EJMHR
ISSN 2786-8524
EUROPEAN JOURNAL OF MEDICAL AND HEALTH RESEARCH
OPEN ACCESS
PEER REVIEWED

Probability Threshold Optimization for Classification of COVID-19 Patients with Higher Mortality Risk: A Case Study from the North-Eastern Region in North Macedonia

Dragana Mogilevska-Gruevska ✉
Radiologist, PHI - General Hospital – Kochani, North Macedonia
Ilija Gruevski
Professor of Economics, Faculty of economics, Goce Delcev University, Stip, North Macedonia
Golubinka Boshevska
Professor of Microbiology, Faculty of Medical Sciences, Goce Delcev University, Stip, North Macedonia
Stevan Gaber
Professor of Economics, Faculty of economics, Goce Delcev University, Stip, North Macedonia

Abstract

Research goal: The goal of this research is to identify, compare and demonstrate some available methodological approaches which are sufficient to draw the optimal probability threshold, particularly in the case of classification of infected patients with increased risk of dying from COVID-19. The presented methodologies generate identical results if the purpose of classification is to maximize the prognosis accuracy from the point of sensitivity. Sample: As part of the whole population, the sample counts 1013 patients from the north-eastern region of the Republic of North Macedonia. **Methodology:** The general methodological frame used to calculate and forecast the probabilities of death outcome from COVID-19 is the binary logistic regression. In extension, we applied the rules of maximum sum and maximum product as well as the so-called Youden Index for the purpose of optimization of the probability threshold. The principals of the ROC curve in addition with the Index of Union was also helpful for the same purpose. Prognosis accuracy was evaluated through the status of patient according to the rules of "golden standard" in which sensitivity, specificity and the general accuracy of prognosis play a crucial role. **Results:** Accordingly, the results from the research indicate that the optimal probability threshold or "cut-off" point that provides maximal accuracy, particularly from the perspective of sensitivity in the prognosis is 0,1. In that point, the coefficient of sensitivity (the percentage of true positively predicted death cases in respect to all death cases from the sample) is measured 85,71%. **Conclusion:** The applied methodological approaches offer scientifically sound foundations in the context of mortality risk evaluation and classification of COVID-19 patients. Then targeted patients will be subject of precaution with strict measures, protocols and more aggressive treatment in order to minimize the chances of death outcome.

More Information

How to cite this article: Mogilevska-Gruevska D, Gruevski I, Boshevska G, Gaber S. Probability Threshold Optimization for Classification of COVID-19 Patients with Higher Mortality Risk: A Case Study from the North-Eastern Region in North Macedonia. Eur J Med Health Res, 2024;2(5):237-46.

DOI: 10.59324/ejmhr.2024.2(5).24

Keywords:

COVID-19,
binary logistic regression,
sensitivity,
specificity,
Youden Index,
Index of Union,
ROC curve,
probability threshold,
North Macedonia.

Introduction

Recapitulation

In this article, are present the benefits of the model of logistic regression used for systematization and classification of patients with higher risk of death by COVID-19. The possibility to classify patients with higher mortality risk is closely related to the possibility to predict the final outcome of a certain disease in each patient. Namely, based on previous research [1], the final outcome from COVID-19 (death or survival) is closely related to few fundamental factors which were detected using the logistic regression, such as age, gender, the presence of comorbidities and secondary complications, as well as the score (or rank) of the primary disease. The logistic regression was performed in Xrealstats, and the results from the basic statistical parameters were summarized in figure No. 1, out of which the most important ones are the odds ratios, shown in the column exp(b). According to these, the patients with comorbidities and complications have the highest chances of death by COVID-19 (OR 16,53 with CI 8,21 – 33,25 and 4,08 with CI 1,34 – 12,38). Also, male patients are subject to insignificantly higher mortality risk with OR 1,55 with CI 0,86 – 2,80. Every year of age increases the odds for death by 1.06 times (CI 1,03 – 1,09), while every additional score of the primary disease leads to higher odds by 1.24 times (CI 1,04 – 1,47) [2].

The software displays the predicted values of the probability an infected person to die from COVID-19, as shown in figure No. 2, in column p-Pred.

Research goal

The goal of this research is to identify, compare and demonstrate some available methodological approaches which are sufficient to draw the optimal probability threshold, particularly in the case of classification of infected patients with increased risk of dying from COVID-19. Presented methodologies are supposed to generate similar results if the purpose of classification is to maximize the prognosis accuracy from the point of sensitivity. The process of classification of patients with higher mortality risk will proceed with their separation in a special targeted group, which will undergo additional precautionary measures, protocols and treatment in order to decrease the risk of death outcome associated with these targeted patients.

The sample

The sample in our research includes 1013 COVID-19 patients who are officially confirmed by PCR test. As a representative of the whole population, the sample reflects the north-east region of the Republic of N. Macedonia, considering that these patients gravitate and were sent in the radio-diagnostic department in Public Health Institution (PHI) "General Hospital – Kochani" in the period between September 2020 - September 2022. The observed timeframe of the pandemics, indicates that alfa, beta and delta variants of the SARS-CoV-2 virus were circulating during that time. We must emphasize that the obtained results regarding the causal relationships between the factor variables and the final outcome from the primary disease, are directly dependent on the characteristics of the observed statistical sample. If, for example, the sample is made up by patients infected with the later variants of the virus which produce a milder clinical manifestation of the primary disease, we can certainly say that the odd ratios from the logistic regression will certainly be different. Nevertheless, the data for our research was collected through a questionnaire and phone calls, and therefore, all the required ethical standards for this trial were completely and strictly met.

