Sys Rev Pharm 2021;12(1):993-1005
A multifaceted review journal in the field of pharmacy

Machine Learning Methods for Predicting Tumor Volume in Rats after Termination of Complex Treatment with a Varying Dose of Cyclophosphamide

Konstantin Sarin[1], Marina Bardamova[1], Mikhail Svetlakov[1], Ilya Hodashinsky[1], Evgeny Kostyuchenko[1,*], Denis Pakhmurin[1,2], Artem Slezkin[1], Victoria Pakhmurina[1], Gennadiy Mikhalchenko[1] and Ivan Sidorov [3]

[1]Tomsk State University of Control Systems and Radioelectronics, 40 Lenina Prospect, 634050 Tomsk, Russia
sks@security.tomsk.ru (K.S.); 722bmb@gmail.com (M.B.); rvvincle@gmail.com (M.S.); hodashn@rambler.ru (I.H.); key@fb.tusur.ru (E.K.); pdo@ie.tusur.ru (D.P.); saotom724@gmail.com (A.S.); pdo@ie.tusur.ru (V.P.);
[2]Siberian State Medical University, 2 Moskovsky trakt, 634055 Tomsk, Russia; pdo@ie.tusur.ru
[3]Irkutsk Supercomputer Center of SB RAS, 134, Lermontova, Irkutsk, 664033, Russia; ivan.sidorov@icc.ru

***Corresponding Author**: Evgeny Kostyuchenko Emaiil: key@fb.tusur.ru

ABSTRACT

Predictive determination of tumor volume is a significant task in cancer treatment. The importance of predictive definition is due to its relevance in research and clinical practice. The use of predictive models makes it possible to evaluate the effectiveness of the treatment used in the early stages and decide on its correction in case of unsatisfactory prognosis. The paper investigates models for predictive determination of tumor volume after applying a complex treatment that combines chemotherapy and hyperthermia. Considered machine learning models are ensembles of decision trees, regression based on Gaussian processes, ridge regression, and Bayesian regression. The raw data for machine learning comes from an experiment on laboratory rats. As part of the experiment, all animals received an injection of the fast-growing Walker-256 carcinoma into the left thigh. After the injection, the animals were subjected to complex treatment with cyclophosphamide and hyperthermia. During the study, the rats were divided into three groups. The division into groups depended on the dose of cyclophosphamide received. The determination of informative features is based on the application of greedy and genetic algorithms. The analysis of the selected features with their use in various types of models. The projections of remission and the depth of remission were calculated based on the tumor volume's predicted values. The applicability of several investigated models for predictive assessment of the therapeutic effect of combined treatment has been shown using Student's and Friedman's statistical tests.

Keywords: Oncological diseases; tumor growth modeling; complex treatment; hyperthermia; machine learning; regression; feature selection; cyclophosphamide dose

Correspondence:
Evgeny Kostyuchenko
Tomsk State University of Control Systems and Radioelectronics, 40 Lenina Prospect, 634050 Tomsk, Russia;
Email: key@fb.tusur.ru

INTRODUCTION

Cancer is one of the leading causes of death in the world. In this regard, numerous attempts are being made to find effective means of treating these diseases. Local hyperthermia is a non-invasive method of cancer treatment. Local hyperthermia involves the effect of high temperatures on a malignant neoplasm, as a result of which the destruction of cancer cells does not damage healthy tissues. This method is used as a stand-alone treatment and improves other treatments such as radiation therapy and chemotherapy.

The construction of predictive models of tumor growth is an integral part of theoretical and practical research in the fight against cancer. For example, the use of predictive models makes it possible to assess the effectiveness of the treatment in the early stages and decide about changing it in case of poor prognosis. These results will clearly save time and increase the chances of finding the patient's right treatment option. At present, analytical models of tumor growth in ordinary differential equations and partial differential equations are mainly used [1-4]. In [5], a model was proposed for studying the random growth of the number of tumor cells in the form of the Langevin equation and the corresponding Fokker – Planck equation.

After the application of local hyperthermia, the change in tumor volume is a very complex process with many factors influencing its course. Therefore, in this case, the construction of accurate analytical models of tumor growth is complicated. For solving this problem, it is proposed to use predictive models based on machine learning. Such models' construction involves the description of an object as a black box, which requires obtaining mutually corresponding sets of its input and output values. The main advantage of using this approach is that theoretical studies are not considered to determine the modeled dependencies. Model development is based only on the available empirical input and output values. However, obtaining these datasets requires laboratory medical research.

This work aimed to investigate the application of machine learning methods to predict tumor size in rats after the termination of complex treatment. Complex treatment includes the use of local hyperthermia and chemotherapy.

The contribution of this work is described in the following paragraphs:

1. Experiments on the use of complex treatment on laboratory rats were implemented, and the observation dataset of tumor volume values for machine learning methods was formed.

2. Models have been synthetized to predict tumor volume and remission after discontinuation of complex treatment.
3. The significance of forecasts was determined using statistical criteria, and the ranking of the obtained forecast models was defined.

Related Work

Machine learning is a relevant tool in biomedical research and medical practice. The active application of computer vision technologies makes it possible to work directly with images, for example, to recognize neoplasms on computed tomography results [6]. But also, analytical and predictive models based on the analysis of numerical and nominal data are in demand. Artificial intelligence can be applied in a variety of applications. Data analysis using machine learning is widely employed in genomics since tasks in this area often require processing a large amount of information. Researchers train algorithms to recognize genomic sequence regions, thereby facilitating the process of annotating genomes [7]. Algorithms help classify different disease phenotypes and identify their potentially important biomarkers [7]. Huang et al describe in [8] many examples of using a support vector machine (SVM) to analyze genomic data to determine the subtypes of cancer, search for drug targets and medicines for various types of cancer. Several studies on intelligent algorithms for analyzing cardiovascular diseases' risk [9-11] and predicting their course [12] are reported and discussed in scientific press. Developers create recommendation systems by training decision algorithms based on anamnestic data (age, weight, smoking habits, previous illnesses, lifestyle, and so on) and the results of diagnostic studies and patient observation. Besides, the created telemedicine applications can check the state of the patient's health "at home" to detect the cardiovascular system's early problems.

Recent research includes examples of the use of machine learning in dentistry and ophthalmology. The article [13] shows a model for predicting the condition of patients with Periodontitis based on oral microbiome profiles was created and investigated. The study [13] concluded that the profiles could identify potential microbial biomarkers for periodontitis diagnosis.

In the study [14], the SVM was applied in three problems:

- differentiation of normal eye conditions from painful ones.
- separation of various eye diseases from each other.
- determining the severity of each ocular condition.

Intelligent systems can facilitate the process of interpreting the readings of diagnostic devices. Monitoring the course of pregnancy and reducing the risk of its possible complications are based on the analysis of the fetal heart rate signal [15], gemoviscosimetric assessment of the blood coagulation state system [16]. Also, data analysis allows creating systems that improve the organization of medical services, including emergency care. The Cheng et al [17] developed a model for assessing the risk of transferring a patient diagnosed with COVID-19 to an intensive care unit within 24 hours after hospitalization. Random forests were used as a decision algorithm. Input variables consisted of vital signs, laboratory data, and electrocardiogram. The article [18] assessed the importance of criteria in determining the need for patient extubating and design a model for predicting the extubating effect in a surgical intensive care unit.

Machine learning can be more accurate than regular patient severity assessment techniques. Studies [19, 20] have confirmed that the models based on machine learning techniques over perform regular conventional methods in predicting the mortality rate in patients with suspected infection. These models are superior in quality to the traditional method of analysis, namely, the qSOFA scale.

From a research point of view, machine learning procedures such as feature extraction and feature selection are of great interest. These procedures help determine the most crucial input data attributes, which are difficult to find statistically. One of the main methods of feature extraction is the principal component analysis. In [21], this technique was used for forming feature vectors in the study of the electroencephalogram signal to detect psychological stress. Feature selection can be performed using filtering algorithms or using wrapper algorithms. Filtering algorithms estimate the relationship between input and output variables. Wrapper algorithms select features depending on the change in the value of target function of the model.

