Safynaz AbdEl-Fattah Sayed
Abeer ElKorany
Sabah Sayed

APPLYING HUNGER GAME SEARCH (HGS) FOR SELECTING SIGNIFICANT BLOOD INDICATORS FOR EARLY PREDICTION OF ICU COVID-19 SEVERITY

Abstract *This paper introduces an early prognostic model for attempting to predict the severity of patients for ICU admission and detect the most significant features that affect the prediction process using clinical blood data. The proposed model predicts ICU admission for high-severity patients during the first two hours of hospital admission, which would help assist clinicians in decision-making and enable the efficient use of hospital resources. The Hunger Game search (HGS) meta-heuristic algorithm and a support vector machine (SVM) have been integrated to build the proposed prediction model. Furthermore, these have been used for selecting the most informative features from blood test data. Experiments have shown that using HGS for selecting features with the SVM classifier achieved excellent results as compared with four other meta-heuristic algorithms. The model that used the features that were selected by the HGS algorithm accomplished the topmost results (98.6 and 96.5%) for the best and mean accuracy, respectively, as compared to using all of the features that were selected by other popular optimization algorithms.*

Keywords ICU severity prediction, COVID-19, clinical blood tests, Hunger Game search (HGS) optimization algorithm, support vector machine (SVM), feature selection

Citation Computer Science 24(1) 2023: 113–136

1. Introduction

In December 2019, the coronavirus disease that was caused by a virus called severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) infected many people around the world and caused many deaths and a global pandemic [72]. The disease is fast-moving and has spread among many people [15]. There are patients with strong immunity and minor infections who may need a general ward and regular health care, while others of advanced ages with weak immunity, strong infections, and associated diseases who are exposed to major health crises such as difficulty breathing, a lack of oxygen, and blood clots and may require critical health care in intensive care units (ICUs) [9, 63]. ICUs are special wards in hospitals for monitoring critical patients all of the time and providing required treatments like providing oxygen and ventilation to maintain a patient's life [13, 56]. The number of ICU beds is very limited and specialized for only those patients with the most dangerous cases [55]. The management of infected patients according to their severity is very important and controversial for saving patients' lives and hospital resources [56].

The World Health Organization (WHO) announced that reverse transcription-polymerase chain reaction (RT-PCR) is the standard way to diagnose COVID-19 patients [20, 64] despite the misdiagnosis of some cases where a PCR may be negative and the patient has positive COVID-19 [7]. The findings of the RT-PCR test may take many hours or days to appear and are expensive [1, 52]. So, many experts recommended using alternative tools that take less time with accurate results, like vital signs, blood lab tests, and chest radiological images [1, 44, 64, 70]. Clinical blood tests are available in all hospitals to provide information about the disease and other pre-existing diseases for a full view of a patient's trajectory [1]. This helps doctors monitor changes in organ functions and detect factors that change with the severity of the patient [40]. The results of blood lab tests and vital signs are very fast and inexpensive [3].

Detecting the most significant indicators is an important matter that may assist doctors by using informative features rather than using noisy or messy features. There are many methods for selecting features like filters, wrappers, and embedded techniques [37, 47, 52]. Nowadays, meta-heuristic algorithms have been used for selecting features as an optimization problem in healthcare and many domains [28, 45, 53].

In this paper, a new prediction model has been proposed for predicting the severity of COVID-19 patients for ICU transfer based on demographic data, previously grouped diseases, blood parameters, and the vital signs of patients. The novelty of this model is that it uses a public time series window data set to achieve the following:

1. identify best time window for performing medical tests correctly;
2. determine most significant features in this time window by applying hybrid model that consists of recent optimization algorithm called Hunger Game search (HGS) for feature-selection step followed by using SVM classifier.

The performance of the HGS-SVM model has been compared with four well-known meta-heuristic algorithms: the flower pollination algorithm (FPA), Harris Hawks optimization (HHO), the whale optimization algorithm (WOA), and the bat algorithm (BA). The proposed prediction model may help hospitals make plans and take the right decisions toward patients by transferring them to other hospitals if there are not enough beds in an ICU or by supplying the hospital with other resources to accommodate the patients within the time frame that they can be saved.

The remaining paper is sectioned as follows: section 2 summarizes related works; section 3 explains the main steps and methods of the proposed model; section 4 contains the experimental results and a discussion; section 5 states our conclusions.

2. Related work

Several studies have used clinical blood data with different machine-learning algorithms to build models with different targets. The objectives most of the published research papers that are based on laboratory blood tests can be categorized into these goals: detecting the COVID-19 disease, predicting the severity of COVID-19 patients for ICU admission, and predicting the risk of a patient's mortality.

For detecting the COVID-19 disease, some research studies have utilized machine-learning classifiers for diagnosing COVID-19 based on blood lab results like [2, 3], and [54]; these achieved 86% accuracy by deep-learning LSTM, 99.8% accuracy by Xtreme gradient boosting (XGBoost), and 75% sensitivity by XGBoost, respectively. The used data is from Hospital Israelita Albert Einstein in São Paulo, Brazil.

For COVID-19 patients' mortality prediction, the research work in [6] proposed a mortality-prediction model with an SVM algorithm and achieved an accuracy of 92% in prognoses that were based on socio-demographic and medical information from South Korean patients; another work [4] built a prediction model by random forest (RF) and achieved an accuracy of 95% using the data of patients that were admitted to King Fahad University Hospital in the Kingdom of Saudi Arabia (KSA).

For predicting the severity of COVID-19 patients for ICU admission, [1] compared some statistical methods with various machine-learning algorithms. They used a public clinical data set with different time windows from Kaggle [31] to predict a patient's severity for ICU transfer. The applied statistical methods in their model were the t-test, chi-square test, and Pearson correlation, which were used for the feature-engineering process. The data set samples were divided into three categories: with noncommunicable disease (with-NCD), without NCD, and all of the combined data. RF and gradient boosting machine (GBM) achieved 93% accuracy for the with-NCD and without NCD data, respectively, and 89% accuracy for all of the combined data. The SHapley Additive exPlanations (SHAP) method [43] was used to select the best parameters according to each category in which respiratory rate, lactate, diastolic blood pressure, systolic blood pressure, neutrophils, and oxygen saturation were common for those groups that included all patients.

Also, the authors in [22] used the same data set to introduce a machine-learning assistance model to help experts determine a patient's priority for ICU admission according to biological laboratory findings. According to their results, the XGBoost classifier achieved the best results (97% accuracy) as compared to other classifiers: SVM, decision tree (DT), k-nearest neighbors (KNN), RF, and artificial neural network (ANN). The SHAP method was used to select the top-most significant statistics like PCR, respiratory rate, lymphocytes, diastolic blood pressure, P02 venous blood gas, GGT, neutrophils, urea concentration, creatinine, glucose, sodium, those aged above 65, TGO, and lactate, which all had important impacts on the performance of the XGBoost prediction model. While [18] used the ANN classifier to predict ICU admission using the same data set, their model achieved an accuracy of 96.9%, and the GSInquire method was used for gaining the quantitative effect of the clinical variables.