Coeff
-9,66817
0,441185
0,056263
2,805109
0,214558
1,406931

LL0	-272,415
LL1	-160,24
Chi-Sq	224,3507
df	5
p-value	1,74E-46
alpha	0,05
sig	yes
R-Sq (L)	0,411781
R-Sq (CS)	0,198661
R-Sq (N)	0,477562
AIC	332,4796
BIC	362,0036

Covariance Matrix						Converge
1,156305	-0,04535	-0,0122	-0,07224	-0,01699	-0,15407	-5,4E-14
-0,04535	0,090479	-0,00026	0,006832	-0,00287	0,013605	-4,4E-14
-0,0122	-0,00026	0,000211	-0,0001	-7,3E-05	-0,00215	-3,8E-12
-0,07224	0,006832	-0,0001	0,127251	-0,00646	-0,00079	-4,9E-14
-0,01699	-0,00287	-7,3E-05	-0,00646	0,008027	-0,00014	-2E-13
-0,15407	0,013605	-0,00215	-0,00079	-0,00014	0,320598	-5,1E-14

	coeff b	*s.e.*	*Wald*	*p-value*	*exp(b)*	*lower*	*upper*
Intercept	-9,66817	1,075316	80,83814	2,45E-19	6,33E-05		
Pol	0,441185	0,300797	2,151269	0,142452	1,554549	0,862115	2,803131
Vozrast	0,056263	0,014538	14,97794	0,000109	1,057876	1,028159	1,088453
Komplikac	2,805109	0,356722	61,83562	3,73E-15	16,52888	8,214871	33,25723
Skor	0,214558	0,089595	5,734827	0,016632	1,239314	1,039723	1,477219
Komorbidi	1,406931	0,566214	6,17426	0,012962	4,083406	1,346048	12,38753

Figure 1: Results from Logistic Regression

Source: [1]

pol-m-1	vozrast	komplik.	STAGE	коморби.	Success	Failure	Total	p-Obs	p-Pred	Suc-Pred	Fail-Pred	LL	% Correct
1	51	0	3	1	0	1	1	0	0,013295	0,013295	0,986705	-0,01338	100
0	33	0	1	0	0	1	1	0	0,000502	0,000502	0,999498	-0,0005	100
1	54	0	2	0	0	1	1	0	0,003142	0,003142	0,996858	-0,00315	100
1	79	0	1	1	0	1	1	0	0,040672	0,040672	0,959328	-0,04152	100
0	60	1	2	1	0	1	1	0	0,160946	0,160946	0,839054	-0,17548	100
0	47	0	1	0	0	1	1	0	0,001102	0,001102	0,998898	-0,0011	100
1	44	0	2	0	0	1	1	0	0,001793	0,001793	0,998207	-0,00179	100
0	63	1	3	1	0	1	1	0	0,219624	0,219624	0,780376	-0,24798	100
1	75	0	3	1	0	1	1	0	0,049424	0,049424	0,950576	-0,05069	100
0	50	0	1	0	0	1	1	0	0,001305	0,001305	0,998695	-0,00131	100
0	64	1	5	1	1	0	1	1	0,313786	0,313786	0,686214	-1,15905	0
0	63	1	3	1	0	1	1	0	0,219624	0,219624	0,780376	-0,24798	100
0	66	0	2	1	0	1	1	0	0,016005	0,016005	0,983995	-0,01613	100
0	63	1	4	1	0	1	1	0	0,258592	0,258592	0,741408	-0,2992	100
1	37	0	1	0	0	1	1	0	0,000976	0,000976	0,999024	-0,00098	100
1	57	0	2	1	0	1	1	0	0,01501	0,01501	0,98499	-0,01512	100
1	39	0	1	0	0	1	1	0	0,001093	0,001093	0,998907	-0,00109	100

Figure 2: Predicted Values of the Probability of Death Outcome

Source: [2]

Theoretical background

Probability vs odds

If there are 4 patients in one department, out of whom one was diagnosed with SARS-CoV-2, the chances to choose the patient with SARS-CoV-2 for an examination purpose are 0,25 or 1/4. This means that the **probability** as a concept is calculated as a number of possible events in relation to the total number of possible outcomes. If there are 2 patients diagnosed with SARS-CoV-2 in the same department, then, the chances of choosing the patient with SARS-CoV-2 for examination are 0,5 or 2/4. If we summarize, the formula (1) used to define probability is [3]:

$$p(e) = \frac{\text{number of favourable outcomes}}{\text{total number of outcomes}} \qquad (1)$$

On the other hand, **chances** are defined as the ratio of the probability of occurrence of a certain event and the probability of non-occurrence of that event. For example, what would the chances be if the chosen patients have SARS-Cov-2 in the first and the second example? The correct answer would be 0,33 (0,25/0,75), and 1 (0,5/0,5), respectively. From this we can conclude that formula (2) that is used to calculate the chances of a certain event is [4]:

$$\mathrm{Odds}(e) = \frac{p(e)}{1-p(e)} \qquad (2)$$

The above expression under number 2 is useful because it can be used to derive the probability of an event, if the chances of that event are previously known [5], as shown in the next formula (3):

$$p(e) = \frac{\mathrm{odds}(e)}{1+\mathrm{odds}(e)} \qquad (3)$$

The relation between the theoretical concepts of probability and chance is clearly seen from these examples. In addition, we will see how we can implement the established relation for predicting the probability of a certain outcome (for example death by COVID, getting cancer, stroke or heart attack) within the model of logistic regression.

