# On Parallel Processing of Machine Learning Based On Big Data and Voronoi Tessellation

Vasyl Martsenyuk, Marcin Bernas, Aleksandra Klos-Witkowska and Tomasz Gancarczy

*University of Bielsko-Biala, 2 Willowa, Bielsko-Biala, 43-309, Poland*

**Abstract**

The paper is devoted to the development of an approach to machine learning for Big Data under epistemic and aleatoric uncertainties which are taken into account with the help of corresponding minimax criteria. The keystone of the method is processing subsets of training data in parallel, using the partitioning based on Voronoi tessellation. Computational complexity is analyzed and compared with sequential data processing. An example from a medical application is considered, where the method is investigated for different learners and resampling strategies.

**Keywords** [1]

Parallel machine learning, Big Data, learner, uncertainty, minimax, Voronoi diagram

## 1. Introduction

The active usage of Big Data technology in various branches [1-8] requires the development of high-performance algorithm for solving the Machine Learning (ML) problems. To cope with Big Data parallel computing is one of the most effective solution in the case of ML. It leads to the necessity of the partitioning on Big Data sets. Voronoi diagrams are traditionally used for such type of problems.

The minimax approach (together with maximin and maximax) is traditionally used for regression problems [9], [10]. In the case of ML, one of the generalized minimax approaches is known as the Minimax Probability Machine (MPM) [11].

It can be argued that MPM is a classic result of studying the reliability of intelligent models [12], [9], which can be considered as a typical method of classifying the reliability of learning. The task of MPM optimization is to minimize the upper limit of the probability of incorrect classification of the study of model parameters.

The upper limit of the probability of incorrect classification can be used as an explicit indicator to assess the reliability of classification models. A version of MPM with parametric reduction was proposed in [11], [13] for nonlinear classification problems. Several advanced MPM algorithms have been presented from different points of view [14], [15], [16], [17]. In [15], [16] it was pointed out that in some cases it is necessary to distinguish the probability of incorrect classification of two classes, as one class may be more important than another. In [18], MPM was extended for regression. In [9], MPM was introduced to prepare the fuzzy classifier for a more transparent and understandable classification model. In addition to MPM, the study of the reliability of intelligent models has been considered from other points of view.

For example, the concepts of "conflict" and "ignorance" were introduced to denote the reliability of classification models in [19], [20].

To make the method of adopting the minimum probabilistic method available for learning additional intelligent models and to implement the study of the reliability of these models, a generalized hidden minimum probability machine (GHM-MPM) is proposed. MPM classification was used as an explicit indicator to characterize the reliability of the classification model.

*Information Technology and Implementation (IT&I-2021), December 01–03, 2021, Kyiv, Ukraine*
EMAIL: vmartsenyuk@ath.bielsko.pl (A. 1); mbernaś@ath.bieslko.pl (A. 2); awitkowska@ath. Bielsko.pl(A. 3); tgan@ath.bielsko.pl (A. 4)
ORCID: 0000-0001 -5622-1038(A. 1); 0000-0002-0099-1647 (A. 2); 0000-0003-2319-5974 (A. 3);0000-0002-9709-0860 (A. 4)

## 2. Description of the method

The problem of supervised ML, which means the prediction of $Y$ with the help of $X$, loss function $L$, and the set of probability distributions $\Gamma$ on $(X, Y)$ can be formulated as minimax problem with the respect to $L$, provided that the maximization is due to all possible distributions $G$ and minimization is with the respect to decision rules $\psi \in \Psi$.

$$arg \min_{\psi \in \Psi} \max_{P \in \Gamma} E[L(Y, \psi(X))] \tag{2.1}$$

where $E[\bullet]$ is expectancy.

The problem (2.1) can be solved with the help of introducing the generalization of the entropy maximum principle. Mathematical description of the problem of supervised ML in systemic medical research was presented in [21], [22]. Here we formulate it in the case of minimax criterion. Mathematically ML problem for the systemic medical research is based on the following data. We have dataset $D$, which includes $N$ tuples

$$D = \left\{ X_i \mid i = \overline{1, N} \right\} \tag{2.2}$$

In order to model aleatory uncertainties, consider supervised ML regarding the distribution of learning tuples. For the class of all subsets of $D$ we introduce $\Gamma$, including the distributions of classes of training and testing datasets

$$\Gamma = \left\{ (D_{train,j}, D_{test,j}) \subset D \times D \mid D_{train,j} \cap D_{test,j} = \varnothing, D_{train,j} \cup D_{test,j} = D, j = \overline{1, 2^N}, \right\}, \tag{2.3}$$

where $D_{train,j}$ and $D_{test,j}$ are all possible datasets for training and testing correspondingly. In practice, resampling strategies are distributions of the classes of tuples which are characterizing aleatoric uncertainties the best. We introduce the resampling strategies $\gamma \subset \Gamma$

$$\gamma = \left\{ (D_{train,k}, D_{test,k}) \subset D \times D \mid D_{train,j} \cap D_{test,j} = \varnothing, \cup_k D_{train,k} = \cup_k D_{test,k} = D \right\}, \tag{2.4}$$

As examples of resampling strategies we can consider $cv3, cv5, cv10$, which correspond to fold cross-validation for different $k$.

