DOI: 10.37990/medr.1226429 Med Records 2023;5(3):500-6

MEDICAL RECORDS-International Medical Journal

Research Article

Prediction of Short or Long Length of Stay COVID-19 by Machine Learning

Muhammet Ozbilen[1], Zubeyir Cebeci[2], Aydin Korkmaz[3], Yasemin Kaya[4], Kaan Erbakan[5]

[1] Ordu University Faculty of Medicine, Department of Internal Medicine, Ordu, Türkiye
[2] Ordu University Faculty of Medicine, Department of Anesthesiology and Reanimation, Ordu, Türkiye
[3] Ordu University Faculty of Medicine, Department of Radiology, Ordu Türkiye
[4] Ordu University Faculty of Medicine, Department of Internal Medicine, Ordu Türkiye
[5] Ordu University Faculty of Medicine, Medical Student 6th Grade, Ordu, Türkiye

Abstract

Aim: The aim of this study is to utilize machine learning techniques to accurately predict the length of stay for Covid-19 patients, based on basic clinical parameters.

Material and Methods: The study examined seven key variables, namely age, gender, length of hospitalization, c-reactive protein, ferritin, lymphocyte count, and the COVID-19 Reporting and Data System (CORADS), in a cohort of 118 adult patients who were admitted to the hospital with a diagnosis of Covid-19 during the period of November 2020 to January 2021. The data set is partitioned into a training and validation set comprising 80% of the data and a test set comprising 20% of the data in a random manner. The present study employed the caret package in the R programming language to develop machine learning models aimed at predicting the length of stay (short or long) in a given context. The performance metrics of these models were subsequently documented.

Results: The k-nearest neighbor model produced the best results among the various models. As per the model, the evaluation outcomes for the estimation of hospitalizations lasting for 5 days or less and those exceeding 5 days are as follows: The accuracy rate was 0.92 (95% CI, 0.73-0.99), the no-information rate was 0.67, the Kappa rate was 0.82, and the F1 score was 0.89 (p=0.0048).

Conclusion: By applying machine learning into Covid-19, length of stay estimates can be made with more accuracy, allowing for more effective patient management.

Keywords: COVID-19, machine learning, R programming, length of stay, accuracy, management

INTRODUCTION

The ongoing COVID-19 pandemic, now approaching its third year, persists in its global dissemination. During the initial wave of the pandemic, all nations have demonstrated a lack of capability in effectively managing the outbreak (1). The current pandemic has resulted in prolonged and unforeseeable hospitalizations due to Covid-19, posing a challenge in the management of patients with pre-existing chronic or severe conditions, as well as those requiring hospitalization and subsequent monitoring (2). The COVID-19 pandemic serves as a reminder of the critical importance of proficient bed capacity management for all illnesses, particularly in the context of outbreaks that require prompt implementation of action plans (3). The challenge of managing the overwhelming demand for hospitalizations during the pandemic, which exceeded the capacity of health facilities, required a proficient integration of scarce data and vague indicators.

In healthcare facilities where there are constraints on bed availability and patient capacity, it is imperative to develop a comprehensive strategy and ascertain whether patients will be admitted for a brief or extended duration. It is advisable to implement pre-hospital interventions to mitigate the duration of hospitalization, commonly referred to as length of stay (LoS) (4). Thus, efficient management can be facilitated by ensuring the effective circulation of all hospital items, including people and logistics, particularly in the context of acute and complex events (5).

Today, machine learning is gaining prominence as a scientific approach capable of effectively forecasting the likelihood of various phenomena or events and their corresponding risk rates

CITATION
Ozbilen M, Cebeci Z, Korkmaz A, et al. Prediction of short or long length of stay COVID-19 by machine learning. Med Records. 2023;5(3):500-6. DOI:1037990/medr.1226429

Received: 10.01.2023 **Accepted:** 16.05.2023 **Published:** 13.07.2023

Corresponding Author: Muhammet Ozbilen, Ordu University, Faculty of Medicine, Department of Internal Medicine, Ordu, Türkiye
E-mail: muhammetozbilen@odu.edu.tr

(6). The employment of machine learning methods and models has become ubiquitous in various domains of the battle against Covid-19, including but not limited to the prediction of disease transmission, mortality rates, and the development of vaccines (7). Consequently, any performance data obtained through the utilization of machine learning models pertaining to Covid-19's length of stay (LoS) will make a substantial contribution to the management of hospitalizations.