Predicting the probability of an event

We reasoned that the logistic regression model is designed to calculate the odds of a particular event, such as death from a disease. If we integrate the general expression of logistic regression into the previous formula, in that case we'll come into the probability of occurrence of a particular event in accordance with the concept of logistic function [6][7], as shown in expression (4):

$$p(e) = \frac{\exp\,(b_0+b_1x_1+b_2x_2+\cdots+b_nx_n)}{1+\exp\,(b_0+b_1x_1+b_2x_2+\cdots+b_nx_n)} \qquad (4)$$

Derived from the model of logistic regression, formula (4) enables the researcher to calculate the probability of a certain event, which would be a subject of interest in his/her clinical research. With it, the researcher does individual evaluation of the associated risk from a certain medical condition in each individual patient, in accordance with the patient's individual characteristics (age, gender, comorbidities, complications, disease score, etc.) Furthermore, the researcher needs to identify and classify the patients with higher risk from the medical condition, in relation to the patients with lower risk from the same medical condition. For the purpose of classification of these patients, it is best to use the so-called probability threshold or cut-off point.

Probability threshold

From the above mentioned, one can assume that the prediction accuracy of a particular event is directly correlated with the concept of probability threshold. Namely, we know from the theory of probability that the probability of a certain event can range from 0 to 1. So, for example, if we throw a coin an infinite number of times, the probability to get heads is 0,5, which is exactly equal to the probability to get tails. The acceptable threshold probability of 0,5 is established in similar way during the classification of the infected patients. So, the software automatically classifies all patients with predicted probability of death higher than 0,5 as **high-risk patients** (patients likely to die from COVID-19), while the patients with death probability lower than 0,5 are classified as **low-risk patients** (patients not likely to die from COVID-19):

Patients **not likely** to die from covid < 0,5 < Patients **likely** to die from covid

Simplified, the patients are divided into these two groups, so that the high-risk patients will be subject to an urgent or more specific medical protocol, or they could be treated with an alternative medical approach. The probability used for binary classification of the predicted events into true or false, in this case that is the probability of 0,5, is known as a cut-off point or probability threshold [8]. Since the prediction accuracy of the final outcome depends directly on this threshold, the margin can be optimized, or adjusted in accordance with the ability of the model to successfully identify the patients with the observed medical condition (sensitivity) or the patients without the observed medical condition (specificity). Even though these calculations seem complicated and complex, they can be calculated relatively automated through one the standard software packages for statistical analysis. In our research, we used Xrealstat software, which is considered as one of the most sophisticated extensions of Excel used for advanced statistical analysis.

Classification of patients in the high-risk group

Results from the initial classification of patients according to the generally accepted probability threshold of 0,5, are given in the classification table in figure No.3. It should be noted that the general (overall) accuracy of prediction is measured 93,78%. There is also high accuracy in prediction of the survival cases which are as high as 98,71%. However, can be noticed, that the accuracy in prediction of death cases is relatively small with 33,76%. Precisely, the software predicted that 38 patients will die, out of whom the death outcome was correctly predicted in 26 patients, while in the other 12 patients who survived, the death outcome was incorrectly predicted. This generates a sensitivity rate of 33,76% (26/77). Furthermore, it is predicted that 975 patients will survive, out of whom 924 patients had a correct prediction, while the other 51 patients died and they had an incorrect prediction. According to these numbers, the specificity rate is 98,71% (924/936), while the overall prediction accuracy rate is 98,78% ((26+924)/1013).

		Classification Table							
Converge			Suc-Obs	Fail-Obs					
-5,4E-14		Suc-Pred	26	12	38		Sensitivity	=	0,337662
-4,4E-14		Fail-Pred	51	924	975		Specificity	=	0,987179
-3,8E-12			77	936	1013		Youden id›	=	0,324842
-4,9E-14									
-2E-13		Accuracy	0,337662	0,987179	0,937808				
-5,1E-14									
		Cutoff	0,5						

Figure 3: Classification of High-Risk Patients
Source: Author's calculations

	Status of patient according to **"gold standard"**		
	Has *the condition*	***Doesn't have*** *the condition*	
Positive	**True positive**	**False positive**	***Positive*** *predictive values* row
Negative	**False negative**	**True negative**	***Negative*** *predictive values row*
	Sensitivity column	***Specificity*** column	

Figure 4: Classification of Patient's Status According to the "Golden Standard"
Source: [8]

Sensitivity vs specificity in prognosis

From the analysis above, it is clear that the focus of prognosis of the final outcome of a certain medical condition is placed on the sensitivity and specificity. For the purpose of defining these important concepts from the aspect of predicting medical conditions in patients, we renamed our classification table in accordance with the terminology from "the golden standard" which is commonly used in describing the patient's status within the clinical trials [9], in a way as shown on figure No.4.

