



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Early prediction of COVID-19 outcome using artificial intelligence techniques and only five laboratory indices

Panagiotis G. Asteris<sup>a</sup>, Styliani Kokoris<sup>b,\*</sup>, Eleni Gavriilaki<sup>c</sup>, Markos Z. Tsoukalas<sup>a</sup>, Panagiotis Houpas<sup>a</sup>, Maria Paneta<sup>d</sup>, Andreas Koutzas<sup>e</sup>, Theodoros Argyropoulos<sup>f</sup>, Nizar Faisal Alkayem<sup>g</sup>, Danial J. Armaghani<sup>h</sup>, Abidhan Bardhan<sup>i</sup>, Liborio Cavaleri<sup>j</sup>, Maosen Cao<sup>g</sup>, Iman Mansouri<sup>k</sup>, Ahmed Salih Mohammed<sup>l</sup>, Pijush Samui<sup>i</sup>, Gloria Gerber<sup>m</sup>, Dimitrios T. Boumpas<sup>n</sup>, Argyrios Tsantes<sup>b</sup>, Evangelos Terpos<sup>o,1</sup>, Meletios A. Dimopoulos<sup>o,1</sup>

<sup>a</sup> Computational Mechanics Laboratory, School of Pedagogical and Technological Education, Athens, Greece

<sup>b</sup> Laboratory of Hematology and Hospital Blood Transfusion Department, University General Hospital "Attikon", National and Kapodistrian University of Athens, Medical School, Greece

<sup>c</sup> Hematology Department – BMT Unit, G Papanicolaou Hospital, Thessaloniki, Greece

<sup>d</sup> Fourth Department of Internal Medicine, University General Hospital "Attikon", National and Kapodistrian University of Athens, Medical School, Greece

<sup>e</sup> Abbott Laboratories (Hellas) SA, Athens, Greece

<sup>f</sup> Gastroenterology Department, Red Cross General Hospital of Athens, Greece

<sup>g</sup> Jiangxi Province Key Laboratory of Environmental Geotechnical Engineering and Hazards Control, Jiangxi University of Science and Technology, Ganzhou 341000, China

<sup>h</sup> Department of Urban Planning, Engineering Networks and Systems, Institute of Architecture and Construction, South Ural State University, 76, Lenin Prospect, Chelyabinsk 454080, Russian Federation

<sup>i</sup> Civil Engineering Department, National Institute of Technology Patna, Bihar, India

<sup>j</sup> Department of Civil, Environmental, Aerospace and Materials Engineering, University of Palermo, Palermo, Italy

<sup>k</sup> Department of Civil and Environmental Engineering, Princeton University Princeton, Princeton, NJ 08544, USA

<sup>l</sup> Engineering Department, American University of Iraq, Sulaimani, Kurdistan-Region, Iraq

<sup>m</sup> Hematology Division, Johns Hopkins University, Baltimore, USA

<sup>n</sup> "Attikon" University Hospital of Athens, Rheumatology and Clinical Immunology, Medical School, National and Kapodistrian University of Athens, Athens, Attica, Greece

<sup>o</sup> Department of Clinical Therapeutics, Medical School, Faculty of Medicine, National Kapodistrian University of Athens, Athens, Greece

### ARTICLE INFO

#### Keywords:

COVID-19  
SARS-CoV2  
Laboratory indices  
Artificial intelligence  
Artificial neural networks

### ABSTRACT

We aimed to develop a prediction model for intensive care unit (ICU) hospitalization of Coronavirus disease-19 (COVID-19) patients using artificial neural networks (ANN). We assessed 25 laboratory parameters at first from 248 consecutive adult COVID-19 patients for database creation, training, and development of ANN models. We developed a new alpha-index to assess association of each parameter with outcome. We used 166 records for training of computational simulations (training), 41 for documentation of computational simulations (validation), and 41 for reliability check of computational simulations (testing). The first five laboratory indices ranked by importance were Neutrophil-to-lymphocyte ratio, Lactate Dehydrogenase, Fibrinogen, Albumin, and D-Dimers. The best ANN based on these indices achieved accuracy 95.97%, precision 90.63%, sensitivity 93.55% and F1-score 92.06%, verified in the validation cohort. Our preliminary findings reveal for the first time an ANN to predict ICU hospitalization accurately and early, using only 5 easily accessible laboratory indices.

### 1. Introduction

The Coronavirus disease-19 (COVID-19) pandemic has

unprecedented impact on all aspects of human activity worldwide. Although vaccination, anti-viral treatment and monoclonal antibodies have positively impacted the course of this pandemic [1], reduction of

\* Corresponding author.

E-mail address: [skokori@med.uoa.gr](mailto:skokori@med.uoa.gr) (S. Kokoris).

<sup>1</sup> These authors have contributed equally.

morbidity and mortality due to severe COVID-19 remains an unmet clinical need. In this context, several groups worldwide have investigated routine laboratory, biochemical and coagulation indices in severe COVID-19. Summarizing the early studies, Terpos et al. have highlighted that lymphopenia, neutrophil/lymphocyte ratio (NLR), inflammatory indices, including lactate dehydrogenase (LDH), and C-reactive protein (CRP), as well coagulation abnormalities, such as prothrombin (PT) time prolongation, are important in COVID-19 [2]. In addition, systematic reviews and meta-analyses have tried to identify biomarkers associated with COVID-19 progression [3,4]. Among them, NLR, ferritin and LDH have been the major markers recognized as predictors of high-risk patients [5,6].

During the last two years, a plethora of studies employing ANNs have also emerged focusing on COVID-19 prediction problems [7–13]. Development of artificial neural networks (ANNs) is based on the concept of the biological neural network of the human brain and was initially used for medicine research purposes to simulate strongly non-linear relationships between numerous input and output parameters [14–20]. ANN models were subsequently introduced into the wider context of engineering disciplines [21–23], which significantly enriched the mathematical background underpinning ANN.

Despite scientific breakthroughs during COVID-19, there is no risk prediction model that can be used in clinical practice to identify COVID-19 severity early in the disease course. Our group has recently described risk prediction models and artificial neural networks (ANN) with high sensitivity and specificity for prediction of both morbidity and mortality from COVID-19 [8]. Nevertheless, these models were based on genetic susceptibility identifying complement genetic variants in COVID-19 patients [7,24,25], requiring genetic screening for several mutations which cannot be applied in the general population.

Therefore, we designed the present study aiming to provide a risk prediction model of COVID-19 outcome using ANN and the minimum number of routine laboratory indices. To increase the value of our model as an early predictor for clinical practice, we utilized laboratory indices from the day of first presentation to the Emergency department that will be further referred to as admission.