The objective of this study was to evaluate the efficacy of machine learning (ML) models in predicting the requirement of short-term or long-term hospitalization of patients based on their basic demographic, biochemical, and imaging data.

MATERIAL AND METHOD

The present study is a retrospective and observational investigation that was carried out at the Ministry of Health Ordu University Training and Research Hospital, following the acquisition of requisite approvals from the pertinent ethics committee and institution.

Patient Selection

The study sample comprises of individuals who are 18 years of age or older and who sought medical attention at our healthcare facility for Covid-19 infection during the period spanning from November 2020 to January 2021, and subsequently required hospitalization. A cohort of patients was identified through a full review of the health information management system (HIMS).

Data Types and Preparation

The demographic and descriptive information pertaining to the patients identified by HIMS was completely documented. This included details such as age, gender, and other relevant factors. The dates pertaining to the hospitalization and subsequent discharge were properly recorded. Based on WHO data spanning the last two decades (1996-2014), the average hospital stay duration in Turkey is 5.1 days (8). Therefore, in our study design, we defined the parameters of low and long-term hospitalization as stays of less than or greater than 5 days, respectively. A new categorical variable named "LOSClass" was introduced into the data set by categorizing individuals based on the duration of their hospitalization. Specifically, those who were hospitalized for 5 days or less were assigned a categorical designation of "1", while those who were hospitalized for more than 5 days were assigned a categorical designation of "2". The length of stay variable, expressed in numeric form, was removed from the dataset. Biochemical markers such as C-reactive protein (CRP), ferritin, and lymphocyte levels were incorporated into the study. In addition, the thoracic computed tomography (CT) scans acquired at the time of submission were recorded, along with the imaging dates and the Covid-19 Reporting and Data System (CORADS) assessments as indicated in the CT reports. The superiority of the CORADS scoring system in predicting the severity of pulmonary involvement, as well as in forecasting prognosis and length of stay (LoS), has been demonstrated in comparison to other proposed scoring systems (9). These findings have been included in the dataset. Upon reevaluation of the public health management system, the PCR outcomes for the detection of COVID-19 were reaffirmed.

Statistical Analysis and Machine Learning

Basic statistical and machine learning analyses were performed using Jamovi version 2.3.16 (10,11) and RStudio version 4.2.0 of the R programming language (11). The caret package (version 6.0-93) in the R programming language was employed for the implementation of a machine learning methodology (12). The caret package available in the R programming language appears to be adequate for the implementation of a machine learning system, as it is capable of performing intricate regression and classification analyses. (13). The train() method is used by the related package to construct a predictive model. The packages "party, physc, tidyverse, and dplyr" were used for ML analysis in addition to caret. A flowchart of the plan for the study is shown in Figure 1.