Alotaibi et al. [5] developed different machine-learning models using ANN, SVM, and RF. The relief algorithm was used for selecting the features by computing and assigning weights for the parameters according to their importance. The models were built based on a data set from Beijing Hospital in China. The best performance was achieved (with an accuracy of 90%) by employing ensemble RF using LPBoost, GentleBoost, and AdaBoostM1 decision trees.

Cheng et al. [16] developed a machine-learning-based risk-prioritization system to predict ICU severity within the first 24 hours with 76.2% accuracy using the RF classifier based on data from Mount Sinai Hospital in New York City. The features were selected by the importance of the Gini metric for training the RF model, which included respiratory rate, oxygen saturation, C-reactive protein, white blood cell count, systolic and diastolic blood pressures, blood urea nitrogen, anion gap, serum creatinine, and lymphocyte count features.

Different meta-heuristic algorithms have been used by different research studies to select the significant features of COVID-19 laboratory blood tests. [26] utilized the particle swarm optimization (PSO) algorithm with a Bayesian network classifier to build a diagnosis model of COVID-19 disease with 95.1% accuracy.

Also, genetic algorithm and AdaBoost classifiers [57] were exercised for a prediction model to distinguish the infected cases of COVID-19 based on clinical features and symptoms. In addition, the butterfly optimization algorithm (BOA) was hybridized with the PSO algorithm in [23] to determine the main attributes and features that provided an efficient decision for manipulating patients; it achieved an accuracy of 91.07%.

3. Proposed system for COVID-19 severity prediction

The details of the proposed prediction model are presented in this section; the structure of the model is introduced in Figure 1.

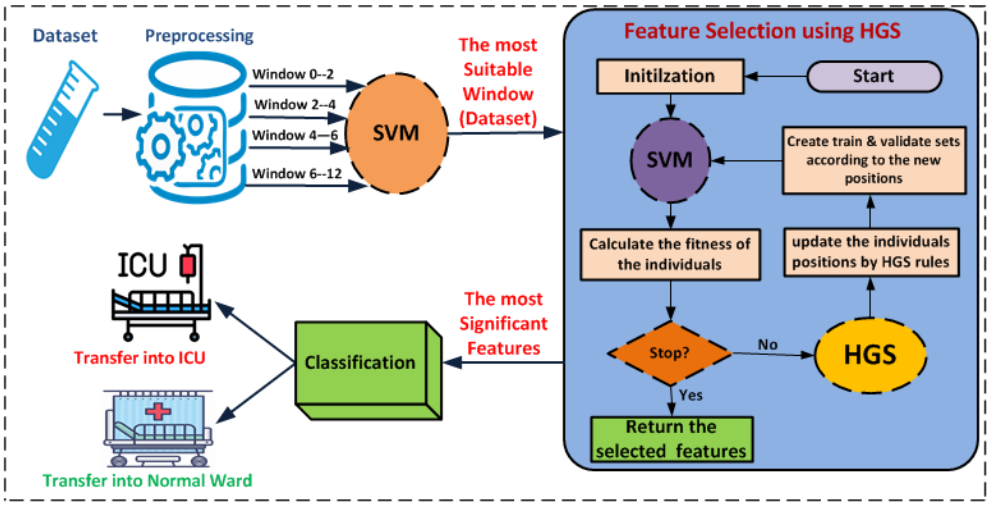

Figure 1. Structure of proposed prediction model

3.1. Data set

The study used a data set called "COVID-19 Clinical Data to Assess Diagnosis" to help specialists and hospitals predict a patient's needs for ICU admission. The data set is public and available on Kaggle [31]; it includes anonymized data from 384 COVID-19-positive patients from Hospital SÍrio-Libanês in São Paulo, Brazil. The data was aggregated by windows in chronological order to reflect the real-life scenarios of events for each patient.

There were five time windows for each patient, where time window "0–2" meant that the data had been taken from between 0 and 2 hours of the patient's admission, time window "2–4" denoted that the data had been obtained from between 2 and 4 hours of admission, and so on for the other time windows ("4–6", "6–12", and "above-12"). Each patient had four main categories of information: demographics, previously grouped diseases, blood lab tests, and vital signs.

According to the data set's description, there was some static information (usually once) like demographic and previous grouped diseases, while vital signs were sampled more frequently (usually hourly) than blood lab tests (usually daily) over the windows.

The description of the data for each category is described in Table 1, where the total was 54 features; when expanded to the mean, median, maximum (max), minimum (min), difference (diff), and relative differences when pertinent, this generated a total of 231 features. The original data set contained 1925 samples with 231 features.

Table 1
Description of data set

Category	Number of features	Data Examples
Demographic	3	Age – Gender
Pre-existing Diseases	9	hypertension; immunocompromised.
Blood lab tests	36	albumin serum, arterial base excess, venous base excess, arterial bicarbonate, venous bicarbonate, bilirubin, blast, calcium, creatintine, FFA, GGT, glucose, hematocrite, hemoglobin, INR, lactate, leukocytes, lymphocytes, neutrophiles, arterial PO2, venous PO2, arterial PCO2, venous PCO2, PCR, arterial pH, venous pH, platelets, potassium, arterial O2 saturation, venous O2 saturation, sodium, aspartate aminotransferase (AST/TGO), alanine aminotransferase (ALT/ TGP), PATT, urea, D-dimer.
Vital Signs	6	systolic and diastolic blood pressure, heart rate, respiratory rate, temperature, oxygen saturation.

3.2. Preprocessing

The publicly available data set required some preparation steps in order to be ready for use by the proposed model, such as the following:

1. Fill in the missing data: some patients may have a clinically stable state that does not have a new measurement in a time window, so the values may be like the values of neighboring windows as described by the data set's authors [31]. Therefore, the data for a specific window is filled using the next or the previous window of the same patient as recommenced in the data set description.
2. Convert the categorical variables into their values, like the age percentile variable (where, for example, the value "60th" is converted to 60, "above 90th" is transformed to 100, and so on).
3. Get rid of the duplicate data. When applying this step, the number of variables was reduced from 231 to 85.
4. Label the sample by the final decision of the patient, where "1" means the patient has ICU admission and "0" does not.

The descriptions of the data set did not recommend using the sample data when the target variable was available, as there were some problems in the order of the events [31]. As a result, the samples from the last event window that was labeled "above-12" was excluded, as were the samples from the patient's ICU admission. Then, the data of each window was separated to be a detached data set; so, the original data set was split into four sub-data sets (their descriptions are presented in Table 2).

Table 2
Description of data set windows

Data set Window	Number of samples	Samples per class
Window 0–2	353	0: 190 1: 163
Window 2–4	326	0: 190 1: 136
Window 4–6	286	0: 190 1: 96
Window 6–12	255	0: 190 1: 65

3.3. Support vector machine (SVM)

A support vector machine (SVM) is a machine-learning algorithm for solving regression and classification problems as optimization problems [19], where it finds an optimal line or hyperplane to separate data with different classes by the highest margin. The optimal line/hyperplane is the best decision boundary where several decision boundaries can separate data samples without error; however, the best one is that which has a maximum margin from the nearest points of the different classes [11].