Accordingly, under **sensitivity**, we understand the proportion of patients with correct positive prognosis of a certain medical outcome in relation to the total number of patients from the sample who have the appropriate outcome or more precisely, sensitivity is measured as a relation between the number of true positive patients and the sum of true positive and false negative patients [10], as shown in expression (5):

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive+False Negative}} \qquad (5)$$

Specificity is used to express the proportion of patients with correct negative prognosis of a certain outcome in relation to the total number of patients without the appropriate outcome from the sample. Precisely, specificity is measured as a relation between the true negative patients and the total number of true negative and false positive patients [11], as shown in formula (6):

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative+False Positive}} \qquad (6)$$

The overall or **general accuracy** of prognosis integrates sensitivity and specificity, and its primary goal is to correctly identify the patients who have, and the patients who do not have the appropriate outcome from the prognosis. It is defined as a proportion between the patients with true positive and true negative prognosis in relation to the total number of patients from the sample. In formula (7) the general accuracy of prognosis is defined as ratio between the sum of true positive and true negative patients in relation to the sum of true positive, true negative, false positive and false negative patients [12]:

$$\text{Accuracy} = \frac{\text{True Positive+True Negative}}{\text{True Positive+True Negative+False Positive+False Negative}} \qquad (7)$$

Applying these formulas in the example from our research, we get the exact same values for sensitivity, specificity and the overall accuracy with 33,76%, 98,72% and 93,78% accordingly.

Optimization of the probability threshold

Obtained results concerning the classification of patients by using the probability threshold of 0,5 in our study are relatively good. Namely, the general accuracy rate in prognosis is good, while specificity shows even better performance results i.e., the prognosis of patients who survived from COVID-19. The only relatively bad performances are found in the spectrum of sensitivity which gives the ratio of correctly predicted patients who died from the virus. However, these results in classification are related to the choice of probability margin on the basis of which the primary systematization of the patients is done. Taking into consideration the fact that the primary goal of the research is identification of the patients with high mortality risk, who should be urgently admitted to hospital or sent to the tertiary healthcare institutions, we cannot qualify the obtained sensitivity rate of 33,76% as a "satisfactory" prediction, despite the high rates of specificity and overall accuracy of prognosis. If the research goal would be to identify the patients with low risk, who would be sent for home treatments, aiming to free the hospital capacities in public health institutions during the pandemic peaks, then the obtained results in prognosis would correspond to the goals established. So, based on the general purpose of prediction, the probability threshold can be changed, and with it, the prediction performances can be changed in aspect of the sensitivity and specificity. Since the aim of our research is identification of high-risk patients, the focus of the prediction performance of the model should be placed towards the sensitivity in prognosis. Therefore, in continuation, few methodological methods are presented in order to demonstrate how to perform an optimization of the probability threshold in order to improve the prognosis from aspect of sensitivity, or, in other words, how to improve the percentage of patients with correct positive prognosis that would likely to die from COVID-19 and thus improve the precision in prognosis and identification of the high-risk patients.

Methodological approach (rules for optimization)

There are many methodological approaches in literature which provide a possibility to optimize the probability threshold or the cut-off point. From the available methodologies for optimization, our choice comes down to the following three: a) the maximum sum rule; b) the maximum product rule; and c) the maximum value of Youden Index rule.

Maximum sum rule

According to the maximum sum rule [13], the probability threshold is considered optimal at the point where the sum of sensitivity and specificity has a maximum value (8):

$$\text{Max SUM} = \text{Max}(\text{Sensitivity} + \text{Specificity}) \quad (8)$$

Maximum product rule

This rule, which is known as the "concordance probability method"[14][15], indicates that the optimal probability margin is achieved at the point at which the product of sensitivity and specificity has a maximum value, as described in expression (9):

$$\text{Max PRODUCT} = \text{Max}(\text{Sensitivity} * \text{Specificity}) \quad (9)$$

Maximum value of the Youden Index rule

As the mere rule shows, optimization of the probability threshold is achieved at the point of maximum value of the so-called Youden Index. If we assume that this index is measured when 1 is subtracted from the sum of sensitivity and specificity, then the Youden Index optimization rule could be defined in the following way [16][17][18], as shown in formula (10):

$$\text{Max YOUDEN} = \text{Max}[(\text{Sensitivity} + \text{Specificity}) - 1] \quad (10)$$