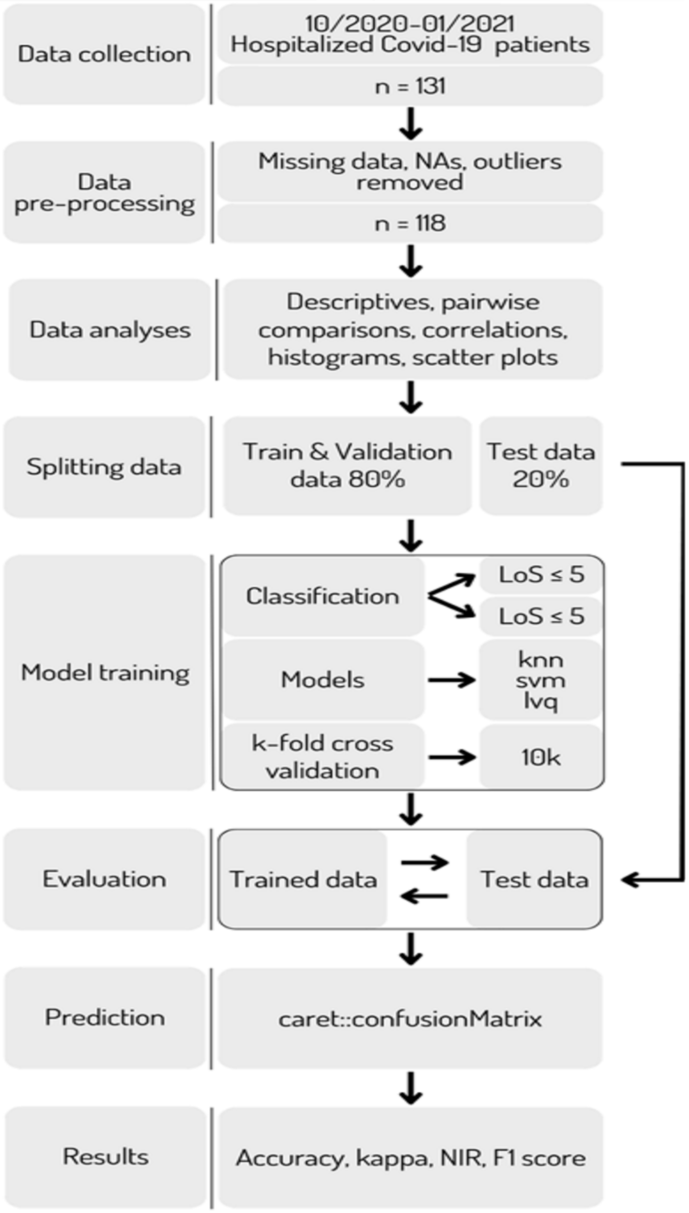

Figure 1. Flow-chart of the study. NA: not available, LoS: length of stay, knn: k-nearest neighbor, lvq: learning vector quantization, svm: support vector machine, NIR: no-information rate

Model Implementation and Evaluation

The labels of all data assigned based on the length of stay categories. The labeled dataset was split into two parts randomly as follows: 80% of the observations were randomly assigned to the training and validation set and the independent test set is 20%.

The k-fold cross validation is a robust and secure internal validation approach to obtain fair results. However, avoiding the potential bias of cross-validation several medical studies use the external validation which is known as independent test (14). We prefer same independent test approach with cross validation. In this paper, the 80% parts of all dataset is used for model training and validation process. The rest of dataset, namely 20% is unseen data and used for external independent testing to evaluate the model performance.

Again, our method applied optimally defined classifier parameters. Cross-validation tuning of hyperparameters requires lengthy training durations and yielded comparable accuracy performance (±1%) for our dataset. Therefore, we preferred to use the system's optimal hyperparameters.

For model training, we used the classification models k-nearest neighbor (method: knn), learning vector quantization (method: lvq), and support vector machine (method: svm). Given the limited size of our dataset, we employed the k-fold cross-validation technique, a widely utilized approach for model evaluation (15) that is favored for its ability to yield unbiased performance outcomes. Specifically, we opted for a 10-fold cross-validation strategy.

Subsequently, the predictive values of each model were calculated, and the performance metrics of each model were assessed through using of the "confusionMatrix" function from the caret package. The evaluation encompassed measures such as accuracy, kappa, no information rate (NIR), and p significance findings. Kappa indicates the degree of agreement between two groups of categorical data, relative to what would be predicted by chance alone. A score of 1 (one) represents total agreement, whereas a score of 0 (zero) represents disagreement (16).

The term NIR, which is also referred to as the base rate, pertains to the proportion of the "most popular state" that is determined through the validation of the test set using machine learning models. It is an accuracy paradox caused by repeatedly "incorrect" estimations of the observational variables (17). It is a crucial requirement for test performance, and when estimating using a good model, the accuracy rate is predicted to be significantly higher than NIR, that is, the p value should be less than 0.05 according to the 95% confidence interval.

RESULTS

A total of 131 patients who met the criteria for the specified time period were identified. During the data preprocessing stage, any missing data and NAs were eliminated. Furthermore, based on the boxplots generated during the initial analysis of the data, it was ascertained that the variables harbored outlier values, and the related observations were eliminated from the study. The final version of the dataset comprised a total of 118 observations. In the final version of the dataset, the category classification of the dependent variable was analyzed, and no class imbalance was found (Script 1).

Script 1. A section showing the process of removing outliers from the dataset, the number of observations obtained after removing all outliers, and the LOSClass table. df: dataset, FERR: ferritin

```
#An example of removing the outliers from a
variable

quartiles <- quantile(df$FERR,
         probs=c(.25, .75), na.rm =
         FALSE)
         IQR <- IQR(df$FERR)
         Lower <- quartiles[1] - 1.5*IQR
         Upper <- quartiles[2] + 1.5*IQR


df <- subset(df, df$FERR > Lower &
df$FERR
        < Upper)



#Dimension of the final dataset
> dim(df)
[1] 118  7

#Table of the LoSClass
> table(df$LOSClass)
 1  2
58 60
```

Tables 1 and 2 present the descriptive statistics of the dataset. The sample's average age was 67.69 years with a standard deviation of 14.63. There were 57 female patients and 61 male patients, as revealed by the distribution of gender. The percentage and quantitative distributions of CORADS scores, whose median value was 5 (five), are shown in Table 3.