SVM can work well with linearly and non-linearly separable data and achieve the best results. When dealing with non-linearly separable data, SVM employs the kernel technique to transform the data from a low-space dimension (nonseparable space) to a high-space dimension (separable space) and then seeks the best decision boundary that divides the data points with the greatest margins [21]. There are different types of kernels: sigmoid, polynomial, linear, and radial basis function (RBF), which are selected according to the type of data and those that yield good results [50]. In addition to kernels, there are two additional hyper-parameters: the cost (C) and gamma (γ) parameters [11] that are tuned by the developer when creating a classifier to achieve the best performance [27, 58]. The cost (C) regularizes the misclassification of the decision boundary by adding a penalty for each misclassified training sample [38]. The gamma γ parameter determines the distance of the effect of a single training data point. The values of these parameters in our model are described in Table 3. Algorithm 1 represents the steps of how the SVM classifier works.

The reason for applying SVM in the proposed model is that SVM has been used in various studies and has achieved promising results for detecting COVID-19 (like in [29, 44]) based on clinical blood results. Studies like [6, 33, 69] have used the SVM classifier and achieved promising results (like 80% accuracy, 81.4% accuracy, and 92.8% sensitivity, respectively) for predicting severity and mortality risks based on clinical blood tests.

Algorithm 1 SVM classifier

Input: train data, test data, and SVM parameters
Output: Accuracy(Acc)
1: **while** (stopping ?) **do**
2: Train SVM classifier using training data samples;
3: Test SVM trained_model using testing data samples;
4: Compute accuracy by comparing predicted test samples with original test samples.
5: **end while**
6: **return** Accuracy(Acc);

3.4. Feature selection by Hunger Game search (HGS) optimization algorithm

Hunger Game search (HGS) is a population-based optimization algorithm that was first presented in 2021 [68]. The optimization concept of the algorithm mimics the social conduct of animals when they are hungry. The idea of "hunger" is the main reason behind the actions and behaviors of animals when making decisions and optimizing food searches to save their lives. In this research, the methodology of the algorithm has been exploited for the feature-selection problem. Feature selection is an operation for finding out the most relevant features that have a significant impact on a machine-learning model. Many researchers have considered this problem to be an optimization problem [28, 45, 53], searching for the best features among all of the features. The steps for using HGS for feature selection are explained as follows:

1. Initialize the population of agents: initially, a group of agents (A) has been created with dimensions that are equal to the number of the data set features (F); then, they loop on these dimensions of each agent and apply the uniform random by Equation 1.

$$a_j^f = lwB^f + (upB^f - lwB^f).u \quad j = 0, .., A; \; f = 0, .., F \tag{1}$$

where a_j^f is the current agent; lwB^f and upB^f denote search space boundaries 0 and 1, respectively; for the feature-selection issue, u is an uniform random number between [0,1]. The result of this operation is a population of random initialized agents (A).

2. SVM classifier to compute fitness values of agents: evaluating the fitness values is the primary reason behind updating the positions of the agents to optimize the solutions. Since the feature-selection problem is a discrete problem, the real values of all solutions should be converted into binary values to deal with this problem. So, Equation 2 is first used to convert the solutions to binary versions.

$$bin_j^f = \begin{cases} 1, & \text{if } bin_j^f > 0.5 \\ 0, & otherwise \end{cases} \tag{2}$$

Then, the features that have "1" values are selected, while those with "0" values are discarded. The quality of the selected features of each solution is evaluated by the fitness value that is calculated by Equation 3.

$$fit_j = \lambda \,.\, (1 - Acc_j) + (1 - \lambda) \,.\, \frac{|bin_j|}{F} \tag{3}$$

where fit_j denotes the fitness value of agent j, $(1 - Acc_j)$ represents the classification error (where Acc_j is the accuracy of agent j computed by SVM), $\lambda = 0.99$ is for making the balance between the classification error and the number of selected features, $|bin_j|$ is the number of solution features that have "1" values, and F is the total number of all of the features.

3. HGS for updating positions The search strategy of HGS follows the following rules for reaching food:
Search individually: the agents in this category search individually for food without any participation from the team. They use Equation 4 to update their positions randomly.

$$\overrightarrow{A(i+1)} = \overrightarrow{A(i)}.(1 + rndm(1)), \qquad rd1 < l \tag{4}$$

where $\overrightarrow{A(i)}$ refers to the agents in the current iteration, $rndm(1)$ is a random number that is uniformly distributed, and $rd1$ is a random value that is within a range of [0,1].

Search with team participation: the agents in this category cooperate with each other, forming a team to search for food. Equations 5 and 6 mimic these types of search rules.

$$\overrightarrow{A(i+1)} = \overrightarrow{Wt_1}.\overrightarrow{A_b} + \overrightarrow{R}.\overrightarrow{Wt_2}.|\overrightarrow{A_b} - \overrightarrow{A(i)}|, \; rd1 > l, rd2 > E \tag{5}$$

$$\overrightarrow{A(i+1)} = \overrightarrow{Wt_1}.\overrightarrow{A_b} - \overrightarrow{R}.\overrightarrow{Wt_2}.|\overrightarrow{A_b} - \overrightarrow{A(i)}|, \; rd1 > l, rd2 < E \tag{6}$$

where $\overrightarrow{A_b}$ refers to the best agent in this iteration, $\overrightarrow{Wt_1}$ and $\overrightarrow{Wt_2}$ are the hunger weights, $\overrightarrow{R}$ is a ranging controller, $rd1$ and $rd2$ are random values that are within a range of [0,1], E is a variation control, and $\overrightarrow{R}$, $\overrightarrow{Wt_1}$, $\overrightarrow{Wt_2}$, and E are variables that are calculated by the equations that are explained in detail in the original paper [68].
As long as the termination iterations are not reached, the search process is conducted, and the agents update their positions by using Equations 4, 5, and 6 to reach the best global solution with the best-selected features. The steps of the HGS algorithm with the SVM classifier for the feature-selection operation are stated in Algorithm 2.

Algorithm 2 Step of HGS algorithm with SVM classifier for feature selection

Input: population size M, maximum number of iterations n_iters
Output: Selected features.

for each agent j ($\forall j = 1, 2, ...M$) **do**
  for each dimension f ($\forall f = 1, 2, ...F$) **do**
    $a_j^f(0) \leftarrow Eq.1$;
  end for
end for
while (i < n_iters) **do**
  Convert solutions to binary version by Eq. 2;
  Calculate fitness of agents using SVM algorithm 1 and Eq. 3;
  Sort agents according to their fitness values;
  Update best agent A_b, best fitness B_{fit}, and worst fitness W_{fit};
  Calculate list of hungry agents;
  for each agent j ($\forall j = 1, 2, ...M$) **do**
    for each dimension f ($\forall f = 1, 2, ...F$) **do**
      Calculate R, E, Wt_1, and Wt_2 variables using equations in [68];
      $rd1 = rand()$;
      $rd2 = rand()$;
      if ($rd1 < l$) **then**
        update solution by Eq. 4;
      else if $rd2 > E$ **then**
        update solution by Eq. 5;
      else
        update solution by Eq. 6;
      end if
    end for
  end for
  Locate best agent A_b that has best features based on highest fitness value;
  $i = i + 1$;
end while
return Best selected features.

