Article

Not peer-reviewed version

An Ensemble Method for Determining the Importance of Selected Social and Health Factors Affecting the Inpatient Treatment Quality

Jelena Mišić , Aleksandar Kemiveš , Milan Ranđelović , Dragan Ranđelovic *

Posted Date: 11 May 2023

doi: 10.20944/preprints202305.0817.v1

Keywords: machine learning; classification algorithms; logistic regression; feature selection; ensemble methods; optimization algorithms

Article

An ensemble method for determining the importance of selected social and health factors affecting the inpatient treatment quality

Jelena Mišić [1,5], Aleksandar Kemiveš [2,3], Milan Ranđelović [4] and Dragan Ranđelović [5,*]

[1] Faculty of Electronic Engineering, University of Niš; jelena.misic@fdb.edu.rs
[2] Department for Postgraduate Studies, Singidunum University, Belgrade; kemives@outlook.com
[3] PUC Infostan technologies, City of Belgrade
[4] Science technology park Niš, Serbia; milan.randjelovic@ntp.rs
[5] University Union-Nikola Tesla Belgrade, Faculty of diplomacy and security; dragan.randjelovic@fdb.edu.rs
* Correspondence: dragan.randjelovic@fdb.edu.rs

Abstract: This study aims to determine the influence of selected social and health factors on the quality of inpatient treatments when regional health organizations use stacking ensemble model. The proposed procedure is based on the logistic regression method, which is used for direct prediction in the case of good fitting data and impossibility of including aggregation of classification algorithms but also in the opposite case, for fine calibration to obtain prediction. In opposite case, the procedure uses classification algorithms and several filter methods which initially rank individual factors according to their importance to reduce the dimensionality of problem and, in such way. obtaining one optimized classification prediction. The proposed procedure is trained using one part and tested using another part of dataset from case study which enables the generalization of the solution to the set goal and is verified by the Kaggle dataset Brest cancer. Case study was conducted using the real data acquired in the region connected to the city of Nis, Republic of Serbia. The obtained results show that the proposed model achieves better results that each of methods included in this stacking ensemble-regression and classification used individually.

Keywords: machine learning; classification algorithms; logistic regression; feature selection; ensemble method

---

1. Introduction

The World Health Organization (WHO) has adopted a program Health21 as a general health policy framework in the WHO European region in the 21st century [1,2]. All planned, education and approach to the health care institutions should be in compliance to the Health21 program, so as the human resources. According to the Health21 all countries in the region should ensure that all health professionals are trained and acquired applicable knowledge, approach and skills necessary for the protection and promotion of the public health [2,3]. According to the Health21 program, bellow stated points 17 and 18, each EU member country or membership candidate country have to fulfil given objectives as well as the obligation to providing health care for citizens and to adopt appropriate health care plan:

- 17. "Funding health services and allocating resources," calling for "sustainable financing and resource allocation mechanisms for health care systems based on the principles of equal access, cost–effectiveness, solidarity, and optimum quality",
- 18. "Developing human resources for health" to ensure that health professionals and others "have acquired appropriate knowledge, attitudes and skills to protect and promote health".

From the 2012, i.e., 2013 until today, the Republic of Serbia, for example, adopted such program [4] as a strategic and operational document of the National Health Insurance Fund. It included implementation of the policy of mandatory health insurance. In order to provide more complete

 doi:10.20944/preprints202305.0817.v1

merge of standard rights (content and scope) and health care arisen from the mandatory health insurance due to needs of end-users. Afore mentioned is the ground plan for implementation of the health care networking in the Republic of Serbia and it's been determined on the basis of the following factors (listed in the Law on Health Care of the Republic of Serbia [5]): plan development, health, population, number and age structure of the population, the number of existing institutions, capacity and distribution of health institutions, range of urbanization, development and transport connectivity of individual areas for equal access to health care. The plan that we have in mind should be implemented in all levels of health care, starting from general practitioner (GP) or GP surgeries all through the specialized clinical centers [6]. This plan also encompassed the necessity to implement health service in all levels: primary—outpatient at the GP surgeries, pharmacies and dispensaries, secondary—general and specialized hospitals and tertiary such as clinics, institutes and clinical centers on third level of expertise.

The Government of Republic of Serbia adopted action plan on the national level for prevention, treatment and control of cardiovascular diseases in the Republic of Serbia until 2020 since the share of cardiovascular diseases is dominant in Serbia. This plan, among other things encompasses the expansion of certain capacities at the clinical centers in the country. Videlicet comparing to the average standardized mortality rate in Europe of 410.1 per 100,000, the Republic of Serbia with the standardized mortality rate 504.3 per 100,000 citizens in 2007 was in the group of countries with a high mortality risk due to the cardiovascular diseases. According to the mortality rate in the Republic of Serbia in 2007, cardiovascular diseases accounted for more than half of all deaths (56.0%). In order to achieve the primary objective of health care in all countries in the region—more successful treatment of patients in the future respectively, it is also necessary, to monitor the impact of various factors on the prevention of different diseases, such as, for example, early diagnostics, changes in the age structure of population and economic power, and certainly monitoring hacking a good approach the of health care institutions organizational network at different levels of expertise, etc. [7,8].

Although one can find papers that deal with the successful organization of treatment in different ways, for example taking into account the influence of factors of different nature in different types of diseases [9–11], sociological determinants [12], impact of political, economic, environmental and other external influences [13], the authors of this paper deals with discussion of influence concrete non-med factors at successful inpatient treatment connected with good organization of health institutions network in one region. Having in mind at the same time the fact that cardiovascular diseases are the most numerous group of patients in today's global world, the authors have taking from mentioned factors available and valid for these kind of patients in account and immersed in considering possibly solution for this problem. In order to check the worthiness of the plan and the organization's network of inpatient health institutions for treatment of cardiovascular patients, the authors opted for studying the influence of various factors from the group of non-medical factors on the successful treatment primarily starting from the level of hospital's competence through the number of days of treatment, age and gender of treated patients etc. Starting from the basic logic hypothesis that the organization in the network of health care institutions can be considered as good only if it turns out that the biggest impact on the level of adherence to the treatment patients is at the hospital where the patient was treated, what kind of environment they live and the number of days of treatment since it essentially means that the plan is practical good coverage at the levels and the timely referral of patients with more serious health issues to institutions of lower level rather than to higher level of expertise. The authors performed training and tests using the data of the Institute of Public Health in city Nis, Republic of Serbia for inpatient treatment of cardiovascular patients in health institutions dividing them into two groups: the tertiary level, as a big, such as Clinical Center—Nis and secondary level, as small which connoting all other hospitals in Nis and Toplica district of the Republic of Serbia, all of that in the period from 2006 by 2009. whereby they used the data for first 2 years to train the procedure, and the procedure was tested on the entire dataset. To solve the problem, the authors can use classical statistical methods such as factor analysis and above all regression analysis, but in today's information society, data mining and machine learning all are more often used. For the sake of comprehensiveness of the study, it should be noted that the whole scientific discipline of multi-criteria decision-making is also concerned with determining the importance of individual factors, but its methodology will not be considered in this paper, because the author's main goal is oriented to find a solution for the set problem which will be realized by a generic procedure based on the stacking ensemble approach to produce an appropriate model. Taking into consideration lots of features and various applications data mining and machine learning and advantages of this technique in comparison to the existing classical statistical methods in healthcare, see [14,15], using of data mining and machine learning is an imperative [16–21]. In modern society based on information technologies (IT), authors of this study decided to perform research on the selected case study using machine learning and data procedures which use mining technology.

Classical statistical methods are often wrong in particularly manual hypothesis testing, not practical and even may not be valid when operating with large numbers of variables, user specifies variables, functional form and type of interaction what may influence resulting models, assumptions on linearity, probability distribution, etc. On the other hand we have the advantages of data mining and machine learning such as: undependability from the number of observed factors and size of data and probability distribution, feature selection algorithm techniques that are used for classification and optimization tasks what gives better possibility to predict the outcome for new samples. It is also known today that in both mentioned group of machine and statistic learning, good results can be found in using suitable so called ensemble methods see [22–25]. Among them the best results give application of so called stacking group of ensemble methods which use aggregation of different algorithms to obtain better predictive models than could be obtained from any of the algorithms individually [26–28]. Ensemble modeling are methods from the group of supervised machine learning and involve running two or more associated but different models and then combining the results into a single result to improve the accuracy of predictive data. Their application in all areas of human activity and especially in healthcare, due to the inestimable importance of human health, is gaining importance due to the advantages it possesses: Better forecasting; A model is constant; Better results and Error reduction.

Taking into account the already established the logical basic hypothesis that the organization in the network of health care institutions can be considered as good only if it turns out that the biggest impact on the level of adherence to the treatment patients from the hospital type and the days of treatment also and what kind of environment are living a patients the authors set as fundamental hypothesis of this research that it is possibly to aggregate more different type algorithms to construct an ensemble procedure which has better characteristics than each of included, individually. This is especially important from standpoint that problem set in this paper could be solved using prediction from the site of classic prediction approach in the form of concrete factors impact in the prediction formula, classification prediction approach in the form of predicted belonging of existing factors to the concrete class or using their aggregation in one ensemble.

Keeping in mind all of the above, as basic fact that each prediction depends not only on the chosen methodology but also on selected dataset and the used software, authors consider the possibility of using one kind of stacking ensemble machine learning generic procedure which produces a models includes both of two mentioned methodologies to take advantage of advantages and eliminate disadvantages of both, for interior validation 10 cross validation an for exterior validation and generalization they chose to use the known dataset from Kaggle platform and software from known manufacturer. In this context they construct one such procedure using filter model which presents a predetermined mining algorithm from a group of feature selection and logistic regression and proved its advantage on one study. So, on the end of this paper after necessary discussion authors propose this procedure for solving problem of determination of individual importance each factor on dependent variable from the set of many factors in many problems of different types and connected with different datasets. For the purpose to meet the set objective authors organized the paper as follows: first chapter consists of mentioned introductory considerations, second chapter is devoted to the background review i.e., bibliography report in discussed area and state of the art in used methodologies for solving stated problem i.e., logistic regression, classification and future selection, third chapter where authors divided their attention

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2023 doi:10.20944/preprints202305.0817.v1

into two subchapters; materials and methods in which authors described used material i.e., dataset considered for training and in case study and in second subchapter authors deal with proposed ensemble algorithm constructed using in the previous chapter already mentioned. Fourth chapter is dedicated to considering of application of proposed model and discussion about obtained results, fifth chapter is reserved for technical solution as implementation of proposed solution and last chapter contents short conclusion remarks. On the end of the paper references are listed.

2. Background review

In this section, the authors dealt with literature review and the state of the art in the field of determining affecting of selected factors on inpatient treatment successfulness. Authors gave the review of existing literature about this problem that is and subject of this paper in first subchapter, including separate subchapter with a critical review of the existing literature on the application of classification and ensemble modelling methods relating to inpatient treatment quality. Authors considered in second subchapter the state of the art in the field of application machine learning classification algorithms in solving this problem. The authors discussed in this second subchapter a two used subgroup of machine learning which they used to propose one new ensemble model, namely classification algorithms in first subsubchapter and a feature selection algorithms in second subchapter.

2.1. Literature review

All countries strive to offer healthcare services to their entire population while maintaining quality, fairness, and accountability. While the approaches may differ, it's typically the central government's responsibility to establish a broad framework for financing and organizing healthcare, while the management of the healthcare system is often shared among regional authorities.

The authors of a thorough study [29] selected several countries including Canada, Denmark, England, Finland, France, Germany, Italy, the Netherlands, and New Zealand. Their aim was to demonstrate a variety of approaches to healthcare financing and organization, as both of these factors have an influence on the methods used for capacity planning [30–32]. Such studies exist for all developed countries worldwide such as USA and Canada [33], Australia [34], China [35], etc. These studies indicate that one of the most important tasks for good healthcare of citizens is good capacity planning in both, number and level of specialization, and competence of network of health care institutions as well as continuous monitoring of the obtained estimates.

Authors of the paper [36–39] for example observed patient satisfaction with hospital care from different standpoints in the veterans affairs health care system in USA: racial, ethnic, and gender equity; potential difference between private and state hospitals, quality of nurses work. The considered subject in the papers [40,41] was patient satisfaction with quality of care at the Kingdom of Saudi Arabia.

In the papers [42–45] general approach is considered respectively to the possibilities of use of the business intelligence in hospitals and the factors that affect a survey on data mining approaches for healthcare using the example of the Hasheminejad hospital in Iran, a survey on data mining approaches for healthcare, data mining applications in healthcare and possibilities of prediction and decision making in healthcare using data mining.

