Original Article

Development of neural network models for prediction of the outcome of COVID-19 hospitalized patients based on initial laboratory findings, demographics, and comorbidities

Mirza Pasic[1], Edin Begic[2,3], Faris Kadic[4], Ali Gavrankapetanovic[5], Mugdim Pasic[1]

[1]Department of Industrial Engineering and Management, Mechanical Engineering Faculty, University of Sarajevo, 71000 Sarajevo, [2]Department of Cardiology, General Hospital „Prim. dr. Abdulah Nakaš“, 71000 Sarajevo, Bosnia and Herzegovina Department of Pharmacology, Faculty of Medicine, University School of Science and Technology, 71000 Sarajevo, [3]Department of Cardiology, General Hospital „Prim. dr. Abdulah Nakaš“, 71000 Sarajevo, Bosnia and Herzegovina Gavrankapetanovic - Department of Surgery, General Hospital „Prim. dr. Abdulah Nakaš“, 71000 Sarajevo, Departments of [4]Internal Medicine and [5]Surgery, General Hospital "Prim.Dr. Abdulah Nakaš", Sarajevo, Bosnia and Herzegovina

Abstract

Background: During the process of the treatment of COVID-19 hospitalized patients, physicians still face a lot of unknowns and problems. Despite the application of the treatment protocol, it is still unknown why the medical status of a certain number of patients worsens and ends with death. Many factors were analyzed for the prediction of the clinical outcome of the patients using different methods. The aim of this paper was to develop a prediction model based on initial laboratory blood test results, accompanying comorbidities, and demographics to help physicians to better understand the medical state of patients with respect to possible clinical outcomes using neural networks, hypothesis testing, and confidence intervals. **Methods:** The research had retrospective-prospective, descriptive, and analytical character. As inputs for this research, 12 components of laboratory blood test results, six accompanying comorbidities, and demographics (age and gender) data were collected from hospital information system in Sarajevo for each patient from a sample of 634 hospitalized patients. Clinical outcome of the hospitalized patients, survival or death, was recorded 30 days after admission to the hospital. The prediction model was designed using a neural network. In addition, formal hypothesis tests were performed to investigate whether there were significant differences in laboratory blood test results and age between patients who died and those who survived, including the construction of 95% confidence intervals. **Results:** In this paper, 11 neural networks were developed with different threshold values to determine the optimal neural network with the highest prediction performance. The performances of the neural networks were evaluated by accuracy, precision, sensitivity, and specificity. Optimal neural network model evaluation metrics are: accuracy = 87.78%, precision = 96.37%, sensitivity = 90.07%, and specificity = 62.16%. Significantly higher values ($P < 0.05$) of blood laboratory result components and age were detected in patients who died. **Conclusion:** Optimal neural network model, results of hypothesis tests, and confidence intervals could help to predict, analyze, and better understand the medical state of COVID-19 hospitalized patients and thus reduce the mortality rate.

Keywords: Classification, covid-19, machine learning, neural networks, prediction

Address for correspondence: Prof. Mirza Pasic, Mechanical Engineering Faculty, University of Sarajevo, 71000 Sarajevo, Bosnia and Herzegovina.
E-mail: mirza.pasic@mef.unsa.ba

Received: 15-01-2022 **Revised:** 07-03-2022
Accepted: 10-03-2022 **Published:** 30-08-2022

Access this article online

Quick Response Code:

Website: www.jfmpc.com

DOI: 10.4103/jfmpc.jfmpc_113_22

How to cite this article: Pasic M, Begic E, Kadic F, Gavrankapetanovic A, Pasic M. Development of neural network models for prediction of outcome of COVID-19 hospitalized patients based on initial laboratory findings, demographics and comorbidities. J Family Med Prim Care 2022;11:4488-95.