Finally, map the selected feature indices of the best solution output from the previous feature-selection step onto the data set to create new train and test sub-sets with the most significant features. Then, the SVM classifier uses the new train and test sets to train and evaluate the model with these features by Algorithm 1.

3.5. Experimental setup

The experiments were evaluated by computing various measurements where Equations 7, 8, 9, and 10 were used for calculating the accuracy, recall, precision, and F1-score, respectively, of the proposed model.

$$acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$recall = \frac{TP}{TP + TN} \tag{8}$$

$$precision = \frac{TP}{TP + FP} \tag{9}$$

$$f1 - score = \frac{2.(Recall \times Precision)}{Recall + Precision} \tag{10}$$

where FP, FN, TP, and TN refer to False Positive, False Negative, True Positive, and True Negative, respectively.

The average performance of the proposed model was implemented by running each algorithm for 30 independent executions and computing the mean value of the mentioned metrics using Equation 11.

$$AVG_{metric} = \frac{\sum_i^{N_r} metric_i}{N_r} \tag{11}$$

The standard deviation of the algorithms over the different executions was computed by 12.

$$STD_{metric} = \sqrt{\frac{1}{N_r}.\sum_i^{N_r}\left(metric_i - AVG_{metric}\right)^2} \tag{12}$$

The hyper-parameters of the experiments, SVM, and the used optimization algorithms are shown in Table 3.

Table 3
Hyper-parameters of experiments

Method	Parameters
General	Number of optimization runs: $N_r = 30$ Population size: $M = 20$ Max iterations: $n\_iters = 500$ Lower bound: lwB = 0 Upper bound: upB = 1
SVM	$kernal = rbf$ $\gamma = 0.01$ $C = 10$
HGS	LH = 10, l = 0.25
FPA	$P = 0.8$, $\lambda = 1.5$
HHO	Levy component $\lambda = 1.5$
WOA	b = 1
BA	$\alpha = 0.9$, $\gamma = 0.9$, $f_{min} = 0$, $f_{max} = 2$, $A = 2$, $r = 1$

4. Experimental results and discussion

The main objectives of the proposed model were to detect the most suitable time for applying the medical tests correctly and to identify the most significant blood

indicators for the early prediction of COVID-19 patients' severity for ICU admission. So, the experiments were divided into the following sections (which corresponded to the study's objectives).

4.1. Best time window

The first objective of this study was to predict the best time to apply clinical tests to a patient in order to determine the severity level and her/his needs for ICU transfer. In order to accomplish this, the SVM classifier was applied to the four data sets (as described in Table 2) that contained diverse patient data with different time windows. The experiment used all of the original features of the data set: demographics, previous grouped diseases, blood lab results, and vital signs. The results are presented in Table 4. It is clear that the data of the (0–2) time window was the best time window for applying medical blood lab tests according to the values of accuracy, precision, recall, and F1-score.

Table 4
Results of different time windows using SVM and all features

Window	Accuracy	Precision	Recall	F1-score
0–2	91.5	92.3	91.1	91.4
2–4	83.3	83	82.7	82.8
4–6	84.4	84.9	79	80.9
6–12	84.3	84.9	71.7	75.1

4.2. Most significant features

Another objective of the study that would help decision-makers easily and correctly handle a patient case was to determine the most significant indicators in the blood data that would affect the performance of the proposed model. Five meta-heuristic optimization algorithms were used and compared for the feature-selection experiment (HGS, FPA, HHO, WOA, and BA), where SVM was used as an objective function. The experiment used the (0–2) time window data set as the outcome of the first experiment. For each optimization algorithm in the experiment, the measures in Section 3.5 like accuracy, recall, precision, and fitness values were computed to compare the performances of the competitive algorithms in selecting the most significant features among the variables of the clinical blood data.

Table 5 presents the accuracy results of the compared algorithms, where the HGS algorithm exceeded the competitors in terms of the mean, std, best, and worst values. HGS achieved a promising mean accuracy value of 96.5%, while the second mean rank value was 93.6%. Also, the best solution over the different executions accomplished a 98.6% topmost accuracy. The performance of the HGS algorithm against the other compared algorithms over the 30 distinct runs is displayed in Figure 2

to show the boxplots of the mean accuracy of the competitive algorithms, where the HGS algorithm recorded excellent performance over all the iterations compared with the others.

Table 5

Mean, Std, Best, and Worst Accuracy values that were obtained from different optimizers

Algorithm	Mean	Std	Best	Worst
HGS	96.5	0.008	98.6	95.8
FPA	93.6	0.012	95.7	91.5
HHO	91.9	0.019	95.8	88.7
WOA	91.9	0.018	95.8	88.7
BA	93.3	0.015	97.2	90.1

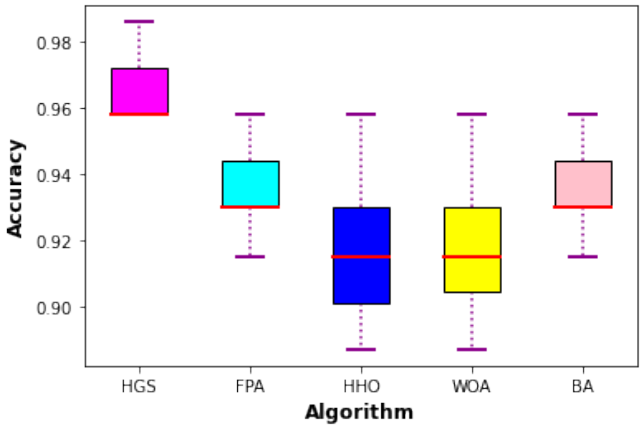

Figure 2. Boxplots of average accuracy levels that were obtained from different optimizers

The results that are illustrated in Tables 6 and 7 indicate that the HGS algorithm outperformed the competitor algorithms in terms of the recall and precision measurements, respectively, according to the mean, std, best, and worst values. HGS accomplished $\approx 96.6\%$ of the measures with the minimal Std value (which proved the stability of the algorithm), while the second rank was FPA with a value of $\approx 93.6\%$.

Table 6

Mean, Std, Best, and Worst Recall values that were obtained from different optimizers

Algorithm	Mean	Std	Best	Worst
HGS	96.4	0.009	98.7	95.5
FPA	93.6	0.012	96	91.3
HHO	91.8	0.019	95.9	89.1
WOA	91.8	0.019	95.9	88.5
BA	93.2	0.016	97.2	89.8

Table 7

Mean, Std, Best, and Worst Precision values that were obtained from different optimizers

Algorithm	Mean	Std	Best	Worst
HGS	96.6	0.008	98.5	95.7
FPA	93.7	0.012	95.8	91.5
HHO	92.1	0.019	95.9	89
WOA	92.1	0.019	95.9	88.7
BA	93.4	0.015	97.2	90.1

The fitness value results of the compared algorithms are stated in Table 8, where HGS surpassed the competitor algorithms by achieving the lowest mean cost value of 0.037, with a Std of 0.008 over the 30 different independent runs with the high-speed convergence (as shown in Figure 3). These obtained results refer to the ability of the HGS algorithm in exploring and exploiting the search space in order to find the best solution with the highest convergence as compared with the other algorithms.