Authors of the paper [46] provided ten key principles for successful health systems integration which enables better conditions for successful treatment of patients but also and their satisfaction with hospital care. In [47] authors studied the role of information systems in healthcare: current research and future trends, in [48] data mining in hospital information system including new fundamental technologies, in [49] determining of factors influencing the success and failure of hospital information system and their evaluation methods and in [50] is described procedure for implementing information systems in health care organizations.

By analyzing literature the authors of this paper discovered different applications of regression analysis [51,52] and data mining and machine learning in the treatment of heart diseases [53,54], in health information exchange-based, risk surveillance system as for example on case of state Main [55],

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2023 doi:10.20944/preprints202305.0817.v1

quality control of complex mixtures, including for example herbal medicines as one ensemble learning application in [56] etc. Using machine learning in medicine authors found in one review article [57], also, one very extensive collection of articles of on special issue of Algorithms journal [58] is a good source of knowledge about the application of different ensemble methods in certain areas of human activity. It is particularly interesting review article [59] which deals with ensemble classification and regression-recent developments, applications.

Yet the authors could not find to many papers which deals with the application of classic statistic technique and/or machine learning algorithms in determining the goodness of network organization of health institutions by level of expertise on one territory also and other selected social and health factors, unless already listed as [16] but they found several examples which deals with determination of influence of different social and health factors on successful inpatient treatment as for example: [60] which deals with effects of demographic and institutional characteristics in patient satisfaction with hospital care, [61] which consider social and health dimensions of measurement and assessment in the performance model for hospitals, [62] dealing with political, social, economic and cultural factors affecting how different global regions initially reacted to the COVID-19 pandemic, [63] determining the impact of social and health incentives on intention to stay in hospital, [64] describing using machine learning in healthcare systems for quality of services, etc.

As a consequence of that established facts the authors have decided to prepare a case study considered in this paper as the basis for discussion in that sense i.e., to propose one novel model for determining the importance of selected social and health factors affecting the successful inpatient care.

2.1.1. Literature review of the application of classification modelling methods relating to inpatient treatment quality

There are several taxonomies of measures [65] used to assess the treatment quality of health care institutions in organizational sense. Themost commonly used model is named the Donabedian model [66]. It has classified those measures as either a structure, process, or outcome measure: structural measures give information about a health care provider's capacity and available systems; process measures indicate what a healthcare organization must do to maintain or improve patients health treatment outcome. These measures typically reflect generally accepted recommendations for clinical practice, services and outcome measures which directly reflect the impact of the health care service on the health status of patients. Also, in practical sense it is important to know that the patient treatment quality has a several different nature factors and we could find in literature their taxonomy on medical, social, economic factors, etc. From application sense, in the same way it could be discussed about different types of current methods for assessing the quality of medical care but from other side, in methodological-technological sense, we could talk about classical statistical and modern methods of machine learning. At the same time, the most up-to-date approach in this sense is represented by different ensemble methods which the authors have used in solving the problem considered in this paper—determining influence of social and medical factors on inpatient treatment quality. The authors chose to apply outcome measure from the standpoint of the influence of social and medical factors on inpatient treatment quality. This methodology uses ensemble machine learning method which is based on classification modelling. One critical, systematic review of the existing literature on the application of classification modelling methods related to general medical application is given by Khan et al. in [67] and the application of classification modelling methods related to the prediction of the hospital length of stay is considered by Zikos et al. in [68] as well as in his doctoral thesis [69]. Samaneh Sheikh-Nia has solved the same problem using standard and ensemble based classification techniques. As we have already mentioned in the previous subchapter 2.1 Literature review, in the literature we could find a lot of work that deals with the application of classification methods in determining the influence of different types of factors, from medical and social to economic, including, in addition to classification, forecasting and prediction [70]. F It could be also used in the diagnosis and prediction of the development of various diseases, such as breast

 doi:10.20944/preprints202305.0817.v1

cancer in [71], HIV in [72], COVID19 in [73] etc. but the authors have not found any application of the proposed generic model.

2.2. State of the art

Beside known and wide applied conventional statistic regression methodology in prediction modelling today is trend the machine learning. Machine learning relies on statistical analysis and artificial intelligence to learn concepts, including models and rules, based on the induction of logical rules that can be understood by humans. This learning process involves dividing a dataset used for learning into a learning set and a test set, where the test set is used to verify the validity of the learned knowledge. Predictive accuracy is the primary measure of the correctness of the learned knowledge, representing the percentage of success in classifying new rules using the learned rules. The goal of prediction is to create a model that can draw conclusions about a unique aspect of a dependent variable based on a combination of independent variables. The selection of variables from the available dataset affects the precision and accuracy of the generated prediction models. Therefore, already from the data preprocessing phase, various techniques are used to select relevant variables and assess their importance on the predictor's output as well as filter feature selection methods, which are used in the proposed prediction model to reduce the number of input variables and in this way to reduce the cost and improve the prediction characteristics of the model. In classification problems, sensitivity quantifies the avoidance of false negatives, while specificity does the same for false positives. The compromise between these measures, which is otherwise difficult to achieve, is shown by the so-called receiver operating characteristic curve.

In this case study are used the logistic regression as basic method from the conventional statistic group of methods where the basic measures of goodness of fit of proposed model with considered data are generated with Hosmer and Lemeshow test. In the case that this test is bad authors suggested on the relevant literature [74–76] using classification test which exists in classification supervised machine learning as so called function method and could be implemented classification discrimination using classification and feature selection algorithms and after that evaluated with most important classification measures as for example Area under the curve, Accuracy, etc.

2.2.1. Classification method based on machine learning

Classification is a widely studied topic in machine learning-based systems, which is used to help domain experts identify knowledge from large datasets. Classification algorithms are predictive methods that use supervised machine learning. This involves having a group of labeled instances in at least two classes (attributes) of objects and predicting the value of a required categorical type of class (attribute) based on the values of other predictive attributes. The classification algorithm analyzes the attribute values and discovers relationships between them to achieve accurate prediction results. Common classification algorithms include regression-based methods (e.g., linear regression, isotonic regression, and logistic regression), decision trees (e.g., J48, ID3, random forest, and C4.5), Bayesian classifiers (e.g., naive Bayes, Bayesian logistic regression, and Bayesian network), artificial neural networks (single-layer perceptron, multi-layer perceptron, and support vector machine), and classifiers based on association rules (e.g., PART, JRip, and M5Rules) [77].

The main task of machine learning from data is to select an appropriate classification algorithm for a specific application. In this study, a classifier that classifies results into two classes, positive and negative, is used. The possible prediction results are presented in the form of a confusion matrix presented in Table 1.

Table 1. The confusion matrix for two-class classifier.

		Label-Predicted	
		Positive	Negative
Label-Actual	Positive	TP	FN
	Negative	FP	TN

Table 2 shows that the total sum of positive and negative cases is equal to the number of members in the set being classified, denoted by N, which can be calculated as TP + FN + FP + TN = N. Known quality evaluation measures for a two-class classifier, include accuracy, precision, recall, and F measure, but the Receiver Operating Characteristic (ROC) curve is commonly used to evaluate the performance of a classifier in predicting outcomes. It is plotted with false positive rate on the x-axis and true positive rate on the y-axis, and certain points on the curve have specific meanings [78,79]. For example, a point at (0,1) represents perfect prediction, while a point at (1,1) represents a classifier that labels everything as positive and a point at (1,0) represents a classifier that labels everything incorrectly. The area under the ROC curve (AUC) is a measure of the diagnostic accuracy of the model, and values greater than 70% indicate a good classification process.

According to reference [80], when using naive Bayes or neural network classifiers, the output in ROC space is a probability or score, whereas with a discrete classifier, only a single point is produced that represents the degree to which an instance belongs to a certain class. In practice, classification is a machine learning and data mining task that involves separating instances in a dataset into predetermined classes based on input variable values, as cited in reference [81]. To achieve this task, the classification procedure involves several steps, including selecting classifiers to apply the classification algorithm, choosing a class attribute (output variable), splitting the dataset into training and test sets, training the classifier on the training code set when the class attribute values are known, and testing the classifier on the test set where the class attribute values are hidden.

The testing phase involves using the classifier to classify the test samples based on their predetermined class attribute classes. If the classifier makes a high percentage of errors in the test data, it can be concluded that a bad and unstable model has been created. In such cases, it is necessary to improve the model by modifying the applied classification process. Previous research has shown that the most commonly used classifiers include Bayes networks, decision trees, neural networks, and others [82].

Authors used for proposed model same of mentioned classification algorithms and because of that it is given short description of those used in continuation of this subsection.

Naive Bayes

Bayes classifier, unlike Bayes networks [83], produces a prediction model that is strongly independent from the assumption, and provides a straightforward and easy-to-understand approach for displaying, using, and inducing probabilistic knowledge [84]. The benefits of using a naive Bayes model are its simplicity, efficiency, ease of interpretation, and suitability for small datasets.

Decisions trees

Decision trees [85] work by dividing the data into nodes and leaves until the entire data set is analyzed and the ID3 [86] and C4.5 [87] algorithms are the most commonly used decision tree algorithms. Advantages of decision tree classifiers include their simplicity, the ability to work with numerical and categorical variables, quick classification of new examples, and flexibility.

LogitBoost

The LogitBoost [88] type of algorithm is widely applied in practice as classifier because it is realized as ensemble boosting algorithm and has measured good values for important classification functions. It is based on using the principle that finding multiple simply rules could be more efficient than finding a single one complex and precise and because of that this algorithm represents one general method for improving the accuracy of machine learning algorithms.

Logistic regression

Calibration is the process of adjusting the output of a classification algorithm's posterior probabilities to match the true prior probability distribution of the target classes. Many authors suggest calibrating machine learning or statistical models to predict the probability that the outcome is 1 for every given data row [89,90]. Calibration is used to transform classifier scores into class membership probabilities in classification. Univariate calibration methods, such as logistic regression, exist to transform classifier scores into class membership probabilities in the two-class case.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2023 doi:10.20944/preprints202305.0817.v1

Logistic regression [91] is a statistical technique that analyzes a dataset where one or more independent variables determine an outcome measured with a dichotomous variable that only contains data coded as 1 or 0. It does not require a linear relationship between the dependent and independent variables and does not require the independent variables to be normally distributed. It is based on the theoretical assumptions and Equations (1)–(5) given below.

Logistic regression methodology aims to identify the most suitable model that can describe the relationship between a dichotomous characteristic of interest (dependent variable or outcome variable) and a set of independent variables (predictor or explanatory variables). The logistic regression algorithm produces coefficients (with their standard errors and significance levels) that can be used to create a formula predicting the logit transformation of the probability of the presence of the characteristic of interest:

$$\mathrm{logit}(p) = b_0 + b_1X_1 + b_2X_2 + \ldots + b_kX_k \tag{1}$$

where p is the probability of presence of the characteristic of interest, b_0, b_1, b_2, ..., b_k are the coefficients of the regression equation and X_1, X_2, ..., X_k, independent variables. The logit transformation is defined as the logged odds:

$$\mathrm{odds} = \frac{p}{1-p} = \frac{\text{probability of characteristics presence}}{\text{probability of characteristics absence}} \tag{2}$$

$$\mathrm{logit}(p) = \ln\left(\frac{p}{1-p}\right) \tag{3}$$

By taking the exponential of both sides of the equations (1) and (3), as given above, we obtain:

$$\mathrm{odds} = \frac{p}{1-p} = e^{b_0} \cdot e^{b_1X_1} \cdot e^{b_2X_2} \cdot e^{b_sX_s} \cdot \ldots \cdot e^{b_kX_k} \tag{4}$$

When a variable X_i increases by 1 unit, with all other factors held constant, the odds will increase by a factor e^{b_i}.

$$e^{b_t(1+X_t)} - e^{b_tX_t} = e^{b_tX_t} = e^{b_t(1+X_t)-b_t X_t} = e^{b_t+b_tX_t-b_t X_t} = e^{b_t} \tag{5}$$

This factor e^{b_i} is the odds ratio (O.R.) for the independent variable X_i and it gives the *relative* amount by which the odds of the outcome increase (O.R. greater than 1) or decrease (O.R. less than 1) when the value of the independent variable is increased by 1 units.

Statistical software programs, such as IBM SPSS [92], offer various methods for performing logistic regression. The authors have used the Enter method for their proposed model as default in used SPSS package.

2.2.2. Future selection techniques based on machine learning

Many classification methods are highly sensitive to data dimensionality and the ratio of instances to features. However, even less sensitive methods can benefit from dimensionality reduction. Attribute ranking evaluates each attribute independently of others and does not consider dependencies between attributes. Subset selection, on the other hand, searches for a set of attributes that together provide the best result. The concept of feature selection is relevant only for classifiers that are highly sensitive to the initial ordering of input features.