Each $i$ th tuple $X_i = (x_1^i, x_2^i, \ldots, x_p^i, Y^i)^T$ consists of input data $(x_1^i, x_2^i, \ldots, x_p^i)^T$ (called by also attributes) and output data $Y^i$.

Let raw $x_j = (x_j^1, x_j^2, \ldots, x_j^N)$ present the value of $j$ th attribute of all $N$ tuples. Output attribute $Y = (Y^1, Y^2, \ldots, Y^N)$ includes all output data. The attributes $x_1, \ldots, x_p$ and $Y$ (depending on the tasks of classification or regression) can accept both numerical and categorial values.

In the simplest case, the supervised ML problem is to predict, using the certain predictor, the value of output attribute $Y^{N+1}$ based on the values of attributes ${x_1}^{N+1}, \ldots, {x_p}^{N+1}$. The predictor should maximize the accuracy of prediction of output attribute, namely the probability $P\{Y^j corresponds\ to\ {x_1}^{N+1}, \ldots, {x_p}^{N+1}\}$ for arbitrary $j \in \{1, \ldots, N\}$[1]. Further, applying minimax approach, we introduce $h \in \Psi$ for the considered class of ML models $h(X, \gamma)$, which can be trained and tuned for the data $X \subset D$ and assessed taking into account certain resampling strategies $\gamma$. Comparing different ML models, the goal is to minimize expected losses. But you also need to consider resampling strategies, which should also assess the loss function. This formulation of the ML problem considers two types of uncertainty. Namely, uncertainty in oversampling is aleatory because it is related to data. At the same time, the uncertainty in the choice of models is epistemic. Mathematically, the minimum problem of MN is described as a search for a model $h$ due to

$$arg \min_{h \in \Psi} \max_{\gamma \in \Gamma} E[L(Y, h(X, \gamma))] \tag{2.5}$$

### 2.1. The problem of the dimension reduction

Because real sets of systemic medical studies include dozens of vital signs, morphological, biochemical, and clinical assessments, it is natural to want to reduce the number of symptoms, leaving

the ones with the greatest differences. The principal components analysis method (PCA) is one of the widely used methods of dimensional reduction. Although it is used for unsupervised ML problems, it helps us refine the results when used for supervised ML, such as a classification or regression problem. The task of reducing the number of attributes is extremely important for medical use in interpreting the results. Below we propose a method of its application in conditions of aleatory uncertainty.

When regarding the Voronoi tessellation, the dimension reduction algorithm has to be applied for each Voronoi cell. Moreover, since in the minimax ML (machine learning) problem the loss function is calculated for all resampling strategies, the dimension reduction algorithm must be applied separately for each strategy $\gamma$. We can present arbitrary resampling strategy $\gamma$ as $U_l^r = I_{training}^l(\gamma)$, where $I_{training}^l(\gamma), l \in \overline{1,r}$ is lth sample of indices from 1 to N, which corresponds to training tuples of the strategy $\gamma$

Let $D_l(\gamma)$ be input data coming from $D$, if sample of indices $I_{training}^l(\gamma)$ were applied. Namely, $D_l(\gamma) = \left\{(x_1^i, x_2^i, \ldots, x_p^i, Y^i)^T\right\}_{i\in\{1\ldots N\}\cap I_{training}^l(\gamma)}$, where $N_l < N$ is the number of training tuples in the lth training sample.

Before training and tuning the model $h(X,\gamma)$, we reduce the dimension $p \in N$ of D with the respect to $\gamma$. For this purpose we offer the modification of PCA method with the respect to Voronoi cell and resampling strategy $\gamma$ (see Algorithm 1).

Algorithm 1: PCA for the resampling strategy $\gamma$

Input data: $D = \left\{(x_1^i, x_2^i, \ldots, x_p^i, Y^i)^T\right\}_{i=1}^N, \gamma$

Output: principle components together with the attributes.

1 transform data D into the matrix A including all numerical entries;

2 apply resampling strategy $\gamma$ for A: $A_l = \left\{(a_1^{l,i}, a_2^{l,i}, \ldots, a_p^{l,i}, Y^{l,i})^T\right\}_{i=1}^{N_l} \in R^{p_1+1\times N_l}, l = \overline{1,r}$;

3 for each $D_l(\gamma), l = \overline{1,r}$ do

4 $\overline{a}_{l,i} := \frac{1}{N_l}\sum_{j=1}^{N_l} a_i^{l,j}, i = \overline{1,p_1}$;

5 calculate $Var(a_{l,i}), i = \overline{1,p_1}$;

6 $Var(A_l) := \sum_{i=1}^{p_1} Var(a_{l,i})$;

7 $A_l^{'} := \{a_j^{l,i} - \overline{a}_{l,i}\}_{i=\overline{1,p_1},j=\overline{1,N_l}} \in R^{p_1\times N_l}$.;

8 $C_l := \frac{1}{p_1} A_l^{'}(A_l^{'})^T \in R^{p_1}$.,

9 calculate eigenvalues: $\lambda_{l,1} \le \lambda_{l,2} \le \ldots \le \lambda_{l,p_1}$

10 calculate eigenvectors $C_l: w_{l,1}, w_{l,2} \ldots w_{l,p_1}$ corresponding to $\lambda_{l,1} \le \lambda_{l,2} \le \ldots \le \lambda_{l,p_1}$ respectively