Introduction

Coronavirus disease-19 (COVID-19) pandemic is associated with high morbidity and mortality,[1] and initial recognition of clinical deterioration is imposed as an imperative in everyday clinical work. It occurs as a consequence of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS CoV-2) infection and has an effect on all organ systems.[1-3] COVID-19 is characterized by a systemic inflammatory process on the one hand, and on the other hand, it is characterized by hypercoagulability. The complete spectrum of laboratory findings, clinical picture, physical examination, anamnestic data, as well as life habits of the patient, is important for the physician in the optimization and deciding on the best treatment for the patient. The question is whether the identification of patients with potential for a bad outcome can be done at the level of the initial admission or whether the worsening of clinical picture can be predicted at this level. The question also arises whether oxygen saturation should be the only reference point for patient admission since it is variable and depends first on the patient's constitution and then on the physical activity before the examination. In addition to the confirmation of the infection, on the territory of Bosnia and Herzegovina, basic laboratory findings are performed and in accordance with the clinical condition of the patient, the existence of comorbidities, and chest X-rays findings, patients received therapy. The use of corticosteroid therapy, and interleukin-6 (IL-6) inhibitors in COVID-19, placed even more emphasis on adequate triage of patients and indicated the need for early recognition of deterioration of the clinical picture.[4] Modern clinical practice is keeping pace with the use of information technology, while neural networks are gaining their clinical significance and becoming part of the diagnostic, but also a present tool to improve the treatment of the patients.[5] The COVID-19 treatment should be based on a symbiosis of laboratory findings, clinical picture, the existence of comorbidities, and time interval of initiation of adequate therapy, all with the aim of preventing lethal outcomes.[6] Treatment of laboratory parameters should not be the primary aim, especially after the acute phase of the disease, but they can give orientation to the clinicians in working with patients. COVID-19 infection is with extremely dynamic changes in clinical and laboratory parameters. Among the most important are inflammatory parameters, haemogram, haemostasis parameters, general hydration status, renal function, which indicate generalized or local hypoxia, and the degree of the inflammatory process (prescribed pharmacological treatment should be taken into consideration, especially since in practice, corticosteroid therapy in the early stages of the disease have been prescribed, which is essentially is not a practice that to be forced). Given the large influx of patients, it is necessary to stratify the patient at the very beginning because the delay in treatment is a negative prognostic factor on the outcome of the patient. Differences in mean age values, erythrocyte, haemoglobin, haematocrit, platelet, leukocyte, neutrophil, lymphocyte, monocyte, basophil, eosinophil, C-reactive protein (CRP), and D-dimer values between survivors and those who survived were verified. At the very beginning of the pandemic, the focus was on the appearance of comorbidities, but as time went on, the impact of diabetes mellitus was still potentiated, while the appearance of other pathology was only a part that contributed to the mosaic of the patient's therapeutic modality. The use of information technologies or artificial neural networks helps us in the synthesis of a large amount of data and allows us to try to understand at the beginning of disease which patient could actually get worse, or which patient could develop a difficult clinical picture. The aim of this paper was to develop a neural network classification model to predict the outcome (survival or death) of hospitalized patients diagnosed with COVID-19 based on initial laboratory findings, comorbidities, and demographics. In addition, hypothesis tests were performed to compare all values of laboratory findings and age between patients who survived and who did not survive COVID-19. Finally, confidence intervals for all values of laboratory findings and age for patients who survived and who died were calculated. The hypothesis of the research was that neural networks could be used to predict the outcome of COVID-19 hospitalized patients with high accuracy using laboratory test results, comorbidities, and demographics as inputs, and in combination with hypothesis testing and confidence intervals could help to better understand the medical state of the patients and contribute to better prediction of the outcome. Also, the network can be a guide for the family physicians, who should refer the patient to a higher level of health care and thus help relieve the burden on hospital capacity.

Materials and Methods

Dataset

This research had a retrospective-prospective, descriptive and analytical character and included patients who were hospitalized in the General Hospital "Prim. dr. Abdulah Nakas" in Sarajevo, Bosnia, and Herzegovina during the period from 01 Sep 2020 to 01 May 2021. From the hospital information system, which was used in clinical work, laboratory parameters at admission were verified, along with demographic data and the existence of comorbidities, while the outcome (survival or death) was recorded thirty days after admission to the hospital. Inclusion criteria in this study were: patients who were polymerase chain reaction (PCR) tested positive for SARS-COV2 (verified COVID-19), patients older than 18 years, and patients who had documented values of complete blood count, differential blood count, CRP, and D-dimer values on admission. The exclusion criteria were met in case a patient had incomplete documentation or was younger than 18 years. Based on these criteria, the hospital provided data for 634 hospitalized patients who met eligibility criteria and were included in the analysis. The study was conducted in accordance with the basic principles of the Declaration of Helsinki from 2013. During the course of this study, the identity and all personal data of patients were protected in accordance with the regulations on the protection of identification data. Ethical approval was obtained from the Ethical Committee of General Hospital "Prim.dr. Abdulah Nakas", Sarajevo, Bosnia, and Herzegovina (approval number 555-50/21). Features included in this research were erythrocytes, leukocytes, haemoglobin, haematocrit, thrombocytes, neutrophil granulocytes, lymphocytes, monocytes, basophil granulocytes, eosinophil granulocytes, CRP, D-dimer, age, gender, hypertension, diabetes

mellitus, hyperlipidaemia, peripheral arterial disease, the oncological process in anamnesis, venous thromboembolism (VTE) in anamnesis and the arterial incident in anamnesis. Reference ranges for parameters were: erythrocytes 4.34–5.72 × 10^{12}/L for males, 3.86–5.08 × 10^{12}/L for females; leukocytes 3.4–9.7 × 10^{9}L for males and females; haemoglobin 137–175 g/L males and 119–157 g/L females; haematocrit 0.41–0.53% males and 0.35–0.47% females; thrombocytes 158–424 × 10^{9}L for males and females; neutrophil granulocytes 44–72%; lymphocytes 20–46%; monocytes 4–8%; basophil granulocytes 0–1%; eosinophil granulocytes 2–4%; CRP up to 5.0 mg/L and D–dimer up to 804 μg/L.

