Neural Network Forecasting of ”Quality of the Day” for the Category of People Sensitive to Weather with Cardiovascular Diseases

Kabildjanov Aleksandr, and Pulatov Giyos

***Abstract*—With the development of neural networks and machine learning, intelligent forecasting is increasingly used in the problems of predicting the risk of occurrence and development of socially significant diseases, the first place among which is occupied by cardiovascular diseases (CVD). A large group is the category of weather-dependent people, for whom the risk of occurrence and development of CVD is correlated with meteorological factors and solar activity. The construction of effective predictive models of the ”quality of the day” for weather-dependent people allows to increase the efficiency of taking preventive measures and the quality of CVD treatment. In this research paper, the issues of using recurrent neural networks (RNN) of various types for forecasting multivariate time series in relation to the problem of predicting the ”quality of the day” for weather-dependent people for CVD are studied. The frequency of admission of patients with CVD to medical institutions served as an evaluation parameter for determining the ”quality of the day”. The input parameters for forecasting were ambient temperature, air humidity, atmospheric pressure, geomagnetic activity index and the rate of their daily change. Forecasting was performed using the sliding time window method using MLP, RBF and LSTM neural networks. The ”quality of the day” was assessed based on the Mamdani fuzzy model. Analysis of the obtained results shows that the accuracy of RNN forecasting is not stable, both in time and in many different types of neural networks. This leads to the conclusion that it is advisable to use ensemble modeling in problems of forecasting the risk of occurrence and development of CVD.**

***Index Terms*—forecasting, meteorological factors, heliogeophysical factors, cardio-vascular diseases, MLP, RBF and LSTM neural networks, neural network architecture.**

I. Introduction

According to the World Health Organization, humanity loses about 2.5 million people annually from cardiovascular diseases (CVD), with more than one third of them being people of working age.

Most people predisposed to or suffering from CVD belong to a fairly large group of weather-dependent people. Their well-being, occurrence, exacerbation and course of the disease are affected by a large group of external factors: meteorological, solar activity, environmental, etc. In this regard, one of the urgent tasks facing the healthcare of various countries is the timely forecasting of CVD depending on many meteorological and heliogeophysical factors. A correct and timely forecast of the frequency of admission of patients with CVD allows solving several problems: preventing the occurrence or exacerbation of CVD by conducting proactive preventive measures; optimizing the organization of the emergency medical care service; increasing the efficiency of medical institutions. In addition, the frequency of admission of patients with CVD can serve as an assessment indicator of the ”quality of the day” for a wide category of people depending on the weather.

In practice, a number of methods are widely used to solve problems of forecasting complex processes: factual, expert, publication, scenario, matrix, mathematical modeling, analogy method, graph, etc. In medicine, the most traditional methods of forecasting diseases until recently remained statistical modeling methods.

The listed forecasting methods have one common drawback - they are ineffective in conditions of weak formalization of the problem, incompleteness and ambiguity of information. Therefore, recently, when forecasting complex processes, greater preference has been given to machine learning methods and deep neural networks [1, 2]. Research by many scientists shows that the accuracy of forecasting CVD using machine learning algorithms such as random forest [3], logistic regression [4, 5], gradient boosting [6] and neural networks [7, 8] significantly exceeds the accuracy of doctors' prediction. Recurrent deep learning neural networks have proven themselves especially well, which are successfully used in forecasting dynamic processes described by multidimensional time series [9].

This article is devoted to solving the problem of creating effective neural network models of various classes that provide high accuracy in predicting the frequency of admission of patients with cardiovascular diseases to medical institutions, and, on their basis, developing production models for assessing the “quality of the day” for weather-sensitive people.

The scientific novelty of the study, the results of which are presented in the article, lies in the development and analysis of new recurrent neural network prognostic models that take into account not only the complex influence of meteorological and heliogeophysical factors, but also the rate of their change with the history of the process.

The paper is structured as follows: this section serves as

Manuscript received July 24, 2025; revised August 28, 2025. Date of publication December 18, 2025. Date of current version December 18, 2025. The associate editor prof. Ivo Stančić has been coordinating the review of this manuscript and approved it for publication.

Authors are with the Tashkent Institute of Irrigation and Agricultural Mechanization Engineers, National Research University, Tashkent, Uzbekistan (e-mails: kabildjanov@yandex.com, giyospulatov1987@gmail.com).

Digital Object Identifier (DOI): 10.24138/jcomss-2025-0025

an introduction, and Section II reviews a number of scientific studies devoted to studying the relationship between meteorological factors and cardiovascular diseases. A comparative table with the analyzed works and the studies conducted in this research is presented. Section III justifies the need to build a model using neural networks to predict cardiovascular diseases. Section IV presents a step-by-step process of training a model based on MLP and RBF to solve the forecasting problem. The final section, Section V, is a conclusion that summarizes the results of the study, main findings, and scientific recommendations for future research.

Recently, deep neural networks [10-11] and machine learning models [12-15] have attracted great interest, showing their high efficiency in relation to time series forecasting tasks.

II. Comparative Analysis of Existing Works on the Problem being solved

Numerous scientific studies have been devoted to the development and application of neural network models in the tasks of diagnosing, predicting the occurrence and progression of cardiovascular diseases, the results of which are most fully reflected in a number of works by foreign authors.