## 2. Materials and methods

### 2.1. Artificial neural networks

The basic building block of ANNs is the artificial neuron, which is a mathematical model mimicking the behavior of the biological neuron. Information is passed onto the artificial neuron as a vector of input parameters and is then processed using a mathematical function to derive an output which determines the behavior of the neuron (similar to the fire-or-not situation of the biological neuron). Before the information enters the neuron, it is weighted in order to process the inputs collectively. More precisely, each artificial neuron has inputs and produces a single output which can be sent to multiple other neurons. The inputs can be the feature values of a sample of external data, such as the values of a set of indicators, or they can be the outputs of other neurons. The outputs of the final output neurons of the neural network accomplish the task, such as classifying the input vectors. A group of such neurons comprise an ANN, similar to the structure of biological neural networks. To define an ANN, (i) the architecture of the ANN, (ii) the training algorithm used during the ANN’s training stage and (iii) the mathematical functions underpinning the mathematical model are required. The architecture of the ANN defines how the artificial neurons are organized and how the information flows within the network. In most applications the neurons are organized in a series of layers comprising a network called a multilayer ANN. The training stage can be considered as a function minimization problem, in which the optimum weight values need to be determined by minimizing an error function which is the sum of squared differences between outputs of each neuron in the last layer to each component of the vector of observations. This is

done by a special learning algorithm that adjusts the values of the weights of each neuron successively for each layer to achieve convergence to the minimum error and corresponding optimal values for the whole of weights.

### 2.2. Study population

We enrolled consecutive adult patients with COVID-19 from Attikon Hospital, Athens, Greece (May 2020–April 2021). All 248 patients met the following inclusion criteria: 1. routine laboratory testing (hematology, biochemical and coagulation profile) at admission, 2. RT-PCR test positive for SARS-COV2, 3. no history of COVID-19 vaccination. Table 1 summarizes demographics according to age, gender and disease severity (requiring or not hospitalization in intensive care unit (ICU)). The study was performed according to the Helsinki Declaration and approved by the Ethics Committee of Attikon Hospital (9.16/6/2020).

### 2.3. Compiled database

Laboratory parameters at admission were used for the creation of the database and the subsequent training and development of the ANN models. More specifically, 25 laboratory indices for each patient were determined which they comprised the input parameters of ANN models while output parameter was defined to be the outcome of the hospitalization for the patient. The output value is a binary parameter that takes value 2 or 1 according to the need for hospitalization or not of the patient in intensive care unit (ICU).

Table 2 presents statistical indices (minimum, average and maximum value as well as the standard deviation) for each of the 25 laboratory indices. The database with the patients’ hematological indices is appended to this paper as supplementary material in the excel file entitled Database – Hematological indices.

### 2.4. Proposed index for the assessment of laboratory indices

The training and development procedure of a mathematical simulation for prediction of disease outcome using the database of 25 laboratory indices, presented in the previous section, is considered challenging, given the huge number of possible combinations among them. More specifically, the possible combinations of the 25 laboratory indices are described by the following equation:

$$\text{Combinations} = \sum_{i=1}^{25} \frac{25!}{i!(25-i)!} = 33.554.431 \quad (1)$$

It should be also noted that for each of the above different combinations it is necessary to investigate 1.500.000 different architectures ANNs. For this purpose, in the present work a new index, the alpha-index is proposed, that has aimed to metrize each one of the 25 laboratory indices according to their impact on COVID-19 severity. Additionally, we aimed that this index is simple and clinically meaningful.

According to these, the below alpha-index is proposed based on the variation of the mean value of the laboratory index of patients that did

**Table 1**

Study population categorized by age, gender and disease severity (requiring or not hospitalization in intensive care unit (ICU)).

Severe	Age	All patients	Male	Female
	Up to 65	142	85	57
	Over 65	106	65	41
All patients (not in ICU and in ICU)	All Ages	248	150	98
	Up to 65	24	17	7
	Over 65	38	27	11
	All Ages	62	44	18
not in ICU	Up to 65	118	68	50
	Over 65	68	38	30
ICU	All Ages	186	106	80

**Table 2**

Database statistics. Minimum, average, maximum and standard deviation of the input parameters (25 hematological indices) and the output parameter (COVID-19 Severe, value equal to 1 means that the effected patient is requiring hospitalization in intensive care unit (ICU) while value equal to 2 means that the effected patient is not requiring hospitalization in intensive care unit (ICU)).

Nr	Variable	Symbol	Units	Statistics			
				Min	Average	Max	STD
1	White Blood Cell count per volume	WBC	4–10.5 kcell/ $\mu$ L	1.16	7.01	38.60	3.82
2	Neutrophil count per volume	Neut	0.9–7 kcell/ $\mu$ L	0.59	5.17	35.30	3.58
3	Lymphocyte count per volume	Lympho	0.7–3.7 kcell/ $\mu$ L	0.25	1.19	16.90	1.16
4	Neutrophil-to-lymphocyte ratio	NLR	0.8–3.6 (ratio)	0.19	6.14	36.31	5.94
5	Immature Granulocyte count per volume	IG	<0.03 kcell/ $\mu$ L	0.00	0.04	1.31	0.11
6	IG-to-lymphocyte ratio	IGLR	0.5–4.8 (ratio x100)	0.00	4.23	124.76	12.09
7	Hemoglobin concentration	Hb	12.0–16.9 g/dL	7.63	13.02	16.70	1.88
8	Mean Cell Volume of RBCs	MCV	79–98 fL	54.40	84.89	98.90	7.71
9	Red cell Distribution Width	RDW	11.5–14%	10.50	13.41	27.40	1.99
10	Percent microcytic RBCs	%MIC	0.2–2.5%	0.45	6.53	75.80	12.24
11	Percent macrocytic BCs	%MAC	0.1–9.1%	0.01	0.25	4.66	0.57
12	Percentage of hypochromic BCs	%HPO	0.1–1.4%	0.36	6.47	82.00	13.22
13	Hemoglobin Distribution Width	HDW	4.0–5.9%	5.03	7.13	15.40	1.85
14	Platelet cell count per volume	PLT	140–440 kcell/ $\mu$ L	40.30	219.90	766.00	94.24
15	D-Dimers	D-Dimers	<500 ng/mL	170.70	1134.04	8901.80	1258.01
16	C-reactive protein	CRP	0.00–6 mg/L	3.11	73.42	1170.00	98.80
17	Fibrinogen	Fib	200–400 mg/dL	199.00	489.39	948.50	149.55
18	Ferritin	Ferritin	13–150 ng/mL	5.00	748.66	17,158.00	1317.50
19	Pro-brain natriuretic peptide	ProBNP	< 125 pg/mL	5.30	1082.27	38,000.00	4259.80
20	Procalcitonin	pCT	<0.5 ng/mL	0.02	11.32	721.00	49.59
21	Lactate Dehydrogenase	LDH	135–225 U/L	103.00	316.22	1652.00	187.61
22	Total bilirubin	Tbil	< 1.20 mg/dL	0.11	0.50	5.02	0.42
23	Albumin	Alb	3.5–5.2 g/dL	1.90	3.92	5.50	0.52
24	Calcium (serum)	Ca + 2	8.6–10.3 mg/dL	6.40	8.74	11.30	0.64
25	Creatinine (serum)	Creat	0.5–0.9 mg/dL	0.30	1.24	11.40	1.49
26	COVID-19 Severe	COVID-19 Severe	1: not in ICU while 2: in ICU	1.00	1.25	2.00	0.43

not require ICU hospitalization compared to the mean value of the laboratory index of the patients that required ICU admission (Fig. 1):