Results and discussion

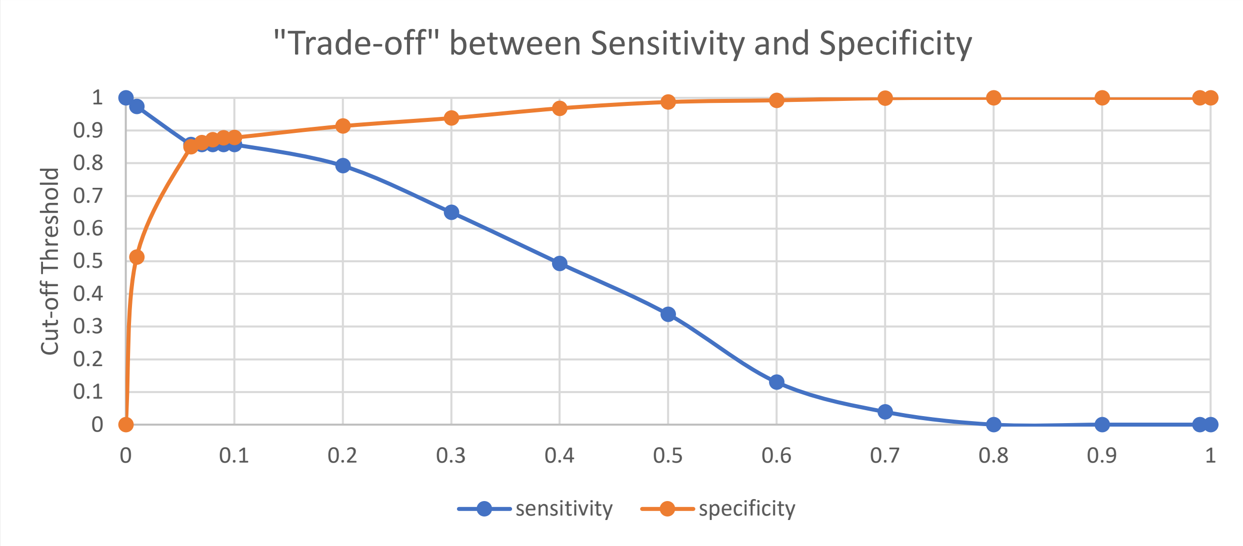
In order to implement these rules for optimization into practice, simulations of the basic model of logistic regression were made, but with using different probability thresholds, whereas from the results of each separate regression, sensitivity and specificity rates were measured and displayed accordingly in separate columns as well as the rate of the overall accuracy of prognosis. The values of the maximum sum and maximum product are additionally calculated and displayed in separated columns, as well as the Youden Index. All of these values, together with the cut-off threshold point are presented in table No.1 titled "optimization of the probability threshold during classification".

Table 1: Optimization of Probability Threshold during Classification

Cut-off threshold	Sensitivity	Specificity	Accuracy	Youden index	Max SUM	Max PROD
0	1	0	0,076012	0	1	0
0,01	0,974	0,512821	0,547878	0,48684649	1,486821	0,499487
0,06	0,857143	0,850427	0,850938	0,70757	1,70757	0,728938
0,07	0,857143	0,863248	0,862784	0,720391	1,720391	0,739927
0,08	0,857143	0,871795	0,870681	0,728938	1,728938	0,747253
0,09	0,857143	0,877137	0,875617	0,73428	1,73428	0,751832
0,1	**0,857143**	**0,878205**	**0,876604**	**0,73534799**	**1,735348**	**0,752747**
0,2	0,792208	0,913462	0,904245	0,70566933	1,705669	0,723651
0,3	0,649351	0,938034	0,916091	0,587385	1,587385	0,609113
0,4	0,493506	0,967949	0,931885	0,461455	1,461455	0,477689
0,5	0,3376	0,9872	0,937808	0,3248	1,3248	0,333279
0,6	0,12987	0,992521	0,92695	0,122391	1,122391	0,128899
0,7	0,038961	0,998932	0,925962	0,037893	1,037893	0,038919
0,8	0	1	0,923988	0	1	0
0,9	0	1	0,923988	0	1	0
0,99	0	1	0,923988	0	1	0
1	0	1	0,923988	0	1	0

Source: Author's calculations

From the data in table No.1. can be noticed that the highest value of sensitivity rate (which is 1) is measured at the probability threshold point of 0, while its lowest value (which is 0) is registered at the probability point of 1. In the case of specificity this is completely opposite. i.e., the lowest value is at the probability threshold of 0, while the highest value comes with the probability point of 1. This means that there is some sort of exchange or "trade-off" between the sensitivity and specificity rates. In other words, as the threshold point increases, the sensitivity tends to decrease while at the same time specificity increases. This is graphically illustrated in graph No.1.



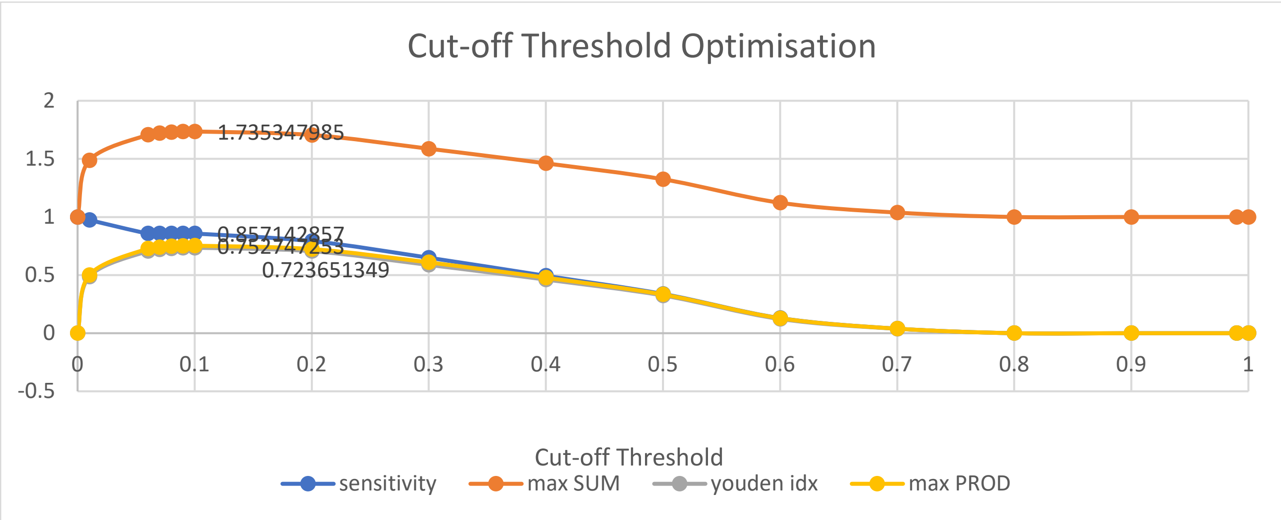
Graph 1: "Trade-off" between Sensitivity and Specificity
Source: Author's calculations