Table 8

Mean, Std, Best, and Worst Fitness values that were obtained from compared optimiziers

Algorithm	Mean	Std	Best	Worst
HGS	0.037	0.008	0.017	0.045
FPA	0.067	0.011	0.045	0.088
HHO	0.084	0.019	0.044	0.116
WOA	0.083	0.018	0.046	0.116
BA	0.071	0.015	0.032	0.102

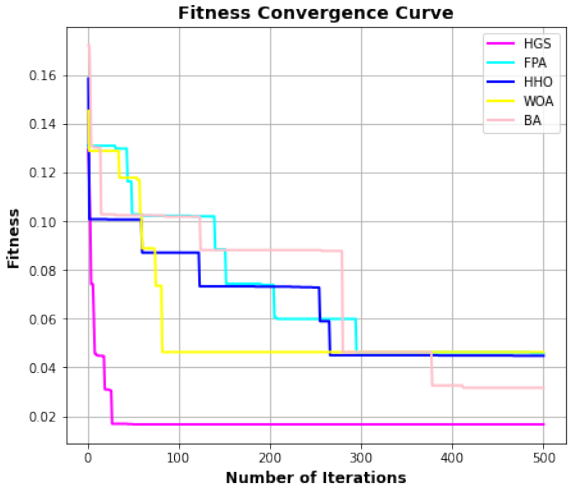

Figure 3. Fitness convergence curves of compared optimiziers

The obtained results in Table 9 revealed that HGS selected the fewer number of features where the mean result was 16.9 features, and the best solution uses the least number of features (23) with the best accuracy and fitness value. The boxplot results show that HGS selected the fewest number of features over the distinct runs against the competitors (as in Figure 4).

Table 9
Mean, Std, and Best number of features values that were obtained from different optimizers

Algorithm	Mean	Std	Best
HGS	16.9	5.05	23
FPA	37.5	3.97	43
HHO	31.4	8.5	25
WOA	29.9	7.8	38
BA	37.8	4.2	32

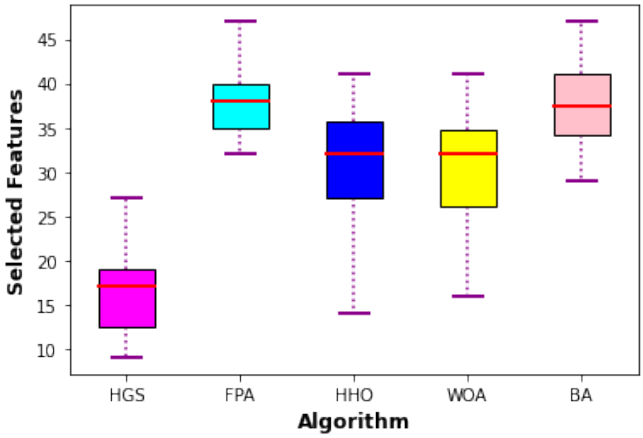

Figure 4. Boxplots of selected features that were obtained from different optimiziers

The HGS algorithm was evaluated statistically using the Wilcoxon signed-rank test. Wilcoxon is a statistical non-parametric test that is meant to assess any significant difference between two different meta-heuristic algorithms. The p-value result of this test indicated that there was a difference between the algorithms when its value was less than .05; otherwise, the algorithms were identical [62]. Table 10 presents a p-value comparison between HGS and the other competitor algorithms; it is obvious that HGS was statistically more significant where all of the p-values were $\leqslant 0.05$.

Table 10
p-values of Wilcoxon test between HGS and competitor algorithms

Algorithm	p-value
HGS vs FPA	4.7E-06
HGS vs HHO	1.5E-06
HGS vs WOA	2.3E-06
HGS vs BA	3.4E-06

Finally, to compare the correctness of the proposed model (HGS-SVM) with the most related studies in the literature, Table 11 reports the classifiers, the feature-selection method (FS-Method), whether a time series was applied or not (TS-Applied), and the accuracy of the related research that used the same public Kaggle data set [31] to predict the severity of COVID-19 patients for ICU admission. It is worth mentioning that most of these studies did not take the time series of the data into account in their work in spite of the description of using the time for the data. The superiority of our model (HGS-SVM) over other published studies is obvious despite the fact that we used the time factor when building our model.

Table 11
Comparison of proposed model (HGS-SVM) with existing studies

Study	Classifier	FS-Method	TS-Applied	Accuracy
[1]	GBM	SHAP	No	89%
[22]	XGBoost	SHAP	No	97%
[18]	ANN	GSInquire	No	96.9%
HGS-SVM	SVM	HGS	Yes	98.6%

4.3. Justifications of selected features

According to the aforementioned results, the HGS algorithm achieved the topmost results, where its best solution accomplished an accuracy of 98.6% with 23 selected features. The selected features were the age variable from the demographic category, 3 types of pre-existing diseases, 13 features from blood lab parameters (venous base excess, creatinine, FFA, hematocrit, lactate, lymphocytes, neutrophils, venous PO2, arterial PH, platelets, potassium, venous O2 saturation, and sodium), and 6 features from vital sign variables (systolic blood pressure, diastolic blood pressure, median value of heart rate, max value of heart rate, respiratory rate, and temperature).

The age feature from the demographic category was selected as one of the most significant features in the model. Many studies [10, 35, 42, 48, 66] have been published that have demonstrated the importance of age in disease progression – particularly in elderly patients.

The increased serum creatinine in COVID-19 patients led to acute kidney injury (AKI) in 5.1% of the patients [17]; this is a serious factor in risk assessment that

suits the outcomes of various models [12, 65]. Raised free fatty acids (FFAs) cause lipo-toxicity, leading to harm to some organs such as the pancreas and liver; they also cause some problems with insulin and proinflammatory [46], which increase the patient's severity. An increase in lactate levels is a factor of cell damage and one of the most vital markers of lung injury [36]. Elevated lactate levels refer to the worst outcome [30]; this feature was selected by many other statistical and machine-learning methods [1, 35, 39, 66, 67].

It has been noted that a low value of lymphocytes is an important indicator for the severity of a patient where lymphocytes are responsible for preserving immune homeostasis and the response of inflammatory cells; this may expose the patient to death [1, 14, 25, 59]. The opposite is true for neutrophils, where a high value of this test reflects the severity of the illness; this is consistent with the findings of the machine-learning models in [1, 35]. Neutrophils are white blood cells, and their activation is related to the high severity of enabling immune cells to defend against attack [8, 60]. Hematocrits and platelets are serious variables that are related to sepsis, and low values of these tests are indices for developing a patient's state [32,34]. The blood gas analysis variables (venous base excess, venous PO2, arterial PH, and venous O2 saturation) are among the leading features that agree with the findings in [1,24,25,39,61]. Also, abnormalities in potassium and sodium have been associated with severe disease [3, 41].