Neural networks

Neural network models were developed using Python 3.8 in Jupyter Notebook. Tensorflow 2.4.0 and Keras were used to create neural network models. Dataset was divided into two sets, training and test set. The training set consisted of 450 samples, while the test set consisted of 184 samples. One sample represents one patient with all laboratory findings, demographics, accompanying comorbidities, and outcome of the hospital treatment. Laboratory findings, accompanying comorbidities, and demographic data were used as neural network inputs, while the outcome of hospital treatment was used as neural network output. The structure of the neural network was designed with 21 neurons in the input layer, one hidden layer with 50 neurons, and one neuron in the output layer. Data were normalized between 0 and 1. Sigmoid activation function was used for both hidden and output layer with Adam optimizer. The results from the sigmoid function are between 0 and 1, but when doing classification problems, results must be either 0 or 1. Tensorflow Keras uses 0.5 as its threshold by default, but it is not necessarily the value that leads to the optimal solution. In order to determine the optimal threshold value, BinaryAccuracy metric was used during the training. BinaryAccuracy is TensorFlow Keras metric that allows to define custom threshold when calculating accuracy. In this research, threshold values from 0 to 1 with step 0.1 were used; thus, 11 neural networks were created with 11 corresponding threshold values. For each neural network, accuracy, precision, true positive rate (TP rate), and false positive rate (FP rate) were calculated, along with corresponding confusion matrices, charts, and Receiver Operating Characteristic (ROC) curves for both training and test sets. Accuracy, precision, sensitivity, and specificity were calculated using Eq. (1–4):

$$accuracy = \frac{TP + TN}{TP + TN + FN + FN} \quad (1)$$

$$precision = \frac{TP}{TP + FP} \quad (2)$$

$$sensitivity = \frac{TP}{TP + FN} \quad (3)$$

$$specificity = \frac{TN}{TN + FP} \quad (4)$$

During the learning process, custom callback function was created to prevent overfitting, which allows a minimum of 50 epochs as well as one of the following conditions: a) to stop the training process when binary accuracy of validation/binary accuracy of training set <0.9 or b) binary accuracy of validation set or binary accuracy of training set <0.1.

Hypothesis tests and confidence intervals

Hypothesis tests were done to compare the mean values of all laboratory findings and ages of patients who survived and those who did not survive COVID-19. First, F-test was performed to check whether variances were equal or unequal and after that one-tailed tests assuming equal or unequal variances were done. Hypothesis tests were performed using $\alpha = 0.05$ level of significance or 95% confidence level. The *P* value for each hypothesis test was calculated and based on the *P* value, a decision was made whether to reject or not reject the null hypotheses H_0. All results with $P < 0.05$ were considered statistically significant. The *P* value is the probability of obtaining statistic equal to or higher than the sample result, given that the null hypothesis H_0 is true.

For all laboratory findings and ages of patients who survived and patients who did not survive COVID-19, 95% confidence interval for the mean was constructed to find upper and lower limits around mean feature values.

Results

Data presentation

Out of these 634 patients, 436 (68.77%) were male, and 199 (31.23%) were female. Based on these numbers, it can be seen that the number of hospitalized male patients was 2.2 times greater than the number of hospitalized female patients. The average age of the hospitalized patients was 64.23 years. The oldest patient was 93 years old, and the youngest patient was 18 years old. Of the total number, 71 patients died [Table 1]. At the initial examination before hospitalization, laboratory parameters were analyzed [Table 2]. In addition to laboratory blood test results, data regarding hypertension, diabetes mellitus, hyperlipidaemia, peripheral arterial disease, the oncological process in anamnesis, venous thromboembolism (VTE) in anamnesis, and the arterial incident in anamnesis (acute coronary syndrome, stroke) were collected for each patient. Table 3 shows the number of patients with these comorbidities, as well as the number and percentages of survivors and death cases for all patients and for male and female patients separately who had these diseases.

Neural network models

A total of 11 neural network models were created. Each neural network model used a different threshold value from 0 to 1

Table 1: Number of survivors and death cases

	Total	Survived	Death cases	Survived (%)	Death cases (%)
Gender					
Male	436	386	50	88.53%	11.47%
Female	198	177	21	89.39%	10.61%
Total	634	563	71	88.80%%	11.20%

with 0.1 steps. The number of epochs for the training of neural network models is depicted in Table 4.

Accuracy, precision, sensitivity, specificity, TP rate, and FP rate were calculated for both training and test set, and results for training set are shown in Table 5, while results for test set are shown in Table 6. TP rate is equal to sensitivity, and FP rate is equal to 1 – specificity.

ROC curves for both training and test sets are shown in Figures 1 and 2, respectively.

In order to determine optimal neural network, two more graphs were created, where accuracy, precision, and sensitivity were plotted for both training and test set against threshold values as shown in Figures 3 and 4, respectively.

Taking into consideration data from Tables 5 and 6 as well as Figures 1-4 it can be seen that neural network with the threshold value of 0.7 had the lowest FP rate (37,84%) with high accuracy (87,78%), precision (96,37%), and sensitivity (90,07%). For the neural network model with the chosen threshold of 0.7, confusion matrices for both training and test set were created and are shown in Tables 7 and 8, respectively, where 0 represents death case, while 1 represents survivor.

From Table 7, it can be seen that the neural network correctly predicted the outcome for 394 patients and made mistakes for 56 patients. However, due to the chosen threshold value of 0.7 and the strategy to minimize FP rate with respect to other metrics, FP number is only 14 out of 450 patients from the test set. The same conclusion applies to the confusion matrix [Table 8] for the test set where only 16 out of 184 patients were FP.

Hypothesis tests

Hypothesis tests were done in order to determine if there is a statistically significant difference between mean values of laboratory test result components and age between patients who survived and died.

It can be seen that there is enough statistical evidence to conclude that the mean values of age, erythrocytes, haemoglobin, haematocrit, leukocytes, neutrophil granulocytes, CRP, and D-dimer of patients who died are higher than the mean values of these features of patients who survived.

Also, it can be seen that mean values of thrombocytes, lymphocytes, monocytes, basophil granulocytes, and eosinophil granulocytes of the patients who survived are significantly higher than the mean values of these features of patients who died.