The authors of [22] used the Monte Carlo feature selection to analyze gene expression profiles. In this work, the extraction of key genes was hold using a SVM with a sequential selection of variables. As a result, nine genes have been identified that are useful for determining the subtype of brain cancer. The presence of many scientific works indicates a significant potential for the use of artificial intelligence in the oncological field. The current trend in cancer diagnostics is image analysis by convolutional neural networks [23]. However, methods that preliminarily extract features from images remain quite reliable. For example, in the study [24], the mammary gland ultrasound images' segmentation is carried out to highlight the list of features of the tumor state. The features are then classified using a neural network to separate the tumor type into benign and malignant.

In [25], a model for breast cancer classification based on the unification of the architecture of a multilayer neural network tuned by a genetic algorithm is proposed. This model achieved 100% accuracy on test data from the Wisconsin Breast Cancer Dataset from the UCI repository. This set includes quantitative and nominal variables describing fine needle aspiration (cytological images) of breast biopsy material.

Machine learning can be applied both in the diagnostic phase and in treatment planning. Scientists managed to simulate a vaccine administration scheme based on dendritic cells during immunotherapy treatment for mice with Wehi164 fibrosarcoma cancer [26]. For this purpose, Scientists used neural networks [26] Application of the obtained scheme made it possible to reduce the growth rate and tumor volume.

A review article [27] examines the models of tumor growth in the treatment of cancer. Ordinary differential equations and partial differential equations are the primary tools for constructing these models.

In our work, we propose intelligent models based on machine learning that predict tumor volume after the termination of complex treatment with various doses of cyclophosphamide and exposure to hyperthermia. This paper contributes to the development of the methods that would have less severe side effects on the body. One of these methods is a combination of heat therapy and chemotherapy. The potential of a combined method lies in the fact that the therapeutic effect can be achieved with the use of smaller doses. This therapeutic effect is comparable to that of monotherapy with a chemotherapy drug. As a result, the toxic load on the body is reduced. Thus, the effectiveness of the treatment is improved.

MATERIALS AND METHODS

Proposed research uses empirical data from a preclinical comprehensive study of combined therapy in cancer

treatment. The distinctive feature of the study is investigation of the effect of controlled local hyperthermia on tumor growth during chemotherapy.
Further data processing was conducted applying machine learning techniques for time series analysis. Several various feature selection algorithms and predictive models were considered and implemented in the course of the study. The computational experiment was designed to validate and compare proposed predictive models.

Data Sources
Authors participated in experimental studies at Tomsk Cancer Research Institute in close cooperation with the specialists affiliated with this organization. In investigating the effect of controlled local hyperthermia on therapeutic effect, a series of animal experiments was conducted. The biological models were female Wistar rats and C57Bl / 6j mice. In total, 730 animals were used in course of the study: 450 rats and 280 mice. Preliminary experimental results showed that a promising reduction in the chemotherapeutic agent dose while achieving a better therapeutic effect is due to the combined use of controlled local hyperthermia with chemotherapy. So, the stimulating effect of local hyperthermia on chemotherapy was concluded.

Outline of the Experimental Setting
Each series of experiments took about one month without considering the preparation for it. Each series had different objectives:

- to develop the method of introducing heaters and their design,
- by assessing the effect of the procedure on the primary tumor focus and metastasis,
- to study the interaction of controlled local hyperthermia with traditional chemotherapy and a dose of chemotherapy.

Each series of experiments were conducted with a specific purpose. The first two series of experiments were devoted to analyzing the effect of controlled local hyperthermia on tumor tissue. Moreover, at this stage, an assessment was made of both the direct (isolated) effect of controlled local hyperthermia and combined with chemotherapy. The third series of experiments was devoted to assessing the effect of local hyperthermia on the dose of a chemotherapy drug. The third series also included a study of the application of the injection of needle heaters using an ultrasonic generator. This procedure was done to assess the effect of ultrasonic vibrations on the effectiveness of the proposed method of controlled local hyperthermia. Namely, the possibility of treating the wound canal at the time of the injection of heaters was investigated, as well as the near and distant consequences of the procedure in terms of increasing the positive effect.

In the fourth series, there is an analysis of the effect of local hyperthermia on the dose of a jointly used chemotherapy drug. All series of experiments were produced with the following technical parameters:

- stabilization temperature 45 °C.
- the distance between the needle heaters is 8–10 mm.
- diameter of needle heaters 0.8 mm.
- the time to reach the stabilization temperature in the tumor 10–15 minutes.
- the duration of the heating session is 30 minutes.
- the maximum power allocated to each heater is 1 W

Regularization of Empirical Data
In the course of the experimental investigation, the results were obtained, reflecting the dynamics of changes in the health indicators of rats (body weight), the dynamics of tumor development (geometric dimensions in three dimensions), as well as the mass of the tumor. The acquired data are presented as a time series depending on the day number. The data vector includes parameters such as tumor volume from 3 to 13 days and the total weight of the rat on each of these days. This vector, together with the dose of cyclophosphamide used in the treatment, comprises the data set supplied to the system's input. The tumor volume outside the time series was included in the data vector as an output value.

Models Design
In academic science, the phenomenon under study is associated with the changes in certain quantitative or qualitative values. Therefore, when designing a model, the value associated with this phenomenon is determined and is its output value. In our case, the phenomenon is a change in the rat tumor volume, and the actual output of the model is the tumor volume. Bodyweight and tumor volumes on the days of the treatment period are factors influencing the outcome. These factors are identified as features, and they are the actual inputs of the model. The difference between the prediction and actual value is normally associated with the quality of the model.

Statement of the Prediction Problem
Any predictive model is formally represented as a functional dependence $f(x_1,x_n)$ between the features x_1, x_n and the value associated with the phenomenon under study. Predictive models were designed using machine learning algorithms, an observation dataset containing the real values of the input and output quantities $\{(\mathbf{x}_i,y_i) \mid i = 1,m\}$, where $\mathbf{x}_i=(x_1^i, \ldots, x_n^i)$ and y_i are the input and output values of the i-th instance of the observation, respectively. The quality criterion normally used to address the quality of the model is a root-mean-square error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m}\left(y_i - f\left(x_1^i,\ldots,x_n^i\right)\right)^2}{m}}.$$

The observation dataset is represented by 262 vectors of 26 features each. Features x_{2j-1} (j=1,12) correspond to the value of the tumor volume on the (j + 2)-th day of the experiment, and $2x_{2j}$ – to the value of the rat's mass on the same day. Feature x_{25} stands for the number of consecutive days after treatment termination. In the experiment, the estimations of the rat's tumor volume and weight were recorded daily for nine days (from 15 to 23 days of the experiment).
Therefore, x_{25} takes an integer value from the set {1, 2, ..., 9}. Feature x_{26} stands for the dose of cyclophosphamide used in treatment. The domain of values of this feature is the set {10 mg / kg, 8 mg / kg, 5 mg / kg}. The output value y corresponds to the tumor volume in cm^3. The functional model based on the described features is shown as a black box in Figure 1.