For example, [16] investigated the prediction of peak days for hospitalization of patients with cardiovascular diseases using air quality and meteorological data in Chengdu, China, for the period from 2015 to 2017 using an artificial neural network and machine learning methods: random forest and extreme gradient boosting.

The paper [17] presents the results of a study to establish a functional relationship between the frequency of admission of patients with CVD, air quality and meteorological data for Hainan Province, China, for the period from 2016 to 2018. The relationship between the factors under study was formalized using a distributed nonlinear model (DLNM).

The study [18] is devoted to the study of forecasting climate-dependent diseases developing as a result of heat or cold stress. Using the LightGBM machine learning algorithm, a model was developed for predicting mortality from cardiovascular diseases, such as myocardial infarction and cerebral infarction, in Tokyo and Osaka (Japan).

In [20], four predictive models were tested, implementing the following methods: linear regression, weighted moving averages, multilayer perceptron (MLP), and support vector regression (SVR). The models used meteorological data for Spain and Chile as input. The results of research studies related to the application of various machine learning methods, such as logistic regression, support vector machine (SVM), random forest, XGBoost, Light-GBM, NGBoost, stacking, and imbalance handling methods, were also analyzed.

The study [21] examines the prediction of the incidence of aneurysmal subarachnoid hemorrhage depending on weather and climate changes. The predictive model in the study is based on data from 2003 to 2020 for Germany and includes 13 different weather and climate parameters. The predictive model is based on deep learning neural networks (LSTM) and CNN.

One of the shortcomings of the studies, the results of which are presented in Table I, is the incomplete representativeness of the many factors influencing CVD and the lack of summary estimates of the frequency of admission of patients with various CVDs. The study [19] demonstrated that only a complete consideration of climatic factors and environmental conditions allows for the effective prediction of socially significant diseases such as CVD, type II diabetes, etc.

In addition, most studies do not take into account the dynamics of changes in the frequency of admission of patients with CVD.

III. Problem Statement

An analysis of the studies, the results of which are presented in the work [22], showed that the frequency of admission of $y(t)$ patients with CVD to medical institutions at each i-th moment in time depends not only on a multitude of meteorological and heliogeophysical factors, but is also correlated with the eigenvalues that occur at previous moments in time $y(t_{i-1}), y(t_{i-2}), \ldots, y(t_{i-k})$

Let the prediction of the values $y^*(t_{i+1})$ be carried out using a recurrent neural network according to an algorithm that implements a multi-step sliding “time window” method:

$$\begin{aligned} y^*(t_{i+1}) &= f\Big(y(t_i),\, y(t_{i-1}),\, \ldots,\, y(t_{i-k}),\\ &\qquad x(t_i),\, x(t_{i-1}),\, \ldots,\, x(t_{i-k})\Big);\\ y^*(t_{i+2}) &= f\Big(y^*(t_{i+1}),\, y(t_i),\, \ldots,\, y(t_{i-k+1}),\\ &\qquad x(t_{i+1}),\, x(t_i),\, \ldots,\, x(t_{i-k+1})\Big);\\ &\ \vdots\\ y^*(t_{i+g}) &= f\Big(y^*(t_{i+g-1}),\, y(t_{i+g-2}),\, \ldots,\, y(t_{i+g-1-k}),\\ &\qquad x(t_{i+g-1}),\, x(t_{i+g-2}),\, \ldots,\, x(t_{i+g-1-k})\Big). \end{aligned} \tag{1}$$

where

$$x(t_i) = \big(x_1(t_i),\, x_2(t_i),\, x_3(t_i),\, x_4(t_i),\, x_5(t_i),\, x_6(t_i),\\ x_7(t_i),\, x_8(t_i),\, x_9(t_i),\, x_{10}(t_i)\big).$$

is the vector of values of input parameters at time t_i; x_1 is the ambient temperature, x_2 is the atmospheric pressure, x_3 is the air humidity, x_4 is the wind speed, x_5 is the geomagnetic activity; $x_6 \div x_{10}$ are the daily rates of change of input parameters $x_1 \div x_5$; k is the length of the sliding “time window”; $y^*(t_{i+g})$ is the predicted value of the output parameter for time t_{i+g}.

This method of predicting the value $y^*(t_{i+1})$ allows making only short and medium-term forecasts, since the accumulation of errors at each forecasting step has a significant impact on accuracy. At the same time, one of the starting points influencing the accuracy of the forecast is the choice of the optimal length of the sliding ”time window”.

The second essential point is the correct choice of type, the determination of the optimal architecture, and high-quality training of the recurrent neural network [23].

TABLE I
A COMPARATIVE STUDY OF WEATHER–CVD ASSOCIATIONS BETWEEN THE RESEARCH PRESENTED HERE AND THE REVIEWED LITERATURE IS CONDUCTED

№	*Derivation*	*Temperature*	*Humidity*	*Pressure*	*Magnetic Storm*	*Wind*	*Ischemia*	*Artemia*	*Hypertension*
1	In this article	+	+	+	+	+	+	+	+
2	[16]	+	+	+			+		
3	[17]	+		+			+		
4	[18]	+	+					+	
5	[19]	+			+			+	
6	[20]	+	+						+
7	[21]	+		+				+	

The formation of an assessment of the “quality of the day” for the category of weather-sensitive people can be carried out on the basis of production rules:

$$\text{if } \alpha_{in} < y^*(t) \leq \alpha_{iv} \text{ then } \beta_i; \quad i = 1, 5 \tag{2}$$

where α_{in}, α_{iv} are the expert values of the lower and upper limits for the i-th assessment of the “quality of the day”; β_i is a linguistic variable with basic values: “very bad”, “bad”, “average”, “good”, “very good”.