Mean values of features for patients who survived and died, along with the H_0 and H_1 and P values are depicted in Table 9.

Formal hypothesis tests were done using $\alpha = 0.05$ level of significance. Thus, the null hypothesis was rejected if the P value was less than α. It can be seen that the null hypothesis H_0 was rejected in all cases, meaning that there is statistical evidence for the claim stated in the research hypothesis H_1 is true.

Table 2: Blood test results for all patients

All patients	Mean±Standard deviation	Maximum	Minimum
Erythrocytes ($\times 10^{12}$/L)	4.48±0.57	6.68	2.80
Haemoglobin (g/L)	133.10±16.56	195.00	54.00
Haematocrit (%)	0.40±0.05	0.59	0.20
Thrombocytes ($\times 10^{9}$/L)	293.49±123.43	1120.00	49.00
Leukocytes ($\times 10^{12}$/L)	9.39±4.29	29.20	2.50
Neutrophil granulocytes (%)	79.00±11.58	97.10	26.70
Lymphocytes (%)	14.23±9.10	66.10	1.50
Monocytes (%)	6.12±3.38	29.30	0.90
Basophil granulocytes (%)	0.23±0.19	2.10	0.00
Eosinophil granulocytes (%)	0.42±0.92	7.50	0.00
C-reactive protein (mg/L)	52.13±65.37	687.00	0.10
D-dimer (μg/L)	4224.18±14375.08	177600.00	110.00

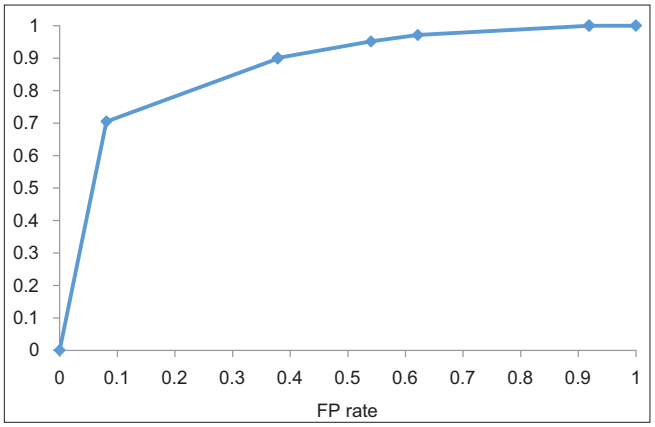

Figure 1: ROC curve for the training set

Table 3: Representation of comorbidities in relation to patient outcome

Comorbidity	Total	Survivors	Death cases	Survived (%)	Death cases (%)
Hypertension	314	258	56	82.17%	17.83%
Diabetes mellitus	130	105	25	80.77%	19.23%
Hyperlipidaemia	73	64	9	87.67%	12.33%
Peripheral arterial disease	6	2	4	33.33%	66.67%
Oncological process in anamnesis	27	17	10	62.96%	37.04%
Venous thromboembolism (VTE) in anamnesis	13	9	4	69.23%	30.77%
Arterial incident in anamnesis	62	47	15	75.81%	24.19%

Table 4: Number of epochs for each neural network model

Neural network (NN) model	Number of epochs
NN model with threshold 0	52
NN model with threshold 0.1	52
NN model with threshold 0.2	56
NN model with threshold 0.3	53
NN model with threshold 0.4	105
NN model with threshold 0.5	95
NN model with threshold 0.6	52
NN model with threshold 0.7	67
NN model with threshold 0.8	53
NN model with threshold 0.9	69

Table 5: Accuracy, precision, sensitivity, specificity, TP rate, and FP rate for training set

Threshold	Accuracy	Precision	Sensitivity	Specificity	TP rate	FP rate
0.0	0.918	0.918	1	0	1	1
0.1	0.918	0.918	1	0	1	1
0.2	0.924	0.924	1	0.081	1	0.919
0.3	0.924	0.924	1	0.081	1	0.919
0.4	0.924	0.924	1	0.081	1	0.919
0.5	0.922	0.946	0.971	0.378	0.971	0.622
0.6	0.911	0.952	0.952	0.459	0.952	0.541
0.7	0.878	0.964	0.901	0.622	0.901	0.378
0.8	0.876	0.964	0.898	0.622	0.898	0.378
0.9	0.722	0.990	0.705	0.919	0.705	0.081
1.0	0.082	/	0	1	0	0

Table 6: Accuracy, precision, sensitivity, specificity, TP rate, and FP rate for test set

Threshold	Accuracy	Precision	Sensitivity	Specificity	TP rate	FP rate
0	0.815	0.815	1	0	1	1
0.1	0.821	0.820	1	0.029	1	0.971
0.2	0.832	0.829	1	0.088	1	0.912
0.3	0.832	0.829	1	0.088	1	0.912
0.4	0.837	0.833	1	0.118	1	0.882
0.5	0.832	0.852	0.960	0.265	0.960	0.735
0.6	0.826	0.860	0.940	0.324	0.940	0.676
0.7	0.815	0.892	0.880	0.529	0.880	0.471
0.8	0.799	0.890	0.860	0.529	0.860	0.471
0.9	0.717	0.945	0.693	0.824	0.693	0.176
1	0.185	/	0	1	0	0

Table 7: Confusion matrix for training set

	Predicted 0	Predicted 1
Actual 0	23	14
Actual 1	41	372

Confidence intervals

Tables 10 and 11 show a 95% confidence interval for the mean values of features for patients who survived and who did not survive COVID-19, respectively, with upper and lower confidence interval (CI) limits.