From all of the above, the question arises: What is the optimal probability threshold for classification of the high-risk patients from the sample in the research, which would produce the most accurate prognosis of the death outcome, without interrupting the precision in predicting of the survival outcome? The application of all 3 rules for optimization gives unified answer and that is the probability threshold point of 0,1. Namely, at this optimal point the sum of sensitivity and specificity has the maximum value of 1,735, the product of specificity and sensitivity also has maximum value of 0,752, as well as the Youden Index, whose value is highest at this critical point measuring 0,735. We can see from the column No.2, which refers to sensitivity, that the sensitivity rate has the same value also at the probability threshold of 0,06, 0,07, 0,08 and 0,09. However, at those points the specificity has lower values when compared to the threshold margin of 0,1 and that is the reason why these points cannot be considered as optimal. At the lowest levels of the threshold, the sensitivity rate is even higher, but the specificity rate lowers additionally, thus moving increasingly away from the critical optimization rules. As far the general accuracy of prognosis is concerned, it is interesting to note that it is not crucially to have maximum value at the point of optimization. After all, in our example, at this threshold point this rate is measured 87,66% , while the maximum value of 93,78% is achieved at probability of 0,5. We can **summarize** from the previous analysis, that if the priority is to obtain maximum accuracy in prognosis of the positive outcome of a certain event (in our case death by COVID-19) and at the same the highest possible accuracy in prognosis of the negative outcome (survival in our case), then any of the above rules can be applied for optimization of the threshold point. At the same time, if the aim is to achieve maximum general accuracy of prediction, our logic dictates that we should implement a probability threshold of 0,5, the one that expresses equal chances or odds for appearance of both, the positive and the negative outcome, Namely, if we look at the numbers, the rate of sensitivity at the optimal value of threshold point of 0,1 is 85,71%, which means that out of 77 deceased patients from the sample, the prognosis was correctly given for 66 patients [(66/77)*100]. At this point, the rate of specificity is 87,82%, which means that correct prognosis was given for 822 patients out of the total number of patients who survived which is 936 [(822/936)*100]. The general accuracy of prognosis comes down to 87,66% [(66+822)/1013]*100.

Optimization can be done visually as shown in graph No.2. The optimal sensitivity rate of 0,8571, which is shown with the blue line is in the vertical of the optimal probability threshold point of 0,1 from the axis, while observed vertically, it connects the maximum values from the peaks of the curves of the 3 criteria for optimization: max SUM, max PROD and Youden Index.

ROC curve analysis

The so-called ROC curve (Receiver Operating Characteristic Curve – ROC curve), is a curve used to map the rate of sensitivity in relation to the difference between 1 and the rate of specificity for all possible points of the probability threshold [19]. ROC curve is most often used to calculate or estimate the diagnostic ability of a certain biomarker, but the same analytical instruments can be implemented successfully to grade the ability to predict and classify the statistical models. In fact, the analysis of the ROC curve offers two methodological advantages: the first one refers to the assessment of the diagnostic accuracy of a certain biomarker or medical test (analogously in our case that would be the prediction accuracy score) and the second one refers to the optimization of the probability threshold.



Graph 2: Optimization of the Probability Threshold
Source: Author's calculations

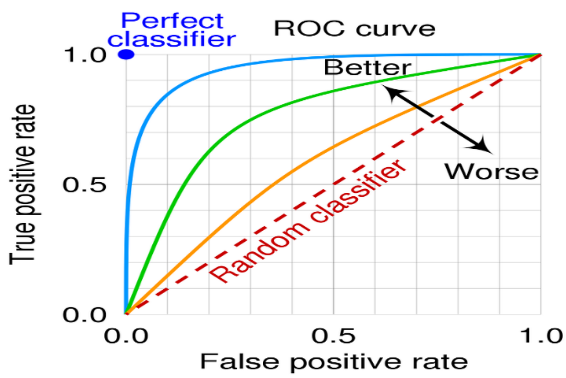
Evaluation of the accuracy of prediction and classification

The basic analytical component which is derived from the ROC curve is the area under the curve (AUC). In the diagnostic (or statistical models) biomarker with value of AUC=1 does perfect classification of the patients into sick and healthy, or dead and survived in our research. On the other hand, biomarker with value of AUC=0,5 means that there are no differences between the patients of both groups which can be distinguished as sick or healthy, dead or survived (coincidental classification). The value of AUC can be any value between these two extremes and the precision of the diagnostic test can be predicted based on the generated value (that is the ability of accurate prediction and classification). The scale on which the test quality and classification accuracy are assessed [20] is given in table No.2, while the shape of the ROC curve according to the different values of AUC [21] is given in graph No.3.