$$\alpha(i) - \text{index} = \frac{\mu_i^{\text{notinICU}} - \mu_i^{\text{inICU}}}{\max(i) - \min(i)} \quad (2)$$

where

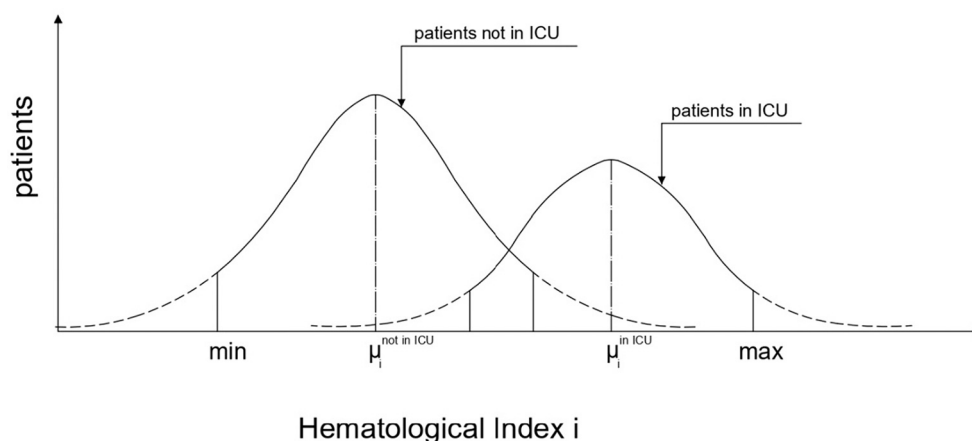
- $i$  is the laboratory index  $i$  ( $i = 1$  to 25),
- $\mu_i^{\text{notinICU}}$  the mean value of laboratory index  $i$  for COVID-19 infected patients who did not require hospitalization in ICU,
- $\mu_i^{\text{inICU}}$  the mean value of laboratory index  $i$  for COVID-19 infected patients who require hospitalization in ICU,
- $\min(i)$  the minimum value of laboratory index  $i$  for all COVID-19 infected patients, and,
- $\max(i)$  the maximum value of laboratory index  $i$  for all COVID-19 infected patients.

According to the above definition, a value of the index greater than zero means that increase of the value of the laboratory index drives into ICU, whereas value of the index less than zero means that decrease of the value of the laboratory index predicts ICU hospitalization. It is noteworthy that the index takes values between  $-100$  and  $100$ .

### 2.5. Performance analysis indices

For assessing the performance of the computational models that were trained and developed in the present study, we used the classical indices that are widely acceptable in the international literature for the assessment of binary classification problems, as shown below [1].

$$\text{Precision} = \frac{\text{TruePositive (TP)}}{\text{TruePositive} + \text{FalsePositive (FP)}} \quad (3)$$



**Fig. 1.** Shape, parameters and terminology of Gaussian function of the study population categorized by the need of requiring or not hospitalization in intensive care unit (ICU).

$$\text{Sensitivity or Recall} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative (FN)}} \quad (4)$$

$$F1 - \text{Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

For the successful application of the indices above, we clarified which are the two states that the computational models are required to estimate. With the term Positive we defined the case of the patient with COVID-19 hospitalized in ICU while the term Negative characterized the patient who did not require ICU hospitalization.

### 3. Results

#### 3.1. Ranking of laboratory indices based on the proposed new alpha-index

With use of the proposed alpha index, the values for each one of the 25 laboratory indices were computed. According to these, the laboratory indices were ranked in decreasing order, as presented in Fig. 2. The first five ranked by importance were Neutrophil-to-lymphocyte ratio (NLR), Lactate Dehydrogenase (LDH), Fibrinogen (Fib), Albumin (Alb) and D-Dimers. The reliability check of the proposed index was documented with the feasibility, the training, and development of ANN models with small number of laboratory indices and additionally those according to the importance ranking (input parameters) that predict reliably if the patient shall be admitted to ICU or not.

#### 3.2. Development of ANN models

In the current section we present the process of training and development of ANN models for the estimation of COVID-19 outcome. More specifically, models were trained and developed with respect to the

number of laboratory indices they use as input parameters. That is 10 different cases of ANN models with 1 to 10 with step 1 laboratory indices as input parameters following the order with which they were ranked according to the alpha index (Fig. 2).

ANN models were trained and designed with the use of the database comprised of 248 records. The 248 records correspond to the 248 patients studied. Each is comprised from 25 values of laboratory indices and the outcome of hospitalization of each patient. The 248 records were split in three parts and specifically, i) 166 (66.94%) records used for the training of the computational simulations (training dataset), ii) 41 (16.54%) records used for documentation of the computational simulations (validation dataset) and iii) 41 (16.54%) records used for the reliability check of the computational simulations (testing dataset).

Namely, for the work presented herein the testing datasets have been used as a cohort study to further support our findings. Table 3 shows a detailed presentation of the key parameters used for the training and development of the optimum ANN models, including the training algorithms, the number of neurons per hidden layer and the cost and transfer functions.

According to parameters used, models were trained and developed 1.500.000 different architectures ANN's for each one of the 10 different cases with respect to the number of laboratory indices as input parameters. Table 4 presents the respective 10 best architectures together with the performance indices for each one of them. According to this, the best ANN architecture model corresponds to the one with only 5 input parameters as follows:

1. Neutrophil-to-lymphocyte ratio (NLR),
2. Lactate Dehydrogenase (LDH),
3. Fibrinogen (Fib),
4. Albumin (Alb), and
5. D-Dimers.

The ANN model with only these five laboratory indices achieves extremely high performances with respect to the outcome of COVID-19. Furthermore, to the best of authors' knowledge, it achieves the best prediction from the totality of the computational models that have been proposed at present in the literature. More specifically it achieves

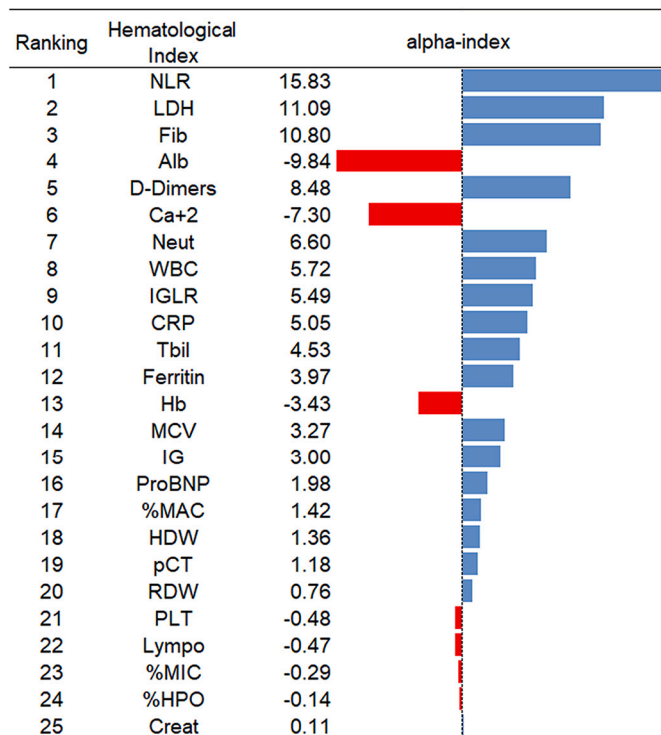


Fig. 2. Ranking of the hematological indices based on the proposed new alpha-index.