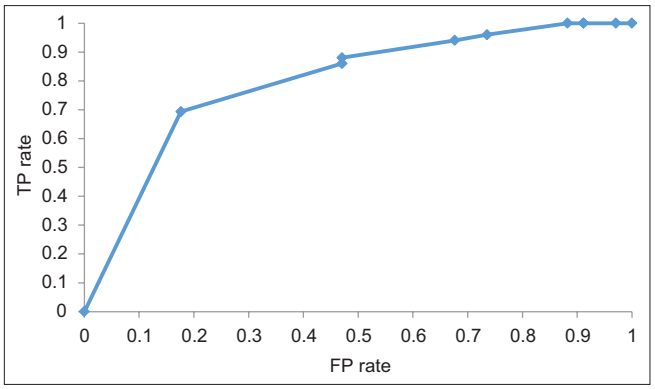

Figure 2: ROC curve for the test set

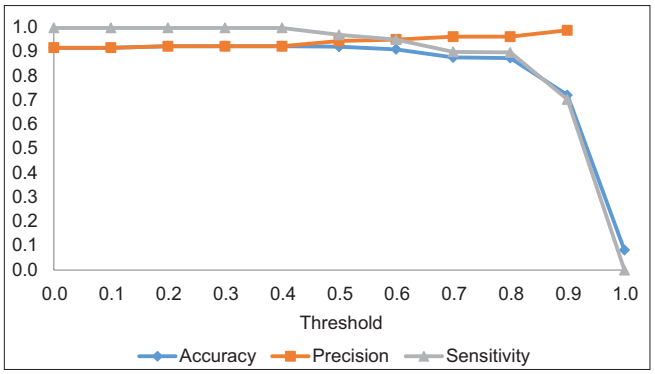

Figure 3: Accuracy, precision, and sensitivity for each neural network model for the training set

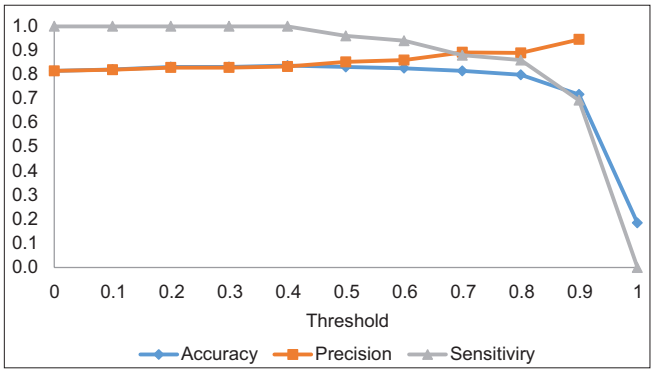

Figure 4: Accuracy, precision, and sensitivity for each neural network model for the test set

Discussion

Development of optimal neural network model, as well as hypothesis testing and confidence intervals analysis, were performed in this research. Combining these three methods could help in the daily work of physicians who treat COVID-19 hospitalized patients. Alotaibi *et al.*[7] made Artificial Neural Network, Support Vector Machine, and Random Forest Regression models for early prediction of illness severity of COVID-19 infected patients at the beginning of the illness so that the patients can be sorted and treated adequately. The prediction was done using the patient's history and laboratory results. The performance of different developed models was

Table 8: Confusion matrix for test set

	Predicted 0	Predicted 1
Actual 0	23	14
Actual 1	41	372

Table 9: Hypothesis tests for the means of features

Feature	μ_1	μ_2	H_0 and H_1	P	Conclusion
Age	70.34	63.46	H_0:μ_1-μ_2≤0 H_1:μ_1-μ_2>0	0.00000	Reject H_0
Erythrocytes	4.64	4.46	H_0:μ_1-μ_2≤0 H_1:μ_1-μ_2>0	0.02060	Reject H_0
Haemoglobin	136.87	132.62	H_0:μ_1-μ_2≤0 H_1:μ_1-μ_2>0	0.04453	Reject H_0
Haematocrit	0.406	0.394	H_0:μ_1-μ_2≤0 H_1:μ_1-μ_2>0	0.03157	Reject H_0
Thrombocytes	219.70	302.80	H_0:μ_2-μ_1≤0 H_1:μ_2-μ_1>0	0.00000	Reject H_0
Leukocytes	11.08	9.18	H_0:μ_1-μ_2≤0 H_1:μ_1-μ_2>0	0.00206	Reject H_0
Neutrophil granulocytes	88.27	77.83	H_0:μ_1-μ_2≤0 H_1:μ_1-μ_2>0	0.00000	Reject H_0
Lymphocytes	7.46	15.08	H_0:μ_2-μ_1≤0 H_1:μ_2-μ_1>0	0.00000	Reject H_0
Monocytes	3.86	6.41	H_0:μ_1-μ_2≤0 H_1:μ_1-μ_2>0	0.00000	Reject H_0
Basophil granulocytes	0.163	0.237	H_0:μ_2-μ_1≤0 H_1:μ_2-μ_1>0	0.00000	Reject H_0
Eosinophil granulocytes	0.245	0.440	H_0:μ_2-μ_1≤0 H_1:μ_2-μ_1>0	0.00457	Reject H_0
CRP	97.56	46.40	H_0:μ_1-μ_2≤0 H_1:μ_1-μ_2>0	0.00007	Reject H_0
D-dimer	8550.66	3678.57	H_0:μ_1-μ_2≤0 H_1:μ_1-μ_2>0	0.01834	Reject H_0