Feature selection methods can be realized using three groups of methods [93]:

- Filter, where belongs as most known Infogain, Gainratio, SU,…
- Wrapper, between these known are BestFirst, LinearForwardSelection,… and
- Embedded, different types of decision trees algorithms as for example J48, PART,…

In their proposed model, the authors used a filter-ranker evaluation approach to detect attributes for selection.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2023 doi:10.20944/preprints202305.0817.v1

Filter-ranker methods

To reduce the number of attributes in the model and determine the optimal subset of attributes that provide good predictive performance, the authors employed a filter-ranker evaluation approach. This approach ranks the attributes based on their importance, helping to identify the most relevant attributes for the model. By using this approach, the authors were able to select a smaller set of attributes that still had strong predictive characteristics. The Weka software [94] is designed to reduce the volume of information by applying various algorithms and techniques. This reduction of information could potentially include the suggested filter-ranker evaluation approach previously. In machine learning, dealing with a large number of attributes can make it difficult to apply techniques such as regression or classification to the collected data. Therefore, the feature selection as an data modelling technique are used to handle the problem of irrelevant and redundant attributes. This approach involves evaluating different attributes using various measures, such as ChiSquare, Relief and GainRatio, to rank them in terms of relevance.Different measures [95] are utilized in the proposed model through appropriate classifiers, which are briefly described in the following sub-section.

GainRatio

Entropy is a measure of the disorder or uncertainty in a system, and it is often used in information theory as a measure of the amount of information contained in a message or data set. In the context of decision trees and attribute selection, entropy is used as a measure of the impurity of a set of examples. The goal is to select the attribute that leads to the greatest reduction in entropy, which in turn leads to a more homogeneous subset of examples. The entropy of Y is:

$$H(Y) = - \sum_{y \subset Y} p(y) \cdot log_2(p(y)) \quad ? \qquad (6)$$

Because entropy is used as a measure of impurity in a training set S, it is possible to create a measure that reflects the amount of additional information about an attribute provided by the class, which indicates the extent to which the entropy of the attribute decreases [96].

InfoGain is a measure that evaluates the worth of an attribute by calculating the amount of information gained about the class when the attribute is known. It is defined as the difference between the entropy of the class before and after splitting on the attribute:

$$InfoGain\left(Class, Attribute\right) = H\left(Class\right) - H\left(Class \mid Attribute\right)? \qquad (7)$$

where symbol H is the information entropy which is calculated using Equation (6).

The GainRatio [97] is a modified version of InfoGain, which is a non-symmetric measure designed to address the bias of InfoGain and the formula for calculating GainRatio [98] is as follows:

$$\text{GainRatio} = \frac{InfoGain}{H\left(Class\right)}? \qquad (8)$$

The equation (7) shows that when the attribute "Attribute" needs to be predicted, the InfoGain is normalized by dividing it with the entropy of "Class", and vice versa. This normalization ensures that the GainRatio values, given with equation (8) always fall within the range of [0, 1]. If the GainRatio is equal to 1, it means that the knowledge of "Class" completely predicts "Attribute", and if the GainRatio is equal to 0, it indicates that there is no relationship between "Attribute" and "Class".

ChiSquaredAttributeEaval

ChiSquaredAttributeEaval is classifier based on chi-square test used to test the independence of two events so that for given data of two variables, we can obtain observed count O and expected count E and using Chi-Square measure [99] how expected count E and observed count O deviates each other what is shown in Equation (9):

$$\chi_c^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} \qquad (9)$$

 doi:10.20944/preprints202305.0817.v1

In equation (9) c is degrees of freedom, O_i is observed value and E_i is expected value whereby degrees of freedom refers to the total number of observations reduced for the number of independent constraints which are imposed with the observations and it can be defined as the total number of observations minus the number of independent constraints imposed on the observations.

Relief

Relief for attribute estimation [100,101], estimates the attribute value by repeatedly sampling the instances and considering the value of the obtained attributes from the nearest instances of the same or different class. This method assigns a weight score to each attribute based on its ability to discriminate between classes and then selects them attributes whose weight exceeds a user-defined threshold as matching attributes.

3. Materials and Methods

As mentioned in Section 1, due to the fast and turbulent development of advanced computer technique-based solutions for different impact prediction and social and health factors and their effect on the inpatient treatment quality, mortality of especially cardio patients caused by non-adequate health care has been a research hop topic in the information field since the beginning of the 21th century. Application of machine learning and especially ensemble methods to the prediction, including technical implementation of obtained solutions in the form of mobile software tools, has been one of the current trends in the field of data prediction. However, as mentioned in Section 2, there have still been fewer studies on ensemble methods that combine machine learning-based methods especially in the field of healthcare which deals with non-medical factors. Therefore, additional research of aggregated methods has been needed, which has been the main motivation of this study.

This study introduces an efficient ensemble stacking machine learning procedure for prediction of impact of selected social and health factors on the inpatient treatment quality. The proposed model performance is trained and tested by case study using the data obtained from the Institute of Public Health in Nis, which were acquired in region connected with city of Nis including Toplica area, Republic of Serbia.

The collected raw data were first classified into two classes, those with the positive outcome of a patient's treatment and those with the false outcome of a patient's treatment also the data wed and normalized. After application of proposed model on case study, authors verified proposed procedure on Kaggle dataset notated Breast cancer.

3.1. Materials

Materials used in this paper are two datasets: one very known dataset called Breast cancer from Kaggle datasets platform which was used for verification and other dataset was used for training and test which is generated in performed case study that the authors considered in solving of the given problem and for checking stated main hypothesis of this paper.

3.1.1. Data acquired during 2006 to 2009, by the Institute of Public Health in Nis

Aiming to evaluate the individual impacts of social and health factors affecting the patient treatment quality, this study considers several parameters, including education level (high level is one group while all other level of education belong to second group), medical institution level (e.g., clinical centers have higher level while all other medical institutions have lower level), place of residence (high level of housing is in urban environment i.e., city of Nis), gender(implies high level for female), age of patients (older than 50 years are notated as high level), and the length of patients treatment (treatment longer from 15 days is noted as high level). These parameters affect the treatment quality of patients with cardiovascular disease and define the treatment outcome, which can be positive or negative.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2023 doi:10.20944/preprints202305.0817.v1

The case study was conducted using data acquired during the period from 2006 to 2009, by the Institute of Public Health in Nis and dispensary medical institutions, including the Clinical Center of Nis, Institute for Prevention and Rehabilitation of Niska Banja, Military Hospital Nis, Special Hospital of Soko Banja, as well as districts (Ozren and Toplica) that included Medical Center of Prokuplje and the Health Center of Kursumlija.

Data analysis was performed using an innovative ensemble machine learning based generic procedure that combines two techniques, namely, conventional logistic regression analysis and classification performed by common application classification and the feature selection algorithms. Selected feature selection algorithms are based on the filter group ranked model and select ranked sub-set of attributes according to prediction accuracy estimation, given by selected classifier.

All of this data are given as supplementary excel file and look as it is shown in the table—Table 2.

Table 2. social and health factors used in the case study.

Variable's serial number	Social and health factor
1	Education
2	HospitalType
3	Gender
4	Age
5	DaysofTreatment
6	UrbanHousing
7	Outcome

In Table 2, the meaning of the listed factors is as follows:

- Education has the value of one for high education of patient;
- HospitalType has a value of one for treatment at the Clinical Center—Nis, and a value of zero for all other hospitals;
- Gender has a value of one for female patients, and a value of zero for male patients';
- Age represents the number of patients years—older than 50 years are notated as high level ;
- DaysofTreatment is the number of days of a patient's hospital stay—longer from 15 days is noted as high level;
- UrbanHousing has the value one for patients living in the city;
- Outcome has a value of "true" for positive outcome of a patient's treatment, but a value of "false" for negative outcome of a patient's treatment.

3.1.2. Breast cancer Kaggle dataset

Breast cancer is a prevalent type of cancer that affects women globally, accounting for a quarter of all cancer cases and impacting over 2.1 million people in 2015 alone. The condition begins when breast cells start to grow uncontrollably, forming tumors that can be detected through an X-ray or by feeling lumps in the breast area. Detecting the type of tumor, whether it is malignant (cancerous) or benign (non-cancerous), presents a significant challenge. The task at hand is to use machine learning, to classify these tumors accurately using the Breast Cancer Wisconsin (Diagnostic) Dataset obtained from Kaggle web platform [102]. This breast cancer databases was obtained from the University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg [103].

The objectives were to clean up and understand the dataset, develop classification models that predict tumor malignancy, fine-tune hyper-parameters, and compare the evaluation metrics of various classification algorithms. In addition to our dataset from considered case study, we have trained the proposed model and evaluated the obtained results and on this dataset having 30 parameters i.e., predictors and diagnosis parameter as outcome which is binary dependent variable.

 doi:10.20944/preprints202305.0817.v1

3.2. Methods

As we mentioned in introduction of this third chapter the application of ensemble machine learning methods in the prediction different functions which solves different problems in different fields of human life, including their technical implementation in the form of useful software application, is an current trend, although is the use of logistic regression in data prediction and classification and consequently for the considered problem in this paper, still the dominant methodology. Namely when using logistic regression in prediction and/or classification, the problem of poor fit of the model and data can often occur, which is usually determined by the Hosmer Lemeshow test, so if its value is less than 0.05, arises the question of the quality of the prediction. It is important, in that case, can a quality prediction be made using and with the help of some other and other methodologies? We can found in literature [64,74,104] that to solve such problem and to improve the accuracy of a regression model next methodologies are useful:

Handling Null/Missing Values;
Data Visualization;
Feature Selection and Scaling;
Use of Ensemble and Boosting Algorithms;
Hyper-parameter Tuning.

Also the authors of articles [74,75] Hosmer and other and [105] Harrell remarked that the Hosmer-Lemeshow test is to obsolete because it requires arbitrary binning of predicted probabilities, has not detection of lack of calibration and does not fully penalize for extreme overfitting of the model. They claimed that the better methods are available such as methods proposed in [74]. More importantly, this kind of assessment just addresses overall model calibration ie. agreement between predicted and observed and does not address lack of fit such as improperly transforming a predictor. For that matter, AUC could be used to compare two models to find one which is more flexible than the other being tested. Practically in this way the stated problem is translated in problem of predictive discrimination, for which AUC and ROC measure, could be more appropriate.

The very selection of the algorithm for the stacking model of machine learning is, generally speaking, conditioned by the following factors:

The type of problem we are solving;
The characteristics of the set of attributes (features);
The volume of data available to us.

In our case study the prediction i.e., classification problem is subject of consideration, on dataset which has 26581 instances and with majority of included factors which are categorical variables.

Because of mentioned facts the authors of this paper chosen stacking ensemble as machine learning methodology to solve considered problem given in the presented case study and for this task they decided to use logistic regression from one side, classification algorithms and that naive Bayes, decision tree and logit boost (boosting) from other side and feature selection algorithms—gain ratio, chi-square and relief classifier as combiner algorithm i.e., final estimator. The proposed algorithm belongs to genetic algorithms' family [106]. This simply means that the algorithm performs a common task by dispatching its input to a particular method that is selected on the basis of the class of the input of the generic function. In general, a generic family of algorithms [107] practically allows reuse of a wide range of different problems with relatively minor reorganization. In other words, the generic modeling is a development of the concept of a model library.

3.2.1. Ensemble prediction methods

Ensemble methods which are used in machine learning [108] are based on the idea that a combination of algorithms of different types can achieve better results than each of included algorithms individually. The simplest form of this type of prediction in the form of decision has ensemble with the application of an odd number of independent models that compare their results and finally determine the solution by a simple majority. Of course that this kind of prediction evolved using different more and more complex ways of aggregating in ensemble whereby some of them use

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2023 doi:10.20944/preprints202305.0817.v1

such obtained model for different, including stochastic based selections of subsets of the considered set of data. Consequently mentioned, it is possibly to find more types of ensemble methods in different kinds of taxonomies of which most familiar are:

1. Bootstrap aggregating (bagging);
2. Boosting;
3. Stacking.

Stacking

Stacking ensemble algorithm involves training a model to combine the predictions using several included machine learning algorithms. Thereby, all of the included algorithms are trained using the available data, then a algorithm which is some their combination is trained to make a final estimation and prediction including all the predictions of these algorithms as the basic estimators and as additional inputs or using cross-validated prediction from these base estimators to prevent overfitting [109].The logistic regression model is often used as the combiner algorithm in practice.

