.1 Data collection

Our COVID-19 patient data were randomly selected using the RAND () function based on the patient from the dataset. sample size consisted of 200 patients selected from 1,000 records", because some patients from (1500) did not have complete information regarding outcomes, demographics, and biomarker measurements. The data were obtained from New York (two hospitals) as: State University of New York (SUNY) Downstate Health Sciences University and Maimonides Medical Centre (Lambert B. et al. [8]).

2.2 Variables definition

The response is binary variable as (Death=0, Discharged=1) and exploratory variables (37 independent variables); Demographics (Age, Sex), Comorbidities (Binary: Yes/No), Laboratory Parameters as:

Liver Function Markers, Kidney Function and Electrolytes, Glucose Metabolism, Hematological Parameters, Red Blood Cell Indices, White Blood Cell Indices, Platelet Indices. Table structure retained as original; formatting corrected only where necessary, as following:

Table 1: Collected laboratory data

Predicate Variable	Unit / Description
Age	years
Sex	male/female
Hypertension	Yes/No
diabetes	Yes/No
Cancer	Yes/No
Copd	Yes/No
AST	U/L (indicator of liver diseases like cirrhosis)
Albumin	g/dL (The most abundant protein in the blood)
Alkaline Phosphatase	U/L (Determines liver and bone diseases)
BASO PCT	% (Percentage of basophils)
BUN	mg/dL (Blood Urea Nitrogen)
Bilirubin Total	mg/dL (Produced during hemolysis, indicates liver disease)
CO2	mmol/L (Evaluates the measure of CO2)
Calcium Level	mg/dL (Mineral indicator of liver, muscle, kidney)
Creatinine	mg/dL (A byproduct of creatine, Kidney function indicator)
EOS PCT	% (Eosinophil percentage)
Glucose Random	mg/dL (The presence of glucose in the blood/urine)
HCT	% (red blood cells proportion in blood)
HGB	g/dL (Hemoglobin concentration in blood)
K Bld (Potassium)	mmol/L (Potassium level in blood)
LUC ABS	10 ³ /μL (Absolute large unstained cells)
LUC PCT	% (Percentage of large unstained cells)
LYM ABS	10 ³ /μL (Absolute lymphocyte count)
LYM PCT	% (Lymphocyte percentage)
MCH	pg (hemoglobin average in RBCs)
MCHC	g/dL (hemoglobin concentration average in RBCs, anemia type)
MCV	fL (Mean cell volume of RBCs)
MONO ABS	10 ³ /μL (Absolute monocyte count)
MONO PCT	% (Monocyte percentage)
MPV	fL (Average volume of platelets)
NEU ABS	10 ³ /μL (Absolute neutrophil count)
NEU PCT	% (Neutrophil percentage)
Na Bld (Sodium)	mmol/L (levels of Sodium may indicate kidney or heart function)
PLAT	10 ³ /μL (Platelet count in blood)
Protein Total	g/dL (Total protein in blood)
RBC	10 ⁶ /μL (count of red blood cells, low levels imply anemia and bleeding)
RDW	% (Red cell distribution width)
Outcome	Response variable (Death=0, Discharged=1)

3. Statistical Analysis

The response variable in our study represents the patient's end-time outcome (Death = 0, Discharge = 1) and is binary (0,1). In such cases, the most appropriate model for data analysis is logistic regression [9,10].

Logistic regression models require several assumptions, including linearity in the logit, independence of observations, absence of multicollinearity, and a sufficiently large sample size. If standard logistic regression is used in the presence of

multicollinearity, the model may fail (e.g., unstable odds ratios).

To address this issue, Robust Statistical Methods such as Neural Network Logistic Regression and Multivariate Ridge Logistic Regression were applied [11,12]. The results were obtained by writing program code in Python version 3.13.2 using libraries such as Scikit-learn, SciPy, and Pandas to conduct statistical data analysis [13,14], you can see the python codes in follow QR



3. Methods

3.1.1 Ridge logistic regression

To address multicollinearity among predictors, we applied Ridge logistic regression, which incorporates an L2 regularization penalty into the standard logistic regression framework. This penalty shrinks regression coefficients toward zero, reducing variance and preventing unstable or extreme odds ratio estimates, with the maximum number of iterations set to 5,000.

Ridge regression retains all variables in the model, ensuring that clinically relevant predictors are not excluded. The model was trained on standardized predictors, with 80% of the data used as the training set and 20% as the test set (seed = 10) [6,15].