11 $PC1_l := A_l^T w_{l,p_1}, PC2_l := A_l^T w_{l,p_1-1}$;

12 calculate variances $Var(PC1_l)$ та $Var(PC2_l)$;

13 $\mathrm{ExplainedVar}(\mathrm{PC1}_l) := \frac{\mathrm{Var}(\mathrm{PC1}_l)}{\mathrm{Var}(D_l)}$, $\mathrm{ExplainedVar}(\mathrm{PC2}_l) := \frac{\mathrm{Var}(\mathrm{PC2}_l)}{\mathrm{Var}(D_l)}$;

14 $\pi(w_{l,p_1})$ та $\pi(w_{l,p_1-1})$

End

16 $\mathrm{ExplainedVar}(\mathrm{PC1}_l) := \frac{1}{r}\sum_{l=1}^{r} \mathrm{ExplainedVar}(\mathrm{PC1}_l)$,

$$\mathrm{ExplainedVar}(\mathrm{PC1}_2) := \frac{1}{r}\sum_{l=1}^{r} \mathrm{ExplainedVar}(\mathrm{PC1}_2)$$

17 return names of attributes з $\pi(w_{l,p_1}), l = \overline{1,r}$ і $\pi(w_{l,p_1-1}), l = \overline{1,r}$

Next, we describe the basic steps of Algorithm 1. In step 1, we convert all categorical attributes, encoding them as a set of boolean inputs, each of which represents one category 0 or 1. We can generate columns with category flags automatically.

Further we repeat the steps 4-14 for each $D_l(\gamma), l = \overline{1,r}$. So, they include the computing of mean values of raws (Step 4), variance $Var(a_{l,i}), i = \overline{1,p_1}$ (Step 5), general variance (sum of sample variances) $Var(A_l)$ (Step 6), deviation matrix $A'_l \in R^{p_1 \times N_l}$ (Step 7), covariance matrix $C_l \in R^{p_1}$ (Step 8), eigenvalues of matrix $C_l$ due to increasing order (Step 9), eigenvectors $C_l$ (Step 10). Here we consider eigenvectors $w_{l,p_1}$ and $w_{l,p_1-1} \in R^{p_1}$, which correspond to $\lambda_{l,p_1}$and $\lambda_{l,p_1-1}$ respectively. At the step 11 we get two principle components $PC1_l := A_l^T w_{l,p_1}$and $PC2_l := A_l^T w_{l,p_1-1}$. We calculate their variances Var (PC1l) and Var (PC2l) at step 12. From this we obtain the percentages of the explained variance corresponding to the first two components ExplainedVar (PC1l) and ExplainedVar (PC2l) respectively (Step 13). Next, we organize the values of the eigenvectors $w_{l,p_1}$i $w_{l,p_1-1}$ in descending order of their absolute values. For this purpose, we use permutations $\pi(w_{l,p_1})$ and $\pi(w_{l,p_1-1})$. We use the denotion $\pi(x)$ for a permutation that organizes the vector x in descending order of the absolute values of its elements.

Next we return the names of the first ExplainedVar (PC1l) 100% attributes in permutation $\pi(w_{l,p_1})$ and the first ExplainedVar (PC2l) 100% attributes in permutation $\pi(w_{l,p_1-1})$ (Step 14).

After completing the main cycle, we calculate the variance of the main components for the resampling strategy \ gamma (Step 16). Finally, we return the names of the first ExplainedVar (PC1l) 100% attributes, which are most common in permutations $\pi(w_{l,p_1}), l = \overline{1,r}$ and the first ExplainedVar (PC2l) 100% attributes that are most common in permutations $\pi(w_{l,p_1-1}), l = \overline{1,r}$ (Step 17).

As a result of reduction of dimension we receive some numerical matrix $A^{red} = \{(a_1^i; a_2^i, \dots, a_{p_2}^i, Y^i)^T\}_{i=1}^N \in R^{p_2+1 \times N}$, $p_2 \le p_1$. These data can then be used as training to solve ML problems based on the minimax approach.

**Note 1.** Stages 2 and 10-14 are modifications of the traditional PCA algorithm. First, in step 1, we convert all categorical attributes that are widely used in systemic medical research into boolean data. Second, when considering the two main components traditionally used for planar presentation of training kits, we propose an approach to selecting some reduced number of attributes for further research (e.g., developing a ML model). This number is related to the number of variations explained. The latter assumption allows us to truly reduce the size of ML problems in systemic medical research under uncertainty.

**Note 2.** Of course, we must take into account the case if the variance due to the first two components is low. In such cases, we need to take into account the components PC3, PC4 and so on to obtain the appropriate dispersion. Steps 10-14 and other algorithms should be changed accordingly.

**Note 3.** It should be noted that the PCA should be calculated depending on the resampling strategy, as the PCA is applied to training tuples. $D_l(\gamma)$ (not for the whole data set D). Therefore, in Step 14, different features may be selected depending on the sample of indices $I^l_{training}(\gamma)$. In turn, this affects the selection of attributes in the last step of the algorithm for the entire resampling strategy.

## 2.2. General flowchart of parallel machine learning with the help of Voronoi diagrams

The general block diagram (Figure 2.1) allows us to obtain a learner of the ML problem based on a minimax approach with the possibility of accurate, acceptable and stable results. The MN model, formulated under conditions of certainty, is presented in [22]. Here, we summarize a flowchart for solving the problem under uncertainty in both the model and the resampling strategy.