Table 2: Scale of Test Quality and Classification Accuracy

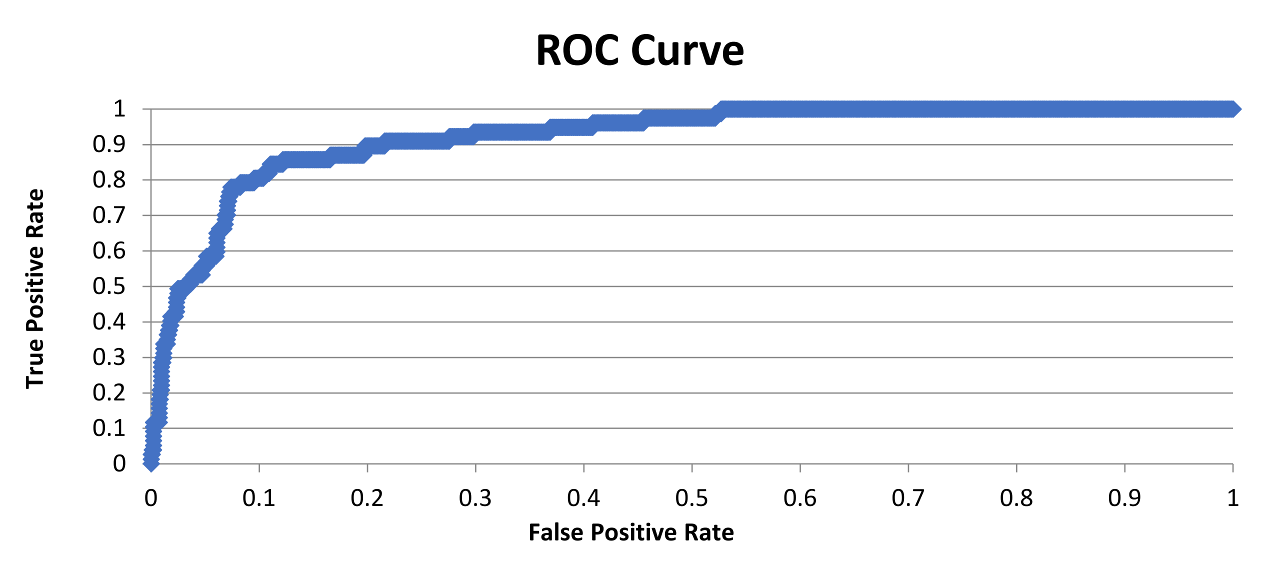
AUC Value	Test quality (classification accuracy)
1	Perfect classifier
0,9 - 1,0	Excellent
0,8 - 0,9	Very Good
0,7 - 0,8	Good
0,6 - 0,7	Satisfactory
0,5 - 0,6	Unsatisfactory
0,5	Random Classifier

Source: [20]



Graph 3: ROC Curve with Different Values of AUC
Source: [21]

To obtain the results that are generated in our model in the aspect of the ROC curve and the AUC value, the software package Xrealstats was used. The value of AUC from the model we use is 0,9224 and according to the quality scale, the model can be given an **excellent grade**, which is correspondent with AUC intervals of 0,9 to 1,0. This means that the chosen model has a **high capacity and accuracy for prediction and classification** of patients with high risk of death by COVID-19. The shape of the ROC curve is shown in the next graph.



Graph 4: Generated ROC Curve from the Source Model

Source: Author's calculations

Optimization by using the Index of Union

Apart from the previously described rules for optimization, the methodological approach for analysis of the ROC curve is also compatible with the ability to bring an optimal choice of the probability threshold by means of additional alternative criteria. One such criterium is the Index of Union - IU which use the absolute difference between the diagnostic measures (sensitivity and specificity) and the values of AUC in order to minimize the rate of wrong classification [22] [23], according to the expression (11):

$$\text{IU} = (\text{Sensitivity} - \text{AUC}) + (\text{Specificity} - \text{AUC}) \quad (11)$$

In other sense, the Index of Union makes it possible to find the point at which the sensitivity and specificity are simultaneously maximal. That is the point or threshold at which the value of IU is maximal (max IU). From the data in table No.3. we can conclude that the optimal threshold point obtained by the Index of Union is consistent with the results from the previous rules for optimization, i.e., that is the threshold of 0,1.