3.1.2 Neural Network Logistic Regression

Neural networks can automatically learn weights without requiring strict assumptions regarding multicollinearity. The model used logistic activation with the following structure [16]:

- Input Layer: all 37 predictors (demographic, clinical, and laboratory)
- Hidden Layer: 10 neurons with logistic activation
- Output Layer: One neuron (probability of death vs. discharge)

As following figures

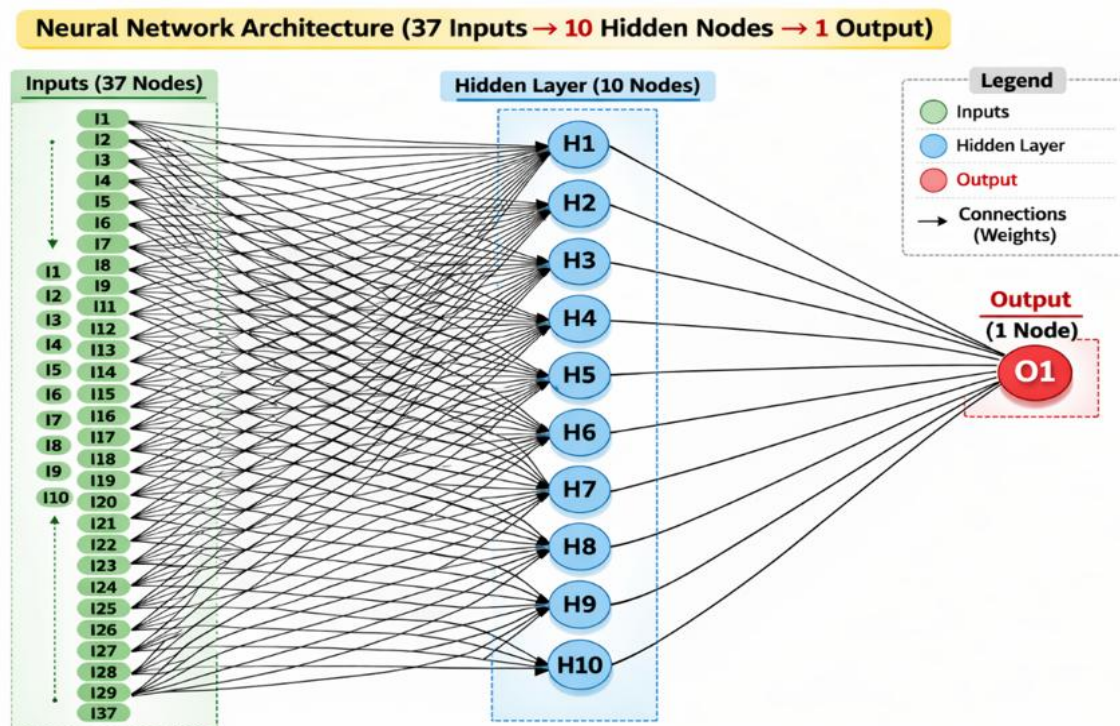


Fig. 1: Diagram of Neural Network Logistic

3.1.3 Performance Criteria

The basic criteria in the statistical analysis of any data are descriptive statistics, which are presented in table (1), and we relied on some model performance evaluation measures as:

- **Accuracy** measures the percentage of cases in which the model correctly predicted either discharge or death out of all cases evaluated.
- **Receiver Operating Characteristic (ROC) Curve** is a graphical representation of a model's diagnostic ability. It plots: True Positive Rate (**Sensitivity**) on the y-axis, False Positive Rate (**1 – Specificity**) on the x-axis.
- **Area Under the Curve (AUC)** refers to the area under the ROC; an AUC value closer to 1 indicates excellent predictive performance, and if 0.5 suggests performance no better than random chance.

- **F-measure (F1-score)** is a metric that combines precision and recall into a single value, providing a balanced assessment of model performance.
- **Statistical Significance** was determined using a threshold of **P-value < 0.05**, meaning that results with p-values below 0.05 were considered statistically significant. [17].

Also, in network analysis, we can rely on the importance of input variables in the network to quantify the contribution of each input variable to the model's output. Therefore, several post hoc interpretability methods—such as Shapley-value-based approaches like SHAP—return signed contribution values per variable as follows: Positive attribution: the variable increases the model's output; Negative attribution: the feature decreases the model's output Jacopo T. et al,2023[18].