Table 10: Confidence intervals (95%) for features of survivors

Feature	Lower CI limit	Upper CI limit	Mean
Age	62.38	64.53	63.46
Erythrocytes ($\times10^{12}$/L)	4.42	4.51	4.46
Haemoglobin (g/L)	131.30	133.95	132.62
Haematocrit (%)	0.39	0.40	0.39
Thrombocytes ($\times10^{9}$/L)	292.66	312.94	302.80
Leukocytes ($\times10^{12}$/L)	8.84	9.52	9.18
Neutrophil granulocytes (%)	76.87	78.79	77.83
Lymphocytes (%)	14.32	15.84	15.08
Monocytes (%)	6.12	6.69	6.41
Basophil granulocytes (%)	0.22	0.25	0.24
Eosinophil granulocytes (%)	0.36	0.52	0.44
C-reactive protein (mg/L)	41.78	51.02	46.40
D-dimer (μg/L)	2548.02	4809.12	3678.57

evaluated and showed that used models could be effective tools to predict the severity of the illness of the patient by using patient history and laboratory results. Atlam *et al.*[8] used machine learning techniques and artificial intelligence for analysing infection risks, survival, and classification. They presented two systems, Cox_COVID_19 and Deep_Cox_COVID_19. Those two systems are based on Cox regression, and they are the combination of autoencoder deep neural network and Cox regression. A clinical dataset of 1085 patients was used for this research. It was concluded that both systems could provide valuable information for medical institutions to make proper decisions, thus reducing mortality.[8] Zoabi *et al.* used machine learning techniques to predict COVID-19 diagnosis. They only used sex, age, whether the person was in contact with another infected person, and the first five symptoms. Their model can be used to prioritize testing if the testing resources are limited.[9] Prada *et al.*[10] developed neural network models to predict the risk of death. First, deep learning convolutional neural network models were developed to predict the risk of death of a patient based on X-ray images. Then, a developed convolutional neural network model was combined with input features such as age and gender in order to improve the goodness of fit and prediction. Results show that the second model showed better results than the convolutional neural network alone. Ma *et al.*[11] proposed the LSTM-Markov model because the Markov model reduced the prediction error of the LSTM model. The prediction results of the LSTM model were combined with the prediction errors of the Markov Model to get the final prediction results. The performance of the developed LSTM-Markov model is better than the LSTM model alone. Using Machine Learning-Integrated Random Forest Algorithm gives us a clearer insight into the importance of individual factors that may contribute to mortality.[12] Sankaranarayanan *et al.*[13] developed a machine learning model to predict mortality within 72 hours of COVID-19 patients' positive test in the Mayo Clinic patient population. In their analysis a recurrent neural network was used, and they stated that the most important features were age, Charlson comorbidity index, minimum oxygen saturation, fibrinogen level, and serum iron level. In essence, the aim of all proposed neural networks is to emphasize the factors that can lead to an increased risk of mortality in a given population, with the aim of preparing effective measures to minimize the number of deaths.[14-18] Villegas *et al.* used deep learning techniques to predict COVID-19 mortality among patients. Recurrent Neural Network models were created, and it was concluded that the time series model could be used for making a decision in healthcare systems. However, the authors suggest that models have their performance increased with more available data.[19] In fact, as much data as possible, as many patients as possible, leads to better conclusions. Prevention is imperative, as is vaccination.[20,21] In[22] machine learning, models were developed to classify coughs recorded on smartphones in order to differentiate COVID-19 positive coughs from COVID-19 negative and healthy coughs. Clinical data and quantity of radiomic features from CT chest images in COVID-19 patients were used to develop machine learning prediction models.[23] Flores *et al.*[24] used two types of machine learning models to predict COVID-19 incidents in three risk categories as well as analysed the relative importance of the features in Chile. Convolutional LSTM classifier showed better performance than support vector machine (SVM). A cloud-based machine learning-based models[25] were developed to predict COVID-19 cases

Table 11: Confidence intervals (95%) for features of survivors

Feature	Lower CI limit	Upper CI limit	Mean
Age	67.83	72.85	70.34
Erythrocytes ($\times 10^{12}$/L)	4.48	4.80	4.64
Haemoglobin (g/L)	132.14	141.61	136.87
Haematocrit (%)	0.39	0.42	0.41
Thrombocytes ($\times 10^{9}$/L)	194.72	244.69	219.70
Leukocytes ($\times 10^{12}$/L)	9.84	12.32	11.08
Neutrophil granulocytes (%)	86.93	89.62	88.27
Lymphocytes (%)	6.41	8.50	7.46
Monocytes (%)	3.39	4.34	3.86
Basophil granulocytes (%)	0.14	0.19	0.16
Eosinophil granulocytes (%)	0.12	0.37	0.25
C-reactive protein (mg/L)	72.39	122.72	97.56
D-dimer (μg/L)	4123.72	12977.61	8550.66