In this way stacking ensemble algorithm typically yields performance better than any single included one of the trained algorithms [110]. It can be successfully used on both supervised [111] (what is the case in this article) and unsupervised [112] learning tasks.

3.2.2. Ensemble prediction method of selected factor effect on inpatient treatment quality

Having in mind explanation and discussion on the beginning of this subsection 3.2 Methods about the possible ways for solving an important problem of bad fitting one model with its data, the authors used machine learning stacking ensemble methodology in proposed model. The authors have decided that in proposed model which is described in this subsection of this paper use model of stacking (sometimes called stacked generalization) which in general observed involves models training to be able to combine the predictions of several other learning algorithms using some combiner algorithm. In this paper proposed stacking ensemble method includes two types machine learning algorithms: one classification algorithm and logistical regression algorithm at the beginning. Finally follows fine calibration of significance of each included factor in prediction i.e., classification process using several algorithms of feature selection for basic ranked classification what enables an optimization whole procedure and which technique is called combiner algorithm.

Authors began this procedure with successive application of the logistic regression and classification on the starting set of data with to determine theirs suitability for application and for regression. They are controlled in the regression using overall percentage in classification table (OPCT) and Hosmer and Lemeshow test of goodness of fit the model with used data using its indicator of significance HLsig) with set condition is ((OPCT)>0.5 and (HLSig>0.05)). The set condition for classification is AUC ((AUC>0.6) what does it mean minimum satisfaction of classifications performances and evaluate accuracy basic prediction formula with defined number of significant factors. After that authors proposed an enhancement of the regression model by including one stacking ensemble machine learning in the procedure, where as the second member of the ensemble, the best of some three classification algorithms suitable for the considered problem is included, and on the end as final combiner algorithm is included the one of some three selected filter algorithms of feature selection that gives the best AUC value. This proposal is according mentioned new conclusions in literature [105] mentioned in the introduction of subchapter 3.2 Methods that the AUC is the measure which is preferred over accuracy as it is a much better indicator of model performance. On the end of this procedure authors include the logistic regression and classification for fine calibration and according on the value of classification accuracy measure AUC as well as the parameter OPCT and HLSig for regression, determine potential smaller number of significant factors than initial and with a better characteristics especially of most important AUC measure. Those factors should be included in the prediction formula. In this way the authors have constructed one optimized generic procedure that is given with algorithm presented with Figure 1, described in Algorithm 1 and showed in the Figure 2.

Algorithm 1: Determining the predictors of social and health factors for inpatient treatment

- Step1. Input data in the form of table with n independent social and health factors and one which is dependent variable and which represents outcome of cardiovascular patients and can be true in the case of successful treatment and false in opposite case. Clean and normalize data.
- Step 2. Perform logistic regression Enter method is used to create a model with n predictors and the dependent variable, which is the treatment outcome.

  In this method, all predictors are included in the model unless there is a problem of collinearity, in which case some predictors may be excluded and in this case we will have l<=n factors. The classification table is used to calculate the OPCT, which should ideally be greater than 0.5 and the Hosmer and Lemeshow tests are used to assess the goodness-of-fit of the model, with a condition that the HLSig indicator should be greater than 0.05, indicating a good calibration of the model to the given data. If both of those two conditions are not satisfied procedure foresees a preprocessing to try to fix that deficiency with some of the given procedures:

  Identifying and handling the missing values.

  Encoding the categorical data.

  Splitting the dataset.

  If the preprocessing is unsuccessful output from procedure is without valid prediction. After that, in this step of algorithm we also realize classification with three selected algorithms of different type, we choose the best and evaluate is its AUC measure greater or equal 0,6. In the case that the condition for AUC is not fulfilled the output from the model is with determined l factors because we already determined in the previous IF block that in this path of the proposed algorithm both regression measures meet the set conditions i.e., HLSig indicator is greater than 0.05 and OPCT is greater than 0.5. This evaluation of fulfillment of these conditions simultaneously in the way which is explained and showed on Figure 1 is crucial for assessing the performance of the proposed ensemble model in subsequent steps because the next steps depend from the fact if they satisfied or not required values for crossing certain thresholds as it is given on Figure 1. Algorithm leads in next step 3 with value of AUC greater or equal 0,6, what means that it is possible to made good prediction for given data set.
- In step 3 of this algorithm used three selected algorithm different type from group of feature selection filter methods with the basic aim to use only the necessary k<=l factors in this classification ensemble algorithm to achieve its optimal features. It is made so that it is with the best of 3 classification algorithms from step 2 is determined value for each of three selected filter algorithms by excluding one factor at a time, starting from the lowest in rank. Using between selected algorithms of feature selection one which gives the best AUC1 value we determine these selected k factors.
- In Step 4 of this algorithm which represents a definitive decision block, we repeat classification with selected best algorithm from the step 2 of this algorithm and evaluate is the new AUC1 measure is equal or greater than AUC. In the case that the condition is not met prediction formula includes all l determined significant factors calculated by logistic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2023 doi:10.20944/preprints202305.0817.v1

regression in step 2. In the opposite case if repeated logistic regression fulfilled the already known conditions OPCT>0,5 and HLSig>0,05 the output is prediction formula with k factors.

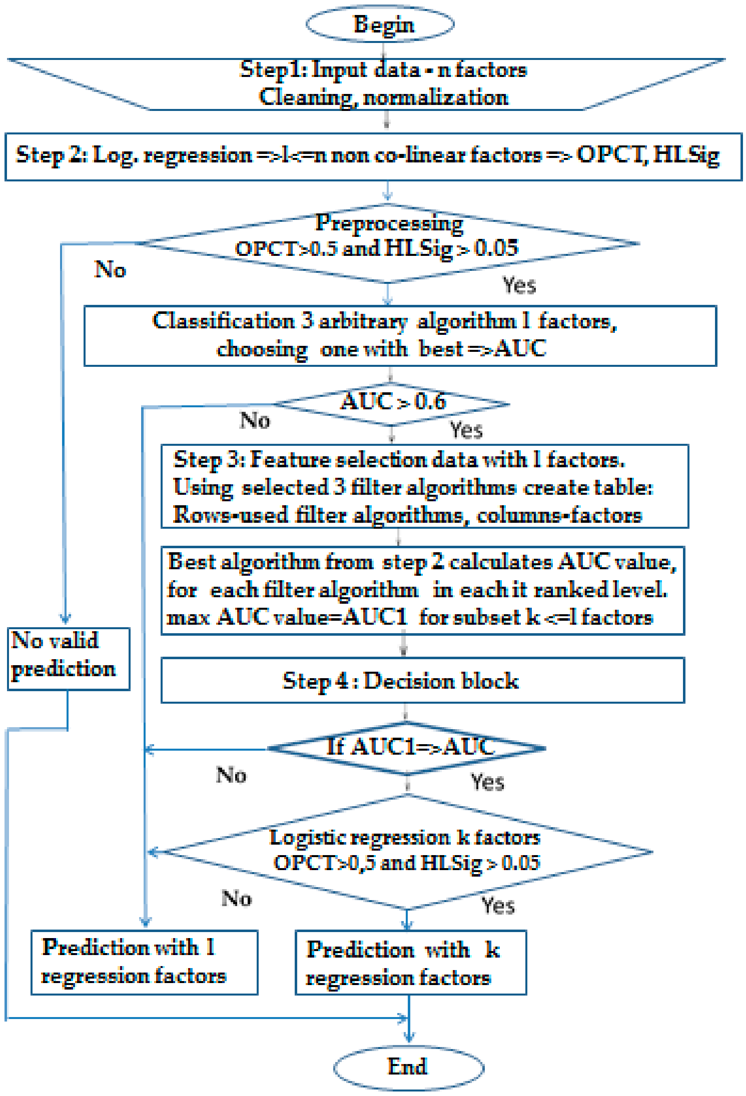

Figure 1. The flowchart of Algorithm 1.

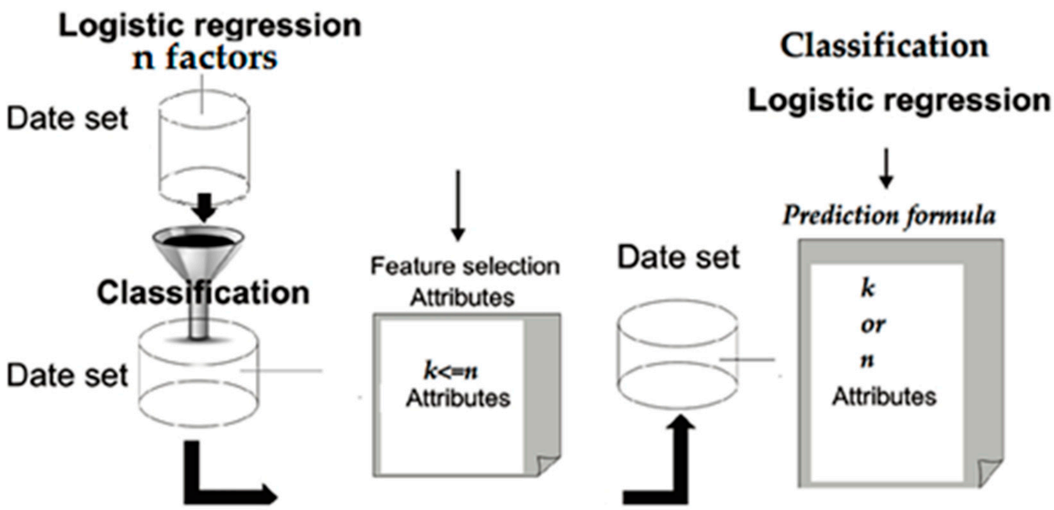

Figure 2. Block schema for the procedure which is described with Algorithm 1.

4. Results and Discussion

In order to assess the impact of selected social and health factors affecting the successful inpatient care, the authors have considered following indicators: education and place of housing of patients, level of the medical institution (clinical centers of higher levels and other medical institutions of lower level), gender and the age of patients, days of patients treatment with cardiovascular disease with positive outcome. The case study is based on the data acquired during the period from 2006 to 2009, from the Institute of Public Health in Nis, Republic of Serbia and dispensary medical institutions which are in health jurisdiction connected for city of Nis, Republic of Serbia, such as Clinical Center—Nis, Institute for Prevention and Rehabilitation—Niska Banja, Military Hospital Nis, Special Hospital—Soko Banja and districts (Ozren and Toplica): Medical Center of Prokuplje and the Health Center—Kursumlija. The data were divided in those for training which are from 2006–2007. with 11833 instances and other for testing which are from 2006–2009. with 26581 instances.

Data analysis was performed using two methodologies organized in one ensemble machine learning model as it is already described in subsection 3.2.2. This section will be divided into subsections to enable clear the steps of proposed procedure and better understanding of obtained results. It will provide a concise and precise description of the experimental results, their interpretation as well as the discussion of experimental results that can be drawn through.

4.1. Step1: Input data for considered case study

Input data in the form of excel table of xlsx and csv type, in the form with n factors which is already described in sub-chapter 3.1 Materials of this paper, cleaning and normalization of this data.

4.2. Step 2: Using logistic regression analysis and classification algorithms

4.2.1. Using logistic regression

Odds Ratio (OR) values and their 95% Confidence Interval (CI) for assessing the impact of the examined factors on the positive outcome of treatment of cardiovascular diseases in inpatient health care institutions Nis and Toplica region during the period from 2006 to 2007, and the results of multivariate logistic regression analysis are listed in the Table 3.

Table 3. OR values and their 95% CI for assessing the impact of the examined factors.

Variables in the Equation									
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	HospitalType	2,030	,596	11,610	1	,001	7,612	2,368	24,464
	UrbanHousing	-1,214	,055	492,897	1	,000	,297	,267	,331
	Education	-,093	,112	,690	1	,406	,911	,731	1,135
	Gender	,102	,054	3,540	1	,060	1,107	,996	1,231
	Age	,289	,071	16,524	1	,000	1,335	1,162	1,535
	DaysOfTreatment	,823	,107	58,707	1	,000	2,277	1,845	2,811
	Constant	1,796	,077	548,193	1	,000	6,027		

a. Variable(s) entered on step 1: HospitalType, UrbanHousing, Education, Gender, Age, DaysOfTreatment.

Notations in the Table 3 and in other tables of this subchapter produced in SPSS 19 are translated as stated below in the process of normalization:

- *UrbanHousing(1)* = Urban place of housing of patient,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2023 doi:10.20944/preprints202305.0817.v1

- *Education (1)*= high level of education of patients
- *HospitalType*(1) = Treatment at the Clinical Center—Nis,
- *Gender*(1) = Female gender (1),
- *Age* = Patient age (years>50,old patients(1)),
- *DaysofTreatment* = Length of hospital stay (days>15(1)),
- *Outcome* (1/0) = Positive/negative outcome of treatment.