3.1. Results

3.2.1 Descriptive Statistics

We can display descriptive statistics and p-value of proportion test as following table.

Table (2): Descriptive Statistics and test of variables

Variable	Yes	No	P. value
Hypertension	110 (55%)	90 (45%)	0.044*
Diabetes	78 (39%)	122 (61%)	<0.0001**
Cancer	13 (7%)	187 (93%)	<0.0001**
COPD	14 (7%)	186 (93%)	<0.0001**
Outcome	Death	Discharge	
	82 (41%)	118 (59%)	0.001**
Gender	Male	Female	
	114 (57%)	86 (43%)	0.005 **

* Low significant, ** high significant

All variables in table 2), showed significance (p-value< 0.05). The strongest statistically significant variables were:

Diabetes, Cancer, COPD

Hypertension was statistically significant but weaker (p = 0.044).

Table (3): Descriptive Statistics (Variable*Outcome) testing

Variable	Outcome		p-value (Cate.)	p-value (V.*Out.)
	Categorical	Death	Discharge	
Gender	Male	57%	43%	0.035*
	Female	60%	40%	0.01*
Hypertension	Yes	69%	31%	<0.0001**
	No	50%	50%	1 ns
Cancer	Yes	60%	40%	0.13ns
	No	46%	54%	0.12ns
COPD	Yes	59%	41%	0.34ns
	No	50	50	1ns

*Is low significant< 0.05, ** is high significant < 0.01, ns No significant

From Table (3), significant differences were found between outcome and certain categorical variables (Sex and Hypertension). No significant differences were observed for other categorical variables.

3.2.2 Ridge Logistic Regression

Before applying logistic models we tested the multicollinearity problem between independent variables in data by VIF criteria (**Variance Inflation Factor**) [19], and we got the following results in table (4).

Table (4): Variance Inflation Factor (VIF) for predictor variables (10)

no	Variable	VIF	no	Variable	VIF
1	MCH	29.7	6	NEU PCT	8.7
2	MCV	13.9	7	LYM PCT	7.2
3	MCHC	12.6	8	NEU ABS	6.5
4	HCT	11.8	9	RBC	6.3
5	HGB	11.64	10	WBC	5.8
All other variables are acceptable					< 5

From above results we note that some variables have the multicollinearity problem because VIF>5, to solve this problem we can depend on remove these variables or depend on robust method as:(*Ridge regression, Neural Network Logistic, Robust Logistic*) [11,12].

From analysis our data by Ridge Logistic Regression we got summary results as:

- a. **High Importance Variables:** if (OR ≥ 2 or OR ≤ 0.5 and p-value < 0.01)
- Strong risk factors
Age (OR = 2.24, p < 0.001), LUC ABS (OR = 2.74, p = 0.005),
WBC (OR = 1.97 ≈ 2, p < 0.001)

- **Powerful protective agents**
EOS PCT (OR = 0.30, p < 0.001), MONO PCT (OR = 0.34, p < 0.001)
LYM PCT (OR = 0.45, p < 0.001)
- b. **Moderate Importance Variables:** if (1.5 ≤ OR < 2 or 0.5 < OR ≤ 0.7 and p < 0.05)
Glucose Random (OR = 1.77, p = 0.035) , Protein Total (OR = 1.94, p = 0.040)
CO₂ (OR = 0.52, p = 0.005), BASO PCT (OR = 0.39, p = 0.045)
MCHC (OR = 0.68, p = 0.050 – borderline)
- c. **Low or Non-significant Variables:** because all other variables have p-value>0.05

Table (5): Estimators of Ridge Logistic method

Variable	Coefficient	Odds Ratio (OR)	p-value
Age	0.807	2.24	<0.001**
EOS PCT	-1.202	0.30	<0.001**
LYM PCT	-0.793	0.45	<0.001**
MONO PCT	-1.092	0.34	<0.001**
WBC	0.68	1.97	<0.001**
CO2	-0.659	0.52	0.005**
LUC ABS	1.008	2.74	0.005**
Glucose Random	0.572	1.77	0.035*
Total Protein	0.665	1.94	0.040*
BASO PCT	-0.934	0.39	0.045*
MCHC	-0.391	0.68	0.050*
all other variables p-value>0.05			No significant effect