IV. PROBLEM SOLUTION

When solving the forecast problem, at the first stage, the optimal length of the sliding ”time window” was selected based on the available a priori information. When selecting the optimal length of the sliding ”time window”, the technique based in the work [24] was used.

At the first step, using the initial data shown in Fig. 1, a graph of the auto-correlation function of the predicted parameter was constructed y(t), the values of which were calculated using the formula:

$$R_y(\tau) = \frac{1}{N-\mu} \sum_{i=1}^{N-\mu} \left[y_i - m_y\right]\left[y_{i+\mu} - m_y\right] \tag{3}$$

where

$$m_y = \frac{1}{N} \sum_{i=1}^{N} y_i$$

is the estimate of the mathematical expectation of the output parameter; $\tau = \dfrac{\mu T}{N}$; T is the observation time; $\mu = 0, 1, 2, \ldots$; μ^* is the number of shifts; $\mu^* = (0.25 - 0.35)\, N$.

The optimal value of the length of the sliding “time window” was selected based on the graph of the autocorrelation function of the predicted parameter, shown in Fig. 2. It can be seen from the graph that for $\tau \geq 5$, the autocorrelation function of the predicted parameter is located under a straight line parallel to the abscissa axis, corresponding to the value $0,05\, R_y(0)$. In this case, with sufficient accuracy for practical calculations, its value can be taken equal to zero.

Thus, it can be assumed that when the condition

$$R_y(\tau) < 0,05\, R_y(0) \quad \forall \tau : \; \tau > 5,$$

is met, there is no correlation between the values of the predicted parameter at times τ and $\tau + 5$, and the length of the sliding “time window” is equal to five days.

When solving the problem of long-term forecasting, it was assumed that the values of the output parameter $y(t_{i+l})$ are correlated only with the values of the input parameters $x(t_i), x(t_{i-1}), \ldots, x(t_{i-k})$.

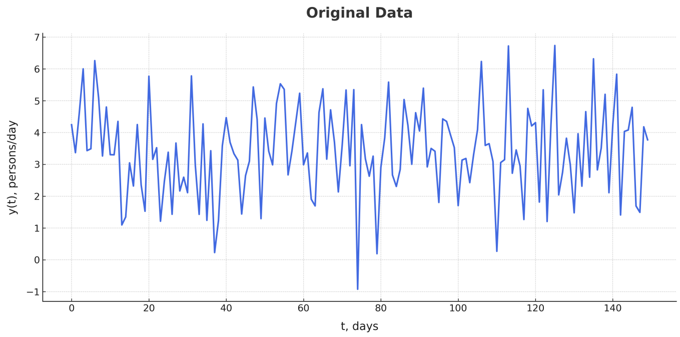

Fig. 1. Values of the frequency of admission of patients with CVD to medical institutions by day

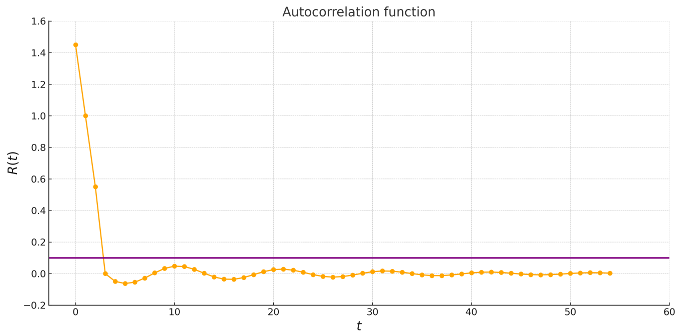

Fig. 2. Graph of the autocorrelation function of the output parameters

Forecasting was carried out according to the following algorithm:

$$y^*(t_{i+l}) = f\big(x(t_i),\, x(t_{i-1}),\, \ldots,\, x(t_{i-k})\big) \tag{4}$$

where l is the discrete step at which the forecast is made.

When choosing the type of neural network for short- and medium-term forecasting, The forecasting problems were addressed based on the existing experience systematized in the works [25-27]. The most frequently used for solving forecasting problems in practice are the multilayer perceptron (MLP) and the radial basis function network (RBF).

Construction of recurrent neural networks for short and medium-term forecasting of values was performed in the neuro package Statistics Neural Networks on a computer with a 12th Gen Intel(R) Core(TM) i9-12900K 3.20 GHz processor and 32 GB of RAM.

TABLE II
ST NEURAL NETWORKS RAW DATA.