We start with the import and preparation of data (feature generation, gap filling, normalization) collected in EMR systems. Methods of importing data sets from EMR systems are presented. Note that the choice of open source EMR systems over commercial ones is extremely important because it allows open access to clinical data that can be processed and selected for subsequent stages of ML [22].

Then we should define the task from the point of view of MN. This can be regression, classification, grouping, and so on. Resampling strategies $\gamma$ are also defined. For example, resampling strategies

supported by the mlr package include: cross-validation (cv), cross-validation (LOO), re-cross-validation (RepCV), color subsampling, also called Monte Carlo cross-validation (Subsample), Holdout method (training / testing) (Holdout) [23]. In a real application, we are dealing with a large number of attributes. Only some of them can be important for the tasks of MN. Therefore, it is natural to try to reduce the dimension by discarding the attributes with the largest deviations.

Next we specify the set Ψ of appropriate methods (learner) of the solution. The most important is the choice of parameters for the methods, which affects the accuracy of the model. In the next cycle, configure the parameters for each model with Ψ based on all resampling strategies $\gamma$, which are used. The original model will satisfy the criterion of the minimax approach (2).

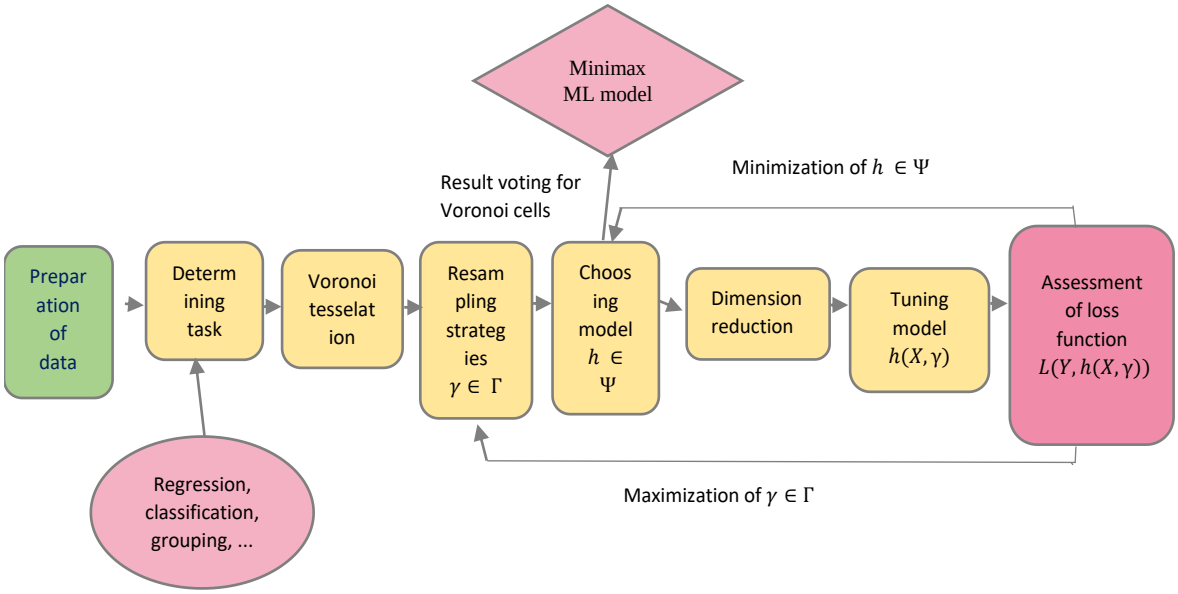


**Figure 2.1**: Development of a parallel ML model for systematic medical research

The capabilities of the mlr package allow us to implement this using appropriate tools designed with certain tasks in mind [24], [25], [26]. The ML model presented above is fully consistent with the mlr package, which offers prototypes of the ML problem cases: task, resampling, learning.

## 2.3. Computational complexity

In order to analyze the computational complexity of the proposed approach, consider an example of a set Ψ, which includes the method of a 4-layer neural network with the number of neurons $i, j, k, l$ on layers based on inverse error propagation and method C5.0 induction of decision tree height $h$.

Assume that the training data sample $D$ includes $\#(D)$ tuples based on $i$ attributes.

Let $v$ be the number of seeds for Voronoi tessellation. Corresponding computational complexity is

$$O_{VD} \coloneqq O\left(v \log v + v^{\left\lceil \frac{i}{2} \right\rceil}\right)$$

The computational complexity of the specified neural network method based on $t$ iterations is **Error! Reference source not found.**:

$$O_{NN} \coloneqq O(t\#(D)(ij + jk + kl)) \tag{2.6}$$

Computational complexity of decision tree induction [28]:

$$O_{C5.0} \coloneqq O(h\#(D) \log \#(D)) \tag{2.7}$$

Computational complexity of resampling based on $k$-fold cross validation is **Error! Reference source not found.**:

$$O_{CV} = O(k\#(D)) \tag{2.8}$$

Thus, the computational complexity of constructing the ML model based on the scheme in Figure 2.1 is

$$O_{ML} = O_{VD} O_{CV} (O_{NN} + O_{C5.0}) \quad (2.9)$$

Since k is constant, then from (2.9) shows that the computational complexity increases by one order of magnitude.