The respiratory rate is one of the most common vital sign tests for detecting severity, where a high value indicates a critical symptom [1, 71]. Also, systolic and diastolic blood pressure, heart rate, and temperature are among the most significant features in severely affected COVID-19 patients [49, 51, 71].

The selection of such features by our model with evidence of their significance from other published research increases the reliability and stability of the proposed model and emphasizes the importance of the selected features.

5. Conclusions

Some COVID-19 patients may be in a good state that requires only minimal treatment in a hospital ward or staying at home, while others are exposed to blood clots and a lack of oxygen (which require critical care in an intensive care unit [ICU]). Making correct early decisions for these patients is an important matter for two reasons: first – to save patients' lives, and second – to conserve hospital resources and give priority according to the level of the severity of patients' health situations. Clinical blood samples are widely available and are used in all hospitals, so this study proposes a machine-learning model for the automatic prediction of the COVID-19 patients' needs for incentive care based on blood lab tests for predicting ICU transfers to assist doctors with decision-making methods that facilitate the efficient use of hospital resources. The proposed model used a public time series window data set to identify the most suitable time window for applying medical tests correctly. Then, the Hunger Game search (HGS) meta-heuristic algorithm and support vector

machine (SVM) classifier were integrated for building the proposed prediction model and selecting the most informative features from the blood test data of the suitable time window. The experiments proved that using HGS for selecting the features with the SVM classifier achieved excellent results as compared to the results of four public meta-heuristic algorithms: the flower pollination algorithm (FPA), Harris Hawks optimization (HHO), the whale optimization algorithm (WOA), and the bat algorithm (BA). According to the results, the proposed model also selected the most informative indicators in the input blood data, where the obtained features were validated with the results of the other published studies to increase the consistency of the proposed model and assert the significance of the selected features.

Despite the high-level results of the proposed model, the study has some limitations that may be addressed in future work. First, the proposed HGS-SVM model was used as a binary classification to predict only a patient's severity for ICU admission, which does not support the mortality or different levels of severity; this may be solved by training the model based on a data set that includes different types of severities. Also, the size of the used data set was small, with some imbalance in the data window samples that can be solved by the availability of more data with different cases to improve the model's results. The proposed model may suffer from a generalization where the model was validated based on a data set that was taken from one hospital and from one country; so, the generalization may be improved when the proposed model is validated externally by using new data sets from more hospitals and countries. The HGS algorithm was integrated with the SVM classifier, which will be executed with more machine-learning classifiers in future work to help compare and achieve the best results.

References

[1] Aktar S., Ahamad M.M., Rashed-Al-Mahfuz M., Azad A.K.M., Uddin S., Kamal A.H.M., Alyami S.A., *et al.*: Machine Learning Approach to Predicting COVID-19 Disease Severity Based on Clinical Blood Test Data: Statistical Analysis and Model Development, *JMIR Medical Informatics*, vol. 9(4), 2021.

[2] Alakus T.B., Turkoglu I.: Comparison of deep learning approaches to predict COVID-19 infection, *Chaos, Solitons & Fractals*, vol. 140, 2020.

[3] AlJame M., Ahmad I., Imtiaz A., Mohammed A.: Ensemble learning model for diagnosing COVID-19 from routine blood tests, *Informatics in Medicine Unlocked*, vol. 21, 2020.

[4] Aljameel S.S., Khan I.U., Aslam N., Aljabri M., Alsulmi E.S.: Machine Learning-Based Model to Predict the Disease Severity and Outcome in COVID-19 Patients, *Scientific Programming*, vol. 2021, 2021.

[5] Alotaibi A., Shiblee M., Alshahrani A.: Prediction of severity of COVID-19-infected patients using machine learning techniques, *Computers*, vol. 10(3), 2021.

[6] An C., Lim H., Kim D.W., Chang J.H., Choi Y.J., Kim S.W.: Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study, *Scientific Reports*, vol. 10(1), pp. 1–11, 2020.

[7] Arevalo-Rodriguez I., Buitrago-Garcia D., Simancas-Racines D., Zambrano-Achig P., Del Campo R., Ciapponi A., Sued O., Martinez-Garcia L., Rutjes A.W., Low N., *et al.*: False-negative results of initial RT-PCR assays for COVID-19: a systematic review, *PloS one*, vol. 15(12), p. e0242958, 2020.

[8] Aschenbrenner A.C., Mouktaroudi M., Krämer B., Oestreich M., Antonakos N., Nuesch-Germano M., Gkizeli K., *et al.*: Disease severity-specific neutrophil signatures in blood transcriptomes stratify COVID-19 patients, *Genome Medicine*, vol. 13(1), pp. 1–25, 2021.

[9] Bajaj V., Gadi N., Spihlman A.P., Wu S.C., Choi C.H., Moulton V.R.: Aging, immunity, and COVID-19: how age influences the host immune response to coronavirus infections?, *Frontiers in Physiology*, vol. 11, 2021.

[10] Bajgain K.T., Badal S., Bajgain B.B., Santana M.J.: Prevalence of comorbidities among individuals with COVID-19: A rapid review of current literature, *American Journal of Infection Control*, vol. 49(2), pp. 238–246, 2021.

[11] Ben-Hur A., Weston J.: A user's guide to support vector machines. In: *Data mining techniques for the life sciences*, pp. 223–239, Springer, 2010.

[12] Bertsimas D., Lukin G., Mingardi L., Nohadani O., Orfanoudaki A., Stellato B., Wiberg H., *et al.*: COVID-19 mortality risk assessment: An international multicenter study, *PloS one*, vol. 15(12), 2020.

[13] Bravata D.M., Perkins A.J., Myers L.J., Arling G., Zhang Y., Zillich A.J., Reese L., *et al.*: Association of intensive care unit patient load and demand with mortality rates in US Department of Veterans Affairs hospitals during the COVID-19 pandemic, *JAMA Network Open*, vol. 4(1), 2021.

[14] Cascella M., Rajnik M., Aleem A., Dulebohn S.C., Di Napoli R.: Features, evaluation, and treatment of coronavirus (COVID-19), *StatPearls*, 2021.

[15] Cascella M., Rajnik M., Aleem A., Dulebohn S.C., Di Napoli R.: Features, evaluation, and treatment of coronavirus (COVID-19), *StatPearls*, 2022.

[16] Cheng F.Y., Joshi H., Tandon P., Freeman R., Reich D.L., Mazumdar M., Kohli-Seth R., *et al.*: Using machine learning to predict ICU transfer in hospitalized COVID-19 patients, *Journal of Clinical Medicine*, vol. 9(6), 2020.

[17] Cheng Y., Luo R., Wang K., Zhang M., Wang Z., Dong L., Li J., *et al.*: Kidney disease is associated with in-hospital death of patients with COVID-19, *Kidney International*, vol. 97(5), pp. 829–838, 2020.