Num	*VAR1*	*VAR2*	*VAR3*	*VAR4*	*VAR5*	*VAR6*	*VAR7*	*VAR8*	*VAR9*	*VAR10*	*VAR11*
1	22.3	697.8	37	2.3	2.8	0.8	-0.3	0.6	0.7	0.5	5.8
2	23.4	698.4	38.2	2.3	2.5	1.1	0.6	1.2	0.0	-0.3	4.1
3	22.6	698.7	37.6	1.6	2.0	-0.8	0.3	-0.6	-0.7	-0.5	4.6
4	22.9	698.2	32.6	2.5	1.7	0.3	-0.5	-5.0	0.9	-0.3	2.6
5	25.0	697.4	25.0	4.2	0.9	2.1	-0.8	-7.6	1.7	-0.8	3.4
6	25.2	697.8	28.2	4.0	1.7	0.2	0.4	3.2	-0.2	0.8	4.8
7	24.5	698.1	29.3	3.4	1.0	-0.7	0.3	1.1	-0.6	-0.7	6.4
8	25.4	697.4	28.0	2.9	1.7	0.9	-0.7	-1.3	-0.5	0.7	4.0
9	26.0	698.2	28.1	2.9	1.0	0.6	0.8	0.1	0.0	-0.7	4.8
10	26.7	697.1	28.8	2.5	1.9	0.7	-1.1	0.7	-0.4	0.9	4.0
11	24.5	698.1	36.4	3.5	1.7	-2.2	1.0	7.6	1.0	-0.2	4.6
12	21.5	700.3	33.0	4.5	2.0	-3.0	2.2	-3.4	1.0	0.3	2.8
...	...	...	...	...	...	...	...	...	...	...	...
154	-8.9	700.9	88.5	3.2	1.2	-7.8	1.9	-1	-0.8	-0.2	8.8

Algorithm 1 Proposed Neural Network Forecasting Procedure (MLP, RBF, LSTM)

Input: Weather parameters and daily change rates, patient admission data
Output: Predicted "Quality of the Day" for cardiovascular patients
Step 1: Data Preparation
Normalize input variables, handle missing values
Split dataset into training, validation, and testing subsets
Step 2: Model Selection
for each neural network $M \in \{MLP, RBF, LSTM\}$ Initialize parameters θ^M, learning rate η, epochs T
endfor
Step 3: Training Phase
for each epoch $t = 1 \dots T$
  for each model M
    Forward propagate inputs through M
    Compute prediction $\hat{y}_t^M$
    Calculate loss $L(y_t, \hat{y}_t^M)$
    Update $\theta^M \leftarrow \theta^M - \eta \nabla_\theta L$
  endfor
endfor
Step 4: Evaluation
for each model M
  Evaluate on test set
  Compute metrics (MSE, MAE, R^2)
endfor
Step 5: Decision
Select best model M^* based on evaluation
Determine "Quality of the Day" using fuzzy rules with M^* predictions
return Predicted "Quality of the Day" =0

The first stage involved input and preparation of initial data. Input of initial data was carried out from a previously prepared file with the extension sta.

The initial data were taken from the official source (Samarkand branch of the Uzbek Republican Scientific Center for Emergency Medical Care) and Internet sites: https://pogoda1.ru/samarkand-2/arkhiv/, https://world-weather.ru/archive/uzbekistan/samarkand/ (meteorological parameters for the city of Samarkand), http://www.theusner.eu/terra/aurora/kparchive.php?year=2013 (solar activity parameters).

The original sample contained 154 rows of data. Table II shows the input and output parameters of the predictive model:

$$\text{VAR1} \leftrightarrow x_1, \quad \dots, \quad \text{VAR10} \leftrightarrow x_{10}, \quad \text{VAR11} \leftrightarrow y.$$

Data from the Samarkand branch of the Republican Scientific Center for Emergency Medicine were obtained during a project commissioned by the Ministry of Health of the Republic of Uzbekistan. The data were presented in a table, with general statistics, and did not include personalized patient data. Therefore, no ethical issues were identified during the study.

Input parameters `VAR1`–`VAR10` were marked with the identifier `Input`, and the output parameter `VAR11` was marked with the identifier `Input / Output` (see Tab.2). The variable type "input–output" was then set as follows: on the header of the variables `VAR1`–`VAR10` in the opened table, the `Input` option was selected from the menu by pressing the right mouse button. Similarly, on the header of the variable `VAR11`, the `Input / Output` option was selected from the menu by pressing the right mouse button. In this case, the name of the output variable was highlighted in blue.

In the second stage, using the automatic network designer (*Intelligent Problem Solver*), the optimal architectures of two classes of networks (MLP and RBF) were selected. In this case, the optimization problem was reduced to choosing the minimum number of layers of a given type of neural network that ensures the required accuracy in predicting $y^*(t_{i+1})$.

As a result of the automatic network designer, a 10-layer MLP network was selected for the full data set. The static data for this network, obtained as a result of its training, are given in Table III.

Due to the fact that the automatic network designer defined too many layers, a manual MLP construction was carried out at the *third stage*. When constructing a recurrent multilayer perceptron, the number of network layers was selected empirically, ranging from 3 to 6.

The neural network was trained using the Levenberg-Marquardt method, which is one of the reliable and fast algorithms for training small-sized networks with one input element and a mean square error function [28].

TABLE III
STATIC DATA AFTER MLP TRAINING.

Metric	*Tr. VAR11*	*Ve. VAR11*	*Te. VAR11*
Data Mean	3.658442	4.059459	3.589744
Data S.D.	1.126633	1.272674	1.257356
Error Mean	0.004367	-0.332	-0.006842
Error S.D.	1.053860	1.139760	1.159129
Abs E. Mean	0.8164054	0.9248384	0.9617875
S.D. Ratio	0.9354066	0.8955627	0.9218777
Correlation	0.3621597	0.4599615	0.3874809

TABLE IV
STATIC DATA AFTER RBF TRAINING.