## 3. Example of the medical data

Modern systemic medical research (evidence-based medicine) is the integration of the best scientific evidence with clinical experience and patient expectations [30]. They are aimed at improving health care in the future. Systematic medical research helps doctors and researchers gain knowledge about human health and disease. They also allow you to find more effective ways to prevent and treat disease. Assessment of health is based on a comprehensive and systematic examination of the patient, which includes history, objective examination of the body, analysis of laboratory blood tests and various secretions, instrumental and interventional studies, including X-ray, CT, MRI, endoscopy, biopsy and others methods.

Nowadays, cardiovascular diseases attract attention because they are "the number one cause of death in the world" [31]. In the study of cardiac diseases, there are quite a number of nuances and indicators that experts pay attention to during diagnosis. Diagnostic criteria include both physical tests and history, as well as laboratory, instrumental research methods. During the survey, the doctor may ask questions about the patient's family members (genetic predisposition), lifestyle and habits. Physical inactivity (sedentary lifestyle), unhealthy diet, alcohol consumption and smoking significantly increase the risk of cardiovascular disease. During laboratory studies, much attention is paid to the assessment of the level of lipids and their fractions (lipid profile). It includes indicators of total cholesterol, triglycerides, high, low, very high and very low lipoprotein density, as well as the level of atherogenicity. Lipid imbalance increases the risk of atherosclerosis. Among other things, the patient's overweight is one of the dangerous risk factors for heart disease. Blood glucose and glycated hemoglobin are among the most important indicators of carbohydrate metabolism in the body and markers of diabetes. Diabetes is a separate disease, but its presence significantly increases the risk of cardiovascular disease. In addition to the risk assessment, the necessary extended hematological, biochemical and instrumental studies are performed. In addition to the general blood test, the patient's blood pressure is measured, the following instrumental methods are used: electrocardiogram (ECG), Holter monitoring, echocardiography, coronary angiography, MR angiography.

This experimental study includes data from 1651 patients diagnosed with myocardial infarction. The target attribute of forecasting is life expectancy. Each patient's data includes 97 attributes that contain both numerical and categorical values. Such information includes data on the type of heart attack (focal or transmural), the location of the heart attack (anterior or posterior). Mortality information (hospital, short-term and long-term) is also used. The presence of concomitant pathologies is described. And here we use a detailed analysis, because such pathologies can be combined. Risk factors typical of cardiovascular diseases are investigated, namely, clinical evaluation includes data on such risk factors as gastritis, gallstone disease, lung disease, nephrological disorders, rheumatic thyroid disease, angiopathology, gastrointestinal diseases, oncology, chronic obstructive pulmonary disease, hypertension, diabetes, smoking.

The considered detailed clinical course includes indicators of vital functions, namely heart rate, systolic blood pressure and diastolic blood pressure, analyzes of heart attack complications in the form of arrhythmias, in particular, detailed heart attack complications developed in the hospital. The data lists all indicators of the general analysis of blood. Special attention is paid to leukocytes (WBC), biochemical analysis of blood is presented, information on medicines which the patient received in hospital is included. After the dimension reduction algorithm, the following features remained: sex, age, re-myocardial infarction (RMI), life expectancy after MI (death_days), body mass index (BMI), leukocyte density (White_blood_cells_count), left ejection fraction ventricle (LVEF).

We consider set Ψ, which includes the models of linear regression (regr.lm), SVM model with radial base kernel (regr.ksvm), and random forest (regr.ranger).

Resampling strategies include cross-validation of cv3, cv5, cv7, cv9, cv10. The loss function $L$ was calculated as the rmsse rmse and the training time. In the case of RMSE as an indicator of efficiency (Table 2.1), the regr.ksvm model is a solution of the ML problem based on the minimax. Namely, we first compare the error values for all the models considered. In the second step, we see that the RMSE value for the ksvm model will be minimal among the maximum. In Figure 2.2 we can see the analysis of the effectiveness of ML models with different resampling strategies for standard deviation as an indicator of efficiency.