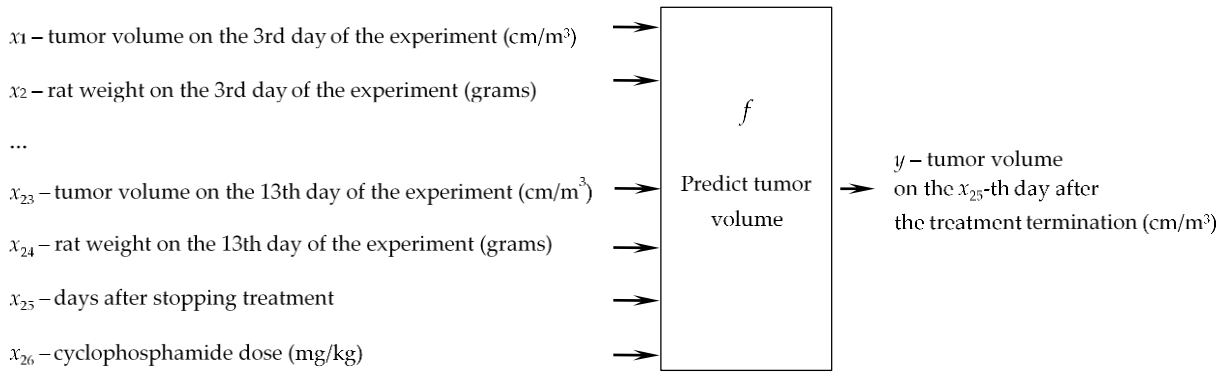

Figure 1. Model for predicting tumor volume as the black box.

The choice of informative features used for forecasting is important when designing models. The choice of the best informative features allows for achieving the following:

- improve the accuracy of the forecast,
- get simpler and less computationally complex models,
- to study more deeply the nature of the observed phenomenon.

Feature selection methods can both be independent algorithms or be a part of the complex algorithms used for a model synthesis [28]. As independent algorithms, these methods are divided into wrappers and filters. Wrappers are the algorithms that use trial and error method. By selecting different options for comprising the groups of features, they find the best option to develop a model. Metaheuristic optimization algorithms for binary search space are widely used here. Filters do not use classificators, unlike wrappers. The selection of features is carried out by analyzing the dependencies between the features' values in the observation dataset. Typical representatives of this approach are algorithms based on correlation and mutual information.

In this work, wrappers were used for designing the models. At stage of feature selection, the RMSE quality criterion was applied. The same criterion was applied for proposed models' evaluation.

Machine Learning Algorithms

Genetic Algorithm (GA). The genetic algorithm is widely used for feature selection when designing predictive machine learning models [29-31]. The main idea of the algorithm focused on modeling the evolutionary process. When solving the problem of feature selection, the search vector (chromosome) is a binary string $s_1, s_2, ..., s_n$, where each bit (gene) s_i is associated with a feature x_i, (i=1,n). The value of bit 1 indicates the feature's presence, and the value 0 indicates its absence. The target function is the RMSE value, which should be minimized. Model f construction and further prediction are performed only with the features selected in the $s_1s_2...s_n$ line. The algorithm iteratively performs three stages of the evolutionary process: selection, crossover, and mutation.

At the selection stage, pairs of individuals are selected, passing into the next generation and participating in crossover. Right after, $(1-p_c)\times N$ individuals with the best values of the target function pass to the next generation (p_c is a parameter of the algorithm, N is the population's size). The rest of the population is used for forming new individuals by crossover. The selection of $p_c\times N$ pairs of parents was produced using the stochastic uniform function.

At the crossover stage, the same number of new individuals is generated from the selected pairs of individuals. A scattered function was used as the crossover operator. A binary vector was randomly created, and genes were selected for a new individual from the first parent where the vector element is 1, and from the second parent where the vector element is 0.

At the stage of mutation, the chromosomes of individuals passing into a new generation undergo random changes. These changes were conducted by bit mutation in this work when each gene takes the opposite value with a probability p_m, where $p_m=n^{-1}$.

The genetic algorithm parameters are.

- population size N.
- the number of generations (iterations) *iter*;
- the proportion of the next generation p_c population generated by crossing.
- the probability of the p_m gene mutation.

Greedy Algorithm (GrA). The major advantage of this algorithm is computational simplicity. Its main idea is that the locally best choice is made at each step [32]. Features are picked over from the current vector of selected features at each iteration of the algorithm. Moreover, the first iteration includes all possible features. This procedure aims to determine the changes in the value of the target function when it is deleted. If the value of the target function deteriorates, then the operation of the algorithm stops. Otherwise, the vector of selected features is removed from the one, in the absence of which the target function's best value is achieved [33].

Predictive Models

Tree Ensemble (TE). Decision tree models are a handy tool for solving classification and regression problems [34, 35]. Combining models into an ensemble improves the predictive potential, as demonstrated in the article [36]. Prediction of rat tumor volume is based on the application of an ensemble of decision trees. Each element of the ensemble of decision trees was constructed using the well-known CART algorithm [35]. When constructing an ensemble, a new element is created and added at each step, which reduces the error between the real output values and the aggregated forecast of all the elements previously created. The selection of the parameters $\mathbf{a}_j = (a^j_1, a^j_k)$ of each j-th element of the ensemble was conducted using the gradient descent method, where the gradient is calculated as follows:

$$\nabla\varphi_j = \left(\frac{\partial\varphi_j}{\partial a_1^j}, \frac{\partial\varphi_j}{\partial a_2^j}, \ldots, \frac{\partial\varphi_j}{\partial a_k^j}\right).$$

Here the target function is the RMSE between the real and aggregated predicted values of the ensemble:

$$\varphi_j = \frac{\sum_{i=1}^{m}\left(y_i - \overline{f}_j\left(x_1^i, \ldots, x_n^i\right)\right)^2}{m},$$

where $\bar{f}_i\ \left(x_1^i, \ldots, x_n^i\right)$ is the aggregated forecast of an ensemble of *j*-elements for the *i*-th instance (*j*-1 elements were created earlier). This approach to constructing an ensemble is known as a gradient boosting [34]. The aggregate prediction of the ensemble is the mean value of all elements. The construction algorithm's main parameters are the number of elements *N* and the maximum element depth L_{max}.

Gaussian Process Regression (GPR). Models based on Gaussian processes are also widely used for solving regression problems and classification [37]. In predicting a rat tumor's volume, the output features are interpreted by this regression model as noisy values of some hidden Gaussian process with zero mean value:

$$f \sim GP\left(0, k(\cdot,\cdot)\right),$$

where $k(\cdot,\cdot)$ is the covariance function.

For the problem under consideration, the sum of the Matern kernel and white noise was chosen as the covariance function. The parameters of the covariance function of the Gaussian process are estimated with the limited-memory Broyden-Fletcher-Goldfarb-Shanno with the simple box constraints (L-BFGS-B) [38]. This method belongs to the class of quasi-Newtonian methods; at each iteration, the Hessian is approximated, and the direction of descent is chosen. Thus, when developing a model, we consider the coefficient μ of the Matern covariance function and the noise level *nl* as the main parameters of the algorithm.

Extremely Randomized Trees Regression (ERTR). This model was proposed in [39]. Its main differences from the usual ensemble of decision trees are that each tree is trained using the entire training set, and tree branches are cut off randomly rather than by calculating the local optimal cut point. When developing this model, gradient boosting is applied. The construction algorithm's main parameters are the number of elements *N* and the maximum element depth L_{max}.

Ridge Regression (RR). Ridge regression is an enhancement to linear regression. Ridge regression is characterized by increased robustness to errors by imposing constraints on the regression coefficients. This method is also called the Tikhonov regularization method. A "ridge" is added to the linear regression model's minimized functional, which regularizes the term. As a result, the coefficients are compressed, and the effective dimension decreases. However, the number of features does not change in this case [40]. The main parameter of the algorithm is the value of the regularization coefficient *alpha.*

Bayesian Ridge Regression (BRR). In this model, it is assumed that there is normally distributed noise in the initial data, and the maximum likelihood method is used for solving the problem [41]. The application of the Tikhonov regularization method is present in this model, as in the previous one. The main parameters of the algorithm are the values of the parameters of the regularizing term distribution (*lambda1, lambda2*), the parameters of noise distribution (*alpha1, alpha2*) and the parameter *tol*, which determines when the optimization algorithm stops.