Table 3  
Training parameters of ANN models.

Parameter	Value
Training Algorithm	Levenberg–Marquardt algorithm (LM) algorithm
	particle swarm optimization (PSO) algorithm
	Grey Wolf Optimization (GWO) Algorithm algorithm
Normalization	Minmax in the range 0.10–0.90, 0.00–1.00
Number of Hidden Layers	1
Number of Neurons per Hidden Layer	1 to 50 by step 1
Control random number generation	10 different random generation
Maximum number of Epochs	250
Cost Function	Mean Square Error (MSE)
	Sum Square Error (SSE)
	Hyperbolic Tangent Sigmoid transfer function (HTS)
	Log-sigmoid transfer function (LS)
	Linear transfer function (Li)
	Positive linear transfer function (PLi)
	Symmetric saturating linear transfer function (SSL)
	Soft max transfer function (SM)
	Competitive transfer function (Co)
	Triangular basis transfer function (TB)
Radial basis transfer function (RB)	
Normalized radial basis transfer function (NRB)	

**Table 4**

Performance indices of the 10 developed Artificial Neural models. The number of models declare the number of hematological indices used as input parameters for the prediction of COVID-19 Patients outcome.

Model	Hematological Indices										Performance Indices		
	CRP	IGLR	WBC	Neut	Ca + 2	D-Dimers	Alb	Fib	LDH	NLR	Accuracy	Sensitivity or Recall	F1 – Score
1										•	50.00	50.00	66.67
2										•	79.03	79.03	88.29
3								•	•	•	87.10	87.10	93.10
4							•	•	•	•	91.94	91.94	95.80
5						•	•	•	•	•	93.55	93.55	96.67
6					•	•	•	•	•	•	90.32	90.32	94.92
7				•	•	•	•	•	•	•	91.94	91.94	95.80
8			•	•	•	•	•	•	•	•	88.71	88.71	94.02
9		•	•	•	•	•	•	•	•	•	88.71	88.71	94.02
10	•	•	•	•	•	•	•	•	•	•	90.32	90.32	94.92

accuracy 95.97%, precision 90.63%, sensitivity or recall 93.55% and F1-score 92.06%. Cross-validation techniques have been also used in our work presented here. It is worth stressing that the same results have been deduced both at the two halves of the database and at the whole. Five-fold cross-validation has been applied based on accuracy. The implementation and the achieved accuracy values are presented in detail in Fig. 3.

In addition, Table 4 documents the reliability of the proposed alpha-index. We also observe that NLR that has the greatest value of the alpha-index achieving significant prediction level which increases as other laboratory indices are added such as LDH that has the second greatest value of the alpha-index. The prediction level continues to increase as other laboratory indices are included with respect to their ranking per alpha-index.

Fig. 4 presents the architecture of the best ANN model including the whole of the parameters of this, required for its design.

In an effort to better understand the role of additional indices, we also performed a correlation analysis of the critical five laboratory indices with respect to the remaining 20 laboratory indices. The results of this calculation that corresponds to the Pearson correlation coefficient are presented at Table 5. Specifically, in this table the values in blue correspond to correlation coefficient greater than 0.30, with  $p$ -value < 0.001.

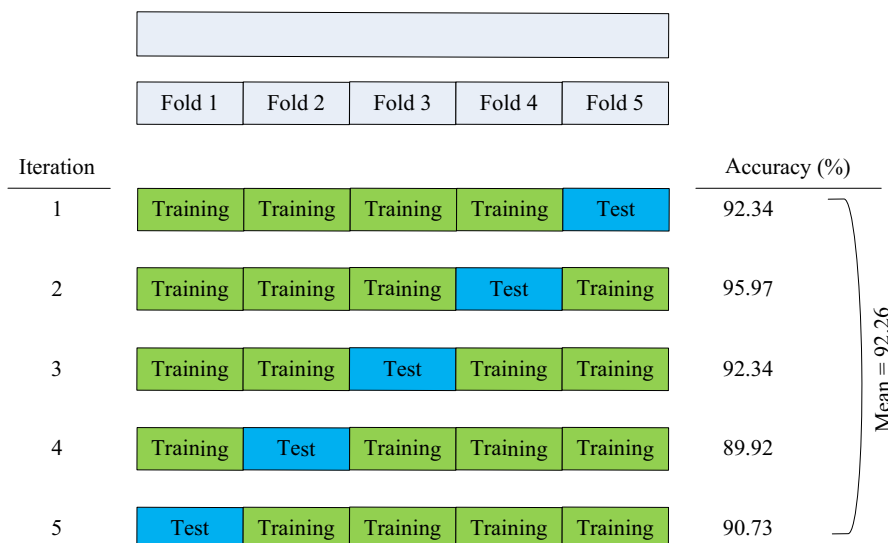
We observed that the 5 laboratory indices determined to be critical are in significant association with 9 other laboratory indices. Specifically,

- Neutrophil-to-lymphocyte ratio (NLR) has a strong correlation with the Neutrophil count per volume (Neut), the Lymphocyte count per volume (Lympo), the Immature Granulocyte count per volume (IG), the IG-to-lymphocyte ratio (IGLR) and the C-reactive protein (CRP),
- Lactate Dehydrogenase (LDH) is strongly correlated with Ferritin and Total bilirubin (Tbil),
- Fibrinogen (Fib) has a strong relation with Neutrophil count per volume (Neut) and C-reactive protein (CRP),
- Albumin (Alb) has a strong relation with White Blood Cell count per volume (WBC), Neutrophil count per volume (Neut), Immature Granulocyte count per volume (IG), the IG-to-lymphocyte ratio (IGLR), Hemoglobin concentration (Hb) and the Calcium (serum) (Ca + 2).

**4. Discussion**

Our study reveals for the first time 5 easily accessible laboratory markers of high importance for predicting COVID-19 severity, that were assessed through a novel alpha-index. Ranking these markers by importance and utilizing a training, validation, and testing dataset, we develop an ANN that can be used in clinical practice to accurately predict ICU hospitalization in COVID-19 early in the hospital course. Importantly, we also describe associations between laboratory indices that may better delineate several pathophysiological processes in COVID-19.