Also, the meaning of the abbreviations in the columns is:

B—Denotes the unstandardized regression weight. It can be simplified in its interpretation, so as each variable increases, the likelihood of scoring a "1" on the dependent variable also increases.

S.E.—Measures how much the unstandardized regression weight can vary by. It is similar to a standard deviation to a mean.

Wald—Denotes the test statistic for the individual predictor variable like multiple linear regression has a t test, logistic regression has a $\chi 2$ test and it is used to determine the Sig. value.

df—This is the number of degrees of freedom for the model. There is one degree of freedom for each predictor in the model.

Sig.—Determines which variables are significant and p value below .050 determines significant.

Exp(B) or OR—Denotes the odds ratio that represents the measurement of likelihood and denotes that for every one unit increase in Variable 1, the odds of a participant having a "1" in the dependent variable increases by a factor of 4.31.

95% C.I. OR—this is the 95% confidence interval (C.I.) for the odds ratio what means that with these values, we are 95% certain that the true value of the odds ratio is between those units. But, if the C.I. does not contain a 1 in it, the Sig. value will end up being less than .050.

Multivariate logistic regression analysis was used to examine the correlation between a positive therapeutic outcome as a dependent variable, on one hand and the age of patients, gender, the number of hospitalization days, level of education of patients, place of housing of patients and type of dispensary health institutions as independent variables, on the other hand. Calculated OR values and the limits of their 95% CI show that the ratio of the probability will perform recovery or improvement in health status and probability that the health condition is unlikely to remain or get worsen. Patient age, number of days of hospitalization, gender of patients, education level, place of housing and types of healthcare institutions were used as categorical variables.

Logistic regression analysis confirmed that the decreasing probability of occurrence of a positive treatment outcomes is associated education level of patient (OR = 0.406 95% CI: 0.731 to 1.135; p = 0.406), as well as housing place (OR = 0.297, 95% CI: 0.267 to 0.331; p < 0.001) and the probability of the occurrence of a positive outcome was increased for female gender (OR = 1.107, 95% CI: 0.996 to 1.231, p =0.060), with patient age (OR = 1.335, 95% CI: 1.162 to 1.535, p <0.001), together with the increased length of hospitalization (OR = 2.277, 95% CI: 1.845 to 2.811, p <0.001) as well as treatment at the Clinical Center—Nis (OR = 7.612, 95% CI: 2.368 to 24.464, p=0.001). Those conclusions only confirm a logical and experiential already known expectation For a comprehensive analysis of the results obtained using logistic regression, further discussion is necessary and information about results of necessary tests for this task are provided in the tables given bellow.

Table 4 provides a summary of the proposed model good performances.

Table 4. Method Enter—beginning regression analysis using all 6 factors.

Hosmer and Lemeshow Test			
	Chi-square	df	Sig.
Step 1	11,001	6	,088

Classification Table[a]					
			Predicted		
			Outcome		Percentage Correct
	Observed		0	1	
Step 1	Outcome	0	0	1735	,0
		1	0	10098	100,0

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2023 doi:10.20944/preprints202305.0817.v1

Overall percentage	85,3
[a] The cut value is ,500	

In the Table 4 part Hosmer and Lemeshow test shows that the value of HLSig=0,088 is greater than requested 0,05, classification table provides exactly that the model predicts on dependent categorical variable for each test case OPCT=85,3% and both of them are acceptable values as required in the procedure. Positive predictive value indicates that the model is identified as successfully treated 85,3% and negative predictor value is 0%, and it shows that the percentage of modelled cases is classified as lacking hallmark—since it is not observed in the group.

The accuracy of the classification by random selection is $(1735/11833)^2 + (10098/11833)^2 = 0.7945$, what is 79.45%, so it can be seen that the model of binary logistic regression analysis with 85.3% has a higher classification accuracy than random selection models. The table of variables in the equation given as part of Table 3 provides information about the importance of each predictor in the Wald column and it cannot be concluded that all predictors, especially predictor the level of education of the patients and predictor Gender of patient have affect on the dependent variable in contrast to all other that have affect. The values given in column B of this table suggest the direction of the relationships from the dependent variables to the independent variable. For example, several of them—the level of expertise of the hospital, days of treatment, age and belonging to the female gender improve the success of treatment carried out by health institutions, and other predictors reduce the success of treatment.

Logistic regression confirm all l=n=6 factors as valid and significant for prediction in this first step.

4.2.2. Using classification algorithms

According to the proposed algorithm, taking into account that in this step logistic regression confirmed the validity of the influence of all l=n=6 considered predictors on the dependent variable outcome, we evaluate the quality of the influence one of the classification algorithm using AUC-ROC measure and we do it with three algorithms of different type—J48 decision tree, NaiveBayes and LoogitBoost. Obtained results are with using 10 cross validation and determine that we can use LoogitBoost classification algorithm as the best, with the highest value of AUC- ROC measure.

Table 5. AUC performance indicators obtained by the classification algorithms using all 6 factors.

	AUC-ROC
Naive Bayes	0,669
Logit Boost	0,671
J48 Decision tree	0,499

4.2.3. Check fulfillment of set conditions

Since the set condition OPCT>0,5 and HLSig>0,05 is fulfilled, we continue, otherwise the algorithm would lead to an output without a possible valid prediction. In the case that following condition—AUC>0,6 is not fulfilled we continue to an output with l=6 prediction factors determined in step 2 because the condition HLSig>0,05 is already determined as fulfilled, otherwise algorithm continues with the next step 3 which leads to an output with k factors where it is possibly to be k<=l i.e., k<=6.

4.3. Step 3: Using feature selection

In this step of procedure the authors have used a selection of relevant attributes, using so called feature selection technique, to reduce the dimensionality of the original space up to the space with lower dimensionality, where determination of the individual factors importance and correlation between the attribute values can be easily identified. They have proposed a filter-ranker evaluation approach for detecting factors and used three randomly selected algoritms different type—

GainRatio(GR), ChiSquaredAttributeEval(CHI) and Relief(REL) instead one. Obtained results in the Table 7 are shown differet ranking but it is easy to conclude that factors Education as Gender are with the smallest significance as it is obtained in regression algorithm.

Table 7. Factors ranking by the feature selection measures—6 factors.

Attribute tag	Attribute Name	GR-Ranking GainRatio	CHI-Ranking ChiSquaredAttributeEval	REL-Ranking ReliefF
1	HospitalType	2	4	1
2	UrbanHousing	1	1	5
3	Education	6	6	3
4	Gender	5	5	6
5	Age	4	3	4
6	DaysofTreatment	3	2	2

Follows the calculation with determined classification algorithm in step 2 which has higher AUC value—LoogitBoost for each filter ranker algorithm and by eliminating one factor at a time, starting from the last ranked one. Using such procedure it is obtained that the maximal value of AUC1 = AUC = 0,599 for LoogitBoost has been achieved with GainRatioalgorithm using only first three ranked attributtes—UrbanHousing, HospitalType, DaysofTreatment as it is given in the Table 8.

Agrafic presentation of this procedure is given in Figure 3. From the Table 8 and the diagram from the Figure 3 it is clear that the optimization usingthe proposed procedure has determined the number of k=5 factors with the Education factor which was excluded.

Table 8. Factors ranking by the feature selection measures.

Algorithm /FactorsNumber	6	5	4	3	2	1
GR	0.671	0.670	0.661	0.651	0.634	0.631
CHI	0.671	0.671	0.661	0.658	0.649	0.631
REL	0.671	0.663	0.560	0.530	0.538	0.501

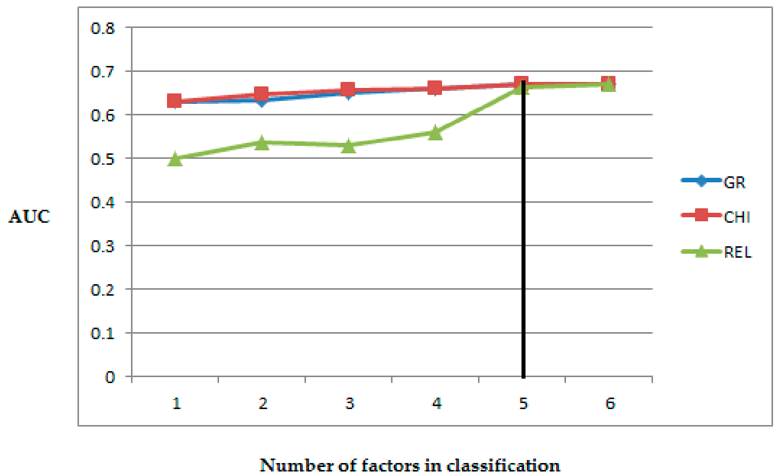

Figure 3. Grafic of determining the highest value of AUC for lowest number of factors.

4.4. Step 4: Decision blok

In this last step using determined k=5 significant factors—DaysofTreatment, UrbanHousing, HospitalType, Age and Gender we first check validity of classification (Table 9) and fulfillment of codition: AUC1=> AUC and if it is not fulfilled output is with l=6 factors in prediction formula determined in step 2. But this condition is fulfilled as it is given in the Table 9 so we continue and check the validity of the logistic regression (Table 10). The authors concluded that if both conditions—OPCT>0.5 and HLSig>0.05 are fullfilled as well as obtained results of Omnibus Tests of Model Coefficients, Hosmer and Lemeshow Test and Classification Table for logistic regression are valid. Because of that facts practically appears as the output of the proposed procedure that fine calibration of classification discrimination using the regression formula will be with k=5 mentioned factors; otherwise output will be with l=6 factors in prediction formula which are determined in the step 2.

Table 9. Performance indicators obtained by the Logit Boost classification algorithm using 5 factors.

	AUC-ROC
6 factors	0,671
5 factors	0,671

Table 10. Method Enter—regression analysis using selected 5 factors.

Hosmer and Lemeshow Test			
	Chi-square	df	Sig.
Step 1	9,606	4	,058

Classification Table[a]					
			Predicted		
			Outcome		Percentage Correct
	Observed		0	1	
Step 1	Outcome	0	0	1735	,0
		1	0	10098	100,0
	Overall percentage				85,3

[a] The cut value is ,500

Variables in the Equation									
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	HospitalType	2,039	,596	11,727	1	,001	7,687	2,392	24,698
	UrbanHousing	-1,209	,054	494,128	1	,000	,298	,268	,332
	Gender	,108	,054	4,012	1	,045	1,114	1,002	1,237
	Age	,294	,071	17,189	1	,000	1,342	1,168	1,542
	DaysOfTreatment	,824	,107	58,885	1	,000	2,280	1,847	2,814
	Constant	1,782	,075	570,063	1	,000	5,940		

[a] Variable(s) entered on step 1: HospitalType, UrbanHousing, Gender, Age, DaysOfTreatment.

We have on exit of application of proposed procedure prediction formula given with Table 10.

4.5. Discussion

As it is represented in this chapter of the paper on the example considered as a case study, for evaluation of the proposed stacking ensemble machine learning model, authors used:

- filter feature selection algorithms different type (GainRatio, ChiSquaredAttributeEval and Relief for dimension reduction of the problem) and combiner algorithm in enseble and
- logistic regression with LogitBoost as the best of selected classification algorithms differen type—LogitBoost, J48 decision tree and NaiveBayes for evaluation i.e., fine calibration in the proposed model.

The obtained results showed that the proposed algorithm provides optimization of the procedure for determining the importance of certain selected social and health factors on the success of hospital treatment through dimensionality reduction with fine calibration by logistic regression and classification algorithm. The results also showed that described procedure leads to a unique prediction formula with good classification characteristics, which qualifies the proposed ensemble method as improved compared to each of the included, aggregated methods individually i.e., logistic regeression and classification algorithms individually applied.

The initial basic hypothesis comprehended in this paper in its introductory part was that it is possible to test the goodness one plan of organization of the health institutions with testing successful treatment depending per their level of professional expertise of health care institutions than among other factors like per days of hospitalization, per education, place of housing, age and gender of the patients. The results of applied analyses show that the success of the treatment of cardiovascular patients dominantly depends on the place of housing than and consequntly connected the type of othe hospital than days of hospitalization, age and gender of patients and don't depends from level of education of patints. However, the influence of these five factors also depends so that not urban place of residence of patients is with a negative sign what decreasing succefullnes of inpatient treatment and the all others factors are with a positive sign what mean that hospital of higher level of expertise and biger number of days of hospitalization bay yanger, female patients increasing succefullnes of treatment.