Metric	*Tr. VAR11*	*Ve. VAR11*	*Te. VAR11*
Data Mean	3.630263	3.815789	3.871795
Data S.D.	1.100184	1.455191	1.136465
Error Mean	-2.98e-16	-0.2478	-0.2636
Error S.D.	1.041969	1.458795	1.059709
Abs E. Mean	0.8219156	1.056192	0.8805509
S.D. Ratio	0.9470857	1.002477	0.9324614
Correlation	0.3209808	0.1394616	0.3636402

The construction of MLP for each number of layers was performed in several steps.

Step 1. In the initial data window "Data Set Editor," the volumes of training (*Training*), control (*Verification*), and test (*Test*) observations were specified. Afterward, the rows were shuffled by using the following sequence of commands: *Edit* → *Cases* → *Shuffle* → *All*. In this study, the final partition was Training = 77, Verification = 38, and Test = 39.

Step 2. Create a new network by successively clicking *File* → *New* → *Network*. In the dialog box that opens, select *Multilayer* as the network type *Perceptron (Multilayer perceptron)*. For this task, the *Steps* parameter was set to 5, and the *Lookahead* parameter was set to 1. Afterward, the *Advice* button was pressed, causing the program to automatically assign the number of neurons in all network layers. Upon pressing the *Create* button, a neural network with the specified architecture was created.

The network was then trained using the Levenberg–Marquardt method, by executing the following sequence of commands: *Train* → *Multilayer Perceptron* → *Levenberg–Marquardt*.

Step 3. The number of epochs for running the algorithm was specified (*Epochs*), and cross-verification was set up (*Cross–Verification*). Training was then performed by pressing the *Train* button. Before starting a new training session, the network weights were reset via the *Reinitialize* button. If the algorithm became stuck in a local minimum, the network weights were "shaken" via the *Jog* button (*Weights*).

Step 4. Go to the *Statistics* menu, select the *Regression* option, and press the *Run* command.

An analysis of the obtained MLP network architectures and their statistical characteristics allowed us to settle on a neural network with a number of layers equal to 4 (see Tab. 5). The reason for this was that it has a relatively average forecasting error, the best correlation coefficient between the actual and predicted values of Correlation and a relatively short training time (10 min. 15 sec.).

TABLE V
STATIC DATA AFTER MANUALLY DESIGNED MLP TRAINING.

Metric	*Tr. VAR11*	*Ve. VAR11*	*Te. VAR11*
Data Mean	3.8	3.718919	3.633333
Data S.D.	1.127593	1.123941	1.475514
Error Mean	-3.068e-16	0.2053179	0.1471995
Error S.D.	1.776e-15	1.034865	1.221717
Abs E. Mean	1.399e-15	0.8803787	1.015176
S.D. Ratio	1.575e-15	0.9207465	0.8279939
Correlation	1	0.5749506	0.5981284

In red in the table of statistical characteristics, and the results obtained for the test set are marked in blue. The first column of the table corresponds to the statistics for the training set.

The sequence of manual construction of an RBF neural network is reduced to the following steps.

Step 1. In the initial data window "Data Set Editor," the volumes of training (*Training*), control (*Verification*), and test (*Test*) observations were specified. The rows were then shuffled by choosing *Edit* → *Cases* → *Shuffle* → *All*. In our case, the final partition was Training = 77, Verification = 38, and Test = 39.

Step 2. Create a new network by selecting *File* → *New* → *Network*. In the dialog box that appears, choose *Radial* as the network type (*Basis Function*). For this task, the *Steps* parameter was set to 5, and the *Lookahead* parameter was set to 1.

Step 3. Next, press the *Advice* button, causing the software to automatically determine the number of neurons in all layers. After pressing the *Create* button, a neural network of the specified architecture was created. The resulting architecture is shown in Table 6.

Step 4. Train the RBF network from the *Train* menu. By pressing the *K–Means* button, The location of the centers of the radial elements is specified at the cluster centers.

Step 5. By clicking the *Isotropic* button, The rule by which the deviations are selected is defined, taking into account both the number of elements and the spread of the training data.

Step 6. By clicking the *Pseudo–inverse* button (*Pseudo–Inverse*), It is specified that the parameters of the linear output layer of the network will be optimized.

Step 7. Go to the *Statistics* menu, select the *Regression* option, and press the *Run* command.

The resulting RBF network architecture and statistical characteristics of the regression are shown in Table 6.

TABLE VI
STATIC DATA AFTER MANUALLY DESIGNED MLP TRAINING.

Metric	*Tr. VAR11*	*Ve. VAR11*	*Te. VAR11*
Data Mean	3.675676	3.694444	3.902564
Data S.D.	1.172144	1.088628	1.400655
Error Mean	-1.26e-16	-0.1156	-0.3027
Error S.D.	1.101101	1.05403	1.407583
Abs E. Mean	0.8431257	0.8594571	1.019833
S.D. Ratio	0.9393908	0.9682188	1.004946
Correlation	0.3428483	0.2883011	0.01021

The calculations performed showed that the RBF approach has one key advantage: it can be trained very quickly. In this problem, its regression quality is comparable to that of a multilayer perceptron. However, a drawback of this method is that the pseudo-inverse algorithm (*Pseudo–Inverse*) can yield large errors when the radial deviations are too small. Therefore, the *Isotropic* and *K–Means* methods should be used with caution.

The final stage of the computational experiment involved varying the prediction step *Lookahead* from 1 to 10. The results of these experiments are shown in Fig. 3–4.