*Is low significant < 0.05, ** is high significant < 0.01

3.2.3 Neural Network Logistic Regression

We relied on a neural network classifier using the MLP Classifier function from scikit-learn in Python, as shown in the previous figure [Fig. (1)], with a maximum of 5,000 iterations. The sample was divided into training (80%) and testing (20%) subsets using stratified sampling with seed = 10 for

reproducibility. The analysis of neural networks in regression models depends on nonlinear models; therefore, the p-value may not be identical to the p-value of other methods. Thus, it is possible to rely on the importance criterion that provides the importance of each variable, as shown in the following table:

Table (6): Estimators of Neural Network Logistic method

Rank	Feature	Mean(SHAP)	p-value	Significant
1	LYM PCT	0.07286	9.77E-06	TRUE
2	NEU ABS	0.06007	6.83E-03	TRUE
3	age	0.04931	2.48E-02	TRUE
4	NEU PCT	0.02680	4.49E-03	TRUE
5	hypertension	0.02007	2.68E-03	TRUE
6	MONO PCT	0.01650	5.59E-04	TRUE
7	RDW	0.01327	1.77E-17	TRUE
8	CO2	0.01057	4.05E-06	TRUE
9	MPV	0.00848	2.40E-04	TRUE
10	PLAT	0.00776	5.31E-11	TRUE
11	MCHC	0.00570	3.88E-13	TRUE
12	diabetes	0.00202	2.98E-07	TRUE
13	Alkaline Phosphatase	0.00057	1.84E-03	TRUE
14	HGB	0.00014	3.08E-03	TRUE
15	LUC PCT	0.00014	3.68E-05	TRUE
16	MCH	0.00005	1.96E-21	TRUE
17	Na Bld	0.00004	1.49E-07	TRUE

*is low significant < 0.05, ** is high significant<0.01

The previous table shows the 17 most important variables (features) influencing the patient's response (death or recovery), ranked according to their importance (mean SHAP). When we compared these results with the ridge logistic regression results (binary case) in Table (5), it is clear that most of the variables had the same significant impact in the Neural Network Logistic model.

3.2.4 Comparison between Ridge Logistic and Neural Network logistic

We can present all performance criteria for the two statistical methods as follows:

Table (7): Comparison between Ridge logistic and Neural Network Logistic

Criteria	Ridge Logistic	Neural Network
Accuracy	0.825	0.85
AUC	0.817	0.951
MSE	0.175	0.150
Confusion Matrix	[[20, 3], [4, 13]]	[[19, 4], [2, 15]]

In ridge logistic regression, we found that of the 23 actual discharges, the model correctly predicted 20 but missed 3 (classified as death); and of the 17 actual deaths, the model

correctly predicted 13 but misclassified 4 as discharge, with 82.5% accuracy overall. It was particularly strong at predicting

deaths (precision = 81.7%), which is clinically important. MSE = 0.175.

In the Neural Network model, we found that of the 23 actual discharges, the model correctly predicted 19 but missed 4 (classified as death); and out of the 17 actual deaths, the model

correctly predicted 15 but misclassified 2 as discharge, with 85% accuracy overall. It was particularly strong at predicting deaths (precision = 95.1%), which is clinically important. MSE = 0.150