within the next seven days in Bangladesh in order to develop assess prevention strategies. Rehman *et al.*[26] performed analysis of machine learning and deep learning techniques used in prediction, diagnosis, classification, and detection of coronavirus in published papers. The authors concluded that these techniques could play a significant role in the COVID-19 pandemic.[27] three clusters of countries and territories with similar COVID-19 dynamics were identified using machine learning. The authors used data visualization and descriptive statistics to help understand COVID-19 spread and impact. A three-step methodology was followed to develop a machine learning model to detect COVID-19 with high accuracy using X-ray images. The authors concluded that the Majority Voting approach is an adequate strategy and may achieve accuracy up to 99.314%.[28] Biochemical tests were used as prediction indicators to develop a machine learning diagnosis and disease severity model. The authors stated that all developed models could assist in the diagnosis and prediction of COVID-19 severity.[29] Alves *et al.*[30] used Random Forest (RF) and Decision Tree to develop machine learning models based on the data obtained from blood parameters. The authors stated that developed models are useful in diagnosis in COVID-19. The optimal neural network model, out of 11 developed models with different threshold values, was determined based on performance evaluation metrics. Optimal neural network model predicted the outcome of COVID-19 hospitalised patients with accuracy = 87.78%, precision = 96.37%, sensitivity = 90.01%, and specificity = 62.16%. The chosen threshold is 0.7. Performed hypothesis tests showed that the mean values of age, erythrocytes, haemoglobin, haematocrit, leukocytes, neutrophil granulocytes, CRP, and D-dimer of patients who died are significantly higher ($P < 0.05$) than mean values of these features of patients who survived, while mean values of thrombocytes, lymphocytes, monocytes, basophil granulocytes, and eosinophil granulocytes of the patients who survived are significantly higher ($P < 0.05$) than mean values of these features of patients who died. Upper and lower 95% confidence intervals for all laboratory findings and age were calculated for patients who survived and died. From the literature review written in the introduction, it can be seen that machine learning techniques were used to predict diagnosis, severity, outcomes, and risk of death, based on inputs, such as demographics, symptoms, frailty, patient history, laboratory findings, X-ray images, etc. This research had a different approach. The idea was to use laboratory findings, information about accompanying comorbidities, and demographics in order to predict the outcome (survival or death) of COVID-19 hospitalized patients. Inflammatory markers, as well as D-dimer values, with relation to the patient's age, can be helpful in stratifying mortality risk on the level of primary health care. In this research, 11 neural network models with different threshold values from 0 to 1 with step 0.1 were developed. The optimal neural network model was determined based on accuracy, precision, sensitivity, and specificity calculated for each out of 11 developed neural network models. Since it is impossible to obtain 100% accuracy of the neural network on both training and test sets, hypothesis testing and 95% confidence interval were performed to better understand the medical state of the patients and help predict the outcome of COVID-19 hospitalised patients. It is likely that the future will bring the use of neural networks in the prediction of outcomes, with the inclusion of preventive measures that have been implemented or the type of vaccination and number of vaccine doses. The limitation of our study is that it did not include X-rays or CT scans of the chest, i.e., pharmacological modality. However, the purpose of the study itself is the prediction of the outcome of hospitalized patients based on laboratory findings, accompanying comorbidities and demographics, using neural networks, hypothesis testing, and 95% confidence intervals. This approach could help physicians in their daily work with COVID-19 hospitalized patients.

Conclusions

This paper demonstrates that the combination of neural networks, hypothesis testing, and confidence intervals could be used to help physicians in their daily work with COVID-19 hospitalized patients. The hypothesis of the research was that neural networks in combination with hypothesis testing and confidence intervals could be used to predict the outcome of COVID-19 hospitalized patients with high accuracy using laboratory blood test results, accompanying comorbidities, and demographics as inputs. The developed optimal neural network model has high accuracy, precision, sensitivity, and low FP rate based on initial laboratory findings, demographics, and comorbidities and can successfully solve this prediction problem. The model has great potential to predict the outcome of the treatment of COVID-19 hospitalized patients. In addition, hypothesis testing and confidence intervals could help to better understand the medical state of the patients and contribute to better prediction of the outcome.