It is very important to note that the authors did not notice any limitation for the use of the proposed method, because the required time for its execution, which is evidently longer than the time required if only one of the algorithms aggregated in the proposed ensemble was used, is not a limitation for working in real-time.

Also it is important to notice that the resalts are evaluted inside the proposed method using 10 cross validation procedure and outside site with using Brest cancer Kaggle dataset. Using this dataset authors checked possible generalization of proposed model as one generic procedure which can be used on each two class classification problem nevertless of their nature and the fields where such generally considered one multivariate or multicriteria problem belongs [113].

Having in mind the above discussion on the evaluation of the proposed model based on the analysis of the results obtained in the presented case study, the authors aim to expand their research by the inclusion of biger number of classification as wellas filter ranking algorithms as a part of the proposed ensemble model that would enable the achievement of additional improvements in its characteristics as wel as including a studies of the possibilities and other more modern methods for assessing the goodness of fit between models and data.

From other site from planed research in technique and methodology for solving considered healthcare problem and problems in healthcare which already was subject of research of authors [114,115] authors will deal in feature with other from many different problems in healthcare and other fields of human life as for example traffic and education which the authors have already done [116,117].

5. Technical solution—code implementation and real-life software platform usage

Model deployment involves taking a model and integrating it into a software application that can be used in real-world scenarios. The purpose of model deployment is to provide a user-friendly interface to interact with the model, allowing users to input new data and get predictions based on the model's output.

Here are the steps involved in deploying a machine learning model:

1. Export the model: Export the trained machine learning model into a file format that can be used by other software applications. This could be a serialized object or a machine learning library-specific format.
2. Set up a server: Create a server to host the model and handle incoming requests from users. This server could be a cloud-based service like Amazon Web Services (AWS) or Microsoft Azure, or it could be set up on a local machine using a software like Flask or Django.
3. Create an API: Create an application programming interface (API) that will handle requests from clients and return responses with predictions from the model. This API can be created using a web framework like Flask or Django, and it will typically use HTTP requests to send and receive data.

 doi:10.20944/preprints202305.0817.v1

4. Create a client application: Create a client application that can be used to interface with the API. This client application can be a web application or a mobile application, and it will typically use HTTP requests to send data to the API and receive predictions from the model.
5. Test and deploy: Test the deployed model using sample data to ensure that it is working as expected. Once testing is complete, deploy the model to a production environment where it can be accessed by users.
6. Monitor and update: Monitor the deployed model to ensure that it is performing as expected and update it as needed with new data or changes to the model itself.

Overall, model deployment involves creating a server that can host the trained machine learning model, setting up an API to handle requests from clients, and creating a client application that can interface with the API. The flow of the data in implemented solution is shown in Figure 4.

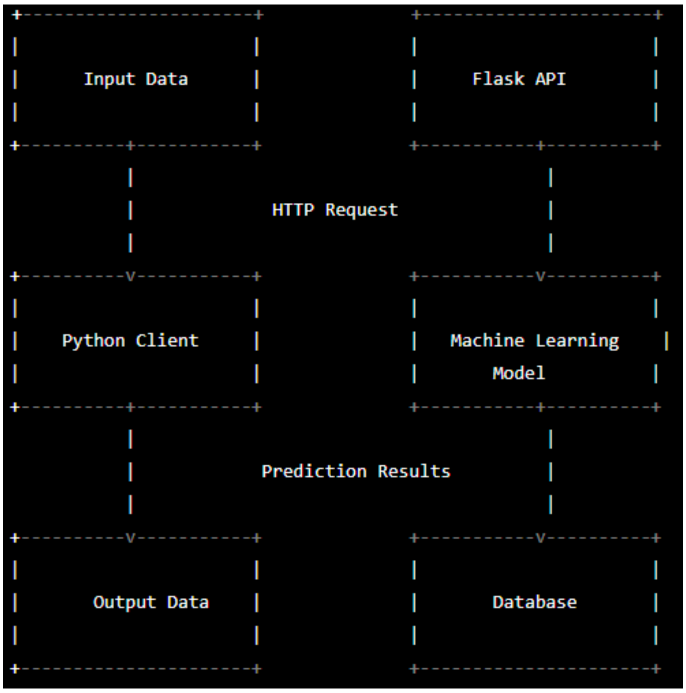

Figure 4. The Data flow in the implemented solution.

As can be seen, the flow of data in this solution starts with the input data, which is collected and sent to the Flask API via the client app. The Flask API receives the input data, passes it to the machine learning model, and returns the predictions to the client app. The client app then stores the input data and predictions in a database for future analysis and reference. This process can be repeated for new input data, allowing the machine learning model to continually improve its predictions over time. A block diagram of the proposed solution is shown in Figure 5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2023 doi:10.20944/preprints202305.0817.v1

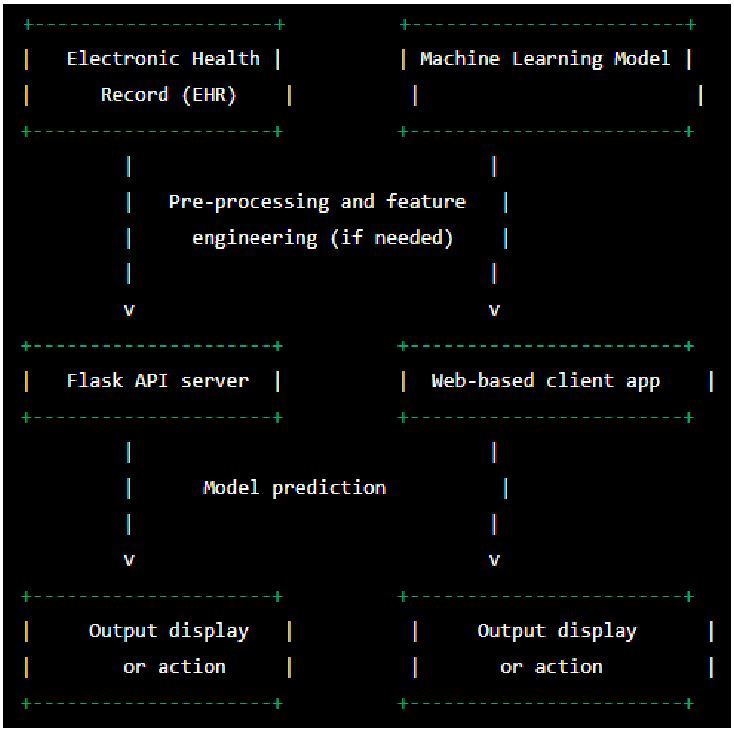

Figure 5. Block diagram of the proposed solution.

As shown in Figure 4, the proposed solution consists of four main components:

1. Electronic Health Record (EHR): This is the source of data for the machine learning model. It could be a database or other storage mechanism that contains information about patients and their treatments.
2. Machine Learning Model: This is the core of the solution, which determines the importance of social and health factors affecting the successful inpatient treatment. The model could be developed using various machine learning algorithms and techniques, depending on the specifics of the problem.
3. Flask API Server: This component serves as the interface between the machine learning model and the client application. It provides a RESTful API that receives input data, performs model prediction, and returns output data.
4. Web-based Client Application: This component provides a user interface for interacting with the machine learning model. It could be a web application that allows users to input data, view model predictions, and take actions based on the predictions.

The input data could be pre-processed and feature-engineered (if needed) before being sent to the Flask API server for model prediction. The output of the model prediction could be displayed or used to take further action, depending on the requirements of the application. Here are the steps involved in deploying a proposed model as a technical solution in order to try and test it:

1. Export the model: We exported the trained machine learning model into a file format that other software applications can use.
2. Set up a server: We developed a server app to host the model and handle incoming requests from users.
3. Create an API: We implemented an application programming interface (API) that will handle requests from clients and return responses with predictions from the model. This API can be created using a web framework like Flask or Django, and it will typically use HTTP requests to send and receive data.
4. Create a client application: We created a client application that can be used to interface with the API. This client application is a web application, and it uses HTTP requests to send data to the API and receive predictions from the model.

5. Test and deploy: The deployed model has been tested using sample data to ensure that it is working as expected. After we finished testing, the model was deployed to a production environment where real-life users could access it.
6. Monitor and update: The monitoring of the deployed model to ensure that it is performing as expected and update it as needed with new data or changes to the model itself has also been supported in this technical solution.

Overall, model deployment has involved creating a server that can host the trained machine learning model, setting up an API to handle requests from clients, and creating a client application that can interface with the API. The implemented solution is accessible to users and used to make predictions based on new data. The source code of the implemented solution is given as supplementary material for this work. The solution is robust, easily expandable and adaptable to any context. In other words, one can use our code and straightforwardly adapt new model and new client scenarios as well as use it as real-world client-server software platform.

6. Conclusions

The ensemble model suggested in this paper is one optimization procedure based on the stacking model in ensemble learning consists from logistic regression and classification technique for the fine evaluation of executed optimization made through dimension reduction of considered problem of estimation the importance of several social and health factors on the successful inpatient treatment. The obtained results on the considered case study inpatient treatments in hospitals of region of city Nis, Republic of Serbia showed that the proposed algorithm provides optimization of the procedure for determining the importance of certain social and health factors on the success of hospital treatment through dimensionality reduction. The results also showed that, in the case of the application of different algorithms for attribute selection, leads to a unique prediction formula with good classification characteristics, which qualifies the proposed ensemble method as better in comparison with each of the methods aggregated in it individually. Thereby, using the proposed procedure authors also gave a positive answer to the main hypothetical question of this paper that it is possible to determine the goodness of health care organization in the observed region i.e., territory using one optimized machine learning approach for determining the importance of social and health factors which affect successful inpatient care. For future work initiated by the results of this article, the authors suggest considering the inclusion and selection of other classification and feature selection algorithms as well as the most recently developed measures for assessing the goodness of fit of the model to its data.

In conclusion, we notice the following important facts:

- Firstly, our and the motivation of others interested for solving considered and similar problems of improving existing methodologies for determining significant factors in a multivariate problems and obtaining more efficious procedures is based on the fact that today, in information society, exist many readily available data in which lies knowledge. On the other hand, with the rapid development of artificial intelligence, an increising number of approaches of different types are available. From the literature cited in the paper, it is known that logistic regression and classification algorithms are the most useful and most frequently used for this type of problem, and that is why they are an integral part of the procedure we proposed in the paper.

Second, it is also known from literature that in machine learning, ensemble methods use of multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone although the evaluating of prediction of an ensemble typically requires more computation than evaluating the prediction of a single model. We could say the ensemble learning may be thought as a way to compensate for poor learning algorithms by performing a lot of extra computation. An ensemble system may be more efficient at improving overall accuracy for the same increase in different compute resources by using that increase on two or more methods, than would have been improved by increasing resource use for single method.Between these enseble methods authors uses in proposed procedure stacking ensemble machine learning approach.The two main benefits of stacking ensemble which determined that

 doi:10.20944/preprints202305.0817.v1

decision were that it can shield the capabilities of a range of well-performing models to solve classification and regression problems and further, it helps to prepare a better model having better predictions than all individual models.

- Thirdly, the conducted research gave a doubly significant result as the final result:
  - From a scientific point of view, the authors proposed a new generic optimization procedure that can be used to solve both classic prediction problems and discriminative classification, both of which basically determine the importance of individual factors in a multivariate problem.
  - From a professional point of view, the authors have developed and made available to the public for use and further development a modern multi-agent application for solving a specific problem in assessing the influence of certain factors on the success of hospital treatment, but it is also usable as such for solving other, similar problems in healthcare but also in other fields of human activity.

Supplementary Materials: The following are available online: Dataset for considered case study https://drive.google.com/file/d/1Mto5zBRFJvgA22g-MUWoyi3JnpL2A4sJ/view and Application at https://drive.google.com/file/d/1Mto5zBRFJvgA22g-MUWoyi3JnpL2A4sJ/view.