From Fig. 3–4, it is clear that the mean square forecasting error (*Error Mean*, shown in blue) increases nonlinearly as the prediction step *Lookahead* increases, while the correlation

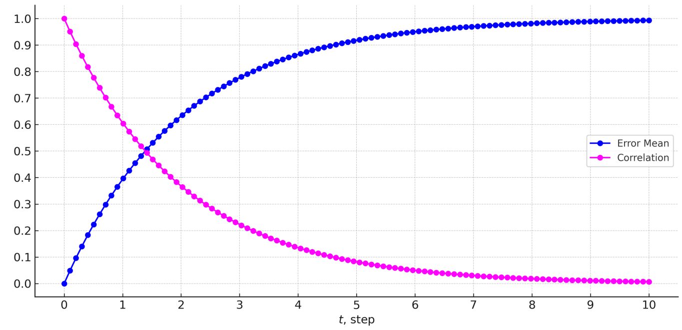

Fig. 3. Graphs of the results of the computational experiment with a recurrent four-layer multiperceptron

coefficient between actual and predicted values (*Correlation*, shown in crimson) decreases.

The 4th figure presents the main stages of constructing MLP and RBF models. It sequentially illustrates all stages of the process, including data preparation, training, forecasting, and others.

The forecasting problem was also solved using the LSTM deep learning neural network, which was implemented using the Python 3.11.5 programming language and its *Keras* library.

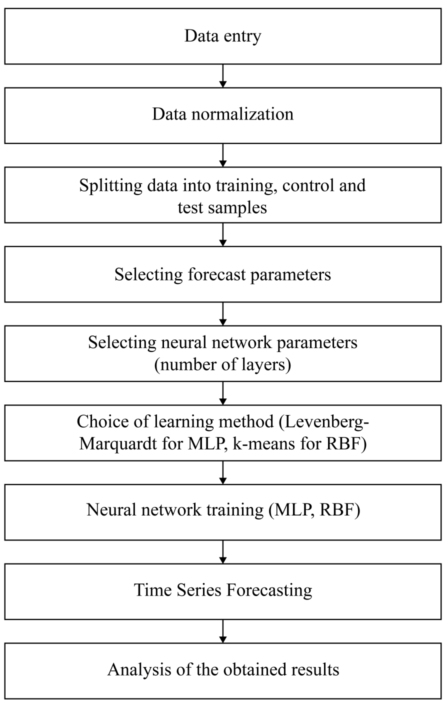

Fig. 4. Methodology flowchart for building MLP and RBF models

LSTM model architecture was used. First layer: 50-neuron LSTM (with ReLU activation function); Dropout = 0.2; Second layer: 50-neuron LSTM; Dropout = 0.2; Output layer: 1 Dense neural layer (dense layer for interpreting predictions). The LSTM deep neural network was trained using the adam method. To prevent gradient explosion, the gradient threshold was set to 1. The following training parameters were set:

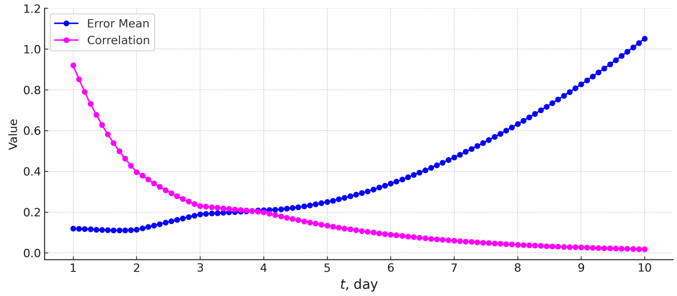

Fig. 5. Graphs of the results of the computational experiment with the RBF recurrent neural network

number of epochs - 50; number of training series - 32. The initial learning rate was set to 0.005. The learning rate was decreased after 25 epochs by multiplying by a factor of 0.2

The computational experiment was performed on the same dataset used with the MLP and RBF networks. For short and medium-term forecasting ($t < 2$), the average standard deviation obtained was 0.346. By contrast, for long-term forecasting ($t > 5$), the average MSE (which also represents the standard deviation) exceeded 3.956 when the forecast horizon varied in the interval $5 < t < 10$.

Numerous computational experiments performed with constructed MLP, RBF and LSTM recurrent neural networks on real and predicted data allowed us to obtain the following picture of the spread of MSE values over adjacent prediction intervals (Fig. 5).

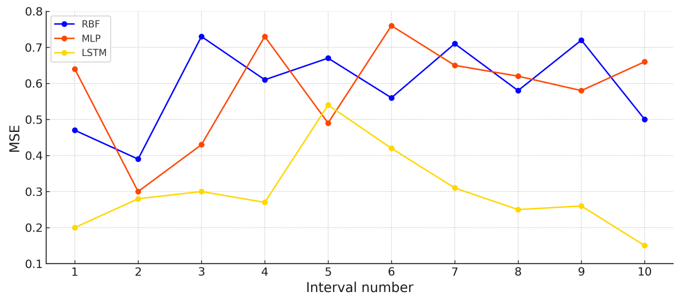

Fig. 6. Graphs of MSE change for different recurrent neural networks

V. Conclusion

The conducted research and development allowed us to obtain three types of neural networks (MLP, RBF and LSTM) designed to predict the frequency of admission of patients with CVD. The constructed neural networks were trained and tested on real data taken from official sources and Internet sites (Samarkand branch of the Uzbek Republican Scientific Center for Emergency Medical Care, https://pogoda1.ru/samarkand-2/arkhiv/, https://world-weather.ru/archive/uzbekistan/samarkand/, http://www.theusner.eu/terra/aurora/kparchive.php?year=2013). The analysis of the results of computational experiments on short-, medium- and long-term forecasting of the "quality of the day" for the category of weather-dependent people according to cardiovascular diseases, conducted with the constructed MLP, RBF and LSTM recurrent neural networks, allows us to draw the following conclusions. The constructed recurrent neural networks provide acceptable accuracy in forecasting patient admission rates for short- and medium-term horizons, where