LASSO model fit with Least Angle Regression and Information Criterion (LLIC). The LASSO (Least Absolute Shrinkage and Selection Operator) model is an enhancement to linear regression. The Lasso model's main idea is to add a term to the minimized functional of linear regression model. However, for this model, a different regularization method is used that is different from that used in ridge regression. In the case of ridge regression, regularization contributes to small values of the model coefficients. In the Lasso model, regularization can reduce some of the model's coefficients to zero. The estimation of the model's coefficients under consideration is based on the application of the LARS optimization algorithm (Least Angle Regression for laSso / Stepwise regression) (the algorithm has the *max_iter* parameter). The choice of the value of the regularization parameter is based on the information criterion (*IC*). In this model, the information criterion of Akaike or Bayes can be applied [42].

Support Vector Machine (SVM). The SVM is characterized by using kernels and by the absence of local minima. The SVM main idea is to construct a hyperplane, separating objects of different classes as much as possible. Its versatility characterizes the method. In addition to classification problems, the SVM expansion is also used for solving regression problems [43]. One of the advantages of support vectors regression is that the quadratic minimization problem is solved to uniquely determine the regression model's parameters. For the problem under consideration, a linear kernel function was chosen. The algorithm's main parameter is the regularization parameter *C* and the error of loss function *ε*.

Computational Experiments

We conducted a series of experiments to develop models for predicting tumor volume after the termination of complex treatment. We used empirical data and described machine learning techniques for the synthesis of predictive models. Table 1 shows the parameters for the models. The quality of the models was assessed according to the scheme of 29-fold cross-validation [34]. Twenty-nine rats underwent a full cycle of treatment. Therefore, there we used 29 subsamples for assessing the quality of the models. Each pair (one for training and another for validation) is characterized by the fact that the test data contain some instances related to a specific rat, while the training data do not. The number of data instances corresponding to one rat is equal to 9, where each record is associated with certain days after the termination of treatment (9 days in total).

Table 2. Values of models' parameters.

Model	Parameters
GA	N=100; *iter*=50; p_c=0.8; p_m=26^{-1}
TE	N=100; L_{max}=20
GPR	μ=0.5; *nl*=1
BRR	*Alpha*1=1*e*-6;*Alpha2*=1*e*-6;*Lambda*1=1*e*-6;*Lambda*2=1*e*-6;*tol*=1*e*-3
ERTR	N=100; L_{max}=20
LLIC	*IC=AIC; max_iter*=500
RR	*Alpha*=1*e*-6
SVM	ε= 0.008;C=1.0

When synthesizing models with feature selection by the GA, the algorithm was run 30 times. If a feature was selected 20 or more times out of 30 replications (i.e., more than 2/3 of the cases), it was used in the final model. Table 3 shows the average *RMSE* values for the designed models.

Table 3. Assessment of the models' quality.

Model	Feature Selection	*RMSE*
TE	GA	3.20
TE	-	4.25
GPR	GA	3.31
GPR	-	3.92
LR[1]	GA	3.48
LR	-	3.53
BRR	GrA	3.21
BRR	-	3.07
ERTR	-	3.04
LLIC	GrA	2.91
LLIC	-	2.90
RR	GrA	2.98
RR	-	3.00
SVM	GrA	3.32
SVM	-	3.47
[1] Linear Regression		

RESULTS AND DISCUSSION

Assesment of the Prediction Quality

Remission and the depth of remission are important characteristics of the results of complex treatment. Remission R is assessed for each rat by the area of the shaded figure (Figure 2). This figure is formed by the curved line of the tumor volume and the abscissa in the time period after the termination of complex treatment. The minimum tumor volume indicates the remission depth D_R during the same period.

The predicted values of R and D_R were calculated based on the predicted values of tumor volume:

$$R = \frac{y_{day=14} + \hat{y}_{day=23}}{2} + \sum_{i=15}^{22} \hat{y}_{day=i},$$

$$D_R = \min_{i=15}^{23} \left(\hat{y}_{day=i} \right),$$

where $y_{day=i}$ и $\hat{y}_{day=i}$ are the real and predicted values of the rat tumor volume on the i-th day of the experiment. Table 4 and Table 5 show the actual (column *Actual*) and predicted values of remission. Table 6 and Table 7 show the remission depth.

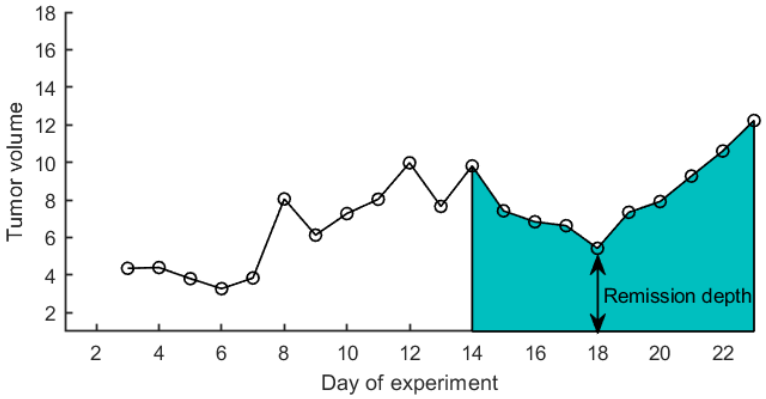

Figure 2. Remission (R) and remission depth (D_R).

Table 4. Actual values and predictions of remission R. The first part.

No	Actual	GA+TE	TE	GA+GPR	GPR	GA+LR	LR	GrA+BRR	BRR
1	72.46	71.76	29.63	71.20	68.04	63.19	62.37	63.53	63.58
2	45.27	34.56	28.35	32.58	75.50	40.81	47.47	40.46	41.26
3	104.73	107.75	87.01	117.59	114.77	99.12	108.08	113.62	112.85
4	152.87	135.21	124.07	132.66	141.53	126.50	125.79	139.04	136.40
5	145.11	174.39	194.60	160.85	122.66	120.05	122.22	144.28	135.22
6	113.92	147.60	119.64	150.72	138.73	120.58	116.94	135.17	131.69
7	89.35	64.30	57.58	66.85	72.76	71.05	68.04	68.33	66.61
8	89.16	84.86	64.15	93.81	116.27	79.13	85.13	85.07	90.08
9	176.94	165.58	141.84	121.63	146.14	133.30	135.31	140.86	148.63
10	115.35	88.00	151.31	133.50	115.61	112.04	112.41	121.65	117.13
11	78.88	69.02	74.71	68.34	81.96	85.14	86.04	94.77	88.08
12	66.30	67.89	82.89	64.30	51.39	62.08	59.41	69.40	72.28
13	77.30	70.21	42.35	84.18	70.60	72.91	75.78	71.30	72.80
14	131.16	150.35	160.61	169.58	166.67	131.92	134.83	156.46	156.17
15	70.65	65.43	59.67	60.71	58.24	65.99	65.83	68.55	74.04
16	124.10	109.17	130.41	126.68	143.61	126.45	131.85	136.04	142.66
17	98.45	66.05	62.44	60.65	85.85	86.47	88.78	90.13	90.90
18	43.51	80.36	86.25	77.74	79.72	81.07	83.07	94.54	87.66
19	58.10	67.89	82.89	71.46	72.16	71.96	77.16	71.64	75.59
20	68.93	57.58	42.87	69.08	68.29	57.71	59.77	53.59	57.50
21	96.33	81.39	82.28	90.51	126.40	107.74	101.80	110.01	110.81
22	36.49	46.29	40.62	60.30	67.21	47.69	43.90	46.18	46.00
23	59.59	47.14	50.45	47.40	62.30	57.59	49.77	52.82	52.92
24	45.26	75.66	90.74	72.92	52.79	65.60	62.32	63.46	63.50
25	86.93	74.24	79.49	92.90	126.61	92.78	84.21	101.31	97.06
26	49.08	99.78	93.83	77.89	146.79	75.56	67.51	58.93	62.18
27	178.01	139.99	129.23	138.20	102.54	114.75	101.76	106.56	103.88
28	65.15	73.35	79.45	113.17	114.88	72.76	72.44	72.77	73.39
29	81.40	58.41	64.86	72.90	65.78	66.26	61.28	66.04	64.57

Table 5. Actual values and predictions of remission R. The second part.