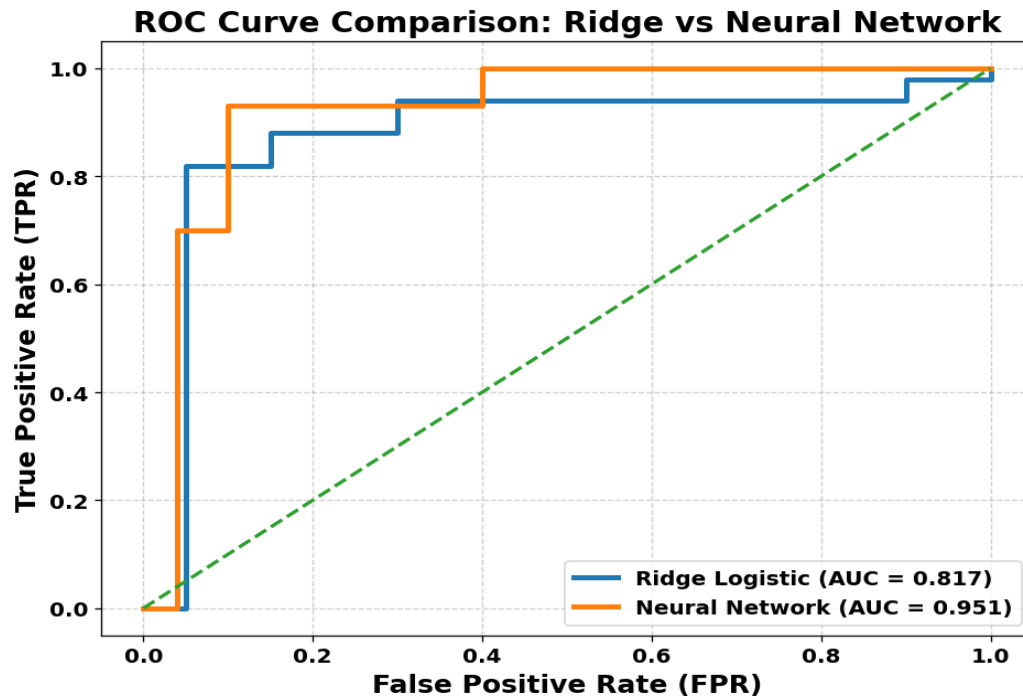


Fig. (2): ROC curve Comparison between Ridge and Neural Network

From the above comparison, we noted that the Neural Network is a good method for analysing COVID-19 data.

4. DISCUSSION

From the results in Table (5), the variables (Age, EOS PCT, LYM PCT, MONO PCT, WBC, MCHC, LUC ABS, Glucose Random, Protein total) were associated with an increased event (death) of the patient because all odds ratios ($OR > 1.00$), and some variables such as (EOS PCT, LYM PCT, MONO PCT, CO₂, BASO PCT, MCHCH) were associated with a decreased event (death), with all odds ratios ($OR < 1.00$). These two groups had statistically significant effects in ridge logistic regression since $p\text{-value} < 0.05$; however, all other variables were not significant ($p\text{-value} > 0.05$).

From the Neural Network logistic regression (Table 6), we showed that some variables (top 17) were more important in model performance according to the mean absolute (SHAP) criterion and also had a significant effect because $p\text{-value} < 0.05$, while the remaining factors showed no statistically significant effects ($p\text{-value} > 0.05$). When we compared our methods (Ridge logistic regression and Neural Network logistic regression) (see Table 7 and Fig. 2), we found that the Neural Network had high efficiency in predicting patient outcomes, achieving an accuracy of 0.85, an area under the curve (AUC =

0.95), and MSE = 0.150. In comparison, the logistic regression method achieved an accuracy of 0.825, an AUC of 0.817, and MSE = 0.175.

Therefore, we suggested using Neural Network logistic regression in analyzing binary responses (0,1).

Our findings are consistent with prior mortality-focused COVID-19 studies showing that advanced age is a strong independent predictor of death, as reported by Rahman A. et al. [20]. Elevated WBC counts and reduced lymphocyte percentages have likewise been associated with non-survivor groups in cohorts such as Al-Shajlawi M. et al. (2022) [21]. The observed link between hyperglycemia and adverse outcomes also agrees with metabolic risk analyses published by Springer Nature (2022). Overall, our results reinforce existing evidence that age, inflammatory dysregulation, and immune imbalance are major determinants of mortality in COVID-19 patients.

5. CONCLUSIONS

In our study we focused on predicting the outcome of infected patients (recovered or died) using demographic, clinical, and laboratory variables. Because many explanatory variables were interrelated (37 predictors), multicollinearity was present. This issue was successfully addressed using robust methods: Ridge Logistic Regression and Neural Network Logistic Regression.

The Neural Network model demonstrated superior performance across all evaluation criteria.

Descriptive analysis showed a mortality rate of 41% and a recovery rate of 59%, with statistically significant differences ($p < 0.001$). Gender differences were statistically significant but weaker ($0.01 < p\text{-value} < 0.05$).

Future research may consider applying deep neural decision forest models to similar medical prediction problems (e.g., liver or heart diseases).