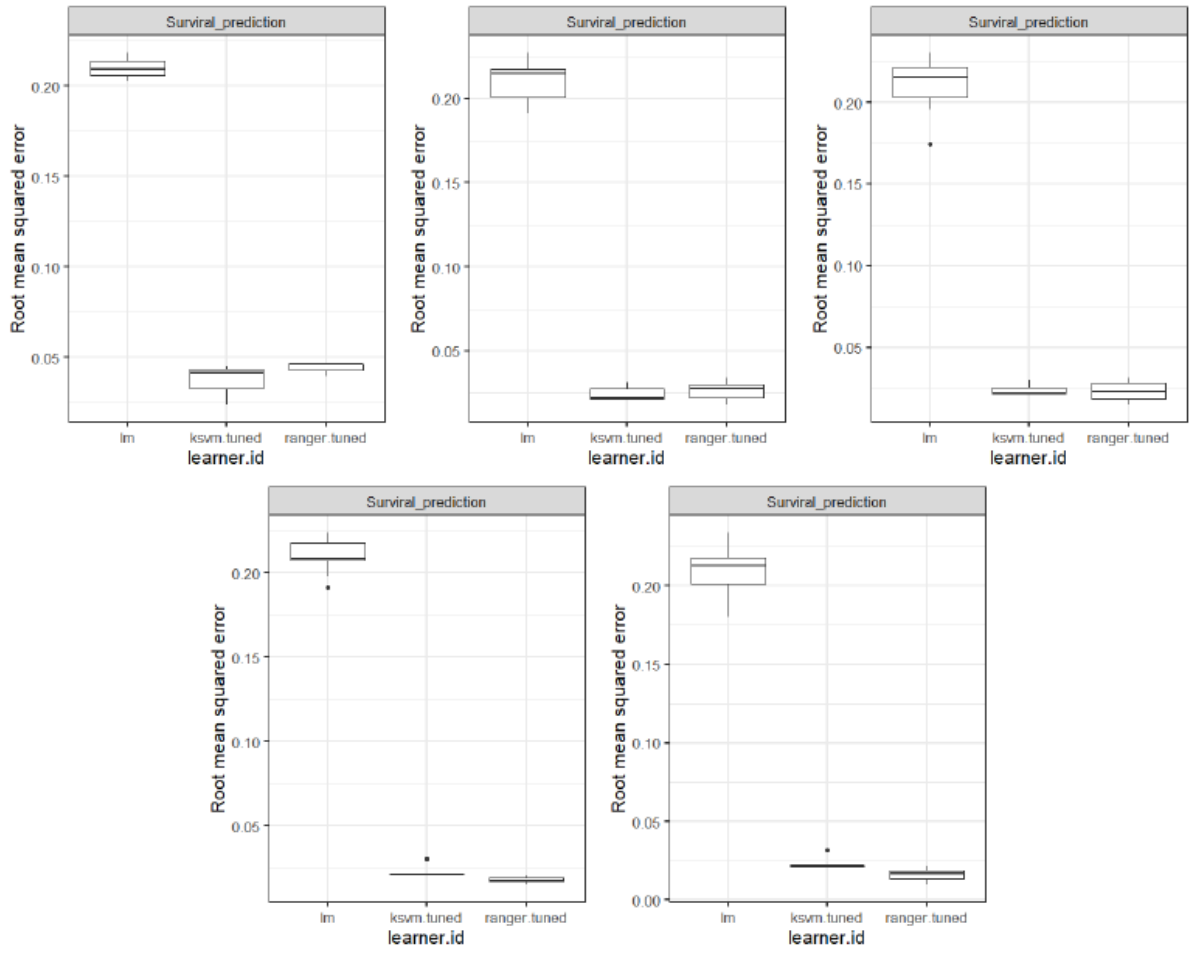


**Figure 2.2:** Comparison of the effectiveness of ML models for different resampling strategies (cv3, cv5, cv7, cv9, cv10) based on RMSE

**Table 2.1.**
RMSE of ML model

| Resampling strategy | Linear regression | SVM | Random forest |
|---|---|---|---|
| **cv3** | 0.2100193167 | 0.03786423738 | 0.04423246658 |
| **cv5** | 0.210649921 | 0.02506639264 | 0.02696577121 |
| **cv7** | 0.2106224629 | 0.02404772839 | 0.02362994412 |
| **cv9** | 0.2098297672 | 0.02371189948 | 0.01803804644 |
| **cv10** | 0.2102195218 | 0.02263394791 | 0.0164288483 |
| **max** | 0.210649921 | 0.03786423738 | 0.04423246658 |
| **minimax** | | **0.03786423738** | |

**Table 2.2.**
Training time of ML model

| Resampling strategies | Linear regression | SVM | Random forest |
|---|---|---|---|
| **cv3** | 0.006666667 | 6.05333333 | 15.62000000 |
| **cv5** | 0.0020000 | 13.98800000 | 35.35600000 |
| **cv7** | 0.0000000 | 18.52428571 | 45.66714286 |
| **cv9** | 0.002222222 | 25.152222 | 65.76222222 |
| **cv10** | 0.0050000 | 28.64400000 | 76.46600000 |
| **max** | 0.006666667 | 28.64400000 | 76.46600000 |
| **minimax** | **0.006666667** | | |