The scatter and histogram graphs of the numerical variables in the final version of the dataset, along with the correlation values calculated using the "pairs.panel" function of the "psych" package, are shown in Figure 2.

The data presented in this figure indicates a weak correlation (r=0.20) between CRP and ferritin levels. This correlation was not statistically significant (p = 0.033). With the exception of lymphocyte levels, Pearson's correlation analysis of the numerical data with CORADS, which is ordinal data, revealed significant relationships with the other three numerical data. We discovered the following correlations: r=-0.343 with age (p<0.001), r=0.252 with ferritin (p<0.001), r=0.252 with CRP (p<0.001), and r=-0.126 with lymphocytes (p=0.174).

Table 1. Descriptive statistics of numeric variables

	Mean	95% CI		SD	Min	Max	Calc W
		Lower	Upper				
AGE	67.69	65.03	70.36	14.63	30	96	0.165
CRP	8.45	7.30	9.61	6.32	0.200	27.4	<0.001
LYM	1211.80	1107.05	1316.54	574.53	240	2740	<0.001
FERR	407.29	353.55	461.04	294.82	13.600	1195.0	<0.001

CRP: c-reactive protein, FERR: ferritin, LYM: lymphocyte, CI: confidence interval, SD: standard deviation, Min: minimum, Max: maximum, Calc W: p value of Shapiro-Wilk test

Table 2. Age-related descriptive data according to gender and length of stay classification

Sex	LOSClass	N	Mean	95% CI		SD	Min	Max
				Lower	Upper			
Female	1	27	63.8	57.6	70.0	15.6	30	96
	2	30	71.9	67.5	76.4	12.0	42	95
Male	1	31	67.2	61.0	73.3	16.7	39	92
	2	30	67.5	62.5	72.5	13.4	31	88
Total	-	118	67.7	65.0	70.4	14.6	30	96

LOSClass: length of stay classification, N: number, CI: confidence interval, SD: standard deviation, Min: minimum, Max: maximum Note: The CI of the mean assumes sample means follow a t-distribution with N - 1 degrees of freedom

Table 3. Frequencies of CORADS

CORADS	Counts	% of Total	Cumulative %
1	1	0.8 %	0.8 %
2	9	7.6 %	8.5 %
3	18	15.3 %	23.7 %
4	21	17.8 %	41.5 %
5	69	58.5 %	100.0 %

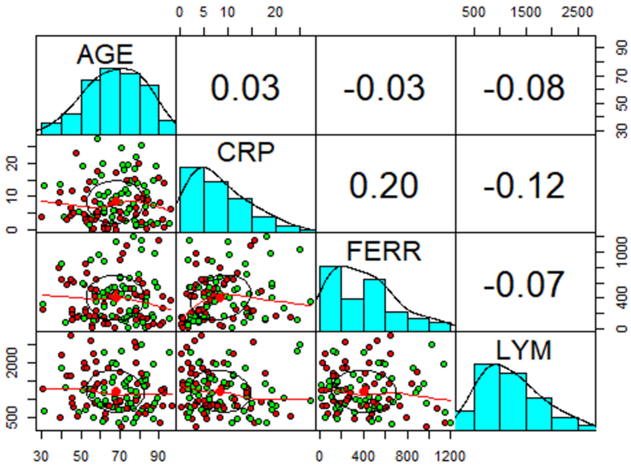

Figure 2. Histograms of numerical data and outcomes of pairwise comparisons. The ordinates and coordinates of each variable's values are shown on the side or top of the box bearing the variable's name. Moreover, scatter plots and correlation coefficients are located inside the boxes where the two variables overlap. CRP: c-reactive protein, FERR: ferritin, LYM: lymphocyte

Following the random splitting of the dataset (referred to as "df" in the analysis) into the two parts of 80% and 20%, two distinct sub-datasets were created, which were respectively named "x_train" and "x_test" in the analysis. The training set consisted of 94 observations, while the test set contained 24 observations. The number of length of stays (LoS) in both groups exhibited comparable. In the training set, there are 50 (fifty) observations for 5 days or less, and 44 (forty-four) observations for longer than 5 days. In the test set, the number of observations for 5 days or less was 8 and the number of observations over 5 days was 16 (Script 2).