So far, risk prediction models for severe COVID-19 are rare and limited by several factors. An existing model widely used to identify



**Fig. 3.** Implementation of 5-fold cross validation based on accuracy index.

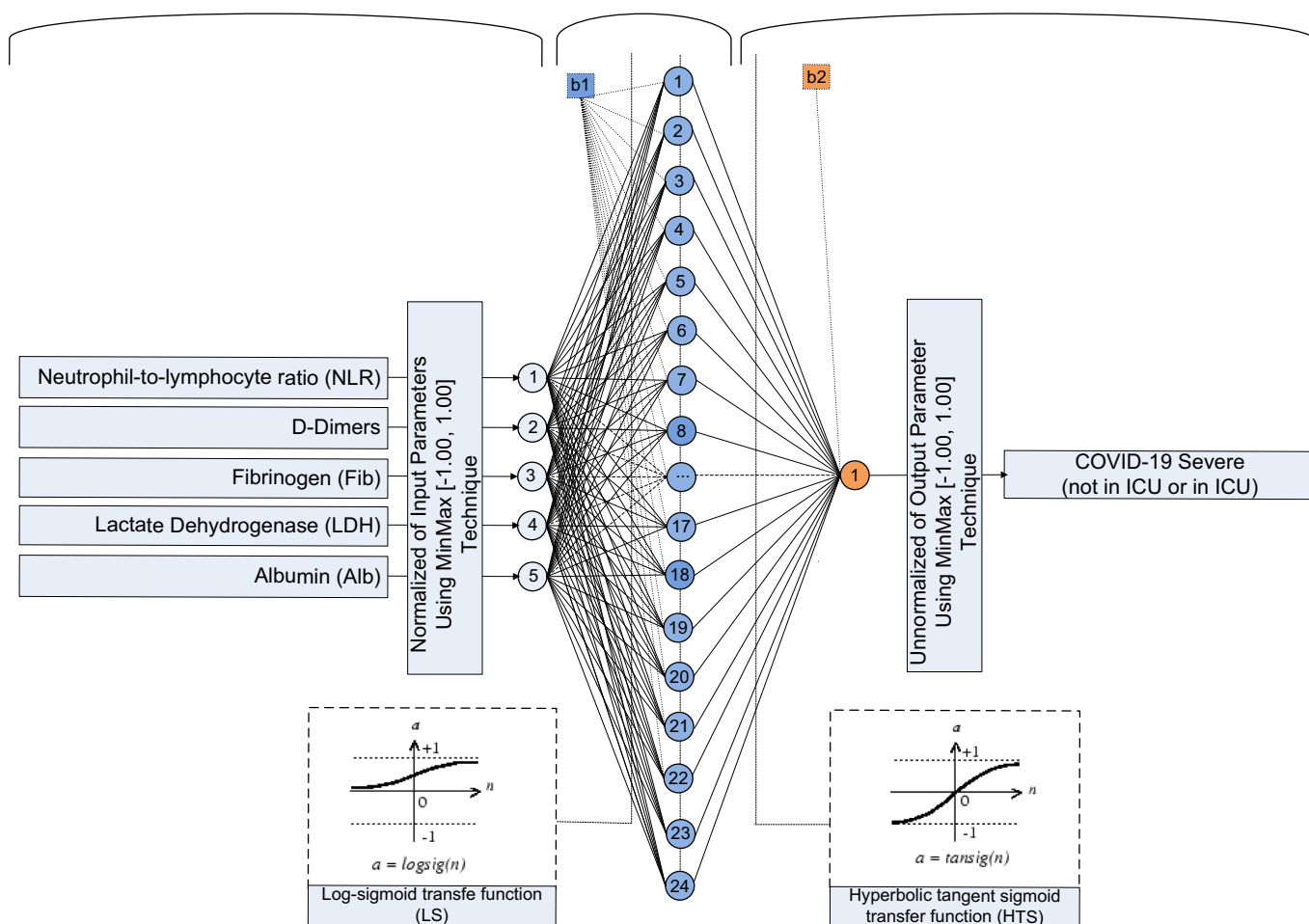


Fig. 4. Architecture of the optimum developed ANN 5-24-1 model.

deteriorating hospitalized patients in the United Kingdom (UK), National Early Warning Score 2 (NEWS2), has shown a low positive predictive value (PPV) 0.53 (0.47–0.59), but good negative predictive value (NPV) 0.96 (0.90–1.00) to identify inpatient COVID-19 deterioration [26]. Only the NPV of NEWS2 monitoring was higher than that of other earlier scores including Modified Early Warning Score (MEWS) and Sepsis Related Organ Failure Assessment (qSOFA) score. Another prediction model using demographics, comorbidities and laboratory tests at the initial presentation yielded moderate but significant accuracy with an area under the curve of 0.74 for ICU hospitalization. Machine learning models have tried to improve predictive values. Some models have focused only on mortality in the ICU or days of ICU hospitalization [27–30].

However, early prediction of ICU hospitalization is equally or even more important for the healthcare systems worldwide. In this context, machine learning was used to predict progression to a score  $\geq 5$  (WHO Clinical Progression Scale) before patients required mechanical ventilation. Strongest predictors of clinical deterioration were arterial blood oxygen partial pressure, followed by age, procalcitonin, LDH, CRP and alterations in blood count and coagulation. In the validation set, the CatBoost AUROC was 0.79, AUPRC 0.21 and Hosmer–Lemeshow test statistic 0.36 [31]. As a result, our ANN prediction model has several strengths not only because it utilizes early and routine laboratory values but also because it has very high accuracy.

It is also noteworthy that several publications have aimed to determine the most critical laboratory indices for ICU admission. Some of these studies suggest laboratory indices that are included in the five

laboratory indices proposed in the present study as critical. In particular, our machine learning approach confirms the predictive value of NLR which has been previously associated with severe COVID-19 [32–34]. Before COVID-19, NLR had already generated a lot of interest as a systemic inflammatory index and a potential prognostic factor in many clinical conditions [35]. In COVID-19, its role has been highlighted by meta-analyses showing an association of NLR with morbidity and mortality of COVID-19 [36–38]. Our approach however clarifies that only one marker is not enough to maximize the predictive value. Except for NLR, our model has identified LDH, fibrinogen and d-dimers as additional critical markers, which are known indicators of severe infections. Their role has been also confirmed in COVID-19, but still as single markers associated with severe COVID-19 [39–41]. Other hematological markers, such as erythrocyte dynamics and platelets, that have not shown critical predictive value in our model, may be also important for the diagnostic approach in severe COVID-19 [42,43]. Their role needs to be further refined.

Interestingly, our approach has documented that albumin is also an important prognostic marker. Although hypoalbuminemia has been largely under-studied in COVID-19 a few reports have indicated its association with disease severity, even independently of age and comorbidities [44–46]. Albumin may be a surrogate marker of the patient’s status at admission (cachexia, drug-binding proteins). Our study also shows inter-correlation between these entities that may better delineate several pathophysiological processes in COVID-19. The strong association that we observed between albumin and calcium is well established from the pathophysiological point of view and explains why the addition

**Table 5**

Correlation of the 5 hematological indices included in the proposed ANN model with the 20 hematological indices not included in the proposed ANN model based on Pearson correlation coefficient (Marked values with blue color correspond to correlation coefficient greater than 0.30, with p-value < 0.001).