REFERENCES

- Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395(10229):1054–62.
- Qin C, Zhou L, Hu Z, et al. Dysregulation of immune response in patients with coronavirus 2019 (COVID-19) in Wuhan, China. *Clin Infect Dis*. 2020;71(15):762–8.
- Codo AC, Davanzo GG, Monteiro LB, et al. Elevated glucose levels favor SARS-CoV-2 infection and monocyte response through a HIF-1 α /glycolysis-dependent axis. *Cell Metab*. 2020;32(3):437–46.
- Zhu JS, Peilin G, Yong Z, Zirun Z, Duong TQ. Deep-learning artificial intelligence analysis of clinical variables predicts mortality in COVID-19 patients. *JACEP Open*. 2020. doi:10.1002/emp2.12205.
- Lin JK, Chien TW, Wang LY, Chou W. An artificial neural network model to predict the mortality of COVID-19 patients using routine blood samples at hospital admission: development and validation study. *Medicine*. 2021;100(28). doi:10.1097/MD.00000000000026532.
- Devi SKC, Ponnusamy S, Ramesh Kumar P, Mishra KK, Tadesse AM, Abbas SH. Robust prediction of COVID-19 mortality with ridge regression and hyperparameter optimization. In: *Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024)*; 2024. doi:10.2139/ssrn.5083709.
- Saeed Reza M, Negin C, Sadra M, et al. Prognosis of COVID-19 using artificial intelligence: a systematic review and meta-analysis. *Front Biomed Technol*. 2025;12(3):658–83.
- Lambert B, Stopard IJ, Momeni-Boroujeni M, Mendoza R, Zuretti A. A dataset of anonymized hospitalized COVID-19 patient data: outcomes, demographics and biomarker measurements for two New York hospitals [dataset]. *Zenodo*; 2022. doi:10.5281/zenodo.6771834.
- Menard S. Standards for standardized logistic regression coefficients. *Soc Forces*. 2011;89(4):1409–28.
- Hosmer DW, Lemeshow S, Sturdivant RX. *Applied logistic regression*. 3rd ed. Hoboken (NJ): Wiley; 2013. doi:10.1002/9781118548387.
- Bianco AM, Yohai VJ. Robust estimation in the logistic regression model. In: *Reider H, editor. Robust statistics, data analysis, and computer intensive methods*. Wiley; 1996. p. 17–34. doi:10.1007/978-1-4612-2380-1_2.
- Ahmed I, Cheng W. The performance of robust methods in logistic regression model. *Open J Stat*. 2020; 10:127–38. doi:10.4236/ojs.2020.101010.
- Topuz S. Machine learning with scikit-learn: predicting disease outcomes. *J Biomed Inform*. 2019; 94:103–20.
- Géron A. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. Sebastopol (CA): O'Reilly Media; 2019.
- Yerushalmy J. Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques. *Public Health Rep*. 1947; 62:1432–49.
- Mostafa S, Raoof N, Arpanahi HK. Design of an artificial neural network to predict mortality among COVID-19 patients. *Inform Med Unlocked*. 2022; 31:100983.
- Cabot JH, Ross EG. Evaluating prediction model performance. *Surgery*. 2023;174(3):723–6. doi:10.1016/j.surg.2023.05.023.
- Jacopo T, Beepul B, Yaniv R, Jeremias S. SHAP-XRT: the Shapley value meets conditional independence testing. *Trans Mach Learn Res*. 2023. doi:10.48550/arXiv.2207.07038.
- Montgomery DC, Peck EA, Vining GG. *Introduction to linear regression analysis*. 5th ed. Hoboken (NJ): Wiley; 2012.
- Sharma J, Rajput R, Bhatia M, Arora P, Sood V. Clinical predictors of COVID-19 severity and mortality: a perspective. *Front Cell Infect Microbiol*. 2021; 11:674277. doi:10.3389/fcimb.2021.674277.
- Al-Shajlawi M, Alsayed AR, Abazid H, Awajan D, Al-Imam A, Basheti I. Using laboratory parameters as predictors for the severity and mortality of COVID-19 in hospitalized patients. *Pharm Pract*. 2022;20(3):2721. Available from: <https://pubmed.ncbi.nlm.nih.gov/36733508/>.

Creative Commons License

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution–Non-commercial–No Derivatives 4.0 International (CC BY-NC-ND 4.0) License. This license permits users to copy and redistribute the material in any medium or format for non-commercial purposes only, provided that appropriate credit is given to the original author(s) and the source. No modifications, adaptations, or derivative works are permitted.

About the Author

Rusul F. Abedi is affiliated with the College of Health and Medical Techniques, Kufa, Al-Furat Al-Awsat Technical University (ATU). Their academic interests include health sciences, medical technologies, clinical practices, and healthcare education, with a focus on advancing medical knowledge and supporting improvements in healthcare services through academic research and professional development.