Script 2. The process of splitting the data set into a training set and a test set by 80 percent and 20 percent, respectively. Additionally, the distribution of newly formed sets' values according to LOSClass

```
#Producing the train and the test samples

set.seed(2022)
split <- sample(1:nrow(df),
   as.integer(0.8*nrow(df)), F)

train1 <- df[split,]
test1 <- df[-split,]

x_train <- train1
x_test <- test1

y_train <- factor(train1[,"LOSClass"],
levels=c(1,2))
y_test <-  factor(test1[,"LOSClass"],
levels=c(1,2))

#The tables according to LOSClass of the train
and test datasets
> table(x_train$LOSClass)
 1  2
50 44
> table(x_test$LOSClass)
 1  2
 8 16
```

The results for the confusion matrix were obtained after model training and prediction computations. Script 3 includes an example model code.

Script 3. The coding and outcomes of knn model's performance

```
set.seed(2022)
knn.fit <- train(x_train, y_train,
      method = "knn",
      preProcess = c("center", "scale"),
      tuneLength = 10,
      trControl = trainControl(method =
      "cv", number=10))
predknn <- predict(knn.fit, x_test)
confknn <-
      confusionMatrix(as.factor(predknn),
      as.factor(y_test), mode="everything")
> confknn
Confusion Matrix and Statistics
            Reference
Prediction    1  2
         1    8  2
         2    0 14
               Accuracy:  0.9167
                 95% CI:  (0.73, 0.9897)
   No Information Rate:  0.6667
   P-Value [Acc > NIR]:  0.004871
                  Kappa:  0.8235
         Mcnemar's Test:  0.479500
                P-Value:
            Sensitivity:  1.0000
            Specificity:  0.8750
         Pos Pred Value:  0.8000
         Neg Pred Value:  1.0000
              Precision:  0.8000
                 Recall:  1.0000
                     F1:  0.8889
             Prevalence:  0.3333
         Detection Rate:  0.3333
   Detection Prevalence:  0.4167
      Balanced Accuracy:  0.9375
       'Positive' Class:   1
```

The performance outcomes of ML models were evaluated with special attention given to accuracy, kappa, NIR, F1 scores, and p significant values. Upon examination of the performance metrics presented in Table 4, it was determined that the accuracy of the knn model was 0.92, with a 95% confidence interval ranging from 0.73 to 1.00 and a p-value of 0.0048. Despite achieving a 0.75 accuracy level in the SVM model, the statistical significance of the p value was not established. The lvq model exhibited poor performance in terms of both accuracy and p values.

The Kappa value, commonly referred to as Cohen's Kappa, is a statistical measure that evaluates the level of agreement between two random categorical variables in terms of their predicted and true labels. Its values range from -1 to 1. Inconsistency is typically attributed to results that fall below 0.6, whereas a higher degree of congruity is associated with values closer to 1 (one). (18). The results of our study indicate that among the models tested, the knn model exhibited the highest Kappa value, which was recorded at 0.82.

On the contrary, the F1 score denotes the ability of the test to predict the true. In cases where there exists an imbalance among classes, the employment of this metric takes precedence over accuracy measurements (19). Although the dataset exhibits no class imbalance, the KNN model's F1 score is notably high, suggesting that it is a reliable model.

The McNemar test is utilized to ascertain whether incorrect data is incidental or if it manifests with a specific frequency. Thus, it is advisable not to reject the null hypothesis as errors are expected to occur consistently (20). The McNemar's test outcome pertaining to our knn model yielded a p-value of 0.48, thereby confirming the null hypothesis.

Table 4. Results of the models by confussionMatrix function

Models	ACC	95% CI	NIR	Kappa	F1 Score	p value [ACC>NIR]	McNemar's Test p value
knn	0.92	0.73-0.99	0.67	0.82	0.89	**0.0048**	0.48
lvq	0.5	0.44-0.84	0.67	0.14	0.33	0.59	0.28
svm	0.75	0.53-0.90	0.67	0.52	0.72	0.26	0.04

knn: k-nearest neighbors, lvq: learning vector quantization, svm: support vector machines, ACC: accuracy, CI: confidence interval, NIR: no information rate. Significant p-values between Accuracy and NIR are shown in bold

DISCUSSION

The present study used a machine learning (ML) approach to categorically evaluate hospitalization duration (≤5 days or >5 days) for Covid-19 patients. The findings revealed a noteworthy accuracy rate exceeding 90%. No statistically significant results were observed in the remaining two machine learning models. Additional findings regarding the performance of the K-Nearest Neighbors (KNN) model suggest that it is efficacious and possesses a favorable predictive capability.