$$\left|y^*(t_i) - y(t_i)\right| < 1, \quad i = 1, \ldots, 10.$$

Therefore, based on rule (2), the assessment of the ”quality of the day” for the category of weather-sensitive people will be informative. In the long-term forecast, the neural networks listed above have a fairly large forecast error, which ensures that only approximate assessments of the ”quality of the day” for the category of weather-sensitive people are obtained.

The accuracy of forecasting by different neural networks in different time intervals is not the same. At the same time, there is no uniform tendency of change in forecasting accuracy in different time intervals. Moreover, from Fig. 5 it is clear that in some time intervals one neural network gives a more accurate forecast, and in others - others.

Thus, the results obtained by the authors of the article allow us to conclude that it is appropriate to use ensemble methods in the tasks of forecasting the ”Quality of the Day” indicator for weather-sensitive people. The recurrent neural network models MLT, RBF and LSTM developed by the authors can be used as basic algorithms in ensemble methods for forecasting the frequency of admission of patients with CVD to medical institutions.

References

[1] N. C. Egejuru, O. Ogunlade and P. A. Idowu. “Adaptive Neuro-Fuzzy Inference Model for Monitoring Hypertension Risk”. International Journal of Healthcare Information Systems and Informatics. vol. 16, no. 4, pp. 1–32, Jan. 2021, doi: 10.4018/IJHISI.295818.

[2] A. S. Kabildjanov, G. G. Pulatov and G. A. Pulatova. “Prediction of the influence of weather conditions on heart blood pressure diseases based on the LSTM model”. Descendants of Al-Farghani electronic scientific journal. vol. 1, no. 4, pp. 173–180, Jun. 2024.

[3] M. S. Azam, M. A. Raihan and H. K. Rana. “An Experimental Study of Various Machine Learning Approaches in Heart Disease Prediction”. International Journal of Computer Applications. vol. 175, no. 21, pp. 16–26, Sep. 2020, doi: 10.5120/ijca2020920741.

[4] A. S. Kabildjanov, G. G. Pulatov and G. A. Pulatova. “Analytical analysis of the influence of weather conditions on cardiovascular diseases”. Descendants of Al-Farghani electronic scientific journal. vol. 1, no. 2, pp. 296–303, Jun. 2024.

[5] R. Navares and J. L. Aznarte. “Deep learning architecture to predict daily hospital admissions”. Neural Computing and Applications. vol. 32, pp. 16235–16244, 2020, doi: 10.1007/s00521-020-04840-8.

[6] A. S. Kabildjanov, G. G. Pulatov and G. A. Pulatova. “Forecasting and optimization of the impact of weather conditions on blood pressure-related diseases using the MLP model”. Descendants of Al-Farghani electronic scientific journal. vol. 1, no. 4, pp. 251–260, Jun. 2024.

[7] S.Subramani, N.Varshney, M. V. Anand, E. M. Soudagar, L. A. Al-keridis, T. K. Upadhyay, N. Alshammari, M. Saeed, K. Subramanian, K. Anbarasu and K. Rohini. “Cardiovascular diseases prediction by machine learning incorporation with deep learning”. Frontiers in Medicine, vol. 10, p. 1150933, Apr. 2023, doi: 10.3389/fmed.2023.1150933.

[8] V.K.Sudha and D.Kumar. “Hybrid CNN and LSTM Network For Heart Disease Prediction”. SN Computer Science. vol. 4, article 172, Jan. 2023, doi: 10.1007/s42979-022-01598-9.

[9] Y. Sakka, D. Qarashai and A. Altarawneh. “Predicting Hypertension using Machine Learning: A Case Study at Petra University”. International Journal of Advanced Computer Science and Applications. vol. 14, no. 3, p. 586, 2023, doi: 10.14569/IJACSA.2023.0140368.

[10] S. V. Kozlov and S. A. Sedenkov, “Analysis of LSTM and GRU Models for Time Series Forecasting,” International Journal of Open Information Technologies, vol. 12, no. 7, pp. 43–50, 2024.

[11] T. V. Oblakova and D. S. Alekseev, “Investigation of Time Series Forecasting Quality Using C++ Implementation of AT-LSTM Model,” Mathematical Modeling and Numerical Methods, vol. 45, no. 1, pp. 80–91, 2025, doi: 10.18698/2309-3684-2025-1-8091.

[12] G. Bontempi, “Local Learning Techniques for Modeling, Prediction and Control,” Ph.D. dissertation, IRIDIA-Université de Libre de Bruxelles, Belgium, 1999, doi: 10.1016/j.ijforecast.2010.09.004

[13] D. C. Montgomery, E. A. Peck, and G. G. Vining, Introduction to Linear Regression Analysis. Hoboken, NJ: John Wiley and Sons, 2021, 30 p.