Hematological indices not included in the proposed ANN model		Hematological indices included in the proposed ANN model				
		NLR	LDH	Fib	Alb	D-Dimers
1	WBC	0.66	0.15	0.27	0.30	0.12
2	Neut	0.79	0.20	0.31	0.34	0.15
3	Lympto	0.31	0.11	0.11	0.09	0.05
4	IG	0.54	0.13	0.12	0.34	0.13
5	IGLR	0.68	0.18	0.16	0.34	0.14
6	Hb	0.06	0.06	0.11	0.39	0.19
7	MCV	0.10	0.07	0.05	0.02	0.03
8	RDW	0.01	0.03	0.03	0.18	0.06
9	%MIC	0.06	0.09	0.06	0.05	0.05
10	%MAC	0.00	0.07	0.10	0.13	0.09
11	%HPO	0.06	0.10	0.13	0.06	0.07
12	HDW	0.05	0.09	0.06	0.17	0.10
13	PLT	0.16	0.01	0.28	0.14	0.00
14	CRP	0.30	0.29	0.43	0.27	0.12
15	Ferritin	0.16	0.51	0.20	0.17	0.07
16	ProBNP	0.05	0.02	0.05	0.08	0.24
17	pCT	0.02	0.08	0.08	0.02	0.03
18	Tbil	0.14	0.30	0.17	0.00	0.09
19	Ca+2	0.28	0.23	0.14	0.61	0.08
20	Creat	0.07	0.06	0.10	0.02	0.20

of calcium in the model does not strengthen the alpha-index. Several underlying mechanisms may be implicated in the correlation of albumin with COVID-19 outcomes. Firstly, endothelial dysfunction is a common denominator in severe COVID-19 [47]. The process of endothelial dysfunction leads to loss of integrity of the epithelial–endothelial damage, contributing to hypoalbuminemia [48]. Secondly, neutrophil extracellular traps (NETs) are also major mediators in COVID-19 [49]. Since albumin inhibits NETs formation, patients with hypoalbuminemia may have a greater risk of morbidity and mortality. Additionally, albumin is essential for balancing hemodynamics and reducing D-dimers, suggesting that hypoalbuminemia is associated with a pro-coagulant phenotype [50].

Our study has some limitations. Despite the prospective patient recruitment, our findings cannot be considered as definitive, indicating the need for future studies. Introduction of additional prognostic factors such as imaging might have improved the predictive accuracy of our model. Nevertheless, our aim was to propose a model with a small number of easily accessible laboratory indices that could be useful in all settings. Furthermore, the suggested ANN needs to be further validated in larger real-world cohorts from different hospitals and patient populations. Although vaccination and different SARS-CoV2 strains are not expected to directly influence our model, it should be noted that our patient population is homogeneous including only unvaccinated patients from the first two waves and alpha to delta variants in Greece.

In conclusion, only 5 routine and easily accessible laboratory indices were able to predict ICU hospitalizations in COVID-19 patients utilizing a novel approach of artificial intelligence based on our preliminary findings. Additional data are needed to ensure the validity of our database and further studies are ongoing. Given that vaccinations and viral mutations are constantly changing the landscape of COVID-19, a prediction tool based on such robust variables is of high importance. Such models could help not only to reduce cost of hospitalization, morbidity and mortality, but also to accurately predict patients at high-risk that would benefit from prophylactic or pre-emptive treatments. Lastly, this artificial intelligence approach paves the way for future applications of

this novel methodology in other clinical entities, since the newly proposed alpha-index may be extremely useful in a plethora of diseases or other entities with classification problems [51–53].

## Funding

This research received no financial support for the research, authorship, and/or publication of this article.

## CRediT authorship contribution statement

**Panagiotis G. Asteris:** Conceptualization, Methodology, Writing – original draft. **Styliani Kokoris:** Conceptualization, Writing – original draft. **Eleni Gavriilaki:** Validation, Writing – original draft. **Markos Z. Tsoukalas:** Methodology. **Panagiotis Houpas:** Methodology. **Maria Paneta:** Data curation. **Andreas Koutzas:** Data curation. **Theodoros Argyropoulos:** Data curation. **Nizar Faisal Alkayem:** Methodology. **Danial J. Armaghani:** Methodology. **Abidhan Bardhan:** Methodology. **Liborio Cavaleri:** Methodology. **Maosen Cao:** Methodology. **Iman Mansouri:** Methodology. **Ahmed Salih Mohammed:** Methodology. **Pijush Samui:** Methodology. **Gloria Gerber:** Writing – review & editing. **Dimitrios T. Boumpas:** Writing – review & editing. **Argyrios Tsantes:** Writing – review & editing. **Evangelos Terpos:** Writing – review & editing. **Meletios A. Dimopoulos:** Writing – review & editing.

## Declaration of Competing Interest

The authors declare no conflict of interest.

## Data availability

Data will be immediately available upon request.

## Acknowledgments

None.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.clim.2022.109218>.

## References

- [1] E.J. Rubin, D.L. Longo, SARS-CoV-2 Vaccination - an Ounce (Actually, Much Less) of Prevention, *N. Engl. J. Med.* 383 (2020) 2677–2678.
- [2] E. Terpos, I. Ntanasis-Stathopoulos, I. Elalamy, E. Kastiris, T.N. Sergentanis, M. Politou, T. Psaltopoulou, G. Gerotziakas, M.A. Dimopoulos, Hematological findings and complications of COVID-19, *Am. J. Hematol.* 95 (2020) 834–847.
- [3] G. Ponti, M. Maccaferri, C. Ruini, A. Tomasi, T. Ozben, Biomarkers associated with COVID-19 disease progression, *Crit. Rev. Clin. Lab. Sci.* 57 (2020) 389–399.
- [4] N. Liu, T. Zhang, L. Ma, H. Zhang, H. Wang, W. Wei, H. Pei, H. Li, The impact of ABO blood group on COVID-19 infection risk and mortality: a systematic review and meta-analysis, *Blood Rev.* 48 (2021), 100785.
- [5] L. Cheng, H. Li, L. Li, C. Liu, S. Yan, H. Chen, Y. Li, Ferritin in the coronavirus disease 2019 (COVID-19): a systematic review and meta-analysis, *J. Clin. Lab. Anal.* 34 (2020), e23618.
- [6] J.W. Martha, A. Wibowo, R. Pranata, Prognostic value of elevated lactate dehydrogenase in patients with COVID-19: a systematic review and meta-analysis, *Postgrad. Med. J.* 98 (2022) 422–427.
- [7] E. Gavriilaki, P.G. Asteris, T. Touloumenidou, E.E. Koravou, M. Koutra, P. G. Papayanni, V. Karali, A. Papalexandri, C. Varelas, F. Chatzopoulou, M. Chatzidimitriou, D. Chatzidimitriou, A. Veleni, S. Grigoriadis, E. Rapti, D. Chloros, I. Kioumis, E. Kaimakamis, M. Bitzani, D. Boumpas, A. Tsantes, D. Sotiropoulos, I. Sakellari, I.G. Kalantzis, S.T. Parastatidis, M. Koopialipour, L. Cavaleri, D.J. Armaghani, A. Papadopoulou, R.A. Brodsky, S. Kokoris, A. Anagnostopoulos, Genetic justification of severe COVID-19 using a rigorous algorithm, *Clin. Immunol.* 226 (2021), 108726.
- [8] P.G. Asteris, E. Gavriilaki, T. Touloumenidou, E.E. Koravou, M. Koutra, P. G. Papayanni, A. Poulereas, V. Karali, M.E. Lemonis, A. Mamou, A.D. Skentou, A. Papalexandri, C. Varelas, F. Chatzopoulou, M. Chatzidimitriou, D. Chatzidimitriou, A. Veleni, E. Rapti, I. Kioumis, E. Kaimakamis, M. Bitzani,