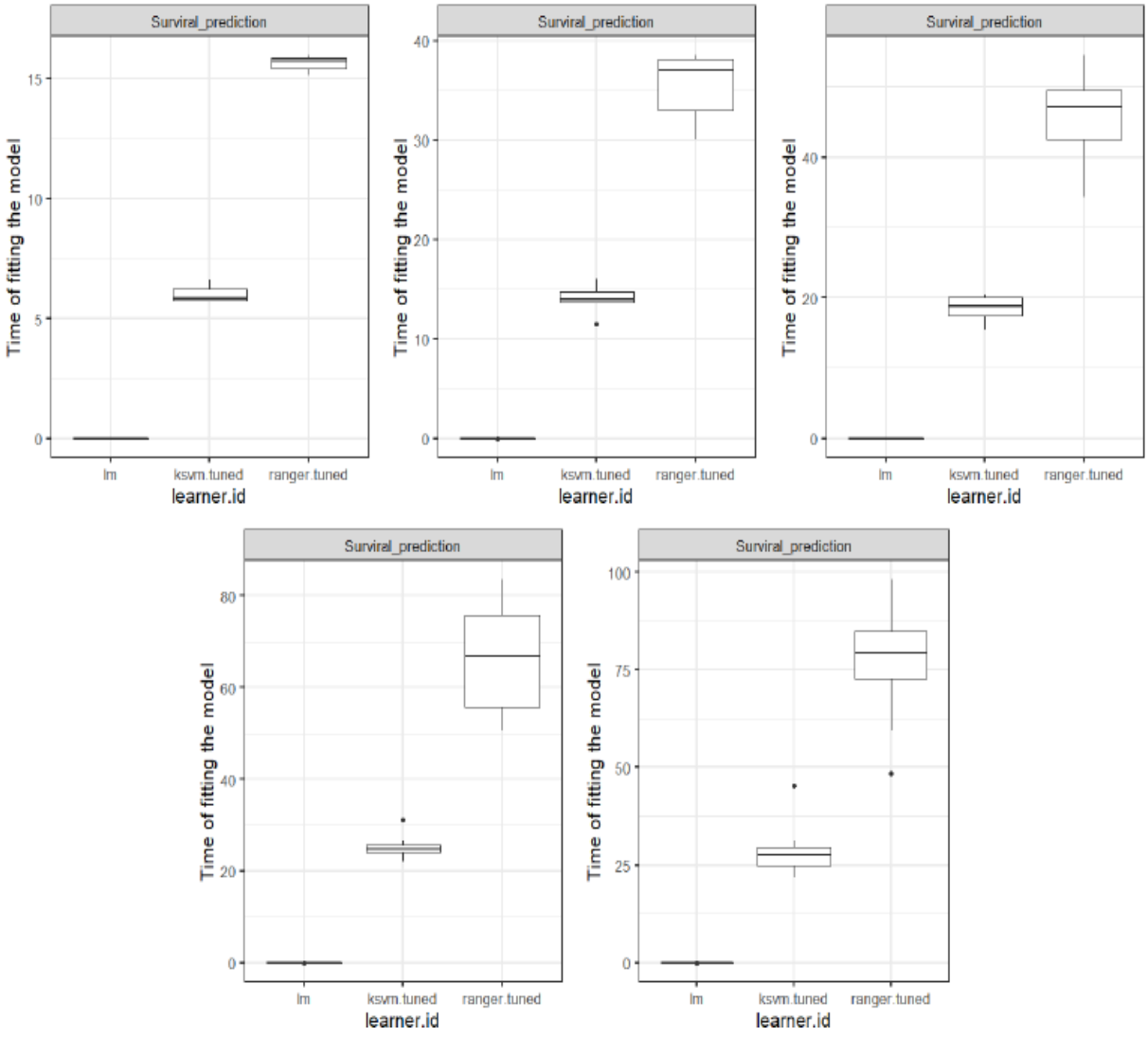


**Figure 2.3:** Comparison of the effectiveness of ML models for different resampling strategies (cv3, cv5, cv7, cv9, cv10) based on training time.

## 4. Conclusions

Based on the above examples, it is established that taking into account the uncertainty in the data (aleatory uncertainty) significantly affects the model based on ML.

For example, as can be seen from Table 2.2, there is a resampling strategy (namely cv10), in which the random forest model shows the lowest value of the RMSE loss function on the set of all models considered.

At the same time, there are resampling strategies (cv3), in which this model shows greater errors compared to the model of SVM. In this situation, the choice of a random forest model would lead to unexpected losses arising from aleatory uncertainty.

Therefore, the minimax approach proposes to establish a resampling strategy with the maximum ("worst") value of the loss function, on which the desired model should behave best (get the minimum value of the loss function).

## 5. References


[1] I.G. Kryvonos, I.V. Krak, O.V. Barmak, A.I. Kulias, Methods to create systems for the analysis and synthesis of communicative information, Cybern. Syst. Anal 53(6), (2017) 847–856. doi:10.1007/s10559-017-9986-7

[2] I. Krak, O. Barmak, E. Manziuk, Using visual analytics to develop human and machine-centric models: A review of approaches and proposed information technology, Computitional Intelligence (2020) 1-26. doi./10.1111/coin.12289

[3] I.G. Kryvonos, I.V. Krak, Modeling human hand movements, facial expressions, and articulation to synthesize and visualize gesture information, Cybernetics and Systems Analysis 47(4) (2011) 501-505. doi:10.1007/s10559-011-9332-4

[4] I.G. Kryvonos, I.V. Krak, O.V. Barmak, A.S. Ternov, O.V. Kuznetsov, Information technology for the analysis of mimic expressions of human emotional states, Cybernetics and Systems Analysis, 51(1) (2015) 25-33. doi:10.1007/s10559-015-9693-1

[5] I.V. Krak, G.I. Kudin, A.I. Kulyas, Multidimensional scaling by means of pseudoinverse operations. Cybernetics and Systems Analysis, 55(1)(2019) 22-29. (2019). doi:10.1007/s10559-019-00108-9

[6] 1. O. Bychkov, K. Merkulova and Y. Zhabska, Software Application for Biometrical Person's Identification by Portrait Photograph Based on Wavelet Transform, 2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT), 2019, pp. 253-256, doi: 10.1109/ATIT49449.2019.9030462

[7] 2. O. Bychkov, K. Merkulova , Y. Zhabska, Information Technology of Person's Identification by Photo Portrait, 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 2020, pp. 786-790, doi: 10.1109/TCSET49122.2020.235542

[8] O. Bychkov, O. Ivanchenko, K. Merkulova, Y. Zhabska, Mathematical methods for information technology of biometric identification in conditions of incomplete data, CEUR Workshop Proceedings of the 7th International Conference "Information Technology and Interactions" (IT&I-2020), 2845, 2021, pp. 336-349. URL: http://ceur-ws.org/Vol-2845/Paper_31.pdf

[9] A. G. Nakonechnyi, A. B. Kachinskiy, Minimax parameter estimators of a linear regression with multiplicative noises, Journal of Automation and Information Sciences 29 (1997) 98–104. doi:10.1615/jautomatinfscien.v29.i2-3.130.