[18] Chung A., Famouri M., Hryniowski A., Wong A.: COVID-Net Clinical ICU: Enhanced Prediction of ICU Admission for COVID-19 Patients via Explainability and Trust Quantification, *arXiv preprint arXiv:210906711*, 2021.

[19] Corinna C., Vapnik V.: Support-Vector Networks, *Machine Learning*, vol. 20, pp. 273–297, 1995. doi: 10.1023/A:1022627411411.

[20] Corman V.M., Landt O., Kaiser M., Molenkamp R., Meijer A., Chu D.K., Bleicker T., *et al.*: Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR, *Eurosurveillance*, vol. 25(3), 2020.

[21] Cover T.M.: Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition, *IEEE Transactions on Electronic Computers*, vol. EC-14(3), pp. 326–334, 1965. doi: 10.1109/PGEC.1965.264137.

[22] Deif M.A., Solyman A.A., Alsharif M.H., Uthansakul P.: Automated Triage System for Intensive Care Admissions during the COVID-19 Pandemic Using Hybrid XGBoost-AHP Approach, *Sensors*, vol. 21(19), 2021.

[23] EL-Hasnony I.M., Elhoseny M., Tarek Z.: A hybrid feature selection model based on butterfly optimization algorithm: COVID-19 as a case study, *Expert Systems*, 2021. doi: 10.1111/exsy.12786.

[24] Elezagic D., Johannis W., Burst V., Klein F., Streichert T.: Venous blood gas analysis in patients with COVID-19 symptoms in the early assessment of virus positivity, *Journal of Laboratory Medicine*, vol. 45(1), pp. 27–30, 2021.

[25] de Fátima Cobre A., Stremel D.P., Noleto G.R., Fachi M.M., Surek M., Wiens A., Tonin F.S., Pontarolo R.: Diagnosis and prediction of COVID-19 severity: can biochemical tests and machine learning be used as prognostic indicators?, *Computers in Biology and Medicine*, 2021.

[26] de Freitas Barbosa V.A., Gomes J.C., de Santana M.A., Albuquerque de J.E., de Souza R.G., de Souza R.E., dos Santos W.P.: Heg. IA: an intelligent system to support diagnosis of Covid-19 based on blood tests, *Research on Biomedical Engineering*, vol. 38, pp. 99–116, 2022.

[27] Friedrichs F., Igel C.: Evolutionary tuning of multiple SVM parameters, *Neurocomputing*, vol. 64, pp. 107–117, 2005.

[28] Ghaemi M., Feizi-Derakhshi M.R.: Feature selection using forest optimization algorithm, *Pattern Recognition*, vol. 60, pp. 121–129, 2016.

[29] Guhathakurata S., Kundu S., Chakraborty A., Banerjee J.S.: A novel approach to predict COVID-19 using support vector machine. In: *Data Science for COVID-19*, pp. 351–364, Elsevier, 2021.

[30] Henry B.M., Aggarwal G., Wong J., Benoit S., Vikse J., Plebani M., Lippi G.: Lactate dehydrogenase levels predict coronavirus disease 2019 (COVID-19) severity and mortality: a pooled analysis, *The American Journal of Emergency Medicine*, vol. 38(9), pp. 1722–1726, 2020.

[31] Hospital Sírio-Libanês: COVID-19 – Clinical Data to assess diagnosis, 2020. https://www.kaggle.com/S%C3%ADrio-Libanes/covid19.

[32] Jiang S.Q., Huang Q.F., Xie W.M., Lv C., Quan X.Q.: The association between severe COVID-19 and low platelet count: evidence from 31 observational studies involving 7613 participants, *British Journal of Haematology*, vol. 190, pp. e29–e33, 2020.

[33] Jiang X., Coffee M., Bari A., Wang J., Jiang X., Huang J., Shi J., *et al.*: Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity, *Computers, Materials & Continua*, vol. 63(1), pp. 537–551, 2020.

[34] Karimi Shahri M., Niazkar H.R., Rad F.: COVID-19 and hematology findings based on the current evidences: a puzzle with many missing pieces, *International Journal of Laboratory Hematology*, vol. 43(2), pp. 160–168, 2021.

[35] Karthikeyan A., Garg A., Vinod P., Priyakumar U.D.: Machine learning based clinical decision support system for early COVID-19 mortality prediction, *Frontiers in Public Health*, vol. 9, 2021.

[36] Kishaba T., Tamaki H., Shimaoka Y., Fukuyama H., Yamashiro S.: Staging of acute exacerbation in patients with idiopathic pulmonary fibrosis, *Lung*, vol. 192(1), pp. 141–149, 2014.

[37] Kojadinovic I., Wottka T.: Comparison between a filter and a wrapper approach to variable subset selection in regression problems. In: *Proceedings European Symposium on Intelligent Techniques (ESIT)*, pp. 311–321, ESIT, Aachen, Germany, 2000.

[38] Lameski P., Zdravevski E., Mingov R., Kulakov A.: SVM parameter tuning with grid search and its impact on reduction of model over-fitting. In: *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing. 15th International Conference, RSFDGrC 2015, Tianjin, China, November 20–23, 2015, Proceedings,*, pp. 464–474, Springer, 2015.

[39] Li X., Ge P., Zhu J., Li H., Graham J., Singer A., Richman P.S., Duong T.Q.: Deep learning prediction of likelihood of ICU admission and mortality in COVID-19 patients using clinical variables, *PeerJ*, vol. 8, 2020.

[40] Li Z., Yi Y., Luo X., Xiong N., Liu Y., Li S., Sun R., *et al.*: Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis, *Journal of Medical Virology*, vol. 92(9), pp. 1518–1524, 2020.

[41] Lippi G., South A.M., Henry B.M.: Electrolyte imbalances in patients with severe coronavirus disease 2019 (COVID-19), *Annals of Clinical Biochemistry*, vol. 57(3), pp. 262–265, 2020.

[42] Liu Y., Mao B., Liang S., Yang J.W., Lu H.W., Chai Y.H., Wang L., *et al.*: Association between age and clinical characteristics and outcomes of COVID-19, *European Respiratory Journal*, vol. 55(5), 2020. doi: 10.1183/13993003.01112-2020.

[43] Lundberg S.M., Lee S.I.: A Unified Approach to Interpreting Model Predictions. In: *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, vol. 30, Curran Associates, Inc., 2017. https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.

[44] de Moraes Batista A.F., Miraglia J.L., Rizzi Donato T.H., Porto Chiavegatto Filho A.D.: COVID-19 diagnosis prediction in emergency care patients: a machine learning approach, *medRxiv*, 2020. doi: 10.1101/2020.04.04.20052092.

[45] Nabil E., Sayed S.A.F., Hameed H.A.: An efficient binary clonal selection algorithm with optimum path forest for feature selection, *International Journal of Advanced Computer Science and Applications*, vol. 11(7), 2020.