- D. Boumpas, A. Tsantes, D. Sotiropoulos, A. Papadopoulou, I.G. Kalantzis, L. A. Vallianatou, D.J. Armaghani, L. Cavaleri, A.H. Gandomi, M. Hajihassani, M. Hasanipannah, M. Koopialipoor, P.B. Lourenco, P. Samui, J. Zhou, I. Sakellari, S. Valsami, M. Politou, S. Kokoris, A. Anagnostopoulos, Genetic prediction of ICU hospitalization and mortality in COVID-19 patients using artificial neural networks, *J. Cell. Mol. Med.* 26 (2022) 1445–1455.
- [9] A.K. Upadhyay, S. Shukla, Correlation study to identify the factors affecting COVID-19 case fatality rates in India, *Diabetes Metab. Syndr.* 15 (2021) 993–999.
- [10] H.R. Niazkhar, M. Niazkhar, Application of artificial neural networks to predict the COVID-19 outbreak, *Glob. Health Res. Policy* 5 (2020) 50.
- [11] C. Mahanty, R. Kumar, P.G. Asteris, A.H. Gandomi, COVID-19 patient detection based on fusion of transfer learning and fuzzy ensemble models using CXR images, *Appl. Sci.* 11 (2021).
- [12] I. Rahimi, A.H. Gandomi, P.G. Asteris, F. Chen, Analysis and prediction of COVID-19 using SIR, SEIQR, and machine learning models: Australia, Italy, and UK cases, *Information* 12 (2021).
- [13] P.-G. Asteris, M.-G. Douvika, C.-A. Karamani, A.-D. Skentou, K. Chlichlia, L. Cavaleri, T. Daras, D.-J. Armaghani, T.-E. Zaoutis, A novel heuristic algorithm for the Modeling and risk assessment of the COVID-19 pandemic phenomenon, *Comput. Model. Eng. Sci.* 125 (2020) 815–828.
- [14] Y. Wu, M.L. Giger, K. Doi, C.J. Vyborny, R.A. Schmidt, C.E. Metz, Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer, *Radiology* 187 (1993) 81–87.
- [15] I. Kononenko, Machine learning for medical diagnosis: history, state of the art and perspective, *Artif. Intell. Med.* 23 (2001) 89–109.
- [16] M.A. Mazurowski, P.A. Habas, J.M. Zurada, J.Y. Lo, J.A. Baker, G.D. Tourassi, Training neural network classifiers for medical decision making: the effects of imbalanced datasets on classification performance, *Neural Netw.* 21 (2008) 427–436.
- [17] D.C. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Mitosis detection in breast cancer histology images with deep neural networks, *Med. Image Comput. Comput. Assist. Interv.* 16 (2013) 411–418.
- [18] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung pattern classification for interstitial lung diseases using a deep convolutional neural network, *IEEE Trans. Med. Imaging* 35 (2016) 1207–1216.
- [19] D. Shen, G. Wu, H.I. Suk, Deep learning in medical image analysis, *Annu. Rev. Biomed. Eng.* 19 (2017) 221–248.
- [20] A.H. Ribeiro, M.H. Ribeiro, G.M.M. Paixao, D.M. Oliveira, P.R. Gomes, J. A. Canazart, M.P.S. Ferreira, C.R. Andersson, P.W. Macfarlane, W. Meira Jr., T. B. Schon, A.L.P. Ribeiro, Automatic diagnosis of the 12-lead ECG using a deep neural network, *Nat. Commun.* 11 (2020) 1760.
- [21] A.H. Gandomi, M.M. Fridline, D.A. Roke, Decision tree approach for soil liquefaction assessment, *TheScientificWorldJournal* 2013 (2013), 346285.
- [22] P.G. Asteris, V.G. Mokos, Concrete compressive strength using artificial neural networks, *Neural Comput. & Applic.* 32 (2020) 11807–11826.
- [23] P.G. Asteris, P.B. Lourenço, P.C. Roussis, C. Elpidia Adami, D.J. Armaghani, L. Cavaleri, C.E. Chaliouris, M. Hajihassani, M.E. Lemonis, A.S. Mohammed, K. Pilakoutas, Revealing the nature of metakaolin-based concrete materials using artificial intelligence techniques, *Constr. Build. Mater.* 322 (2022), 126500.
- [24] L. Valenti, S. Griffini, G. Lamorte, E. Grovetti, S.C. Uceda Renteria, F. Malvestiti, L. Scudeller, A. Bandera, F. Peyvandi, D. Prati, P. Meroni, M. Cugno, Chromosome 3 cluster rs11385942 variant links complement activation with severe COVID-19, *J. Autoimmun.* 117 (2021), 102595.
- [25] V. Ramlall, P.M. Thangaraj, C. Meydan, J. Foox, D. Butler, J. Kim, B. May, J.K. De Freitas, B.S. Glicksberg, C.E. Mason, N.P. Tatonetti, S.D. Shapira, Immune complement and coagulation dysfunction in adverse outcomes of SARS-CoV-2 infection, *Nat. Med.* 26 (2020) 1609–1615.
- [26] K.F. Baker, A.T. Hanrath, I. Schim van der Loeff, L.J. Kay, J. Back, C.J. Duncan, National Early Warning Score 2 (NEWS2) to identify inpatient COVID-19 deterioration: a retrospective analysis, *Clin. Med. (Lond. Engl.)* 21 (2021) 84–89.
- [27] P. Pan, Y. Li, Y. Xiao, B. Han, L. Su, M. Su, Y. Li, S. Zhang, D. Jiang, X. Chen, F. Zhou, L. Ma, P. Bao, L. Xie, Prognostic assessment of COVID-19 in the intensive care unit by machine learning methods: model development and validation, *J. Med. Internet Res.* 22 (2020), e23128.
- [28] G. Lorenzoni, D. Azzolina, A.S. Acar, L. Silvestri, P. Berchiulla, D. Gregori, Understanding the factors affecting COVID-19 mortality in Italy: does a relationship exist with a sharp increase in intensive care unit admissions? *Disaster Med. Public Health Prep.* (2021) 1–6.
- [29] S.S. Lorenzen, M. Nielsen, E. Jimenez-Solem, T.S. Petersen, A. Perner, H. C. Thorsen-Meyer, C. Igel, M. Sillesen, Using machine learning for predicting intensive care unit resource use during the COVID-19 pandemic in Denmark, *Sci. Rep.* 11 (2021) 18959.
- [30] D.A. Alabbad, A.M. Almuhaideb, S.J. Alsunaidi, K.S. Alqudaihi, F.A. Alamoudi, M. K. Alhobaishi, N.A. Alaqeel, M.S. Alshahrani, 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* 30 (2022), 100937.