[14] G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning. New York, NY: Springer, 2013, 440 p, doi: 10.1007/978-1-4614-7138-7

[15] Z. H. Zhou, Ensemble Methods: Foundations and Algorithms. New York, NY: CRC Press, 2012, 234 p.

[16] M. Zhang, Y. Zhang, J. Zhang and S. Lin. “Time Series Analysis of the Impact of Meteorological Conditions and Air Quality on the Number of Medical Visits for Hypertension in Haikou City, China”. Atmosphere. vol. 15, no. 3, p. 370, Mar. 2024, doi: 10.3390/atmos15030370.

[17] Y. Ohashi, T. Ihara, K. Oka, Y. Takane and Y. Kikegawa. “Machine learning analysis and risk prediction of weather-sensitive mortality related to cardiovascular disease during summer in Tokyo, Japan”. Scientific Reports. vol. 13, article 17020, 2023, doi: 10.1038/s41598-023-44181-9.

[18] H. C. W. Hsiao, S. H. F. Chen and J. J. P. Tsai. “Deep Learning for Risk Analysis of Specific Cardiovascular Diseases Using Environmental Data and Outpatient Records”. In 2026 IEEE 16th International Conference on Bioinformatics and Bioengineering (BIBE). pp. 369–372, 2016, doi: 10.1109/BIBE.2016.75.

[19] A. De Vita, A. Belmusto and F. Di Perna. “The Impact of Climate Change and Extreme Weather Conditions on Cardiovascular Health and Acute Cardiovascular Diseases”. Journal of Clinical Medicine. vol. 13, article 759, Jan. 2024, doi: 10.3390/jcm13030759.

[20] S. G. V. Zirbo, B. S. Hoszu, L. S. Dios, A. M. Coroiu and A. E. Croitoru. “Predicting Health Outcomes using Weather Data: A Dual Machine Learning Approach”. Procedia Computer Science. vol. 246, pp. 1399–1408, 2024, doi: 10.1016/j.procs.2024.09.721.

[21] M. Helsper, A. Agarwal and A. Aker. “The Subarachnoid Hemorrhage–Weather Myth: A Long-Term Big Data and Deep Learning Analysis”. Frontiers in Neurology. vol. 12, article 653483, May 2021, doi: 10.3389/fneur.2021.653483.

[22] H. Qiu, L. Luo, Z. Su, L. Zhou, L. Wang and Y. Chen. “Machine learning approaches to predict peak demand days of cardiovascular admissions considering environmental exposure”. BMC Medical Informatics and Decision Making. vol. 20, article 83, 2020, doi: 10.1186/s12911-020-1101-8.

[23] S. Albrecht, D. Broderick, K. Dost, I. Cheung, N. Nghiem, M. Wu, J. Zhu, N. Poonawala-Lohani, S. Jamison, D. Rasanathan, S. Huang, A. Trenholme, A. Stanley, S. Lawrence, S. Marsh, L. Castelino, J. Paynter, N. Turner, P. McIntyre, P. Riddle, C. Grant, G. Dobbie and J. S. Wicker. “Forecasting severe respiratory disease hospitalizations using machine learning algorithms”. BMC Medical Informatics and Decision Making. vol. 24, article 293, 2024, doi: 10.1186/s12911-024-02702-0.

[24] A. S. Kabildjanov, A. Nazarov, H. Umerov and B. Ibrokhimov. “Statistical models for forecast of ‘quality of the day’ for the category of facilities of meteodependent people”. In International Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities (ICISCT), Tashkent, 2021. [Online]. Available: http://www.icisct2021.org/Icisct

[25] A. I. Nazarov. “Algorithms for forecasting and supporting management decision-making in organizing the functioning of medical institutions”. Ph.D. dissertation abstract, Institute of Mathematics and Information Technologies of the Academy of Sciences of the Republic of Uzbekistan, Tashkent, 2012. [Online]. Available: https://www.inlibrary.uz/index.php/autoabstract/article/view/46878

[26] A. K. Jameson and P. R. Brennan. “Ensemble deep learning for early detection of heart failure from real-time clinical data”. IEEE Journal of Biomedical and Health Informatics. vol. 25, no. 6, pp. 1891–1899, Jun. 2021, doi: 10.1109/JBHI.2021.3065470.

[27] H. Zhang, T. Wu, S. Li and J. Li. “Multi-factor machine learning approach to hypertension risk prediction with wearable devices”. Artificial Intelligence in Medicine. vol. 120, article 102159, Jan. 2022, doi: 10.1016/j.artmed.2021.102159.

Kabildjanov Aleksandr is a candidate of technical sciences, associate professor of the department of ”Automation and control of technological processes” of the ”Tashkent Institute of Irrigation and Agricultural Mechanization Engineers” National Research University. His research interests include artificial intelligence in healthcare, artificial intelligence in finance, artificial intelligence in robotics, artificial intelligence in agriculture and computer modeling, among others.

Pulatov Giyos is a PhD student at the National Research University ”Tashkent Institute of Irrigation and Agricultural Mechanization Engineers”. Among his scientific interests are artificial intelligence in the field of healthcare, artificial intelligence in the field of finance, artificial intelligence in the field of robotics, Big Data, computer networks. Currently, he is conducting research in his scientific field.