Author Contributions: author (1): Methodology, Software, Writing—review & editing; author (2) Project administration, Resources, Validation, Formal analysis; author (3) Conceptualization, Formal analysis, Investigation, Writing—review & editing; author (4) Methodology, Software, Validation, Writing—original draft, Supervision.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank the Science and Technology Park in Nis, Republic of Serbia for their help so that this paper could be done and published**.**

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Assembly Resolution WHA51.7. 1998 Health for all policy for the twenty-first century Geneva: World Health Organization (Accessed 12 February 2023) http://legacy.library.ucsf.edu/documentStore/g/w/o/gwo93a99/Sgwo93a99.pdf
2. Health21: an introduction to the health for all policy framework for the WHO European Region. 1998 (European Health for All series; no. 5.) Copenhagen: World Health Organization, Regional Office for Europe (Accessed 12 February 2023) http://www.euro.who.int/__data/assets/pdf_file/0004/109759/EHFA5-E.pdf
3. Health21: the health for all policy framework for the WHO European Region 1999 (European Health for All series; no. 6.) Copenhagen: World Health Organization Regional Office for Europe (Accessed 12 February 2023) http://www.euro.who.int/__data/assets/pdf_file/0010/98398/wa540ga199heeng.pdf
4. Plan zdravstvene zastite iz obaveznog zdravstvenog osiguranja u Republici Srbiji za 2012 https://www.rfzo.rs/download/plan%20zz/planZZ-2012.pdf (Accessed 12 February 2023)
5. Zakon o zdravstvenoj zastiti Republike Srbije http://www.zdravlje.gov.rs/tmpmzadmin/downloads/zakoni1/zakon_zdravstvena_zastita.pdf (Accessed 12 February 2023)
6. Uredba o Nacionalnom programu prevencije, lecenja i kontrole kardiovaskularnih bolesti u Republici Srbiji do 2020 (Accessed 12 February 2023) https://www.pravno-informacionisistem.rs/SlGlasnikPortal/eli/rep/sgrs/vlada/uredba/2010/11/5 (Accessed 12 February 2023)
7. Meijden, V.D.; Tange, M. J.; Troost, H.J.; Hasman, J.A. Determinants of success of inpatient clinical information systems: a literature review. *J. Am. Med. Inform. Assoc.* **2003,** 10(3), 235–243.

 doi:10.20944/preprints202305.0817.v1

8. Delpierre, C.; Cuzin, L.; Fillaux, J.; Alvarez, M.; Massip, P.; Lang, T. A systematic review of computer-based patient record systems and quality of care: more randomized clinical trials or a broader approach?. *Int. J. Qual. Health Care* **2004**, 16(5), 407-416.
9. Valaitis, R., Meagher-Stewart, D., Martin-Misener, R. et al. Organizational factors influencing successful primary care and public health collaboration. *BMC Health Serv Res.* **2018**, 18, 420. https://doi.org/10.1186/s12913-018-3194-7
10. Mosadeghrad AM. Factors influencing healthcare service quality. *Int J Health Policy Manag.* **2014**, 26;3(2):77-89. doi: 10.15171/ijhpm.2014.65. PMID: 25114946; PMCID: PMC4122083.
11. Truglio-Londrigan et al. A qualitative systematic review of internal and external influences on shared decisionmaking in all health care settings. *JBI Database of Systematic Reviews & Implementation Reports*. **2014**, 12(5) 121 - 194 doi:10.11124/jbisrir-2014-1414
12. Itulua-Abumere, Flourish. (2013). Social factors that determine health. Education > College & Uni versity(Articlesbase.com). (Accessed 20 February 2023)
https://www.researchgate.net/publication/259759029_Social_factors_that_determine_health
13. The impact of political, economic, socio-cultural, environmental and other external influences https://www.healthknowledge.org.uk/public-health-textbook/organisation-management/5b-understanding-ofs/assessing-impact-external-influences (Accessed 20 February 2023)
14. Statistical Methods in Healthcare. Frederick W. Faltin (Editor), Ron S. Kenett (Editor), FabrizioRuggeri (Editor), Wiley, 2012. ISBN: 978-0-470-67015-
15. 6Häyrinen, K.; Saranto, K.; Nykänen, P. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. *Int. J. Med. Inform.* **2008,** 7(5), 291-304.
16. El-Sappagh, S.H.; El-Masri, S.; Riad, A.M.; Elmogy, M. Data Mining and Knowledge Discovery: Applications, Techniques, Challenges and Process Models in Healthcare. *International Journal of Engineering Research and Applications* **2013**, 3(3), 900-906.
17. Hand, D.J. Data mining: statistics and more?. *The American Statistician* **1998**, 52(2), 112-118.
18. Koh, H.C.;Tan, G. Data Mining Applications in Healthcare. *Journal of Healthcare Information Management* **2005**, 19(2), 64-72.
19. Milley, A. Healthcare and data mining. *Health Management Technology* **2000,** 21(8), 44-47.
20. Silver, M.; Sakata, T.; Su, H.C.; Herman, C.; Dolins, S.B.; O'Shea, M.J. Case study: how to apply data mining techniques in a healthcare data warehouse. *Journal of Healthcare Information Management* **2001**,15(2),155-164/
21. Trybula, W.J. Data mining and knowledge discovery. *Annual Review of Information Science and Technology* **1997**, 32, 197-229.
22. Rokach, L. Ensemble-based classifiers. *Artificial Intelligence Review* **2010,** 33 (1-2), 1–39.
23. Opitz, D.; Maclin, R. Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research* **1999**, 11, 169–198.
24. Nguyen DK, Lan CH, Chan CL. Deep Ensemble Learning Approaches in Healthcare to Enhance the Prediction and Diagnosing Performance: The Workflows, Deployments, and Surveys on the Statistical, Image-Based, and Sequential Datasets. Int J Environ Res Public Health. 2021 Oct 14;18(20):10811. doi: 10.3390/ijerph182010811. PMID: 34682554; PMCID: PMC8536161.
25. Alekhya, B., Sasikumar, R. An ensemble approach for healthcare application and diagnosis using natural language processing. Cogn Neurodyn 16, 1203–1220 (2022). https://doi.org/10.1007/s11571-021-09758-y
26. Breiman, L.; Stacked Regression. *Machine Learning* **1996**,24
27. Ozay, M.; Vural, F.T.Y. A New Fuzzy Stacked Generalization Technique and Analysis of its Performance. *arXiv 1204.0171*, (**2012**)
28. Smyth, P., Wolpert, D. H.: Linearly Combining Density Estimators via Stacking. Machine Learning Journal **1999**,36, 59-83.
29. Wolpert, D.H.; Macready, W.G. An Efficient Method to Estimate Bagging's Generalization Error. *Machine Learning Journal* **1999**,35, 41-55.
30. Yahia, M. E.; El-Taher, M. E. A New Approach for Evaluation of Data Mining Techniques. International *Journal of Computer Science Issues* **2010**,7(5), 181-186.
31. France, G.; Taroni, F.; Donatini, A. The Italian health-care system. *Health Economics* **2005**, 14, 187–202.
32. Müller, U.; Offermann, M. Krankenhausplanung im DRG-System Expertenbefragung des Deutschen Krankenhausinstituts Düsseldorf Deutsches Krankenhausinstitut. Düsseldorf Deutschland (2004)

33. Figueras, J. Effective health care planning – the role of financial allocation mechanisms. University of London (1993)
34. Protti, D. US Regional Health Information Organizations and the Nationwide Health Information Network: Any Lessons for Canadians?. *ElectronicHealthcare* **2008**, 6(4), 96-103.
35. Country Health Care Plan Taskforce (Accessed 12 October 2014) http://www.countryhealthsa.sa.gov.au/Publications.aspx
36. Kahler, C. China's Healthcare Reform: How Far Has It Come? (Accessed 12 October 2014) http://www.chinabusinessreview.com/chinas-healthcare-reform-how-far-has-it-come/
37. Zickmund, S.L.; Burkitt, K.H.; Gao, S.; Stone, R.A.; Jones, A.L.; Hausmann, L.R.M.; Switzer, G.E.; Borrero, S.; Rodriguez, K/L.; Fine, M.J. Racial, Ethnic, and Gender Equity in Veteran Satisfaction with Health Care in the Veterans Affairs Health Care System. *J Gen Intern Med.* **2018**, 33(3),305-331. doi: 10.1007/s11606-017-4221-9.
38. Kressin, N.R.; Skinner, K.; Sullivan, L.; Miller, D.R.; Frayne, S.; Kazis, L.; Tripp, T. Patient satisfaction with Department of Veterans Affairs health care: do women differ from men? *Mil Med.* **1999** , 164(4),283-8.
39. Rollins, R.J. Patient satisfaction in VA medical centers and private sector hospitals: a comparison. *Health Care Superv.* **1994** , 12(3),44-50.
40.. McCleery, E.; Christensen, V.; Peterson, K.; Humphrey, L.; Helfand, M. Evidence Brief: The Quality of Care Provided by Advanced Practice Nurses [Internet]. Washington (DC): Department of Veterans Affairs (US); 2014.
41. Almass, A.; Aljohani, H.M.; Alhaqbani, R.M.; Alromih, A.M.; Hadal, S.; Abozaid, H.S. Patient Satisfaction With Quality of Care at the Kingdom of Saudi Arabia. *Cureus.* **2022**,1,14(12):e32102. doi: 10.7759/cureus.32102.
42. Almugti, H.S. et. al. Quality of Clinical Counseling About Lifestyle-Related Diseases: An Analysis of Daily Practice From the National Guard Primary Health Care Center at Jeddah, Saudi Arabia. *Cureus*. **2022**, 15,14(11):e31551. doi: 10.7759/cureus.31551
43. Naderinejad, M.; Tarokh, M.J.; Poorebrahimi, A. Recognition and Ranking Critical Success Factors of Business Intelligence in Hospitals - Case Study: Hasheminejad Hospital. *International Journal of Computer Science & Information Technology* **2014**, 6(2), 121-129.
44. Tomar, D.; Agarwal, S. A survey on data mining approaches for healthcare. *International Journal of Bio-Science and Bio-Technology* **2013**, 5(5), 241-266.
45. Koh, H. C.; Tan, G. Data Mining Applications in Healthcare. *International Journal of Computer Science & Information Technology* **2014**,6(2),121-129.
46. Milovic, B.; Milovic, M. Prediction and Decision Making in Health Care using Data Mining . International *Journal of Public Health Science* 2012,1(2),69-78.
47. Suter, E.; Oelke, N. D.; Adair, C.E.; Armitage, G.D. Ten Key Principles for Successful Health Systems Integration. *Healthc. Q.* **2009**, 13(Spec No), 16–23.
48. Fichman, R. G.; Kohli, R.; Krishnan, R. Editorial Overview "The Role of Information Systems in Healthcare: Current Research and Future Trends. *Information Systems Research* 2011, 22(3) 419–428.
49. Li, J.S.; Yu, H.Y.; Zhang, X.G. Data Mining in Hospital Information System. New Fundamental Technologies in Data Mining. Prof. Kimito Funatsu ed InTech (Shanghai, China) (2011) Available from: http://www.intechopen.com/books/new-fundamental-technologies-in-data-mining/data-miningin-hospital-information-system (Accessed 12 October 2014)
50. Sadoughi, F.; Kimiafar, K.; Ahmadi, M.; Shakeri, M.T. Determining of Factors Influencing the Success and Failure of Hospital Information System and Their Evaluation Methods: A Systematic Review. *Iran Red Crescent Med. J.* **2013**, 15(12) 11716.
51. Berg, M. Implementing information systems in health care organizations: Myths and challenges. *Int. J. Med. Inform.* **2001**, 64,143–56.
52. Wilhelmsen, L.; Wedel, H.; Tibblin, G. Multivariate Analysis of Risk Factors for Coronary Heart Disease. Circulation **2015**, 1973, 48, 950-958. doi: 10.1161/01.CIR.48.5.950
53. Yan, L. The Effect of Risk Factors on Coronary Heart Disease: An Age-Relevant Multivariate Meta Analysis. Electronic Theses Treatises and Dissertations. Paper 1428, (2010). http://diginole.lib.fsu.edu/etd/1428 (Accessed 12 February 2023)