- [31] S. Garcia-Gutierrez, C. Esteban-Aizpiri, I. Lafuente, I. Barrio, R. Quiros, J. M. Quintana, A. Uranga, C.-R.W. Group, Machine learning-based model for prediction of clinical deterioration in hospitalized patients by COVID 19, *Sci. Rep.* 12 (2022) 7097.
- [32] A.P. Yang, J.P. Liu, W.Q. Tao, H.M. Li, The diagnostic and predictive role of NLR, d-NLR and PLR in COVID-19 patients, *Int. Immunopharmacol.* 84 (2020), 106504.
- [33] S. Jimeno, P.S. Ventura, J.M. Castellano, S.I. Garcia-Adasme, M. Miranda, P. Touza, I. Lllana, A. Lopez-Escobar, Prognostic implications of neutrophil-lymphocyte ratio in COVID-19, *Eur. J. Clin. Invest.* 51 (2021), e13404.
- [34] W. Shang, J. Dong, Y. Ren, M. Tian, W. Li, J. Hu, Y. Li, The value of clinical parameters in predicting the severity of COVID-19, *J. Med. Virol.* 92 (2020) 2188–2192.
- [35] T. Iba, T. Miki, N. Hashiguchi, Y. Tabe, I. Nagaoka, Is the neutrophil a 'prima donna' in the procoagulant process during sepsis? *Crit. Care* 18 (2014) 230.
- [36] X. Li, C. Liu, Z. Mao, M. Xiao, L. Wang, S. Qi, F. Zhou, Predictive values of neutrophil-to-lymphocyte ratio on disease severity and mortality in COVID-19 patients: a systematic review and meta-analysis, *Crit. Care* 24 (2020) 647.
- [37] R.K. Mahat, S. Panda, V. Rathore, S. Swain, L. Yadav, S.P. Sah, The dynamics of inflammatory markers in coronavirus disease-2019 (COVID-19) patients: a systematic review and meta-analysis, *Clin. Epidemiol. Glob. Health* 11 (2021), 100727.
- [38] D.M. Simadibrata, J. Calvin, A.D. Wijaya, N.A.A. Ibrahim, Neutrophil-to-lymphocyte ratio on admission to predict the severity and mortality of COVID-19 patients: a meta-analysis, *Am. J. Emerg. Med.* 42 (2021) 60–69.
- [39] H. Zhan, H. Chen, C. Liu, L. Cheng, S. Yan, H. Li, Y. Li, Diagnostic value of D-dimer in COVID-19: a meta-analysis and meta-regression, clinical and applied thrombosis/hemostasis : official journal of the international academy of, *Clin. Appl. Thromb. Hemost.* 27 (2021), 10760296211010976.
- [40] M.E. Duz, A. Balci, E. Menekse, D-dimer levels and COVID-19 severity: systematic review and meta-analysis, *Tuberk Toraks* 68 (2020) 353–360.
- [41] F. Del Sole, A. Farcomeni, L. Loffredo, R. Carnevale, D. Menichelli, T. Vicario, P. Pignatelli, D. Pastori, Features of severe COVID-19: a systematic review and meta-analysis, *Eur. J. Clin. Invest.* 50 (2020), e13378.
- [42] M.M. Mendonca, K.R. da Cruz, D.D.S. Pinheiro, G.C.A. Moraes, P.M. Ferreira, M. L. Ferreira-Neto, E.S. da Silva, R.V. Goncalves, G.R. Pedrino, J.O. Fajemiroye, C. H. Xavier, Dysregulation in erythrocyte dynamics caused by SARS-CoV-2 infection: possible role in shuffling the homeostatic puzzle during COVID-19, *Hematol. Transfus. Cell. Ther.* 44 (2022) 235–245.
- [43] G. Bomhof, P. Mutsaers, F.W.G. Leebeek, P.A.W. Te Boekhorst, J. Hofland, F. N. Croles, A.J.G. Jansen, COVID-19-associated immune thrombocytopenia, *Br. J. Haematol.* 190 (2020) e61–e64.
- [44] J. Wagner, V. Garcia-Rodriguez, A. Yu, B. Dutra, S. Larson, B. Cash, A. DuPont, A. Farooq, Elevated transaminases and hypoalbuminemia in Covid-19 are prognostic factors for disease severity, *Sci. Rep.* 11 (2021) 10308.
- [45] J. Huang, A. Cheng, R. Kumar, Y. Fang, G. Chen, Y. Zhu, S. Lin, Hypoalbuminemia predicts the outcome of COVID-19 independent of age and co-morbidity, *J. Med. Virol.* 92 (2020) 2152–2158.
- [46] M.C. Viana-Llamas, R. Arroyo-Espiguero, J.A. Silva-Obregon, G. Uribe-Heredia, I. Nunez-Gil, B. Garcia-Magallon, C.G. Toran-Martinez, A. Castillo-Sandoval, E. Diaz-Caraballo, I. Rodriguez-Guinea, J. Dominguez-Lopez, Hypoalbuminemia on admission in COVID-19 infection: an early predictor of mortality and adverse events. A retrospective observational study, *Med. Clin.* 156 (2021) 428–436.
- [47] E. Gavriilaki, P. Anyfanti, M. Gavriilaki, A. Lazaridis, S. Douma, E. Gkaliagkousi, Endothelial dysfunction in COVID-19: lessons learned from coronaviruses, *Curr. Hypertens. Rep.* 22 (2020) 63.
- [48] V. Zerbato, G. Sanson, M. De Luca, S. Di Bella, A. di Masi, P. Caironi, B. Marini, R. Ippodrino, R. Luzzati, The impact of serum albumin levels on COVID-19 mortality, *Infect. Dis. Rep.* 14 (2022) 278–286.
- [49] P. Skendros, A. Mitsios, A. Chrysanthopoulou, D.C. Mastellos, S. Metallidis, P. Rafailidis, M. Ntinopoulou, E. Sertaridou, V. Tsiironidou, C. Tsigalou, M. Tektonidou, T. Konstantinidis, C. Papagoras, I. Mitroulis, G. Germanidis, J. D. Lambris, K. Ritis, Complement and tissue factor-enriched neutrophil extracellular traps are key drivers in COVID-19 immunothrombosis, *J. Clin. Invest.* 130 (2020) 6151–6157.
- [50] G. Ramadori, Albumin infusion in critically ill COVID-19 patients: Hemodilution and anticoagulation, *Int. J. Mol. Sci.* 22 (2021).
- [51] M.-S. Barkhordari, D.-J. Armaghani, P.-G. Asteris, Structural Damage Identification Using Ensemble Deep Convolutional Neural Network Models, *Computer Modeling in Engineering & Sciences*, {pages}, 2022.
- [52] P.G. Asteris, F.I.M. Rizal, M. Koopialipoor, P.C. Roussis, M. Ferentinou, D. J. Armaghani, B. Gordan, Slope stability classification under seismic conditions using several tree-based intelligent techniques, *Appl. Sci.* 12 (2022) 1753.
- [53] S. Lu, M. Koopialipoor, P.G. Asteris, M. Bahri, D.J. Armaghani, A novel feature selection approach based on tree models for evaluating the punching shear capacity of steel Fiber-reinforced concrete flat slabs, *Materials* 13 (2020) 3902.