The COVID-19 pandemic has resulted in a noteworthy duration of hospitalization for patients, and we will look at and discuss certain outcomes pertaining to this phenomenon. One of the most intriguing studies among these is a systematic review. In that review, encompassed a sample size of approximately 430,000 patients from 126 distinct investigations. The average length of stay (LoS) for Covid-19 was found to be 14.49 days, with a standard deviation of 7.92. The minimum and maximum durations of hospitalization were 3.5 and 53.8 days, respectively (21).

The use of blood tests and imaging scans can aid in the early determination of the probability of short or extended hospitalization. In addition to Covid-19, various factors including age, lung involvement scores, CRP, ferritin, and lymphocyte levels have been identified as significant predictors of disease severity and consequently, length of hospitalization (22). The research conducted by Oksuz et al. revealed a significant association between abnormal blood parameters and increased hospitalization duration and costs (23).

We observed that machine learning analyses of the variables of the severity of lung involvement in The existing literature on Covid-19 and the estimation of length of stay (LoS) was found to be limited. In a study involving 254 patients conducted by Chamberlin et al., a score other than CORADS was used to predict hospitalization and an accuracy of 0.77 was determined (24). In a different study by Purkayastha et al. that used a different scoring system than CORADS, the accuracy levels of the ML analysis results using the "boosting" and "bayesian" methods were 0.68 (p=0.023) and 0.77 (p=0.023), respectively (25).

The study conducted by Olivato et al. involved a sample size of 6000 patients diagnosed with Covid-19. The researchers used a machine learning approach to forecast the duration of hospitalization, regardless of the severity of lung involvement. The approach was designed to predict hospitalization periods of less than seven days or more. The achieved F1 score was 0.76; however, the values for kappa and p were not disclosed (26).

Saadatmand et al. used ML to determine the length of stay (LoS) in hospital, identifying between stays of less than seven days and those exceeding this duration. The study conducted on 112 critical care patients demonstrated that the CART model attained the highest accuracy score of 0.82 (27). In the related research, the statistical significance level was not specified, and the kappa coefficient was observed to be low, measuring at 0.48.

Finally, Alabbas et al. showed that the rf model had an accuracy score of 94.16% in a study involving 895 patients that did not include lung involvement levels (28). In the linked study, attempts were made to estimate the number of hospitalized days for seven distinct five-day groups.

CONCLUSION

In our study, we used ferritin, CRP, lymphocyte, and lung involvement scores, which are frequently requested basic tests for Covid-19, as well as age and gender to predict whether the length of stay would be short or long. Numerous biomarkers were incorporated into prediction models in the published literature, whereas lung involvement scores were rarely used. In addition, crucial prediction performance scores such as "kappa" and "p significance" were absent from the analysis outcomes. Alternatively, we believe that our relatively small sample size has an effect on the prediction models. As evidenced by the literature and other methodologies, using the ML method to determine the risk of short- and long-term hospitalization appears feasible and reasonable. We believe that the incorporation of ML coding of such data into HIMS interfaces will make substantial contributions to the management of Covid-19 patients through the assessment of short or long hospitalization risk rates.

Financial disclosures: *The authors declared that this study has received no financial support.*

Conflict of Interest: *The authors have no conflicts of interest to declare.*

Ethical approval: *Ordu University Clinical Research Ethics Committee, No: 2023/05, 6.01.2023.*