No	Actual	ERTR	GrA+LLIC	LLIC	GrA+RR	RR	GrA+SVM	SVM
1	72.46	67.85	65.99	68.16	63.56	63.15	75.00	75.66
2	45.27	56.97	49.37	49.39	46.26	42.53	43.33	45.20
3	104.73	108.67	106.00	105.88	109.17	111.12	114.72	115.04
4	152.87	135.36	133.79	132.02	139.90	137.25	138.10	127.13
5	145.11	125.88	133.75	132.63	134.76	131.96	144.11	141.16
6	113.92	124.81	136.57	135.93	134.20	129.43	139.91	142.07
7	89.35	68.05	72.00	72.05	73.70	69.26	76.85	77.77
8	89.16	92.74	88.80	88.87	88.16	91.11	85.81	86.08
9	176.94	141.35	143.21	143.21	143.69	147.69	140.21	138.15
10	115.35	123.37	114.63	114.51	119.35	115.97	125.33	126.63

11	78.88	72.80	84.49	84.55	89.63	86.63	102.29	101.57
12	66.3	68.49	81.99	82.25	75.76	74.42	59.18	52.36
13	77.3	76.55	75.22	75.41	75.41	74.67	71.52	72.81
14	131.16	152.50	142.26	140.70	147.62	153.18	145.93	153.95
15	70.65	76.83	71.79	71.98	73.81	75.98	58.29	60.37
16	124.1	133.39	125.47	124.84	136.98	142.57	134.14	136.51
17	98.45	80.68	84.76	84.76	93.88	91.69	89.95	90.18
18	43.51	82.77	80.22	80.88	86.81	86.83	93.10	93.03
19	58.1	78.64	76.64	76.74	72.13	76.16	77.32	76.10
20	68.93	67.40	65.35	65.54	57.93	59.96	53.72	53.86
21	96.33	114.30	111.47	111.38	107.90	108.93	102.46	102.46
22	36.49	52.14	50.25	52.10	46.31	45.66	41.07	39.92
23	59.59	53.49	53.87	54.76	55.25	53.57	48.23	41.82
24	45.26	61.44	67.08	68.35	64.51	63.47	54.51	54.56
25	86.93	93.25	98.86	98.82	96.02	95.71	100.42	100.42
26	49.08	85.61	61.75	61.77	70.86	63.83	56.61	59.15
27	178.01	114.39	107.53	110.17	101.52	101.60	110.73	110.73
28	65.15	75.94	71.98	71.58	76.44	74.56	76.51	76.39
29	81.4	59.24	66.58	66.62	63.40	64.13	64.75	66.71

Table 6. Actual values and predictions of remission depth D_R. The first part.

No	Actual	GA+TE	TE	GA+GPR	GPR	GA+LR	LR	GrA+BRR	BRR
1	5.42	6.91	2.07	6.68	7.02	5.59	5.73	5.81	5.86
2	3.85	2.50	1.83	2.51	7.84	2.97	3.78	2.09	2.28
3	7.87	8.93	6.82	10.86	10.47	9.58	10.76	9.54	9.64
4	10.25	12.73	10.81	10.03	10.38	12.99	13.12	11.32	11.32
5	12.03	14.18	15.44	14.25	11.10	11.80	12.15	12.00	11.31
6	8.42	11.54	10.15	11.28	13.03	12.48	12.10	10.82	10.92
7	6.89	5.61	4.02	5.30	7.08	6.75	6.56	5.90	5.79
8	8.55	7.44	5.60	7.60	10.45	7.31	8.15	7.03	7.29
9	13.73	12.65	13.23	9.77	12.51	13.62	13.98	11.76	12.11
10	9.08	7.11	9.36	10.40	10.07	11.22	11.51	9.98	9.62
11	5.09	6.34	6.92	5.59	6.95	7.98	8.15	7.84	7.62
12	5.66	6.44	5.98	2.80	4.19	5.33	5.08	5.89	6.45
13	6.66	6.14	2.69	8.79	6.85	6.87	7.35	6.74	6.42
14	11.16	11.10	11.47	12.42	14.62	13.46	14.00	11.46	10.79
15	5.17	5.65	5.39	5.12	5.10	5.98	6.13	6.40	6.43
16	8.65	9.07	9.90	10.65	13.23	12.86	13.65	11.36	11.75
17	7.15	5.79	4.52	5.34	8.35	8.39	8.78	7.64	7.86
18	3.47	7.99	7.77	6.56	6.61	7.69	8.09	7.67	7.53
19	4.99	6.20	7.51	5.99	6.32	6.67	7.32	6.18	6.69

Table 6. Continuation.

No	Actual	GA+TE	TE	GA+GPR	GPR	GA+LR	LR	GrA+BRR	BRR
20	5.60	5.19	3.11	5.61	6.37	5.04	5.46	4.92	5.16
21	8.10	7.40	4.99	8.39	11.44	10.44	9.83	9.08	9.38
22	2.75	4.49	1.44	4.38	6.83	3.73	3.25	3.45	3.67
23	4.44	3.99	4.06	4.48	6.02	5.13	4.34	4.05	4.22
24	3.98	6.37	8.56	6.90	4.80	5.84	5.38	5.34	5.53
25	7.09	7.22	7.29	8.04	9.55	9.02	8.17	8.06	7.67
26	3.97	8.30	6.36	7.50	16.46	6.98	6.01	3.12	3.88
27	12.07	12.39	10.29	10.54	9.70	11.48	10.08	9.14	8.85
28	5.80	7.12	6.93	11.70	12.60	6.60	6.55	6.24	6.34
29	7.20	4.99	5.85	7.43	6.47	6.18	5.68	5.51	5.60

Table 7. Actual values and predictions of remission depth D_R. The second part.

No	Actual	ERTR	GrA+LLIC	LLIC	GrA+RR	RR	GrA+SVM	SVM
1	5.42	5.79	6.25	6.43	5.96	5.93	5.74	6.19
2	3.85	4.55	3.89	3.89	3.67	2.94	2.68	2.72
3	7.87	9.59	9.33	9.28	9.22	9.45	9.22	9.23
4	10.25	11.23	10.95	10.84	11.61	11.40	12.44	10.97

5	12.03	10.24	10.78	10.78	11.27	11.06	13.66	12.37
6	8.42	9.51	10.13	10.12	11.08	10.60	9.86	10.66
7	6.89	6.21	6.12	6.12	6.47	6.09	4.48	5.39
8	8.55	8.43	8.02	8.02	7.41	7.58	6.93	7.29
9	13.73	10.84	11.51	11.51	11.69	12.03	12.46	12.34
10	9.08	9.78	9.05	9.05	9.82	9.52	10.10	10.34
11	5.09	6.73	7.29	7.31	7.78	7.49	8.45	8.66
12	5.66	6.70	7.09	7.31	6.95	6.86	3.66	2.93
13	6.66	6.81	6.68	6.68	6.98	6.74	6.55	6.46
14	11.16	11.31	10.49	10.02	10.53	10.86	7.39	8.96
15	5.17	6.39	6.45	6.45	6.77	6.87	3.73	4.68
16	8.65	10.69	10.72	10.62	10.95	11.59	10.82	10.57
17	7.15	6.85	7.34	7.34	8.15	8.00	6.68	7.10
18	3.47	7.42	6.68	6.87	7.51	7.51	7.32	7.20
19	4.99	7.80	7.06	7.15	6.38	6.86	7.41	5.53
20	5.60	6.13	6.05	6.05	5.49	5.63	4.17	4.16
21	8.10	10.01	9.65	9.56	8.88	9.07	8.40	8.40
22	2.75	4.80	3.70	3.70	3.60	3.70	1.68	1.42
23	4.44	4.79	4.76	4.76	4.59	4.40	1.61	1.99
24	3.98	5.54	5.98	6.06	5.93	5.85	3.58	3.47
25	7.09	7.90	8.63	8.62	7.76	7.72	7.41	7.41
26	3.97	7.31	5.37	5.37	5.56	4.43	0.42	1.52
27	12.07	9.72	9.27	9.30	8.63	8.64	9.41	9.41
28	5.80	6.86	6.23	6.42	6.73	6.56	6.14	6.40
29	7.20	5.62	6.09	6.09	5.53	5.61	5.12	5.92

Figure 3 shows the actual and predicted tumor volume values for the first six rats of each group. The forecasts of the LLIC and GPR models with the choice of GA features are presented.