Acknowledgements

The authors would like to thank General Hospital "Prim. dr. Abdulah Nakas" from Sarajevo, Bosnia, and Herzegovina, for allowing the use of data for this research.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Yi Y, Lagniton PN, Ye S, Li E, Xu RH. COVID-19: What has been learned and to be learned about the novel coronavirus disease. Int J Biol Sci 2020;16:1753–66.
2. Begic E, Naser N, Begic N. Hypercoagulability in COVID-19 and post-COVID patients-Characteristics and current treatment guidelines. e-J Cardiol Pract 2021;21.
3. Sharifpour M, Rangaraju S, Liu M, Alabyad D, Nahab FB, Creel-Bulos CM, *et al.* Emory COVID-19 quality & clinical research collaborative: C-Reactive protein as a prognostic indicator in hospitalized patients with COVID-19. PLoS One 2020;15:e0242400.
4. Coomes EA, Haghbayan H. Interleukin-6 in COVID-19: A systematic review and meta-analysis. Rev Med Virol 2020;30:1-9.
5. Ye Q, Zhou J, Wu H. Using information technology to manage the COVID-19 pandemic: Development of a technical framework based on practical experience in China. JMIR Med Inform 2020;8:e19515. doi: 10.2196/19515.
6. Prnjavorac B, Mujaković A, Prnjavorac L, Bego T, Jusufović E, Begić E, *et al.* Chest x-ray resolution after SARS-CoV-2 infection. Med Glas (Zenica) 2021;18:370-7.
7. Alotaibi A, Shiblee M, Alshahrani A. Prediction of severity of COVID-19-infected patients using machine learning techniques. Computers 2021;10:31.
8. Atlam M, Torkey H, El-Fishawy N, Salem H. Coronavirus disease 2019 (COVID-19): Survival analysis using deep learning and Cox regression model. Pattern Anal Appl 2021;24:1-13.
9. Zoabi Y, Deri-Rozov S, Shomron N. Machine learning-based prediction of COVID-19 diagnosis based on symptoms. NPJ Digit Med 2021;4:1-5.
10. Prada J, Gala Y, Sierra AL. COVID-19 mortality risk prediction using x-ray images. Int J Interact Multi Artif Intell 2021;6:7-14.
11. Ma R, Zheng X, Wang P, Liu H, Zhang C. The prediction and analysis of COVID-19 epidemic trend by combining LSTM and Markov method. Sci Rep 2021;11:1-14.
12. Cornelius E, Akman O, Hrozencik D. COVID-19 mortality prediction using machine learning-integrated random forest algorithm under varying patient frailty. Mathematics 2021;9:2043.
13. Sankaranarayanan S, Balan J, Walsh JR, Wu Y, Minnich S, Piazza A, *et al.* Covid-19 mortality prediction from deep learning in a large multistate electronic health record and laboratory information system data set: Algorithm development and validation. J Med Internet Res 2021;23:e30157.
14. Adhinata FD, Rakhmadani DP. Prediction of Covid-19 daily case in indonesia using long short term memory method. Teknika 2021;10:62-7.
15. Jarndal A, Husain S, Zaatar O, Al Gumaei T, Hamadeh A. GPR and ANN based prediction models for COVID-19 death cases. 2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI), Sharjah, United Arab Emirates, 3-5 Nov. 2020, IEEE. p. 1-5.
16. Ko H, Chung H, Kang WS, Park C, Kim SE, Chung CR, *et al.* 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. J Med Internet Res 2020;22:e25442. doi: 10.2196/25442.
17. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, *et al.* Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. BMJ 2020;369. doi: 10.1136/bmj.m1328.
18. Gong J, Ou J, Qiu X, Jie Y, Chen Y, Yuan L, *et al.* A tool for early prediction of severe coronavirus disease 2019 (COVID-19): A multicenter study using the risk nomogram in Wuhan and Guangdong, China. Clin Infect Dis 2020;71:833-40.
19. Villegas M, Gonzalez-Agirre A, Gutiérrez-Fandiño A, Armengol-Estapé J, Carrino CP, Fernández DP, *et al.* Predicting the evolution of COVID-19 mortality risk: A recurrent neural network approach. medRxiv 2021:2020-12. doi: 10.1101/2020.12.22.20244061.
20. Rahmani AM, Mirmahaleh SYH. Coronavirus disease (COVID-19) prevention and treatment methods and effective parameters: A systematic literature review. Sustain Cities Soc 2021;64:102568.
21. Xing K, Tu XY, Liu M, Liang ZW, Chen JN, Li JJ, *et al.* Efficacy and safety of COVID-19 vaccines: A systematic review. Zhongguo Dang Dai Er Ke Za Zhi 2021;23:221-8.
22. Pahar M, Klopper M, Warren R, Niesler T. COVID-19 cough classification using machine learning and global smartphone recordings. Comput Biol Med 2021;135:104572.
23. Shiri I, Sorouri M, Geramifar P, Nazari M, Abdollahi M, Salimi Y, *et al.* Machine learning-based prognostic modeling using clinical data and quantitative radiomic features from chest CT images in COVID-19 patients. Comput Biol Med 2021;132:104304.
24. Flores C, Taramasco C, Lagos ME, Rimassa C, Figueroa R. A feature-based analysis for time-series classification of covid-19 incidence in Chile: A case study. Appl Sci 2021;11:7080.
25. Satu M, Howlader KC, Mahmud M, Kaiser MS, Shariful Islam SM, Quinn JM, *et al.* Short-term prediction of COVID-19 cases using machine learning models. Appl Sci 2021;11:4266.
26. Rehman A, Iqbal MA, Xing H, Ahmed I. COVID-19 detection empowered with machine learning and deep learning techniques: A systematic review. Appl Sci 2021;11:3414.
27. António N, Rita P, Saraiva P. COVID-19: Worldwide profiles during the first 250 days. Appl Sci 2021;11:3400.
28. Ben Jabra M, Koubaa A, Benjdira B, Ammar A, Hamam H. COVID-19 diagnosis in chest x-rays using deep learning and majority voting. Appl Sci 2021;11:2884.
29. Cobre AF, Stremel DP, Noleto GR, Fachi MM, Surek M, Wiens A, *et al.* Diagnosis and prediction of COVID-19 severity: Can biochemical tests and machine learning be used as prognostic indicators? Comput Biol Med 2021;134:104531.
30. Alves MA, Castro GZ, Oliveira BAS, Ferreira LA, Ramírez JA, Silva R, *et al.* Explaining machine learning based diagnosis of COVID-19 from routine blood tests with decision trees and criteria graphs. Comput Biol Med 2021;132:104335.