54. Shouman, M.; Turner, T.; Stocker, R. Using data mining techniques in heart disease diagnosis and treatment . In Proc. Conference on Electronics, *Communications and Computers -JEC-ECC (Japan-Egypt)* **2012**, pp. 173 – 177.
55. Hachesu, P. R.; Ahmadi, M.; Alizadeh, S.; Sadoughi, F. Use of Data Mining Techniques to Determine and Predict Length of Stay of Cardiac Patients. *Healthc. Inform. Res.* **2013**, 19(2), 121–129.
56. Hu, Z.; at al.: Real-Time Web-Based Assessment of Total Population Risk of Future Emergency Department Utilization: Statewide Prospective Active Case Finding Study. *Interact. J. Med. Res.* **2015**, 4(1):e2 () doi:10.2196/ijmr.4022
57. Chen, H. at al.: Ensemble learning for prediction of the bioactivity capacity of herbal medicines from chromatographic fingerprints. *BMC Bioinformatics* **2015**, 16(Suppl 12):S4
58. Rahmani, A.M.; Yousefpoor, E.; Yousefpoor, M.S.; Mehmood, Z.; Haider, A.; Hosseinzadeh, M.; Ali Naqvi, R.Machine Learning (ML) in Medicine: Review, Applications, and Challenges. *Mathematics* **2021**, 9, 2970. https://doi.org/10.3390/math9222970
59. Panagiotis, P. and Livieris, I.E. Special Issue on Ensemble Learning and Applications. Reprinted from: *Algorithms* **2020**, 13, 140, doi:10.3390/a13060140
60. Ren, Y.; Zhang, L.; Suganthan, P.N. Ensemble Classification and Regression-Recent Developments, Applications and Future Directions. *IEEE Computational Intelligence Magazine* **2016**, vol. 11, no. 1, pp. 41-53. doi: 10.1109/MCI.2015.2471235.
61. Kludacz-Alessandri, M. Non-financial dimensions of measurement and assessment in the performance model for hospitals. *Managerial Economics* **2016**, 17. 93. doi:10.7494/manage.2016.17.1.93.
62. Tang, J.W. et al. An exploration of the political, social, economic and cultural factors affecting how different global regions initially reacted to the COVID-19 pandemic. Interface Focus, 2022, 12:20210079 http://doi.org/10.1098/rsfs.2021.0079
63. Ballou, B.; Heitger, D.; Tabor, R. Nonfinancial Performance Measures in the Healthcare Industry. *Management accounting quarterly*, **Fall 2003**,5(1).
64. Dibyahash, B. et al. Deep Learning in Healthcare System for Quality of Service, Journal of Healthcare Engineering, 2022, Article ID 8169203, 11 pages, 2022. https://doi.org/10.1155/2022/8169203
65. Types of Health Care Quality Measures. Content last reviewed July 2015. Agency for Healthcare Research and Quality, Rockville, MD. https://www.ahrq.gov/talkingquality/measures/types.html (Accessed 20 February 2023)
66. Donabedian A. Evaluating the quality of medical care. 1966. Milbank Q. 2005;83(4):691-729. doi: 10.1111/j.1468-0009.2005.00397.x.
67. Khan, H., Srivastav, A. and Mishra, A.K. (2020), "Use of Classification Algorithms in Health Care", Tanwar, P., Jain, V., Liu, C.-M. and Goyal, V. (Ed.) Big Data Analytics and Intelligence: A Perspective for Health Care, Emerald Publishing Limited, Bingley, pp. 31-54. https://doi.org/10.1108/978-1-83909-099-820201007
68. Zikos, Dimitrios & Tsiakas, Konstantinos & Qudah, Fadiah & Athitsos, Vassilis & Makedon, Fillia. (2014). Evaluation of Classification Methods for the Prediction of Hospital Length of Stay using Medicare Claims data. ACM International Conference Proceeding Series. 2014. 10.1145/2674396.2674430.
69. https://dam-oclc.bac-lac.gc.ca/download?is_thesis=1&oclc_number=1032984386&id=a857f382-5577-46a6-ba30-4891ba279c7b&fileName=Sama_Thesis.pdf
70 Fontalvo-Herrera, T.; Delahoz-Dominguez, E.; Fontalvo, O. Methodology of classification, forecast and prediction of healthcare providers accredited in high quality in Colombia. Int. J. Productivity and Quality Management 2021, Vol. 33,No. 1, pp.1–20.
71. Vijayalakshmi, G. V.; Mahesh, M.; Mohan K. An ensemble classification based approach for breast cancer prediction. IOP Conf. Series: Materials Science and Engineering 1065 (2021) 012049 doi:10.1088/1757-899X/1065/1/012049
72. Brandt, P. et al. An Investigation of Classification Algorithms for Predicting HIV Drug Resistance without Genotype Resistance Testing. Foundations of Health Information Engineering and Systems. Lecture Notes in Computer Science. 8315 2014, 236-253. doi:10.1007/978-3-642-53956-5_16.
73. Rodrigues et al. Predicting the outcome for COVID-19 patients by applying time series classifcation to electronic health records. BMC Medical Informatics and Decision Making 2022, 22:187 https://doi.org/10.1186/s12911-022-01931-5
74. Hosmer, D.W., Hosmer, T., le Cessie, S. and Lemeshow, S. 'A comparison of goodness-of-fit

tests for the logistic regression model', Statistics in Medicine, 16, 965-980 (1997).
75. Hosmer, D.W. and Lemeshow, S. 'A goodness-of-fit test for the multiple logistic regression model', Communications in Statistics, A1 0, 1043-1069 (1980).
76. Hosmer, D.W. and Lemeshow, S. Applied Logistic Regression: Second Edition, John Wiley and Sons Inc., New York, NY (2000).
77. Fawcett, T. ROC graphs: notes and practical considerations for data mining researchers; Technical Report HP Laboratories: Palo Alto, CA,USA, 2003;
78. Vuk, M. and Curk, T. ROC curve, lift chart and calibration plot. *Metodoloski zvezki* **2006**, Vol. 3,No. 1,pp. 89-108.
79. Dimić, G.; Prokin, D.; Kuk, K. and Micalović, M. Primena Decision Trees i Naive Bayes klasifikatora na skup podataka izdvojen iz Moodle kursa. In Proceedings of the conference INFOTEH, Jahorina, Bosnia and Herzegovina, 21-23 March 2012; Vol 11,pp. 877-882.
80. Witten, H. and Eibe, F. Data mining: Practical machine learning tools and techniques", 2nd ed., Morgan Kaufmann, 2005.
81. Benoît, G. Data Mining. *Annual Review of Information Science and Technology* **2002**, Vol. 36, pp. 265-310.
82. Romero, C.; Ventura, S.; Espejo, P. and Hervas, C. Data mining algorithms to classify students. Proceedings for the 1st IC on Educational Data Mining (EDM08), Montreal, Canada , 2008;pp. 20-21.
83. Pearl, J. Probabilistic reasoning in intelligent systems: networks of plausible inference", Morgan Kaufman Publishers, 1988.
84. Harry, Z. The Optimality of Naive Bayes", FLAIRS conference, 2004.
85. Rokach, L.; Maimon, O. Decision Trees. In book: The Data Mining and Knowledge Discovery Handbook 2005 (pp.165-192) 10.1007/0-387-25465-X_9.
86. Xiaohu, W.; Lele, W.; Nianfeng, L. An Application of Decision Tree Based on ID3. *Physics Procedia* **2012**, 25. 1017-1021. 10.1016/j.phpro.2012.03.193.
87. Quinlan, J. R. C4.5: Programs for Machine Learning", San Mateo, Morgan Kaufmann, 1993.
88. Friedman, J.; Hastie, T. and Tibshirani, R. Additive Logistic Regression: A Statistical View of Boosting. *The Annals of Statistics* **2000**, Vol. 28, No. 2, pp. 337-407.
89. Bella, A.; Ferri, C.; Hernández-Orallo, J.; Ramírez-Quintana, M.J. Calibration of machine learning models. In: Handbook of research on machine learning applications; IGI Global: Hershey, USA, 2009;
90. Park, H.A. An Introduction to Logistic Regression: From Basic Concepts to Interpretation with Particular Attention to Nursing Domain. *J Korean Acad Nurs* **2013**, Vol.43, No.2, 154-164, https://doi.org/10.4040/jkan.2013.43.2.154
91. Rajendra,P;Latifi, S. Prediction of diabetes using logistic regression and ensemble techniques, *Computer Methods and Programs in Biomedicine Update* **2021**,Volume 1,100032, https://doi.org/10.1016/j.cmpbup.2021.100032.
92. IBM SPSS Statistics. https://www.ibm.com/products/spss-statistics (accessed on 15.02.2023.)
93. Zadrozny, B. and Elkan, C. Obtaining calibrated probability estimates from decision trees and naive bayesian classifiers. In Proceedings of the Eighteenth International Conference on Machine Learning, 2001; Morgan Kaufmann Publishers, Inc.; pages 609–616.
94. Weka (University of Waikato: New Zealand) Available from: http://www.cs.waikato.ac.nz/ml/weka (accessed on 20.03.2023.)
95. Liu, H.; Motoda, H. Feature Selection for Knowledge Discovery and Data Mining. Kluwer Academic Publishers, London, 1998.
96. Hall, M. A.; Smith, L.A. Practical feature subset selection for machine learning. In Proc. of the 21st Australian Computer Science Conference, pp 181–191, 1998.
97. Moriwal, R.; Prakash, V. An efficient info-gain algorithm for finding frequent sequential traversal patterns from web logs based on dynamic weight constraint. In Proceedings of the CUBE International Information Technology Conference (CUBE '12) (New York: ACM), pp 718-723(2012)
98. Pravena Priyadarsini. R. et. al.: Gain ratio based feature selection method for privacy preservation. ICTACT Jounal on soft computing, 2011, Vol. 1(4), doi: 10.21917/ijsc.2011.0031

99. Turhan, N.S. Karl Pearson's chi-square tests. *Educational Research and Reviews*, **2020**, Vol. 15(9), pp. 575-580, September, 2020, DOI: 10.5897/ERR2019.3817

100. Robnik-Šikonja, Marko, and Igor Kononenko. Theoretical and empirical analysis of ReliefF and RReliefF. *Machine learning*, **2003**, 53.1: 23–69.
101. Xie, Y.; Li, D.; Zhang, D.; Shuang, Ha. *An Improved Multi-label Relief Feature Selection Algorithm for Unbalanced Datasets*. **2018**, 141-151. doi: 10.1007/978-3-319-69096-4_21.
102. https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset (accessed on 20.03.2023.)
103. Mangasarian, O. L.; Setiono, R. and Wolberg, W.H. Pattern recognition via linear programming: Theory and application to medical diagnosis, in: Large-scale numerical optimization, Thomas F. Coleman and Yuying Li, editors, SIAM Publications, Philadelphia 1990, pp 22-30.
104. https://towardsdatascience.com/how-to-improve-the-accuracy-of-a-regression-model-3517accf8604 (accessed on 20.03.2023.)
105. Frank Harrell (https://stats.stackexchange.com/users/4253/frank-harrell), Hosmer-Lemeshow vs AIC for logistic regression, URL (version: 2023-03-22): https://stats.stackexchange.com/q/18772 (accessed on 20.03.2023.)

106. Chikhachev, S.A. Generic models. Algebra and Logic 14, 214–218 (1975).

https://doi.org/10.1007/BF0166855

107. Shelah, S. (1972). A note on model complete models and generic models. Proceedings of the American Mathematical Society, 34(2), 509-514.
108. Smyth, P.; Wolpert, D. *Machine Learning*, **1999**, 36: 59–83. doi:10.1023/A:1007511322260.
109. Bennett, K. P. and Mangasarian, O. L. Robust linear programming discrimination of two linearly inseparable sets, *Optimization Methods and Software*, **1992**, 1, 23-34.
110. https://scikit-learn.org/stable/modules/ensemble.html#stacking. (accessed on 20.03.2023.)
111. Wolpert, D. Stacked Generalization, *Neural Networks*, **1992**, 5 (2): 241–259. doi:10.1016/s0893-6080(05)80023-1
112. Breiman, L. Stacked regressions. *Machine Learning*, **1996**, 24: 49–64. doi:10.1007/BF00117832
113. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection, 27-Jul2009.*Statistics Surveys 4* **2010**, 40-79 https://doi.org/10.1214/09-SS054
114. Aleksić, A.; Nedeljković, S.; Jovanović, M.; Ranđelović, M.; Vuković, M.; Stojanović, V.; Radovanović, R.; Ranđelović, M.; Ranđelović, D. Prediction of Important Factors for Bleeding in Liver Cirrhosis Disease Using Ensemble Data Mining Approach. Mathematics 2020, 8, 1887. https://doi.org/10.3390/math8111887
115. Ranđelović, D.; Ranđelović, M.; Čabarkapa, M. Using Machine Learning in the Prediction of the Influence of Atmospheric Parameters on Health. Mathematics 2022, 10, 3043. https://doi.org/10.3390/math10173043
116. Aleksić,A.; Ranđelović,M.; Ranđelović,D. Using Machine Learning in Predicting the Impact of Meteorological Parameters on Traffic Incidents, *Mathematics* **2023**, 11(2), 479; DOI: 10.3390/math11020479
117. Ranđelović, M.; Aleksić, A.; Radovanović, R.; Stojanović, V.; Čabarkapa, M.; Ranđelović, D. One Aggregated Approach in Multidisciplinary Based Modeling to Predict Further Students' Education. *Mathematics* **2022**, 10, 2381. https://doi.org/10.3390/math10142381