Figure 4 shows the feature selection frequencies for TE, GPA, and linear regression models. The analysis of the obtained selection results made it possible to distinguish 13th and 25th features (tumor volumes on 9th and 15th days respectively). These features were selected for each model and in each run, resulting from their high information value for tumor volume prediction. The 19th feature (the tumor volume on the 12th day of the experiment) was selected just in two runs. So we can conclude that this feature has low information content. Some features are selected for synthesizing models of one type but are not selected when synthesizing another model. For instance, features with numbers 1 and 7 are the tumor volume on the third and sixth day of the experiment. These features were selected for the GPA and LR models, but never selected for the TE model. Feature 15 is the tumor volume on the 10th day of the experiment. It was selected only for the TE model. This feature selection behavior is explained by the nuances of the predictive model architecture. It confirms the thesis that no single feature set would be suitable for all types of models.

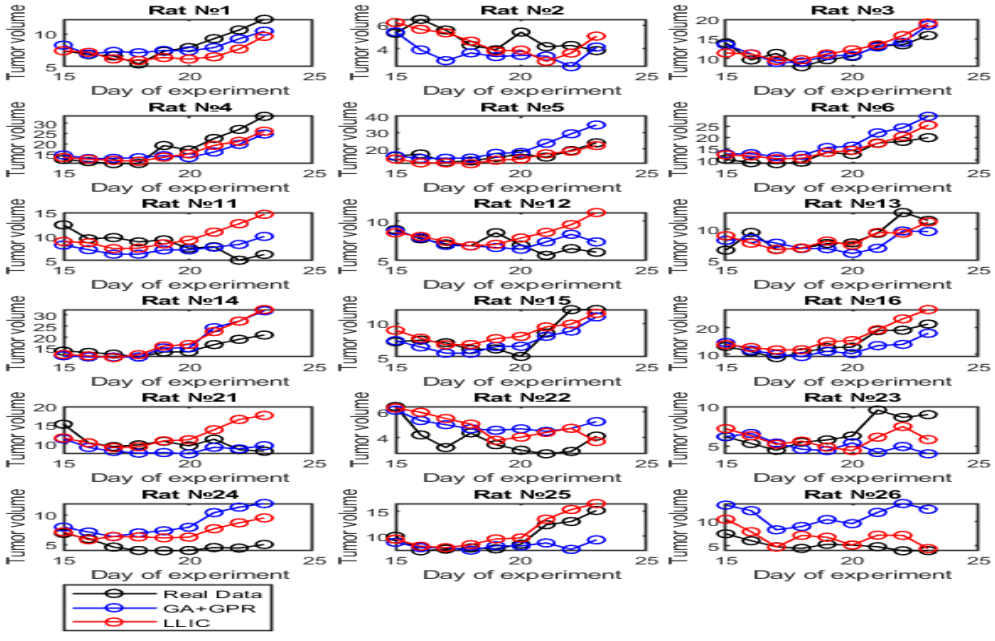

Figure 3. Actual and predicted tumor volume values.

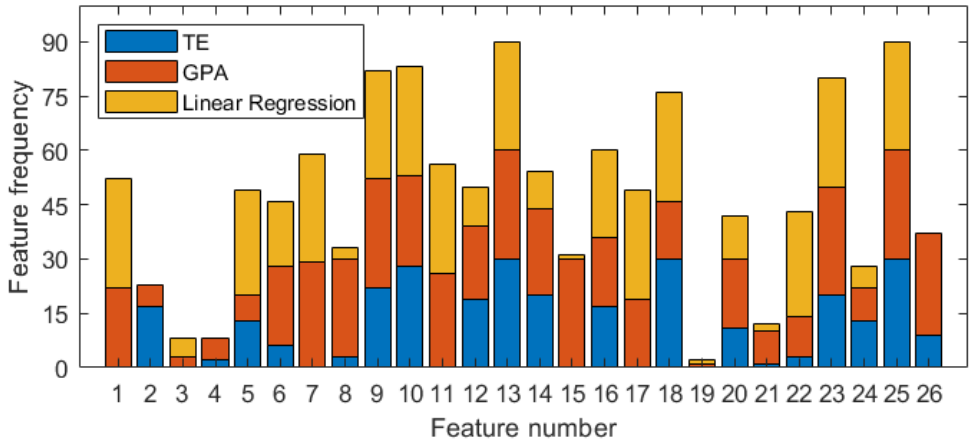

Figure 4. Frequencies of feature selection by the genetic algorithm.

Statistical Analysis

In order to formally assess the significance of the differences between real and predicted values, we use the paired Student's t-test [44]. As comparison values, we used remission R, remission depth D_R, and tumor volumes for each of the nine predicted days. Dividing the volume comparisons by day will indicate the models' predictive ability to determine the tumor volume for each day after termination of treatment.

Table 8. Comparison results of predicted and real values. The first part.

		GA+TE	TE	GA+GPR	GPR	GA+LR	LR	GrA+BRR	BRR
R	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.696	0.587	0.553	0.175	0.303	0.263	0.894	0.898
D_R	**Hypothesis**	Accepted	Accepted	Accepted	**Rejected**	**Rejected**	**Rejected**	Accepted	Accepted
	***p*-value**	0.088	0.688	0.138	0.000	0.000	0.001	0.196	0.144
V_1	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	**Rejected**	**Rejected**	Accepted	Accepted
	***p*-value**	0.934	0.494	0.934	0.138	0.000	0.000	0.742	0.934
V_2	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.967	0.647	0.967	0.077	0.546	0.567	0.825	0.887
V_3	**Hypothesis**	Accepted	Accepted	Accepted	**Rejected**	**Rejected**	Rejected	Accepted	Accepted
	***p*-value**	0.861	0.870	0.861	0.014	0.005	0.006	0.729	0.908
V_4	**Hypothesis**	Accepted	Accepted	Accepted	**Rejected**	**Rejected**	**Rejected**	Accepted	Accepted
	***p*-value**	0.995	0.858	0.995	0.019	0.000	0.001	0.685	0.763
V_5	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.763	0.582	0.763	0.301	0.261	0.380	0.920	0.937
V_6	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.397	0.371	0.397	0.069	0.308	0.426	0.920	0.933
V_7	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.613	0.606	0.613	0.552	0.120	0.101	0.968	0.958
V_8	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	**Rejected**	**Rejected**	Accepted	Accepted
	***p*-value**	0.633	0.689	0.633	0.623	0.049	0.040	0.944	0.890
V_9	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	**Rejected**	**Rejected**	Accepted	Accepted

Table 9. Comparison results of predicted and real values. The second part.