[10] J. Michálek, O. Nakonechny, Minimax estimates of a linear parameter function in a regression model under restrictions on the parameters and variance-covariance matrix, Journal of Mathematical Sciences, 102(1) (2000) 3790–3802. doi:10.1007/bf02680236.

[11] G.R.G. Lanckriet, L.E. Ghaoui, C. Bhattacharyya, and M.I. Jordan, A robust minimax approach to classification, J. Mach. Learn. Res,3(3) (2003), 555-582.

[12] A. G. Nakonechny , V. P. Marzeniuk, Uncertainties in Medical Processes Control, Lecture Notes in Economics and Mathematical Systems 581(2006) 185-192.

[13] G.R.G Lanckriet. , L.E. Ghaoui , C. Bhattacharyya , et al. , Minimax probability machine, Neural Information Processing Systems (NIPS) (2001) 801–807 .

[14] Z. Deng , L. Cao , Y. Jiang , S. Wang , Minimax probability tsk fuzzy system classifier: a more transparent and highly interpretable classification model, IEEE Trans. Fuzzy Syst. 23 (4) (2015) 813–826 .
[15] K. Huang. , H. Yang , I. King , et al. , The minimum error minimax probability machine, J. Mach. Learn. Res, 5 (4) (2004) 1253–1286 .
[16] K. Huang , H. Yang , I. King , M.R. Lyu , Imbalanced learning with a biased minimax probability machine, IEEE Trans. Syst. Man Cybern. Part B 36 (4) (2006) 913–923 .
[17] T. Strohmann , G.Z. Grudic , A formulation for minimax probability machine regression, in: 2002 Neural Information Processing Systems (NIPS), (2002) 769–776 .
[18] T. Strohmann , G.Z. Grudic , A formulation for minimax probability machine regression, in: 2002 Neural Information Processing Systems (NIPS)(2002) 769–776 .
[19] E. Lughofer , Single-pass active learning with conflict and ignorance, Evolving Syst 3 (4) (2012) 251–271 .
[20] E. Lughofer , O. Buchtala , Reliable all-pairs evolving fuzzy classifiers, IEEE Trans. Fuzzy Syst 21 (4) (2013) 625–641.
[21] V. Martsenyuk, L. Babinets, Y. Dronyak, O. Paslay, O. Veselska, K. Warwas, I. Andrushchak, and A. Klos-Witkowska, On development of machine learning models with aim of medical differential diagnostics of the comorbid states," in 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), IEEE, Sep. 2019,pp313-318 doi:10.1109/idaacs.2019.8924345.
[22] V. Martsenyuk, V. Povoroznyuk, A. Semenets, and L. Martynyuk, On an approach of the solution of machine learning problems integrated with data from the open-source system of electronic medical records: Application for fractures prediction, Artificial Intelligence and Soft Computing, Springer International Publishing, 2019, pp. 228-239.doi:10.1007/978-3-030-20915-5_21.
[23] Resampling, https://mlr.mlr-org.com/articles/tutorial/resample.html, (Accessed on 10/08/2020)
[24] Jun Ma, Liming Yang, Yakun Wen, and Qun Sun, Twin minimax probability extreme learning machine for pattern recognition, Knowledge-Based Systems 187, (2020) 104806. doi:10.1016/j.knosys.2019.06.014.
[25] Zhaohong Deng, Junyong Chen, Te Zhang, Longbing Cao, Shitong Wang, Generalized Hidden-Mapping Minimax Probability Machine for the training and reliability learning of several classical intelligent models, Information Sciences 436–437 (2018) 302-319. doi:10.1016/j.ins.2018.01.034.
[26] Jun Ma ,Jumei Shen, A novel twin minimax probability machine for classification and regression, Knowledge-Based Systems 196(2020) 105703. doi.org/10.1016/j.knosys.2020.105703.
[27] Khadiev K. The Quantum Version Of Classification Decision Tree Constructing Algorithm C5.0 / K. Khadiev, I. Mannapov, L. Safina. – 2019.
[28] Z. Pawlak. Rough sets. International Journal of Information and Computer Sciences 11 (1982) 341–356.
[29] W. Xizhao . Learning with Uncertainty / W. Xizhao, Z. Junhai., 2016. – (Taylor & Francis Groups).
[30] D.L. Sackett , W.M. Rosenberg, J.A Gray , R.B. Haynes RB, W.S Richardson , Evidence based medicine: what it is and what it isn't, BMJ, 312(7023),(1996) 71–72.
[31] Cardiovascular diseases, https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_1, 366 (Accessed on 11/24/2020).