REFERENCES

1. Akcan FA, Onec K, Annakkaya A, et al. Düzce University Hospital in the pandemic process: from the perspective of chief physician. Konuralp Medical Journal. 2020;12:354–7.
2. Costa GJ, Júnior H de AF, Malta FC, et al. The impact of the COVID-19 pandemic on tertiary care cancer center: Analyzing administrative data. Semin Oncol. 2022;49:182–8.
3. Abdollahi F, Ghanyan S, Asadi F. COVID-19 pandemic and management on hospital length of stay: A review. Healthcare in Low-Resource Settings. 2021;9:10057.
4. Scanlon C, Cheng R, McRobb E, Ibrahim M. In-house testing for COVID-19: effects on length of stay, isolation and the need for inpatient rehabilitation. Aust Health Review. 2022;46:273–8.
5. Alwafi H, Naser AY, Qanash S, et al. Predictors of length of hospital stay, mortality, and outcomes among hospitalised covid-19 patients in Saudi Arabia: a cross-sectional study. J Multidiscip Healthc. 2021;14:839-52.
6. Javaid M, Haleem A, Pratap Singh R, et al. Significance of machine learning in healthcare: features, pillars and applications. International Journal of Intelligent Networks. 2022;3:58-73.
7. Vaishya R, Javaid M, Khan IH, Haleem A. Artificial Intelligence (AI) applications for COVID-19 pandemic. Diabetes Metab Syndr. 2020;14:337-9.
8. WHO European health information at your fingertips. https://gateway.euro.who.int/en/indicators/hfa_540-6100-average-length-of-stay-all-hospitals/visualizations/#id=19635&tab=table access date 04.05.2023.
9. Zayed NE, Bessar MA, Lutfy S. CO-RADS versus CT-SS scores in predicting severe COVID-19 patients: retrospective comparative study. Egypt J Bronchol. 2021;15:13.
10. The jamovi project. https://www.jamovi.org access date 25.12.2022
11. R: A Language and environment for statistical computing. https://cran.r-project.org access date 25.12.2022
12. A Short Introduction to the caret package. https://cran.r-project.org/web/packages/caret/vignettes/caret.html acces date 06.12.2022.
13. Kumar A. Pre-processing and modelling using caret package in R. International Journal of Computer Applications. 2018;181:39–42.

14. Al-Areqi F, Konyar MZ. Effectiveness evaluation of different feature extraction methods for classification of covid-19 from computed tomography images: A high accuracy classification study. Biomed Signal Process Control. 2022;76:103662.

15. Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. ArXiv. 2020:/abs/1811.12808

16. Kuhn M. Building predictive models in R using the caret package. Journal of Statistical Software. 2008;28:1–26.

17. Bond RR, Novotny T, Andrsova I, et al. Automation bias in medicine: the influence of automated diagnoses on interpreter accuracy and uncertainty when reading electrocardiograms. J Electrocardiol. 2018;51:6–11.

18. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: measures of agreement. Perspect Clin Res. 2017;8:187.

19. Grandini M, Bagli E, Visani G. Metrics for multi-class classification: an overview. ArXiv. 2020:/abs/2008.05756

20. Dalianis H. Evaluation Metrics and Evaluation. In: Dalianis H, editor. Clinical Text Mining: Secondary Use of Electronic Patient Records. Cham: Springer International Publishing; 2018. p. 45–53.

21. Alimohamadi Y, Yekta EM, Sepandi M, et al. Hospital length of stay for COVID-19 patients: a systematic review and meta-analysis. Multidiscip Respir Med. 2022;17:856.

22. Savrun A, Aydin IE, Savrun ST, Karaman U. The predictive role of biomarkers for mortality in COVID-19 patients. Trop Biomed. 2021;38:366–70.

23. Oksuz E, Malhan S, Gonen MS, et al. COVID-19 healthcare cost and length of hospital stay in Turkey: retrospective analysis from the first peak of the pandemic. Health Econ Rev. 2021;11:39.

24. Chamberlin JH, Aquino G, Schoepf UJ, et al. An interpretable chest CT deep learning algorithm for quantification of COVID-19 lung disease and prediction of inpatient morbidity and mortality. Acad Radiol. 2022;29:1178–88.

25. Purkayastha S, Xiao Y, Jiao Z, et al. Machine learning-based prediction of COVID-19 severity and progression to critical illness using CT imaging and clinical data. Korean J Radiol. 2021;22:1213–24.

26. Olivato M, Rossetti N, Gerevini AE, et al. Machine learning models for predicting short-long length of stay of COVID-19 patients. Procedia Comput Sci. 2022;207:1232–41.

27. Saadatmand S, Salimifard K, Mohammadi R, et al. Using machine learning in prediction of ICU admission, mortality, and length of stay in the early stage of admission of COVID-19 patients. Ann Oper Res. 2022;1-29.

28. Alabbad DA, Almuhaideb AM, Alsunaidi SJ, et al. Machine learning model for predicting the length of stay in the intensive care unit for COVID-19 patients in the eastern province of Saudi Arabia. Inform Med Unlocked. 2022;30:100937.