		ERTR	GrA+LLIC	LLIC	GrA+RR	RR	GrA+SVM	SVM
R	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.834	0.993	0.962	0.833	0.915	0.977	0.980
D_R	**Hypothesis**	**Rejected**	**Rejected**	**Rejected**	**Rejected**	Accepted	Accepted	Accepted
	***p*-value**	0.022	0.035	0.038	0.041	0.066	0.493	0.566
V_1	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.663	0.695	0.667	0.775	0.856	0.802	0.946
V_2	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.800	0.744	0.934	0.831	0.907	0.469	0.736
V_3	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.701	0.726	0.603	0.823	0.924	0.744	0.566

V_4	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.922	0.944	0.908	0.700	0.766	0.555	0.694
V_5	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.966	0.893	0.992	0.881	0.933	0.877	0.809
V_6	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.929	0.906	0.895	0.985	0.948	0.804	0.733
V_7	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.814	0.996	0.996	0.914	0.984	0.806	0.734
V_8	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.878	0.882	0.918	0.874	0.929	0.892	0.951
V_9	**Hypothesis**	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
	***p*-value**	0.943	0.936	0.959	0.853	0.912	0.957	0.924

The data for comparing *R*are in Table 4 and Table 5. The data for comparing D_R are in Table 6 and Table 7. For comparing tumor volumes, actual data from the observation dataset and the models' predictive values are used. The results are divided into groups related to one day (i.e., the same values of the characteristic x_{25}). The null hypothesis of the H_0 criterion shows the absence of statistically significant differences between the predicted and real values, the alternative H_1 shows about the presence of statistically significant differences. Table 8 and Table 9 show the acceptance of the null hypothesis at the significance level α = 0.05. Comparison of tumor volumes at 1,2, ..., 9 days after stopping treatment is designated as V_1,V_2...,V_9.

The above comparison shows that the predictive estimates of all characteristics (R, D_R, V_1,,V_9) turned out to be statistically indistinguishable from the real data for the GA + TE, TE, GA + GPR, GrA + BRR, BRR, RR, GrA + SVM and SVM. A further selection of a model from those listed above can be conducted using the Friedman test for linked samples [44]. The *p*-values in Table 8 and Table 9 are selected as comparison values. The null hypothesis of the H_0 test shows the absence of statistically significant differences between the *p*-values of the GA + TE, TE, GA + GPR, GrA + BRR, BRR, RR, GrA + SVM models, and SVM. Alternative hypothesis H_1 indicates the presence of statistically significant differences. At the significance level α = 0.05, the null hypothesis is accepted with a *p*-value = 0.074. Table 10 shows the obtained ranks, where the higher value corresponds to the more accurate model. Interpreting the result, we can talk about statistical indistinguishability in forecasting models, which means that any given model can be used for forecasting.

Table 10. Ranks of models in statistical comparison.

Model	Rank
GA+TE	3.82
TE	2.91
GA+GPR	3.73
GrA+BRR	4.55
BRR	5.45
RR	5.91
GrA+SVM	4.45
SVM	5.18

CONCLUSION

This study is devoted to the assessment of remission after termination of cancer treatment. The study included a comparison of different predictive models using different feature selection algorithms.

Observations of biological models (female rats) in the laboratory provided data for the models' training and validation. Animals were exposed to fast-growing carcinoma and were differentiated into three groups depending on the administered chemotherapy dosage in combination with controlled local hyperthermia. In observations of rats, data on tumor volumes were recorded daily.

Further data processing was conducted using time series. As a result, 26 component feature vectors were formed, which were used in training. Validation of predictive models and assessing their accuracy was performed using reserved data obtained in the same experiment. The projections and observations were compared in terms of such parameters as remission value and the depth of remission.

The computational experiment showed the statistical correspondence of the obtained forecasts with the estimates of the value and depth of remission observed during the experiment. This result allows stating that each of the models considered in the study can predict treatment effectiveness. For further comparative assessment of the accuracy of the models, they were ranked using the Friedman criterion. The use of Friedman's criterion has shown the advantage of regression models that do not use special algorithms for feature selection.

The results obtained can be used for predicting remission upon the treatment procedure's termination at some points, depending on the specific parameters of the biological model. This model makes it possible to draw up a patient-oriented treatment plan, determine the dose of chemotherapy drug and hyperthermic exposure duration. In the future, it is planned to continue research for other types of tissues and tumors and consider the potential possibility of switching from biological models to situations more like the treatment of real patients.

FUNDING

This research was funded by the Ministry of Education and Science of the Russian Federation within the framework of scientific projects carried out by teams of research

laboratories of educational institutions of higher education subordinate to the Ministry of Science and Higher Education of the Russian Federation, project number FEWM-2020-0042.

ACKNOWLEDGEMENT

The authors would like to thank Irkutsk Supercomputer Center of SB RAS for providing the access to HPC-cluster «Akademik V.M. Matrosov» [45].

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Jalalimanesh, A.; Haghighi, H.S.; Ahmadi, A.; Soltani, M. Simulation-based optimization of radiotherapy: Agent-based modeling and reinforcement learning. *Math. Comput. Simul.* **2017**, *133*, 235-248, doi:10.1016/j.matcom.2016.05.008.
2. Claret, L.; Girard, P.; Hoff, P.M.; Van Cutsem, E.; Zuideveld, K.P.; Jorga, K.; Fagerberg, J.; Bruno, R. Model-Based Prediction of Phase III Overall Survival in Colorectal Cancer on the Basis of Phase II Tumor Dynamics. *J. Clin. Oncol.* **2009**, *27*, 4103-4108, doi:10.1200/jco.2008.21.0807.
3. Feng, Y.; Wang, X.N.; Suryawanshi, S.; Bello, A.; Roy, A. Linking Tumor Growth Dynamics to Survival in Ipilimumab-Treated Patients with Advanced Melanoma Using Mixture Tumor Growth Dynamic Modeling. *Cpt-Pharmacometrics & Systems Pharmacology* **2019**, *8*, 825-834, doi:10.1002/psp4.12454.
4. Nagase, M.; Aksenov, S.; Yan, H.; Dunyak, J.; Al-Huniti, N. Modeling Tumor Growth and Treatment Resistance Dynamics Characterizes Different Response to Gefitinib or Chemotherapy in Non-Small Cell Lung Cancer. *Cpt-Pharmacometrics & Systems Pharmacology* **2020**, *9*, 143-152, doi:10.1002/psp4.12490.
5. Beigmohammadi, F.; Masoudi, A.A.; Khorrami, M.; Fatollahi, A.H. Mathematical modeling of tumor growth as a random process. *J. Theor. Appl. Phys.* **2020**, *14*, 245-249, doi:10.1007/s40094-020-00384-3.
6. Moazemi, S.; Khurshid, Z.; Erle, A.; Lutje, S.; Essler, M.; Schultz, T.; Bundschuh, R.A. Machine Learning Facilitates Hotspot Classification in PSMA-PET/CT with Nuclear Medicine Specialist Accuracy. *Diagnostics (Basel, Switzerland)* **2020**, *10*, 622, doi:10.3390/diagnostics10090622.
7. Libbrecht, M.; Noble, W. Machine learning applications in genetics and genomics. *Nat. Rev. Genet.* **2015**, *16*, 321–332, doi:10.1038/nrg3920.
8. Huang, S.; Cai, N.; Pacheco, P.P.; Narrandes, S.; Wang, Y.; Xu, W. Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. Cancer Genomics Proteomics **2018**, *15*, 41–51, doi:10.21873/cgp.20063.
9. Sinha, A.; Mathew, R. Machine Learning Algorithms for Early Prediction of Heart Disease. *Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2019), Madurai, IN, 19–20 December 2009*; Pandian, A., Palanisamy, R., Ntalianis, K.; Springer: Cham, Switzerland, **2020**; pp. 162–168, doi:10.1007/978-3-030-43192-1_18.
10. Mienye, D.I.; Sun, Y.; Wang, Z. Improved sparse autoencoder based artificial neural network approach for prediction of heart disease. *Inf. Med. Unlocked* **2020**, *18*, 100307, doi:10.1016/j.imu.2020.100307.
11. Sowmiya, C., Sumitra, P. A hybrid approach for mortality prediction for heart patients using ACO-HKNN. *J. Ambient. Intell. Human Comput.* **2020**, doi:10.1007/s12652-020-02027-6.
12. Chang, W.; Liu, Y.; Xiao, Y.; Yuan, X.; Xu, X.; Zhang, S.; Zhou, S. A Machine-Learning-Based Prediction Method for Hypertension Outcomes Based on Medical Data. *Diagnostics* **2019**, *9*, 178, doi:10.3390/diagnostics9040178.
13. Na, H.S.; Kim, S.Y.; Han, H.; Kim, H.-J.; Lee, J.-Y.; Lee, J.-H.; Chung, J. Identification of Potential Oral Microbial Biomarkers for the Diagnosis of Periodontitis. *J. Clin. Med.* **2020**, *9*, 1549, doi:doi.org/10.3390/jcm9051549.
14. Alam, M.; Le, D.; Lim, J.I.; Chan, R.V.; Yao, X. Supervised Machine Learning Based Multi-Task Artificial Intelligence Classification of Retinopathies. *J. Clin. Med*. **2019**, *8*, 872, doi:10.3390/jcm8060872.
15. Zhao, Z.; Zhang, Y.; Deng, Y. A Comprehensive Feature Analysis of the Fetal Heart Rate Signal for the Intelligent Assessment of Fetal State. *J. Clin. Med*. **2018**, *7*, 223, doi:10.3390/jcm7080223.
16. Hodashinsky, I.A.; Bardamova, I.B.; Bardamova, M.B. Complex Assessment of Coagulation Parameters in Pregnant Women Using a Fuzzy Classifier. *Biomed. Eng.* **2017**, *51*, 223–228, doi:10.1007/s10527-017-9719-2.
17. Cheng, F.-Y.; Joshi, H.; Tandon, P.; Freeman, R.; Reich, D.L.; Mazumdar, M.; Kohli-Seth, R.; Levin, M.A.; Timsina, P.; Kia, A. Using Machine Learning to Predict ICU Transfer in Hospitalized COVID-19 Patients. *J. Clin. Med*. **2020**, *9*, 1668, doi:10.3390/jcm9061668.
18. Tsai, T.-L.; Huang, M.-H.; Lee, C.-Y.; Lai, W.-W. Data Science for Extubation Prediction and Value of Information in Surgical Intensive Care Unit. *J. Clin. Med*. **2019**, *8*, 1709, doi:10.3390/jcm8101709.
19. Kwon, Y.S.; Baek, M.S. Development and Validation of a Quick Sepsis-Related Organ Failure Assessment-Based Machine-Learning Model for Mortality Prediction in Patients with Suspected Infection in the Emergency Department. *J. Clin. Med*. **2020**, *9*, 875, doi:10.3390/jcm9030875.
20. Perng, J.-W.; Kao, I.-H.; Kung, C.-T.; Hung, S.-C.; Lai, Y.-H.; Su, C.-M. Mortality Prediction of Septic Patients in the Emergency Department Based on Machine Learning. *J. Clin. Med*. **2019**, *8*, 1906, doi:10.3390/jcm8111906.
21. Attallah, O. An Effective Mental Stress State Detection and Evaluation System Using Minimum Number of Frontal Brain Electrodes. *Diagnostics* **2020**, *10*, 292, doi:10.3390/diagnostics10050292.
22. Cai, Y.-D.; Zhang, S.; Zhang, Y.-H.; Pan, X.; Feng, K.; Chen, L.; Huang, T.; Kong, X. Identification of the Gene Expression Rules That Define the Subtypes in Glioma. *J. Clin. Med.* **2018**, *7*, 350, doi:10.3390/jcm7100350.
23. Samhitha, B.K.; Mana, S.C.; Jose, J.; Vignesh, R; Deepa, D. Prediction of Lung Cancer Using Convolutional Neural Network (CNN) *Int. J. Adv. Trends Comput. Sci. Eng*. **2020**, *3*, 3361–3365, doi:10.30534/ijatcse/2020/135932020.
24. Rani, V.M.K.; Dhenakaran, S.S. Classification of ultrasound breast cancer tumor images using neural learning and predicting the tumor growth rate. *Multimed. Tools Appl.* **2020**, *23*, 16967–16985, doi:10.1007/s11042-019-7487-6.