[46] Nakeshbandi M., Maini R., Daniel P., Rosengarten S., Parmar P., Wilson C., Kim J.M., *et al.*: The impact of obesity on COVID-19 complications: a retrospective cohort study, *International Journal of Obesity*, vol. 44(9), pp. 1832–1837, 2020.

[47] Onan A., Korukoğlu S.: A feature selection model based on genetic rank aggregation for text sentiment classification, *Journal of Information Science*, vol. 43(1), pp. 25–38, 2017.

[48] Pan A., Liu L., Wang C., Guo H., Hao X., Wang Q., Huang J., *et al.*: Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China, *Jama*, vol. 323(19), pp. 1915–1923, 2020.

[49] Peer N., Lombard C., Steyn K., Levitt N.: Elevated resting heart rate is associated with several cardiovascular disease risk factors in urban-dwelling black South Africans, *Scientific Reports*, vol. 10(1), 2020.

[50] Phienthrakul T., Kijsirikul B.: Evolutionary strategies for multi-scale radial basis function kernels in support vector machines. In: *GECCO'05: Proceedings of the 7th annual conference on Genetic and evolutionary computation*, pp. 905–911, 2005.

[51] Ran J., Song Y., Zhuang Z., Han L., Zhao S., Cao P., Geng Y., *et al.*: Blood pressure control and adverse outcomes of COVID-19 infection in patients with concomitant hypertension in Wuhan, China, *Hypertension Research*, vol. 43(11), pp. 1267–1276, 2020.

[52] Sayed S.A.F., Elkorany A.M., Mohammad S.S.: Applying different machine learning techniques for prediction of COVID-19 severity, *IEEE Access*, vol. 9, pp. 135697–135707, 2021.

[53] Sayed S.A.F., Nabil E., Badr A.: A binary clonal flower pollination algorithm for feature selection, *Pattern Recognition Letters*, vol. 77, pp. 21–27, 2016.

[54] Schwab P., Schütte A., Dietz B., Bauer S.: predCOVID-19: A systematic study of clinical predictive models for coronavirus disease 2019, *arXiv preprint arXiv:200508302*, vol. 76, 2020.

[55] Sen-Crowe B., Sutherland M., McKenney M., Elkbuli A.: A closer look into global hospital beds capacity and resource shortages during the COVID-19 pandemic, *Journal of Surgical Research*, vol. 260, pp. 56–63, 2021.

[56] Shang Y., Pan C., Yang X., Zhong M., Shang X., Wu Z., Yu Z., *et al.*: Management of critically ill patients with COVID-19 in ICU: statement from front-line intensive care experts in Wuhan, China, *Annals of Intensive Care*, vol. 10(1), 2020.

[57] Soui M., Mansouri N., Alhamad R., Kessentini M., Ghedira K.: NSGA-II as feature selection technique and AdaBoost classifier for COVID-19 prediction using patient's symptoms, *Nonlinear Dynamics*, vol. 106, pp. 1453–1475, 2021.

[58] Syarif I., Prugel-Bennett A., Wills G.: SVM parameter optimization using grid search and genetic algorithm to improve classification performance, *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 14(4), pp. 1502–1509, 2016.

[59] Tan L., Wang Q., Zhang D., Ding J., Huang Q., Tang Y.Q., Wang Q., Miao H.: Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study, *Signal Transduction and Targeted Therapy*, vol. 5(1), 2020.

[60] Uhl B., Vadlau Y., Zuchtriegel G., Nekolla K., Sharaf K., Gaertner F., Massberg S., Krombach F., Reichel C.A.: Aged neutrophils contribute to the first line of defense in the acute inflammatory response, *Blood*, vol. 128(19), pp. 2327–2337, 2016.

[61] Wang K., Zuo P., Liu Y., Zhang M., Zhao X., Xie S., Zhang H., Chen X., Liu C.: Clinical and laboratory predictors of in-hospital mortality in patients with COVID-19: a cohort study in Wuhan, China, *Clinical Infectious Diseases*, vol. 71(16), pp. 2079–2088, 2020.

[62] Wilcoxon F.: Individual comparisons by ranking methods. In: *Breakthroughs in statistics. Volume II. Methodology and Distribution*, pp. 196–202, Springer, 1992.

[63] World Health Organization: Clinical management of COVID-19: interim guidance, 27 May 2020. Technical report, 2020.

[64] World Health Organization: Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases: interim guidance, 2 March 2020. Technical report, 2020.

[65] Wu J., Zhang P., Zhang L., Meng W., Li J., Tong C., Li Y., *et al.*: Rapid and accurate identification of COVID-19 infection through machine learning based on clinical available blood test results, *MedRxiv*, 2020.

[66] Xu W., Sun N.N., Gao H.N., Chen Z.Y., Yang Y., Ju B., Tang L.L.: Risk factors analysis of COVID-19 patients with ARDS and prediction based on machine learning, *Scientific Reports*, vol. 11(1), 2021.

[67] Yan L., Zhang H.T., Goncalves J., Xiao Y., Wang M., Guo Y., Sun C., *et al.*: An interpretable mortality prediction model for COVID-19 patients, *Nature Machine Intelligence*, vol. 2(5), pp. 283–288, 2020.

[68] Yang Y., Chen H., Heidari A.A., Gandomi A.H.: Hunger games search: Visions, conception, implementation, deep analysis, perspectives, and towards performance shifts, *Expert Systems with Applications*, vol. 177, 2021. doi: 10.1016/j.eswa.2021.114864.

[69] Yao H., Zhang N., Zhang R., Duan M., Xie T., Pan J., Peng E., *et al.*: Severity detection for the coronavirus disease 2019 (COVID-19) patients using a machine learning model based on the blood and urine tests, *Frontiers in Cell and Developmental Biology*, vol. 8, 2020.

[70] Yasin R., Gouda W.: Chest X-ray findings monitoring COVID-19 disease course and severity, *Egyptian Journal of Radiology and Nuclear Medicine*, vol. 51(1), 2020.

[71] Yu L., Halalau A., Dalal B., Abbas A.E., Ivascu F., Amin M., Nair G.B.: Machine learning methods to predict mechanical ventilation and mortality in patients with COVID-19, *PloS One*, vol. 16(4), 2021.

[72] Zhu N., Zhang D., Wang W., Li X., Yang B., Song J., Zhao X., *et al.*: A novel coronavirus from patients with pneumonia in China, 2019, *New England Journal of Medicine*, vol. 382(8), pp. 727–733, 2020.

Affiliations

Safynaz AbdEl-Fattah Sayed
Beni-Suef University, Faculty of Computers and Artificial Intelligence, 62521, Egypt,
safyfci@gmail.com, safynaz@bsu.edu.eg

Abeer ElKorany
Cairo University, Faculty of Computers and Artificial Intelligence, 12613, Egypt,
a.korani@fci-cu.edu.eg

Sabah Sayed
Cairo University, Faculty of Computers and Artificial Intelligence, 12613, Egypt,
s.sayed@fci-cu.edu.eg

Received: 21.01.2022
Revised: 01.07.2022
Accepted: 09.07.2022