Table 3: Optimization with the Criterium IU

Cut-off threshold	Sensitivity	Specificity	AUC	Index of Union (IU)
0	1	0	0,92248	-0,844960
0,01	0,974	0,512821	0,92248	1,48682
0,06	0,857143	0,850427	0,92248	1,70757
0,07	0,857143	0,863248	0,92248	1,72039
0,08	0,857143	0,871795	0,92248	1,72893
0,09	0,857143	0,877137	0,92248	1,73428
0,1	**0,857143**	**0,878205**	**0,92248**	**1,73534**
0,2	0,792208	0,913462	0,92248	1,70566
0,3	0,649351	0,938034	0,92248	1,58738
0,4	0,493506	0,967949	0,92248	1,46145
0,5	0,3376	0,9872	0,92248	1,3248
0,6	0,12987	0,992521	0,92248	1,12239
0,7	0,038961	0,998932	0,92248	1,03789
0,8	0	1	0,92248	1
0,9	0	1	0,92248	1
0,99	0	1	0,92248	1
1	0	1	0,92248	1

Source: Author's calculations

Conclusion

In this research we applied the rules of maximum sum and maximum product as well as the so-called Youden Index for the purpose of optimization of the probability threshold. The principals of the ROC curve in addition with the Index of Union was also helpful for the same purpose. Prognosis accuracy was evaluated through the status of patient according to the “golden standard” in medical trials, in which sensitivity, specificity and the general accuracy of prognosis were specifically evaluated. The results from the simulation of the probability threshold, revealed that all presented methodologies generate identical results, if the purpose of classification is to maximize the prognosis accuracy from the point of sensitivity. According to them, the critical (optimal) probability threshold point was measured 0,1. We may conclude that the applied methodological approaches offer scientifically sound foundations in the context of mortality risk evaluation and classification of COVID-19 patients. Targeted patients will be subject of precaution with strict measures, protocols and more aggressive treatment in order to minimize the chances of death outcome.

References

[1] Mogilevska-Gruevska D, Gruevski I, Boshevska G, Gjoreski K. Covid-19 outcome prediction model by using radio-diagnostic methods. Eur J Med Health Res. 2024;2(3).

[2] Mogilevska-Gruevska D, Gruevski I, Boshevska G, Gjoreski K. Covid-19 outcome prediction model by using radio-diagnostic methods. Eur J Med Health Res. 2024;2(3).

[3] Lane DM. Probability of a single event: Basic concepts – Introduction to probability. linestatbook.com. 2024.

[4] Harris JK. Primer on binary logistic regression. PMC PubMed Central. 2021. doi: 10.1097/01.ehp.0000483147.19985.b7

[5] Rotella J. Probability, log-odds and odds. Internal material, Montana State University. 2024.

[6] Harris JK. Primer on binary logistic regression. PMC PubMed Central. 2021. doi: 10.1097/01.ehp.0000483147.19985.b7

[7] Zaiontz C. Basic concepts of logistic regression. Real-statistics.com. 2024.

[8] DataCadamia. Threshold/cut-off of binary classification. datacadamia.com. 2024.

[9] Trevethan R. Sensitivity, specificity and predictive values: Foundations, pliabilities and pitfalls in research and practice. PMC PubMed Central. 2017. doi: 10.1155/2017/8952732

[10] Mourosi JT, Anwar S, Hosen MJ. The sex and gender dimensions of COVID-19: A narrative review of the potential underlying factors. Infect Genet Evol. 2022;95:105043. doi: 10.1016/j.meegid.2022.105043

[11] Baratloo A, Hosseini M, Negida A, El Ashal G. Part 1: Simple definition and calculation of accuracy, sensitivity and specificity. PMC PubMed Central. 2015. doi: 10.4103/2277-9175.151868

[12] Baratloo A, Hosseini M, Negida A, El Ashal G. Part 1: Simple definition and calculation of accuracy, sensitivity and specificity. PMC PubMed Central. 2015. doi: 10.4103/2277-9175.151868

[13] Kaivanto K. Maximization of the sum of sensitivity and specificity as a diagnostic cut-point criterion. J Clin Epidemiol. 2008;61(5):517-8. doi: 10.1016/j.jclinepi.2007.10.008

[14] Liu X. Classification accuracy and cut point selection. Stat Med. 2012;31(24):2676-2686. doi: 10.1002/sim.4509

[15] Unal I. Defining an optimal cut-point value in ROC analysis: An alternative approach. PMC PubMed Central. 2017. doi: 10.12669/pjms.336.13198

[16] Perkins NJ, Schisterman EF. The Youden index and the optimal cut-point corrected for measurement error. Biometrical J. 2005;47(4):428-441. doi: 10.1002/bimj.200410135

[17] Unal I. Defining an optimal cut-point value in ROC analysis: An alternative approach. PMC PubMed Central. 2017. doi: 10.12669/pjms.336.13198

[18] Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden index and optimal cut-off point estimated from observations affected by a lower limit of detection. PMC PubMed Central. 2008. doi: 10.1097/01.epi.0000310942.40399.da

[19] Unal I. Defining an optimal cut-point value in ROC analysis: An alternative approach. PMC PubMed Central. 2017. doi: 10.12669/pjms.336.13198

[20] Trifonova OP, Lokhov PG, Archakov AI. Metabolic profiling of human blood. Russian Academy of Medical Sciences. 2012.

[21] Kilic I. ROC curve and AUC: Evaluating model performance. medium.com. 2023.

[22] Hassanzad M, Hajian-Tilaki K. Methods of determining optimal cut-point of diagnostic biomarkers with application of clinical data in ROC analysis: An update review. BMC Med Res Methodol. 2024. doi: 10.1186/s12874-024-01944-6

[23] Unal I. Defining an optimal cut-point value in ROC analysis: An alternative approach. PMC PubMed Central. 2017. doi: 10.12669/pjms.336.13198