25. Agrawal, S.; Tiwari, A.; Goel, I. (2020) Genetically Optimized Deep Neural Learning for Breast Cancer Prediction. In *Soft Computing for Problem Solving*; Nagar, A., Deep, K., Bansal, J., Das, K., Eds.; Springer: Singapore, 2019; pp. 127–139, doi:10.1007/978-981-15-3287-0_10.
26. Mirsanei, Z.; Habibi, S.; Kheshtchin, N.; Mirzaei, R.; Arab, S.; Zand, B.; Jadidi-Niaragh, F.; Safvati, A.; Sharif-Paghaleh, E.; Arabameri, A.; Asemani, D.; Hajati, J. Optimized Dose of Dendritic Cell-Based Vaccination in Experimental Model of Tumor Using Artificial Neural Network. *Iran. J. Allergy Asthma Immunol.* **2020**, *19*, 172–182, doi:10.18502/ijaai.v19i2.2770.
27. Al-Huniti, N.; Feng, Y.; Yu, J.J.; Lu, Z.; Nagase, M.; Zhou, D.S.; Sheng, J. Tumor Growth Dynamic Modeling in Oncology Drug Development and Regulatory Approval: Past, Present, and Future Opportunities. *Cpt-Pharmacometrics & Systems Pharmacology*, doi:10.1002/psp4.12542.
28. Hodashinsky, I.A.; Sarin, K.S. Feature Selection for Classification through Population Random Search with Memory. *Autom. Remote Control* **2019**, *80*, 324–333, doi:10.1134/s0005117919020103.
29. Ghareb, A.S.; Bakar, A.A.; Hamdan, A.R. Hybrid feature selection based on enhanced genetic algorithm for text categorization. *Expert Systems Appl.* **2016**, *49*, 31–47, doi:10.1016/j.eswa.2015.12.004.
30. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M. Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. *Energies* **2018**, *11*, 1636, doi:10.3390/en11071636.
31. Kabir, M. M.; Shahjahan, M.; Murase, K. A new local search-based hybrid genetic algorithm for feature selection. *Neurocomputing* **2011**, *74*, 2914–2928, doi:10.1016/j.neucom.2011.03.034.
32. Cormen, T.H.; Leiserson, C.E.; Rivest, R.L.; Stein, C. Greedy Algorithms. In *Introduction to Algorithms*, 2nd ed.; The MIT Press: Cambridge, UK, 2001; pp. 317 – 347, ISBN 0-262-03293-7.
33. Guyon, I.; Elisseeff, A. An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182, doi:10.1162/153244303322753616.
34. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*, 2nd ed.; Springer: New York, NY, USA, 2008; ISBN 978-0-387-84858-7.
35. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. *Classification and Regression Trees*, 1st ed.; Chapman & Hall\CRC: Boca Raton, FL, 1984; ISBN 9780412048418.
36. Breiman, L. Random Forests. *Machine Learning* **2001**, *45*, 5–32, doi:10.1023/a:1010933404324.
37. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*. The MIT Press: Cambridge, UK, 2006; ISBN 0-262-18253-X.
38. Zhu, C.; Byrd, R. H.; Lu, P.; Nocedal, J. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans. Math. Softw*. 1997, *23*, 550–560, doi:10.1145/279232.279236.
39. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Machine Learning* **2006**, *63*, 3–42, doi:10.1007/s10994-006-6226-1.
40. Rifkin, R.M.; Lippert, R.A. Notes on Regularized Least Squares. Massachusetts Institute of Technology, Tech. Rep. MIT-CSAIL-TR-2007- 025, 2007.
41. MacKay, D. J. C. Bayesian Interpolation. *Neural Comput*. 1992, *4*, 415–447, doi:10.1162/neco.1992.4.3.415.
42. Zou, H.; Hastie, T.; Tibshirani, R. On the degrees of freedom of the lasso. *Ann. Statist*. 2007, *35*, 2173–2192, doi:10.1214/009053607000000127.
43. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000, doi:10.1017/CBO9780511801389.
44. Glantz, S. A. *Primer of Biostatistics*, 7th ed; McGraw-Hill: New York, NY, USA, 2011; ISBN 9780071781503.
45. Irkutsk Supercomputer Center of SB RAS. Available online: http://hpc.icc.ru (accessed on 24 November 2020).